INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

ANALYTICAL CHEMISTRY DIVISION*

UNCERTAINTY ESTIMATION AND FIGURES OF MERIT FOR MULTIVARIATE CALIBRATION

(IUPAC Technical Report)

Prepared for publication by
ALEJANDRO C. OLIVIERI[1,‡], NICOLAAS (KLAAS) M. FABER[2], JOAN FERRÉ[3],
RICARD BOQUÉ[3], JOHN H. KALIVAS[4], AND HOWARD MARK[5]

[1]*Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, Rosario S2002LRK, Argentina;* [2]*Chemometry Consultancy, Rubensstraat 7, 6717 VD Ede, The Netherlands;* [3]*Department of Analytical and Organic Chemistry, Rovira i Virgili University, 43007 Tarragona, Spain;* [4]*Department of Chemistry, Idaho State University, Pocatello, ID 83209, USA;* [5]*Mark Electronics, 69 Jamie Court, Suffern, NY 10901, USA*

‡Corresponding author: E-mail: aolivier@fbioyf.unr.edu.ar

Uncertainty estimation and figures of merit for multivariate calibration

(IUPAC Technical Report)

Abstract: This paper gives an introduction to multivariate calibration from a chemometrics perspective and reviews the various proposals to generalize the well-established univariate methodology to the multivariate domain. Univariate calibration leads to relatively simple models with a sound statistical underpinning. The associated uncertainty estimation and figures of merit are thoroughly covered in several official documents. However, univariate model predictions for unknown samples are only reliable if the signal is sufficiently selective for the analyte of interest. By contrast, multivariate calibration methods may produce valid predictions also from highly unselective data. A case in point is quantification from near-infrared (NIR) spectra. With the ever-increasing sophistication of analytical instruments inevitably comes a suite of multivariate calibration methods, each with its own underlying assumptions and statistical properties. As a result, uncertainty estimation and figures of merit for multivariate calibration methods has become a subject of active research, especially in the field of chemometrics.

Keywords: multivariate calibration; standard error of prediction; prediction interval; figures of merit; limit of detection; results reporting; chemometrics; IUPAC Analytical Chemistry Division.

CONTENTS

1. INTRODUCTION

The International Union of Pure and Applied Chemistry (IUPAC) states that "In general, calibration is an operation that relates an output quantity to an input quantity for a measuring system under given conditions" [1], while according to the *International Vocabulary of Basic and General Terms in Metrology* (*VIM*), calibration is an "operation establishing the relation between quantity values provided by measurement standards and the corresponding indications of a measuring system, carried out under specified conditions and including evaluation of measurement uncertainty" [2]. Unless otherwise mentioned, we will assume in the remainder of this paper that analyte concentration(s) is the output quantity, while the indication of the measuring instrument is the input quantity. Although the common case in analytical chemistry is to have spectral values as inputs [1], signals can be of other types, e.g., electrochemical. In univariate calibration, the concentration of a single analyte is predicted from a single instrumental signal. Provisions concerning uncertainty estimation in univariate calibration are offered by several official documents [3–5], reflecting the fact that they are an important aspect of chemical measurement and that, in a formal sense, an analytical result is not complete until a statement about its uncertainty is formulated. The following compelling statement is taken from an editorial by de Bièvre in *Accreditation and Quality Assurance* [6]: "So, a result without reliability (uncertainty) statement cannot be published or communicated because it is not (yet) a result. I am appealing to my colleagues of all analytical journals not to accept papers anymore which do not respect this simple logic." In the usual single-component calibration scenario, expressions for sample-specific standard error of prediction are available from IUPAC [1], as well as figures of merit, detection capabilities, and results reporting [7].

When deploying a univariate method to predict for unknown samples, one essentially assumes that the signal is highly *selective* for the analyte of interest. According to the latest IUPAC recommendation, "selectivity refers to the extent to which the method can be used to determine particular analytes in mixtures or matrices without interferences from other components of similar behavior" [8]. Consequently, the signal may contain a significant contribution from other components only if this contribution is well approximated by a particularly simple form. In case it is nearly constant, for example, one may effectively eliminate it by a background correction or model it as an intercept (Fig. 1). Traditionally, analytical chemists attempt to generate sufficient selectivity by improving the method, either by *experimental* means (e.g., sample clean-up or chromatographic separation) or by *instrumental* means (e.g., selecting a wavelength channel in a high-resolution instrument). However, the traditional approach may not always work well for complex samples. Method development may also be impractical due to excessive costs. The increasing concern about food safety, clean environment, and fair sports—to mention just a few challenging areas of application—currently leads to an accumulation of analytical problems for which convenient traditional approaches are lacking. Consequently, nontraditional approaches are called for.

Multivariate calibration can be seen as a promising mathematical approach to the ubiquitous selectivity problem. Chemometricians have fostered multivariate calibration methods ever since the foundation of chemometrics as an independent branch of chemistry in the early 1970s. As explained below, multivariate calibration methods aim to construct a valid predictive model on the basis of (possibly highly) unselective multiple predictor variables. Valcárcel et al. [9] summarize the potential utility of chemometrics as follows: "The best selectivity levels can be obtained by applying chemometrics in the various physico-chemical methods for discrimination of analytes." In other words, by (correctly) applying chemometrics one may efficiently extract selective information from unselective data. This immediately raises the question of how one should assess the degree of selectivity of intrinsically unselective multivariate data. Many proposals can be found in the literature; see refs. [10–16] for reviews of the subject. However, it is argued in ref. [17] that the selectivity criterion independently put forth by Lorber [18,19] and Bergmann, von Oepen, and Zinn [20], henceforth referred to as the LBOZ criterion, is the most suitable one for two reasons. First, it relates to prediction uncertainty in a transparent way. Second, it allows for a consistent generalization to so-called multiway data analysis. A consistent gen-
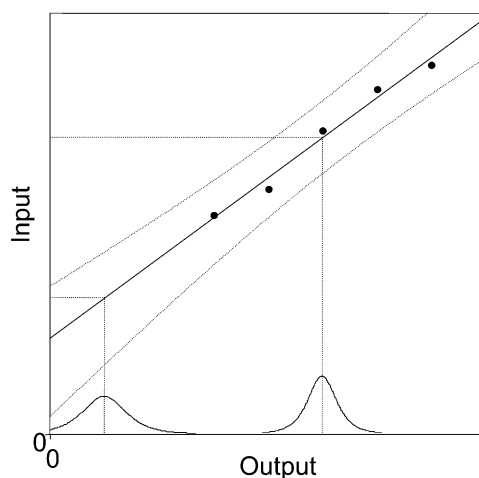
Fig. 1 Prediction of analyte concentration from a straight line fit with intercept. The uncertainty in the straight line (——) fitted to the univariate data (●) is indicated by the familiar 95 %-confidence prediction bands (···). These bands define intervals for individual predictions. Prediction intervals are relatively narrow close to the mean, where the model is most precise. The shape of the confidence prediction bands further implies that extrapolation should be avoided.

eralization is, for example, required when the data arise from hyphenated techniques such as gas chromatography with mass spectral detection (GC–MS or GS–MS–MS).

It stands to reason that, with respect to figures of merit and uncertainty estimation, no conceptual difference should exist between univariate and multivariate calibration methods. Consequently, descriptions of multivariate methods must also include the corresponding estimated figures of merit, and, likewise, reports of analyte concentrations predicted by multivariate methods must include the corresponding uncertainty estimates. However, while there is general consensus on univariate prediction uncertainty expressions derived from basic statistics [1], cf. Fig. 1, their generalization to the multivariate domain has been lacking for a long time. The common practice when applying multivariate calibration strategies leads to a so-called root-mean-square error of prediction (RMSEP), which is an average over a number of samples employed for model testing [21]. An average error, however, characterizes the model rather than individual (unknown) samples in the prediction phase. Hence, although RMSEP is a correct summary statistic for guiding the model selection process (e.g., optimal data pretreatment), it cannot lead to prediction intervals with good coverage probabilities. This rather conspicuous difference in the handling of prediction uncertainty is a serious gap between univariate and multivariate calibration methodologies. We pursue a two-fold purpose with this paper, namely to (1) review the considerable progress that has been made toward closing this gap and (2) provide guidelines for the proper use of this newly developed methodology. In fact, uncertainty estimation is currently a major trend in the chemometrics-oriented literature. It is worth pointing out that this trend was initiated by two seminal papers more than a decade ago [22,23]. In our opinion, much of the currently reviewed material may directly supplement the thoughtful ideas advanced in, for example, the IUPAC guidelines for multicomponent calibration [24] and some American Society for Testing Materials (ASTM) standards [25,26].

The remainder of this paper is organized as follows. In Section 2, we discuss various aspects of multivariate calibration from a chemometrics perspective. This tutorial-style section aims at making the subject more accessible to non-chemometricians, since many of the methods and a great deal of terminology have been developed within chemometrics. Section 3 reviews generally accepted univariate methodology and previously proposed multivariate extensions and gives specific guidelines as to the ex-

perimental parameters required for uncertainty estimation. The discussion in Section 4 of three types of multivariate limit of detection estimators serves to illustrate the potential utility of these extensions. We conclude with Section 5.

2. MULTIVARIATE CALIBRATION FROM A CHEMOMETRICS PERSPECTIVE

2.1 Why multivariate calibration?

Univariate calibration is concerned with the prediction of analyte concentration from a *single* spectral value [1]. Usually, it is synonymous with the term "single-component analysis". As a powerful extension of univariate calibration, multivariate calibration constructs a predictive model on the basis of *multiple* spectral values. For the most common procedures (see below), the predicted analyte concentration is obtained as

$$\hat{c} = \sum_{j=1}^{J} r_j \hat{b}_j = \boldsymbol{r}^{\mathrm{T}} \hat{\boldsymbol{b}} \tag{1}$$

where c denotes the (true) analyte concentration, a "hat" signifies prediction (of a random variable) or estimation (of a parameter), r_j is the signal measured at wavelength j ($j = 1,\dots,J$), b_j is the (true) model parameter associated with wavelength j, $\boldsymbol{r}$ is the $J \times 1$ prediction sample spectrum with elements r_j ($j = 1,\dots,J$), the superscript "T" signifies the transpose of a matrix or vector, and $\boldsymbol{b}$ is the $J \times 1$ (true) model parameter vector with elements b_j ($j = 1,\dots,J$). The estimate of $\boldsymbol{b}$ is obtained from calibration data, similar to fitting a straight line. Measuring a multivariate signal enables one to compensate for varying contributions of interferences in the prediction sample; see Fig. 2. It is this compensating aspect that allows one to conduct a truly quantitative analysis of intrinsically unselective multicomponent systems. Algebraically, a "good" model vector copes with varying interferences by being approximately *orthogonal* to their pure spectra. This near-orthogonality with respect to interfering signals implies that the inner product on the right-hand side of eq. 1 is primarily dominated by the analyte spectrum. Multicomponent analysis (MCA) is a topic of central importance to the analytical community [24]. It is therefore expected that the sophistication of instrumental techniques will continue to increase, leading to a wealth of multivariate data (and suitable methods for their analysis).
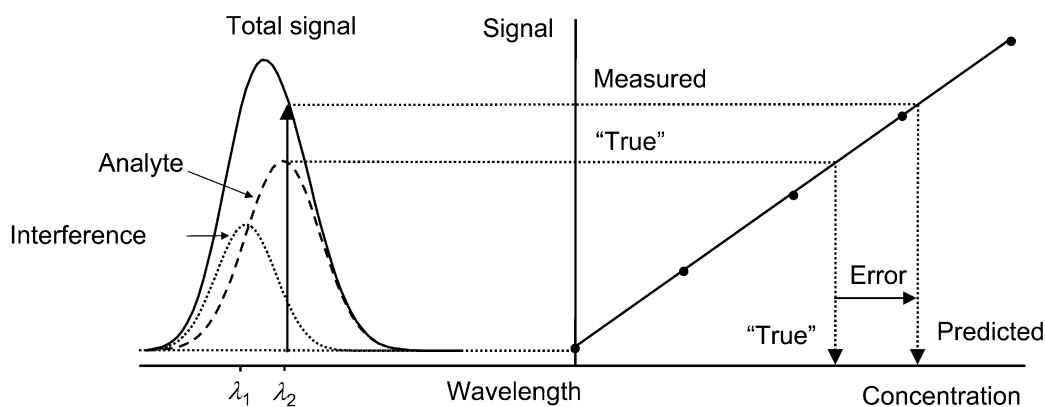
Fig. 2 Illustration of how a univariate model will lead to severely biased predictions when unsuspected interferences give a variable contribution to the signal, whereas multiple measurements may permit accurate prediction in such a situation. To adequately model interferences, one must measure at least as many signals as there are independently varying sources of variation in the spectra. Consequently, measurements at channels λ_1 and λ_2 are sufficient for this simple two-component system. The resulting unit-concentration "spectra" for analyte and interference (in arbitrary units) are $s_a = (2.5, 4.2)$ and $s_i = (2.8, 1.1)$, respectively. A "good" model vector for the analyte is $b = (-0.12, 0.31)$, since applying the prediction eq. 1 to the pure spectra of analyte and interference would approximately yield unity [= $s_a^T b = 2.5 \times (-0.12) + 4.2 \times 0.31$] and zero [= $s_i^T b = 2.8 \times (-0.12) + 1.1 \times 0.31$], respectively. It is seen that the model (b) compensates for the varying contribution of the interference through a negative coefficient for λ_1. The symbols are explained in the text.

2.2 Classical vs. inverse model

The most common calibration procedures belong to either one of two classes, namely, depending on whether the problem is cast into the *classical* or *inverse* model formulation. In the univariate classical model, a spectral value (r) is modeled as a function of analyte concentration (c), i.e., $r = f(c)$. Assuming a linear additive signal for the various constituents (Beer–Lambert) leads to the multivariate classical model for the prediction sample

$$r = c_1 s_1 + \cdots + c_K s_K + e = Sc + e \qquad (2)$$

where r is as defined before, the c_k's ($k = 1,\ldots,K$) denote the concentrations of the K constituents, the $J \times 1$ s_k's ($k = 1,\ldots,K$) are the pure-component spectra at unit concentration, S ($J \times K$) is the spectral matrix that has the s_k's as columns, c ($K \times 1$) is the vector that holds the c_k's ($k = 1,\ldots,K$), and e is a $J \times 1$ vector of noise. (We use matrix-vector notation to avoid excessive use of indices.) Equation 2 is illustrated in Fig. 3 for a simulated three-component mixture. The concentrations of all K constituents are obtained by fitting the pure-component spectra (S) to the spectrum of the prediction sample (r). Usually, an ordinary least-squares (OLS) fit is employed:

$$\hat{c} = S^+ r \qquad (3)$$

where the superscript "+" denotes the pseudo-inverse operation. The pseudo-inverse matrix is the generalization of the regular inverse of a square, full rank matrix to non-square, possibly rank-deficient matrices. Above all, it allows one to solve overdetermined systems of equations ($K < J$). The solution eq. 3 is known as classical least-squares (CLS) or MCA. The parameter vector for the analyte of interest ($\hat{b}$) is obtained as the corresponding column of $(S^+)^T$. This is clarified in Fig. 4 for the artificial example of Fig. 3. The scope of the classical model is rather limited, since it requires the spectra of all contributing species to be measured (so-called *direct* mode) or estimated from mixture spectra (so-called *indirect* mode). In other words, the full S must be available. Successful application has been reported for
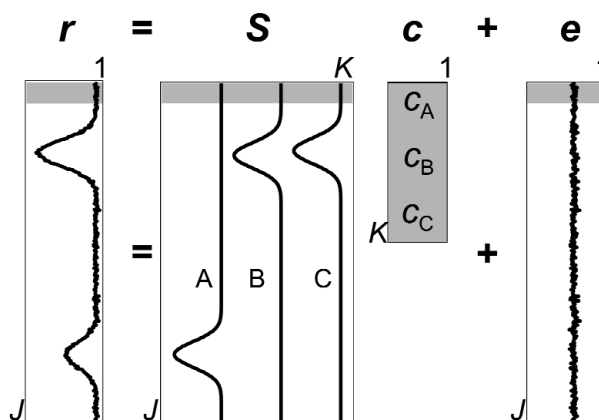
Fig. 3 Representation of eq. 2 for the simulated Gaussians ($K = 3$) adapted from Bergmann et al. [20]. Component A is free from overlap, whereas components B and C strongly interfere. The instrumental signal of the prediction sample is a linear combination of the corresponding elements of the pure-component spectra, superimposed by noise. The concentrations are the weights in the linear combination. The symbols are explained in the text.
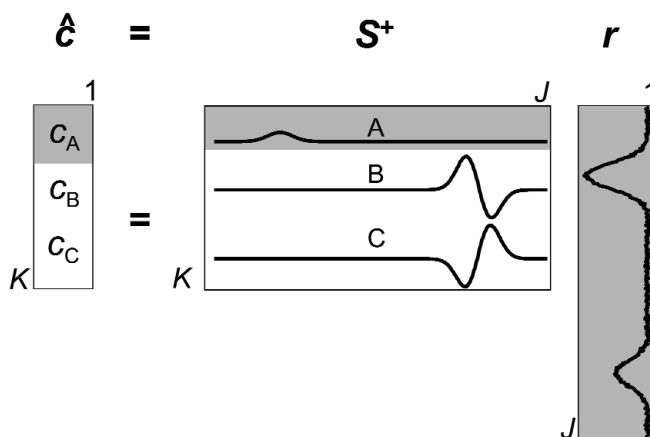
Fig. 4 Representation of eq. 3 for the simulated Gaussians ($K = 3$) adapted from Bergmann et al. [20]. It shows that the prediction of a particular analyte is obtained by taking the inner product of the corresponding row of S^+ and the vector of instrumental signals. The parameter vector of component A is proportional to the pure spectrum, because component A is free from spectral overlap. By contrast, the parameter vectors for components B and C have large negative portions to compensate for the interference. The symbols are explained in the text.

the calibration of *atomic* spectra, e.g., obtained by inductively coupled plasma–optical emission spectrometry (ICP–OES) [27–29].

The calibration of molecular spectra, however, usually requires a different approach, simply because identifying all constituents in the prediction sample is impractical or virtually impossible (e.g., petrochemical or biological matrices). Moreover, owing to interactions (matrix effects) the pure-component spectra may no longer adequately describe the spectrum of the test mixture. Inverse models may conveniently solve these problems by treating analyte concentration as a function of spectral values [21]. They are preferred for the quantitation from, for example, Raman and near-infrared (NIR) (transmission, reflection) spectroscopy. An added bonus of inverse modeling has been recently demonstrated: the resulting model better copes with noise in the spectra, hence leading to significantly better predic-

tions when the noise in the spectra is appreciable [30–32]. The multivariate inverse model for the prediction sample is

$$c = \sum_{j=1}^{J} r_j b_j + e_j = \boldsymbol{r}^T \boldsymbol{b} + \boldsymbol{e} \tag{4}$$

where c is the (true) concentration of the analyte of interest, e is noise, and the remaining symbols are as defined before. An estimate for $\boldsymbol{b}$ is obtained from calibration data, which is expressed as

$$c = \boldsymbol{R}b + \boldsymbol{e} \tag{5}$$

where c ($I \times 1$) contains the analyte concentrations for I calibration samples, $\boldsymbol{R}$ ($I \times J$) contains the spectra, and $\boldsymbol{e}$ is an $I \times 1$ vector of noise. Note that interferences do not play an explicit role in eq. 5, which enables one to calibrate for single analytes—a major advantage over the classical model eq. 2. This is the reason why the inverse and classical models are also known as *partial* and *total* calibration, respectively. Often, the number of spectral values exceeds the number of calibration samples, i.e., $I < J$. As a result, the calibration step requires solving an underdetermined system of equations to eliminate $\boldsymbol{R}$ from the right-hand side of eq. 5. Consequently, OLS cannot be used for estimating the regression vector, unless the number of predictor variables is reduced. The combination of variable subset selection (to obtain an overdetermined system of equations) and OLS is known as inverse least-squares (ILS). Many algorithms exist for selection of the best subset, e.g., generalized simulated annealing, artificial neural networks (ANNs), and genetic algorithms [33–35]. To avoid picking up variables by chance, the final choice is best supported either by specific spectroscopic knowledge or by testing the predictive ability of models calculated with different spectral subsets. Owing to the large number of possible subsets, this approach is often cumbersome. Hence, a popular data-driven alternative is to calculate "scores" or "surrogate variables" as linear combinations of the original predictors and perform the regression step on a limited number of these scores ($F < I$). This leads to the so-called "full-spectrum" methods of which principal component regression (PCR) and partial least-squares regression (PLSR) are the prime examples. As an aside, it is noted that score-based methods can be applied in the classical mode too. This has already been demonstrated by Frank et al. [36] for PLSR in connection with the generalized standard addition method (GSAM). A successful combination of CLS and PLSR has recently been reported as well [37].

2.3 Framework for calibration

Instrumental data come in varying complexity. Sanchez and Kowalski [38–40] proposed to classify them according to tensor algebra. For our purpose, it is convenient to view a tensor as an object that holds data, see Fig. 5. In a loose sense, the order of a tensor equals the minimum number of indices required to organize the data in a meaningful way. Scalars and vectors are zeroth- and first-order tensors, respectively. A (mathematical) matrix for which the elements obey certain (meaningful) relationships is a second-order tensor, and so on. The framework of Sanchez and Kowalski is general in that it allows for any number of indices, although in practice the data complexity will be second-order at most. An added bonus of this classification is that it also pertains to the instruments that deliver the data, as well as the methods for their analysis. Figure 5 further clarifies that multivariate calibration covers data structures with multiple sets of predictor variables. In other words, multiway calibration is a subdivision of multivariate calibration. It is important to make this distinction, because often the terms "multivariate" and "first-order" are used interchangeably. Here, we will use the term "multivariate" to denote all data structures of higher complexity than scalars.
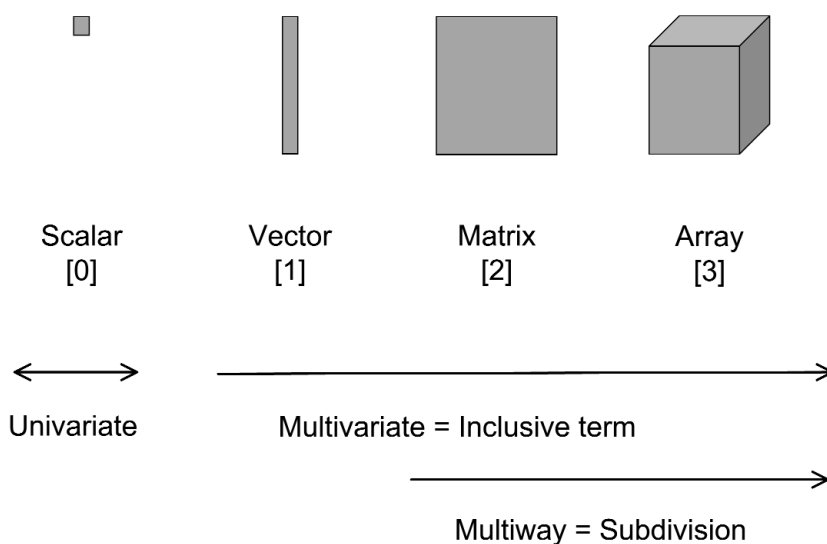
Fig. 5 Representation of tensors with the associated order given in parentheses. The order of a tensor should not be confused with, for example, the order of the familiar polynomial model based on scalars (zeroth-order tensors) or the order of the derivative of spectra (first-order tensors). It is seen that the term "multivariate" covers data structures of increasing complexity.

It should be noted that the term "order", as employed in the present paper, is closely related to the term "dimensionality of analytical experiments", as described in a previous IUPAC Technical Report [41], although the meaning of the latter term appears to be more general. In the framework of ref. [41], thus, zeroth-order calibration is equivalent to a one-dimensional analytical experiment, first-order calibration to a two-dimensional analytical experiment, etc.

2.3.1 Zeroth-order (univariate) calibration

Consider an instrument that yields a single ultraviolet (UV) absorbance, or a single electrochemical potential. In other words, the data (datum) produced for a single chemical sample consist(s) of a scalar. A scalar is a zeroth-order tensor, since an index is superfluous. Hence, the UV instrument can be classified as a zeroth-order instrument, and straight-line fitting using OLS can be viewed as a zeroth-order method. The number of zeroth-order methods is extremely limited. A major drawback of zeroth-order calibration is its inability to handle varying amounts of spectrally active interferences. Perhaps even more disconcerting is the fact that this situation cannot be diagnosed because there is simply no way to break up the various contributions to a zeroth-order signal, cf. Fig. 2.

2.3.2 First-order (multivariate) calibration

An NIR spectrometer produces spectra (Fig. 6). Now a single variable index suffices to organize the predictor data (wavelength axis). Hence, the instrument is first-order, and standard (i.e., linear) PCR and PLSR are first-order methods. The number of first-order methods already greatly exceeds the number of zeroth-order methods. However, one may still keep an overview by classifying a first-order method as solving either the classical or inverse model. First-order methods can cope with varying amounts of spectrally active interferences (if they are sufficiently represented in the calibration phase). In addition, first-order data enable one to diagnose a sample containing unsuspected spectral interferences as an *outlier*, because its spectrum will fit significantly worse to the model than the calibration spectra. These properties have been aptly referred to as the *first-order advantage* [14]. However, it should be noticed that although outliers can be detected, predictions cannot be corrected for the bias caused by the unsuspected interferences.
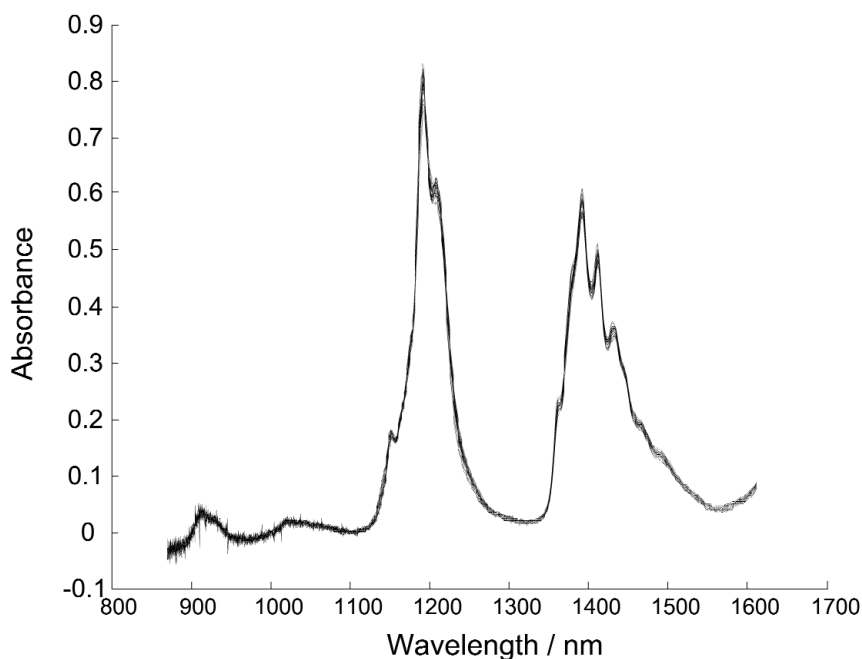
Fig. 6 NIR spectra of 29 gasoline samples.

2.3.3 Second-order (multivariate) calibration and beyond

If the order of the predictor data exceeds unity, one enters the field that is traditionally known as multiway analysis. Figure 5 illustrates that multiway analysis can be viewed as the natural extension of first-order analysis. With increasing order of the predictor data, the number of possibilities for data analysis seems to explode. This can be understood as follows. Second-order data are usually recorded in two ways, namely, from two "hyphenated" first-order instruments (e.g., GC–GC, GC–MS, MS–MS, etc.) or from a single second-order instrument (e.g., a spectrofluorometer capable of registering excitation–emission matrices (EEMs) or a diode-array spectrophotometer where a chemical reaction takes place). The predictor data for a single chemical sample fill a (mathematical) matrix (rows × columns), which can be visualized as a two-dimensional surface or "landscape" (Fig. 7). For the proper selection of a higher-order calibration method, it is of considerable interest to know whether special relationships exist between the various orders. For example, GC–MS produces a (second-order) data matrix for a *pure component* that, in the absence of noise, can be expressed as an outer product of two vectors, i.e., $M = xy^T$ where x and y denote the GC profile and mass spectrum, respectively. The outer product concisely describes that, ideally, the mass spectrum does not depend on elution time and, likewise, the elution profile is identical for all masses. This kind of data has been termed *bilinear* (linear in one variable when the other one is fixed and vice versa) to distinguish it from, for example, MS–MS data. MS–MS data are not bilinear, because a daughter spectrum depends on its position in the parent spectrum. Consequently, an MS–MS data matrix cannot be described, in the absence of noise, as $M = xy^T$. Kiers and Smilde [42] have correctly pointed out the existence of a more rigorous terminology, namely, complexity-one and mixed complexity for bilinear and non-bilinear, respectively. This terminology reflects whether, in the absence of noise, a pure-component data matrix has mathematical rank unity (complexity-one) or higher (mixed complexity). However, we will use the term "bilinear" here, because, in our opinion, it does not lead to confusion and also because it has been adopted in other authoritative texts [43–45]. Second-order bilinear calibration methods are of considerable interest since many instruments produce data that, ideally, follow the bilinear model [46–48]. Suitable methods for analyzing
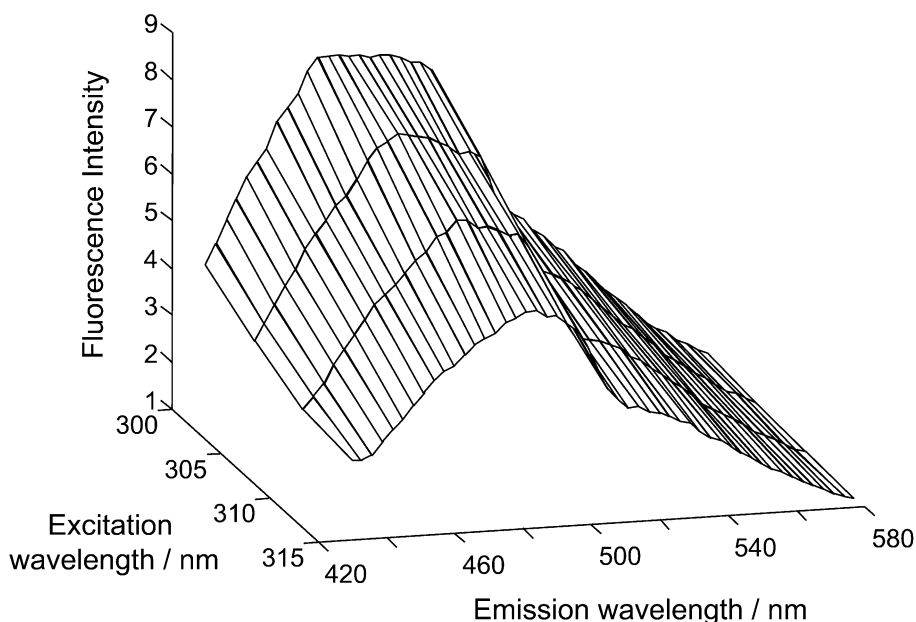
Fig. 7 EEM fluorescence landscape of a seawater sample. The vertical axis is in arbitrary fluorescence units (FU). (Reproduced by permission of The Royal Society of Chemistry.)

bilinear data are, for example, parallel factor analysis (PARAFAC) [49], the generalized rank annihilation method (GRAM) [38], bilinear least-squares (BLLS) with calibration using pure standards [50] or mixtures [44,45], and multivariate curve resolution–alternating least-squares (MCR–ALS) [51]. Since MS–MS data are not bilinear, other methods should be used for calibration [42,52,53]. Bilinear data structures are of prime relevance to the analysis of complex mixtures, because they permit quantification even in the presence of unsuspected sample constituents. This property of second-order bilinear data is known as the *second-order advantage* [14]. It is important to note that the second-order advantage is already obtained using a *single calibration sample* and methods such as PARAFAC [49] or GRAM [38]. (This calibration sample is conveniently obtained by a single addition of all analytes of interest [38].) In this way, one avoids a major drawback of first-order methods, namely, the requirement of a large and diverse calibration set. This practical benefit can hardly be overstated. In case no relationship between the various orders can be assumed, the alternative is to rearrange the higher-order data arrays into vectors, which are first-order tensors, and then apply a first-order method. This procedure leads, for example, to unfold-PCR and unfold-PLSR [54]. A promising alternative to these unfolding methods is N-PLS [55], which is a genuine N-way method. It has been shown that N-PLS leads to simpler models [55] and more stable predictions when the noise in the predictor variables is high [55–57]. However, for low-noise predictor data, the "unfolding" methods may be preferable when focus is on prediction [56]. In the multivariate situation, one may therefore sometimes benefit from simultaneously applying competing methods to perform the various tasks. Booksh and Kowalski [14] have given a comprehensive overview of what can be achieved depending on the order and inherent structure of the data. A summary, adapted from ref. [14], is given in Table 1.

Table 1 Characteristics of calibration of instruments of increasing complexity.

Order	Analytes	Standards	Unsuspected interferences	Statistics
0[a]	1	≥1	Not diagnosed; prediction bias	Fairly simple; well developed
1[b]	>1	≥1 Per species	Diagnosed; prediction bias	Relatively complex; under active research
2[c]	>1	≥1	Diagnosed; no prediction bias	Very complex; poorly developed

[a]Linear model.
[b]Linear additive model.
[c]Bilinear model.

3. UNCERTAINTY ESTIMATION AND FIGURES OF MERIT

3.1 Accepted methodology in univariate calibration

Guidelines for univariate calibration are discussed in detail in several official documents [3–5]. They include provisions that clearly express that all laboratories should possess and apply procedures enabling the estimation of the uncertainty of measurement for all calibrations [3]. More importantly, official regulations state that when estimating this uncertainty of measurement, all error sources which are of importance in a given situation should be taken into account using accepted methods of analysis. All this highlights the need of estimating standard errors for predicted concentrations as part of the chemical measurement process. The following is a summary of accepted univariate methodology, occasionally supplemented with new developments:

- *Standard error of prediction.* For the common single-component calibration performed by straight-line fitting using OLS, expressions for the standard errors in the regression parameters and also in the predicted analyte concentrations are available in the official literature [1]. Underlying these expressions is the assumption that the uncertainty in the concentration or reference values is negligible during model estimation. The resulting standard errors of prediction are sample-specific, since the model is relatively more precise for samples close to the center of the calibration standards. Recently, model estimation has been considered with non-negligible error in the concentration or reference values, employing the bivariate least-squares (BLS) method [58]. For a refinement of this work, see ref. [59].

- *Prediction interval.* Expressions for the prediction interval are available in the IUPAC literature [1]. It follows from the application of *t* statistics and the estimated standard error of prediction as

$$\hat{c} - t_{v,\alpha/2} \cdot \hat{\sigma}(\hat{c} - c) \le c \le \hat{c} + t_{v,\alpha/2} \cdot \hat{\sigma}(\hat{c} - c) \tag{6}$$

  where $t_{v,\alpha/2}$ is the upper $\alpha/2$-percentage point of the *t*-distribution with v degrees of freedom and $\sigma(\hat{c} - c)$ denotes the standard error of prediction (square root of the variance of the prediction error $\hat{c} - c$). Prediction intervals for BLS are considered in [58].

- *Figures of merit.* Figures of merit for the univariate case are thoroughly covered by the IUPAC literature [1]:
  - *Sensitivity.* Sensitivity for a given analyte is defined as the slope of the analytical calibration curve [60,61]. Its dimensions are (signal units × concentration^{-1}). Although the sensitivity may vary with concentration, it is usually constant at low concentration values, where the calibration curve becomes a straight line, cf. Fig. 1. In general terms, it is sensible to consider sensitivity as the change in the (net) response of the instrument divided by the corresponding change in the stimulus (the concentration of the analyte of interest). A method

is said to be sensitive if a small change in concentration of the analyte causes a large change in the (net) response.

- *Selectivity.* A pairwise selectivity index has been defined as the ratio of the slopes of the calibrations lines of the analyte of interest and a particular interference [62], i.e.,

$$\xi_{i,a} = \frac{s_a}{s_i} \tag{7}$$

where s_a and s_i denote the sensitivities of analyte and interference, respectively. Ideally, the selectivity indices should be evaluated for each important interference likely to be present in varying amounts, since this will lead to biased predictions. In Fig. 2, the pairwise comparison of sensitivities at λ_2 yields $\xi_{i,a} = 4.2/1.1 = 3.8$. In other words, the measurement at λ_2 is about four times more sensitive to the analyte than the interference. Interestingly, selectivity coefficients have been mainly evaluated for ion-selective electrodes [63–65].

- *Signal-to-noise ratio.* It is the ratio of the useful analytical signal to the background noise, with the latter one identified as a measure of the statistical fluctuations in a blank signal.

- *Analytical sensitivity.* This parameter is defined as the ratio between sensitivity and instrumental noise. It appears to be more useful than "plain" sensitivity, because it is independent of the specific technique, equipment, and scale employed [66,67]. This parameter has unit concentration^{-1}, and its reciprocal value defines the minimum concentration difference that can be appreciated across the linear range of the employed technique.

- *Limit of detection.* The detection capability is considered by IUPAC as a fundamental performance characteristic of a chemical measurement process [7]. In the new IUPAC and International Organization for Standardization (ISO) documents, detection limits (minimum detectable amounts) are derived from the theory of hypothesis testing and the probabilities of false positive (α), and false negative (β) [7,68,69]. For an illustration of the relevant concepts, see Fig. 8. Its computation has been described in [70]. For a recent review, see ref. [71]. With non-negligible error in the concentration or reference values, BLS may substantially improve the more common OLS and weighted least-squares (WLS) procedures, see ref. [72].

- *Limit of discrimination.* It is the minimum increase in analyte concentration in a prediction sample which ensures that the analytical signal is significantly different from that corresponding to the original analyte concentration [73]. It is intimately related to the analytical sensitivity and may be considered as a generalization of the concept of limit of detection.

- *Results reporting.* IUPAC recommends to always report both the predicted value and its uncertainty, even in case of a decision "not detected". Otherwise, there is needless loss of information, and, of course, the impossibility of averaging a series of results.
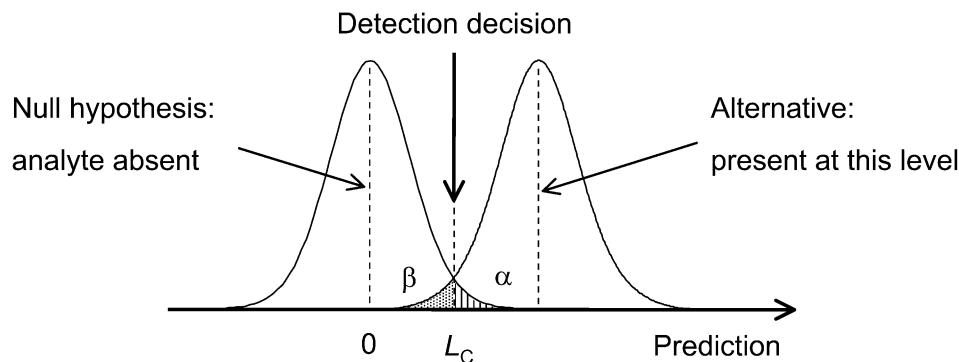
Fig. 8 The limit of detection is the analyte level that with sufficiently high probability $(1 - \beta)$ will lead to a *correct* positive detection decision. The detection decision amounts to comparing the prediction $\hat{c}$ with the critical level (L_c). This level is estimated to allow for a positive detection decision with probability α when, in reality, the analyte is absent. The critical level is only determined by the distribution of the prediction under the null hypothesis (H_0: not present). By contrast, the limit of detection is also determined by the distribution of $\hat{c}$ under the alternative hypothesis (H_A: present at certain level). The probabilities of false positive and false negative detection decisions are obtained by integrating the probability density function of $\hat{c}$ found under the null-hypothesis H_0 (right tail) and alternative hypothesis H_A (left tail), respectively, i.e., $\alpha = \int\limits_{L_c}^{\infty} p\left(\hat{c}|H_0\right) d\hat{c}|H_0$ and $\beta = \int\limits_{-\infty}^{L_c} p\left(\hat{c}|H_A\right) d\hat{c}|H_A$.

3.2 Previously proposed methodology in multivariate calibration

Table 1 shows that the information content increases dramatically when moving from a simple univariate (zeroth-order) model to its complex multivariate (first- and higher-order) counterparts. This increase is primarily of a *qualitative* nature in the sense that more difficult analytical problems can be solved using multivariate models. (In addition, an increasing volume of data will also lead to better noise averaging, which is a favorable *quantitative* aspect.) Because of the intrinsically different information content, one should not expect a straightforward generalization of all univariate quantities listed in the previous section. Surprisingly, only an equally intuitive definition of multivariate selectivity appears to be problematic. The following is a summary of previous proposals:

• *Standard error of prediction.* A recent review on calibration [74] states that there are two basic ways of estimating uncertainty, namely, error propagation [75], which is also employed in the univariate case for BLS [58], or resampling strategies, such as jackknife or bootstrap [76]. We primarily focus on error propagation because it leads to closed-form expressions. These expressions have been shown to offer many benefits, e.g., they
  - are highly convenient for estimating detection and discrimination limits,
  - give fundamental insight into the statistical properties of modeling procedures [77,78],
  - provide the rationale for spectral pretreatment, i.e., "desensitizing" calibration models for artifacts in the data such as wavelength instability or baseline shifts [79],
  - motivate calibration sample selection procedures, such as "local centering" [80], and
  - enable an in-depth discussion of the utility of sensor selection [81].

By contrast, resampling is essentially a "black box" approach, which, however, is often more accurate because fewer assumptions and approximations are made. Consequently, resampling can be extremely useful for testing the adequacy of formulas resulting from error propagation. A literature survey shows that deriving formulas using the method of error propagation has been a major research topic. When employing first-order multivariate data, most publications are concerned with standard (i.e., linear) PLSR [22,23,82–101] (the parameter results derived in refs. [102–104] lead directly to a formula

for standard error of prediction), although some work deals with alternative methods such as PCR [22,56,87,90,105,106], ILS [22], ANNs [107], nonlinear PLSR [108], CLS [18,27,109], and GSAM [110,111]. Considerable progress has been achieved for some higher-order methods, namely, rank annihilation factor analysis (RAFA) [112–114], GRAM [109,115–120], BLLS with calibration using pure standards [50] and mixtures (as well as some alternatives) [44,45], N-PLS [56,121,122], and PARAFAC [123,124]. It is noted that the expressions proposed for N-PLS and PARAFAC are rather crude approximations that can likely be refined using the parameter results derived in [125,126].

A general mathematical expression which has been employed to estimate standard errors $s(c)$ in predicted concentrations by several first- [90] and second-order [124] multivariate models is

$$\left[s(c)\right]^2 = h s_c^2 + h(s_r/S_n)^2 + (s_r/S_n)^2 \tag{8}$$

where h is the leverage, a quantity which places the test sample relative to the calibration space [21], s_c is the standard error in calibration concentrations, S_n is the multivariate sensitivity appropriate for analyte n (see below) and s_r is the instrumental noise level. Equation 8 generalizes the univariate expression to the multivariate context, and concisely describes the error propagated from three uncertainty sources to the concentration standard error: calibration concentration errors (first term on the right-hand side), errors in calibration instrumental signals (second term), and errors in test sample signals (third term). However, for certain methodologies, eq. 8 is only an approximation (see above), although it is likely that analogous, both sample- and analyte-specific expressions will be derived in the future.

- *Prediction interval.* As in univariate calibration, prediction intervals can be constructed from the estimated standard error of prediction (square root of a variance estimate) and the relevant t statistics. However, an important issue should be dealt with in the multivariate case. In the official literature concerning univariate calibration, it is usually assumed that only the instrument signal carries an uncertainty, although a number of works have emphasized the importance of taking into account the errors in both axes when studying a single component [72,127–131]. This leads to exact variance expressions and exact t statistics. By contrast, multivariate models are often constructed using calibration concentrations that are not error-free (see eq. 8). As a result, the variance expressions obtained using the method of error propagation are only approximate, not exact, likewise the assumed distributional properties of the test statistic, i.e., the ratio of prediction error and estimated standard error of prediction

$$\rho = \frac{\hat{c} - c}{\hat{\sigma}(\hat{c} - c)} \tag{9}$$

  where the symbols are defined as in eq. 6. However, an appeal to the central limit theorem helps to justify the required normality assumption for the prediction error (numerator) when the number of spectral values (J) is large: For large J, the prediction $\hat{c}$, hence its error, will be approximately normally distributed no matter the distribution of the individual $r_j \hat{b}_j$ values in eq. 1. Furthermore, the multivariate standard error of prediction (denominator) may have contributions from different sources. A direct consequence of this is that the required χ^2-distribution is only approximately obeyed and the number of degrees of freedom should be established as a compromise between the degrees of freedom corresponding to each error source. It has been proposed in ref. [90] to calculate an overall number of degrees of freedom using the well-known procedure of Satterthwaite [132].

- *Figures of merit.* Figures of merit have become very important in order to characterize, compare, and develop new multivariate methods. Many of them are closely related to the concept of multivariate net analyte signal (NAS), as introduced by Lorber [18]. (Very recently, Brown [32] has noted that "Although rarely cited, Morgan [133] had actually discussed a similar concept some years earlier, giving expressions similar to those provided by Lorber, although readers should note

that Morgan's work contains some errors.") The NAS concept arises quite naturally in multi-variate calibration from the fact that a prediction sample spectrum may have varying contributions from other sample components. Hence, it is logical to decompose the spectrum in two orthogonal parts: a part that can be uniquely assigned to the analyte of interest (the NAS), and the remaining part that contains the (possibly varying) contribution from other components (Fig. 9). This decomposition is carried out by the various regression algorithms in such a way that the NAS is proportional to the concentration of the analyte of interest. Since the NAS is the only part of the spectrum that is employed for prediction, no information is lost when transforming the vector NAS into a scalar. The natural choice is to take the Euclidean norm (i.e., its length), so that the scalar NAS is obtained as $r^* = \|\boldsymbol{r}^*\|$. In the remainder of the paper, if not otherwise mentioned, the term "NAS" will refer to the scalar quantity. Using the NAS, a multivariate calibration model can be represented in a pseudo-univariate plot [134,135]. It is important to note that this representation is exact, not approximate. (For refinements of the original calculation method, see refs. [136–138].) In other words, the calibration curve is obtained by replacing the measured instrumental signal used in univariate calibration (e.g., absorbance at a single wavelength) by the NAS (r^*). An additional distinction can be made between the classical model, i.e., $r^* = f(c)$, and the inverse model, i.e., $c = f(r^*)$. These models lead to two different regression graphs, which are illustrated in Fig. 10. The figures of merit discussed above for univariate calibration can be derived, often in a very similar way, for multivariate calibration:

- *Sensitivity.* This parameter is defined as the NAS generated by an analyte concentration equal to unity [18]. For a classical model, sensitivity is the slope of the (pseudo-univariate) calibration graph. Conversely, sensitivity is the inverse of the slope of this graph for an inverse model. Hence, in terms of the vector of regression coefficients, the sensitivity is given by $1/\|\boldsymbol{b}\|$. For the consistent generalization to higher-order data, see refs. [124,133,140].

- *Selectivity.* In univariate calibration, selectivity coefficients ($\xi_{a,i}$) are fundamental to know how variable amounts of interferences in the sample can bias the predictions, for example, in the application of ion-selective electrodes. By contrast, interferences can be adequately modeled using multivariate data. This essential difference explains why the numerical assessment of multivariate calibration selectivity has always been approached differently. Numerous criteria have been published in the past. It has been explained in the Introduction that the criterion independently proposed by Lorber [18,19] and Bergmann, von Oepen, and Zinn [20] is (currently considered to be) the most suitable one. The LBOZ selectivity criterion focuses on describing what part of the measured signal remains for quantitation, i.e., the NAS. In the absence of interferences, the NAS would just equal the length of the total analyte signal. A natural criterion for assessing the impact of all spectral interferences *simultaneously* follows as

$$\xi_a = \frac{\left\| \boldsymbol{s}^*_a \right\|}{\left\| \boldsymbol{s}_a \right\|} \tag{10}$$

where s_a denotes the pure analyte spectrum and s^*_a is the associated non-overlapping part, cf. Fig. 9. Since the NAS is only a part of the total analyte signal, it follows that the selectivity is a dimensionless parameter ranging from 0 to 1, and that a lower selectivity is associated with a higher degree of spectral overlap between the analyte and other sample constituents. In Fig. 2, the multivariate measurement at λ_1 and λ_2 yields $s^*_a = 1/\|\boldsymbol{b}\| = 3$ and consequently $\xi_a = \dfrac{3}{4.89} = 0.61$ (see the caption to Fig. 2), hence, the LBOZ selectivity criterion has the simple interpretation that 39 % of the analyte signal is lost due to overlap. Compared to the ideal case without overlap ($\xi_{a \, \mathrm{def}} 1$), increased b-coefficients are required

in eq. 1 to extract the useful signal from the prediction sample spectrum (**r**). Increased b-coefficients, however, lead to more spectral error propagation (compared to the ideal case). Specifically, assuming a spectral error variance $V(r_j)$ for wavelength j ($j = 1,...,J$), yields a contribution $\sum_{j=1}^{J} \hat{b}_j^2 V(r_j)$ to prediction uncertainty. The overlap-induced increase of b-coefficients is further illustrated in Fig. 4: Clearly, component A, which corresponds to the ideal case, will experience much less spectral error propagation than the heavily overlapped components B and C.

It is important to note that in inverse calibration, a problem arises because the pure spectrum for the analyte is usually not available. In that case, a pessimistic selectivity estimate is obtained by replacing the total analyte signal in eq. 10 by the length of the mixture spectrum [139]. For the consistent generalization of eq. 10 to higher-order data, see refs. [124,133,140].

Finally, it is reiterated that the LBOZ selectivity criterion (ξ_a) simultaneously accounts for *all* interferences in the mixture (just like the model itself). A variance-decomposition has been developed as an approach to test for the presence of spectral overlap in a multicomponent system (degree of component-wise selectivity) and simultaneously assess the extent to which concentration estimates may degrade [141]. A result of the process is also identification of which chemical species are involved in respective spectral overlaps. The process can be implemented pairwise as well. Only very recently has larger focus been directed toward quantifying the impact of individual interferences on multivariate analyte predictions, leading to a definition of (pairwise) multivariate selectivity coefficients $\xi_{a,i}$ [142–144]. The radically different standpoint taken in that work may lead to a critical re-examination of multivariate selectivity assessment.

- *Signal-to-noise ratio.* It is defined for first-order data as the ratio between NAS and the background noise [18]. For the consistent generalization to higher-order data, see refs. [124,133,136].
- *Analytical sensitivity.* It is defined for first-order [145–147] and for second-order bilinear calibration [148], using a direct analogy with the univariate parameter.
- *Limit of detection.* Garner and Davidson [149] note in 1988 that "There is currently no generally accepted multivariate model of instrumental signals incorporating detection limit estimators but there are no major reasons why such models cannot now be developed." They continue their discussion of multivariate models with some suggestions and further state that "It is highly recommended that multivariate models and estimators be developed and used. Until this is done, decision and detection limits for multiple-signal instruments may be inappropriately estimated." Boqué and Rius [150] have reviewed the progress made in the multivariate area until 1996. (For a general review of more recent date, see ref. [71].) For considerations of space, we will only list the various contributions. Three illustrative types of estimators are discussed to some detail in Section 4. Several approaches have been advanced in the literature for the first-order classical [18,28,151,152] and inverse models [19,153,154]. (The proposal of [154] is directly applicable to higher-order data.) Typical method-specific approaches for higher-order data can be found in [124,155–162]. (It is noted that the method presented in [155,156] is only applicable for multiple signals that are fully selective for the analyte of interest.) A generally applicable nonparametric approach is to train a neural classifier to optimize the false positive and false negative error rates [163].
- *Limit of discrimination.* The consistent multivariate generalization is developed in ref. [164].

- *Results reporting.* Analyte concentrations predicted by multivariate methods should be reported together with the corresponding estimated standard error.
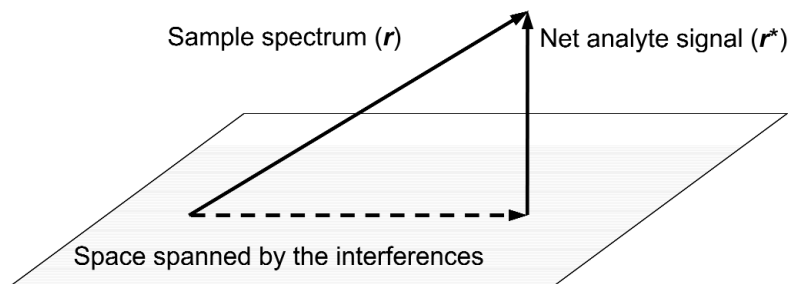

Fig. 9 Representation of the NAS concept. The NAS vector is the part of the sample spectrum that is used for prediction. Hence, it is orthogonal to the space spanned by the spectra of the interferences.
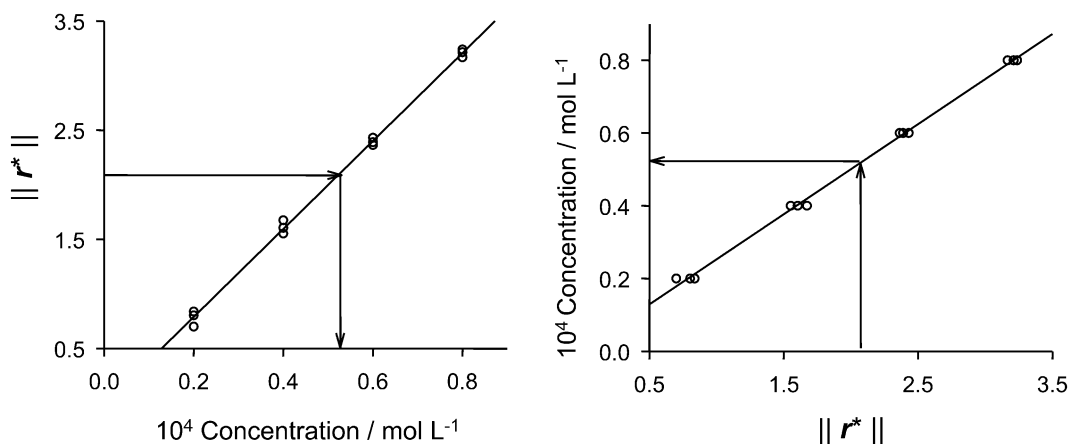

Fig. 10 Pseudo-univariate plot of (a) classical calibration and (b) inverse calibration. The arrow indicates how prediction is calculated from the NAS of a test sample.

3.3 Specific guidelines for estimating prediction errors

When a multivariate model is employed to predict analyte concentrations or reference properties, the estimation of uncertainties in prediction by using analytical expressions requires the knowledge of several parameters. Usually, these expressions reflect the combined effect of the propagation of uncertainties in instrumental signals for the test sample, and also in the model parameters, stemming, in turn, from uncertainties in calibration values (see ref. [90] and refs. therein).

The crudest approximation only involves effects arising from the test sample signals, which are given by the ratio between instrumental noise and sensitivity [90]. The latter parameter is provided by the model as the inverse of the length of the regression coefficients (see above), and thus it is model-dependent. Instrumental noise can be estimated, in turn, in two ways: (1) by sample replication experiments, and (2) from the inspection of the residuals of the prediction step. In the second case, two different scenarios can be envisaged. In the first-order multivariate domain, residuals are comparable to the instrumental noise level if the calibration set of samples is sufficiently representative of the test sample, and can then be inserted in uncertainty formulas. However, if the test sample presents uncalibrated components or spectral artifacts, the residuals will be abnormally large in comparison with the instrumental noise. In this latter case, the sample is classified as an outlier, and the analyte concentra-

tion cannot be predicted by the current model [14]. This constitutes the basis of the excellent outlier detection capabilities of first-order multivariate methodologies.

When second- and higher-order multivariate calibrations exploiting the second-order advantage are employed, interferences that are not present in the calibration standards can in principle be modeled, and prediction residuals will in general reflect the instrumental noise level [44,45,50,56,109,115–126]. This is true provided the data follow the required bi- or tri-linear structure. For methods not showing this advantage, such as unfold-PCR, unfold-PLSR, and multi-way N-PLS regressions, the situation is similar to that discussed above for first-order calibrations, even if the data follow the required structure [54,55].

In any case, the estimated uncertainty in predicted concentrations using the above crude approach provides an idea of the expected standard error for low-leverage samples, i.e., those which are close to the center of the calibration, since the leverage serves to place the test sample relative to the calibration space [21].

The values provided by this simple approach ignore the model uncertainties, and are thus overoptimistic. In order to obtain more realistic values, it is necessary to consider the effect of calibration errors, which stem from two error sources: (1) calibration concentrations, and (2) calibration instrumental signals. Several analytical expressions are known, both for first-order and second-order multivariate methods, which include these model effects (see Section 3.2). In general, they strongly depend on the sample leverage, which should be calculated for each particular model. Concentration errors are also required, which are usually available to experienced analysts from the details in the preparation of the calibration samples, or from the uncertainty in the method employed to determine the reference concentrations or property values. They can also be estimated as the RMSEP on a test set of samples of known analyte concentration. In some cases, the calibration set is split into two subsets, one for calibration and the remaining one for prediction. In this latter case, since calibration samples are sometimes extreme relative to the center of the calibration space, a pessimistic estimate of the concentration error is obtained, and the prediction uncertainty is overestimated. Consequently, it is advisable to build several models, one for prediction and the other one(s) to estimate ingredients of prediction uncertainty [99].

It should be noticed that the presence of unsuspected interferences may constitute an additional source of uncertainty. Usually, a prediction sample with unsuspected interferences will have a prediction uncertainty that is larger than that expected for normal samples, unless a multivariate methodology achieving the second-order advantage is employed. This source of uncertainty should not play a role for normal samples that pass the check on spectral residuals.

4. ILLUSTRATIVE EXAMPLES FROM THE LITERATURE

In this section, we present selected examples of multivariate detection limit estimators. Limit of detection is directly tied to uncertainty estimation, but at very low concentration levels. Concretely, at a concentration level that can be differentiated, statistically, from zero concentration, with preselected probabilities of false detection decisions. Limit of detection estimation therefore constitutes a severe test for any developed uncertainty estimation methodology.

4.1 Formula-based

A rather straightforward approach to estimate the limit of detection is to apply an error propagation-based formula for standard error of prediction to zero concentration level. This was the approach taken by Boqué et al. [153] in the determination of the aromatic content of gasolines by NIR spectroscopy. The selected formula accounts for all sources of errors in the data (signals and concentrations) of calibration and prediction samples [87]. Figure 6 shows the NIR spectra recorded for the calibration set of samples. No part in the spectra is selective of the analyte of interest. Not even trends in the spectra can

be related to an increase/decrease of the analyte concentration. A multivariate limit of detection estimator is clearly justified for this type of data. The limit of detection estimator applied to these data was sample-specific (i.e., dependent on the level of interferences in the sample). The validation with real samples gave promising results; see Table 1 in ref. [153].

Additional selected examples involving the use of error-propagation for estimating concentration uncertainties and limit of detection are the spectrofluorometric determination of the antibiotic tetracycline in human serum [98,145] and the spectrophotometric analysis of bromhexine in a decongestant syrup [98,165]. In the former case, the analyte signal is completely overlapped by the stronger and intrinsically variable fluorescence bands of human serum, see Fig. 11. Therefore, several first-order multivariate models based on PLS regression were employed. A detailed analysis of the various uncertainty sources present in this system led to the conclusion that the dominating factor in estimating prediction uncertainties was the instrumental uncertainty for unknown samples, rather than the calibration modeling parameters [98]. The error propagation approach was compared with Monte Carlo simulations based on noise addition, and also with an empirical approach involving replicate analysis of spiked samples. The concentration uncertainties, estimated by the formula-based error propagation approach, by statistical analysis of Monte Carlo noise addition, and by the empirical model were all in mutual agreement. These three approaches were also successfully applied to the estimation of the limit of detection. The empirical approach included the analysis of samples of progressively decreasing analyte concentration [98,145].
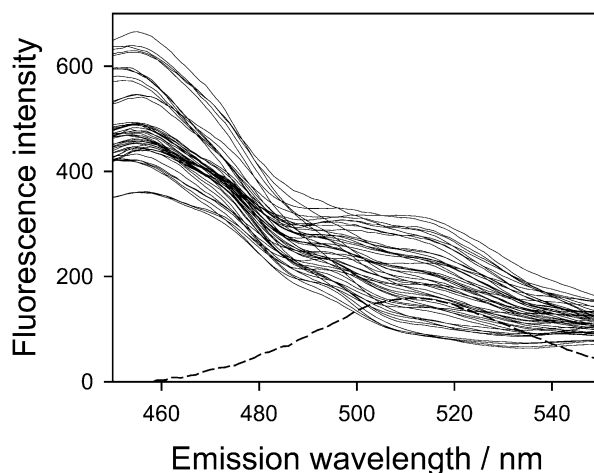

Fig. 11 Fluorescence emission spectra, recorded in the synchronous scanning mode with $\Delta\lambda = 100$ nm, for the set of 50 calibration human sera spiked with concentrations of the antibiotic tetracycline in the range 0–4 ppm (solid lines). The spectrum of the pure analyte at a concentration of 1 ppm is shown with a dotted line. The vertical scale is in arbitrary fluorescence intensity units (FU).

In the analysis of bromhexine in pharmaceutical syrups [165], the strong absorption of the background syrup does also seriously overlap with the analyte signal, see Fig. 12. In contrast to the previous case, here the factor, which dominates the concentration variance, is the uncertainty in analyte calibration concentrations [98]. Analogous to the univariate case, error-propagation expressions indicate that in the latter case the concentration uncertainty is strongly dependent on the sample leverage. (The leverage quantifies the distance of the unknown sample to the calibration center, measured in the multivariate calibration space.) Therefore, since blank samples present varying amounts of background components, leverages for unknown samples differ from sample to sample, making the limit of detection become sample-specific.
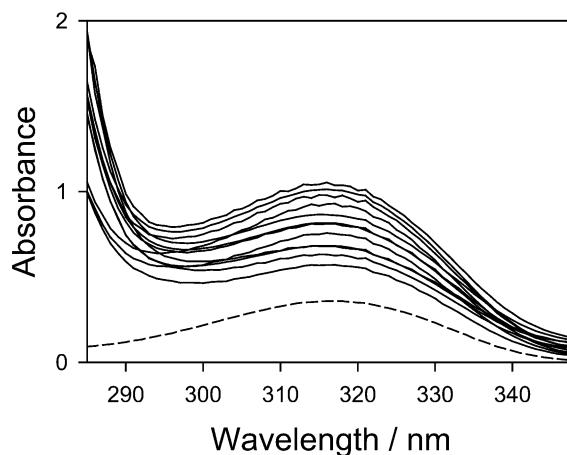
Fig. 12 UV–vis absorption spectra, recorded for the set of 12 calibration syrups spiked with concentrations of bromhexine in the range 1.55×10^{-4} to 2.66×10^{-4} mol dm^{-3} (solid lines). The spectrum of the pure analyte at a concentration of 1 mol dm^{-3} is shown with a dotted line.

Recently, sample-specific standard errors for the concentrations predicted by a three-way PARAFAC model exploiting the second-order advantage have been estimated [166]. A simple error propagation equation based on NAS theory was shown to be useful for this purpose. The results were supported by noise addition Monte Carlo calculations of variance inflation factors and also by experimental data concerning the determination of a therapeutic drug in human urine samples by fluorescence excitation–emission matrices.

4.2 Transforming multivariate models to univariate

An approach that has been suggested by several authors is to perform standard univariate regression using a "surrogate" signal variable, obtained from the sample multivariate signal and directly related to the concentration of the analyte. From this univariate regression line (signal vs. analyte concentration) it is possible to derive figures of merit such as accuracy, sensitivity, and limit of detection. The approach will work well provided the multivariate model is able to efficiently extract the signal corresponding to the selected analyte. In the context of MCR, Saurina et al. [159] have proposed to use the areas of the recovered pure signal profiles of the analyte, which do not depend on the signal of interferents. The authors applied this methodology to calculate the limit of detection, among other figures of merit, for the analysis of triphenyltin, $SnPh_3(1+)$ (IUPAC name triphenylstannylium), in seawater samples by excitation–emission fluorescence, see Fig. 7 for a fluorescence "landscape" of a real sample. The pure emission profile of the analyte is resolved by MCR–ALS, see Fig. 13. The area of the resolved emission profiles is then used as "surrogate" signal. Ortiz et al. [154] have applied a similar approach for multivariate (and multiway) PLS calibration. In this case, the surrogate variable is the concentration of analyte in the calibration samples predicted by the optimal PLS model.
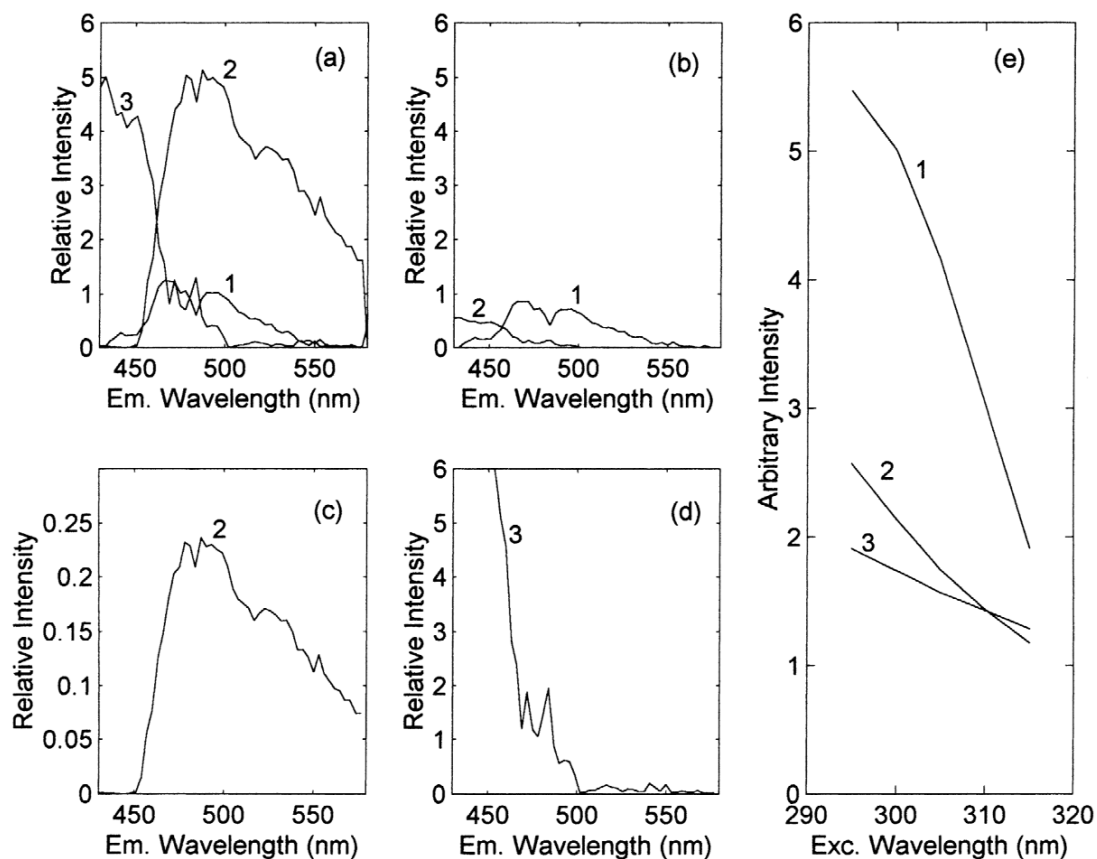
Fig. 13. Resolution of an augmented data matrix with MCR–ALS: (a) emission spectra for the unknown sea water sample; (b) emission spectra for an analyte standard; (c) emission spectra for flavonol [3-hydroxy-2-phenyl-4*H*-chromen-4-one]; (d) emission spectra for the sea water background matrix; (e) excitation spectra. Species assignment: (1) $SnPh_3(1+)$-flavonol complex; (2) flavonol; (3) sea water background. Triphenyltin reference samples supplied as triphenyltin chloride. (Reproduced by permission of Elsevier.)

4.3 Neural classifier

Sarabia et al. [163] have advanced a nonparametric approach that consists of training a neural network to simultaneously optimize the α and β probabilities of giving false detects and false non-detects, respectively, for a fixed threshold concentration. This approach can be extended to any concentration level, and it requires a representative training set of samples with concentration levels well selected above and below the threshold limit. The authors applied this procedure to the polarographic determination of Tl^I/Pb^{II} mixtures and tenoxicam/indomethacin mixtures. Figure 14 shows the polarograms recorded for the Tl^I/Pb^{II} mixtures. The signals show two peaks and a severe overlap between the two analytes, so a classical univariate approach to estimate the capability of detection cannot be applied here. The results with validation samples showed that the four analytes could be detected at less than 10 % of the threshold level with α and β probabilities of error of around 5 %.
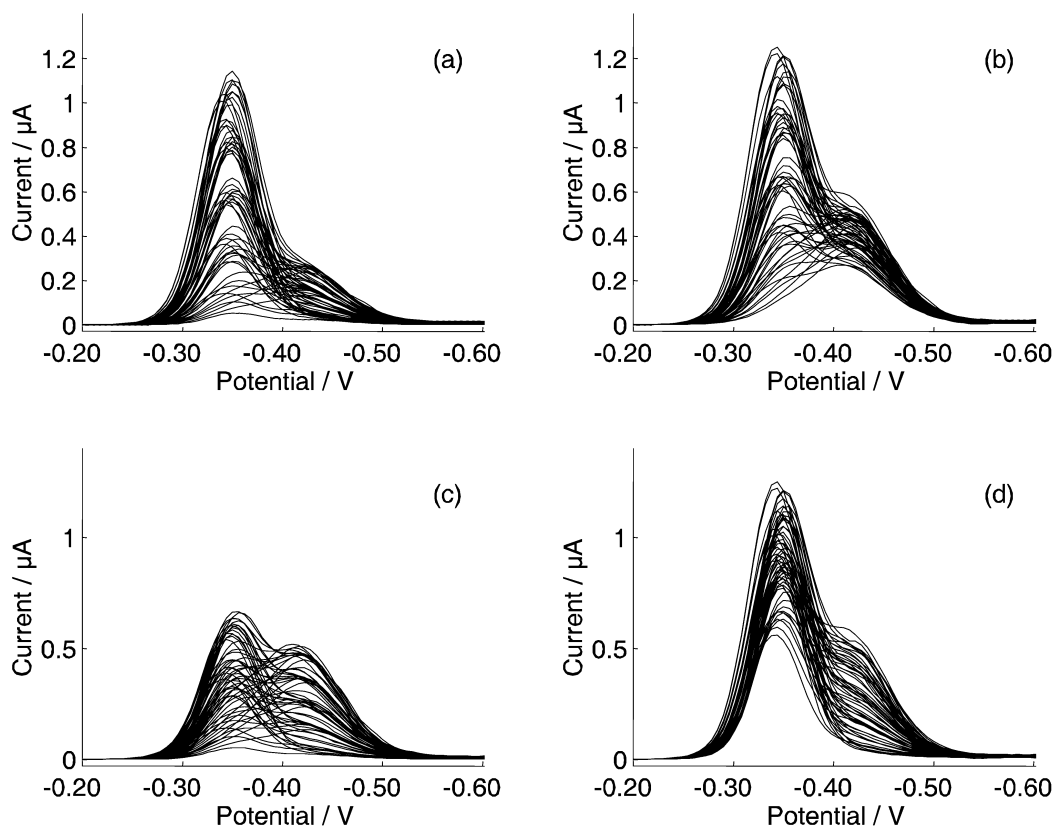
Fig. 14 Polarograms recorded for raw data sets consisting of mixtures of thallium and lead. The peaks corresponding to thallium appears on the right of the overlapping signal, and that for lead on its left. The concentrations of the metal ions are such that (a) contains thallium below and (b) above the threshold values. Likewise, the concentration of lead is below the threshold value in (c) and above in (d). (Reproduced by permission of Elsevier.)

5. CONCLUSIONS AND OUTLOOK

To the best of our knowledge, this paper presents the first comprehensive overview of uncertainty estimation and figures of merit for multivariate calibration. The reader should be aware that we have treated the subject from a chemometrics perspective. The reason for this is that most contributions originate from the chemometrics-oriented literature. Three types of multivariate limit of detection estimators illustrate the potential utility of these contributions. In short, the results are as good as might be expected when applying a common univariate procedure to highly *selective* data. In other words, multivariate methods may allow one to extract truly quantitative information from highly *unselective* data. We hope that this observation will stimulate further testing and application of the methodology covered in this review.

LIST OF SYMBOLS, ACRONYMS, AND ABBREVIATIONS

ANN	artificial neural network
ASTM	American Society for Testing and Materials
b	regression coefficient
$\boldsymbol{b}$	vector of regression coefficients
BLLS	bilinear least-squares
BLS	bivariate least-squares
c	concentration
$\boldsymbol{c}$	concentration vector
CLS	classical least-squares
e	residual value
$\boldsymbol{e}$	vector of residuals
EEM	excitation–emission matrix
F	number of scores
GC	gas chromatography
GRAM	generalized rank annihilation method
GSAM	generalized standard addition method
I	number of samples
i	running index for samples
ICP–OES	inductively coupled plasma–optical emission spectrometry
ILS	inverse least-squares
ISO	International Organization for Standardization
IUPAC	International Union of Pure and Applied Chemistry
J	number of wavelengths
j	running index for wavelengths
K	number of analytes
k	running index for analytes
LBOZ	Lorber, Bergmann, von Oepen, and Zinn
$\boldsymbol{M}$	matrix of second-order signals
MCR–ALS	multivariate curve resolution–alternating least-squares
MS	mass spectrometry
NAS	net analyte signal
NIR	near-infrared
N-PLS	multiway partial least-squares
OLS	ordinary least-squares
PARAFAC	parallel factor analysis
PCR	principal component regression
PLSR	partial least-squares regression
$\boldsymbol{R}$	matrix of first-order calibration signals
r	instrumental signal
$\boldsymbol{r}$	vector of instrumental signals
RAFA	rank annihilation factor analysis
RMSEP	root-mean-square error of prediction
$\boldsymbol{S}$	matrix of pure analyte spectra
s	sensitivity
$\boldsymbol{s}$	pure analyte spectrum
t	percentage point of the t-distribution
UV	ultraviolet
V	variance

WLS	weighted least-squares
x	profile in the first dimension of second-order signals
y	profile in the second dimension of second-order signals
α	nominal coverage of statistical interval or probability of false positive detection decision
β	probability of false negative detection decision
χ^2	statistical χ^2-distribution
λ	wavelength
v	degrees of freedom
ρ	test statistic
σ	standard error of prediction
ξ	pairwise selectivity index

ACKNOWLEDGMENTS

REFERENCES

1. K. Danzer, L. A. Currie. *Pure Appl. Chem.* **70**, 993 (1998).
2. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. *International Vocabulary of Basic and General Terms in Metrology (VIM)*, revision of the 1993 ed., International Organization for Standardization (ISO), Geneva (2004).
3. ISO/IEC 17025. *General Requirements for the Competence of Testing and Calibration Laboratories*, International Organization for Standardization (ISO), Geneva (1999).
4. ISO-GUM. *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization (ISO), Geneva (1995).
5. *Eurachem/CITAC Guide, Quantifying Uncertainty in Analytical Measurement*, 2nd ed., Eurachem/CITAC (2000).
6. P. De Bièvre. *Accred. Qual. Assur.* **2**, 269 (1997).
7. L. A. Currie. *Pure Appl. Chem.* **67**, 1699 (1995).
8. J. Vessman, R. I. Stefan, J. F. van Staden, K. Danzer, W. Lindner, D. T. Burns, A. Fajgelj, H. Müller. *Pure Appl. Chem.* **73**, 1381 (2001).
9. M. Valcárcel, A. Gómez-Hens, S. Rubio. *Trends Anal. Chem.* **20**, 386 (2001).
10. M. Otto, W. Wegscheider. *Anal. Chim. Acta* **180**, 445 (1986).
11. D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte, L. Kaufman. *Chemometrics: A Textbook*, Elsevier, Amsterdam (1988).
12. J. H. Kalivas. *Appl. Spectrosc. Rev.* **25**, 229 (1989–90).
13. J. H. Kalivas, P. M. Lang. *Mathematical Analysis of Spectral Orthogonality*, Marcel Dekker, New York (1994).
14. K. S. Booksh, B. R. Kowalski. *Anal. Chem.* **66**, 782 (1994).
15. J. H. Kalivas, P. M. Lang. *Chemom. Intell. Lab. Syst.* **32**, 135 (1996).
16. K. Danzer. *Fresenius' J. Anal. Chem.* **369**, 394 (2001).
17. N. M. Faber, J. Ferré, R. Boqué, J. H. Kalivas. *Trends Anal. Chem.* **22**, 352 (2003).
18. A. Lorber. *Anal. Chem.* **58**, 1167 (1986).
19. A. Lorber, A. Harel, Z. Goldbart, I. B. Brenner. *Anal. Chem.* **59**, 1260 (1987).
20. G. Bergmann, B. von Oepen, P. Zinn. *Anal. Chem.* **59**, 2522 (1987).

21. H. Martens, T. Næs. *Multivariate Calibration*, John Wiley, Chichester (1989).
22. A. Lorber, B. R. Kowalski. *J. Chemom.* **2**, 93 (1988).
23. A. Höskuldsson. *J. Chemom.* **2**, 211 (1988).
24. K. Danzer, M. Otto, L. A. Currie. *Pure Appl. Chem.* **76**, 1215 (2004).
25. ASTM D6122-99 "Standard Practice for Validation of Multivariate Process Infrared Spectrophotometers", American Society for Testing and Materials International, West Conshohocken, PA, USA.
26. Annual Book of ASTM Standards, Vol. 03.06, E1655, "Standard Practices for Infrared, Multivariate, Quantitative Analysis", American Society for Testing and Materials International, West Conshohocken, PA, USA (1998).
27. G. Bauer, W. Wegscheider, H. M. Ortner. *Spectrochim. Acta* **46B**, 1185 (1991).
28. G. Bauer, W. Wegscheider, H. M. Ortner. *Spectrochim. Acta* **47B**, 179 (1992).
29. E. H. van Veen, S. Bosch, M. T. C. de Loos-Vollebrecht. *Spectrochim. Acta* **52B**, 321 (1997).
30. V. Centner, D. L. Massart, S. de Jong. *Fresenius' J. Anal. Chem.* **361**, 2 (1998).
31. R. Marbach. *J. Biomed. Opt.* **7**, 130 (2002).
32. C. D. Brown. *Anal. Chem.* **76**, 4364 (2004).
33. H. Swierenga, F. Wülfert, O. E. de Noord, A. P. de Weijer, A. K. Smilde, L. M. C. Buydens. *Anal. Chim. Acta* **411**, 121 (2000).
34. R. Todeschini, D. Galvani, J. L. Vilchez, M. del Olmo, N. Navas. *Trends Anal. Chem.* **18**, 93 (1999).
35. R. Leardi. *J. Chemom.* **15**, 559 (2001).
36. I. E. Frank, J. H. Kalivas, B. R. Kowalski. *Anal. Chem.* **55**, 1800 (1983).
37. D. M. Haaland, D. K. Melgaard. *Appl. Spectrosc.* **55**, 1 (2001).
38. E. Sanchez, B. R. Kowalski. *Anal. Chem.* **58**, 496 (1986).
39. E. Sanchez, B. R. Kowalski. *J. Chemom.* **2**, 247 (1988).
40. E. Sanchez, B. R. Kowalski. *J. Chemom.* **4**, 29 (1990).
41. K. Danzer, J. F. Van Staden, D. T. Burns. *Pure Appl. Chem.* **74**, 1479 (2002).
42. H. A. L. Kiers, A. K. Smilde. *J. Chemom.* **9**, 179 (1995).
43. S. Leurgans, R. T. Ross. *Stat. Sci.* **7**, 289 (1992).
44. M. Linder, R. Sundberg. *Chemom. Intell. Lab. Syst.* **42**, 159 (1998).
45. M. Linder, R. Sundberg. *J. Chemom.* **16**, 12 (2002).
46. T. Hirschfeld. *Anal. Chem.* **52**, 297 (1980).
47. M. Kubista. *Chemom. Intell. Lab. Syst.* **7**, 273 (1990).
48. I. D. Wilson, U. A. Th. Brinkman. *J. Chromatogr., A* **1000**, 325 (2003).
49. R. Bro. *Chemom. Intell. Lab. Syst.* **38**, 149 (1997).
50. I. M. Warner, E. R. Davidson, G. D. Christian. *Anal. Chem.* **49**, 2155 (1977).
51. A. DeJuan, E. Casassas, R. Tauler. In *Encyclopedia of Analytical Chemistry*, Vol. 11, R. A. Myers (Ed.), pp. 9800–9837, John Wiley, Chichester (2002).
52. B. E. Wilson, W. Lindberg, B. R. Kowalski. *J. Am. Chem. Soc.* **111**, 3797 (1989).
53. Y. Wang, O. Borgen, B. R. Kowalski, M. Gu, F. Turecek. *J. Chemom.* **7**, 117 (1993).
54. S. Wold, P. Geladi, K. Esbensen, J. Øhman. *J. Chemom.* **1**, 41 (1987).
55. R. Bro. *J. Chemom.* **10**, 47 (1996).
56. N. M. Faber, R. Bro. *Chemom. Intell. Lab. Syst.* **61**, 133 (2002).
57. R. Bro, Å. Rinnan, N. M. Faber. *Chemom. Intell. Lab. Syst.* **75**, 69 (2005).
58. F. J. del Río, J. Riu, F. X. Rius. *J. Chemom.* **15**, 773 (2001).
59. M. Galea-Rojas, M. V. de Castilho, H. Bolfarine, M. de Castro. *Analyst* **128**, 1073 (2003).
60. IUPAC. *Compendium of Analytical Nomenclature*, Web edition <http://www.iupac.org/publications/analytical_compendium/>.

61. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML. *International Vocabulary of Basic and General Terms in Metrology (VIM)*, 2nd ed., International Organization for Standardization (ISO), Geneva (1993).
62. M. Thompson, S. L. R. Ellison, R. Wood. *Pure Appl. Chem.* **74**, 835 (2002).
63. Y. Umezawa, P. Bühlmann, K. Umezawa , K. Tohda, S. Amemiya. *Pure Appl. Chem.* **72**, 1851 (2000).
64. Y. Umezawa, K. Umezawa, P. Bühlmann, N. Hamada, H. Aoki, J. Nakanishi, M. Sato, K. P. Xiao, Y. Nishimura. *Pure Appl. Chem.* **74**, 923 (2002).
65. Y. Umezawa, P. Bühlmann, K. Umezawa, N. Hamada. *Pure Appl. Chem.* **74**, 995 (2002).
66. J. Mandel, R. D. Stiehler. *J. Res. Natl. Bur. Stand.* **53**, 155 (1954).
67. L. Cuadros Rodríguez, A. M. García Campaña, C. Jiménez Linares, M. Román Ceba. *Anal. Lett.* **26**, 1243 (1993).
68. P. Wilrich (Chairman). ISO 11843 series of standards about capability of detection; Part 1: Terms and definitions (1997) + Cor 1:2003, IDT; Part 2: Methodology in the linear calibration case (2000 IDT); Part 3: Methodology for determination of the critical value for the response variable when no calibration data are used (2003, IDT); Part 4: Methodology for comparing the minimum detectable value with a given value (2003, IDT); Part 5: Methodology in the linear and non-linear calibration cases (2005, AWI ISO TC 69/SC6).
69. L. A. Currie. *Chemom. Intell. Lab. Syst.* **37**, 151 (1997).
70. R. Boqué, F. X. Rius. *Trends Anal. Chem.* **16**, 432 (1997).
71. H. van der Voet. In *Encyclopedia of Environmetrics*, Vol. 1, A. H. El-Shaarawi, W. W. Piegorsch (Eds.), pp. 504–515, John Wiley, Chichester (2002).
72. F. J. del Río Bocio, J. Riu, R. Boqué, F. X. Rius. *J. Chemom.* **17**, 413 (2003).
73. R. Ferrús, M. R. Egea. *Anal. Chim. Acta* **287**, 119 (1994).
74. P. Geladi. *Chemom. Intell. Lab. Syst.* **60**, 211 (2002).
75. J. R. Magnus, H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley, Chichester (1988).
76. J. Shao, D. Tu. *The Jackknife and Bootstrap*, Springer, New York (1995).
77. N. M. Faber. *J. Chemom.* **15**, 743 (2001).
78. N. M. Faber, J. Ferré, R. Boqué, J. H. Kalivas. *Chemom. Intell. Lab. Syst.* **63**, 107 (2002).
79. R. DiFoggio. *Appl. Spectrosc.* **54**, 94A (2000).
80. A. Lorber, K. Faber, B. R. Kowalski. *J. Chemom.* **10**, 215 (1996).
81. A. Lorber, B. R. Kowalski. *J. Chemom.* **2**, 67 (1988).
82. M. C. Denham. Ph.D. Thesis, University of Liverpool (1991).
83. T. V. Karstang, J. Toft, O. M. Kvalheim. *J. Chemom.* **6**, 177 (1992).
84. A. Phatak, P. M. Reilly, A. Penlidis. *Anal. Chim. Acta* **277**, 495 (1993).
85. A. Phatak. Ph.D. Thesis, University of Waterloo (1993).
86. S. De Vries, C. J. F. Ter Braak. *Chemom. Intell. Lab. Syst.* **30**, 239 (1995).
87. K. Faber, B. R. Kowalski. *Chemom. Intell. Lab. Syst.* **34**, 283 (1996).
88. A. J. Berger, M. S. Feld. *Appl. Spectrosc.* **51**, 725 (1997).
89. M. C. Denham. *J. Chemom.* **11**, 39 (1997).
90. K. Faber, B. R. Kowalski. *J. Chemom.* **11**, 181 (1997).
91. N. M. Faber, D. L. Duewer, S. J. Choquette, T. L. Green, S. N. Chesler. *Anal. Chem.* **70**, 2972 (1998).
92. M. Høy, K. Steen, H. Martens. *Chemom. Intell. Lab. Syst.* **44**, 123 (1998).
93. T. Morsing, C. Ekman. *J. Chemom.* **12**, 295 (1998).
94. M. C. Denham. *J. Chemom.* **14**, 351 (2000).
95. N. M. Faber. *J. Chemom.* **14**, 363 (2000).
96. N. M. Faber. *Chemom. Intell. Lab. Syst.* **52**, 123 (2000).

97. X. -H. Song, N. M. Faber, P. K. Hopke, D. T. Suess, K. A. Prather, J. J. Schauer, G. R. Cass. *Anal. Chim. Acta* **446**, 329 (2001).
98. A. C. Olivieri. *J. Chemom.* **16**, 207 (2002).
99. J. A. Fernández Pierna, L. Jin, F. Wahl, N. M. Faber, D. L. Massart. *Chemom. Intell. Lab. Syst.* **65**, 281 (2003).
100. N. M. Faber, X.-H. Song, P. K. Hopke. *Trends Anal. Chem.* **22**, 330 (2003).
101. S. Serneels, P. Lemberge, P. J. Van Espen. *J. Chemom.* **18**, 76 (2004).
102. R. Kleinknecht. *J. Chemom.* **10**, 687 (1996).
103. P. Stoica, T. Söderström. *Scand. J. Stat.* **25**, 17 (1998).
104. A. Phatak, P. M. Reilly, A. Penlidis. *Lin. Algebra Appl.* **354**, 245 (2002).
105. T. Næs, H. Martens. *J. Chemom.* **2**, 155 (1988).
106. W. J. Egan, W. E. Brewer, S. L. Morgan. *Appl. Spectrosc.* **53**, 218 (1999).
107. G. Chryssolouris, M. Lee, A. Ramsey. *IEEE Trans. Neural Networks* **7**, 229 (1996).
108. G. Baffi, E. Martin, J. Morris. *Chemom. Intell. Lab. Syst.* **61**, 151 (2002).
109. N. M. Faber, L. M. C. Buydens, G. Kateman. *J. Chemom.* **7**, 495 (1993).
110. C. Jochem, P. Jochem, B. R. Kowalski. *Anal. Chem.* **53**, 85 (1981).
111. M. G. Moran, B. R. Kowalski. *Anal. Chem.* **56**, 562 (1984).
112. C.-N. Ho, G. D. Christian, E. R. Davidson. *Anal. Chem.* **52**, 1071 (1980).
113. C. J. Appellof, E. R. Davidson. *Anal. Chim. Acta* **146**, 9 (1983).
114. E. R. Malinowski. *Factor Analysis in Chemistry*, 3rd ed., John Wiley, New York (2002).
115. E. Sanchez. Ph.D. Thesis, University of Washington (1987).
116. K. Booksh, B. R. Kowalski. *J. Chemom.* **8**, 45 (1994).
117. N. M. Faber, L. M. C. Buydens, G. Kateman. *J. Chemom.* **8**, 181 (1994).
118. K. Faber, A. Lorber, B. R. Kowalski. *J. Chemom.* **11**, 95 (1997).
119. N. M. Faber, J. Ferré, R. Boqué. *Chemom. Intell. Lab. Syst.* **55**, 67 (2001).
120. N. M. Faber, R. Boqué, J. Ferré. *Chemom. Intell. Lab. Syst.* **55**, 91 (2001).
121. S. Serneels, M. Moens, P. J. Van Espen, F. Blockhuys. *Anal. Chim. Acta* **516**, 1 (2004).
122. S. Serneels, N. M. Faber, P. J. Van Espen. In preparation.
123. C. J. Appellof, E. R. Davidson. *Anal. Chem.* **53**, 2053 (1981).
124. A. C. Olivieri, N. M. Faber. *Chemom. Intell. Lab. Syst.* **70**, 75 (2004).
125. P. Paatero. *Chemom. Intell. Lab. Syst.* **38**, 223 (1997).
126. X. Liu, N. Sidiropoulos. *IEEE Trans. Signal Processing* **49**, 2074 (2001).
127. F. J. del Río, J. Riu, F. X. Rius. *Anal. Chim. Acta* **446**, 49 (2001).
128. À. Martínez, J. Riu, F. X. Rius. *Anal. Chim. Acta* **446**, 145 (2001).
129. À. Martínez, J. Riu, F. X. Rius. *Chemom. Intell. Lab. Syst.* **54**, 61 (2000).
130. À. Martínez, F. J. del Río, J. Riu, F. X. Rius. *Chemom. Intell. Lab. Syst.* **49**, 179 (1999).
131. J. Riu, F. X. Rius. *Anal. Chem.* **68**, 1851 (1996).
132. F. E. Satterthwaite. *Psychometrika* **6**, 309 (1941).
133. D. R. Morgan. *Appl. Spectrosc.* **51**, 404 (1977).
134. K. Faber, A. Lorber, B. R. Kowalski. *J. Chemom.* **11**, 419 (1997).
135. N. M. Faber. *J. Chemom.* **12**, 405 (1998).
136. J. Ferré, S. D. Brown, F. X. Rius. *J. Chemom.* **15**, 537 (2001).
137. J. Ferré, N. M. Faber. *Chemom. Intell. Lab. Syst.* **69**, 123 (2003).
138. R. Bro, C. M. Andersen. *J. Chemom.* **17**, 646 (2004).
139. A. Lorber, K. Faber, B. R. Kowalski. *Anal. Chem.* **69**, 1620 (1997).
140. N. J. Messick, J. H. Kalivas, P. M. Lang. *Anal. Chem.* **68**, 1572 (1996).
141. J. H. Kalivas. *J. Chemom.* **3**, 409 (1989).
142. M. A. Arnold, G. W. Small, D. Xiang, J. Qui, D. W. Murhammer. *Anal. Chem.* **76**, 2583 (2004).
143. C. D. Brown, T. D. Ridder. *Appl. Spectrosc.* **59**, 787 (2005).
144. T. D. Ridder, C. D. Brown, B. J. Versteeg. *Appl. Spectrosc.* **59**, 804 (2005).

145. H. C. Goicoechea, A. C. Olivieri. *Anal. Chem*. **19**, 4361 (1999).
146. M. Bélen Sanz, L. A. Sarabia, A. Herrero, M. Cruz Ortiz. *Anal. Chim. Acta* **489**, 85 (2003).
147. J. W. B. Braga, R. J. Poppi. *J. Pharm. Sci*. **93**, 2124 (2004).
148. A. Muñoz de la Peña, A. Espinosa Mansilla, D. González Gómez, A. C. Olivieri, H. Goicoechea. *Anal. Chem*. **75**, 2640 (2003).
149. F. C. Garner, G. L. Robertson. *Chemom. Intell. Lab. Syst*. **3**, 53 (1988).
150. R. Boqué, F. X. Rius. *Chemom. Intell. Lab. Syst*. **32**, 11 (1996).
151. K. Faber, B. R. Kowalski. *Fresenius' J. Anal. Chem*. **357**, 789 (1997).
152. R. Boqué, N. M. Faber, F. X. Rius. *Anal. Chim. Acta* **423**, 41 (2000).
153. R. Boqué, M. S. Larrechi, F. X. Rius. *Chemom. Intell. Lab. Syst*. **45**, 397 (1999).
154. M. C. Ortiz, L. A. Sarabia, A. Herrero, M. S. Sánchez, M. B. Sanz, M. E. Rueda, D. Giménez, M. E. Meléndez. *Chemom. Intell. Lab. Syst*. **69**, 21 (2003).
155. M. F. Delaney. *Chemom. Intell. Lab. Syst*. **3**, 45 (1988).
156. A. Singh. *Anal. Chim. Acta* **277**, 205 (1993).
157. C. Leal, M. Granados, J. L. Beltrán, R. Campañó, M.-D. Prat. *Analyst* **122**, 1293 (1997).
158. H. van der Voet, W. J. de Boer, W. G. de Ruig, J. A. van Rhijn. *J. Chemom*. **12**, 79 (1998).
159. J. Saurina, C. Leal, R. Campañó, M. Granados, M.-D. Prat, R. Tauler. *Anal. Chim. Acta* **432**, 241 (2001).
160. F. M. Fernández, M. B. Tudino, O. E. Troccoli. *Anal. Chim. Acta* **433**, 119 (2001).
161. A. Herrero, S. Zamponi, R. Marassi, P. Conyi, M. C. Ortiz, L. A. Sarabia. *Chemom. Intell. Lab. Syst.* **61**, 63 (2002).
162. R. Boqué, J. Ferré, N. M. Faber, F. X. Rius. *Anal. Chim. Acta* **451**, 313 (2002).
163. L. A. Sarabia, M. Cruz Ortiz, M. Julia Arcos, M. Sagrario Sánchez, A. Herrero, S. Sanllorente. *Chemom. Intell. Lab. Syst.* **61**, 89 (2002).
164. M. B. Sanz, L. A. Sarabia, A. Herrero, M. C. Ortiz. *Anal. Chim. Acta* **446**, 297 (2001).
165. H. C. Goicoechea, A. C. Olivieri. *Talanta* **49**, 793 (1999).
166. A. C. Olivieri. *J. Chemom.* **18**, 363 (2004).