



This is the **submitted version** of the article:

Faura, Georgina; González-Calabuig, Andreu; Valle Zafra, Manuel del. «Analysis of amino acid mixtures by voltammetric electronic Tongues and artificial neural networks». Electroanalysis, Vol. 28, issue 8 (Aug. 2016), p. 1894-1900. DOI 10.1002/elan.201600055

This version is available at https://ddd.uab.cat/record/251194 under the terms of the $\bigcirc^{\mbox{\footnotesize{IN}}}$ license

Analysis of Amino Acid Mixtures by Voltammetric Electronic Tongues and Artificial Neural Networks

Georgina Faura, Andreu González-Calabuig, Manel del Valle *

Sensors & Biosensors Group, Department of Chemistry, Universitat Autònoma de Barcelona, Edifici Cn, 08193 Bellaterra, Spain * e-mail: manel.delvalle@uab.es

Received: Accepted:

Abstract

A new voltammetric electronic tongue with graphite-epoxy composite electrodes modified with metal-oxide nanoparticles is presented for the quantification of tryptophan, tyrosine and cysteine aminoacid mixtures. The signals were obtained by cyclic voltammetry, and data was processed using two different chemometrical techniques, artificial neural networks and partial least squares regression, for comparison purposes. Before performing artificial neural networks data was compressed by fast Fourier transform or discrete wavelet transform. The best results were attained using artificial neural networks with previous fast Fourier transform compression of the data with a normalized root-mean-square error of 0.032 (n=15) for the external test subset. The present method shows results comparable to other similar approaches, but with a much easier sampling process for the training set and new electrode modifiers to form the voltammetric sensors.

Keywords: Amino acid, Artificial neural network, Electronic tongue, Metal oxide, Voltammetric sensors

DOI: 10.1002/elan

1. Introduction

The analysis of amino acids (AAs) is usually performed using techniques such as gas [1,2] and liquid [3,4] chromatography or capillary electrophoresis [5-7]. Though, the great demand of AA determinations in many important fields, such as clinical analysis or food and pharmaceutical industry [8,9], leads to the need of simpler and faster procedures, such as sensor-based methods. These electroanalytical approaches have been already exposed as good methods for the analysis of specific AAs thanks to the electroactivity presented by aromatic and thiol groups [8].

Materials traditionally used in the construction of sensors for this kind of applications are carbon, platinum and gold. But these materials are usually modified in order to increase the sensitivity and introduce new features due to electrocatalitic activity, which leads to an increase in the range of detectable AAs [10].

The use of metal oxide-modified sensors (MOS) has increased during the last years due to its great potential in the analysis of many electroactive substances [11,12] or even the detection of the presence of some bacteria in complex systems [13]. Moreover, the recent availability of these materials in the form of nanopowder has facilitated enormously the possibility of using them as electrode modifiers. Potential of using MOS materials has widened its scope even more by the use of

semiconducting MOS arrays as gas sensors forming electronic noses (ENs). ENs have gained recognition in many fields like food, aroma or medical diagnosis [14]. The large amount of data provided by ENs increases the sensitivity and selectivity of the method just by multivariate data processing, which avoids tedious derivatization or optimisation of the sensors for each analyte [15,16]. However, the use of MOS is only a possible option to develop electronic tongues (ETs) [17].

ETs have been used for more than a decade, and its use has grown considerably as they present a large range of applications, including environmental analysis, food industry and agriculture [18-20]. Due to the good performance of voltammetric techniques to analyze certain AAs, voltammetric ETs have already been presented as a good approach for their analysis, in this case for the ones with oxidizable properties, such as cysteine (CYS), tyrosine (TYR) and tryptophan (TRP) [21-23].

An important weakness of the voltammetric ETs resides in the complex response obtained, as there is a need to process each voltammogram from each sensor. However, huge amounts of data derived from voltammetric analysis have usually been modelled

Table 1 Composition of the amino acid mixtures standards used to train and evaluate performance the ANN model

	Train set (10 ⁻⁴ M)																						
#	1	2	3	4	5	6	7	8	9	10) 1	1	12	13	14	1	L5	16	17	18	19	20	21
TRP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.	.5 ().5	0.5	0.5	0	.5	0.5	0.5	0.5	3.0	3.0	3.0
TYR	0.0	0.0	0.0	0.5	0.5	0.5	3.0	3.0	3.0	0.0	0.	.0 (0.0	0.5	0.5	0	.5	3.0	3.0	3.0	0.0	0.0	0.0
CYS	0.0	0.5	3.0	0.0	0.5	3.0	0.0	0.5	3.0	0.0	0.	.5 3	3.0	0.0	0.5	3	.0	0.0	0.5	3.0	0.0	0.5	3.0
Train set (10 ⁻⁴ M)					Test set (10 ⁻⁴ M)																		
#	22	23	24	25	26	27		1	2	3	4	5	6	7	' :	8	9	10	11	12	13	14	15
TRP	3.0	3.0	3.0	3.0	3.0	3.0		2.7	2.7	1.3	2.7	1.3	1.	3 1	.3 (0.0	0.7	0.7	1.7	0.7	0.7	1.3	0.3
TYR	0.5	0.5	0.5	3.0	3.0	3.0		2.7	2.3	1.3	1.7	1.3	2.	3 2	.7 2	2.3	1.3	1.3	2.3	0.0	0.0	2.0	2.3
CYS	0.0	0.5	3.0	0.0	0.5	3.0		1.7	1.3	0.7	1.7	2.0	1.	3 2	.3 2	2.7	2.7	0.3	2.7	0.7	0.0	1.3	1.7

successfully by chemometric techniques which interpret and extract meaningful data, offset any matrix effect and may allow the resolution of interfering agents and drifts. The most widely used multivariate data analysis techniques for voltammetric ETs are principal component analysis (PCA), partial least squares regression (PLS) and artificial neural networks (ANNs) [24,25].

PCA is commonly used for the revisualization and compression of the data by its transformation into new orthogonal coordinate axes (called principal components), defined in a way to describe almost all the variance contained in the data. The orthogonality presented by the new directions facilitates the use of linear regression models, such as principal components regression [24-27].

PLS, like PCA, represents the data in new coordinate axes, but in this case it maximises the covariance instead of the variance, and takes into account not just the input data matrix, but also the output matrix [25,28].

Both PCA and PLS are especially suited for linear systems, and although some modifications can be introduced in order to work with non-linear systems, they still do not present as good results as other chemometric techniques, such as ANNs, which excel in the modelling of non-linear systems.

ANNs [21,25,29] present an iterative modelling scheme which imitates the learning neuronal system of the brain, and permit modelling both linear and nonlinear data. As in the brain, ANNs comprise neurons, units arranged in layers. The input data is sent to each neuron, where it is multiplied by a weight plus a bias value and sent through a transfer function. The error between the output and target output values is reduced via iteration by changing the weight values, until the desired error is reached. The ANNs may be optimised by changing its specific configuration (number of neurons, number of layers, transfer functions, error...), which leads to a great amount of possible models for a single study case. Fortunately, there are many algorithms and optimisation methods described in the literature in order to simplify the process [24].

In this work we present a novel voltammetric ET consisting in graphite-epoxy composite electrodes modified with metal-oxide nanoparticles, aimed to the simultaneous determination of CYS, TYR and TRP in mixtures. Overlapping voltammetric responses have been modelled by both ANNs (which have been demonstrated to be one of the most powerful data processing strategies) and PLS, for comparing purposes. Given the multidimensionality and complexity of the generated data (3D data matrix *current x sensors x potential*), before ANN modelling, the data has been compressed by either discrete Wavelet transform (DWT) or fast Fourier transform (FFT).

2. Experimental

2.1. Reagents, apparatus and software

All employed solutions were prepared using ultrapure water from MilliQ System (Millipore, Billerica, MA, USA). All chemicals for electrolytes and AA solutions were supplied by Merck (Darmstadt, Germany). ZnO <50 nm nanoparticles (NPs), Bi₂O₃ 90-210 nm NPs, Sn₂O₃ <100 nm NPs and WO₃ <100 nm NPs were purchased at Sigma-Aldrich (St. Louis, MO, USA).

Cyclic voltammograms were obtained with a PGSTAT 30 Autolab potentiostat (EcoChemie, The Netherlands), working in multichannel configuration configuration using the GPES Multichannel 4.7 software package (EcoChemie).. A combination electrode formed by metallic Pt plus a Ag/AgCl reference electrode (Crison 5261, Barcelona, Spain) was used in the voltammetric measurements. A Crison micro pH 2002 pH-meter (Crison instruments, Barcelona, Spain) was used for the pH adjustments.

Data processing was performed using MATLAB R2013b for windows (MathWorks, Natick, MA). Routines were programmed by the authors, using the Neural Networks toolbox (MathWorks).

2.2. Electrolyte and data sets.

The support electrolyte solution consisted of a 0.1 M potassium monohydrogenphosphate, 0.1 M potassium dihydrogenphosphate and 0.1 M KCl solution adjusted to pH 7.5 with a 37% (w/w) hydrochloric acid solution. The train set consisted of 27 synthetic samples defined as a full factorial design for the three AAs and three different concentration levels (0 M, 5·10⁻⁵ M and 3·10⁻⁴ M). To estimate performance of the developed model, a test set was used, formed by 15 synthetic samples with the three AAs of interest in random concentrations (with the values rounded in order to get easy attainable micropipette addition volumes) included in the train range (from 0 M to 3·10⁻⁴ M). All the samples used for the model building and validation are summarised in Table 1.

2.3. Electronic tongue

Measurements with the voltammetric ET were taken using a multichannel voltammetric cell formed by the reference electrode, auxiliary electrode and five working electrodes, in this case, modified epoxy-graphite composite sensors. One of these was a bare epoxy-graphite sensor constructed following the conventional methodology established in our research group [30]. The other four were bulk-modified epoxy-graphite sensors with a 4% (w/w) of metal oxide nanopowder added in the mixture before curing (metal oxides described in Section 2.1.).

2.4. Voltammetric procedure

The cyclic voltammetries were carried out at room temperature and the measurement cell was composed by the voltammetric ET described at 2.3. The measurements were performed in a range of potential from -1 V to +1.2 V at a scan rate of $0.1~\rm V\cdot s^{-1}$. Before the whole experiment, sensors were properly regenerated by polishing and cleaning, and 3-5 stabilisation cycles in buffer solution (electrolyte described at 2.2.) were done in order to stabilise the voltammetric response. Between measures, a cleaning stage consisting in three full cycles with Millipore water were performed. Samples were measured randomly and only the last scan of each sample was used.

2.5. Data processing details

The voltammetries from the train set were used in order to build the ANN model. Several parameters defining the ANN configuration were taken into account (Figure 1), in a way that a number of 5760 different models were built and compared.

Given the high dimensionality conditions, data was first compressed by using DWT or FFT transforms, and subsequently unfolded in a single vector to form the input to the ANN model. The training algorithm (Bayesian regularization) and the DWT (Daubechies) and FFT parameters (decomposition levels and number of coefficients) are stated on Figure 1; these were chosen based on previous experience with similar conditions [31]. For the ANN training, convergence errors were chosen at least 5 orders of magnitude lower than the higher concentration of AA tested, and the maximum of training epochs was taken as 200.

For comparison purposes, also a PLS2 model was evaluated. Unlikely other linear chemometric techniques, such as PLS or PCR, PLS2 permits the modelling of several variables together [32]. Thus, the linear model obtained by PLS2 would equate the multiple output ANN model. The optimisation of the PLS2 model consisted in choosing the number of latent variables based on a compromise between this number and the percentage of predictive significance, in order to avoid overfitting.

To select the best model (for both ANN and PLS2), the predicted concentrations for the three AAs were compared with the expected ones for both, the train and the test set. The best regression parameters (slope (m) and correlation value (R) close to 1 and intercept (b) close to 0) and the lowest root-mean-square error (RMSE) and normalized root-mean-square error (NRMSE) were considered as indicators of the optimum.

You may get deeper into the chemometric routines of our research group by checking some of our related works [21,25,30,33,34].

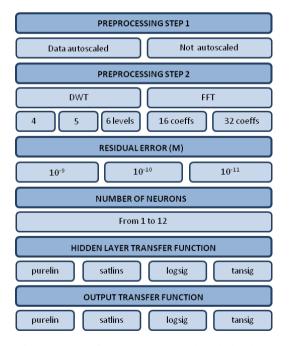


Figure 1. Parameters for the ANN model optimisation. DWT, Discrete Wavelet transform; FFT, Fast Fourier transform.

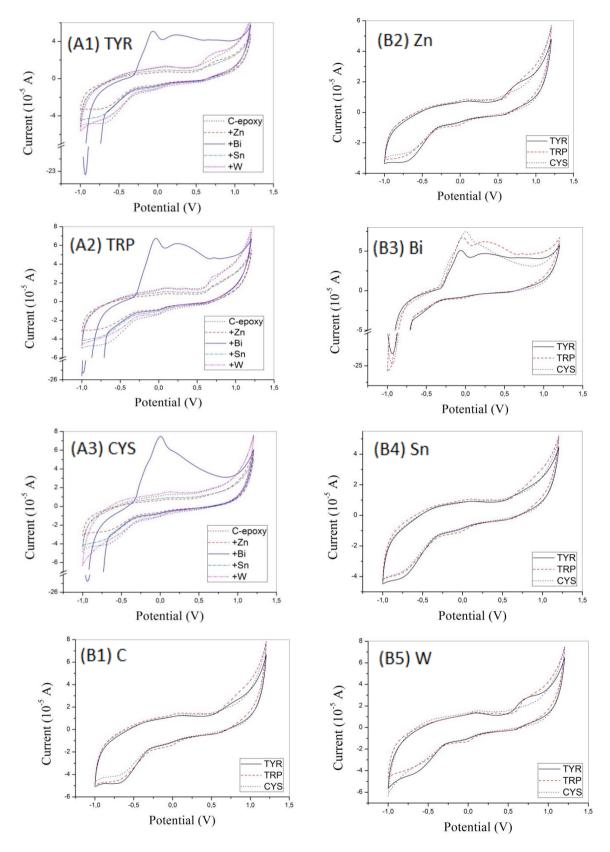


Figure 2. Samples of the cyclic voltammetries obtained. (A1-A3) Voltammetric responses of the three AAs $(3\cdot10^{-4} \text{ M})$ for each sensor (B1-B5) Voltammetric responses of the MOS-modified sensor array for each AA $(3\cdot10^{-4} \text{ M})$.

3. Results and discussion

3.1. Voltammetric responses

When we construct an ET it is very important to use an array of sensors with cross-response, namely poorly selective sensors with partial specificity to the different target compounds in a solution (just as happens in the taste buds of some animals). The low selectivity and specificity are easily solved by the posterior chemometric treatment of the data [35,36].

As it can be seen in figure 2A, the signals offered by the sensors for a specific AA present different shapes. Hence, there is no redundancy between the different information generated by the sensors.

Moreover, as shown in figure 2B, every sensor provides a slightly different signal for each AA. Thus, this suggests the sensor array chosen presents the desired cross-response for the desired ET application.

The higher signals provided by the Bi MOS can be related to the capacity of Bi in forming complexes with some AAs [37,38]. Therefore, a proper established cleaning step between measures results mandatory.

3.2. Chemometrics data processing

3.2.1. ANN

The best ANN model was designated following the criteria described in section 2.5., and was achieved without any autoscaling of the data, by FFT compression and using 160 input neurons, 1 single hidden layer formed by 7 neurons, the purelin hidden layer transfer function and satlins output transfer function.

Figure 3 shows the comparison of the obtained values vs. expected results (the known concentrations) with both the training model (blue dashed line and squares) and the external test subset (red dotted line and circles) for the three AAs. Also the ideality (r=1, slope=1, interception=0) is shown (black solid line). A clear evidence from the comparison figures is that the PLS variant displays a higher dispersion of the data than the ANN case.

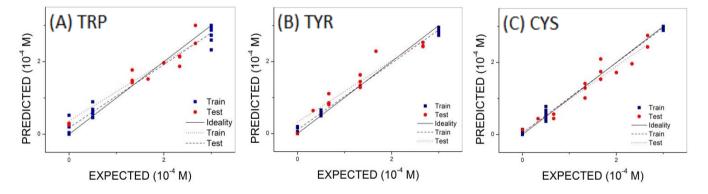


Figure 3 Comparison of the expected vs. the predicted concentrations for the optimized FFT-ANN model.

Table 2 Linear regression parameters obtained comparing the expected and predicted concentration values by the ANN model.

	AA	Correlation	Slope	Intercept (10 ⁻⁴ M)	RMSE (10 ⁻⁵ M, n=27))	NRMSE (n=15)
TRAIN	TRP	0.999	0.924±0.004	0.105±0.007		
	TYR	0.989	0.872 ± 0.010	0.176 ± 0.017	1.654	0.018
	CYS	0.998	0.973 ± 0.012	0.046 ± 0.008		
TEST	TRP	0.960	0.853±0.035	0.315±0.054		
	TYR	0.952	0.807 ± 0.036	0.363 ± 0.070	2.570	0.032
	CYS	0.967	0.920 ± 0.068	0.053 ± 0.118		

The optimisation of the PLS model was performed following the protocol described in section 2.5., and the best performance was achieved with 10 latent variables. In this case, no compression was performed at all.

On figure 4 are shown the plots which compare the expected results vs. the obtained concentration values with both the training model (blue dashed line and squares) and the test subset (red dotted line and circles) for the three AAs. Also the ideality (r=1, slope=1, intercept=0) is shown (black solid line). In table 3 the correlation, slope and interception for every set and AA (confidence level 95%), and RMSE and NRMSE values for the training model and the test subset are presented.

If we compare the values from table 2 and 3, we can notice that the slope and interception values for the train set from the PLS model are better (closer to one and zero respectively) than the ones got with the ANN model, but with the test set occurs the opposite. We also can appreciate that the correlation values got with the ANN model are closer to one than the ones achieved with the PLS model, as well as lower RMSE and NRMSE values. Taking into account the obtained results we could affirm that the ANN model presents a better prediction of the concentration of TRP, TYR and CYS in AA mixtures than the PLS model.

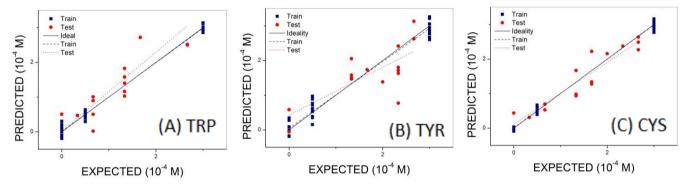


Figure 4 Comparison of the expected vs. the predicted concentrations for the optimized PLS2 model.

Table 3 Linear regression parameters obtained comparing the expected and predicted concentration values obtained with use of the PLS2 model.

	AA	Correlation	Slope	Intercept (10 ⁻⁴ M)	RMSE (10 ⁻⁵ M)	NRMSE
TRAIN	TRP	0.995	0.991±0.020	0.011±0.013		
	TYR	0.979	0.961±0.015	0.046 ± 0.026	1.768	0.020
	CYS	0.995	0.995 ± 0.005	0.006 ± 0.025		
TEST	TRP	0.867	1.136±0.087	0.020±0.133		
	TYR	0.718	0.690 ± 0.088	0.426 ± 0.168	0.426 ± 0.168 5.040	
	CYS	0.926	0.895 ± 0.097	0.115±0.086		

4. Conclusions

In this work we have presented a novel voltammetric ET consisting in a simple five metal oxide sensor nanopowder modified array capable of predicting the concentration of TRP, TYR and CYS mixtures in solution. The complex voltammetric data has been successfully processed by the coupling of FFT and ANN.

Thanks to the different redox catalytic properties of the modifiers, slightly different oxidation and reduction peaks were obtained for each sensor and AA (cross response), making possible the ET application

The processing of the raw data by FFT and ANN allowed the removal of noise, matrix effect and redundant information. Comparison with PLS2 model showed that the ANN model presents better results with the samples studied in this work. Although the involved computing times are extended, the protocols for the optimisation of ANN models become increasingly efficient daily, and they are needed only during the optimisation and calibration step. Obtained results, although perhaps not matching perfectly the

expected ones are counterbalanced by the rather simple procedure used for training. With fixed factorial experimental designs, richness or variability is precluded, and with its limited extension only a quick screening methodology is finally derived. It is evident that simply by enlarging the training data set, or by using some other experimental design with higher variability obtained results would surely be improved.

Future works will try to quantify the three oxidizable AAs in real samples using the methodology described. The method presented would be especially useful in the analysis of samples in which the speed in obtaining the results is more critical than the accuracy of the analysis, avoiding this way the use of more expensive, tedious and time consuming techniques, such as chromatography.

5. Acknowledgements

Authors acknowledge financial support from the Spanish Ministry of Science and Innovation (MINECO, through projects CTQ2013-41577-P. M. del Valle acknowledges support by the Catalonia program ICREA Academia.

6. References

- [1] K. Dettmer, A. P. Stevens, S. R. Fagerer, H. Kaspar, P. J. Oefner, Amino Acid Analysis, Methods in Molecular Biology. Vol. 828. Humana Press, 2012, pp. 165-181.
- [2] E. Jiménez-Martín, J. Ruiz, T. Pérez-Palacios, A. Silva, T. Antequera, J. Agric. Food Chem. 2012, 60 (10), pp. 2456-2463
- [3] C. Singh, C. S. Sharma, P. R. Kamble, *Int. J. Pharmacogn.* **2014**, pp. 756-762
- [4] P. Krumpochova, B. Bruyneel, D. Molenaar, A. Koukou, M. Wuhrer, W.M.A. Niessen, M. Giera, J. Pharmaceut. Biomed. 2015, 114, pp. 398-407
- [5] V. Poinsot, M. Lacroix, D. Maury, G. Chataigne, B. Feurer, F. Couderc, *Electrophoresis* 2006, 27 (1), pp. 176-194
- [6] A. Hirayama, T. Soga, *Amino Acid Analysis, Methods and protocols. Vol.* 828. Humana Press, **2012**, pp. 77-82.
- [7] T. Soga, D. N. Heiger, Anal. Chem. 2000, 72 (6). pp. 1236-1241
- [8] L. Moreno, A. Merkoçi, S. Alegret, S. Hernández-Cassou, J. Saurina, Anal. Chim. Acta. 2004, 507, pp. 247-253
- [9] J. Saurina, S. Hernández Cassou, E. Fàbregas, S. Alegret, Anal. Chim. Acta 2000, 405, pp.153-160
- [10] M. Nooredeen Abbas, A. Ali Saeed, B. Singh, A. A.Radowan, E. Dempsey, *Anal. Methods* 2015, 7, pp. 2529-2536
- [11] N. Barsan, D. Koziej, U. Weimar, Sensor Actuat. B-Chem. 2007, 121, pp. 18-35
- [12] L. A. Ramos, E. T. G. Cavalheiro, C. C. S. Cavalheiro, Braz. J. Food Technol. 2005, 8 (1), pp. 1-5

- [13] E. Gobbi, M. Falasconi, G. Zambotti, V. Sberveglieri, A. Pulvirenti, G. Sberveglieri, *Sensor Actuat. B-Chem.* **2015**, 207, pp. 1104-1113
- [14] A. Berna, Sensors, **2010**, 10 (4), pp. 3882-3910
- [15] R. Dutta, E. L. Hines, J. W. Gardner, K. R. Kashwan, M. Bhuyan, *Sensors Actuat. B-Chem.* **2003**, 94 (2), pp. 228-237
- [16] R. Dutta, K. R. Kashwan, M. Bhuyan, E. L. Hines, J. W. Gardner, *Neural Networks* **2003**, 16 (5-6), pp. 847-853
- [17] N. Liu, Y. Liang, J. Bin, Z. Zhang, J. Huang, R. Shu, K. Yang, Food Anal. Methods 2014, 7, pp. 472-480
- [18] M. del Valle, Electroanalysis 2010, 22 (14), pp. 1539-1555
- [19] M. del Valle, Int. J. Electrochem. 2012, 2012, pp. 1-11
- [20] F. Winquist, Microchim. Acta 2008, 163, pp. 3-10
- [21] L. Moreno-Barón, R. Cartas, A. Merkoçi, S. Alegret, M. del Valle, L. Leija, P. R. Hernandez, R. Muñoz, Sensors Actuat. B-Chem. 2006, 113, pp. 487-499
- [22] Y. Miyanaga, A. Tanigake, T. Nakamura, Y. Kobayashi, H. Ikezaki, A. Taniguchi, K. Matsuyama, T. Uchida, *Int. J. Pharm.* **2002**, 248, pp. 207-218
- [23] J. F. Folmer-Andersen, M. Kitamura, E. V. Anslyn, *J. Am. Chem. Soc.* **2006**, 128(17), pp. 5652-5653
- [24] E. Richards, C. Bessant, S. Saini, *Electroanalysis* **2002**, 14 (22), pp. 1533-1542
- [25] X. Cetó, F. Céspedes, M. del Valle, *Microchim. Acta* 2013, 180, pp. 319-330
- [26] S. Wold, K. Esbensen, P. Geladi, *Chemometr. Intell. Lab.* 1987, 2 (1-3), pp.37-52
- [27] H. Abdi, L. J. Williams, Computation Stat. 2010, 2 (4), pp. 433-459
- [28] P. Geladi, B. R. Kowalski, Anal. Chim. Acta 1986, 185, pp.1-17
- [29] G. Hanrahan, Anal. Chem. 2010, 82, pp. 4307-4313
- [30] X. Cetó, J. M. Gutiérrez, L. Moreno-Barón, S. Alegret, M. del Valle, *Electroanalysis* 2011, 23 (1), pp. 72-78
- [31] L. Moreno-Barón, R. Cartas, A. Merkoçi, S. Alegret, J. M. Gutiérrez, L. Leija, P. R. Hernandez, R. Muñoz, M. del Valle, *Anal. Lett.* 2005, 38 (13), pp. 2189-2206
- [32] M. Blanco, A. Peguero, Talanta 2008, 77, pp. 647-651
- [33] X.Cetó, J.M. Gutiérrez, M. Gutiérrez, C.Jiménez, F.Céspedes, J.Capdevila, S.Minguez, M. del Valle, Anal. Chim. Acta. 2012, 732, pp. 172-179
- [34] A. C. De Sá, A. Cipri, A. González-Calabuig, N. R. Stradiotto, M. del Valle, *Sensors Actuat. B-Chem.* **2016**, 222, pp. 645-653
- [35] D. Ha, Q. Sun, K. Su, H. Wan, H. Li, N. Xu, F. Sun, L. Zhuang, N. Hu, P. Wang, Sensors Actuat. B-Chem. 2015, 207, pp. 1136-1146
- [36] N. Prieto, P. Oliveri, R. Leardi, M. Gay, C. Apetrei, M. L. Rodriguez-Méndez, J. A. De Saja, *Sensors Actuat. B-Chem.* **2013**, 183, pp. 52-57
- [37] S. K. Ghatak, D. Dey, S. Sen, K. Sen, *Analyst* **2013**. 138, pp. 2308-2314
- [38] M. Busse, E. Border, P. C. Junk, R. L. Ferrero, P. C. Andrews, *Dalton Trans.* 2014, 43, pp. 17980-17990