Chemometrics and Intelligent Laboratory Systems 146 (2015) 437–446

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Advanced and tailored applications of an efficient electrochemical approach assisted by AsLSSR–COW–rPLS and finding ways to cope with challenges arising from the nature of voltammetric data

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

[a] Department of Analytical Chemistry, 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 & Technology, Department of Food Science, Faculty of Science, University of Copenhagen, Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 14 April 2015
Received in revised form 29 June 2015
Accepted 30 June 2015
Available online 4 July 2015

Keywords:
COW
icoshift
rPLS
Simultaneous quantification
Potential-shift correction
Variable selection

ABSTRACT

The use of chemometric data processing is becoming an important part of modern voltammetry. The most challenges arising from voltammetric data are interactions among analytes and the background interferents which may cause signal changes in comparison with pure analyte profiles, and sample-to-sample potential shifts in the analyte profiles. These disadvantages can be tackled by baseline- and potential shift-correction Regarding the above commented problems, performances of asymmetric least squares spline regression (AsLSSR) algorithm for baseline correction and two well-known chemometric tools including interval correlation optimized shifting (icoshift) and correlation optimized warping (COW) for potential shift correction were examined. Finally, the COW was chosen for potential shift correction before applying recursive weighted partial least squares (rPLS) for simultaneous quantification of dopamine (DP), serotonin (ST), acetaminophen (AC) and noradrenaline (NA). In contrast to many other variable selection methods, the rPLS method has the advantage that only the number of latent factors used in the PLS needs to be estimated. A multivariate calibration (MVC) model was developed as a quaternary calibration model in a blank human serum sample (drug-free) provided by a healthy volunteer to regard the presence of a strong matrix effect which may be caused by the possible interferents present in the serum, and it was validated and tested with two independent sets of analytes mixtures in blank and actual human serum samples, respectively. Fortunately, the AsLSSR–COW–rPLS approach was successful in simultaneous quantification of DP, ST, AC, and NAD in both blank and actual human serum samples.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

'Shift' is a common occurrence in chemistry. Many analytical techniques yield data where the same phenomena may yield variations at different positions or may have different 'durations' depending on the specific analytical conditions. Analogously, the measurements for the single samples can have different scales, or the sample vectors may have different lengths (e.g. different batch lengths in industrial processes).

Electroanalytical data have been analyzed by many chemometricians during the past four decades. However, during the early development of these methods, the difficulty of the electrochemical hard modeling and the lack of linearity between the current and concentration restricted the application of chemometrics in electroanalytical chemistry [1]. Fortunately, the rapid development of artificial neural networks and dramatic improvements in soft modeling methods, have enhanced the application of chemometrics to electroanalytical data during the last several years [2].

In the recent years, multivariate calibration methods applied to absorptive spectral and electrochemical data are being increasingly used for the analysis of complex mixtures [3,4]. Several tools have been reported in the literature for processing these data [5], although the most popular are principal component regression (PCR) [6] and partial least squares regression (PLS) [7]. All these techniques have the advantage of using the full spectral information and not only a characteristic peak value. Moreover, they allow a rapid determination of mixture components, often with no prior separation, and the calibration can be performed ignoring the concentrations of all components except the analyte of interest in complex samples.

In PLS regression [8] modeling, it is possible to give the X-variable weights [9]. This can result in models that are easier to handle and easier to interpret, and in some cases, it may even lead to models with a better predictivity than models based on unweighted variables. Other purposes of variable weighting may be to speed up calculations, or simply to select a limited number of important variables, that is, wavelength selection. Despite the high redundancy in spectroscopic data, the multivariate methods can often be improved by variable selection such as forward selection [10] or by regional selection such as interval PLS

* Corresponding author. Tel.: +98 8334274557; fax: +98 8334274559.
E-mail address: mbgholivand2013@gmail.com (M.-B. Gholivand).

http://dx.doi.org/10.1016/j.chemolab.2015.06.017
0169-7439/© 2015 Elsevier B.V. All rights reserved.

[11]. The primary reason for the improvements is the reduced number of interferences in the reduced set of variables. Perhaps even more important than improved predictions and parsimonious models is the improved interpretation as the selected variables or region(s) can be assigned by the spectroscopist and referred to a chemical or physical phenomenon. When an acceptable full range PLS model has been established, the normal procedure is to inspect the model parameters, for example, the regression coefficients. High absolute regression coefficients are considered important, and small regression coefficients are considered less important.

Monitoring of monoamines such as serotonin (ST), noradrenaline (NA), and dopamine (DP) has become an important research tool in the development of advanced analytical technologies [12]. In this respect, the techniques enabling simultaneous analysis of monoamines, e.g. ST, NA and DP, are of particular interest.

Acetaminophen (AC) is a widely used anti-pyretic and analgesic drug with actions similar to aspirin. It is an effective and safe agent for the relief of mild to moderate pain associated with headache, arthritis and postoperative pain. Its ready access has resulted in its increased use in attempted suicide [13]. Among different methods for determination of AC, electrochemical methods maybe the most widely applied because of high sensitivity, simplicity and reproducibility of this approach [14–18]. AC electrophysiological [19] effects support the idea that this potent analgesic drug can act in the central nervous system. Animal model studies have shown that AC might protect neurons from degeneration. For example, AC can protect primary rat embryonic DA neurons from glutamate toxicity [20]. Also, AC administration at antinociceptive doses affects serotonin and DP levels in various brain areas and the spinal cord in rats [21]. Additionally, important drugs such as AC will interfere with DP measurements in biological samples [22]. Thus, it is necessary to develop low-cost, simple, reproducible and reliable methods for simultaneous determination of DP and AC, which is a major goal of electroanalytical research. Therefore, simultaneous determination of DP, ST, AC, and NA is very important, but there is not any electrochemical report about simultaneous determination of these four compounds. Electroanalytical methods have attracted more attention in recent years for environmental and biological compounds determination due to their sensitivity, accuracy, and simplicity, but the treatment involved is time consuming and the associated cost is high. Thus it is important to develop new methodologies for the simultaneous determinations of these analytes. An attractive possibility is the use of chemometric and multivariate calibration methods.

For the first time, in this work we made a comparison between two well-known chemometric tools including COW and icoshift for voltammetric data alignment and finally AsLSSR–COW–rPLS was chosen for the simultaneous determination of DP, ST, AC and NA in human serum samples. Literature survey revealed that no attempt has been made till date to use the AsLSSR–COW–rPLS in an electrochemical approach.

2. Theoretical and experimental details

2.1. Theoretical details

2.1.1. COW

The COW algorithm is also based on a piece-wise linear correction function, but, unlike icoshifts, it is continuous and made up of segments whose slope is allowed to take a limited number of discrete values determined by the length ℓ of the interval in which the voltammograms are divided and the maximum number of scan points, s, by which the length of each interval is allowed to change [23]. When the slope of a segment of the correction function $f(t_y)$ is not one, the corresponding intervals in sample and target contain a different number of points and linear interpolation is used so that the interval in the sample is compressed or expanded to the same length as the corresponding interval in the target. The optimized cost function is the sum of the Pearson's correlation coefficient for all segments after interpolation and dynamic programming is used to attain the global maximum given the constraints. Typically, a maximum allowed correction is set to further reduce the feasible region for $f(t_y)$. One known problem of the standard COW method is that, close to the endpoints, the maximum correction allowed by the slope constraints is reduced. While it is possible to modify the algorithm to account for this, a computationally more intensive but equally effective solution is to attach zeroes at both ends of the signals so that the necessary flexibility is guaranteed (namely $w_{max} \not\ell s^{-1}$ zeroes should be attached at each end). It is worth mentioning, that, while COW also allows for custom intervals, here its commonest format is used in which sample and target are divided in segments of equal length.

2.1.2. icoshift

The icoshift program allows for different ways of aligning signals primarily due to different NMR alignment procedures [24]. In this method, one or more intervals are defined, either manually or automatically, on the potential-axis and the segments in the sample to be aligned are shifted in order to maximize their cross-correlation with the corresponding segment of the target. The local maximum cross-correlation upon shift is calculated using a Fast Fourier Transform computation engine in which the optimal correction for all the samples is computed together and it is possible to fill in the inserted part using missing values [24]. Compared to other FFT based alignment methods (e.g., PAFFT and RAFFT), endpoint contamination during the calculation of cross-correlation (aliasing) is avoided by padding the segments with a number of zeroes corresponding to the maximum allowed correction w_{max} and improved efficiency is obtained by treating all the samples at once [24]. The icoshift only implements simple segmentation strategies, e.g., one can define a number of segments of equal length, but accepts interval definitions provided by the user, as is the case here. The icoshift uses a greedy algorithm for the optimization [25], which means that the intervals are optimized separately, without accounting for any global effects. Note that, like other alignment methods but unlike COW, the icoshift optimization criterion is not well defined mathematically unless missing values are used for the insertion, because the cross-correlation function is not calculated after the correction is applied and the segment boundary point replicated. With respect to w_{max}, it is possible to let icoshift automatically increase it, for each interval separately, if the optimal shift for any of the signals is found to be exactly on the boundary of the interim w_{max} value. The iterative procedure starts at 5% of the segment length and increases by an additional 5% at each iteration, up to 50% of the segment length. The icoshift program has the built-in option of automatically picking the signal with the largest area under the curve as a target for each interval. This option was tested against the more standard approach of using as a single target either the average voltammogram or a representative sample with average potential shift.

2.1.3. Recursive weighted PLS or rPLS

The recursive weighted PLS or rPLS introduces a novel variable selection method where the univariate response variable y is used to guide the weighting in an iterative or recursive manner [26]. In rPLS, the regression coefficients are recursively used as weights on the original data matrix. Using the regression vector as a base for the weights is a reasonable idea as the regression vector reflects the importance of the variables. Regression vector weights near 0 indicate unimportant variables, and weights with large absolute values indicate important variables. The method suggested here is very simple, focusing on the regression coefficients and using them as recursive weights. One of the advantages of the method is that, in its simplest form, it only requires the user to determine the number of factors. No additional parameters such as threshold values or confidence limits are required.

The rPLS is related to the method proposed by Forina et al. [27] and jack-knifing [28], but instead of iteratively eliminating variables, the rPLS method iteratively uses the regression coefficients to magnify important variables and thus relatively down-weight less important ones. rPLS is based on a process of repeated PLS models; the current regression coefficients are used as cumulative weights on **X**:

$$X_R = X_{R-1} \times \mathrm{diag}(b) \quad (1)$$

where $\mathbf{X}_{R-1}$ is the previous updated weighted **X**, b is the regression coefficient from the last model (based on $\mathbf{X}_{R-1}$), and $\mathbf{X}_R$ is the "new **X**" to be used in the subsequent PLS model. If repeating the reweighting, the rPLS model has the fascinating property that it, under normal conditions, will converge to a very limited number of variables (good for interpretation), but that it will exhibit optimal regression performance at an earlier stage, normally including covarying neighbor variables.

2.1.4. 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 MVC models, each model was validated for the prediction of validation and test sets, evaluating root mean square errors of prediction (RMSEP), and relative error of prediction (REP):

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

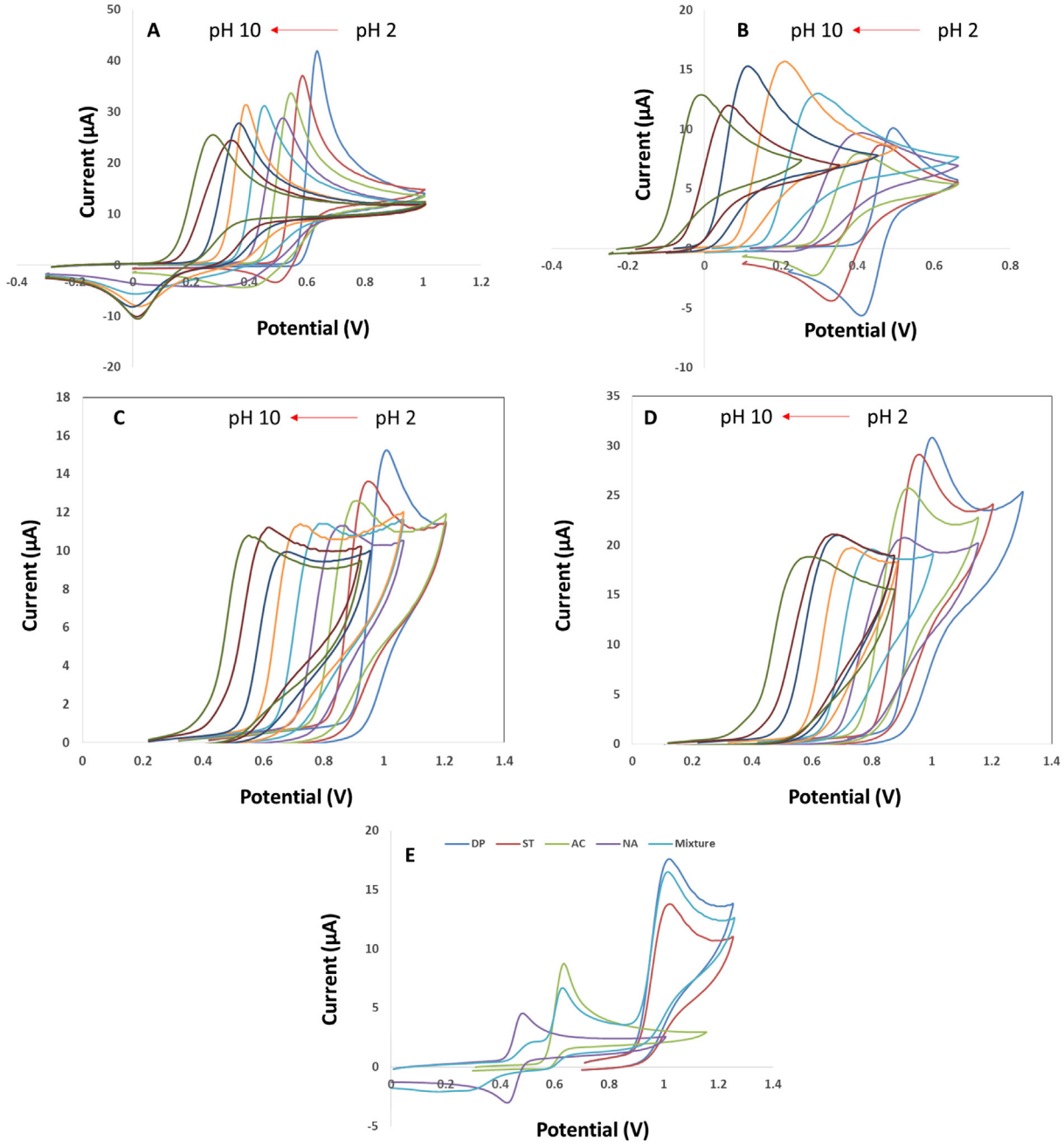

Fig. 1. Cyclic voltammograms of (A) AC (10^{-3} mol L^{-1}), (B) NA (2.5×10^{-3} mol L^{-1}), (C) DP (10^{-3} mol L^{-1}), and (D) ST (2.0×10^{-4} mol L^{-1}) in PBS (0.05 mol L^{-1}) at different pHs. (E) Cyclic voltammograms of AC (25×10^{-4} mol L^{-1}), NA (1.2×10^{-3} mol L^{-1}), DP (10^{-3} mol L^{-1}), ST (10^{-4} mol L^{-1}), and their mixture in PBS (0.05 mol L^{-1}, pH 2.0).

$$\mathrm{REP}(\%) = \frac{100}{y_{\mathrm{mean}}}\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{\mathrm{pred}} - y_{\mathrm{act}}\right)^2} \quad (3)$$

where y_{act} and y_{pred} are actual and predicted concentrations of each component, respectively, and n is the number of samples in validation or test set.

2.2. Experimental backgrounds

2.2.1. Chemicals and solutions

Dopamine (DP), serotonin (ST), acetaminophen (AC) and noradrenaline (NA) were purchased from Sigma-Aldrich. A phosphate buffered solution (PBS, 0.05 mol L^{-1}) of pH 2.0 was prepared from chemicals (analytical grade) including NaH_2PO_4 and Na_2HPO_4 from Merck (pH adjustment was carried out by adding H_3PO_4 or NaOH). All other chemicals used in the investigation were of analytical grade obtained from regular sources and used without further purification. Stock standard solutions of DP, ST, AC, and NA were prepared by exact weighing and dissolving of their solid powder in PBS, 0.05 mol L^{-1}, pH 2.0 with a concentration level of 0.5 mol L^{-1} and were stored at dark in a refrigerator until analysis time. Working solutions were prepared by appropriate dilution of the stock standard solutions with PBS, 0.05 mol L^{-1}, pH 2.0. All the solutions were prepared by doubly distilled water (DDW).

2.2.2. Apparatus and software

Electrochemical experiments were performed using a μ-Autolab controlled by the NOVA software (Version 1.8). A conventional three-electrode cell was used with a saturated calomel electrode (SCE) as reference electrode, a Pt wire as counter electrode and a GCE as working electrode. The pH of the solutions was adjusted using a JENWAY-3510 pH meter equipped by a combined glass electrode. All the recorded electrochemical data were smoothed, when necessary, and converted to data matrices by the use of several home-made m-files in MATLAB environment (Version 7.14, MathWorks, Inc.). All the computations for baseline correction, data alignment and multivariate calibration were performed in MATLAB environment. All the computations were performed on a DELL XPS laptop (L502X) with Intel Core i7-2630QM 2.0GHz, 8 GB of RAM and Windows 7-64 as its operating system.

2.2.3. Preparation of the serum samples

A blank human serum sample (drug-free) was provided by a healthy volunteer who not exposed to any drug for at least 10.0 months. An actual human serum samples were collected from a patient under AC treatment, kindly provided by a Medical Diagnostic Laboratory in Kermanshah, Iran. The following methodology was used to prepare all the serum samples: to eliminate protein and other substances, 5.0 mL of human serum sample was placed in a10.0 mL glass tube and 5.0 mL of 15.0% (w/v) zinc sulfate solution–acetonitrile (50/40,v/v) was added. The glass tube was vortexed for 20.0 min, maintained at 4.0 °C for 15.0 min followed by centrifugation at 4000.0 rpm for 5.0 min. Then, the supernatant was collected in the same tube and this solution was used for subsequent analyses.

2.2.4. Electrochemical procedure

Prior to electrochemical experiments, the GCE was successively polished to a mirror using 0.3 and 0.05 μm Alumina slurry. Afterward,

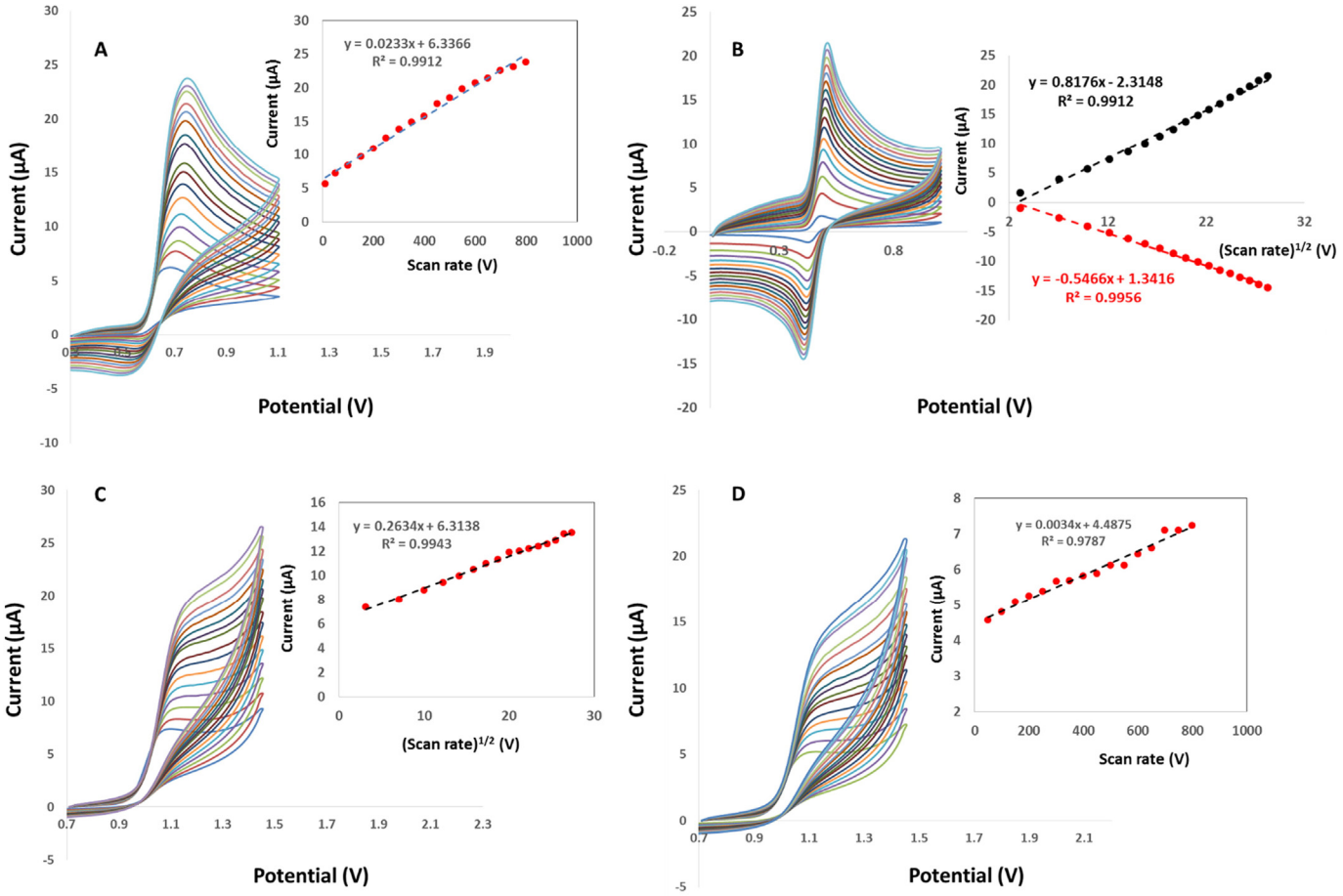

Fig. 2. Cyclic voltammograms of (A) AC (10^{-4} mol L^{-1}), (B) NA (10^{-3} mol L^{-1}), (C) DP (5.0×10^{-4} mol L^{-1}), and (D) ST (3.3×10^{-5} mol L^{-1}) in PBS (0.05 mol L^{-1}, pH 2.0) at different scan rates. Insets: variation of peak currents versus scan rate or $(\text{scan rate})^{1/2}$.

the electrode was washed thoroughly with ethanol and DDW and dried at room temperature. All electrochemical experiments were carried out at room temperature. The DPV measurements were carried out at the following operating conditions for the five studied analytes: step potential 0.035 V, modulation amplitude 0.055 V, modulation time 0.05 s, interval time 0.5 s, and scan rate 0.05 V s^{-1}.

3. Results and discussion

3.1. Electrochemical studies

3.1.1. pH dependence study

To select the best pH for the simultaneous determination of DP, ST, AC, and NA, the effect of pH on their cyclic voltammograms was investigated. Fig. 1A–D shows the influence of the pH of the PBS, 0.05 mol L^{-1}, in the range of 2.0–10.0, on the signal intensities of 0.1 mM AC, NA, DP, and ST. Taking into account that for analytical purposes both maximal and stable currents are necessary, a pH value of 2.0 was selected for further experiments. The oxidation peak potential of all studied analytes shifted to less positive values as the pH of the buffer solution was increased (see Fig. 1A–D).

3.1.2. Effect of scan rate

The influences of scan rate (υ) on the peak currents of AC, NA, DP, and ST at the GCE in PBS (0.05 mol L^{-1}, pH 2.0) were studied by cyclic voltammetry (Fig. 2A–D). In the studied range of υ, a linear relationship was established between I_p and υ for AC and ST, indicating the adsorption controlled mechanism, and a linear relationship was established between I_p and $\upsilon^{1/2}$ for NA and DP, indicating the diffusion controlled mechanism.

3.2. Chemometric analysis

3.2.1. A glance to necessity of MVC

Fig. 1E shows the CVs of AC, NA, DP, and ST and their mixture in PBS (0.05 mol L^{-1}, pH 2.0). In all conditions evaluated, a strong signal overlapping was observed for the simultaneous analysis of AC, NA, DP, and ST at the GCE. 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 MVC.

3.2.2. Calibrations

3.2.2.1. Univariate calibrations. Prior to multivariate calibration experiments, univariate calibration experiments were performed (Fig. 3A–D)

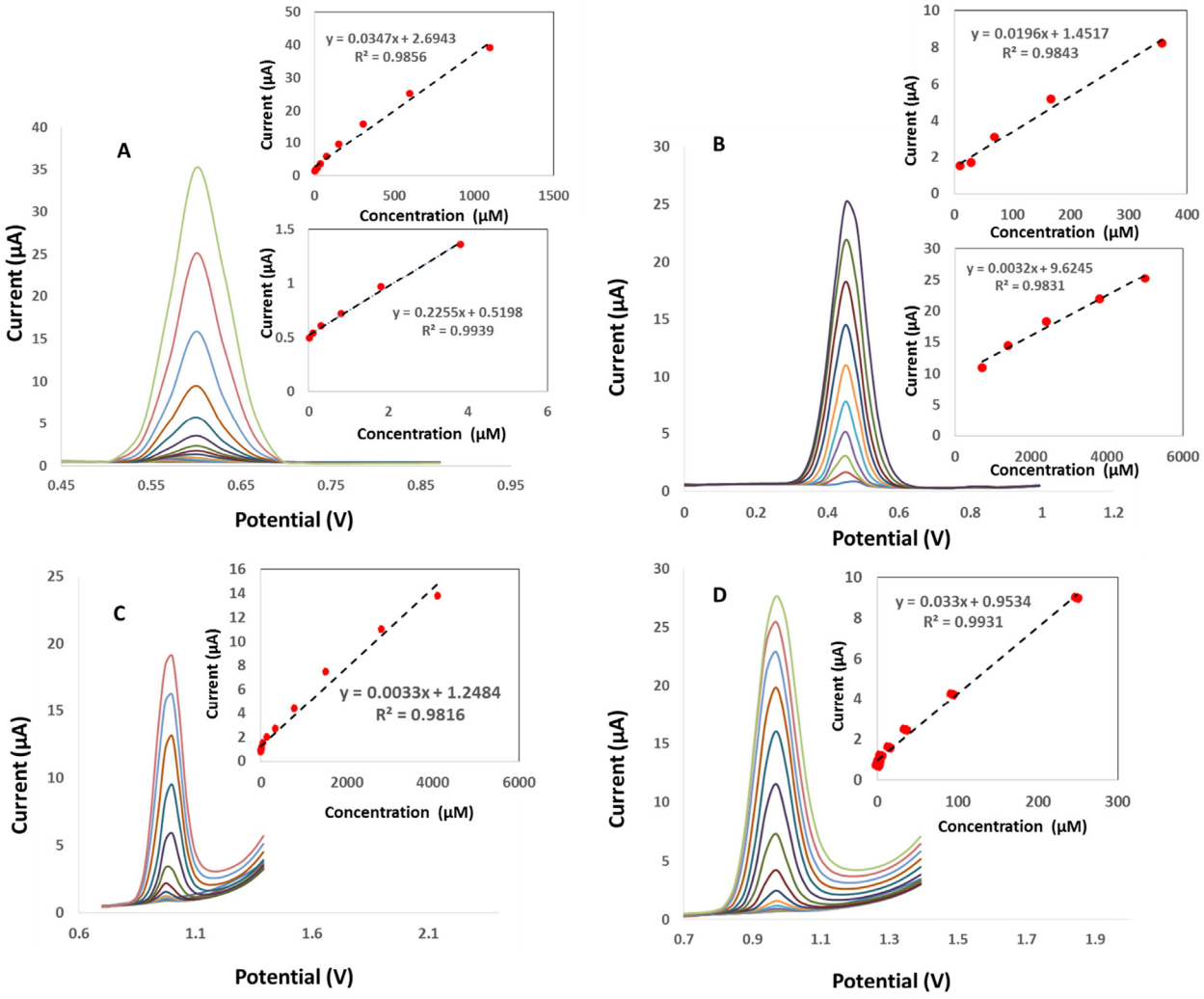

Fig. 3. Representative differential pulse voltammograms of (A) AC, (B) NA, (C) DP, and (D) ST in PBS (0.05 mol L^{-1}, pH 2.0) at different concentrations. Insets: dependence of peak current with concentration.

and calibration curves were constructed with several points as peak current versus analyte concentration in the ranges 0.02 to 3.82 and 3.82 to 1100 × 10^{-6} mol L^{-1} for AC (insets of Fig. 3A), 9.9 to 357.0 and 357.0 to 5000.0 × 10^{-6} mol L^{-1} for NA (inset of Fig. 3B), 0.02 to 4100.0 × 10^{-6} mol L^{-1} for DP (inset of Fig. 3C), and 0.1 to 248.0 × 10^{-6} mol L^{-1} for ST (inset of Fig. 3D), and evaluated by linear regression which were the limiting assayed concentrations in subsequent analyses. All analytes showed linear dependences between peak current and concentration at different concentration intervals.

3.2.2.2. Multivariate calibrations. When the analytes are analyzed in the presence of interferents, the electrochemical profile revealed additional changes to those observed in the absence of interferences. The main changes observed were minor alterations in the baseline 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 response and for this reason, the calibration and validation sets were prepared in a blank human serum sample (drug-free).

3.2.2.2.1. 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, the calibration set was prepared in a blank human serum sample (drug-free) which was collected from a healthy volunteer to regard the complex matrix of the serum sample which may contain a lot of unexpected interferences. This strategy was applied in order to provide rPLS enough information concerning the signals of the analytes when they are embedded into the real background. All the calibration mixtures (the compositions of the calibration mixtures were selected according to a central composite design (CCD), Table 1) were prepared in the blank human serum sample spiked with an appropriate amount of each analyte of interest considering the linear calibration ranges (previously established from univariate calibrations for each analyte). All samples were diluted with PBS (0.05 mol L^{-1}, pH 2.0) to adjust the pH and then appropriate amounts of these diluted samples were transferred to the electrochemical cell, and the solutions were measured in random order. Final concentration of each analyte was obtained by multiplying the detected value by the appropriate dilution factor.

3.2.2.2.2. Validation set. To check the prediction ability of the model after optimizing all calibration parameters, a validation set of ten quaternary mixtures (Table 1) was prepared in the blank human serum sample (drug-free). The concentrations of four analytes were selected at random from the corresponding calibration ranges. All samples were diluted with PBS (0.05 mol L^{-1}, pH 2.0) to adjust the pH and then appropriate amounts of these diluted samples were transferred to the electrochemical cell, and the solutions were measured in random order. Final concentration of each analyte was obtained by multiplying the detected value by the appropriate dilution factor.

3.2.2.2.3. Test set. With the purpose of evaluating the proposed method in a very interfering environment such as human serum, a test set of ten quaternary mixtures (Table 1) was prepared in an actual serum 1 (see Section 2.2.3) with random amount of each analyte of interest in the same concentration range used for calibration. All samples were diluted with PBS (0.05 mol L^{-1}, pH 2.0) to adjust the pH and then appropriate amounts of these diluted samples were transferred to the electrochemical cell, and the solutions were measured in random order. Final concentration of each analyte was obtained by multiplying the detected value by the appropriate dilution factor. It should be noted that all the samples related to the test set were prepared in a serum which contains AC therefore, the exact concentrations of the spiked AC were computed by a previous knowledge about the initial amount of AC in the serum which was obtained by analyzing the serum with HPLC method by a medical diagnostic laboratory prior to analyzing by the proposed method in this study.

3.2.3. Solution to challenges arising from voltammetric data

Besides the problem arising from the presence of severely overlapping analyte profiles, in the present study two additional complications may occur: (1) interactions among analytes and the background

Table 1
Composition of the samples used in the calibration, validation, and test sets.

	Calibration	(×10^{-6} mol L^{-1})				Validation	(×10^{-6} mol L^{-1})		
Samples	AC	DP	ST	NA	Samples	AC	DP	ST	NA
S_1	1300	1000	1500	1200	V_1	162.5	58. 7	70.0	360.9
S_2	1300	1000	1500	3600	V_2	568.7	93.3	40.0	192.5
S_3	1300	1000	4500	1200	V_3	243.7	106. 7	22.0	129.9
S_4	1300	1000	4500	3600	V_4	325.0	29.3	88.0	91.4
S_5	1300	3000	1500	1200	V_5	650.0	21.3	42.0	291.1
S_6	1300	3000	1500	3600	V_6	178.7	133.3	112.0	101.0
S_7	1300	3000	4500	1200	V_7	1300.0	266. 7	150	187.7
S_8	1300	3000	4500	3600	V_8	812.5	666. 7	68.0	216.5
S_9	3900	1000	1500	1200	V_9	81.2	56.0	84.0	182.8
S_{10}	3900	1000	1500	3600	V_{10}	536.2	21.3	44.0	132.3
S_{11}	3900	1000	4500	1200					
S_{12}	3900	1000	4500	3600		Test	(×10^{-6} mol L^{-1})		
S_{13}	3900	3000	1500	1200	Samples	AC	DP	ST	NA
S_{14}	3900	3000	1500	3600	T_1	1625	56.0	44.0	842.1
S_{15}	3900	3000	4500	1200	T_2	3412.5	61.3	22.0	986.4
S_{16}	3900	3000	4500	3600	T_3	812.5	120.0	16.0	79.4
S_{17}	0.02	2000	3000	2400	T_4	1056.2	144.0	152.0	26.5
S_{18}	5200	2000	3000	2400	T_5	715.0	88.0	62.0	16.8
S_{19}	2600	0.02	3000	2400	T_6	520.0	29.3	36.0	156.4
S_{20}	2600	4100	3000	2400	T_7	178.75	21.3	28.0	129.9
S_{21}	2600	2000	0.1	2400	T_8	1430	18.7	44.0	187.7
S_{22}	2600	2000	5900	2400	T_9	1462.5	10. 7	64.0	52.9
S_{23}	2600	2000	3000	9.99	T_{10}	893.7	64.0	112.0	84.2
S_{24}	2600	2000	3000	5000					
S_{25}	2600	2000	3000	2400					
S_{26}	2600	2000	3000	2400					
S_{27}	2600	2000	3000	2400					
S_{28}	2600	2000	3000	2400					

interferents present in the serum, which may cause signal changes in comparison with pure analyte profiles, and (2) sample-to-sample potential shifts in the analyte profiles, which are common in voltammetric studies. For tackling the first problem, it was necessary to include the possible interferents in the calibration set in order to allow rPLS to model the analyte–background interactions before prediction on new samples. Concerning the second commented problem, some preprocessing alternatives were independently applied on the electrochemical responses before rPLS model building and validation.

In our previous works [29–31] we pointed out that voltammetric performance can be enhanced by eliminating noise and background components therefore, baseline elimination is a crucial step for reducing both complexity and number of the unexpected components. Moreover, it was demonstrated that the use of signal pre-treatments such as baseline- and potential shift-corrections improve the quality of voltammetric signals and, as a consequence, the performance of resolution by chemometric algorithms. In the next sections, some strategies are examined to achieving the mentioned aims.

3.2.3.1. Baseline correction. Baseline correction has been considered as a critical step for enhancing the signals and reducing the complexity of the analytical data [32]. Considering this aim, we used the method proposed by Eilers at al. [33] for background elimination in two-dimensional signals based on asymmetric least squares splines regression approach. For an understanding of the assumptions and conditions necessary to successfully apply baseline correction the reader is referred to Refs. [33,34].

3.2.3.2. Potential shift correction. For chemometric model building, several strategies have been proposed to align shifted signals such as icoshift and COW [23]. However, they have been scarcely described for electrochemical signals [29–31]. 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, among others [35].

A basic assumption for application of a first-order multivariate calibration model is the data bi-linearity, which may be compromised by the above commented potential shifts. Therefore, the DPV signals were aligned towards a target using icoshift and COW. For applying COW, first, the segment and slack were optimized based on the Simplicity concept, using a simplex-like optimization routine. Then, mean voltammogram was selected as target "signal". The voltammograms were also aligned using icoshift based on five intervals toward the average voltammogram as a reference.

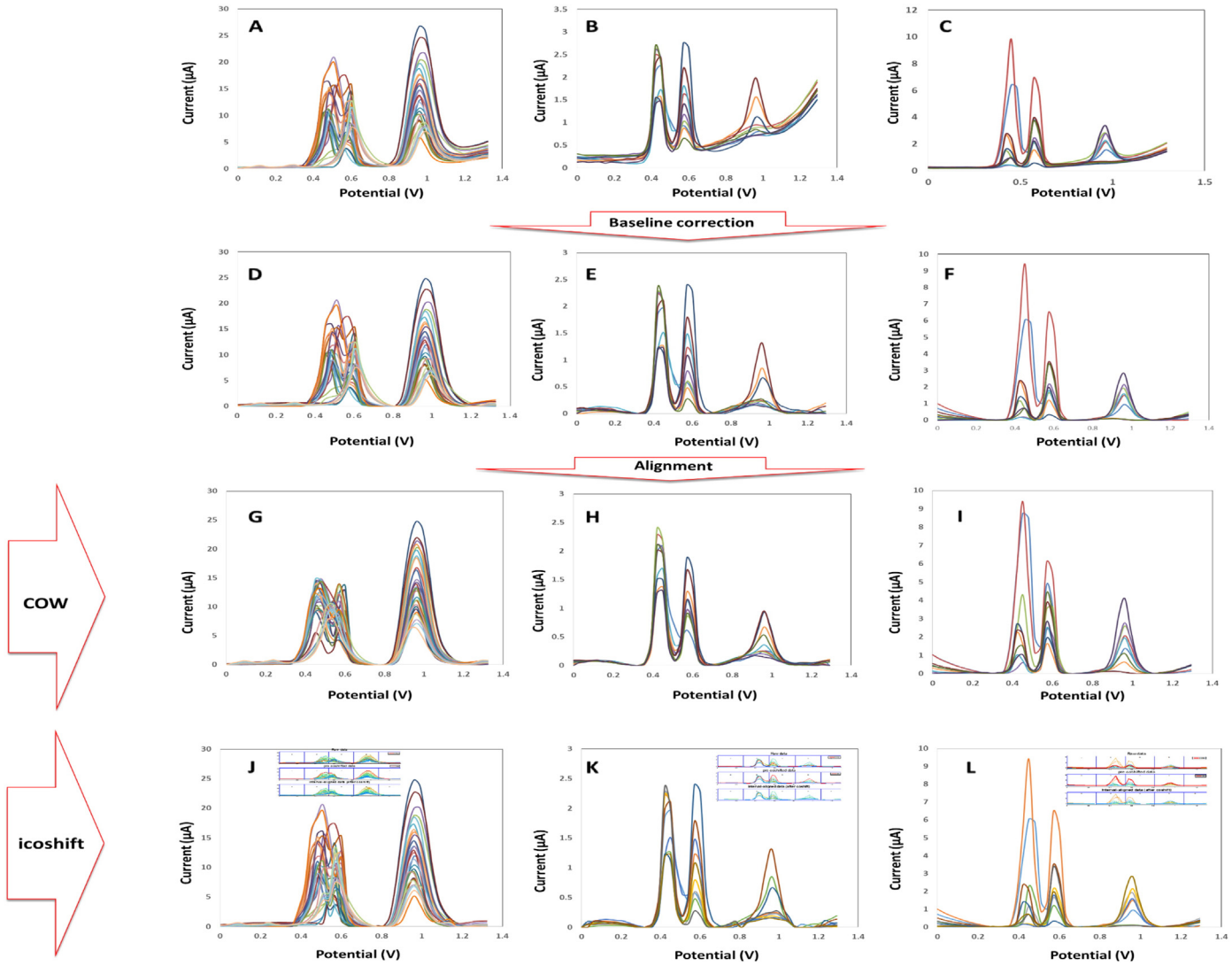

Fig. 4. (A)–(C): Raw data related to calibration, validation, and test sets, respectively. (D)–(F): Baseline corrected data by AsLSSR related to calibration, validation, and test sets, respectively. (G)–(I): Shift-corrected data by COW related to calibration, validation, and test sets, respectively. (J)–(L) Shift-corrected data by icoshift related to calibration, validation, and test sets, respectively.

Table 2
Error estimation for AsLSSR–COW–rPLS and AsLSSR–icoshift–rPLS approaches for simultaneous determination of AC, DP, ST, and NA.

Calibration approach	Analyte	rPLS components	RMSEP[a](μM)	REP[a](%)	RMSEP[b] (μM)	REP[b] (%)
AsLSSR–COW–rPLS	AC	4	3.01×10^{-6}	1.1	2.11×10^{-6}	1.2
	DP	4	4.18×10^{-6}	1.8	4.10×10^{-6}	1.9
	ST	4	5.11×10^{-6}	1.3	3.08×10^{-6}	1.3
	NA	4	7.88×10^{-6}	1.4	4.23×10^{-6}	1.7
AsLSSR–icoshift–rPLS	AC	6	9.34×10^{-5}	4.8	–	–
	DP	7	9.75×10^{-5}	5.2	–	–
	ST	5	8.35×10^{-5}	5.5	–	–
	NA	6	9.10×10^{-5}	5.1	–	–

[a] Validation set.
[b] Test set.

The results of baseline- and shift-corrections are shown in Fig. 4. Fig. 4A–C shows the raw DPV data recorded for the calibration, validation, and test sets, respectively. Fig. 4D–F shows the baseline-corrected data of calibration, validation, and test sets, respectively, and as can be seen the baselines are satisfactorily corrected. Fig. 4G–I and J–L shows the results of applying COW and icoshift for aligning the data, respectively.

3.3. Performance evaluation of AsLSSR–COW–rPLS and AsLSSR–icoshift–rPLS in blank and actual serum samples

In multivariate calibration, it is usual to have two data sets: a calibration set, employed to build the regression model, and a validation set to check the prediction ability of the model after the number of latent variables has been optimized, but in this study we are going to determine the analytes' concentrations in human serum samples which have a very complex matrix due to the presence of many interferents present in the serum therefore, both calibration and validation sets were prepared in a blank human serum (drug-free) and with the purpose of evaluating performance of the method, a test set was also prepared in an actual human serum sample.

Both shift-corrected datasets by COW or icoshift were assigned for processing by rPLS to identify which approach (AsLSSR–COW–rPLS or AsLSSR–icoshift–rPLS) offers the best predictions. Prior to any modeling the presence of outliers was checked in order to focus the analysis on the variable selection method itself, and not to confuse it with the complex outlier detection problem, which definitely may occur during variable selection. An eight segmented cross-validation using Venetian blinds according to the reference value was used. During the recursive progress, the root mean squared error of prediction (RMSEP) value for

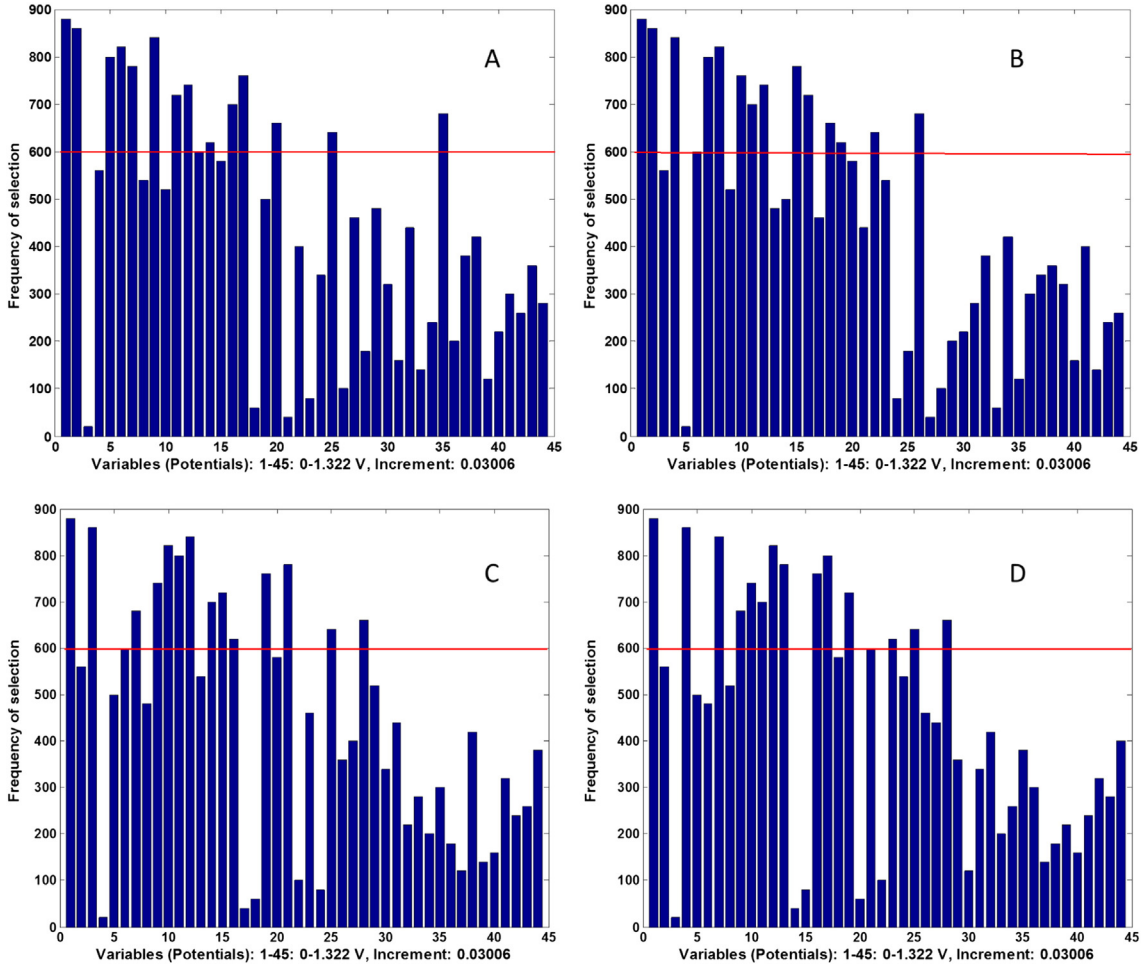

Fig. 5. Frequency of selection of different variables (potentials) by AsLSSR-COW-rPLS for (A) AC, (B) DP, (C) ST, and (D) NA.

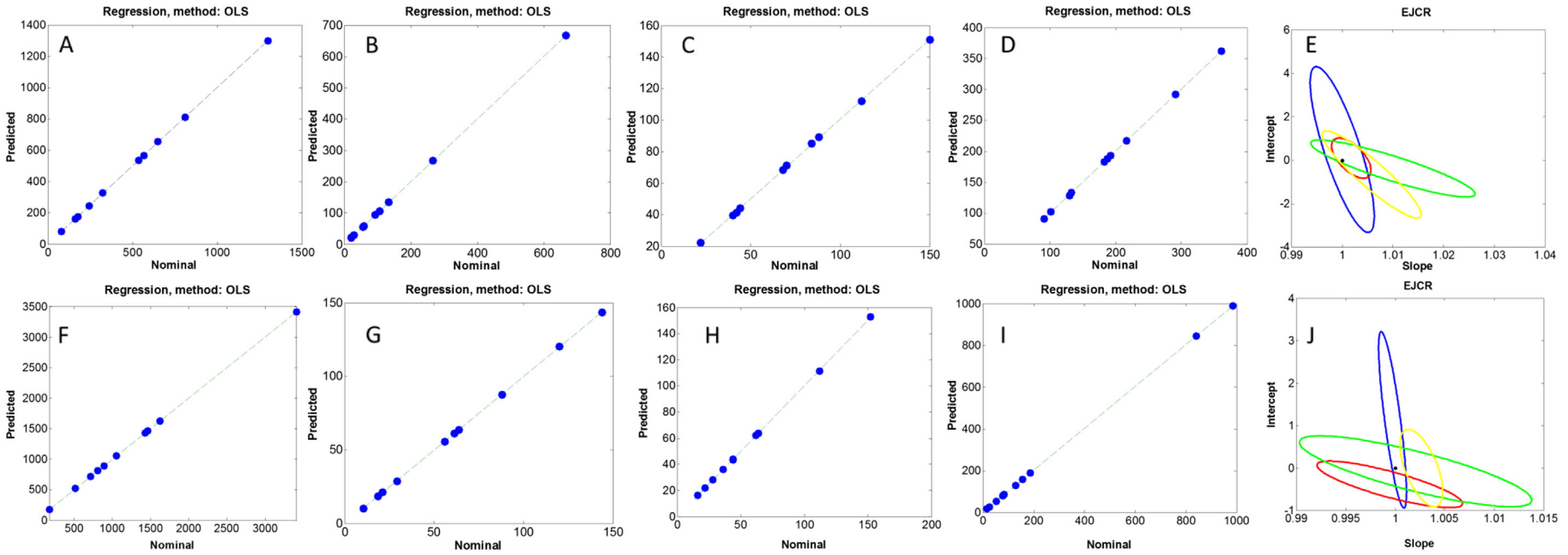

Fig. 6. Plots for predicted concentrations by AsLSSR-COW-rPLS as a function of nominal values for (A) AC, validation set, (B) DP, validation set, (C) ST, validation set, (D) NA, validation set, (F) AC, test set, (G) DP, test set, (H) ST, test set, and (I) NA, test set. (E) and (J) Elliptical joint regions (at 95% confidence level) related to the predictions by AsLSSR-COW-rPLS for AC (blue ellipse), DP (red ellipse), ST (green ellipse), and NA (yellow ellipse) for validation and test set, respectively. The black point marks the theoretical (0,1) point. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

each iteration was saved for identification of the optimal model (lowest RMSEP). In addition, the selection frequency of the raw variables was also recorded to investigate the consistency of the variable selection. The estimated number of factors was based on root mean squared errors of cross-validation (RMSECV) values from an eight segmented cross-validation on the full range dataset.

In order to evaluate the performance of the above-mentioned approaches, each model was validated for prediction of the validation set, evaluating RMSEP and REP. The results obtained for each of the above approaches are presented in Table 2. Notice that the optimum number of AsLSSR–icoshift–rPLS latent variables in all cases is larger than the theoretically expected value of four, probably because of inefficiency of icoshift to correct the above commented potential shifts in the studied signals. According to the obtained RMSEP and REP values, for the four studied analytes the best results were obtained when the AsLSSR–COW–rPLS approach was used. Therefore, AsLSSR–COW–rPLS will be applied for the analysis of actual human serum samples. Fig. 5A–D shows the selected variables by AsLSSR-COW-rPLS for AC, DP, ST, and NA, respectively by processing the potential shift-corrected data by COW. As can be seen, fourteen potentials have been selected for each analyte.

With the purpose of analyzing the potentiality of the evaluated methodology based on DPV data processed by AsLSSR–COW–rPLS, a test set of ten quaternary mixtures (see Table 1) was prepared in the actual serum with random amount of each analyte of interest in the same concentration range used for calibration. All samples were diluted with PBS (0.05 mol L^{-1}, pH 2.0) to adjust the pH and then appropriate amounts of these diluted samples were transferred to the electrochemical cell. Final concentration of each analyte was obtained by multiplying the detected value by the appropriate dilution factor. According to the results presented in Table 2 for the test set, satisfactory values for RMSEP and REP for the four analytes of interest, it is apparent that the AsLSSR–COW–rPLS approach has found the correct answer.

For the sake of a further investigation into the accuracy of the proposed method, the predicted concentrations of both validation and test sets were regressed on the nominal concentrations. In this case an ordinary least squares (OLS) analysis of predicted concentrations versus nominal concentrations was applied [36]. The calculated intercept and slope were compared with their theoretically expected values (intercept = 0, slope = 1), based on the elliptical joint confidence region (EJCR) test. If the ellipses contain the values 0 and 1 for intercept and slope (ideal point), respectively, showing the predicted and nominal values do not present significant difference at the level of 95% confidence and the elliptic size denotes precision of the analytical method, smaller size corresponds to higher precision [37]. Fig. 6A–D and F–I shows the regression of predicted concentrations on nominal values based on OLS method corresponding to the validation and test sets, respectively, and Fig. 6E and J shows the corresponding ellipses of the EJCR analyses. As can be concluded from Fig. 6A–D and F–I, the predictions for AC, DP, ST, and NA in both validation and test sets are in good agreement with the nominal values. If the EJCRs for the test set are analyzed (Fig. 6J), it is notable that while the ellipses for AC, and ST include the theoretically expected point (ideal point), indicating accuracy of the developed methodology for these analytes, the ideal point is fallen outside the ellipses of DP and NA, denoting slightly poorer prediction accuracy for DP and NA.

4. Conclusions

In the present work, integration of electrochemistry with chemometrics led us to introduce an efficient analytical method for simultaneous determination of AC, NA, DP, and ST at a GCE in complex matrices. A strong voltammetric overlapping was observed for the simultaneous analysis of these compounds. The overlapping was successfully resolved using AsLSSR–COW–rPLS. The baseline of the DPV signals was successfully removed by AsLSSR as an efficient chemometric algorithm. Because of the non-bilinearity observed in the experimental data, the potential shift correction was carried out by icoshift and COW, and the COW was chosen as the best algorithm for data alignment. To regard the presence of a strong matrix effect which may be caused by the possible interferents present in the human serum sample, the MVC model was built and validated in a blank human serum sample (drug-free) provided by a healthy volunteer which allowed us to exploit first-order advantage for the simultaneous determination of the studied compounds in very interfering media such as human serum samples. Finally, the application of the developed method to simultaneously assay the concentrations of AC, NA, DP, and ST in an actual human serum sample allowed to obtain satisfactory results. This study allows one to propose the present method as a promissory, cheap and accessible alternative for simultaneous determination of AC, NA, DP, and ST in human serum samples.

Acknowledgments

The authors gratefully acknowledge the financial support of this study by Razi University Research Council.

References

[1] A. Alberich, J.M. Diaz-Cruz, C. Arino, M. Esteban, Analyst 133 (2008) 112.
[2] M. Esteban, C. Arino, J.M. Diaz-Cruz, Crit. Rev, Anal. Chem. 36 (2006) 295.
[3] J.A. Arancibia, G. Martinez-Delfa, C.E. Boschetti, G.M. Escandar, A.C. Olivieri, Anal. Chim. Acta 553 (2005) 141.
[4] J.A. Arancibia, A. Rullo, A.C. Olivieri, S. Di Nezio, M. Pistonesi, A. Lista, B.S. Fernandez, Band. Anal. Chim. Acta 512 (2004) 157.
[5] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, 1989.
[6] H. Martens, M. Martens, Multivariate Analysis of Quality—An Introduction, Wiley, Chichester, 2000.
[7] S. Wold, M. Sjostrom, L. Eriksson, Chemom. Intell. Lab. Syst. 58 (2001) 109.
[8] S. Wold, H. Martens, H. Wold, Lect. Notes Math. 973 (1983) 286–293.
[9] C.L. Stork, B.R. Kowalski, Chemom. Intell. Lab. Syst. 48 (1999) 151–166.
[10] L. Wilkinson, G.E. Dallal, Technometrics 23 (1981) 377–380.
[11] L. Nørgaard, A. Saudland, J. Wagner, J.P. Nielsen, L. Munck, S.B. Engelsen, Appl. Spectrosc. 54 (2000) 413–419.
[12] T. Yoshitake, S. Yoshitake, K. Fujino, H. Nohta, M. Yamaguchi, J. Kehr, J. Neurosci. Methods 140 (2004) 163–168.
[13] C. Radovan, C. Cofan, D. Cinghita, Electroanalysis 20 (2008) 1346–1353.
[14] H. Beitollahi, I. Sheikhshoaie, J. Electroanal. Chem. 661 (2011) 336–342.
[15] C.X. Xu, K.J. Huang, Y. Fan, Z.W. Wu, J. Li, J. Mol. Liq. 165 (2012) 32–37.
[16] H. Xiong, H. Xu, L. Wang, S. Wang, Microchim. Acta 167 (2009) 129–133.
[17] M.R. Akhgar, H. Beitollahi, M. Salari, H. Karimi-Maleh, H. Zamani, Anal. Methods 4 (2012) 259–264.
[18] Q. Wan, X. Wang, F. Yu, X. Wang, N. Yang, J. Appl. Electrochem. 39 (2009) 785–790.
[19] H. Karimi-Maleh, M. Keyvanfard, K. Alizad, M. Fouladgar, H. Beitollahi, A. Mokhtari, F. Gholami-Orimilnt, J. Electrochem. Sci. 6 (2011) 6141–6150.
[20] K.H. Carlsson, I. Jurna, Neurosci. Lett. 77 (1987) 339–343.
[21] D. Casper, U. Yaparpalvi, N. Rempel, P. Warner, Neurosci. Lett. 289 (2000) 201–204.
[22] J.P. Courade, F. Caussade, K. Martin, D. Besse, C. Delchambre, N. Hanoun, M. Hamon, A. Eschalier, A. Cloarec, Naunyn Schmiedeberg's Arch. Pharmacol. 364 (2001) 534–537.
[23] G. Tomasi, F. Savorani, S.B. Engelsen, J. Chromatogr. A 1218 (2011) 7832–7840.
[24] F. Savorani, G. Tomasi, S.B. Engelsen, J. Magn. Reson. 202 (2010) 190.
[25] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, Introduction to Algorithms, The MIT Press, USA, 2001.
[26] Å. Rinnan, M. Andersson, Carsten ridder, S. B. Engelsen J. Chemom. 28 (2014) 439–447.
[27] F.M, Casolini, P. Millan, J. Chemom. 13 (1999) 165-184.
F.M. Casolini, P. Millan, J. Chemom. 13 (1999) 165–184.
[28] H. Martens, M. Martens, Food Qual. Prefer. 11 (2000) 5–16.
[29] M.B. Gholivand, A.R. Jalalvand, H.C. Goicoechea, Th. Skov, Talanta 119 (2014) 553–563.
[30] M.B. Gholivand, A.R. Jalalvand, H.C. Goicoechea, R. Gargallo, Th. Skov, G. Paimard, Talanta 131 (2015) 26–37.
[31] A.R. Jalalvand, M.B. Gholivand, H.C. Goicoechea, Th. Skov, Talanta 134 (2015) 607–618.
[32] M.M. De Zan, M.D. Gil García, M.J. Culzoni, R.G. Siano, H.C. Goicoechea, M. Martínez Galera, J. Chromatogr. A 1179 (2008) 106–114.
[33] P.H.C. Eilers, Parametric time warping, Anal. Chem. 76 (2004) 404–411.
[34] P.H.C. Eilers, I.D. Currie, M. Durban, Comput. Stat. Data An. 50 (2006) 61–76.
[35] A.J. Bard, L.R. Faulkner, Electrochemical Methods: Fundamentals and Applications, John Wiley & Sons, Inc., New York, 2001.
[36] A.G. Gonzalez, M.A. Herrador, A.G. Asuero, Talanta 48 (1999) 729–736.
[37] J.A. Arancibia, G.M. Escandar, Talanta 60 (2003) 1113–1121.