# Recent applications of first- and second-order multivariate calibration to analytical chemistry

**JUAN A. ARANCIBIA, PATRICIA C. DAMIANI, GABRIELA A. IBAÑEZ, ALEJANDRO C. OLIVIERI***

Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Instituto de Química Rosario (IQUIR-CONICET), Suipacha 531, Rosario (2000), Argentina

**In this report, we review the recent literature on the application of multivariate calibration techniques to both first- and second-order data, aimed at the analytical determination of analytes of interest or sample properties in a variety of industrial, pharmaceutical, food and environmental samples, including examples of process control. The most employed data processing tools will be briefly described, with emphasis on the advantages which can be obained by applying specific combinations of multivariate data and algorithms. The main focus will be on works devoted to first-order data (i.e., spectra, chromatograms, etc.) combined with partial least-squares (PLS) regression, which has become the standard for this type of analytical research. A brief discussion on recent work on second-order data and algorithms is also included, as this field is rapidly growing, although it does not show at present the general applicability of the first-order counterparts.**

Classical analytical calibration is based on fully selective instrumental signals (1). Any foreign component producing a signal under the same experimental conditions as the analyte constitutes an interference, and leads to a bias in its determination (2). The presence of interferences calls for the employment of complementary techniques, such as: a) sample pre-treatment or cleanup in order to remove the interferences previous to the analyte determination, b) masking the interference effect, leaving the analyte signal as the only measured one, or c) physical separation of analyte and using chromatography, capillary electrophoresis, etc.

There are instances where these approaches would not work, for a variety of reasons, such as cost, time, or simply because none of them is able to free the analyte from the interference. In these cases, a valid alternative is the measurement and processing of multiple signals, some of which are only partially selective regarding the analyte of interest. The whole subject started when near infrared (NIR) spectroscopy was applied to the determination of components of intact materials, such as protein in seeds (3). NIR spectra are composed of the superposition of many different signals, arising from the several constituents (most of them possibly unknown) of the sample. The subsequent development of multivariate techniques for processing these data led to the outburst of a miryad of algorithms, starting the discipline of chemometrics in the decade of 1960 (4).

In order to place classical and multivariate calibration in a broader scenario, a consistent nomenclature of different data types is required. One possibility is to employ the concept of ‘order’, widely used in analytical chemistry (5,6). The order is a tensorial property of the data measured for a single sample: scalars are zeroth-order tensors, and thus univariate calibration is also known as zeroth-order. If spectra are measured (or other unidimensional vectors per sample), the calibration becomes first-order. Increasing the number of data dimensions per sample leads to correspondingly complex data structures, which give rise to

second- and higher-order multivariate calibration. An alternative nomenclature is based on the number of ways, which is the number of modes of a data array for a group of samples (7). In this context, univariate calibration is based on one-way data, first-order calibration on two-way data, etc. In what follows, the concept of order is employed, because it is more familiar to analytical chemists, and is linked to the expression 'second-order advantage', popular in the analytical community (8). Figure 1 pictorially shows the relationship among data orders.

Figure 1 near here.

First-order multivariate calibration is mainly dominated by partial least-squares (PLS) regression, a technique developed around 1970 by Hermann Wold and coworkers, particularly his son Svandte Wold (4). PLS suitably combines two characteristics which makes it appealing for NIR analysis of complex materials: a) an inverse calibration phase, in which concentration information is regressed onto signals rather than signals onto concentrations, and b) dimensionality reduction, a technique which adequately compresses a data matrix of size samples × hundreds (or thousands) of variables (wavelengths) into a much smaller matrix, of size samples × a small number of the so-called latent variables. The latter number is close to the number of chemically responsive components in the spectral region of interest.

Today PLS is the most applied first-order multivariate calibration method, with the exception of non-linear analytical systems, for which algorithms such as artificial neural networks or support vector machines appear to be preferable (9). Concerning the analyzed data, they include not only spectra, i.e., NIR, mid-IR (MIR), UV-visible absorption, luminescence, but other type of vectorial signals as well, such as chromatograms, electrochemical data (voltammograms), and responses associated with arrays of sensors (electronic tongues or noses) or sensometric parameters.

First-order multivariate calibration demands preparing a sufficiently representative set of calibration samples. All chemical components expected to be found in future test samples

should be present in the calibration set, leading in certain cases to the collection of a large number of training samples. The latter problem may be greatly alleviated if second- or higher-order signals can be measured per sample (5,6), because in this case the relevant algorithms are able to separate the contribution of the interferences from that of the analyte. This property is known as 'second-order advantage' (8), and permits the building of considerably simpler calibration sets; sometimes simply composed of a handful of pure analyte samples.

In the present review, progress in the application of first- and second-order multivariate calibration which has taken place in the last years will be reviewed, with emphasis experimental analytical works.

### *First-order multivariate calibration algorithms*

The first step in the multivariate processing of first-order data is to build a reference data set with the recorded signals for as many samples as required to span the chemical variability expected in future samples, and the reference values of the analyte concentrations or properties of samples (10). The various algorithms are designed to correlate the recorded signals with the concentrations, so that a model is produced which can then be applied to signals for an unknown sample in order to predict its analyte content or property value.

Collection of spectra (or other vectorial signals) leads to the building of a calibration matrix of signals. Usually this matrix has many more signal values than samples, since instruments deliver measurements at a large number of sensors, and signals may be correlated, since some samples may have similar composition. This poses some challenges to first-order calibration algorithms, with the result than some of them are of limited applicability, and have been surpassed with the advent of more powerful tools.

The simplest first-order algorithm is classical least-squares (CLS), which resorts to two basic chemical laws: (1) the direct proportionality between signal and concentration and

(2) the additivity of signals of the sample components. CLS belongs to a group of methods known as direct, because they are based on the assumption that signals are proportional to concentrations. In order to be able to apply CLS, however, one needs to know all the components in the calibration set of samples and their concentrations, something which is not possible for complex or natural samples, and seriously limits its applicability (10).

Inverse methods are preferable to overcome this problem, because by assuming that concentration is linearly related to signal, they are able to build a calibration model, ignoring the concentrations of most sample components except for a handful of analytes of interest, or even a single one. Figure 2 illustrates a group of NIR spectra collected for a set of seeds, intended to calibrate for the determination of fat, humidity, protein and starch. Only inverse methods can be applied for this purpose.

Figure 2 near here.

The simplest version of this type of algorithms is inverse least-squares (ILS) (10), which however shows a major drawback: it requires less sensors than samples. This calls for complex procedures for sensor selection, with the result that important information has to be removed from the calibration signals.

Two popular algorithms for processing full-sensor first-order data using inverse calibration are principal component regression (PCR) and PLS (10,11). The first step in these methodologies is to decompose the calibration data matrix in the so-called latent variables: loadings and scores. They are called ‘latent’ (meaning occult) is opposed as ‘explicit’, which are the raw instrumental variables. The decomposition is accomplished using an algorithm whose aim is obtaining loadings and scores representing as much as possible the spectral variance in the calibration matrix (PCR), and at the same time the maximal correlation between signals and the concentrations of an analyte of interest in each of the calibration samples (PLS) (10). The latter characteristic has made PLS the *de facto* standard for first-

order multivariate calibration. One should note that there are two PLS versions: PLS-1, which focuses on a single analyte at a time, and seems to be the preferred variant, since the model is optimized towards a given analyte, and PLS-2, which allows one to calibrate for several analytes simultaneously, but is less employed (4).

It is important to mention that the raw data matrix is sometimes subjected to a mathematical pre-treatment procedure, in order to remove spectral variations due to artifacts or to physical phenomena but not to sample chemical composition (11). This is particularly important in NIR analysis of solid or semi-solid materials, where dispersion phenomena may be unrelated to the measured property. Good accounts on the subject can be found in the literature (11). In the case of chromatographic data, the application of PLS requires that the elution profiles of different samples are first aligned or synchronized, because this phenomenon is not related to the chemical composition of samples. Otherwise, the method will not work or an unnecessary number of latent variables will be needed in order to compensate for shifts, broadenings and peak shape changes among chromatograms.

As explained above, PLS is a linear regression method. It may cope with slight deviations of linearity in the signal-concentration relationship by including additional latent variables. However, it cannot be applied to strongly non-linear problems, where one must resort to truly non-linear algorithms such as artificial neural networks, support vector machines, radial basis function networks (9) or non-linear PLS versions (12).

### *Second-order multivariate calibration algorithms*

It is not the purpose of the present work to discuss the many available second-order algorithms, although a brief account will be given. One group of methods comprises the so-called trilinear algorithms, which require the data to fulfil the condition of trilinearity. In plain

terms, this basically implies that: a) the constituent signals are proportional to concentrations, and b) their profiles along both instrumental modes are unique and equal for all samples (13).

The most popular trilinear algorithm is parallel factor analysis (PARAFAC) (14). The latter has become the algorithm of choice due to its efficiency, robustness, ability to process multiple samples, and availability of constraints to be applied during the modelling phase, which aids in reaching physically interpretable results (14). PARAFAC achieves the second-order advantage, often leading to uniquely defined models, in which the contribution of the potential interferences and the analytes to the total signal are adequately decomposed.

When component profiles change from sample to sample (e.g., chromatograms) PARAFAC cannot be applied, and non-trilinear algorithms are needed, with the most popular alternative being multivariate curve resolution-alternating least-squares (MCR-ALS) (15). In the so-called extended mode (16), the latter operates by decomposing an augmented data matrix created from the set of signals. Decomposition is achieved through ALS under a series of sensible constraints, which provide physical interpretability to the final solution and limit their possible number. MCR-ALS needs initial estimates of component profiles, which can be efficiently computed using several methods (17,18).

Figure 3 pictorially shows the different between PARAFAC, which builds a three-way data structure with the individual data matrices, and assumes a trilinear model for decomposition, and MCR-ALS, which constructs an augmented data matrix in a given direction and assumes a bilinear model for decomposition.

Figure 3 near here.

For additional analytical systems deviating from the trilinear model, latent-variable regression methods may be useful, such as unfolded and multi-dimensional PLS (U-PLS and N-PLS respectively) (19,20). Here the achievement of the second-order advantage is a post-calibration activity, a procedure called residual bilinearization (RBL), which separates the

portion of the signal explained by calibration from the contribution of the interferences. The result is the flexible U-PLS/RBL and N-PLS/RBL methods (21,22).

Table 1 provides a summary of second-order algorithms, and Figure 4 shows a three-dimensional plot of an excitation-emission fluorescence matrix for a sample of an antibiotic. The latter can be quantitated in human urine even in the presence of a fluorescent background from the biological matrix, illustrating the achievement of the second-order advantage.

Figure 4 near here.

### *Software for multivariate calibration*

A variety of free software is available on the internet from a large variety of sources. One particular example is the MVC1 toolbox (23) which employs useful MATLAB graphical user interface (GUI) for making easy the data loading and manipulation without knowledge of MATLAB programming.

For second-order calibration, the most employed software is freely available as MATLAB codes, including useful graphical interfaces, as shown in Table 2 (24-26). Almost no commercial software exists, which in itself is an indication that second-order analysis is still in its infancy regarding its popularization among analytical chemists.

### *First-order multivariate calibration examples*

The number of literature references available for review is too large to be thoroughly reviewed in the present report. We have thus selected a representative number of works published in the last years, with emphasis on covering several different research areas. In the cited papers, PLS has been applied for data processing, in many cases along with other first-order multivariate variants for comparison, including non-linear techniques such as artificial

neural networks. Our view is that PLS is the de facto standard for these analyses, and that only in cases of strong non-linearities it is justified to move to alternative methodologies.

### *Food analysis*

Probably because of historic reasons, spectroscopy/multivariate calibration has been widely employed for food analysis, with NIR/PLS-based calibrations already established in official documents (27). Hence, this discussion occupies a prominent place in the present review.

NIR spectroscopy is ubiquitous in the analysis of dairy products. Coupled to spectral pre-treatment and PLS or other chemometric models, it was used to measure fat, proteins and carbohydrate in milk powder samples (28), fat, protein, lactose, urea and somatic cell count in milk (29), and fat, dry matter, protein and fat/dry matter contents in cheeses (30).

Beverages can be conveniently analyzed by IR/PLS. In wines and related samples, MIR/PLS was used to determine tartaric, malic, lactic, succinic, citric and acetic acids (31) and several anthocyanins (32). NIR/PLS allowed the measurement of haloanisoles and halophenols, responsible for musty taint defects in barrel aged red wines (600 wines of different ageing time and from four different geographic zones) with reference values obtained by gas chromatography with mass spectral detection (GC-MS) (33), and calcium, potassium, magnesium, phosphorus, sodium, sulphur, iron, boron and manganese (34). In the case of juices, glucose, fructose and sucrose were determined in bayberry juice by NIR/PLS (35). The sensorial attributes acidity, bitterness, flavor, cleanliness, body and overall quality of coffees were assessed by NIR/PLS (36).

Food adulteration has been the subject of several works. Olive oils possibly adulterated with different vegetable oils were analyzed using MIR/PLS (37,38), and soya bean products, widely used in the animal feed industry as a protein based feed ingredient,

were checked for adulteration with melamine using NIR/PLS to detect the latter in dehulled soya, soya hulls and toasted soya (39). The fraudulent addition of barley to roasted and ground coffee samples was identified and quantitated (40), and added sugar, total soluble solids and real juice content in fresh and commercial mango juice were measured by MIR/PLS in order to predict the adulteration of mango juice by added sugar (41)

Interesting developments involve NIR/PLS study of microbial contamination. Microbial spoilage was detected on Atlantic salmon (*Salmo salar*) by a PLS model which predicts the number of bacteria that would be present after nine days of storage (42). Also, aflatoxin B1 was measured in red chili powder (43), and the fungal toxins fumonisin B1 and B2 in corn meal (44).

Comparatively less employed are PLS-processed Raman and UV-visible spectral data. Some Raman examples which deserve to be cited are the determination of fat in liquid homogenized milk (45), furazoline and malachite green in fish samples (using the surface-enhanced Raman mode) (46), sulfonamide residues in muscle foods (47) and glucose in sport drinks by visible micro-Raman/PLS (48). Concerning UV-visible data, caramel was measured in spirits (cachaca, whiskies and brandies) aged in oak casks (49),cobalt and nickel were determined in water, food and geological certified reference materials after ionic liquid-based dispersive liquid-liquid microextraction and complexation with 1-(2-pyridylazo)-2-naphthol (50), two herbicides (atrazine and cyanazine) in rice, mealie, soybean, pear, apple, cauliflower and cabbage, after reaction with *p*-aminoacetophenone (51), the dyes Allura Red, Sunset Yellow and Tartrazine in powdered soft drinks (52), and monosodium glutamate, guanosine-5'-monophosphate and inosine-5'-monophosphate in stock cube samples, without any previous extraction step, using stopped flow injection (53).

Kinetic UV-visible spectrophotometric data, i.e., time profiles at a single wavelength, have been employed, via PLS calibration, for food analysis. The dyes Amaranth, Ponceau 4R,

Sunset Yellow, Tartrazine and Brilliant Blue were determined in fruit juices, tea and jellies, by oxidation with iron (III) acid solution, followed by reaction of the generated iron (II) with hexacyanoferrate (III) to yield Prussian blue. Several algorithms besides PLS were employed, including non-linear methods (54). Acesulfame-K, sodium cyclamate and saccharin sodium were measured in artificial sweeteners relying on the kinetic differences in the analyte oxidations by $KMnO_4$ (55).

Electrochemical data processed with PLS usually correspond to promising devices such as electronic tongues, which are arrays of partially selective electrodes based on voltammetric or other techniques. They allowed to determine polyphenols in wine (56), theaflavins and thearubigins in black tea (57), in both cases including data pre-treatment and comparison with non-linear methods. Similar systems were proposed for measuring chloride, nitrite and nitrate in minced meat (58), and fructose and glucose in soft drinks, in the latter case using a potentiometric electronic tongue containing 36 lipo/polymeric membranes (59).

Chromatographic data showing partial peak resolution is being gradually incorporated into multivariate calibration protocols. PLS usually allows to cope with coeluting chromatograms, successfully predicting individual analyte concentrations or sample properties. For example, beer sensorial parameters such as bitterness and grain taste were predicted from unresolved GC-MS data (60) and pine nuts and Pecorino in Pesto Genovese by headspace sorptive extraction and GC-MS data (61). The content of olive oil in blends was assessed from GC-MS (62) and high performance liquid chromatographic (HPLC) (63) profiles of triacylglycerol, and also from CG-flame ionization detection (FID) profiles of fatty acids methyl esters (64). Protein HPLC profiles with UV-visible detection were used to control the types of milk used in the elaboration of dairy products, detecting adulterations in milk mixtures and cheeses (65).

Miscellaneous techniques include visible light scattering for measuring fat and total protein in milk (66), impedance spectroscopy for salt content in minced meats, cured hams and pork loins (67), time domain reflectometry for content in extra virgin olive oils (68), differential scanning calorimetry for fatty acid composition (palmitic, stearic, oleic and linoleic acids, saturated, mono and polyunsaturated, oleic/linoleic and unsaturated/saturated ratios) in vegetable oils (69), nuclear magnetic resonance (NMR) for organic acids (acetic, citric, lactic, malic, pyruvic and succinic) in beer (70), fluorescence for riboflavin and the aromatic amino acids tryptophan, tyrosine and phenylalanine in beer (71), and matrix-assisted laser desorption ionization-time of flight-mass spectrometry (MALDI-TOF-MS) for quantitation of bacterial spoilage in milk and pork meat (72).

### *Pharmaceutical and biomedical analyses*

Of particular interest to the pharmaceutical industry is the combination of multivariate calibration to NIR, MIR and Raman spectroscopies, because they allow for the non-destructive analysis of intact pharmaceutical forms (73-75). This includes the determination of both active pharmaceutical ingredients and excipients, which can be performed remotely through optical fibers. The results are often comparable to those obtained by official HPLC, and are being incorporated into official protocols (76).

As examples, hydrochlorothiazide was determined in powdered manufactured tablets (77), chondroitin, glucosamine, and methyl sulfonyl methane in tablets (78), amoxicillin in powdered drugs (79) and oral suspensions (80), ranitidine during the manufacturing process (in granulates and cores) and in the final step (in coated tablets) (81), azithromycin in tablets (82), and particle size distribution in paracetamol powders (83), all resorting to the successful combination NIR/PLS. Glucose, fructose and maltose were measured in depilatory formulations without any sample pretreatment, based on MIR/PLS, with similar results to

HPLC with on-line infrared detection (84). Attenuated total reflectance-Fourier transformed-MIR and Raman spectra allowed to detect the presence of acetaminophen in over-the-counter pharmaceutical formulations (85).

Raman spectroscopy is also suitable for quality control assays and to identify polymorphic forms or monitor phase changes in pharmaceutical products. Coupled to PLS, it allowed in situ anhydrate concentration measurements during the crystallization and phase transition processes of citric acid in water (86).

A recently introduced technique for non-destructive characterization of pharmaceutical materials is the collection of hyperspectral data: NIR or Raman spectra are collected for many different pixels in the surface of a solid, and a chemical image (CI) is produced reflecting the component distribution using PLS. The model can be built on the average spectrum of the images (first-order data) and can be used to predict bulk concentrations in unknown images (from new image average spectra) or concentrations at a pixel level (from new individual pixel spectra), which provides chemical distribution maps. This helped to quantitate piroxicam polymorphs (87) and carbamazepine by NIR-CI (88) and to evaluate atorvastatin calcium formulations by Raman-CI (89).

Additional pharmaceutically relevant data have been processed by PLS, such as thermogravimetric profiles for the resolution of paracetamol and codeine phosphate (90) and thin layer chromatography-densitometric data for atenolol, chlorthalidone and their degradation products (91). For completeness, we include the PCR/PLS processing of UV-visible data for simultaneous determinations in multi-component pharmaceutical forms after dissolution, such as mixtures of ambroxol and doxycycline (92), aspirin, paracetamol, caffeine, and chlorphenamine (93), paracetamol, ibuprofen and caffeine (94), two antiparkinsonians (95), paracetamol, propiphenazone, caffeine and thiamine (96), diprophylline, guaiphenesin, methylparaben and propylparaben or clobutinol, orciprenaline,

saccharin sodium and sodium benzoate (97),chlorpheniramine maleate and phenylpropanolamine hydrochloride with ibuprofen and caffeine (mixture 1) or with propyphenazone (98),dienogest (DIE) and estradiol valerate (EST) in sugar-coated tablets (99)

In this regard, for some pharmaceutical mixtures of simple composition, the basic CLS algorithm was shown to be useful, as in the UV-visible spectrophotometric study of binary, ternary and quaternary mixtures of the water-soluble vitamins thiamine, pyridoxine, riboflavin and cyanocobalamin (100), and in the determination of ezetimibe and simvastatin in bulk powder, laboratory-prepared mixtures and a combined dosage form (101). The results were compared with those provided by PCR and PLS respectively.

In the biomedical field, NIR data have been applied for direct in situ analysis in complex biological systems. In this context, the non-invasive glucose monitoring in blood using NIR/PLS or MIR/PLS is perhaps the single most important contribution of the combination of spectroscopy and chemometrics to biomedical analysis, because of its impact in the monitoring of diabetic patients (102-104). Despite the efforts devoted to this project in recent decades, however, the final aim appears to be elusive (105).

Following earlier ideas on the application of IR spectroscopy combined to multivariate techniques for the rapid screening of clinically relevant parameters in human blood (106), there has been a recent revival on the subject. Promising research works are the determination of albumin, immunoglobulin, total globulin, and albumin/globulin by MIR/PLS (107), glucose, glucose-6-phosphate and pyruvate by NIR/PLS (108) and glucose, human serum albumin and γ-globulin by NIR/PLS (109). The results may find applications in the rapid screening of clinical parameters with blood samples in the microliter range.

NIR/PLS has also enabled the determination of triglycerides and high-density lipoprotein in rat plasma (110), and fish sperm deoxyribonucleic acid in solution after

adsorption pre-concentration (111), while dispersive near-infrared Raman spectroscopy has been applied for diagnosing toxoplasmosis, quantitating anti-Toxoplasma gondii (IgG) antibodies in blood sera from domestic cats, with comparable results to the enzyme linked immunosorbent assay (ELISA) (112).

### *Industrial and miscellaneous applications*

There are many examples of multivariate calibration with impact on industrial fields. The driving force has been the replacement of classical analytical methodologies by techniques capable of producing reliable results in a short time, with simple equipment, and amenable for automation. A nice example is the estimation of the octane number in fuels, a measure of the resistance to autoignition in internal combustion engines. The most common type of octane rating is the Research Octane Number (RON), determined by running the fuel in a test engine with a variable compression ratio under controlled conditions, and comparing the results with those for mixtures of iso-octane (RON = 100) and *n*-heptane (RON = 0). NIR spectroscopy and PLS allows for an accurate measurement of octane number, in a considerably much simpler and faster way than with the engine method (113).

Other relevant fuel properties can be adequately measured using the PLS-assisted analysis of the indicated data: a) in diesels, flash point and cetane number (114) and quality parameters (115) by NIR, residual oil by fluorescence (116), ethanol and specific gravity by distillation curves (117), vegetable oils and fats adulterants by liquid chromatography with UV-visible detection (118), b) in biodiesels, methyl esters content by NIR (119), density, kinematic viscosity, methanol and water content by MIR (120), adulteration with vegetable oil by NIR and MIR (121), c) in gasolines, blending control by NIR (122), adulteration with diesel oil, kerosene, turpentine spirit or thinner by MIR (123), pyrolytic diene values by UV-visible (124), d) fatty acid methyl ester in jet fuel by mass spectrometry (125), e) constituents

of heavy fuel oil by $^{1}$H NMR (126), and f) adulteration of ethanol fuel with methanol by MIR (127).

Additional determinations of industrial relevance are the total acid number (128), quality (129) and sulfur content (130) in crude oil, glucose and ethanol in bioethanol (131), butene in ethylene/propylene/1-butene terpolymers (132), all of them by MIR, saturates, aromatics, resins, and asphaltenes in crude oil by fluorescence (133), nickel and chromium in steel by laser UV (134), and eleven pesticides in agrochemical formulations, a NIR/PLS-based quality control with results comparable to HPLC (135). In a typical real-time NIR/PLS application, the natural antimalarial artemisinin was determined in dried leaves of *Artemisia annua* L. using a hand-held device (136).

Some environmentally related analytical determinations, conducted with the aid of PLS, are the determination of the pesticide mistures parathion-methyl, chlorpyrifos-methyl and vinclozolin, parathion-ethyl, chlorpyrifos and triadimefon, and endosulfan sulfate and carbophenothion by GC data (137), Fe(III), Al(III), and Zr(IV) in water samples by kinetic-potentiometric data ([138]), monitoring of trace elements in estuarine sediments by NIR/MIR (139) and X-ray fluorescence (140), and in water samples by UV-visible spectrophotometry (141), carbamate pesticides in waters by UV-visible data using flow injection analysis (142,143), anions in water using a voltammetric electronic tongue (144) and UV-visible data (145), phenol derivatives in air by fluorescence (146), and the pesticides aminocarb and carbaryl in vegetables and waters by kinetic-spectrophotometric data based on their differential oxidation rates with potassium ferricyanide (147).

### *Process control*

Industrial process control is a relevant field for pharmaceutical and other industrial applications. The classical approach to handling process control through univariate-control

charting techniques, such as the Shewart approach, suffers from several drawbacks. In the presence of complex relationships and correlations among process variables, the simple graphical methods cannot be employed as performance indicators based on independent assumptions. To overcome these problems, PLS has been proposed for optimizing the quality of products produced in process industries with a large number of process and quality variables (148).

Specifically in the pharmaceutical industry, process analytical technology (PAT) was proposed by the United States Food and Drug Administration (US-FDA) (), in order to develop more technically and scientifically rigorous production processes (). NIR spectroscopy is one of the most useful spectroscopic techniques for applying the PAT concepts to the analysis of pharmaceutical products, because of its ability to measure a number of physical and chemical properties of samples, and its suitability for on-line and in-line measurements. Some of the better known uses of NIR in the production of solid pharmaceutical forms include chemical raw material identification, blend uniformity assessment, granulation monitoring, roller compaction monitoring, drying end-point determination and coating end-point and uniformity determinations.

As examples, me may mention the development of a fast and reliable quality control system by building different quantitative calibrations of a pharmaceutical key quality parameters, which allowed to derive a general classification model capturing the over-all product grade (149), and the development of fast and non-invasive quantitative methods for real-time prediction of critical quality attributes of pharmaceutical granulates (active principle content, pH, moisture, flowability, angle of repose and particle size) (150).

*Second-order multivariate calibration examples*

This field has been dominated in the past by matrix data such as chromatography with spectral detection and excitation-emission fluorescence matrix (EEFM) spectrometry. EEFM have a number of advantages: a) they can be easily measured in most modern spectrofluorimeters, and, with the advent of fast-scanning monochromators or multi-wavelegnth charge-coupling detectors, they can be recorded in a very short time, b) fluorescence signals are highly selective and sensitive, allowing to reach low detection limits for the analytes of interest, and c) the specific data structure of EEFM is very simple, since excitation and emission spectra for each sample component do not vary from sample to sample. This latter aspect is important when compared with other second-order signals for which at least one of the component profiles suffer important changes for different experimental runs, such as chromatograms, pH gradients or kinetic profiles. Recent applications in this field involve the development of luminescent signals on solid surfaces, either in batch mode or as detecting system for flow injection analysis (FIA). For example, nylon membranes in batch mode (151-153), molecularly imprinted polymers (154), C18 particles (155), and nylon powder (156,157) as sensing phases for FIA systems. For recent reviews on the subject, see refs. (5,6,158).

Second-order data can also be produced by connecting two instruments, each of which provides a data mode to the finally collected data. This is called hyphenation, and has given rise to a large number of second-order data works. For example, if chromatography is followed by UV-visible detection with a diode array device, a fast-scanning spectrofluorimeter or a mass spectrometer, then second-order data are obtained, in which one data mode is the elution direction and the second data mode is the spectral direction. These data pose some challenges to data processing algorithms, mainly because the elution profiles are not constant from run to run. This demands two alternative solutions: a) correcting for the

temporal changes, by synchronizing the chromatograms (159-161), or b) applying a calibration methodology which is able to model the elution profile changes, such as MCR-ALS (162).

Second-order liquid chromatography with diode array detection (LC-DAD) has been profusely employed to analyze complex samples of different origins, mostly in the presence of unexpected interferences, i.e., achieving the second-order advantage. Some recent examples from the prolific research group of Hunan University, China, are the following determinations: eleven antihypertensives in isocratic mode in only ten minutes (163), twelve quinolones in honey (164), atrazine, ametryn and prometryne in soil, river sediment and wastewater (165) and levodopa, carbidopa and methyldopa in human plasma samples (166).

When spectrofluorimetric detection is employed in chromatographic analysis, lower detection limits are possible, as in the following determinations using MCR-ALS as processing algorithm: ten polycyclic aromatic hydrocarbons in waters (167), and the marker pteridins neopterin, biopterin, pterin, xanthopterin and isoxanthopterin in urine samples (168). When determining eight fluoroquinolones in urine, on the other hand, PARAFAC was employed after chromatographic alignment (169). In all of these cases, the second-order advantage had to be exploited, because the test samples contained unexpected interferences.

Two-dimensional GC-GC with FID detection constitutes an additional possibility for second-order data generation, as reported for the quantitative analysis of essential oils in perfume using MCR-ALS (170), and for gasoline analyses using N-PLS (171,172).

## Conclusions

Multivariate calibration is a powerful tool available to the analytical chemist, which allows for the determination of a variety of analytes or properties in a wide range of sample types, from pharmaceuticals to foodstuff, and from fuels to natural waters. The processed data

are mainly of spectroscopic origin, although other techniques are being gradually incorporated, such as chromatography, mass spectrometry and electrochemistry. It is likely that the present trend will continue in the next years, as instrumental advances add new dimensions to the data, and chemometrician uncover new algorithmic data processing techniques.

## Acknowledgments

Universidad Nacional de Rosario, CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas, Project No. PIP 1950), ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica, Project No. PICT-2010-0084) are gratefully acknowledged for financial support.

## References


(1) Danzer, K., & Currie, L.A. (1998) *Pure Appl. Chem*. **70**, 993–1014.

(2) Van der Linden, W.E. (1989) *Pure Appl. Chem*. **61**, 91–95.

(3) Burns, D.A., & Ciurczak, E.W. (2008) *Handbook of near-infrared analysis*, 3rd Ed., Practical Spectroscopy Series, Vol. 35, CRC Press, Boca Raton, USA.

(4) Wold, S., Sjöström, M., & Eriksson, L. (2001) *Chemom. Intell. Lab. Syst*. **58**, 109–130.

(5) Olivieri, A.C. (2012) *Anal. Methods* **4**, 1876–1886.

(6) Escandar, G.M., Faber, N.M., Goicoechea, H.C.; Muñoz de la Peña, A., Olivieri, A.C., & Poppi, R.J. (2007) *Trends Anal. Chem.* **26**, 752–765.

(7) Smilde, A., Bro, R., & Geladi, P. (2004) *Multi-way analysis with applications in the chemical sciences*, John Wiley & Sons, Chichester, UK.

(8) Booksh, K.S., & Kowalski, B.R. (1994) *Anal. Chem*. **66**, 782A–791A.

(9) Haykin, S. (1999) *Neural networks. A comprehensive foundation*, 2nd. Edition, Prentice-Hall, Upper Saddle River, NJ, USA.

(10) Haaland, D.M., & Thomas, E.V. (1988) *Anal. Chem*. **60**, 1193–1202.

(11) Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., De Jong, S., Lewi, P.J., & Smeyers-Verbeke, J. (1997) *Handbook of chemometrics and qualimetrics*, Elsevier, Amsterdam.

(12) Czekaj, T., Wu, W., & Walczak, B. (2005) *J. Chemometrics* **19**, 341–354.

(13) Olivieri, A.C., Escandar, G.M., & Muñoz de la Peña, A. (2011) *Trends Anal. Chem*. **30**, 607–617.

(14) Bro, R. (1997) *Chemom. Intell. Lab. Syst*. **38**, 149–171.

(15) Tauler, R. (1995) *Chemom. Intell. Lab. Syst*. **30**, 133–146.

(16) Tauler, R., Maeder M., & de Juan, A. (2009) *Multiset data analysis: Extended multivariate curve resolution*, in Brown, S., Tauler, R., & Walczak, B. (Eds.), *Comprehensive chemometrics*, Elsevier, Oxford, Vol. 2, pp. 473–505.

(17) Windig, W., & Guilment, J. (1991) *Anal. Chem*. **63**, 1425–1432.

(18) Antunes, M.C., Simao, J.E.J., Duarte, A.C., & Tauler, R. (2002) *Analyst* **127**, 809–817.

(19) Wold, S., Geladi, P., Esbensen, K., & Øhman, J. (1987) *J. Chemometrics* **1**, 41–56.

(20) Bro, R. (1996) *J. Chemometrics* **10**, 47–61.

(21) Öhman, J., Geladi, P., & Wold, S. (1990) *J. Chemometrics* **4**, 79–90.

(22) Olivieri, A.C. (2005) *J. Chemometrics* **19**, 253–265.

(23) Olivieri, A.C., Goicoechea, H.C., & Iñón, F.A. (2004) *Chemom. Intell. Lab. Syst*. **73**, 189–197.

(24) Jaumot, J., Gargallo, R., de Juan, A., & Tauler, R. (2005) *Chemom. Intell. Lab. Syst*. **76**, 101–110.

(25) Gemperline, P.J., & Cash, E. (2003) *Anal. Chem*. **75**, 4236–4243.

(26) Olivieri, A.C., Wu, H.L., & Yu, R.Q. (2009) *Chemom. Intell. Lab. Syst*. **96**, 246–251.

(27) Second Committee Draft of a Recommendation on Protein Measuring Instruments for Cereal Grain and Oil Seeds, International Organization of Legal Metrology, Bureau International de Métrologie Légale, Paris (2010).

(28) Wu, D., He, Y., & Feng, S. (2008) *Anal. Chim. Acta* **610**, 232–242.

(29) Melfsen, A., Hartung, E., & Haeussermann, A. (2012) *J. Near Infrared Spec.* **20**, 477–482.

(30) Oca, M.L., Ortiz, M.C., Sarabia, L.A., Gredilla, A.E., & Delgado, D. (2012) *Talanta* **99**, 558–565.

(31) Regmi, U., Palma, M., & Barroso, C.G. (2012) *Anal. Chim. Acta* **732**, 137–144.

(32) Romera-Fernandez, M., Berrueta, L.A., Garmon-Lobato, S., Gallo, B., Vicente, F., & Moreda, J.M. (2012) *Talanta* **88**, 303–310.

(33) Garde-Cerdán, T., Lorenzo, C., Zalacain, A., Alonso G.L., & Salinas, M.R. (2012) *LWT-Food Sci. Technol*. **46**, 401–405.

(34) Cozzolino, D., Kwiatkowski, M.J., Dambergs, R.G., Cynkar, W.U., Janik, L.J., Skouroumounis, G., & Gishen, M. (2008) *Talanta* **74**, 711–716.

(35) Xie, L., Ye, X., Liu D., & Ying, Y. (2009) *Food Chem.* **114**, 1135–1140.

(36) Ribeiro, J.S., Ferreira, M.M.C., & Salva, T.J.G. (2011) *Talanta* **83**, 1352–1358.

(37) Maggio, R.M., Cerretani, L., Chiavaro, E., Kaufman T.S., & Bendini, A. (2010) *Food Control* **21**, 890–895.

(38) Rohman, A., & Che Mana, Y.B. (2010) *Food Res. Int.* **43**, 886–892.

(39) Haughey, S.A., Graham, S.F., Cancouët, E., & Elliott, C.T. (2012) *Food Chem.* (in press, DOI http://dx.doi.org/10.1016/j.foodchem.2011.11.136).

(40) Ebrahimi-Najafabadi, H., Leardi, R., Oliveri, P., Casolino, M.C., Jalali-Heravi, M., & Lanteri, S. (2012) *Talanta* **99**, 175–179.

(41) Jha, S.N., & Gunasekaran, S. (2010) *J. Food Eng.* **101**, 337–342.

(42) Tito, N.B., Rodemann, T., & Powell, S.M. (2012) *Food Microbiol.* **32**, 431–436.

(43) Tripathi S., & Mishra, H.N. (2009) *Food Control* **20**, 840–846.

(44) Gaspardo, B., Del Zotto, S., Torelli, E., Cividino, S.R., Firrao, G., Della Riccia, G., & Stefanon, B. (2012) *Food Chem.* **135**, 1608–1612.

(45) El-Abassy, R.M., Eravuchira, P.J., Donfack, P., von der Kammer, B., & Materny, A. (2011) *Vib. Spectrosc.* **56**, 3–8.

(46) Zhang, Y., Huang, Y., Zhai, F., Du, R., Liu, Y., & Lai, K. (2012) *Food Chem.* **135**, 845–850.

(47) Lai, K., Zhai, F., Zhang, Y., Wang, X., Rasco, B.A., & Huang, Y. (2011) *Sens. Instrum. Food Qual. Saf.* **5**, 91–96.

(48) Delfino, I., Camerlingo, C., Portaccio, M., Della Ventura, B., Mita, L., Mita, D.G., & Lepore, M. (2011) *Food Chem.* **127**, 735–742.

(49) Boscolo, M., Andrade-Sobrinho, L.G., Lima-Neto, B.S., Franco, D.W. & Castro Ferreira, M.M. (2002) *J. AOAC Int.* **85**, 744–750.

(50) Khani, R., & Shemirani, F. (2012) *Food Anal. Method.* (in press, DOI 10.1007/s12161-012-9449-8).

(51) Zhang, G., Pan, J. (2011) *Spectrochim. Acta A* **78**, 238–242.

(52) Al-Degs, Y.S. (2009) *Food Chem.* **117**, 485–490.

(53) Acebal, C.C., Lista, A.G., & Fernández Band, B.S. (2008) *Food Chem.* **106**, 811–815.

(54) Ni, Y., Wang, Y., & Kokot, S. (2009) *Talanta* **78**, 432–441.

(55) Ni, Y., Xiao, W., & Kokot, S. (2009) *Food Chem.* **113**, 1339–1345.

(56) Cetó, X., Gutiérrez, J.M. , Gutiérrez, M., Céspedes, F., Capdevila, J., Mínguez, S., Jiménez-Jorquera, C., & del Valle, M. (2012) *Anal. Chim. Acta* **732**, 172–179.

(57) Ghosh, A., Tudu, B., Tamuly, P., Bhattacharyya, N., & Bandyopadhyay R. (2012) *Chemom. Intell. Lab. Syst.* **116**, 57–66.

(58) Campos, I., Masot, R., Alcañiz, M., Gil, L., Soto, J., Vivancos, J.L., García-Breijo, E., Labrador, R.H., Barat, J.M., & Martínez-Mañez, R. (2010) *Sens. Actuators B* **149**, 71–78.

(59) Dias, L.G., Peres, A.M., Barcelos, T.P., Sá Morais, J., & Machado, A.A.S.C. (2011) *Sens. Actuators B* **154**, 111–118.

(60) da Silva, G.A., Maretto, D.A., Bolini, H.M.A., Teófilo, R.F., Augusto, F., & Poppi, R.J. (2012) *Food Chem.* **134**, 1673–1681.

(61) Zunin, P., Leardi, R. & Boggia, R. (2009) *J. AOAC Int.* **92**, 1526–1530.

(62) Ruiz-Samblás, C., Marini, F., Cuadros-Rodríguez, L., & González-Casado, A. (2012) *J. Chromatogr. B* (in press, DOI http://dx.doi.org/10.1016/j.jchromb.2012.01.026).

(63) de la Mata-Espinosa, P., Bosque-Sendra, J.M., Bro, R., & Cuadros-Rodriguez, L. (2011) *Talanta* **85**, 177–182.

(64) Monfreda, M., Gobbi, L., & Grippa, A. (2012) *Food Chem.* **134**, 2283–2290.

(65) Rodríguez, N., Ortiz, M.C., L. Sarabia, L., & Gredilla, E. (2010) *Talanta* **81**, 255–264.

(66) Bogomolov, A., Dietrich, S., Boldrini, B., & Kessler, R.W. (2012) *Food Chem*. **134**, 412–418.

(67) Masot, R., Alcañiz, M., Fuentes, A., Schmidt, F.C., Barat, J. M., Gil, L., Baigts, D., Martínez-Máñez, R., & Soto, J. (2010*) Sens. Actuators A* **158**, 217–223.

(68) Ragni, L., Berardinelli, A., Cevoli, C., & Valli, E. (2012) *J. Food Eng.* **111**, 66–72.

(69) Cerretani, L., Maggio, R.M., Barnaba, C., Gallina Toschi, T., & Chiavaro, E. (2011) *Food Chem.* **127**, 1899–1904.

(70) Rodrigues, J.E.A., Erny, G.L., Barros, A.S., Esteves, V.I., Brandão, T., Ferreira, A.A., Cabrita, E., & Gil, A.M. (2010) *Anal. Chim. Acta* **674**, 166–175.

(71) Sikorska, E., Gliszczynska-Swiglo, A., Insinska-Rak, M., Khmelinskii, I., De Keukeleired, D., & Sikorski, M. (2008) *Anal. Chim. Acta* **613**, 207–217.

(72) Nicolaou, N., Xu, Y., & Goodacre, R. (2012) *Anal. Chem.* **84**, 5951–5958.

(73) Luypaert, J., Massart, D.L., & Heyden, Y.V. (2007) *Talanta* **72**, 865–883.

(74) Jamrógiewicz, M. (2012) *J. Pharm. Biomed. Anal*. **66**, 1–10.

(75) Strachan, C.J., Rades, T., Gordon, K.C., & Rantanen, J. (2007) *J. Pharm. Pharmacol*. **59**, 179–192.

(76) U.S. Pharmacopeial Convention, Rockville, MD, Chapter USP-1119 (2012).

(77) Ferreira, M.H., Braga, J.W.B., & Sena, M.M. (2012) *Microchem. J.* (in press, DOI http://dx.doi.org/10.1016/j.microc.2012.03.008).

(78) El-Gindy, A., Attia, K.A.S., Nassar, M.W., Seda, H.H.A. & Al-Shabrawi, M. (2012) *J. AOAC Int.* **95**, 1035–1042.

(79) Qu, N., Zhu, M., Mi, H., Dou, Y., & Ren, Y. (2008) *Spectrochim. Acta A* **70**, 1146–1151.

(80) Silva, M.A.M., Ferreira, M.H., Braga, J.W.B., & Sena, M.M. (2012) *Talanta* **89**, 342–351.

(81) Rosa, S.S., Barata, P.A., Martins, J.M., & Menezes, J.C. (2008) *Talanta* **75**, 725–733.

(82) Li, P., Du, G., Cai, W., & Shao X. *J. Pharm. Biomed. Anal.* (in press, DOI http://dx.doi.org/10.1016/j.jpba.2012.07.013).

(83) Sarraguça, M.C., Cruz, A.V., Amaral, H.R., Costa, P.C., & Lopes, J.A. (2011) *Anal. Bioanal. Chem.* **399**, 2137–2147.

(84) Cascant, M., Kuligowski, J., Garrigues, S., & de la Guardia, M. (2011) *Talanta* **85**, 1721–1729.

(85) Komsta, L., Czarnik-Matusewicz, H., Szostak, R., Gumieniczek, A. Pietraš, R., Skibiński, R. & Inglot, T. (2011) *J. AOAC Int.* **94**, 743–749.

(86) Chen, Z.-P., Fevotte, G., Caillet, A., Littlejohn, D., & Morris, J. (2008) *Anal. Chem.* **80**, 6658–6665.

(87) de Carvalho Rocha, W.F., Sabin, G.P., Março, P.H., & Poppi, R.J. (2011) *Chemom. Intell. Lab. Syst.* **106**, 198–204.

(88) Sabin, G.P., de Carvalho Rocha, W.F., & Poppi, R.J. (2011) *Microchem. J.* **99**, 542–547.

(89) Breitkreitz, M.C., Sabin, G.P., Polla, G., & Poppi, R.J. *J. Pharm. Biomed. Anal.* (in press, DOI http://dx.doi.org/10.1016/j.jpba.2012.03.054).

(90) Khanmohammadia, M., Soleimania, M., Morovvata, F., Bagheri Garmarudia, A., Khalafbeigic, M., & Ghasemia, K. (2012) *Thermochim. Acta* **530**, 128–132.

(91) Abdelwahab, N.S. (2010) *Anal. Methods* **2**, 1994–2001.

(92) Hadad, G.M., El-Gindy, A., & Mahmoud, W.M.M. (2008) *Spectrochim. Acta A* **70**, 655–663.

(93) Mo, A.C., Soponar, F., Medvedovici, A., & Sârbu, C. (2010) *Anal. Lett.* **43**, 804–813.

(94) Khoshayand, M.R., Abdollahi, H., Shariatpanahi, M., Saadatfard, A., & Mohammadia, A. (2008) *Spectrochim. Acta A* **70**, 491–499.

(95) Grünhut, M., Centurión, M.E., & Fernández Band, B.S. (2007) *Anal. Lett.* **40**, 2016–2031.

(96) Dinc¸ E., Baleanuc, D., Ioele, G., De Luca, M., & Ragno, G. (2008) *J. Pharm. Biomed. Anal.* **48**, 1471–1475.

(97) Hadad, G.M., El-Gindy, A. & Mahmoud, W.M.M. (2008) *J. AOAC Int.* **91**, 39–51.

(98) Hadad, G.M., El-Gindy, A. & Mahmoud, W.M.M. (2007) *J. AOAC Int.* **90**, 957–970.

(99) Aglayan, M.G., Palabiyik, I.M. & Onur, F. (2010) *J. AOAC Int.*, **93**, 862–868.

(100) Mohamed, A.M.I., Mohamed, H.A., Mohamed, N.A. & El-Zahery, M.R. (2011) *J. AOAC Int.* **94**, 467–481.

(101) Lotfy, H.M., Alamein, A.M.A. & Hegazy, M.A.M. (2010) *J. AOAC Int.* **93**, 1844–1855.

(102) Ferrante do Amaral, C.E., & Wolf, B. (2008) *Med. Eng. Phys.*, **30**, 541–549.

(103) Chuah, Z.-M., Paramesran, R., Thambiratnam, K., & Poh, S.-C. (2010) *Chemom. Intell. Lab. Syst.* **104**, 347–351.

(104) Li, Q.-B., Li, L.-N., & Zhang, G.-J. (2010) *Infrared Phys. Techn.* **53**, 410–417.

(105) Pickup, J.C., Hussain, F., Evans, N.D., & Sachedina, N. (2005) *Biosens. Bioelectron.* **20**, 1897–1902.

(106) Heise, H.M., Marbach, R., Koschinsky, Th., & Gries, F.A. (1994) *Appl. Spectrosc.* **48**, 85–95.

(107) Perez-Guaita, D., Ventura-Gayete, J., Pérez-Rambla, C., Sancho-Andreu, M., Garrigues, S., & de la Guardia, M. (2012) *Anal. Bioanal. Chem.* **404**, 649–656.

(108) Liu, L., & Arnold. M.A. (2009) *Anal. Bioanal. Chem.* **393**, 669–677.

(109) Kasemsumran, S., Du, Y.P., Murayama, K., Huehn, M., & Ozaki, Y. (2004) *Anal. Chim. Acta* **512**, 223–230.

(110) de Oliveira Neves, A.C., de Araújo, A.A., Silva, B.L., Valderrama, P., Henrique Março, P., & Gomes de Lima, K.M. (2012) *J. Pharm. Biomed. Anal.* **66**, 252–257.

(111) Yang, Y., Tu, J., Cai, W., & Shao, X. *Talanta* (in press, DOI http://dx.doi.org/10.1016/j.talanta.2012.07.049).

(112) Duarte, J., Pacheco, M.T.T., Villaverde, A.B., Machado, R.Z., Zângaro, R.A., & Silveira, L. (2010) *J. Biomed. Opt.* **15**, 047002–047007.

(113) Boschetti, C.E., & Olivieri, A.C. (2004) *J. NIR Spectrosc.* **12**, 85–91.

(114) Alves, J.C.L., Henriques, C.B., & Poppi, R.J. (2012) *Fuel* **97**, 710–717.

(115) Alves, J.C.L. & Poppi, R.J. (2012) *J. Near Infrared Spec.* **20**, 419–425

(116) Corgozinho, C.N.C., Pasa, V.M.D., & Barbeira, P.J.S. (2008) *Talanta* **76**, 479–484.

(117) Aleme, H.G., & Barbeira, P.J.S. (2012) *Energ. Fuel.* (in press, DOI 10.1021/ef3008757).

(118) Paiva Brandão, L.F., Batista Braga, J.W., & Ziani Suarez, P.A. (2012) *J. Chromatogr. A* **1225,** 150– 157.

(119) Baptista, P., Felizardo, P., Menezes, J.C., & Neiva Correia, M.J. (2008) *Anal. Chim. Acta* **607**, 153–159.

(120) Flôres Ferrão, M., de Souza Viera, M., Panta Pazos, R.E., Fachini, D., Gerbase, A.E., & Marder, L. (2011) *Fuel* **90**, 701–706.

(121) Gaydou, V., Kister, J., & Dupuy, N. (2011) *Chemom. Intell. Lab. Syst.* **106**, 190–197.

(122) Martínez, E., Huertas, S., Ménez, H., Peña, J.L. & Alcalde A. (2008) *J. Near Infrared Spec.* **16**, 297–303.

(123) Teixeira, L.S.G., Oliveira, F.S., dos Santos, H.C., Cordeiro, P.W.L., & Almeida, S.Q. (2008) *Fuel* **87**, 346–352.

(124) Hilgemann, M., Nascimento, P.C., Dias, D., Guterres, M.V., Stringhini, F.M., de Carvalho, L.M., & Bohrer, D. (2011) *J. Petrol. Sci. Eng.* **78**, 283–287.

(125) Eide, I., Neverdal, G., & Westad, F. (2010) *Energ. Fuel.* **24**, 3661–3664.

(126) Nielsen, K.E., Dittmer, J., Malmendal, A., & Nielsen, N.C. (2008) *Energ. Fuel.* **22**, 4070–4076.

(127) Carneiro, H.S.P., Medeiros, A.R.B., Oliveira, F.C.C., Aguiar G.H.M., Rubim, J.C., & Suarez, P.A.Z. (2008) *Energ. Fuel.* **22**, 2767–2770.

(128) Jingyan, L., Xiaoli, C., & Songbai, T. (2012) *Energ. Fuel.* (in press, DOI 10.1021/ef3002372).

(129) Abbas, O., Rebufa, C., Dupuy, N., Permanyer, A., & Kister, J. (2012) *Fuel* **98**, 5–14.

(130) Müller, A.L.H., Sogari Picoloto, R., Azevedo Mello, P., Flores Ferrão, M., Pereira dos Santos, M.F., Lourenço Guimarães, R.C., Müller, E.I., & Moraes Flores, E.M. (2012) *Spectrochim. Acta A* **89**, 82– 87.

(131) Liebmann, B., Friedl, A., & Varmuza, K. (2009) *Anal. Chim. Acta* **642**, 171–178.

(132) Marengo, E., Longo, V., Bobba, M., Robotti, E., Zerbinati, O., & Di Martino, S. (2009) *Talanta* **77**, 1111–1119.

(133) Pantoja, P.A., Lopez-Gejo, J., Le Roux, G.A.C., Quina, F.H., & Nascimento, C.A.O. (2011) *Energ. Fuel.* **25**, 3598–3604.

(134) Barbieri Gonzaga, F., & Pasquini, C. (2012) *Spectrochim. Acta B* **69**, 20–24.

(135) Armenta, S., Garrigues, S. & de la Guardia, M. (2008) *J. Near Infrared Spec.* **16**, 129–137.

(136) Camps, C., Toussirot, M., Quennoz, M. & Simonnet X. (2011) *J. Near Infrared Spec.* **19**, 191–198.

(137) Gil García, M.D., Martínez Galera, M., Pablos Espada, M. del C., Garrido Frenich, A. & Martínez Vidal, J.L. (2000) *J. AOAC Int.* **83**, 1068–1075.

(138) Karimi, M.A., Ardakani, M.M., Ardakani, R.B., Mashhadizadeh, M.H., Zandmonfared, M.R. & Tadayon, M. (2010) *J. AOAC Int.* **93**, 327–334.

(139) Moros, J., Fernández-Ortiz de Vallejuelo, S., Gredilla A., de Diego, A., Madariaga, J.M., Garrigues, S., & de la Guardia, M. (2009) *Environ. Sci. Technol.* **43**, 9314–9320.

(140) Moros, J., Gredilla, A., Fernández-Ortiz de Vallejuelo, S., de Diego, A., Madariaga, J.M., Garrigues, S., & de la Guardia, M. (2010) *Talanta* **82**, 1254–1260.

(141) Ghasemi, J.B., & Zolfonoun, E. (2010) *Talanta* **80**, 1191–1197.

(142) Chu, N., & Fan, S. (2009) *Spectrochim. Acta A* **74**, 1173–1181.

(143) Ostra, M., Ubide, C., & Zuriarrain, J. (2007) *Anal. Chim. Acta* **584**, 228–235.

(144) Winquist, F., Olsson, J., & Eriksson, M. (2011) *Anal. Chim. Acta* **683**, 192–197.

(145) Sarlak, N., & Anizadeh, M. (2011) *Sensors Actuat. B* **160**, 644–649.

(146) Pistonesi, M.F., Di Nezio, M.S., Centurión, M.E., Lista, A.G., Fragoso, W.D., Pontes, M.J.C., Araújo, M.C.U., & Fernández Band, B.S. (2010) *Talanta* **83**, 320–323.

(147) Ni, Y., Xiao, W., & Kokot, S. (2009) *J. Hazard. Mater.*, **168**, 1239–1245.

(148) Song, K., Jang, P.Y., Cho, Y. & Juun, C.H. (2002) *IIE Trans.* **34**,881–890.

(149) Märk, J., Andre, M., Karner, M. & Huck, C.W. (2010) *Eur. J. Pharm. Biopharm.* **76**, 320–327.

(150) Rosas, J.G., Blanco, M., Gonzalez , J.M. & Alcala, M. (2012) *Talanta* **97**, 163–170.

(151) Bortolato, S.A., Arancibia, J.A., & Escandar, G.M. (2008) *Anal. Chim. Acta* **613**, 218–227.

(152) Bortolato, S.A., Arancibia, J.A., & Escandar, G.M. (2008) *Anal. Chem.* **80**, 8276–8286.

(153) Chiarandini, J.P., & Escandar, G.M. (2012) *Anal. Bioanal. Chem.* **402**, 2221–2225.

(154) Valero-Navarro, A., Damiani, P.C., Fernández-Sánchez, J.F., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2009) *Talanta* **78**, 57–65.

(155) Bortolato, S.A., Arancibia, J.A., & Escandar, G.M. (2011) *Environ. Sci. Technol.* **45**, 1513–1520.

(156) Piccirilli, G., & Escandar, G.M. (2007) *Anal. Chim. Acta* **601**, 196–203.

(157) Piccirilli, G., & Escandar, G.M. (2009) *Anal. Chim. Acta* **646**, 90–96.

(158) Andersen, C.M., & Bro, R. (2003) *J. Chemometrics* **17**, 200–215.

(159) Amigo, J.M., Skov, T., & Bro, R. (2010) *Chem. Rev.* **110**, 4582–4605.

(160) Pierce, K.M., Hoggard, J.C., Mohler, R.E., & Synovec, R.E. (2008) *J. Chromatogr. A* **1184**, 341–352.

(161) Skov, T., van den Berg, F., Tomasi, G., & Bro, R. (2006) *J. Chemometrics* **20**, 484–497.

(162) Hantao, L.W., Aleme, H.G., Pedroso, M.P., Sabin, G.P., Poppi, R.J., & Augusto, F. (2012) *Anal. Chim. Acta* **731**, 11–23.

(163) Zhao, J., Wu, H.-L., Niu, J.-F., Yu, Y.-J., Yu, L.-L., Kang, C., Li, Q., Zhang, X.-H., & Yu, R.-Q. (2012) *J. Chromatogr. B* **902**, 96–107.

(164) Yu, Y-J., Wu, H.-L., Shao, S.-Z., Kang, C., Zhao, J., Wang, Y., Zhu, S.-H., & Yu, R.-Q. (2011) *Talanta* **85**, 1549–1559.

(165) Li, Y.-N., Wu, H.-L., Qing, X.-D., Li, Q., Li, S.-F., Fu, H.-Y., Yu, Y.-J., & Yu, R.-Q. (2010) *Anal. Chim. Acta* **678**, 26–33.

(166) Li, S.-F., Wu, H.-L., Yu, Y.-J., Li, Y.-N., Nie, J.-F., Fu, H.-Y., & Yu, R.-Q. (2010) *Talanta* **81**, 805–812.

(167) Bortolato, S.A., Arancibia, J.A., & Escandar, G.M. (2009) *Anal. Chem*. **81**, 8074–8084.

(168) Mancha de Llanos, A., de Zan, M.M., Culzoni, M.J., Espinosa Mansilla, A., Cañada Cañada, F., Muñoz de la Peña, A., & Goicoechea, H.C. (2011) *Anal. Bioanal. Chem*. **399**, 2123–2135.

(169) Cañada Cañada, F., Arancibia, J.A., Escandar, G.M., Ibañez, G.A., Espinosa Mansilla, A., Muñoz de la Peña, A., & Olivieri, A.C. (2009) *J. Chromatogr. A* **1216**, 4868–4876.

(170) de Godoy, L.A.F., Wang, L., Pedroso, M., Poppi, R.J., & Augusto, F. (2011) *Anal. Chim. Acta* **699**, 120–125.

(171) de Godoy, L.A.F., Pedroso, M.P., Ferreira, E.C., Augusto, F. & Poppi, R.J. (2011) *J. Chromatogr. A* **1218**, 1663–1667.

(172) Pedroso, M.P., de Godoy, L.A.F., Ferreira, E.C., Poppi, R.J., & Augusto, F. (2008) *J. Chromatogr. A* **1201**, 176–182.

Table 1. Description of usual algorithms employed for processing second-order data.

| Algorithm | Brief description |
|---|---|
| PARAFAC | Assumes a trillinear model for the data array built with a set of second-order signals, i.e., an element ($i$,$j$,$k$) is the sum of contributions of the form ($a_i \times b_j \times c_k$), where $a_i$ is the relative concentration of a component in the $i$th. sample, and $b_j$, $c_k$ are the values of the instrumental profiles at the $j$th., $k$th. channel in each data mode. Values of $a_i$ are employed for analyte quantitation using a pseudo-univariate calibration graph. |
| MCR-ALS | Places $I$ data matrices (size $J \times K$) on top of each other along one of the data modes, and assumes that the augmented matrix follows a bilinear model, i.e., a matrix element ($m$,$k$) is the sum of contributions of the form ($a_m \times b_k$), where $a_m$ describes the profile for each sample in the augmented mode, and $b_k$ in the common mode ($m$ runs from 1 to $I \times J$). For analyte quantitation, areas under each of the sample profiles in the augmented mode are computed, and employed to build a pseudo-univariate calibration graph. |
| PLS/RBL | After calibrating a PLS model, the interference signals in the test sample are assumed to follow a bilinear model, i.e., an element ($j$,$k$) of the interference matrix is the sum of contributions of the form ($y_j \times z_k$), where $y_j$, $z_k$ are abstract loadings in each mode. Analyte scores are produced by modelling the residuals of the fit of the test sample data array to the bilinear model, hence the name residual bilinearization. In U-PLS, data arrays are unfolded into vectors; in N-PLS, the original data structure is maintained. |

Table 2. Software for first- and second-order calibration.

| Name | Main algorithm(s) | Ref. | Web page |
|---|---|---|---|
| The N-way toolbox | PARAFAC<br>U-PLS<br>N-PLS | 14 | http://www.models.life.ku.dk/algorithms |
| MCR graphical interface | MCR-ALS | 24 | http://www.mcrals.info/ |
| GUIPRO graphical interface | MCR-ALS | 25 | http://personal.ecu.edu/gemperlinep |
| MVC2 graphical interface | PARAFAC<br>U-PLS/RBL<br>N-PLS/RBL | 26 | www.chemometry.com |

## Figure captions

**Figure 1**: Pictorial representation of different data types from zeroth- to second-order, for a single sample and for a set of samples, giving rise to various analytical orders.

**Figure 2**: A set of near-infrared spectra collected for 100 samples of sunflower seeds, which can be employed to build a first-order multivariate model for the determination of fat, humidity, protein and starch. None of the wavelenghts is specific for performing any of these determinations, but the multivariate model is able to provide a suitable answer.

**Figure 3**: Schematic representation of the main second-order algorithms. From a set of data matrices (in this case chromatographic elution time-spectral matrices), PARAFAC builds a three-way array, which is decomposed into spectral and elution profiles of individual samples components (elution profiles are assumed to be identical in all samples if the trilinear model holds). MCR-ALS, on the other hand, builds an augmented data matrix, which is decomposed into spectral and augmented elution profiles of sample components using a bilinear model. The latter better reflects the variability of chromatograms from sample to sample.

**Figure 4**: Three-dimensional surface showing the fluorescence emission intensity of a urinary antibiotic as a function of emission and excitation wavelengths. From these second-order data it is possible to quantitate the concentration of the antibiotic in a biological fluid such as urine, even in the presence of uncalibrated interferences (for

example, the urine background constituents or other pharmaceuticals). The vertical scale is in arbitrary fluorescence intensity units.