Talanta 119 (2014) 553–563

Contents lists available at ScienceDirect

Talanta

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

Chemometrics-assisted simultaneous voltammetric determination of ascorbic acid, uric acid, dopamine and nitrite: Application of non-bilinear voltammetric data for exploiting first-order advantage

Mohammad-Bagher Gholivand [a,*], Ali R. Jalalvand [a,b], Hector C. Goicoechea [b], Thomas Skov [c]

[a] Faculty of Chemistry, Razi University, Kermanshah 671496734, Iran
[b] Laboratorio de Desarrollo Analítico y Quimiometría (LADAQ), Cátedra de Química Analítica I, Universidad Nacional del Litoral, Ciudad Universitaria, CC 242, S3000ZAA Santa Fe, Argentina
[c] Quality and Technology group, Department of Food Science, Faculty of Life Sciences, University of Copenhagen, Rolighedsvej 30, DK-1958 Frederiksberg, Denmark

ARTICLE INFO

Article history:
Received 25 July 2013
Received in revised form
7 November 2013
Accepted 8 November 2013
Available online 27 November 2013

Keywords:
Ascorbic acid
Uric acid
Dopamine
Nitrite
Simultaneous determination
Linear and non-linear multivariate calibration models

ABSTRACT

For the first time, several multivariate calibration (MVC) models including partial least squares-1 (PLS-1), continuum power regression (CPR), multiple linear regression-successive projections algorithm (MLR-SPA), robust continuum regression (RCR), partial robust M-regression (PRM), polynomial-PLS (PLY-PLS), spline-PLS (SPL-PLS), radial basis function-PLS (RBF-PLS), least squares-support vector machines (LS-SVM), wavelet transform-artificial neural network (WT-ANN), discrete wavelet transform-ANN (DWT-ANN), and back propagation-ANN (BP-ANN) have been constructed on the basis of non-bilinear first order square wave voltammetric (SWV) data for the simultaneous determination of ascorbic acid (AA), uric acid (UA), dopamine (DP) and nitrite (NT) at a glassy carbon electrode (GCE) to identify which technique offers the best predictions. The compositions of the calibration mixtures were selected according to a simplex lattice design (SLD) and validated with an external set of analytes' mixtures. An asymmetric least squares splines regression (AsLSSR) algorithm was applied for correcting the baselines. A correlation optimized warping (COW) algorithm was used to data alignment and lack of bilinearity was tackled by potential shift correction. The effects of several pre-processing techniques such as genetic algorithm (GA), orthogonal signal correction (OSC), mean centering (MC), robust median centering (RMC), wavelet denoising (WD), and Savitsky–Golay smoothing (SGS) on the predictive ability of the mentioned MVC models were examined. The best preprocessing technique was found for each model. According to the results obtained, the RBF-PLS was recommended to simultaneously assay the concentrations of AA, UA, DP and NT in human serum samples.

1. Introduction

Dopamine (DP), ascorbic acid (AA), uric acid (UA) and nitrite (NT) usually coexist in biological matrixes, and they were considered as crucial molecules for physiological processes in human metabolism. For instance, DP is one of the important natural catecholamine neurotransmitters for message transmission in the central nervous system, which plays a critical role in the function of central nervous, hormonal, and cardiovascular systems. Abnormal levels of DP will lead to Huntington's disease and neurodegenerative disorders, such as Alzheimer's and Parkinson's [1–3]. AA is another important component in human diet, and it plays a vital role in neurochemistry, bioelectrochemistry and clinical diagnostics applications [4]. More importantly, it has been used for prevention and treatment of scurvy, mental illness and cancer [5]. UA is a primary end product of purine metabolism.

Abbreviations: MVC, multivariate calibration; PLS-1, partial least squares-1; CPR, continuum power regression; MLR, multiple linear regression; SPA, successive projections algorithm; RCR, robust continuum regression; PRM, partial robust M-regression; PLY-PLS, polynomial-partial least squares; SPL-PLS, spline-partial least squares; RBF-PLS, radial basis function-partial least squares; LS-SVM, least squares-support vector machines; WT-ANN, wavelet transform-artificial neural network; DWT-ANN, discrete wavelet transform-artificial neural network; BP-ANN, back propagation-artificial neural network; SWV, square wave voltammetry; AA, ascorbic acid; UA, uric acid; DP, dopamine; NT, nitrite; GCE, glassy carbon electrode; SLD, simplex lattice design; AsLSSR, asymmetric least squares splines regression; COW, correlation optimized warping; GA, genetic algorithm; OSC, orthogonal signal correction; MC, mean centering; RMC, robust median centering; WD, wavelet denoising; SGS, Savitsky–Golay smoothing; LOO-CV, leave one out cross-validation; LVs, latent variables; rPCA, robust principal component analysis, MLP, multilayer perceptron; PDC, percentage of data contamination; RMSECV, root mean squared errors of cross-validation; RMSEP, root mean square errors of prediction; REP, relative error of prediction; Q^2, the square correlation coefficient of cross-validation; PRESS, prediction residual error sum of squares

* Corresponding author. Tel.: +98 8314274557; fax: +98 8314274559.
E-mail address: mbgholivand@yahoo.com (M.-B. Gholivand).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.11.028

Abnormal concentration levels of UA will lead to some diseases, such as gout and hyperuricaemia [6]. In recent years, many papers reported that NO could act as a neurotransmitter or a neuromodulator in the central nervous system. Although the physiological results of NO for DP release in the striatum are controversial, it is undisputed that NO can be oxidized to NT in biological circumstance as fast as in a few seconds [7–9]. Therefore, simultaneous determination of AA, DP, UA and NT is important for investigating their physiological functions and diagnosing diseases.

Whereas zeroth-order univariate calibration cannot detect sample components producing an interfering signal, first-order MVC, which operates using a vector of data per sample, may compensate for these potential interferents, provided they are included in the calibration set, a property known as the "first-order advantage". The MVC methods are increasingly used to extract relevant information from different types of absorptive spectral and electrochemical data to predict analyte concentrations or properties of complex samples [10–12]. Several tools have been reported in the literature for processing these data [13], and the most important linear calibration method is PLS [14]. One problem which restricted the application of chemometrics in electroanalytical chemistry is the non-linearity of electrochemical data [15]. Several strategies have been used for the calibration of non-linear data systems. They are: data pretreatment (such as data alignment); the use of linear methods (for slight nonlinearities only); the use of local modeling; the addition of extra variables; the use of non-linear calibration techniques [16–18]. Among these strategies, non-linear calibration techniques are able to build robust calibration models.

In this work, we are going to compare the performance of classical linear (PLS-1, CPR, and MLR), robust linear (PRM, and RCR), and non-linear (PLY-PLS, SPL-PLS, RBF-PLS, LS-SVM, WT-ANN, DWT-ANN, and BP-ANN) MVC models for predicting the concentration of AA, UA, DP and NT in a synthetic sample with a complex matrix to choose the best MVC model for determining the concentration of the mentioned analytes in human serum samples which have a very complex matrix. Literature survey revealed that no attempt has been made till date to the simultaneous voltammetric determination of AA, UA, DP and NT with the aid of Chemometrics.

2. Experimental

2.1. Chemicals and solutions

The AA, UA, and DP were purchased from Sigma-Aldrich Chemie, Steinheim, Germany. Sodium nitrite was obtained from Riedel-de Haën (Sigma-Aldrich Chemie, Steinheim, Germany). Sodium dihydrogen phosphate (NaH_2PO_4), and disodium hydrogen phosphate (Na_2HPO_4) were obtained from Merck. All other materials were used of the highest quality available and purchased from regular sources. The human serum samples used in this study were obtained from a Medical Diagnostic Laboratory in Kermanshah, Iran. Phosphate buffered solution (PBS, 0.1 M, pH2) was prepared using NaH_2PO_4, and Na_2HPO_4 and titrated with H_3PO_4 to pH2. All working and sample solutions were analyzed in the PBS. All solutions were prepared with double-distilled water (ddH_2O). Pure nitrogen was passed through all the experimental solutions.

Stock standard solutions (0.01 M) of the analytes were prepared daily by exact weighing and dissolving their solid powder in a PBS (0.1 M, pH2). Working solutions were prepared immediately before their use by taking appropriate aliquots of each stock standard solution and diluting with PBS to the desired concentrations.

2.2. Apparatus and softwares

Electrochemical experiments were performed using a μ-Autolab TYPE III, Eco Chemie BV, Netherlands, and driven by the NOVA software (Version 1.8). A conventional three-electrode cell was used with a saturated Ag/AgCl as reference electrode, a Pt wire as counter electrode and a GCE as working electrode. The pH values were measured using a JENWAY-3345 pH-meter equipped with a combined glass electrode. The recorded experimental data was smoothed, when necessary, and converted to matrices by means of several homemade mfiles. Baseline correction was performed using AsLSSR [19], and signal alignment was performed using correlation optimized warping (COW) [20] employing MATLAB software (Version 7.14 from MathWorks, Inc.) [21]. PLS-1, PLY-PLS, SPL-PLS, GA, MC, and OSC analyzes were performed using PLS-Toolbox (Version 3.5, Eigenvector Research Inc., USA [22]). All ANN modellings were implemented employing MATLAB. Computations based on CPR, PRM, RCR, rPCA, and RBF-PLS were performed in MATLAB environment using a series of m-files written by Walczak et al. [23,24]. Computations based on SGS, SPA, and MLR were performed in MATLAB environment using a series of m-files written by Paiva et al. [25]. All calculations were run on a DELL XPS laptop (L502X) with Intel Core i7-2630QM 2.0 GHz, 8 GB of RAM and Windows 7-64 as its operating system.

2.3. Model optimization

To truly compare the different MVC models, the efficiency of the best possible model should be found. Because of the dependence of the calibration model efficiency on its parameters, the following parameters were varied (optimized):

- *PLS-1*: number of latent variables (LVs).
- *CPR*: number of LVs, and power.
- *MLR*: number of LVs.
- *PRM*: number of LVs, and percentage of data contamination (PDC).
- *RCR*: number of LVs, PDC, and delta parameter (δ).
- *PLY-PLS*: number of LVs and degree of polynomial (D).
- *SPL-PLS*: number of LVs, number of knots (K), and D.
- *RBF-PLS*: number of LVs, and sigma parameter (σ).
- *LS-SVM*: number of LVs, regularization parameter (γ), and kernel-related parameter (σ^2, here, RBF kernel function was selected).
- *ANNs*: Number of input neurons (IN), number of hidden neurons (HN), number of output neurons (ON), and transfer functions of the hidden and output layers.

2.4. Model efficiency estimation

Whether a model can be applied to analysis of human serum samples or not, model validation is possibly the most important step in the model building sequence. In order to evaluate the performance of the previously mentioned MVC models, each model was validated for the prediction of the validation set, evaluating root mean squared errors of cross-validation (RMSECV), cross-validated correlation coefficient (Q^2), root mean square errors of prediction (RMSEP), and relative error of prediction (REP).

$$\mathrm{RMSECV} = \sqrt{\frac{1}{m}\sum_{1}^{m}\left(y_{pred} - y_{act}\right)^2} \quad (1)$$

$$\mathrm{RMSEP} = \sqrt{\frac{\sum_{1}^{n}\left(y_{pred} - y_{act}\right)^2}{n}} \quad (2)$$

$$Q^2 = 1 - \frac{\sum_{1}^{n} (y_{pred} - y_{act})^2}{\sum_{1}^{n} (y_{act} - y_{mean})^2} \quad (3)$$

$$\mathrm{REP}(\%) = \frac{100}{y_{mean}} \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_{pred} - y_{act})^2} \quad (4)$$

where y_{act} and y_{pred} are actual and predicted concentrations of each component, respectively, and y_{mean} refer to the mean of the actual concentrations. m and n are the number of samples in calibration and validation sets, respectively. Another measure of the model fitting to the training data is R^2, defined as:

$$R^2 = 1 - \frac{\mathrm{SSR}}{\mathrm{SSY}} \quad (5)$$

where SSR is the sum of squares of the residual, and SSY is the sum of squares of the response variable.

2.5. Real sample preparation

Two different human serum samples were selected as real samples for analysis. Human serum samples were centrifuged before the experiment. All samples were diluted with PBS (0.1 M, pH2) and then appropriate amounts of these diluted samples were transferred to the electrochemical cell for the determination of each analyte.

2.6. Electrochemical procedure

All electrochemical experiments were carried out at room temperature. The SWV measurements were carried out at the

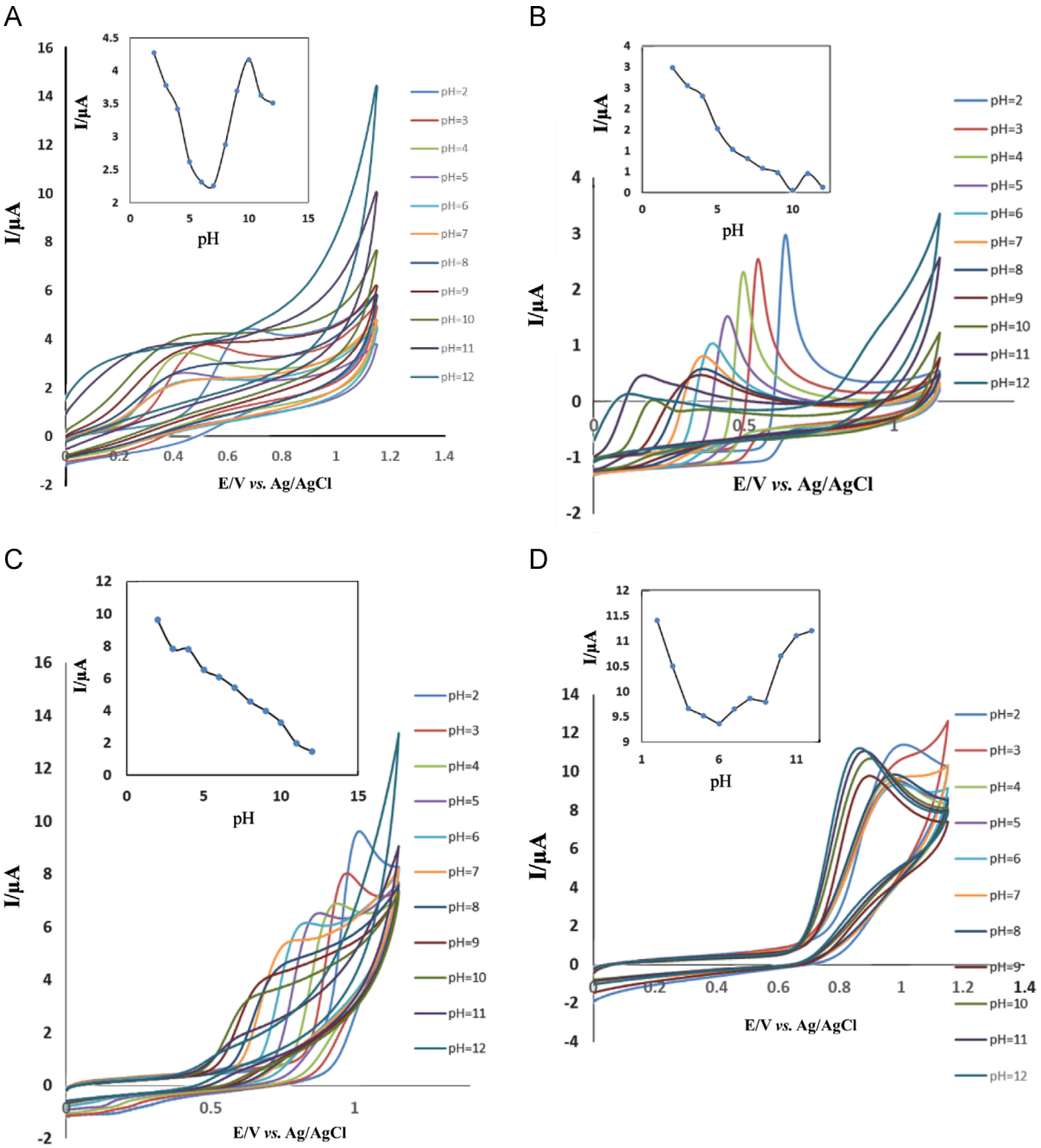

Fig. 1. Cyclic voltammograms of (A) AA (1 mM), (B) UA (1 mM), (C) DP (1 mM), and (D) NT (1 mM) in 0.1 M PBS at different pHs. Insets: variation of I_p vs. pH.

following operating conditions for the four studied analytes: step potential 0.002 V, amplitude 0.02 V, frequency 25 Hz, and scan rate 0.05 V/s. Prior to each measurement, the GCE was polished with 0.50 and 0.05 μm alumina slurries and then rinsed thoroughly with ddH_2O. The electrode was then sonicated in water for 5 min to remove the adsorbed alumina particles.

3. Results and discussion

3.1. Electrochemical studies

3.1.1. pH dependence study

Taking into account that for analytical purposes maximal currents are necessary, the effect of the solution pH on the electrochemical response of the GCE towards the simultaneous determination of AA, UA, DP, and NT was studied using cyclic voltammetry in the pH range from 2 to 12. As observed in Fig. 1, all peak currents (I_P) of the studied analytes have a maximum value at pH2. Therefore, in order to obtain a high sensitivity, a pH value of 2 was selected for further experiments. The oxidation peak potential (E_P) of all studied analytes shifted to less positive values as the pH of the buffer solution was increased (Fig. 1).

3.1.2. Effect of scan rate

The influences of scan rate (υ) on the oxidation peak potential (E_p) and, peak current (I_p) of AA, UA, DP, and NT at the GCE in PBS (0.1 M, pH2) were studied by cyclic voltammetry (not shown). In the range of 10–500 mV s^{-1}, a linear relationship was established between I_P and $\upsilon^{1/2}$, for all of the studied analytes, indicating the diffusion controlled mechanism. Additionally, the E_P of these four molecules shifted positively as increasing υ which confirmed that the electrode reaction was irreversible, as confirmed from the lack of a reduction peak in the cyclic voltammograms.

3.1.3. Why MVC is necessary?

Fig. 2 shows the cyclic voltammograms of AA (curve a), UA (curve b), DP (curve c), NT (curve d) and their mixture (curve e) at pH2. In all conditions evaluated, a strong signal overlapping was observed for the simultaneous analysis of AA, UA, DP, and NT at the GCE (see Fig. 2). Thus, the quantification of any of these analytes will be biased if univariate calibration is used as analytical method, and for tackling this problem it was necessary to use multivariate calibration. Since SWV has a much higher current sensitivity than cyclic voltammetry, it has been used to simultaneous determination of the studied analytes.

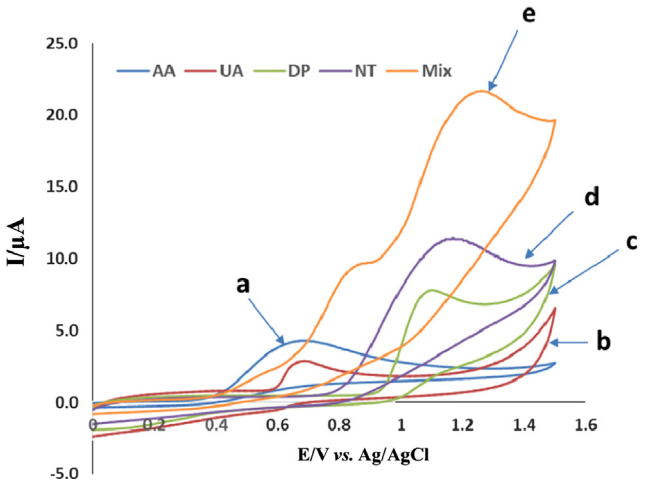

Fig. 2. Cyclic voltammograms of (a) AA, (b) UA, (c) DP, (d) NT, and (e) their mixture in PBS (0.1 M, pH2). Concentration of each analyte is 1 mM.

3.2. Chemometrical studies

When the analytes are analyzed in the presence of interferences, the electrochemical profile revealed additional changes to those observed in the absence of interferences. The main changes observed were minor alterations in the base line and displacement of peak potential, probably due to modification in viscosity of the solution and consequently the diffusion coefficient of the analytes. This effect produces alterations in the chemometrics responses and for this reason, the calibration and validation sets were prepared by including the common interferences in all mixtures.

3.2.1. Calibrations

3.2.1.1. Individual calibrations. Individual calibration curves were constructed with several points (Fig. S1, Supplementary information) as peak current versus analyte concentration in the range 22.5–366 μM, 6.95–363 μM, 18.6–287 μM and 41.7–369 μM for AA, UA, DP, and NT, respectively, and evaluated by linear regression. All analytes showed linear dependences between peak current and concentration at different concentrations intervals.

3.2.1.2. Multivariate calibrations

3.2.1.2.1. Calibration set. The human serum has a complex matrix and may contain a lot of unexpected interferences therefore, if the presence of these interferences was not considered during calibration, a first order MVC model would give biased predictions of the concentration of the analytes of interest. Therefore, a PBS (0.1 M, pH2) spiked with several compounds from common co-existing interferences (exploiting first-order advantage) including NaCl, KCl, $CaCl_2$, $MgSO_4$, $ZnCl_2$, glucose, L-cysteine, L-tyrosine, L-tryptophan, citric acid, lysine and valine having randomized concentrations from 20 to 100 μM was chosen for preparing the calibration set. All the calibration mixtures (the compositions of the calibration mixtures were selected according to a SLD, Table S1) were prepared in the mentioned PBS spiked with an appropriate amount of each analyte of interest considering the linear calibration ranges (previously established from univariate calibrations for each analyte). Finally, the solutions were measured in random order.

3.2.1.2.2. External validation set. An external validation set of ten quaternary mixtures (Table S1) was prepared in the pretreated PBS described in the previous section with random amount of each analyte of interest in the same concentration range used for calibration. The solutions were measured in random order.

3.2.2. Linear MVC models

For chemometric model building, several strategies have been proposed to align shifted signals such as chromatograms, electropherograms or NIR spectra. One of the most popular ones is COW [26,27]. However, this situation has been scarcely described for electrochemical signals. According to the literature, the shift in electrochemical responses can be originated from adsorptive phenomena on the electrode surface, pH variations in the cell or fluctuations in the composition of cell solution [28].

A basic assumption for application of linear MVC models is the data bilinearity, which may be compromised by the above commented potential shifts. Therefore, data alignment was performed before applying linear MVC models including PLS-1, CPR, MLR, RCR, and PRM.

3.2.2.1. PLS-1. The PLS is a well-known first-order MVC methodology. It has been widely applied for different kind of instrumental

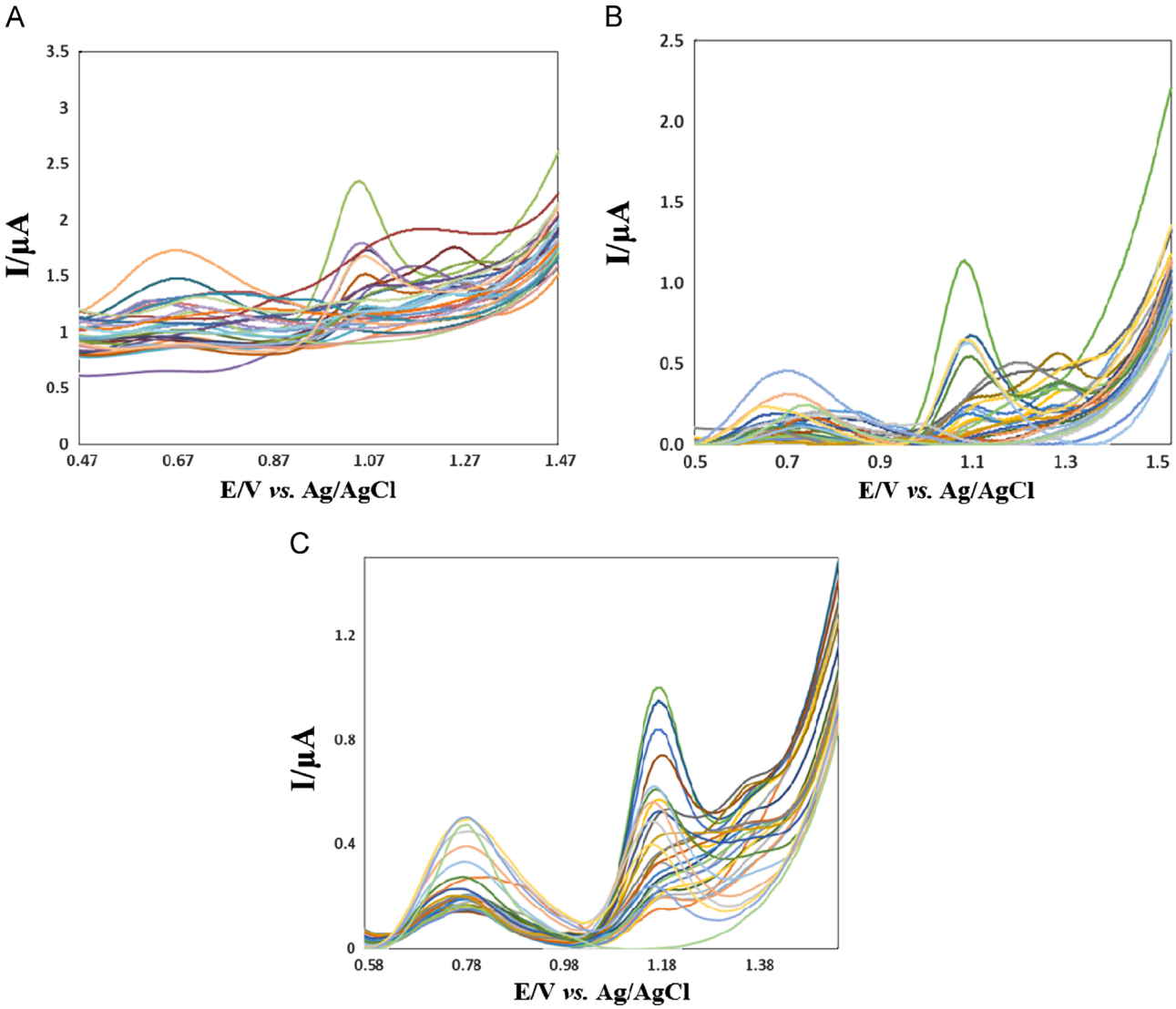

Fig. 3. Square wave voltammograms corresponding to the calibration set. (A) Raw data, (B) Baseline corrected data with AsLSSR, and (C) Potential shift corrected data with COW.

data (i.e. spectroscopic, electrochemical or chromatographic) with satisfactory results [13,29]. A detailed description of PLS-1 has been implemented in Supplementary information.

Before calibration, it is usual to assess the optimum number of LVs in order to avoid overfitting, by applying the well-known leave one out cross-validation (LOO-CV) method described by Haaland and Thomas [29].

Besides the problem arising from the presence of severely overlapping analytes' profiles, in this study two additional complications may occur: (1) the baseline of SWV signals, and (2) sample-to-sample potential shifts in the analyte profiles, which are common in voltammetric studies. The first one was tackled by baseline correction based on an adaptation of the method (AsLSSR) described by Eilers et al. [19]. Subsequently, the SWV signals were aligned towards a target signal using COW. The COW algorithm was introduced by Nielsen et al. [26] as a method to correct for shifts in discrete data signals. It is a piecewise or segmented data preprocessing technique that uses dynamic programming to align a sample signal towards a reference signal by stretching or compression of sample segments using linear interpolation. First, the segment and slack were optimized using a simplex-like optimization routine and then mean voltammogram was selected as target "signal".

The results of baseline- and shift correction are shown in Fig. 3. Fig. 3(A) shows the raw data recorded for the calibration set. Fig. 3 (B) shows the results of baseline corrected data, and as can be seen the baseline is satisfactorily corrected. Fig. 3(C) shows the results of applying COW for data alignment and it confirms the capability of COW for aligning the data.

After baseline correction and data alignment, linear models were applied to predict the concentrations of the studied analytes in the validation set.

Results of the PLS-1 model application to validation set obtained by us with a view to predict analytes' concentrations are given in Table 1. The figures of merit obtained for PLS-1 model are presented in Table 1. As can be seen, the application of the PLS-1 model shows an inadequate predictive ability for the simultaneous quantification of AA, UA, DP, and NT. Therefore, the effect of several preprocessing techniques such as GA, OSC, and MC on the predictive ability of PLS-1 was examined.

3.2.2.2. GA. Constructing the PLS-1 model after selecting the optimal variables (potentials) improves the prediction capacity of the model [30,31]. The GA can be used successfully for variable selection in PLS-1 calibration. The GA was run using a PLS-1

Table 1
Results of PLS-1 and CPR models applied to the validation set.

Model	Analyte	LVs	RMSECV	RMSEP	REP (%)	R^2	Q^2
PLS-1	AA	8	1.6946	1.1544	11.0213	0.7865	0.7781
	UA	7	2.0554	1.0897	10.8765	0.8123	0.8081
	DP	8	1.6674	1.1234	10.7891	0.7989	0.7898
	NT	6	2.4457	0.9984	10.8808	0.8298	0.8283
GA-PLS-1	AA	8	0.8589	0.9564	10.0102	0.8405	0.8381
	UA	6	1.0362	0.9324	10.2134	0.8335	0.8278
	DP	7	0.9686	0.9876	9.8901	0.8123	0.8035
	NT	7	1.0491	0.9764	9.7321	0.8109	0.7995
OSC-PLS-1	AA	6	1.1243	1.0224	10.7561	0.8011	0.7901
	UA	7	1.3454	0.9934	10.2341	0.8099	0.8032
	DP	8	1.0876	1.0456	10.2134	0.8123	0.8001
	NT	7	1.2154	0.9574	9.7668	0.8167	0.8145
GA–OSC-PLS-1	AA	6	0.3124	0.2198	9.6108	0.8699	0.8674
	UA	7	0.2325	0.2645	9.3249	0.8807	0.8732
	DP	7	0.1885	0.2134	9.2131	0.8812	0.8704
	NT	8	0.1643	0.1964	9.0981	0.8805	0.8795
MC-PLS-1	AA	7	1.5465	1.3021	10.9931	0.7789	0.7708
	UA	8	1.8499	1.0946	10.8431	0.7809	0.7798
	DP	8	1.5545	1.0549	10.7901	0.7899	0.7806
	NT	6	2.3443	0.9876	10.2411	0.7867	0.7732
CPR	AA	8	2.4567	2.0187	11.4765	0.7432	0.7325
	UA	9	2.3421	2.1123	11.0381	0.7328	0.7298
	DP	9	1.9806	2.0432	10.9807	0.7409	0.7301
	NT	8	2.2845	2.0987	11.0032	0.7398	0.7313

regression method with maximum number of factors allowed is the optimal number of components determined by cross-validation on the model containing all the variables, and the selected variables were used for running of PLS-1. For obtaining the optimum set of potential for determination of each analyte, the GA procedure was repeated 10 times. Finally a potential was selected if the percent of selection for that variable exceed from a critical value. The thresholds of 70 were obtained for all analytes according to minimum error of prediction for each analyte. The selected potentials were 870, 823, 807, 785, 756, 733, 701, 689, 628, 622, 618, 612, 605, 597, and 564 for AA, 845, 833, 801, 775, 728, 705, 688, 631, 620, 589, and 560 for UA, 1210, 1188, 1178, 1154, 1131, 1112, 1081, 1033, 1001, 988, 967, 941, and 933 for DP and 1261, 1228, 1221, 1209, 1201, 1188, 1151, 1109, 1099, 1077, 1041, 1018, 1008, and 985 for NT. Results of the GA-PLS-1 model application to validation set are given in Table 1. As can be seen, the results obtained by GA-PLS-1 were better than PLS-1, but it does not have adequate predictive ability for the simultaneous quantification of AA, UA, DP, and NT either.

3.2.2.3. OSC. This technique removes, from the X matrix (voltammetric data), factors that are not correlated to the y matrix (concentrations). Loadings of these factors are later subtracted from the X matrix before calibration development. The subtraction of the influential subspace from the original X matrix (X_0) is performed by:

$$X_0^* = X_0 - C \quad (6)$$

where C is a correction matrix. X_0^* is then used to develop prediction models.

The OSC uses NIPALS algorithm [32] to decompose X into scores t (by principal component analysis (PCA)). These scores are orthogonalized to y to obtain t_{new} and a weight factor w is calculated ($w = X_0^- t_{new}$) with X_0^- being a generalized inverse. New scores t are determined with $t = Xw$ and the loading vector p is estimated ($p^T = t^T X_0/(tt_{new})$). The correction matrix C is finally estimated by $C = tp^T$. The operation can be repeated until the desired number of corrections has been applied to X_0. More details about OSC algorithm can be found in Ref. [30]. Results of the OSC-PLS-1 model application to validation set are given in Table 1. Furthermore, by coupling GA to OSC-PLS-1 (GA–OSC-PLS-1) better results were obtained and they are given in Table 1.

The PLS-1 model was also built using MC and results of the MC-PLS-1 model application to validation set are given in Table 1. Comparison of the values of RMSECV, RMSEP, REP, R^2, and Q^2 shows that the accuracy of the GA–OSC-PLS model is better than other models. However, substantially high values of RMSECV, RMSEP, and REP do not characterize the GA–OSC-PLS model positively. Therefore, GA–OSC-PLS model is not the best model for predicting the concentrations of the studied analytes, even though some superiorities can be observed.

3.2.2.4. CPR. For a detailed in-depth theoretical background on CPR, see Ref. [31]. No math will be described here; see Ref. [31] for all necessary equations and formalism. The MC and LOO-CV were used for data preprocessing and determination of the number of LVs, respectively. The results of CPR model application to validation set are given in Table 1. According to the values of RMSECV, RMSEP, and REP, R^2, and Q^2, it can be concluded that CPR model showed an ineffective model for the simultaneous quantification of AA, UA, DP, and NT.

3.2.2.5. MLR. The MLR is the simplest approach for calibration model creation. It is found on an assumption of a linear "signal-property" connection [33,34]. It is possible to say that this approach is the basic method for experimental data processing in analytical chemistry. The SPA is a variable selection technique aimed at reducing collinearity problems in MLR modeling [35]. However, this method has been scarcely described for electrochemical signals. In addition to providing simpler models, SPA often leads to better prediction results compared to full voltammogram calibration method. For this purpose, variable selection by SPA was used in order to obtain simple MLR models based on a small subset of potentials. The results of MLR-SPA models application to validation set are given in Table 2. The effects of WD, and SGS as preprocessing techniques on the predictive ability of MLR-SPA were examined and the results are given in Table 2. It is also possible to employ SGS followed by WD for data preprocessing. The results of SGS–WD–MLR-SPA model application to validation set are given in Table 2. According to the results reported in Table 2, all MLR-SPA models showed low prediction ability and cannot be recommended to simultaneously assay the concentrations of AA, UA, DP and NT.

In conclusion, the methodology proposed based on SWV data processed with classical linear MVC models was not able to quantify simultaneously AA, UA, DP, and NT in the presence of interferences. Therefore, interest in the development of robust linear MVC models for the simultaneous determination of the studied analytes continues.

3.2.3. Developing robust linear MVC models: RCR, and PRM

The probability that the data can be exactly originated from normal distribution is practically close to zero. Particularly, it is reasonable to expect that prevalent data would be contaminated with outliers inconsistent with the majority of observations and unlikely generated by the same model, because one often has less control over the execution of the experiment. Outliers incorporated into a classical MVC model can significantly degrade the

Table 2
Results of MLR-SPA models applied to the validation set.

Model	Analyte	LVs	RMSECV	RMSEP	REP (%)	R^2	Q^2
MLR-SPA	AA	14	3.0114	3.6874	13.1121	0.6903	0.6897
	UA	12	2.9436	3.4102	12.9801	0.6699	0.6654
	DP	14	3.4175	3.8895	13.0901	0.6705	0.6601
	NT	13	3.1121	3.4411	12.789	0.6966	0.6901
SGS[a]–MLR-SPA	AA	12	2.9901	3.2123	12.2113	0.7112	0.7001
	UA	11	2.9031	3.0815	12.4307	0.6898	0.6851
	DP	13	3.3389	3.2155	11.8801	0.6804	0.6754
	NT	11	3.0437	2.7892	11.679	0.6912	0.6898
WD–MLR-SPA	AA	14	3.1168	3.2178	12.349	0.6816	0.68
	UA	11	2.9154	3.034	12.4506	0.6717	0.6698
	DP	14	3.3878	3.3321	12.0001	0.6711	0.6699
	NT	12	3.0231	3.3891	12.098	0.7118	0.7021
SGS[a]–WD–MLR-SPA	AA	11	2.4501	2.1128	11.9135	0.7611	0.7554
	UA	9	2.2134	2.031	11.6534	0.7608	0.7568
	DP	10	2.6807	2.4357	11.2101	0.7499	0.7341
	NT	8	2.1889	2.2431	11.3508	0.7416	0.7399

[a] No derivative.

performance of the model, since the classical multivariate linear regression is non-robust because of its extreme sensitivity to outliers in the data. Therefore, it was necessary to develop robust linear MVC models.

Here, robust principal component analysis (rPCA) [36] was used for detecting the outliers. Briefly, using robust distances and robust orthogonal distances, a distance–distance plot (an outlier map which was constructed to identify outliers) can be made, facilitating identification of outlying samples (see Fig. S2). For the ith sample, the robust distance, RD_i, is defined as follows:

$$RD_i = \sqrt{\sum_{j=1}^{n} \left(\frac{t_{ij}^R}{S_j^R}\right)^2} \quad (7)$$

where t_{ij}^R are the elements of the robust score matrix and S_j^R is the squared root of the jth robust eigenvalue. The orthogonal distances are obtained as:

$$OD_i = \|x_i - P_f t_{i,f}^T\| \quad (8)$$

where $t_{i,f}$ is the ith score vector with f elements and P_f is a matrix (n,f) of f robust loadings. To detect outlying observation one can use z-transformed distances, i.e., to center every vector with distances around median and to divide all elements by corresponding Qn-scale [37] of the distances. For z-transformed distances, a general cutoff value is 3, it means that all objects with a z-transformed distance above 3 can be considered as outliers (here, six outliers were detected).

It is well known that least squares techniques suffer from outlying observations, making the models weak when outliers are present in the data. Therefore, RCR, and PRM as robust linear MVC models which search for an outlier-free subset of the data (for a detailed in-depth theoretical background on RCR, and PRM, see Refs. [38,39]) have been chosen for predicting concentrations of the studied analytes in the validation set. Robust median centering (RMC) and LOO-CV were used for data preprocessing and determining the number of LVs, respectively, and maximum percentage of data contamination was fixed at 15. Results of application of RCR, and PRM models for predicting concentrations of the studied analytes in validation set are shown in Table 3.

As can be seen, more appropriate results were obtained by applying robust linear MVC models in comparison to linear ones,

Table 3
Results of RCR and PRM models applied to the validation set.

Model	Analyte	LVs	δ	RMSECV	RMSEP	REP (%)	R^2	Q^2
RCR	AA	6	0.75	0.1754	0.1431	6.9808	0.9388	0.9341
	UA	6	1.00	0.1121	0.1644	6.4501	0.9445	0.9456
	DP	7	0.50	0.1310	0.1531	6.5421	0.9501	0.9499
	NT	7	0.50	0.1011	0.1248	6.1435	0.9578	0.9561
PRM	AA	6	–	0.2431	0.2110	7.4109	0.9209	0.9108
	UA	7	–	0.1871	0.2402	7.2103	0.9244	0.9294
	DP	7	–	0.1501	0.1821	7.1453	0.9409	0.9391
	NT	8	–	0.1341	0.1461	7.4506	0.9408	0.9328

but according to the values of RMSECV, RMSEP, REP, R^2, and Q^2, it can be concluded that RCR, and PRM models showed an ineffective model for the simultaneous quantification of AA, UA, DP, and NT. Therefore, interest in the development of non-linear MVC models for the simultaneous determination of the studied analytes continues.

3.2.4. Non-linear MVC models

A high degree of non-linearity in current-concentration dependence leads to the need of non-linear treatment of system. The most important non-linear calibration models are the non-linear variants of PLS (e.g., PLY-PLS, SPL-PLS, and RBF-PLS), LS-SVM and ANNs. Several comparative studies on these techniques have been conducted using various data sets. In some studies, the ANNs performed better than PLS when the data were non-linear, in some studies, ANN and non-linear PLS gave equally good results [40] and in some studies LS-SVM performed better than ANNs [41]. It is possible that the different conclusions obtained from the various studies resulted from differences in the nature of the non-linearities [42,43].

For all non-linear models only baseline correction was performed on the raw data and then baseline corrected data was used for next computations.

3.2.4.1. PLY-PLS, SPL-PLS, and RBF-PLS. The PLS method itself is a linear method of data analysis. However, there exists a number of

Table 4
Results of non-linear variants of PLS applied to the validation set.

Model	Analyte	σ	LVs		D	K	RMSECV	RMSEP	REP (%)	R^2	Q^2
PLY-PLS	AA	–	8	–	5	–	0.2867	0.214	8.7509	0.8896	0.8899
	UA	–	7	–	3	–	0.2019	0.2546	8.1891	0.8834	0.8812
	DP	–	7	–	5	–	0.1605	0.2001	8.001	0.8899	0.8991
	NT	–	7	–	4	–	0.1608	0.1795	8.2809	0.8912	0.8816
SPL-PLS	AA	–	8	–	3	5	0.2761	0.2115	8.8102	0.8799	0.8789
	UA	–	7	–	3	4	0.2101	0.2598	8.2103	0.8871	0.8861
	DP	–	8	–	3	5	0.161	0.2043	7.9895	0.8865	0.8711
	NT	–	8	–	3	5	0.1628	0.1809	8.3656	0.889	0.8754
RBF-PLS	AA	0.94	6	0.94	–	–	0.0109	0.0368	1.2451	0.9815	0.9786
	UA	0.97	5	0.97	–	–	0.0259	0.0489	1.0897	0.9763	0.9654
	DP	0.91	6	0.91	–	–	0.0101	0.0411	1.1132	0.9899	0.9808
	NT	0.91	6	0.91	–	–	0.0241	0.0551	1.0101	0.9831	0.971

K is the number of knots; D is the degree of polynomial.

its non-linear modifications: PLY-PLS, SPL-PLS, and RBF-PLS. The only difference between PLY-PLS, and SPL-PLS and the linear PLS is in one step, in which linear function is changed by polynomial one for PLY-PLS or spline function, a piecewise polynomial function, for SPL-PLS. The polynomial and spline function can have any order.

The RBF-PLS [23,24] model is a non-linear version of PLS. In this approach, a non-linear inner relationship is adopted instead of the linear inner relation in PLS. The non-linear inner relationship may be achieved using the RBF. In case of the RBF-PLS, the RBFs are used to carry out the non-linear transformation of X to form an activation matrix, X_A. The elements of which are defined as:

$$a_{ij} = r_j(x_i) \quad i,j = 1,2,...,m \tag{9}$$

where a_{ij} is the element of X_A at the ith row and the jth column, r_j is the jth RBF, and x_i is a vector consisting the values of independent variables taken from the ith observation. The Gaussian function is the most commonly used RBF, which takes the form:

$$a_{ij} = \exp\left(\|c_j - x_i\|^2/\sigma_j^2\right) \quad i,j = 1,2,...,m \tag{10}$$

in which $\|\|$ denotes the Euclidean distance when the argument is a difference of two vectors, c_j and σ_j are two parameters of jth Gaussian function, the center and the width. The parameter c_j of jth Gaussian function is computed by:

$$c_j = x_j \quad j = 1,2,...,m \tag{11}$$

and the elements of the parameter σ_j of jth Gaussian function is obtained as:

$$\sigma_{j1} = \sigma_{j2} = ... = \sigma_{jm} = \frac{e}{m}\sum_{i=1}^{m}\|x_i - x_j\| \tag{12}$$

where e is a constant assigned a value $e > 0$. Thus, the diagonal elements of activation matrix X_A have the value 1. Thus, the number of RBFs is m, and the RBFs themselves are vector functions, the dimensions of which are all n.

Now in PLS-1, the variance of the prediction is minimized, while maximizing the covariance of X_A and y and the PLS-1 model is set up as:

$$y = TB + E = (X_A X_A^T)B + E \tag{13}$$

where T is the low dimensional score matrix of X_A with the dimension of $m \times n_T$, B is the regression coefficient matrix with the dimension of $n_T \times l$, X_A^T is the transformation matrix of X_A with the dimension of $n \times n_T$, and E is the residuals matrix with the dimension of $m \times l$. The optimal value of n_T used in the score matrix, T can be determined by the LOO-CV method. Since, T is the linear combination of the Gaussian function (row vector of X_A) that will maximize the variance between X_A and y. Subsequently, the RBF-PLS model is obtained and used for prediction purposes. Now, for a new data set, X_{new} which is not included in RBF-PLS model, the dependent variable, y_{new} can be computed as:

$$y_{new} = X_{Anew} X_A^T B \tag{14}$$

where X_{Anew} is the activation matrix of the new dataset (X_{new}) and is preprocessed in an identical manner as X of the data set used for modeling and the new activation matrix (X_{Anew}) is calculated keeping the values of centers and widths of the Gaussian functions.

Table 4 represents the results of application of SPL-PLS, PLY-PLS, and RBF-PLS models to the validation set. As is clear from Table 4, RBF-PLS exceeds both methods of SPL-PLS and PLY-PLS with respect to its effectiveness.

Fig. S3(A–D) shows the comparative graphs between the expected and predicted concentrations by the RBF-PLS model for the four studied analytes. As can be seen, good correlations were found between expected and predicted concentrations which confirm the good performance of RBF-PLS model in predicting the concentrations of the studied analytes in the presence of interferences.

3.2.4.2. LS-SVM. The details of LS-SVM algorithm could be found in the literature [44]. The basic concept of LS-SVM regression is mapping the non-linearly the original data X into a higher dimensional feature space. The transformation into higher dimensional space is implemented by a kernel function and RBF kernel function was selected in this research. In order to reduce the dimensionality and compress the voltammetric data, the optimal principal components of LS-SVM regression was also determined according to the lowest RMSECV values.

In order to obtain a good performance, γ and σ^2 of the kernel function in LS-SVM regression model have to be optimized. Parameter γ determines the trade-off between minimizing the training error and minimizing model complexity. Parameter σ^2 implicitly defines the non-linear mapping from input space to some high dimensional feature space. Before the parameters were optimized, an initial value was set, the range of parameters optimization was based on the initial value setting. Cross-validation in the calibration set was used to direct the optimization process. For each combination of γ and σ^2 parameters, the RMSECV was calculated and the optimum parameters were selected when produced smaller RMSECV. The optimizing processes are shown in Fig. S4(A–D) for AA, UA, DP, and NT,

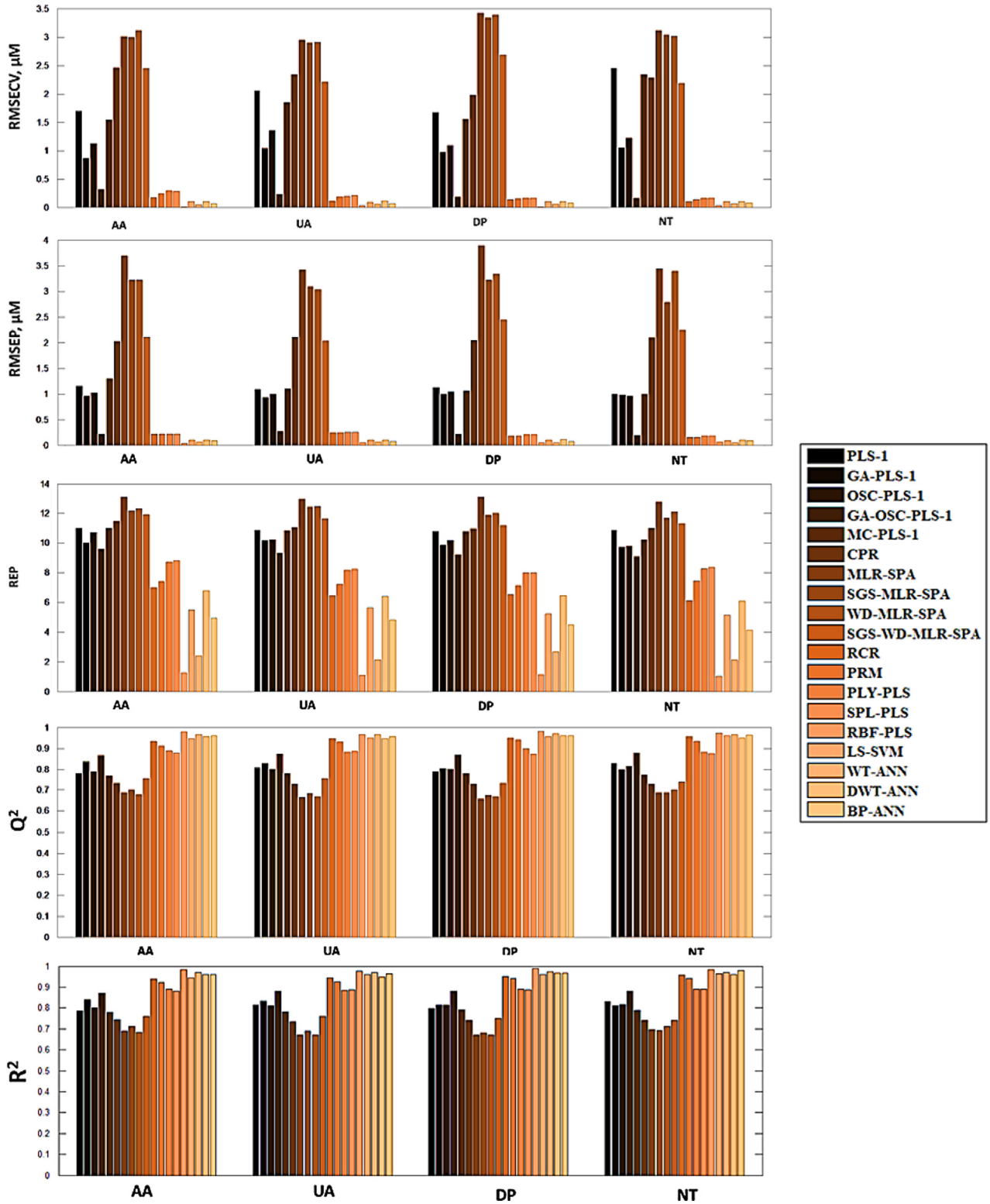

Fig. 4. The comparison of the performance of linear and non-linear calibration models for the prediction of concentrations of AA, UA, DP, and NT in the external validation set.

respectively. The grids "." in the first step are 10×10, and the searching steps in the first step are large. The optimal search area is determined by error contour line. The grids " $\times$ " in the second step are 10×10, and the searching steps in the second step are smaller. Finally, the optimal pairs of (γ,σ^2) for all the studied analytes were found at the values reported in Table S2.

Table S2 represents the results of application of LS-SVM model to the validation set. One can see that LS-SVM model showed results better than those of SPL-PLS, and PLY-PLS models, but not better than the results of RBF-PLS model. Therefore, LS-SVM is not recommended to simultaneously assay the concentrations of the studied analytes.

3.2.4.3. WT-ANN, and DWT-ANN. A detailed description of the theory behind WT-ANN, and DWT-ANN has been implemented in Supplementary information associated with this article.

In all calculations the best performing ANNs were feed forward back propagation. The obtained results are summarized in Table S3.

For DWT-ANN all the computations were the same as WT-ANN but input data matrices were pre-processed by DWT and the obtained results are summarized in Table S3. As can be seen, WT-ANN showed better predictive ability than DWT-ANN, but worse than those of RBF-PLS. Therefore, WT-ANN is not recommended to simultaneously assay the concentrations of the studied analytes.

3.2.4.4. BP-ANN. A feed-forward neural network with back-propagation training algorithm (BP-ANN) was employed for the treatment of data and its results were compared with the results of WT-ANN and DWT-ANN as well. The network was trained employing maximal number of epochs of 100,000. The number of IN was computed by LOO-CV, the output layer consists of a single neuron of value corresponding to concentration of one analyte at a time. The HN value, which leads to the lowest cross-validation error, was chosen as the optimal one. The obtained results are summarized in Table S3. According to the results summarized in Table S3, BP-ANN is not recommended to simultaneously assay the concentrations of the studied analytes.

3.3. Comparison of models

The results of application of classical linear (PLS-1 and CPR, MLR), robust linear (PRM and RCR), and non-linear (PLY-PLS, SPL-PLS, RBF-PLS, LS-SVM, WT-ANN, DWT-ANN, and BP-ANN) MVC models on voltammetric data are shown in Fig. 4. One can conclude the following:

(1) Nineteen different MVC models were used to predict the analytes' concentrations in an external validation set containing ten mixtures. The SWV data were used in all cases.
(2) The RBF-PLS model is the most effective for creation of an express method for multi-component analysis, and the worst model is MLR-SPA. Computation accuracy can be arranged in the following order:
RBF-PLS > WT-ANN > BP-ANN > LS-SVM > DWT-ANN ≈ RCR > PRM > PLY-PLS ≈ SPL-PLS > GA–OSC-PLS-1 > GA-PLS-1 > OSC-PLS-1 > MC-PLS-1 ≈ PLS-1 > CPR > SGS-WD–MLR-SPA > SGS-MLR-SPA > WD–MLR-SPA ≈ MLR-SPA
(3) Surprisingly, the accuracy of the RCR, a robust linear model, is comparable with the accuracy of the DWT-ANN, and also is better than PLY-PLS, and SPL-PLS.
(4) The RBF-PLS model is recommended for practical implementation. The regression model based on RBF-PLS is sufficiently accurate and reliable.
(5) It should be noted that calibration models were also characterized by simplicity for investigation (comprehensibility of main algorithms, availability of software, etc.) and by the volume of required calculations (capacity of computers for realization, time of a model creation, etc.). With respect to these parameters, the above-stated models can be arranged in the following order: Computation time (all computation times were computed in MATLAB environment):

WT-ANN ≈ DWT-ANN > SPL-PLS > BP-ANN > GA–OSC-PLS-1 ≈ PLY-PLS > LS-SVM ≈ RBF-PLS > GA-PLS-1 > CPR ≈ RCR ≈ PRM ≈ OSC-PLS-1 ≈ MC-PLS-1 > SGS-WD–MLR-SPA > WD–MLR-SPA ≈ SGS-MLR-SPA > PLS-1 > MLR-SPA.

ANN training is more than 100 times more time consuming than MLR-SPA model building approach.

Ease of use:

MLR-SPA > WD–MLR-SPA ≈ SGS-MLR-SPA > SGS-WD–MLR-SPA > RBF-PLS > PLS-1 ≈ CPR ≈ RCR ≈ PRM > OSC-PLS-1 ≈ MC-PLS-1 > GA-PLS-1 > GA–OSC-PLS-1 > LS-SVM > BP-ANN > SPL-PLS ≈ PLY-PLS > WT-ANN ≈ DWT-ANN.

Therefore, according to the obtained results, the RBF-PLS model was chosen as the best model for the simultaneous determination of the studied analytes in human serum samples.

3.4. Simultaneous determination of AA, UA, DP, and NT in human serum samples

Two different human serum samples were selected as real samples for the determination of AA, UA, DP, and NT by the use of SWV data processed by RBF-PLS. All samples were diluted with PBS (0.1 M, pH2) and then appropriate amounts of these diluted samples were transferred to the electrochemical cell for the determination of each analyte. The results are presented in Table 5. The proposed method showed excellent recoveries suggesting that this method could be used for the determination of AA, UA, DP, and NT in real samples.

4. Conclusion

Having analyzed results of application of several methods for calibration model creation using voltammetric data, one can make the following conclusions:

(1) The baseline of SWV signals was corrected by AsLSSR as an efficient chemometric algorithm.

Table 5
Determination of AA, UA, DP, and NT in human serum samples by RBF-PLS.

Sample	Analyte	Detected[a] (μM)	Added (μM)	Found[a] (μM)	Recovery (%)	Total value[b] (mM)
Serum 1	AA	–	45	45.2 ± 0.2	100.4	–
	UA	10.2 ± 1.1	25	36.1 ± 0.5	102.5	0.510 ± 0.110
	DP	–	50	49.3 ± 1.1	98.6	–
	NT	–	65	67.7 ± 1.5	104.1	–
Serum 2	AA	–	30	29.7 ± 0.8	99	–
	UA	11.3 ± 1.3	60	73.4 ± 1.4	102.9	0.565 ± 0.108
	DP	–	25	25.2 ± 0.7	100.8	–
	NT	–	50	48 ± 1.2	96	–

[a] Mean value ± standard deviation.
[b] Total value was obtained by multiplying the detected value by 50 (dilution factor).

(2) The potential shifts observed in the voltammograms affecting the strict data bilinearity were tackled by the use of COW as an efficient chemometric algorithm.
(3) By regarding the presence of the common interferences in both calibration and validation sets we could apply the first order data for exploiting first-order advantage for the simultaneous determination of the studied analytes in human serum samples which have a complex matrix.
(4) The results obtained by linear methods were fairly good after data alignment and using suitable preprocessing techniques, but non-linear methods proved their superiority over linear ones.
(5) The RBF-PLS turned out to be the most suitable method for making a calibration model.
(6) The RBF-PLS model was successfully applied to simultaneous determination of the studied analytes in human serum samples.

We hope that results obtained by us will help both further chemometrical and electrochemical investigations of multi-component systems.

Acknowledgments

Gholivand M.B., and Jalalvand A.R. wish to express their sincere appreciation to Razi University Research Council for providing fund and time to this project. Goicoechea H.C. thanks UNL, CONICET and ANPCyT for financial support. The authors would like to thank the reviewers for their insightful comments which led to an improvement of the work.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2013.11.028.

References

[1] J. Zen, P. Chen, Anal. Chem. 69 (1997) 5087–5093.
[2] R.M. Wightman, L.J. May, A.C. Michael, Anal. Chem. 60 (1998) 769–779.
[3] C. Martin, Chem. Br. 34 (1998) 40–42.
[4] I. Koshiishi, T. Imanari, Anal. Chem. 69 (1997) 216–220.
[5] M. Noroozifar, M.K. Motlagh, Talanta 61 (2003) 173–179.
[6] V.V.S.E. Dutt, H.A. Mottola, Anal. Chem. 46 (1974) 1777–1781.
[7] J. Garthwaite, Trends Neurosci. 14 (1991) 60–67.
[8] G. Lonart, K.L. Cassels, K.M. Johnson, J. Neurosci. Res. 35 (1993) 192–198.
[9] Y. Zhang, R. Yuan, Biosens. Bioelectron. 26 (2011) 3977–3980.
[10] L. Xu, J.H. Jiang, H.L. Wu, G.L. Shen, R.Q. Yu, PLS Variable-weighted, Chem. Intell. Lab. Syst. 85 (2007) 140–143.
[11] S. Wold, H. Martens, H. Wold, The Multivariate Calibration Method in Chemistry Solved by the PLS Method, Springer-Verlag, Heidelberg, 1983.
[12] S. Wold, J. Cheney, N. Kettaneh, C. McCready, Chem. Intell. Lab. Syst. 84 (2006) 159–163.
[13] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, 1989.
[14] S. Wold, M. Sjostrom, L. Eriksson, Chemom. Intell. Lab. Syst. 58 (2001) 109–130.
[15] A. Alberich, J.M. Diaz-Cruz, C. Arino, M. Esteban, Analyst 133 (2008) 112–125.
[16] H. Yang, P.R. Griffiths, J.D. Tate, Anal. Chim. Acta 489 (2003) 125–136.
[17] V. Centner, J. Verdú-Andrés, B. Walczak, D. Jouan-Rimbaud, F. Despagne, L. Pasti, D.L. MassartO.E. de Noord, Appl. Spectrosc. 54 (2000) 608–623.
[18] T. Næs, T. Isaksson, T. Fearn, T. Davies, A User-Friendly Guide to Multivariate Calibration and Classification, NIR. Publications, Chichester, 2002.
[19] P.H.C. Eilers, Anal. Chem. 76 (2004) 404–411.
[20] T. Skov, F. van den Berg, G. Tomasi, R. Bro, J. Chemom. 20 (2006) 484–497.
[21] ⟨http://www.mathworks.com/⟩.
[22] ⟨http://www.eigenvector.com/⟩.
[23] B. Walczak, D.L. Massart, Anal. Chim. Acta 331 (1996) 177–185.
[24] B. Walczak, D.L. Massart, Anal. Chim. Acta 331 (1996) 187–193.
[25] H.M. Paiva, S.F. Carreiro Soares, R.K. Harrop Galvão, M.C.Ugulino Araújo, Chem. Intell. Lab. Syst. 118 (2012) 260–266.
[26] N.P.V. Nielsen, J.M. Carstensen, J. Smedsgaard, J. Chromatogr. A 805 (1998) 17–35.
[27] G. Tomasi, F. Van den Berg, C. Andersson, J. Chemom. 18 (2004) 231–241.
[28] A.J. Bard, L.R. Faulkner, Electrochemical Methods: Fundamentals and Applications, Second ed., John Wiley & Sons, Inc., New York, 2001.
[29] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193–1202.
[30] S. Wold, H. Antti, F. Lindgren, J. Ohman, Chem. Intell. Lab. Syst. 44 (1998) 175–185.
[31] S. de Jong, R.W. Farebrother, Chem. Intell. Lab. Syst. 25 (1994) 179–181.
[32] H. Wold, Path Models with Latent Variables: The NIPALS Approach, Academic Press, New York, 1975.
[33] Y. Ni, L. Wang, S. Kokot, Anal. Chim. Acta 439 (2001) 159–168.
[34] K. Bessant, S. Saini, J. Electroanal. Chem. 489 (2000) 76–83.
[35] M.C.U. Araujo, T.C.B. Saldanha, R.K.H. Galvão, T. Yoneyama, H.C. Chame, V. Visani, Chem. Intell. Lab. Syst. 57 (2001) 65–73.
[36] G. Li, Z.L. Chen, J. Am. Stat. Assoc. 381 (1985) 754–759.
[37] P.J. Rousseeuw, C. Croux, J. Am. Stat. Assoc. 88 (1993) 1273–1283.
[38] S. Serneels, P. Filzmoser, C. Croux, P.J. Van Espen, Chem. Intell. Lab. Syst. 76 (2005) 197–204.
[39] S. Serneels, C. Croux, P. Filzmoser, P.J. Van Espen, Chem. Intell. Lab. Syst. 79 (2005) 55–64.
[40] S. Sekulic, M.B. Seasholtz, Z. Wang, B.R. Kowalski, S.E. Lee, B.R. Holt, Anal. Chem. 65 (1993) 835–845.
[41] N. Xiaoying, Z. Zhilei, J. Kejun, L. Xiaoting, Food Chem. 133 (2012) 592–597.
[42] R.M. Balabin, R.Z. Safieva, E.I. Lomakina, Chemometr. Intell. Lab. 88 (2007) 183–188.
[43] R.M. Balabin, R.Z. Safieva, E.I. Lomakina, Chemometr. Intell. Lab. 93 (2008) 58–62.
[44] H. Guo, H.P. Liu, L. Wang, J. Syst. Simul. 18 (2006) 2033–2036.