



# Workshop de Quimiometría

Burgos, 15 y 16 de septiembre de 2008  
Facultad de Ciencias. Universidad de Burgos



UNIVERSIDAD  
DE BURGOS



Junta de  
Castilla y León



MINISTERIO  
DE EDUCACIÓN  
Y CIENCIA



## PRESENTACIÓN

Con esta tercera edición del Workshop de Quimiometría se consolida el esfuerzo continuado de los grupos de investigación españoles de esta disciplina por mantener un cauce de comunicación fluida para nuestra actividad investigadora y contribuir a la difusión de nuestra disciplina a nivel nacional. En este sentido queremos agradecer, en nombre de los quimiómetros españoles, a la coordinadora de la red, Dra. Anna de Juan, por su innegable implicación en el proyecto y al Dr. Raimundo Gargallo su dedicación a la difusión de la Quimiometría manteniendo la página web de la SEQyC.

Cumpliendo con su vocación, el III Workshop dedica una especial atención a los jóvenes investigadores en Quimiometría. Son ellos los encargados de las comunicaciones científicas y sobre todo pensando en ellos se pretende un espacio de comunicación mutua que garantice, para el futuro, la continuidad de las buenas relaciones entre los grupos españoles de esta disciplina.

Como ya es tradición de las ediciones anteriores, el workshop se abrirá mediante una conferencia plenaria impartida por un conferenciante extranjero de especial relevancia en el ámbito europeo. En esta edición, esa invitación ha sido aceptada por la Dra. Beata Walczak. Posteriormente, se presentarán y debatirán once comunicaciones orales sobre el estado actual de la Quimiometría y, asimismo, se debatirán los dieciséis carteles que permanecerán expuestos en el vestíbulo de la Facultad de Ciencias de la Universidad de Burgos que acoge el evento.

De otro lado no se puede perder la oportunidad que brinda un encuentro de estas características para crear también un espacio de conocimiento. Por ello se ha organizado, como actividad específica dentro del III Workshop, un seminario formado por cuatro conferencias sobre clasificación y modelado, impartidas por el Dr. Michele Forina, catedrático de la Universidad de Génova (Italia) y Doctor Honoris Causa por la Universidad de Burgos; el Dr. Roberto Todeschini, catedrático de la Universidad de Milán (Italia) y doctor "ad honorem" de la Universidad de Azuay (Cuenca, Ecuador); la Dra. Beata Walczak, catedrática de la Universidad de Silesia (Polonia) y el Dr. Luis Sarabia, catedrático de la Universidad de Burgos. Este grupo de técnicas de *Pattern Recognition* están siendo revisadas y cobrando gran actualidad en los nuevos retos de la quimiometría, como pueden ser la caracterización y tipificación de alimentos, medioambiente, genómica, proteómica, control de calidad, etc.

Deseo que este espacio de conocimiento y comunicación que es el III Workshop sea vivido, por cada uno de vosotros, como un evento fructífero. Su éxito ya está garantizado con vuestra presencia y el impulso joven que ella entraña.

Bienvenidos a Burgos.

M.C. Ortiz

Este evento cuenta con financiación del Ministerio de Educación y Ciencia (CTQ2007-28940-E/BQU) y de la Junta de Castilla y León (ORDEN EDU/671/2008, de 21 de abril, BOCyL de 30 de abril de 2008). Los organizadores desean agradecer también el soporte prestado por la Universidad de Burgos.



## RED TEMÁTICA DE QUIMIOMETRÍA

Institut de Química Ambiental, CSIC, Barcelona

Romà Tauler

Universidad de Almería

Antonia Garrido

Universidad de Burgos

M<sup>a</sup> Cruz Ortiz, Luis Sarabia, Ana Herrero, M<sup>a</sup> Sagrario Sánchez

Universidad de Córdoba

Ana Garrido

Universidad de Granada

Luis Cuadros, Juan Manuel Bosque

Universidad de la Rioja

Consuelo Pizarro

Universidad Politécnica de Valencia

Alberto Ferrer

Universidad de Santiago de Compostela

Rafael Cela

Universidad de Valencia

Guillermo Ramis

Universidade de A Coruña

José Manuel Andrade

Universitat Autònoma de Barcelona

Marcelo Blanco, Jordi Coello, Santiago Maspoch

Universitat de Barcelona

Anna de Juan, Raimundo Gargallo

Universitat de Girona

Manuela Hidalgo, Victòria Salvadó

Universitat de Lleida

Jaume Puy

Universitat Ramon Llull, Institut Químic de Sarrià, Barcelona

Xavier Tomàs

Universitat Rovira i Virgili, Tarragona

Ricard Boqué, Pilar Callao, Joan Ferré, María Soledad Larrechi, Itziar

Ruisánchez

## COMITÉ ORGANIZADOR

M<sup>a</sup> Cruz Ortiz (Presidenta)

Anna de Juan

Luis A. Sarabia

Ana Herrero

Celia Reguera

M<sup>a</sup> Sagrario Sánchez



# PROGRAMA

## Día 15 de septiembre

8:45-9:15 **Recepción** de participantes, salón de actos de la Facultad de Ciencias

9:15-9:30 **Inauguración del Workshop** con la presencia del Vicerrector de Investigación de la Universidad de Burgos, Dr. Jordi Rovira Carballido.

9:30-10 **Conferencia inaugural** "Use and abuse of Chemometrics", Dra. Beata Walczak

10:15-11:15 **Comunicaciones orales** (Preside: M<sup>a</sup> Sagrario Sánchez)

- O-1. Measurement error structure influence in multivariate curve resolution / Influencia de la estructura de error en resolución multivariante, *L. Blanchet*, U. Barcelona
- O-2. Quality assessment of Multivariate Curve Resolution (MCR) methods for the investigation of environmental pollution patterns, *M. Terrado*, CSIC, Barcelona
- O-3. Usefulness of excitation-emission fluorescence spectrophotometry and PARAFAC in the analytical determination where a severe matrix effect exists, *N. Rodríguez*, U. Burgos

11:15-11:45 **Pausa para café**

11:45-12:45 **Comunicaciones orales** (Preside: M<sup>a</sup> Cruz Ortiz)

- O-4. Sequential estimation of FTIR-ATR spectra of the intermediate compounds to resolve consecutive and side reactions in cationic curing process by MCR-ALS. A strategy based on convergence criterion by global phase angle, *N. Spegazzini*, U. Rovira i Virgili
- O-5. Monitorización del envejecimiento de vertidos de hidrocarburos mediante rotaciones de procrustes de tres vías, *R. Fernández*, U. A Coruña
- O-6. maSigPro and ASCA-genes: tools for the analysis of Time-Course microarray data, *M<sup>a</sup> J. Nueda*, U. Politécnica de Valencia
- O-7. Nuevas aproximaciones a la optimización computerizada de separaciones en LC-MS, *J. García*, U. Santiago de Compostela

12:45-13:45 **Comunicaciones orales** (Preside: Luis A. Sarabia)

- O-8. Classification of tumour cells from gene expression data using Probabilistic Discriminant Partial Least Squares with reject option, *C. Botella*, U. Rovira i Virgili
- O-9. Selección de píxeles, basada en información NIRS relevante, como estrategia de discriminación de imágenes-NIRS de harinas cárnicas de diferentes especies animales, *C. Riccioli*, U. Córdoba
- O-10. Application of StepwiseLDA to Classify Pottery Sherds Belonging to the Southeast of the United States, *S. Rodríguez*, U. La Rioja
- O-11. Dependencia entre los términos "suave/intenso" del aceite de oliva andaluz y diferentes parámetros físico-químicos, *P. de la Mata*, U. Granada

13:45-15:30 **Comida** en el comedor de la residencia universitaria "Camino de Santiago"

15:30-16:15 **Sesión de presentación de pósteres** (Preside: Ana Herrero):

P1-P8

16:15-17:00 Discusión de pósteres y **café**

17:00-17:45 **Sesión de presentación de pósteres** (Preside: M<sup>a</sup> Sagrario Sánchez):

P9-P16

17:45-18:00 Discusión de pósteres

18:00 **Reunión** de la red de Quimiometría

21:30 **Cena** en el comedor de la residencia universitaria "Camino de Santiago"

## **Día 16 de septiembre**

### **Seminario sobre clasificación y modelado** (Preside: Anna de Juan)

9:30-10:15 Conferencia "Classification and class modeling techniques: problem, sampling, method, evaluation", Dr. Michele Forina

10:30-11:15 Conferencia "Classification method CAIMAN and classification of sequential data", Dr. Roberto Todeschini

11:30-12:00 Pausa para **café**

12:00-12:45 Conferencia "Sensitivity and specificity as criteria to model categories", Dr. Luis A. Sarabia

13:00-13:45 Conferencia "Modeling methods based on SVMs and related techniques", Dra. Beata Walczak

14:00 **Clausura del Workshop**

14:15 **Comida** en el comedor de la residencia universitaria "Camino de Santiago"





Conferencia inaugural  
Comunicaciones orales



## USE AND ABUSE OF CHEMOMETRICS

Beata Walczak

*Institute of Chemistry, Silesian University,  
9 Szkolna Street, 40-006 Katowice, Poland*

Chemometrics is considered as part of analytical chemistry. In its arsenal we can find methods that help analytical chemists to deal with all steps of analytical procedures, starting from the design of an experiment to extraction of information and to the final decision making. A majority of chemometrics methods are general, i.e. they can be applied to any type of analytical experiment and to any type of instrumental signal. However, there are problems associated with the specific types of instrumental signals or with the particular analytical techniques that need special treatment. Only with the knowledge of the system studied and of the principles of the performed measurements, the well suited methods can be chosen.

For instance, in separation science, many new possibilities are associated with the fact, that chromatograms, electropherograms etc., can be treated as analytical signals and use for comparative analysis of different samples or systems in the same manner as spectra. In the arsenal of chemometrics methods we have all methods for signal enhancement, mixture analysis, warping, compression, visualization, calibration, classification, etc.

Although the main goal of chemometrics is to provide analytical chemists with efficient tools of data processing, it seems that these methods are mainly used by chemometricians and we are still miles away from their daily use in analytical laboratories. On the other hand, if they are used, they are not always used properly. Typical examples of misuse and abuse of chemometrics that can occur are common for many fields (separation science, food chemistry, etc.). Among them there are:

- Application of improper methods;
- Improper (or lack of) model validation;
- Problems with data representativity;
- Unfair comparison of different methods;
- Chemometrics approaches as black-box tools;
- Usless efforts.



# MEASUREMENT ERROR STRUCTURE INFLUENCE IN MULTIVARIATE CURVE RESOLUTION / INFLUENCIA DE LA ESTRUCTURA DE ERROR EN RESOLUCIÓN MULTIVARIANTE

L. Blanchet<sup>1,2</sup> C. Ruckebush<sup>2</sup> J.P. Huvenne<sup>2</sup> R. Tauler<sup>3</sup> A. de Juan<sup>1</sup>

<sup>1</sup> Grup de Quimiometria, Facultat de Química, Universitat de Barcelona, Diagonal 647, 08028 Barcelona, Spain

<sup>2</sup> Laboratoire de Spectrochimie Infrarouge et Raman (LASIR CNRS UMR 8516), Centre d'Etudes et de Recherches Lasers et Applications (CERLA CNRS FR 2416), Université des Sciences et Technologies de Lille(USTL), bât C5, 59655 Villeneuve d'Ascq, France, lionel.blanchet@ed.univ-lille1.fr

<sup>3</sup> IIQAB-CSIC, Jordi Girona 18, Barcelona 08034, Spain

**Keywords:** Error covariance, noise, WALs, time-resolved spectroscopy

The error affecting multivariate measurements is often assumed to be an independent and identically distributed noise (*iid*-normal noise). The reality of multichannel experiments might be very different, exhibiting heteroscedastic errors or correlated errors among the channels [1]. The error structure can be related to instrumentation, to environmental factors and to methodological aspects. Their overall influence on the signal must be estimated. Proper data analysis necessarily implies knowledge of the nature of error, with all the contributions that configure its correlated or uncorrelated structure [2]. Once this is known, the whole error structure can be incorporated in the data analysis through weighting schemes. Indeed curve resolution methods are known to be affected by some overfitting issues. This means that the noise or some measurement errors may be interpreted by the algorithm as a significant signal and so incorporated in the resolution. Data affected by correlated errors are particularly sensitive to this problem.

Time-resolved pump-probe absorption spectroscopy is used as a case study. This technique is widely used in photochemistry to investigate electronic excited states of the molecules [3]. The 2-way data set  $\mathbf{D}$  recorded is a collection of measurements performed simultaneously at many wavelengths for different delays with respect to the exciting pulse. Among different potential error sources, the values measured at different wavelengths are suspected to be statistically dependant, with covariance matrices to be evaluated, as a whole spectrum is measured simultaneously.

A comprehensive analysis of the error structure of pump-probe spectroscopic data is proposed using replicates of a time-resolved experiment and the Wentzell's methodology [2]. The cube  $\underline{\Sigma}$ , formed by all the error covariance matrices  $\Sigma_i$  may be analysed using 3-way methods, e.g. PARAFAC and the loadings obtained will give insight on the potential correlated contributions to the error. It is also possible to use 2-way algorithms on the averaged covariance matrix, the mean of the cube, when the different slices are equivalent. For this purpose, target testing has been proposed to identify the physical factors influencing the error structure [2].

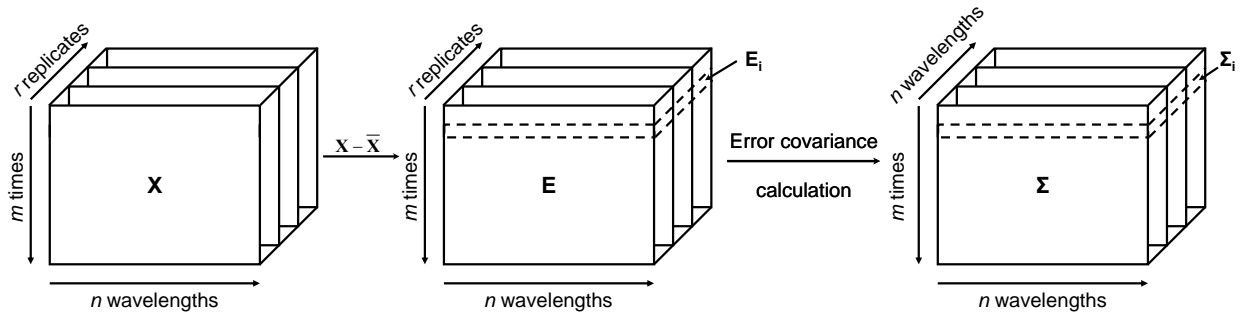


Figure 1 – Calculation of the error covariance matrix  $\Sigma$  in the spectral direction

The outputs of the analysis of the error structure, in both time and wavelength directions, are then used in a Multivariate Curve Resolution – Alternating Least Squares (MCR-ALS) [4] resolution through the suitable weighting error procedure [5]. MCR-ALS algorithm relies on the alternated least-squares decomposition of the 2-way data set  $\mathbf{D}$  in the two matrices of concentration profiles  $\mathbf{C}$  and pure spectra  $\mathbf{S}^T$ . The data matrix  $\mathbf{D}$  can be weighted according to Eq. 1 to incorporate the noise structure expected, for example, in the wavelength direction.

$$\hat{\mathbf{d}}_i = \mathbf{d}_i \cdot \Sigma_i^{-1} \cdot \mathbf{S}^T (\mathbf{S} \cdot \Sigma_i^{-1} \cdot \mathbf{S}^T)^{-1} \mathbf{S} \quad (1)$$

where  $\mathbf{d}_i$  represent the  $i^{\text{th}}$  row of  $\mathbf{D}$ . With the weighted matrix in this form, the matrix  $\mathbf{C}$  is calculated according to Eq. 2.

$$\mathbf{C} = \mathbf{D} \cdot \mathbf{S}^T (\mathbf{S} \cdot \mathbf{S}^T)^{-1} \quad (2)$$

The same operation may be done weighting according to the noise structure in the time direction using the error covariance matrix  $\Psi$ .

$$\hat{\mathbf{d}}_j = \mathbf{C} (\mathbf{C}^T \cdot \Psi_j^{-1} \cdot \mathbf{C})^{-1} \mathbf{C}^T \Psi_j^{-1} \cdot \mathbf{d}_j \quad (3)$$

where  $\mathbf{d}_j$  represent the  $j^{\text{th}}$  column of  $\mathbf{D}$ . The matrix  $\mathbf{S}$  may then be evaluated using the Eq. 4.

$$\mathbf{S} = (\mathbf{C}^T \cdot \mathbf{C})^{-1} \mathbf{C}^T \cdot \mathbf{D} \quad (4)$$

The error structure is thus incorporated in the multivariate curve resolution procedure.

## References

- [1] M. N. Leger, V. M. Montoto, P. D. Wentzell. Methods for systematic investigation of measurement error covariance matrices. *Chemom. Intell. Lab. Syst.*, 77(1-2):181-205, 2005.
- [2] M. N. Leger, V. M. Montoto, P. D. Wentzell. Methods for systematic investigation of measurement error covariance matrices. *Chemom. Intell. Lab. Syst.*, 77(1-2):181-205, 2005
- [3] G. Buntinx, O. Poizat, in: M. Mostafavi, T. Gustavsson (Eds.), Réactions ultrarapides en solutions, CNRS Editions, Paris, 2006, p. 163.
- [4] R. Tauler, Multivariate curve resolution applied to second order data. *Chemom. Intell. Lab. Syst.*, 30(1):133-146, 1995.
- [5] P. D. Wentzell, T. K. Karakach, S. Roy, M. J. Martinez, C. P. Allen, M. Werner-Washburne. Multivariate curve resolution of time course microarray data. *BMC Bioinformatics*, 7:343, 2006

# QUALITY ASSESSMENT OF MULTIVARIATE CURVE RESOLUTION (MCR) METHODS FOR THE INVESTIGATION OF ENVIRONMENTAL POLLUTION PATTERNS

Marta Terrado, Damià Barceló, Romà Tauler

Chemometric data analysis methods (1-5) provide powerful tools for analysis and interpretation of large environmental data sets generated within environmental monitoring studies. These methods assume that chemical concentrations and parameters measured in a particular sample are affected by multiple contributions coming from independent sources of different origin. These sources may be defined by pattern profiles describing their chemical composition as well as their geographical and temporal or compartmental distributions.

Principal Component Analysis (PCA) (3) is a multivariate statistical method which has been frequently used in exploratory environmental data analysis (6-8). The PCA method uses constraints like orthogonality and maximum explained variance, which are very helpful for data exploration since obtained solutions are unique, but not easily interpretable in environmental terms. An alternative approach to PCA is Multivariate Curve Resolution using Alternating Least Squares (MCR-ALS) (9), a method using more natural constraints than those used in PCA, such as non-negativity, which allows obtaining contamination profiles more easily interpretable and more similar to the ones causing the observed data variance. MCR-ALS application has already been proposed in previous works (10-12) to investigate pollution patterns from the analysis of different environmental data sets. However, results obtained by MCR methods in the analysis of two-way data matrices may be rotation and scale ambiguous (9, 13). This means that MCR solutions may not be unique, and a set of feasible solutions fulfilling postulated constraints and fitting the experimental data equally well are possible. Apart from ambiguity, noise propagation constitutes another effect to consider in relation to the parameters estimated by MCR-ALS. If there is no noise in data, only rotation and intensity ambiguities are present. But when noise exists, as it is the general case in all environmental data sets, both effects appear intermixed, being therefore difficult to discern among these two possible indeterminacy sources. Rotation ambiguities can be totally solved for data following a trilinear model. In this case, MCR-ALS solutions with trilinearity constraints can be easily compared to those obtained by PARAFAC (14). The assumption of a trilinear model simplifies data interpretation in a considerable way, since not only the description of the identified contamination patterns is the same for the different periods or environmental compartments, but also the spatial distribution of the contamination patterns follows the same trend during the different sampling campaigns (time). Nevertheless, a total trilinear data structure is a highly demanding requirement which is rarely totally fulfilled in environmental data sets. The main advantage of using MCR-ALS over other methods based on trilinear models like PARAFAC is that trilinearity constraints can be implemented in a very flexible way (15).

The present study proposes the application of MCR-ALS to experimental environmental data for identification of the main contamination patterns of pesticides existing in the Ebro river delta channels (Catalunya, NE Spain) during the rice-growing season in year 2005. Three independent contamination patterns of pesticides with different spatial and temporal behaviour are clearly identified. A pattern of pesticides coming from the Ebro river main flow is clearly distinguished from the other two patterns, originated in the delta channels due to rice crop practices. An exploratory data analysis applying PCA was already performed in a previous study to identify the general trends existing in data (16). Validation of MCR-ALS by means of simulated data set is also presented in this work. The purpose of this simulation

is, on one hand, to prove that the application of the MCR-ALS method to environmental data provides a good resolution of the main patterns of data variation and, consequently, allows the investigation of the “true” contamination sources existing in the area of study. On the other hand, the simulation contributes to an improved understanding of the MCR-ALS method possibilities and limitations, taking into account the different effects that can influence profiles resolution and recovery. Thus, the consequence of choosing a specific data pretreatment (comparing results obtained using raw, scaled and unscaled data) over the identified profiles is assessed. The influence that noise and rotation ambiguities have over profiles recovery are also studied using different simulated data sets (with and without noise) as well as applying different constraints to data (trilinearity and non-trilinearity). Finally, the role that data structure (trilinear or not trilinear) plays in the determination of the more suitable model for the analysis is also assessed. Validation of the method has the aim of establishing possible recommendations for MCR-ALS application, depending on the characteristics of the data set to be analysed.

## References

- (1) Einax, J. W.; Zwanzinger, H. W.; Geiss, S. *Chemometrics in environmental analysis*; VCH-Wiley: Germany, 1997.
- (2) Geladi, P. Analysis of multiway data; *Chemometrics Intell. Lab. Syst.* **1989**, *7*, 11-30.
- (3) Joliffe, I. T. *Principal Component Analysis*; Springer: New York, 1986.
- (4) Malinowski, E. D. *Factor Analysis in chemistry*; John Wiley & Sons: New York, 2002.
- (5) Massart, D. L.; Vandergisnte, B. G. M.; Buydens, L. M. C.; de Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. *Handbook of chemometrics and qualimetrics: parts A and B*; Elsevier: Amsterdam, 1998.
- (6) Berzas, J. J.; García, L. F.; Rodríguez, R. C.; Martín-Alvarez, P. J. *Water Res.* **2000**, *34*, 3161-3170.
- (7) St-Hilaire, A.; Brun, G.; Courtenay, S. C.; Ouarda, T. B. M. J.; Boghen, A. D.; Bobee, B. *J. Am. Water Resources Assoc.* **2004**, *40*, 691-703.
- (8) Teixidó, E.; Olivella, L.; Figueras, M.; Ginebreda, A.; Tauler, R. *Int. J. Environ. Analyt. Chem.* **2001**, *81*, 295-313.
- (9) Tauler, R.; Smilde, A.; Kowalski, B. *J. Chemometrics* **1995**, *9*, 31-58.
- (10) Felipe-Sotelo, M.; Gustems, L.; Hernández, I.; Terrado, M.; Tauler, R. *Atmos. Environ.* **2006**, *40*, 7421-7436.
- (11) Peré-Trepát, E.; Ginebreda, A.; Tauler, R. *Chemometrics Intell. Lab. Syst.* **2007**, *88*, 69-83.
- (12) Tauler, R.; Barceló, D.; Thurman, E. M. *Environ. Sci. Technol.* **2000**, *34*, 3307-3314.
- (13) Jaumot, J.; Gargallo, R.; Tauler, R. *J. Chemometrics*, **2004**, *18*, 327-340.
- (14) Bro, R. *Chemometrics Intell. Lab. Syst.*, **1997**, *38*, 141-171.
- (15) Jaumot, J.; Gargallo, R.; de Juan, A.; Tauler, R. *Chemometrics Intell. Lab. Syst.*, **2005**, *76*, 101-110.
- (16) Terrado, M.; Kuster, M.; Raldúa, D.; López de Alda, M.; Barceló, D.; Tauler, R. *Analyt. Bioanalyt. Chem.*, **2007**, *387*, 1479-1488.
- (17) Walczak, B.; Massart, D. L. *Chemometrics Intell. Lab. Syst.*, **2001**, *58*, 15-42.

## Acknowledgements

This work is supported by the European Union FP6 Integrated Project AQUATERRA (GOCE 505428) and the Spanish Ministry of Education and Science (CTQ2006-15052).



# USEFULNESS OF EXCITATION-EMISSION FLUORESCENCE SPECTROPHOTOMETRY AND PARAFAC IN THE ANALYTICAL DETERMINATION WHEN A SEVERE MATRIX EFFECT EXISTS

N. Rodríguez<sup>a</sup>, B.D. Real<sup>a</sup>, M.C. Ortiz<sup>a</sup>, L.A. Sarabia<sup>b</sup>, M.S. Sánchez<sup>b</sup> y A. Herrero<sup>a</sup>  
(<sup>a</sup>) Dto. Química, (<sup>b</sup>) Dto. Matemáticas y Computación, Fac. Ciencias, Universidad de Burgos  
Pza. Misael Bañuelos s/n, 09001 Burgos

Molecular fluorescence spectrophotometry is a technique widely used in the analysis of residues because of its high sensitivity, low cost and wide linear ranges. One of its principal disadvantages lies in the fact that fluorescence emission is influenced by factors such as pH, temperature, chemical environment, etc. In many cases, this means a limitation to the application of the technique since, when complex matrices must be analysed, it is necessary to pretreat the sample and/or to use chromatographic techniques.

However, excitation-emission fluorescence spectrophotometry provides, in a simple way, a data matrix for each sample (two-order data). For this reason, calibration methods based on three-way data are suitable to be used, i.e. data cubes made by the excitation-emission matrix (EEM) with a third way added, which is the increasing amounts of analyte. The interest of the models based on three-way data (or n-way, in general) lies in the fact that, if data are trilinear, it is possible to have a unique decomposition of the data cube, which allows one to determine an analyte in the presence of non-calibrated interferents. This is the so-called second-order advantage, which makes these methods more and more requested in routine analysis<sup>1,2</sup>. Parallel factor analysis (PARAFAC)<sup>3</sup> is among the methods that has this property.

In this work, some applications of PARAFAC when determining residues in complex matrices are shown. Through three case-studies, some of the problems that in several occasions are inherent in the excitation-emission fluorescence itself are shown, as well as the strategies proposed by our research group for solving them.

## Case I. Determination of ciprofloxacin in urine

Second-order property is very useful when working with biological samples since they usually have a high complex composition. Further, this composition varies among individuals and even within the same individual, which hinders the obtaining of signals specific of the analyte. In this work, the possibility of determining ciprofloxacin in human urine without previous extraction stages is studied. For that, a global PARAFAC model is fitted with the excitation-emission fluorescence data. It has to be taken into account that human urine consists of numerous substances, some of them are fluorescent. PARAFAC decomposition is made in the usual way with calibration samples spiked with ciprofloxacin, and with other four series of samples also spiked. Samples correspond to different individuals and some of them contain in addition a certain amount of another drug as interferent since they come from a patient that was treated with mesalazine. The use of the PARAFAC model has made possible

to identify and quantify the ciprofloxacin present in urine in a fast and precise way, with independence of the individual and on the presence of other fluorescent species in the matrix.

### **Case II. Determination of tetracycline in tea**

In this case, quantification in the presence of a strong quenching caused by matrix is studied. This phenomenon does not imply any modification of the shape of the signals of the analytes, but a decrease of their magnitude. The analysis of EEM in the determination of tetracycline in tea solutions, which acts as quencher, points out the ineffectiveness of three-way PARAFAC and standard addition method, in spite of the fitting of the trilinear model to the experimental tensor. Physically, the quencher causes a multiplicative effect in the EEM which can be modelled using four-way PARAFAC decomposition. Therefore, a fourth way is added to the tensor of data by varying the amount of quencher, tea, in such a way that this four-way PARAFAC decomposition gives successful results<sup>4</sup>.

### **Case III. Determination of tetracycline in milk**

Another quite general situation is the presence of other species in the matrix that affect the fluorescence signals. In this case, the interferent effect of some inorganic species (Ca(II), K(I), P(V), Fe(III), Na(I), Zn(II) and Mg(II)) on the fluorescence emission of tetracycline is studied, effects that produce, for instance, that the tetracycline signal is enhanced, inhibited and/or modified in a significant way. This study concludes that the tetracycline signal lost trilinearity in the presence of the studied interferents at the concentrations at which they are present in a natural way in milk. Although the standard addition method is the tool frequently used to obviate this interference, its use with PARAFAC does not give successful results due to the severe matrix effect that strongly modifies both the size than the shape of the tetracycline signal. For solving this problem, we propose an experimental strategy that consists of obtaining several dilutions of the sample, previously deproteinized, and spiked them with increasing amounts of tetracycline. The idea is to include in the three-way PARAFAC model enough variability of the substances present in the matrix. This strategy has been essential to successfully solve the determination of tetracycline in milk.

### **Acknowledgements**

This work has been partially supported by Ministerio de Ciencia e Innovación (Project CTQ2008-02264/BQU), and JCyL (Project BU024A07), both with FEDER funds. N. Rodríguez thanks the JCyL (Orden EDU/1165/2007) by her FPI grant.

### **References**

1. G.M. Escandar, A.C. Olivieri, N.M. Faber, H.C. Goicoechea, A. Muñoz de la Peña, R.J. Poppi, *TrAC Trends Anal. Chem.*, 26 (2007) 752-765.
2. D. Giménez, L.A. Sarabia, M.C. Ortiz, *Analyst*, 130 (2005) 1639-1647.
3. R. Bro, *Chemom. Intell. Lab. Systems*, 38 (1997), 149-171.
4. N. Rodríguez, M.C. Ortiz, L.A. Sarabia, *Talanta* (in press, August 2008)

# SEQUENTIAL ESTIMATION OF FTIR-ATR SPECTRA OF THE INTERMEDIATE COMPOUNDS TO RESOLVE CONSECUTIVES AND SIDES REACTIONS IN CATIONIC CURING PROCESS BY MCR-ALS. A STRATEGY BASED ON CONVERGENCE CRITERION BY GLOBAL PHASE ANGLE

Nicolás Spegazzini, Itziar Ruisánchez, M. Soledad Larrechi

Chemometrics Qualimetrics and Nanosensors Group. Department of Analytical and Organic Chemistry, Rovira i Virgili University. Marcel·lí Domingo s/n, 43007 Tarragona, Spain.

[nicolas.spegazzini@urv.cat](mailto:nicolas.spegazzini@urv.cat); [itziar.ruisanchez@urv.cat](mailto:itziar.ruisanchez@urv.cat); [marioledad.larrechi@urv.cat](mailto:marioledad.larrechi@urv.cat)

## Abstract

Of the different methods of polymerization, the ring-opening mechanism leads to lower shrinkage introduced the term *expanding monomers*, which refers to monomers that lead to zero shrinkage or even positive expansion during polymerisation. In this work a representative reaction of this mechanism is studied: the reaction between diglycidyl ether of bisphenol A (DGEBA) and  $\gamma$ -valerolactone ( $\gamma$ -VL). In its curing process several and competitive reactions can be done. In a previous work [1] these reactions has been established using 2D correlation spectroscopy of the spectra obtained in the monitoring by Fourier Transform Infrared (FTIR-ATR) Spectroscopy.

The present work is focussed in the quantitative resolution of the FTIR-ATR spectra using multivariate resolution method (MCR-ALS) [2] and using the error criterion based on global phase angle for to identified the optimal number of iterations in ALS [3]. A previous and very important step, in the optimization process of MCR-ALS, is to overcome the rank deficiency often found in experimental data obtained from a chemical reaction. An usual strategy consist in incorporating spectral information of the reactants and final products [4]. In a copolymerization process, as the studied in the present work, it is not possible to obtain a spectrum representative of the final product and also there are several and parallel reactions. So, a strategy to estimated the spectrum of the intermediates and final compounds is proposed based in the sequential application of MCR-ALS to the time intervals identified by EFA [1], where the number de reactions involved are lower. This information is incorporated to the data matrix in order to overcome the deficiency rank. The spectra estimated as representatives of each species during this sequential process, are used as initial estimates in the final step of optimization using ALS with all data.

The results are evaluated by studying the residuals and parameters such as lack of fit, the percentage of explained variance and the coefficient of similarity ( $r$ ) between the recovered spectra and the spectra of the pure species, if is possible. For intermediates species a qualitative evaluation of the correspondence between the ALS solution and the chemistry knowledge of this species is also made.

## References

1. N. Spegazzini, I. Ruisánchez, M. S. Larrechi, A. Serra, A. Mantecón, J. Polym. Sci. Part A: Polym. Chem. 46 (2008) 3886.
2. R. Tauler, A. Smilde, B. R. Kowalski, J. Chemometrics 9 (1995) 31.
3. H. Shinazawa, M. Iwahashi, I. Noda, Y. Ozaki J. Mol. Struct. 883-884 (2008) 73.
4. M. Amrhein, B. Srinivasan, D. Bonvin, M. Shumacher, Chemometrics Intell. Lab. Syst. 33 (1996) 17.

## MONITORIZACIÓN DEL ENVEJECIMIENTO DE VERTIDOS DE HIDROCARBUROS MEDIANTE ROTACIONES DE PROCUSTES DE TRES VÍAS

**R. Fernández Varela, J.M. Andrade, S. Muniategui, D. Prada**

Depto. Química Analítica. Universidade da Coruña. Campus da Zapateira, s/n, E-15071, A Coruña, España. [rfernandezv@udc.es](mailto:rfernandezv@udc.es), tel: +34-981-167000, fax: +34-981-167065

Aunque existen emisores naturales de hidrocarburos al medio natural (filtraciones de bolsas petrolíferas, volcanes, incendios, descomposición orgánica, etc.), su presencia habitual en la mayor parte de los ecosistemas se debe a fuentes antropogénicas. Típicamente, escorrentías urbanas, vertidos de residuos, calefacción, etc. Aunque estas fuentes difusas general en torno al 64% de los hidrocarburos que llegan al medio natural, la fuente más “llamativa” es la relacionada con vertidos de accidentes de buques-tanque (aporta un 5% de los hidrocarburos que llegan al medio natural).

El estudio de los hidrocarburos vertidos al medio ambiente es muy complejo y exige técnicas Químico-Analíticas especializadas (HPLC-MS, GC-MS, RMN, etc.), las cuales buscan, en general, abordar una de las dos cuestiones clave: (1) elucidar la fuente contaminante y (2) evaluar la evolución del vertido en el tiempo (o pauta de envejecimiento). Este punto, a su vez, se relaciona con la toxicidad a medio-largo plazo y el establecimiento de estrategias de actuación. La presente comunicación aborda el segundo aspecto citado.

Tras el accidente del buque *Prestige-Nassau* (13-Nov-2002) se diseñó un estudio comparativo entre el proceso de envejecimiento del fuel vertido y otros cinco productos seleccionados por su especial interés: un fuel oil con características físicas similares al vertido desde el buque y cuatro crudos de petróleo seleccionados de una colección más amplia. Se realizaron vertidos controlados en recipientes especialmente diseñados y se monitorizaron los procesos de envejecimiento a lo largo de cuatro meses empleando CG-MS (modo SIM). El objetivo es determinar la existencia de pautas de envejecimiento comunes entre los vertidos y estudiar qué variables pueden resultar más relevantes para su seguimiento.

La matriz tridimensional de datos era de dimensiones 17 muestras (“filas”) x 85 variables (“columnas”) x 6 productos (“capas”). Para simplificar la interpretación se separaron los compuestos alifáticos (35 columnas) de los aromáticos (50 columnas). Los cubos de datos se estudiaron mediante PARAFAC, PCA de matriz aumentada (MA-PCA) y una nueva opción desarrollada en nuestro grupo basada en las rotaciones de Procustes.

La rotación de Procustes es una técnica intuitiva, sencilla y rápida diseñada para comparar varios conjuntos de datos a través de los subespacios de sus componentes principales. Se definen unos nuevos factores (vectores consenso) que describen las pautas generales y mediante una estrategia similar a la desarrollada para MA-PCA se estudia la distribución promedio de las muestras en el tiempo y la influencia promedio del producto en el modelo. Los resultados encontrados demuestran que los resultados son totalmente comparables a los obtenidos mediante las técnicas de PARAFAC y MA-PCA, con lo cual se puede postular como una alternativa de interés para tratar datos de tres vías.

### AGRADECIMIENTOS

Se agradece el apoyo económico de los gobiernos español y gallego (PROFIT-050100-2003-4, PROFIT-310200-2004-106, PROFIT-310200-2005-37, PGIDIT03RMA16401PR), con la participación de los fondos FEDER.



## MASIGPRO AND ASCA-GENES: TOOLS FOR THE ANALYSIS OF TIME-COURSE MICROARRAY DATA

María José Nueda<sup>1</sup>, Ana Conesa<sup>2</sup>, and Alberto Ferrer<sup>1</sup>

<sup>1</sup>Departamento de Estadística e Investigación Operativa Aplicadas y Calidad. Universidad Politécnica de Valencia. Cno. Vera s/n. Edificio I-3. Apartado 46022. Valencia (SPAIN).

<sup>2</sup>Bioinformatics Department, Centro de Investigación Príncipe Felipe Departamento, Valencia (SPAIN).

Gene expression profiling has imposed as a frequent approach to study the molecular basis of a wide variety of biological phenomena. In multiple series of time course experiments (MTC) a large number of genes are measured on different conditions, which are combinations of time and different experimental groups. Identifying genes with different trends over time for the experimental groups and the levels of the factors that produce the differences are the main objectives in these studies. In this talk we introduce two novel statistical methods for the analysis of MTC: maSigPro and ASCA-genes.

The method maSigPro [1] (microarray Significant Profile) applies a regression model per gene to address the problem of evaluating statistically significant profile differences and selects genes for which the model is statistically significant. The analysis of the regression coefficients of the fitted model for each significant gene permits the identification of the conditions for which the gene shows statistically significant profile changes. The maSigPro methodology has been implemented in the statistical language R and it is freely available for the Bioconductor contributed packages repository <http://www.bioconductor.org/>.

ASCA (ANOVA-SCA, Analysis of Variance - Simultaneous Component Analysis) is a statistical approach for the analysis of designed experiments. Basically, ASCA applies PCA (Principal Components Analysis) to the estimated parameters in each source of variation of an ANOVA model. In other words, ASCA splits the data on the different existing effects (time, experimental groups, their interactions and the residuals) and explores the underlying variation in each one. The application of ASCA to MTC has resulted in the development of a specific methodology, called ASCA-genes [2], to select the genes of interest.

[1] Conesa A., Nueda, M.J., Ferrer, A. and Talón, M. (2006) maSigPro: a Method to Identify Significantly Differential Expression Profiles in Time-Course Microarray Experiments. *Bioinformatics*, 22 (9), 1096-1102.

[2] Nueda, M.J.; Conesa, A.; Westerhuis, J.A.; Hoefsloot, H.C.J.; Smilde, A.K.; Talón, M. and Ferrer, A. (2007) Discovering gene expression patterns in Time Course Microarray Experiments by ANOVA-SCA. *Bioinformatics*, 23 (14), 1792-1800.





## NUEVAS APROXIMACIONES A LA OPTIMIZACIÓN COMPUTERIZADA DE SEPARACIONES EN LC-MS

**J. García-Lavandeira, B. Losada, J.A. Martínez-Pontevedra, M. Lores and R. Cela**

Instituto de Investigación y Análisis Alimentarios. Dpto. Química Analítica, Nutrición y Bromatología. Universidad de Santiago de Compostela.

El desarrollo de métodos asistido por computador en HPLC ha sido un campo de investigación permanentemente abierto desde la década de los 70, con aportaciones intermitentes tanto desde la comunidad académica como desde los desarrolladores de software comercial. En general, estas propuestas consideran la resolución de los picos como un elemento clave en el concepto de la aproximación y por ello, la inmensa mayoría de las funciones objetivo desarrolladas utilizan de alguna forma la resolución de picos. Mientras el detector por excelencia en rutina ha sido el UV-Vis, esta aproximación es perfectamente lógica y el cromatografista está acostumbrado de decidir cuando una separación es mejor o peor fijándose de forma preferente en la resolución de los componentes que forman la mezcla a separar.

Más recientemente, la espectrometría de masas ha entrado en los laboratorios analíticos como un sistema de detección en LC de extraordinarias posibilidades. Cada día es mayor el número de laboratorios que utilizan LC-MS como una herramienta para desarrollo y también para análisis rutinario. Esta nueva situación no solo proporciona ventajas al cromatografista, sino también alguna pregunta, como puede ser: ¿es razonable transferir sin más, los procedimientos ya optimizados y en rutina para LC-UV a LC-MS? ¿No se podría conseguir una separación más rápida, y por tanto más eficiente, puesto que la herramienta de detección proporciona mucha más selectividad? ¿Sería posible desarrollar métodos de separación en sistemas LC-UV, mucho más baratos y luego transferirlos a sistemas LC-MS, listos para su aplicación?

Parece obvio que una transferencia directa de los “viejos” procedimientos no tiene porque ser lo más recomendable. Un MS con su enorme selectividad puede permitirnos ignorar fuertes solapamientos entre picos, siempre y cuando las señales de sus espectros de masa sean suficientemente diferentes (selectivas). Por tanto, una solución a este tipo de preguntas sería desarrollar un sistema que:

- Utilice la potencia de discriminación del detector hasta donde sea posible
- Ajuste la separación en la columna allí donde el detector no es capaz de proporcionar suficiente selectividad, pero no más, de modo que el tiempo de análisis pueda reducirse sensiblemente.
- Haga todo esto, de forma automática y fiable para el cromatografista

En esta ponencia se presenta el desarrollo y resultados de un nuevo sistema que trata de cumplir los objetivos anteriormente mencionados. Se parte de un marco genérico de sistema de desarrollo de métodos asistido por computador (PREGA) que ha venido siendo desarrollado durante años en nuestro laboratorio. Esto permite encajar el problema de las separaciones por LC-MS como un problema más de optimización de separaciones en LC. Además, permite reutilizar todas las herramientas ya en uso con estos fines, de modo que si las separaciones habían sido ya desarrolladas con la herramienta PREGA, disponemos de todos los datos necesarios y no es preciso efectuar nuevos experimentos para acometer el problema en LC-MS.

Es importante hacer notar que una herramienta genérica como la mencionada es suficientemente versátil como desarrollar un modelo de retención que después se aplica a condiciones operacionales particulares según las necesidades del cromatografista. Por tanto, no solo estamos hablando de “reutilizar separaciones ya optimizadas”, sino también de utilizar los equipos de menor coste y ocupación para el desarrollo mientras que los sistemas LC-MS se mantienen completamente operativos durante el trabajo de desarrollo de los modelos de retención, etc.

El mecanismo, tal y como ha sido implementado en la práctica parte del modelo de retención para los componentes de interés en la mezcla y en una columna cromatográfica particular. A partir de los espectros de masa de esos componentes, se calcula una función objetivo que se ha denominado *matriz de selectividad*, que implica mapear las condiciones de interferencia entre señales de los espectros de masa, en términos de la resolución requerida entre pares de picos particulares. La matriz de selectividad y el tiempo de elución total, se utilizan para proceder a la optimización Pareto de la separación utilizando algoritmos evolutivos. El resultado es el frente Pareto óptimo dentro del que se elige la solución o soluciones más convenientes en términos prácticos para el cromatografista. Todo el proceso se ha facilitado mediante el desarrollo de una herramienta de software específica denominada PREGA-LC-MS, cuyos detalles y resultados se presentan.

# CLASSIFICATION OF TUMOUR CELLS FROM GENE EXPRESSION DATA USING PROBABILISTIC DISCRIMINANT PARTIAL LEAST SQUARES WITH REJECT OPTION

Cristina Botella, Joan Ferré\* and Ricard Boqué

*Chemometrics, Qualimetrics and Nanosensors Group. Department of Analytical Chemistry and Organic Chemistry. Rovira i Virgili University. Tarragona, Catalonia, Spain. \*e-mail address: joan.ferre@urv.cat*

A classifier must provide reliable classifications. This is especially important in clinical diagnosis, because of the potentially large impact of the classification results when these methods are applied in practice.

The overall reliability of the classifications can be increased by allowing the classifier to reject classification if the probability of error is too high, and to classify only when there is a large certainty that the classification is correct. By classifying only those objects that have the largest certainty of being correctly classified, classification with reject option improves the accuracy of the classifier and yields classifications with increased reliability. The benefits of the reject option are obtained for (a) objects that do not belong to any of the classes, (b) objects that belong to one of the classes but they are very different from those used for training the classifier, and (c) objects that lie in the boundary region between classes. In order to solve these unclear classification situations, the reject option has been introduced in various ways [1-2]. Concretely, Chow [2] derived a reject criterion for Bayesian classification, which is suitable for dealing with situation (c). Shortly stated, the Bayes rule assigns the object to the class for which it has the highest *a posteriori* probability to belong to [3]. And Chow's rule states that the object should only be assigned if the highest *a posteriori* probability is above some threshold value. The rejection of objects in situations (a) and (b) must be dealt with criteria different than the Bayes rule and the Chow's rule. In addition, the application of these rules is often limited in practice by the need of knowing the probability density functions of each class.

In this work, we show how DPLS can be used to implement a classification system based on Bayesian classification with the possibility of rejection. For that purpose, a new probabilistic version of DPLS has been used. This probabilistic version takes into account the uncertainty of the predicted value ( $\hat{y}_i$ ). This allows evaluating the probability density functions of each class and hence the *a posteriori* probabilities of the classification, based on which the Bayes rule for classification is applied. In addition to Bayes rule, reject constraints are applied to solve situations (a)-(c). This approach has been tested to discriminate between healthy and different tumour tissues (ovary, colon, lung, liver, stomach, brain, pancreas, prostate...) using microarray gene expression data [4]. The results have confirmed the importance of the reject option, since some rejected samples would have been classified wrongly without reject option.

- 
1. Tortorella, F., *A ROC-based reject rule for dichotomizers*. Pattern Recognition Letters, 2005. **26**: p. 167-180.
  2. Chow, C.K., *On optimum recognition error and reject tradeoff*. IEEE -Transactions on information theory, 1970. **16**(1): p. 41-46.
  3. Duda, R.O., P.E. Hart, and D.G. Store, *Pattern Classification (2nd edition)*, ed. W. Interscience. 2001, New York.
  4. Lu, J., et al., *MicroRNA expression profiles classify human cancers*. Nature Letters, 2005. **435**: p.834-838.



## SELECCIÓN DE PÍXELES RELEVANTES PARA INCREMENTAR LA CAPACIDAD DE DISCRIMINACIÓN EN IMÁGENES-NIRS DE HARINAS CÁRNICAS DE DIFERENTES ESPECIES ANIMALES

C. Riccioli, P. Vallesquino-Laguna, D. Pérez-Marin, J.E. Guerrero-Ginel., A. Garrido-Varo  
*Dpt. Animal Production, ETSIAM, University of Cordoba, Campus Rabanales, Ed. Prod. Animal 14071  
Córdoba (Spain) [z62rirc@uco.es](mailto:z62rirc@uco.es)/[es2valap@uco.es](mailto:es2valap@uco.es)/[pa1gava@uco.es](mailto:pa1gava@uco.es)/[pa2pemad@uco.es](mailto:pa2pemad@uco.es)*

Con la crisis de la Encefalopatía Espongiforme Bovina, la Comisión Europea (CE ) prohibió totalmente el uso de proteínas de origen animal en piensos compuestos [1]. Posteriormente, la CE aprobó el Reglamento 1774/2002/CE, que establecía las normas sanitarias aplicables a los subproductos animales no destinados al consumo humano [2]. Este documento señala de forma explícita que sólo se permite el uso en alimentación animal de los subproductos procedentes de animales aptos para el consumo humano y que reúnan ciertos requisitos de seguridad. Su publicación reabrió el debate sobre la posibilidad de volver a utilizar proteínas animales elaboradas en la alimentación animal, aunque solamente para monogástricos, y evitando siempre el reciclado intra-especie o canibalismo. No obstante, el levantamiento de la prohibición total de uso de harinas animales quedaba condicionado al desarrollo y validación de métodos de análisis capaces de detectar dichos ingredientes en piensos. Aunque la microscopía clásica es el método oficial reconocido para tal objetivo, sus limitaciones para diferenciar entre especies animales y para cuantificar, han hecho necesario el estudio y desarrollo de nuevos métodos alternativos, entre los que se encuentra la tecnología de imagen-NIR, la cual frente a la Espectroscopía NIR tradicional proporciona la oportunidad de analizar “microrregiones” localizadas dentro de un producto [2]. Un instrumento NIR-Imaging típico recoge simultáneamente miles de espectros NIR espacialmente distintos y almacena los datos creando matrices de miles de espectros [3]. En el desarrollo de una aplicación NIR-imagen robusta, es esencial el optimizar numerosos aspectos (ej. aspectos de iluminación, control de temperatura, niveles de saturación, soporte de muestra, control interno del instrumento, colocación del objetivo, captura y tratamiento de imágenes [4]). El objetivo de disminuir el número de píxeles por muestra tiene su fundamento en tres aspectos. El primero es la necesidad de acelerar los cálculos, ya que al obtener 76.800 espectros por muestra se crean imágenes de gran tamaño (26 MB), lo que dificulta y ralentiza el manejo de dichas bases de datos. El segundo motivo está relacionado con la calidad de los espectros asociados a los píxeles, para eliminar fenómenos de no-linealidad y de zonas con iluminación irregularmente distribuida. El tercer aspecto que hay que tener en cuenta analizando una imagen NIR es la falta de especificidad que ciertos píxeles pueden tener en relación a la especie animal, ya que

los píxeles y los espectros asociados pueden corresponder a partículas singulares de la muestra medida o a otras equivalentes en otras muestras. Por esta razón, hay que encontrar una metodología que permita seleccionar sólo los píxeles que realmente permitan discriminar entre especies diferentes. El objetivo de este trabajo es evaluar la capacidad discriminadora de la combinación imagen-NIR utilizando diferentes estrategias basadas en la identificación de longitudes de onda de absorción de compuestos orgánicos (grasas y proteínas) para la discriminación entre harina de carne de rumiantes y harina de ave. Se utilizaron 4 muestras de harina de carne y huesos de bovino (BBM) y 4 de harina de ave (ABM), no sedimentadas ambas, que fueron analizadas en un espectrómetro MatrixNIR (Malvern Instruments, Olney, Maryland) equipado con un detector “focal-plane array” activo en la región espectral 900-1750 nm. Para cada imagen, se compararon dos estrategias diferentes basadas en la elección de espectros con valores de absorbancia similares a longitudes de onda específicas (1200 y 1520 nm), que representan zonas de absorción de grasas y proteínas, respectivamente. Una vez seleccionados los píxeles representativos, se desarrollaron modelos predictivos PLSDA. Los mejores resultados para la discriminación de ave y vacuno se obtuvieron utilizando la librería formada por los píxeles más cercanos a la moda en 1200 nm, relacionada con la absorción de grasa, alcanzando un poder discriminatorio cercano al 100%.

[1] CE. 2000c. Decisión del Consejo 2000/766/CE de 4 de diciembre de 2000 relativa a determinadas medidas de protección contra las encefalopatías espongiformes transmisibles y la utilización de proteínas animales en la alimentación animal. Diario Oficial No. L 306 de 7.12.2000, pp. 32-33.

[2] CE. 2002. Reglamento (CE) N° 1774/2002 del Parlamento Europeo y del Consejo de 3 de octubre de 2002 por el que se establecen las normas sanitarias aplicables a los subproductos animales no destinados al consumo humano. Diario Oficial n° L 273 de 10.10.2002, pp. 0001 – 0095.

[3] Lewis, E.N., Kidder, L.H., Lee, E., Haber, K.S.: *Near-infrared spectral imaging with focal plane array detectors*. In *Spectrochemical Analysis using Infrared Multichannel Detectors*. (Bhargava R. and Levin, W., eds.), Blackwell Publishing, pp. 25-55, 2005.

[4] De la Haba, M.J., Garrido-Varo, A., Pérez-Marín, D., Guerrero, J.E. (2007). Near infrared análisis as a first-line screening technique for identifying animal species in rendered animal by-product meals. *Journal of Near Infrared Spectroscopy* 15, 237-245.

## APPLICATION OF STEPWISELDA TO CLASSIFY POTTERY SHERDS BELONGING TO THE SOUTHEAST OF THE UNITED STATES

C. Pizarro, S. Rodríguez-Tecedor, N. Pérez-del-Notario, C. Sáenz-González, J. M. González-Sáiz

*Departamento de Química. Universidad de La Rioja, C/Madre de Dios 51, 26006 Logroño (La Rioja)*

The compositional information obtained from the Instrumental Neutron Activation Analysis (INAA) of 186 pottery sherds belonging to 21 different regions across the Southeast of United States available in Steponaitis data set at MURR database was combined with Chemometric techniques to establish a detailed geographical classification.

Analytical techniques used in chemical composition analysis supply large sets of data and variables. This information is used to establish a compositional “fingerprint” of samples, which is a powerful tool in a provenance studies. Therefore, the application of multivariate treatments, variable selection, and pre-processing data is generally necessary in order to recognize compositionally homogeneous groups of samples that may be associated with geographical areas.

The data set selected has been previously studied by Steponaitis et al. (1996) [1]. The authors of this work classified the sherds in four geographical regions (Northern, Southern, Eastern and Western) based on compositional information but some questions were unsolved. We consider two of them, first the fact that the 30 per cent of the objects were ungrouped and second, that the subdivision of the regions into finer areas were not studied.

The present work aims improving this contribution by subjecting the elemental composition data obtained by INAA to different Chemometric tools such as Linear Discriminant Analysis (LDA) [2] combined with pretreatments and stepwiseLDA [2] analysis.

This study allows not only to confirm previous conclusions for the classification in four geographical areas but also for subdividing the samples into finer units obtaining high quality validated values for classification (89.74 - 100.00%) and prediction (86.21– 97.67%).

LDA was applied after column pre-treatment as a classification method working on PCs when were necessary due to the large number of original variables, and using leave-one-out cross-validation for the validation of results. In many cases classification and prediction results were improved by applying StepwiseLDA as feature selection technique based on minimum Wilks Lambda to eliminate non informative variables. All calculations were carried out with V-Parvus software [3].

Classification and prediction rates obtained after applying LDA, with only 6 sherds not assigned, to the place where were founded, are reported in Table1.

In the first step, LDA was performed on the column centered data of the whole dataset, without any variable selection and using leave-one-out cross-validation to validate the results.

Secondly, each group was separately studied and subdivided by applying StepwiseLDA followed of the multivariate treatments used in the first step. Fig. 1 shows one of these separations.

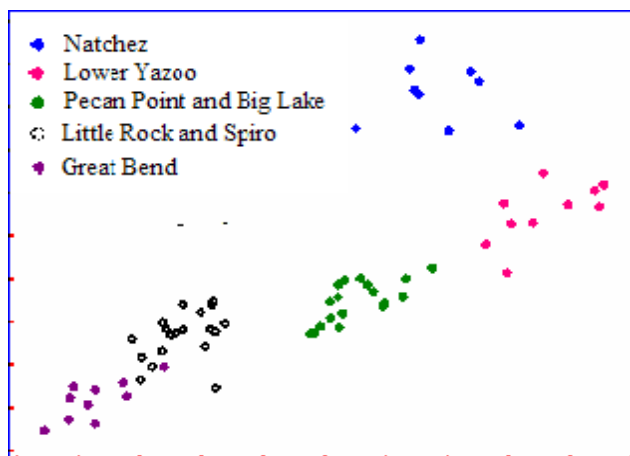


Fig 1. Projection on the two first canonical variables of LDA for Western region

Finally the categories that overlapped in the second step were studied again in order to separate them.

Table 1: LDA results: percentage of correct classifications and predictions

	<b>Subdivisions achieved</b>	<b>Classification ability (%)</b>	<b>Prediction ability (%)</b>
<b>Full set</b>	Northern; Southern; Eastern; Western	94.70	86.56
<b>North</b>	Nashville; Tellico & Servieville; Harpeth	97.06	90.00
	Tellico & Servieville	99.75	90.00
<b>South</b>	Mobile Delta & Bay; Black Warrior & Gainesville; Wheeler Lake, Wetumpka	100.00	97.67
<b>East</b>	Carters Lake; Fort Gaines; Eufaula	96.43	72.41
<b>West</b>	Lower Yazoo, Natchez, Pecant Point & Big Lake; Great Bend, Little Rock & Spiro	100.00	87.88
	Pecant Point & Big Lake	94.74	89.47
	Little Rock & Spiro	95.06	88.89

The successful results obtained for the geographical discrimination of most of the sherds are an improvement on those obtained in a previous study, thus proving the potential of the multivariate data analysis applied. A large number of samples would be necessary to include Natchitoches region in the study and for the separation of the overlapped regions in Southern region.

## References

- 1) V. P. Steponaitis, M. J. Blackman, H. Neff, Large-Scale patterns in the chemical composition of Mississippian pottery, *American Antiquity*, 6(3), (1996), 555-572
- 2) C. Pizarro, I. Esteban-Diéz, C. Saénz-Gonzalez, J.M. Gonzalez-Saiz, Vinegar classification based on feature extraction and selection from headspace solid-phase microextraction/gas chromatography volatile analyses: A feasibility study, *Anal. Chim. Acta*, 608(1), (2008), 38-47
- 3) M. Forina, S. Lanteri, C. Armanino, M. C. Cerrato-Oliveros, C. Casolino, V-Parvus. An extendable package of programs for explorative data analysis, classification and regression analysis, Dipartimento di Chimica e Tecnologie Farmaceutiche ed Alimentari, Università di Genova, Italy (2004), free available at <http://www.parvus.unige.it>



# DEPENDENCIA ENTRE LOS TÉRMINOS "SUAVE/INTENSO" DEL ACEITE DE OLIVA ANDALUZ Y DIFERENTES PARÁMETROS FÍSICO-QUÍMICOS

P. de la Mata Espinosa <sup>1</sup>, L. Cuadros Rodríguez <sup>2</sup>, J.M. Bosque Sendra <sup>3</sup>, C. Ruiz Samblás <sup>4</sup>  
Unidad de Cualimetría y Metrología Química (CMQ)  
Departamento de Química Analítica, Universidad de Granada.  
Email: 1 pmata@correo.ugr.es; 2 lcuadros@ugr.es; 3 jbosque@ugr.es; 4 crsamblas@ugr.es

## INTRODUCCIÓN

El Reglamento (CE) No. 640/2008 que modifica el No. 2568/91 considera la terminología opcional para el etiquetado de los aceites de oliva vírgenes, en donde indica que los términos "intenso" y "suave" (ligero) se definen en función a la intensidad y a la percepción de los atributos positivos obtenidos en análisis sensorial.

Estas mismas denominaciones se utilizan para calificar algunos aceites refinados comerciales y es habitual encontrar en la etiqueta declaraciones como "suave" e "intenso". Sin embargo, no existen una regla objetiva que motive tal declaraciones y la industria lo decide en base a criterios propios. En este trabajo se pretende establecer una relación de estos términos con diferente parámetros físicoquímicos característicos de los aceites de oliva y relacionados con el color y sabor.

## METODOLOGÍA

El estudio se realizó sobre los datos de un conjunto de 279 muestras de aceites de oliva andaluces (mezcla de oliva virgen y refinado). Dentro de estas muestras se pueden encontrar dos subconjuntos de objetos, el primero compuesto de 129, caracterizados por 17 parámetros físicoquímicos: absorbancia a 415 nm, absorbancia a 454 nm, absorbancia a 483 nm, absorbancia a 583 nm, absorbancia a 671 nm, índice general de color (IGC), luminancia, índice K a 225 nm, índice K a 270 nm, índice  $\Delta K$ , acidez libre (ACID), estabilidad Rancimat, contenido en  $\alpha$ -tocoferol, contenido en  $\beta$ -tocoferol, contenido en  $\gamma$ -tocoferol, contenido en  $\delta$ -tocoferol. Por su parte el segundo subconjunto esta constituido de 150 objetos y caracterizados únicamente por 6 parámetros físicoquímicos: IGC, índice K a 225 nm, índice K a 270 nm, índice  $\Delta K$ , ACID, estabilidad Rancimat.

El estudio se realizó aplicando herramientas de tratamiento de datos multivariantes para obtener la clasificación de los objetos en 2 grupos de acuerdo a los términos suave e intenso como:

- Análisis exploratorio de componentes principales (PCA)
- Análisis de la correlación
- Análisis discriminante lineal (LDA)

El análisis de los datos se ha realizado utilizando los siguientes programas:

- Statgraphics Plus, ver 5.1
- The Unscrambler, ver 8.0

## **RESULTADOS**

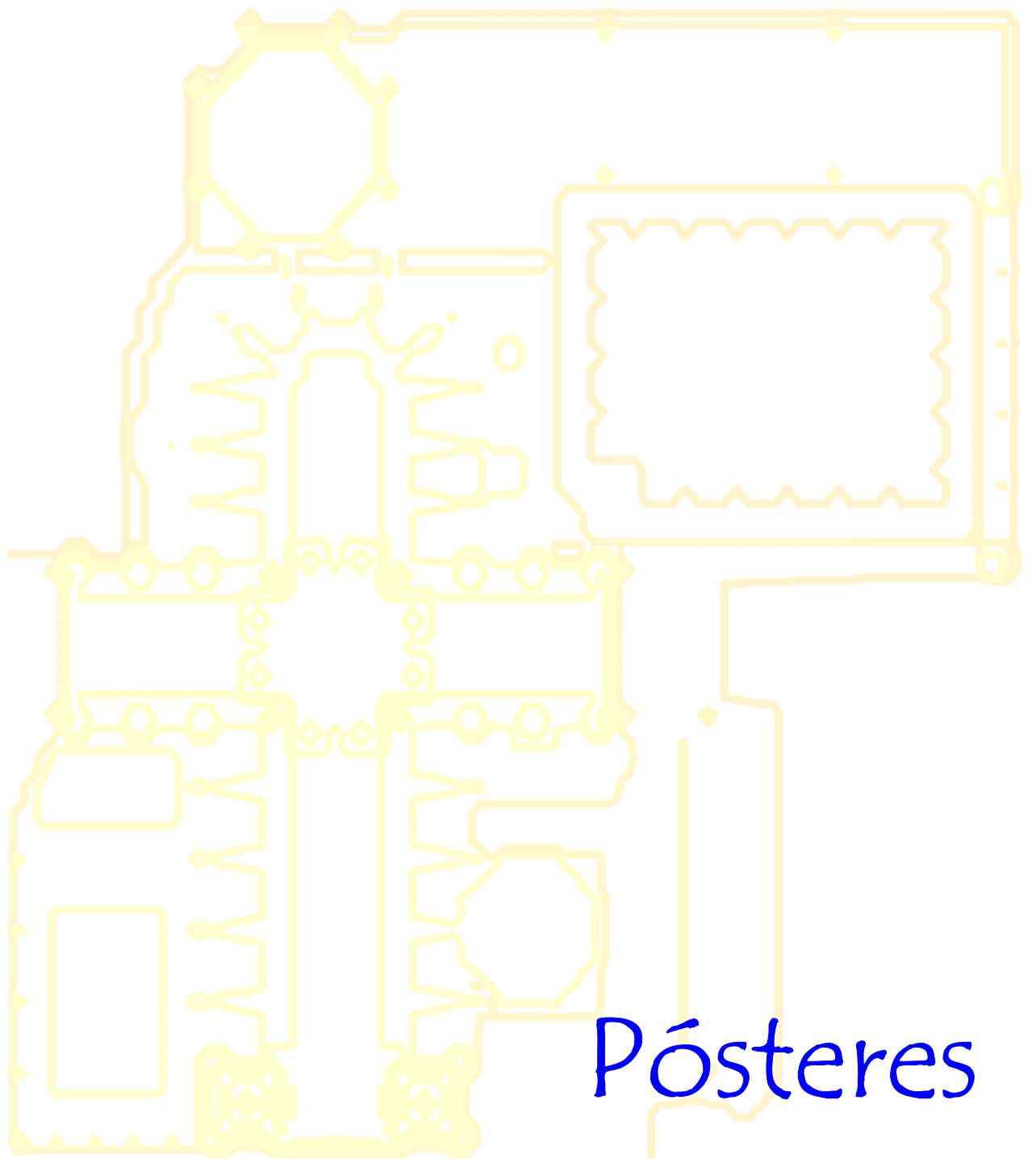
Realizando el análisis de las componentes principales (PCA) sobre las 129 primeras muestras, tomando en consideración las 17 variables fisicoquímicas, se pone de manifiesto que la primera componente principal es la que permite diferenciar las muestras "suaves" de las "intensas".

Se realiza un análisis discriminante de las 279 muestras, utilizando los 6 parámetros fisicoquímicos, ya que todas ellas tienen tabulado el valor de dichas variables. La función de discriminación resultante es capaz de separar correctamente 263 muestras (94,3%) y sólo comete 16 errores (5,7%).

Una mayor discusión sobre las variables significativas y los modelos obtenidos serán presentadas en la comunicación

## **CONCLUSIONES**

El término "suave/intenso" del aceite de oliva está directamente relacionado con algunas variables fisicoquímicas. Utilizando una función de discriminación definida a partir de dichas variables, es posible (con una fiabilidad del 95%): i) o bien asignar un valor de dicho término a un aceite desconocido de las mismas características que las que componen la muestra en estudio; ii) o bien verificar si la declaración del producto se corresponde con las características reales del aceite.



Pósteres



# POSSIBLE OPTIMIZATION OF DRINKING-WATER DISINFECTION PROCESSES FOR CONTROLLING THE FORMATION OF DISINFECTION BY-PRODUCTS BY LINEAR AND NON-LINEAR MODELLING

Stefan Platikanov <sup>a</sup>, Jordi Martín <sup>b</sup>, Romà Tauler <sup>a</sup>

<sup>a</sup> *Department of Environmental Chemistry, IIQAB-CSIC, Jordi Girona, 18-26,*

*Barcelona 08026, Spain*

<sup>b</sup> *AGBAR, Av. Diagonal 2090-211, 08018 Barcelona, Spain*

Formation and occurrence of trihalomethanes ( $\text{CHCl}_3$ ,  $\text{CHBr}_3$ ,  $\text{CHCl}_2\text{Br}$ , and  $\text{CHBr}_2\text{Cl}$ ) are investigated in water chlorination disinfection processes in the Barcelona's water works plant (WWP). Forty-one WWP variables were measured in the entire plant process and were investigated for correlation with trihalomethane formation. Linear and non-linear regression techniques including Multilinear Regression (MLR), Partial Least Squares Regression (PLSR) and Support vector machine regression have been used and compared to model and predict the complex behavior observed for the measured trihalomethane concentrations. The results, obtained by evaluation of the statistical significance of the coefficients in the linear regression vectors, together with an investigation of the VIP scores using PLSR, revealed that the most important WWP variables for trihalomethane formation were: water temperature, total organic carbon, pre- and post- added chlorine concentrations, UV absorbance and turbidity at different sites of the WWP, as well as other variables like wells supply flow levels and carbon filters age. Overall, SVM performed the best and gave good predictive properties. Best results were obtained for the total sum of trihalomethane concentrations, TTHM, with average modeling and prediction relative errors of 2.5%-12% and 14%-18% respectively. Among the individual trihalomethanes, the concentrations of  $\text{CHBr}_3$  were the worst predicted ones with average modeling and prediction relative errors between 11-21% and 21-29% respectively, followed by  $\text{CHCl}_2\text{Br}$  with 6-20% and 17-25%. Better predictions were obtained for the concentrations of  $\text{CHBr}_2\text{Cl}$  with relative modeling and prediction errors varying between 7-14% and 16-21%, and for the concentrations of  $\text{CHCl}_3$  with 9-19% and 17-21% errors respectively.



## ***MIA CODE. SOFTWARE PARA ANÁLISIS MULTIVARIANTE DE IMÁGENES***

José Manuel Prats-Montalbán, Alberto Ferrer

*Applied Statistics Department, Technical University of Valencia*

*Cno. de Vera s/n, Edificio I3, 46022 Valencia, Spain*

El análisis de imágenes se engloba dentro de un área de conocimiento más general conocida como mejora y tratamiento de imagen. Básicamente, trata por una parte de optimizar el aspecto visual y/o la definición de los objetos existentes en una imagen, así como la compresión de dichas imágenes maximizando la calidad de las mismas; y por otro lado trata el análisis (y síntesis) de las imágenes, con el fin de establecer la adecuación de los objetos a unos estándares determinados, detectar fallos de posición o defectos en piezas, o clasificar piezas atendiendo a características diversas de calidad.

Estas tareas pueden llevarse a cabo trabajando sobre el espacio de la imagen (comparación entre píxeles) o bien transformando las imágenes en un vector de características definidas *a priori* que resumen la información existente en las mismas.

Tradicionalmente, estas técnicas han estado ligadas al análisis de imágenes en escala de grises. Con la aparición de las cámaras RGB, el número de canales de color a analizar aumentó, y la correlación entre dichos canales aumentó la complejidad de estos análisis.

Es en este contexto donde surge, hace ya 20 años, el análisis multivariante de imágenes, o *Multivariate Image Analysis (MIA)*, con el fin primero de comprimir de manera eficiente la información existente en las imágenes en color, y detectar los fenómenos de interés (mediante el análisis de agrupaciones de píxeles) existentes en las mismas.

Estas técnicas se basan en la aplicación de modelos multivariantes (básicamente PCA y PLS) a estructuras de datos derivados de imágenes, ya sean en escala de grises, RGB, multi- o, últimamente, hiper-espectrales (si bien el análisis de este tipo de imágenes, muy ligadas a composiciones químicas, debería posiblemente analizarse por medio de otros modelos quimiométricos, con el fin de un análisis posterior mediante técnicas *MIA*).

A pesar de la relativa antigüedad del *MIA*, el *software* desarrollado asociado a estas técnicas ha sido muy limitado (funciones básicas en el PLS\_Toolbox, MACC software y el creado por Jim Burger y Paul Geladi). El objetivo de este póster es ilustrar las posibilidades de análisis que se pueden llevar a cabo a partir del código de libre disposición desarrollado por el Grupo de Ingeniería Estadística Multivariante (GIEM).

Las funciones básicas van ligadas al análisis exploratorio de la imagen, con el objeto de detectar valores extremos en la misma o alejados del modelo creado (Imágenes de  $T^2$  y RSS) o bien segmentar la imagen en función de la cantidad de objetos existentes en la misma.

El código también permite llevar a cabo labores de clasificación de píxeles dentro de la imagen a partir de modelos PCA o PLS-DA, o una combinación de ambos, con el fin de cuantificar y localizar los fenómenos de interés (tales como enfermedades en cítricos, defectos en superficies, bacterias, etc.) existentes en la imagen.

Una de las principales finalidades del código inicialmente desarrollado fue servir de soporte para llevar a cabo tareas de monitorización estadística multivariante por medio del análisis de imágenes. Este hecho ha posibilitado el desarrollo de técnicas que permiten establecer cuándo una imagen (o la pieza existente en ella) muestra algún tipo de desviación respecto del comportamiento general del modelo creado a partir de imágenes consideradas como “típicas” del proceso bajo control; así como detectar en la imagen las zonas que muestran dichas desviaciones, tales como manchas, rayas, etc.

A partir de este código inicial se han ido implementando técnicas de extracción de características propias del campo de análisis de imágenes y del reconocimiento de patrones, por medio de las cuales una imagen pasa a convertirse en un vector de características que son utilizadas como variables del proceso, y que son analizadas por medio de modelos PCA y PLS en un contexto tradicional de *MSPC (Multivariate Statistical Process Control)*.



## DETERMINATION OF FOOD ADULTERATION BY CLASSIFICATION TECHNIQUES

<sup>1</sup>Carolina Vanesa D i Anibal, <sup>2</sup>M.Pilar Callao and <sup>3</sup>Itziar Ruisánchez.

Departament de Química Analítica i Química Orgànica. Universitat Rovira i Virgili. Campus Sescelades, C/ Marcel·lí Domingo, s/n 43007 Tarragona, SPAIN

<sup>1</sup>carolinadianibal@hotmail.com; <sup>2</sup> mariapilar.callao@urv.cat ; <sup>3</sup>itziar.ruisanchez@urv.cat

Sudan dyes (I to IV) are red dyes that are used for colouring fatty and oily materials such as solvents, oils, waxes, petrol, and shoe and floor polishes. They have also been discovered (2003) in some imported chilli powder from India and most recently finding their way into relishes, chutneys and seasonings containing chilli powder.

Sudan dyes are categorized as Class 3 carcinogens by the International Agency for Research on Cancer (IARC) and, therefore, are illegal as food additives according to both the FDA and the European Union. Cargoes of dried and crushed or ground chilli and curry powders coming into any country in the EU must have a certificate to shown they have been tested and found to be free of Sudan dyes (1,2).

The four Sudan dyes are shown in the figure 1:

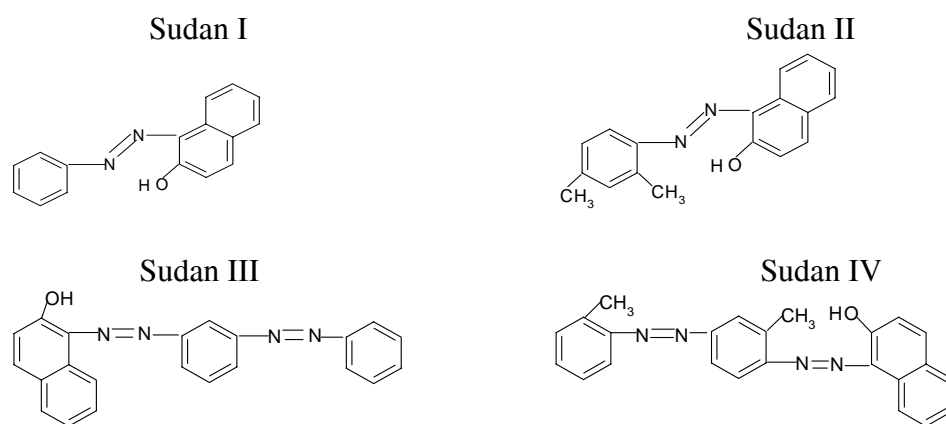


Figure 1: Chemical Structure of Sudan I, II, III and IV

The Sudan dyes are usually analysed using Liquid Chromatography with a diode array detector and confirmed using Liquid Chromatography / Mass Spectrometry (LC/MS) methods which provide high selectivity and sensitivity (3).

A rapid methodology for detecting Sudan dyes in species is presented. It uses spectroscopic measurements and classification techniques as SIMCA and KNN. We have found no references to classification techniques being used in screening problems to determine whether dyes are present or not in food.

Among the several spectroscopic techniques available, UV-Visible, Nuclear Magnetic Resonance (NMR), Infrared (IR), Fluorescence and Raman may be used to detect the mentioned dyes. UV-visible is the technique most commonly used to analyse dyes in food and industrial products. One of the goals of the present work is to characterize the samples using another spectroscopic technique in an attempt to obtain information that complements the information already obtained by UV-Vis, thus improving the screening process. To improve the classification results, data from several spectroscopic sources will be fused (4).

The algorithms are written in MATLAB code and tested with experimental data from conventional HPLC-DAD instruments.

We work with 50 samples distributed as follows: 25 samples without dye and 25 samples spiked with Sudan I. The group of samples is constituted by sweet and spicy paprika.

## References

- (1) Commission Decision of 20 June 2003 on emergency measures regarding hot chilli and hot chilli products, notified under document number C(2003) 1970, (2003/460/EC), OJ L. 154/114, 21.6.2003.
- (2) Implementation of Commission Decision 2003/460/EC of 21 January 2004.
- (3) Limin H., Yijuan S., Binghu F., Xiangguang S., Zhenling Z., Yahong L., *Analytica Chimica Acta*, 594 (2007) 139-146.
- (4) D.L. Hall, J. Llinas, *Proceedings of the IEEE*, 85 (1997) 6-23.

**OPTIMIZATION OF MULTI-RESPONSE ANALYTICAL PROCEDURES USING PARETO-  
OPTIMAL FRONT METHODOLOGY.  
APPLICATION TO THE DETERMINATION OF VETERINARY DRUGS BY FIA SYSTEM.**

C. Reguera<sup>1</sup>, M. S. Sánchez<sup>2</sup>, M. C. Ortiz<sup>1</sup> and L. A. Sarabia<sup>2</sup>

<sup>1</sup>Department of Chemistry (Analytical Chemistry)

<sup>2</sup>Department of Mathematics and Computation

Faculty of Sciences, University of Burgos

Pza. Misael Bañuelos s/n, 09001 Burgos, Spain

**Abstract**

The response surface methodology in the optimization of analytical procedures provides efficient experimental designs to separately model each response. However, in many occasions, these procedures are multi-response either because several analytes are simultaneously determined or because it is desired to have great signals with small variability, that is, more sensitivity and better repeatability are looked for. In these cases the experimental conditions that correspond to the optimum of each response in the experimental domain are rather different so that it is necessary to find a compromise solution (unique experimental conditions).

A methodological approach for the optimization of several responses, each one adequately described by the corresponding validated model, consists of determining the set of experimental conditions that leads to optimal values in at least one of responses, is the Pareto-optimal front [1,2,3]. The exploration of this region of the experimental domain (the experimental conditions that constitute the Pareto-optimal front) allows the analyst to consider advantages and disadvantages related to the experimental factors and to maintain its criteria on which considers a sufficiently valid response, and further it does it "a posteriori", that is to say, after having obtained the possible optimal solutions to the problem. In addition, once we have seen the achievable expected values of the responses in the experimental domain, we can explore the interactions among the experimental factors or search for the experimental conditions where the analytical procedure is robust.

To show the methodology, two real problems of determination of veterinary drugs by FIA system with amperometric detection have been analysed. Its relevance is in the framework of the European regulations [4,5].

The first one is about the determination of sulfadiazine (SD) and the aim of the optimization is to maximize the peak height (analytical signal) and to minimize its variability by analyzing three experimental factors (flow rate, percentage of methanol and pH of the carrier solution). The

exploration of the Pareto-optimal solutions shows that the simultaneous increase of the three experimental factors produce the increase of both responses. However, to constant flow rate, the peak height and pH increase in almost an identical way and the standard deviation behaves like the percentage of methanol. Therefore, to obtain high signals, the flow rate should be increased but for high flow rates the percentage of methanol and pH should be carefully controlled to avoid an excessive increase of the standard deviation.

The second case tackles the determination of malachite green (MG) and leucomalachite green (LMG) when we need to simultaneously optimize four responses (peak height and standard deviation for both analytes) by analyzing two experimental factors (flow rate and percentage of acetonitrile of the carrier solution). What we can say from the analysis of the Pareto-optimal solutions is that the simultaneous increase of the percentage of acetonitrile and flow rate makes the current of MG increase and the current of LMG decrease and that the flow rate itself when the percentage of acetonitrile is constant is the responsible of the increase of both peak heights.

### **Acknowledgments**

This work has been partially supported by Spanish Ministerio de Ciencia e Innovación (under project number CTQ2008-02264/BQU) and regional Consejería de Educación de la Junta de Castilla y León under project BU024A07, both co-financed with European FEDER funds.

### **References**

- [1] M.C. Ortiz, L.A. Sarabia, A. Herrero, M.S. Sánchez, *Chemom. Intell. Lab. Syst.* 83 (2006) 157.
- [2] R. Díez, L.A. Sarabia, M.S. Sánchez, M.C. Ortiz, *Chemom. Intell. Lab. Syst.* 92 (2008) 71.
- [3] C. Reguera, M.S. Sánchez, M.C. Ortiz, L.A. Sarabia, *Anal. Chim. Acta* 624 (2008) 210–222.
- [4] European Commission, Council Regulation 2377/90/EC of 26 June 1990 laying down a Community procedure for the establishment of maximum residue limits of veterinary medicinal products in foodstuffs of animal origin, *Off. J. Eur. Union*, L224 (18 August 1990).
- [5] Commission Decision (EC) No 2004/25/EC of 22 December 2003 amending Decision 2002/657/EC as regards the setting of minimum required performance limits (MRPLs) for certain residues in food of animal origin, *Official Journal L* 6, 10/1/2004, pp. 38-39.

## **APLICACIÓN DE COMPOSICIÓN ESTRATÉGICA DE MUESTRAS PARA EL CONTROL DE COLORANTES PROHIBIDOS EN MATERIALES TEXTILES**

**J. García-Lavandeira, E. Blanco, M.C. Salgado, I. Rodríguez, M. Llompart, R. Cela**

Instituto de Investigación y Análisis Alimentarios. Dto. Química Analítica, Nutrición y Bromatología. Universidad de Santiago de Compostela.

La sociedad de la globalización en la que nos encontramos en el momento actual, ha dado una vital importancia en la industria a la búsqueda del máximo beneficio como principal objetivo empresarial, convirtiendo en un factor dominante a la reducción de costes en todos los campos de los procesos de manufactura de cualquier producto, incluyendo los procedimientos de control de riesgos derivados del empleo de sustancias prohibidas, tóxicas, o de productos químicos potencialmente alergénicos, que si bien están controlados legalmente en la mayoría de los países desarrollados, en aquellos donde la mano de obra es más barata, y por tanto donde se desarrolla la mayoría de estos procesos de manufactura, no existe este tipo de control; por lo que consecuentemente, estos procedimientos analíticos de control adquieren si cabe una mayor importancia para poder ofrecer a los consumidores productos inocuos.

El modo de conseguir reducir costes en los procesos analíticos pasa por, además de intentar conseguir métodos analíticos de screening más rápidos y robustos, ser capaces de reducir el número de muestras a controlar, hecho que podremos lograr aplicando la composición de muestras.

La composición de muestras convencional es útil cuando se busca una respuesta de tipo cualitativa (SI/NO), los límites de detección de los métodos analíticos empleados son lo suficientemente bajos como para conseguir un riesgo bajo de falsos negativos por dilución, y el número de especímenes positivos esperados es realmente bajo en relación al total, puesto que la aparición de un positivo implica el análisis individual de cada una de las especies.

Una alternativa práctica es la composición estratégica de muestras, procedimiento desarrollado por R. Cela y colaboradores, que permite evaluar aproximadamente la concentración de analitos en los especímenes que forman la muestra compuesta sin necesidad de analizarlos individualmente. Para ello se emplean matrices supersaturadas que definen la estructura del experimento, en términos del número de muestras compuestas a analizar y la posición de cada una de las muestras individuales en cada una de las muestras compuestas preparadas. Una de estas muestras compuestas es equivalente a una muestra compuesta convencional (mezcla de todos los especímenes individuales en iguales o diferentes proporciones) y se analiza en primer lugar. Si el resultado es negativo es posible afirmar que ninguno de los especímenes originales era positivo y por tanto, el proceso concluye con el ahorro de tiempo y trabajo consiguiente. Por el contrario, cuando el resultado de esta muestra compuesta es positivo, al menos uno de los especímenes originales era a su vez positivo, pero puede que dicho número sea mayor. Puesto que el sistema de ecuaciones es sobresaturado, se aplica un sistema de regresión que hace uso de algoritmos genéticos para resolver entonces la

matriz de diseño, de modo que se obtiene una estimación semicuantitativa de los contenidos en los analitos de interés para cada uno de los especímenes originales, sin que sea necesario re-analizar dichos especímenes.

Esta técnica presenta diversas ventajas adicionales, especialmente desde el punto de vista de logística y organización de los ensayos rutinarios de laboratorio cuando se trabaja en campañas con gran número de muestras y cuando se sospecha que la probabilidad de muestras positivas es realmente baja (principio Pareto). Adicionalmente, es posible automatizar todo el proceso, tanto en la preparación de las muestras compuestas como en el cálculo de los resultados de modo que resulte una aproximación eficiente y simple para el operador. De hecho, se han desarrollado las herramientas necesarias para ello en nuestro laboratorio.

En este caso particular, se presentan los resultados para el control de colorantes prohibidos en materiales textiles que por reducción pueden dar lugar a aminas cancerígenas. Un problema que requiere control de producto finalizado, en tiempos muy reducidos y en grandes partidas de producción y que implica procedimientos analíticos lentos y costosos, que justifican plenamente recurrir a técnicas de composición de muestras.

## **Referencias**

- Cela R., Martínez E., Carro A.M., **Chemometr. Intell. Lab. 52** 2000 167-182, and **Chemometr. Intell. Lab. 57**, 2001, 75-92
- E. Martínez, P. Landin, A.M. Carro, M.P. Llompart, R. Cela, **J. Environ. Monit. 4** 2002 490-497.
- E. Martínez, R. Cela, A.M. Carro, J.C. Cobas, B. García., **J. Anal. Atom. Spectros. 17**, 2002. 1373-1380
- E. Martinez, M. Pazos, A.M. Carro, M. Llompart, R. Cela., **J. of AOAC Int. 86** 2003. 846-855
- J.B. Quintana, E. Martinez, A.M. Carro, R.A. Lorenzo, R. Cela. **Int.J. Environ. Anal. Chem. 83** 2003. 269-284
- R.Rodil, E. Martinez, A.M. Carro, R.A. Lorenzo, R. Cela, **LC-GC North America. 22** 2004. 272-286.
- J.Carpinteiro, J.B. Quintana, E. Martinez, I. Rodríguez, A.M. Carro, R.A. Lorenzo, R. Cela., **Anal. Chim. Acta. 524**, 2004. 63-71
- L. Pensado, E. Blanco, M.C. Casais, M.C. Mejuto, E. Martinez, A.M. Carro, R. Cela, **Jour. Chromatogr. A. 1054** 2004. ,121-130

# **MONITORIZACIÓN DE LA FERMENTACIÓN ACÉTICA DE ALCOHOL DE CEBOLLA MEDIANTE ESPECTROSCOPIA NIR: VALORIZACIÓN DE SUBPRODUCTOS DE CEBOLLA COMO MATERIA PRIMA PARA LA PRODUCCIÓN DE VINAGRE**

J.M. González-Sáiz, I. Esteban-Díez and C. Pizarro

*Departamento de Química. Universidad de La Rioja, C/Madre de Dios 51, 26006 Logroño  
(La Rioja)*

Los residuos y subproductos de procesamiento en la industria de la cebolla plantean un problema de eliminación y medioambiental cada vez más grave, al mismo tiempo que representan una pérdida de valiosas fuentes de nutrientes. El presente trabajo forma parte de un ambicioso proyecto, llevado a cabo en cooperación con otros grupos de investigación, destinado a investigar la viabilidad de nuevas alternativas para la conversión de residuos de cebolla en productos de valor añadido. En particular, este estudio se centra en la producción de vinagre a partir de desechos de cebolla como una posible ruta de valorización que podría proporcionar una solución viable a numerosos problemas medioambientales y de eliminación, ofreciendo simultáneamente la posibilidad de convertir residuos en un producto alimenticio de calidad y de explotar las propiedades únicas y los beneficios para la salud de las cebollas.

El primer paso esencial en la producción de vinagre de cebolla, es decir, la producción efectiva de etanol a partir de residuos de cebolla y la necesaria monitorización de la fermentación alcohólica de zumo de cebolla para poder controlar dicho proceso en tiempo real, ya ha sido estudiada y descrita en detalle en un estudio anterior recientemente publicado [J.M. González et al., (2007) J Agric Food Chem 55:2930-2936]. El actual estudio es la continuación natural del mencionado trabajo, ya que se centra en la valorización de residuos de cebolla para promover su transformación en productos de valor añadido (tales como el vinagre de cebolla) y, más específicamente, en el segundo y definitivo paso del sistema de fermentación analizado: la producción eficaz de vinagre de cebolla a partir de alcohol de cebolla (previamente obtenido por fermentación alcohólica de zumo de cebolla) mediante fermentación acética. Con el fin de encontrar una metodología adecuada para monitorizar este proceso en tiempo real, se construyeron modelos de regresión PLS separados en base a los espectros NIR de muestras tomadas durante la fermentación acética para modelar los parámetros clave implicados en la fermentación acética del alcohol de cebolla. Asimismo, un

método de selección de longitudes de onda basado en la regresión PLS como es el ‘pesado iterativo de predictores’ (*iterative predictor weighting* (IPW)) fue aplicado para testar la hipótesis de que un mínimo número (máxima parsimonia) de predictores informativos (bandas NIR de absorción íntimamente relacionadas con la respuesta de interés) podían predecir una determinada variable respuesta con al menos la misma exactitud, o incluso con fiabilidad y robustez mejoradas, que al emplear el rango completo de longitudes de onda NIR. Las concentraciones de biomasa, sustrato (etanol) y producto (ácido acético) de la fermentación acética de alcohol de cebolla fueron predichas con gran exactitud empleando los modelos IPW-PLS desarrollados, con una raíz del error cuadrático medio de los residuos en predicción externa (RMSEP) menor del 2.5% tanto para etanol como para ácido acético, y un RMSEP del 6.1% en el caso de la concentración total de biomasa (un resultado sumamente satisfactorio considerando la baja precisión y exactitud asociadas con el método de referencia empleado en su determinación). El uso potencial como herramienta rutinaria de monitorización de los simples y fiables modelos de calibración propuestos (desarrollados a partir de un número mínimo de predictores significativos) podría reducir sustancialmente el tiempo de análisis, los costes y el trabajo requeridos en la determinación de los parámetros críticos de fermentación acética de alcohol de cebolla. De esta forma, una determinación no destructiva, certera y prácticamente en tiempo real, fundamental para el control y optimización global del proceso, sería posible.



# VALIDACIÓN CRUZADA PARA LA SELECCIÓN DEL DESPLIEGUE Y NÚMERO DE SUB-MODELOS EN EL MODELADO PCA DE PROCESOS POR LOTES.

José Camacho, Jesús Picó, Alberto Ferrer

## Resumen:

Existe un amplio número de propuestas de modelado de procesos por lotes basadas en el Análisis por Componentes Principales (PCA). En todos los casos, las propuestas siguen un conjunto de pasos establecidos a priori. Las características del modelo dependerán de estos pasos y se ajustarán en mayor o menor medida a los datos dependiendo de la naturaleza específica del proceso: ¿presenta el proceso dinámicas lentas, rápidas o combinación de ambas? ¿son las relaciones dinámicas constantes o sufren cambios a lo largo del proceso?.

Una de las decisiones críticas del modelado con PCA es cómo transformar la matriz tridimensional de datos de un proceso por lotes en dos dimensiones: desde el despliegue en la dirección de los lotes hasta el despliegue en la dirección de las variables y desde el uso de un único modelo hasta el uso de múltiples modelos. Distintas soluciones se ajustan mejor a distintas dinámicas. Por ejemplo, el uso de varios modelos es conveniente para capturar dinámica variable en el tiempo. Por tanto, parece necesaria la definición de un procedimiento que a partir de los datos de un proceso nos permita elegir la transformación de 3-D a 2-D óptima. La propuesta presentada en este póster está basada en el algoritmo de validación cruzada propuesto en Camacho et al. (2007). El algoritmo permite calcular la suma de cuadrados de error de predicción (prediction error sum of squares o PRESS) asociada a un modelo. Así, de igual manera que el PRESS puede ser utilizado para decidir el número de componentes en PCA, en la propuesta presentada permite además elegir el despliegue y número de sub-modelos en el modelado de un proceso por lotes. Los datos de varios procesos por lotes de distinta naturaleza (polimerización, fermentación y tratamiento de aguas) son utilizados para ilustrar la propuesta.

Camacho, J., Picó, J., Ferrer, A.J. New Cross-Validation Methods in Principal Component Analysis. 10th Scandinavian Symposium on Chemometrics, 2007.



# RESOLUTION METHODS APPLIED TO THE ANALYSIS OF SPECTROSCOPICAL DATA FROM BIOANALYTICAL PROCESSES

Joaquim Jaumot and Raimundo Gargallo

Chemometrics Group. Analytical Chemistry Dept. Universitat de Barcelona.  
Av. Diagonal 647, E-08028 Barcelona, Spain

Multiequilibria reactions and interactions between different conformations of biomolecules like nucleic acids can be studied using several spectroscopic methods like UV molecular absorption, circular dichroism, fluorescence and NMR. Investigations are performed in experiments where temperature, pH or salt concentrations are changed and where the process evolution is monitored with time.

In order to extract chemical information from these spectroscopic data, different hard- and/or soft- modelling methods have been demonstrated to be very effective and powerful tools. Hard modelling methods<sup>[1-2]</sup> are based on a proposed chemical model defined by: (1) the stoichiometries of the species present, (2) on the fulfillment of the mass-action law and (3) approximate values of their equilibrium constants. This procedure provides excellent results when applied to the study of known chemical equilibria with a detailed model. If the number of species and/or the complexity of the system increase, the selection of a correct model is more difficult and the use of soft-modeling methods<sup>[3]</sup>, like MCR-ALS<sup>[4-5]</sup>, may be more appropriate. Multivariate Curve Resolution allows the resolution of species profiles (concentrations, spectra) in complex mixture systems. Algorithms based on Alternating Least Squares accommodate easily the application of different type of constraints for an optimal resolution of spectra and concentration profiles and/or allow also easily the analysis of multiple experiments to overcome rank deficiency, rotational and intensity ambiguities ubiquitously present in Factor Analysis based methods.

In this work, some examples of application of Multivariate Curve Resolution are shown in the study of complex multiequilibria biomolecular processes. Selected examples showed the analysis of conformational equilibria of a Triplex Forming Oligonucleotides by UV molecular absorption and temperature dependent transitions of G-quadruplexes by UV molecular absorption, fluorescence emission and circular dichroism.

- [1] Maeder M, Zuberbuhler AD. *Analytical Chemistry* **1990**; 62:2220-2224.
- [2] Dyson RM, Kaderli S, Lawrance GA, Maeder M, Zunderbuhler AD. *Analytica Chimica Acta* **1997**; 353:381-393.
- [3] Massart DL, Buydens LMC, Vandegiste BGM. *Handbook of Chemometrics and Qualimetrics* (1st edn. ed), Data Handling in Science and Technology. 20. **1997**; Elsevier: Amsterdam, The Netherlands.
- [4] Tauler R, Smilde A, Kowalski B. *Journal of Chemometrics* **1995**; 9:31-58.
- [5] Tauler R. *Chemometrics and Intelligent Laboratory Systems* **1995**; 30:133-146.



# APLICACIÓN DE MAPAS AUTO-ORGANIZATIVOS DE KOHONEN EN EL ESTUDIO DE CONTAMINACIÓN DE SUELOS POR METALES

**M.P. Gómez-Carracedo<sup>(1)</sup>, J. Aires de Sousa<sup>(2)</sup>,  
J.M. Andrade<sup>(1)</sup>, A. Carlosena<sup>(1)</sup>, D. Prada<sup>(1)</sup>**

(1) Dept. Química Analítica, Universidade da Coruña, Campus da Zapateira s/n, E-15071, A Coruña, España

(2) Dept. Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal.

En este trabajo se estudia la aplicación de redes neuronales de Kohonen (o mapas auto-organizativos de Kohonen, SOM) como una herramienta de reconocimiento de pautas en el estudio de contaminación por metales pesados en muestras de suelo urbano. El período de muestreo comprende una campaña de muestreo por cada estación (primavera, verano, otoño e invierno), en un año. En cada campaña se toman 93 muestras en el entorno metropolitano de la ciudad de A Coruña (Galicia, España). En cada muestra se estudian 12 variables analíticas: 9 metales pesados y 3 parámetros físico-químicos (pérdida por ignición, pH y humedad). Los metales se cuantificaron mediante espectrometría de absorción atómica: Cu, Fe, Mn, Pb y Zn (empleando llama) y Cd, Co, Cr y Ni (empleando horno de grafito).

El primer objetivo del trabajo es determinar si los SOM aportan información cuya interpretación sea comparable a otras técnicas clásicas, tales como PCA. *A priori*, las muestras podrían clasificarse en cuatro modalidades, dependiendo de su procedencia: de autopista (tomadas al borde de la autopista AP9 en el tramo Coruña-Santiago de Compostela), de transectos de autopista (tomadas perpendicularmente a la autopista), de jardines (urbanos) y de avenida (tomadas en el entorno de una avenida con gran circulación rodada). El segundo objetivo del trabajo es determinar si esta división subjetiva se corresponde a la situación real de contaminación de los suelos. La topología de los SOM se ha optimizado para cada una de las estaciones del año empleando el programa JATOON (Java tools for neural networks). Se ha observado una buena clasificación de las muestras para las diferentes estaciones del año empleando topologías sencillas y resultados totalmente análogos a los obtenidos mediante PCA. Las muestras se distribuyeron en el mapa formando grupos muy homogéneos según su procedencia (jardines, avenida, autopista y transectos). Con objeto de simplificar el estudio se ha aplicado una nueva metodología SOM, denominada MolMap, que permite abordar simultáneamente todos los muestreos, de forma similar a otras herramientas para el análisis de datos de tres vías.

## AGRADECIMIENTOS

MPGC agradece a la Xunta de Galicia (Dirección Xeral de Investigación, Desenvolvemento e Innovación de la Consellería de Innovación e Industria) la ayuda financiera concedida en forma de Contrato de Investigación, Programa Angeles Alvaríño.



## **EVALUACIÓN DE MODELOS DE VARIACIÓN DE ÓXIDO DE NITRÓGENO DURANTE EL PERIODO 2001-2006 EN CATALUÑA UTILIZANDO MÉTODOS DE ANÁLISIS MULTIVARIANTE.**

M. Alier, M. Felipe-Sotelo, I. Hernández, R. Tauler  
Dept. of Environmental Chemistry. IIQAB-CSIC. Barcelona

Los estudios de contaminación atmosférica están adquiriendo actualmente importancia debido a sus interacciones con el clima y con los seres humanos. Es por esta razón por la que se están aumentando las redes monitorizadas de control de contaminantes atmosféricos. La mejor manera de analizar estos datos es a partir del análisis multivariante ya que nos permite trabajar con matrices de datos muy grandes, no requiere excesivos tiempos de computación y nos permite alcanzar conclusiones útiles.

En este trabajo se han usado herramientas quimiométricas como PCA (Principal Component Analysis) y MCR-ALS (Multivariate Curve Resolution Alternating Least Squares) para estudiar las variaciones geográficas y temporales del óxido de nitrógeno en diferentes estaciones de muestreo en Cataluña durante los años 2001-2006. Se dispone de medidas semihorarias de óxido de nitrógeno junto con otros contaminantes atmosféricos y parámetros ambientales en un total de 50 estaciones de medida.

En primer lugar se ha aplicado PCA (utilizando el autoescalado como pretratamiento) para tener una visión general de la varianza explicada, y en segundo lugar se ha aplicado MCR-ALS (sin pretratamiento) aplicando diferentes restricciones tales como la no-negatividad y la trilinealidad. La restricción de no negatividad nos permite hacer los resultados más comprensibles en términos medioambientales y la de trilinealidad nos permite ver si el comportamiento de las diferentes estaciones es constante a lo largo de los diferentes años.

Se han realizado diferentes análisis con las matrices de datos experimentales obtenidas en los diferentes años y en diferentes estaciones de muestreo con tipos de contaminación diversa. A partir del análisis de estos datos experimentales, se han estimado las variaciones diarias y anuales de las diferentes pautas de contaminación por NO, las cuales nos han permitido ver la importancia de la contaminación urbana e industrial en las diferentes localizaciones así como su interacción con otros contaminantes atmosféricos como el ozono.





# ESTRUCTURA DE COVARIANZA DINÁMICA EN EL MODELADO PCA DE PROCESOS POR LOTES

José Camacho, Jesús Picó, Alberto Ferrer

## Resumen:

La literatura contempla un amplio número de propuestas de modelado de procesos por lotes basadas en el Análisis por Componentes Principales (PCA). En muchos casos, la propuesta de un nuevo método se justifica con un buen funcionamiento para una aplicación concreta, como monitorización estadística, y para un conjunto específico de procesos. Sin embargo, es difícil encontrar razonamientos sólidos que apoyen ese buen funcionamiento desde un punto de vista general. Muchos autores están de acuerdo en que el modelado de procesos por lotes basado en PCA no es claramente entendido. Por tanto, se hace necesario ahondar en las propiedades teóricas que permitan identificar las diferencias entre distintas propuestas.

¿Por qué el modelado basado en el despliegue en la dirección de los lotes (batch-wise) permite capturar dinámica variante en el tiempo? ¿Qué consecuencias tiene el despliegue en la dirección de las variables (variable-wise)? ¿Existe un método de despliegue óptimo? ¿Cuándo el modelado de un proceso debe llevarse a cabo con varios modelos PCA? El objetivo de este póster es dar respuesta a éstas y otras preguntas relacionadas a partir del análisis de la matriz de varianzas-covarianzas utilizada en PCA. Distintas propuestas de modelado son analizadas, incluyendo modelos batch-wise, variable-wise, batch dynamic, modelos locales y envolventes y modelos multi-fase. La estructura de la matriz de varianzas-covarianzas permite analizar la conveniencia de una u otra propuesta a partir de la naturaleza propia de los datos del proceso y teniendo en cuenta la parsimonia de los modelos. Este análisis permite clarificar cómo los modelos PCA capturan la dinámica del proceso, qué porción de los datos es representada por qué parámetros del modelo y qué implicaciones conlleva esta asignación.



# FRONT FACE FLUORESCENCE AND PLS CALIBRATION TO STUDY THE EFFECT OF THERMICAL TREATMENT IN MILK QUALITY

R. Diez<sup>a</sup>, M.C. Ortiz<sup>a</sup>, L. Sarabia<sup>b</sup>, I. Birlouez-Aragon<sup>c</sup>

(<sup>a</sup>) Dpt. of Chemistry. (<sup>b</sup>) Dpt. of Mathematics and Computation

Faculty of Sciences. Pza. Misael Bañuelos s/n, 09001 Burgos, Spain.

(<sup>c</sup>) Institut Supérieur d'Agriculture de Beauvais, Rue Pierre Waguet - BP 30313

F-60026 Beauvais Cedex, France.

## 1 Introduction

Heat treatment of milk is commonly used to ensure microbial safety and to supply milk with longer shelf-life. Commercial sterilization and pasteurization are the most frequent processes in the treatment of milk. There are many technological alternatives that can be used with this purpose, but some of these techniques may cause different degrees of damage of the milk nutritional properties.

Three good parameters of the modifications caused during the heat treatment are: FAST index, concentrations of vitamin C and soluble protein. FAST index is a simple and sensitive indicator of the global advancement of the Maillard reaction, while vitamin C and soluble protein are good indicators of the nutritional quality. Conventional methodologies for determining these parameters are tedious and time-consuming, in addition all these techniques destroy the sample for its analysis. Front face fluorescence spectroscopy presents several advantages against traditional techniques, providing information of the presence of fluorescent molecules in intact samples in a fast and simple way.

This work shows that front face fluorescence spectroscopy together with PLS calibration is an appropriate method for the evaluation of the thermal treatment of milk samples by the quantification of FAST index, vitamin C and soluble protein [1].

## 2 Experimental

A formulation based on pure whey proteins to mimic whey-based infant milk formulas was proposed by MONTAIGU. The formulation consists basically of a mineral buffer, lactowhey Prolacta 90, lactose in powder, a premix of 13 vitamins and a preparation with different oils. Milk samples were incubated for six different times, from 2 to 9.5 minutes into an oil bath heated at seven different temperatures (72°C, 80°C, 87°C, 95°C, 100°C, 110°C and 115°C).

For the chemical analysis, two emission fluorescence intensities were measured in each sample to calculate the FAST index: the fluorescence of Tryptophane ( $F_{\text{Trp}}$ ) at an excitation of 290 nm and an emission of 340 nm and the fluorescence of the advanced Maillard product ( $F_{\text{AMP}}$ ) at an excitation of 330 nm and emission of 420nm. FAST index was calculated as  $(F_{\text{AMP}}/F_{\text{Trp}})*100$ , as is described in references [2, 3].

Vitamin C was analysed by HPLC-fluorescence in the milk samples after stabilization of the sample with metaphosphoric acid 10%, immediately after cooling. The method is described in reference [4]. Quantification of soluble protein has been made by means the calibration of tryptophan fluorescence with PROLACTA microfiltrated whey proteins.

### 2.1 Instrumentation and software

$F_{\text{Trp}}$  and  $F_{\text{AMP}}$  intensities at the maximum were measured in diluted milk samples by conventional fluorescence in a Spex spectrofluorimeter. Spectra of intact milk samples were measured by front face fluorescence in a Varian spectrofluorimeter (model Cary Eclipse). Excitation and emission slits are both set to 5 nm and the scan speed is 600 nm min<sup>-1</sup>. Emission spectra for the tryptophane are recorded between 312 and 398 nm (regular steps of 1 nm) at an

emission wavelength of 290 nm and for the advanced Maillards product, emission wavelength is fixed at 350 nm and they are recorded between 402 and 498 nm.

PLS regressions were computed with the PLS Toolbox [5] for MATLAB 6.5. Regression models "calculated concentration vs. true concentration" were carried out with PROGRESS [6]. The statistical analysis of the data and regression models were done with STATGRAPHICS [7].

### 3 Results and discussion

For quantification of FAST index, spectra of Tryptophane and the advanced Maillard product have been considered as predictor variables and for vitamin C and soluble protein only the spectrum of tryptophan has been taken. In all the cases spectra have been centered.

The logarithm of the different parameters has been taken as response. The response has always been autoscaled.

The number of latent variables has been determined by leave-one-out cross-validation, choosing that number which minimizes the root mean square error in cross-validation (RMSECV).

A PLS model with all the samples has been carried out for each response. In order to evaluate the stability of the prediction, samples have been divided into two sets: a test set with 5 samples and a calibration test with the rest of the samples. Eight models have been carried out by changing the test set.

The construction of each calibration model was made as follows: i) Q residual and  $T^2$  Hotelling were calculated, removing those objects whose values exceeded the corresponding threshold at 99% confidence level. ii) The objects with absolute values of the standardised residuals higher than 2.5 were removed. iii) Steps i) and ii) were repeated until none of the objects could be considered outlier neither in signal nor in response. The absence of outliers is guaranteed by using a least median squares regression between the concentrations calculated with PLS and the true concentrations.

Regarding to the stability of the prediction, in all the models between 70 and 85% of the variance of the response is explained with 99% of the variance of the predictor variables. The mean of the absolute value of the relative errors both in calibration (11% for vitamin C and around 18% for FAST index and soluble protein) and in prediction (around 18%) show also good stability and capability of prediction of the models for the three parameters studied.

### Acknowledgments

The project was performed in the scope of the European ICARE project. The authors thank Spanish Ministerio de Ciencia e Innovación (under project number CTQ2008-02264/BQU) and regional Consejería de Educación de la Junta de Castilla y León under project BU024A07 for financial support partially with FEDER funds. R.Díez also thanks Junta de Castilla y León for her FPU grant.

### References

- [1] R. Díez, M.C. Ortiz, L.A. Sarabia, I. Birlouez-Aragon, *Anal. Chim. Acta* 606 (2008) 151-158.
- [2] Birlouez-Aragon I. et al., *Int Dairy Journal* 8 (1998) 771-777.
- [3] Birlouez-Aragon I., Sabat P., Gouti N., *Int. Dairy Journal* 12 (2002) 59-67.
- [4] Gliguem H., Birlouez-Aragon I., *J. Dairy Sci.* 88 (2005) 891-899.
- [5] Wise BM, Gallagher NB, Bro R, Shaver JM, PLS Toolbox 3.5, Eigenvector Research Inc., Manson, WA.
- [6] Rousseeuw PJ, Leroy AM, *Robust Regression and Outlier Detection*. John Wiley and Sons Inc., Hoboken, New Jersey, 2001.
- [7] STATGRAPHICS Plus for Windows, Ver 5.1, Statistical Graphics Corp., Rockville, Maryland.

# COMPARACIÓN DE DOS MÉTODOS PARA CUANTIFICAR CLONIXINATO DE LISINA EN SOLUCIONES INYECTABLES MEDIANTE ESPECTROMETRÍAS UV-VIS Y NIR-PLS

Erika Alinne Santander García<sup>1</sup>, Guadalupe Alvarez Ávila<sup>1</sup>, José Antonio Garduño Rosas<sup>1</sup>, José Manuel Andrade Garda<sup>2</sup>, **Raquel López Arellano**<sup>1</sup>

1. Laboratorio de Ensayos Farmacéutico (LEDEFAR). Facultad de Estudios Superiores Cuautitlán UNAM
2. Departamento de Química analítica, Facultad de Ciencias, Universidad A Coruña

La industria farmacéutica necesita métodos analíticos rápidos y fiables para controlar eficientemente las materias primas y los productos terminados. La espectrometría NIR resulta una técnica interesante debido a que proporciona información química muy versátil de forma rápida sin destruir la muestra ni precisar tratamiento previo. Otras técnicas clásicas sí la destruyen (p.ej., HPLC) o necesitan tratar la muestra (p.ej. UV-VIS). NIR también compete con UV-VIS en el bajo coste de los análisis. El objetivo del estudio consiste en comparar dos métodos de cuantificación del Clonixinato de lisina en soluciones inyectables; empleando espectrometría UV y NIR, este último combinado con regresión multivariante mediante PLS.

Para la técnica de UV se utilizó un espectrofotómetro Cary 100 y para la técnica de NIR un espectrofotómetro Rapid Content Analyser de FOSS NIRSystem Model 6500 operando con el software Vision 3.1.0.0. Para ambos métodos se optimizaron las condiciones de análisis y se validaron los parámetros de desempeño: linealidad, precisión, exactitud y reproducibilidad. Para el análisis UV se realizó un calibrado univariante clásico mediante 5 disoluciones en el rango de concentración 50 - 150% de clonixinato de lisina, en medio NaOH 0.1N. Se realizaron las lecturas a 283 nm, termostaticando a 25°C.

Para el análisis NIR se optimizó el número de barridos mediante el estudio de la señal/ruido a 1686 y 2302 nm; posteriormente se midieron 15 soluciones patrón en el rango de concentración 60% - 110% de clonixinato de lisina, empleando reflectancia difusa (rango 1100 - 2500 nm) y restando a cada una el blanco de reactivos correspondiente. Los espectros resultantes se suavizan mediante Savitzski-Golay (ventana de 11 puntos) y, a continuación, se someten a diversos pretratamientos: MSC en todo el rango espectral, selección de tramos espectrales y primera derivada (en todo el rango).

El modelo predictivo elegido emplea únicamente los rangos espectrales 1116-1278 y 1672-1838 nm, corregidos a la línea base. El número de factores (variables latentes) seleccionado mediante *leave-one-out-cross validation* fue de cinco. El error cuadrático medio de predicción (RMSEC) fue 0.38 y el error cuadrático medio de predicción (empleando muestras diferentes a las del calibrado), RMSEP fue 0.40; los cuales se consideran excelentes para las necesidades de la industria. Los parámetros de desempeño del método NIR-PLS se compararon con el de UV-VIS. La repetibilidad y reproducibilidad intermedias se compararon por la prueba F-Fisher, seleccionando dos muestras tipo, sin diferencia significativa ( $p > 0.05$ ).

Como conclusión, puede establecerse que el método NIR-PLS constituye una alternativa válida a otros métodos analíticos existentes en la bibliografía para determinar clonixinato de lisina en soluciones inyectables.

**AGRADECIMIENTOS:** RLA y JMAG agradecen el apoyo económico de la Universidad Nacional Autónoma de México (Facultad de Estudios Superiores de Cuautitlán y al grupo de investigación LEDEFAR, respectivamente).



# PHOTODEGRADATION STUDY OF BROMOPHENOLS USING HYBRID HARD AND SOFT MODELLING APPROACH ON FUSED SPECTROSCOPY AND CHROMATOGRAPHIC DATA

**Autores: Sílvia Mas<sup>1,2</sup>, Anna de Juan<sup>2</sup>, Sílvia Lacorte<sup>1</sup> and Romà Tauler<sup>1</sup>**

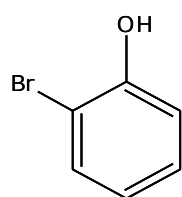
**<sup>1</sup> Dept. of Environmental Chemistry. IIQAB-CSIC. Barcelona**

**<sup>2</sup> Chemometrics group. Universitat de Barcelona**

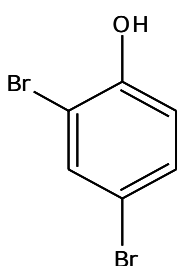
Flame retardants are used to protect the public from accidental fires, by reducing the flammability of combustible materials such as plastics and synthetic polymers. The most important group of flame retardants is the brominated flame retardants (BFRs), which contain a diversity of chemicals such as polybrominated biphenyls (PBBs), tetrabromobisphenol A (TBBPA), polybrominated diphenyl ethers (PBDEs), heabromocyclododecane (HBCD) and bromophenols (BP).

During the past years, concerns have arisen due to the increasing occurrence of these compounds in the environment and to the proven toxic and endocrine disruptor activity of lowly brominated congeners [1]. There is some evidence that highly brominated congeners can degrade by the action of light into lowly brominated congeners in the environment. Hence, the interest in understanding the photodegradation process of BFRs.

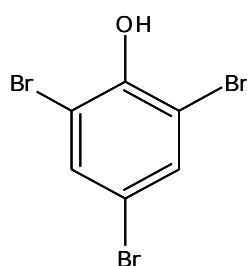
In this work, the photodegradation of four BP congeners with different degree of bromination have been monitored by UV spectroscopy (classical on-line spectroscopic reaction monitoring) and HPLC-DAD (aliquots collected at different reaction times have been injected).



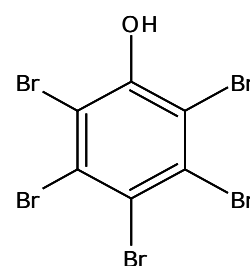
**2-bromophenol**



**2,4-dibromophenol**



**2,4,6-tribromophenol**



**2,3,4,5,6-pentabromophenol**

A recently developed Hybrid Hard- and Soft-modelling Multivariate Curve Resolution-Alternating Least Squares (HS-MCR) [2] approach was applied to the fused spectroscopic-chromatographic data in order to elucidate the mechanism of the degradation process, to resolve kinetic profiles and pure spectra of the photodegradation products and to evaluate the related rate constants. This method is a modification of the iterative soft-modelling MCR-ALS that combines the qualities of deterministic (rate law) hard-modelling and self or soft-modelling methods (no physicochemical law assumption), allowing: a) reduction of the rotational ambiguities associated with the resolved concentration profiles and pure spectra, b) application of kinetic fitting to all or to some of the concentration profiles, while the others are left unconstrained, c) providing of rate constants as additional information, d) application of different individual or global models in a multiexperiment analysis and e) simultaneous analysis of model based and model free experiments in a multiexperiment analysis.

From the preliminary results of this study we can conclude that:

1. An approximate global description of the kinetics linked to the photochemical degradation process of the four bromophenol compounds, with successive steps and related rate constants could, has been provided.
2. The interpretation of the successive steps of the photodegradation experiments has shown how the highly brominated congeners can end up its environmental fate transformed in more lowly brominated congeners (more toxic compounds) and many other by-products of different chemical nature.
3. HS-MCR analysis of fused spectroscopic data (model-based experiments) and chromatographic data (model-free experiments) improved results versus those obtained using only one of the two experiments. Thus, the presence of chromatographic data breaks the rank-deficiency of the spectroscopic data set. On the other hand, the use of spectroscopic data provides a high resolution process time axis and the possibility to fit a kinetic model and obtain the related rate constants.

[1] De Wit, C. A. *Chemosphere*. 2002, 46 (5), 583-624.

[2] De Juan, A.; Maeder, M.; Martínez, M.; Tauler, R. *Chemom. Intell. Lab. Syst.* 2000, 54(2) 123-141.



# STUDY OF PHOTOCATALYTIC DEGRADATION OF TEXTILE DYES IN WASTEWATER USING UV-VIS SPECTROSCOPY AND MULTIVARIATE CURVE RESOLUTION - ALTERNATING LEAST SQUARES (MCR-ALS).

<sup>1</sup>Cristina Fernández; <sup>2</sup>M. Soledad Larrechi; <sup>3</sup>M. Pilar Callao

Department of Analytical and Organic Chemistry, Rovira i Virgili University, Marcel·lí Domingo  
s/n Campus Sescelades, E-43007 Tarragona, Spain

<sup>1</sup>cristina.fernandez@urv.cat; <sup>2</sup>maria-soledad.larrechi@urv.cat; <sup>3</sup>mariapilar.callao@urv.cat

Textile dyes and other industrial dyestuffs constitute one of the largest group of organic compounds that represent an increasing environmental danger<sup>1</sup>. The discharge of highly coloured waste is not only aesthetically displeasing, but it also impedes light penetration, thus upsetting biological processes within stream. In addition, many dyes are toxic to some organisms and may cause direct destruction of the aquatic communities. Hence, removal of dyes from such wastewaters is a major environmental problem and it is necessary because they are visible even at low concentrations.

In this work, the photodegradation process of dyes is considered. Through this process the dyes decompose completely to CO<sub>2</sub> and H<sub>2</sub>O, but some intermediates can be produced during their photodegradation. Nowadays the development of catalysts used and the optimization of the experimental conditions are subject of interest in the field of the catalysis. The kinetic degradation of polluting agents and the yield of their degradation process, demands the correct quantitative determination of the polluting agents studied. For simplicity reasons, various research articles have been published in these fields carrying out the study of the process quantifying by univariate form using a UV-visible detector. The results obtained by this way can lead to erroneous conclusions of the process reality, mainly of their quantification.

We have studied three dyes. Table 1 shows the structure and the name of these compounds. Dyes were obtained from Trumpler Española, S.A. and used without further purification.

**Table 1** Structure of the dyes

Dye	Structure
Acid Red 97	
Acid Orange 61	
Acid Brown 425	

We have developed and compared two analytical methods for the simultaneous determination of dyes concentration. Both methodologies are based on the application of multivariate curve resolution method specifically multivariate curve resolution with alternating least squares (MCR-ALS) to the UV-vis spectra recorded. One way to record UV-vis spectra is using a sequential injection analyzer (SIA)<sup>2</sup>, the other way to obtain information is recording the UV-vis spectra using the spectrophotometer.

---

<sup>1</sup> O. Prieto, J. Feroso, Y. Nuñez, J.L. del Valle, R. Irusta. *Solar Energy*. 79 (2005): 376 – 383.

<sup>2</sup> V. Gómez, J. Font, M.P. Callao. *Talanta*. 71 (2007): 1393-98

# CHEMOMETRIC RESOLUTION OF NIR SPECTRA DATA OF A MODEL AZA-MICHAEL REACTION WITH A COMBINATION OF LOCAL RANK EXPLORATORY ANALYSIS AND MULTIVARIATE CURVE RESOLUTION-ALTERNATING LEAST SQUARES (MCR-ALS) METHOD

Vanessa del Río<sup>a</sup>, M. Pilar Callao<sup>a</sup>, M. Soledad Larrechi\*

Lucas Montero de Espinosa<sup>b</sup>, J. Carles Ronda<sup>b</sup>, Virginia Cádiz<sup>b</sup>

(a) Chemometrics, Qualimetrics and Nanosensors Research Group (b) Polymer Research Group Analytical Chemistry and Organic Chemistry Department, Rovira i Virgili University.

Marcel·li Domingo s/n, Campus Sescelades, 43007 Tarragona, Spain.

\* mariasoledad.larrechi@urv.cat

The aza-Michael reaction, a variation of the Michael reaction in which an amine acts as the nucleophile, permits the synthesis of sophisticated macromolecular structures with potential use in many applications such as drug delivery systems, high performance composites and coatings [1]. The aza-Michael product can be affected by a retro-Mannich-type fragmentation [2]. A way of determining the reactions that are taking place and evaluate the quantitative evolution of the chemical species involved in the reactions is presented. The aza-Michael reaction between a modified fatty acid ester with  $\alpha,\beta$ -unsaturated ketone groups (enone containing methyl oleate (eno-MO)) and aniline (1:1) was studied isothermally at 95 °C and monitored *in-situ* by near-infrared spectroscopy (NIR).

The number of reactions involved in the system was determined analyzing the rank matrix of NIR spectra data recorded during the reaction. Singular value decomposition (SVD) and evolving factor analysis (EFA) adapted to analyze full rank augmented data matrices have been used [3]. In the experimental conditions, we found that the resulting aza-Michael adduct undergoes a retro-Mannich-type fragmentation, but the final products of this reaction were present in negligible amounts. This was confirmed by recording the <sup>1</sup>H NMR spectra of the final product. Applying Multivariate Curve Resolution – Alternating Least Squares (MCR-ALS) to the NIR spectra data obtained during the reaction, it has been possible to obtain the concentration values of the species involved in the aza-Michael reaction [4]. The performance of the model was evaluated by two parameters: ALS lack of fit ( $lof = 1.31\%$ ) and explained variance ( $R^2 = 99.92\%$ ). Also, the recovered spectra were compared with the experimentally recorded spectra for the reagents (aniline and eno-MO) and the correlation coefficients ( $r$ ) were 0.9997 for the aniline and 0.9578 for the eno-MO.

---

[1] B.D. Mather, K. Viswanathan, K.M. Miller, T.E. Long, Prog. Polym. Sci. 31 (2006) 487.

[2] S.G. Lim, C.H. Jun, Bull Korean Chem. Soc. 11 (2004) 1623.

[3] A. de Juan, S. Navea, J. Diewok, R. Tauler, Chem. Intell. Lab. Syst, 70 (2004) 11.

[4] M. Garrido, M.S. Larrechi, F.X. Rius, Appl. Spectrosc. 60 (2006) 174.





# Seminario Clasificación y modelado



# **CLASSIFICATION AND CLASS MODELING TECHNIQUES: PROBLEM, SAMPLING, METHOD, EVALUATION**

*Michele Forina, Monica Casale, Paolo Oliveri  
University of Genova*

Chemometrics began about forty years ago with the application of multivariate classification techniques to chemical problems. The first tools, linear discriminant analysis and linear learning machines, arrived from statistics and from the newborn artificial intelligence.

In 1976-1977 the first class modeling technique born in Chemometrics, SIMCA, was presented, with special attention to classification.

Also today many people working with Chemometrics technique don't understand the difference between classification and class modeling.

In this lecture the hierarchy of data analysis in the case of the use of classification or class modeling is presented, with special attention to the problem, to the related collection of samples, to the uncertainty of the evaluation parameters.





# **CLASSIFICATION METHOD CAIMAN AND CLASSIFICATION OF SEQUENTIAL DATA**

*Roberto Todeschini*

*Universidad de Milán*

In this lecture, the methodology of classification based on CAIMAN will be presented. It is implemented to do discriminant, modelling or asymmetric classification.

In addition, a distance for sequential data, as spectra, peptides, sequences of DNA, etc., will be also presented. This measure allows comparing and classifying these types of data. Examples for food classification are given. The food is characterized by measures with an electronic nose together with the detection with more discriminant sensors.



## **SENSITIVITY AND SPECIFICITY AS CRITERIA TO MODEL CATEGORIES**

Luis A. Sarabia<sup>2</sup>, M.Cruz Ortiz<sup>1</sup>, M<sup>a</sup> Sagrario Sánchez<sup>2</sup>

<sup>1</sup>Dpto. de Química (Química Analítica), <sup>2</sup>Dpto. de Matemáticas y Computación  
Facultad de Ciencias, Universidad de Burgos,  
Pza. Misael Bañuelos s/n 09001 Burgos, España

### **Abstract**

Class-modelling represents an important part into the pattern recognition and is newly emerging above all for characterization of products.

The most interesting characteristic of modelling methods as against discriminating methods is the possibility of estimating the sensitivity and specificity for each class-model. These two parameters qualify the goodness of the models computed and allow evaluation of the confusion among categories and detection of foreign data. On the other hand, we wish class-models with high sensitivity and specificity. However, in general, these two parameters show a conflicting behaviour: when one increases the other one decreases and vice versa.

This lecture revises the possibilities of looking for 'balanced' class-models and, alternatively, the search for a complete set of class-models that show the different achievable values in sensitivity and specificity. Two procedures for this task are presented: one of them based on the direct construction of the decision function associated to each class, by solving the problem of a multicriteria optimization to obtain the Pareto-optimal front of the solutions; the other one is based on the definition of a class-modelling problem as a hypothesis test to decide between the probability distribution for each class. In practice, the PLS regression models with discrete response are shown to be the best method to tackle the problem. In this way, the user may choose the class-model that best adapt to a specific need.

### **Acknowledgements**

This work has been partially supported by Spanish Ministerio de Ciencia e Innovación (under project number CTQ2008-02264/BQU) and regional Consejería de Educación de la Junta de Castilla y León under project BU024A07, both co-financed with European FEDER funds.



## MODELING METHODS BASED ON SVMs AND RELATED TECHNIQUES

Beata Walczak

*Institute of Chemistry, Silesian University, 9 Szkolna Street, 40-006 Katowice, Poland*

Modern instrumental techniques yield multidimensional data (spectra, chromatograms, microarray data etc). Processing of a limited number of samples, described by thousands of variables, requires careful choice of the modeling approaches. According to the learning theory, a classification or calibration model, constructed for a limited number of data samples, can have good generalization properties (i.e., predictive ability) only, when the model complexity is controlled.

Linear modeling techniques, well-suited for ill-posed (i.e., collinear) data with good generalization properties such, as Principal Component Regression, Partial Least Squares and Reduced Rank Regression can be considered as special cases of Continuum Regression, and extended to the powerful non-linear techniques. This extension can be achieved by replacement of the standard inner product with an appropriate (chosen) 'kernel' function and it can be thought of as of a non-linear mapping to a high-dimensional feature space (provided, that the kernel computes the inner product of the feature vectors, corresponding to the two inputs). The kernel can be exploited, whenever the weight vector is expressed as a linear combination of the training objects. The use of kernels is an attractive computational short-cut, as all the necessary computations are performed directly in the input space. It allows applications of linear algorithms to work in a high-dimensional feature space, but in practice, it does not involve any computations in that high-dimensional space. We even do not need to know the underlying feature map in order to calculate the kernel matrix.

Within the presented framework, the Support Vector Machines and the kernel latent variables approaches will be discussed.



# ÍNDICE DE AUTORES

Aires de Sousa, J.	P-9
Alier Pedemonte, M.	P-10
Álvarez Ávila G.	P-13
Andrade Garda, J.M.	O-5, P-9, P-13
Barceló, D.	O-2
Birlouez-Aragon, I.	P-12
Blanchet, L.	O-1
Blanco, E.	P-5
Boqué Martí, R.	O-8
Bosque Sendra, J.M.	O-11
Botella, C.	O-8
Callao, M.P.	P-3, P-15, P-16
Camacho Páez, J.	P-7, P-11
Carlosena, A.	P-9
Casale, M.	S-1
Cela Torrijos, R.	O-7, P-5
Conesa, A.	O-6
Cuadros Rodríguez, L.	O-11
Di Anibal, C.V.	P-3
Díez Azofra, R.	P-12
Esteban-Díez, I.	P-6
Felipe-Sotelo, M.	P-10
Fernández Barrat, C.	P-15
Fernández Varela, R.	O-5
Ferré, J.	O-8
Ferrer, A.	O-6, P-2, P-7, P-11
Forina, M.	S-1
García Lavandeira, J.	O-7, P-5
Garduño Rosas, J.A.	P-13
Gargallo, R.	P-8
Garrido-Varo, A.	O-9
Gómez Carracedo, M.P.	P-9

González-Sáiz, J.M.	O-10, P-6
Guerrero-Ginel, J.E.	O-9
Hernández, I.	P-10
Herrero Gutiérrez, A.	O-3
Huvenne, J.P.	O-1
Jaumot Soler, J.	P-8
Juan, Anna de	O-1, P-14
Lacorte, S.	P-14
Larrechi, M.S.	O-4,P-15, P-16
Llompart, M.	P-5
López Arellano, R.	P-13
Lores, M.	O-7
Losada, B.	O-7
Martín, J.	P-1
Martínez-Pontevedra, J.A.	O-7
Mas Garcia, S.	P-14
Mata Espinosa, P. de la	O-11
Muniategui, S.	O-5
Nueda Roldán, M.J.	O-6
Oliveri, P.	S-1
Ortiz Fernández, M.C.	O-3, P-4, P-12, S-3
Pérez del Notario, N.	O-10
Pérez-Marin, D.	O-9
Picó, J.	P-7, P-11
Pizarro Millán, C.	O-10, P-6
Platikanov, S.	P-1
Prada, D.	O-5,P-9
Prats Montalbán, J.M.	P-2
Real García, B.D.	O-3
Reguera Alonso, C.	P-4
Riccioli, C.	O-9
Río Sánchez, V. del	P-16
Rodríguez Contreras, N.	O-3
Rodríguez Tecedor, S.	O-10
Rodríguez, I.	P-5



Ruckebush, C.	O-1
Ruisánchez, I.	O-4,P-3
Ruiz Samblás, C.	O-11
Sáenz-González, C.	O-10
Salgado, M.C.	P-5
Sánchez Pastor, M.S.	O-3, P-4, S-3
Santander García, E.A.	P-13
Sarabia Peinador, L.A.	O-3, P-4, P-12, S-3
Spegazzini, N.	O-4
Tauler, R.	O-1, O-2, P-1, P-10, P-14
Terrado, M.	O-2
Todeschini, R.	S-2
Vallesquino-Laguna, P.	O-9
Walczak, B.	CI, S-4



## LISTA DE PARTICIPANTES

<b>Nombre</b>	<b>Apellidos</b>	<b>Universidad</b>	<b>e-mail</b>
Marta	Alier Pedemonte	CSIC	martaquimica@yahoo.es
José Manuel	Andrade Garda	Coruña	andrade@udc.es
Lionel	Blanchet	Barcelona	lionel.blanchet@ed.univ-lille1.fr
Ricard	Boqué Martí	Rovira i Virgili	ricard.boque@urv.cat
Juan Manuel	Bosque Sendra	Granada	jbosque@ugr.es
Cristina	Botella	Rovira i Virgili	cristina.botella@urv.cat
José	Camacho Páez	Politécnica de Valencia	jcamacho@isa.upv.es
Rafael	Cela Torrijos	Santiago de Compostela	rafael.cela@usc.es
Luis	Cuadros	Granada	lcuadros@ugr.es
Carolina Vanesa	Di Anibal	Rovira i Virgili	carolinavanesa.dianibal@urv.cat
Rocío	Díez Azofra	Burgos	rdazofra@ubu.es
Isabel	Esteban Díez	La Rioja	isabel.esteban@unirioja.es
Cristina	Fernández Barrat	Rovira i Virgili	cristina.fernandez@urv.cat
Raquel	Fernández Varela	Coruña	rfernandezv@udc.es
Alberto	Ferrer	Politécnica de Valencia	aferrer@eio.upv.es
Michele	Forina	Génova (Italia)	forina@dictfa.unige.it
Juan	García Lavandeira	Santiago de Compostela	
M <sup>a</sup> Paz	Gómez Carracedo	Coruña	pmgomez@udc.es
José María	González Sáiz	La Rioja	josemaria.gonzalez@unirioja.es
Ana	Herrero Gutiérrez	Burgos	aherrero@ubu.es
Joaquim	Jaumot Soler	Barcelona	joaquim@apolo.ubi.es
Anna de	Juan	Barcelona	anna.dejuan@ub.edu
Raquel	López Arellano	Coruña	rlajjd@yahoo.com.mx
Sílvia	Mas Garcia	CSIC	smgqam@iiqab.csic.es
Paulina de la	Mata Espinosa	Granada	pmata@correo.ugr.es
Arsenio	Muñoz de la Peña	Extremadura	arsenio@unex.es
Cristina	Nebreda Gómez	Burgos	cng0004@alu.ubu.es

<b>Nombre</b>	<b>Apellidos</b>	<b>Universidad</b>	<b>e-mail</b>
María José	Nueda Roldán	Politécnica de Valencia	mj.nueda@ua.es
M <sup>a</sup> Cruz	Ortiz Fernández	Burgos	mcortiz@ubu.es
Rafael	Pardo Almudí	Valladolid	rpardo@qa.uva.es
Nuria	Pérez del Notario	La Rioja	nuria.perez-del-notario@unirioja.es
Consuelo	Pizarro Millán	La Rioja	consuelo.pizarro@unirioja.es
Stefan	Platikanov	CSIC	splqam@iiqab.csic.es
José Manuel	Prats Montalbán	Politécnica de Valencia	jopramon@eio.upv.es
Blanca Delia	Real García	Burgos	blanki_d@hotmail.com
Celia	Reguera Alonso	Burgos	cregue@ubu.es
Cecilia	Riccioli	Córdoba	ceciliariccioli@hotmail.com
Vanessa del	Río Sánchez	Rovira i Virgili	vanessa.delrio@urv.cat
Noelia	Rodríguez Contreras	Burgos	nrcontreras@ubu.es
Sofía	Rodríguez Tecedor	La Rioja	sofia.rodriguez@alum.unirioja.es
Cristina	Ruiz Samblás	Granada	
Cristina	Sáenz González	La Rioja	cristina.saenz@unirioja.es
M <sup>a</sup> Sagrario	Sánchez Pastor	Burgos	ssanchez@ubu.es
Luis A.	Sarabia Peinador	Burgos	lsarabia@ubu.es
Nicolás	Spegazzini	Rovira i Virgili	nicolas.spegazzini@urv.cat
Romà	Tauler	CSIC	rtaqam@iiqab.csic.es
Marta	Terrado	CSIC	mtcqam@iiqab.csic.es
Roberto	Todeschini	Milán (Italia)	roberto.todeschini@unimib.it
M <sup>a</sup> del Sol	Vega Alegre	Valladolid	solvega@qa.uva.es
Beata	Walczak	Silesia (Polonia)	beata@us.edu.pl
Juan	Zuriarrain Ocio	País Vasco	juan.zuriarrain@ehu.es