Review

Does chemometrics enhance the performance of electroanalysis?

Yongnian Ni\textsuperscript{a,b,}*, Serge Kokot\textsuperscript{c}

\textsuperscript{a} State Key Laboratory of Food Science and Technology, Nanchang University, Nanchang, Jiangxi 330047, China
\textsuperscript{b} Department of Chemistry, Nanchang University, Nanchang, Jiangxi 330047, China
\textsuperscript{c} Inorganic Materials Research Program, School of Physical and Chemical Sciences, Queensland University of Technology, Brisbane, Queensland 4001, Australia

\textbf{Abstract}

This review explores the question whether chemometrics methods enhance the performance of electroanalytical methods. Electroanalysis has long benefited from the well-established techniques such as potentiometric titrations, polarography and voltammetry, and the more novel ones such as electronic tongues and noses, which have enlarged the scope of applications. The electroanalytical methods have been improved with the application of chemometrics for simultaneous quantitative prediction of analytes or qualitative resolution of complex overlapping responses. Typical methods include partial least squares (PLS), artificial neural networks (ANNs), and multiple curve resolution methods (MCR-ALS, N-PLS and PARAFAC). This review aims to provide the practising analyst with a broad guide to electroanalytical applications supported by chemometrics. In this context, after a general consideration of the use of a number of electroanalytical techniques with the aid of chemometrics methods, several overviews follow with each one focusing on an important field of application such as food, pharmaceuticals, pesticides and the environment. The growth of chemometrics in conjunction with electronic tongue and nose sensors is highlighted, and this is followed by an overview of the use of chemometrics for the resolution of complicated profiles for qualitative identification of analytes, especially with the use of the MCR-ALS methodology. Finally, the performance of electroanalytical methods is compared with that of some spectrophotometric procedures on the basis of figures-of-merit. This showed that electroanalytical methods can perform as well as the spectrophotometric ones. PLS-1 appears to be the method of practical choice if the %relative prediction error of $\sim$$\pm$10% is acceptable.

\textcopyright 2008 Elsevier B.V. All rights reserved.
1. Introduction

Electroanalysis in practice, is subject to the same quality assurance (QA) criteria and common method preferences, such as versatility, rapid throughput, simplicity of operation, and lower costs, as the other analytical sub-disciplines such as spectroscopy in its various forms. In general, over the last twenty years, the development of instrumentation has been significantly influenced by the rapid developments in electronics and particularly desktop computers. This is especially the case for electrochemistry. Such progress has enabled faster measurement and efficient storage of data, and in turn this has provided increasing opportunities for the application of multivariate methods of analysis, i.e., chemometrics in the context of analytical chemistry, for rapid qualitative and quantitative data interpretation. This also facilitated potential cost savings because, for example, it became possible to predict several different variables from a single response provided properly constructed multivariate calibrations were available.

In general, it is well known that electrochemical analysis has benefited from the electronics revolution in at least two important ways (i) the development of neater, faster, more simple and arguably, competitively affordable instrumentation, and (ii) the potential for rapid analysis with the aid of chemometrics applications in electroanalytical chemistry, illustrating the great interest in the subject by the research community, and the potential for future applications [7,8]. Richards et al. [8] indicated that many chemometrics data analysis techniques were already available, and voltammetric methods were the main focus of interest for electrochemical data analysis as they could analyse for several analytes simultaneously. Analysis of amperometric data typically involved only signal conditioning to remove noise and increase accuracy, although when an electrode array was used, there was scope for multivariate modelling. Esteban et al. presented two significant reviews [9,10], which considered the application of chemometrics to electroanalytical data with special attention to the contributions of the last decade. The first paper [9] is a useful overview of the most commonly used chemometrics methods in electroanalysis, and cites 65 articles, and the second is a deeper approach with 183 references. In these two reviews, the authors noted that the self-modelling curve resolution (MCR) and other multivariate analysis methods were very powerful tools for the interpretation of electroanalytical data, especially for multianalyte calibration and modelling.
in multicomponent dynamic systems. They also pointed out that modern electrochemical instrumentation can provide reliable and reproducible data, which are the basis of analysis with very low Limits of Detection. The combined use of modern electroanalytical instrumentation with chemometrics data modelling was envisaged to improve significantly the capabilities of electroanalysis.

In this work, we address an important topic for the practising analyst, i.e., the comparison of electrochemical methods and applications with similar methodologies for the same or similar analytes determined by other techniques such as spectrophotometry. Richards et al. [8] have only generally addressed this issue, probably because at the time of writing (2002), there was only limited information of this kind available in the literature. Since then, there has been a significant increase in publications, which address at least some, if not all aspects of analysis, required for quality assurance purposes. Also, now there are sufficient studies available to compare the performance of some electrochemical methods for some analytes with those from other techniques such as UV-visible spectrophotometry.

This review explores the question whether chemometrics methods enhance the quantitative and qualitative performance of electroanalytical methods. Also, in this context, the aim is to provide the practising analyst with a broad overview of the application of chemometrics to electroanalysis. After a general overview of the use of several electroanalytical techniques such as potentiometric titrations, polarography and voltammetry, aided by chemometrics methods, such as PLS, ANNs and MCR, there follow focused overviews of the applications of chemometrics in some common, important areas of analysis such as food, pharmaceuticals, pesticides and the environment. The important developing fields of application of electronic tongues and noses are also considered. This is followed by a brief overview of the use of chemometrics for the resolution of complicated profiles measured for the purpose of qualitative identification of analytes, especially highlighting the usefulness of the MCR-ALS methodology. The review is rounded off with a brief comparison of the performance of electroanalytical methods with spectrophotometric ones on the basis of some figures-of-merit taken from recent literature. Some suggestions are provided for the practising analysts to facilitate the selection of a method of analysis, which is fit-for-purpose for their application.

2. Gran methods, potentiometric titrations and chemometrics

In the early 1950s, Gran [11,12] proposed two mathematical procedures for evaluating the end-points of potentiometric titration in terms of linear functions, which intersect at the end-point. Later, Ingman et al. [13,14] and Johansson [15,16] developed methods based on chemical equilibrium equations, from which it was possible to determine the acid constituents in their mixture by submitting the measured data to regression analysis. In 1981, Gran et al. [17] developed a linear precipitation titration method based on equilibrium equations, which was applied for the determination of chloride. Potentiometric titration methods for complex systems were significantly improved after the multivariate calibration approaches were introduced [18]. Recently, Michalowski et al. [19] provided an overview (52 references) on the Gran and other mathematical methods involving potentiometric titrations.

2.1. Acid–base potentiometric titration

In the titration of an acid–base system, the detection of end-points usually depends on the use of visual indicators or potentiometric methods. These methods are essentially based on the inflection point at which there is a maximum change in pH or potential. In general, for binary or ternary mixed acid systems, if \( \Delta pK \) (acid-dissociation constant) between any two acids is less than 4, the titration steps of the acids overlap, and it is difficult to determine the concentration of each acid in such cases.

Gran’s [11,12] and Ingman and Johanssons’ [13–16] methods mentioned above have two main disadvantages. The first one is that the thermodynamic dissociation constants of each acid have to be known for the derivation of the equilibrium equations. However, published constants can be unreliable for a particular mixture of acids, and as a result, errors can occur in the analysis. The second disadvantage is that the potential titration equations involved with these methods, are based on thermodynamic constants, and are usually very complex. Thus, accurate measurement of pH is essential in order to eliminate systematic errors, and to achieve this, accurate calibration of the electrode systems is required, which is quite time consuming.

Lindberg and Kowalski [18] suggested the use of PLS calibration for the determination of acids in mixtures. In this approach, no explicit model is assumed and thus, the model error is significantly reduced. The chemical equilibrium of mixtures of acids in a potential titration was derived as follows [20]:

\[
V_0 \text{ mL of solution of a polyprotic acid, } H_nA, \text{ is titrated with a strong base (e.g., sodium hydroxide). When } V \text{ mL of titrant is added, the mass equilibrium is:}
\]

\[
\text{[Na]} = \frac{c_V V}{V_0 + V} \quad \text{(for simplicity, ionic charges are omitted)} \tag{1}
\]

and the electroneutrality condition is:

\[
[H] + [Na] = [OH] + n[A] + (n - 1)[HA] + \cdots + [H_{n-1}A] \tag{2}
\]

where \( c_V \) is the concentration of the titrant. The fraction of each species present can be written as:

\[
[A] = \frac{c_V V}{V_0 + V} \cdot \frac{K_1K_2 \cdots K_n}{[H]^n + K_1[H]^{n-1} + \cdots + K_1K_2 \cdots K_n} = \frac{c_V V}{V_0 + V} \cdot \alpha_n \tag{3}
\]

\[
[HA] = \frac{c_V V}{V_0 + V} \cdot \frac{K_1K_2 \cdots K_{n-1}}{[H]^n + K_1[H]^{n-1} + \cdots + K_1K_2 \cdots K_{n-1}^{n-1}} = \frac{c_V V}{V_0 + V} \cdot \alpha_{n-1}
\]

\[
[H_{n-1}A] = \frac{c_V V}{V_0 + V} \cdot \frac{K_1}{[H]^n + K_1[H]^{n-1} + \cdots + K_1K_2 \cdots K_{n-1}} = \frac{c_V V}{V_0 + V} \cdot \alpha_1
\]
where $K_1$, $K_2$, ..., $K_n$ are the stepwise acid-dissociation constants and $q_1$, $q_2$, ..., $q_n$ denote the fractional composition of the acids. As detailed in literature [20], the combination and rearrangement of the above equations lead to a relationship for $M$ different acids in a mixture, provided the relationship between the volume of titrant added and the mole concentration of acids is approximately additive:

$$V_j = k_{j0} + k_{j1}c_1 + \cdots + k_{jm}c_m = k_{j0} + \sum_{m=1}^{M} k_{jm}c_m \quad (j = 1, 2, \ldots, J)$$  \hspace{1cm} (4)$$

where $k_{j0}$ is the non-zero intercept. Thus, the equation for the potentiometric titration of mixed acids is linear, and can be simplified. Let $c_0 = 1$ and Eq. (4) can be written as:

$$V_j = \sum_{m=0}^{M} k_{jm}c_m \quad (j = 0, 1, \ldots, J)$$  \hspace{1cm} (5)$$

and in matrix form:

$$v^T = c^T K$$  \hspace{1cm} (6)$$

From the above description, it is clear that, although the chemical equilibrium model for the potentiometric titration is quite complex, it can be reduced to a simple form and the data collected from such analyses can be made available in matrix form for chemometrics interpretation.

Song et al. [21] applied a three-layer ANN-BP (back propagation) model to potentiometric titration data from polybasic acid mixtures. They pointed out that the mixed acid system is non-linear because of the complex solution chemistry equilibrium involved, and the interaction among components of the mixture. Satisfactory prediction results were obtained for three-component samples containing maleic, propandioic and succinic acids with an average relative error of 4.5%. ANN-BP was also used to model the complex non-linear relationship between the concentration of anthranilic, nicotinic, picolinic and succinic acids, as well as ternary mixtures of tartaric, citric, and ascorbic acids, titrated with sodium hydroxide. The model was then used to predict the concentrations of components in unknown samples in the concentration range of $4 \times 10^{-4}$ to $2 \times 10^{-3}$ mol L$^{-1}$, and acceptable results were obtained. Aktas and Yasar [26] used ANN-BP method to model a complex non-linear relationship between the concentration of p-coumaric, sinapinic, vanillic, and isovanillic acids. The principal components of the measured data matrix were used as the input of the network and the optimized model was used to predict the concentrations of the acids in synthetic mixtures with an average %RPE $< 4.18$. The same method was also applied to resolve the potentiometric titration data of ibuprofen, indomethacin and naproxen with an average %RPE $< 2.30$ [27].

### 2.2. Complexometric and precipitation titrations

In a similar manner, it has been shown that it is possible to analyse: (i) by potentiometric complexometric titration a mixture of metal ions, and (ii) by potentiometric precipitation titration a mixture of halide ions, $X$, and the thiocyanate ion as well [19,20,28,29].

(i) The potentiometric complexometric titration model for the determination of a mixture of metal ions has been described in detail elsewhere [28]. Its basis is as follows: consider that a small amount of the mercury complex, $\text{HgY}_2^-$, $Y$ (ethylenediaminetetraacetic acid, EDTA) was added to a solution of $m$ different metal ions, $M_i^{L^+} (i = 1, 2, \ldots, m)$. If the differences between the values of the formation constants of the $M_i$-EDTA complexes formed were small, then the following equilibria would apply:

$$\text{H}_2\text{Y}_2^- + \text{Hg}^{2+} = \text{HgY}_2^- + 2\text{H}^+,$$  \hspace{1cm} (7)$$

and

$$\text{H}_2\text{Y}_2^- + M_i^{L^+} = \text{M}_i\text{Y}^{L^-} + 2\text{H}^+,$$  \hspace{1cm} (8)$$

where $M_i^{L^+}$ denotes the metal ion, $i$, in the solution, and $K_{\text{HgY}}$ and $K_{\text{M}_i\text{Y}}$ are the formation constants of $\text{HgY}_2^-$ and $\text{M}_i\text{Y}^{L^-}$ ($K_{\text{HgY}} > K_{\text{M}_i\text{Y}}$), respectively.

Ultimately, it can be shown [28] that the titration volume, $V_i$, of the EDTA titrant at a selected potential point $j$ ($j = 1, 2, \ldots, J$) is given by:

$$V_i = k_iV_0c_{M_i} = k_iN_{M_i}$$  \hspace{1cm} (9)$$

where $k_i$ is a constant, $V_0$ is the mixed metals sample volume, $c_{M_i}$ is the initial concentration of $M_i$, and $N_{M_i}$ is the number of moles of metal ion, $M_i^{L^+}$, in the solution.

It follows that for all metal ions, $m$, the total volume of titrant, EDTA, consumed is:

$$V = k_0 + k_1N_{M_1} + k_2N_{M_2} + \cdots + k_mN_{M_m}$$  \hspace{1cm} (10)$$
where $V$ is the total titration volume, $k_i$ is the proportionality coefficient, and $k_0$ is the non-zero intercept, which accounts for the noise and the background. Thus, by measuring the volume of EDTA at $r$ potential points during the titration, $r$ equations are obtained, which can be expressed in matrix form as:

$$v^T = n^T K$$  \hspace{1cm} (11)

This method was used for the determination of twenty synthetic mixtures with two-, three- and four-component combinations of lead, zinc and cobalt with concentration levels from $5 \times 10^{-4}$ to $1.5 \times 10^{-3}$ mol L$^{-1}$. A PLS calibration model was built and acceptable prediction results were obtained [28]. Calvo and Valle [30] applied ANN to the potentiometric EDTA titrations of Ca$^{2+}$, Mg$^{2+}$ and Sr$^{2+}$ mixtures, up to 3.3 mM total ion concentration. Good comparisons were observed between the obtained and expected concentrations for the three cations with the external validation samples ($n = 17$, with root mean square error (RMSE) values of 0.18, 0.28 and 0.32 mM for Ca$^{2+}$, Mg$^{2+}$ and Sr$^{2+}$, respectively). The proposed procedure was applied to mineral waters and compared with reference methods.

(ii) The potentiometric titration model involving precipitation is generally applied for the determination of the halide (X) and thiocyanate ions with use of a silver indicator electrode. Its basis may be understood from the reaction equilibria involved. Thus, consider the case when $V_0$ mL of a solution containing $m$ different kinds of anions X with concentration $c_m$ mol L$^{-1}$, is titrated with $V$ mL of concentration, $c_T$ (mol L$^{-1}$) AgNO$_3$, and assume the formation of complexes, AgX$^j$ ($j = 1, 2, \ldots, J$). If the hydroxo-complexes Ag(OH)$_n$ are neglected, then at a given potential the amount of the formed precipitate, $P$, with respect to component, X, can be obtained [29]:

$$P = c_V 0 - ([X] + \sum_{j=1}^J \beta_j [Ag][X]^j)(V_0 + V)$$  \hspace{1cm} (12)

and similarly for $P$ with respect to the titrant, AgNO$_3$:

$$P = c_T V_0 - ([X] + \sum_{j=1}^J \beta_j [Ag][X]^j)(V_0 + V)$$  \hspace{1cm} (13)

where $\beta_j$ is the successive formation constant of X with Ag$^+$ ion, and $J$ denotes the maximum number of the formed silver complexes. By combining and rearranging Eqs. (12) and (13), the following equation is obtained:

$$V = \frac{c_V 0 + V_0([Ag] - [X] - \sum_{j=1}^J (j - 1)\beta_j [Ag][X]^j)}{c_T - [Ag] + [X] + \sum_{j=1}^J (j - 1)\beta_j [Ag][X]^j}$$  \hspace{1cm} (14)

When the solution is titrated to a selected potential point from total of $N$ potential points ($n = 1, 2, \ldots, N$), both $[X]$ and $[Ag]$ are constant. Then, if the volume of the added AgNO$_3$ at a potential, $n$, is denoted by $V_n$, the above relationship may be written as:

$$V_n = k_{i0} + k_{i1}c$$  \hspace{1cm} (15)

where $k_{i0}$ is $\frac{V_0([Ag] - [X] - \sum_{j=1}^J (j - 1)\beta_j [Ag][X]^j)}{c_T - [Ag] + [X] + \sum_{j=1}^J (j - 1)\beta_j [Ag][X]^j}$ and $k_{i1}$ is $\frac{V_0}{c_T - [Ag] + [X] + \sum_{j=1}^J (j - 1)\beta_j [Ag][X]^j}$.

For $M$ kinds of anion, $X_m$ ($m = 1, 2, \ldots, M$), if the concentration of components, $m$, is represented by $c_m$, and let $c_0 = 1$, then the above equation expands to:

$$V_n = k_{i0} + k_{i1}c_1 + k_{i2}c_2 + \ldots + k_{iM}c_M$$  \hspace{1cm} (16)

which in matrix form is:

$$v^T = c^T K$$  \hspace{1cm} (17)

This model was applied for 27 synthetic mixtures for the determination of three- and four-analyte component sets with combinations of chloride, bromide, iodide and thiocyanate. The analyte concentrations were quite low in the range of $1.8 \times 10^{-4}$ to $6.2 \times 10^{-4}$ mol L$^{-1}$ [29]. Shamsipur et al. [31] applied a three-layer ANN-BP to model the non-linear potentiometric titration data of silver, copper and mercury with potassium thiocyanate as the titrant and a carbon paste electrode as the indicator electrode. The proposed method was applied for the simultaneous determination of these metal ions in synthetic mixtures with low values of RPE (in the range of 0.1–0.55%). Brodnjak-Voncina et al. [32] applied ANN to resolve the potentiometric precipitation data from sulfate ions in river and drinking waters (titrant: Ba$^{2+}$) and calcium in wine samples (titrant: oxalate). The results indicated that ANN could successfully predict the concentration of compounds from the titration curves with an error of ±10%, which is acceptable for rapid screening of waters and wines.

### 3. Quantitative applications of electroanalysis and chemometrics in important samples

Over the last decade, we have often commented in our work on the utility of the electrometric and spectrophotometric methods of analysis aided by chemometrics in the developing countries. The novel methods we and, indeed other colleagues, have described for the above two techniques require relatively simple, cheap instrumentation. They are fit-for-purpose because after method validation, the applicability of the method for the analysis of real off-the-shelf products was successfully demonstrated. Thus, in this review, we provide an overview of methods of electroanalysis, which may be generally applied in important fields such as food, pharmaceuticals, pesticides and the environment.

Electronic nose and tongue sensors attempt to mimic animals’ smell and taste senses to detect odours and chemical compounds [33]. In recent reviews on electronic sensors, their applications showed commercial potential in the fields such as biomedicine [34], environment [35], food and security. It was
noted that in most cases chemometrics, especially the pattern recognition methods, were generally required for data interpretation. The role of electronic sensors and chemometrics is further addressed in the following section.

3.1. Food analysis

Adulteration and authentication of food products are important topics in the food industry [36]. It is a major concern not only for consumers, but also for producers and distributors [37]. Indeed, regulatory authorities, food processors, retailers and consumer groups all have a stake in ensuring that foods are produced within the prescribed specifications. With the formation and expansion of the European Common Market, there has been, in general, a harmonization of the agricultural policy, which necessarily resulted in greater vigilance to pursue authentication of foodstuffs. Similar trends are evident in other parts of the globe where other international markets have emerged. In all instances, the goal has been to protect and support the quality of local produce facilitating its entry into the wider markets [38].

In general, most food is analysed in laboratories by methods based on instrumental analysis. The equipment is diverse and typically includes methods with the use of techniques such as UV–visible spectrophotometry, atomic absorption spectroscopy, inductively coupled plasma-mass spectrometry, X-ray fluorescence analysis, gas and liquid chromatography, near infra-red spectroscopy and mass spectrometry. However, with the exception of the important Karl–Fisher coulometric titration for trace water analysis, potentiometry with the use of ion selective electrodes and more recently, electronic sensors of various kinds, electroanalytical methods are rarely found in routine use for food analysis. On the other hand, electroanalytical methods, and especially the voltammetric, methods are of particular use for trace metal analysis, and as independent reference validation procedures given their applicability for analysis of a wide range of organic and inorganic analytes. Many of these analytes are often able to be analysed simultaneously, and detection limits, in general, are in the μg–mg L⁻¹ range. Electrochemical methods involving for example, voltammetry and polarography, are readily available for quantitative food analysis, including that of additives as well as of the nutritional and spolit contents. In most cases, the responses from such analyses overlap but these complex signals can be often resolved by the application of chemometrics. The main methods and applications are listed in Table 1.

It is important to mention the developments and applications in the field of electronic devices, such as the “electronic tongue” and the “electronic nose”. The two techniques have grown in interest and usage over the past two decades in parallel with the development of the chemometrics methods. Generally, these devices have the same basic structure, which consists of three main components, viz., an array of sensors, a signal processor and a facility for pattern recognition. The sensors in the array are usually non-specific with the result that the measured response is complex with many overlapping electrochemical signals, e.g., in the case of the voltammetric technique the net response would consist of a range of large and small amplitude pulses. This is an ideal situation for the application of chemometrics or multivariate data analysis so that useful information may be extracted [59–62].

Winquist et al. [63] produced an interesting report on the application of the electronic tongue and chemometrics in food analysis. It was demonstrated how such a device was able to classify various samples of fruit juices, still drinks and milk with the use of PCA. In another study [64], a hybrid electronic tongue utilized six working electrodes made of different metals (gold, iridium, palladium, platinum, rhodium and rhodium) and an Ag/AgCl reference electrode. The obtained voltammetric and potentiometric data were submitted to PCA and ANN and these two methods were used for classification of six different types of fermented milk with satisfactory results. Similar devices were also applied for the discrimination of tea samples [65–67]. A voltammetric electronic tongue consisting of four working electrodes (gold, platinum, rhodium and stainless steel) was used to follow milk from different sources during its processing, and then to monitor the clean-up procedures. PCA was applied for the electroanalytical data interpretation [68]. This study clearly showed that milk from different sources, and thus, of different grades could be separated successfully. In another study [69], quantitative information from DPV (differential pulse voltammetry) voltammograms obtained from the direct measurements of three amino acids in animal feed samples by an electronic tongue was also extracted by employing the discrete wavelet transform (DWT) and ANNs. Best results were obtained when two hidden layers were used in a BP neural network. An electronic tongue was applied for milk recognition [70]. Five chemometrics methods were used: K-nearest neighbours (KNN), PLS, soft independent modelling of class analogy (SIMCA), ANN-BP and learning vector quantization (LVQ). The latter method, LVQ, was the preferred one. Electronic tongue was also applied for qualitative analysis of mineral waters and apple juices with the aid of PCA and ANN [71].

The other sensor technique—the electronic nose technology, referred to above, is based on the detection of volatile compounds present in the headspace of a food sample. The measurements are carried out with an array of semi-selective gas sensors. Advantages of electronic nose technology include the relatively minimal sample preparation, and rapid analysis. However, this technique employs sensors that are not very selective for any signature compound of a food sample, and this precludes any possibility of specific identification or quantification of the individual compounds. Such a drawback has obvious implications for food authentication, as an adulterant cannot be identified definitively.

Electronic nose technology in conjunction with chemometrics analysis has been successfully applied for differentiation of wine samples. White, red and rosé Spanish wines were differentiated with the use of a combination of an electronic nose and PCA [72]. Italian wines of different geographical and varietal origin were also successfully separated with a correct classification rate as high as 100% [73]. Authentication of honey samples [74] and determination of the geographical origin of Valencia orange juices [75] and Emmental cheese [76] have also been successfully carried out with the use of a similar device. There are also occasional reports on the application of the electronic noses with non-specific arrays, which were used to analyse, by DPV and multivariate analysis,
occasionally, the electronic tongue, nose and chemometrics remain an area of research that holds considerable potential for future development and application. It is a rapid form of analysis, which can be easily applied, and as indicated above, it is able to assist with the authentication of a wide range of foods. Occasionally, the electronic tongue, nose and chemometrics were used together for food analysis. Linear discriminant analysis (LDA) was applied to study the differences of age-determination of extra virgin olive oils. This new approach could offer a valid alternative for difficult and time-consuming traditional analytical methods and could be a useful tool for on-line or routine determination of olive oil storage conditions. Another work reported the application of the electronic nose and tongue in clinical analysis (urine) and food analysis. Linear discriminant analysis (LDA) was applied to study the differences of age.

3.2. Pharmaceutical and pesticide residues analyses

With the rapid growth of world population, there is an ever increasing demand for agricultural products, and a consequent need for pesticides. Because these substances are generally highly toxic, it is essential to have analytical methodology for monitoring their levels in the environment. The worldwatch institute estimates that over 700 different organic compounds, particularly pesticides and their breakdown products, surfactants, phenols and polycyclic aromatic hydrocarbons, may be found in environment [82]. Also, there are about 70,000 synthetic chemicals in everyday use, with between 500 and 1000 new chemicals being added to the list each year.

In general, electroanalytical methods can determine the pesticides quantitatively, as well as providing information about their degradation mechanisms. Other advantages include good precision, accuracy, selectivity and relatively low costs. In a DPV analysis of pesticide residues, Guiberteau Cabanillas et al. [83] resolved the overlapping signal from a mixture of carbamate pesticides, carbaryl and carbofuran, with the aid of chemometrics. Different chemometric methods (PLS-1, PLS-2 and CLS (classical least squares)) were applied for the resolution of the overlapped peaks of both compounds and their determination in mixtures. It was found that the best results were obtained with the PLS-1 method, which gave a %RPE of 8.2 and 5.1 for the synthetic binary mixtures of carbaryl and carbofuran, respectively. The developed method was applied for the determination of both compounds in river water samples. Reviejo et al. [84] used differential pulse polarography (DPP) to determine organochlorine pesticides, such as dieldrin, heptachlor, endosulfan and endosulfan sulfate. For each mixture, the total concentration of pesticides was 1–9 μM. A PLS model was built with a calibration data set of 35 pesticide samples including single, binary, ternary and quaternary mixtures. Guiberteau Cabanillas et al. [85] reported that the strongly overlapping polarographic signals from binary mix-
tures of atrazine–simazine and terbutryn–prometryn could be resolved, and the two analytes were analysed quantitatively with the use of PLS and ANN calibration models. The results obtained showed that lower standard deviations and better prediction results were achieved with the ANN models. The recoveries of the pesticides in the analysis of spiked river water samples were 92–110% by ANN, compared with 78–122% for PLS-1. Diaz et al. [86] applied DPV and PLS for the determination of the herbicides, paraquat and diquat after their degradation in strong alkaline solution. The method was based on the observed oxidation peak at a glassy carbon electrode. Both paraquat and diquat were determined in the range between 1.0 and 10μg mL⁻¹. This method was applied for the determination of these herbicides in a commercial product and in spiked river water samples with satisfactory results.

A differential pulse stripping voltammetry (DPSV) method of analysis was described for the determination of parathion-methyl, fenitrothion, parathion and isocarboxphos at a hanging mercury drop electrode (HMDE) [87,88]. Quantitative analysis for each of the pesticide compounds in a mixture was investigated with the aid of chemometrics and the prediction performance of the different methods such as CLS, PCR, PLS, KF and radial basis function-artificial neural networks (RBF-ANN; RBF=radial basis function), was compared. A similar method was applied successfully for analysis of the pesticide residues from farm vegetables, and the reliability of the methodology was supported by good recoveries in range of 94–105%. Four carbamate pesticides, propoxur, isoprocarb, carbaryl and carbofuran in dilute perchloric acid electrolyte, were analysed by the DPV method [89]. The pesticide mixtures were found to have well defined voltammetric oxidation waves. Serious overlapping of individual voltammetric peaks was observed from mixtures of these four compounds in aqueous solutions, and different calibration models such as CLS, PCR, PLS and RBF-ANN, were applied for the prediction of unknown pesticides in water. The analytical performance of the prediction models was characterized with the use of various figures-of-merit, and the best results were obtained with the PC-RBF-ANN model (%RPE_T (global or total) = 5.6 and %Recovery (average) = 100).

The development of analytical technologies and methodologies for pharmaceutical analysis is integral for the success of drug manufacture, distribution and application. Pharmaceutical analysis is an essential field to ensure safe and efficient drug administration, and is a classical example, which illustrates the global need and application of quality assurance practices. Such analytical work is carried out with the use of simple or sophisticated instrumentation in areas such as pharmacokinetics, metabolic profiling, therapeutic monitoring, and stability studies. Many electrochemical methods, especially the voltammetric and polarographic techniques, have been described for application in analysis of pharmaceuticals and related materials. Pharmaceutical preparations are complex, often consisting of several drugs mixed in different benign matrices, and thus, the response profiles from voltammetric and polarographic measurements often consist of many overlapping signals. The use of chemometrics becomes essential in order to resolve the composite signals and facilitate the prediction of the component drugs in a preparation. In this overview, we illustrate the contribution of electroanalysis to this field with a summary of methods, which have been successfully applied for such analyses (Table 2).

### 3.3. Environmental analysis

Quality of environment is currently a central topic globally, and has been of increasing importance over the last twenty years. Environmental analysis encompasses the properties of air, water and soil, and is a hugely important topic in its own right. It should be noted that the application of chemometrics to data concerned with the quality of the environment is extensive and significant in its impact, and has been previously discussed elsewhere, e.g. [108]. In this overview, we concentrate only on the contribution of electroanalytical and chemometrics, providing an indication of their modest role in the field. In general, electroanalytical methods are quite competitive for metal analysis with other common methods such as ICP-AES, ICP-MS and ETAAS. This is because the voltammetric methods can be portable, relatively fast, cheap, responsive to simultaneous analysis of groups of elements, and have low Limits of Detection. On the other hand, for organic analysis, electroanalytical methods struggle to compete with HPLC or the hyphenated MS techniques. Thus, the literature abounds with electroanalytical procedures for the analysis of metal ions over a wide range of concentrations and types of method [109]. Examples of quantitative analysis of mixtures of metal ions include the application of MCR-ALS for the interpretation of voltammetric data [110]. The performance of this method was evaluated by the quality of resolution of and prediction from the overlapped ASV peaks obtained in the analysis of binary and quaternary mixtures of Cu, In, Pb and Tl metal ions. MCR-ALS results were validated and compared with those obtained by applying the PLS and CLS methods.

For other multi-way data analysis, Chow et al. [111] used four modified gold electrodes for the simultaneous determination of mixtures of Cu, Cd and Pb at trace concentrations (100 nM to 10μM), and with the use of N-PLS calibration. Other examples, which described the use of electroanalytical procedures for metals in various kinds of environmental samples, are listed in Table 3. They are illustrations of the possible different applications in this class.

Interestingly, a few reports on the application of electronic tongues for environmental monitoring have been published. These included measurements of a number of different chemical species, such as Cu, Cd, Fe, Cr, Zn, Cl, SO₄ and H, in environmental samples by such devices with the aid of nonlinear least squares and ANN models for data interpretation [112]. A large amount of qualitative data was processed, and a significant improvement in prediction of the analytes was obtained. In another study, a voltammetric electronic tongue was used to discriminate the different rinses from 20 machine wash runs in a household washing machine with the aid of PCA and SIMCA [113]. With the PCA model, only one of the rinses was misclassified, and with SIMCA all rinses were classified correctly. Another relatively recent, interesting work
Table 2 – Application of chemometrics and electroanalysis for pharmaceuticals

<table>
<thead>
<tr>
<th>Technique</th>
<th>Chemometrics</th>
<th>Object</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWAdSV</td>
<td>CLS, PLS and ANNs</td>
<td>Nalidixic acid and its main metabolite (OHNA)</td>
<td>[90,91]</td>
</tr>
<tr>
<td>DPP</td>
<td>PCR</td>
<td>Captopril in preparation</td>
<td>[92]</td>
</tr>
<tr>
<td>AdSV</td>
<td>Experimental design</td>
<td>Kynurenic acid</td>
<td>[93]</td>
</tr>
<tr>
<td>CA</td>
<td>PLS</td>
<td>Propionaldehyde</td>
<td>[94]</td>
</tr>
<tr>
<td>CV</td>
<td>PCA</td>
<td>Epinephrine and norepinephrine</td>
<td>[95]</td>
</tr>
<tr>
<td>LSV</td>
<td>CLS, PCR and PLS</td>
<td>Dopamine and ascorbic acid</td>
<td>[96]</td>
</tr>
<tr>
<td>DPAdSV</td>
<td>PCR</td>
<td>Enrofloxacin and ciprofloxacin</td>
<td>[97]</td>
</tr>
<tr>
<td>DPP</td>
<td>PLS</td>
<td>Furazolidone and furaldtadone</td>
<td>[98]</td>
</tr>
<tr>
<td>Amperometry</td>
<td>MLR</td>
<td>Aascobic acid, dopamine, epinephrine and dipryone</td>
<td>[99]</td>
</tr>
<tr>
<td>LSV</td>
<td>MCR-ALS</td>
<td>Aascobic acid, epinephrine and pyrocatechol</td>
<td>[100]</td>
</tr>
<tr>
<td>LV</td>
<td>ANNs</td>
<td>Adenine and cytosine</td>
<td>[101]</td>
</tr>
<tr>
<td>DPV</td>
<td>PLS</td>
<td>Anti-inflammatory drugs</td>
<td>[102]</td>
</tr>
<tr>
<td>LSV</td>
<td>GA-PLS and PC-ANN</td>
<td>Isoniazid and hydrazine</td>
<td>[103]</td>
</tr>
<tr>
<td>LSSV</td>
<td>PCR, PLS and ANN</td>
<td>Ofloxacine, norfloxacin and ciprofloxacin</td>
<td>[104]</td>
</tr>
<tr>
<td>DPSV</td>
<td>PCR and PLS</td>
<td>Panacetamol and phenobarbital</td>
<td>[105]</td>
</tr>
<tr>
<td>DPSV</td>
<td>PCR and PLS</td>
<td>Chlorpromazine and promethazine hydrochloride</td>
<td>[106]</td>
</tr>
<tr>
<td>LSV</td>
<td>WNN</td>
<td>Aascobic acid, 4-aminophenol and paracetamol</td>
<td>[107]</td>
</tr>
</tbody>
</table>

AdSV = adsorptive stripping voltammetry; CA = chronoamperometry; DPAdSV = differential pulse adsorptive stripping voltammetry; DPSV = differential pulse stripping voltammetry; LSSV = linear sweep stripping voltammetry; MCR-ALS = multivariate curve resolution-alternating least squares; MLR = multiple linear regression; SWAdSV = square wave adsorptive stripping voltammetry; PC-ANN = principal component-artificial neural networks; WNN = wavelet neural network; OHNA = 7-hydroxymethylnalidixic acid. Methods not listed appear in Table 1.

involved the use of a formic acid biosensor for air monitoring and the response data were evaluated using a combination of PCA and MLR [114]. The most important factor for the biosensor performance was found to be the enzyme concentration, and the optimum operation conditions for the biosensor were determined with the use of the information from the chemometrics analysis. The steady state currents were increased by 18–30% and the initial two response rates by 47–89% compared with a biosensor, which had not been optimized.

Table 3 – Application of chemometrics and electroanalysis on environmental samples

<table>
<thead>
<tr>
<th>Technique</th>
<th>Chemometrics</th>
<th>Object</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPV</td>
<td>PCR and PLS</td>
<td>Nitrobenzene and nitro-substituted phenols</td>
<td>[115]</td>
</tr>
<tr>
<td>ACSV</td>
<td>PCA and HCA</td>
<td>Elements in seawater</td>
<td>[116]</td>
</tr>
<tr>
<td>DPASV</td>
<td>PCR, PLS and ANN</td>
<td>Tl and Pb in synthetic mixtures</td>
<td>[117,118]</td>
</tr>
<tr>
<td>DPAdSV</td>
<td>PLS</td>
<td>Sb(III) and Sb(V)</td>
<td>[119]</td>
</tr>
<tr>
<td>DPP</td>
<td>PCR and PLS</td>
<td>Zn and Ni in water samples</td>
<td>[120]</td>
</tr>
<tr>
<td>DPV</td>
<td>WPT-ERNN</td>
<td>Ni, Zn and Cd in synthetic mixtures</td>
<td>[121]</td>
</tr>
<tr>
<td>DPAdSV</td>
<td>PLS</td>
<td>Sb in pharmaceutical preparations</td>
<td>[122]</td>
</tr>
<tr>
<td>ACV</td>
<td>HPCR and HPLS</td>
<td>Copper electroplating</td>
<td>[123]</td>
</tr>
<tr>
<td>DPP and DPASV</td>
<td>PLS</td>
<td>Zn, Cd, Pb and Cu in tap and river water</td>
<td>[124,125]</td>
</tr>
<tr>
<td>DPP</td>
<td>PLS</td>
<td>Pb, Sn and Cd in synthetic samples</td>
<td>[126]</td>
</tr>
<tr>
<td>DPP</td>
<td>GA-NNs and CPR</td>
<td>Tl and Pb in synthetic mixtures</td>
<td>[127,128]</td>
</tr>
<tr>
<td>DPASV</td>
<td>PLS</td>
<td>Cu in presence of Fe</td>
<td>[129]</td>
</tr>
<tr>
<td>DPAdSV</td>
<td>PLS</td>
<td>Cr(VI) and Cr(III) in water</td>
<td>[130,131]</td>
</tr>
<tr>
<td>ASV</td>
<td>PCR and PLS</td>
<td>Pb, In, Cd and Tl in synthetic mixtures</td>
<td>[132]</td>
</tr>
<tr>
<td>DPV</td>
<td>ANNs</td>
<td>Cd and Pb in synthetic mixtures</td>
<td>[133]</td>
</tr>
<tr>
<td>ASV and PSA</td>
<td>PCR</td>
<td>Zn and Cu in brass samples</td>
<td>[134]</td>
</tr>
<tr>
<td>ASV</td>
<td>MCR-ALS</td>
<td>Cu, In, Pb and Tl in synthetic mixtures</td>
<td>[135]</td>
</tr>
<tr>
<td>DPASV</td>
<td>PC-ANN and WNN</td>
<td>Cu and Cu-PAN (Cu-XG)</td>
<td>[136]</td>
</tr>
<tr>
<td>ASV</td>
<td>CWT</td>
<td>Cu and Bi in water and human hair</td>
<td>[137]</td>
</tr>
<tr>
<td>DPSS</td>
<td>ANNs</td>
<td>Cu, Pb and Cd</td>
<td>[138]</td>
</tr>
<tr>
<td>DPSS</td>
<td>PC-ANN</td>
<td>Cu-Mo in river, tap water and alloy</td>
<td>[139]</td>
</tr>
<tr>
<td>DPP</td>
<td>PCR, PLS and ITTFA</td>
<td>Pb, Cu, V, Cd and Ni in water</td>
<td>[140]</td>
</tr>
</tbody>
</table>

ACV = alternating current voltammetry; ASV = anodic stripping voltammetry; DPASV = differential pulse anodic stripping voltammetry; PSA = potentiometric stripping analysis; CPR = continuum power regression; CWT = continuous wavelet transform; GA-NN = genetic algorithms-neural networks; HCA = hierarchical cluster analysis; HPCR = hierarchical PCR; HPLS = hierarchical PLS; WPT-ERNN = wavelet packet transforms-Elman recurrent neural network; WNN = wavelet neural network; PAN = 1-(2-pyridylazo)-2-naphthol; XO = xylene orange. Methods not listed appear in Tables 1 and 2.
A range of illustrative electroanalytical applications for which complex response data were interpreted by chemometrics is listed in Table 3.

4. Application of chemometrics and electroanalysis for qualitative analysis

Qualitative analysis is an important part of analytical practice. Such analyses often form the initial part of the work, and its reliability and the subsequent conclusions are dependent on the quality of the initial qualitative analysis.

In this section, we provide an overview of the impact of chemometrics on the resolution and extraction of qualitative information from complex, highly overlapping electrochemical responses. Most studies are concerned with biochemical analytes—metallic or organic in nature, and indicate the efficacy of the electroanalytical techniques to sample complex biological systems, and resolve their responses with the aid of chemometrics. This enables the identification of the individual analyte components.

Esteban et al. [7] applied the MCR-ALS method to data collected by different voltammetric techniques for the analysis of processes involving metal complexes. They showed that this factor analysis approach is a powerful tool for solving such problems. In another study, the complexity of Cd²⁺ by glutathione (GSH) was investigated by DPP [141], and the simultaneous analysis of the titration of the peptide with the metal, and of the metal with the peptide. This allowed the resolution of the Cd²⁺/Cys-Gly system with the aid of MCR-ALS [142]. The metal-phytochelatin interactions were studied with the use of DPV [143], while the binding properties of Cd²⁺-GSH were investigated by CV [144], and the results were resolved by MCR-ALS. The complexation between Zn²⁺ and GSH and their fragments γ-Glu-Cys and Cys-Gly were studied by DPV and again with the aid of MCR-ALS [145,146]. The formation of 1:1 and 1:2 Cd:GSH complexes was detected simultaneously with the assistance of the extracted concentration profiles of the two analyte complexes. Elsewhere, this chemometrics method was also applied for the resolution of voltammograms of Zn⁴⁺-peptide and Zn⁴⁺-glycine complex systems [147,148]. The stability constants of these complexes were estimated from the concentration profiles obtained by means of the ALS optimization. The results were satisfactory, which suggested that soft modelling can provide good estimates of the complexation parameters, and can be a very useful tool in cases where hard modelling cannot be applied.

Electroanalytical techniques such as SWV and DPP [149–152] were applied to resolve the formation of complexes of Cd and Zn ions, which were present in a mixture with the ligand, Lys-Cys-Thr-Cys-Cys-Ala (56-61) (FT). The biological effects of these two metal ions are very different with the former being regarded as toxic, whilst the latter is an essential element. On the other hand, the two ions also have some similar chemical properties. The complex measured responses were satisfactorily resolved with the use of the MCR-ALS method. This allowed both the voltammograms of individual components and their concentration profiles to be extracted. On the basis of these results, the electrochemical reduction processes and some structures were obtained for the chemical species in the mixed metal Cd-Zn-FT system. The complex formation between fulvic acid and Cd²⁺ was studied by ASV, and the voltammetric data obtained during the titration of Cd²⁺ and fulvic acid at different concentrations were analysed by MCR methods [153]. Apart from the free Cd²⁺ work, two more papers related to the formation of complexes were found, and the corresponding pure voltammograms and concentration profiles were resolved and calculated, respectively. Three other interesting investigations involved metal complex formation. The first applied the constant current chronopotentiometric stripping analysis with the use of adsorptive accumulation (AdSCP) on an HMDE to study complex formation between Zn²⁺ and GSH and (γ-Glu-Cys)₂-Gly (PC2) (note: PC—proprotein convertase; ‘2’ represents two (γ-Glu-Cys) potential ligands). MCR-ALS was successfully applied to interpret the DP voltammetric and chronopotentiometric measurements of the Zn-PC2 [154]. The second and third quite recent contributions were concerned with the complex formation of Cd²⁺ and Zn²⁺ with phytochelatin (γ-Glu-Cys)₄-Gly (PC4) [155] as well as the complex formation of Pb²⁺ with phytochelatins involving Cys-Gly, γ-Glu-Cys, (γ-Glu-Cys)₂-Gly (PC2) and (γ-Glu-Cys)₃-Gly (PC3) as ligands [156]. These chemical systems were investigated by DPP, and the polarograms were interpreted by MCR-ALS models. The results established the complex formation sequence and final stoichiometries of the complex species.

5. Comparison of method performance—choices of techniques and methods

Research and development of analytical methods ultimately aims not only to progress science and technology, but also focuses on the potential applications of the new methods in practice. In this context, it is particularly useful to have some comparative data on the analytical performance of techniques and methods. Thus, in this section, the analytical performance of some common electrochemical methods is compared with that of some spectrophotometric methodologies. The discussion is not exhaustive but rather intends to explore and provide some guidelines of the performance of electroanalytical methods in relation to:

(1) various chemometrics calibration models;
(2) comparison with method performance of some spectrophotometric methods.

A number of papers based on spectrophotometry have compared the performance of several multivariate calibration models. It should be noted that in these papers calibration and validation data sets were based on the statistical orthogonal factorial design [157]. For example, ten different models were applied for the simultaneous prediction of five metal ions (Co(II), Cu(II), Ni(II), Fe(III) and Cr(III)) in industrial electroplating solutions [158]. The methods includedCLS, PCR, PLS, ANN-BP and ANN-RBF and were applied to the complex first order spectra in the 370–760 nm range. In addition, the
five methods were applied to the first derivative spectra. The limit of detection range for the analytes was ca. 1–50 mg L\(^{-1}\) and \%RPE\(_T\) varied from 3.7\% (DCLS) to 9.2\% (DANN-BP). The PCR (%RPE\(_T\) = 8.4) and the PLS (%RPE\(_T\) = 8.4) performed adequately given the overall operational concentration range (ca. 5–250 mg L\(^{-1}\)). A spectrophotometric study targeting completely different analytes, namely carbamate pesticides [159], included a comparison of some eight multivariate calibration models NPLS1 (%RPE\(_T\) = 6.7) and NPLS2 (%RPE\(_T\) = 9.3), PARAFAC (%RPE\(_T\) = 11), ANN-BP (%RPE\(_T\) = 10), ANN-RBF (%RPE\(_T\) = 3.2) and several variations on the latter including the PC-ANN-RBF (%RPE\(_T\) = 3.8) approach. The last calibration model is an interesting approach which utilizes data reduction and the extraction of principal components as a pretreatment. The ANN operations were then performed on the scores of the statistically significant PCs, i.e., using the data matrix free of residuals. The Limits of Detection range for the three pesticide analytes was 0.12–0.26 mg L\(^{-1}\). As is evident from the above \%RPE\(_T\) information, the ANN-RBF and the PC-ANN-RBF stood out from the other models in performance.

A further important method selection criterion is the %Recovery value. For both studies described above, the %Recovery values were generally satisfactory for most multivariate calibrations being usually in the range ca. 90–110\%. Also, for the two studies, it was possible to investigate the overall performance of the methods on the basis of the three method performance criteria: %RPE\(_T\), %Recovery and %RPE\(_T\). This was carried out with the use of the non-parametric PROMETHEE and GAIA ranking routines [160]. This work indicated that the DPLS and the DPCR methods were to be preferred for the simultaneous determination of metal ions, while for the carbamate pesticide analysis, ANN-RBF and the PC-ANN-RBF models performed best.

The above summary of the two different spectrophotometric studies was intended to indicate the scope of calibration modelling available to the practising analyst and to utilize the discussion and the conclusions as a baseline for a comparative method selection process. Other spectrophotometric studies in which a smaller number of calibration models have been investigated, indicate general support for the above performance of the models. Thus, where ANN modelling is involved there seems to be a trend for PC-ANN-RBF or ANN-RBF calibration models to perform best overall; when the more conventional calibration models are used CLS tends to perform poorly (often RPE\(_T\) is 20–60\% and the %Recoveries are poor and erratic), and PLS-1 models are somewhat better than PLS-2 and PCR ones. For the PLS-1 calibrations, RPE\(_T\) values are generally in the range of about 5–10\%, and the %Recovery values are satisfactory as described above [161–167].

The range of electroanalytical studies with the use of multivariate modelling does not appear to be as extensive as that for spectrophotometry. However, there is sufficient information available to make some comparisons. Thus, the work on simultaneous analysis of the four carbamate pesticides with the use of the DPV method [89], offers a number of multivariate calibration models for performance comparison. The five methods considered included PCR (%RPE\(_T\) = 13.2), PLS (%RPE\(_T\) = 11.8), ANN-RBF (%RPE\(_T\) = 8.4), and PC-ANN-RBF (%RPE\(_T\) = 5.6). The Limits of Detection were in the range of 0.4–0.8 mg L\(^{-1}\) and the %Recovery values were satisfactory.

A similar analysis but one involving the organophosphorus pesticides [87] included the methods: CLS (%RPE\(_T\) = 40.1), PCR (%RPE\(_T\) = 17.1), PLS (%RPE\(_T\) = 17.1), ANN-RBF (%RPE\(_T\) = 8.4), and the Kalman filter (%RPE\(_T\) = 41.2). The %Recovery values were satisfactory for the PCR, PLS and the ANN-RBF methods but were very erratic for the three analytes based on the CLS and the Kalman filter models. The Limits of Detection were in the range of 4–5 μg L\(^{-1}\). Other electroanalytical studies with less number of multivariate models explored, reflect similar trends although sometimes the ANN-RBF (%RPE\(_T\) = 8.1) models perfumed significantly better than the PLS one (%RPE\(_T\) = 18.4) [68,104,105]. Thus, in general, the first two electroanalytical studies [87,89] reflect effectively the same general trends observed with the spectrophotometric methods.

In the context of comparison of the analytical performance of methods from different techniques, i.e., electroanalysis and spectrophotometry, the above comments are quite important because they are supported by figures-of-merit obtained with different analyte mixtures from studies spread over several years. What they indicate is that irrespective of the analytical technique or method, the ANN-RBF calibration modelling, preferably prefaced by PC data pretreatment, is likely to produce the best set of figures-of-merit. Otherwise, of the more conventional calibration methods, PLS-1 seems to be the method-of-choice. Its performance may be improved if, where relevant, the response profile is converted to the first derivative format. Thus, arguably, within the μg–mg L\(^{-1}\) concentration range of analyte mixtures, the analytical performance of the spectrophotometric and the electroanalytical methodologies is very similar.

From this perspective, the practising analyst then has to decide if there is the capability within the laboratory to develop ANN calibration models. These can be quite sensitive to modelling conditions and provide considerable challenges to the user. Importantly, such models are as yet not readily available through convenient commercial packages, which most commonly include PLS as the method-of-choice for prediction.

For the practising analyst, the choice of the analytical method would be determined by other criteria, perhaps the foremost of which is the familiarity with the technique and its background, operator training, sample throughput and costs. We would suggest that the first of these criteria is the critical one because many potential scientists and technologists find electrochemistry, in general, and electroanalysis in particular, a difficult field. This is a challenge for the electrochemical fraternity to overcome. If only because, unquestionably, electrochemistry and -technology will continue to grow in importance in their application in our global society in the foreseeable future.

### 6. Conclusion

This review explores the question whether chemometrics methods enhance the performance of electroanalytical methods. The sheer number of referenced papers, which applied electroanalysis to collect qualitative and quantitative data and used chemometrics methods for data interpretation, is evidence of a strong affirmative reply to this question.
It has been demonstrated that both the well-established electroanalytical techniques such as potentiometric titrations, polarography and voltammetry, and the more novel ones such as electronic tongues and noses either benefit significantly from or indeed require the application of chemometrics for data analysis.

Chemometrics is now broadly used in electroanalysis for modelling multivariate calibrations in a wide range of important areas of application such food, pesticides, pharmaceuticals and the environment.

Chemometrics is essential for the extraction of qualitative information in the form of unique profiles of analytes found in complex mixtures of biological and metal complex systems.

There now is sufficient evidence available, which enables the performance of electroanalytical methods to be undertaken on the basis of figures-of-merit. In this context, in general, it appears that electroanalytical methods perform as well as spectrophotometric ones under comparable analytical requirements. It is also apparent that the PLS-1 calibrations are satisfactory to apply so long as the Relative Error of Prediction of about ±10% is acceptable. However, in many respects, PLS may be regarded as the first-stop method-of-convenience, principally because it is well known and used. It should also be noted that ANNs methods have consistently shown better figures-of-merit performance, and could be considered as useful alternatives to PLS if tighter %RPEs were required.

### Acknowledgements

The authors gratefully acknowledge the financial support of this study by the National Natural Science Foundation of China (NSFC20562009), the State Key Laboratories of Electroanalytical Chemistry of Changchun Applied Chemical Institute (SKLEAC2004-3) and the Jiangxi Provincial Natural Science Foundation (JXNSF062004).

### Appendix A

#### A.1. Summaries of chemometrics methods

In this section, essential aspects of the main chemometrics methods referred to in the above manuscript are summarized.

**A.1.1. Classical least squares [168,169]**

In voltammetric measurement of multicomponent mixtures, it is assumed that there is a proportional relationship between the peak current and the concentration:

\[ r_i = k_2c_2 + k_1c_1 + \sum_{j=0}^{m} k_j c_j \quad (i = 1, 2, \ldots , n) \]  

(18)

where \( r_i \) is the voltammetric current of the mixture at the \( i \)th potential point; \( c_j \) is the concentration of the \( j \)th component in the mixture; \( k_j \) is the proportional coefficient of the \( j \)th component at the \( i \)th potential point, and \( k_0c_0 (c_0 = 1) \) is the bias term responsible for the residual current. It varies with the applied potential and, in general, cannot be neglected. By measuring the current at \( n \) potential points, \( n \) equations are obtained, which can be expressed in matrix format as:

\[ r^T = c^T K \]  

(19)

where \( r \) and \( c \) are vectors. The coefficient matrix, \( K \), and, \( c \), may be obtained experimentally. Specifically, in voltammetric analysis, a calibration data set may be represented by extending Eq. (19), as:

\[ R = CK \]  

(20)

\( R \) is the current matrix corresponding to the voltammograms, \( C \) is the concentration matrix from the standards, and \( K \) can be obtained by applying the least squares method of Eq. (20). If \( K \) is known, the concentration of unknown sample can easily be obtained according to the following equation:

\[ c_{\text{unk}}^T = r_{\text{unk}}^T K^T (K^T K)^{-1} \]  

(21)

However, the inversion of \( K^T K \) will result in almost singular matrices and no statistical solution can be obtained when many complex mixtures are analysed, and consequently this method has limited satisfactory application in multivariate prediction.

**A.1.2. Principal component regression and partial least squares**

Principal component regression and partial least squares are factor analysis multivariate methods of data interpretation and have been successfully applied to electrochemical analysis of multicomponent mixtures [170,171]. The PCR decomposition of a data matrix is based entirely on current response variations without regard the concentrations of the analytes. For PLS, the response decomposition is weighted to the concentration. The major difference in the predictive abilities of these two methods is that PLS seems to predict better than PCR when there are random linear baselines or independently varying major response components, which overlap with the signal’s features [172].

PCR involves decomposition of \( R \) response matrix into latent matrices:

\[ R = TP^T \]  

(22)

where \( T \) is the score matrix of \( R \) with \( p \) rows and \( d \) columns (\( d \) is the number of principal components of the system) and \( P^T \) is the loading matrix of \( R \), and has \( d \) rows and \( n \) columns. The regression equation involving \( C \) and \( T \) can be written as:

\[ C = TG \]  

(23)

where \( G \) is the regression coefficient matrix for the regression between \( C \) and \( T \). The concentration of the unknown can then be obtained from:

\[ c_{\text{unk}}^T = r_{\text{unk}}^T PG \]  

(24)
The PLS method is carried out by decomposing the matrices $R$ and $C$ at the same time, and in addition to Eq. (22) the following process is also effected:

$$ C = UQ^T $$

(25)

where $U$ is the $m \times d$ concentration score matrix and $Q^T$ is the $d \times (p + 1)$ loading matrix. Relating the score matrices $U$ and $T$, gives the following equation:

$$ U = TB $$

(26)

where $B$ is the diagonal matrix of the regression coefficients and is then used in the prediction step for the estimation of the composition of the unknowns from the measured current data, $\hat{\tau}_{unk}^T$, of the unknown sample:

$$ c_{muk}^T = \hat{\tau}_{unk}^T (U^T R^T)^T B Q^T $$

(27)

The matrices $U$, $Q$ and $R$ are known from the PLS calibration model. The prediction of an unknown sample can also be carried out by an iterative method [173].

A.1.3. Artificial neural networks

The neural networks perform a non-linear iterative fit of data [174]. The structure of a network consists of three node layers: an input layer, a hidden layer and an output layer. The nodes in the input layer transfer the input data to all nodes in the hidden layer. These nodes calculate a weighted sum of the inputs that is subsequently subjected to a non-linear transformation

$$ s_j = f \left( \sum_{i=1}^{I} s_i w_{ij} \right) $$

(28)

where $s_i$ is the input to the node $i$ in the input layer; $I$ is the number of nodes in the input layer; $w_{ij}$ (weights) are the connections between each node $i$ in the input layer; each node $j$ in the hidden layer, $o_j$, is the output of node $j$ in the hidden layer, and $f$ is a non-linear function (usually a sigmoid function is used):

$$ f(x) = \frac{1}{1 + e^{-x}} $$

(29)

The output of the network is a weighted sum of the outputs of the hidden layer, and it is the calculated concentration. During the training process (i.e., calibration) the weights are iteratively calculated in order to minimize the sum of squares of the residuals. The weight correction, $\Delta w_{ij}$, is defined as follows:

$$ \Delta w_{ij}(n + 1) = \eta \delta_j o_i + \alpha \Delta w_{ij}(n) $$

(30)

where $\delta_j$ is the error term, $\eta$ is the learning rate, $\alpha$ is the momentum and $n$ is the iteration number. The iteration is completed when the error of prediction reaches a minimum.

A.1.4. Multiple curve resolution methods

A.1.4.1. PARAFAC and NPLS. A straightforward approach to three-way data decomposition, which avoids some of the inherent problems of three-way analysis, is to unfold the three-way data array into a two-way array, i.e., a rectangular data table, which can be processed according to the methods of general chemometrics, such as BP-ANN, RBF-ANN, PLS, etc. The operation of the unfolding process is analogous to the spreading of a three dimensional stack of cards into a two-dimensional layout. However, the unfolding of such a stack can be carried out in two different ways, i.e., an $m \times p \times k$ array spread into an $(m \times p) \times k$ or $(m \times k) \times p$ rectangular data table. Application of chemometrics methods to each of these two unfoldings will generally produce a different result. In particular, one obtains two different biplots for one and the same three-way table, which may be undesirable.

The decomposition of a three-way matrix, $A$, can also be achieved by means of the parallel factor analysis (PARAFAC) model in terms of three two-way loading matrices (one for each mode) $D$, $B$ and $C$ [175,176]:

$$ A_{mnpk} = \sum_{f=1}^{F} d_{mf} b_{pf} c_{pkf} + \epsilon_{mnpk} $$

(31)

where $\epsilon_{mnpk}$ represents the residual term. This decomposition is also referred to as the trilinear model. The $m \times p \times k$ three-way matrix $A$ is decomposed into the $m \times f$, $p \times f$ and $k \times f$ loadings matrices, $D$, $B$ and $C$ for the row-, column- and layer-items of $A$, respectively. In this model, the three loading matrices $D$, $B$ and $C$ are not necessarily orthogonal [177]. However, the solution of the PARAFAC model is unique and does not suffer from the indeterminancy that arises with principal components and factor analysis in the unfolding method. The number of factors in each model is the same. This number is chosen to be much smaller than the original number of dimensions of the data matrix in order to achieve a considerable data reduction. The elements of the loading matrices $D$, $B$ and $C$ are computed so as to minimize the sum of squares of the residuals.

The trilinear PLS algorithms are straightforward extensions of the PLS algorithm [178]. For the trilinear PLS regression, PARAFAC-like trilinear structure of the independent data is used. However, the trilinear components are calculated such that the scores are predictive for the dependent variables as in ordinary two-way regression. Successful applications can be found in references [179,180].

A.1.4.2. Multivariate curve resolution with alternating least squares (MCR-ALS). Alternate least squares (ALS) is a well known soft modelling algorithm developed by Tauler et al. [181,182], which has been applied to multiequilibria chemical systems [7,141,183]. ALS facilitates the estimation of concentration profiles of each component in the reaction as well as the extraction of the corresponding pure responses.

For the multi-equilibria response data matrix, $X$, it is necessary to nominate the number of chemical species, $N$, which when estimated correctly, will minimize the residual term, $E$. The estimate of $N$ is commonly facilitated by methods such as singular value decomposition (SVD) [184,185], evolving factor analysis [186], or pure-variable detection methods, e.g., SIM-
PLISMA [187]. The rank of $X$ calculated by any of these methods is assumed to be the number of chemical species, $N$.

The ALS process is started by initializing the concentration profiles, which leads to a constrained ALS optimization and eventually extracts the correct set of concentration profiles and pure individual responses. This extraction process is based on the assumption that the instrumental responses of the chemical species are bilinear and can be expressed by the equation:

$$ X = CS^T + E $$  \hspace{0.8cm} (32)

where $X$ is the data matrix with NR (number of spectra) rows and NC (number of wavelengths) columns, $C$ is the NR x $N$ dimensional concentration matrix, $S^T$ is the $N$ x NC dimensional pure spectral matrix, and $E$ is the residual matrix. Given $N$, the initial estimation of the individual response may be obtained by the application of, e.g., the EFA method. Then, the ALS algorithm is performed to calculate the component matrix describing the data ($C$ and $S^T$ in Eq. (32)) by repeatedly alternating between the following two calculations until convergence. For example, if $S$ is the initial estimation used to start the iterative process, in the first step, an estimation of the concentrations $C$ is obtained by least squares regression as:

$$ C = XS(S^TS)^{-1} $$  \hspace{0.8cm} (33)

In the second step, this new estimation of the $C$ matrix can then be used to recalculate by least squares (LS) a new estimation of the species responses, $S$:

$$ S = X^TC(C^TC)^{-1} $$  \hspace{0.8cm} (34)

The LS solutions so obtained are purely mathematical, and may not be appropriate from the chemical perspective, e.g., they may have negative concentrations, and the spectral shapes may be unreasonable. Thus, each time Eqs. (33) and (34) are applied, they are submitted to constraints which require compliance with (i) all negative values of concentrations and responses are discarded, (ii) unimodality, (iii) selectivity and (iv) closure [181,188–190].

REFERENCES