

UNIVERSIDADE FEDERAL DA PARAÍBA  
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA  
DEPARTAMENTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

**TESE DE DOUTORADO**

**ESTUDO SOBRE O EMPREGO DE QUIMIOMETRIA EM  
ESPECTROSCOPIA DE IMPEDÂNCIA ELETROQUÍMICA**

**DAYVISON RIBEIRO RODRIGUES**

*SAPIENTIA AEDIFICAT*

João Pessoa – PB – Brasil  
Agosto/2019

## **TESE DE DOUTORADO**

# **ESTUDO SOBRE O EMPREGO DE QUIMIOMETRIA EM ESPECTROSCOPIA DE IMPEDÂNCIA ELETROQUÍMICA**

**Dayvison Ribeiro Rodrigues**

Tese de doutorado submetida ao  
Programa de Pós-Graduação como parte  
dos requisitos para obtenção do título de  
Doutor em Química pela Universidade  
Federal da Paraíba

**Orientador: Dr. Sherlan Guimarães Lemos**

**Co-orientador: Dr. Wallace Duarte Fragoso**

**\*Bolsista Capes**

**João Pessoa – PB – Brasil  
Agosto/2019**

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

R696e Rodrigues, Dayvison Ribeiro.

Estudo sobre o emprego de quimiometria em  
espectroscopia de impedância eletroquímica / Dayvison  
Ribeiro Rodrigues. - João Pessoa, 2019.

137 f. : il.

Orientação: Sherlan Guimarães Lemos.

Coorientação: Wallace Duarte Fragoso.

Tese (Doutorado) - UFPB/CCEN.

1. espectroscopia de impedância eletroquímica. 2.  
números complexos. 3. café. 4. MLR. 5. PLS. 6. PLS-DA.  
7. SIMCA. I. Lemos, Sherlan Guimarães. II. Fragoso,  
Wallace Duarte. III. Título.

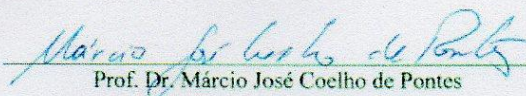
UFPB/BC

## ESTUDO SOBRE O EMPREGO DE QUIMIOMETRIA EM ESPECTROSCOPIA DE IMPEDÂNCIA ELETROQUÍMICA

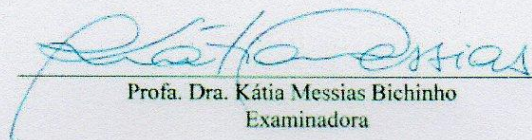
Tese de Doutorado apresentada pelo aluno Dayvison Ribeiro Rodrigues e aprovada pela banca examinadora em 28 de agosto de 2019.



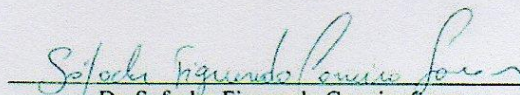
Prof. Dr. Sherlan Guimarães Lemos  
Orientador/Presidente



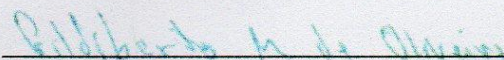
Prof. Dr. Márcio José Coelho de Pontes  
Examinador



Profa. Dra. Kátia Messias Bichinho  
Examinadora



Dr. Sofacless Figueredo Carreiro Soares  
Examinador



Prof. Dr. Gildiberto Mendonça de Oliveira  
Examinador

**“Se não houver frutos, valeu a beleza das flores. Se não houver flores, valeu a sombra das folhas. Se não houver folhas, valeu a intenção da semente.”**

(Autor desconhecido)

Dedico este trabalho à Gláucia Ribeiro, minha mãe (*in memoriam*), que nos deixou precocemente durante o curso desse trabalho. A Francisco Rodrigues, meu pai. A meus irmãos, David e Sara. Também dedico à Luana, minha esposa querida, e a Davi, meu filho amado.

## AGRADECIMENTOS

- ✧ A **Deus** que me deu forças para superar as dificuldades que surgiram durante o desenvolvimento desse trabalho.
- ✧ À minha família por sempre estar comigo em todos os momentos, especialmente aos meus pais Francisco e Gláucia (*in memoriam*) por investir na minha formação. A meu irmão David pelos conselhos, amizade e amor. Às minhas cunhadas Priscylla e Luciana pelas amizade. À minha amada sobrinha Ana Beatriz pelas brincadeiras com o titio.
- ✧ Ao **Prof. Dr. Sherlan Guimarães Lemos** por me acolher como aluno de iniciação científica em 2010, e me orientar durante o desenvolvimento do mestrado e doutorado. Agradecer a imensurável influência acadêmica e pessoal, e principalmente pela amizade, liberdade e suporte que concede a seus orientandos.
- ✧ Ao **Prof. Dr. Wallace Duarte Fragoso** pela coorientação durante o desenvolvimento desse trabalho e pela amizade ao longo dessa caminhada.
- ✧ Aos Professores **Dr. Edvan Cirino da Silva** e **Dr. Márcio José Coelho de Pontes** pelas contribuições durante a banca de qualificação.
- ✧ Aos examinadores externos, **Dr. Gildiberto Mendonça de Oliveira** e **Dr. Sófacles Figueredo Carreiro Soares**, e examinadores internos, **Dra. Kátia Messias Bichinho** e **Dr. Márcio José Coelho de Pontes**, pelas contribuições valiosas durante a banca de defesa.
- ✧ Aos amigos do IFPB, especialmente **Francisco Antônio Borges**, **Prof. Dr. Sabiniano Araújo Rodrigues**, **Prof. Dr. Inakã Barreto**, **Jonathan Silva**, **Renata Patriota**, **Tharciana Mirella**, **Ailton Araújo** pela amizade, incentivo, ensinamentos, confiança e experiências compartilhadas.
- ✧ Aos amigos **Rômulo Sampaio**, **Carol Mascarenhas**, **Caio**, **Erika**, **Lucas**, **Ângela**, **Taís Moraes**, **Allesya**, **Cleilson**, **Suellen**, **Israel**, **José Ferreira**, **Clarissa**, **Gabi**, **Levi**, **Edvaldo**, **Josiele**, **Rafael**, **Ivson**, **Eduardo Antônio**, **Jefferson**, **Laís**, **Sófacles**, **Arnayra**, **Stéfany**, **Marcelo Batista** pela amizade.
- ✧ À **CAPES** pela bolsa concedida.

## SUMÁRIO

|                                                                                                      |                                                |
|------------------------------------------------------------------------------------------------------|------------------------------------------------|
| RESUMO.....                                                                                          | i                                              |
| ABSTRACT .....                                                                                       | ii                                             |
| LISTA DE FIGURAS .....                                                                               | iii                                            |
| LISTA DE TABELAS.....                                                                                | vii                                            |
| LISTA DE ABREVIATURAS E SIGLAS .....                                                                 | ix                                             |
| CAPÍTULO I – APRESENTAÇÃO.....                                                                       | 10                                             |
| 1.1 Introdução.....                                                                                  | 11                                             |
| 1.2 Objetivos.....                                                                                   | 13                                             |
| 1.2.1 Objetivos Gerais .....                                                                         | 13                                             |
| CAPÍTULO II – CONSTRUÇÃO DE MODELOS QUANTITATIVOS EMPREGANDO<br>REGRESSÃO LINEAR MÚLTIPLA (MLR)..... | 14                                             |
| 2.1 Introdução.....                                                                                  | 15                                             |
| 2.2 Objetivos.....                                                                                   | 16                                             |
| 2.2.1 Objetivos Específicos.....                                                                     | 16                                             |
| 2.3 Fundamentação Teórica.....                                                                       | 17                                             |
| 2.3.1 Espectroscopia de Impedância Eletroquímica.....                                                | 17                                             |
| 2.3.2 Modelagem por Circuito Elétrico Equivalente.....                                               | 19                                             |
| 2.3.3 Resistor.....                                                                                  | 19                                             |
| 2.3.4 Capacitor .....                                                                                | 20                                             |
| 2.3.5 Elemento de Fase Constante (CPE) .....                                                         | 22                                             |
| 2.3.6 Circuito de Randles .....                                                                      | 23                                             |
| 2.3.7 Validação das medidas .....                                                                    | 24                                             |
| 2.3.8 Construção de Modelos Quantitativos .....                                                      | 25                                             |
| 2.3.9 Calibração Univariada - Regressão Linear Simples.....                                          | 26                                             |
| 2.3.10 Trabalhos que Aplicam Regressão Linear Simples aos Dados de EIE .....                         | 27                                             |
| 2.3.11 Regressão Linear Múltipla (MLR).....                                                          | 29                                             |
| 2.3.12 Trabalhos que Aplicam Regressão Linear Múltipla aos Dados de EIE .....                        | <b>Erro!</b><br><b>Indicador não definido.</b> |
| 2.4 Metodologia.....                                                                                 | 31                                             |
| 2.4.1 Reagentes e Amostras .....                                                                     | 31                                             |
| 2.4.2 Medidas de Impedância.....                                                                     | 31                                             |
| 2.4.3 Construção de modelos de Regressão Linear Múltipla.....                                        | 32                                             |
| 2.5 Resultados e Discussão.....                                                                      | 34                                             |
| 2.5.1 Quantificação de ferrocianeto .....                                                            | 34                                             |
| 2.5.2 Quantificação de catecol e hidroquinona .....                                                  | 39                                             |

|                                                                                                              |    |
|--------------------------------------------------------------------------------------------------------------|----|
| 2.6 Conclusões do Capítulo .....                                                                             | 47 |
| CAPÍTULO III – CONSTRUÇÃO DE MODELOS QUANTITATIVOS EMPREGANDO REGRESSÃO POR MÍNIMOS QUADRADOS PARCIAIS ..... | 48 |
| 3.1 Introdução .....                                                                                         | 49 |
| 3.2 Objetivos .....                                                                                          | 51 |
| 3.2.1 Objetivos Específicos .....                                                                            | 51 |
| 3.3 Fundamentação Teórica .....                                                                              | 52 |
| 3.3.1 Regressão por Mínimos Quadrados Parciais (PLS) .....                                                   | 52 |
| 3.3.2 Trabalhos que aplicam PLS a dados de EIE .....                                                         | 53 |
| 3.4 Metodologia .....                                                                                        | 55 |
| 3.4.1 Construção de modelos usando a Regressão por Mínimos Quadrados Parciais .....                          | 55 |
| 3.5 Resultados e Discussão .....                                                                             | 56 |
| 3.5.1 Determinação de ferrocianeto empregando o PLS para números complexos .....                             | 56 |
| 3.5.2 Determinação de catecol e hidroquinona usando o PLS para número complexo ....                          | 70 |
| 3.6 Conclusões do Capítulo .....                                                                             | 76 |
| CAPÍTULO IV – CONSTRUÇÃO DE MODELOS PARA CLASSIFICAÇÃO DE AMOSTRAS DE CAFÉS .....                            | 77 |
| 4.1 Introdução .....                                                                                         | 78 |
| 4.2 Objetivos .....                                                                                          | 79 |
| 4.2.1 Objetivos Específicos .....                                                                            | 79 |
| 4.3 Fundamentação Teórica .....                                                                              | 80 |
| 4.3.1 Modelagem independente e Flexível por Analogia de Classes (SIMCA) .....                                | 80 |
| 4.3.2 Análise Discriminantes por Mínimos Quadrados Parciais .....                                            | 80 |
| 4.3.3 Trabalhos que utilizam dados de EIE no reconhecimento de padrões .....                                 | 81 |
| 4.4 METODOLOGIA .....                                                                                        | 83 |
| 4.4.1 Soluções e Amostras .....                                                                              | 83 |
| 4.4.2 Medidas de Impedância .....                                                                            | 83 |
| 4.4.3 Procedimento Analítico .....                                                                           | 84 |
| 4.4.4 Cálculo dos parâmetros dos modelos .....                                                               | 84 |
| 4.5 Resultados e Discussão .....                                                                             | 86 |
| 4.5.1 Construção dos modelos PLS-DA e SIMCA .....                                                            | 86 |
| 4.6 Conclusões do Capítulo .....                                                                             | 91 |
| CONCLUSÕES Gerais .....                                                                                      | 92 |
| Conclusões Gerais .....                                                                                      | 93 |
| Propostas Futuras .....                                                                                      | 93 |
| REFÊRENCIAS BIBLIOGRÁFICAS .....                                                                             | 95 |

|                                                        |     |
|--------------------------------------------------------|-----|
| ANEXOS .....                                           | 105 |
| Anexo I – Gráficos referentes aos modelos PLS-DA ..... | 106 |
| Anexo II – Gráficos referentes aos modelos SIMCA ..... | 112 |

## RESUMO

A aplicação de um potencial senoidal a um sistema eletroquímico provoca o surgimento de uma corrente que também é senoidal, e cujas características dependem dos processos impedimétricos que ocorrem na interface eletrodo-solução e da frequência da perturbação aplicada. Esse é o princípio da Espectroscopia de Impedância Eletroquímica. Os modelos quantitativos baseados nessa técnica são geralmente construídos relacionando-se a variação da concentração do analito aos valores de um único elemento elétrico de um circuito equivalente usado para modelar o comportamento dos espectros. A primeira etapa do trabalho propõe uma estratégia para otimizar a construção desses modelos quantitativos ao correlacionar mais de um elemento elétrico do circuito equivalente a variação da concentração do analito. No trabalho foram construídos modelos multivariados para a quantificação de ferrocianeto, catecol e hidroquinona empregando a regressão linear múltipla. O resultado mais expressivo foi observado com o modelo multivariado da hidroquinona que apresentou Erro Padrão Médio da previsão (RMSEP) de  $0,79 \mu\text{mol L}^{-1}$  contra  $2,51 \mu\text{mol L}^{-1}$  alcançado com o modelo univariado, ambos na quantificação de hidroquinona em amostras de água de torneira fortificada. Uma metodologia alternativa para construção de modelos quantitativos, e que não necessita recorrer aos circuitos equivalentes, foi proposta na segunda etapa desse trabalho. A estratégia foi construir modelos de Regressão Parcial por Mínimos Quadrados (PLS) empregando a impedância expressa como número complexo ( $Z' + jZ''$ ). Modelos quantitativos para o ferrocianeto, catecol e hidroquinona foram construídos e os resultados obtidos por esses modelos foram comparados aos fornecidos pelos modelos PLS construídos com a parte real da impedância ( $Z'$ ), a parte imaginária ( $Z''$ ), o módulo da impedância ( $|Z|$ ) e o ângulo de fase ( $\varphi$ ). A terceira etapa do trabalho foi a construção de modelos para classificação de amostras de cafés puros e adulterados com cascas e paus, empregando dados de impedância expressos como números complexos. A Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) e a Modelagem Independente e Flexível por Analogia de Classes (SIMCA) foram as técnicas empregadas. Verificou-se nos casos avaliados que a construção de modelos quantitativos e qualitativos com a impedância expressa como número complexo conduziu a resultados similares ou superiores aos modelos construídos com apenas uma das informações relacionadas a impedância ( $Z'$ ,  $Z''$ ,  $|Z|$  e  $\varphi$ ). Isso é uma vantagem, pois não é necessário avaliar todas as matrizes, e possivelmente, o uso dos dados complexos conduzirá aos melhores resultados.

Palavras-chave: espectroscopia de impedância eletroquímica, números complexos, café, MLR, PLS, PLS-DA e SIMCA.

## ABSTRACT

The application of a sinusoidal potential to an electrochemical system causes the rise of a current that is also sinusoidal, whose characteristics depend on the impedimetric processes that occur at the electrode-solution interface and of the applied signal frequency. This is the principle of Electrochemical Impedance Spectroscopy. Quantitative models based on this technique are generally constructed by relating the change in analyte concentration to the values of a single electrical element of an equivalent circuit used to model the behavior of spectra. The first step of the work proposed a strategy to optimize the construction of these quantitative models by correlating more than one electrical element of the circuit equivalent to the variation of analyte concentration. Multivariate models for ferrocyanide, catechol and hydroquinone quantification using multiple linear regression were built. The most significant result was observed with the multivariate hydroquinone model which presented an RMSEP of  $0.79 \mu\text{mol L}^{-1}$  against  $2.51 \mu\text{mol L}^{-1}$  achieved with the univariate model, both in the quantification of hydroquinone in spiked tap water samples. An alternative methodology for the construction of quantitative models, which does not need to use equivalent circuits, was proposed in the second stage of this work. The strategy was to construct partial least square models (PLS) using the impedance expressed as complex number ( $Z' + jZ''$ ). Quantitative models for ferrocyanide, catechol and hydroquinone were constructed and the results obtained by these models were compared to those provided by the PLS models constructed with the real impedance part ( $Z'$ ), imaginary part ( $Z''$ ), absolute impedance ( $|Z|$ ) and the phase angle ( $\phi$ ). The third stage of the work was the construction of models to classify samples of pure and adulterated coffees with shells and sticks, using impedance data expressed as complex numbers. The techniques employed were PLS-DA and SIMCA. It was found in all studied cases that the construction of quantitative and qualitative models with the impedance expressed as a complex number led to similar or better results than the models constructed with only one of the impedance related information ( $Z'$ ,  $Z''$ ,  $|Z|$  and  $\phi$ ). This is an advantage as it is not necessary to evaluate all arrays as the use of complex data will lead to better or similar results.

Keywords: Electrochemical impedance spectroscopy, complex numbers, coffee, MLR, PLS, PLS-DA and SIMCA.

## LISTA DE FIGURAS

|                                                                                                                                                                                                                                                                                                                                                                                           |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 1</b> - Plano complexo de Argand-Gauss.....                                                                                                                                                                                                                                                                                                                                     | 18 |
| <b>Figura 2</b> - (a) Comportamento da corrente elétrica resultante da aplicação de uma tensão senoidal a um resistor com resistência de 500 $\Omega$ . (b) Variação do módulo da impedância e ângulo de fase em função da aplicação de frequências diferentes (gráfico de Bode). (c) Comportamento da impedância real ( $Z'$ ) e impedância imaginária (diagrama de Nyquist). ...        | 20 |
| <b>Figura 3</b> - (a) Comportamento da corrente elétrica resultante da aplicação de uma tensão senoidal a um capacitor com capacitância de 1,0 $\mu\text{F}$ . (b) Variação do módulo da impedância e ângulo de fase em função da aplicação de frequências diferentes (gráfico de Bode). (c) Comportamento da impedância real ( $Z'$ ) e impedância imaginária (diagrama de Nyquist). ... | 21 |
| <b>Figura 4</b> – Diagramas de Nyquist referentes a processos faradaicos controlado por cinética de transferência da carga e controlados por processo misto de cinética de transferência de carga e transferência de massa (difusão) Adaptado de Carvalho (2006). Valores dos elementos: $R_s = 250 \Omega$ ; $R_{ct} = 2 \text{ k}\Omega$ ; $C_{dl} = 1 \mu\text{F}$ ; .....             | 24 |
| <b>Figura 5</b> - Diagramas de Nyquist referentes à aquisição dos espectros de impedância em diferentes concentrações de ferrocianeto. $E_{dc} = 0,22 \text{ V vs Hg}_2\text{Cl}_2 (\text{sat})$ ; amplitude = 5 mV. Inserção: circuito equivalente ajustado aos espectros. ....                                                                                                          | 34 |
| <b>Figura 6</b> – Processo de seleção entre os circuitos usados no ajuste dos dados do ferrocianeto exemplificado no espectro referente à concentração de 200 $\mu\text{mol L}^{-1}$ . ....                                                                                                                                                                                               | 35 |
| <b>Figura 7</b> - Mudanças nos valores normalizados dos elementos elétricos do circuito equivalente em função da variação da concentração de ferrocianeto. ....                                                                                                                                                                                                                           | 36 |
| <b>Figura 8</b> - Diagramas de Nyquist referentes a aquisição de espectros de impedância em concentrações diferentes de (a) catecol e (b) hidroquinona. ....                                                                                                                                                                                                                              | 40 |
| <b>Figura 9</b> - Mudanças nos valores normalizados dos elementos elétricos do circuito equivalente em função da variação da concentração de (a) catecol e (b) hidroquinona. ....                                                                                                                                                                                                         | 41 |
| <b>Figura 10</b> – Concentração das amostras de catecol previstas pelos modelos II e III com destaques para os resíduos deixados pela modelagem. ....                                                                                                                                                                                                                                     | 43 |
| <b>Figura 11</b> - Concentração das amostras de água de torneira fortificadas com catecol previstas pelos modelos II e III. ....                                                                                                                                                                                                                                                          | 44 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 12</b> - Previsão das amostras de calibração e amostras de água de torneira fortificadas com hidroquinona usando o modelo II. ....                                                                                                                                                                                                                                                                                                                                                                                                     | 45 |
| <b>Figura 13</b> - Previsão das amostras de calibração e amostras de água de torneira fortificadas com hidroquinona usando o modelo construído com os elementos W e $CPE_1$ .....                                                                                                                                                                                                                                                                                                                                                                | 46 |
| <b>Figura 14</b> - Diagrama de Nyquist (a,c,e) e bode (b,d,f) referentes aos espectros registrados na determinação de ferrocianeto em eletrodo de carbono vítreo aplicando potenciais diferentes: $E_a = 0,100$ V (a) e (b), $E_a = 0,220$ V (c) e (d), $E_a = 0,450$ V, (e) e (f).....                                                                                                                                                                                                                                                          | 57 |
| <b>Figura 15</b> - Escores da primeira componente principal real e imaginária (T1) realizada nos dados complexos obtidos durante a quantificação de ferrocianeto aplicando potenciais e amplitudes diferentes.....                                                                                                                                                                                                                                                                                                                               | 59 |
| <b>Figura 16</b> - Escores da primeira (T1) e segunda (T2) componente principal realizada na matriz de dados $Z'$ (a), $-Z''$ (b), $ Z $ (c) e $\phi$ (d) obtidas com os dados da quantificação de ferrocianeto aplicando potenciais e amplitudes diferentes. ....                                                                                                                                                                                                                                                                               | 60 |
| <b>Figura 17</b> - Cartas de Pareto referente à ANOVA realizada nas respostas RMSECV (a) e número de variáveis latentes (b), onde: A = potencial aplicado; B = amplitude; C = apresentação dos dados; D = pré-tratamento dos dados. Variação das respostas RMSECV (c) e número de variáveis latentes (d) em relação aos efeitos significativos. ....                                                                                                                                                                                             | 64 |
| <b>Figura 18</b> - Comportamento das partes real e imaginária dos valores previstos com o modelo complexo em função do número de variáveis latentes. ....                                                                                                                                                                                                                                                                                                                                                                                        | 67 |
| <b>Figura 19</b> – Comportamento dos pesos da primeira (a) e segunda (b) variáveis latentes usadas nos modelos PLS construídos com: a impedância expressa como número complexo ( $Z' + jZ''$ ), parte real da impedância ( $Z'$ ) e a parte imaginária da impedância ( $Z''$ ). ....                                                                                                                                                                                                                                                             | 68 |
| <b>Figura 20</b> - (a) Resíduo relativo resultante da diferença entre a impedância real medida ( $Z'_{med}$ ) e a impedância real estimada em cada frequência a partir de $Z''_{med}$ usando a transformação de Kronig-Kramers. (b) Resíduo relativo deixado pela diferença entre a impedância imaginária medida ( $Z''_{med}$ ) e a impedância imaginária estimada a partir de $Z'$ (medido) usando a transformação de Kronig-Kramers. A impedância registrada sob aplicação do potencial de pico ( $E_a = 0,220$ V) e amplitude de 10 mV. .... | 69 |
| <b>Figura 21</b> - Gráficos de nyquist (a,c,e) e bode (b,d,f) referente aos espectros registrados na determinação de hidroquinona em eletrodo de pasta de nanotubo de carbono aplicando potenciais diferentes: $E_a = 0,110$ V (a) e (b), $E_a = 0,225$ V (c) e (d), $E_a = 0,450$ V, (e) e (f)...                                                                                                                                                                                                                                               | 71 |

|                                                                                                                                                                                                                                                                                                                                                               |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Figura 22</b> - Cartas de Pareto referente a ANOVA realizada nas respostas RMSECV (a) e número de variáveis latentes (c) na determinação de hidroquinona, onde: A = potencial aplicado; B = amplitude; C = apresentação dos dados; D = pré-tratamento dos dados. Variação do RMSECV relacionado a interação entre potencial aplicado e amplitude (b). .... | 74  |
| <b>Figura 23</b> – Concentrações das amostras de calibração previstos por validação cruzada (preto) e resultados das previsões (vermelho) de hidroquinona (a) e catecol (b) em água da torneira.                                                                                                                                                              | 75  |
| <b>Figura 24</b> - (a) Diagrama de Nyquist referente aos espectros das amostras de cafés adulterados e puros. (b) Escores da primeira componente principal real e imaginária (T1) realizada nos dados complexos das amostras de cafés, puros e adulterados com cascas e paus, centrados na média. ....                                                        | 87  |
| <b>Figura 25</b> - <i>Loading</i> da componente real dos três primeiros fatores do modelo PLS-DA construído com os dados de impedância expressos como números complexos ( $Z' + jZ''$ )...                                                                                                                                                                    | 106 |
| Figura 26 - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da impedância expressos como números complexos número complexo ( $Z' + jZ''$ ). ....                                                                                                                                                                         | 106 |
| <b>Figura 27</b> - <i>Loading</i> dos quatro primeiros fatores do modelo PLS-DA construído com os dados da parte real da impedância ( $Z'$ ).....                                                                                                                                                                                                             | 107 |
| <b>Figura 28</b> - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da parte real da impedância ( $Z'$ ).....                                                                                                                                                                                                             | 108 |
| <b>Figura 29</b> - <i>Loading</i> dos quatro primeiros fatores do modelo PLS-DA construído com os dados da parte imaginária da impedância ( $Z''$ ). ....                                                                                                                                                                                                     | 108 |
| <b>Figura 30</b> - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da parte imaginária da impedância ( $Z''$ ). ....                                                                                                                                                                                                     | 109 |
| <b>Figura 31</b> - <i>Loading</i> dos três primeiros fatores do modelo PLS-DA construído com os dados do módulo da impedância ( $ Z $ ). ....                                                                                                                                                                                                                 | 109 |
| <b>Figura 32</b> - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados do módulo da impedância ( $ Z $ ). ....                                                                                                                                                                                                               | 110 |
| <b>Figura 33</b> - <i>Loading</i> dos quatro primeiros fatores do modelo PLS-DA construído com os dados da fase da impedância ( $\varphi$ ). ....                                                                                                                                                                                                             | 110 |

|                                                                                                                                                                                   |     |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| <b>Figura 34</b> - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da fase da impedância ( $\varphi$ ).....                                  | 111 |
| <b>Figura 35</b> - <i>Loading</i> das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados da impedância expressa como número complexo ( $Z' + jZ''$ )..... | 112 |
| <b>Figura 36</b> - <i>Loading</i> das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados da parte real da impedância ( $Z'$ ).....                        | 112 |
| <b>Figura 37</b> - <i>Loading</i> das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados da parte imaginária da impedância ( $Z''$ ). ....                | 113 |
| <b>Figura 38</b> - <i>Loading</i> das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados do módulo da impedância ( $ Z $ ). ....                          | 113 |
| <b>Figura 39</b> - <i>Loading</i> das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados do ângulo de fase da impedância ( $\varphi$ ).....               | 114 |

## LISTA DE TABELAS

|                                                                                                                                                                                                                                                                                                                                                                                                          |    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Tabela 1</b> - Matriz com os valores de correlação entre os quatro elementos do circuito equivalente ajustado para os dados do ferrocianeto. ....                                                                                                                                                                                                                                                     | 37 |
| <b>Tabela 2</b> – Modelos de regressão usados para quantificação de ferrocianeto (* $\gamma = 95\%$ ). ....                                                                                                                                                                                                                                                                                              | 38 |
| <b>Tabela 3</b> - Matrizes de correlações dos elementos dos circuitos equivalentes ajustados aos espectros do catecol e hidroquinona. ....                                                                                                                                                                                                                                                               | 42 |
| <b>Tabela 4</b> - Coeficientes de regressão dos modelos construídos para quantificar catecol. (* $\gamma = 95\%$ ). ....                                                                                                                                                                                                                                                                                 | 42 |
| <b>Tabela 5</b> – Coeficientes de regressão dos modelos construídos para quantificar hidroquinona .....                                                                                                                                                                                                                                                                                                  | 45 |
| <b>Tabela 6</b> - Valores dos RMSECV obtidos com os modelos PLS construídos com dados de impedância gerados na determinação de ferrocianeto aplicando potenciais e amplitudes diferentes ao eletrodo de carbono vítreo. A quantidade de variáveis latentes selecionadas encontra-se entre parênteses. A faixa de concentração de ferrocianeto empregada foi de 40 a 140 $\mu\text{mol L}^{-1}$ . ....    | 62 |
| <b>Tabela 7</b> - Valores das previsões das concentrações de ferrocianeto realizados com os dados de impedância registrados durante a aplicação do potencial de pico e amplitude de 10 mV sobre eletrodo de carbono vítreo. Os valores dos RMSECV e número de variáveis latentes usadas no modelo encontram-se entre parênteses. ....                                                                    | 66 |
| <b>Tabela 8</b> - Valores dos RMSECV obtidos com os modelos PLS construídos com dados de impedância gerados na determinação de hidroquinona aplicando potenciais e amplitudes diferentes ao eletrodo de carbono vítreo. A quantidade de variáveis latentes selecionadas encontra-se entre parênteses. A faixa de concentração empregada para a hidroquinona foi de 3 a 20 $\mu\text{mol L}^{-1}$ . ....  | 72 |
| <b>Tabela 9</b> – Melhores resultados da classificação de cafés para os conjuntos de validação e teste obtidos com os modelos PLS-DA construídos, com os dados da: impedância expressa como número complexo ( $Z' + jZ''$ ), parte real da impedância ( $Z'$ ), parte imaginária da impedância, módulo da impedância ( $ Z $ ) e ângulo de fase da impedância ( $\varphi$ ). Os números entre parênteses |    |

referem-se à quantidade de fatores usados na modelagem. O índice 1 representa a classe dos cafés adulterados com cascas e paus e o índice 2 representa a classe dos cafés puros. .... 89

**Tabela 10** - Melhores resultados da classificação de cafés para os conjuntos de validação e teste obtidos com os modelos SIMCA, construídos com os dados da: impedância expressa como número complexo ( $Z' + jZ''$ ), parte real da impedância ( $Z'$ ), parte imaginária da impedância, módulo da impedância ( $|Z|$ ) e ângulo de fase da impedância ( $\varphi$ ). Os números entre parênteses referem-se à quantidade de fatores usados na modelagem. O índice 1 representa a classe dos cafés adulterados com cascas e paus e o índice 2 representa a classe dos cafés puros. .... 90

## LISTA DE ABREVIATURAS E SIGLAS

|           |                                                                                                                  |
|-----------|------------------------------------------------------------------------------------------------------------------|
| PLS       | Regressão por mínimos quadrados parciais (do inglês, Partial Least Square)                                       |
| MLR       | Regressão linear múltipla (do inglês, Multiple Linear Regression)                                                |
| PCA       | Análise de componentes principais (do inglês, Principal Component Analysis)                                      |
| SIMCA     | Modelagem independente e flexível por analogias de Classe (do inglês Soft Independent Modeling by Class Analogy) |
| PLS-DA    | Análise discriminante por mínimos quadrados parciais                                                             |
| ANSA      | amino-2- naphthol-4-sulphonic acid                                                                               |
| $Z'$      | Componente real da impedância                                                                                    |
| $Z''$     | Componente imaginária da impedância                                                                              |
| $ Z $     | Módulo da impedância                                                                                             |
| $\varphi$ | Ângulo de deslocamento de fase                                                                                   |

## **CAPÍTULO I – APRESENTAÇÃO**

## 1.1 Introdução

A espectroscopia de impedância eletroquímica (EIE) é uma das técnicas mais sensíveis a alterações na interface eletrodo-solução, e por esse motivo tem sido amplamente empregada nos estudos de reações heterogêneas de transferência eletrônica (RANDVIIR, 2018), processos de adsorção (PRIETO et al., 2017), elucidação de mecanismo de reações (WU *et al.*, 2013), estudos cinéticos (VIDAKOVIC-KOCH et al., 2013), corrosão de materiais (WANG et al., 2017; AYAGOU et al., 2018), células solares (YADAV et al., 2017; SACCO, 2017), biocombustíveis (KASHYAP et al., 2014) e várias outras aplicações. A EIE mede a impedância elétrica de uma interface em estado estacionário quando se aplica sobre ela uma tensão alternada numa frequência particular e se mede a corrente resultante (LI et al., 2019). O registro dessa grandeza é geralmente realizado em várias frequências resultando em um espectro de impedâncias.

A aplicação de um potencial senoidal sobre o sistema de interesse provoca o surgimento de uma corrente que também varia no tempo de maneira senoidal (PAJKOSSY; JURCZAKOWSKI, 2017), e quando o valor máximo dessa corrente é atingido em momento diferente ao valor máximo do potencial aplicado, surge um deslocamento entre as fases da senóides, cujas características dependem dos processos eletroquímicos que estavam ocorrendo na superfície do eletrodo no momento da medida. A forma mais comum de descrever o comportamento do potencial é através de uma função seno, do tipo  $V_0 \sin(\omega t)$ , onde  $V_0$  é o valor máximo (amplitude) do potencial. A função  $I_0 \sin(\omega t + \varphi)$  representa o comportamento da corrente resultante, onde  $I_0$  é o valor máximo atingido por essa corrente, diferenciando do potencial pelo termo  $(\varphi)$  que representa o deslocamento entre a fase da senoide do potencial e da corrente (RAMOS et al., 2016). Por se tratar da razão entre o potencial e a corrente, a impedância é descrita mais facilmente quando a relação de Euler é aplicada. Assim, a descrição da impedância passa a ser feita por um número complexo do tipo  $(Z' + jZ'')$ , onde  $Z'$  representa a parte real, e  $Z''$  representa a parte imaginária da impedância. Também é possível descrever a impedância usando seu módulo ( $|Z| = \sqrt{Z'^2 + Z''^2}$ ) e ângulo de fase  $(\varphi)$ .

A análise da impedância em uma única frequência não fornece tantas informações sobre o sistema quanto a análise de impedâncias medidas em várias frequências. Por exemplo, a impedância medida em altas frequências fornece informações do sistema eletroquímico mais relacionadas aos processos resistivos, que estão ocorrendo na interface eletrodo-solução e contribuindo para a componente real da impedância ( $Z'$ ), enquanto a impedância medida em baixas frequências fornece informações mais relacionadas aos processos reativos (capacitivo e

indutivo), que estão ocorrendo na interface eletrodo-solução e contribuindo para a componente imaginária da impedância ( $Z''$ ). Nesse aspecto é natural afirmar que todo o potencial que a técnica tem a ofertar somente poderá ser extraída se a investigação for realizada em várias frequências.

Apesar de toda a informação fornecida durante a aquisição de um espectro de impedâncias, a abordagem mais usada na extração de informações utiliza o ajuste de um modelo ao espectro experimental, tomando como base um circuito elétrico equivalente. A maior parte dos trabalhos da literatura recorrem a essa abordagem quando necessitam construir modelos quantitativos. Diante do que já foi exposto pode-se destacar algumas desvantagens que surgem quando se utiliza essa abordagem. A primeira desvantagem é a perda de informações sobre o sistema investigado quando elementos elétricos são usados para modelar os sistemas eletroquímicos, como está detalhado no capítulo III. A segunda desvantagem é a não utilização de toda informação modelada pelo circuito equivalente, quando modelos quantitativos são construídos correlacionando apenas um único elemento elétrico do circuito equivalente ajustado à variação da concentração do analito. Essa é a prática comum dos trabalhos encontrados na literatura, como será mostrado no capítulo II.

Nesse contexto, o capítulo II desse trabalho propõe o uso de mais de um elemento do circuito equivalente ajustado, ao invés de um único elemento, na construção de modelos quantitativos empregando a regressão linear múltipla. Modelos multivariados e univariado foram construídos para o ferrocianeto, catecol e hidroquinona e seus resultados foram comparados entre si. Os capítulos III e IV apresentam abordagens alternativas que consistem na utilização direta dos espectros de impedância, que são expressos por números complexos, sem recorrer ao ajuste por circuito equivalente, na construção de modelos multivariados quantitativos (capítulo III), usando a regressão por mínimos quadrados parciais, e qualitativos de caráter classificatório (capítulo IV), usando a análise discriminante por mínimos quadrados parciais (PLS-DA) e a modelagem independente por analogias de classes (SIMCA). Ambas as propostas empregam os dados de impedância expressos como números complexos ( $Z' + jZ''$ ). Os modelos PLS do capítulo III foram construídos para o ferrocianeto, catecol e hidroquinona e seus resultados comparados aos resultados fornecidos pelos modelos PLS construídos com a parte real da impedância ( $Z'$ ), parte imaginária da impedância ( $Z''$ ), módulo da impedância ( $|Z|$ ) e ângulo de fase ( $\varphi$ ). Os modelos PLS-DA e SIMCA do capítulo IV foram construídos com o objetivo de classificar amostras de cafés quanto à presença de adulteração – amostras de cafés adulterados com cascas e paus – e sem adulteração (cafés puros).

## 1.2 Objetivos

### 1.2.1 Objetivos Gerais

- Propor uma estratégia multivariada que emprega mais de um elemento do circuito elétrico equivalente ajustado aos dados de impedância para construir modelos quantitativos usando a regressão linear múltipla (MLR), como alternativa aos modelos univariados construídos com um único elemento do circuito equivalente, empregando a regressão linear simples. Aplicar a estratégia na quantificação de ferrocianeto, catecol e hidroquinona, separadamente, em amostras sintéticas e de água de torneira fortificada;
- Propor uma estratégia multivariada para a construção de modelos quantitativos empregando a regressão por mínimos quadrados parciais (PLS) aos dados de impedância na forma de números complexos ( $Z' + jZ''$ ), sem a necessidade de aplicar o ajuste por circuito elétrico equivalente aos espectros. Aplicar a estratégia na quantificação de ferrocianeto, catecol e hidroquinona, separadamente, em amostras sintéticas e de água de torneira fortificada;
- Propor estratégias multivariadas para a construção de modelos qualitativos, aliando os dados de impedância escritos na forma de números complexos ( $Z' + jZ''$ ) a duas técnicas de reconhecimento de padrão supervisionado (PLS-DA e SIMCA); aplicando-as na construção de modelos para classificação de amostras de cafés puros e adulterados com cascas e paus.

## **CAPÍTULO II – CONSTRUÇÃO DE MODELOS QUANTITATIVOS EMPREGANDO REGRESSÃO LINEAR MÚLTIPLA (MLR)**

## 2.1 Introdução

A obtenção de informações a partir do espectro de impedância do sistema investigado é uma tarefa difícil e, por esse, motivo alguns métodos foram desenvolvidos para essa finalidade. A análise por circuito elétrico equivalente é uma das estratégias mais usadas na investigação e extração de informações dos espectros de impedância. Esse método é usado na descrição da interface eletrodo-solução por intermédio de um circuito elétrico composto por elementos passivos (resistores, capacitores e indutores). Nessa abordagem, a impedância do circuito proposto é ajustada a impedância obtida experimentalmente, e os arranjos dos elementos elétricos e seus valores no circuito são empregados para descrever os macroprocessos envolvidos no transporte de massa e carga (BOUKAMP, 1985).

Neste capítulo é apresentada uma estratégia para melhorar a construção de modelos quantitativos, costumeiramente construídos relacionando-se a variação da concentração de uma espécie de interesse, presente na interface eletrodo-solução, aos valores de um único elemento elétrico do circuito equivalente ajustado. A proposta inicial visa construir modelos quantitativos de regressão linear múltipla utilizando mais de um elemento do circuito elétrico equivalente ajustado, desde que os elementos adicionados obedeçam aos requisitos da modelagem, como será discutido mais adiante.

## **2.2 Objetivos**

### **2.2.1 Objetivos Específicos**

1. Encontrar um circuito equivalente que se ajuste aos espectros de impedância registrados em diferentes níveis de concentração do analito;
2. Avaliar as correlações entre a concentração do analito e os valores dos elementos elétricos do circuito ajustado, e as correlações entre os elementos do próprio circuito elétrico;
3. Construir modelos de regressão linear múltipla usando como variáveis explicativas os elementos elétricos do circuito equivalente ajustado e verificar a significância de cada elemento elétrico para os modelos construídos;
4. Comparar os resultados do modelo MLR com o modelo de regressão univariada construído com o componente elétrico mais correlacionado com a variação da concentração da espécie investigada.

## 2.3 Fundamentação Teórica

### 2.3.1 Espectroscopia de Impedância Eletroquímica

A impedância é a capacidade de um material/interface resistir à passagem de uma corrente elétrica alternada, e por isso é medida na mesma unidade que a resistência elétrica ( $\Omega$ ). Na EIE, um potencial alternado, geralmente senoidal, é aplicado à uma célula eletroquímica, obtendo-se como resposta uma corrente que também é alternada, possui a mesma frequência do potencial aplicado, porém pode apresentar fase deslocada em relação a fase do potencial aplicado. Esse deslocamento de fase ou ângulo de fase ( $\varphi$ ) depende das características resistivas e capacitivas do sistema eletroquímico, e por esse motivo pode ser usado para monitorar processos físicos na superfície do eletrodo (LASIA, 2014).

As Equações Equação 1 e Equação 2 descrevem o potencial senoidal aplicado  $V(\omega, t)$  e a corrente resultante  $I(\omega, t)$ , respectivamente, onde  $t$  é o tempo,  $\omega$  é frequência angular ( $\omega = 2\pi f$ , em que  $f$  representa a frequência linear),  $V_0$  é o valor máximo (amplitude) alcançado pelo potencial; e  $I_0$  é o valor máximo da corrente resultante, que está com a fase deslocada por um ângulo  $\varphi$ .

$$V_{(\omega, t)} = V_0 \text{sen}(\omega t) \quad \text{Equação 1}$$

$$I_{(\omega, t)} = I_0 \text{sen}(\omega t + \varphi) \quad \text{Equação 2}$$

A impedância do sistema ( $Z$ ) pode ser encontrada usando uma expressão análoga a Lei de Ohm, conforme a equação Equação 3.

$$Z_{(\omega, t)} = \frac{V_{(\omega, t)}}{I_{(\omega, t)}} = \frac{V_0 \text{sen}(\omega t)}{I_0 \text{sen}(\omega t + \varphi)} \quad \text{Equação 3}$$

Uma notação amplamente utilizada para simplificar os cálculos consiste na transformação das funções trigonométricas (seno e cosseno) em funções exponenciais complexas, empregando a relação de Euler, descrita na equação Equação 4, onde  $j^2$  representa a unidade imaginária (-1).

$$e^{j\varphi} = \cos(\varphi) + j \text{sen}(\varphi) \quad \text{Equação 4}$$

Aplicando a relação de Euler é possível expressar a impedância descrita pela Equação 3 em termos de números complexos. Com base nessa notação, pode-se reescrever a Equação 3, conforme mostrado na Equação 5:

$$Z_{(\omega,t)} = |Z(\omega)|e^{j\varphi} = |Z|[\cos(\varphi) + j\sin(\varphi)] = Z' + jZ'' \quad \text{Equação 5}$$

, onde a parte real da impedância ( $Z'$ ) é igual a  $|Z|\cos(\varphi)$  e a parte imaginária da impedância ( $Z''$ ) é igual a  $|Z|\sin(\varphi)$ . O módulo e a fase da impedância, Equações 6 e 7, podem ser obtidos a partir da aplicação de relações trigonométricas do triângulo retângulo exibido no plano complexo da Figura 1.

$$Z(\omega) = \sqrt{Z'^2(\omega) + jZ''^2(\omega)} \quad \text{Equação 6}$$

$$\varphi(\omega) = \tan^{-1} \left( \frac{Z''(\omega)}{Z'(\omega)} \right) \quad \text{Equação 7}$$

No plano complexo fica explícito que para descrever a impedância ( $Z$ ) são necessárias duas informações: Pode-se: descrevê-la, usando as partes real ( $Z'$ ) e imaginária ( $Z''$ ), ou usando o módulo da impedância ( $|Z|$ ) e o ângulo de fase ( $\varphi$ ).

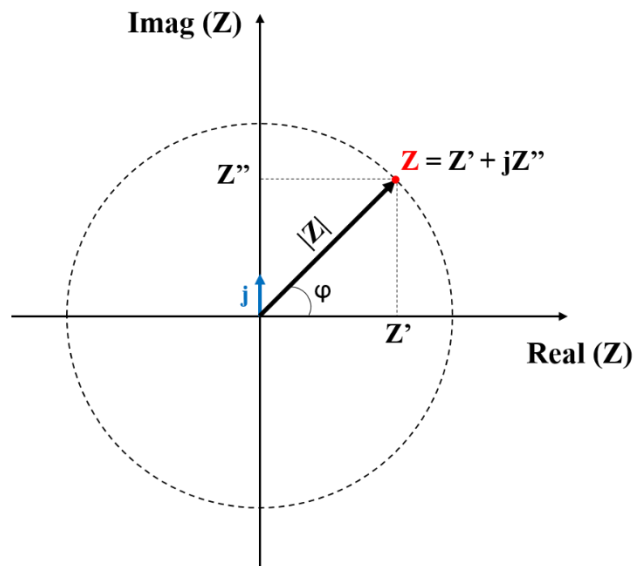


Figura 1 - Plano complexo de Argand-Gauss

### 2.3.2 Modelagem por Circuito Elétrico Equivalente

A EIE fornece dados no domínio da frequência/tempo e isso dificulta a obtenção de informações eletroquímicas diretamente do espectro. O procedimento mais utilizado para extrair essas informações é a modelagem do processo que está ocorrendo na célula eletroquímica por meio de um circuito elétrico equivalente derivado da combinação de resistores (R), capacitores (C) e/ou indutores (L). Nessa abordagem, uma curva teórica, resultante da combinação dos elementos elétricos do circuito equivalente escolhido para simular o processo eletroquímico, é ajustada ao espectro de impedância obtido experimentalmente, chegando-se aos valores dos elementos elétricos do circuito (XU et al., 2016; LASIA, 2005).

Antes de buscar compreender a modelagem empregando circuito elétrico equivalente é importante entender a relação entre o potencial e a corrente elétrica AC nos elementos passivos, visto que esses elementos são usados nos circuitos equivalentes que descrevem os dados.

### 2.3.3 Resistor

O resistor é o elemento mais simples de um circuito elétrico, no qual, a relação entre a tensão e a corrente é descrita pela lei de Ohm, conforme a Equação Equação 8. Se um potencial senoidal for aplicado a esse elemento, a corrente resultante também será senoidal e os valores máximos do potencial aplicado e da corrente resultante serão alcançados no mesmo instante (Figura 2a). Isso significa que a corrente senoidal resultante possui a mesma fase que o potencial senoidal aplicado, não existindo diferença (deslocamento) entre as fases das ondas.

$$Z_R(\omega, t) = \frac{V_{(\omega, t)}}{I_{(\omega, t)}} = \frac{(V_0 \text{sen}(\omega t))}{(I_0 \text{sen}(\omega t + \varphi))} = \frac{V_0}{I_0} = R \quad \text{Equação 8}$$

A impedância de um resistor independe da frequência do sinal aplicado, como pode ser interpretado pela Equação Equação 8 e visto na Figura 2b. Outra característica da impedância de um resistor é que sua parte imaginária é nula (Figura 2c), sendo representado somente pela a parte real da impedância, que é a parte que descreve a impedância resistiva.

Nos sistemas eletroquímicos o resistor é empregado para modelar, principalmente, a resistência da solução que está em contato com a interface do eletrodo/material e a resistência à transferência de carga da reação eletroquímica que se desenvolve na interface.

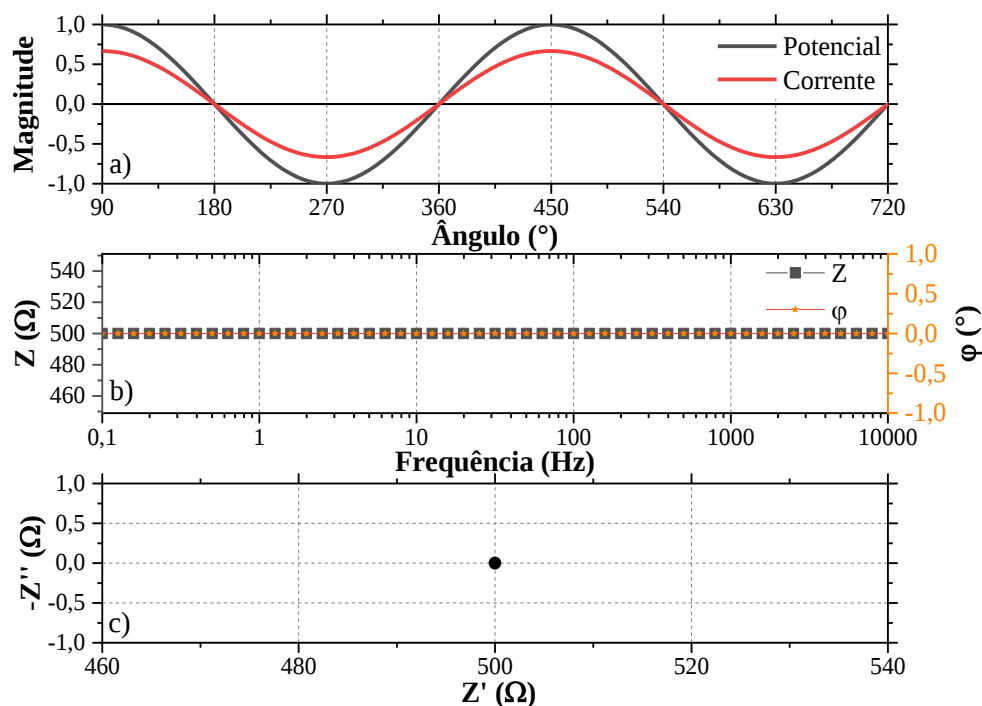


Figura 2 - (a) Comportamento da corrente elétrica resultante da aplicação de uma tensão senoidal a um resistor com resistência de 500  $\Omega$ . (b) Variação do módulo da impedância e ângulo de fase em função da aplicação de frequências diferentes (gráfico de Bode). (c) Comportamento da impedância real ( $Z'$ ) e impedância imaginária (diagrama de Nyquist).

### 2.3.4 Capacitor

O capacitor é o elemento cuja principal função consiste no armazenamento de energia potencial elétrica. Ele é composto por dois ou mais condutores isolados entre si por meio de um material dielétrico (isolante). A geometria, distância entre os condutores e a natureza do material dielétrico definem a propriedade fundamental do capacitor: a capacitância. Essa propriedade determina a quantidade de carga elétrica armazenada por um capacitor quando uma dada diferença de potencial é aplicada.

A corrente resultante em um capacitor de capacitância  $C$ , submetido a um potencial senoidal  $V(\omega, t)$ , pode ser expressa segundo a Equação 9:

$$I(\omega, t) = C \frac{dV(\omega, t)}{dt} = C\omega V_0 \cos(\omega t) = C\omega V_0 \sin\left(\omega t + \frac{\pi}{2}\right) \quad \text{Equação 9}$$

Através da Equação 9 percebe-se que a corrente resultante em um capacitor está com a fase deslocada  $90^\circ$  à frente em relação à fase do potencial aplicado. Da Equação 9 pode-se inferir que a amplitude da corrente  $I_0$  é  $C\omega V_0$ . Usando a Equação Equação 5 é possível calcular a impedância de um capacitor, conforme descreve a Equação Equação 10.

$$Z_C = \frac{V(\omega, t)}{I(\omega, t)} = \frac{V_0}{I_0} e^{j\varphi} = \frac{V_0}{C\omega V_0} e^{j\frac{\pi}{2}} = \frac{1}{j\omega C} \quad \text{Equação 10}$$

Se um potencial senoidal for aplicado a um capacitor, a corrente resultante será senoidal, mas o valor máximo da tensão aplicada estará  $90^\circ$  adiantado do valor máximo da corrente resultante (Figura 3a). Através da Equação Equação 10 é possível verificar que a magnitude da impedância ( $Z_C$ ) depende da frequência do potencial aplicado ( $\omega=2\pi f$ ). Da Equação Equação 10 também é possível inferir que um capacitor submetido a frequências pequenas, tendendo a zero, atuará como uma chave aberta, pois o valor de sua impedância tenderá ao infinito. Já em altas frequências, o capacitor tende a se comportar como um curto-circuito, visto que o valor de sua impedância se aproxima de zero à medida que a frequência tende ao infinito (Figura 3b).

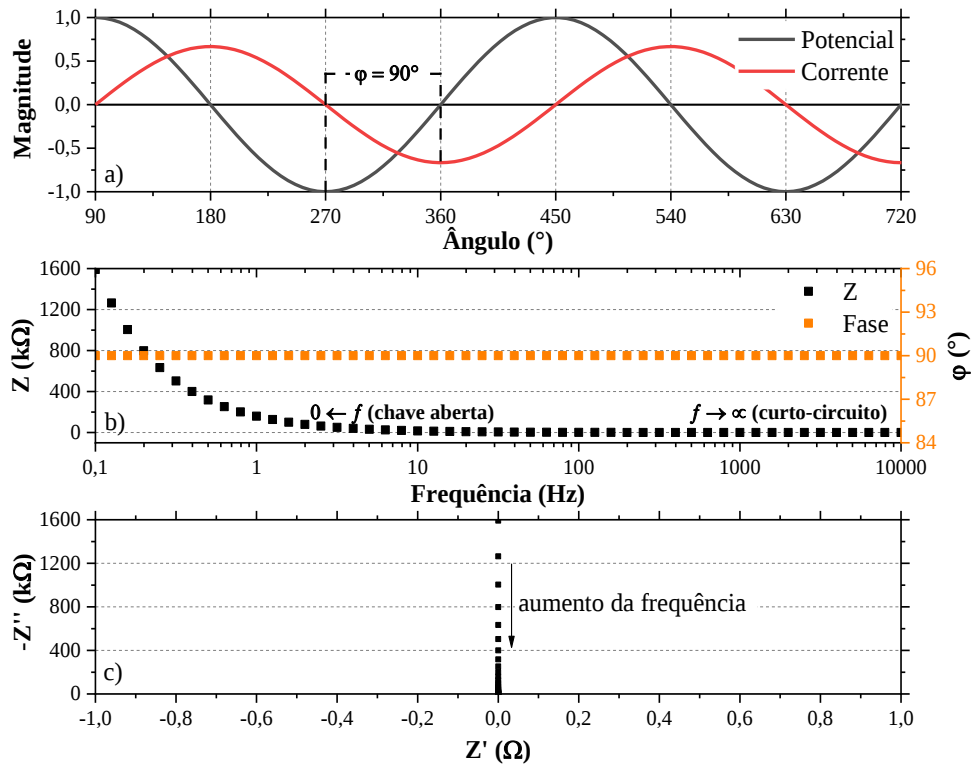


Figura 3 - (a) Comportamento da corrente elétrica resultante da aplicação de uma tensão senoidal a um capacitor com capacitância de  $1,0 \mu\text{F}$ . (b) Variação do módulo da impedância e

ângulo de fase em função da aplicação de frequências diferentes (gráfico de Bode). (c) Comportamento da impedância real ( $Z'$ ) e impedância imaginária (diagrama de Nyquist).

Outra característica da impedância de um capacitor é que sua parte real ( $Z'$ ) é nula (Figura 3c), sendo representado somente pela parte imaginária da impedância ( $Z''$ ), que é a parte que descreve as reatâncias capacitivas e indutivas.

O comportamento do indutor não será detalhado tendo em vista que o circuito empregado para explicar a maior parte dos processos eletroquímicos não utilizam esse elemento elétrico. Todavia, ele apresenta comportamento inverso do capacitor, visto que a corrente resultante que o circula está  $90^\circ$  atrasada em relação ao potencial aplicado, e sua impedância é diretamente proporcional à frequência do potencial aplicado, como mostrado pela Equação Equação 11, onde L representa a indutância do indutor.

$$Z_C = j\omega L \quad \text{Equação 11}$$

Nos sistemas eletroquímicos os capacitores são utilizados para modelar, principalmente, a dupla camada elétrica e camadas superficiais dielétricas (ex. eletrodo recoberto).

### 2.3.5 Elemento de Fase Constante (CPE)

A aplicação de uma diferença de potencial ao eletrodo provoca a polarização de cargas na sua superfície e, como consequência forma-se uma camada de íons de carga oposta na interface eletrodo-solução. Essa dupla camada elétrica (dl), formada pelas cargas do eletrodo e por uma camada de íons de carga oposta, é idealmente representada por um capacitor de placas paralelas. Contudo, a presença de irregularidades na superfície do eletrodo (ex. rugosidade, distribuição dos sítios ativos não uniformes) faz o comportamento dessa interface se afastar do comportamento de um capacitor ideal. A maneira encontrada para descrever melhor o comportamento da dupla camada foi generalizando a Equação Equação 10, conforme descrito na Equação Equação 12.

$$Z_{CPE} = \frac{1}{Q(j\omega)^n} \quad \text{Equação 12}$$

O expoente  $n$  pode assumir valores que variam de  $-1 \leq n \leq 1$ . O CPE se comporta como: um indutor ideal, quando  $n = -1$ ; um resistor, quando  $n = 0$ ; um elemento de Warburg quando  $n = 0,5$  e um capacitor ideal quando  $n = 1$  (LASIA, 2014; AMIRUDIN; THIERRY, 1995). O ângulo de fase do CPE é constante e só depende do valor de  $n$ , como exibido na Equação Equação 13.

$$\varphi_{CPE} = \frac{-n\pi}{2} \quad \text{Equação 13}$$

### 2.3.6 Circuito de Randles

O circuito de Randles é a representação comum, mais simples, dos sistemas eletroquímicos que apresentam processos controlados por cinética. Se o espectro de impedância é dominado pelas características de um semicírculo, isso significa que o processo eletroquímico que está acontecendo na superfície do eletrodo é limitado pela velocidade de transferência eletrônica. Graficamente, isso pode ser verificado no diagrama de Nyquist, que é um gráfico de  $Z'$  vs.  $Z''$ , como um semicírculo com duas intersecções no eixo real: uma em altas, e outra em baixas frequências (Figura 4/curva preta). A interseção em baixas frequências representa a resistência da solução ( $R_s$ ) e em alta frequência representa a soma entre a resistência da solução e a resistência à transferência de carga ( $R_s + R_{ct}$ ). Assim, a resistência à transferência de carga corresponde ao diâmetro do semicírculo. O capacitor ideal  $C_{dl}$  representa a dupla camada elétrica, mas em sistemas reais é comum substituí-lo por um CPE (HIRSCHORN et al., 2010).

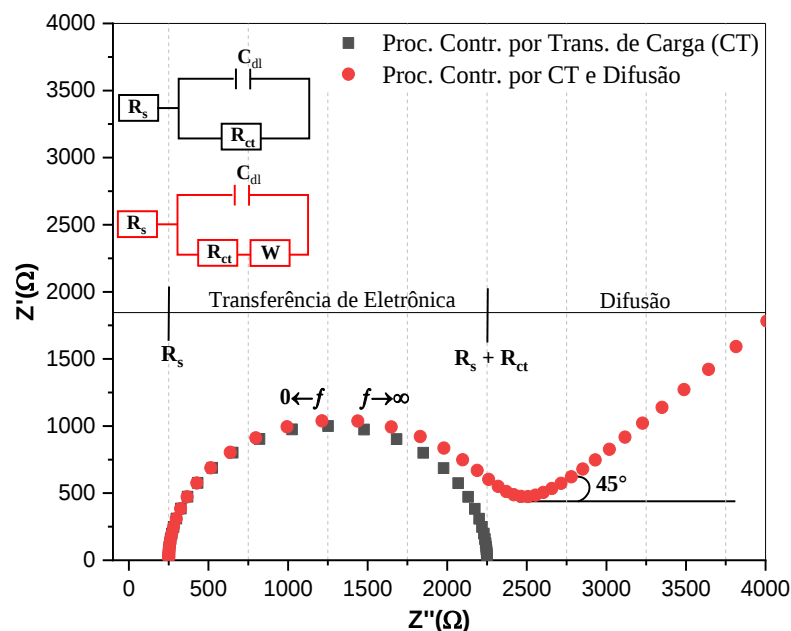


Figura 4 – Diagramas de Nyquist referentes a processos faradaicos controlado por cinética de transferência da carga e controlados por processo misto de cinética de transferência de carga e transferência de massa (difusão) Adaptado de Carvalho (2006). Valores dos elementos:  $R_s = 250 \Omega$ ;  $R_{ct} = 2 \text{ k}\Omega$ ;  $C_{dl} = 1 \mu\text{F}$ ;

Nos processos onde a corrente faradaica é limitada pela difusão da espécie eletroativa, pois a transferência eletrônica é rápida, é comum o aparecimento de uma região linear nas frequências baixas. Esse perfil indica que o sistema eletroquímico apresenta processos controlados por difusão e transferência de carga (BISQUERT, 2002). Esse comportamento linear pode ser modelado com o uso de um elemento de Warburg ( $Z_w$ ).

### 2.3.7 Validação das medidas

Os dados de EIE devem atender quatro requisitos principais para serem considerados confiáveis: linearidade, casualidade, estabilidade e finitude (GINER-SANZ, 2015). Esses critérios são fundamentais e juntos indicam se as medidas realizadas são válidas. Na eletroquímica, sabe-se que a relação entre a corrente e o potencial não é linear e tende a assumir um comportamento exponencial, como descrito na equação de Butler-Volmer. Para a lei de Ohm ser válida é necessário que a relação entre essas grandezas seja linear, condição que pode ser obtida por aplicação de amplitudes pequenas. Nesse caso, diz-se que esta relação é pseudo-linear (LAISA, 2014).

Se a intenção é avaliar a resposta de um sistema através da Espectroscopia de Impedância Eletroquímica, que emprega um estímulo AC, não faz sentido que essa resposta seja resultado da contribuição de outros estímulos. O critério de casualidade descreve que a resposta do sistema deve depender, exclusivamente, do sinal de excitação (AC) e independe de outros estímulos externos ao sistema em estudo. O sistema também deve ter estabilidade. Isso significa que o sistema analisado deve permanecer em estado estacionário até a perturbação ser implementada pela fonte externa (AC), e ser capaz de retornar ao estado estacionário assim que a perturbação for descontinuada. A implicação imediata desse critério é que o sistema não pode apresentar desvios durante a aquisição do espectro. O último critério é a finitude. Esse critério descreve que os valores das impedâncias real e imaginária devem ser finitos ao longo de toda a faixa de frequência.

Alguns métodos podem ser usados na avaliação da confiabilidade dos dados de impedância. A figura de Lissajous é um dos métodos mais empregados e tem como finalidade principal avaliar a existência de distorção na linearidade dos dados, verificando o comportamento da corrente AC em função do potencial AC (EROL, 2015). Outro método bastante utilizado é a relação de Kronig-Kramers (BOUKAMP, 1986). A aplicação dessa relação permite a obtenção de um dos dois componentes da impedância (real ou imaginária) em função do outro. Em outras palavras, pode-se calcular o valor da componente imaginária da impedância utilizando a componente real, e vice-versa. Para que seja possível calcular o valor de um componente com base no valor do outro, é necessário que os dados obedeçam aos critérios descritos acima. Caso contrário, as medidas podem ser consideradas inválidas.

### **2.3.8 Construção de Modelos Quantitativos**

Os principais problemas da química analítica giram em torno da identificação e quantificação de uma ou mais espécies presentes em uma matriz, e a gravimetria é um dos métodos clássicos mais empregados para esta finalidade. Nesse método, o analito é isolado da matriz, geralmente, após alguns processos físico-químicos, e a quantificação feita através da aferição da massa do analito resultante do processo. Com o advento da instrumentação surgiram novos métodos e através deles tornou-se possível construir metodologias que permitem realizar a identificação/quantificação de maneira mais rápida, e em alguns métodos sem destruir a amostra. Nesses métodos é explorado algumas propriedades físicas do analito (absorção ou

emissão de radiação, corrente elétrica, diferença de potencial, impedância elétrica, magnetismo, índice de refração e outras) para inferir sua presença e estimar sua quantidade.

Nos métodos que utilizam a voltametria, por exemplo, pode-se identificar a presença de uma espécie e estimar sua quantidade por meio do potencial e da corrente elétrica. Nesse método, aplica-se uma quantidade definida de potenciais à superfície do eletrodo de trabalho e mede-se a corrente elétrica resultante em cada potencial aplicado. O potencial onde a reação de oxirredução é máxima identifica o analito, e a intensidade de corrente elétrica gerada no processo está relacionada à concentração desse analito na superfície do eletrodo. Nesse caso é possível encontrar a concentração do analito por intermédio da corrente elétrica gerada no processo, porém, é necessário construir o modelo matemático que relacionará essas duas grandezas. Assim, a escolha da abordagem que culminará na construção do modelo dependerá de uma resposta: a concentração será correlacionada à corrente de pico ou às correntes obtidas em todos os potenciais analisados?

### 2.3.9 Calibração Univariada - Regressão Linear Simples

Com o intuito de correlacionar a concentração do analito à corrente de pico, uma alternativa é construir a curva analítica utilizando a regressão linear simples. Nessa abordagem, uma variável dependente  $Y$  (corrente, no exemplo dado) é correlacionada a uma variável independente  $X$  (concentração, no exemplo dado) através da Equação 14.

$$Y = a + bX \quad \text{Equação 14}$$

, onde  $a$  é o intercepto e  $b$  é a inclinação da reta de regressão linear.

Os coeficientes  $a$  e  $b$  são estimados durante a etapa de calibração, que consiste na etapa de construção do modelo. Nessa etapa são realizadas  $n$  observações variando-se os níveis de concentração e medindo-se o valor da corrente elétrica gerada em cada nível. De posse desses conjuntos de observações é possível encontrar os valores dos coeficientes  $a$  e  $b$ , usando o método dos mínimos quadrados, conforme descrito nas Equações 15 e (16).

$$b = \frac{n \sum XY - (\sum X)(Y)}{n(\sum X^2) - (\sum X)^2} \quad \text{Equação 15}$$

$$a = \frac{\sum Y - b \sum X}{n} \quad (16)$$

Esse tipo de abordagem é chamado de calibração de ordem zero porque um único valor de medida experimental é correlacionado à concentração (ou outra propriedade) da espécie por amostra (VALDERRAMA, 2009). Esse método funciona quando se garante que somente o analito contribui para a resposta e, que a mesma não é influenciada por fatores como efeito de matriz e presença de interferentes (que também é um efeito de matriz). Por isso, a aplicação desse método implica na utilização de sensor específico, caso contrário, o valor estimado pelo modelo não é confiável.

### 2.3.10 Trabalhos que Aplicam Regressão Linear Simples aos Dados de EIE

A Espectroscopia de Impedância Eletroquímica é predominantemente empregada no desenvolvimento de trabalhos qualitativos, especialmente em caracterizações eletroquímicas (YADAV et al., 2019; KASHYAP et al., 2014), mas quando o objetivo é o desenvolvimento de trabalhos quantitativos, a maioria construíram modelos quantitativos empregando diferentes abordagens aliadas a regressão linear simples.

SINGH e colaboradores desenvolveram um sensor impedimétrico para discriminar bactérias Gram-positivas e Gram-negativas, funcionalizando um eletrodo interdigitado de ouro com um filme fino de óxido de tungstênio modificado com vancomicina, um antibiótico glicopeptídeo conhecido pela interação específica com a camada de peptidoglicano de bactérias Gram-positivas. Os pesquisadores perceberam durante o estudo que a resistência a transferência de carga ( $R_{ct}$ ) do circuito equivalente ajustados aos espectros aumentavam linearmente com o aumento da quantidade de bactérias Gram-positivas (SINGH et al., 2019).

Silva e colaboradores desenvolveram um eletrodo de ouro modificado com nanopartículas de ANSA que ao longo dos estudos apresentou sensibilidade a tiramina, uma monoamina derivada da tirosina. O aumento da concentração de tiramina em meio tamponado provocava um aumento exponencial no  $\Delta R_{ct}$  ( $R_{ct(tampão)} - R_{ct(tiramina)}$ ), que foi modelado linearmente aplicando-se a função logarítmica (SILVA et al., 2019).

Ito e colaboradores construíram camadas automontadas (SAM) de dinucleótido de flavina adenina (FAD) e glicose -6-fosfato desidrogenase (GHB) sobre a superfície de um eletrodo de ouro. Verificou-se uma dependência entre a impedância e a variação da

concentração de glicose, que foi confirmada pela mudança na resistência a transferência eletrônica ( $R_{ct}$ ). A linearização entre essa relação se deu aplicando-se a função inversa sobre os valores de resistência ( $1/R_{ct}$ ) (Ito *et al*, 2019). Outro exemplo da abordagem supracitada pode ser verificado no trabalho de Carvalho (CARBALLO *et al.*, 2016).

Elshafey e colaboradores desenvolveram um sensor impedimétrico para detecção do herbicida alaclor (2-cloro-N-[2,6-dietil fenil]-N-metoximetil-acetanilida). Um filme de polímero impresso molecularmente, criado por eletrodeposição de o-fenilenodiamina na presença de alaclor, foi imobilizado na superfície de um eletrodo de carbono vítreo. O herbicida imobilizado foi lixiviado deixando cavidades específicas na superfície do eletrodo modificado.  $R-R_0/R_0$  (Elshafey *et al*, 2018). Outro exemplo da abordagem supracitada pode ser verificado no trabalho de Diouf (DIOUF *et al.*, 2012).

Day e colaboradores apresentaram um sensor condutimétrico para quantificação de íons potássio. Eles recobriram um eletrodo interdigitado planar de ouro com uma matriz polimérica composta por 35% policloreto de vinila (PVC), 5% do ionóforo valinomicina e 65% do plastificante dioctil sebacato (DOS). O módulo da impedância foi monitorado na frequência de 100 Hz, sendo verificado um decaimento exponencial dessa grandeza ao longo tempo e em função da concentração do íon potássio. O modelo de regressão linear simples foi construído correlacionando o logaritmo da concentração do íon potássio à razão  $Z-Z_0/Z_0$ , onde  $Z_0$  era o módulo da impedância registrada na ausência de íons potássio e na presença de diferentes concentrações de íons potássio ( $Z$ ) (DAY *et al.*, 2018). Outro exemplo da abordagem supracitada pode ser verificado no trabalho de Mahmoud (MAHMOUD *et al.*, 2019).

Aicher e colaboradores apresentaram um sensor impedimétrico íon-seletivo para Cálcio construído funcionalizando um eletrodo interdigitado de cobre, recoberto com ouro e com o ionóforo de cálcio ETH 1001. O modelo de regressão foi construído correlacionado à diferença  $1/Z''-Z_0''$ , onde  $Z''$  é a impedância imaginária registrada na presença do íon cálcio e  $Z_0''$  sem o cálcio, à concentração de cálcio na solução (AICHER *et al.*, 2017).

Zlatev e colaboradores desenvolveram um sensor impedimétrico para quantificar a atividade industrial de lipase. Uma camada fina de um nanocomposito sensível foi depositado sobre um eletrodo construído com placa de circuito impresso. O nanocomposito catalisava a hidrólise da lipase e com isso sofria um desgaste que provocava uma diminuição na impedância registrada na frequência de 10 Hz. O deslocamento de fase da corrente AC ( $\varphi$ ) causado pela mudança de impedância do eletrodo foi a resposta modelada em função da atividade da lipase (ZLATEV *et al.*, 2019).

Além da escolha da regressão linear simples, outra similaridade apresentada pelos trabalhos descritos acima é a modificação do eletrodo para tornar a resposta do mesmo mais seletiva, sendo o objetivo torná-la específica. O procedimento de modificação do eletrodo é geralmente laborioso e dispendioso, mas essa modificação pode ser vista como obrigatória, não somente porque a modificação é realizada para garantir mais seletividade na resposta, como também para garantir que a previsão de amostras externas fornecidas pelo modelo calibrado será válida.

### 2.3.11 Regressão Linear Múltipla (MLR)

Uma outra alternativa de calibração seria utilizar as correntes medidas em todos os potenciais aplicados ao mesmo sensor e correlacioná-las à concentração do analito. Esse tipo de abordagem é classificado como calibração multivariada de primeira ordem porque correlaciona a propriedade de interesse (concentração, no exemplo descrito) a mais de um valor da medida experimental (corrente elétrica, no exemplo descrito) (VALDERRAMA et al., 2009) coletada no mesmo sensor, por amostra. Em outras palavras, na regressão linear simples, a concentração é correlacionada a um único valor (um escalar), a corrente de pico. Na regressão linear múltipla, a concentração é correlacionada aos valores de correntes coletados em todos os potenciais aplicados em um mesmo eletrodo (um vetor).

Suponha que  $X$  é uma matriz que guarda os valores das correntes  $x_1, x_2, x_3 \dots x_k$ , registradas nos potenciais  $p_1, p_2, p_3 \dots p_k$  quando as concentrações do analito na interface eletrodo-solução eram  $y_1, y_2, y_3 \dots y_n$ . Cada concentração  $y_k$  está relacionada uma amostra diferente, e  $y$  é o vetor que guarda os valores dessas concentrações (propriedade de interesse). A regressão linear simples é um caso específico da regressão linear múltipla, onde uma única resposta instrumental ( $x_1$ ) é correlacionada ao parâmetro de interesse ( $y$ ). No caso geral,  $k$  variáveis podem ser correlacionadas a resposta  $y$  através de um modelo como descreve a Equação Equação 17.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k \quad \text{Equação 17}$$

$\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_k$  são vetores que guardam os valores da variável  $x$  registrados nos  $n$  níveis de concentrações diferentes.  $\beta_0$  é a interseção e  $\beta_1, \beta_2, \beta_3 \dots \beta_k$  são os coeficiente de regressão relacionados a cada variável do modelo.

O vetor que guarda os coeficientes de regressão  $\boldsymbol{\beta}$  é estimado através do método dos mínimos quadrados, como na regressão simples. Assim, durante a modelagem, Equação 18, procura-se encontrar o ajuste matemático que conduz ao menor somatório da diferença entre o valor estimado ( $\hat{y}$ ) e valor experimental ( $y$ ) ao quadrado, chamado de soma quadrática dos resíduos ou  $SSE = (\sum (\hat{y} - y)^2)$ .

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y}) \quad \text{Equação 18}$$

$\mathbf{X}$  é a matriz que guarda as respostas instrumental e  $\mathbf{y}$  é o vetor com os valores dos níveis do parâmetro de interesse (concentração) usados para calibrar o modelo.

Partindo do princípio que as amostras desconhecidas estão em ambiente químico semelhante as amostras usadas para calibrar o modelo, e voltando ao exemplo mencionado acima, é possível estimar a concentração ( $y$ ) de amostras desconhecidas aplicando os coeficientes de regressão estimados na Equação Equação 18, e as correntes elétricas apresentadas por essas amostras ( $\mathbf{x}$ ), na equação Equação 17. Todavia, antes de construir o modelo MLR é necessário garantir que os seguintes critérios estão sendo atendidos (SHARMA et al., 2018):

1. Existe um relacionamento linear entre as variáveis dependente e independente.
2. O resíduo deixado deve ser homocedástico. Isso significa que a variância do resíduo da modelagem deve independer do nível da variável dependente.
3. Não existe multicolinearidade entre as variáveis independentes ( $\mathbf{x}_k$ ). Esse critério é muito importante e minimiza as possibilidades da matriz quadrada  $\mathbf{X}^T \mathbf{X}$  ser singular.
4. O número de amostras (observações) tem que ser maior ou igual ao número de variáveis.

Apesar da abordagem na construção de modelos de Regressão Linear Múltipla ser similar à usada na construção de modelos Regressão Linear Simples, não foram encontrados trabalhos que modelam dados de Espectroscopia de Impedância Eletroquímica usando o MLR.

## 2.4 Metodologia

### 2.4.1 Soluções e Amostras

Os reagentes ferrocianeto de potássio ( $K_4[Fe(CN)_6]$ ), ácido acético glacial ( $CH_3COOH$ ), acetato de sódio ( $CH_3COONa$ ), grafite, nanotubo de carbono de parede múltiplas ( $\geq 98\%$  carbono, O.D.  $\times$  I.D.  $\times$  L  $10\text{ nm} \pm 1\text{ nm} \times 4.5\text{ nm} \pm 0.5\text{ nm} \times 3\text{-}6\text{ }\mu\text{m}$ ) e os isômeros hidroquinona (1,4 dihidroxibenzeno) e catecol (1,2 dihidroxibenzeno) foram adquiridos da Sigma-Aldrich com pureza superior a 99%. Todas as soluções foram preparadas com água ultrapura (resistividade  $\geq 18\text{ M}\Omega\text{ cm}$ ) obtida em um sistema de purificação Milli-Q (Millipore, Bedford, MA, USA). O óleo mineral foi adquirido da Specsol.

As medidas foram realizadas em tampão acetato  $0,10\text{ mol L}^{-1}$  (pH 4,66), preparado por meio da mistura de volumes iguais de uma solução de ácido acético e acetato de sódio, ambos, a  $0,20\text{ mol L}^{-1}$ . As soluções de ferrocianeto de potássio, hidroquinona e catecol foram preparadas por meio da dissolução da respectiva substância na solução tampão acetato  $0,1\text{ mol L}^{-1}$ . As amostras de água de torneira fortificadas foram preparadas misturando-se volumes iguais de água de torneira, fortificada com hidroquinona ou catecol, com solução tampão acetato  $0,2\text{ mol L}^{-1}$  (pH 4,66).

Um dos objetivos iniciais do trabalho era aplicar as metodologias desenvolvidas na determinação simultânea de quatro compostos fenólicos (catecol, hidroquinona, resorcinol e p-cresol), e por isso os ensaios com o catecol e a hidroquinona foram realizados no mesmo pH. A escolha do pH usado na construção do método de determinação simultânea foi baseada no estudo feito por Gomes, que selecionou o pH 5 para desenvolver o seu método de determinação simultânea (Gomes, 2015). Partindo dessa escolha, e diante da disponibilidade dos reagentes, simplicidade no preparo e durabilidade da solução, preferiu-se trabalhar com o tampão acetato, que ao ser preparado conforme descrito acima produzia uma solução cujo pH era 4,66.

### 2.4.2 Medidas de Impedância

As medidas de impedância foram realizadas em um potenciostato Autolab PGSTAT M101 (Metrohm Autolab, The Netherlands) controlado pelo software Nova 1.11 numa célula eletroquímica convencional de três eletrodos. Os eletrodos auxiliar e de referência foram uma haste de grafite e o eletrodo de calomelano saturado. O eletrodo de trabalho de carbono vítreo foi empregado durante as medidas realizadas na solução de ferrocianeto de potássio. As

medidas nas soluções de catecol e hidroquinona foram realizadas empregando um eletrodo de trabalho de pasta de nanotubos de carbono, cujo compósito foi preparado através da mistura de nanotubo de carbono de parede múltipla (70%) e óleo mineral (30%).

Os espectros de impedância foram registrados na faixa de frequência de 10 kHz a 0.1 Hz coletando o sinal numa taxa de 10 pontos por década de frequência. O potencial do pico de oxidação da espécie a ser quantificada era aplicado durante o registro dos espectros de impedância. Esse potencial de oxidação era encontrado mediante uma varredura de onda quadrada realizada em solução tampão acetato 0,10 mol L<sup>-1</sup> contendo 1 mmol L<sup>-1</sup> da espécie investigada, tomando como padrão a velocidade de varredura de 100 mV s<sup>-1</sup>. A amplitude de oscilação aplicada durante a aquisição dos espectros foi 5 mV.

#### **2.4.3 Construção de modelos de Regressão Linear Múltipla**

O procedimento iniciava-se com a obtenção dos espectros de impedância. Para isso, a concentração da espécie de interesse era variada crescentemente mediante adição de uma alíquota fixa da solução estoque e um espectro era registrado em cada nível de concentração. As espécies avaliadas foram o ferrocianeto de potássio, hidroquinona e catecol, empregando as faixas de trabalho de 60 a 220 µmol L<sup>-1</sup> para o ferrocianeto, de 3,0 a 15,0 µmol L<sup>-1</sup> para hidroquinona e de 35 a 150 µmol L<sup>-1</sup> para o catecol. Essas análises foram realizadas em tampão acetato 0,1 mol L<sup>-1</sup> (pH 4,66).

A próxima etapa consistia no ajuste do circuito equivalente aos espectros registrados, priorizando circuitos mais simples que conduziam ao menor RMSECV. Ao escolher o circuito que deveria ser ajustado ao espectro, um algoritmo do software NOVA 1.11 realizava o ajuste calculando os valores dos elementos elétricos do circuito escolhido, que só eram considerados válidos quando a convergência do cálculo era alcançada. Assim, só foram considerados os espectros cujos circuitos ajustados atingiram a convergência. Tendo em vista que o mesmo circuito era ajustado aos espectros registrados em níveis de concentrações diferentes, construía-se uma matriz onde cada coluna guardava os valores para cada elemento do circuito equivalente registrados ajustados em um dado nível de concentração (linha).

O modelo de regressão linear múltipla era construído usando os valores ajustados em cada elemento elétrico como variável explicativa. Essas variáveis eram normalizadas antes da construção do modelo, empregando a Equação Equação 19.

$$x_{ij} = \frac{\varepsilon_{ij} - \frac{\max(\varepsilon_j) + \min(\varepsilon_j)}{2}}{\frac{\max(\varepsilon_j) - \min(\varepsilon_j)}{2}} \quad \text{Equação 19}$$

Onde “max” e “min” representam os valores máximos e mínimos, respectivamente, das variáveis explicativas  $\varepsilon_j$  (componentes elétricos do circuito equivalente). As variáveis que possuíam alta colinearidade entre si foram identificadas mediante avaliação da matriz de correlação calculada através da Equação Equação 20.

$$r_{x_j}, r_{x_k} = \frac{\sum_{i=1}^n (x_{ji} - \bar{x}_j)(x_{ki} - \bar{x}_k)}{(n-1)s_{x_i}s_{x_k}} \quad \text{Equação 20}$$

Os valores  $s_{x_i}$  e  $s_{x_k}$  representam o desvio-padrão das variáveis  $x_{ij}$  e  $x_k$ , respectivamente, enquanto  $\bar{x}_j$  e  $\bar{x}_k$  representam as médias dessas variáveis.

O teste da hipótese nula ( $H_0: \beta_j = 0$ ) foi aplicado aos coeficientes de regressão para avaliar se eles eram realmente significativos para o modelo. Assim, se  $|t_0| > t_c$ , a hipótese nula era rejeitada sendo  $\beta_j$  significativo para o modelo. A Equação Equação 21 exibe o cálculo de  $t_0$ :

$$t_0 = \frac{\hat{b}_j}{\sqrt{s^2 C_{jj}}} \quad \text{Equação 21}$$

Onde  $\hat{b}_j$  é o coeficiente de regressão avaliado,  $C_{jj}$  é o j-ésimo elemento da diagonal da matriz inversa  $(\mathbf{X}^T \mathbf{X})^{-1}$ ,  $t_0$  é o valor da distribuição de t-Student calculado para o coeficiente de regressão  $b_j$  e  $t_c$  é o ponto da distribuição t tabelado usando um nível de confiança de 95%.

A raiz quadrada do erro quadrático médio obtido durante a validação cruzada (RMSECV), juntamente com a raiz quadrada do erro quadrático médio de previsão para as amostras de água de torneira fortificadas, foram os critérios empregados para avaliar o desempenho dos modelos construídos.

## 2.5 Resultados e Discussão

### 2.5.1 Quantificação de ferrocianeto

A primeira etapa do trabalho foi avaliar a aplicação do MLR na obtenção de modelos com dados de impedância, empregando o eletrodo de carbono vítreo e ferrocianeto com espécie de interesse. Os dados usados nessa etapa são os mesmos gerados no desenvolvimento dos modelos PLS exibidos no capítulo III. Os espectros de impedância registrados com o eletrodo de carbono vítreo em várias concentrações diferentes de ferrocianeto, assim como o circuito equivalente que se ajustou melhor aos espectros, estão exibidos na Figura 5.

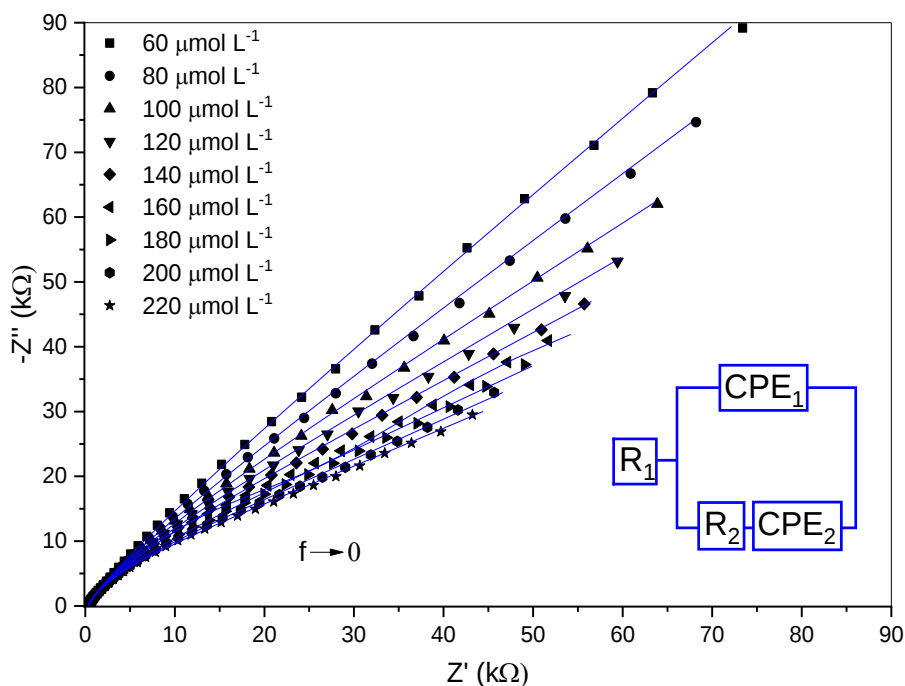


Figura 5 - Diagramas de Nyquist referentes à aquisição dos espectros de impedância em diferentes concentrações de ferrocianeto.  $E_{dc} = 0,22 \text{ V vs Hg}_2\text{Cl}_2 \text{ (sat)}$ ; amplitude = 5 mV. Inserção: circuito equivalente ajustado aos espectros.

Os espectros da Figura 5 exibem duas regiões: i) um semicírculo malformado na região de alta frequência, e ii) uma parte linear na região de baixas frequências. O semicírculo e a parte linear são usualmente associados ao processo de transferência de carga e ao controle difusional, respectivamente. O circuito equivalente exibido na Figura 5 conduziu ao melhor ajuste, como será discutido adiante. Esse circuito é composto por quatro elementos elétricos: duas resistências ( $R_1$  e  $R_2$ ) e dois elementos de fase constante ( $CPE_1$  e  $CPE_2$ ). Os elementos elétricos

$R_1$ ,  $CPE_1$ ,  $R_2$  e  $CPE_2$  estão relacionados, respectivamente, à: resistência da solução, impedância da dupla camada elétrica, impedância do processo de transferência eletrônica e à impedância do processo difusional (transferência de massa). O espectro do ferrocianeto está discutido detalhadamente no capítulo III. A Figura 6 exibe os circuitos equivalentes testados durante o ajuste dos dados do ferrocianeto, exemplificado na concentração de  $200 \mu\text{mol L}^{-1}$ .

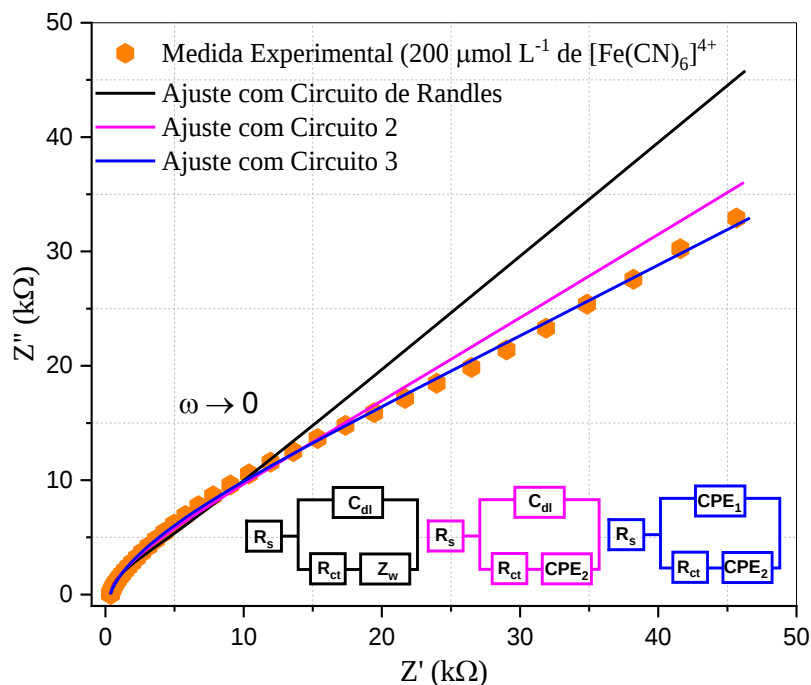


Figura 6 – Processo de seleção entre os circuitos usados no ajuste dos dados do ferrocianeto exemplificado no espectro referente à concentração de  $200 \mu\text{mol L}^{-1}$ .

Inicialmente, selecionou-se o circuito de Randles para ser ajustado aos dados, visto que o mesmo pode descrever um processo ideal de transferência eletrônica controlada por difusão (circuito de Randles - Figura 6), semelhantemente ao processo apresentado pelo ferrocianeto. Apesar da convergência ter sido alcançada durante o ajuste utilizando esse circuito, percebeu-se que ele deixava muito resíduo nas regiões de média e baixa frequência. A convergência indicou que o circuito escolhido era compatível com os processos que estavam ocorrendo na interface eletrodo-solução, porém o resíduo deixado pelo ajuste indicou que esses processos não estavam ocorrendo de maneira ideal. A baixa concentração de ferrocianeto empregada possivelmente estava contribuindo para o afastamento da idealidade, e por isso o elemento de Warburg do circuito de Randles foi substituído por um elemento de fase constante ( $CPE_2$ ), dando origem ao circuito 2 da Figura 6. A substituição conduziu a uma melhora

significativa, contudo o melhor ajuste só foi alcançado quando o capacitor também foi substituído por um elemento de fase constante, resultando no circuito 3 da Figura 6. O valor de  $n$  ajustado (equação Equação 12)) para o  $CPE_1$  foi 0,89, que é um valor próximo ao esperado para um capacitor ideal ( $n = 1$ ), enquanto o valor de  $n$  ajustado para o  $CPE_2$  foi 0,36, que é próximo ao valor um elemento de Warburg ideal ( $n = 0,5$ ). Esse mesmo circuito foi ajustado aos demais espectros.

A próxima etapa após o ajuste do circuito equivalente foi avaliar o comportamento dos valores ajustados à cada elemento elétrico em função da variação da concentração. A normalização da equação Equação 19 foi aplicada às variáveis para remover o efeito da magnitude, visto que os valores dos elementos elétricos variavam em diferentes ordens de intensidade. Por exemplo, os valores de  $R_1$  variavam na escala de Ohm ( $\Omega$ ) enquanto os valores de  $R_2$  variavam na escala de quilo-Ohm ( $k\Omega$ ). A Figura 7 exibe o comportamento dos valores normalizados de cada componente elétrico do circuito ajustado aos espectros de impedância obtidos em concentrações diferentes de ferrocianeto. De maneira geral, observou-se que os componentes  $CPE_1$  e  $CPE_2$  apresentaram valores crescentes com o aumento da concentração, enquanto os componentes  $R_1$  e  $R_2$  apresentaram valores de resistências que tendiam a diminuir.

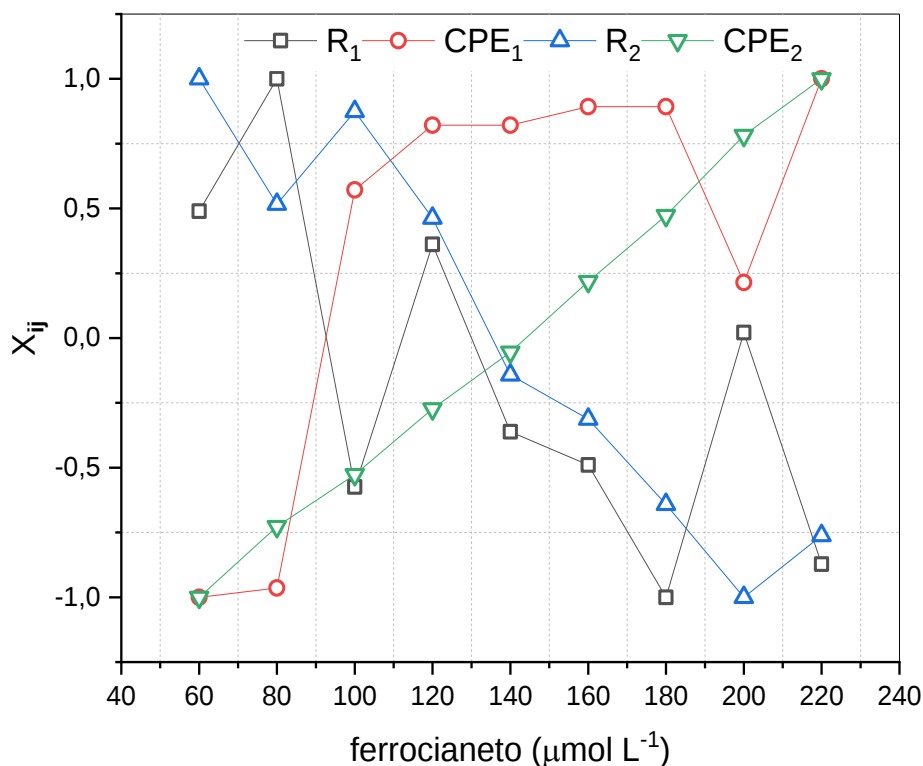


Figura 7 - Mudanças nos valores normalizados dos elementos elétricos do circuito equivalente em função da variação da concentração de ferrocianeto.

O comportamento de  $R_2$ ,  $CPE_1$  e  $CPE_2$  estão correlacionados com a variação da concentração do ferrocianeto, apresentando coeficientes de correlação de Pearson de -0,95, 0,70 e 0,99, respectivamente. Esses valores de coeficientes são maiores que o valor crítico de 0,67 (ASUERO et al., 2006; Link), considerando um teste bicaudal com 95% de confiança. Por outro lado,  $R_1$  apresentou um coeficiente de -0,66, que implica afirmar que a variação exibida por ele não é estatisticamente diferente de uma variação aleatória. Em uma situação ideal, espera-se que os valores ajustados para esse elemento sejam constantes, visto que a concentração do eletrólito é muito alta e a resistência da solução varia infimamente com o aumento da concentração do ferrocianeto. Porém, apesar do comportamento ser diferente do esperado, observa-se que os valores ajustados são aleatórios ao longo de toda a faixa de concentração, e sua variação real (349–379  $\Omega$ ) é menor que 9%. As correlações entre a concentração do ferrocianeto e os elementos  $R_2$ ,  $CPE_1$  e  $CPE_2$  indicam, respectivamente, a diminuição da resistência a transferência de carga, aumento da capacitância da dupla camada, e aumento da impedância faradaica, referente ao transporte de massa, em função do aumento da concentração do ferrocianeto.

A ausência de multicolinearidade entre as variáveis explicativas é um dos principais requisitos para construção de um modelo MLR, garantindo menores chances de o cálculo da matriz quadrada ( $X^T X$ ) conduzir a uma matriz singular. A avaliação dessa multicolinearidade foi realizada com base nas correlações de Pearson entre os elementos do circuito equivalente usado na determinação do ferrocianeto, como exhibe a matriz de correlação (Tabela 1 ). Os valores em negrito destacam as correlações estatisticamente significativas.

Tabela 1 - Matriz com os valores de correlação entre os quatro elementos do circuito equivalente ajustado para os dados do ferrocianeto.

| Elementos | $R_1$       | $CPE_1$ | $R_2$        | $CPE_2$ |
|-----------|-------------|---------|--------------|---------|
| $R_1$     | 1,00        | -       | -            | -       |
| $CPE_1$   | <b>0,81</b> | 1,00    | -            | -       |
| $R_2$     | 0,54        | -0,54   | 1,00         | -       |
| $CPE_2$   | 0,66        | -0,67   | <b>-0,95</b> | 1,00    |

Na Tabela 1 observa-se a existência de correlação positiva entre  $R_1$  e  $CPE_1$ , e negativa entre  $R_2$  e  $CPE_2$ . Partindo do princípio que os valores ajustados para o elemento  $R_1$  deveriam idealmente serem constantes, como foi discutido anteriormente, esperava-se que a correlação

entre  $R_1$  e  $CPE_1$  fosse a mínima possível. Todavia isso não foi observado e essa correlação possivelmente foi fruto de ajuste matemático. O conhecimento sobre o sistema investigado já seria suficiente para descartar o elemento  $R_1$ , porém, supondo-se não ter esse conhecimento, e com a finalidade de avaliar as outras ferramentas empregadas, decidiu-se mantê-lo na avaliação. O elemento  $CPE_2$ , que modela a impedância faradaica, exibiu correlação negativa forte com  $R_2$ , elemento que modela a transferência de carga. Isso significa que à medida que a transferência de massa aumentava, impulsionada pelo aumento da concentração do ferrocianeto, a resistência a transferência de carga diminuía.

A correlação forte entre  $R_2$  e  $CPE_2$  indica que somente um desses elementos deverá compor o modelo MLR, pois os dois elementos estão descrevendo as mesmas informações. Entretanto, os dois elementos serão mantidos e avaliados pela outra ferramenta, assim como  $R_1$ . A Tabela 2 exibe três modelos de regressão construídos para quantificar o ferrocianeto, com seus coeficientes de regressão e suas respectivas significâncias estatísticas. O modelo I foi construído com todos os elementos do circuito equivalente,  $CPE_1$  e  $R_2$  foram usados na construção do modelo II, e o elemento que apresentou maior correlação com a concentração ( $CPE_2$ ) do ferrocianeto foi usado na construção do modelo III (univariado).

Tabela 2 – Modelos de regressão usados para quantificação de ferrocianeto (\* $\gamma = 95\%$ ).

| <b>Modelo</b> | <b><math>b_0</math> (<math>t_0</math>)</b> | <b><math>b_{R1}</math> (<math>t_0</math>)</b> | <b><math>b_{R2}</math> (<math>t_0</math>)</b> | <b><math>b_{CPE1}</math> (<math>t_0</math>)</b> | <b><math>b_{CPE2}</math> (<math>t_0</math>)</b> | <b><math>t_c</math> (gl)*</b> |
|---------------|--------------------------------------------|-----------------------------------------------|-----------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------|
| I             | 140 (408)                                  | -1,53 (-2,02)                                 | -3,64 (-2,44)                                 | 3,14 ( <b>4,69</b> )                            | 72,6 ( <b>39,1</b> )                            | 2,77 (4)                      |
| II            | 133 (26,1)                                 | -                                             | -60,2 ( <b>-7,89</b> )                        | 17,9 ( <b>2,53</b> )                            | -                                               | 2,45 (6)                      |
| III           | 141 (178)                                  | -                                             | -                                             | -                                               | 79,8 ( <b>65,2</b> )                            | 2,36 (7)                      |

No modelo I, observa-se que somente os coeficientes de regressão ( $b$ ) relacionados a média ( $b_0$ ) e aos elementos  $CPE_1$  e  $CPE_2$  apresentaram t-calculados ( $t_0$ ) maiores que o t-crítico ( $t_c$ ), ou seja, eles são os únicos que contribuem de forma estatisticamente significativa para o modelo construído para a quantificação do ferrocianeto. Isso significa que os demais elementos podem ser excluídos do modelo sem prejuízo para quantificação, especialmente o elemento  $R_1$ . Em outras palavras, a explicação dada pelo modelo I deve-se exclusivamente aos elementos  $CPE_1$  e  $CPE_2$ . Apesar de o elemento  $R_2$  estar fortemente relacionado com a concentração, como foi verificado anteriormente, ele não apresentou contribuição estatisticamente significativa ao modelo I, tendo em vista que a variância explicada por ele é a mesma já explicada pelo elemento  $CPE_2$ , pois ambos são colineares. O modelo II foi construído para verificar essa afirmação e por

isso o elemento  $CPE_2$  foi trocado pelo elemento  $R_2$ , e como esperado, sua contribuição foi estatisticamente significativa. O modelo II apresentou o  $F_{\text{calculado}}$  superior ao  $F_{\text{crítico}}$ .

Quanto aos parâmetros de desempenho, o modelo II explicou 92% da variância dos dados ( $R^2=0,92$ ) e alcançou um RMSECV de  $14,8 \mu\text{mol L}^{-1}$ , enquanto o modelo I ( $b_0$ ,  $b_{CPE1}$  e  $b_{CPE2}$ ) forneceu um RMSECV de  $1,34 \mu\text{mol L}^{-1}$  – dez vezes menor que o obtido pelo modelo II – e explicou 99% da variância dos dados ( $R^2=0,99$ ). O modelo III foi construído com o elemento mais correlacionado a concentração de ferrocianeto ( $CPE_2$ ). Ele foi estatisticamente significativo, explicou 99% da variância dos dados e forneceu um RMSECV de  $2,69 \mu\text{mol L}^{-1}$ , duas vezes maior que o modelo I ( $\beta_0$ ,  $\beta_{CPE1}$  e  $\beta_{CPE2}$ ). Os modelo I e III apresentaram RMSECVs de  $0,53$  e  $2,09 \mu\text{mol L}^{-1}$ , respectivamente.

### 2.5.2 Quantificação de catecol e hidroquinona

O eletrodo de pasta de nanotubos de carbono foi usado na determinação de catecol e hidroquinona empregando as faixas de concentração de  $35,0 - 150,0 \mu\text{mol L}^{-1}$  e  $3,0 - 15,0 \mu\text{mol L}^{-1}$ , respectivamente. A Figura 8 exibe os diagramas de Nyquist referente à determinação de catecol e hidroquinona, e os circuitos elétricos que se ajustaram melhor aos espectros de calibração e aos espectros das amostras de água de torneira fortificada empregadas na avaliação dos modelos. O circuito equivalente usado na determinação de catecol (Figura 8a) era composto pelos elementos  $R_1$ ,  $R_2$ ,  $CPE_1$ ,  $CPE_2$  e  $W$  que modelavam, respectivamente, à: resistência do eletrólito, impedância da transferência de carga, impedância da dupla camada elétrica, impedância capacitiva e à impedância do processo de difusão do catecol. A descrição dos elementos do circuito equivalente ajustado aos dados da hidroquinona é a mesma.

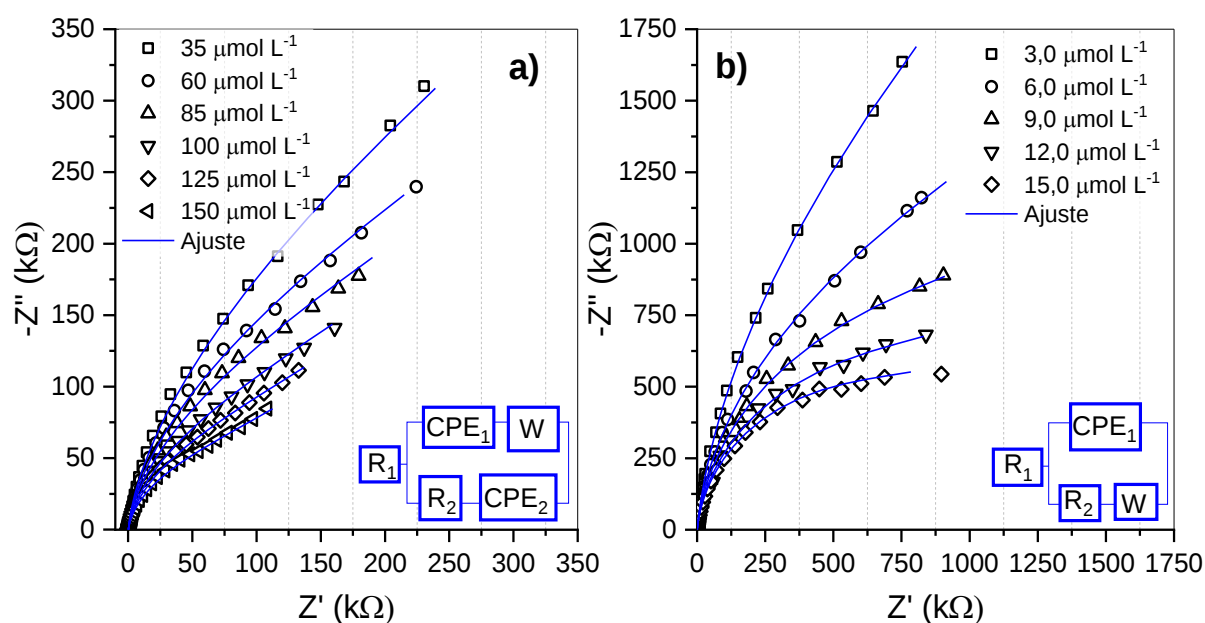


Figura 8 - Diagramas de Nyquist referentes a aquisição de espectros de impedância em concentrações diferentes de (a) catecol e (b) hidroquinona.

A Figura 9 exibe o comportamento dos valores normalizados de cada elemento do circuito equivalente ajustado aos espectros de impedância obtidos em concentrações diferentes de catecol (Figura 9a) e hidroquinona (Figura 9b). O elemento  $R_1$  não participou das avaliações da hidroquinona e catecol, sendo removido em virtude de estar modelando a resistência da solução, fenômeno que independe da concentração do analito dentro do desenho experimental adotado, como já foi mencionado anteriormente. Na Figura 9a, observa-se que o aumento da concentração de catecol provoca o aumento dos valores de  $CPE_1$  e  $CPE_2$ , diminuição dos valores de  $R_2$  e um comportamento aleatório assumido por  $W$ .  $R_2$ ,  $CPE_1$  e  $CPE_2$  exibiram correlação com a concentração de catecol, apresentando coeficientes de correlação de -0,91, 0,94 e 0,96, respectivamente. Por outro lado,  $W$  apresentou uma correlação sem significância estatística, pois o valor de seu coeficiente (-0,43) foi menor que o valor crítico de 0,81.

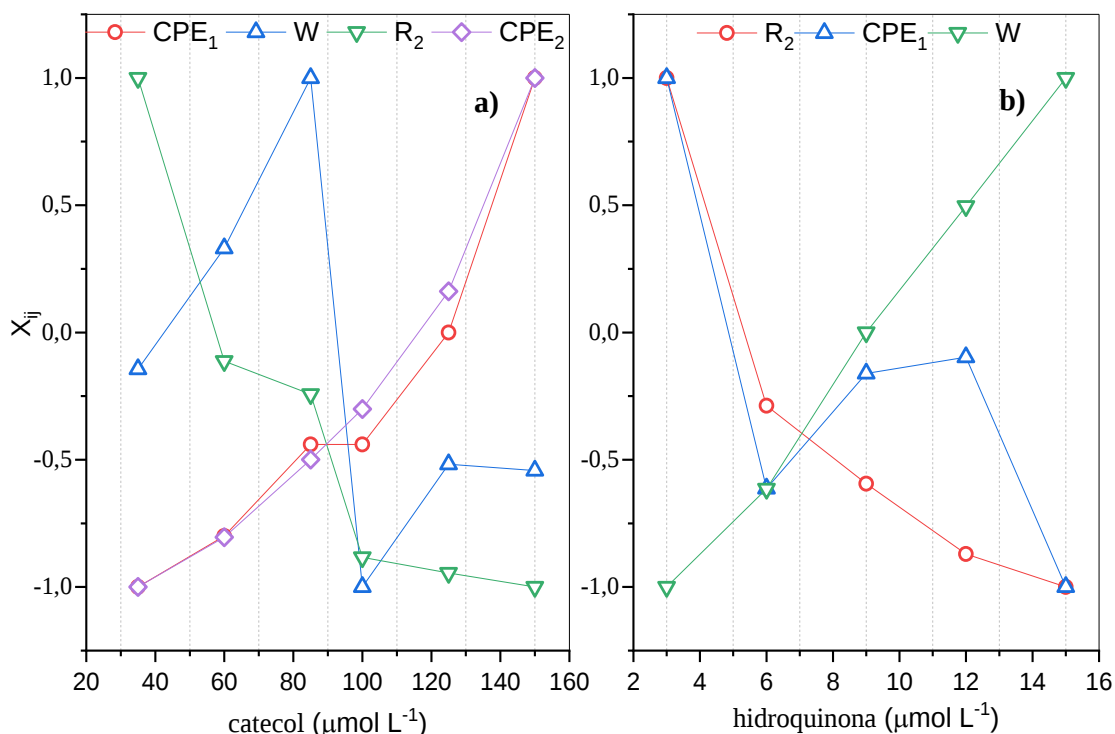


Figura 9 - Mudanças nos valores normalizados dos elementos elétricos do circuito equivalente em função da variação da concentração de (a) catecol e (b) hidroquinona.

Na Figura 9b observa-se que o aumento da concentração de hidroquinona provocou diminuição dos valores de  $R_2$  e aumento dos valores de  $W$ .  $R_2$  e  $W$  exibiram correlação estatisticamente significativa com a concentração de hidroquinona, exibindo coeficiente de correlação de -0,90 e 0,99, respectivamente. Os valores de  $CPE_1$  exibiram comportamento aleatório com o aumento da concentração de hidroquinona, exibindo coeficiente de correlação (0,73) inferior ao valor crítico a 95% de confiança (0,87), o que indica a não existência de correlação com a concentração de hidroquinona.

A Tabela 3 exibe os coeficientes de correlação resultantes da correlação entre os elementos dos circuitos ajustados aos espectros do catecol e hidroquinona. Os valores destacados em **negrito** apresentaram significância estatística visto que suas correlações foram superiores aos valores críticos do catecol (0,81) e hidroquinona (0,88) a 95% de confiança. Do circuito ajustado aos espectros do catecol, verificou-se que a correlação entre  $CPE_1$  e  $CPE_2$  é a única estatisticamente significativa. Essa correlação não indica que a capacitância da dupla camada aumenta por causa do aumento da impedância capacitiva, e sim porque a direção da variação dos valores desses elementos em resposta à variação da concentração do catecol é a mesma. Os elementos  $CPE_1$  e  $CPE_2$  também exibiram coeficientes de correlação negativos com o elemento  $R_2$ . Neste caso as correlações não foram estatisticamente significativas, indicando

que as chances de introdução de instabilidades no cálculo da matriz inversa, durante a construção do modelo MLR são menores. Assim, a correlação entre W e R<sub>2</sub> foi a única significativa entre os elementos do circuito ajustado para hidroquinona, embora R<sub>2</sub> e CPE<sub>1</sub> também tenha apresentado um valor de correlação considerado alto (0,85), porém abaixo do valor crítico (0,87).

Tabela 3 - Matrizes de correlações dos elementos dos circuitos equivalentes ajustados aos espectros do catecol e hidroquinona.

| Catecol          |                  |       |                |                  | Hidroquinona     |                |                  |      |
|------------------|------------------|-------|----------------|------------------|------------------|----------------|------------------|------|
| Elementos        | CPE <sub>1</sub> | W     | R <sub>2</sub> | CPE <sub>2</sub> | Elementos        | R <sub>2</sub> | CPE <sub>1</sub> | W    |
| CPE <sub>1</sub> | 1,00             | -     | -              | -                | R <sub>2</sub>   | 1,00           | -                | -    |
| W                | -0,35            | 1,00  | -              | -                | CPE <sub>1</sub> | 0,85           | 1,00             | -    |
| R <sub>2</sub>   | -0,75            | 0,43  | 1,00           | -                | W                | <b>-0,88</b>   | -0,69            | 1,00 |
| CPE <sub>2</sub> | <b>0,99</b>      | -0,44 | -0,78          | 1,00             |                  |                |                  |      |

Três modelos foram construídos para a determinação do catecol, e seus coeficientes de regressão e significância estatística são apresentadas na Tabela 4. O modelo I foi construído com todos os elementos elétricos do circuito equivalente ajustado aos espectros do catecol, exceto R<sub>1</sub>. Observa-se neste modelo que somente o coeficiente de regressão relativo à média (b<sub>0</sub>) apresentou t<sub>0</sub> maior que t<sub>c</sub>. O F<sub>calculado</sub> para esse modelo foi 62,87, abaixo do F<sub>tabelado</sub> (225), indicando que ele não apresentava significância estatística na descrição da variação do catecol.

Tabela 4 - Coeficientes de regressão dos modelos construídos para quantificar catecol. (\*γ = 95%)

| Modelo | b <sub>0</sub> (t <sub>0</sub> ) | b <sub>R2</sub> (t <sub>0</sub> ) | b <sub>CPE1</sub> (t <sub>0</sub> ) | b <sub>CPE2</sub> (t <sub>0</sub> ) | b <sub>w</sub> (t <sub>0</sub> ) | T <sub>c</sub> (gl)* |
|--------|----------------------------------|-----------------------------------|-------------------------------------|-------------------------------------|----------------------------------|----------------------|
| I      | 95,7 (21,8)                      | -15,5 (-1,96)                     | -82,6 (-1,21)                       | 127 (1,71)                          | 9,68 (1,24)                      | 12,7 (1)             |
| II     | 93,5 (24,7)                      | -21,9 (-4,21)                     | -                                   | 37,4 (6,84)                         | -                                | 3,18 (3)             |
| III    | 105 (19,4)                       | -                                 | -                                   | 55,5 (7,19)                         | -                                | 2,77 (4)             |

A inclusão de variáveis com baixa capacidade explicativa, como é o caso do elemento W, e altamente correlacionadas, como CPE<sub>1</sub> e CPE<sub>2</sub>, introduziram instabilidade ao modelo, refletida na baixa capacidade explicativa. Por esse motivo decidiu-se retirar a variável W, e

entre as variáveis  $CPE_1$  e  $CPE_2$ , que apresentaram correlação forte entre si, optou-se pela retirada de  $CPE_1$  em virtude da correlação entre  $CPE_2$  e a concentração do catecol ser maior. O resultado foi o modelo II, que apresentou o  $F_{\text{calculado}}$  (142,8) maior que o  $F_{\text{tabelado}}$  (9,55), coeficientes de regressão com significância estatística, variância explicada acima de 98% e RMSECV de  $11,0 \mu\text{mol L}^{-1}$ .

O modelo III foi construído usando o elemento que apresentou maior correlação com a concentração do catecol ( $CPE_2$ ). Esse modelo univariado apresentou um  $F_{\text{calculado}}$  (51,7) maior que o tabelado (7,71), RMSECV de  $22,3 \mu\text{mol L}^{-1}$  e conseguiu explicar 91% da variância dos dados. A Figura 10 exibe as amostras de calibração do catecol previstas pelos modelos II e III, e os resíduos deixado por cada modelo. Os resíduos deixados pelo modelo II foram menores e exibiram uma distribuição aleatória ao longo do zero, diferentemente do comportamento quadrático apresentado pelos resíduos do modelo III.

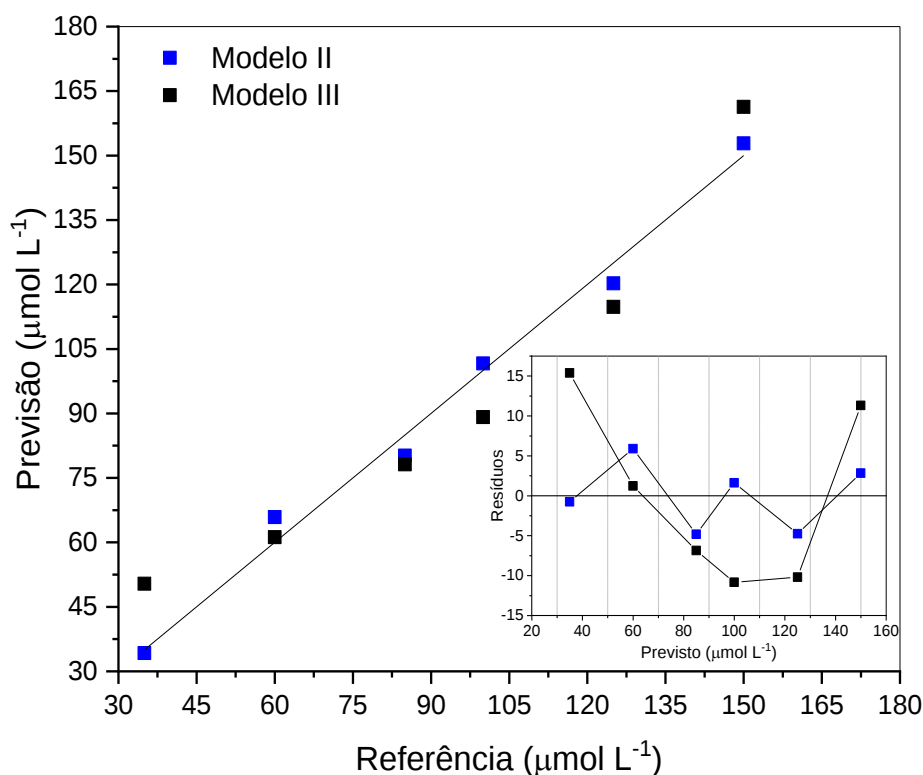


Figura 10 – Concentração das amostras de catecol previstas pelos modelos II e III com destaques para os resíduos deixados pela modelagem.

Os modelos II e III foram usados na análise de amostras de água de torneira fortificadas com catecol. O procedimento empregado nessas amostras fortificadas foi o mesmo: ajustar o mesmo circuito equivalente ajustado as amostras de calibração e empregar os valores dos

elementos no modelo MLR construído. Reiterando a informação mencionada na metodologia, só foram usados os dados dos ajustes que convergiram durante o cálculo do circuito equivalente. A Figura 11 exibe o resultado das amostras de água de torneira fortificadas com catecol previstas pelos modelos II e III. A aplicação das amostras de água de torneira fortificadas no modelo II forneceu RMSEP de  $11,0 \mu\text{mol L}^{-1}$  e recuperações variando de 85,7% a 108,3%. Já no modelo III forneceu RMSEP de  $27,7 \mu\text{mol L}^{-1}$  e recuperações variando de 57,6% a 108,2%. A inclusão do elemento  $R_2$  ao modelo de regressão construído com  $\text{CPE}_2$  originou um modelo (modelo II) com maior capacidade preditiva, reduzindo em 50% os erros de validação cruzada e da previsão das amostras de água de torneira.

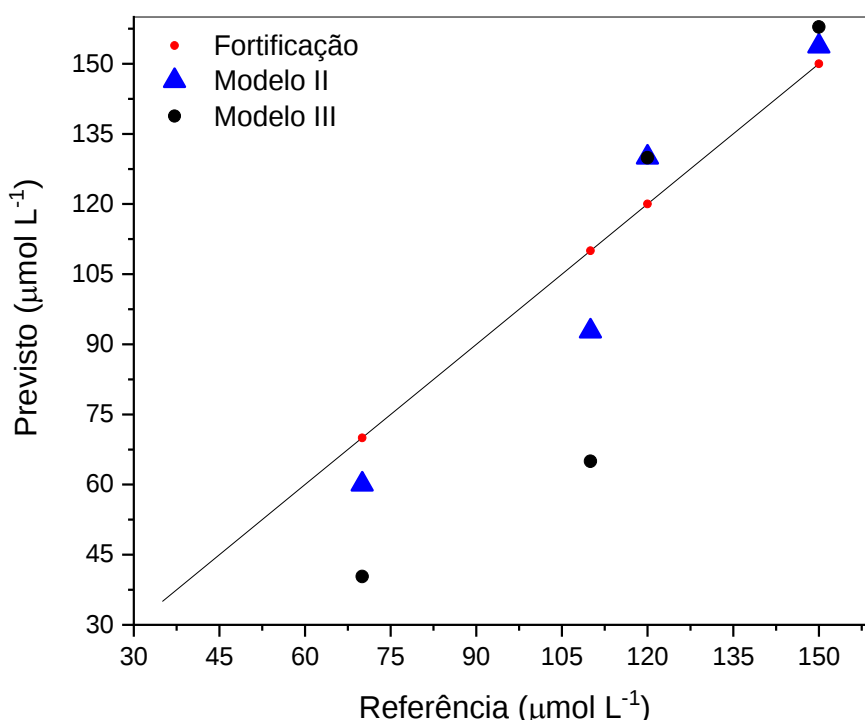


Figura 11 - Concentração das amostras de água de torneira fortificadas com catecol previstas pelos modelos II e III.

A Tabela 5 exibe os coeficientes de regressão dos modelos construídos para quantificar hidroquinona usando os elementos dos circuitos equivalentes ajustados aos espectros de impedância. O modelo I foi construído com todos os elementos, à exceção de  $R_1$ . Observa-se que  $W$  foi o único elemento que exibiu significância estatística na explicação da variação de concentração de hidroquinona, e por isso foi o único elemento usado na construção do modelo II. O resultado desse modelo univariado foi uma explicação superior a 99% da variância e um RMSECV de  $0,43 \mu\text{mol L}^{-1}$ .

Tabela 5 – Coeficientes de regressão dos modelos construídos para quantificar hidroquinona

| Modelo | $b_0 (t_0)$ | $b_{R2} (t_0)$ | $b_W (t_0)$ | $b_{CPE1} (t_0)$ | $t_c (gl)$ |
|--------|-------------|----------------|-------------|------------------|------------|
| I      | 9,00 (61,1) | -0,15 (-0,38)  | 5,46 (19,3) | -0,40 (-1,49)    | 12,7 (1)   |
| II     | 9,14 (63,2) | -              | 5,84 (29,3) | -                | 3,18 (3)   |

O modelo II foi aplicado à análise das amostras de água de torneira fortificadas com hidroquinona e o resultado foi uma recuperação variando entre 71% a 88% e um RMSEP de  $2,51 \mu\text{mol L}^{-1}$ , que está alto quando comparado a menor concentração da curva ( $3,0 \mu\text{mol L}^{-1}$ ). A Figura 12 exibe os valores das amostras de calibração obtidos por validação cruzada e os valores das previsões das amostras de água de torneira fortificadas com hidroquinona, ambas usando o modelo II. Observa-se que os valores previstos pelo modelo II nas amostras de água de torneira fortificadas se encontram abaixo dos valores de referência. Este comportamento é um forte indício da presença de efeito de matriz, caracterizada pela mudança na inclinação.

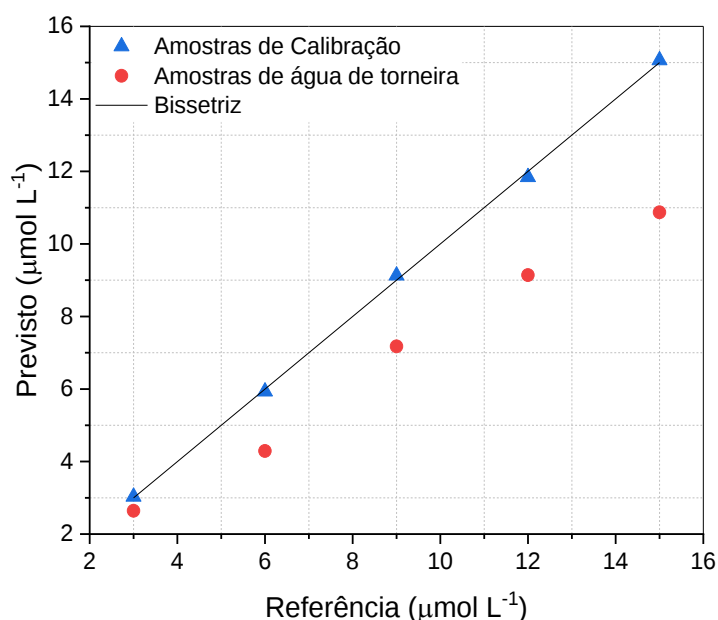


Figura 12 - Previsão das amostras de calibração e amostras de água de torneira fortificadas com hidroquinona usando o modelo II.

O elemento  $CPE_1$  foi adicionado ao modelo II conduzindo a uma melhora significativa da previsão das amostras de água de torneira, como pode ser verificado pelo RMSEP,  $0,79 \mu\text{mol L}^{-1}$ , e recuperações, que variaram de 94% a 136%. Nenhuma melhora foi verificada com a

adição do elemento  $R_2$ . A Figura 13 exibe os valores das amostras de calibração e os valores das previsões das amostras de água de torneira feitas pelo modelo construído com os elementos  $W$  e  $CPE_1$ . A avaliação dos elementos do circuito ajustado aos dados da hidroquinona indicou um forte correlação entre  $R_2$  e  $W$ . Por outro lado, os valores dos coeficientes de correlação desses elementos com a concentração expressaram que esses fenômenos são extremamente dependentes da mesma, e como se sabe,  $R_2$  modela a transferência de carga e  $W$  a transferência de massa. O elemento  $CPE_1$  por sua vez não apresentou correlação estatisticamente significativa com  $R_2$  e  $W$ , sugerindo que ele descreve uma variância diferente. Teoricamente, sabe-se que esse elemento modela a impedância da dupla camada elétrica, que depende diretamente do ambiente iônico da vizinhança do eletrodo. Esse ambiente é pouco afetado pela adição de hidroquinona, visto que a concentração da mesma na solução é mais de 6 mil vezes menor que a concentração do eletrólito, sendo menor ainda na interface eletrodo-solução. Contudo, o ambiente iônico pode ser alterado na presença de um efeito de matriz e o elemento  $CPE_1$  seria, teoricamente, o responsável por capturar essa variância. Por esse motivo sua entrada no modelo II melhora significativamente a previsão das amostras de água de torneira fortificadas, mas não contribui semelhantemente na previsão das amostras sintéticas.

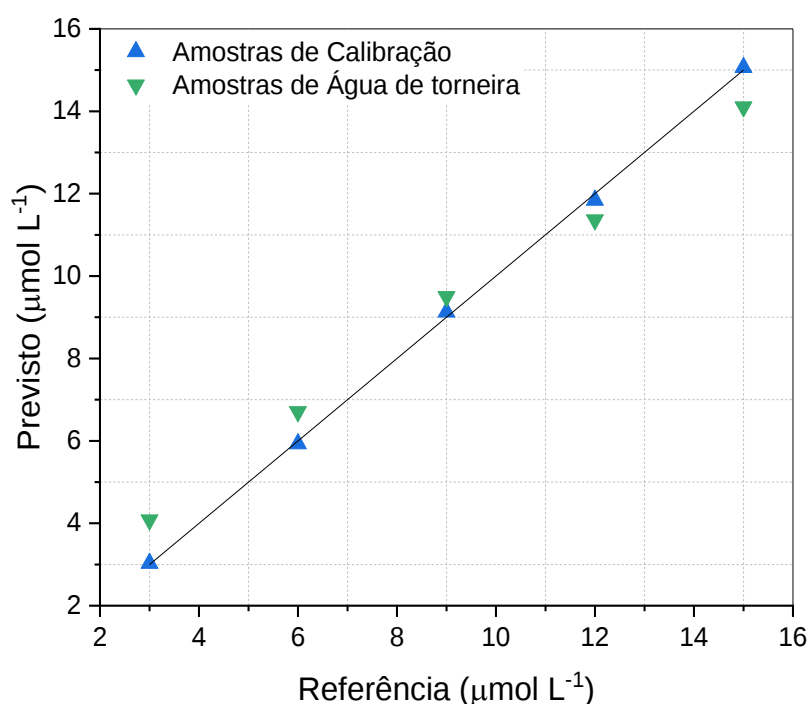


Figura 13 - Previsão das amostras de calibração e amostras de água de torneira fortificadas com hidroquinona usando o modelo construído com os elementos  $W$  e  $CPE_1$ .

## 2.6 Conclusões do Capítulo

- ✧ A construção de modelos quantitativos construídos correlacionando-se a variação da concentração do analito aos valores de mais de um elemento do circuito elétrico equivalente, para os casos avaliados nesse trabalho, conduziram a resultados similares ou melhores que os obtidos com os modelos univariado;
- ✧ Os modelos multivariados demonstraram maior robustez que os modelos univariado diante da previsão de amostras de água de torneira fortificadas, fornecendo menores RMSEPs e melhores recuperações quando comparados aos modelos univariados.

# **CAPÍTULO III – CONSTRUÇÃO DE MODELOS QUANTITATIVOS EMPREGANDO REGRESSÃO POR MÍNIMOS QUADRADOS PARCIAIS**

### 3.1 Introdução

A extração de informações químicas é realizada com base no arranjo de um circuito elétrico ajustado matematicamente para descrever a realidade física da interface. Assim, embora a modelagem de sistemas eletroquímicos reais usando circuito equivalente seja útil, pode-se perder informações importantes sobre a interface visto que na maioria das vezes o processo é uma simplificação dos processos reais (MISZCZYK; DAROWICKI, 2014). MACDONALD (1998) aponta outras diferenças inerentes da modelagem de sistemas eletroquímicos usando um sistema elétrico: i) as interfaces eletroquímicas frequentemente contêm elementos ativos, diferente dos sistemas elétricos compostos por elementos passivos; ii) sistemas eletroquímicos tendem a mudar com o tempo devidos as instabilidades, o que não acontece nos sistemas elétricos passivos; iii) o relacionamento entre a tensão e a corrente nos sistemas eletroquímicos são geralmente não lineares (equação de Butler-Volmer), ao contrários dos sistemas elétricos que tendem a linearidade (Lei de Ohm) (MACDONALD et al., 1998). Outro fato importante é a necessidade de ter o conhecimento prévio do comportamento do sistema na hora de escolher o circuito elétrico usado no ajuste.

Apesar de o espectro de impedância ser formado pelas impedâncias medidas nas diferentes frequências avaliadas, e as respostas em cada uma dessas frequência serem independentes entre si, quando o assunto é a construção de modelos quantitativos empregando dados de impedância eletroquímica, a prática comum é a utilização de uma única variável para ser relacionada a concentração da espécie investigada. A maioria das estratégias constroem modelos relacionando os valores de um único elemento do circuito equivalente, ajustado aos dados de impedância, a variação da concentração da espécie na interface eletrodo-solução. Essa é a estratégia mais empregada no desenvolvimento de sensores impedimétricos (NAG et al., 2019; BONANNI et al., 2012; PRODRONIDIS, 2010). Comparada à estratégia que extrai informações do circuito equivalente para construir modelos, existem poucos trabalhos que constroem modelos quantitativos ou qualitativos diretamente a partir dos dados de impedância. Entre essa minoria, uma parte representativa emprega métodos multivariados.

A regressão por mínimos quadrados parciais (PLS) tem sido a escolha quando se opta pela construção de modelos multivariados com dados de impedância. Essa preferência é facilmente justificada quando se conhece as principais vantagens apresentadas pelo PLS, como, por exemplo, a: capacidade de lidar com dados contendo variáveis multicolineares, desvios de linearidade e ruídos (CIAVOLINO; EL-NASSER, 2009). Nessa abordagem, ao invés de

construir o modelo relacionando uma única variável instrumental a concentração, o modelo é construído relacionando uma matriz de dados instrumental (X) a uma matriz de respostas (Y). O objetivo do PLS é prever um conjunto de variáveis de respostas a partir de um conjunto de variáveis preditoras por meio de variáveis latentes. Isso é geralmente conseguido via processo iterativo, onde vetores dos escores (variáveis latentes) das matrizes X e Y são extraídos e usados no cálculo dos coeficientes de regressão do modelo (MOU et al., 2018). Os trabalhos que aplicam o PLS aos dados de impedância geralmente empregam como matrizes de dados instrumentais (X), a parte real da impedância ( $Z'$ ; MONZÓN et al., 2018), a parte imaginária da impedância ( $-Z''$ ; GARCIA-HERNANDEZ et al., 2018), o módulo da impedância ( $|Z|$ ; CONESA et al., 2015) e o deslocamento de fase ( $\phi$ ; MARTÍNEZ-BISBAL et al., 2019), registrados em diferentes frequências nas diferentes amostras.

Em 2007, Geladi e colaboradores introduziram a decomposição em valores singulares (SVD, do inglês “*Singular Value Decomposition*”) na análise de espectros de impedância com os dados escritos na forma de números complexos ( $Z' + jZ''$ ), em vez de utilizar as partes real ( $Z'$ ) e imaginária ( $Z''$ ) como números reais independentes (GELADI et al., 2007). Com essa estratégia eles conseguiram obter informações dos dados originais através das componentes principais extraídas, e mostrar uma melhora na redução da dimensionalidade dos dados quando comparadas a redução de dimensionalidade obtida com os dados na forma de números reais. Eles também propuseram a construção de modelos quantitativos com a estratégia dos números complexos, construindo modelos correlacionados os escores extraídos por SVD a concentração do analito (TESFALIDET et al., 2016; PHAL et al., 2017; KRUPINSKA et al., 2017; PHAL et al., 2018).

Tendo em vista que as metodologias de construção de modelos de classificação e quantificação pesquisadas na literatura exploram apenas uma parte da informação dos dados de impedância coletados, quando optam por trabalhar com uma das matrizes de dados, ou um dos componentes elétricos do circuito equivalente ajustado, que, aliás, já é uma simplificação do sistema eletroquímico investigado; e, diante da abordagem proposta por Geladi e colaboradores, e da potencialidade dos métodos multivariados, o presente trabalho pretende construir modelos de quantificação usando a regressão por mínimos quadrados parciais para números complexos.

## 3.2 Objetivos

### 3.2.1 Objetivos Específicos

1. Avaliar a influência do potencial e amplitudes aplicadas durante o registro dos espectros de impedância em função dos erros de validação cruzada fornecidos pelo modelo, e escolher a melhor condição experimental na quantificação dos analitos investigados;
2. Comparar os resultados fornecidos pelos modelos PLS construídos com os dados de impedância na forma de número complexos ( $Z' + j Z''$ ) aos resultados fornecidos pelos modelos PLS construídos com a parte real da impedância ( $Z'$ ), a parte imaginária da impedância ( $Z''$ ), o módulo da impedância ( $|Z|$ ) e o ângulo de fase ( $\varphi$ );
3. Avaliar o desempenho dos modelos selecionados em função recuperação de amostras de água da torneira fortificadas com catecol ou hidroquinona;

### 3.3 Fundamentação Teórica

#### 3.3.1 Regressão por Mínimos Quadrados Parciais (PLS)

A aplicação do MLR conduz a modelos robustos, desde que os critérios mencionados sejam cumpridos. Contudo, com o aperfeiçoamento tecnológico, o número de variáveis fornecidas pelos instrumentos, por observação, vem crescendo e isso constitui uma limitação à implementação do MLR, visto que o número de observações é geralmente pequeno comparado ao número de variáveis. Outro problema é a forte correlação comumente apresentada por essas variáveis. Esses problemas podem ser superados para o MLR através da aplicação de técnicas de seleção de variáveis ou aplicando uma técnica de análise multivariada mais robusta.

A regressão por mínimos quadrados parciais tem sido uma das técnicas mais usadas quando o assunto é o tratamento de dados de primeira ordem. Ela possui as vantagens de conseguir trabalhar com dados que contêm muitas variáveis, presença de multicolinearidade, desvios de linearidade e ruídos (WOLD, 2001). Além disso, apresenta a vantagem de primeira ordem, que é a capacidade de realizar a previsão na presença de interferentes, desde que eles também estejam presentes durante a calibração do modelo (VALDERRAMA et al., 2009).

Diferentemente do MLR que constrói o modelo que relaciona X e Y de maneira direta, no PLS, o modelo só é construído após a realização de etapas que permitirão a eliminação de informações desnecessárias das matrizes X e Y. Durante a modelagem são calculadas as variáveis latentes ( $T_n P_n'$ ) que resumirão as variâncias contidas nas matrizes X e Y, conforme equações (22) e (23).

$$X = T_1 P_1' + T_2 P_2' + T_n P_n' \dots + E \quad (22)$$

$$Y = T_1 C_1' + T_2 C_2' + T_n C_n' \dots + F \quad (23)$$

T é a matriz que guarda os escores de X calculados para as n variáveis, P e C são matrizes que guardam os valores dos pesos referentes as matrizes X e Y, respectivamente, e as matrizes E e F guardam a variância não explicada das matrizes X e Y, respectivamente. O algoritmo de NIPALS (“*Nonlinear Iterative Partial Least Square*”) é um dos métodos usados para realizar o cálculo do PLS (WOLD, 2001). Ele, basicamente, calcula a primeira variável latente e extrai essa variabilidade ( $T_1 P_1'$ ) da matriz original, restando o resíduo não explicado

pela primeira variável latente ( $E_1$ ), conforme equação (24). Em seguida, a segunda variável latente é calculada tomando como base a matriz contendo a variância não explicada pela primeira variável latente, ao invés da matriz  $X$ , como descrito na equação Equação 19. Esse procedimento também é realizado na matriz  $Y$ , e continua até atingir o critério de convergência.

$$E_1 = X - T_1 P_1' \quad (24)$$

$$E_2 = E_1 - T_2 P_2' \quad (25)$$

A matriz com os escores de  $X$  ( $T$ ) são preditores de  $Y$ , sendo a mesma estimada através da combinação linear das variáveis originais de  $X$  com coeficientes  $W$ , conforme descreve a equação Equação 20.

$$T = XW^* \quad (26)$$

Substituindo a equação (26) na equação (23), chega-se a uma expressão geral igual a relação do MLR, que relaciona  $X$  e  $Y$ , como está descrito na equação (27).

$$Y = TC + F = XW * C + F = XB + F \quad (27)$$

### 3.3.2 Trabalhos que aplicam PLS a dados de EIE

Martínez-Bisbal e colaboradores monitoraram o processo de degradação de microalgas após o estágio de concentração no final do processo de produção. Esse monitoramento foi realizado com duas agulhas duplas de aço inoxidável, posicionadas em paralelos, aplicando uma tensão alternada com frequência variando de 100 MHz a 10 MHz. O modelo PLS foi construído relacionando os dados da fase ( $\varphi$ ) em função do tempo de produção (dias) (MARTÍNEZ-BISBAL et al., 2019).

Cortina-Puig e colaboradores recobriram um micro-eletrodo de ouro com uma camada de polipirrol com ionóforos e usaram o mesmo na quantificação de íons sódio, potássio e amônio. Os dados da impedância imaginária ( $Z''$ ) foram relacionadas as concentrações dos três

analitos, separadamente, usando o PLS1, chegando a resultados satisfatórios nos modelos para o sódio e potássio (CORTINA-PUIG et al., 2007).

Conesa e colaboradores usaram um eletrodo composto por duas agulhas duplas de aço inoxidável para quantificar o nível de açúcares fermentáveis presentes na segunda fase enzimática da produção de bioetanol de abacaxi de segunda geração. Foram construídos modelos PLS1 relacionando a concentração de glicose, frutose e sacarose ao módulo da impedância ( $Z$ ) e a fase ( $\varphi$ ), separadamente (CONESA et al., 2015). Outro exemplo de abordagem empregando uma estratégia semelhante usada pelo mesmo grupo (CONESA et al., 2016).

Ulrich e colaboradores desenvolveram uma metodologia para quantificar o pH e a concentração de fluído de corte em águas residuais de indústrias. Eles usaram o módulo da impedância ( $Z$ ) e a fase ( $\varphi$ ) para construir modelos PLS que possibilitaram o monitoramento (ULRICH et al., 2012).

Monzón e colaboradores determinaram a taxa de corrosão de concreto reforçado através de modelos PLS, como alternativa aos métodos polarização linear, teste do pulso e circuito equivalente. Os modelos PLS1 foram construídos com o módulo da impedância ( $Z$ ), impedância real ( $Z'$ ) e impedância imaginária ( $Z''$ ) (MONZÓN et al., 2018).

Garcia-Hernandez e colaboradores desenvolveram um multisensor composto por três sensores cujo substrato era um eletrodo à base de óxido de índio-estanho (ITO) recoberto com poli(3,4-etilenodioxitiofeno) poli estireno sulfonato (PEDOT:PSS). Em um dos três sensores foi depositada uma camada eletrocatalítica de nanopartículas de ouro, em outro uma camada eletrocatalítica de bifalocianinato de lutécio (III), e no terceiro não foi feita deposição. Esse multisensor foi usado na análise de vinhos, e através das matrizes dos dados obtidos do espectro de impedância ( $\varphi$ ,  $Z'$  e  $Z''$ ) ou da matriz formada pelos valores dos componentes elétricos do circuito ( $CPE_c$ ,  $CPE_n$ ,  $R_{ct}$ ,  $W_R$ ,  $W_c$  e  $W_N$ ) ajustado aos espectros, foi possível estimar alguns parâmetros (índice de polifenóis totais, acidez,  $SO_2$ , pH e grau alcoólico) de qualidade dos vinhos usando o PLS (GARCIA-HERNANDEZ et al., 2018).

### 3.4 Metodologia

#### 3.4.1 Construção de modelos usando a Regressão por Mínimos Quadrados Parciais

Todos os reagentes e soluções usadas encontram-se descritas na seção reagentes e amostras do capítulo II. O procedimento iniciava-se com a obtenção dos espectros de impedância. Para isso, a concentração da espécie de interesse era variada crescentemente mediante adição de uma alíquota fixa da solução estoque e um espectro era registrado em cada nível de concentração. As espécies avaliadas foram o ferrocianeto de potássio, empregando a faixa de trabalho de 40 a 140  $\mu\text{mol L}^{-1}$ , hidroquinona, empregando a faixa de trabalho de 3,0 a 15,0  $\mu\text{mol L}^{-1}$ , e catecol, empregando a faixa de trabalho de 40 a 220  $\mu\text{mol L}^{-1}$ . Essas análises foram realizadas em tampão acetato 0,1 mol  $\text{L}^{-1}$  com pH 4,66

Os espectros foram registrados conforme descrito no capítulo II. Cada registro fornecia quatro vetores de saída, obtidos em relação a cada frequência aplicada: o vetor de valores da componente real da impedância ( $Z'$ ), o vetor de valores da componente imaginária da impedância ( $Z''$ ), o vetor de valores da fase ( $\varphi$ ) e o vetor de valores do módulo da impedância ( $|Z|$ ). Um único espectro de impedância, registrado em um único nível de concentração, fornecia os quatro vetores descritos acima. O agrupamento desses vetores em função dos diferentes níveis de concentração dava origem as quatro matrizes de dados instrumentais,  $Z'$ ,  $Z''$ ,  $\varphi$  e  $|Z|$ . A matriz dos dados complexos era obtida a partir das matrizes  $Z'$  e  $Z''$ , onde cada elemento dessa matriz era escrito sob a forma  $Z' + jZ''$ .

A regressão por mínimos quadrados parciais foi aplicada às cinco matrizes instrumentais ( $Z' + jZ''$ ,  $Z'$ ,  $Z''$ ,  $|Z|$  e  $\varphi$ ). Os diferentes níveis de concentração da espécie investigada deram origem ao vetor de respostas ( $Y$ ) usado na regressão. A avaliação dos modelos foi realizada com base nos valores do RMSECV e coeficiente de determinação ( $R^2$ ). O mesmo algoritmo empregado no cálculo de modelos PLS usando números reais pode ser empregado no cálculo de modelos PLS usando número complexos, desde que o transposto Hermitiano (transposto conjugado) seja usado durante o cálculo. O cálculo do transposto Hermitiniano é feito no matlab colocando-se o acento agudo (') após a variável que representa a matriz que sofrerá a transposição. A escolha dos números de fatores foi feita com base no critério de Haaland-Thomas empregando um nível de significância de 25% ( $\gamma = 75\%$ ).

### 3.5 Resultados e Discussão

#### 3.5.1 Determinação de ferrocianeto empregando o PLS para números complexos

A primeira etapa do trabalho consistiu na avaliação da aplicabilidade de números complexos na obtenção de modelos PLS com dados de EIE e ao mesmo tempo na avaliação das condições experimentais envolvidas na obtenção dos dados de EIE, especificadamente, o potencial DC aplicado durante o registro dos espectros e a amplitude de oscilação. Essa etapa do trabalho foi desenvolvida com o ferrocianeto, uma espécie que possui reação redox comumente considerada de transferência eletrônica de esfera externa, tendo em vista à rápida transferência de elétrons controlada pelo transporte de massa (KOŠICEK et al., 2016)

Os espectros de impedância do ferrocianeto foram obtidos aplicando-se três potenciais diferentes: o potencial de +0,100 V, no início do pico, denominado de “onset”; o potencial de +0,220 V, correspondente ao potencial do pico de oxidação; e um potencial +0,450 V depois do pico, denominado de sobrepotencial. A amplitude do potencial também foi avaliada durante esse estudo, aplicando-se 1 mV, 5 mV e 10 mV. A faixa de concentração do ferrocianeto foi mantida fixa entre 40-140  $\mu\text{mol L}^{-1}$  durante a avaliação. A Figura 14 apresenta os diagramas de Nyquist e Bode referentes aos dados adquiridos nos três potenciais avaliados aplicando-se amplitude de 10 mV ao eletrodo de carbono vítreo.

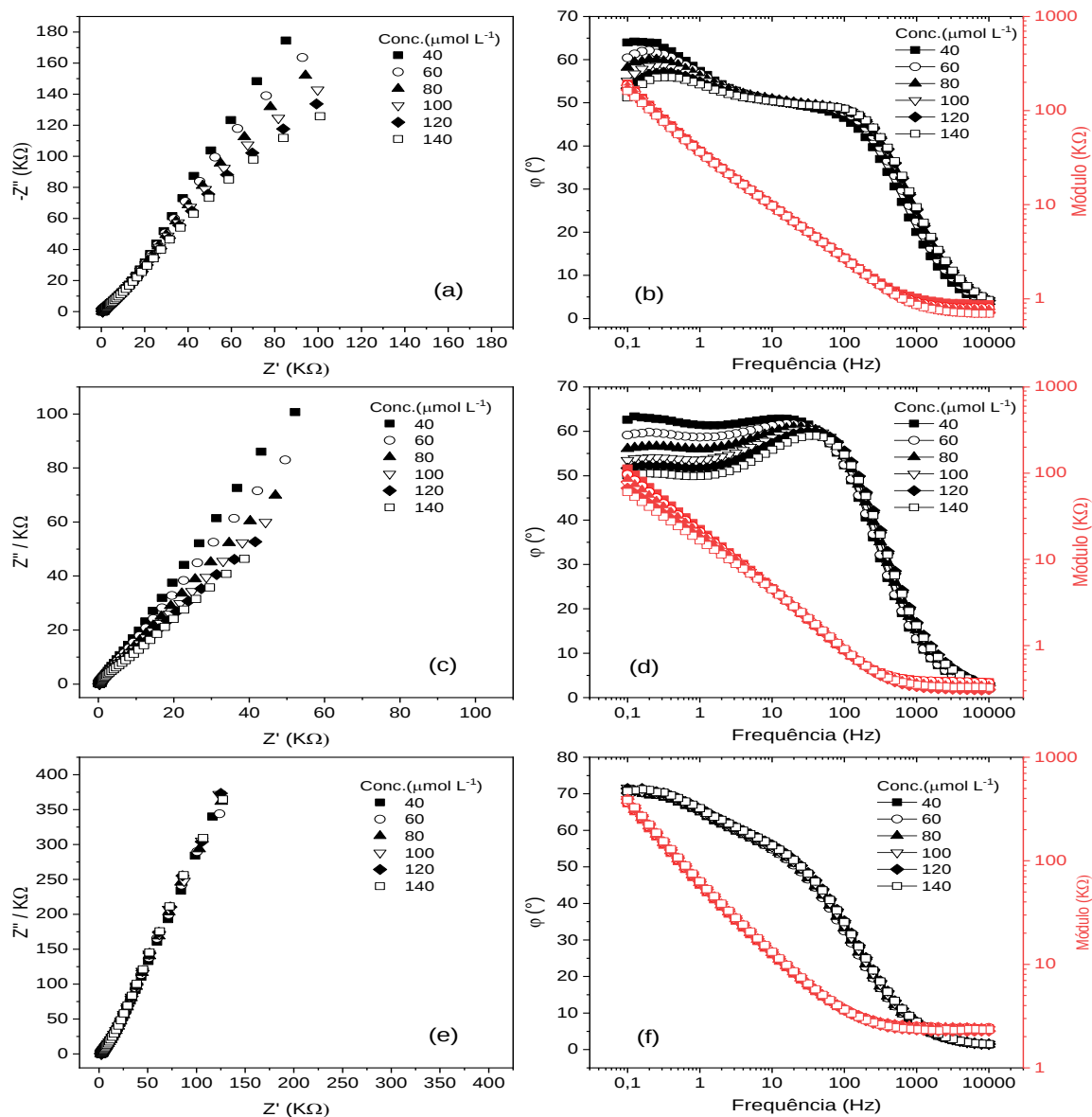


Figura 14 - Diagrama de Nyquist (a,c,e) e bode (b,d,f) referentes aos espectros registrados na determinação de ferrocianeto em eletrodo de carbono vítreo aplicando potenciais diferentes:  $E_a = 0,100$  V (a) e (b),  $E_a = 0,220$  V (c) e (d),  $E_a = 0,450$  V, (e) e (f).

Na Figura 14 é possível observar uma dependência entre o comportamento dos espectros e as situações de transferência de carga e massa resultantes da aplicação de potenciais diferentes. No diagrama de Nyquist, registrados em  $E_a = +0,100$  (Figura 14a), verifica-se a presença de linhas retas com ângulos de  $45^\circ$  no domínio de frequências altas, característico de processo difusional em superfícies rugosas. Essas linhas retas são acompanhadas por um semicírculo mal definido no domínio de baixa frequência, representando uma alta resistência à

transferência de carga. As alterações nos espectros em função da concentração de ferrocianeto ocorrem no domínio de baixa frequência, indicando uma diminuição da resistência à transferência de carga com o aumento da concentração.

Os espectros de impedância registrados em  $E_a = 0,220$  V (Figura 14c) refletem um comportamento controlado por transferência de massa dentro de toda a faixa de frequência. Isso pode ser verificado através das retas com declives que se aproximam de  $45^\circ$  à medida que a concentração do ferrocianeto aumenta. Os espectros registrados em regime de sobrepotencial (Figura 14e),  $E_a = +0,450$  V, permaneceram inalterados com a mudança da concentração de ferrocianeto, exibindo uma impedância capacitiva não ideal.

Os três conjuntos de espectros estão de acordo com os pressupostos esperados para uma reação faradaica influenciada pelo transporte da espécie reagente ao eletrodo (TRAN et al., 2017), ou seja, a impedância de difusão é negligenciável quando a corrente é controlada pela cinética ( $E_a = +0,100$  V); a resistência de transferência de carga é insignificamente pequena para potenciais onde a corrente é limitada por transferência de massa (em  $E_a = +0,220$  V), e no caso de um potencial suficientemente grande ( $E_a = +0,450$  V), a concentração de ferrocianeto na interface pode ser considerada aproximadamente nula quando comparada com o *bulk* da solução.

Os gráficos de Bode (Fig. 1) exibem perfis de magnitude similares para  $E_a = +0,100$  V e  $E_a = +0,450$  V, denotando respostas resistivas ( $\log |Z|$  inclinação  $\approx 0$ ) a capacitivas ( $\log |Z|$  inclinação  $\approx -1$ ), das frequências altas para as baixas. Também não se percebe mudanças significativas nos espectros em função do aumento da concentração de ferrocianeto. O mesmo comportamento é observado das altas as médias frequências em  $E_a = +0,220$  V. Já em frequências mais baixas, verifica-se que as inclinações das relação freq. vs  $\log |Z|$  tendem a diminuir para  $-0,5$  à medida que a concentração de ferrocianeto aumenta. O perfil do ângulo de fase evidencia a mudança de uma resposta resistiva para uma capacitiva em  $E_a = +0,450$  V, comportamento típico de impedância relacionada a resistência de eletrólitos e da dupla camada, sem alterações espectrais em função do aumento da concentração de ferrocianeto. Em  $E_a = +0,220$  V observa-se que o ângulo de fase tende a diminuir para  $45^\circ$  com o aumento da concentração de ferrocianeto, em baixas frequências. Comportamento típico de uma reação faradaica controlada por difusão semi-infinita. Mudanças no ângulo de fase em função da variação da concentrações de ferrocianeto, observadas em  $E_a = +0,220$  V, também ocorreram em  $E_a = +0,100$  V, porém, somente em frequências abaixo de 1 Hz (Figura 14b).

Os dados de impedância (brutos) escritos na forma de números complexos ( $Z' + jZ''$ ), referentes à avaliação dos potenciais e amplitudes exibidos acima, foram compilados em uma matriz de 54 x 51, e submetidos a análise por componentes principais (PCA) para averiguar se a mudança nos espectros causada pela variação dos níveis dos parâmetros, produzia algum padrão (Figura 15). As 54 linhas correspondem às 6 concentrações de ferrocianeto nas 9 combinações dos fatores potencial (0,110V, 0,220V e 0,450 V) e amplitude (1 mV, 5 mV e 10 mV) nos três níveis avaliados, enquanto as 51 colunas correspondem as frequência monitoradas.

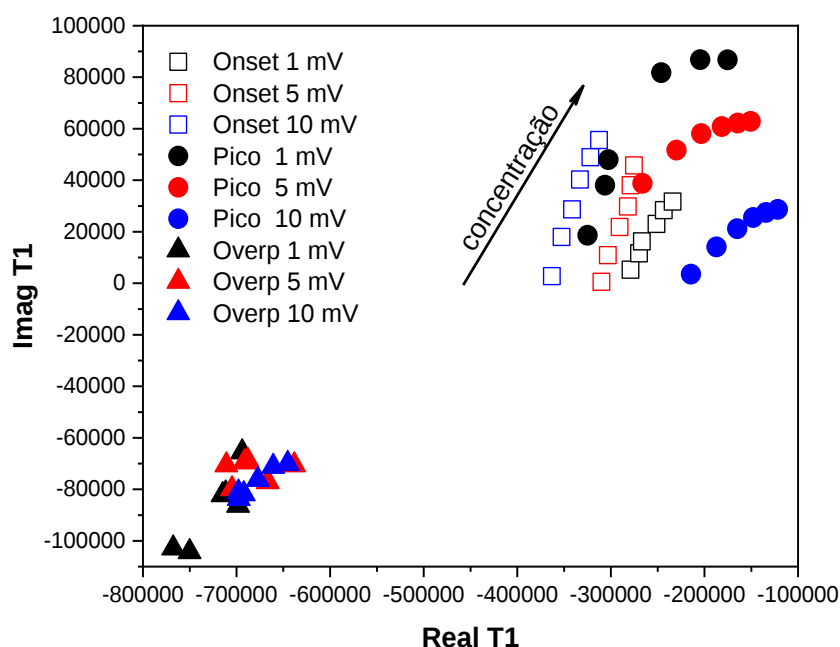


Figura 15 - Escores da primeira componente principal real e imaginária (T1) realizada nos dados complexos obtidos durante a quantificação de ferrocianeto aplicando potenciais e amplitudes diferentes.

As duas primeiras PCs foram capazes de descrever 99,6% e 0,3% da variância dos dados. Essa alta concentração de informações na primeira PC também foi verificada por Geladi e colaboradores (GELADI et al., 2007). O plano com os escores complexos da primeira PC (real vs. imaginária) encontra-se na Figura 15. Observa-se seis padrões de distribuição localizados nos escores positivos da componente imaginária. Os escores da componente imaginária de cada uma dessas distribuições descreveu um comportamento de crescimento, aproximadamente linear, com o aumento da concentração do ferrocianeto. Esses padrões foram apresentados pelos dados registrados aplicando-se o potencial no início do pico

(quadrados/Figura 15) e o potencial de pico (bola /Figura 15). Por outro lado, verificou-se que os dados registrados em regime de sobrepotencial não apresentaram padrão que pudesse ser relacionado a variação da concentração.

A análise por componente principais dos espectros obtidos durante o estudo do potencial aplicado e da amplitude também foi realizada nos dados da parte real da impedância ( $Z'$ ), parte imaginária ( $Z''$ ), módulo da impedância ( $|Z|$ ) e ângulo de fase ( $\varphi$ ) (Figura 16), e comparadas a análise realizada nos dados complexos, exibido na Figura 15.

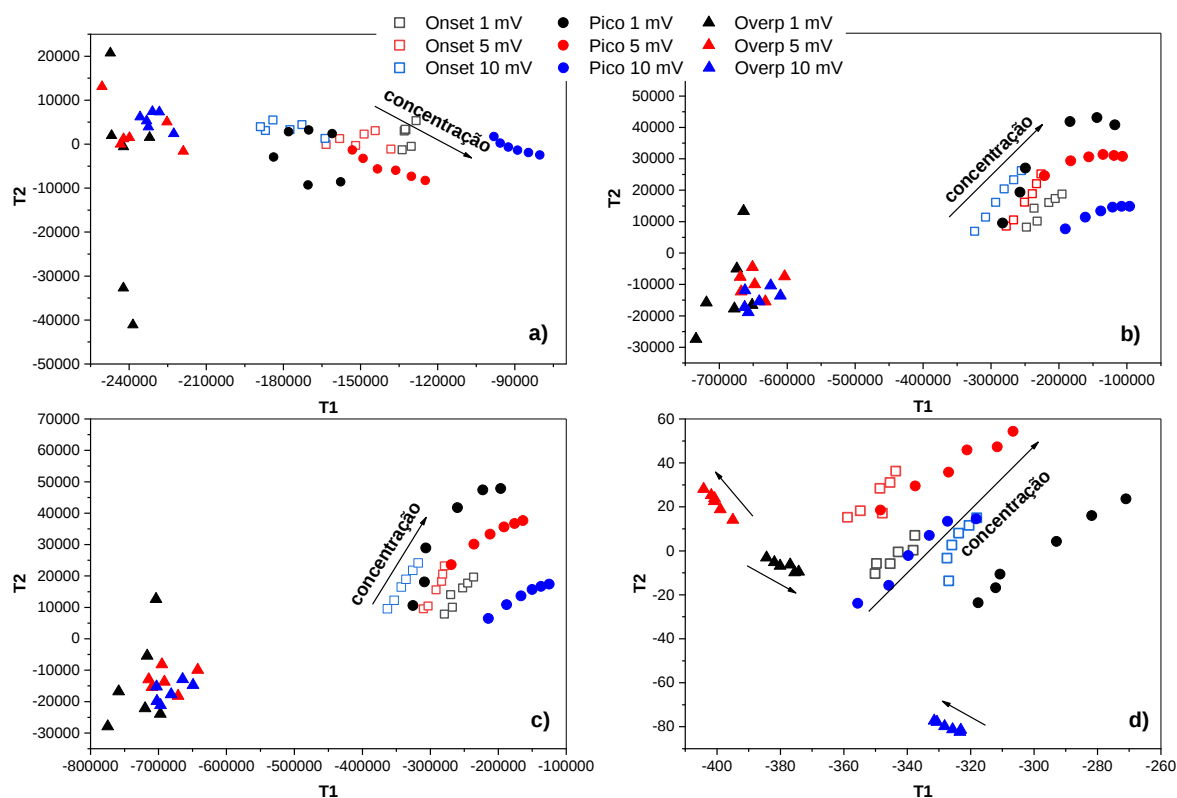


Figura 16 - Escores da primeira (T1) e segunda (T2) componente principal realizada na matriz de dados  $Z'$  (a),  $-Z''$  (b),  $|Z|$  (c) e  $\varphi$  (d) obtidas com os dados da quantificação de ferrocianeto aplicando potenciais e amplitudes diferentes.

As duas primeiras componentes principais da PCA realizada nos quatro conjunto de dados explicaram 99,9% da variância dos dados, semelhante a explicação alcançada com os dados complexos. Verificou-se que apenas os dados registrados sob aplicação do potencial de pico, nas amplitudes de 5 mV e 10 mV, exibiram uma distribuição de escores relacionadas ao aumento da concentração do ferrocianeto (bolas azuis e vermelhas/Figura 16a). Pode-se dizer

que a PCA realizada nos dados da parte real da impedância exibe um padrão resultante dos processos (principalmente a resistência da solução e resistência a transferência de carga) que contribuem para a parte resistiva da impedância (Figura 16a). O aumento de concentração está sendo descrito pela diminuição dos escores da primeira PC (T1), que possivelmente está capturando a variância referente a resistência a transferência de carga, visto que esse processo tende a diminuir com o aumento da concentração da espécie. A maior parte dos escores dos dados descritos pela segunda PC (T2) estão variando entre  $\pm 10.000$  e praticamente não mudam com a variação da concentração da espécie, como pode ser observado quando o escore é projetado sobre a PC. Essa PC possivelmente está descrevendo a variância referente a resistência da solução, pois os escores desse componente varia pouco com a variação da concentração da espécie.

A PCA realizada nos dados da parte imaginária exibe um padrão que é resultado dos processos que contribuem para a reatância, principalmente a capacitiva. A Figura 16b e Figura 16c exibem um padrão similar ao apresentado pela PCA realizada nos dados complexos. Até a direção da distribuição dos escores tem o mesmo sentido, sugerindo que algum efeito capacitivo domina o comportamento. Como o processo é controlado por difusão, é possível que a primeira componente (T1) esteja descrevendo a impedância faradaica, visto que existe discriminação entre os escores em função da concentração nessa componente. A segunda componente possivelmente está descrevendo a impedância da dupla camada elétrica, visto que as projeções nessa componente não conduzem a uma discriminação clara entre os escores referentes aos diferentes níveis de concentração. O padrão dos escores apresentado pelos dados da fase, registrados no pico e antes do pico, foram similares aos apresentados por  $Z'$ ,  $Z''$  e  $|Z|$  (Figura 16d). Contudo, também foi verificado que os escores referentes aos dados registrados em regime de sobrepotencial estavam descrevendo o aumento da concentração, fato que não ocorreu nas outras avaliações. A direção de crescimento dos escores, nos dados da fase registrados em sobrepotencial, foram iguais, mas os dados registrados em 1 mV apresentaram escores com o sentido de crescimento oposto ao sentido apresentado pelos dados registrados em 5 mV e 10 mV.

A próxima etapa foi avaliar os resultados fornecidos pelo PLS construídos com os números complexos e compará-los aos fornecidos pelos modelos PLS construídos com os dados de impedância real ( $Z'$ ), imaginária ( $-Z''$ ), fase ( $\varphi$ ) e o módulo da impedância ( $|Z|$ ). **A Fonte de referência não encontrada.** apresenta os valores do RMSECV e o número de variáveis latentes selecionados para cada modelo usando dados brutos e centrados na média.

Verifica-se que os modelos construídos com os dados gerados na aplicação de sobrepotencial apresentam baixa capacidade preditiva, como pode ser observado pelos RMSECV, que chegam a valores 10 vezes maiores que os fornecidos pelos modelos construídos com os dados registrados sob aplicação do potencial de pico ou no início do pico. Além disso, também se percebe o pequeno número de variáveis latentes selecionadas pelos modelos construídos com os dados obtidos em regime de sobrepotencial. Isso implica que a adição de mais variáveis latentes não aumenta a variância explicada de forma significativa, e consequentemente não proporciona diminuição significativa do RMSECV. O modelo que forneceu o melhor resultado foi construído com os dados complexos gerados sob a aplicação do potencial de pico e amplitude de 10 mV. Os modelos construídos com  $Z'$ ,  $|Z|$  e  $\phi$ , nessa mesma condição experimental, também conduziram a bons resultados.

Tabela 6 - Valores dos RMSECV obtidos com os modelos PLS construídos com dados de impedância gerados na determinação de ferrocianeto aplicando potenciais e amplitudes diferentes ao eletrodo de carbono vítreo. A quantidade de variáveis latentes selecionadas encontra-se entre parênteses. A faixa de concentração de ferrocianeto empregada foi de 40 a 140  $\mu\text{mol L}^{-1}$ .

| Trat. dos dados | Forma dos dados | $E_a = +0.100 \text{ V}$ |         |         | $E_a = +0.220 \text{ V}$ |         |                | $E_a = +0.450 \text{ V}$ |        |        |
|-----------------|-----------------|--------------------------|---------|---------|--------------------------|---------|----------------|--------------------------|--------|--------|
|                 |                 | 1 mV                     | 5 mV    | 10 mV   | 1 mV                     | 5 mV    | 10 mV          | 1 mV                     | 5 mV   | 10 mV  |
| CM*             | $Z'+jZ''$       | 6,2 (1)                  | 3,8 (3) | 4,9 (1) | 9,9 (1)                  | 3,6 (2) | <b>1,1 (2)</b> | 17 (2)                   | 17 (1) | 15 (1) |
|                 | $Z'$            | 18 (3)                   | 3,4 (1) | 11 (2)  | 20 (2)                   | 25 (1)  | <b>1,7 (2)</b> | 44 (1)                   | 19 (1) | 18 (1) |
|                 | $Z''$           | 5,9 (1)                  | 6,1 (2) | 3,7 (1) | 9,7 (1)                  | 5,6 (2) | <b>0,9 (3)</b> | 15 (3)                   | 19 (1) | 15 (1) |
|                 | $ Z $           | 8,1 (1)                  | 12 (1)  | 2,9 (1) | 11 (1)                   | 6,0 (2) | <b>2,0 (2)</b> | 13 (4)                   | 18 (1) | 15 (1) |
|                 | $\phi$          | 4,0 (1)                  | 1,3 (3) | 2,9 (2) | 13 (2)                   | 2,7 (3) | 2,7 (3)        | 28 (1)                   | 27 (1) | 12 (2) |
| Orig. **        | $Z'+jZ''$       | 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) |
|                 | $\phi$          | 5,3 (3)                  | 3,8 (3) | 7,8 (3) | 7,4 (3)                  | 5,6 (4) | 3,5 (4)        | 40 (1)                   | 27 (2) | 14 (3) |

\*Dados centrados na média, \*\* Dados originais.

Uma avaliação estatística dos valores apresentados pela Tabela 6 foi realizada para averiguar se a variação das condições experimentais, pré-tratamento e apresentação dos dados ( $Z' + j Z''$ ,  $Z'$ ,  $-Z''$ ,  $|Z|$  e  $\varphi$ ) eram significativas para o desempenho analítico dos modelos PLS. Assim, uma análise de variância (ANOVA) foi realizada nos valores de RMSECV e número de variáveis latentes considerando os seguintes fatores: i) potencial aplicado, em três níveis: +1 (+0,100 V), 0 (+0,220 V) e -1 (0,450 V); ii) amplitude, em três níveis: -1 (1 mV), 0 (5 mV) e 1 (10 mV); iii) apresentação dos dados, em cinco níveis: -1 ( $Z' + jZ''$ ), -0,5 ( $Z'$ ), 0 ( $Z''$ ), +0,5 ( $|Z|$ ) e +1 ( $\varphi$ ); e pré-tratamento dos dados, em dois níveis: -1 (dados brutos) e +1 (dados centrados na média). A Figura 17 exibe os resultados dessa avaliação através das cartas de Pareto para o RMSECV e para o número de variáveis latentes.

A carta de Pareto é uma visão do teste estatístico t para cada efeito, onde cada barra representa o efeito padronizado – efeito estimado dividido pelo seu erro padrão – que é comparado ao valor t-crítico representado pela linha vertical. Cada barra que se estende além do valor crítico corresponde a efeitos estatisticamente significativos. As barras cinza claro correspondem a efeitos que são positivos na resposta, enquanto as pretas correspondem aos efeitos que são negativos. Ser positivo significa que o aumento do nível do fator avaliado provoca o aumento da resposta, enquanto um efeito negativo significa que o aumento do nível provoca a redução da resposta. Os fatores potencial aplicado, amplitude, apresentação dos dados e pré-tratamento foram representados pelas letras A, B, C e D, respectivamente, além de servir de identificação para os efeitos principais e interações representadas pelas combinações de letras. A Figura 17 também mostra a mudança na resposta média em função da mudança dos níveis dos fatores individuais, enquanto os outros fatores são mantidos constantes nos níveis centrais.

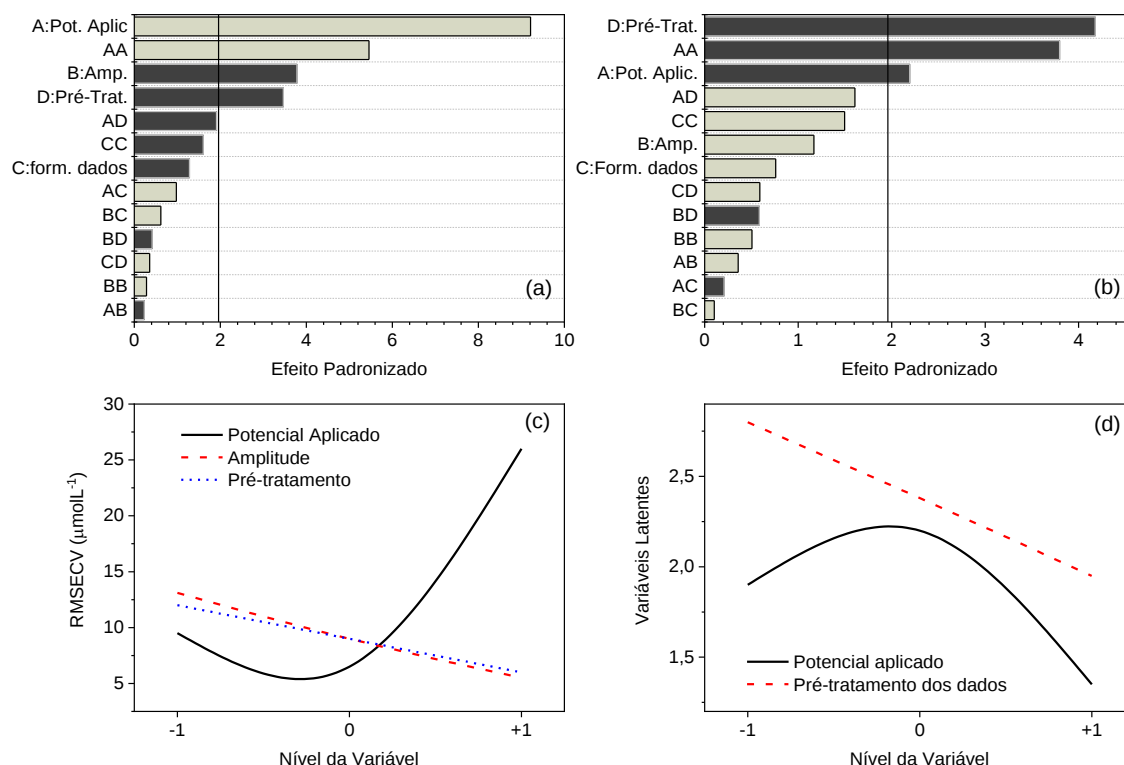


Figura 17 - Cartas de Pareto referente à ANOVA realizada nas respostas RMSECV (a) e número de variáveis latentes (b), onde: A = potencial aplicado; B = amplitude; C = apresentação dos dados; D = pré-tratamento dos dados. Variação das respostas RMSECV (c) e número de variáveis latentes (d) em relação aos efeitos significativos.

A carta de Pareto da Figura 17a exibiu quatro efeitos estatisticamente significativos para o RMSECV: potencial aplicado, amplitude, pré-tratamento dos dados e o efeito de interação quadrático AA. Os efeitos principais amplitude e pré-tratamento foram negativos, enquanto o efeitos do potencial aplicado e da interação AA foram positivos. Isso significa que os menores valores de RMSECV foram obtidos com o nível máximo da amplitude (10 mV) ou pré-tratamento dos espectros (dados centrados na média) (Figura 17c). Por outro lado, valores altos de RMSECV foram obtidos com os dados gerados em regime de sobrepotencial ( $E_a = +0,450$  V). O efeito de interação AA indica o aumento não linear do RMSECV em função da variação do menor para o maior nível do potencial aplicado, com o ponto mínimo correspondendo ao nível intermediário ( $E_a = 0,220$  V).

Quanto ao número de variáveis latentes, verificou-se que três efeitos foram estatisticamente significativos e todos eles apresentaram valores negativos (Figura 17b): os efeitos principais potencial aplicado e pré-tratamento dos dados, e o efeito de interação

quadrática AA. Assim, modelos que usavam poucas variáveis latentes foram construídos com os espectros registrado em regime de sobrepotencial e centrados na média. O efeito de interação AA, na presente situação, indica uma diminuição não linear no número de variáveis latentes selecionadas em função da mudança do nível inferior para o superior do potencial aplicado, alcançando o ponto máximo no nível intermediário. Isso significa que os modelos construídos com os espectros registrados sob a aplicação do potencial de pico tendem a selecionar mais variáveis latentes. Isso porque, nesses dados, a adição de mais variáveis latentes ao modelo produz um aumento significativo da variância explicada, e consequentemente uma redução significativa do RMSECV.

A confirmação do potencial aplicado como variável importante para o RMSECV e número de variáveis latentes reitera a inferência realizada a partir da observação da Tabela 6, que uma perda significativa da capacidade de previsão é observada nos modelos PLS construídos com os dados obtidos em regime de sobrepotencial. Contudo, uma conclusão importante que não pode ser inferida diretamente dessa tabela é que os espectros obtidos no potencial de pico conduzem a modelos com melhores previsões, o que pode estar diretamente associado a maior variação da impedância do espectro em função da mudança de concentração do ferrocianeto (Figura 14). Esse aumento de sensibilidade provavelmente explica o porquê de os espectros obtidos em amplitude de 10 mV fornecerem modelos com menores valores de RMSECV. Quanto aos pré-tratamento dos dados, a centragem na média fornece modelos com melhores capacidades preditivas e poucas variáveis latentes quando comparadas com os modelos construídos com os dados originais.

A Tabela 7 exhibe os valores das previsões do ferrocianeto obtida por validação cruzada do tipo “leave one out” com os modelos PLS construídos com os espectros obtidos no potencial de pico, amplitude de 10 mV e centrados na média. Um teste t pareado foi realizado para comparar os resultados exibidos pelo modelo PLS construído com os dados de natureza complexa aos resultados exibidos pelos demais modelos construídos com dados de natureza real.

Assim como um solvente nivelador atua nivelando a força de ácidos e bases fortes, os resultados similares exibidos na Tabela 7 podem estar relacionados ao comportamento simples da reação redox do ferrocianeto em baixas concentrações. Talvez uma reação redox mais complexa provocasse uma diferenciação significativa nos resultados, mostrando a robustez do modelo complexo, assim como o solvente diferenciador atua diferenciando a força dos ácidos e bases fortes. Mesmo assim, o uso de modelos PLS construídos com números complexos ainda

é vantajoso, visto que não há necessidade de avaliar quais componentes da impedância ( $Z'$ ,  $Z''$ ,  $|Z|$  e  $\varphi$ ) está mais relacionado a variação da concentração do analito.

Tabela 7 - Valores das previsões das concentrações de ferrocianeto realizados com os dados de impedância registrados durante a aplicação do potencial de pico e amplitude de 10 mV sobre eletrodo de carbono vítreo. Os valores dos RMSECV e número de variáveis latentes usadas no modelo encontram-se entre parênteses.

| valores de referência<br>( $\mu\text{mol L}^{-1}$ ) | $Z' + jZ''$<br>(1.1; 2) | $Z'$<br>(1.7; 2) | $-Z''$<br>(0.9; 5) | $ Z $<br>(2.0; 2) | $\varphi$<br>(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                   |

Os coeficientes do modelo PLS são calculados relacionando-se a matriz de dados instrumental (X) à matriz com os parâmetros experimentais (Y). Tendo em vista que a matriz instrumental, na abordagem proposta, é composta por elementos complexos, o modelo gerado também possui coeficientes complexos, e, conseqüentemente, os valores dos parâmetros previstos serão números complexos. Todavia, a concentração é uma grandeza real e fisicamente não tem sentido tentar interpretá-la como um número complexo. Por outro lado, do ponto de vista formal, sabe-se que os conjunto dos números reais estão contidos dentro do conjunto dos números complexos (AGUILAR; DIAS, 2015). Em outras palavras, todo número real é um número complexo cuja parte imaginária é nula. Isso significa que as concentrações da matriz Y também são números complexos, cuja parte imaginária é nula. Assim, idealmente, durante a modelagem espera-se que a parte real da concentração prevista seja igual a concentração analítica da espécie e a parte imaginária seja nula.

A Figura 18 exhibe o comportamento da parte real e imaginária da concentração prevista em função do número de variáveis latentes usadas no modelo PLS construídos com os dados complexos registrados no potencial de pico ( $E_a = 0,220$  V) e amplitude de 10 mV.

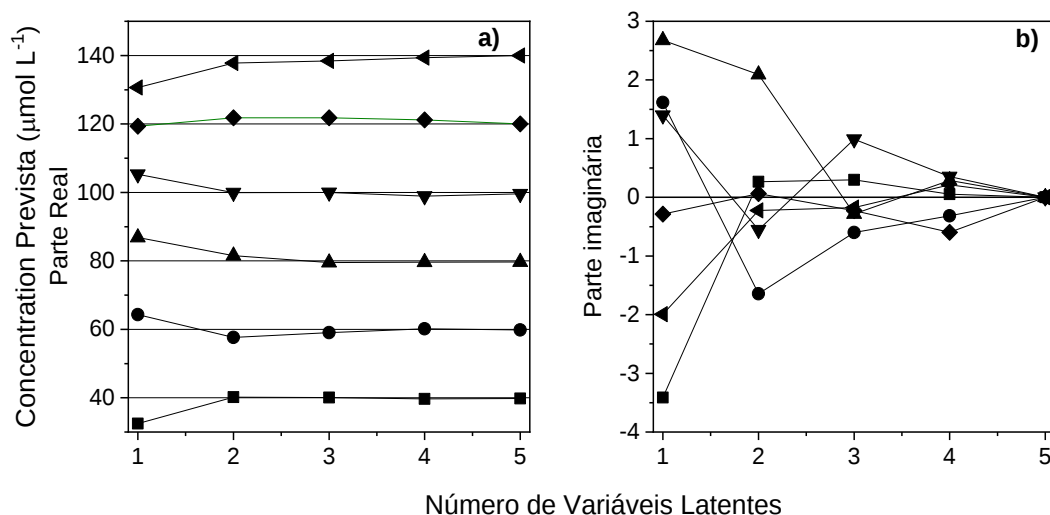


Figura 18 - Comportamento das partes real e imaginária dos valores previstos com o modelo complexo em função do número de variáveis latentes.

As partes real e imaginária da concentração apresentaram o comportamento esperado. Verificou-se que a concentração prevista, representada pela parte real da concentração complexa (Figura 18a), se aproximou da concentração analítica à medida que as variáveis latentes foram adicionadas. Por outro lado, a parte imaginária da concentração tendeu a zero com a adição de variáveis latentes (Figura 18b). Por esse motivo, a parte real do RMSECV, que também é um número complexo, é usada durante o cálculo da escolha do número de variáveis latentes (critério de Haaland-Thomas). O número complexo não é usado no cálculo em virtude da realização do teste F durante o cálculo da seleção do número de fatores. Não teria sentido aplicar números complexos em testes estatístico desenvolvido com base em números reais. A escolha do número de fatores usando a parte imaginária poder ser possível, mas nesse caso, possivelmente, outro critério dever ser usado. O uso do critério de Haaland-Thomas no cálculo do número de fatores, usando como base o valor da parte imaginária do RMSECV, só seria válido quando a convergência das partes real e imaginária ocorressem com o mesmo número de fatores selecionados.

A Figura 19 exhibe o comportamento dos pesos das duas primeiras variáveis latentes dos modelos construídos com: a impedância expressa como números complexos, a parte real da impedância ( $Z'$ ) e a parte imaginária da impedância ( $Z''$ ). Em ambos os modelos se verifica que as impedâncias que mais contribuíram foram registradas em frequências abaixo de 10 Hz. Outra observação importante é que as componentes real e imaginária dos pesos ( $P_z' + jP_z''$ ) do

modelo complexo ( $Z' + jZ''$ ) são diferentes dos pesos dos modelos construídos com a parte real ( $Z'$ ) e imaginária ( $Z''$ ) da impedância.

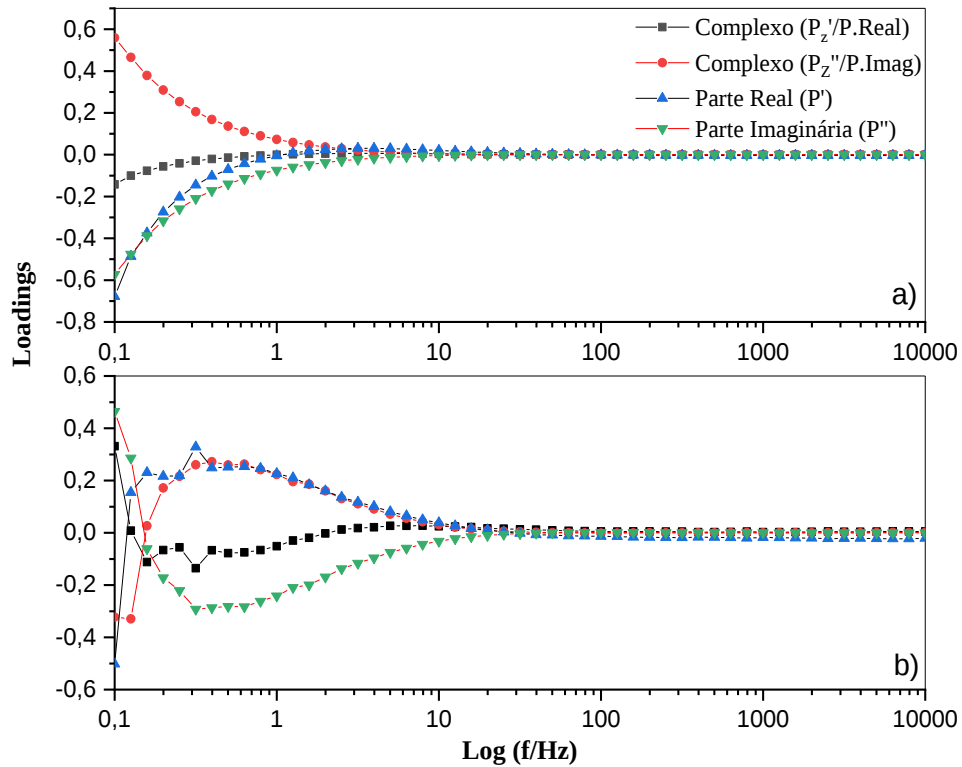


Figura 19 – Comportamento dos pesos da primeira (a) e segunda (b) variáveis latentes usadas nos modelos PLS construídos com: a impedância expressa como número complexo ( $Z' + jZ''$ ), parte real da impedância ( $Z'$ ) e a parte imaginária da impedância ( $Z''$ ).

A Figura 20 (a e b) exibe os resíduos deixados em cada frequência após a aplicação da transformação de Kronig-Kramers. Na transformação de Kronig-Kramers, a parte real da impedância é estimada a partir da parte imaginária da impedância, e vice-versa. O ideal é que os valores estimados a partir de uma das partes da impedância sejam iguais aos valores da outra parte. Se um dos critérios que validam a medida falhar (linearidade, estabilidade casualidade e finitude), a estimativa de uma das partes da impedância em função da outra não será satisfatória e o resultado será o aumento dos resíduos deixados pela diferença entre o valor previsto e o valor medido experimentalmente para aquela parte.

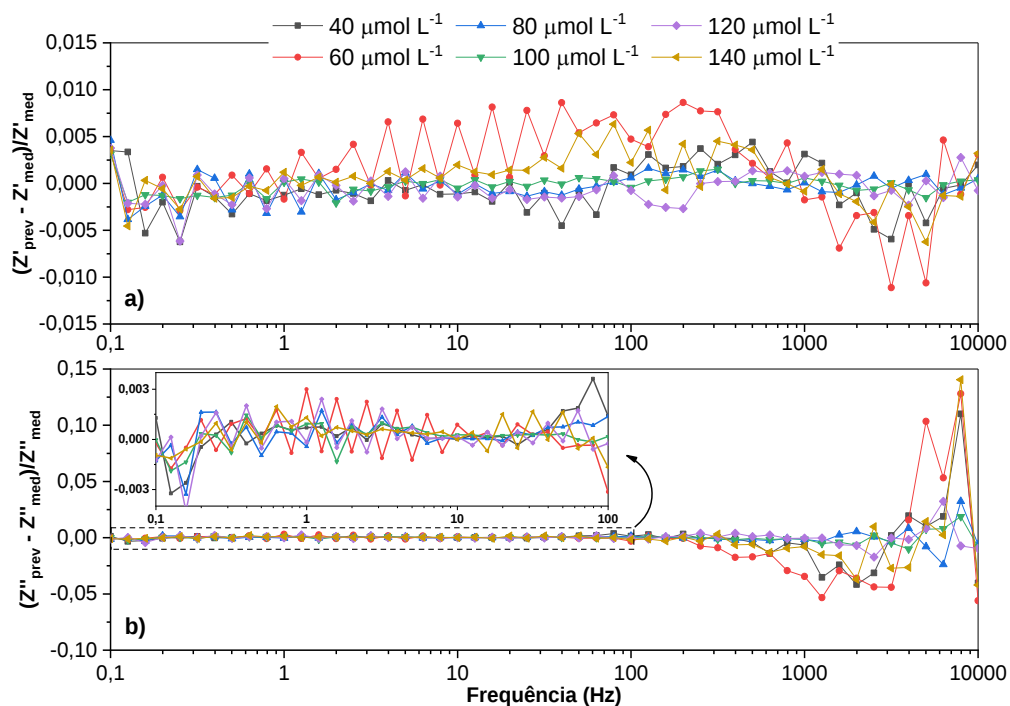


Figura 20 - (a) Resíduo relativo resultante da diferença entre a impedância real medida ( $Z'_{med}$ ) e a impedância real estimada em cada frequência a partir de  $Z''_{med}$  usando a transformação de Kronig-Kramers. (b) Resíduo relativo deixado pela diferença entre a impedância imaginária medida ( $Z''_{med}$ ) e a impedância imaginária estimada a partir de  $Z'$  (medido) usando a transformação de Kronig-Kramers. A impedância registrada sob aplicação do potencial de pico ( $E_a = 0,220$  V) e amplitude de 10 mV.

Os resíduos resultantes de ambos ajustes exibiram aleatoriedade em torno do zero, com os resíduos deixados pela parte real prevista a partir da parte imaginária variando entre  $\pm 0,005$ , e os resíduos deixados pela parte imaginária prevista a partir da parte real variando entre  $\pm 0,003$ . A soma dos quadrados dos resíduos relativos ( $\chi^2$ ) alcançados com a transformação de Kronig-Kramers situou-se na faixa de  $10^{-7}$  a  $10^{-6}$ , o que é indicativo de ajuste bom/excelente, indicando que as medidas realizadas estavam obedecendo os critérios pré-estabelecidos.

### 3.5.2 Determinação de catecol e hidroquinona usando o PLS para número complexo

A hidroquinona e o catecol também foram quantificados através de modelos PLS construídos com dados complexos de EIE. O catecol e a hidroquinona são isômeros naturalmente encontrados no meio ambiente frutos de atividades antropogênica. Ambos são compostos eletroativos com mecanismos de oxidações bem definidos e por isso são muito usados como analito de teste na validação de métodos analíticos novos (LAKSHMI et al., 2009; DUVALL; MCCREERY, 1999). A mesma metodologia usada na determinação do ferrocianeto foi aplicada na determinação de hidroquinona. Os espectros foram adquiridos aplicando o potencial de + 0,110 V (início do pico), +0,225 V (potencial de pico) e + 0,450 V (sobrepotencial) ao eletrodo de nanotubos de pasta de carbono. As amplitudes de 1 mV, 5mV e 10 mV também foram avaliadas. A faixa de concentração usada na quantificação da hidroquinona foi de 40-220  $\mu\text{mol L}^{-1}$ . A Figura 21 apresenta os diagramas de Nyquist e Bode referente aos espectros obtidos em 5 mV. Resultados similares foram obtidos aplicando-se 1 e 10 mV.

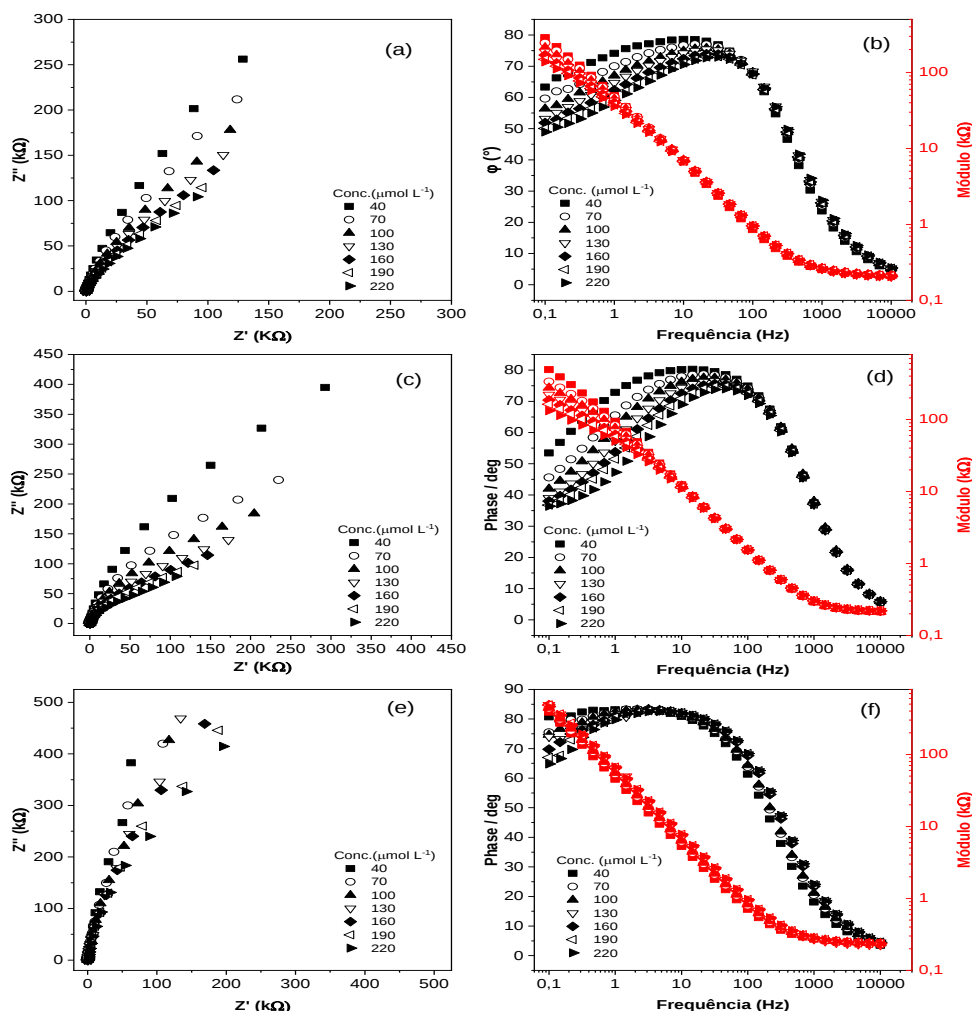


Figura 21 - Gráficos de nyquist (a,c,e) e bode (b,d,f) referente aos espectros registrados na determinação de hidroquinona em eletrodo de pasta de nanotubo de carbono aplicando potenciais diferentes:  $E_a = 0,110$  V (a) e (b),  $E_a = 0,225$  V (c) e (d),  $E_a = 0,450$  V, (e) e (f).

Assim como foi observado para o ferrocianeto, verificou-se dependência entre o comportamento dos espectros e o potencial aplicado. Os espectros registrados aplicando-se os potenciais de início do pico e pico apresentaram características similares nos diagramas de Nyquist e Bode (Figura 21a-d). Os espectros apresentaram duas regiões: um semicírculo em frequências altas, correspondente ao processo de transferência de carga, e uma região linear nas frequências baixas, correspondendo ao controle difusional. Esse comportamento de reação faradaica controlada por difusão semi-infinita também pode ser visto nas frequências baixas do diagrama de Bode ( $|Z|$ ) com inclinações de aproximadamente  $-0,5$  e ângulo de fase de aproximadamente  $45^\circ$ . Essas características são mais pronunciadas nos espectros adquiridos aplicando o potencial de pico. Nessa condição também foi observada uma mudança maior nos espectros em função da variação da concentração de hidroquinona. Os espectros registrados em

regime de sobrepotencial (Figura 21e-f) não exibiram mudanças expressivas decorrentes da variação da concentração da hidroquinona. Consequentemente, os modelos PLS construídos a partir dos espectros registrados nessa condição apresentaram baixas capacidades preditivas. A Tabela 8 apresenta os valores de RMSECV e o número de variáveis latentes selecionadas para os modelos construídos com os espectros registrados aplicando-se o potencial de início do pico ( $E_a = +0,100$  V) e potencial de pico ( $+0,225$  V).

Uma ANOVA foi realizada nos resultados da Tabela 8 considerando os seguintes fatores e níveis: i) potencial aplicado, nos níveis -1 =  $+0,110$  V e +1 =  $+0,225$  V; ii) amplitude, nos níveis, -1 = 1 mV, 0 = 5 mV e +1 = 10 mV; apresentação dos dados, nos níveis -1 = número complexo ( $Z' + jZ''$ ), -0,5 = parte real da impedância ( $Z'$ ), 0 = parte imaginária da impedância ( $Z''$ ), +0,5 = módulo da impedância ( $|Z|$ ), +2 = ângulo de fase ( $\phi$ ); e iv) pré-tratamento dos dados, nos níveis -1 = bruto e +1 = centrados na média. A Figura 22 apresenta as cartas de Pareto obtidas a partir da análise do RMSEC (Figura 22a) e número de variáveis latentes (Figura 22c). As barras cinzas correspondem aos efeitos positivos e as pretas aos negativos.

Tabela 8 - Valores dos RMSECV obtidos com os modelos PLS construídos com dados de impedância gerados na determinação de hidroquinona aplicando potenciais e amplitudes diferentes ao eletrodo de carbono vítreo. A quantidade de variáveis latentes selecionadas encontra-se entre parênteses. A faixa de concentração empregada para a hidroquinona foi de 3 a  $20 \mu\text{mol L}^{-1}$ .

| Tratamento dos dados | Forma dos dados | $E_a = +0.110$ V |         |         | $E_a = +0.225$ V |                |         |
|----------------------|-----------------|------------------|---------|---------|------------------|----------------|---------|
|                      |                 | 1 mV             | 5 mV    | 10 mV   | 1 mV             | 5 mV           | 10 mV   |
| Centrados na média   | $Z' + jZ''$     | 12 (1)           | 9.0 (3) | 8.1 (2) | 4.5 (4)          | <b>2.7 (2)</b> | 16 (3)  |
|                      | $Z'$            | 26 (6)           | 5.7 (2) | 16 (1)  | 11 (3)           | 4.6 (2)        | 13 (3)  |
|                      | $Z''$           | 11 (1)           | 2.9 (6) | 9.4 (2) | 7.4 (2)          | 4.8 (4)        | 19 (3)  |
|                      | $ Z $           | 8.3 (1)          | 11 (5)  | 9.0 (2) | 1.5 (6)          | 3.6 (2)        | 22 (1)  |
|                      | $\phi$          | 7.4 (3)          | 18 (1)  | 8.8 (3) | 13 (1)           | 2.5 (4)        | 15 (2)  |
| Originais            | $Z' + jZ''$     | 1.5 (6)          | 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 (6) |
|                      | $ Z $           | 12 (5)           | 31 (4)  | 9.0 (2) | 5.0 (4)          | 8.3 (4)        | 1.2 (6) |
|                      | $\phi$          | 54 (2)           | 26 (3)  | 8.8 (3) | 14 (2)           | 5.9 (5)        | 20 (2)  |

Três efeitos foram estatisticamente significativos para o RMSECV (Figura 22a): os efeitos de primeira ordem potencial aplicado e forma dos dados, e o efeito de segunda ordem AB referente a interação entre potencial e amplitude. O potencial aplicado apresenta um efeito de primeira ordem negativo, indicando que os valores menores de RMSECV foram alcançados no potencial de pico. Por outro lado, o efeito de primeira ordem relacionado a apresentação dos dados teve um sinal positivo, significando que valores menores de RMSECV foram observados nos dados de impedância escritos sob a forma de números complexos. Esse resultado reforça a hipótese de melhor eficiência dos dados sob a forma de números complexos, a qual contém toda informação relacionada a variação da concentração do analito. O pré-tratamento dos dados e amplitude não mostraram influência significativa sob o RMSECV. Assim, considerando apenas os efeitos principais na escolha da condição experimental ideal, o menor RMSECV seria conseguido com os espectros registrados sob a aplicação do potencial de pico usando as amplitudes de 1 mV ou 10 mV, empregando os dados na forma de números complexos sem centrá-los na média. Entretanto, a ocorrência do efeito de interação de segunda ordem AB deve ser considerado, e nesse caso a interpretação do efeito principal relacionado ao potencial aplicado não é tão direta.

A Figura 22b mostra a influência do efeito da interação AB sobre o RMSECV, apresentando os valores estimados para ele em diferentes combinações dos níveis dos fatores potencial aplicado e amplitude. As linhas do gráfico exibem o comportamento do RMSECV em função da variação do potencial aplicado quando as amplitudes foram mantidas fixas. Os demais fatores foram mantidos no nível central. Pode-se observar que o RMSECV é baixo no maior nível do potencial aplicado quando se utiliza as amplitudes de 1 mV ou 5 mV; semelhantemente ao observado na interação de primeira ordem AB da Figura 22a. Por outro lado, o RMSECV aumenta significativamente no nível superior do potencial aplicado quando a amplitude escolhida é 10 mV, diferentemente do que foi observado para o ferrocianeto. Embora a amplitude comumente reportada pela literatura seja 10 mV, não há motivos para esperar que ela seja a melhor em todos os sistemas experimentais.

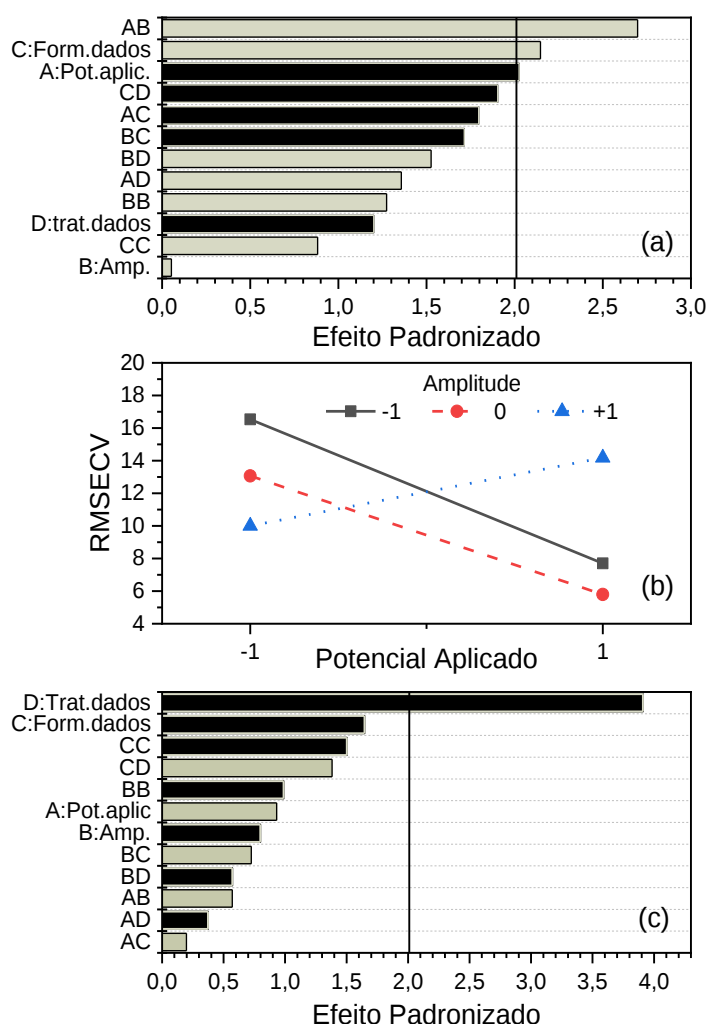


Figura 22 - Cartas de Pareto referente a ANOVA realizada nas respostas RMSECV (a) e número de variáveis latentes (c) na determinação de hidroquinona, onde: A = potencial aplicado; B = amplitude; C = apresentação dos dados; D = pré-tratamento dos dados. Variação do RMSECV relacionado a interação entre potencial aplicado e amplitude (b).

A Figura 22c exibe a carta de Pareto obtida da ANOVA realizada sobre o número de variáveis latentes. Como pode-se ver, o efeito de primeira ordem relacionado ao pré-tratamento dos dados foi estatisticamente significativo para definir a quantidade de variáveis selecionadas. O sinal negativo desse efeito indica que os dados de impedância centrados na média produzem modelos PLS que tendem a utilizar poucas variáveis latentes. Assim, considerando a avaliação realizada, o melhor modelo PLS para determinação de hidroquinona será construído com os espectros registrados sob aplicação do potencial de pico, amplitude de 5 mV, dados escritos na forma de números complexos e centrados na média. O resultado fornecido pelo modelo com

melhor desempenho analítico, obtido nas condições supracitadas, encontra-se destacado em negrito na Tabela 8.

A determinação de hidroquinona em água de torneira foi realizada para avaliar a robustez do método. Para isso, uma curva analítica nova foi construída empregando as condições experimentais estabelecidas no estudo anterior, utilizando uma faixa de concentração de 3 - 25 mmol L<sup>-1</sup> e realizando as análises diretamente em oito amostras de água de torneira fortificadas. A Figura 23a exibe os valores das previsões das amostras de calibração obtidos na validação cruzada do tipo “*leave one out*” e as concentrações de hidroquinona previstas nas amostras de água fortificadas. A determinação de catecol em água de torneira também foi realizada por meio da EIE e do PLS usando a impedância na forma de número complexo. Os espectros foram obtidos em eletrodo de pasta de nanotubo de carbono, aplicando o potencial do pico de oxidação e amplitude de 5 mV. A curva analítica foi construída na faixa de 65 – 175 µmol L<sup>-1</sup> e cinco amostras de água de torneira fortificadas com catecol foram analisadas. A Figura 23b exibe os valores das previsões das amostras de calibração obtidos por validação cruzada do tipo “*leave one out*” e as concentrações de catecol previstas nas amostras de água fortificadas. As previsões foram satisfatórias em ambas determinações, considerando que as análises foram realizadas diretamente nas amostras reais fortificadas usando a curva analítica construída em água deionizada. As recuperações das amostras fortificadas com hidroquinona variaram de 87,6% a 116% e das amostras fortificadas com catecol variaram de 95,5% a 108,9%

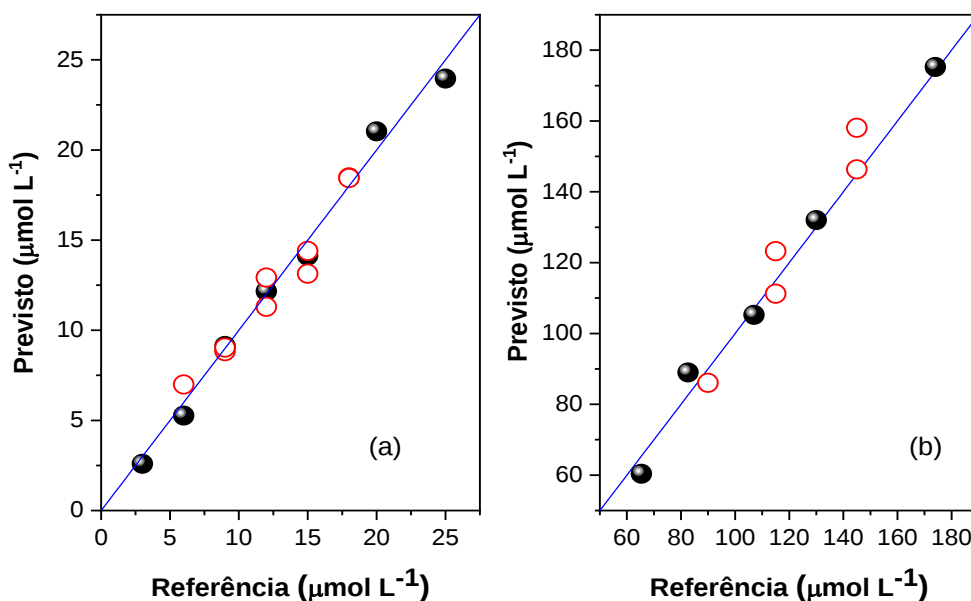


Figura 23 – Concentrações das amostras de calibração previstos por validação cruzada (preto) e resultados das previsões (vermelho) de hidroquinona (a) e catecol (b) em água da torneira.

### 3.6 Conclusões do Capítulo

- ✦ Os melhores resultados foram obtidos com os modelos construídos com os espectros registrados sobre aplicação do potencial de oxidação da espécie empregando amplitude de oscilação de 5 mV ou 10 mV;
- ✦ Os modelos construídos com as impedâncias expressas como número complexo conduziram a resultados similares ou melhores que resultados obtidos com os modelos construídos com a parte real ( $Z'$ ), imaginária ( $Z''$ ), módulo ( $|Z|$ ) e fase ( $\varphi$ ) da impedância;
- ✦ As recuperações encontradas para as amostras de água de torneira fortificadas, empregando a metodologia proposta, foram satisfatórias, com valores variando de 87,6% a 116,3%.

## **CAPÍTULO IV – CONSTRUÇÃO DE MODELOS PARA CLASSIFICAÇÃO DE AMOSTRAS DE CAFÉS**

## 4.1 Introdução

O café é um dos produtos alimentares mais consumidos no mundo e o segundo produto mais comercializado (SEZER et al., 2018). Seu mercado mundial cresce aproximadamente 2% ao ano, enquanto o segmento de cafés especiais aumenta entre 10% e 15% (FIGUEIREDO et al., 2019). No Brasil é a segunda bebida mais consumida, depois da água, e segundo a Empresa Brasileira de Pesquisa Agropecuária, em 2018, a produção de café do Brasil foi superior a um terço da produção mundial (EMBRAPA, 2018). Esses dados demonstram quão atrativo e lucrativo é o comércio de café e justificam porque a adulteração de cafés torrados é uma prática comum. Somente no Brasil, estima-se que aproximadamente 25% das marcas disponíveis no mercado adulteram o café torrado. (DANIEL et al., 2018).

A prática de adulteração de cafés torrados é comum, podendo estar relacionada à qualidade do grão; quando adicionam-se grãos defeituosos de espécies de cafés diferentes ou de origem geográfica diferente da informada no produto; e a adição de outras substâncias, como cascas, paus e caule do próprio cafeeiro, café vencido, milho, cevada, chicória, farelo de trigo, açúcar mascavo, soja, centeio entre outros (MORAIS et al., 2019; TOCI et al., 2016; PAULI et al., 2014). Algumas normas vêm sendo criadas para tentar combater essa prática. No Brasil, a Associação Brasileira da Indústria de Café (ABIC) em parceria com o Conselho Deliberativo da Política do Café (CDPC) criaram o programa de qualidade do café (PQC), cuja finalidade principal é certificar a pureza dos cafés torrados. Este programa é regulamentado pelo Ministério da Agricultura, Pecuária e abastecimento, que através da Instrução Normativa 16 de 24 de maio de 2010, limita a quantidade de impurezas totais presentes no café torrado a 1% (MAPA, 2010).

Vários métodos vêm sendo desenvolvidos para auxiliar na detecção de adulteração de café, dentre eles pode-se destacar os métodos voltamétricos (MORAIS et al., 2019), cromatográficos (PAULI et al., 2014; DOMINGUES et al., 2014), espectroscópicos (EL-ABASSY et al., 2011) e os biológicos baseados no DNA (COMBES et al., 2018). O objetivo do presente trabalho é desenvolver um método para identificação de adulterações em amostras de cafés utilizando a espectroscopia de impedância eletroquímica, com os dados de impedância expressos como números complexos, aliada a técnicas de reconhecimento de padrão supervisionados, especificadamente, SIMCA e PLS-DA.

## **4.2 Objetivos**

### **4.2.1 Objetivos Específicos**

1. Encontrar a condição experimental que será empregada durante a aquisição dos espectros das amostras de cafés;
2. Construir modelos PLS-DA e SIMCA com a impedância expressa sob a forma de número complexo;
3. Comparar os resultados dos modelos PLS-DA e SIMCA obtidos com os números complexos aos resultados alcançados com os modelos construídos com a parte real da impedância ( $Z'$ ), a parte imaginária da impedância ( $Z''$ ), o módulo da impedância ( $|Z|$ ) e o ângulo de fase ( $\varphi$ ).

### 4.3 Fundamentação Teórica

#### 4.3.1 Modelagem independente e Flexível por Analogia de Classes (SIMCA)

O SIMCA é uma técnica de modelagem de classes baseada na análise por componentes principais (PCA). Nesse método são realizadas uma PCA em cada classe do conjunto de dados. A PCA realizada em cada classe é usada na construção de uma fronteira (“caixa ou envelope”) que delimita os elementos pertencentes a essa classe. Para construir a fronteira de cada classe é necessário ter um conjunto de treinamento/aprendizagem, composto exclusivamente de elementos identificados como pertencentes a classe que se deseja modelar. Após a construção dos modelos é possível identificar amostras cujas classes são desconhecidas, projetando-as no subespaço (envelope) criado em cada modelo, e comparando a distância da projeção a cada modelo. O número de componentes principais empregadas por cada modelo SIMCA é comumente calculada por validação cruzada e a distância limite (“threshold”) de cada classe é encontrada mediante o aumento da sensibilidade e especificidade da classe (REID et al., 2007).

#### 4.3.2 Análise Discriminantes por Mínimos Quadrados Parciais

A Análise Discriminante por mínimos Quadrados Parciais (PLS-DA, do inglês, *Partial Least Square Discriminant Analysis*) é um método de classificação que une as características do PLS (descrito no capítulo III) ao poder da análise discriminante. Teoricamente, o PLS-DA combina a redução de dimensionalidade ao poder discriminatório, em um mesmo algoritmo, e assim consegue modelar conjunto de dados com muitas variáveis (LEE, 2018). O objetivo é obter melhor discriminação entre as classes buscando a covariância máxima entre as matrizes de dados instrumental (X) e categóricas (Y).

No PLS-DA, a matriz Y contém G colunas e n linhas. A quantidade de colunas depende da quantidade de classes da modelagem, enquanto o número de linhas depende da quantidade de amostras usadas na modelagem. Cada entrada  $y_{ig}$  de Y é codificada com os valores 0 ou 1. Assim, em uma modelagem contendo duas classes, por exemplo, as entradas  $y_{1g}$  das amostras da primeira coluna, que pertencem a classe 1, receberão o valor 1 e as demais receberão o valor 0. As entradas  $y_{2g}$  das amostras da segunda coluna, que pertencem a classe 2, receberão o valor 1 e as demais receberão o valor 0. O PLS é executado e retorna os valores  $y_g^{calc}$  estimados para cada amostra em cada classe. Quanto mais próximo de 1 for o valor estimado  $y_g^{calc}$ , significa que a amostra pertence à classe. Não pertencendo a ela quanto mais próximo de 0 for o valor

estimado. Todavia, a atribuição da classe de uma amostra é realizada com base em um critério mais objetivo. Uma alternativa é calcular um valor limiar (“Threshold”) para cada classe da modelagem e usá-lo como parâmetro para definir a qual classe pertence a amostra com valor  $y_g^{calc}$  estimado. Uma das maneiras de calcular o limiar é usando o teorema de Bayes. Por esse método, o limiar selecionado será aquele que o número de falsos positivos e falsos negativos são minimizados (Ballabio, 2013).

#### **4.3.3 Trabalhos que utilizam dados de EIE no reconhecimento de padrões**

Chen e colaboradores desenvolveram uma metodologia usando dados de impedância para diferenciar peitos de frango frescos daqueles que foram congelados e descongelados. Os eletrodos utilizados foram duas agulhas que eram inseridas na direção longitudinal do musculo. A magnitude da impedância ( $|Z|$ ) e ângulo de fase ( $\varphi$ ) foram registrados em 15 frequências de 50 kHz a 200 kHz. As classificações foram realizadas a partir de três matrizes de dados: magnitude da impedância, ângulo de fase e a matriz formada pela fusão dos dados da magnitude da impedância e ângulo de fase. Dois métodos supervisionados foram aplicados: rede neural (aprendizado por quantização vetorial) e PLS-DA (CHEN et al, 2016).

SINGH e colaboradores propuseram um método para classificação de amostras de chá verde aromatizados. Os espectro de impedância dos extratos foram registrados de 1 Hz a 100 kHz numa célula com três eletrodos, platina (auxiliar), ouro (trabalho) e Ag/AgCl. A estratégia usada foi ajustar um circuito elétrico equivalente aos espectros de impedância e utilizar os valores do elementos elétrico na classificação. A técnica utilizada na classificação foi a aprendizagem de máquina de vetores de suporte (SINGH et al, 2015).

Conesa e colaboradores desenvolveram uma metodologia para identificar e quantificar açúcares fermentáveis (glicose, frutose e sacarose) presentes no processo de sacarificação enzimática do abacaxi, durante a produção de bioetanol de segunda geração. Os valores dos ângulos de fase, registrados de 596 kHz a 747 kHz, usando duas agulhas de aço como eletrodos, foram usados na modelagem. A quantificação foi realizada com base em um modelo PLS e a classificação com base em um modelo de redes neurais artificiais. A percentagem de amostras classificadas corretamente foi superior a 93% (CONESA et al, 2015).

Scandurra e colaboradores propuseram um método baseado em análise por circuito elétrico equivalente para determinação da origem méis florais. Nesse método, os parâmetros dos circuitos equivalentes ajustados foram associados características físico-químicas dos

diferentes méis. No mesmo trabalho, os pesquisadores conseguiram agrupar os méis de diferentes origem realizando uma PCA nos dados dos ângulos de fase, coletados entre 1kHz a 1 MHz (SCANDURRA et al, 2013).

Garcia-Hernandez e colaboradores propuseram uma língua eletrônica para discriminar vinhos tintos com características similares (mesma região, safra e método de envelhecimento) usando um arranjo de eletrodos com diferentes modificações e dados de EIE registrados de 10 kHz a 1 Hz. A discriminação foi alcançada mediante PCA realizada numa matriz obtida da fusão dos dados de impedância real ( $Z'$ ), impedância imaginária ( $Z''$ ) e fase ( $\varphi$ ) registradas em 1,25 Hz (Garcia-Hernandez et al, 2018).

Não foram encontrados trabalhos que aplicam o Simca em classificações com dados de espectroscopia de impedância. Esses trabalhos representam como são aplicados os dados de impedância, atualmente. As maioria das técnicas aplicadas são não supervisionadas.

## **4.4 METODOLOGIA**

### **4.4.1 Soluções e Amostras**

As medidas foram realizadas em tampão acetato  $0,10 \text{ mol L}^{-1}$  (pH 4,66), preparado através da mistura de volumes iguais de uma solução de ácido acético e acetato de sódio, ambos a  $0,20 \text{ mol/L}^{-1}$ .

Noventa amostras de cafés torrados, certificadas pela Associação Brasileira de Indústrias de café (ABIC), e disponibilizadas pelo Núcleo Global de Análise e Pesquisa (NUGAP) e o Serviço Nacional de Aprendizagem Industrial (SENAI), foram separadas em dois grupos, um contendo 60 amostras de cafés puros e o outro contendo 30 amostras de cafés adulterados com cascas e paus de cafeeiro. A porcentagem de adulteração variou de 1,6% a 6,5%, tomando como referência o que estabelece a Instrução Normativa nº 16/2010 do ministério da agricultura, pesca e abastecimento. As amostras de cafés empregadas eram misturas das variedades arábica e robusta, que são as duas espécies mais produzidas pela indústria Brasileira.

### **4.4.2 Medidas de Impedância**

As medidas de impedância foram realizadas em um potenciostato Autolab PGSTAT M101 (Metrohm Autolab, The Netherlands) controlado pelo software Nova 1.11 numa célula eletroquímica convencional de três eletrodos. Os eletrodos auxiliar e de referência foram, respectivamente, uma haste de grafite e o eletrodo de calomelano saturado. Um eletrodo de pasta de grafite foi o eletrodo de trabalho. Tal foi preparado através da mistura de pó de grafite (70%) com o óleo mineral (30%).

Os espectros de impedância foram registrados na faixa de frequência de 10 kHz a 0,1 Hz coletando o sinal numa taxa de 10 pontos por década de frequência. O potencial e amplitude aplicados durante o registro dos espectros de impedância foram 0,47 V e 5 mV, respectivamente.

#### 4.4.3 Procedimento Analítico

O procedimento de extração se deu da seguinte maneira: 1,0 g de pó de café era transferido para um filtro de papel posicionado em um funil de vidro; 50 mL de água purificada aquecida a  $90\text{ }^{\circ}\text{C} \pm 5\text{ }^{\circ}\text{C}$  era adicionada sobre o pó; o extrato era aparado em um Becker, onde permanecia até atingir a temperatura ambiente. A análise de cada extrato foi realizada da seguinte maneira: registrava-se o espectro na solução do tampão acetato  $0,1\text{ mol L}^{-1}$  (branco) e em seguida registrava-se o espectro após a adição de 1 mL do extrato da amostra de café. O registro do espectro era feito sob aplicação  $0,47\text{ V}$  de potencial com amplitude de  $5\text{ mV}$ .

O total de 90 amostras foram analisadas, sendo 60 amostras de cafés puros e 30 amostras de cafés adulterados com cascas e paus de cafeeiro. O registro espectral de cada amostra fornece como saída quatro vetores principais, e cada vetor guarda as respostas instrumental ( $Z'$ ,  $Z''$ ,  $|Z|$  e  $\varphi$ ) medidas em cada frequência analisada, como já foi mencionado nos capítulos anteriores. Os modelos de classificação PLS-DA e SIMCA foram construídos com a impedância expressa sob a forma de número complexos ( $Z' + jZ''$ ) e os resultados dessa abordagem foram comparados aos modelos construídos com os dados das respostas reais  $Z'$ ,  $Z''$ ,  $|Z|$  e  $\varphi$ .

Os modelos foram construídos com dados centrados na média. O algoritmo de Kennard-Stone foi usado para selecionar as amostras, separando-se 60% para o conjunto de treinamento/validação e 40% para o conjunto de teste. O número de variáveis latentes usadas no PLS-DA e PCs usadas no SIMCA foram selecionadas com base na taxa do erro da validação cruzada do tipo “*leave one out*”. O método de Bayes foi usado no cálculo do limiar do PLS-DA. Os cálculos foram executados nas rotinas do *classification toolbox*, versão 5.2, desenvolvida pelo professor Davide Ballabio. As rotinas foram executadas na versão 9.3.0 do Matlab® (The MathWorks™, Inc.).

#### 4.4.4 Cálculo dos parâmetros dos modelos

A precisão representa a capacidade do modelo de classificação não incluir amostras de outras classes à classe considerada. Ela compreende a razão entre as amostras classificadas na classe correta (verdadeiro positivo) e o número total de amostras classificadas como pertencentes a classe (verdadeiro positivo + falso positivo). A equação x descreve o cálculo da

precisão de uma classe  $g$ , onde,  $V_p$ , representa o número de verdadeiros positivos, e  $F_p$ , representa o número de falsos positivos.

$$Precisão_{(g)} = \frac{V_p}{V_p + F_p} \quad (28)$$

A sensibilidade representa a habilidade do modelo reconhecer, corretamente, as amostra pertencentes a classe  $g$ . Ela é calculada como a razão entre os verdadeiros positivo da classe  $g$  e o número total de amostras dessa classe ( $NT_g$ ).

$$Sensibilidade_{(g)} = \frac{V_p}{NT_g} \quad (29)$$

A especificidade representa a habilidade de uma classe  $g$  rejeitar as amostras de outras classes. Ela é calculada como a razão entre verdadeiros negativos ( $V_n$ ), que representa o número de amostras que não pertencem a classe avaliada e não foram classificadas nela, e a soma dessas amostras com os falsos positivos ( $F_p$ ), que representa o número de amostras que foram classificadas na classe avaliada, mas não pertencem a ela.

$$Especificidade_{(g)} = \frac{V_n}{V_n + F_p} \quad (30)$$

A taxa de erro é calculada como a diferença entre um (1) e a razão entre a sensibilidade média, que é o somatório das sensibilidades de cada classe dividido pelo número de classes ( $g$ ),

$$Taxa\ de\ erro = 1 - \left( \frac{\sum_1^g Sens.\ de\ cada\ classe}{número\ de\ classes} \right) \quad (31)$$

## 4.5 Resultados e Discussão

### 4.5.1 Construção dos modelos PLS-DA e SIMCA

Morais *et al.* (2019) propôs uma língua eletrônica voltamétrica para identificação de adulteração em amostras de extratos de cafés. Uma das estratégias do trabalho foi um modelo baseado em Análise Discriminante Linear (LDA) que utilizava os valores das correntes de 10 potenciais, selecionados pelo Algoritmo Genético (GA), para discriminar amostras de cafés puros e adulterados. O potencial de 0,47 V foi o único entre os 10 potenciais selecionados pelo GA que se situava numa região de alta densidade eletrônica e ao mesmo tempo proporcionava a discriminação entre os cafés puros e adulterados. Por isso, o potencial de 0,47 V foi aplicado durante as análises e empregou-se a mesma condição experimental utilizada por Moraes.

Na abordagem empregada neste trabalho optou-se por não modificar a superfície do eletrodo e utilizar a impedância faradaica durante a aquisição dos espectros dos extratos de cafés. A modificação da superfície do eletrodo é um recurso frequentemente empregado para tornar a análise mais seletiva ao analito. Por outro lado, o processo envolvido nessa modificação tende a ser laborioso, dispendioso e de difícil reprodução, o que pode prejudicar a modelagem. Em vez de modificar a superfície do eletrodo para deixá-la sensível aos componentes da adulteração presente no café, a estratégia desse trabalho foi empregar a impedância faradaica com o eletrodo descoberto, aplicando um potencial cuja corrente tivesse a contribuição das espécies eletroativas naturais da matriz e das espécies introduzidas pela adulteração, e promover a discriminação entre cafés puros e adulterados com o uso da quimiometria.

Sabe-se que algumas espécies que estão presentes no extrato de café apresentam pico de oxidação próximos a 0,47 V em eletrodos carbonáceos. Os grupos hidroxilas de catecóis de isômeros do ácido cafeoilquínico/dicafeoilquínico oxidam reversivelmente próximos a  $E_a = +0,35$  V em processos que envolve a transferência de dois prótons e dois elétrons (SERUGA *et al.*, 2014). Isômeros do ácido feruloilquínico sofrem oxidação nos grupos fenoxilas em potenciais próximos a  $E_a = +0,30$ , apresentando um mecanismo EC (eletroquímico-químico) envolvendo a transferência de 1 próton e 1 elétron (TOMAC *et al.*, 2016). Todos esses composto são ácidos clorogênicos e, além de apresentar os processos descritos anteriormente, comumente sofrem oxidação nas hidroxilas dos grupos catecóis, por volta de  $E_a = +0,55$  V. O ácido tânico é outro constituinte do café que apresenta pico de oxidação próximo a  $E_a = +0,30$  V (VU *et al.*, 2015). Esses compostos são exemplos de espécies conhecidas, cujas oxidações podem contribuir com a corrente gerada em 0,47V.

A Figura 24a exibe os diagramas de Nyquist das 60 amostras de cafés puros e 30 amostras de cafés adulterados, e seus respectivos escores no plano complexo formado pela primeira componente principal (T1), real versus imaginária. Os espectros das amostras de cafés não exibiram o semicírculo, típico de processos controlados por cinética, na região de alta frequência. Essa ausência de semicírculo também foi verificada por Bhondekar *et al.* (2010) em análises de chás pretos, matriz similar ao café, usando o eletrodo de carbono vítreo. Também não foi verificado a reta com inclinação de 45°, típica de processos controlados por difusão, nas regiões de médias e baixas frequências. Um circuito de Randles composto por um resistor em série com dois elementos de fase constante (CPE) em paralelo se ajustou aos espectros. Um dos CPE exibiu comportamento de um resistor não ideal, apresentando baixas resistências a transferência de carga e valores de  $n$  variando entre 0,18 e 0,23, enquanto o outro CPE exibiu o comportamento de um capacitor não ideal, apresentando valores de  $n$  variando entre 0,8 e 0,85. Apesar de os espectros médios exibirem as diferenças médias entre as amostras das classes diferentes, a olho nu não é possível discriminar se o espectro é de uma amostra pura ou adulterada, e por isso se faz necessário a aplicação de técnicas quimiométricas.

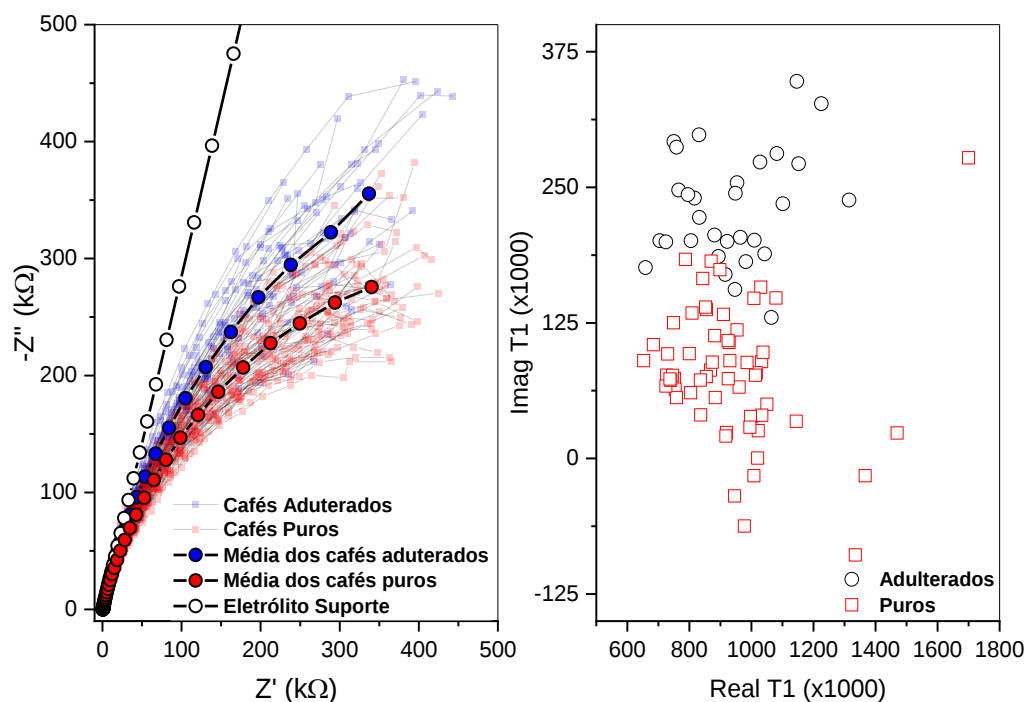


Figura 24 - (a) Diagrama de Nyquist referente aos espectros das amostras de cafés adulterados e puros. (b) Escores da primeira componente principal real e imaginária (T1) realizada nos dados complexos das amostras de cafés, puros e adulterados com cascas e paus, centrados na média.

Uma análise exploratória baseada em PCA foi executada nos dados centrados na média. A Figura 24b exibe os escores da primeira componente principal (T1) do primeiro fator complexo, real versus imaginário. Observa-se que a componente imaginária consegue discriminar as amostras pertencentes as classes diferentes. As amostras de cafés puros encontram-se nos escores menores e as amostras de cafés adulterados encontram-se nos escores mais altos. Os modelos de reconhecimento de padrão supervisionados foram construídos e a seleção do número de fatores, empregado pelo PLS-DA, e PCs, empregado pelo SIMCA, foi realizado com base na taxa de erro, calculada por meio da equação (31). No anexo I encontram-se os gráficos da taxa de erro em função do número de fatores dos modelos PLS-DA construídos a partir dos dados das diferentes informações relacionadas à impedância. A Tabela 9 resume os melhores resultados da classificação PLS-DA realizada a partir de cada conjunto de dados relacionados à impedância.

Os modelos que conduziram aos melhores resultados encontram-se destacados na Tabela 9. O modelo construído com a impedância expressa como números complexos apresentou uma taxa de erro de 0,097 durante a validação cruzada, selecionando três variáveis latentes. Sete amostras da classe 2 foram classificadas como pertencentes à classe 1 durante a validação cruzada, e no conjunto de teste apenas uma amostra da classe 2 foi classificada na classe 1. Todas as amostras da classe 1 foram classificadas corretamente. O modelo construído com a fase da impedância apresentou uma taxa de erro de 0,083 durante a validação cruzada, selecionando apenas duas variáveis latentes. Seis amostras da classe 2 foram classificadas como pertencentes à classe 1 durante a validação cruzada. Já no conjunto de teste apresentou quatro erros de classificação, errando duas amostras de cada classe. Os parâmetros de desempenho apresentados por ambos os modelos foram satisfatórios.

A Tabela 10 resume os melhores resultados da classificação SIMCA realizada a partir de cada conjunto de dados avaliados. O modelo construído com a impedância expressa como número complexo apresentou uma taxa de erro de 0,16 durante a validação cruzada. As variâncias acumuladas, explicadas pelas duas PCs de cada classe, foram superiores a 94%. Esse modelo apresentou cinco erros de classificação na previsão das amostras de teste. O melhor modelo, com base na taxa de erro da validação cruzada (0,08), foi construído com os dados da parte imaginária da impedância, usando duas PCs para descrever cada classe. Apenas três amostras dos conjuntos de testes foram classificadas incorretamente. Os parâmetros de desempenho de ambos os modelos foram similares.

Tabela 9 – Melhores resultados da classificação de cafés para os conjuntos de validação e teste obtidos com os modelos PLS-DA construídos, com os dados da: impedância expressa como número complexo ( $Z' + jZ''$ ), parte real da impedância ( $Z'$ ), parte imaginária da impedância, módulo da impedância ( $|Z|$ ) e ângulo de fase da impedância ( $\varphi$ ). Os números entre parênteses referem-se à quantidade de fatores usados na modelagem. O índice 1 representa a classe dos cafés adulterados com cascas e paus e o índice 2 representa a classe dos cafés puros.

| Índice de Classe Predito          |         |    |                   |      |               |      |                |      |                |      |                                 |      |
|-----------------------------------|---------|----|-------------------|------|---------------|------|----------------|------|----------------|------|---------------------------------|------|
|                                   |         |    | <b>Z + jZ (3)</b> |      | <b>Z' (4)</b> |      | <b>Z'' (3)</b> |      | <b> Z  (3)</b> |      | <b><math>\varphi</math> (2)</b> |      |
| Conjuntos                         | Índices | N  | 1                 | 2    | 1             | 2    | 1              | 2    | 1              | 2    | 1                               | 2    |
| Validação                         | 1       | 18 | 18                | 0    | 14            | 4    | 16             | 2    | 17             | 1    | 18                              | 0    |
|                                   | 2       | 36 | 7                 | 29   | 9             | 27   | 4              | 32   | 6              | 30   | 6                               | 30   |
| Teste                             | 1       | 12 | 12                | 0    | 12            | 0    | 11             | 1    | 3              | 9    | 10                              | 2    |
|                                   | 2       | 24 | 1                 | 23   | 6             | 18   | 2              | 22   | 3              | 21   | 2                               | 22   |
| Sensibilidade (Validação)         |         |    | 1                 | 0,81 | 0,78          | 0,75 | 0,89           | 0,89 | 0,94           | 0,83 | 1                               | 0,83 |
| Especificidade (Validação)        |         |    | 0,81              | 1    | 0,75          | 0,78 | 0,89           | 0,89 | 0,83           | 0,94 | 0,83                            | 1    |
| Precisão (Validação)              |         |    | 0,72              | 1    | 0,61          | 0,87 | 0,8            | 0,94 | 0,74           | 0,97 | 0,75                            | 1    |
| Taxa de erro da validação cruzada |         |    | 0,097             |      | 0,23          |      | 0,11           |      | 0,11           |      | 0,083                           |      |

Tabela 10 - Melhores resultados da classificação de cafés para os conjuntos de validação e teste obtidos com os modelos SIMCA, construídos com os dados da: impedância expressa como número complexo ( $Z' + jZ''$ ), parte real da impedância ( $Z'$ ), parte imaginária da impedância, módulo da impedância ( $|Z|$ ) e ângulo de fase da impedância ( $\varphi$ ). Os números entre parênteses referem-se à quantidade de fatores usados na modelagem. O índice 1 representa a classe dos cafés adulterados com cascas e paus e o índice 2 representa a classe dos cafés puros.

| Índice de Classe Predito          |         |    |                |      |            |      |             |      |             |      |                 |      |
|-----------------------------------|---------|----|----------------|------|------------|------|-------------|------|-------------|------|-----------------|------|
|                                   |         |    | $Z + jZ$ (2,2) |      | $Z'$ (2,3) |      | $Z''$ (2,2) |      | $ Z $ (2,2) |      | $\varphi$ (1,1) |      |
| Conjuntos                         | Índices | N  | 1              | 2    | 1          | 2    | 1           | 2    | 1           | 2    | 1               | 2    |
| Validação                         | 1       | 18 | 14             | 4    | 8          | 10   | 16          | 2    | 11          | 7    | 16              | 2    |
|                                   | 2       | 36 | 4              | 32   | 7          | 29   | 2           | 34   | 2           | 34   | 5               | 31   |
| Teste                             | 1       | 12 | 8              | 4    | 10         | 2    | 10          | 2    | 8           | 4    | 10              | 2    |
|                                   | 2       | 24 | 1              | 23   | 9          | 15   | 1           | 23   | 1           | 23   | 3               | 21   |
| Sensibilidade (Validação)         |         |    | 0,78           | 0,88 | 0,44       | 0,81 | 0,88        | 0,94 | 0,61        | 0,94 | 0,88            | 0,86 |
| Especificidade (Validação)        |         |    | 0,88           | 0,78 | 0,81       | 0,44 | 0,94        | 0,88 | 0,94        | 0,61 | 0,94            | 0,76 |
| Precisão (Validação)              |         |    | 0,78           | 0,88 | 0,53       | 0,74 | 0,88        | 0,94 | 0,84        | 0,83 | 0,76            | 0,94 |
| Taxa de erro da validação cruzada |         |    | 0,16           |      | 0,37       |      | 0,08        |      | 0,22        |      | 0,12            |      |

## 4.6 Conclusões do Capítulo

- ✎ Modelos para classificação de amostras de café adulteradas e puras apresentaram resultados similares quando as matrizes de dados utilizadas foram: a impedância expressa como número complexo, a parte imaginária e a fase da impedância;
- ✎ Diante dos resultados, pode-se inferir que a matriz que melhor representa os dados é composta pelas impedâncias expressas como número complexo, tendo em vista que a partir dela é possível alcançar resultados similares aos encontrados com as demais matrizes, sem a necessidade de avaliá-las.

## **CONCLUSÕES GERAIS**

## Conclusões Gerais

Nesse estudo foram apresentadas novas estratégias de tratamento de dados de espectroscopia de impedância eletroquímica empregando métodos quimiométricos. Verificou-se que modelos multivariados construídos com os elementos do circuito equivalente ajustados aos dados do ferrocianeto, catecol e hidroquinona foram mais robustos que os modelos univariados construídos com apenas um único elemento do circuito. Durante a construção de modelos PLS empregando dados de impedância, escritos na forma de números complexos, foi verificado que a aplicação do potencial de pico da espécie juntamente com uma amplitude de 5 ou 10 mV, no momento da aquisição dos espectros, conduz a modelos com maior capacidade preditiva. O emprego dos dados de impedância expressos na forma de números complexos produziu modelos com resultados melhores ou iguais aos fornecidos pelo modelos construídos com os dados reais. Quanto a classificação, verificou-se que o modelo PLS-DA construído com os dados expressos na forma de números complexos, e o modelo PLS-DA construído com a fase da impedância, forneceram resultados bons e similares entre si. Os modelos SIMCA construídos com os dados expressos na forma de números complexos, parte imaginária e fase da impedância, também forneceram resultados bons e similares entre si. Ambos apresentaram robustez na predição das amostras do conjunto de teste, mas os modelos construídos com dados escritos na forma de números complexos se destacaram por, geralmente, utilizar poucas variáveis latentes/componentes principais na obtenção de predições similares.

## Propostas Futuras

- Construir modelos PLS usando dados de impedância escritos como número complexos para modelar analitos que apresentam comportamento eletroquímico que envolvem reações com mais de uma etapa, e comparar os resultados desses modelos com os fornecido por modelos construídos com dados reais;
- Avaliar a aplicação de métodos de seleção de variáveis na escolha das frequências que conduzem a melhores modelos de classificação ou quantitativos;
- Estender a abordagem para métodos de ordem superior como o N-PLS;



## **REFÊRENCIAS BIBLIOGRÁFICAS**

AICHER, M.; GROTHE, H.; WOLF, B. A novel thin film impedance Ca ion sensor for drinking water. *Sensors and Actuators B*, v.244, p.1103-1112, 2017.

AMIRUDIN, A.; THIERRY D. Application of electrochemical impedance spectroscopy to study the degradation of polymer-coated metals. *Prog. Org. Coat.*, v.26, p.1-28, 1995.

ASUERO, A. G.; SAYAGO, A.; GONZÁLEZ, A. G. The Correlation Coefficient: An Overview. *Critical Reviews in Analytical Chemistry*, v.36, p.41–59, 2006.

AYAGOU, M. D. D.; TRAN, T. T. M.; TRIBOLLET, B.; KITTEL, J.; SUTTER, E.; FERRANDO, N.; MENDIBIDE, C.; DURET-THUAL, C. Electrochemical impedance spectroscopy of iron corrosion in H<sub>2</sub>S solutions. *Electrochimica Acta*, v.282, p.775-783, 2018.

BALLABIO, D.; CONSONI, V. Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal. Methods*, v.5, P.3790-3798, 2013.

BISQUERT, J. Theory of the impedance of electron diffusion and recombination in a thin layer. *Journal of Phys. Chem. B*, v.106, p.325-333, 2002.

BONANNI, A.; LOO, A. H.; PUMERA, M. Graphene for impedimetric biosensing. *Trends in Analytical Chemistry*, V.37, p.12-21, 2012.

BOUKAMP, B. A. A nonlinear least squares fit procedure for analysis of immittance data of electrochemical systems. *Solid State Ionics*, V.20 p. p.31-44, 1986.

BRERETON, R. G. Pattern recognition in chemometrics. *Chemometrics and Intelligent Laboratory Systems*, v.149, p.90–96, 2015.

CARBALLO, R.; RINALDI, A. L.; REZZANO, I. Electrochemical study of azide bridged heterobimetallic films of porphyrins: Application as an impedimetric sensor. *Electrochimica Acta*, v.222, p.1700-1708, 2016.

CARVALHO, L. A.; ANDRADE, A. R. Espectroscopia de impedância eletroquímica aplicada ao estudo das reações heterogêneas em ânodos dimensionalmente estáveis. *Quim. Nova*, V.29, p.796-804, 2006.

CHEN, T.; ZHU, Y.; HAN, P. M.; WEI, R.; XU, X.; GUANG-HONGZHOU. The use of the impedance measurements to distinguish between fresh and frozen-thawed chicken breast muscle. *Meat Science.*, v.116, p.151–157, 2016.

CHOKKAREDDY, R.; REDHI, G. G.; KARTHICK, T. A lignin Polymer nanocomposite based electrochemical sensor for the sensitive detection of chlorogenic acid in coffee samples. *Heliyon.*, v.5, 1457, 2019.

COMBES, M. C., JOËT, T., LASHERMES, P. Development of a rapid and eficiente DNA-based method to detect and quantify adulterations in coffee (Arabica versus Robusta). *Food Control.*, v.88, p.198-206, 2018.

CONESA, C.; GARCÍA-BREIJO, E.; LOEFF, E.; SEGUÍ, L.; FITO, P.; Laguarda-Miró, N. An Electrochemical Impedance Spectroscopy-Based Technique to Identify and Quantify Fermentable Sugars in Pineapple Waste Valorization for Bioethanol Production. *Sensors.*, v.15, p.22941-22955, 2015.

CONESA, C.; GARCÍA-BREIJO, E.; LOEFF, E.; SEGUÍ, L.; FITO, P.; Laguarda-Miró, N. An Electrochemical Impedance Spectroscopy System for Monitoring Pineapple Waste Saccharification. *Sensors.*, v.16, p.188-196, 2016.

CONESA, C.; GARCÍA-BREIJO, E.; LOEFF, E.; SEGUÍ, L.; FITO, P.; Laguarda-Miró, N. An Electrochemical Impedance Spectroscopy-Based Technique to Identify and Quantify Fermentable Sugars in Pineapple Waste Valorization for Bioethanol Production. *Sensors.*, v.15, p.22941-22955, 2015.

CORTINA-PUIG, M.; NOZ-BERBEL, X. M.; DELVALLE, M.; MUNOZ, F. J.; ALONSO-LOMILLO, M. A. Characterization of an ion-selective polypyrrole coating and application to the joint determination of potassium, sodium and ammonium by electrochemical impedance spectroscopy and partial least squares method. *Analytica Chimica Acta.*, v.597, p.231–237, 2007.

DANIEL, D.; LOPES, F. S.; DOS SANTOS, V. B.; LAGO, C. L. Detection of coffee adulteration with soybean and corn by capillary electrophoresis-tandem mass spectrometry. *Food Chemistry.*, v.243, p.305-310, 2018.

DAY, C.; SØPSTAD, S.; MA, H.; JIANG, C.; NATHAN, A.; ELLIOTT, S.R.; FRANKL, F. E. K.; HUTTER, T. Impedance-based sensor for potassium ions. *Analytica Chimica Acta.*, v.1034, p.39-45, 2018.

DIOUF, A.; BOUCHIKHIA, B.; EL BARI, N. A nonenzymatic electrochemical glucose sensor based on molecularly imprinted polymer and its application in measuring saliva glucose. *Materials Science & Engineering C.*, v.98, p.1196–1209, 2019.

DOMINGUES, D. S.; PAULI, E. D.; DE ABREU, J. E.; MASSURA, F. W.; CRISTIANO, V.; SANTOS, M. J.; NIXDORF, S. L. Detection of roasted and ground coffee adulteration by HPLC by amperometric and by post-column derivatization UV–Vis detection. *Food Chemistry.*, v.146, p.353–362, 2014.

EL-ABASSY, R. M.; DONFACK, P.; MATERNY, A. Discrimination between Arabica and Robusta green coffee using visiblemicro Raman spectroscopy and chemometric analysis. *Food Chemistry.*, v.126, p.1443-1448, 2011.

ELSHAFEY, R.; RADI, ABD-ELGAWAD. Electrochemical impedance sensor for herbicide alachlor based on imprintedpolymer receptor. *Journal of Electroanalytical Chemistry.*, v.813, p.171-177, 2018.

EROL, S.; ORAZEM, M. E. The influence of anomalous diffusion on the impedance response ofLiCoO<sub>2</sub>/C batteries. *Journal of Power Sources.*, v.293, p.57-64, 2015.

ESCANDAR, G. M.; OLIVIERI, A. C. Multi-way chromatographic calibration — A review. *Journal of Chromatography A.*, v.1587, p.2–13, 2019.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S.; VOLPE, P. L. O. Quimiometria I: calibração multivariada, um tutorial. *Quím. Nova.*, v.22, p.724-731, 1999.

FIGUEIREDO, L. P.; BORÉM, F. M.; ALMEIDA, M. R.; DE OLIVEIRA, L. F. C.; ALVES, A. P. C.; SANTOS, C. M.; RIOS, P. A. Raman spectroscopy for the differentiation of Arabic coffee genotypes. *Food Chemistry.*, V.288, p.262-267, 2019.

GARCIA-HERNANDEZ, C.; COMINO, C. S.; MARTÍN-PEDROSA, F.; RODRIGUEZ-MENDEZ, M. L.; GARCIA-CABEZON, C. Impedimetric electronic tongue based on

nanocomposites for the analysis of red wines: Improving the variable selection method. *Sensors & Actuators: B. Chemical.*, v.277, p.365-372, 2018.

GELADI, P.; NELSON, A.; LINDHOLM-SETHSON. Complex numbers in chemometrics Examples from multivariate impedance measurements on lipid monolayers *B. Anal Chim Acta.*, v.595, p.152-159, 2007.

GINER-SANZ, J. J.; ORTEGA, E.M.; PÉREZ-HERRANZ, V. Montecarlo based quantitative Kramers-Kronig test for PEMFC impedance spectrum validation. *International journal of hydrogen energy.*, v.40, p.279-293, 2015.

HIRSCHORN, B.; ORAZEM, M.E.; TRIBOLLET, B.; VIVIER, V.; FRATEUR, I.; MUSIANI, M. Determination of effective capacitance and film thickness from constant-phase-element parameters. *Electrochim. Acta.*, v.55, p.6218-6227, 2010.

ITO, Y.; SHIMAZAKI, J. O.; TSUGAWA, W.; LOEW, N.; SHITANDA, I.; LINE, C.; LABELLE, J.; SODE, K. Third generation impedimetric sensor employing direct electron transfer type glucose dehydrogenase. *Biosensors and Bioelectronics.*, v.129 p.189-197, 2019.

JACYNA, J.; KORDALEWSKA, M.; MARKUSZEWSKI, M. J. Design of Experiments in metabolomics - related studies: An overview. *Journal of Pharmaceutical and Biomedical Analysis.*, v.164, p.598-606, 2019.

KASHYAP, D.; DWIVEDI, P. K.; PANDEY, J. K.; KIM, Y. H.; KIM, G. M.; SHARMA, A.; GOEL, S. Application of electrochemical impedance spectroscopy in bio-fuel cell characterization: A review. *International journal of hydrogen energy.*, v.39, p.20159-20170, 2014.

KASHYAP, DIWAKAR.; DWIVEDI, P. K.; PANDEY, J. K.; HO KIM, Y.; KIM, G. M.; SHARMA, A.; GOEL, S. Application of electrochemical impedance spectroscopy in bio-fuel cell characterization: A review. *international journal of hydrogen energy.*, v.39, p.20159-20170, 2014.

KOSCIELNIAK, P.; WIECZOREK, M. Univariate analytical calibration methods and procedures. A review. *Analytica Chimica Acta.*, v.944, p.14-28, 2016.

KOŠICEK, K. M.; KVASTEK, K.; HORVAT-RADOŠEVIC, V. Different charge storage mechanisms at some carbon electrodes in redox active electrolyte revealed by electrochemical impedance spectroscopy. *Electrochimica Acta.*, v.195, p.77–84, 2016.

KRUPINSKA, K.; GELADI, P.; VIKHOLM-LUNDIN, I.; LINDHOLM-SETHSON, B. Detection of low levels of *Escherichia coli* by electrochemical impedance spectroscopy and singular value decomposition. *Chemometrics and Intelligent Laboratory Systems.*, v.163, p.49–54, 2017.

LASIA, A. *Electrochemical Impedance Spectroscopy and its Applications, Modern Aspects of Electrochemistry*. Springer., p.143-248, 2002.

LEE, L. C.; LIONG, C. Y.; JEMAIN, A. A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps. *Analyst.*, v.143, p.3526-3539, 2018.

LI, F.; YU, Z.; HAN, X.; LAI, R. Y. Electrochemical aptamer-based sensors for food and water analysis: A review. *Analytica Chimica Acta.*, v.1051, p.1-23, 2019.

LIU, C.; YANG, N.; YANG, Q.; AYED, C.; LINFORTH, R.; FISK, I. D. Enhancing Robusta coffee aroma by modifying flavour precursors in the green coffee bean. *Food Chemistry.*, v.281, p.8-17, 2019.

MACDONALD, D. D.; SIKORA, E.; ENGELHARDT, G. Characterizing electrochemical systems in the frequency domain. *Electrochimica Acta.*, V.43, p.87-107, 1998.

MAHMOUD, A.; ECHABAANE, M.; OMRI, K.; EL MIR, L.; CHAABANE, R. B. Development of an impedimetric non enzymatic sensor based on ZnO and Cu doped ZnO nanoparticles for the detection of glucose. *Journal of Alloys and Compounds.*, v.786, p.960-968, 2019.

MARTÍNEZ-BISBAL, M. C.; MESTRE, N. C.; MARTÍNEZ-MÁÑEZA, R.; BAUZÁ, J.; FILLOL, M. A. Microalgae degradation follow up by voltammetric electronic tongue impedance spectroscopy and NMR spectroscopy. *Sensors & Actuators: B. Chemical.*, v.281, p.44–52, 2019.

MISZCZYK, A.; DAROWICKI, K. Multivariate analysis of impedance data obtained for coating systems of varying thickness applied on steel. *Progress in Organic Coatings.*, v.77, p.2000–2006, 2014.

MONZÓN, P.; RAMÓN, J. E.; GANDÍA-ROMERO, J. M.; VALCUENDE, M.; SOTO, J.; PALACÍ-LÓPEZ, D. PLS multivariate analysis applied to corrosion studies on reinforced concrete. *Journal of Chemometrics.*, v.33, p.3096-3108, 2018.

MOU, Y.; ZHOU, L.; CHEN, W.; FAN, J.; ZHAO, X. Maximum correntropy criterion partial least squares. *Optik.*, v.165, p.137–147, 2018.

NAG, A.; AFSRIMANESH, N.; MUKHOPADHYAY, S. C. Impedimetric microsensors for biomedical applications. *Current Opinion in Biomedical Engineering.*, v.9, p.1–7, 2019.

PAJKOSSY, T.; JURCZAKOWSKI, R. Electrochemical impedance spectroscopy in interfacial studies. *Current Opinion in Electrochemistry.*, p.53–58, 2017.

PAULI, E. D.; BARBIERI, F.; GARCIA, P. S.; MADEIRA, T. B.; JUNIOR, V. R. A.; SCARMINIO, I. S.; CAMARA, C. A. P, NIXDORF, S. L. Detection of ground roasted coffee adulteration with roasted soybean and wheat. *Food Research International.*, v.61, p.112-119, 2014.

PHAL, S.; LINDHOLM-SETHSON, B.; GELADI, P.; SHCHUKAREV, A.; TESFALIDET, S. Determination of methotrexate in spiked human blood serum using multi-frequency electrochemical immittance spectroscopy and multivariate data analysis. *Analytica Chimica Acta.*, V.987, p.15-24, 2017.

PHAL, S.; SHATRI, B.; BERISHA, A.; GELADI, P.; LINDHOLM-SETHSON, B.; TESFALIDET, S. Covalently electrografted carboxyphenyl layers onto gold surface serving as a platform for the construction of an immunosensor for detection of methotrexate. *Journal of Electroanalytical Chemistry.*, v.812, p.235–243, 2018.

PRIETO, P.; RUEDA, M.; ALVAREZ-MALMAGRO, J. Electrochemical Impedance Spectroscopy analysis of na adsorption process with a coupled preceding chemical step. *Electrochimical Acta.*, v.232, p.164-173, 2017.

PRODROMIDIS, M. I. Impedimetric immunosensors - A review. *Electrochimica Acta.*, v.55, p.4227–4233, 2010.

RAMOS, P. M.; JANEIRO, F. M.; GIRÃO, P. S. Uncertainty evaluation of multivariate quantities: A case study on electrical impedance. *Measurement.*, v.78, p.397–411, 2016.

RANDVIIR, E. P. A cross examination of electron transfer rate constants for carbon screen-printed electrodes using Electrochemical Impedance Spectroscopy and cyclic voltammetry. *Electrochimica Acta.*, v.286, p.179-186, 2018.

SACCO, A. Electrochemical impedance spectroscopy: Fundamentals and application in dye-sensitized solar cells. *Renewable and Sustainable Energy Reviews.*, v.79, p.814–829, 2017.

SEZER, B.; APAYDIN, H.; BILGE, G.; BOYACI, I. H. Coffee arabica adulteration: Detection of wheat, corn and chickpea. *Food Chemistry.*, v.264, p.142-148, 2018.

SHARMA, V.; KUMAR, R. Trends of chemometrics in bloodstain investigations. *Trends in Analytical Chemistry.*, v.107, p.181-195, 2018.

SILVA, W.; GHICA, M. E.; AJAYI, R.F.; IWUOHA, E. I.; BRETT, C. M. A. Impedimetric sensor for tyramine based on gold nanoparticle doped-poly(8-anilino-1-naphthalenesulphonic acid) modified gold electrodes. *Talanta.*, v.195, p.604-612, 2019.

SINGH, S.; MOUDGIL, A.; MISHRA, N.; DAS, S.; MISHRA, P. Vancomycin functionalized WO<sub>3</sub> thin film-based impedance sensor for efficient capture and highly selective detection of Gram-positive bacteria. *Biosensors and Bioelectronics.*, v.136, p.23-30, 2019.

SINGH, M.; SEMWAL, S.; KUMAR, A.; SINGH, S. Impedance modeling for classification of flavored green teas. *Turk J Elec Eng & Comp Sci.*, v.23, p.2208–2214, 2015.

TESFALIDET, S.; GELADI, P.; SHIMIZU, K.; LINDHOLM-SETHSON, B. Detection of methotrexate in a flow system using electrochemical impedance spectroscopy and multivariate data analysis. *Analytica Chimica Acta.*, V.914, p.1-6, 2016.

TOCI, A. T.; FARAH, A.; PEZZA, H. R.; PEZZA, L. Coffee Adulteration: More than Two Decades of Research. *CRITICAL REVIEWS IN ANALYTICAL CHEMISTRY.*, v.46, p.83–92, 2016.

TRAN, M. T. T.; TRIBOLLET, B.; VIVIER, V.; ORAZEM, M. E. On the impedance response of reactions influenced by mass transfer. *Russian Journal of Electrochemistry.*, v.53, p.932–940, 2017.

VALDERRAMA, P.; BRAGA, J. W. B.; POPPI, R. J. Estado da arte de figuras de mérito em calibração multivariada. *Quím. Nova.*, v.32, p.1278-1287, 2009.

VIDAKOVIC-KOCH, T.; MITTAL, V. K.; DO, T. Q. N.; VARNICIC, M.; SUNDMACHER, K. Application of electrochemical impedance spectroscopy for studying of enzyme kinetics. *Electrochimica Acta.*, v.110, p.94–104, 2013.

WANG, Z.; LI, J.; WANG, Y.; WANG, Z. An EIS analysis on corrosion resistance of anti-abrasion coating. *Surfaces and Interfaces.*, v.6, p.33-39, 2017.

WOLD, S.; SJOSTROM, M.; ERIKSSON, L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems.*, v.58, p.109–130, 2001.

WU, K.; YANG, J.; QIU, X.; XU, JIN-MEI.; ZHANG, QIAN-QIAN.; JIN, J.; HUANG, QUAN-CHAOZ. Study of spinel  $\text{Li}_4\text{Ti}_5\text{O}_{12}$  electrode reaction mechanism by electrochemical impedance spectroscopy. *Electrochimica Acta.*, v.108, p.841–851, 2013.

XU, Y.; XIE, X.; DUAN, Y.; WANG, L.; CHENG, Z.; CHENG, JING. A review of impedance measurements of whole cells, *Biosens. Bioelectron.*, v.77, p.824-836, 2016.

YADAV, P.; PANDEY, K.; BHATT, V.; KUMAR, M.; KIM, J. Critical aspects of impedance spectroscopy in silicon solar cell characterization: A review. *Renewable and Sustainable Energy Reviews.*, V.76, p.1562–1578, 2017.

YADAV, P.; PANDEY, K.; BHATT, V.; KUMAR, M.; KIM, J. Critical aspects of impedance spectroscopy in silicon solar cell characterization: A review. *Renewable and Sustainable Energy Reviews.*, v.76, p.1562–1578, 2017.

YU, P.; LOW, M. Y.; ZHOU, W. Design of experiments and regression modelling in food flavour and sensory analysis: A review. *Trends in Food Science & Technology.*, v.71, p.202–215, 2018.

YUN-YING, S.; JING-YAN, L.; XIAO-LI, C. Progress and Applications of Multivariate Calibration Model Transfer Methods. Chinese J. Anal. Chem., v.47, p.479-487, 2019.

ZLATEV, R.; STOYTCHEVA, M.; VALDEZ, B.; MONTERO, G.; TOSCANO, L. Simple impedimetric sensor for rapid lipase activity quantification. Talanta., v.203, p.161-167, 2019.

## **ANEXOS**

## Anexo I – Gráficos referentes aos modelos PLS-DA

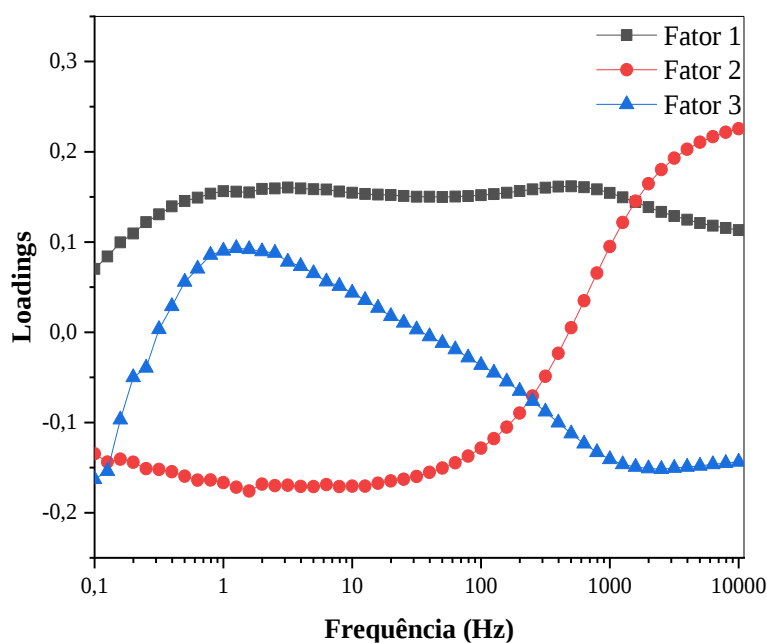


Figura 25 - *Loading* da componente real dos três primeiros fatores do modelo PLS-DA construído com os dados de impedância expressos como números complexos ( $Z' + jZ''$ ).

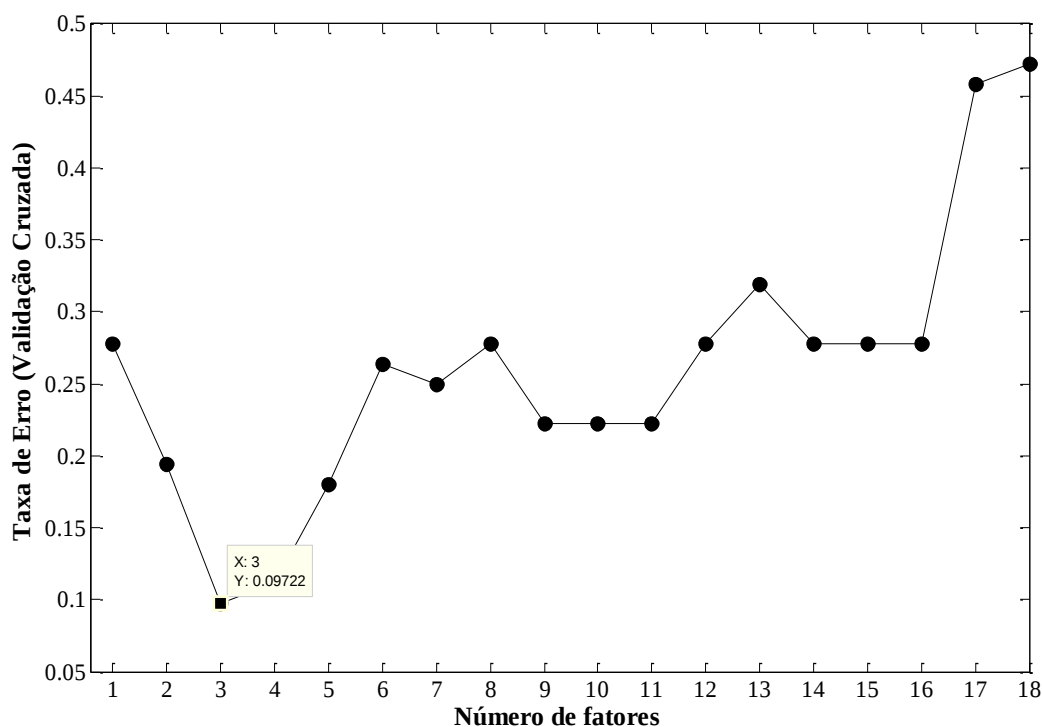


Figura 26 - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da impedância expressos como números complexos ( $Z' + jZ''$ ).

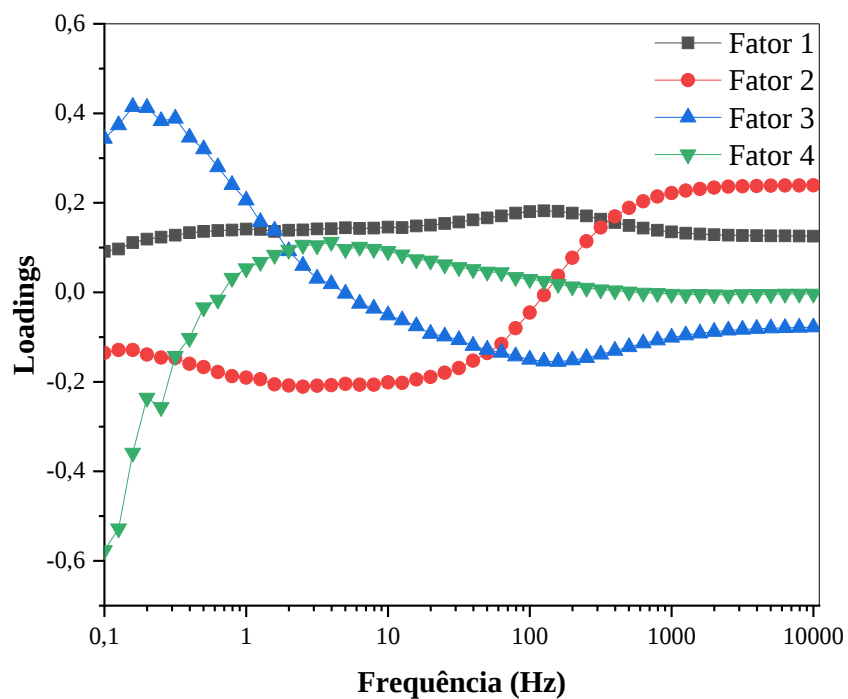


Figura 27 - *Loading* dos quatro primeiros fatores do modelo PLS-DA construído com os dados da parte real da impedância ( $Z'$ ).

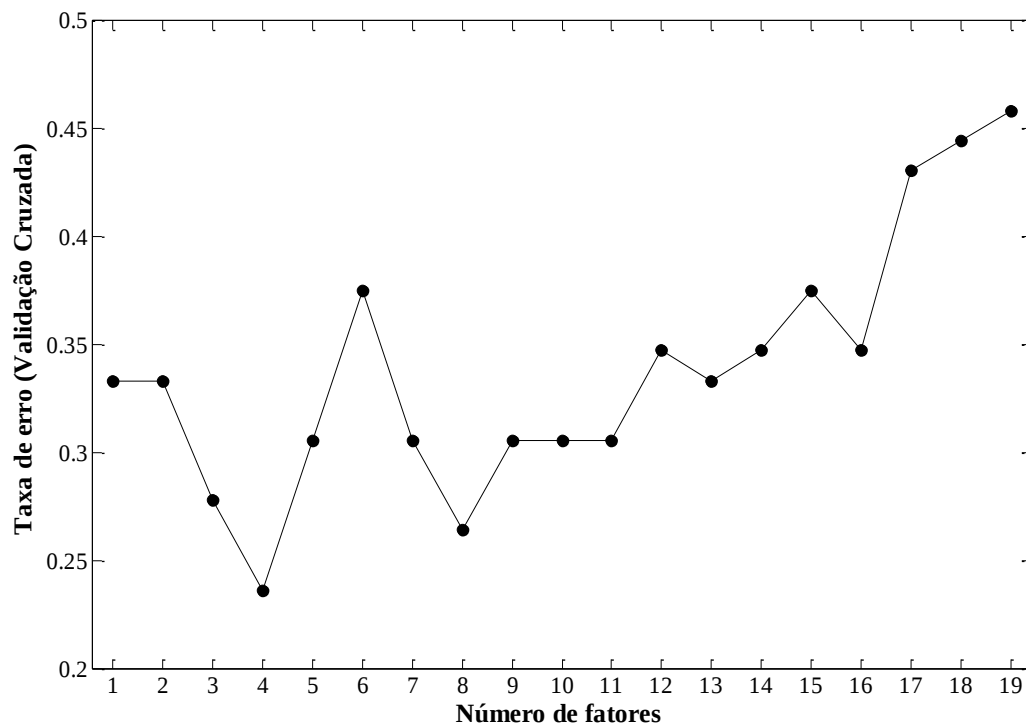


Figura 28 - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da parte real da impedância ( $Z'$ ).

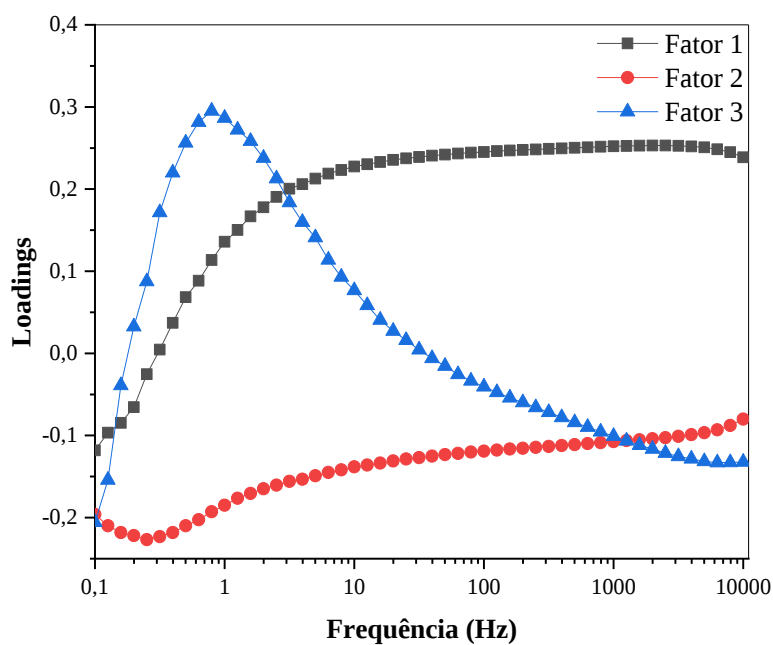


Figura 29 - *Loading* dos quatro primeiros fatores do modelo PLS-DA construído com os dados da parte imaginária da impedância ( $Z''$ ).

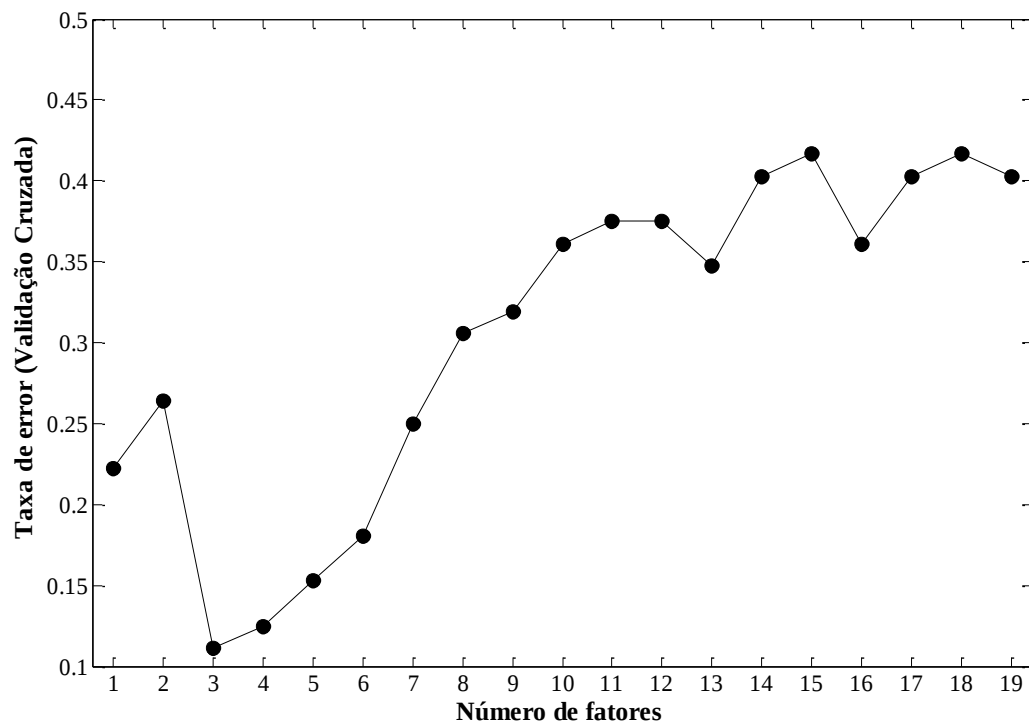


Figura 30 - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da parte imaginária da impedância ( $Z''$ ).

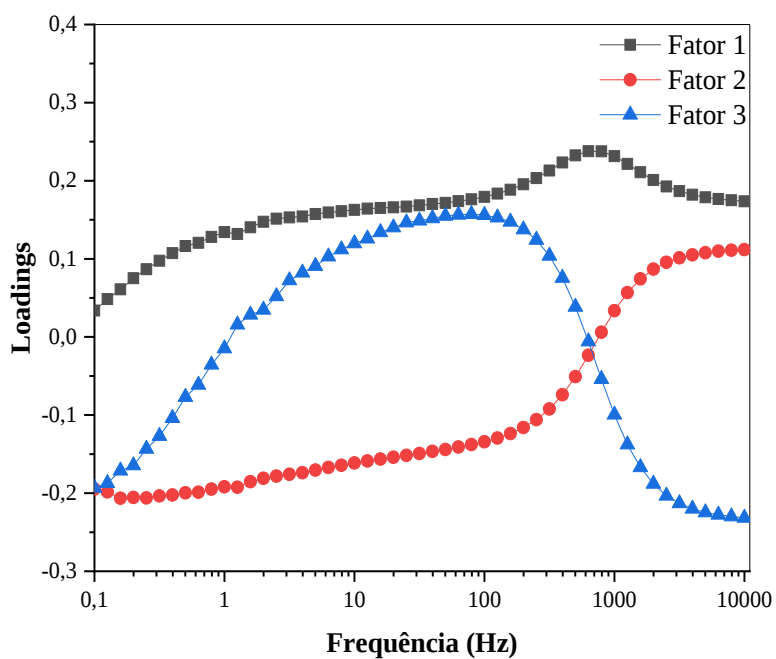


Figura 31 - *Loading* dos três primeiros fatores do modelo PLS-DA construído com os dados do módulo da impedância ( $|Z|$ ).

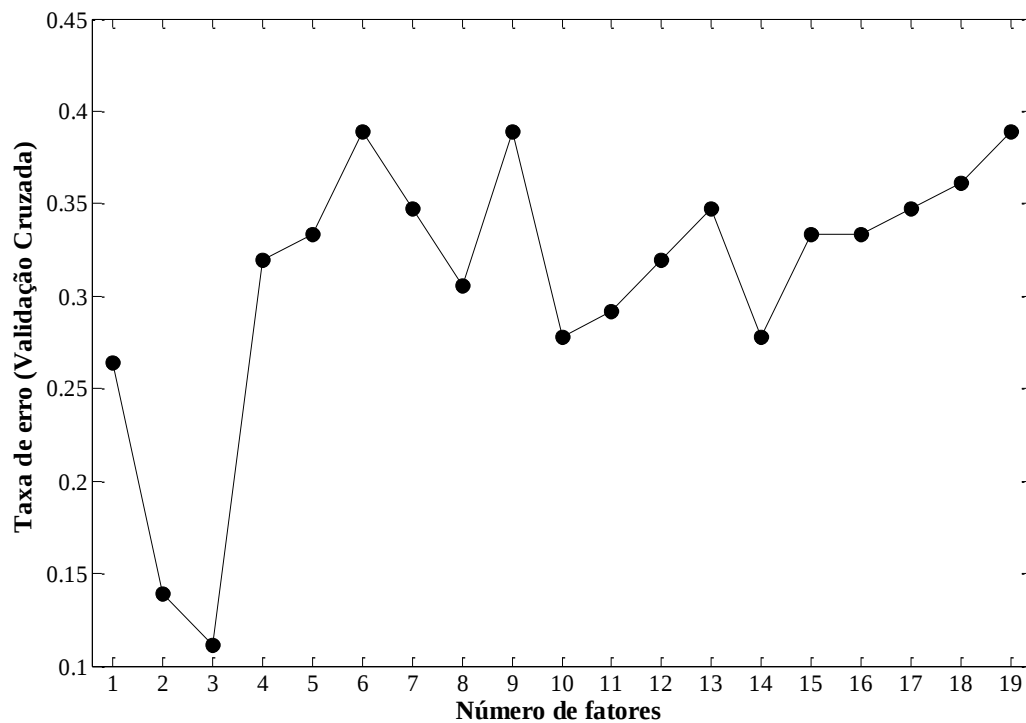


Figura 32 - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados do módulo da impedância ( $|Z|$ ).

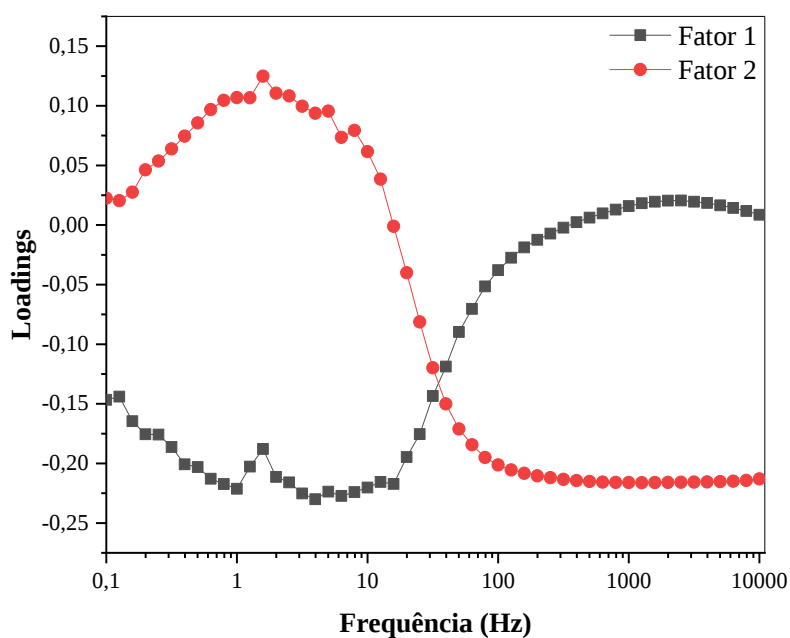


Figura 33 - *Loading* dos quatro primeiros fatores do modelo PLS-DA construído com os dados da fase da impedância ( $\varphi$ ).

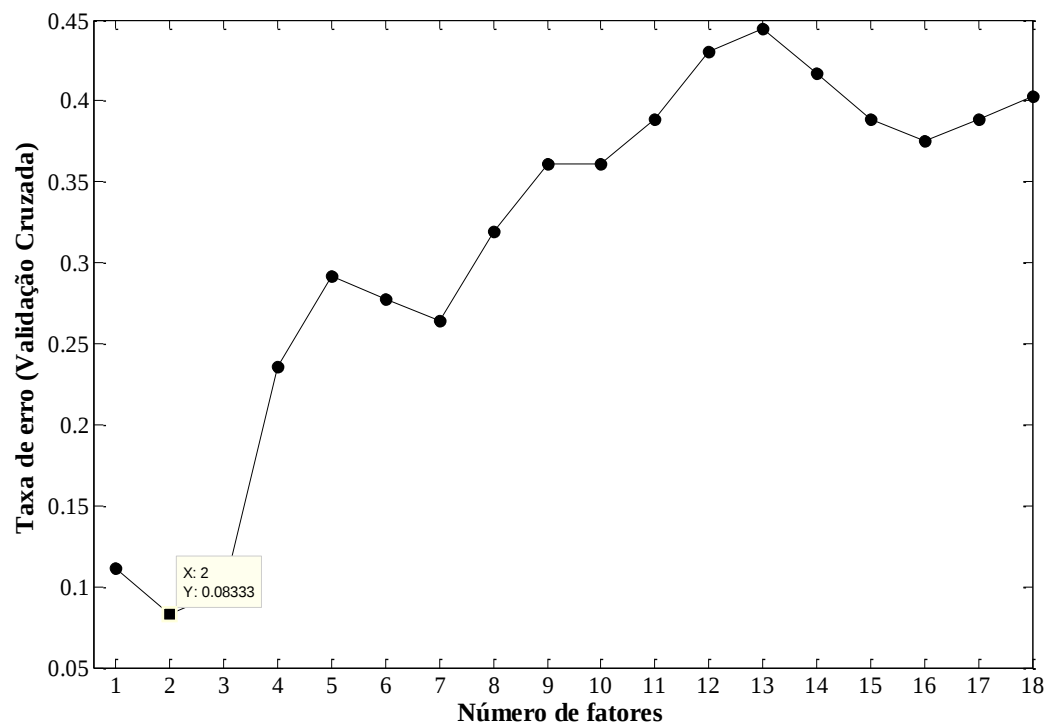


Figura 34 - Taxa de erro em função do número de fatores do modelo PLS-DA construído com os dados da fase da impedância ( $\varphi$ ).

## Anexo II – Gráficos referentes aos modelos SIMCA

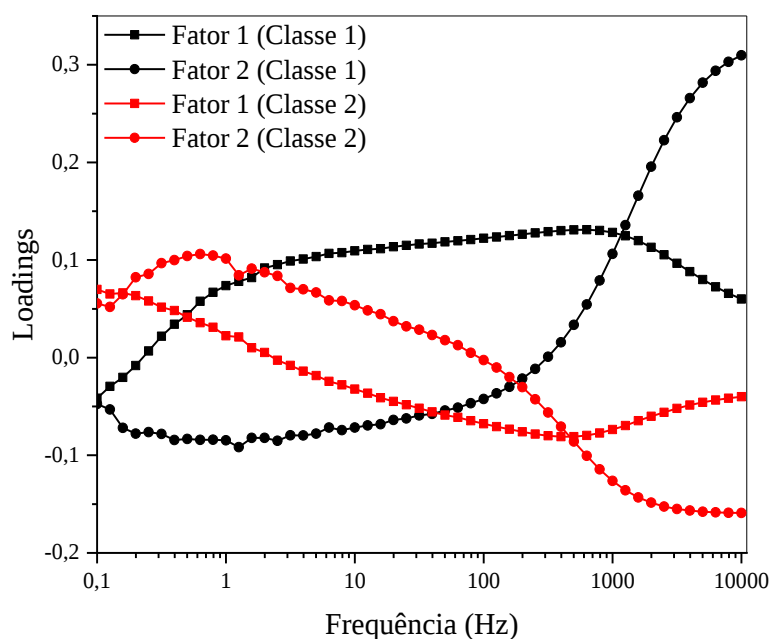


Figura 35 - *Loading* das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados da impedância expressa como número complexo ( $Z' + jZ''$ ).

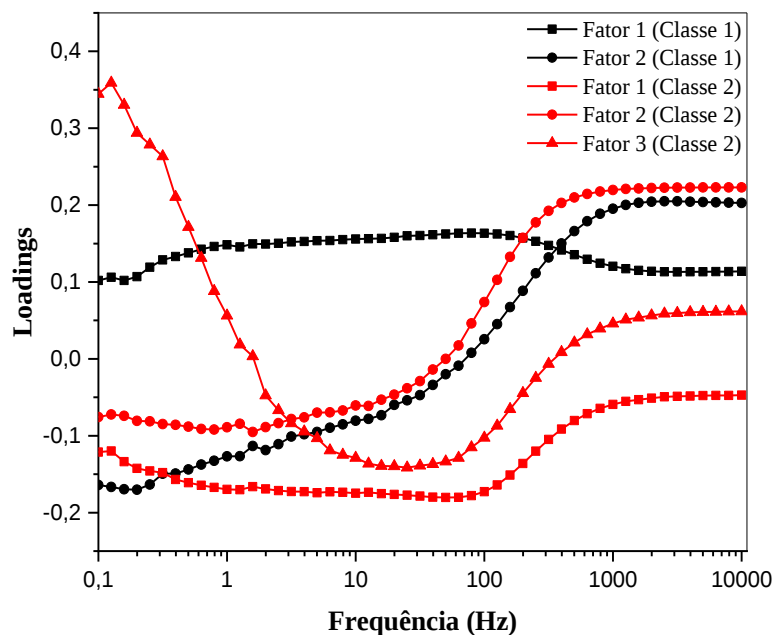


Figura 36 - *Loading* das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados da parte real da impedância ( $Z'$ ).

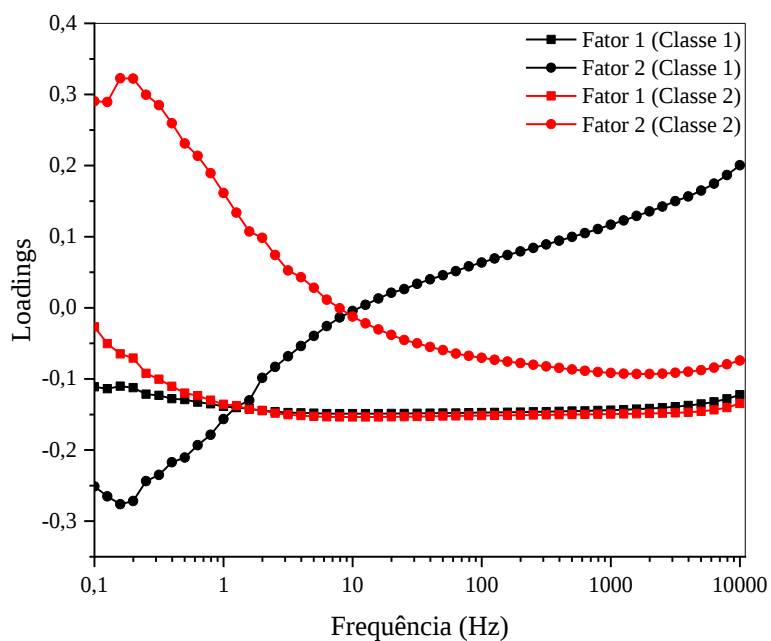


Figura 37 - *Loading* das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados da parte imaginária da impedância ( $Z''$ ).

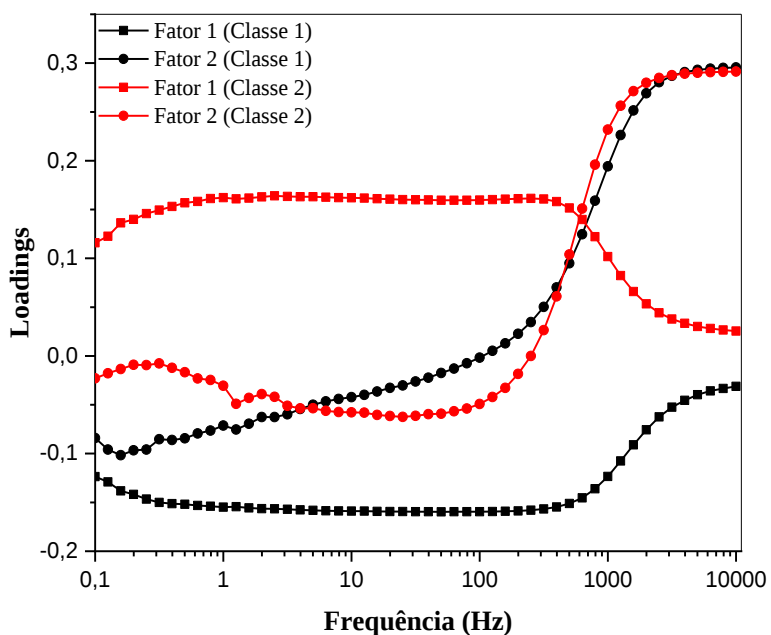


Figura 38 - *Loading* das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados do módulo da impedância ( $|Z|$ ).

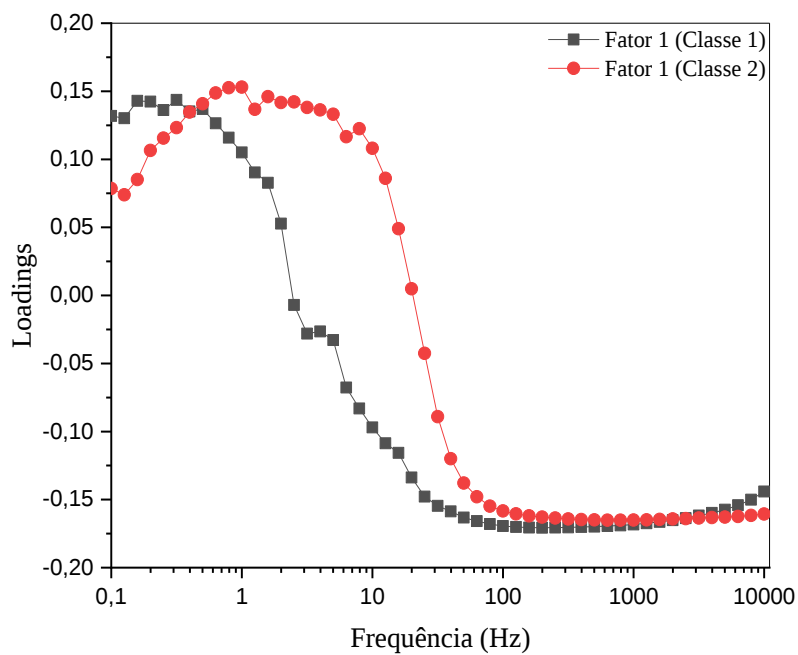
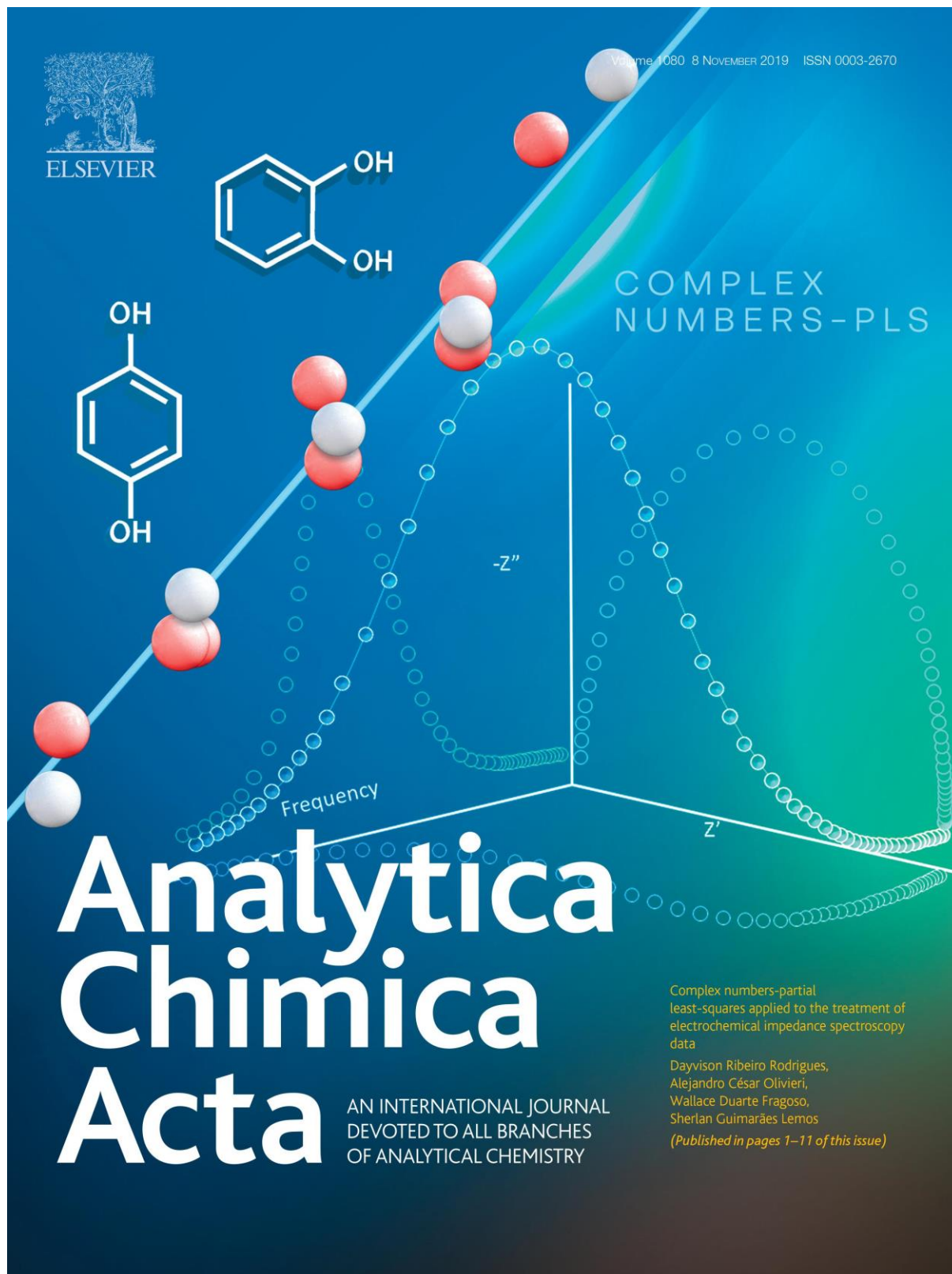


Figura 39 - *Loading* das PCs usadas na construção do modelo SIMCA de cada classe a partir dos dados do ângulo de fase da impedância ( $\varphi$ ).



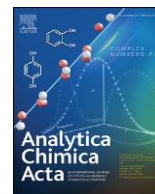


Contents lists ScienceDirect

Analytica Chimica Acta

Journal

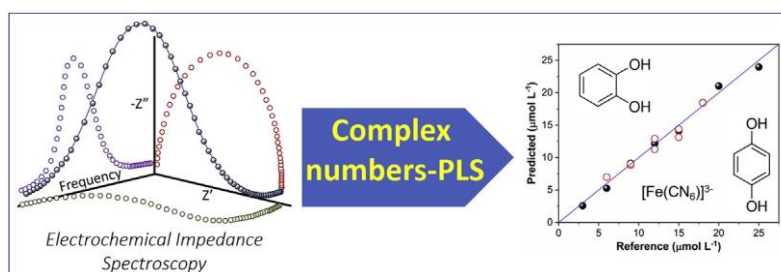
www.elsevier.com/locate/acta



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

Dayvison Ribeiro Rodrigues<sup>a</sup>, Alejandro Cesar Olivieri<sup>b</sup>, Wallace Duarte Fragoso<sup>a</sup>, Sherlan Guimaraes Lemos<sup>a,\*</sup>

<sup>a</sup>Grupo de Estudos Avançados em Química Analítica, Departamento de Química, Universidade Federal da Paraíba, 58051-970, Joao Pessoa, PB, Brazil~<sup>b</sup>Departamento de Química Analítica, Facultad de Ciencias Bioquímicas y Farmaceuticas, 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 dihydroxybenzenes determination.

### abstract 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

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 numbersePLS was evaluated in the individual determination of two well-known redox probes: ferrocyanide and hydroquinone. The predictive ability of complex numbersePLS 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 ePLS 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 numbersePLS.

© 2019 Published by Elsevier B.V.

## 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

\* Corresponding author.

E-mail addresses: [sherlan@quimica.ufpb.br](mailto:sherlan@quimica.ufpb.br), [sgleamos@pq.cnpq.br](mailto:sgleamos@pq.cnpq.br) (S.G. Lemos).

<https://doi.org/10.1016/j.aca.2019.07.047>

0003-2670/© 2019 Published by Elsevier B.V.

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 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  $e_k$  one for each  $k$  frequency tested and obtained for  $n$  samples, the investigation of this type of data array ( $n \times k$  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 e logarithm of the absolute impedance ( $\log|Z_j|$ ) or the phase angle plotted with log frequency on the X-axis e 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 leastsquares 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:

$$X \approx TP' + E \quad (2) \quad Y \approx UQ' + F \quad (3)$$

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 [15e17]. The predictive ability of complex numbers-PLS models was assessed in comparison to PLS applied to EIS data as real numbers  $e$  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  $e$  applied potential and amplitude  $e$  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 numbersePLS 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  $\pm$  1 nm 4.5 nm  $\pm$  0.5 nm 3e6 mm, Sigma-Aldrich), and mineral oil. All chemicals were of analytical grade, used without further purification, and purchased from SigmaeAldrich, except for the mineral oil, which was obtained from Specsol. Tap water samples were collected from our lab.

Ultrapure water (18 M $\Omega$  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 ( $25 \pm 1$  °C) in a 20 mL cell. The electrochemical impedance spectroscopy was carried out within the frequency range 10 kHz to 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  $4.0 \times 10^{-4}$  to  $1.4 \times 10^{-1}$  mmol L<sup>-1</sup>, 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 established with square wave voltammetry: i) the oxidation peak potential; ii) the onset potential; iii) a 200 to 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<sup>-1</sup> 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 a  $\alpha$  of 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 ( $E_a$ ) e at  $\pm 0.100$  V (onset potential), at  $\pm 0.220$  V (peak potential) and at  $\pm 0.450$  V (overpotential) e by employing three amplitudes: 1 mV, 5 mV and 10 mV. Ferrocyanide concentrations were in the range  $40\text{e}140\text{ mmol L}^{-1}$ . 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.

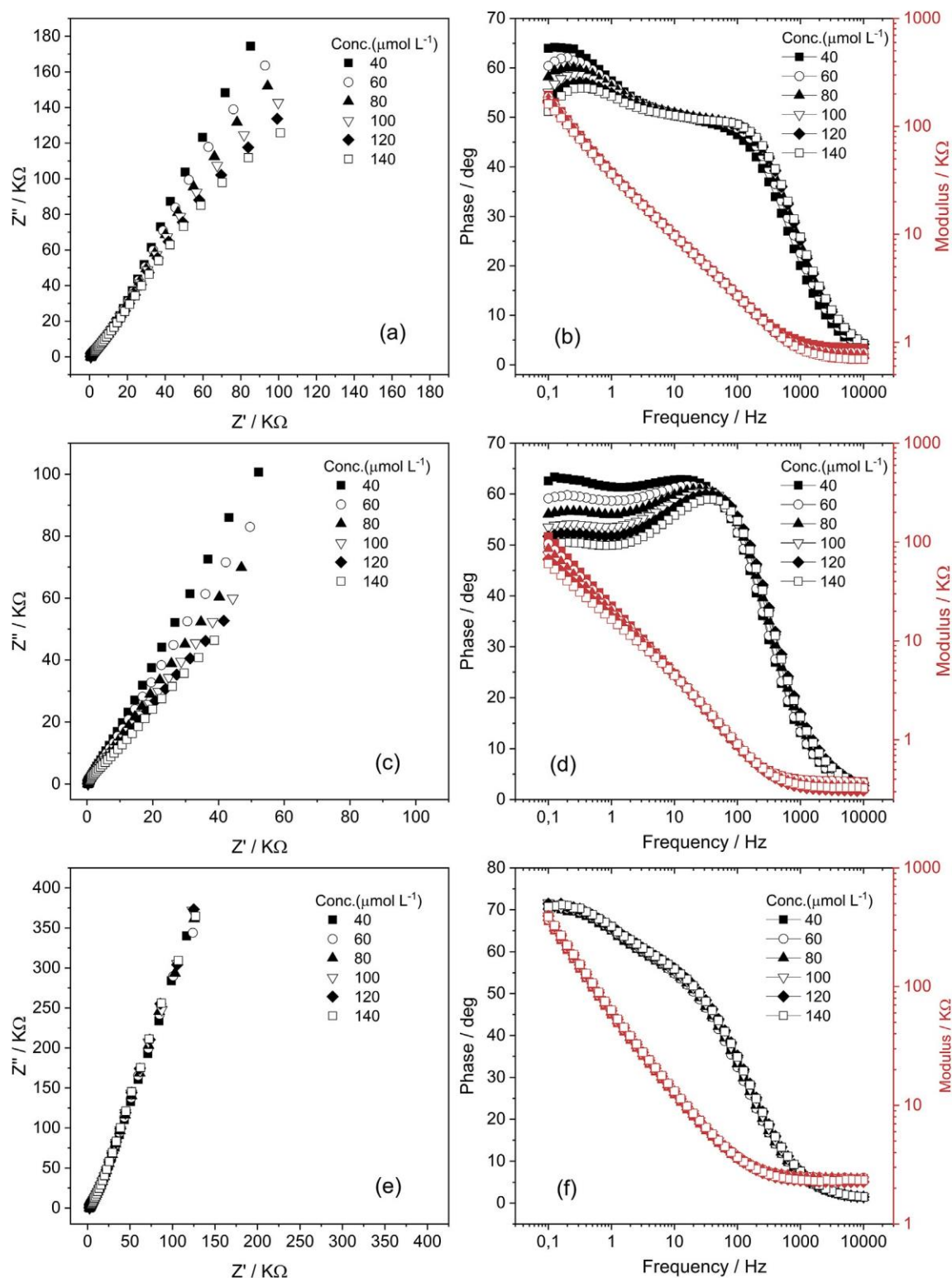
As can be observed, there is a dependence on the behavior of spectra with applied potential resulting from the different charge transfer and mass-transfer scenarios. At  $E_a \approx \pm 0.100$  V (Fig. 1a), the Nyquist plot presents straight lines of approximately 45° in the high frequency domain, characteristic 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  $E_a \approx \pm 0.220$  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  $E_a \approx \pm 0.450$  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  $E_a \approx \pm 0.100$  V), the charge-transfer resistance is negligibly small for potentials at which the current is limited by mass transfer (at  $E_a \approx \pm 0.220$  V), and in the case of sufficiently large overpotential ( $E_a \approx \pm 0.450$  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/quasireversible redox reaction such as ferrocyanide oxidation [19].

Regarding Bode plots in Fig. 1, one can see similar magnitude profiles for  $E_a \approx \pm 0.100$  V and  $E_a \approx \pm 0.450$  V, showing resistive ( $\log|Z_j|$  slope  $\approx 0$ ) to capacitive responses ( $\log|Z_j|$  slope  $\approx 1$ ) 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  $E_a \approx \pm 0.220$  V. However, at lower frequencies  $\log|Z_j|$  slopes tend to decrease towards 0.5 as ferrocyanide concentration increases. Phase-angle ( $\phi$ ) profiles confirm the resistive to capacitive response change behavior ( $\phi \approx 0^\circ/90^\circ$ ) at  $E_a \approx \pm 0.450$  V, what is typical for electrolyte resistance and double-layer capacitive impedance responses,<sup>19</sup> also showing no changes on spectra as ferrocyanide concentration increases. At  $E_a \approx \pm 0.220$  V, a single time constant is observed with characteristic frequency situated at  $10^1$ – $10^2$  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  $45^\circ$ , typical for a semi-infinite diffusion controlled faradaic reaction [19]. A single time constant is also observed for phase-angle profiles at  $E_a \approx \pm 0.100$  V, but with characteristic frequency found at  $10^2$ – $10^3$  Hz. As observed for  $E_a \approx \pm 0.220$  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.  $E_a \approx \pm 0.100$  V, (a) and (b);  $E_a \approx \pm 0.220$  V, (c)



and (d);  $E_a \approx 0.450$  V, (e) and (f). Amplitude  $\approx 10$  mV.

$E_a \approx 0.100$  V, but only at frequencies below 1 Hz (Fig. 1b).

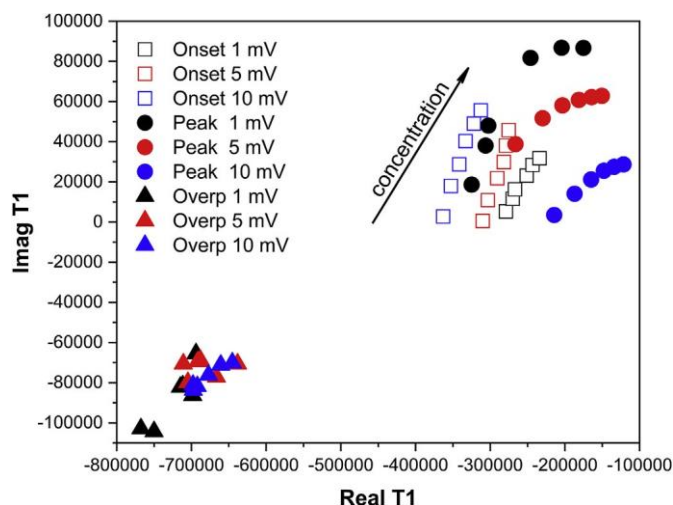
The 54  $\times$  51 complex impedance matrix and impedance spectra obtained at six ferrocyanide concentrations, and at nine different combinations of three applied potentials and three amplitudes were mean-centered

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 ( $\theta$ ). Bolds values in Table 1 indicate the models with better analytical performance e 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 ¼ 0.100 V; 0 ¼ 0.220 V; p1 ¼ 0.450 V; ii) amplitude, at three levels: -1 ¼ 1 mV; 0 ¼ 5 mV; p1 ¼ 10 mV; iii) data presentation, at five levels: -1 ¼ complex numbers ( $Z' + jZ''$ ); 0.5 ¼ real part of complex impedance ( $Z'$ ); 0 ¼ imaginary part of complex impedance



( $Z''$ );  $\rho 0.5 \frac{1}{4}$  absolute impedance ( $jZ_j$ );  $\rho 1 \frac{1}{4}$  phase angle (4); and iv) data pretreatment, at two levels: -1  $\frac{1}{4}$  none;  $\rho 1 \frac{1}{4}$  mean-centering.

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  $e$  the estimated effect divided by its standard error  $e$  that is compared to a vertical line  $e$  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 meancentering of spectra (Fig. 3c). Conversely, higher RMSECV values are obtained by applying the overpotential condition  $e$   $E_a \frac{1}{4} \rho 0.450$  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 nonlinear decrease of the number of latent variables as the applied potential is changed from the low level to the high level, with a

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

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 e Shapiro-Wilk, Kolmogorov-Smirnov, and D'AgostinoK squared e 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 ( $n \approx 90$ ). 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 mmol L<sup>-1</sup> [Fe(CN)<sub>6</sub>]<sup>4-</sup>.

| Data treatment | Data form       | E <sub>a</sub> ¼ 0.100 V |         |         | E <sub>a</sub> ¼ 0.220 V |         |         | E <sub>a</sub> ¼ 0.450 V |        |        |
|----------------|-----------------|--------------------------|---------|---------|--------------------------|---------|---------|--------------------------|--------|--------|
|                |                 | 1 mV                     | 5 mV    | 10 mV   | 1 mV                     | 5 mV    | 10 mV   | 1 mV                     | 5 mV   | 10 mV  |
| Mean-center    | Z'pZ"           | 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 <sub>0</sub>  | 18 (3)                   | 3.4 (1) | 11 (2)  | 20 (2)                   | 25 (1)  | 1.7 (2) | 44 (1)                   | 19 (1) | 18 (1) |
|                | Z <sub>00</sub> | 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 <sub>j</sub>  | 8.1 (1)                  | 12 (1)  | 2.9 (1) | 11 (1)                   | 6.0 (2) | 2.0 (2) | 13 (4)                   | 18 (1) | 15 (1) |
|                | j <sub>j</sub>  | 4.0 (1)                  | 1.3 (3) | 2.9 (2) | 13 (2)                   | 2.7 (3) | 2.7 (3) | 28 (1)                   | 27 (1) | 12 (2) |
|                | 4               |                          |         |         |                          |         |         |                          |        |        |
| None           | Z'pZ"           | 17 (3)                   | 3.6 (3) | 6.8 (2) | 12 (3)                   | 9.4 (3) | 4.9 (4) | 26 (2)                   | 37 (1) | 24 (4) |
|                | Z <sub>0</sub>  | 40 (1)                   | 5.0 (3) | 18 (2)  | 14 (2)                   | 5.4 (3) | 2.6 (4) | 42 (1)                   | 36 (1) | 16 (3) |
|                | Z <sub>00</sub> | 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) |
|                | 4               | 5.3 (3)                  | 3.8 (3) | 7.8 (3) | 7.4 (3)                  | 5.6 (4) | 3.5 (4) | 40 (1)                   | 27 (2) | 14 (3) |
|                |                 |                          |         |         |                          |         |         |                          |        |        |

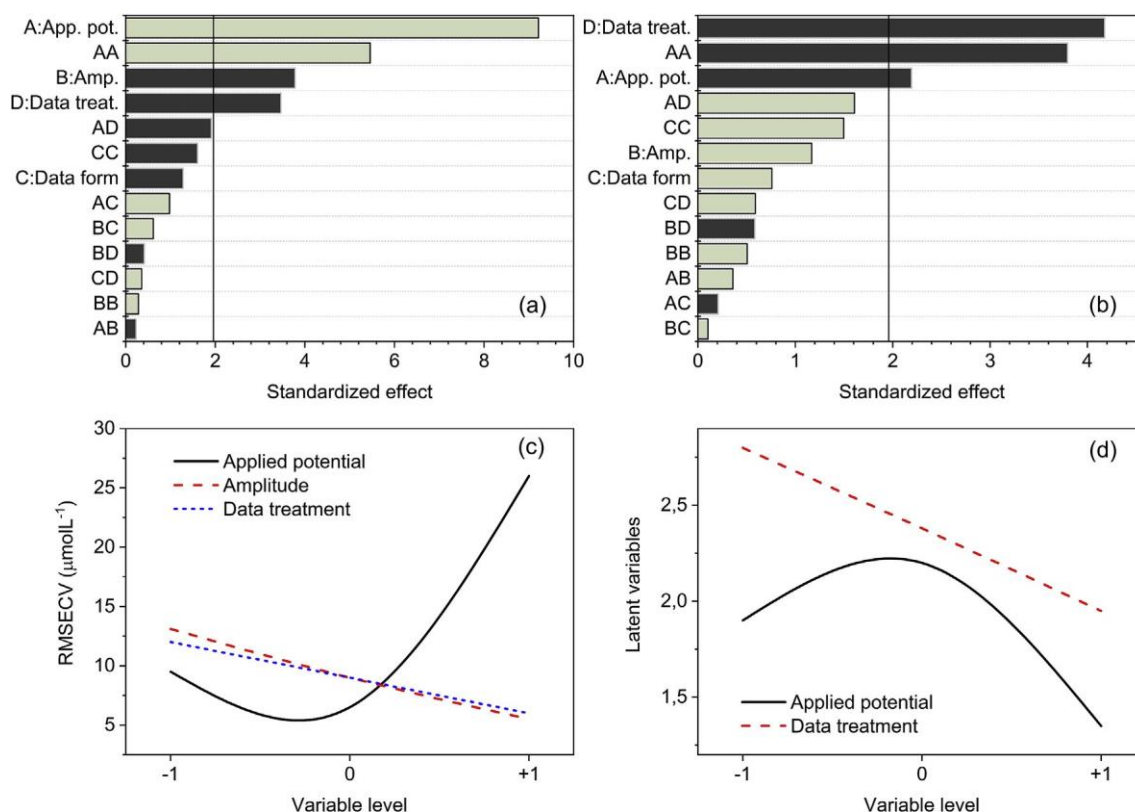


Fig. 3. (a) and (b): Pareto charts of the ANOVA performed on the responses RMSECV and number of latent variables, respectively. Gray bar: positive effect; black bar: negative effect. Where: A: applied potential; B % amplitude; C % data presentation; D % data pretreatment. (c) and (d): variation of responses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

parametric tests equivalent to ANOVA e Kruskal-Wallis and Mood's median e 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, meancentering provides models with better predictions and fewer latent variables when compared to the models based on unprocessed spectra. The better performance obtained with meancentering 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 meancentered 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  $30\text{e}150\text{ mmol L}^{-1}$ . 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  $\text{Fe}(\text{CN})_6^{3/4-}$  redox couple have already been reported for different electrode surface morphologies

such as other carbon surfaces and nanostructured porous electrodes [19,22e24] 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 wellknown 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  $\pm 0.110$  V (onset potential), at  $\pm 0.225$  V (peak potential), and at  $\pm 0.450$  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  $40\text{e}220\text{ mmol L}^{-1}$ . 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. 5aed). 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

lowfrequency 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 (mmol L <sup>-1</sup> ) | Z' þ jZ'' (1.1; 2) | Z' (1.7; 2) | Z'' (0.9; 5) | jZj (2.0; 2) | 4 (2.7; 3) |
|------------------------------------------|--------------------|-------------|--------------|--------------|------------|
| 40                                       | 40 þ 1j            | 44          | 39           | 38           | 42         |
| 60                                       | 58 e 2j            | 59          | 58           | 58           | 57         |
| 80                                       | 81 þ 2j            | 80          | 80           | 82           | 81         |
| 100                                      | 100 e 1j           | 100         | 100          | 101          | 101        |
| 120                                      | 121 þ 0j           | 120         | 120          | 119          | 118        |
| 140                                      | 139 e 1j           | 140         | 140          | 137          | 145        |

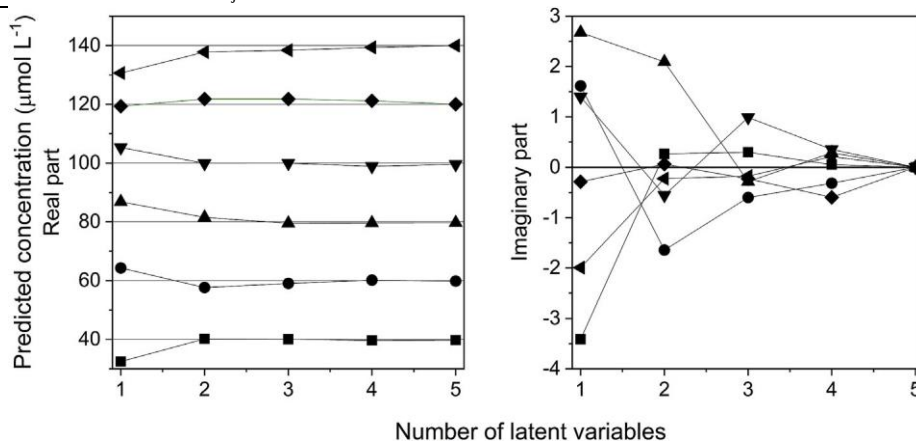


Fig. 4. Behavior of the real and imaginary parts of the prediction results with the number of latent variables included in the complex numbers-PLS model.

condition. Again, spectra obtained at overpotential (Fig. 5eef) 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 ( $E_a \approx 0.110$  V) and peak potential ( $E_a \approx 0.225$  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:

-10  $\pm$  0.1105 mV; V; þ11  $\pm$  0.22510 mV; iii) data presentation, at V; ii) amplitude, at three levels: -1 five levels; 1 mV;

-1  $\pm$  complex numbers ( $Z' þ jZ''$ ); 0.5  $\pm$  real part of complex impedance ( $Z'$ ); 0  $\pm$  imaginary part of complex impedance ( $Z''$ ); þ0.5  $\pm$  absolute impedance

(jZj);  $p1 \frac{1}{4}$  phase angle (4); and iv) data pretreatment, at two levels:  $-1 \frac{1}{4}$  none;  $p1 \frac{1}{4}$  mean-centering. Fig. 6 presents the Pareto charts obtained from analysis of RMSECV (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 firstorder 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 meancentering of spectra. However, the occurrence of the AB secondorder 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  $3 \times 10^{-5}$  mmol L<sup>-1</sup> was obtained at the adequate experimental conditions previously established 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

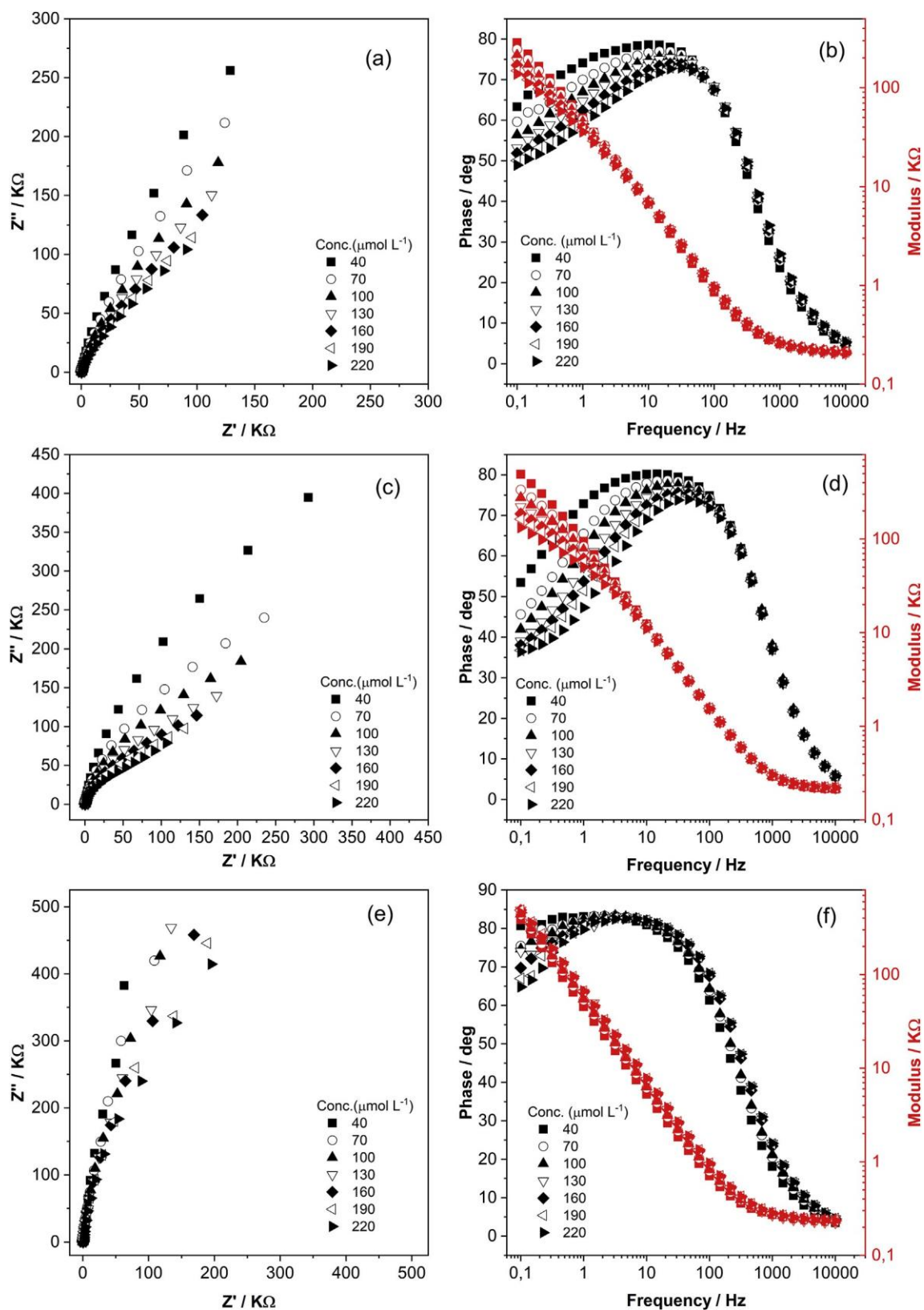


Fig. 5. Nyquist (a, c, e) and Bode (b, d, f) plots for the determination of hydroquinone at carbon nanotube paste electrode and different applied potentials.  $E_a$   $\approx$  0.110 V, (a) and (b);  $E_a$   $\approx$  0.225 V, (c) and (d);  $E_a$   $\approx$  0.450 V, (e) and (f). Amplitude  $\approx$  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: 40e220 mmol L<sup>-1</sup>.

| Data treatment | Data form       | E <sub>s</sub> % p0.110 V |         |         | E <sub>s</sub> % p0.225 V |         |         |
|----------------|-----------------|---------------------------|---------|---------|---------------------------|---------|---------|
|                |                 | 1 mV                      | 5 mV    | 10 mV   | 1 mV                      | 5 mV    | 10 mV   |
| Mean-center    | Z' p 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 <sub>00</sub> | 11 (1)                    | 2.9 (5) | 9.4 (2) | 7.4 (2)                   | 4.8 (4) | 19 (3)  |
|                | j <sub>j</sub>  | 8.3 (1)                   | 11 (5)  | 9.0 (2) | 1.5 (5)                   | 3.6 (2) | 22 (1)  |
| None           | 4               | 7.4 (3)                   | 18 (1)  | 8.8 (3) | 13 (1)                    | 2.5 (4) | 15 (2)  |
|                | Z' p 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 <sub>00</sub> | 24 (4)                    | 9.9 (4) | 9.8 (4) | 8.8 (4)                   | 5.1 (5) | 6.6 (5) |
|                | j <sub>j</sub>  | 12 (5)                    | 31 (4)  | 9.0 (2) | 5.0 (4)                   | 8.3 (4) | 1.2 (5) |
|                | 4               | 54 (2)                    | 26 (3)  | 8.8 (3) | 14 (2)                    | 5.9 (5) | 20 (2)  |

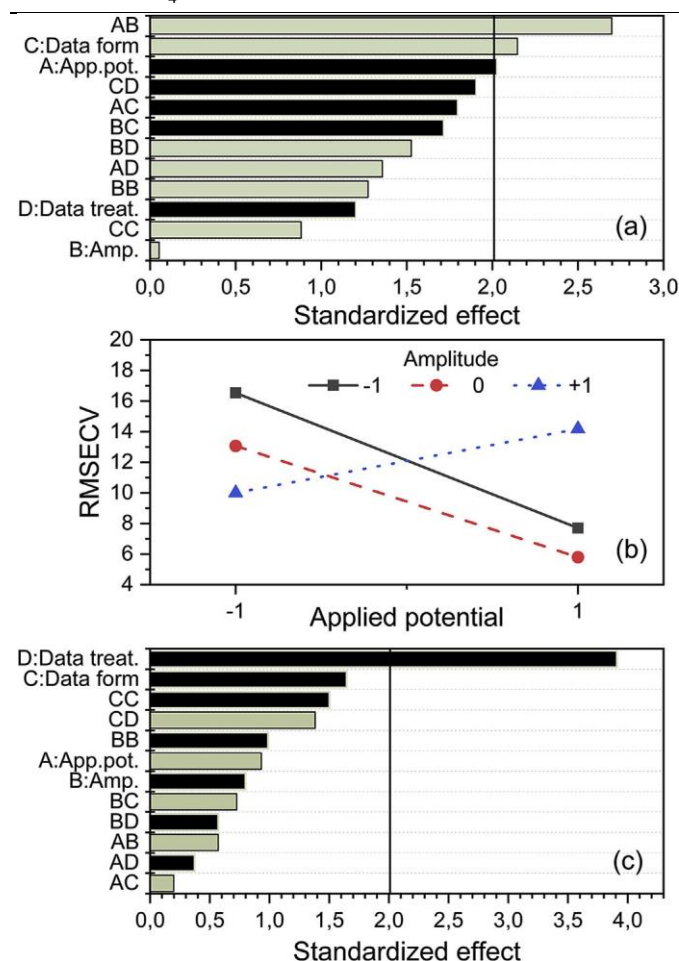


Fig. 6. (a) and (c): Pareto charts of the responses RMSECV and number of latent variables, respectively, for hydroquinone determination. Gray bar: positive effect; black bar: negative effect. Where: A: applied potential; B % amplitude; C % data presentation; D % data pretreatment. (b) Change of RMSECV 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 65e175 mmol L<sup>-1</sup> 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.

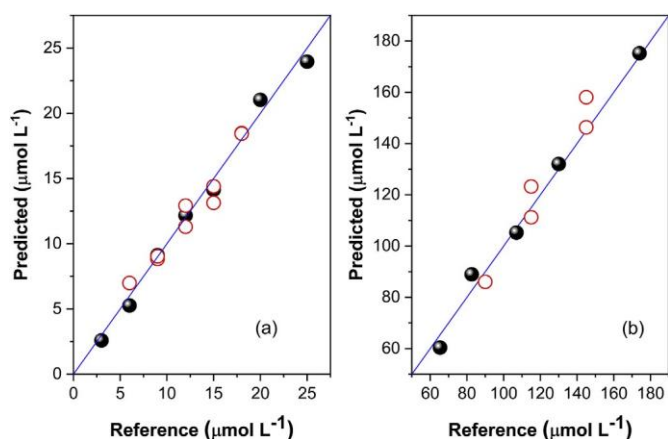


Fig. 7. Analytical curves and prediction results of the determination of hydroquinone (a) and catechol (b) in tap waters. Black: calibration; Red/white: prediction. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

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 carbontype electrodes. Since the predicted concentrations are also complex numbers, the deviation from the expected value of the imaginary part (0j) could also be used as an indication of the quality of complex numbers-PLS models, although additional investigations are required. Complex

numbersePLS 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 e Brazil (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico e Brazil (CNPq, grant numbers 456569/2014-6, 301662/ 2015-0, and 310380/2017-0), Agencia Nacional de Promoción Científica y Tecnológica - Argentina (ANPCyT, grant number PICT 2016-1122), and Consejo Nacional de Investigaciones Científicas 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. Pebere, B. Tribollet, J. Electrochem. Soc. 153 (2006) B129eB136.
- [2] E.L. Anderson, P. Bühlmann, Anal. Chem. 88 (2016) 9738e9745.
- [3] A. Miszczyk, K. Darowicki, Prog. Org. Coat. 77 (2014) 2000e2006.
- [4] B. Lindholm-Sethson, P. Geladi, A. Nelson, Anal. Chim. Acta 446 (2001) 121e131.
- [5] C.A. Olivati, A. Riul Jr., D.T. Balogh, O.N. Oliveira Jr., M. Ferreira, Bioproc. Biosyst. Eng. 32 (2009) 41e46.
- [6] G. Kim, J.-H. Moon, M. Morgan, Anal. Methods 5 (2013) 4074e4080.

- [7] C. Conesa, E. García-Breijo, E. Loeff, L. Seguí, P. Fito, N. Laguarda-Miro, *Sensors* 15 (2015) 22941e22955.
- [8] A.V. Sidelnikov, D.M. Bikmeev, D.I. Dubrovskii, F. Kh Kudasheva, V.N. Maistrenko, *J. Anal. Chem.* 70 (2015) 837e842.
- [9] S. Phal, B. Lindholm-Sethson, P. Geladi, A. Shchukarev, S. Tesfalidet, *Anal. Chim. Acta* 987 (2017) 15e24.
- [10] K. Bundy, M. Karlsson, G. Lindbergh, A. Lundqvist, *J. Power Sources* 72 (1998) 118e125.
- [11] Ulrich, H. Petersson, H. Sundgren, F. Bjorefors, C. Krantz-Rülcker, *Sens. Actuators B Chem.* 127 (2017) 613e618.
- [12] P. Geladi, A. Nelson, B. Lindholm-Sethson, *Anal. Chim. Acta* 595 (2007) 152e159.
- [13] A. Bonanni, D. Calvo, M. del Valle, *Electroanalysis* 20 (2008) 941e948.
- [14] H. Wold, *Research Paper in Statistics: Festschrift for J. Neyman*, Wiley, New York, 1966, pp. 411e444.
- [15] K. Pitman, J. Nerut, E. Lust, S. Franssila, M. Raud, T. Kikas, *ECS Trans.* 77 (2017) 1771e1782.
- [16] F. Sundfors, J. Bobacka, *J. Electroanal. Chem.* 572 (2004) 309e316.
- [17] S. Vogt, Q. Su, C. Gutierrez-Sanchez, G. Noël, *Anal. Chem.* 88 (2016) 4383e4390.
- [18] D.M. Haaland, E.V. Thomas, *Anal. Chem.* 60 (1988) 1193e1202.
- [19] K.M. Kosicek, K. Kvastek, V. Horvat-Radosevic, *Electrochim. Acta* 195 (2016) 77e84.
- [20] A.S. Adekunle, K.I. Ozoemena, *Electrochim. Acta* 53 (2008) 5774e5782.
- [21] M.T.T. Tran, B. Tribollet, V. Vivier, M.E. Orazem, *Russ. J. Electrochem.* 53 (2017) 932e940.
- [22] P. Heiduschka, J. Dittrich, *Electrochim. Acta* 37 (1992) 2573e2580.
- [23] H. Zanin, P.W. May, D.J. Fermin, D. Plana, S.M.C. Vieira, W.I. Milne, E.J. Corat, *ACS Appl. Mater. Interfaces* 6 (2014) 990e995.
- [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) 4373e4385.
- [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) 3576e3584.
- [26] S.H. DuVall, R.L. McCreery, *Anal. Chem.* 71 (1999) 4594e4602.