Electrochimica Acta 127 (2014) 68–78

Contents lists available at ScienceDirect

Electrochimica Acta

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

Multivariate linear regression with variable selection by a successive projections algorithm applied to the analysis of anodic stripping voltammetry data

Paola D. Marreto[a], Alexsandro M. Zimer[a], Ronaldo C. Faria[a], Lucia H. Mascaro[a], Ernesto C. Pereira[a], Wallace D. Fragoso[b], Sherlan G. Lemos[b,*]

[a] Laboratório Interdisciplinar de Eletroquímica e Cerâmica (LIEC), Universidade Federal de São Carlos (UFSCar), CP 676, 13565-905, São Carlos, SP, Brazil
[b] Departamento de Química, Universidade Federal da Paraíba (UFPB), CP 5093, 58051-970, João Pessoa, PB, Brazil

ARTICLE INFO

Article history:
Received 19 October 2013
Received in revised form 5 February 2014
Accepted 7 February 2014
Available online 19 February 2014

Keywords:
Anodic stripping voltammetry
Mercury film electrode
Chemometrics
Multivariate calibration
Variable selection

ABSTRACT

Multivariate linear regression aided by a successive projections algorithm (SPA-MLR) was applied in the evaluation of anodic stripping voltammetry data obtained in the simultaneous determination of metals under conditions where there were significant complications due to interference processes such as the formation of intermetallic compounds and overlapping peaks. Using simulated data, modeled from complex interactions experimentally observed in samples containing Cu and Zn, as well as Co and Zn, it was demonstrated that SPA-MLR selected variables that allow chemical interpretation. This feature was used to make inferences about the underlying electrochemical processes during the simultaneous determination of four metals (Cu, Pb, Cd, and Co) in a concentration range where all responses were complicated by interference processes (10-100 ng mL^{-1}). Additionally, the analytical performances of MLR models for quantitative predictions were excellent despite the complexity of the system under study.

1. Introduction

In electroanalytical chemistry, the traditional approach is the use of univariate models, i.e., calibration curves constructed from a single characteristic of the sample. In voltammetry, for example, the intensity of the peak current is the most commonly used variable. A different approach called multivariate modeling, which is less sensitive to the presence of interfering substances, uses more than one variable simultaneously. This is equivalent, for example, to construct a calibration model with the current intensity measured at various potentials in the voltammogram. Multivariate linear regression (MLR) is a multivariate natural expansion of univariate linear regression and is the simplest procedure to perform a multivariate calibration. However, this technique is not efficient if there is significant collinearity in the data matrix (as for voltammetric data), and if the number of variables is greater than the number of samples. The most common way to avoid this problem is the use of latent variables or latent structures methods. These methods use linear combinations of the original variables as new variables. Commonly, a small number of orthogonal or almost orthogonal latent variables can be obtained and used for multivariate calibration.

The selection of predictor variables prior to multivariate regression is a practice that can provide significant improvement in the prediction results when compared to the use of full data (spectra or voltammogram), mainly by discarding variables not related to the analytical response and which only incorporate noise into the regression model [1–3]. Another feature that has been highlighted as an advantage of variable selection is preservation of the original variables domain when the calibration is based on MLR. Consequently, it is easy to propose the physical interpretation of mathematical models, in contrast to what is obtained with methods based on latent structures such as partial least squares (PLS), and principal component regression (PCR). In practice, this advantage has not been greatly explored mainly due to focus on the predictive ability of the models and their comparison with different methods [4]. Additionally, there is little demand for the physical interpretation of variables obtained using spectroscopy techniques because the important spectral ranges are commonly known [5].

Voltammetric measurements have been associated with multivariate calibration to provide simultaneous determinations in a multicomponent system [6–15]. Although this is a growing trend,

* Corresponding author. Tel.: +55 83 3216 7438; fax: +55 83 3216 7438.
E-mail addresses: sherlan03@yahoo.es, sherlan@quimica.ufpb.br (S.G. Lemos).

http://dx.doi.org/10.1016/j.electacta.2014.02.029
0013-4686/

there are fewer successful cases described in the literature compared with spectrometric data. An important difference between these two techniques is that, in voltammetric results, interference processes can introduce new voltammetric peaks such as those related to intermetallic compounds in the simultaneous determination of metallic ions by anodic stripping voltammetry (ASV). Since this kind of interference is related to the metal concentrations themselves, new signals must be included in the calibration model, which necessarily becomes a multivariate model. It is, however, very difficult to obtain information about the signals of intermetallic species. For this reason, many researchers prefer to build multivariate calibration models with all voltammograms, using latent structure methods and including many signals without correlation with the metal concentration [16].

Wang [17], in his famous book, asserts that interference processes in ASV usually occur when metallic thin films are used as the working electrode. These substrates offer many advantages over bulk metal electrodes such as an improved scope for different cell configurations and for chemical modification of their surface, lower cost (requires only minute quantities of the metal to assemble the film), a larger surface-to-volume ratio, and mechanically stability (in comparison with mercury drops, for example) [18]. However, due to the larger surface-to-volume ratio, metallic films are more susceptible to the formation of intermetallic compounds. In this case, the common approaches for eliminating or correcting such interference include removing the interfering response by adding a masking substance [19,20] or the single standard addition method [21–23].

In this work, we intended to evaluate MLR coupled to variable selection performed by the successive projections algorithm [24] (SPA-MLR) in multivariate calibration for voltammetric data using simulated voltammograms and real measurements in the simultaneous determination of metals by anodic stripping voltammetry. The focus was on observing if variable selection could reveal variables apparently not related to a specific analyte, denoting unknown electrochemical phenomena such as coupled reactions, the formation of intermetallic compounds, matrix effects, etc. It was expected that this guided soft modeling semi-empirical approach would provide a very useful qualitative interpretation of the analyzed electrochemical system in addition to quantitative predictions in the same way as hard modeling [25,26].

First, simulated voltammograms were employed in order to explore and discuss the performance of SPA-MLR under a variety of conditions of intermetallic compound formation and peak overlap. Two scenarios were studied. The first one involved the simultaneous determination of Cu and Zn considering the formation of an intermetallic compound CuZn. Pb was also introduced in this simulation in order to evaluate whether the selection algorithm would select variables not related to the analytes, once we considered that there is no interaction of Pb with Cu and Zn. It is important to point out that Cu-Pb intermetallic effects have been reported in ASV [27,28]. However, no intermetallic formation was considered in the present simulations. The second scenario involved the simultaneous determination of Cu, Zn and Co, considering the formation of the intermetallic compounds CuZn and CoZn.

SPA-MLR was also submitted to the analysis of a real system containing five metals that interact to varying degrees in a single sample. In the present work, the simultaneous DPASV analysis of Cu, Co, Pb, Cd, and Zn in water samples with a mercury thin-film electrode (MTFE) was investigated. All analytes were studied in the range from 10 to 100 ng mL^{-1}. Thus, several sorts of interactions are expected such as intermetallic compounds between Cu and Zn, Cu and Cd, and between Cu and Co, to name a few. They result in current-concentration relationships that are difficult to represent using a simple univariate regression model.

2. SPA-MLR background

The SPA is a variable selection algorithm that uses sets of calibration and validation data containing instrumental responses (**X**) and parameter values measured by a reference method (**y**). It initially defines the $\mathbf{X_{CAL}}$ matrix ($K_C \times J$), where rows correspond to K_C samples of the calibration set and columns correspond to J variables, which in the context of this work are the intensities of current at each potential of the full voltammogram. From a $\mathbf{X_{CAL}}$ column matrix, arbitrarily chosen and named $\mathbf{x_0}$, SPA determines which of the other columns has the largest projection in the subspace S_0 orthogonal to $\mathbf{x_0}$. This column is named $\mathbf{x_1}$ and can be interpreted as containing the largest amount of information not included in $\mathbf{x_0}$. In the next iteration, the SPA uses $\mathbf{x_1}$ as the new reference column, and proceeds as above to select $\mathbf{x_2}$. The algorithm follows with projections until a certain number of variables potential minimally collinear with each other is selected.

The maximum number of variables that can be selected is K_C, since the dimension of the column space of $\mathbf{X_{CAL}}$ is reduced by one after each iteration, i.e., one degree of freedom is removed. Therefore, after K_C iterations, all column vectors of $\mathbf{X_{CAL}}$ have been projected on the origin of the space and $\mathbf{X_{CAL}}$ will become a matrix of null rank. Several variable chains are built upon the selection of each one of J variables as the initialization vector $\mathbf{x_0}$ in successive projections, and varying the length N of the chains, typically between 2 and K_C. The best variable chain is selected by building a MLR model for each chain, and validated with a validation samples set constituted by new samples that did not enter the calibration set. The chain of choice is one that corresponds to the MLR model that has the lowest root mean square error of validation.

The next stage of SPA-MLR is necessary because the construction of the variable chains takes place solely based on minimizing the collinearity between the variables and does not take into account the correlation between each variable and the response, i.e., concentration of the analyte contained in the vector **y**. Thus, a procedure that eliminates the variables has been selected but not correlate with the concentrations is performed in order to obtain a simpler model.

3. Experimental

3.1. Simulated data

Data simulation was based on the approach used by Stromberg and Gorodovykh [29], who modeled the Cu-Zn system in which the intermetallic compound is apparently insoluble. Eq. 1 shows the expression used for simulating responses characteristic of redox species in a thin-layer [30]:

$$i = \frac{n^2F^2\nu VC}{RT}\frac{exp\left[\frac{nF}{RT}(E-E^0)\right]}{\left\{1+exp\left[\frac{nF}{RT}(E-E^0)\right]\right\}} \quad (1)$$

In this equation, n is the number of electrons involved in the reaction, F is the Faraday constant, ν is the scan rate, V is the volume of the solution and R the gas constant, E is the applied potential and E° is the formal potential of the species. Then, it is possible to calculate the current, i, for a concentration C of the analyte, which in anodic stripping voltammetry is represented by the concentration of the mercury electrode–C_{Hg}–given by eq. 2.

$$C_{Hg} = \gamma C \quad (2)$$

In eq. 2, γ is the accumulation coefficient and is dependent on deposition time, mass transport of the analyte to the electrode surface, concentration of the analyte in the bulk solution, C, and electrode area. As stated previously, the behavior of the Cu-Zn system has been characterized and can be related to the concentrations

of Cu and Zn in solution (C_{Cu} and C_{Zn}, respectively) by the following expression [29]:

$$i_{p,Zn(Cu)} = -\frac{i_{p,Zn}}{2\gamma_{Zn}C_{Zn}}\left(\gamma_{Zn}K'C_{Cu} - \gamma_{Zn}C_{Zn} - \sqrt{\gamma_{Zn}^2/K'^2C_{Cu}^2 - 2\gamma_{Zn}^2K'^2C_{Cu}^2C_{Zn} + \gamma_{Zn}^2C_{Zn}^2 - 4K_{sp}}\right) \quad (3)$$

In eq. 3, $i_{p,Zn(Cu)}$ is the peak current of Zn in the presence of Cu, $i_{p,Zn}$ is the peak current for Zn in the absence of Cu, and γ is the accumulation coefficient (γ_{Cu} = 3.23; γ_{Zn} = 3.42) [29]. In addition, K' is the ratio of the accumulation coefficients of Cu to Zn, and K_{sp} is the solubility product for the CuZn intermetallic compound in the mercury electrode ($K_{sp} = 5 \times 10^{-8}$) [29].

Data simulation for the first scenario (simultaneous determination of Cu and Zn) was performed using eq. 1-3. Initially, for a given combination of Cu and Zn solution concentrations, the maximum peak current in the absence of Cu, $i_{p,Zn}$, is calculated from eq. 1 considering the concentration of Zn in the drop as determined by eq. 2, where γ_{Zn} = 3.42. The $i_{p,Zn}$ value is substituted in eq. 3 and the current after accounting for intermetallic compound formation is determined. Subsequently, the equilibrium concentration of Zn in the mercury electrode after intermetallic compound formation is calculated from eq. 1 when E = E°(Zn). The difference between this concentration and the maximum concentration in the drop determined from eq. 2 (i.e. in the absence of intermetallic compound formation) is the concentration of the CuZn intermetallic compound. After subtracting this value from the maximum concentration of Cu in the mercury electrode, also determined from eq. 2, it was possible to calculate the equilibrium concentration of Cu in the electrode after intermetallic compound formation. Finally, using the equilibrium concentrations of Cu, Zn, and CuZn and their formal potentials, a stripping voltammogram was simulated using eq. 1.

For the second scenario (simultaneous determination of Co, Cu and Zn), data simulation was performed considering the simultaneous influences of Co and Cu in the Zn peak. However, unlike CuZn, CoZn was considered as soluble species in the mercury film [31]. Eq. 3 was modified to express the additional influence of Co. Hovsepian and Shain [31] measured several peak currents for the stripping of Zn from mixed Co-Zn amalgams formed from solution concentrations of Zn and Co in the range of 1–10 ppm. Data obtained from this work were used to develop an empirical model expressing the decrease in the Zn peak current for a particular combination of Co and Zn solution concentrations:

$$\Delta i_{p,Zn(Co)} = 0.0468407C_{Zn} + 0.121311C_{Co} \quad (4)$$

In eq. 4, $\Delta i_{p,Zn(Co)}$ is the drop in the peak current of Zn by the presence of Co related to the maximum peak current for a given concentration of Zn (i.e., in the absence of Co). C_{Zn} and C_{Co} are solution concentrations of Zn and Co, respectively. Thus, the peak current of Zn after considering both intermetallic compounds formation, $i_{p,Zn(Cu,Co)}$, is given by:

$$i_{p,Zn(Cu,Co)} = -\frac{i_{p,Zn}}{2\gamma_{Zn}C_{Zn}}\left(\gamma_{Zn}K'C_{Cu} - \gamma_{Zn}C_{Zn} - \sqrt{\gamma_{Zn}^2/K'^2C_{Cu}^2 - 2\gamma_{Zn}^2K'^2C_{Cu}^2C_{Zn} + \gamma_{Zn}^2C_{Zn}^2 - 4K_{sp}}\right) - \Delta i_{p,Zn(Co)} \quad (5)$$

Now, as performed for data simulation in the first scenario, the maximum peak current in the absence of Cu and Co, $i_{p,Zn}$, is calculated from eq. 1 and used in eq. 5 to obtain the current after accounting for intermetallic compound formation. Again, the equilibrium concentration of Zn in the mercury electrode after intermetallic compound formation is calculated from eq. 1. The equilibrium concentrations of Co, Cu, CoZn, and CuZn were estimated by performing a systematic treatment of equilibrium.

Considering the dissociation constants for CoZn and CuZn as $K_1 = 1.3 \times 10^{-2}$ [31] and $K_2 = 1.9 \times 10^{-3}$ [32], respectively:

$$CoZn \leftrightarrow Co + Zn \quad K_1 = \frac{[Co][Zn]}{[CoZn]} \quad (6)$$

$$CuZn \leftrightarrow Cu + Zn \quad K_2 = \frac{[[Cu][Zn]}{[CuZn]} \quad (7)$$

Rearranging eq. 6 to solve for [Zn] and substituting the resulting expression in eq. 7 gives:

$$\frac{[CoZn]}{[Co]K_1} = \frac{[CuZn]}{[Cu]K_2} \quad (8)$$

The difference between the maximum concentration of Zn in the drop determined from eq. 2 and its equilibrium concentration obtained from eq. 1, $[Zn]_i$, can be determined and is in the form

$$[CoZn] \text{ and } [CuZn]. \text{ Thus: } [CoZn] + [CuZn] = [Zn]_i. \quad (9)$$

Conversely, all of the Co or Cu delivered to the drop, C_{Co} and C_{Cu} respectively, can also be determined from eq. 2–considering $\gamma_{Co} = \gamma_{Zn}$ [31] – and are either in the form of free metal or associated to Zn to form the intermetallic, following:

$$[Co] + [CoZn] = C_{Co} \quad (10)$$

$$[Cu] + [CuZn] = C_{Cu} \quad (11)$$

Now, with equations 8 to 11, the four unknowns ([Co], [Cu], [CoZn], and [CuZn]) can be determined. Finally, using the equilibrium concentrations of Cu, Zn, Co, CoZn, and CuZn and their formal potentials, a stripping voltammogram was simulated using eq. 1.

Calibration and validation sets of voltammograms were simulated using a complete factorial design. For the first scenario, ten levels of concentration for each metal were used for the calibration set, and six for the validation set. For the second scenario, six levels were used for the calibration set, and four for the validation set. Solution concentrations of Co, Cu and Zn were used within the range of 10–100 ppm for the calibration sets, and within the range of 20–90 ppm for the validation sets. The formal potentials of Cu, Pb, Co and Zn were set at +0.022, −0.550, −0.270, and −0.994 V vs. Ag/AgCl, respectively, while the formal potential for the intermetallic compounds was changed when necessary to explore the performance of the method [33]. Simulated stripping voltammograms were obtained using a standard spreadsheet program (Microsoft Excel).

3.2. Experimental five-component system

Twenty-six experiments were performed in the simultaneous DPASV analysis of Cu, Co, Pb, Cd, and Zn in water samples with MTFE: 16 obtained from an orthogonal array OA16.5.4.2 [34] plus 10 experiments prepared with random concentration values within the same concentration range (10 to 100 ng mL^{-1}). Each experiment was performed in triplicate (independent replicates), resulting in a data set with 78 voltammograms.

3.3. Chemicals, materials and electrochemical setup

The reagents used were KCl (Merck), KNO_3 (Merck), HNO_3 (Merck), CH_3COOH (Mallinckrodt), CH_3COONa (Merck), $Hg(NO_3)_2.2H_2O$ (Merck), and AgCl (Merck). All reagents were of analytical grade and used without prior purification. Ultrapure water (18.3 MΩcm^{-1}) was used throughout. Working metal ion

solutions were prepared from atomic absorption certified standard solutions (1000 mgL^{-1} of Cu^{2+}, Pb^{2+}, Cd^{2+}, Zn^{2+} and Co^{2+}, SpecSol) after appropriate dilution with ultrapure water. An acetic-acetate buffer stock solution (pH 4.0), prepared by mixing proper amounts of acetic acid and sodium acetate, was used as the supporting electrolyte.

A silver chloride electrode (Ag/AgCl) in saturated KCl solution was used as the reference electrode. A Pt wire was used as the auxiliary electrode. The working electrode was a mercury film modified carbon fiber microelectrode. The carbon fiber electrode consisted of a bundle of fibers with a total thickness of 50 μm embedded in a glass tube and sealed with polyester resin. Modification of the carbon fiber microelectrode was carried out by *ex situ* deposition of the mercury film at −1.2 V for 300 s in a solution containing 10 mmolL^{-1} $Hg(NO_3)_2$ and 1.0 molL^{-1} KNO_3 at pH 1.0, and in the absence of dissolved oxygen. The electrochemical measurements were carried out using an Autolab PGSTAT 30 potentiostat (EcoChemie, The Netherlands) and GPES 4.9 software (EcoChemie, The Netherlands) for control and data acquisition. All measurements were carried out at room temperature (25 ± 1 °C) in a 10 mL voltammetric cell.

For DPASV measurements, a deposition potential of −1.2 V was applied to the working electrode for 60 s. At the end of this accumulation period, the stirrer was switched off and a positive-going scan was performed in the −1.2 V to +0.2 V range, with a pulse amplitude of 25 mV and scan rate of 10 mV s^{-1}.

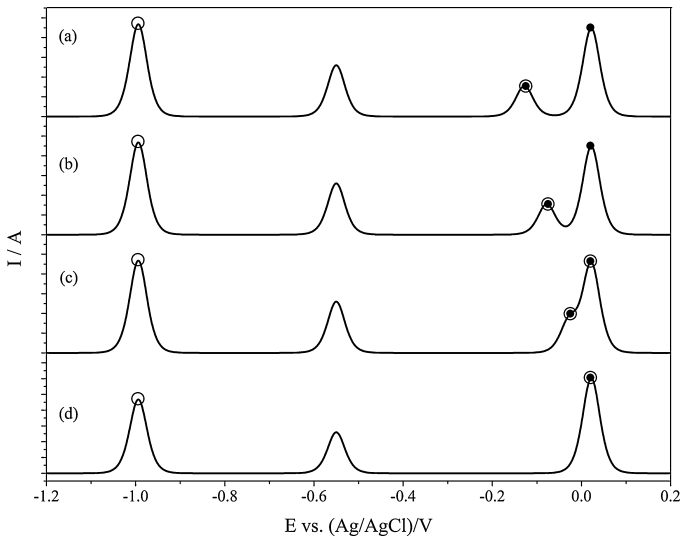

Fig. 1. General simulated voltammetric responses for a system containing Cu, Zn, and Pb, where the response due to the intermetallic compound, E°(CuZn), ranged from: (a) −0.128 V; (b) −0.078 V; (c) −0.058 V; (d) +0.022 V. E_1° = −0.994 V, E_2° = −0.550 V, and E_3° = +0.022 V correspond to the stripping peaks of Zn, Pb, and Cu, respectively. Symbols indicate the SPA selected variables for Cu (filled circles) and Zn (open circles) MLR models.

3.4. Data analysis

The experimental data set was split into appropriate calibration and validation sets with the algorithm SPXY (Sample Partitioning based on joint X-Y short distances) [35]. SPA-MLR regression and SPXY sample selection were carried out with routines written in Matlab® 7.8.0.347 (The MathWorks™, Inc.) [36]. For comparison, PLS regression was applied to the full voltammograms. PLS regression was performed with the software Pirouette® 4.0 (Infometrix Inc., USA). The figures of merit correlation coefficient (r), the prediction error sum of squares (PRESS), the standard deviation of validation errors (SDV), and the number of latent variables and pure variables selected (n) were used to evaluate the quality of the models. All calculations were performed on the mean-centered voltammograms and submitted to 15-point window Savitzky-Golay smoothing [37]. Peak alignment was accomplished by using the *icoshift* algorithm [38], also performed in the Matlab environment.

4. Results and Discussion

4.1. Simulated data

Fig. 1 presents the simulated voltammograms for a three component system (E_1° = −0.994 V, E_2° = −0.550 V, and E_3° = +0.022 V, corresponding to the stripping peaks of Zn, Pb, and Cu) considering the formation of a single intermetallic compound (CuZn) and no interaction of Pb with Cu or Zn. The formal potential of the intermetallic compound, E°(CuZn), was changed: −0.128 V (Fig. 1a), −0.078 V (Fig. 1b), −0.028 V (Fig. 1c), and +0.022 V (Fig. 1d). Circles indicate the SPA selected variables for the Cu (filled circles) and Zn (open circles) MLR models. For each metal, SPA-MLR calibration was compared with univariate calibration performed with the characteristic peak potential (E_1° = −0.994 V or E_3° = +0.022 V). Fig. 2 presents the reference vs. predicted y values for all models of case (a) (E°(CuZn) = −0.128 V), representatively. The correspondent figures of merit are shown in Table 1.

Case (a) (Fig. 1a) represents the situation where the intermetallic peak is completely resolved. For both metals, SPA selected two variables: the specific metal peak potential and the intermetallic peak potential. As one can verify in Fig. 2 and Table 1, SPA models are better than univariate models, presenting higher R-values and significantly smaller *PRESS* and *SDV*. Additionally, univariate models are clearly heteroscedastic (Fig. 2). Obviously, the metal concentration depends on the current obtained from its peak potential, and from the intermetallic peak, since there is a stoichiometric relationship in intermetallic compound formation. Therefore, models built with both peaks led to better results. The SPA-MLR algorithm identifies this stoichiometric relationship when it chooses exactly the variables of metal and intermetallic peaks. Similar considerations can also be applied to case (b) (Fig. 1b), where a small overlap was observed.

Case (c) (Fig. 1c) presents a situation where the resolution between Cu and CuZn peaks decreased significantly. In this case, univariate and multivariate models for Cu presented behaviors similar to that observed in cases (a) and (b) (see Table 1). On the other hand, SPA-MLR chooses three variables for the determination of Zn: the peak potentials of Zn, CuZn and Cu peaks. This occurs because there is a partial overlap between Cu and CuZn peaks and, consequently, there is a strong interference of Cu peak on the intermetallic peak. Then, considering the Zn model, to solve this interference and include correctly the intermetallic contributions, information from the Cu peak is required. Table 2 presents equations for all Zn and Cu models. We can see for the SPA-MLR model for Zn in case (c) that the Cu peak coefficient is negative. This supports the hypothesis that information from the Cu peak was included to correct the peak overlap interference, since it generates an amount to be deducted from the estimated concentration of Zn.

Case (d) (Fig. 1d) presents the extreme situation where the intermetallic peak is completely overlapped by the Cu peak. In this case, SPA-MLR selected only one variable–the Cu peak potential. Therefore, the Cu model is univariate, since all Cu–free Cu or in the intermetallic form–is associated with the same peak. In contrast, the SPA-MLR model for Zn was built with two variables, with a

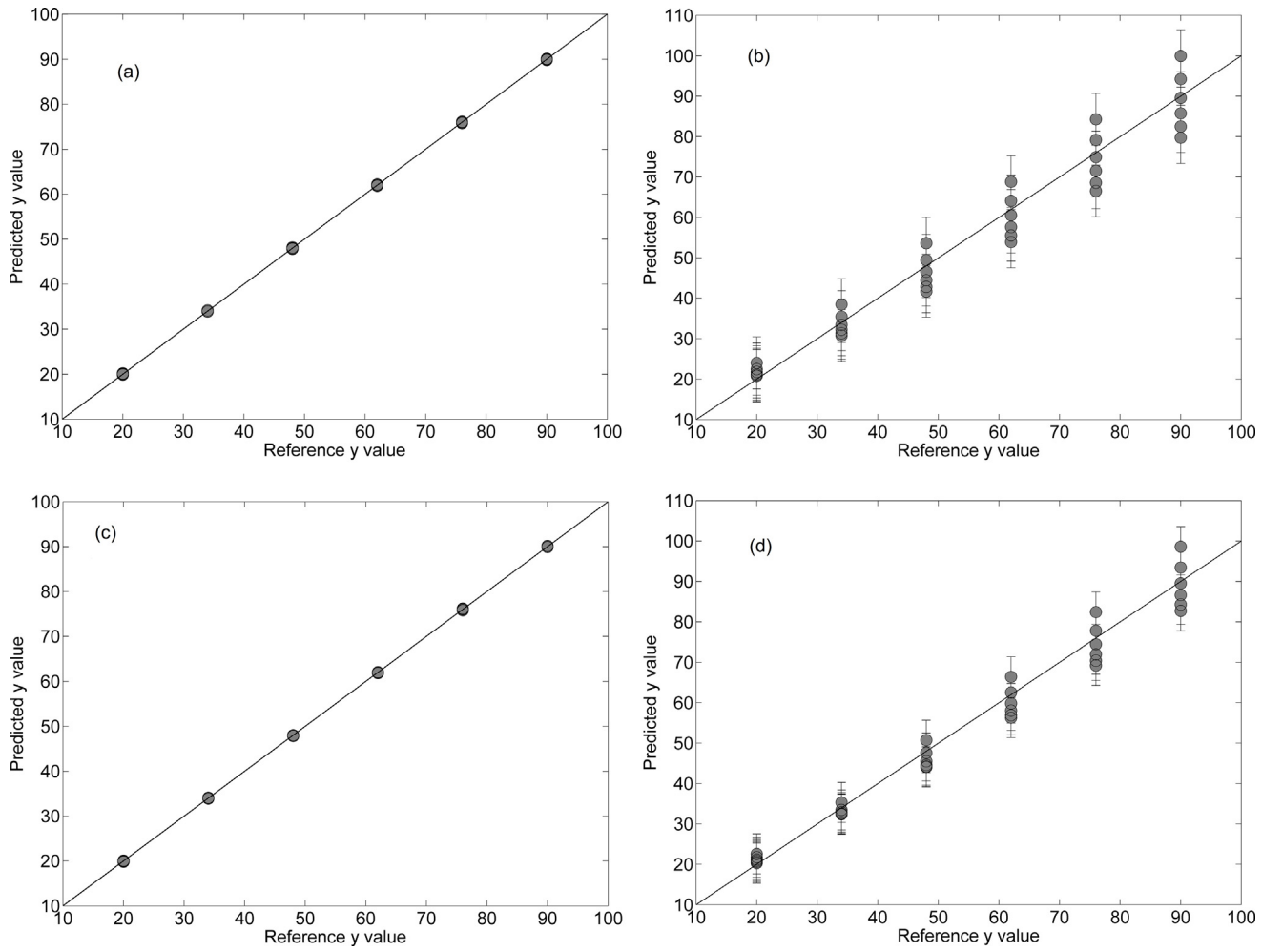

Fig. 2. Reference *vs.* predicted values for determinations where the intermetallic response occurred at E°(CuZn) = −0.128 V. (a) SPA-MLR for Cu; (b) univariate regression for Cu; (c) SPA-MLR for Zn; (d) univariate regression for Zn.

Table 1
Figures of merit obtained in the first scenario. Cases (a), (b), (c), and (d) correspond to simulations where E°(CuZn) used was − 0.128 V, − 0.078 V, − 0.028 V, and + 0.022 V, respectively. Bold values correspond to plots shown in Fig. 2.

Case	Calibration	Cu			Zn		
		PRESS	*SDV*	*R*	*PRESS*	*SDV*	*R*
(a)	**MLR-SPA**	**0.38**	**0.10**	**0.9999**	**0.39**	**0.11**	**0.9999**
	Univariate	**911**	**5.03**	**0.9785**	**504**	**3.74**	**0.9885**
(b)	MLR-SPA	0.26	0.08	0.9999	0.38	0.10	0.9999
	Univariate	909	5.03	0.9786	491	3.69	0.9887
(c)	MLR-SPA	0.26	0.08	0.9999	0.49	0.12	0.9999
	Univariate	748	4.56	0.9822	499	3.64	0.9886
(d)	MLR-SPA	0.28	0.09	0.9999	264	2.57	0.9944
	Univariate				503	3.74	0.9885

Table 2
Analytical curves obtained in the first scenario. Cases (a), (b), (c), and (d) correspond to simulations where E°(CuZn) used was − 0.128 V, − 0.078 V, − 0.028 V, and + 0.022 V, respectively.

Case	Calibration	Cu *Equation*	Zn *Equation*
(a)	MLR-SPA	$[Cu] = 7420 \times I_{(+0.022)} + 7478 \times I_{(-0.128)}$	$[Zn] = 7150 \times I_{(-0.995)} + 7261 \times I_{(-0.128)}$
	Univariate	$[Cu] = 8.992 + 7843 \times I_{(+0.022)}$	$[Zn] = 3.922 + 8329 \times I_{(-0.995)}$
(b)	MLR-SPA	$[Cu] = 7403 \times I_{(+0.022)} + 7463 \times I_{(-0.078)}$	$[Zn] = 7164 \times I_{(-0.995)} + 7158 \times I_{(-0.078)}$
	Univariate	$[Cu] = 8.972 + 7842 \times I_{(+0.022)}$	$[Zn] = 3.922 + 8329 \times I_{(-0.995)}$
(c)	MLR-SPA	$[Cu] = 6745 \times I_{(+0.022)} + 6855 \times I_{(-0.028)}$	$[Zn] = 7156 \times I_{(-0.995)} + 7972 \times I_{(-0.995)} - 1130 \times I_{(+0.022)}$
	Univariate	$[Cu] = 7.924 + 7837 \times I_{(+0.022)}$	$[Zn] = 3.922 + 8329 \times I_{(-0.995)}$
(d)	MLR-SPA	$[Cu] = 0.006981 + 7474 \times I_{(+0.022)}$	$[Zn] = -2.994 + 8438 \times I_{(-0.995)} + 842.7 \times I_{(+0.022)}$
	Univariate		$[Zn] = 3.922 + 8329 \times I_{(-0.995)}$

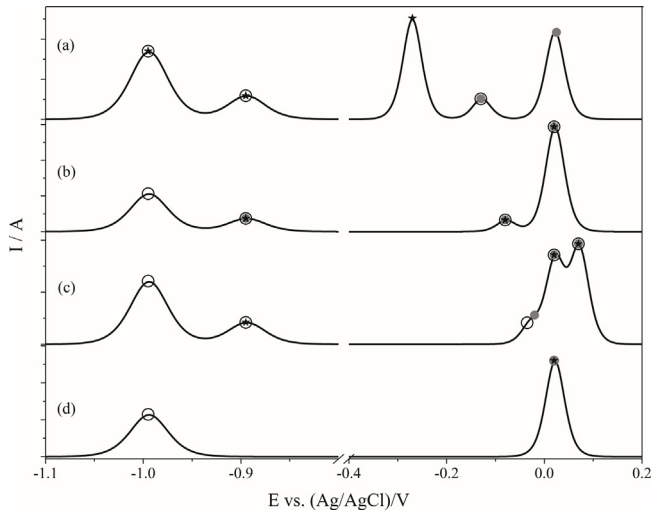

Fig. 3. General simulated voltammetric responses for a system containing Cu, Zn, and Co. $E_1^\circ = -0.994$ V, $E_2^\circ = -0.270$ V, and $E_3^\circ = +0.022$ V correspond to the stripping peaks of Zn, Co, and Cu, respectively. E°(CuZn) ranged from: (a) −0.128 V; (b) −0.078 V; (c) −0.028 V; (d) +0.022 V. E°(CoZn) ranges from: (a), (b), and (c) −0.895 V; (d) −0.995 V. Symbols indicate the SPA selected variables for Cu (filled gray circles), Zn (open circles), and Co (stars) MLR models.

decrease in the predictive power (see Table 1) because the interference of the Cu peak on the CuZn intermetallic peak was completely unsolved. However, the analytical performance of SPA-MLR was still better than univariate regression, which is also heteroscedastic. It is noteworthy that SPA-MLR did not select any variable from the Pb peak for any multivariate model, suggesting the good ability of the algorithm to use only variables with relevant chemical information.

A second scenario was also investigated involving the simultaneous determination of three analytes: Zn, Cu, and Co. The results are presented in Fig. 3 and in Tables 3 and 4. In this situation, besides the formation of an insoluble intermetallic CuZn, the formation of an intermetallic between Co and Zn has been included. The considered intermetallic CoZn is soluble in mercury, and its dissociation constant is not extremely low. The literature reports a significant effect of Zn on the determination of Co in MTFE, where an increasing amount of Zn from an equimolar Zn:Co ratio results in the elimination of the Co peak at −0.27 V, and the occurrence of a new peak at about 0.1 V more positive to the Zn peak potential [39]. Consequently, E^o(CoZn) = −0.894 V was the formal potential used for intermetallic CoZn in the simulated voltammograms, once the formal potential of Zn had been set at −0.994 V.

The determination of Co is also influenced by the concentration of Cu in MTFE. The presence of Cu exceeding an equimolar ratio results in the overlap of the Co and Cu peaks [39]. Thus, in some simulations, E^o(Co) was changed from −0.270 V to +0.022 V and to +0.072 V to include complete and partial overlapping of the Co peak by the Cu peak, respectively. Additionally, the behavior of the CuZn peak was similar to that employed in the first scenario, with changes from −0.128 V to +0.022 V. As one can see, this scenario is clearly much more complex than the previous one, mixing two different intermetallic formation processes with similar equilibrium constants, resulting in a competition between Co and Cu for Zn ions and the loss of a characteristic stripping peak by overlapping.

Fig. 3 shows the voltammograms of the simulated three-component system containing Cu, Co and Zn. Just as performed in the previous scenario, four distinct cases were investigated. In case (a), all analytes exhibit their well-defined characteristic stripping peaks, and the peaks of the intermetallics CuZn and CoZn. In the remaining cases, the Co peak is completely overlapped–cases (b) and (d)–or partially overlapped–case (c)–by the Cu peak. The formal potential of the intermetallic CuZn was also changed from −0.128 V (Fig. 3a) to +0.022 V (Fig. 3d). Case (d) presents a situation where the Co and CuZn peaks are overlapped by the Cu peak, and the CoZn peak is overlapped by the Zn peak. For each metal, SPA-MLR was also compared with univariate regression performed with the characteristic peak potential ($E_1^\circ = -0.994$ V, $E_2^\circ = -0.270$ V, or $E_3^\circ = +0.022$ V). Importantly, the univariate determination of Co in cases (b), (c), and (d) is impossible, since there is no characteristic Co peak to perform this task. The figures of merit PRESS, SDV and *r* are shown in Table 3. In Fig. 3, symbols indicate the SPA selected variables for the Cu (filled gray circles), Zn (open circles), and Co (stars) MLR models.

In case (a) of Fig. 3, when all metal peaks are apparent and intermetallic peaks are also completely resolved, three variables were selected for the determination of Zn, and two variables were selected for the determination of Co and Cu. The Zn peak potential and the peak potentials of the CoZn and CuZn processes were used in the determination of Zn. For Cu, only the CuZn peak potential was required besides the Cu peak potential. The same situation was obtained for the determination of cobalt–the Co peak potential plus the CoZn peak potential. The multivariate models were better than the respective univariate ones, as observed in Table 3.

Case (b) (Fig. 3b) presents the first situation where the Co peak is completely overlapped by the Cu peak. In this case, SPA-MLR used three variables for the determination of Co. In the absence of a Co peak, the Cu peak potential was selected since this is the most representative variable of the peak that has the highest amount of information regarding the concentration of Co. The peak potential of the intermetallic compound CuZn was also included due to a small overlap between Cu and CuZn peaks, in order to compensate for this interference. Table 4 presents all equations for Zn, Co and Cu models obtained in the second scenario. One can see that the coefficient for the CuZn peak is negative in the MLR model for Co in case (b). The CoZn peak potential was the third variable selected. The figures of merit show a decrease in the predictive power of the SPA-MLR model for Co compared to case (a). However, the results are still excellent, especially if one considers the situation imposed by the impossibility of determination of Co via univariate calibration.

Table 3
Figures of merit obtained in the second scenario. Cases (a), (b), (c), and (d) correspond to simulations where E°(CuZn) used was − 0.128 V, − 0.078 V, − 0.028 V, and + 0.022 V, respectively. E°(CoZn) used was −0.895 V for cases (a), (b), and (c); E°(CoZn) = −0.995 V for case (d).

Case	Calibration	Cu			Zn			Co		
		PRESS	*SDV*	*R*	*PRESS*	*SDV*	*R*	*PRESS*	*SDV*	*R*
(a)	MLR-SPA	0.78	0.11	0.9999	1.22	0.14	0.9999	0.55	0.09	0.9999
	Univariate	561	2.96	0.9912	773	3.41	0.9884	224	1.86	0.9966
(b)	MLR-SPA	2916	6.25	0.9608	0.94	0.12	0.9999	2211	5.45	0.9713
	Univariate	19675	17.7	0.6209	762	3.39	0.9886	-	-	-
(c)	MLR-SPA	0.43	0.08	0.9999	1.03	0.13	0.9999	0.52	0.09	0.9999
	Univariate	971	3.90	0.9850	771	3.41	0.9885	-	-	-
(d)	MLR-SPA (Univariate)	17139	16.5	0.6815	595	3.07	0.9907	14989	15.4	0.7295

Table 4
Analytical curves obtained in the second scenario. Cases (a), (b), (c), and (d) correspond to simulations where E°(CuZn) used was − 0.128 V, − 0.078 V, − 0.028 V, and + 0.022 V, respectively. E°(CoZn) ranges from: (a), (b), and (c) −0.895 V; (d) −0.995 V.

Case	Calibration	**Cu** *Equation*	**Zn** *Equation*	**Co** *Equation*
(a)	MLR-SPA	$[Cu] = 7470 \times I_{(+0.022)} + 7446 \times I_{(-0.128)}$	$[Zn] = 8523 \times I_{(-0.995)} + 7199 \times I_{(-0.128)} + 7172 \times I_{(-0.895)}$	$[Co] = 6440 \times I_{(-0.270)} + 6455 \times I_{(-0.895)}$
	Univariate	$[Cu] = 1.988 + 8078 \times I_{(+0.022)}$	$[Zn] = 4.504 + 9859 \times I_{(-0.995)}$	$[Co] = 2.329 + 6781 \times I_{(-0.270)}$
(b)	MLR-SPA	$[Cu] = 3843 \times I_{(+0.022)} - 27471 \times I_{(-0.895)} + 31466 \times I_{(-0.078)}$	$[Zn] = 7142 \times I_{(-0.995)} + 7199 \times I_{(-0.078)} + 7184 \times I_{(-0.895)} - 12.26 \times I_{(+0.022)}$	$[Co] = 3116 \times I_{(+0.022)} - 21021 \times I_{(-0.078)} + 30412 \times I_{(-0.895)}$
	Univariate	$[Cu] = 9.998 + 3140 \times I_{(+0.022)}$	$[Zn] = 4.500 + 8263 \times I_{(-0.995)}$	—
(c)	MLR-SPA	$[Cu] = 6388 \times I_{(+0.022)} + 7519 \times I_{(-0.020)} - 452 \times I_{(+0.072)}$	$[Zn] = 7140 \times I_{(-0.995)} + 7798 \times I_{(-0.035)} + 7129 \times I_{(-0.895)} - 374.6 \times I_{(+0.022)} + 24.32 \times I_{(+0.072)}$	$[Co] = 6520 \times I_{(+0.072)} + 6423 \times I_{(-0.895)} - 590.6 \times I_{(+0.022)}$
	Univariate	$[Cu] = - 1.629 + 7856 \times I_{(+0.022)}$	$[Zn] = 4.500 + 8262 \times I_{(-0.995)}$	$[Co] = - 2.986 + 6573 \times I_{(+0.072)}$
(d)	MLR-SPA	$[Cu] = 3.359 + 3412 \times I_{(+0.022)}$	$[Zn] = 1.948 + 7709 \times I_{(-0.995)}$	$[Co] = - 1.056 + 5545 \times I_{(+0.022)}$
	Univariate			—

The SPA-MLR model for the determination of Cu in case (b) used three variables: the Cu and CuZn peak potentials, as well as the peak potential of the CoZn peak. As expected, the first two variables have positive coefficients. Conversely, the coefficient for CoZn peak potential is negative, indicating that this information is used to deduct the total concentration of Cu provided by the other two variables. Once the Cu peak current suffers a major influence due to overlap with the Co peak, the concentration of Cu would be overestimated if the multivariate model used only the current obtained from the Cu and CuZn peaks. Thus, the only region of the voltammogram that presents additional information to correct this interference is the CoZn peak. However, as this peak contains only partial information regarding the concentration of Co, the correction is not effective and the figures of merit are not as good as in case (a). Nevertheless, it is important to stress out that SPA-MLR model has excellent analytical performance despite the strong interference of Co, especially compared with univariate calibration. Regarding the determination of Zn in case (b), the model was very similar to that observed in case (a). Here, the Zn, CuZn and CoZn peak variables were also selected. Additionally, the Cu peak was selected and shows a negative MLR model coefficient, correcting the interference related to the small overlap between this peak and the CuZn peak.

Case (c) represents the situation where the Co peak is partially overlapped by the Cu peak. For both metals, SPA selected three variables: the peak potential of the metal and their intermetallic compounds; the peak potential of the interfering metal partially overlapped. In both models, peak potentials of metal and their intermetallic compound have positive coefficients, and the coefficient of the interfering metal peak is negative. The variables selected solve this strong interference between Cu and Co and correctly include the intermetallic contributions. Comparing these results with those obtained for Cu and Co in case (b), it can be concluded that multivariate calibration could be performed properly if the metal peaks are apparent, even with poor resolution. SPA selected five variables for the determination of Zn: Zn, CuZn, CoZn, and Cu peak variables–as in case (b)–plus the peak variable of the Co peak, which is poorly resolved from the Cu peak. The only variable that presented a negative signal in this MLR model is related to the Cu peak due to the strong interference from the CuZn peak.

Case (d) (Fig. 3d) presents a situation where the peaks of the intermetallics are completely overlapped by the Zn and Cu peaks, and the Co peak is overlapped by Cu peak. In this case, SPA-MLR selected only one variable for each model: the Zn peak potential for the determination of Zn, and the Cu peak potential for the determination of Cu and Co. Consequently, all models are univariate. It is important to mention that the univariate model for Zn has predictive power better than the univariate models previously obtained, since the Zn peak contains more information in case (d) than in cases (a), (b), and (c). Although the Cu peak also contains information related to the concentration of Zn, it was not selected due to the strong interference of Cu and Co on the CuZn peak. The models for Cu and Co presented SDV values greater than 15.0, due to the complete overlap between the Cu and Co peaks and lack of complementary information related to the intermetallic peaks. Similarly, as observed in the evaluation of the first scenario, the SPA models were always better than the univariate models, and showed no heteroscedastic behavior.

As a general conclusion, the variables selected by SPA include in the MLR models the highest amount of chemical information representing the best linear relationship between current and concentration. Thus, for each analyte, the variables selected normally belong to the characteristic stripping peak of the metal–usually the peak potential, the most representative variable–and to other peaks where additional information concerning the metal concentration can be found. Additionally, SPA-MLR did not select any variable from the regions of the voltammogram that had no relevant chemical information.

Using simulated data, modeled from complex interactions experimentally observed in samples containing Cu, Zn, Pb and Co, we demonstrated that SPA-MLR selected variables allow for chemical interpretation. Thus, we can use this approach to make inferences about the electrochemical processes in a real experimental five-component system.

4.2. Experimental five-component system

Initially, we investigated the individual electroanalytical behavior of the five metallic ions in the range from 10 to 100 ng mL^{-1}. This preliminary study is indispensable since the chemometric tools used in the simultaneous determination are based on linear relationships between independent and dependent variables. The stripping voltammograms were characterized by just one current peak centered at −1.05 V, −0.65 V, −0.47 V, −0.10 V, and 0.00 V for Zn^{2+}, Cd^{2+}, Pb^{2+}, Co^{2+}, and Cu^{2+}, respectively (Fig. 4). Linear relationships between peak current and concentration were observed with correlation coefficients greater than 0.990.

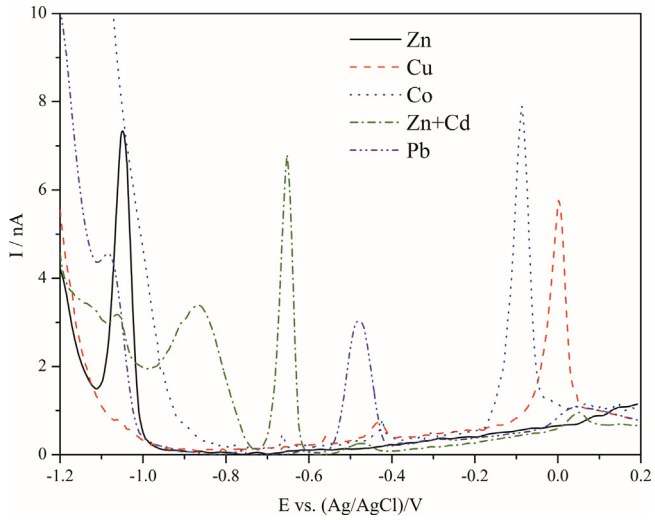

Fig. 4. Anodic stripping voltammograms of aqueous solutions of Cu (80 ng mL^{-1}), Zn (80 ng mL^{-1}), Pb (80 ng mL^{-1}), Co (100 ng mL^{-1}), and a mixture of Cd (80 ng mL^{-1}) and Zn (80 ng mL^{-1}).

Fig. 5 presents the average voltammograms obtained in triplicate from each experiment of the experimental design. The use of average voltammograms was made for clarity, since there were no significant differences in the independent replicates for each experiment. However, it is noteworthy that the calculations were performed using all 78 voltammograms, as described in Section 3.2. In this figure, one can see five peaks. However, except for the Cd and Pb peaks (at about −0.67 V and −0.47 V, respectively) the remaining peaks do not represent the sum of the individual stripping processes of the remaining metals, as considered for the simulated voltammograms. Therefore, there are strong interactions between these metals that significantly influence electrochemical deposition and stripping. The main changes are related to the absence of the Co peak at −0.10 V and the occurrence of a new peak between −0.9 V and −0.8 V. It is possible that the Co peak could be overlapped by the Cu peak in the same way as occurred in the theoretical data. As can be observed in the inset of Fig. 5, the range between -0.06 V and +0.06 V shows the occurrence of up to three peaks. One of those peaks could be associated to the stripping of Co in the same region as the Cu peak [39]. The third peak in this region could

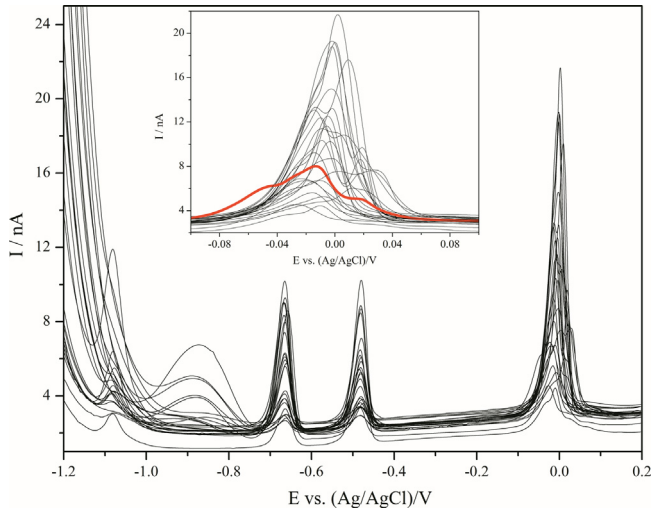

Fig. 5. Differential pulse anodic stripping voltammograms of solutions containing Cu, Cd, Co, Pb, and Zn in concentrations within the range 10-100 μg L^{-1}. Each voltammogram is the average of an experiment carried out in triplicate.

Table 5
Figures of merit obtained for calibration models in the analysis of the real multicomponent system.

Element	Data	Method	*n*	PRESS	*SDV*	*r*
Cd	Raw	SPA-MLR	1	971	6.13	0.9886
		PLS	5	914	5.93	0.9944
	Aligned	**SPA-MLR**	**4**	**430**	**3.72**	**0.9943**
		PLS	19	890	5.85	0.9893
Co	Raw	SPA-MLR	11	22847	27.9	0.6385
		PLS	8	26108	31.7	0.5214
	Aligned	**SPA-MLR**	**9**	**3496**	**10.5**	**0.9113**
		PLS	4	5060	14.0	0.9081
Cu	**Raw**	**SPA-MLR**	**13**	**778**	**5.39**	**0.9843**
		PLS	5	1376	7.27	0.9701
	Aligned	SPA-MLR	14	1732	8.30	0.9729
		PLS	13	2015	8.80	0.9665
Pb	Raw	SPA-MLR	8	896	5.90	0.9860
		PLS	20	300	3.40	0.9959
	Aligned	**SPA-MLR**	**10**	**217**	**2.93**	**0.9966**
		PLS	23	456	4.19	0.9944

be related to a Cu-Zn intermetallic compound [40,41]. This new peak in the −0.9 V to −0.8 V range could be attributed to a Cd-Zn intermetallic compound [42,43], as can be seen in Fig. 4. In this figure, the addition of 80 ng mL^{-1} of Cd to a solution containing the same concentration of Zn causes a decrease in the Zn peak and the occurrence of a new peak at about −0.86 V.

The stripping peaks of Pb and Cd do not appear to be strongly affected by the presence of the other metals studied. However, these peaks show non-systematic shifts that could impair the predictive power of multivariate models. To work properly, multivariate calibration algorithms require that the same underlying process be associated with the same variables in all the samples. Then, it is reasonable to propose that peak shifting could be attributed to the degradation of the reference electrode over time or to intermetallic bonding between target analytes [44]. If this is the case, it could be resolved using an alignment algorithm. Peak alignment could decrease the variation related to peak shifting, adjusting voltammetric data in such a way that the multivariate calibration algorithms have their maximum efficiency. The peak alignment tool adds an offset to the potential axis for each voltammogram such that the positions of the metal peaks correspond with those on a reference voltammogram. In this work, peak alignment was accomplished by using the *icoshift* algorithm [38].

Table 5 presents PRESS, SDV, and *R*-values, as well as the number of real or latent variables (*n*) employed in the SPA-MLR and PLS models, respectively. The values of *r* and SDV are very impressive considering the complexity of the system. The multivariate models for the determination of Zn achieved no satisfactory results (higher PRESS and SDV values, as well as $r < 0.900$), and for that reason they are not presented in Table 5. In fact, the determination of Zn is a very difficult task in this multi-component system. Besides the strong influence of Cd, the Zn peak also decreases with an increase of the concentration of Cu (not shown here), which is indicative of the formation of an intermetallic Cu-Zn compound [32,45,46]. Other important interference occurs due to the presence of Co, since an increased concentration causes the anticipation of the solvent breakdown reaction (Fig. 4), thus masking the Zn signal at potentials more negative than −1.0 V in most experiments (Fig. 5).

Two main conclusions can be obtained from the simultaneous determination of Cd, Co, Cu and Pb by multivariate calibration: i) SPA-MLR and PLS have similar predictive power and ii) in a general way, the use of peak alignment leads to models with better predictions. However, peak alignment did not improve the prediction results for the determination of Cu, due to the large number of electrochemical processes involved in the Cu peak region. In this region, the Cu peak is probably confounded by the Co and Cu-Zn peaks, and

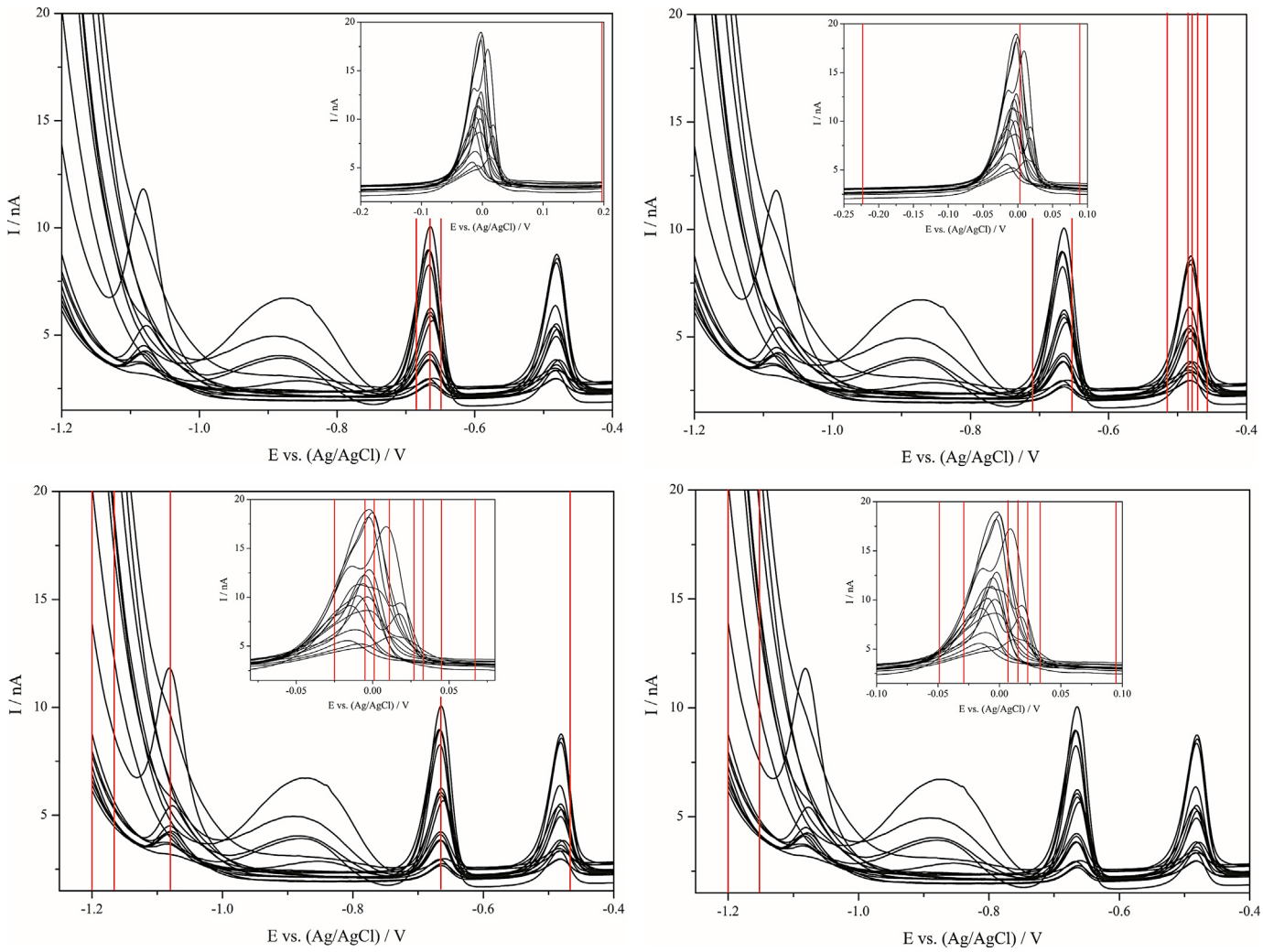

Fig. 6. Variables selected and employed in the MLR-SPA models. a) Cd; b) Pb; c) Cu; d) Co.

the establishment of a reference peak to carry out the alignment led to the loss of important discriminative information. This situation is similar to that observed in case (c) of the second scenario of simulated data, where a satisfactory multivariate calibration was obtained with poorly resolved peaks.

It is important to stress the similar predictive powers of SPA-MLR and PLS, since PLS is a widely used and well-established multivariate regression technique, whose application to the simultaneous determination of metals by voltammetric techniques has been previously demonstrated [6–15]. However, MLR brings an important advantage: the easier interpretation of the current-concentration relationship based on real variables (not latent variables as in PLS) directly selected from particular regions of the voltammogram. Each variable–with the respective MLR model coefficient–provides information to understand the underlying electrochemical processes occurring in the analyzed system. A total of 4, 9, 13, and 10 variables were used in the MLR models for Cd, Co, Cu, and Pb, respectively. These variables are depicted in Fig. 6. Each voltammogram was split into two sections. The main view shows the range from −1.2 V up to −0.4 V. The inset shows an expanded view of the Cu peak region.

Fig. 6a presents the variables required for the determination of Cd: three variables from the stripping peak of Cd and an additional variable related to baseline information. The variables selected from the Cd peak model the direct increase of this peak with the concentration of Cd, and some peak shifts caused by the changing concentration of other metals. In view of the selected variables, we can infer that Cu, Co, Pb, and Zn have a small influence on the deposition and stripping processes of Cd. No variable related to the Zn-Cd peak was used in this model, despite the occurrence of this peak is dependent on the concentration of Cd. The information contained in Zn-Cd peak was not significant to complement and improve the prediction ability of the MLR model. The univariate determination of Cd using the current at the characteristic peak potential ($E_p = -0.650$ V) was performed, but a small decrease of the analytical performance (SDV = 6.65; r = 0.9926) was observed. The better results with MLR were due to the modeling of changes in the Cd peak and baseline offset by the additional variables.

The SPA-MLR model for the determination of Pb employed variables from the Pb peak (five variables) to describe the increase of this peak with the increase of the concentration of Pb and some peak shifting, in the same way as observed for the determination of Cd. The higher number of variables selected could reflect the difficulty found by the alignment algorithm to correct the original peak shifting. The remaining variables model how an increase in the concentration of Pb changes the baseline offset (two variables) and mutual interactions involving Cu [47] (one variable) and Cd [48] (two variables). A univariate calibration model using the peak current at $E_p = -0.470$ V led to a significant decrease in analytical performance (SDV = 10.0; r = 0.9827), showing that important information had been lost by the adoption of a single variable.

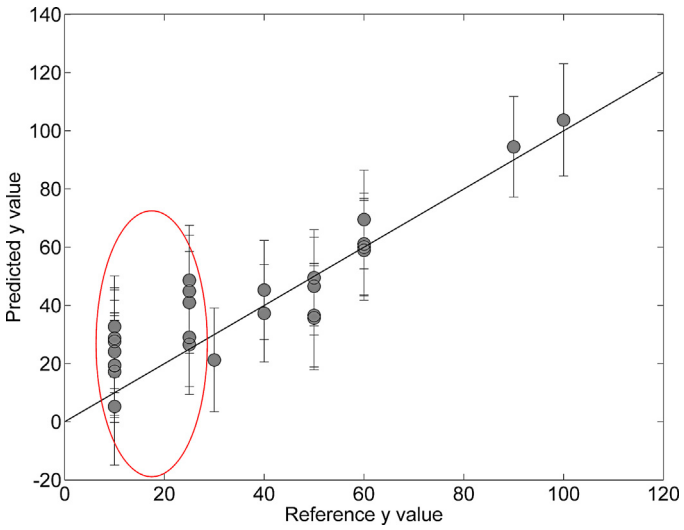

Fig. 7. Reference *vs.* predicted results obtained for the determination of Co in prediction set. The marked area shows the region with the highest prediction errors.

The MLR model for the determination of Cu used a higher number of variables (Fig. 6c), a reflection of its complex deposition and stripping processes in this multi-metal system. Five variables were selected in the range of potentials where the stripping of Cu is expected to occur (inset of Fig. 6c): −0.025 V, −0.005 V, 0.00 V, +0.011 V, and +0.027 V. The variable at the left-hand side of the peak (−0.025 V) have positive MLR model coefficients and could be related to the oxidation of an intermetallic Cu-Zn compound that usually occurs close to the Cu peak [46,47]. The contribution of stripping of the intermetallic Cu-Zn compound to the increase in the Cu peak has been widely reported in the literature [46,47,49].

SPA also selected the peak variable observed in the individual determination of Cu (0.00 V), and the model coefficient was positive. The variable at the left of this peak (E_p = −0.005 V) and the variable immediately to the right of the peak potential (+0.011 V) describe, respectively, peak shifting and the occurrence of a third peak in this region, probably the stripping of Co. The remaining variables selected in this range add information from baseline and peak shifting. Two variables selected at the beginning of the voltammogram have coefficients with negative values. These variables model the anticipation of the solvent breakdown reaction in order to use this information to deduct the contribution of Co in the calculated concentration of Cu. Other variables were selected in the stripping peaks of Zn, Pb, and Cd, with one variable for each metal. All three potentials exhibit MLR model coefficients with a negative sign, indicating the reduction of the respective peaks with an increasing concentration of Cu.

Nine variables were selected for the determination of Co by SPA-MLR. Six were chosen in the stripping range of Cu (Fig. 6d), with four related to the unresolved third peak at more positive potentials to the Cu peak: +0.007 V, +0.015 V, +0.023 V, and +0.033 V. The variable +0.023 V presents a positive model coefficient with a higher absolute value–three times higher the value of the other coefficients in this region–and could be attributed to the most probable position of Co peak. The other variables selected at this range present coefficients that model overlapping between the Co and Cu peaks. The variables −0.049 V and −0.029 V are used to model the interference of the intermetallic Cu-Zn compound on the determination of Co. The remaining variables model information from baseline, with one variable related to the change of the baseline offset at the region after Cu and Co oxidation, and two variables with positive model coefficients associated with the anticipation of the hydrogen evolution reaction.

Although the analytical performance of the MLR model in the determination of Co was satisfactory given the complexity of the studied system, an observation of the reference vs. predicted values plot (Fig. 7) could help to identify the measurements with higher prediction errors. In this figure, it can be observed that the MLR model failed in the prediction of concentrations lower than 30 ng mL^{-1}, which could be attributed to an increase in the limit of quantification of Co when all analytes were simultaneously determined. Thus, samples with concentrations below 30 ng mL^{-1} were removed, and the remaining samples were further subjected to analysis using the MLR model. A significant improvement of the prediction results was observed, with new values of the figures of merit: PRESS = 635, SDV = 6.99 and r = 0.9629.

5. Conclusions

In this work, MLR aided by variable selection (SPA) was applied in the analysis of simulated and experimental anodic stripping voltammetry data, and its ability to provide a useful qualitative interpretation of the analyzed electrochemical systems and quantitative predictions was evaluated. In general, variables selected by SPA usually belong to the characteristic stripping peak of the metal, and to other regions of the voltammogram where additional information concerning the metal concentration can be found–peaks of intermetallic compounds, overlapping interfering processes, and changes in the baseline. Variables with no apparently relevant chemical information were not selected. The interpretation of MLR models allowed the observation of known (Cu over Zn and Co) and unusual interferences (Cd over Zn), and the verification that simultaneous determination at a mercury thin-film microelectrode by anodic stripping is strongly influenced by the relative concentration of the analytes in the solution. This soft modeling semi-empirical approach could also be applied to simultaneous determinations using other thin-film electrodes (such as bismuth films), or also other techniques such as cathodic stripping voltammetry.

SPA-MLR was compared to PLS and similar predictive powers were observed. The analytical performance of MLR and PLS models was considerable given the complexity of the studied system. A peak alignment algorithm was applied in the pretreatment of the voltammograms and was fundamental to provide adequate prediction results in multivariate calibration, except when partial overlap of two or more peaks (with discriminative information) occurred.

Acknowledgments

The authors acknowledge research support from FAPESP, CAPES and CNPq.

References

[1] M.J. Arcos, C. Alonso, M.C. Ortiz, Electrochim. Acta 43 (1998) 479–485.
[2] M.A.A. Lomillo, O.D. Renedo, M.J.A. Martínez, Anal. Chim. Acta 449 (2001) 167–177.
[3] M.R. Majidi, A. Jouyban, K. Asadpour-Zeynali, Electroanal 17 (2005) 915–918.
[4] M. Friedel, C.D. Patz, H. Dietrich, Food Chem. 141 (2013) 4200–4207.
[5] M.F. Pistonesi, M.S. Di Nezio, M.E. Centurión, A.G. Lista, W.D. Fragoso, M.J.C. Pontes, M.C.U. Araújo, B.S.F. Band, Talanta 83 (2010) 320–323.
[6] A. Henrion, R. Henrion, G. Henrion, F. Scholz, Electroanal. 2 (1990) 309–312.
[7] D. Jagner, L. Renman, S.H. Stefansdottir, Anal. Chim. Acta 281 (1993) 315–321.
[8] A. Herrero, M.C. Ortiz, Anal. Chim. Acta 348 (1997) 51–59.
[9] G.D. Pierini, N.E. Llamas, W.D. Fragoso, S.G. Lemos, M.S. Di Nezio, M.E. Centurión, Microchem J. 106 (2013) 347–350.
[10] A. Herrero, M.C. Ortiz, Talanta 46 (1998) 129–138.
[11] A. Herrero, M.C. Ortiz, Electroanal. 10 (1998) 717–721.
[12] A. Herrero, M.C. Ortiz, Anal. Chim. Acta 378 (1999) 245–259.
[13] M.C. Antunes, J.E. Simão, A.C. Duarte, Electroanal. 13 (2001) 1041–1045.
[14] L. Pinto, S.G. Lemos, Electroanal 26 (2014) 299–305.
[15] G.M.S. Alves, J.M.C.S. Magalhães, H.M.V.M. Soares, Electroanal. 23 (2011) 1410–1417.
[16] M. Esteban, C. Ariño, J.M. Díaz-Cruz, TRAC-Trend. Anal. Chem. 25 (2006) 86–92.

[17] J. Wang, Stripping Analysis: Principles, Instrumentation, and Applications, VCH, Michigan, 1985.
[18] A. Economou, P.R. Fielden, Analyst 128 (2003) 205–212.
[19] F.L. Coco, L. Ceccon, L. Ciraolo, V. Novelli, Food Control 14 (2003) 55–59.
[20] E. Shams, A. Babaei, M. Soltaninezhad, Anal. Chim. Acta 501 (2004) 119–124.
[21] C. Colombo, C.M.G. van den Berg, Anal. Chim. Acta 337 (1997) 29–40.
[22] E. Shams, H. Abdollahi, M. Yekehtaz, R. Hajian, Talanta 63 (2004) 359–364.
[23] P. Koscielniak, Chemom. Intell. Lab. Syst. 47 (1999) 275–287.
[24] S.F.C. Soares, A.A. Gomes, M.C.U. Araujo, A.R. Galvão Filho, R.K.H. Galvão, TRAC-Trend Anal. Chem. 42 (2013) 84–98.
[25] R. Gulaboski, V. Mircoeeski, S. Komorsky-Lovric, Milivoj Lovric, Electroanal. 16 (2004) 832–842.
[26] V. Mirceski, S.B. Hocevar, B. Ogorevc, R. Gulaboski, I. Drangov, Anal. Chem 84 (2012) 4429–4436.
[27] H. Gunasingham, R.R. Dalangin, Anal. Chim. Acta 246 (1991) 309–313.
[28] S. Dong, Y. Wang, Talanta 35 (1988) 819.
[29] A.G. Stromberg, V.E. Gorodovykh, Russ. J. Inorg. Chem. 8 (1963) 1234–1236.
[30] A.J. Bard, L.R. Faulkner, Electrochemical Methods: Fundamentals and Applications, John Wiley & Sons, New York, 1980.
[31] B.K. Hovsepian, I. Shain, J. Electroanal. Chem. 14 (1967) 1–16.
[32] M.S. Shuman, G.P. Woodward Jr., Anal. Chem. 48 (1976) 1979–1983.
[33] H. Chan, A. Butler, D.M. Falck, M.S. Freund, Anal. Chem. 69 (1997) 2373–2378.
[34] A.S. Hedayat, N.J.A. Sloane, J. Stufken, Orthogonal arrays: theory and applications, Springer-Verlad, New York, 1999.
[35] R.K.H. Galvão, M.C.U. Araujo, G.E. Jose, M.J.C. Pontes, E.C. Silva, T.C.B. Saldanha, Talanta 67 (2005) 736–740.
[36] Matlab Matlab in The Mathworks.
[37] A. Savitzky, M.J.E. Golay, Anal. Chem. 36 (1964) 1627–1639.
[38] G. Tomasi, F. Savorani, S.B. Engelsen, J. Chromatogr. A 1218 (2011) 7832–7840.
[39] H. Bloom, B.N. Noller, D.E. Richardson, Anal. Chim. Acta 109 (1979) 157–160.
[40] L. Mogensen, L. Kryger, Electroanal. 10 (1998) 1285–1287.
[41] G. Piccardi, R. Udisti, Anal. Chim. Acta 202 (1987) 151–157.
[42] A.J. Baca, Y. Garcia, A.L. Briseno, F. Zhou, J. Electroanal. Chem. 513 (2001) 25–35.
[43] V.D. Jovic, V. Jevtic, Electrochim. Acta 43 (1998) 63–68.
[44] M. Cauchi, C. Bessant, S. Setford, Electroanal 20 (2008) 2571–2577.
[45] J.M. Zen, H.Y. Lin, H.H. Yang, Electroanal 13 (2001) 505–508.
[46] J.A. Wise, D.A. Roston, W.R. Heineman, Anal. Chim. Acta 154 (1983) 95–104.
[47] R.R. Dalangin, H. Gunasingham, Analyst 119 (1994) 2187–2191.
[48] A. Nur Onar, A. Temizer, Analyst 112 (1987) 227–229.
[49] C.M.A. Brett, M.B.Q. Garcia, J.L.F.C. Lima, Anal. Chim. Acta 339 (1997) 167–172.