Analytica Chimica Acta 1080 (2019) 1–11

Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Complex numbers-partial least-squares applied to the treatment of electrochemical impedance spectroscopy data

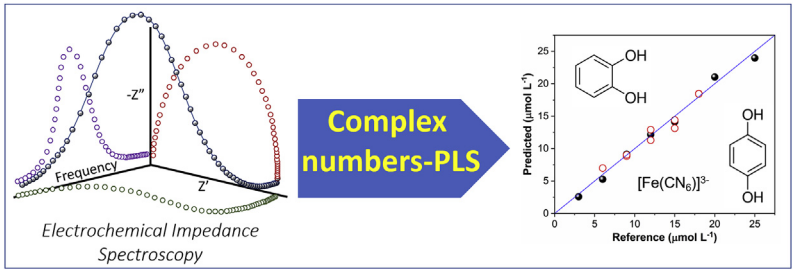
Dayvison Ribeiro Rodrigues [a], Alejandro César Olivieri [b], Wallace Duarte Fragoso [a], Sherlan Guimarães Lemos [a, *]

[a] Grupo de Estudos Avançados em Química Analítica, Departamento de Química, Universidade Federal da Paraíba, 58051-970, João Pessoa, PB, Brazil
[b] Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Instituto de Química de Rosario (IQUIR-CONICET), Suipacha 531, Rosario, S2002LRK, Argentina

HIGHLIGHTS

- Partial least-squares applied to complex numbers.
- Predictive models obtained directly from EIS complex data.
- No required use of EIS complex spectra as real numbers.
- Complex numbers-PLS is more efficient for more complex electrochemical systems.
- Direct analysis of tap water samples for dihydroxybenezes determination.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 May 2019
Received in revised form
17 July 2019
Accepted 22 July 2019
Available online 29 July 2019

Keywords:
Electrochemical impedance
Multivariate analysis
Complex numbers
Partial least-squares
Dihydroxybenzenes

ABSTRACT

This work investigated the application of partial least-squares regression of complex numbers on multivariate data obtained by electrochemical impedance spectroscopy (EIS). The use of complex numbers–PLS was evaluated in the individual determination of two well-known redox probes: ferrocyanide and hydroquinone. The predictive ability of complex numbers–PLS was evaluated for EIS spectra obtained at different applied potentials and perturbation amplitudes, and was also compared to that obtained with PLS applied to EIS data presented as real numbers - only the real or imaginary part of the complex impedance, or the absolute impedance or the phase angle. It is shown that complex numbers–PLS is more efficient (better prediction models) when more complex electrochemical systems (hydroquinone) are probed. Excellent predictions were obtained for the determination of hydroquinone and catechol in the direct analysis of spiked tap water samples with EIS and complex numbers–PLS.

1. Introduction

Electrochemical impedance spectroscopy (EIS) is known to be one of the most versatile electrochemical techniques since it allows that electrochemical processes with several time constants at the electrode/electrolyte interface could be probed in a single experiment. The most usual way of measuring electrochemical impedance involves the application of an alternating voltage signal of small amplitude at the interface, with the frequency varying over a wide range. Recording the resulting current signal at each selected frequency generates a register with tens of complex impedance values (impedance spectrum). Several visual representations of the spectra are known, and some interfacial processes can be elucidated directly from them [1,2]. However, the most popular

* Corresponding author.
E-mail addresses: sherlan@quimica.ufpb.br, sglemos@pq.cnpq.br (S.G. Lemos).

https://doi.org/10.1016/j.aca.2019.07.047
0003-2670/

approach to analyzing impedance spectra consists of rationalizing the interface as an equivalent circuit consisting of resistances, capacitances, and inductances representing the dielectric properties and the electrochemical processes occurring at the interface, as well as the alignment between these elements (parallel and serial connections) representing the physical reality of the interface. Conclusions can be obtained, and the system can be characterized from the results of the mathematical fit of the chosen circuit to the experimental data. On the other hand, modelling real electrochemical systems using an equivalent circuit can be complicated and, while useful, is often a simplification [3]. The consequence is that important information about the interface can be lost.

An alternative way of handling EIS data could be pondered. If one considers the electrochemical impedance spectra as a set of complex numbers – one for each *k* frequency tested – obtained for *n* samples, the investigation of this type of data array (*n* x *k* matrix) can be easily performed with multivariate analysis. Such approach allows the analysis of all data simultaneously, without any subjective assumptions about equivalent circuits, by considering the covariance between frequencies the most important phenomenon. Another important feature of the multivariate approach is that even the slightest change in a relaxation process is not discarded [4]. Examples of the use of multivariate analysis in the treatment of impedance data include the detection of phenolic compounds [5], microorganisms [6], sugars [7], surfactants [8], and drugs [9]. In this sense, different ways of how impedance spectra are used in multivariate analysis have been observed. The most usual is to change from complex numbers to real numbers, since the use of multivariate analysis has been reported almost exclusively for the last ones. The simplest and most common way is to use only the real or the imaginary part of the complex impedance spectra [10,11] or by performing the low-level data fusion of both parts [12], i.e., by carrying-out the juxtaposition of the real and imaginary parts in the composition of the sample vector. Additionally, the representations found in the Bode plot – logarithm of the absolute impedance (log|Z|) or the phase angle plotted with log frequency on the X-axis – are also employed [7,13]. However, it is obvious to conclude that the use of impedance data as real numbers can lead to the loss of important information about the processes that occur at the electrochemical interface, which could lead to quantitative models with unsatisfactory predictive power. The premise is that EIS data contain information on both real and complex parts of the complex numbers and should ideally be analyzed by complex data analysis.

In 2007, Geladi et al. [12] introduced the use of singular-value decomposition (SVD) in the analysis of impedance spectra as complex numbers, allowing the execution of principal component analysis (PCA) directly in the original data. In summary, SVD of complex numbers gives real singular values, but complex scores and loadings. Geladi and coworkers showed that the proposed approach is better at reducing data dimensionality than when real numbers are used. They also proposed the use of SVD in multivariate calibration [9], correlating the obtained scores with the analyte concentration through multiple linear regression, constituting a principal component regression (PCR) approach expressed as [9]:

$$y = [1U]b + f \quad (1)$$

where **y** is the vector of concentrations, **U** is the matrix containing the real part of the complex scores obtained with SVD, **1** is a vector where all elements are equal to 1, **b** is the regression coefficients vector, and **f** is the residue vector. This approach took an important step towards the use of all information available in electrochemical impedance spectra, especially in studies of the electrode/electrolyte interface. However, the full-potential of EIS in the development of regression models remains unexplored, since only the real part of the complex scores has been used. Therefore, to effectively employ all the information contained in EIS data for multivariate calibration, the present work proposes the evaluation of partial least-squares regression (PLS) performed directly on complex impedance spectra for the development of EIS-based calibration methods.

PLS regression is based on the properties of the nonlinear iterative partial least-squares algorithm (NIPALS) developed by Herman Wold [14]. The principle of PLS regression is to decompose both the matrix of instrumental responses **X** and the matrix of responses **Y** (considering the case of multiple analytes) in their respective scores, loadings and residues matrices:

$$X = TP' + E \quad (2)$$

$$Y = UQ' + F \quad (3)$$

where **T** and **U** are score matrices ($i \times a$), **P** and **Q** are loading matrices ($k \times a$), **E** and **F** are residue matrices ($i \times k$). In this case, the' index corresponds to the common nonconjugate transpose. In PLS, the decompositions of **X** and **Y** are not done independently as in PCR, but considering the information of both matrices. After decomposition, the regression between **T** and **U** is performed. The application of PLS to complex numbers is simple and straightforward, by simply using the Hermitian transpose instead of the common nonconjugate transpose. In the case of the **Y** matrix, where the values are often real numbers, the complex conjugate transpose produces the same results as the nonconjugate transpose. In the software used to perform the calculations (Matlab), the complex conjugate transpose is the standard operation of matrix transposition. Consequently, both SVD and PLS are naturally performed in this environment considering the Hermitian transpose, with no need for changes in the routines available in the literature and often used for the treatment of real numbers. An important consequence of PLS is the use of the complex scores in the construction of the regression models, not a representation as real numbers (eq. (1)), allowing the effective use of all available information present in the impedance data.

In this work, complex numbers-PLS was initially evaluated in the EIS-based determination of ferrocyanide, a redox probe with well-reported electrochemical behavior [15–17]. The predictive ability of complex numbers-PLS models was assessed in comparison to PLS applied to EIS data as real numbers – by using only the real or imaginary parts of the complex impedance, or the absolute impedance or the phase angle. It was also evaluated the influence of the experimental parameters of EIS – applied potential and amplitude – on the predictive ability of PLS models. The same methodology was applied in the determination of a more complex redox probe: hydroquinone. Finally, hydroquinone and catechol and were determined with EIS and complex numbers–PLS in the direct analysis of spiked tap water samples.

2. Experimental

2.1. Reagents and materials

The following reagents and materials were used: sodium acetate, acetic acid glacial, potassium ferrocyanide, catechol, hydroquinone, multi-walled carbon nanotubes (≥98% carbon basis, O.D. × I.D. × L 10 nm ± 1 nm × 4.5 nm ± 0.5 nm × 3–6 μm, Sigma-Aldrich), and mineral oil. All chemicals were of analytical grade, used without further purification, and purchased from Sigma–Aldrich, except for the mineral oil, which was obtained from Specsol. Tap water samples were collected from our lab.

Ultrapure water (≥18 MΩ cm), produced by a Millipore water purification system, was used throughout the experiments.

2.2. Apparatus and measurements

Electrochemical measurements were performed in a potentiostat Autolab PGSTAT M101 (Metrohm Autolab, The Netherlands) with a frequency response analyzer module controlled by software Nova 1.11 in a conventional three-electrode electrochemical cell. The reference and auxiliary electrodes were a saturated calomel electrode and a pencil lead, respectively. Two working electrodes were used: i) glassy carbon; ii) carbon nanotubes paste electrode prepared with 60% multi-walled carbon nanotubes (Sigma-Aldrich) and 40% mineral oil (Specsol®) w/w. All measurements were carried out at room temperature (25 ± 1 °C) in a 20 mL cell. The electrochemical impedance spectroscopy was carried out within the frequency range 10 kHz–100 mHz by collecting the responses at 10 points per frequency decade.

The study was conducted throughout two stages. The first stage aimed at evaluating complex numbers-PLS in the determination of ferrocyanide in the range 40–140 μmol L^{-1}, and the analyses were performed with both glassy carbon electrode and carbon nanotubes paste electrode. The predictive ability of complex numbers-PLS and PLS applied to real numbers were evaluated, considering that the experimental parameters applied potential and perturbation amplitude could influence the predictive ability. Three potentials were evaluated, which were stablished with square wave voltammetry: i) the oxidation peak potential; ii) the onset potential; iii) a 200–250 mV overpotential. Additionally, three amplitudes were evaluated: 1, 5 and 10 mV.

In the second stage of this work, complex numbers-PLS was evaluated in the determination of substances of major analytical importance: hydroquinone and catechol. Initially, the methodology applied for ferrocyanide was reproduced for hydroquinone, by using only carbon nanotube paste electrode. The best conditions found in such study were used in the direct determination of hydroquinone in spiked tap water samples. Next, complex numbers-PLS was applied in the determination of catechol by EIS in test samples and in fortified tap water samples. Spectra of catechol solutions were obtained by applying the oxidation peak potential and 5 mV amplitude. All measurements were performed in 0.1 mol L^{-1} acetate buffer at pH 4.7 in a quiescent solution without convection.

2.3. Data analysis

PLS regression models were obtained with raw and mean-centered EIS data. Leave-one-out cross-validation was carried out to select the number of factors and validate the models, based on the lowest root mean-squared error of cross-validation (RMSECV) set by the F-test criterion of Haaland and Thomas with $\alpha = 0.25$ [18]. Final evaluation of PLS models was based on the comparison of RMSECV values and number of latent variables. PLS was carried out with a routine written in Matlab® 9.3.0 (The MathWorks™, Inc.).

3. Results and discussion

3.1. Determination of ferrocyanide

The first stage of this work was to evaluate the applicability of complex numbers-PLS in obtaining calibration models with EIS data, as well as the predictive ability of such models regarding the experimental conditions of EIS. In this sense, the determination of potassium ferrocyanide was chosen as model system. Ferrocyanide redox reaction is commonly considered typical outer-sphere electron transfer reaction, inherently associated with fast electron transfer controlled by mass transport of redox active species toward/from the electrode surface [19,20] Impedance spectra were obtained at three different applied potentials (E_a) – at +0.100 V (onset potential), at +0.220 V (peak potential) and at +0.450 V (overpotential) – by employing three amplitudes: 1 mV, 5 mV and 10 mV. Ferrocyanide concentrations were in the range 40–140 μmol L^{-1}. As an example of the results, Fig. 1 presents the Nyquist and Bode plots acquired with a 10-mV amplitude at a bare glassy carbon electrode. Similar results were obtained for different amplitudes.

As can be observed, there is a dependence on the behavior of spectra with applied potential resulting from the different charge-transfer and mass-transfer scenarios. At $E_a = +0.100$ V (Fig. 1a), the Nyquist plot presents straight lines of approximately 45° in the high frequency domain, characteristic of diffusional process on rough surfaces. These straight lines are followed by an ill-defined semicircle in the low frequency domain, representing a high resistance to charge transfer. The changes on spectra as a function of ferrocyanide concentration occurs in the low frequency domain, indicating a decrease of charge-transfer resistance with increasing concentration. These profiles characterize a system controlled by mass transport at high frequencies and by electronic transport at low frequencies. At $E_a = +0.220$ V (Fig. 1c), the impedance spectra reflect pure mass-transfer controlled behavior within the whole frequency range, indicated by straight lines with slopes starting at about 45° and increasing as the concentration decreases. Fig. 1e shows all spectra presenting a non-ideal capacitive impedance response at $E_a = +0.450$ V, which remains unchanged with change of ferrocyanide concentration. These three spectra sets are in accordance with the corresponding assumptions for the impedance response of a faradaic reaction influenced by transport of the reacting species to the electrode [21], i.e., the diffusion impedance is negligibly small when the current is controlled by kinetics (at $E_a = +0.100$ V), the charge-transfer resistance is negligibly small for potentials at which the current is limited by mass transfer (at $E_a = +0.220$ V), and in the case of sufficiently large overpotential ($E_a = +0.450$ V), ferrocyanide concentration at the interface can be considered approximately null when compared to that of the bulk solution. It could be inferred that the charge-transfer resistance reaches a minimum when E_a reaches E^0 and assumes higher values far from E^0, which agrees with the expected for a reversible/quasi-reversible redox reaction such as ferrocyanide oxidation [19].

Regarding Bode plots in Fig. 1, one can see similar magnitude profiles for $E_a = +0.100$ V and $E_a = +0.450$ V, showing resistive (log|Z| slope ≈ 0) to capacitive responses (log|Z| slope ≈ −1) at high to low frequencies. No significant changes on spectra were observed with change of ferrocyanide concentration. The same behavior is observed at high to medium frequencies for $E_a = +0.220$ V. However, at lower frequencies log|Z| slopes tend to decrease towards −0.5 as ferrocyanide concentration increases. Phase-angle (φ) profiles confirm the resistive to capacitive response change behavior ($\varphi \approx 0° \rightarrow 90°$) at $E_a = +0.450$ V, what is typical for electrolyte resistance and double-layer capacitive impedance responses,[19] also showing no changes on spectra as ferrocyanide concentration increases. At $E_a = +0.220$ V, a single time constant is observed with characteristic frequency situated at 10–100 Hz and phase angle nearly constant at medium to low frequencies for the lowest ferrocyanide concentration. As the concentration increases, phase angle at low frequencies tends to decrease to 45°, typical for a semi-infinite diffusion controlled faradaic reaction [19]. A single time constant is also observed for phase-angle profiles at $E_a = +0.100$ V, but with characteristic frequency found at 100–1000 Hz. As observed for $E_a = +0.220$ V, changes on phase angle with ferrocyanide concentrations also occurred at

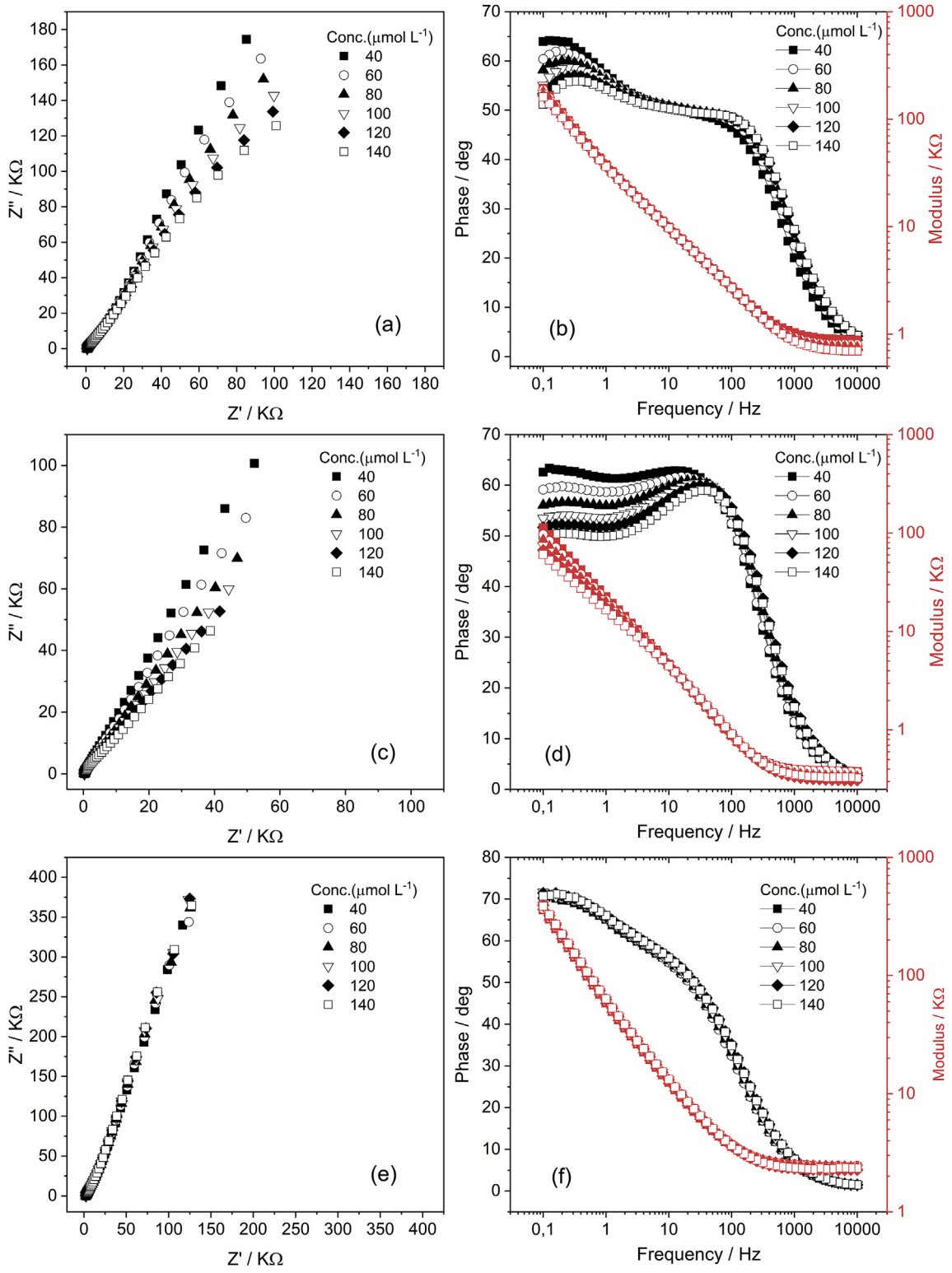


Fig. 1. Nyquist (a, c, e) and Bode (b, d, f) plots for the determination of ferrocyanide at glassy carbon and different applied potentials. E_a = +0.100 V, (a) and (b); E_a = +0.220 V, (c) and (d); E_a = +0.450 V, (e) and (f). Amplitude = 10 mV.

E_a = +0.100 V, but only at frequencies below 1 Hz (Fig. 1b).

The 54 × 51 complex impedance matrix – impedance spectra obtained at six ferrocyanide concentrations, and at nine different combinations of three applied potentials and three amplitudes – was mean-centered and submitted to a principal component analysis (PCA) in order to evaluate if the changes on spectra could be related to visual patterns associated to changes in ferrocyanide concentration. Two components explained 99.6% and 0.3% of the total sum of squares. The high concentration of information in the first principal component was also observed by Geladi and Nelson

[12] and seems to be a characteristic of PCA for complex numbers. Fig. 2 shows the complex plane with the complex scores of the first principal component (real and imaginary parts). It can be seen that the spectra obtained at overpotential cluster together far from the spectra obtained at onset and peak potentials, which are more similar. These spectra are also distributed according to apparently linear or quasi-linear relationships between the real and imaginary parts of the scores related to the increase of ferrocyanide concentration (arrow in Fig. 2), which enable to construct calibration models. Such concentration-dependent behavior is not observed among the spectra obtained at overpotential, in accordance with results shown in Fig. 1.

Complex numbers-PLS was applied to EIS data of the determination of ferrocyanide. Table 1 presents the root mean-squared errors of cross-validation (RMSECV) and the number of latent variables for the models constructed by using raw and mean-centered complex spectra (Z'+jZ", Table 1). We also present the results obtained with PLS for impedance data as real numbers, i.e., by using only the real (Z') or imaginary (Z") parts of the complex impedance, the absolute impedance |Z|, and the phase angle (φ). Bolds values in Table 1 indicate the models with better analytical performance – lower RMSECV and number of latent variables (shown between parentheses).

Some inferences can be made from the observation of the data in Table 1. The first is that the models constructed from spectra obtained at overpotential present, in general, a lower prediction ability, with RMSECV values that can be up to 10 times higher than the models built with spectra obtained at onset and peak potentials. One can see that the model with the best analytical performance was obtained using PLS for complex numbers. Good results were also obtained with PLS applied to impedance spectra as real numbers. Four of the five models were obtained at peak potential. However, a statistical evaluation of the values presented in Table 1 was necessary to verify if the analytical performance of the PLS models are defined by the variation of the EIS experimental conditions, the data presentation and the data pre-treatment. Thus, an analysis of variance (ANOVA) was performed on RMSECV values and on the number of latent variables considering the influence of the following factors: i) applied potential, at three levels: -1 = 0.100 V; 0 = 0.220 V; +1 = 0.450 V; ii) amplitude, at three levels: -1 = 1 mV; 0 = 5 mV; +1 = 10 mV; iii) data presentation, at five levels: -1 = complex numbers (Z'+jZ"); −0.5 = real part of complex impedance (Z'); 0 = imaginary part of complex impedance (Z"); +0.5 = absolute impedance (|Z|); +1 = phase angle (φ); and iv) data pretreatment, at two levels: -1 = none; +1 = mean-centering.

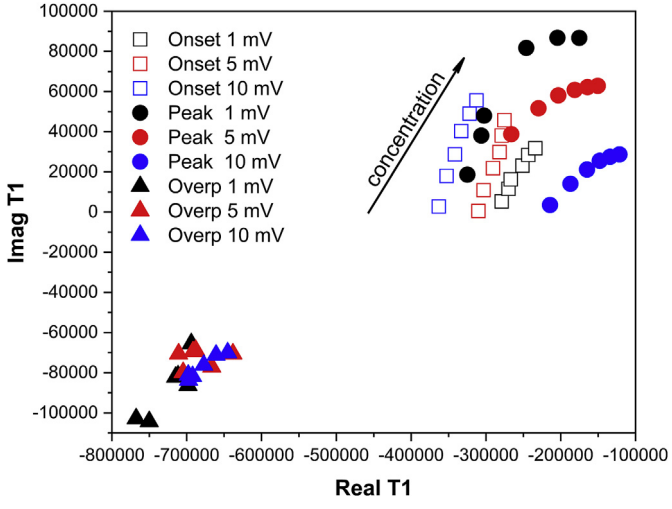


Fig. 2. First principal component real and imaginary scores T1 obtained by complex numbers PCA of ferrocyanide oxidation at various applied potentials and amplitudes.

Fig. 3 shows the Pareto charts obtained from analysis of RMSCEV (Fig. 3a) and number of latent variables (Fig. 3b) data. The Pareto chart is a graphical view of a *t*-statistical test for each effect, where each bar represents the standardized effect – the estimated effect divided by its standard error – that is compared to a vertical line – a *t*-critical value. Any bars which extend beyond the line correspond to effects which are statistically significant. Light gray bars correspond to positive effects, whereas black bars correspond to negative ones. In the figure, each factor was represented by a capital letter. The factors applied potential, amplitude, data presentation, and data pretreatment were represented by the letters A, B, C, and D, respectively, which are also used to identify their main effects. Interaction effects were represented by the combination of the letters representing each variable. For example, an interaction effect between variables applied potential and amplitude is identified by the combination AB. Fig. 3 also shows the change on the response as a function of the change of the factor from its low level to its high level (average response at each level), while all other factors are held constant at their central values (Fig. 3c and d).

Four effects were statistically significant to RMSECV (Fig. 3a): the main effects applied potential, amplitude and data pretreatment, and the quadratic interaction effect AA regarding applied potential. Main effects amplitude and data pretreatment presented a negative sign (antagonistic behavior), while the main effect applied potential and the interaction effect AA presented positive values (synergic behavior). This means that lower RMSECV values were obtained with 10 mV amplitude or by performing mean-centering of spectra (Fig. 3c). Conversely, higher RMSECV values are obtained by applying the overpotential condition – E_a = +0.450 V. AA interaction effect indicates a non-linear increase of RMSECV as a function of the change of the applied potential from the low level to the high level, with a minimum point corresponding to the intermediate level, i.e., RMSECV reaches the lowest value using the spectra obtained at peak potential (Fig. 3c).

Regarding the number of latent variables, three effects were statistically significant, all of them presenting negative values (Fig. 3b): the main effects applied potential and data pretreatment, and the quadratic interaction effect AA. Thus, models with fewer latent variables were built from mean-centered spectra obtained at overpotential. AA interaction effect, in this case, indicates a non-linear decrease of the number of latent variables as the applied potential is changed from the low level to the high level, with a maximum point corresponding to the intermediate level, which means that models based on spectra obtained at peak potential are built with more latent variables than at the other conditions evaluated. However, despite there is a statistically significant increase in the number of latent variables used in the models derived from spectra obtained at peak potential, this increase means in practice the adoption of an additional latent variable (Fig. 3d), which does not represent a significant decrease of robustness.

It is important to point out that the values of latent variables are integers between 1 and 4, which could characterize the data as not taken from a normally distributed population, casting doubt on the validity of ANOVA. Normality tests were performed on latent variables data – Shapiro-Wilk, Kolmogorov-Smirnov, and D'Agostino-K squared – and the results were inconclusive. The first two indicated that normality should be rejected, while the latter did not. On the other hand, based on the central limit theorem, even if hypothesis tests are formally based on the assumption of normality, reliable results can be obtained with non-normal data when the sample is large enough (n > 20). Therefore, the assumption of normality is not critical since the sample is large enough, which is the case for latent variables data (n = 90). Nevertheless, non-

Table 1
RMSECV values of PLS models for ferrocyanide determination at glassy carbon electrode with EIS spectra obtained at three applied potentials and three amplitudes. Number of latent variables between parentheses. Concentration range: 40 a 140 μmol L^{-1} $[Fe(CN)_6]^{4-}$.

Data treatment	Data form	E_a = +0.100 V			E_a = +0.220 V			E_a = +0.450 V		
		1 mV	5 mV	10 mV	1 mV	5 mV	10 mV	1 mV	5 mV	10 mV
Mean-center	Z'+jZ"	6.2 (1)	3.8 (3)	4.9 (1)	9.9 (1)	3.6 (2)	**1.1 (2)**	17 (2)	17 (1)	15 (1)
	Z′	18 (3)	3.4 (1)	11 (2)	20 (2)	25 (1)	**1.7 (2)**	44 (1)	19 (1)	18 (1)
	Z″	5.9 (1)	6.1 (2)	3.7 (1)	9.7 (1)	5.6 (2)	**0.9 (3)**	15 (3)	19 (1)	15 (1)
	|Z|	8.1 (1)	12 (1)	2.9 (1)	11 (1)	6.0 (2)	**2.0 (2)**	13 (4)	18 (1)	15 (1)
	φ	4.0 (1)	**1.3 (3)**	2.9 (2)	13 (2)	2.7 (3)	2.7 (3)	28 (1)	27 (1)	12 (2)
None	Z'+jZ"	17 (3)	3.6 (3)	6.8 (2)	12 (3)	9.4 (3)	4.9 (4)	26 (2)	37 (1)	24 (4)
	Z′	40 (1)	5.0 (3)	18 (2)	14 (2)	5.4 (3)	2.6 (4)	42 (1)	36 (1)	16 (3)
	Z″	6.6 (4)	3.3 (3)	6.4 (2)	14 (3)	5.5 (3)	13 (3)	21 (2)	37 (1)	38 (1)
	Z	13 (3)	9.9 (2)	5.0 (2)	13 (2)	7.0 (3)	6.0 (3)	25 (2)	37 (1)	38 (1)
	φ	5.3 (3)	3.8 (3)	7.8 (3)	7.4 (3)	5.6 (4)	3.5 (4)	40 (1)	27 (2)	14 (3)

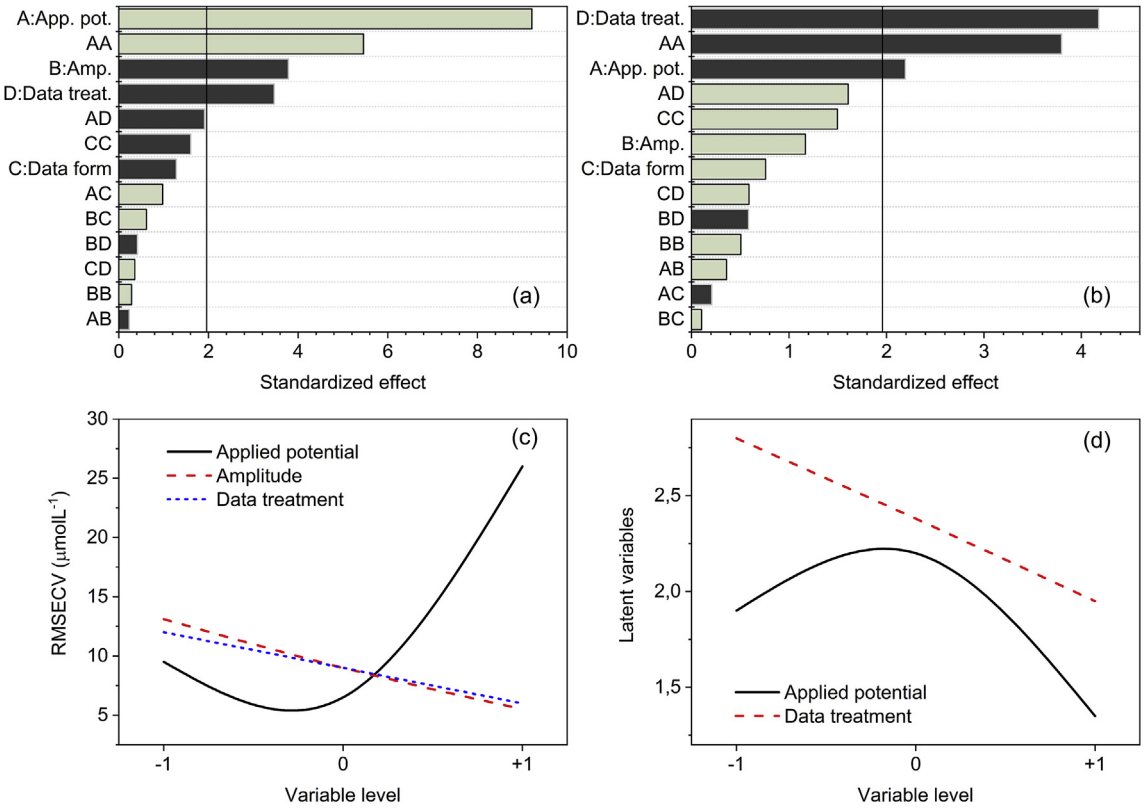


Fig. 3. (a) and (b): Pareto charts of the ANOVA performed on the responses RMSECV and number of latent variables, respectively. Gray bar: positive effect; black bar: negative effect. Where: A: applied potential; B = amplitude; C = data presentation; D = data pretreatment. (c) and (d): variation of responses RMSECV and number of latent variables, respectively, regarding the significant effects.

parametric tests equivalent to ANOVA – Kruskal-Wallis and Mood's median – were performed on latent variables data and the results obtained by both tests were similar to those obtained by ANOVA.

The confirmation of the applied potential as an important variable for RMSECV and number of latent variables indorse the previous inference carried out from the observation of Table 1, that a significant lost in the predictive ability of PLS models is observed with spectra obtained at overpotential. However, an important conclusion that could not be directly inferred from Table 1 is that obtaining spectra at peak potential leads to the construction of PLS models with better predictions, which could be directly associated to the greater change of impedance observed on spectra obtained at this condition following the change on ferrocyanide concentration (Fig. 1). Consequently, one can infer the greater sensitivity to changes in concentration of the impedance controlled by mass transfer. The increase on sensitivity is also the probable reason that explains why spectra obtained at 10 mV amplitude gives models with lower RMSECV values. Regarding data pretreatment, mean-centering provides models with better predictions and fewer latent variables when compared to the models based on unprocessed spectra. The better performance obtained with mean-centering could be also related to the sensitization of pretreated spectra to changes in ferrocyanide concentration, mainly at medium to low frequencies (spectra not shown).

Data presentation did not show a statistically significant effect on RMSECV and number of latent variables. It means that no significant differences were observed on the quality of the models constructed from data presented as complex or real numbers, although the best analytical performance was observed for complex numbers-PLS (Table 1). These results could be related to the

simple nature of the redox behavior of ferrocyanide at low concentrations, which exhibits a rapid charge transfer with no preceding or post-chemical reactions. Nevertheless, the use of complex numbers-PLS is advantageous as it does not require prior research on which impedance representation best relates to the change in analyte concentration.

Table 2 shows the excellent prediction values of the ferrocyanide concentration obtained by leave-one-out cross-validation of the PLS models constructed with spectra obtained at peak potential using a glassy carbon electrode, a 10-mV amplitude and mean-centered spectra. Ferrocyanide concentrations predicted from complex spectra are also expressed as complex numbers, the real and imaginary parts corresponding to the concentration of ferrocyanide and the predicted value of the imaginary part, respectively, where the last one should be 0 j, since the concentration of ferrocyanide is a real number. Thus, the deviation from the expected value of the imaginary part could also be an indication of the PLS model quality. Fig. 4 shows the predicted values of the real and imaginary parts as a function of the number of latent variables included in the PLS model. One can verify that the predicted values of both the real and the imaginary parts approach the reference value as latent variables are added to the model, reaching the overfitting as expected. In the evaluation of the prediction results of the models developed in this study, it was observed that, in general, large deviations in the prediction of the imaginary part followed bad predictions of the real part. On the other hand, when the prediction of the real part was adequate, it was usually accompanied by a good prediction of the imaginary part.

The determination of ferrocyanide by EIS was also evaluated using a carbon nanotube paste electrode (Fig. S1). The spectra were obtained at peak potential considering the good results obtained with glassy carbon electrode. However, the influence of the amplitude was evaluated, and the spectra were obtained at 1 mV, 5 mV and 10 mV amplitudes. The range of concentrations of the calibration curve was 30–150 μmol L^{-1}. Only the Nyquist and Bode plots acquired with a 5-mV amplitude are presented in Fig. S1, once no significant differences were observed for the other amplitudes. The behavior of the spectra obtained on carbon nanotubes is very similar to that observed on glassy carbon. Some differences are observed in the behavior of the phase angle, which presents a smaller change with the concentration at medium to low frequencies. Similar impedance spectra of $Fe(CN)_6^{3-/4-}$ redox couple have already been reported for different electrode surface morphologies such as other carbon surfaces and nanostructured porous electrodes [19,22–24] suggesting that redox reaction of ferrocyanide ions occurs mostly at the outer sites of the porous surface.

Table S1 presents the RMSECV values for the models constructed by using complex spectra and its representations as real numbers. Now, only mean-centered spectra were considered. One can observe RMSECV values higher than those obtained with glassy carbon electrode, which could be related to the lower sensitivity to changes on ferrocyanide concentration observed in the spectra obtained with carbon nanotubes. An ANOVA was also performed on the results of Table S1 considering the influence of the factors amplitude and data presentation at the same levels as performed with glassy carbon. Only the quadratic interaction effect regarding amplitude was statistically significant for RMSECV, with positive sign (results not shown). It indicates that no differences were observed for RMSECV from low to high level of the factor amplitude, but with a significant decrease of RMSECV occurring at the intermediate level, i.e., RMSECV reaches the lowest value using the spectra obtained at 5 mV amplitude. No other effect was significant for the RMSECV or the number of latent variables. Therefore, different data presentations provide similar analytical performance in the same way as observed with glassy carbon, although the PLS model performed directly on complex data with a 5 mV amplitude showed the smallest RMSECV and number of latent variables. Table S2 presents the prediction values of ferrocyanide concentration with PLS models performed on mean-centered impedance spectra obtained with a carbon nanotube paste electrode, at peak potential and 5 mV amplitude. Good predictions were achieved for the real part of the complex number regarding ferrocyanide concentration, also with adequate predictions of the imaginary part.

3.2. Determination of hydroquinone and catechol

Complex numbers-PLS was also evaluated in the determination of hydroquinone and catechol by EIS. Catechol and hydroquinone are 1,2 and 1,4 dihydroxybenze isomers, respectively, and are naturally present in the environment or as a result of anthropic activities. Both isomers are electroactive compounds with well-known oxidation mechanisms, being commonly used as test analytes to validate new electroanalytical methods [25,26]. The same methodology used in the determination of ferrocyanide was employed in the determination of hydroquinone. Impedance spectra were acquired at +0.110 V (onset potential), at +0.225 V (peak potential), and at +0.450 V (overpotential) with a carbon nanotube paste electrode. Also, the same three amplitudes were applied: 1 mV, 5 mV and 10 mV. Hydroquinone concentrations were in the range 40–220 μmol L^{-1}. Fig. 5 presents the Nyquist and Bode plots obtained with a 5 mV amplitude. Similar results were obtained for 1 mV and 10 mV amplitudes.

In the same way as observed for ferrocyanide, there is a dependence on the behavior of spectra with applied potential. Spectra obtained at onset potential and peak potential presented similar features for both Nyquist and Bode plots (Fig. 5a–d). Complex plane impedance spectra show two regions: a semicircle in the high-frequency region corresponding to the charge-transfer process and a linear part in the low-frequency region corresponding to diffusion control. This fast, semi-infinite linear diffusion controlled faradaic reaction behavior is also indicated in the low-frequency region of Bode plots by log |Z| with slopes of about −0.5 and phase angle of about 45°. These features are more pronounced for spectra obtained at peak potential. Also, the greatest changes in spectra as a function of the change of hydroquinone concentration were observed for spectra obtained at that

Table 2
Predictions of ferrocyanide concentration with spectra obtained at peak potential and 10 mV amplitude, at glassy carbon electrode. In parentheses: RMSECV; number of latent variables.

Reference values (μmol L^{-1})	Z' + jZ'' (1.1; 2)	Z' (1.7; 2)	Z'' (0.9; 5)	|Z| (2.0; 2)	φ (2.7; 3)
40	40 + 1j	44	39	38	42
60	58 − 2j	59	58	58	57
80	81 + 2j	80	80	82	81
100	100 − 1j	100	100	101	101
120	121 + 0j	120	120	119	118
140	139 − 1j	140	140	137	145

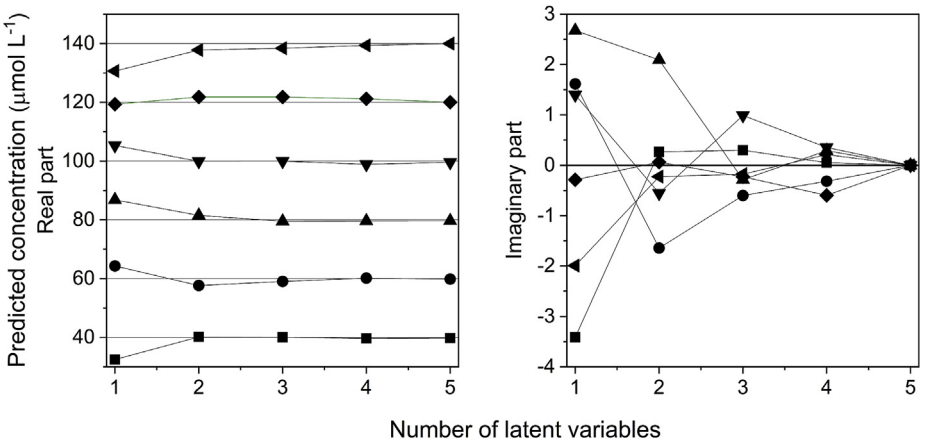


Fig. 4. Behavior of the real and imaginary parts of the prediction results with the number of latent variables included in the complex numbers-PLS model.

condition. Again, spectra obtained at overpotential (Fig. 5e–f) show no significant changes related to the changes on the analyte concentration. Consequently, PLS models constructed from spectra obtained at overpotential presented a lower prediction ability (results not shown). Table 3 presents the RMSECV values and the number of latent variables for the models constructed from spectra obtained at onset potential (E_a = +0.110 V) and peak potential (E_a = +0.225 V), represented as complex and real numbers.

An ANOVA was performed results of Table 3 considering the following factors: i) applied potential, at two levels: -1 = 0.110 V; +1 = 0.225 V; ii) amplitude, at three levels: -1 = 1 mV; 0 = 5 mV; +1 = 10 mV; iii) data presentation, at five levels: -1 = complex numbers (Z'+jZ"); −0.5 = real part of complex impedance (Z'); 0 = imaginary part of complex impedance (Z"); +0.5 = absolute impedance (|Z|); +1 = phase angle (φ); and iv) data pretreatment, at two levels: -1 = none; +1 = mean-centering. Fig. 6 presents the Pareto charts obtained from analysis of RMSCEV (Fig. 6a) and number of latent variables (Fig. 6c). Gray bars correspond to positive effects, and black bars correspond to negative effects.

Three effects were statistically significant to RMSECV (Fig. 6a): the first-order effects applied potential and data form, and the second-order effect AB regarding interaction between applied potential and amplitude. Applied potential presented a negative first-order effect, indicating that lower RMSECV values were observed for PLS models obtained with spectra carried out at peak potential. On the other hand, first-order effect related to data presentation has a positive sign. This means that lower RMSECV values were observed for PLS performed on complex impedance data. This result reinforces the hypothesis of higher efficiency of complex numbers-PLS performed on complex impedance spectra, where all information related to the change in analyte concentration is contained. Data treatment and amplitude showed no significant influence on RMSECV. Thus, if one could consider only the main effects, the ideal experimental condition for the lowest RMSECV would be carry out the spectra at peak potential using any amplitude between 1 mV and 10 mV, followed by complex numbers-PLS carried out in the complex numbers-dataset with no mean-centering of spectra. However, the occurrence of the AB second-order interaction effect must be considered and makes the interpretation of the main effect related to applied potential not so straightforward.

Fig. 6b shows the influence of AB interaction effect on RMSECV, by presenting RMSECV values estimated in different combinations of the studied levels of applied potential and amplitude. In that plot, the three lines present the change on RMSECV from the low level to the high level of the applied potential at each studied level of amplitude. All other factors besides the two involved in the interaction are at their central levels. As one can observe, RMSECV is lower at the higher level of the applied potential when using the 1 mV and 5 mV amplitudes. It recalls the behavior presented by the first-order effect of the applied potential. However, RMSECV increases significantly at the higher level of the applied potential when using an amplitude of 10 mV. Thus, although the first order effect of the amplitude is not significant for RMSECV, its interaction with the applied potential does not guarantee that lower RMSECV values will always be obtained for PLS models constructed from spectra acquired at potential peak. The use of a 10-mV amplitude probably added non-linearity to EIS experiments at some degree that influenced the predictive ability of PLS, which is a linear regression technique. Although the use of a 10-mV amplitude is commonly reported in the literature, there is no reason to expect, given the wide range of electrochemical properties investigated with EIS, that this amplitude may be optimal for most experimental systems.

Fig. 6c presents the Pareto chart obtained for ANOVA performed on the number of latent variables. As one can see, only the first-order effect related to data treatment was statistically significant to define the number of latent variables. The negative sign of this effect indicates that mean-centered impedance spectra produced PLS models with lower number of latent variables. Therefore, considering the evaluation of the results on Table 3, hydroquinone determination with EIS and PLS could provide adequate predictions with spectra acquired at peak potential and applying 5 mV amplitude, and performing PLS on previously mean-centered complex impedance data. The model with best analytical performance on Table 3 (bold values) was obtained with these experimental conditions. Indeed, after the determinations of ferrocyanide and hydroquinone, one can conclude that these are the most suitable experimental conditions for EIS and complex numbers-PLS regression.

The determination of hydroquinone in tap water was evaluated. A new analytical curve in the range 3–25 μmol L^{-1} was obtained at the adequate experimental conditions previously stablished and applied to the direct analysis of eight samples of tap water fortified with hydroquinone. Fig. 7a shows the analytical curve obtained after leave-one-out cross-validation, and the concentrations of hydroquinone predicted for the tap water samples. The method presented a relative error percentage (REP) of 7%, with apparent recoveries ranging from 90% to 115%. Catechol determination was

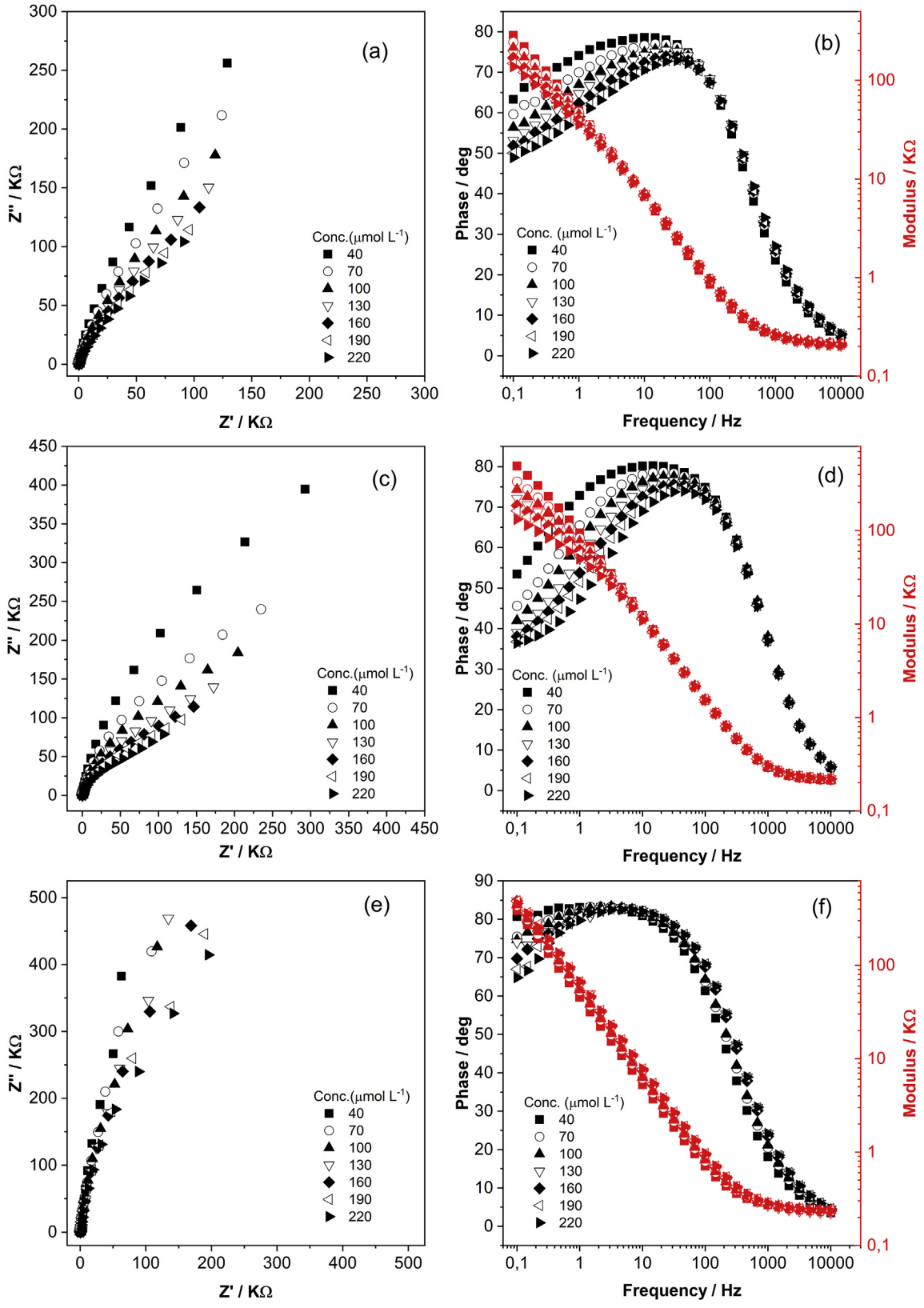


Fig. 5. Nyquist (a, c, e) and Bode (b, d, f) plots for the determination of hydroquinone at carbon nanotube paste electrode and different applied potentials. E_a = +0.110 V, (a) and (b); E_a = +0.225 V, (c) and (d); E_a = +0.450 V, (e) and (f). Amplitude = 5 mV.

Table 3
RMSECV and number of latent variables (in parentheses) of PLS models for hydroquinone determination at carbon nanotube paste electrode. Concentration range: 40–220 μmol L^{-1}.

Data treatment	Data form	E_a = +0.110 V			E_a = +0.225 V		
		1 mV	5 mV	10 mV	1 mV	5 mV	10 mV
Mean-center	**Z' + jZ"**	12 (1)	9.0 (3)	8.1 (2)	4.5 (4)	**2.7 (2)**	16 (3)
	Z′	26 (5)	5.7 (2)	16 (1)	11 (3)	4.6 (2)	13 (3)
	Z″	11 (1)	2.9 (5)	9.4 (2)	7.4 (2)	4.8 (4)	19 (3)
	|Z|	8.3 (1)	11 (5)	9.0 (2)	1.5 (5)	3.6 (2)	22 (1)
	φ	7.4 (3)	18 (1)	8.8 (3)	13 (1)	2.5 (4)	15 (2)
None	**Z' + jZ"**	1.5 (5)	13 (5)	2.1 (4)	3.5 (4)	9.4 (5)	19 (4)
	Z′	9.1 (5)	4.2 (2)	19 (5)	8.3 (5)	11 (5)	10 (4)
	Z″	24 (4)	9.9 (4)	9.8 (4)	8.8 (4)	5.1 (5)	6.6 (5)
	|Z|	12 (5)	31 (4)	9.0 (2)	5.0 (4)	8.3 (4)	1.2 (5)
	φ	54 (2)	26 (3)	8.8 (3)	14 (2)	5.9 (5)	20 (2)

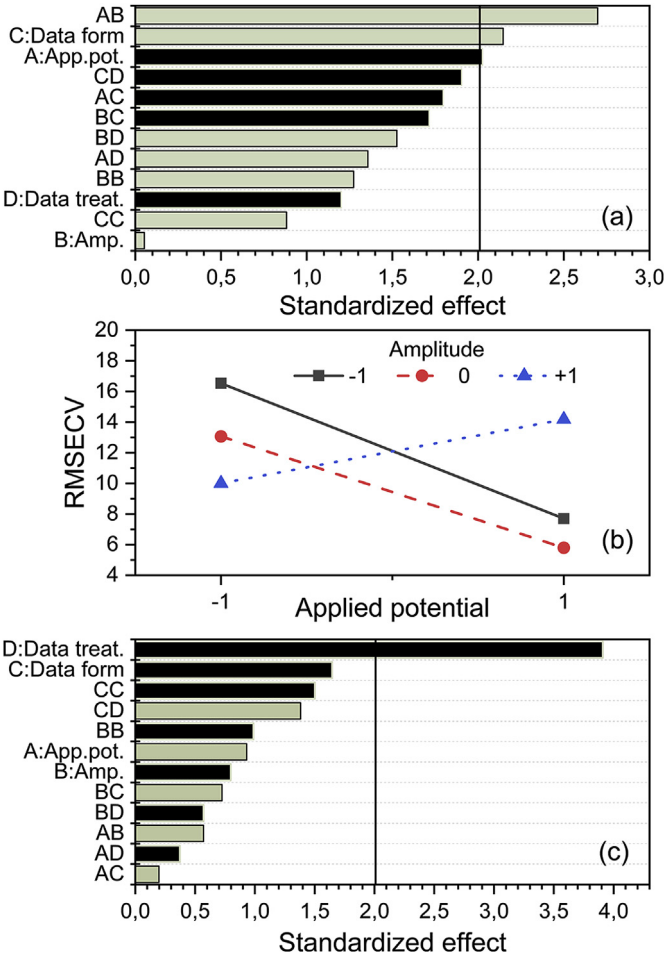


Fig. 6. (a) and (c): Pareto charts of the responses RMSECV and number of latent variables, respectively, for hydroquinone determination. Gray bar: positive effect; black bar: negative effect. Where: A: applied potential; B = amplitude; C = data presentation; D = data pretreatment. (b) Change of RMSECV related to the interaction between applied potential and amplitude.

also performed with EIS and complex numbers-PLS. Impedance spectra was also obtained with a carbon nanotube paste electrode at peak potential and 5 mV amplitude. The analytical curve was in the range 65–175 μmol L^{-1} and five tap water samples fortified with catechol were analyzed. Fig. 7b presents the results for the analytical curve obtained after leave-one-out cross-validation, and for the predicted concentrations of catechol in tap water samples.

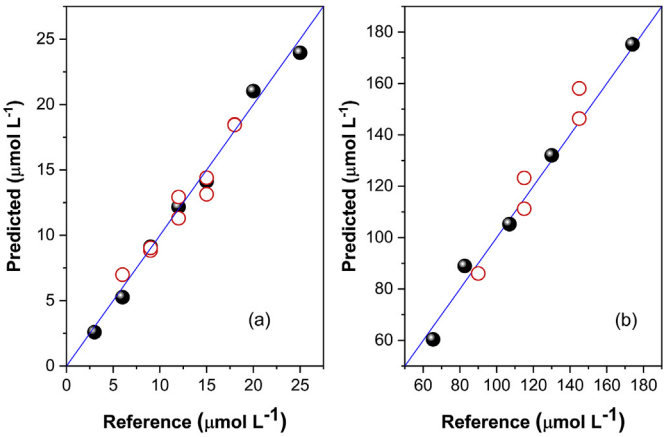


Fig. 7. Analytical curves and prediction results of the determination of hydroquinone (a) and catechol (b) in tap waters. Black: calibration; Red/white: prediction. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The method presented a REP of 6%, with apparent recoveries ranging from 96% to 109%. The predictions were excellent for both determinations, considering the direct analysis of real samples with an analytical curve prepared in deionized water.

4. Conclusions

In this study, it was demonstrated a novel approach for the treatment of electrochemical impedance complex spectra in order to obtain quantitative prediction models based on complex numbers-PLS. This is the first application of PLS directly to electrochemical impedance spectra as vectors of complex numbers in the construction of calibration models for electroanalytical methods. Complex numbers-PLS avoids the reduction of impedance spectra to the real number's domain before performing PLS, procedure usually employed, which can cause loss of important information for calibration. Regarding the influence of the experimental parameters of EIS on PLS modelling, better predictions were obtained with impedance spectra acquired at peak potential. Significant changes on impedance spectra by changing the analyte concentration are not observed at 1 mV amplitude, resulting in PLS models with poor prediction ability. Predictive models were obtained directly from EIS complex data, with no need for reducing complex spectra to real numbers for a typical outer-sphere electron transfer reaction (ferrocyanide) carried out at two different carbon-type electrodes. Since the predicted concentrations are also complex numbers, the deviation from the expected value of the imaginary part (0j) could also be used an indication of the quality of complex numbers-PLS models, although additional investigations are required. Complex numbers–PLS also produced better prediction results when a more complex redox behavior analytes (hydroquinone and catechol) were determined in test samples, including their direct analysis on tap water samples. In summary, the application of PLS to complex spectra is more efficient in the production of calibration models for simple and complex electrochemical systems, because it provides better predictions and because it lacks the need for investigation of the impedance representation as real numbers that provides better results.

Conflicts of interest

Author Dayvison R. Rodrigues declares that he has no conflict of interest.

Author Alejandro C. Olivieri declares that he has no conflict of interest.

Author Wallace D. Fragoso declares that he has no conflict of interest.

Author Sherlan G. Lemos declares that he has no conflict of interest.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are grateful for the financial support of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brazil (CNPq, grant numbers 456569/2014-6, 301662/2015-0, and 310380/2017-0), Agencia Nacional de Promoción Cientifica y Tecnológica - Argentina (ANPCyT, grant number PICT 2016-1122), and Consejo Nacional de Investigaciones Cientificas y Técnicas - Argentina (CONICET).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2019.07.047.

References

[1] M.E. Orazem, N. Pébère, B. Tribollet, J. Electrochem. Soc. 153 (2006) B129–B136.
[2] E.L. Anderson, P. Bühlmann, Anal. Chem. 88 (2016) 9738–9745.
[3] A. Miszczyk, K. Darowicki, Prog. Org. Coat. 77 (2014) 2000–2006.
[4] B. Lindholm-Sethson, P. Geladi, A. Nelson, Anal. Chim. Acta 446 (2001) 121–131.
[5] C.A. Olivati, A. Riul Jr., D.T. Balogh, O.N. Oliveira Jr., M. Ferreira, Bioproc. Biosyst. Eng. 32 (2009) 41–46.
[6] G. Kim, J.-H. Moon, M. Morgan, Anal. Methods 5 (2013) 4074–4080.
[7] C. Conesa, E. García-Breijo, E. Loeff, L. Seguí, P. Fito, N. Laguarda-Miró, Sensors 15 (2015) 22941–22955.
[8] A.V. Sidelnikov, D.M. Bikmeev, D.I. Dubrovskii, F. Kh Kudasheva, V.N. Maistrenko, J. Anal. Chem. 70 (2015) 837–842.
[9] S. Phal, B. Lindholm-Sethson, P. Geladi, A. Shchukarev, S. Tesfalidet, Anal. Chim. Acta 987 (2017) 15–24.
[10] K. Bundy, M. Karlsson, G. Lindbergh, A. Lundqvist, J. Power Sources 72 (1998) 118–125.
[11] Ulrich, H. Petersson, H. Sundgren, F. Bjorefors, C. Krantz-Rülcker, Sens. Actuators B Chem. 127 (2017) 613–618.
[12] P. Geladi, A. Nelson, B. Lindholm-Sethson, Anal. Chim. Acta 595 (2007) 152–159.
[13] A. Bonanni, D. Calvo, M. del Valle, Electroanalysis 20 (2008) 941–948.
[14] H. Wold, Research Paper in Statistics: Festschrift for J. Neyman, Wiley, New York, 1966, pp. 411–444.
[15] K. Pitman, J. Nerut, E. Lust, S. Franssila, M. Raud, T. Kikas, ECS Trans. 77 (2017) 1771–1782.
[16] F. Sundfors, J. Bobacka, J. Electroanal. Chem. 572 (2004) 309–316.
[17] S. Vogt, Q. Su, C. Gutiérrez-Sánchez, G. Nöll, Anal. Chem. 88 (2016) 4383–4390.
[18] D.M. Haaland, E.V. Thomas, Anal. Chem. 60 (1988) 1193–1202.
[19] K.M. Košicek, K. Kvastek, V. Horvat-Radoševic, Electrochim. Acta 195 (2016) 77–84.
[20] A.S. Adekunle, K.I. Ozoemena, Electrochim. Acta 53 (2008) 5774–5782.
[21] M.T.T. Tran, B. Tribollet, V. Vivier, M.E. Orazem, Russ. J. Electrochem. 53 (2017) 932–940.
[22] P. Heiduschka, J. Dittrich, Electrochim, Acta 37 (1992) 2573–2580.
[23] H. Zanin, P.W. May, D.J. Fermin, D. Plana, S.M.C. Vieira, W.I. Milne, E.J. Corat, ACS Appl. Mater. Inerfaces 6 (2014) 990–995.
[24] K. Honda, M. Yoshimura, R. Uchikado, T. Kondo, T.N. Rao, D.A. Tryk, A. Fujishima, M. Watanabe, K. Yasui, H. Masuda, Electrochim. Acta 47 (2002) 4373–4385.
[25] D. Lakshmi, A. Bossi, M.J. Whitcombe, I. Chianella, S.A. Fowler, S. Subrahmanyam, E.V. Piletska, S.A. Piletsky, Anal. Chem. 81 (2009) 3576–3584.
[26] S.H. DuVall, R.L. McCreery, Anal. Chem. 71 (1999) 4594–4602.