Original Paper

Use of Artificial Neural Networks, Aided by Methods to Reduce Dimensions, to Resolve Overlapped Electrochemical Signals. A Comparative Study Including other Statistical Methods

J. M. Palacios-Santander[1], **A. Jiménez-Jiménez**[2], **L. M. Cubillana-Aguilera**[1], **I. Naranjo-Rodríguez**[1], and **J. L. Hidalgo-Hidalgo-de-Cisneros**[1,*]

[1] Departamento de Química Analítica, Universidad de Cádiz, Polígono Río San Pedro, Apartado 40, Puerto Real 11510, Cádiz, Spain
[2] Servicio de Informática Científica y Estadística, Universidad de Cádiz, Polígono Río San Pedro, Apartado 40, Puerto Real 11510, Cádiz, Spain

Abstract. A method using artificial neural networks (ANNs) combined with Fourier Transform (FT) and Wavelet Transform (WT) was used to resolve overlapping electrochemical signals. This method was studied as a powerful alternative to traditional techniques such as principal component regression (PCR) and partial least square (PLS), typically applied to this kind of problems.

WT and FT were applied to experimental electrochemical signals. These are two alternative methods to reduce dimensions in order to obtain a minimal recomposition error of the original signals with the least number of coefficients, which are utilized as input vectors on neural networks. Tl^+ and Pb^{2+} mixtures were used as a proof system.

In this paper, neural networks with a simple topology and a high predictive capability were obtained, and a comparative study using PLS and PCR was done, producing the neural models with the lowest RMS errors. By comparing the error distributions associated with all the different models, it was established that models based on FT and WT (with 11 coefficients) neural networks were more efficient in resolving this type of overlapping than the other chemometric methods.

The problem of determining two or more species with similar analytical signals has been a matter of substantial interest since the former developments in instrumental techniques of analysis. Nowadays, instrumental techniques combined with the suitable chemical procedures are able to resolve this problem in most situations, but inefficient in others. For example, electroanalytical techniques have found numerous applications [1, 2] due to their simplicity and low cost, but selectivity problems occur frequently. Overlapped peaks occur more commonly in voltammetry than in chromatography or most spectral methods, because the width of a voltammetric peak is an appreciable fraction of the accessible potential range. It is in these situations that statistical techniques and methods based on signal processing play an important role, allowing the separation of signals including the most serious overlapping cases.

* Author for correspondence. E-mail: jluis.hidalgo@uca.es

Between the statistical techniques used most frequently for deconvolution and simultaneous evaluation of overlapped signals, independently of the type of signal, are the signal ratio resolution method [3], multivariate calibration methods (PLS, PCR, etc) [4, 5], derivation of signals [6], curve fitting [7], Fourier transform (FT) [8], wavelet transform (WT) [9] and ANNs [10–12].

FT is one of the most well-known and most frequently used tools in any scientific discipline. Using this kind of transform, a signal $f(t)$ is represented as a combination of basis functions, normally sines and cosines. This means that the original signal is decomposed in a sum of sinusoidal signals at different frequencies. The FT has two fundamental advantages: it is simple, and it can be explained in physical terms (undulatory nature of many signals).

WT now plays a special role in the field of signal processing: data pre-processing [9], de-noising and compression [13, 14], and overlapped signals processing [15] among others. While the basis functions of FT are unlimited in duration and smooth and periodic, the component functions of WT are finite, asymmetric and non-periodic. This provides WT with the advantageous characteristic of compressing information, as it allows to approximate signals with features that change over time and signals that have jumps and other non-smooth features. Another important characteristic of WT is that this type of transform can represent a non-stationary signal, where, in contrast to FT, the frequency depends on the variable t,.

In wavelet analysis, linear combinations of wavelet functions are used to represent signals $f(t)$. This representation allows you to separate a signal into multi-resolution components. The fine and coarse resolution components capture the high-frequency and the low-frequency parts of a signal, respectively. In the same way, these representations are useful in a broad range of applications, such as data compression, signal and image processing, non-parametric statistical estimation, numerical analysis, chemistry, astronomy and oceanography.

Nowadays ANNs, understood as black box models, are accepted by the scientific society and extensively used to perform multicomponent analysis [16] and pattern recognition [17]. ANNs have already been utilized in combination with most analytical techniques, and the number of applications in electroanalysis is also increasing [18, 19]. There are many general books about the application of this method in chemistry [20, 21].

Although the theory and principles of ANNs have been dealt with in detail in the literature [22, 23], a small summary about the issue will be given here.

Similar to the biological structures of living organisms, ANNs consist of a set of processing units called neurones (cells or nodes) which are capable of sharing information. In particular, there are neural models that can be used as supervised models with a predictive feature. These models have a disposition in several layers: an input layer, one or more hidden layers, and an output layer, all of them connected to each other between adjacent layers which determine the structure or the topology of the network. The number of nodes in the input and output layer are defined by the problem being solved. The input layer receives the experimental information (such as experimental parameters conveniently or not pre-processed which constitute the training set) while the output layer delivers the response function. Regarding the hidden layers, they encode and organize the information received from the input layer and deliver it to the output layer. A bias is used to calculate the net input of a neuron from all the neurons connected to it. The neuron calculates a weighed sum for each signal. The objective is basically the adequate estimation of a set of parameters called weights by an iterative process named network training. These weights establish the importance of the connections between the neurones and are able to generate a neural network model with a minimum error rate. The model obtained can be validated using another set of samples called test set.

The statistical validation of the training and the test sets guarantees a satisfactory neural model when the error function (RMS error: root mean square error) for the training and the test samples is sufficiently small. Moreover, this type of validation is of general character and allows you to compare with other supervised models.

The process of calculating a predictive neural model is based on a type-gradient algorithm of convergence which tries to obtain the configuration of weights that gives the minimum RMS error.

In this paper, a procedure for resolving hard overlapped electrochemical signals is proposed. This procedure uses Fourier and wavelet transforms as methods to reduce dimensions in connection with

artificial neural networks, instead of multivariate calibration techniques such as PLS and PCR. The well-known Tl^+ and Pb^{2+} mixture, which has already been resolved (but with higher concentrations of the ions) by the traditional methods PLS and PCR, is employed as the proof system [24].

Although the new procedure is more complicated from the mathematical viewpoint, its use should be justified whether its ability of prediction was better than that of PLS and PCR or it opened a new way of treating this kind of signals statistically. Furthermore, it must be taken into account that PLS and PCR can only be used in those cases where linearity exists between the initial data and the values of the response function; by contrast, the proposed procedure is applicable to highly non-linear signal/concentration relationships as well as to linear ones, since these transforms lend a non-linear character to the reduced data.

Moreover, the justification to use artificial neural networks as a method to resolve hard overlapped signals versus PLS and PCR is based on the fact that previous pre-processing of the initial information (voltammograms) was done by applying Fourier and wavelet transforms, both suitable to detect small changes in the resultant signal frequency associated with a hard overlapping. By contrast, there is no guarantee that the PLS approach theoretically extracts the most reliable information [25], and the use of this method depends in part on what is known about the data, the nature of noise and signals and so on. This can also be applied to PCR, since PLS is more powerful than PCR.

Both transforms applied as methods to reduce dimensions will allow you to show all information contained in the signal as a vector with a few coefficients. In this paper, the two types of transforms will be compared in order to find out which has the best ability to reduce and retain information. The pre-processing of the signals by FT and WT allows you to find simple and stable neural models with three layers to resolve the problem of overlapped signals. The problem of overfitting should be avoided using the least number of coefficients possible to determine the network parameters. Furthermore, a comparative study of the results of these methods with those of PCR and PLS will be carried out. The plots of the RMS errors for each model will give an idea about the best model from all those assayed. And finally, by using box and whiskers plots, a comparison between

the error distributions will be carried out for all models applied.

Experimental

The DPASV (Differential Pulse Anodic Stripping Voltammetry) measurements were carried out with an Autolab®/PGSTAT20 system coupled to a Metrohm VA 663 Stand. An electrochemical three electrode cell, with a platinum auxiliary electrode, a silver/silver chloride, 3 M potassium chloride reference electrode and an HMDE (Hanging Mercury Drop Electrode, Metrohm) was employed.

Analytical reagent grade chemicals were used throughout the experiments. Voltammograms were recorded at room temperature. All solutions were de-aerated with nitrogen, when necessary, for at least 10 min prior to carrying out the experiments.

A 2 M acetic acid/2 M ammonium acetate buffer solution was utilized as supporting electrolyte (pH = 4.8–5.0). Lead and thallium solutions were prepared from nitrate salt stock solutions of $250 \, mg \, l^{-1}$.

The voltammetric parameters were as follows: deposition potential = $-1.3 \, V$; deposition time = 120 s; rest period = 20 s; initial potential = $-1.3 \, V$; end potential = $0 \, V$; scan rate = $8.5 \, mV \, s^{-1}$; pulse amplitude = $0.10 \, V$; pulse time = $0.07 \, s$; pulse repetition time = $0.6 \, s$. The drop surface was approximately $0.52 \, mm^2$.

For signal processing and statistical treatment the following software packages were used: MATLAB® 5.1, Statistica® 5.1, Unscrambler® 7.01 and EXCEL® 97 Pro. Qnet® 2000 neural network software was utilized to obtain the neural models.

Results and Discussion

In this paper, a supervised neural model is used as a direct pattern recognition method. In some studies, all points the instrument gives for each signal (in the present case, 80 points) are utilized as input vectors [10, 11]. However, this means working with a large number of dimensions, much redundant information, a relatively high computation time and possibly over-parametrization of the model. To avoid these problems, we applied the methods to reduce dimensions (to compress information) as a previous step to the estimation of the neural models.

We used a feedforward-type (connections must connect to the next layer) and multilayered neural network with an improved faster back propagation (BP) algorithm. There are two adaptive parameters (learning rate and momentum) for each weight in a BP neural network. The improved BP algorithm will make the learning process faster and avoid a local minimum in the surface of the RMS error. These types of networks operate in a supervised mode.

Backpropagation training is accomplished using the following logic sequence: 1) data is supplied to the input layer as a normalized vector and combined in the next (hidden) layer(s). 2) Each node of a given

hidden layer(s) processes the vector received multiplying it by a weight vector and adding bias value. 3) The resulting value is then processed through a transfer function. This transfer function serves to define the response of each neuron as a value that ranges from 0 (absence of response) to 1 (maximum intensity of the response). The transfer functions used here are the sigmoid function, $f(x) = 1/(1 + \exp(-x))$, the Gaussian function, $f(x) = \exp(-x^2)$ and the hyperbolic tangent, $f(x) = (\tanh(x) + 1)/2$. 4) Each node's output value is combined in the current hidden or output layer to form the layer's output vector. This output vector becomes the input vector for the next layer. 5) Processing proceeds to the next layer until arriving at the output layer where the final output vector is obtained. 6) The final output vector is combined with the training target vector to obtain the output layer's error vector. This is obtained for each hidden layer implying that the error is backpropagated through the network (thus the name for the paradigm). 7) Next, the weight vectors for each node must be updated using learning rate and momentum factor. The momentum term helps to keep the training process stable by damping weight change oscillations. 8) All input vectors (patterns) are processed through the network to adjust the weights for a given iteration. 9) The RMS error between the network response and the training targets is computed after each iteration. The RMS error is also equivalent to the standard deviation of the error in the network's response. 10) If a dispositive to modify and control the learning rate parameter is active for the run, a new learning rate (eta) is computed based on the change in the RMS error value. 11) The entire process cycle starts again with next training iteration.

One of the objectives of this paper was to obtain topologies as simple as possible to resolve overlapped signals of two species by a pattern recognition model. The key to this is using methods (FT and WT) to reduce dimensions of the original data that allow to simplify the input information without losing the non-linear character associated with the overlapping, so that the posterior neural model is as simple as possible.

To apply the proposed procedure, forty different samples were determined experimentally. These samples were distributed as follows: ten samples of pure thallium, ten samples of pure lead and twenty mixtures of both ions. The range of concentration was from 0.1 to 1.0 mg L^{-1}. The disposition and concentration of the samples were as represented in Table 1.

Table 1. Concentrations of samples used

| Sample	$[Tl^+]^*$	$[Pb^{2+}]^*$	Sample	$[Tl^+]^*$	$[Pb^{2+}]^*$
T1	0.1	0.0	T1L1	0.1	0.1
T2	0.2	0.0	T1L6	0.1	0.6
T3	0.3	0.0	T2L2	0.2	0.2
T4	0.4	0.0	T2L7	0.2	0.7
T5	0.5	0.0	T3L3	0.3	0.3
T6	0.6	0.0	T3L8	0.3	0.8
T7	0.7	0.0	T4L4	0.4	0.4
T8	0.8	0.0	T4L9	0.4	0.9
T9	0.9	0.0	T5L5	0.5	0.5
T10	1.0	0.0	T5L10	0.5	1.0
L1	0.0	0.1	T6L1	0.6	0.1
L2	0.0	0.2	T6L6	0.6	0.6
L3	0.0	0.3	T7L2	0.7	0.2
L4	0.0	0.4	T7L7	0.7	0.7
L5	0.0	0.5	T8L3	0.8	0.3
L6	0.0	0.6	T8L8	0.8	0.8
L7	0.0	0.7	T9L4	0.9	0.4
L8	0.0	0.8	T9L9	0.9	0.9
L9	0.0	0.9	T10L5	1.0	0.5
L10	0.0	1.0	T10L10	1.0	1.0

L = Pb^{2+}; T = Tl$^+$; * in mg L^{-1}.

We intended to choose two mixtures for both thallium and lead concentration values. Apart from that, another eight samples were measured later than the others and used as an external test set (tst) to check the predictive ability of the models.

The analysed mixtures showed very hard overlapping between the signals of both ions as shown in Fig. 1.

The samples produced a discrete data set composed of the full voltammograms of the samples, each one consisting of 80 points, the potential ranging from -0.3 to -0.7 V.
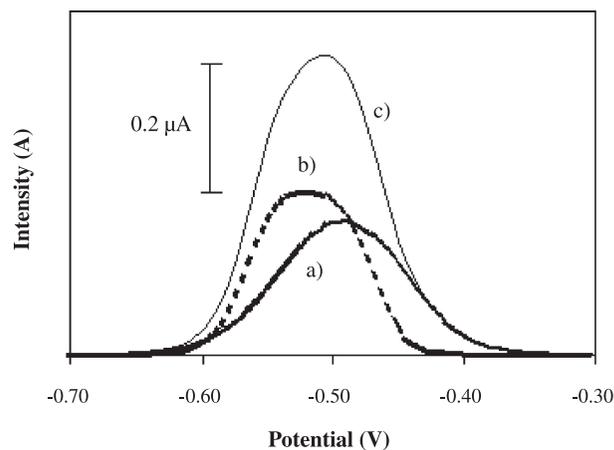
Fig. 1. Superposition of voltammograms: (a) 0.7 mg L^{-1} of Tl$^+$ (T7); (b) 0.7 mg L^{-1} of Pb^{2+} (L7); (c) mixture of 0.7 mg L^{-1} of Tl$^+$ and 0.7 mg L^{-1} of Pb^{2+} (T7L7)
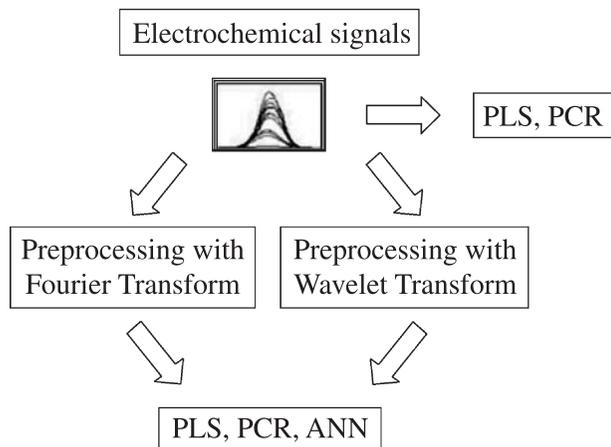
Fig. 2. Procedure used in the preprocessing and the statistical treatment of the electrochemical signals carried out with different chemometric tools

A comparison of the neural models and other statistical methods was performed. A scheme of the procedure applied using these chemometric tools is shown in Fig. 2.

Reduction of Dimensions Using Transforms

Fourier Transform

A subroutine of MATLAB was designed to obtain the FT of all voltammograms and reduce the number of initial data (dimensions). In this way, the signals are transformed from the time domain to the frequency domain. Afterwards, a cut frequency is chosen and a low-pass filter is applied. This filter removes high frequencies (usually noise) and maintains exclusively low frequencies (high amplitudes) which contain the information related to the signals. Immediately after this, the filtered signals are reconstructed at the time domain using the inverse Fourier transform in order to estimate the recomposition errors by the next equation:

$$\varepsilon = \sqrt{\sum_{i=1}^{80} \frac{(e_i - e_i^*)^2}{e_i^2}} \cdot 100 \qquad (1)$$

where e_i and e_i^* represent the points of the original and reconstructed signals, respectively.

The final objective is to obtain a suitable cut frequency that gives a number of Fourier coefficients as small as possible and, at the same time, the smallest recomposition error of the signals. Varying the parameter of the cut frequency, a chart was obtained with the different recomposition errors for each signal, as

well as their respective Fourier coefficients. The "best" value for the cut frequency was selected establishing a strategy of commitment between the recomposition error and the number of Fourier coefficients. So, this cut frequency was $\omega = 4\,\mathrm{Hz}$ for a voltammogram of 128 frequencies. The minimum number of dimensions (Fourier coefficients) corresponding to this frequency was $N = 7$, i.e. the seven fundamental amplitudes equivalent to the first three harmonics of the Fourier series. In this manner, every signal with 80 points was reduced to only 7 components of frequency, resulting in a recomposition error lower than 3% in all cases (calculated by expression (1)). To summarize, a dimension reduction of 91.25% was obtained for the original signals, maintaining at least 97% of the information.

Wavelet Transform

The objective is the same as with FT: obtaining a number of wavelet coefficients as small as possible with the smallest recomposition error of the signals.

Different kinds of wavelet basis were tested: Haar, Daublet 'n' ($n = 3, 4, 6, 8, 10, 12, 14, 16, 18$ and 20), Symmlet 'm' ($m = 2$–8) and Coiflet 'g' ($g = 1$–5). These wavelets were applied to all the signals. The differences between these types of wavelets are well explained in the bibliography [26].

A MATLAB[®] program was used to examine the reductions of the dimensions. Five decomposition levels and different thresholdings were applied. Considering the recomposition error of the signals as well as the number of coefficients obtained, Symmlet 3 and Symmlet 4 wavelets were selected (typical wavelets to represent symmetric signals as in these cases).

After de-noising and compression of the signals, the wavelet coefficients, which represented the reduced signals, were obtained: 9 and 11 for Symmlet 3 and Symmlet 4, respectively. Likewise, in each case the percentage of zero coefficients and the recomposition percentage of the signals after their reconstruction were obtained as well. Thus, the complete cycle of WT application was as follows: decomposition, denoising, compression and signals reconstruction.

Table 2 presents the results obtained in the processes of reduction with the two types of transforms. This table shows that the reduction percentage of coefficients obtained with FT is slightly greater. Likewise, the recomposition minimum (and the maximum) error percentages in FT were less than in WT.

Table 2. Comparison of efficacy of the methods to reduce dimensions: FT and WT

Dimension reduction methods	Number of coefficients	Decomposition level (ω = Fourier)	Percentage of reduction of coefficients	Recomposition minimum error (%)	Recomposition minimum error (%)
Symlet 3 (WT)	9	4	90.72	1.42	4.75
Symlet 4 (WT)	11	4	89.62	0.71	3.77
Fourier (spectrum of 128 frequencies)	7	ω = 4 Hz	91.25	0.27	2.62

Neural Network Analysis

Neural Networks Based on Fourier Coefficients

A set of network topologies with three layers (perceptron) was designed to resolve the mixtures of analytes by using the Fourier coefficients obtained for each signal after having applied the process to reduce dimensions. These models had 7 input nodes (the seven Fourier coefficients for each signal) and 2 output nodes (the concentrations for each ion). A low number of hidden neurons was used to avoid overfitting and overparametrization of the system and to obtain a model as simple as possible.

The characteristics of these models were as follows:

– Topology: 7-X-2, where X $\in$ {2, 3}.
– Activation functions: linear for input layer, and all the possible combinations of gaussian, sigmoid and hyperbolic tangent functions for the rest of the layers.
– Training set (trn): 32 samples.
– Monitoring set (mon): selected randomly at first and then fixed for the remaining methods, consisted of 8 samples: T1, T8, T1L1, T2L2, T8L3, T4L9, T6L1 and T1L6.
– External test set (tst): T1L9, T2L10, T3L5, T5L3, T6L8, T8L6, T9L1 and T10L2.
– Starting weights values: as they were randomly generated, three ANN runs were made and the resulting RMS's were averaged.

The root mean square (RMS) error of each set (training, monitoring and test set) was obtained for all cases. The RMS error was used as a decision parameter to find the best model. This parameter is defined by the following equation:

$$\text{RMS} = \sqrt{\frac{\sum_i (y_i - y_i^*)^2}{n}} \qquad (2)$$

where y_i represents the measured concentration for the ions; y_i^* represents the predicted concentration for the ions; and n is the number of samples.

The best neural models, i.e. with the least RMS values for both the training and the test sets obtained, were: 7-3-2 linear-tangential-gaussian (ltg) (7 nodes in the input layer with linear activation functions, 3 nodes in the hidden layer with hyperbolic tangent activation functions and 2 nodes in the output layer with gaussian activation functions) and 7-2-2 linear-tangential-gaussian (ltg) (7 nodes in the input layer with linear activation functions, 2 nodes in the hidden layer with hyperbolic tangent activation functions and 2 nodes in the output layer with gaussian activation functions). The RMS values for each set are shown in Table 3.

Both neural models were trained and validated again by varying two adaptive parameters, the learning rate (η) and the momentum (α). Both parameters optimize the rate at which a network learns. The results obtained in this way did not improve the RMS errors significantly. This indicates that the optimal neural models can be considered the previous ones.

Table 3. Comparison of RMS errors obtained with each model

Optimal model	RMS$_{(\text{trn})}$	RMS$_{(\text{mon})}$	RMS$_{(\text{tst})}$
PCR full	0.0334	0.0373	0.0498
FT + PCR	0.0382	0.0227	0.0511
WT(9) + PCR	0.0504	0.0466	0.0608
WT(11) + PCR	0.0481	0.0481	0.0592
PLS full	0.0328	0.0368	0.0496
FT + PLS	0.0378	0.0226	0.0514
WT(9) + PLS	0.0502	0.0464	0.0605
WT(11) + PLS	0.0480	0.0468	0.0589
7-3-2 ltg	0.0235	0.0264	0.0447
7-2-2 ltg	0.0254	0.0269	0.0426
9-3-2 lss	0.0225	0.0297	0.0774
9-2-2 lgt	0.0299	0.0256	0.0475
11-3-2 lsg	0.0229	0.0169	0.0486
11-2-2 lsg	0.0230	0.0164	0.0474

$RMS_{(\text{trn})}$ = RMS error for training set; $RMS_{(\text{mon})}$ = RMS error for monitoring set; $RMS_{(\text{tst})}$ = RMS error for external test set; l = lineal activation function; g = gaussian activation function; s = sigmoid activation function; t = hyperbolic tangent activation function.

The predicted values for the concentrations of both thallium and lead ions were obtained by utilizing the best models, and the relative percentage errors were below 6% on average.

Neural Networks Based on Wavelet Coefficients

In this case, the wavelet coefficients that were obtained with the Symmlet 3 and Symmlet 4 wavelets (9 and 11 coefficients, respectively) were employed as input vectors in the neural networks.

Neural models with three layers were developed, similarly to the neural networks based on Fourier coefficients. Their topologies were as follows: 9-X-2 for Symmlet 3 (9 input nodes, one for each coefficient; 2 output nodes, one for each concentration of the ions; $X \in \{2, 3\}$ hidden neurones) and 11-X-2 for Symmlet 4 (11 input nodes, one for each coefficient; 2 output nodes, one for each concentration of the ions; $X \in \{2, 3\}$ hidden neurones).

The same training parameters, and the same training, monitoring and test sets as used for the Fourier models, were applied, and overfitting was avoided in the same way.

A set of different neural models was tested for each case: 9 and 11 wavelet coefficients from Symmlet 3 and Symmlet 4, respectively. The best models using Symmlet 3 coefficients, with the lowest RMS errors, had the following topologies: 9-3-2 linear-sigmoid-sigmoid (lss) and 9-2-2 linear-gaussian-tangential (lgt). The RMS error values appear in Table 3. The relative percentage errors were on average below 8% for the two models.

The best models using Symmlet 4 coefficients produced these topologies: 11-3-2 linear-sigmoid-gaussian (lsg) and 11-2-2 linear-sigmoid-gaussian (lsg). The RMS errors are shown in Table 3 as well. The RMS values were less than in the case of models with 9 wavelet coefficients and quite similar to those obtained with Fourier models. Here, the relative percentage errors were also on average below 6%.

As in the Fourier case, after trying to refine these models using the same procedure, the results obtained were not significantly better than the previous ones. That is why the neural models, shown in Table 3, are considered the best ones for this proof binary system.

Comparative Study of Both Kinds of Neural Networks

To compare the ability of reducing and retaining information, we obtained the improving percentages of each neural model based on wavelet coefficients compared to the models constructed using Fourier coefficients. The comparison was established between wavelet and Fourier models with the same number of hidden neurons. The percentage of improvement was remarkably better in the case of the $RMS_{(mon)}$ values, 17% on average; i.e. when using a wavelet procedure of compression, the results for the monitoring set improve by approximately 17% in comparison to a Fourier procedure. For the $RMS_{(trn)}$ errors, the results were very similar in all cases and, finally, as shown in Table 3, the best $RMS_{(tst)}$ values were obtained with Fourier transforms, which indicates that there is no improvement in this case when using a wavelet model.

In general, the neural models with 11 coefficients allow greater improvements than models with 9 coefficients. There is reason to assume that neural models using 11 wavelet coefficients are better than those with 9 coefficients. Using more wavelet coefficients to compress and reconstruct the signals, and consequently maintaining more information, may affect this situation. However, almost all mixtures were predicted with concentration errors lower than a hundredth of $mg\,L^{-1}$, independently of the model used.

It has been proved that the prediction capability as well as the recomposition percentage obtained in this paper were slightly better with Fourier models, despite the more advantageous characteristics of WT as a pre-processing tool. The reason is the form of the signal which is usually gaussian (their fitting errors are lower than 2%). Since all gaussian functions can decompose in a sum of sines and cosines functions, it is not surprising that FT adapts better than WT in the representation of this type of signals and in the process of reducing dimensions. Furthermore, the use of Fourier coefficients allows you to construct simpler neural models than using wavelet coefficients.

Multivariate Calibration Methods

PCR Analysis

The PCR analysis was carried out with the aid of the software package Unscrambler® 7.01 and the variables chosen as follows: independent variables were 1) the 80 points of all voltammograms of the samples (PCR full), and 2) the Fourier and wavelet coefficients obtained after pre-processing the initial data with the respective transforms (FT + PCR and WT + PCR); dependent variables were the concentration values of each ion for each sample. All data was centered, and
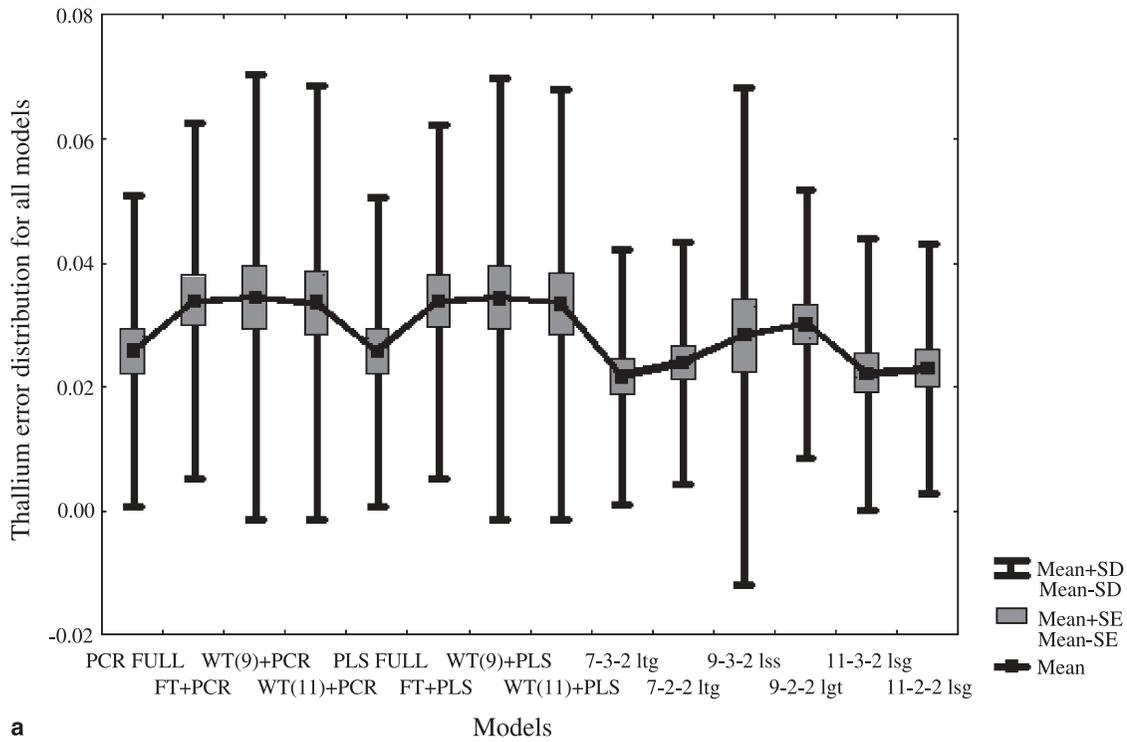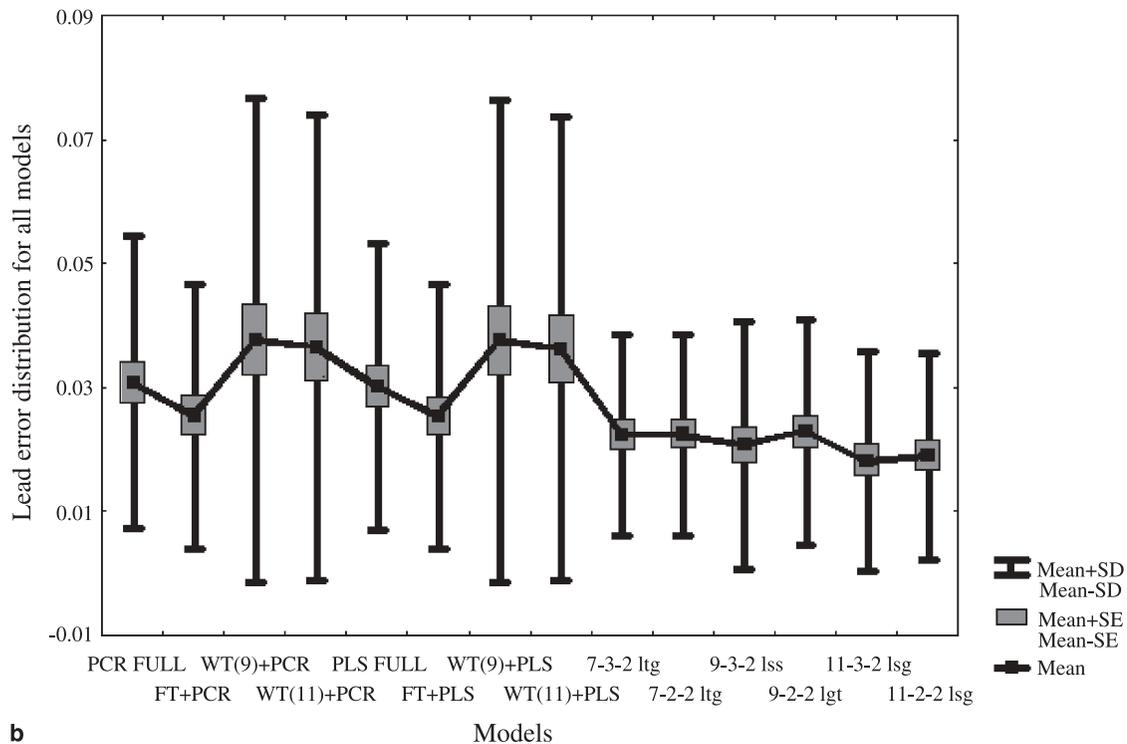
Fig. 3. Box & whiskers plot of thallium (*a*) and lead (*b*) error distributions for all models: PCR (principal component analysis); PLS (partial-least square); FULL (using all the initial data without reduction of dimensions); FT (using Fourier coefficients); WT (using wavelet coefficients: (9) for Symmlet 3 and (11) for Symmlet 4); 7-3-2 and 7-2-2 (Fourier neural models topologies); 9-3-2, 9-2-2, 11-3-2 and 11-2-2 (wavelet neural models topologies); ltg (linear-tangential-gaussian); lss (linear-sigmoid-sigmoid); lgt (linear-gaussian-tangential); lsg (linear-sigmoid-gaussian)

the model was validated using the same monitoring set as for neural models. The rest of samples (32) made up the training set (trn).

The best number of components, which explained the greatest variance, was 3 in PCR full model and 2 in FT + PCR and WT + PCR models. This means that a pre-processing step reduces the number of principal components. The RMS errors obtained for both the training and the monitoring set are shown in Table 3. The external test set was used to check the predictive ability of the models as well. The RMS error values can also be found in Table 3.

All models obtained make predictions in a similar way, but RMS errors are lower in PCR full and FT + PCR models.

PLS Analysis

Using the same software and developing the analysis as PCR (with the same variables, procedure, validation method and training, monitoring and external sets), the best number of components was 3 in PLS full model and 2 in FT + PLS and WT + PLS models, as in the previous analysis. The RMS errors for each set of samples can be seen in Table 3. As shown, the results are not much different than the PCR ones.

As in PCR analysis, RMS errors are lower in PLS full and FT + PLS models.

Comparison Between all Statistical Methods

The lowest $RMS_{(trn)}$ and $RMS_{(mon)}$ errors appear with wavelet models using 11 coefficients, and Fourier neural models show the lowest $RMS_{(tst)}$ errors (Table 3). In general, PCR and PLS models have a worse predictive ability with the external test set. Nevertheless, PCR full, PLS full, FT + PCR and FT + PLS models have $RMS_{(tst)}$ values very similar to Fourier and wavelet neural models. This means that the use of Fourier transforms as a pre-processing procedure produces less RMS errors independently of the multivariate calibration technique used to resolve the mixtures. Besides, in the case of PCR and PLS full models without pre-processing, their good performance can be attributed to the linearity of the binary system of mixtures resolved here. In general, the full methods that give a better response to solve the two overlapped peaks are neural models with FT and WT preprocessing. They perform slightly better than PLS and PCR, the traditional models which would be the

most suitable methods for resolving this kind of problem, even when the initial data presents linearity.

The box and whiskers plots of Fig. 3 were obtained by using the thallium and lead error distributions obtained for all different models with the same training, monitoring and test sets. As shown, Fourier and wavelet neural models (Fourier and wavelet transform as previous techniques applied to neural networks) provide a more sensitive precision than linear models PLS and PCR, except for model 9-3-2 lss. WT + PLS and WT + PCR models offer the worst precision with both ions.

Conclusions

The models combining transforms and neural networks presented here are able to predict the concentrations of the two ions in the mixtures slightly better than the traditional techniques, PLS and PCR, even when the initial data presents high linearity. This is due to the ability of Fourier and wavelet transforms to detect information of high frequency due to the hard overlapping between the two signals of the ions, which is what linear models (PLS and PCR) are not able to do.

Both FT and WT have demonstrated to be tools of similar performance with voltammetric signals, but the topologies obtained with FT are simpler. For this reason, the use of FT may be considered more adequate for the signals studied in this paper: it is possible that the use of WT would be preferable for another type of signals, more asymmetric or further than the gaussian form.

References

[1] P. T. Kissinger, W. R. Heineman, *Laboratory Techniques in Electroanalytical Chemistry*, 2nd Edn., Marcel Dekker, New York, 1996.
[2] I. Naranjo-Rodríguez, J. L. Hidalgo-Hidalgo-de-Cisneros, Organic Analysis in Environmental Samples by Electrochemical Methods (Encyclopedia of Analytical Chemistry: Instrumentation and Applications). In: R. A. Meyers (Ed.) John Wiley & Sons, United Kingdom, 2000, vol. 4, p. 3035–3064.
[3] Z. Grabarić, B. S. Grabarić, M. Esteban, E. Cassasas, Determination of Small Amounts of Analytes in the Presence of a Large Excess of One Analyte From Multivariate Global Signals of Differential-Pulse Voltammetry and Related

Techniques With the Signal Ratio Resolution Method, *Analyst* **1996**, *121*, 1845.

[4] R. M. De Carvalho, C. Mello, L. T. Kubota, Simultaneous Determination of Phenol Isomers in Binary Mixtures by Differential Pulse Voltammetry Using Carbon Fibre Electrode and Neural Networks With Pruning as a Multivariate Tool, *Anal. Chim. Acta* **2000**, *420*, 109.

[5] F. Navarro-Villoslada, L. V. Pérez-Arribas, M. E. León-González, L. M. Polo-Díez, Selection of Calibration Mixtures and Wavelengths for Different Multivariate Calibration Methods, *Anal. Chim. Acta* **1995**, *313*, 93.

[6] J. R. Morrey, On Determining Spectral Peak Positions From Composite Spectra With a Digital Computer, *Anal. Chem.* **1968**, *40*, 905.

[7] W. Huang, T. L. E. Henderson, A. M. Bond, K. B. Oldham, Curve Fitting to Resolve Overlapping Voltammetric Peaks: Model and Examples, *Anal. Chim. Acta* **1995**, *304*, 1.

[8] I. Pižĕta, Deconvolution of Non-Resolved Voltammetric Signals, *Anal. Chim. Acta* **1994**, *285*, 95.

[9] C. Cai, P. de B. Harrington, Wavelet Transform Preprocessing for Temperature Constrained Cascade Correlation Neural Networks, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 874.

[10] A. Cladera, J. Alpízar, J. M. Estela, V. Cerdà, M. Catasús, E. Lastres, L. García, Resolution of Highly Overlapping Differential Pulse Anodic Stripping Voltammetric Signals Using Multicomponent Analysis and Neural Networks, *Anal. Chim. Acta* **1997**, *350*, 163.

[11] C. Bessant, S. Saini, Simultaneous Determination of Ethanol, Fructose and Glucose at an Unmodified Platinum Electrode Using Artificial Neural Networks, *Anal. Chem.* **1999**, *71*, 2806.

[12] W. Hongmei, W. Lishi, X. Wanli, Z. Baogui, L. Chengjun, F. Jianxing, An Application of Artificial Neural Networks. Simultaneous Determination of the Concentration of Sulphur Dioxide and Relative Humidity With a Single Coated Piezoelectric Crystal, *Anal. Chem.* **1997**, *69*, 699.

[13] U. Depczynski, K. Jetter, K. Molt, A. Niemöller, The Fast Wavelet Transform on Compact Intervals as a Tool in Chemometrics. II. Boundary Effects, Denoising and Compression, *Chemom. Intell. Lab. Syst.* **1999**, *49*, 151.

[14] E. Llobet, J. Brezmes, R. Ionescu, X. Vilanova, S. Al-Khalifa, J. W. Gardner, N. Bârsan, X. Correig, Wavelet Transform and Fuzzy ARTMAP-Based Pattern Recognition for Fast Gas Identification Using a Micro-Hotplate Gas Sensor, *Sens. Actuators B* **2002**, *83*, 238.

[15] X. Shao, W. Cai, P. Sun, Determination of the Component Number in Overlapping Multicomponent Chromatograms Using Wavelet Transform, *Chemom. Intell. Lab. Syst.* **1998**, *43*, 147.

[16] P. J. Gemperline, J. R. Long, V. G. Gregoriou, Non-Linear Multivariate Calibration Using Principal Components Regression and Artificial Neural Networks, *Anal. Chem.* **1991**, *63*, 2313.

[17] D. Wienke, G. Kateman, Adaptive Resonance Theory Based Artificial Neural Networks for Treatment of Open-Category Problems in Chemical Pattern Recognition – Application to UV-Vis and IR Spectroscopy, *Chemom. Intell. Lab. Syst.* **1994**, *23*, 309.

[18] C. Krantz-Rülcker, M. Stenberg, F. Winquist, I. Lundström, Electronic Tongues for Environmental Monitoring Based on Sensor Array and Pattern Recognition: A Review, *Anal. Chim. Acta* **2001**, *426*, 217.

[19] E. Cukrowska, L. Trnková, R. Kizek, J. Havel, Use of Artificial Neural Networks for the Evaluation of Electrochemical Signals of Adenine and Cytosine Mixtures Interfered With Hydrogen Evolution, *J. Eletroanal. Chem.* **2001**, *503*, 117.

[20] J. Zupan, Neural Networks and Pattern Recognition. Academic Press, USA, 1998.

[21] J. Zupan, J. Gasteiger, Neural Networks for Chemists: An Introduction. VCH, Weinheim, 1992.

[22] F. Despagne, D. L. Massart, Neural Networks in Multivariate Calibration, *Analyst* **1998**, *123*, 157R.

[23] J. Zupan, J. Gasteiger, Neural Networks: A New Method for Solving Chemical Problems or Just a Passing Phase?, *Anal. Chim. Acta* **1991**, *248*, 1.

[24] M. C. Ortíz, J. Arcos, L. Sarabia, Using Continuum Regression for Quantitative Analysis With Overlapping Signals Obtained by Differential Pulse Voltammetry, *Chemom. Intell. Lab. Syst.* **1996**, *34*, 245.

[25] R. G. Brereton, Chemometrics in Analytical Chemistry, *Analyst* **1987**, *112*, 1635.

[26] R. Todd Ogden, Essential Wavelets for Statistical Applications and Data Analysis. Birkhäuser, Boston, 1997.