ESANN'2001 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 347-352

A two steeps method: Non Linear Regression and Pruning Neural Network for Analyzing Multicomponent Mixtures

César Hervás Martínez*, José A. Martínez Heras*,
Sebastián Ventura Soto*, Manuel Silva Rodríguez**

* Department of Computer Science, University of Córdoba.
C2 Building, Campus of Rabanales. E-14071 Córdoba, Spain

** Department of Analytical Chemistry, University of Córdoba.
C3 Building, Campus of Rabanales. E-14071 Córdoba, Spain

Abstract. This work deals with the use of pruning ANNs in conjunction with genetic algorithms for resolving nonlinear multicomponent systems based on oscillating chemical reactions. The singular analytical response provides by this chemical system after its perturbation was fitted to a gaussian curve by least-square regression and the estimates were used as inputs to the ANNs. The proposed methodology was validated by the simultaneous determination of pyrogallol and gallic acid (two strong related phenol derivatives) in mixtures on the basis of their perturbation effects on the classical Belousov-Zhabotinskii reaction. The trained network estimates concentrations of pyrogallol and gallic acid with a standard error of prediction for the testing set of ca. 4% and 5.7% respectively or 4.4%, 9% for different sets of train/test patterns. This result is much smaller than those provided by a classical parametric method such as non-linear regression.

1 Introduction

The differential kinetic methodology is an effective way for resolving mixtures of closely related chemical species with no prior physical separation. Methods for kinetic multicomponent determinations based on different chemometric tools have been proliferated in the last few years [1]. These methods do not require the prior knowledge of the reaction rates involved in the analytical system and they eliminate or reduce synergistic effects as well as other unknown sources of non-linearity. However, the nonlinear chemical phenomena known

The authors wish to thank Spain's (DGICyT) for funding this research within the framework of Project ALI98-0676-CO2-02 and Project INTAS grant INTAS-OPEN-97-1094.

ESANN'2001 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 347-352

as oscillating chemical reactions, which exhibit several non-monotonic regimes such as regular oscillations, periodic doubling, quasi-periodicity and deterministic chaos, among others, has been the focus of much research in the area of chemical kinetic in the last few years. On account of the great degree of non-linearity showed by these system, only powerful multivariate calibration techniques can offer the suitable accuracy for the resolution of these mixtures. We choose to use ANNs in this work considering the suitability of their features to the proposed chemical problem.

ANNs based on different versions of standard back-propagation (BP) learning algorithm have been used as highly powerful tools to solve a great variety of problems in analytical chemistry [2, 3]. Recently, a current issue in this field is the design of ANNs with a minimum size to solve real-world problems. One way for reducing the chance that a fully connected complex model is formed during the training process is to use regularization to constrain or eliminate the network weights. A regularization function used is the sum of squares of the weight magnitudes, but recently another regularization function are proposed [4, 5].

This work deals with the use of pruning ANNs in conjunction with genetic algorithms for resolving nonlinear multicomponent systems based on oscillating chemical reactions. The singular analytical response provides by this chemical system after its perturbation was fitted to a gaussian curve by least-square regression and the estimates were used as inputs to the ANNs. Several pruning neural network models were tested and compared and from the proposed model the subsequent equations were derived, which allow the direct determination of the concentration of the components in the mixture in a hardware implementation. The proposed methodology was validated by the simultaneous determination of pyrogallol and gallic acid (two strong related phenol derivatives) in mixtures on the basis of their perturbation effects on the classical Belousov-Zhabotinskii reaction, the most widely known and studied oscillating chemical system.

2 Basis of the method

Our approach to the resolution of mixtures of species based on their perturbation effects on oscillating chemical reactions involves the development of a two steps procedure in order to obtain nonlinear models for predicting the concentration of the components in such mixtures. The approach was tested on the simultaneous determination of two related phenol derivatives, such as pyrogallol (P) and gallic acid (GA), through their perturbation effects on the classical oscillating chemical reaction that involves the oxidation of malonic acid by bromate ion in a sulfuric acid medium catalyzed by cerium(IV) salts, which is known as the Belousov-Zhobotinskii reaction. This reaction exhibits periodic changes in the concentration of some species that reflect in potential changes or cyclic color, which involves color changes between yellow [cerium(IV)] and colorless [cerium(III)] (see regular oscillations in Figure 1).

ESANN'2001 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 347-352

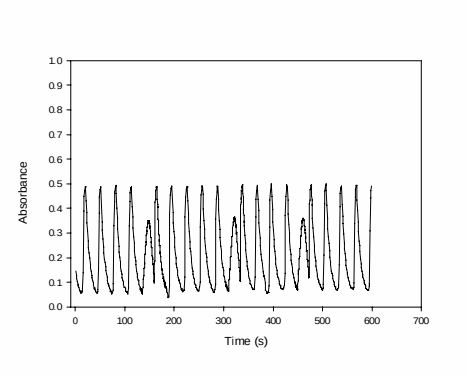

Figure 1: Analytical response obtained when the system is perturbed by injecting a microvolume of the mixture sample

The first step of our approach for extracting the information in order to select the inputs to the ANNs is based on the kind of the signal provided by oscillating reaction after its perturbation. Thus, by inspection of this global response, we have observed that the signals set (t_i, S_{t_i}) was accurately fitted by least-square regression to a gaussian curve if the time domain used ranged from $t_i = 0$ (injection time of the mixture in the oscillating system) to $t_i = t_m + 0.4t_m$, where t_m is a centralization parameter associated with the time corresponding to the maximum of the response curve. In this case, S_{t_i} is the response variable, which is proportional to the concentration of the components in the mixture, and t_i the independent variable. It assumed that the change of S_{t_i} with time (reaction rate) is proportional to its value at such time, the proper time t_i and a parameter, s, associated with the dispersion of the S_{t_i} values with respect to that corresponding to t_m. From these, it follows that the following differential equation can be obtained once the variable t_i was changed for $t_i - t_m$:

$$\frac{\partial S_{t_i}}{\partial (t_i - t_m)} = -\frac{1}{s^2}(t_i - t_m)S_{t_i} \tag{1}$$

The integration of this equation considering that S_{t_i} at time t_m is given by a_m (the maximum of the response curve) corresponds to a three parameters gaussian. If additive errors are assumed, the nonlinear model is given by

$$S_{t_i} = a_m e^{-\frac{1}{2}\frac{(t_i - t_m)^2}{s^2}} + \varepsilon_i \tag{2}$$

The least squares principle was used to estimate the parameters by Levenberg-Marquardt method, The estimates obtained, $\hat{a}_m$, $\hat{s}$ and $\hat{t}_m$ are used as inputs to the ANN, in such way that all proposed neural network models have three variables in the input layer.

A second step involves a network minimization design by successively removing weights after the network has been trained to satisfactory performance.

ESANN'2001 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 347-352

This learning process is carried out by using the back-propagation learning procedure EDBD [6], where the pruning algorithm is independent of the particular training procedure. We used a variant of the algorithm proposed in [3]. In that algorithm we deals with the design of regularization-pruning ANNs in conjunction with genetic algorithms where our approach to network minimization involves successively removing weights after the network has been trained to satisfactory performance. We achieve this design objective by creating a fully connected multilayer neural network with a number of weigths big enough for our nonlinear regression model, then prune it by eliminating certain synaptic weights using a complexity regularization procedure where the complexity penalty function defined by [5]:

$$\lambda n_w log \sum_{k=1}^{n_w} |w_k| \tag{3}$$

represents the complexity of the network as a function of the absolute value of the weight magnitudes, w_k, where n_W is the number of weights. The evaluated function is based on the hypothesis that the "a priori" distribution of network weights follows a Laplacian distribution. The λ parameter is a regularization term, which represents the relative importance of the complexity-penalty term with respect to the performance-measure term, defined as:

$$f(w) = \sum_{p=1}^{n_v} (y_p - o_p)^2 + \lambda_{WE} \sum_{k=1}^{W} |w_k| \tag{4}$$

This term is the sum of the squared errors between the actual output values (o_p) and the target output values (y_p), and n_T is the number of patterns for the training set. Sigmoid and linear functions were used for hidden nodes and output nodes, respectively; in addition, in order to avoid saturation problems when sigmoid functions are used, the values of the inputs and outputs nodes were normalized over the range from 0.1 to 0.9. Thus, the normalized input values, $\tilde{a}_m^*$, $\tilde{s}^*$ and $\tilde{t}_m^*$, and those for the output nodes corresponding to the concentration of pyrogallol, $[P]^*$, and gallic acid, $[GA]^*$.

3 Experimental section

The Levenberg-Marquardt algorithm was used in order to obtain the three estimated coefficients. The convergence of the iterative process was achieved with a tolerance of 0.0001 and a maximum number of iterations of 100. The algorithm software for ANN [3], written in C language, was run on an IRIS Release 6.5 in an Origin 2000.

Overall 27 synthetic samples (by triplicate) containing uniformly distribution concentrations of the analytes (pyrogallol and gallic acid) were prepared as described below. We used randomly two replicates for the training set and the other for the generalization set. The performance of the algorithm was

ESANN'2001 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 347-352

network topology	data set train/test	pyrogallol					
		mean			SD		
		SEP train	SEP test	Connec.	SEP train	SEP test	Connec.
3:3:1	44/22	5.07	4.83	11.5	0.60	0.34	2.83
3:2:1	54/27	5.86	5.64	9.0	0.46	0.44	1.41
3:4:2	44/22	4.73	4.39	21.7	0.27	0.41	3.19
		best			worst		
		SEP train	SEP test	Connec.	SEP train	SEP test	Connec.
3:3:1	44/22	3.79	4.19	14	5.67	5.19	14
3:2:1	54/27	4.76	4.65	8	6.22	6.12	9
3:4:2	44/22	5.22	4.08	19	4.83	5.27	22
3:3:2	54/27	5.57	4.40	18	5.97	5.17	19
		gallic acid					
		mean			SD		
		SEP train	SEP test	Connec.	SEP train	SEP test	Connec.
3:2:1	44/22	5.44	5.66	9.7	0.36	0.17	0.67
3:3:1	54/27	10.80	10.01	11.1	0.44	0.85	1.91
3:4:2	44/22	5.13	6.01	21.7	0.29	0.30	3.19
3:3:2	54/27	9.78	9.52	18.2	0.21	0.38	0.91
		best			worst		
		SEP train	SEP test	Connec.	SEP train	SEP test	Connec.
3:2:1	44/22	5.50	5.38	9	5.89	6.01	10
3:3:1	54/27	10.01	8.51	15	11.37	11.07	10
3:4:2	44/22	4.92	5.70	25	5.21	6.65	24
3:3:2	54/27	9.67	8.99	16	10.09	10.24	17

Table 1: Accuracy of the Algorithm Used with Various Network Topologies and Data Set Sizes as Applied to the Resolution of Mixtures of Pyrogallol and Gallic Acid

tested with various network topologies, which were running ten times in each case. The accuracy for each model were evaluated from the results obtained for both data sets by using the relative standard error of prediction (SEP)

$$SEP = \frac{100}{\overline{A}_i}\sqrt{\frac{\sum_{i=1}^{n}(A_i - \tilde{A}_i)^2}{n}} \tag{5}$$

where $\tilde{A}_i$ and A_i are the found an expected values for the analyte concentration in the mixture, $\bar{A}_i$ is its average value, and n is the number of patterns used. Analysis of variance (ANOVA) and the Student-Newman-Keuls (SNK) test were used to evaluate statistically the performance of the different models in order to propose a network topology and data set sizes for the training and testing sets. As can be seen from Table 1 the models with are better SEP test and a less number of connections were a 3:4:2 network topology and a 44/22 data set train/test, and a 3:3:2 network topology which a 54/27 data set train/test. The equations of the last proposed model can be seen from Table 2.

ESANN'2001 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 25-27 April 2001, D-Facto public., ISBN 2-930307-01-3, pp. 347-352

Model 6 (Topology: 3:3:2; Connections: 18; Data Set: 54 train/27 test)
$[P]^* = 1.54h_1 + 0.37h_2 - 0.78h_3$
$[GA]^* = -0.69 - 1.74h_1 + 1.89h_2 + 1.22h_3$
$h_1 = (1 + exp(3.69 - 0.07a_m - 3.54s))^{-1}$
$h_2 = (1 + exp(0.74 - 3.71a_m + 1.5s - 0.38t_m))^{-1}$
$h_3 = (1 + exp(0.8 - 1.93a_m + 7.05s + 2.3t_m))^{-1}$

Table 2: Derived Equations From the Best ANN Model Obtained

4 Conclusion

That regularization pruning is an effective way of reducing network complexity with a network topology where the estimates by NLR are included in the input layer and three or four nodes are used in the hidden layer. The use of estimated parameters of NLR as input, provide the information required by the ANNs to discriminate among several kinetic curves for differents $[P]^*$ and $[GA]^*$ values. In summary, pruning ANNs have been shown to posses a high potential for kinetic analysis, and, in general, as an analytical tool for deriving quality information with substantial savings in time and in experimental and computational costs.

References

[1] Toledo,R.; Silva, M.; Khavrus, W.O.; Strizhak, P.E. Potential of the analyte pulse perturbation technique for the determination of polyphenols based on the Belousov-Zhabotinskii reaction, *Analyst*, **2000**, *125*, 2118-2124.

[2] Zupan J.; Gasteiger J. *Neural networks for chemists. An introduction*; VCH: Weinheim, 1993.

[3] Hervás, C.; Algar, J.A.; Silva M.; Correction of temperature variations in kinetic-based determinations by use of pruning computational neural networks in conjunction with genetic algorithms; *J. Chem. Inf. Comput. Sci.*, **2000**, *40*, 724-731.

[4] Weigend, A.; Rumelhart, D.; Huberman, B.; Generalization by weight elimination with application to forecasting; *Adv. Neural Inform. Process. Syst.*, **1991**, *3*, 875-882.

[5] Williams, P. M.; Bayesian regularization and pruning using a Laplace prior; *Neural Comput.*, **1995**, *7*, 117-143.

[6] Minai, A. A.M; Williams, R. J.; Back-propagation heuristics: A study of the extended delta-bar-delta; *IEEE International Joint Conference on Neural Networks*, San Diego, CA, 1990, pp 595-600.