Application of the wavelet transform coupled with artificial neural networks for quantification purposes in a voltammetric electronic tongue

L. Moreno-Barón a, R. Cartas a, A. Merkoçi a, S. Alegret a, M. del Valle a, b, L. Leija b, P.R. Hernandez b, R. Muñoz b

a Grup de Sensors i Biosensors, Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalunya, Spain
b Sección Bioelectrónica, Departamento de Ingeniería Eléctrica, CINVESTAV, Ciudad de México, México

Received 1 November 2004; received in revised form 9 March 2005; accepted 14 March 2005
Available online 22 April 2005

Abstract

This work describes a voltammetric electronic tongue, in which the quantitative information contained in voltammograms obtained from amperometric sensors is firstly extracted employing the discrete wavelet transform (DWT) and then processed employing artificial neural networks (ANNs). The analytical case studied is the direct determination of the oxidizable aminoacids tryptophan, cysteine and tyrosine, and its application in the direct measurement of these amino acids in animal feed samples. A conventional voltammetry cell with a Pt working electrode is the experimental set-up and differential pulse voltammetry the selected technique. Due to the complexity of the obtained signals, the DWT pre-treatment was needed in order to eliminate noise components and compress voltammograms by selecting and extracting significant information. The ANN was subsequently used to model the system departing from the reduced information, and obtaining the concentrations of the considered species. Best results were obtained when using two hidden layers in a backpropagation neural network trained with the Bayesian regularization algorithm.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Voltammetric electronic tongue; Wavelet transform; Artificial neural networks; Oxidizable amino acids

1. Introduction

The lack of selectivity is a key point which in practice limits many applications of sensors and biosensors currently being developed. While the main progress line has been to further improve the selectivity of existing sensors or to add masking steps or convert species to more tolerant forms, a trend from last years is to extract multivariable information and to apply chemometric methods to deal with the response model. This has been successfully applied in the case of voltammetric sensors, by processing the entire voltamgram [1], or, in the case of optical sensors [2], by processing the entire absorption, reflection or fluorescence spectra. In the case of potentiometric sensors, a different scheme with the same underlying idea has been applied, which is the use of an electrode array intended to correct the insufficient selectivity, instead of one discrete, more ideal sensor. In all cases, the underlying concept is a multisensor system based on the coupling of multichannel information generated from non-specific sensors and pattern recognition tools [3,4].

This concept has coined the specific terms electronic nose and electronic tongue, denoting sensors for gases or sensors for liquid media, respectively. The electronic nose [5], firstly conceived and applied, has gained recognition in fields like food, aroma or medical diagnosis. The electronic tongue [6] has special interest in the quantification of multiple analytes. A correlation between their output and human sensory assessments made by taste panels opens up interesting approaches for the food industry. The electronic tongue may be the underlying concept in new analytical systems for quantitative and qualitative analysis of complex samples.

An important part of an electronic tongue is obviously the sensor array. Or, in the case of the application presented in...
this study, it would be the electrochemical technique itself, which provides a complete voltammogram for each experiment [7]. In a multicomponent environment, the sensor array produces complex signals (patterns), which contain information about different compounds plus other features. Hence, the second part in multisensor approaches is the signal processing stage. In this sense, different approaches, like principal component analysis, partial least squares and artificial neural networks have been used [8–10]. The choice of the data processing technique for a particular case depends on the task to be solved and the structure of the data (nonlinearity, correlations, etc.). Neural networks are an interesting processing alternative, because they try to reproduce the logical operations performed by an animal brain. They accomplish this feature by the use of an entity, known as artificial neuron, which processes the obtained signals as animal neurons do, in a biomimetic approach [11]. From here, a new group of bioinspired instrumental techniques has arisen with the use of multiple sensor signals and artificial neural networks (ANNs). This trend is of main concern, consisting in an artificial duplicate of the learning process and the response ability of the human nervous system, trying to replicate the function of our senses and our brain.

Electronic tongues use different multivariate tools for subsequent prediction of several analytes in overlapped, and cross-response signals. They have proved to be new and represent interesting alternatives. The reliable performance of electronic tongues in recognition tasks (classification, identification or discrimination) has been demonstrated along the last few years. Potentiometric sensor arrays still are the most widely used type of electronic tongue systems [12–15]. Because of their novelty, such systems still need to gain confidence as quality control tools in the food industry, medicine or environmental fields.

An appropriate sensor system will be the first requirement in order to attempt this approach. Additionally, the systems responses must cover the different chemical species and the dynamic range of concentrations expected. These are considered the departure point to build the model needed to create an intelligent system. In this way, the set-up will be able to predict responses of samples not processed initially, as well as to classify them as the human brain does. When ANNs are used, the process includes a transformation of n-dimensional vectors that represent the signals, into an m-dimensional vector. This vector contains the main features of the sample. The networks have the capability to learn from a set of representative samples of the actual problem to be modeled.

In some cases the departure information is too complex to be handled by the modeling tools, for example, when a complete spectra or voltammogram with ANNs needs to be processed. The signal is so complex that the network performance is not satisfactory. This is, a network architecture with hundreds of input neurons, as required by a conventional voltammogram, quickly becomes unpractical [16]. Therefore, it is required to perform some pre-processing of the incoming system information [17,18]. This has been traditionally accomplished by the use of principal component analysis, or by the use of a subset of coefficients obtained from a Fourier transformation [16,19–23]. Whatever technique is used, it must bring two objectives: to reduce the complexity of the input signal and to preserve the amount of relevant information. In order to perform this task, the wavelet transformation was applied in the present work to solve the problem, operating as a tool to extract the significant information from the voltammetric signals [24,25].

Wavelet transform (WT), a tool of applied mathematics, is a high performance signal processing technique developed from the Fourier transform (FT) during the late 1980s [26–28]. This technique is used for signal decomposition onto a set of basis functions. These functions are obtained by dilations and translations of a unique function called mother wavelet [29]. Dilation, also known as scaling, compresses or stretches the mother wavelet and translation shifts it along the time (or x) axis. Thus, WT decomposes the signal in labelled and located contributions for one scale and one position parameter. Every contribution represents the information of frequency bands contained in a specific segment of the analysed signals. An additional advantage of this approach is the facility to remove the noise present in signals. Alternatively, the WT can be used for data compression if only the most significant decomposition coefficients are kept for further treatment.

In the last decade, numerous applications of the WT have been proposed for chemical analysis, including denoising and smoothing, data compression, baseline correction, resolution of overlapping signals, electroanalytical chemistry and chromatography [30]. Shao and Sun applied the WT to the resolution of overlapping chromatograms, and studied the ability of different wavelet base functions at different noise levels [31]. Coelho et al. presented a work addressed to optimize a methodology to maximize the wavelet compression ability of ICP-AES data for the multicomponent determination of Mn, Mo, Cr, Ni and Fe in steel samples [32]. Depczynski et al. demonstrated the use of the wavelet coefficient regression in combination with a genetic algorithm for spectral calibration [33].

Frequently, a degree of overlapping among different components is present if different species undergo oxidation or reduction at similar potentials, when using voltammetry. WT has resulted efficient in this situation, since it offers the advantage of performing data number reduction, feature extraction and noise reduction at the same time. The overlapping in voltammograms from mixtures of metals were satisfactorily resolved qualitatively and quantitatively [34]. Interesting contributions in this field have been made by the group of Winquist in Sweden [7,35–37], who used an array of different metallic working electrodes, in what they term voltammetric electronic tongue. Their work uses a specific polarization strategy, in which a series of potential pulses of decreasing or increasing amplitude are applied to each metallic electrode and current transients are collected. In this sense, Holmin et al. used electrodes made of platinum metals and
copper, glassy carbon, and silver showing promising capabilities for classification when applied in multianalyte samples, like washing formulations [37]. Recently, Cocchi et al. described a close approach to the work presented here, in which WT is used for feature selection prior to quantitative calibration using ANNs. Compressed voltammetric signals of Pb(II)/Tl⁺ mixtures were processed with WT and modelled with ANNs [38,39].

In this work, we apply the principles of the voltammetric electronic tongue to solve a mixture of three components by direct voltammetric analysis. The overlapped voltammetric signal was obtained from the differential pulse voltammetric response of the three oxidizable amino acids tryptophan, cysteine and tyrosine. The signal is first compressed and then its features are extracted employing WT. For this purpose, the base function and the compression level have been studied in order to reconstruct the original signal. Next, the compressed information is employed to construct a calibration model employing ANNs. The final application developed is the direct measurement of these amino acids in animal feed samples. With this approach, a successful combination of chemometric techniques for the use of complex signals in voltammetric electronic tongues is demonstrated.

2. Theory of the wavelet transform (WT)

Wavelet transform (WT) was developed for the analysis of non-stationary signals. By WT, a function \( f(t) \) can be represented by a family of functions called wavelets, generated by translating and dilating a single base function called mother wavelet. Dilation, also known as scaling, compresses the mother wavelet. Dilation, also known as scaling, compresses the mother wavelet, and specifies its position along time axis. Technically, we can say that scale parameter is related with the spectral content and specifies its position along time axis. We can say that scale parameter is related with the spectral content of function \( f(t) \) at a definite position (translation parameter).

The family of wavelet functions is related with the mother wavelet by:

\[
\Psi_s(t) = \frac{1}{\sqrt{s}} \Psi \left( \frac{t - \tau}{s} \right)
\]  

(1)

In Eq. (1), \( s \) is the scaling factor, \( \Psi \) the mother wavelet, \( \tau \) the translation parameter, and \( s^{-1/2} \) a factor used to ensure that all the wavelets at every scale have the same energy of the mother wavelet. Continuous wavelet transform (CWT) of a function \( f(t) \) is defined as \( (\ast) \) indicates complex conjugate:

\[
W(t, s) = \int_{-\infty}^{\infty} f(t)\Psi^*_s(t)dt
\]  

(2)

The CWT, as described in Eq. (2), cannot be used in practice. In Eq. (2) translation and scale parameters are continuous variables, which mean that a function \( f(t) \) might be decomposed in an infinite number of wavelet functions. To overcome this problem, discrete wavelet transform (DWT) has been introduced. In DWT translation and dilation parameters take discrete values instead of continuous. DWT can be expressed as Eq. (3) from modifying Eq. (1) to:

\[
\Psi_{j,k}(t) = \frac{1}{\sqrt{s^j}} \Psi \left( \frac{t - \tau}{s^j} \right)
\]  

(3)

In Eq. (3), \( j \) and \( k \) are integers, \( s_j \) the dilation factor and \( \tau_j \) the translation factor. By doing \( s_j \) and \( \tau_j \) equal to 2 and 1 respectively, we will have dyadic scales and positions permitting the implementation of efficient algorithms for the processing of discrete signals.

\[
\Psi_{j,k}(t) = 2^{-j/2} \Psi(2^{-j}t - k)
\]  

(4)

The DWT is implemented using Mallat’s pyramidal algorithm [40], as sketched in Fig. 1. This signal processing technique operates over a single discrete signal of length \( M \) by decomposing it into orthogonal sub-spaces of length ca. \( M/2 \). Decomposition is made by applying two digital filters, which involves low-pass (LPF) and high-pass (HPF) versions and downsampling. The result of such decomposition is a series of approximation coefficients \( c_A \) and detail coefficients \( c_D \). The \( c_A \) set and \( c_D \) set retain the low-frequency and high-frequency content of the signal, respectively. Moreover, this decomposition can be iteratively applied to approximation components to get components of lower resolution and obtain what is known as multiresolution analysis.

For each decomposition level \( j \) a faithful reconstruction of the original signal is possible using the inverse discrete wavelet transform (IDWT) and the set of approximation coefficients obtained at level \( j-1 \) and all sets of detail coefficients from level \( j-1 \) until level 1. Depending on the main goal pursued (noise suppression, data compression, singularities detection, etc.) a set of wavelet components is chosen as the result. In this work, DWT is applied to voltamograms for their compression and extraction of their main features in order to facilitate the training of an ANN used as calibration model.
3. Experimental

3.1. Reagents and solutions

All chemicals for electrolyte and the stock amino acid solutions tryptophan (Trp), cysteine (Cys) and tyrosine (Tyr) were purchased from Merck. The buffer electrolyte solution consisted of 0.1 M potassium chloride + 0.1 M phosphate solution (pH was adjusted to 7.5). Synthetic mixtures for a preliminary evaluation of the voltammetric method were prepared from 0.1 M stock solutions of each amino acid.

3.2. Samples

Three series of synthetic solutions were prepared for Trp, Cys and Tyr analysis. For each analyte, six concentration levels were considered as follows: 5.0, 10, 20, 25, 30 and 35 mM for Cys and Tyr; 2.0, 6.0, 10, 14, 17 and 21 mM for Trp. Interferences were studied at two levels: 10 and 25 μM for Trp and Cys; 5.0 and 34 μM for Tyr. As a result, each analyte series was composed by 24 mixture solutions, and the set of voltammograms processed 72.

Animal feed samples were kindly provided by Cooperativa Agropecuaria de Guissona, Lleida, Spain. Free (non-proteic) oxidizable amino acids present in feed samples were collected following a leaching procedure as follows: 5 g of feed sample were treated with 50 ml of 0.1 M HCl solution for 30 min in a conical flask with magnetic stirring. The resulting extract solution was subsequently filtered through a nylon membrane of 0.45 μm of pore size and refrigerated. Free amino acids are added to animal feeds as nutritional fortifiers.

3.3. Apparatus and measurement procedure

A PGSTAT 20 Autolab potentiostat with a Pt working electrode (Crisson, Barcelona) was used for differential pulse voltammetric measurements. The cell was completed by a second Pt counterelectrode (Crisson, Barcelona) together with a Ag/AgCl reference electrode (Crisson, Barcelona). The resulting voltammetric data consisted of current intensities registered in the range of potentials from 0.4 to 1.0 V in steps of 0.00365 V. Hence, 164 data points per sample were registered which constituted the multivariate signal for further analysis. Each voltammogram was recorded by using the differential pulse voltammetric technique. The modulation amplitude was 0.025 V, the modulation time 70 ms and the pulse interval 300 ms. No preconditioning was performed.

For the series of prepared samples, microvolumes of each amino acid stock solution were added to 25 ml of the buffer electrolyte solution, constituted by 0.1 M Na2HPO4 + 0.1 M KCl + 5 ml of animal feed extract, to match real application samples. Hence, the samples considered were hydrolysed animal feed in 0.1 M HCl medium, with added quantities of Trp, Cys and Tyr amino acids.

3.4. Software

WT as well as ANN modelling were implemented employing Matlab 6.1 (MathWorks, Natick, MA) with the aid of its Neural Network (version 4.0) and Wavelet (version 2.0) toolboxes.

4. Results and discussion

In the proposed coupling of WT with ANNs for the modelling of complex multivariate electrochemical signals, the smoothing and compression of the data is performed prior to the building of the calibration model. The data that represent the universe of information to be compressed, prior to their processing employing ANNs, were obtained from the set of voltammograms of different prepared amino acid samples. Fig. 2 shows the nature of the experimental voltammograms recorded, observing a high degree of overlapping. The three oxidizable amino acids were the targets to be quantified. Thus, the departure data formed a matrix with 72 pattern vectors, 164 current elements each, plus a target matrix with 72 triplets of concentrations (of the three amino acids considered Trp, Cys and Tyr). The matrix forming the patterns contained therefore a total of 11,808 elements. The dimension of each pattern vector of 164 current values per sample, represent an impractical input to be applied to ANNs. In this way, after compression, the reduced voltammograms were the patterns fed to the ANN. The proposed choice is then the compression of these input data, trying simultaneously to smooth and denoise this signal, along with the preservation of any significant information on each vector.

4.1. Selection of the mother Wavelet and level of decomposition

The compression percentage in a WT processed voltammogram establishes the amount of information preserved: the greater percentage, the less information is kept. Since linear sweep electrochemical signals contain their significant information at the low frequency scale (frequency) region [41], hence the approximation coefficients cA are used as the compressed information.

The mother wavelet and decomposition level, for best compression and smoothing of our voltammograms, were chosen taking into account the degree of similarity between the original voltammogram and the one reconstructed from approximation coefficients after compression. To quantify the similarity we propose a comparison factor named f. This factor that considers the area under both signals when superimposed. The discrimination capability is compared to the conventional correlation coefficient R.

The f is defined as the ratio of the area intersected by both curves to the total area under both curves. From the set theory, with A and B as the areas under each curve the f can
be expressed as $f_c = \frac{(A \cap B)}{(A \cup B)}$. This factor ranges from 0 to 1 depending on similarity; it values 0 when two signals have nothing in common and increases its value as similarity does. In this sense, $f_c$ computes similarity in a way related to a correlation coefficient $R$.

Fig. 3 shows graphically the terms used to calculate $f_c$. The shadowed area on the left part of the figure represents the area intersected by both curves, while the shadowed area on the right represents the total area covered by both curves.

Considering discrete signals of length $M$, the $f_c$ factor is expressed by:

$$f_c = \frac{A \cap B}{A \cup B} = \frac{\sum_{i=1}^{M} (\max(a_i, b_i) - \min(a_i, b_i))}{\sum_{i=1}^{M} (\min(a_i, b_i) - |a_i - b_i|)}$$

(5)

where $a_i$ is the $i$th data point of the original voltammogram, $b_i$ the $i$th data point of the reconstructed signal from approximation coefficients.

4.2. Selection of the ANN

To choose the adequate ANN for our application, several factors were considered:

- The number of neurons in the input layer is equal to the number of approximation coefficients retained after application of DWT.
- The number of neurons in output layer depends on the number of species to be predicted (the three amino acids Trp, Cys and Tyr). Under this context, we considered two cases in ANN structure: (i) one network with three neurons in its output layer, one per aminoacid, (ii) one network per aminoacid, we mean, three parallel ANNs with one output neuron each.
- The number of neurons and layers in ANN architecture is normally determined by trial and error. We first considered one hidden layer with a number of neurons equal to the geometric mean of inputs and outputs. Afterwards, the number of neurons in the hidden layer was increased in order to improve the modelling performance. Once the number of neurons in this structure was set, an additional check was made to test the performance of a two-hidden layer network. In such case, the sum of neurons in both hidden layers was made equal to the number of neurons in the one hidden layer structure.
- ANNs tested were feedforward backpropagation networks with supervised learning. Learning algorithms tested were...
those already preprogrammed in the MATLAB environment: conjugate gradient, descend gradient, descend gradient with adaptive learning rate, descend gradient with momentum, descend gradient with momentum and adaptive learning rate, Levenberg–Marquardt optimization algorithm and Bayesian regularization algorithm.

Matrices of approximation coefficients and targets were normalized to the range \([-1, 1]\), and, split into training and test subsets (75% of data set taken for training). The generalization capability of the trained network was computed for the test subset through the relative absolute error (RAE), corresponding to estimated versus expected target values:

\[
\text{RAE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{V_t(i) - V_e(i)}{V_t(i)} \right| \times 100
\]

where \(V_t\) is the true value, \(V_e\) the estimated value and \(N\) the number of points in the test subset. RAE values greater than 10% made us consider its network as with low generalization ability.

Fig. 4 shows the flow chart used to implement the ANN. The goal in this process is to get a balance between network performance to model training data and its generalization capability to data not used in training process. For tracking the network error during training we used the sum of squared errors (SSE), defined as the sum of squared differences between the outputs of the network and corresponding expected values. For testing evaluation we used the RAE. Goal values for convergence were set at SSE = 0.001 and initial value for RAE = 15%. The process stopped when 200 training epochs were completed or when SSE achieved the initial goal. The next stage is then the interpolation of the test subset and the calculation of its RAE value. If this new RAE is smaller than the starting value, then the trained network is temporally saved, it is iteratively retrained till RAE is not further improved or if a week of calculation time has elapsed. The described algorithm was implemented in Matlab v6.3 using a Celeron processor desktop computer @ 1 GHz and 128 Mbytes of RAM. Typical calculation times ranged from 1 to 7 days.

4.3. Smoothing and compression

The goal in compression is to represent the original signal with the lowest number of wavelet coefficients. In addition to compression, the signal was smoothed with the DWT by removing the noise. The number of data points obtained from a compressed voltammogram will depend on the mother wavelet and decomposition level used. To perform a voltammogram compression, only the approximation coefficients obtained with DWT were considered. Therefore, the reconstructed signal will be smoothed and could be slightly distorted in some cases. One hundred and sixty-four data points originally composed our voltammograms. The mother wavelets evaluated were coiflets, daubechies, biorthogonal and symlets. After the application of the DWT, we obtained from 96 approximation coefficients for wavelet coiflets with order 5 and decomposition level 1, to seven approximation coefficients for wavelet biorthogonal of order 2.6 and decomposition level 5. Fig. 5 shows four bar graphs that compare the number of approximation coefficients obtained with DWT when four different mother wavelets of different order and under different decomposition levels were evaluated.

In our work, a factor \(f_c\) to evaluate the similarity between the original and the reconstructed signals is proposed. It is
Fig. 5. Bar graphs showing the number of approximation coefficients obtained with five decomposition levels using wavelets (A) coiflets, (B) daubechies, (C) biorthogonal and (D) symlets of different orders. Normally, voltammograms were reduced to vectors having 7–30 approximation coefficients.

well known that this task is usually performed by the correlation coefficient $R$. For this reason, the factor $f_c$ was referenced to the factor $R$ for validation. The value of $R$ resulted greater than $f_c$ for all the wavelets and decomposition levels used but less detail sensitivity in the reconstructed signals was observed. For example, using coiflets at level 3, the mean values were 0.9688 ± 0.0252 and 0.9867 ± 0.0169 for $f_c$ and $R$, respectively. Mean values for $f_c$ and $R$ for nine orders of Daubechies’ wavelets at five decomposition levels applied to the set of voltammograms, are shown in Fig. 6. In this case, $f_c$ and $R$ factors had similar trends, but the factor $f_c$ presented better discrimination capability than $R$, specially for the first levels, as can be seen by the distance between the markers (○) and (□).

Voltammograms were compressed using 165 combinations obtained with 33 different mother wavelets and five different decomposition levels. Each time the matrix of voltammograms was compressed, it was reconstructed and compared against the original data using the factor. The mother wavelet and decomposition level finally selected were

Fig. 6. Mean values of $f_c$ (left side) and $R$ (right side) factors obtained from comparing the raw voltammograms with reconstructed signals. The daubechies wavelet family (db2–db10) was used to compact signals at five levels: (●) level 1, (□) level 2, (×) level 3, (+) level 4 and (*) level 5.
chosen considering that: (i) similarity between original signals and reconstructed were greater than \( f_c 0.95 \), and (ii) the least number of approximation coefficients were used.

The number of approximation coefficients obtained with wavelets evaluated at decomposition level 5 varies from 7 to 33. However, not all reconstructions performed with this level showed an \( f_c \) greater than 0.95. Nevertheless, approximation coefficients obtained with decomposition level 4 produces a better reconstruction of voltammograms. For this level, the number of approximation coefficients obtained goes from 13 to 37 and \( f_c \) is greater than 0.96. In order to get a balance between the least possible number of approximation coefficients and a \( f_c \) greater than 0.95, voltammogram compression using the fourth order Daubechies’ wavelet and a decomposition level 4, obtaining 16 approximation coefficients and a \( f_c \) of 0.965 \((R=0.984)\), was performed.

As the decomposition level increases, the frequency band of the low-pass filter (LPF) is shifted to the lower frequency and applying a downsampling to the filtered signal, the number of approximation coefficients is reduced. The method used in the Mallat’s algorithm to obtain the approximation coefficients is similar to the decimation technique utilized in the sampling rate conversion. The performance of the fourth order Daubechies’ wavelet observed in voltammograms compression, was compared to the conventional method of decimation (downsampling 1 out of 10 original values). In this case, each raw voltammogram was reduced from 164 to 16 data points.

### 4.4. WT–ANN modelling

In order to choose the ANN that best fulfilled our purpose, different structures with supervised learning were considered. All these used an input layer with 16 neurons, equal to the compressed voltammogram (16 approximation coefficients obtained after the DWT pre-processing). Structures evaluated had three or one output neurons and one or two hidden layers with variable number of neurons. In the structure with one hidden layer and three outputs, neurons varied from 20 to 40 in steps of 10 neurons. In the structure of two hidden layers, neurons varied from 5 to 10 in the first layer and from 25 to 20 in the second layer; here, the sum of neurons in the two hidden layers was fixed as 30. Also, different combinations of transfer functions in hidden and output layer, in addition to training algorithms, were tested in order to get the best modelling network. All the combinations evaluated are summarized in Table 1.

For all these ANN architectures, different learning algorithms were used in order to select the one with better training and prediction abilities. In all calculations, as graphically summarized in Fig. 7, the best performing ANNs were feed-forward backpropagation with Bayesian regularization algorithm.

In the structure with one output neuron we trained three parallel ANNs, one per aminoacid. The structure that best performed had two hidden layers without any linear transfer function, 6 neurons in the first hidden layer, 24 in the second one and linear transfer function in output layer (network L). Other combinations listed in Table 1 did not reach at least the initially programmed RAE. Each training procedure lasted approximately 5 days to get the lowest RAE and it was considered done when no improvements in RAE were obtained after 48h of extra training. Minimum RAE values for these networks were 6.768%, 5.901% and 8.545% for Trp, Cyr and Tyr aminoacids, respectively.

Network performances were also evaluated by comparing obtained versus expected values with linear regression analysis. Fig. 8 shows this comparison for the three parallel networks with one output. Table 2 contains the values for the slope and intercept of the regression lines corresponding to these comparisons, and their corresponding uncertainties at the 95% confidence level, for training and external test subsets. In all cases, regression lines were indistinguishable from the theoretical comparison, line having 1 value for the slope.

<table>
<thead>
<tr>
<th>Network</th>
<th>Number of neurons in hidden layer 1</th>
<th>Number of neurons in hidden layer 2</th>
<th>Transfer functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>40</td>
<td>–</td>
<td>logsig–logsig</td>
</tr>
<tr>
<td>B</td>
<td>40</td>
<td>–</td>
<td>logsig–tansig</td>
</tr>
<tr>
<td>C</td>
<td>40</td>
<td>–</td>
<td>tansig–tansig</td>
</tr>
<tr>
<td>D</td>
<td>40</td>
<td>–</td>
<td>tansig–logsig</td>
</tr>
<tr>
<td>E</td>
<td>20</td>
<td>–</td>
<td>logsig–purelin</td>
</tr>
<tr>
<td>F</td>
<td>30</td>
<td>–</td>
<td>logsig–purelin</td>
</tr>
<tr>
<td>G</td>
<td>40</td>
<td>–</td>
<td>logsig–purelin</td>
</tr>
<tr>
<td>H</td>
<td>20</td>
<td>–</td>
<td>tansig–purelin</td>
</tr>
<tr>
<td>I</td>
<td>30</td>
<td>–</td>
<td>tansig–purelin</td>
</tr>
<tr>
<td>J</td>
<td>40</td>
<td>–</td>
<td>tansig–purelin</td>
</tr>
<tr>
<td>K</td>
<td>5</td>
<td>25</td>
<td>tansig–tansig–purelin</td>
</tr>
<tr>
<td>L</td>
<td>6</td>
<td>24</td>
<td>tansig–tansig–purelin</td>
</tr>
<tr>
<td>M</td>
<td>7</td>
<td>23</td>
<td>tansig–tansig–purelin</td>
</tr>
<tr>
<td>N</td>
<td>8</td>
<td>22</td>
<td>tansig–tansig–purelin</td>
</tr>
<tr>
<td>O</td>
<td>9</td>
<td>21</td>
<td>tansig–tansig–purelin</td>
</tr>
<tr>
<td>P</td>
<td>10</td>
<td>20</td>
<td>tansig–tansig–purelin</td>
</tr>
</tbody>
</table>

Networks are labelled ‘A’ to ‘P’. The last column indicates the combinations of transfer functions used in the layers: hidden1–output or hidden1–hidden2–output.
Fig. 7. Calculated RAE values for outputs Trp (A), Cys (B) and Tyr (C). Labels contained under ‘Training Functions’ correspond to the different functions available in Matlab used to train the networks, traincgb is the conjugate gradient backpropagation with Powell–Beale restarts; traincgp the conjugate gradient backpropagation with Polak–Ribiere updates; traingd the gradient descent backpropagation; traingda the gradient descent with adaptive learning rate backpropagation; traingdm the gradient descent with momentum backpropagation; traingdx the gradient descent with momentum and adaptive learning rate backpropagation; trainlm the Levenberg–Marquardt backpropagation; trainoss the one-step secant backpropagation.
and zero value for the intercept. From these results, it can be seen that the parallel ANNs permitted an excellent modelling of the voltammograms pretreated by DWT technique.

The ANN structure with three outputs was also tested, as it could simplify and reduce the data treatment. However, this structure never showed comparable results to the three parallel networks. The best performing ANN with three outputs had one hidden layer with 40 neurons with non-linear transfer functions, plus a linear function in the output layer (network J). This structure was harder to train, reaching the minimum RAE after 3 weeks of iterative training. Obtained RAE values for this network were 10.607%, 10.809% and 12.560% for Trp, Cys and Tyr amino acids, respectively (mean RAE 11.328%). Fig. 9 shows the comparison lines obtained with this network for the training and external test subsets. Table 3 contains the corresponding values of slope and intercept of the respective comparison lines, together with their uncertainties at the 95% confidence level. The table clearly shows a slightly worse modelling performance, specially for the test subset.

In order to compare the obtained results, a decimation procedure for data reduction was done as a reference. The processing of the voltammograms compacted by decimation, was accomplished using a similar model to that obtained with the best performing applied to the cA4–db4: three parallel ANNs (network L, Table 1) and Bayesian regularization.

<table>
<thead>
<tr>
<th>Aminoacid</th>
<th>Training set</th>
<th>External test set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m</td>
<td>b</td>
</tr>
<tr>
<td>Trp</td>
<td>0.999 ± 0.0022</td>
<td>1.6E–8 ± 2.6E–8</td>
</tr>
<tr>
<td>Cys</td>
<td>0.999 ± 0.0014</td>
<td>0.36E–8 ± 3.7E–8</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.998 ± 0.0006</td>
<td>3.1E–8 ± 1.2E–8</td>
</tr>
</tbody>
</table>

Values split for training and external test subsets, using the three parallel neural networks with one output each (network L). Uncertainty intervals calculated at the 95% confidence level.
Fig. 9. Comparison of the obtained vs. expected results for the three considered substances obtained using the ANN with three concentration outputs. The dashed line corresponds to ideality ($y = x$) and the solid line is the regression of the comparison data. Each substance tested occupies one column in the order tryptophan, cysteine and tyrosine. The two rows correspond to the training (top) and external test subsets (bottom).

Table 3

<table>
<thead>
<tr>
<th>Aminoacid</th>
<th>Training set</th>
<th>External test set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$m$</td>
<td>$b$</td>
</tr>
<tr>
<td>Trp</td>
<td>0.998 ± 0.0021</td>
<td>1.8E−8 ± 2.7E−8</td>
</tr>
<tr>
<td>Cys</td>
<td>0.999 ± 0.0037</td>
<td>1.7E−8 ± 3.5E−8</td>
</tr>
<tr>
<td>Tyr</td>
<td>0.994 ± 0.134</td>
<td>0.71E−8 ± 3.0E−8</td>
</tr>
</tbody>
</table>

Values split for training and external test subsets, using the neural network featuring three concentration outputs (network J). Uncertainty intervals calculated at the 95% confidence level.

5. Conclusions

We have shown in this communication how the coupling of DWT and ANNs is a successful option for the modelling of complex sensor signals, such as those of voltammetry, in order to make possible the voltammetric electronic tongue. The DWT processing of the voltammograms has permitted the reduction of the amount of information needed to represent their content in a factor of ca. 10, being the case studied of high difficulty, as the signal corresponded to the overlapped combination of three compounds plus noise and the oxidation of containing media. To estimate the degree of representation of the original data, a comparison factor $f_c$ has been proposed, which computes the similarity between the original and the reconstructed signal with better results than with the correlation coefficient. The calibration models to quantify the three oxidizable amino acids have been successfully obtained employing ANNs and the DWT that reduced the representation of the data. The same treatment can be extended to the use of a larger number of sensor signals, per example, from electrodes made of different metals, in a more typical...
voltammetric electronic tongue. Although the involved computing times are large, they are needed only for calibration, not for the application. Simpler tentatives such as direct modelling of downsampled voltammograms with ANNs did not yield acceptable models. This problem illustrates the need of treatments applying the DWT coupled with ANN technique, as the one proposed in this work.

The methodology outlined above is not uniquely applicable to voltammetric signals. A general approach can be equally applied to any spectral or dynamic signal, with information contained along the time, if the complexity of the information constrains the capability to model it directly with ANNs. Even though, the DWT treatment could be applied to other multiparametric analysis tools, such as principal component analysis (PCA) for identification purposes, or partial least squares (PLS) for calibration.

The study case reported here represents a clear simplification of the analytical procedures usually employed, as for the separate quantification of aminoacids in this study, techniques are needed like liquid chromatography with pre- or post-column labeling and spectroscopic detection of the corresponding derivatives (e.g. UV–vis spectrophotometry or fluorescence). Whereas sensors alone cannot resolve the applications, the performance accomplished with the coupled DWT–ANN chemometric treatment, a very simple and fast procedure is feasible, in order to satisfy the demand of rapid amino acid determinations in fields such as food, biochemistry and pharmaceutical and clinical analysis.

Acknowledgements

Financial support for this work was provided by the MCyT (Madrid, Spain) trough project CTQ2004-08134, by CONACYT (Mexico) trough project 43553 and by the Department of Universities and the Information Society (DURSI) from the Generalitat de Catalunya. Exchange of researchers was made through the ACI programme (DURSI, Generalitat de Catalunya). The study case reported here represents a clear simplification of the analytical procedures usually employed, as for the separate quantification of aminoacids in this study, techniques are needed like liquid chromatography with pre- or post-column labeling and spectroscopic detection of the corresponding derivatives (e.g. UV–vis spectrophotometry or fluorescence). Whereas sensors alone cannot resolve the applications, the performance accomplished with the coupled DWT–ANN chemometric treatment, a very simple and fast procedure is feasible, in order to satisfy the demand of rapid amino acid determinations in fields such as food, biochemistry and pharmaceutical and clinical analysis.

References

Historical context and overview...

New directions and challenges...

Recent developments and applications...

Future directions and outlooks...

Biographies:

Laura Moreo-Barón received the MSc degree in chemistry in 2005 from the Autonomous University of Barcelona, where she is at the moment pursuing his PhD in analytical chemistry. His main research topics are volumetric sensors, electronic tongues, and chemometric tools for chemical sensors.

Roberto Muñoz received the BS degree in electronic engineering in 1999 from the Technological Institute of Veracruz in Mexico and the MSc degree in electrical engineering from CINVESTAV-IPN in 2002. Now, he is beginning his PhD in Electronic Engineering at the Autonomous University of Barcelona, in the field of nonlinear signal processing for chemical sensors.

Arben Merkoçi was awarded the PhD in chemistry from the University of Tirana, Albania, in 1991 and then did post-doctoral researches in Greece, Hungary, Italy, Spain and USA. His main interests have been electroanalytical methods for several applications in sensors and biosensors. Currently he is “Ramón y Cajal” researcher and professor at the Sensors & Biosensor Group, Chemistry Department, Autonomous University of Barcelona. His main research interests concern the design of composites, bio composites and nanocomposite materials for enzyme, immuno and DNA based electrochemical sensors.

Salvador Alegría was awarded the PhD in chemistry in 1978 from the Autonomous University of Barcelona, and got his Full Professor position of analytical chemistry at Autonomous University of Barcelona in 1991, where he previously was an Associate Professor. He is head of the Sensors & Biosensors Group in the Chemistry Department. He is involved in the concept and development of integrated analytical systems, mainly electrochemical sensors and biosensors. The resultant sensor devices are typically applied in automated analytical systems based on biosensorization or biomimetic concepts, and used for process control and monitoring in fields like medicine, environment and chemical industry.

Manuel del Valle received the PhD in chemistry in 1992 from the Autonomous University of Barcelona, and got a position of Associate Professor in analytical chemistry in 1997 at the same university. He is member of the Sensors & Biosensors Group where he is specialist in instrumentation and electrochemical sensors. He has initiated there the research lines of sensor arrays and electronic tongues. Other interests of his work are the use of impedance measurements for sensor development and the design of automated flow systems.

Lorenzo Leija is Senior Lecturer-Researcher at the Biorecognition Section of the Department of Electrical Engineering at the CINVESTAV in Mexico DF. He is Electronic Engineer (1976) from the San Luis Potosí Technological Institute (México), MSc (1981) in Biotechnology from the CINVESTAV (México) and PhD (1989) in “Genie Biologique” from the Nancy I University (France). His main research line follows the development of instrumentation for cardiac pacemaker and the use of RF on health.

Pablo Rogelio Hernandez received a BSc degree in Electronics in 1978 from the Instituto Politécnico Nacional. He was awarded the PhD degree from CINVESTAV-IPN, Mexico, in 1995. He is a Lecturer-Researcher at the CINVESTAV-IPN at Mexico city on the design of chemical sensors and the instrumentation for electrochemical measurements.

Roberto Munoz received the BSc degree in electronics engineering from the National Polytechnic Institute of Mexico in 1983, the MSc in bio electronics in 1986 and PhD in Electrical Engineering in 1999, both at the Center for Research and Advanced Studies of IPN (CINVESTAV). Since 1990, he has been a Lecturer-Researcher at the Department of Electrical Engineering of CINVESTAV on nonlinear biomedical signal processing and has made recent contributions in the field of potentiometric and volumetric electronic tongues.