Use of Pruned Computational Neural Networks for Processing the Response of Oscillating Chemical Reactions with a View to Analyzing Nonlinear Multicomponent Mixtures

César Hervás*

Department of Computer Science, Edificio C2, Campus of Rabanales, University of Córdoba, E-14071 Córdoba, Spain

Rocío Toledo and Manuel Silva

Analytical Chemistry Division, Department of Analytical Chemistry and Ecology, Edificio C3-Anexo, Campus of Rabanales, University of Córdoba, E-14071 Córdoba, Spain

Received February 16, 2001

The suitability of pruned computational neural networks (CNNs) for resolving nonlinear multicomponent systems involving synergistic effects by use of oscillating chemical reaction-based methods implemented using the analyte pulse perturbation technique is demonstrated. The CNN input data used for this purpose are estimates provided by the Levenberg−Marquardt method in the form of a three-parameter Gaussian curve associated with the singular profile obtained when the oscillating system is perturbed by an analyte mixture. The performance of the proposed method was assessed by applying it to the resolution of mixtures of pyrogallol and gallic acid based on their perturbating effect on a classical oscillating chemical system, viz. the Belousov−Zhabotinskyi reaction. A straightforward network topology (3:3:2, with 18 connections after pruning) allowed the resolution of mixtures of the two analytes in concentration ratios from 1:7 to 6:2 with a standard error of prediction for the testing set of 4.01 and 8.98% for pyrogallol and gallic acid, respectively. The reduced dimensions of the selected CNN architecture allowed a mathematical transformation of the input vector into the output one that can be easily implemented via software. Finally, the suitability of response surface analysis as an alternative to CNNs was also tested. The results were poor (relative errors were high), which confirms that properly selected pruned CNNs are effective tools for solving the analytical problem addressed in this work.

INTRODUCTION

Differential kinetic methodology is an effective choice for resolving mixtures of closely related chemical species with no prior physical separation. Methods for kinetic multicomponent determinations based on various chemometric tools have proliferated in the past few years.[1] Thus, linear and extended Kalman filtering[2−5] and, recently, multivariate calibration techniques based on factor analysis and artificial intelligence such as principal component regression (PCR),[6,7] partial least squares (PLS),[8−11] and CNNs[12−18] have found increasing application in this field. These methods require no prior knowledge of the reaction rates for the species in the analytical system; also they avoid or reduce synergistic effects and other unknown sources of nonlinearity. These differential kinetic approaches have so far been used exclusively in connection with analytical systems exhibiting monotonic regimes—the most widely ones used in kinetic methods for chemical analysis. However, the nonlinear chemical phenomena known as "oscillating chemical reactions", which exhibit various nonmonotonic regimes such as regular oscillations, periodic doubling, quasi-periodicity, and deterministic chaos, among others, have been the subjects of much research in chemical kinetics in the past few years.[19,20] To our knowledge, only one recent paper about the resolution of mixtures based on an oscillating chemical reaction appears to exist.[21] A special mathematical procedure based on the singular kinetic behavior of the mixture components was used to resolve mixtures in which only one component was present in excess amounts. This methodology is very restrictive, so further improvement is required if multicomponent systems are to be effectively resolved by using oscillating chemical reactions. Because of its high nonlinearity, this type of mixture can be accurately resolved only by using powerful multivariate calibration techniques. We chose to use CNNs for this purpose taking into account their suitability to the chemical problem addressed.

CNNs based on various versions of the standard back-propagation (BP) learning algorithm have proved highly powerful tools for solving a wide variety of problems in analytical chemistry.[22−26] One recent trend in this field is the development of CNNs of minimal size to solve real-world problems, i.e. networks with a small number of weights or target parameters but possessing a reasonably high generalization capacity. A small-size neural network will be less prone to overtraining noise or the structure of the data in the training set, which will increase its generalization capacity over a new set data. One way of reducing the chance of a fully connected complex model forming during the training process is to use regularization in order to restrict

* Corresponding author phone: +34-957-218349; e-mail: ma1hemac@uco.es.

10.1021/ci010012j CCC: $20.00
Published on Web 07/06/2001

or suppress network weights. One regularization function used for this purpose is the sum of squares of the weights; alternative regularization functions have recently been proposed.[27−29] The main disadvantage of regularization is the difficulty in selecting the appropriate magnitude in a regularization term for a given problem. One method used to set the regularization term is to train a given CNN a number of times, using a different regularization term in each training run;[29] this is computationally very expensive, however. Weight suppression is necessary to construct models with a small number of parameters in order to avoid using too many training patterns and hence reduce experimental costs. In this work, we used the network learning process as a statistical procedure for estimating a regression function. The accuracy thus achieved depends on the particular hypothesis on which the Fisher information matrix for a multilayer perceptron (MLP) model is nonsingular because it plays a prominent role in the asymptotic behavior of the least squares error estimator.[30,31] However, the Fisher information matrix for a MLP is singular if the network has redundant hidden units. With a view to solving this problem, we used a genetic algorithm for suppression of learning weights and removal of input and hidden units—essential inputs and hidden units should be preserved because excessive removal increases model bias.

This work deals with the use of pruned CNNs in conjunction with genetic algorithms for resolving nonlinear multicomponent systems based on oscillating chemical reactions. The singular analytical response provided by this type of chemical system upon perturbation was fitted to a Gaussian curve using least-squares regression and the estimates thus obtained were used as inputs to the CNNs. Various pruned neural network models were tested and compared in order to adopt the most suitable one, which was used to derive easily solved equations for the mathematical transformation of the input vector (estimated from the Guassian regression) into the output vector (concentrations of the components in the mixture). The performance of response surface analysis based on the estimates provided by the Gaussian fitting was also tested and compared with the results yielded by the pruned CNNs. The proposed methodology was validated with the simultaneous determination of pyrogallol and gallic acid (two closely related phenol derivatives) in mixtures by their perturbing effects on the classical Belousov−Zhabotinskyi reaction (the most widely known and studied oscillating chemical system[19,20]), which was implemented in a continuous-flow stirred-tank reactor (CSTR), using the recently reported analyte pulse perturbation technique.[20]

THEORETICAL BACKGROUND

The proposed approach to the resolution of mixtures of species based on their perturbing effect on oscillating chemical reactions uses a two-step procedure to construct nonlinear models for predicting the concentration of each mixture component. In the first step, information is extracted from the analytical response with a view to selecting the inputs to the CNNs. The signal set (t_i, S_{t_i}) can be accurately fitted to a Gaussian curve by using least-squares regression; S_{t_i} is the response variable, which is proportional to the concentration of the components in the mixture, and t_i is the independent variable. If the change in S_{t_i} with time (reaction rate) is assumed to be proportional to its value at such time, time t_i and a parameter s associated with the dispersion of the S_{t_i} values with respect to that corresponding to the time of the peak of the response curve, t_m, then the following differential equation is obtained following substitution of t_i by $t_i - t_m$:

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

Because S_{t_i} at time t_m coincides with a_m (the maximum of the response curve), integration of this equation yields

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

which corresponds to a three-parameter Gaussian curve. If additive errors ϵ_i are assumed to exits, then the nonlinear model will be represented by

$$S_{t_i} = a_m e^{-1/2\ (t_i - t_m)^2/s^2} + \epsilon_i \tag{3}$$

The parameters in this nonlinear model were estimated by using the least squares principle. The least-squared estimates of a_m, s, and t_m, labeled $\hat{a}_m$, $\hat{s}$, and $\hat{t}_m$, respectively, were those that minimized the sum of squared residuals:

$$SS(Res) = \sum_{i=1}^{n} [S_{t_i} - a_m\, e^{-1/2\ (t_i - t_m)^2/s^2}]^2 \tag{4}$$

These minima were obtained by calculating the partial derivatives of *SS(Res)* with respect to each parameter and equalling them to zero in order to derive the three normal equations:

$$\frac{\partial SS(Res)}{\partial a_m} = \sum_{i=1}^{n} (S_{t_i} - a_m\, e^{-1/2\ (t_i - t_m)^2/s^2})(e^{-1/2\ (t_i - t_m)^2/s^2}) = 0 \tag{5}$$

$$\frac{\partial SS(Res)}{\partial s^2} = \sum_{i=1}^{n} (S_{t_i} - a_m e^{-1/2\ (t_i - t_m)^2/s^2}) a_m \frac{1}{2s^4} (t_i - t_m)^2 (e^{-1/2\ (t_i - t_m)^2/s^2}) = 0 \tag{6}$$

$$\frac{\partial SS(Res)}{\partial t_m} = \sum_{i=1}^{n} (S_{t_i} - a_m e^{-1/2\ (t_i - t_m)^2/s^2}) a_m \frac{1}{s^2} (t_i - t_m) e^{-1/2\ (t_i - t_m)^2/s^2} = 0 \tag{7}$$

There are no explicit solutions to these equations, which thus call for iterative numerical. We used a Levenberg−Marquardt method,[32] which tends toward the Gauss−Newton fitting[33] if the residual sum of the squares is reduced at each step, or toward the steepest descent fitting[34] if the residual sum of squares increases in any step. The estimates obtained, $\hat{a}_m$, $\hat{s}$, and $\hat{t}_m$, were used as inputs to the CNN, so all the neural network models proposed in this work had three variables in the input layer.

In the second step, a procedure was used to construct various models based on the learning of the CNNs from the patterns of the training set by using a variant of an algorithm

recently developed by ourselves.[35] The algorithm is used to design regularization-pruning CNNs in conjunction with genetic algorithms; our approach to network minimization involves successively removing weights after the network has been trained to satisfactory performance.

At every iteration, a new population is formed using genetic operators of reproduction, crossover, and mutation on the old population. Each member of the population is decoded as a feedforward network, and then its aptitude is obtained. The roulette selection method is used with a fitness linear scaling function to prevent premature convergence, which is obtained using a stop criteria based in the average improvement of the average fitness or a *maxgen* parameter. Together with this selection method two heuristics were raised: the elitism strategy, since a specific group of the best individuals is maintained in the following population; and the gap, G_p, of the generation, through which the percentage of the population which will not undergo any change is controlled. Thus, the $N_p(1-G_p)$ population elements are selected, where N_p is the population size, so that they remain intact in the following generation. The crossover procedure, involving choosing both network elements to be crossed with P_{cross} probability, is performed by applying an exchange of the nodes of the first layer. This method is a particular case of parametrized uniform crossover in which an exchange takes place with a probability P_{recom} at each bit. Finally, a mutation operator was applied on the networks, which permitted a random search. The mutation was produced for each individual with a P_{mut} probability. If there is a mutation, this occurs by the addition to each group of weights of a random amount in the form

$$w_j' = w_j + 2(X - 0.5)R_{mut} \quad j = 1,..., n_w \tag{8}$$

where w_j is the jth weight, X a uniform random variable in the interval [0,1], R_{mut} a parameter, and n_w the number of weights.

We accomplish this design goal by developing a thoroughly connected multilayer neural network with a number large enough of nodes in the hidden layer and then prune it by removing certain synaptic weights using a complexity regularization procedure[36] where the complexity penalty function, E_1, defined by[28]

$$E_1 = \lambda n_w \log \sum_{k=1}^{n_w} |w_k| \tag{9}$$

represents the complexity of the network as a function of the absolute value of the weight magnitudes, w_k. The evaluated function is based on the hypothesis that the a priori distribution of network weights follows a Laplacian distribution. Parameter λ is a regularization term that represents the relative significance of the complexity-penalty term with respect to the performance-measure term, E_2, defined as

$$E_2 = \frac{1}{2} n_T \log \sum_{p=1}^{n_T} (y_p - o_p)^2 \tag{10}$$

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.

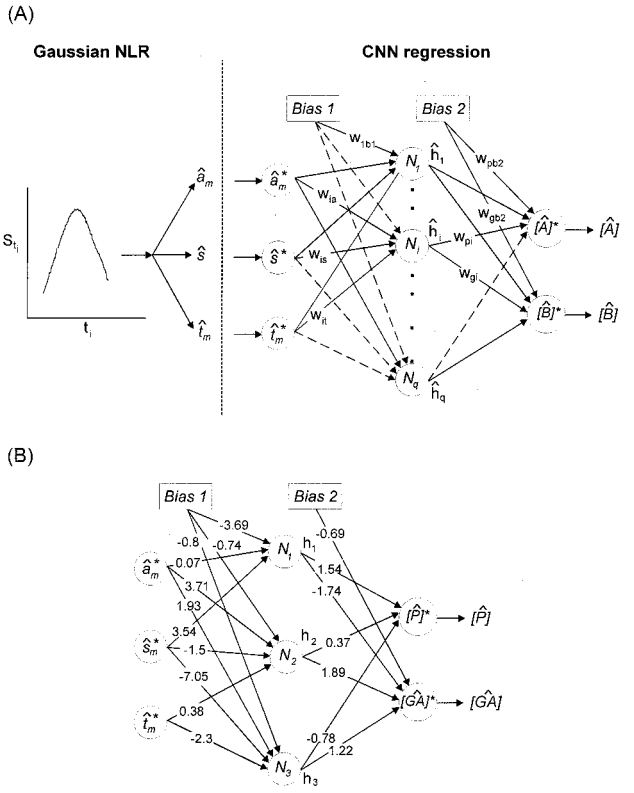

Figure 1. (A) Schematic diagram of the foundation and topology of the pruned CNN used to resolve nonlinear multicomponent systems by oscillating chemical reaction-based methods. Dashed lines correspond to the removed connections after pruning. (B) Proposed pruned CNN architecture used to estimate the concentration of pyrogallol (*P*) and gallic acid (*GA*) in the studied mixtures.

Sigmoidal and linear functions were used for hidden nodes and output nodes, respectively; to avoid saturation problems with the sigmoidal functions, the values of the input and output nodes were scaled over the range from 0.1 to 0.9. Thus, the scaled input values, $\widehat{a_m^*}$, $\widehat{s^*}$, and $\widehat{t_m^*}$ those for the output nodes corresponding to the concentrations of the analytes in the mixture, *[A]** and *[B]**, were calculated by using the following general expression:

$$X^* = 0.8 \frac{X - X_{min}}{X_{max} - X_{min}} + 0.1 \tag{11}$$

Once the network architecture was optimized by the pruning algorithm (see Figure 1A) a set of equations was obtained by using the hidden layer and input layer weights provided by it. Such an equation can be written as follows for the ith output hidden node:

$$\hat{h}_i = \frac{1}{1 + exp(w_{ib1} + w_{ia}\hat{a}_m^* + w_{is}\hat{s}^* + w_{it}\hat{t}_m^*)} \tag{12}$$

The scaled values of the concentrations of the analyte *A*, and similarly for *B*, in the mixture can be determined by using the following equation which is a linear combinations of the h_i functions:

$$[\hat{A}]^* = w_{pb2} + w_{p1}\hat{h}_1 + \cdots + w_{pi}\hat{h}_{pi} + \cdots + w_{pm}\hat{h}_{pm} \tag{13}$$

Table 1. Parameter Values Used by the Algorithm

algorithm	parameters
training EDBD[38]	$\theta = 0.7$, $\kappa_\eta = 0.095$, $\gamma_\eta = 0.1$, $\phi_\eta = 0.1$, $\kappa_\mu = 1.0$, $\gamma_\mu = 0.05$, $\phi_\mu = 0.01$
pruning Williams[28]	ghiw = 0.9, ghider = 1.0
genetic	$\gamma_{frozen} = 0.9$, $G_P = 0.9$, $P_{cross} = 0.8$, $P_{recomb} = 0.5$, $P_{mut} = 0.1$, $R_{mut} = 0.1$, maxgen = 200, $\lambda = 0.75$, $N_P = 100$

Finally, the analytical concentration is calculated from their scaled value, using the following equation:

$$[\hat{A}] = \frac{([A]_{max} - [A]_{min})([\hat{A}]^* - 0.1)}{0.8} + [A]min \quad (14)$$

EXPERIMENTAL SECTION

We used the Levenberg−Marquardt algorithm to solve eqs 5−7 and obtained the three estimated coefficients. 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 CNN,[35] in C language, was run on IRIS Release 6.5 on an Origin 2000. Table 1 shows the parameter values used to ensure optimal application of the algorithm. In this case, the value of the γ_{frozen} parameter is the most influential over the algorithm's performance [as indicated by the lower standard error of prediction (SEP) number of connections and number of generations achieved until convergence, for a network topology]; this parameter is the percentage of frozen connections in the networks of the initial population, *i.e.*, the individuals of the initial population have random weights and a γ_{frozen} percent of them are zero.[35]

An overall 27 synthetic samples (in triplicate) containing uniformly distributed concentrations of the analytes (pyrogallol and gallic acid) were prepared as described below. Of the three replicates, two (randomly chosen) were used to design the training set and the other the generalization test. The performance of the algorithm was tested with various network topologies that were run 10 times. The accuracy of each model was assessed in terms of the SEP for the results obtained for both data sets

$$SEP = \frac{100}{\bar{C}_i}\sqrt{\frac{\sum_{i=1}^{n}(C_i - \hat{C}_i)^2}{n}} \quad (15)$$

where C_i and $\hat{C}_i$ are the found expected values for the analyte concentration in the mixture, $\bar{C}_i$ is the average value, and n is the number of patterns used. Analysis of variance (ANOVA) and the Student-Newman-Keuls (SNK) test were used for evaluation of the performance of the different models in order to adopt a network topology for each data set size for the training and testing sets used.

Experimental data were obtained by spectrophotometrically monitoring the response provided by the perturbation of pyrogallol and gallic acid on the Belousov−Zhabotinskyi reaction, using the following procedure: in a continuous-flow stirred-tank reactor thermostated at 35 °C were placed 9.0 mL of 0.45 M malonic acid, 10.0 mL of 0.2 M potassium bromate, 1.0 mL of 6.5×10^{-3} M ammonium cerium(IV) nitrate (all in 1.0 M sulfuric acid), and bidistilled water up to a final volume of 30 mL, and the mixture was homogenized by magnetic stirring. Then, a photometric probe was inserted, and, without delay, a peristaltic pump was started to supply the reactants—the overall feed stream was 2.1×10^{-3} M in cerium(IV), 0.15 M in malonic acid, and 6.6×10^{-2} M in potassium bromate and delivered at a constant flow rate of 1.5 mL/min. Once the steady state (constant oscillating period and amplitude) was reached, the system was perturbed by injecting microvolumes of mixtures containing variable amounts of pyrogallol (1.0−6.0 μmol) and gallic acid (1.0−7.0 μmol). Absorbance changes at 400 nm in the oscillations following perturbation were acquired at a rate of 1.0 point/s, using an instrumental setup consisting of (1) a photometric probe connected to a Metrohm 662 spectrophotometer; (2) a Gilson Minipuls-3 four-channel peristaltic pump to keep oscillations under steady-state conditions (three channels were used to deliver the reactant an the fourth to keep the volume of the reaction mixture constant); (3) a Metrohm Dosimat 665 autoburet to inject microvolumes of the mixture; and (4) a Pentium 100-MHz PC-compatible computer equipped with a PC-Multilab PCL-812PG 12-bit analog-to-digital converter.

RESULTS AND DISCUSSION

The growing interest of analytical chemists in nonlinear oscillating chemical reactions has been aroused, among others, by the advent of the analyte pulse perturbation (APP) technique,[20] which provides a rapid, simple method for performing many determinations on the same oscillating system tanks to the fact that the steady state is rapidly regained after each perturbation with a microvolume of the analyte (see Figure 2). The analytical goal of this work was to develop a general methodology for the resolution of mixtures of species based on their perturbation on oscillating chemical reactions by using pruned CNNs to deal with the high nonlinearity of these chemical systems. The approach was tested on the simultaneous determination of two related phenol derivatives, viz. pyrogallol (*P*) and gallic acid (*GA*); their perturbation on the classical oscillating chemical reaction involving the oxidation of malonic acid by bromate ion in a sulfuric acid medium, which is catalyzed by cerium(IV) salts and known as the Belousov−Zhabotinskyi reaction,[19,20] was used for this purpose. This reaction exhibits periodic changes in the concentration of some species that reflect in potential changes or cyclic color, the latter being directly related to the catalytic cycle, which involves color changes between yellow [cerium(IV)] and colorless [cerium(III)] (see regular oscillations in Figure 2). The analytical response obtained when this nonlinear system is perturbed by injecting a microvolume of the mixture (inset in Figure 2) is the combination of the signals provided mainly by the following chemical processes:

(a) The oscillating reaction, which consists of six reactions for five kinetically independent species and involves the autocatalytic compound $HBrO_2$, the control species Br^-, catalyst Ce^{4+}, HOBr, and $BrO_2^\bullet$. The conversion rates of the

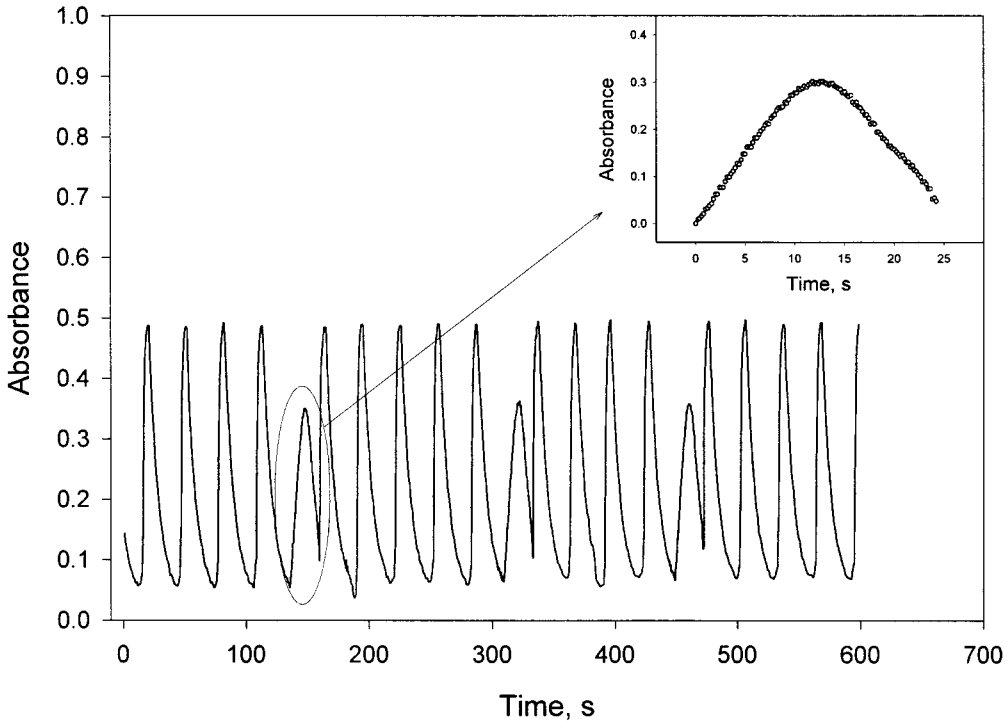

Figure 2. Typical profiles for the proposed oscillating reaction in the presence and absence of pyrogallol/gallic acid perturbation. The inset illustrates the kind of the response obtained upon injection of the mixture into the oscillating system.

chemical species in the CSTR are described by the following differential equations[37]

$$\frac{d[HBrO_2]}{dt} = f_1([HBrO_2], [Br^-], [BrO_2^\bullet], [HOBr], [Ce^{4+}]) - k_0[HBrO_2] \quad (I)$$

$$\frac{d[Br^-]}{dt} = f_2([HBrO_2], [Br^-], [HOBr], [Ce^{4+}]) + k_0([Br^-]_0 - [Br^-]) \quad (II)$$

$$\frac{d[BrO_2^\bullet]}{dt} = f_3([HBrO_2], [BrO_2^\bullet], [Ce^{4+}]) - k_0[BrO_2^\bullet] \quad (III)$$

$$\frac{d[HOBr]}{dt} = f_4([HBrO_2], [Br^-], [HOBr], [Ce^{4+}]) - k_0[HOBr] \quad (IV)$$

$$\frac{d[Ce^{4+}]}{dt} = f_5([HBrO_2], [BrO_2^\bullet], [HOBr], [Ce^{4+}]) + k_0([Ce]_{tot} - [Ce^{4+}]) \quad (V)$$

where f_i is a reaction-rate term and k_0 a flow one (flow rate). In addition, $[Br^-]_0$ and $[Ce]_{tot}$ provide the inflow concentrations of bromide and cerium(IV) ions, respectively, in the CSTR.

(b) The oxidation of phenol derivatives by oxybromine compounds according to the following kinetic scheme, which in the case of the gallic acid can be written as[37]

$$HOBr + GA \rightarrow Br^- + Q + H^+ + H_2O \qquad k = 2.0\ M^{-3}\ s^{-1}[H^+]^2 \quad (VI)$$

$$BrO_3^- + GA + H^+ \rightarrow HBrO_2 + Q + H_2O \qquad k = 2.0 \times 10^{-6}\ M^{-2}\ s^{-1}[H^+] \quad (VII)$$

$$BrO_2^\bullet + Q \rightarrow HBrO_2 + Q^\bullet \quad k = 25\ M^{-2}\ s^{-1}[H^+] \quad (VIII)$$

$$2Q^\bullet + H_2O \rightarrow Q + \text{products} \qquad k = 6.2 \times 10^4\ M^{-2}\ s^{-1}[H^+] \quad (IX)$$

where Q is the corresponding quinone and $Q^\bullet$ its radical form.

(c) The potential presence of synergistic effects, which is the main source of error in kinetic analysis. As shown below, this phenomenon has been experimentally confirmed in the resolution of the studied mixtures.

The relative contribution of these chemical processes to the analytical signal depends on the reaction time and on the relative concentration of both components in the mixture, the latter being associated with synergistic effects. Thus, at reaction times smaller than that corresponding to the maximum of the response curve, t_m, the oxidation of the phenol derivatives exhibits a major contribution to this response; at longer times (after the maximum is reached) the oscillating chemical reaction contributes more markedly. The maximum in the response curve was thus obtained as a compromise between both contributions.

Figure 3 shows the responses extracted from the overall oscillating signals provided by the APP technique, which were obtained by perturbating the oscillating system with

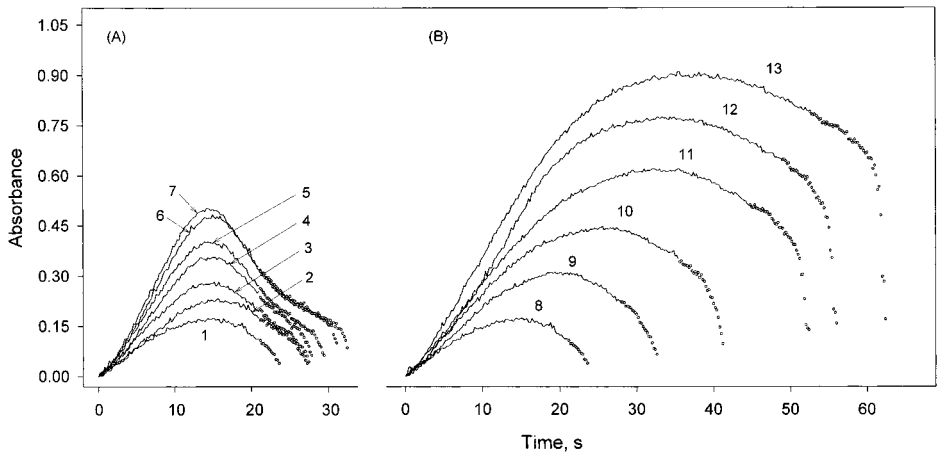

Figure 3. Variation of the absorbance with time for pyrogallol and gallic acid in mixtures. (A) Curves 1−7 correspond to mixtures with a fixed amount of pyrogallol (1 μmol) and variable amounts of gallic acid (1−7 μmol). (B) Curves 8−13 correspond to mixtures in which the amount of gallic acid was fixed at 1 μmol and that for pyrogallol was changed from 1 to 6 μmol.

mixtures of the phenols. The responses are normalized with respect to the injection time: the analytical signal (absorbance) and time at this point were both set to zero. As can be seen, the two phenols exhibit a rather different kinetic behavior. Thus, mixtures with the same concentration of *P* (1.0 μmol) and variable concentrations of *GA* over the range 1−7 μmol exhibit a gradual increase in the signal corresponding to the maximum, a_m, with increase in the concentration of the latter (Figure 3A); on the other hand, the time corresponding to the maximum, t_m, remains virtually constant. The opposite behavior is observed in mixtures with a constant concentration of *GA* (1 μmol) and variable concentrations of *P* from 1 to 6 μmol (Figure 3B): both a_m and t_m increase with an increase in the concentration of *P*—the former to a greater extent than in the previous case owing to the increased molar absorptivity of the oxidation product of *P*. This behavior is consistent with a stronger dependence of the analytical response on the concentration of *P*, which, as shown below, results in increased accuracy in its joint determination with *GA*. In developing a new methodology for resolving multicomponent mixtures, it is interesting to previously determine the presence of synergistic effects, which are a major source of error in kinetic analysis. To this end, the a_m and t_m values obtained in several experiments with real mixtures were compared with those provided by "theoretical" values representing the sum of the real kinetic data for the individual phenols. a_m was found to be scarcely additive: its real value exceeded the "theoretical" one by ca. 20−50%, depending on the particular concentrations of *P* and *GA* in the mixture. On the other hand, t_m remained virtually constant, which confirms the presence of synergistic effects in the studied mixtures.

The principal issue to be addressed in multicomponent kinetic determinations using CNNs is selecting of the type of inputs to be used and their number. A small number of inputs reduces CNN complexity and learning time as well as the number of patterns needed to train the network, which is of great practical significance. Initially, we used up to 34 S_{t_i} values that were uniformly distributed on both sides of the maximum of the response curve for the mixture as inputs; however, the average SEP for the test set (10 runs of the 34:5:2 model) was very high for both components: 21.9% for *P* and 32.3% for *GA*.

Based on these results, and, taking into account the above-mentioned dependence of *P* and *GA* on a_m and t_m, an equation based on these parameters was developed to model the response profile provided by the oscillating chemical system upon perturbation. The equation corresponding to a three-parameters Gaussian curve, in which the additional parameter *s* was associated with the dispersion of the analytical signal (absorbance) values from that corresponding to the maximum, was found to be the best choice. As can be seen from Figure 3, this parameter depends directly on the concentration of *P* in the mixture. A compromise situation must be adopted with a view to fitting. In fact, the optimal time region resulting in the best fit should be determined, especially for mixtures with high contents of *P* (see Figure 3B). The best results were obtained by using a time domain spanning the range from $t_i = 0$ (the time at which the mixture was injected into the oscillating system) to $t_i = t_m + 0.4t_m$ (R^2 ranged from 0.9733 to 0.9989). At longer times, R^2 decreased through a stronger influence of regular oscillations on the analytical response as they tended to regain the steady state.

Selection of Network Topology and Data Set Size. In this work, to further investigate the prediction ability of pruning CNNs, besides global model, neural networks models for individual component were also made with respect to the output layer considered as a single node corresponding to the analyte to be analyzed. Two different data sets were tested for this purpose: (1) the whole data set (81 samples) obtained by preparing 27 synthetic samples (in triplicate as described under Experimental Section), which was designated (train/test 54:27) and (2) a partial data set consisting of 22 synthetic samples (also prepared in triplicate), where the highest concentration of *P* in the mixture was 4 μmol. This data set was designated (train/test 44:22) and was processed by considering the above-described kinetic behavior of *P* in the mixture as one way of increasing the accuracy of the method, albeit at the expense of reducing the range of [*P*]/[*GA*] ratios determinable in the mixture.

Table 2 shows the results provided by the network topologies tested in all experiments (over 10 runs). In our analysis of the experimental data we used the analysis of variance (ANOVA) technique to ascertain the statistical significance of observed differences among the means, assuming that the SEP values obtained have a normal

Table 2. 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

			pyrogallol					
			mean			SD		
model	network topology	data set train/test	SEP train	SEP test	connections	SEP train	SEP test	connections
1	3:3:1	44/22	5.07	4.83	11.1	0.60	0.34	2.72
2	3:2:1	54/27	5.86	5.64	9.1	0.46	0.44	1.37
5	3:4:2	44/22	3.82	3.48	21.7	0.27	0.41	3.19
6	3:3:2	54/27	5.01	4.01	18.2	0.19	0.26	0.91
			best			worst		
model	network topology	data set train/test	SEP train	SEP test	connections	SEP train	SEP test	connections
1	3:3:1	44/22	3.79	4.19	10	5.67	5.19	14
2	3:2:1	54/27	4.76	4.65	8	6.22	6.12	9
5	3:4:2	44/22	3.40	3.17	19	3.92	4.36	24
6	3:3:2	54/27	3.79	3.63	18	5.20	4.40	17
			gallic acid					
			mean			SD		
model	network topology	data set train/test	SEP train	SEP test	connections	SEP train	SEP test	connections
3	3:2:1	44/22	5.44	5.66	9.7	0.36	0.17	0.67
4	3:3:1	54/27	10.80	10.01	11.1	0.44	0.85	1.91
5	3:4:2	44/22	5.13	5.81	21.7	0.29	0.30	3.19
6	3:3:2	54/27	9.98	8.98	18.2	0.21	0.38	0.91
			best			worst		
model	network topology	data set train/test	SEP train	SEP test	connections	SEP train	SEP test	connections
3	3:2:1	44/22	5.50	5.38	9	5.89	6.01	10
4	3:3:1	54/27	10.01	8.51	15	11.37	11.07	10
5	3:4:2	44/22	4.92	5.50	19	5.21	6.45	24
6	3:3:2	54/27	9.87	8.45	18	10.29	9.70	17

Table 3. Homogeneous Subsets of Models Provided by the SNK Test

			pyrogallol					
			av SEP for training set subsets for $\alpha = 0.05$			av SEP for test set subsets for $\alpha = 0.05$		
model	network topology	data set train/test	1	2	3	1	2	3
5	3:4:2	44/22	3.82			3.48		
6	3:3:2	54/27			5.01		4.01	
1	3:3:1	44/22	5.07				4.83	
2	3:2:1	54/27		5.86				5.64
	p-value		0.429	1.000	1.000	1.000	0.792	1.000
			gallic acid					
			av SEP for training set subsets for $\alpha = 0.05$				av SEP for test set subsets for $\alpha = 0.05$		
model	network topology	data set train/test	1	2	3	4	1	2	3
5	3:4:2	44/22	5.13				5.81		
3	3:2:1	44/22		5.44			5.66		
6	3:3:2	54/27			9.98			8.98	
4	3:3:1	54/27				10.80			10.01
	p-value		1.000	1.000	1.000	1.000	0.118	1.000	1.000

distribution. The ANOVA involves a linear regression model in which % SEP is the dependent variable, and the independent variable is the network architecture. The comparisons were done in terms of a critical level for Snedecor's *F* values with 3 and 36 degrees of freedom. If the signification level, α, is higher than this critical level, *p*, the hypothesis of identical SEP's means were rejected. If this hypothesis was rejected a SNK test for ranking mean values is applied in order to obtain the corresponding homogeneous subsets shown in Table 3. Finally, based on these results and on analytical chemical considerations, the model to be used to resolve the mixture was adopted.

The $F_{3,36}$ values for the training and test sets of *P* were 20 and 99.12, respectively; both corresponded to a critical level $p = 0.000$, so the hypothesis of identical means was rejected. As can be seen from Table 3, the SNK test provided three homogeneous training and generalization subsets: α was equal to 0.05 in all cases and hence smaller than the *p* value for each topology subset. Based on the results of the SNK test, 3:4:2 (44/22) CNN was the best topology (model 5) as it provided relative SEP values for the training and testing sets that were smaller than those yielded by the other topologies. However, mixtures with higher [P]/[GA] ratios, which are of great practical interest, can be resolved with

the 3:3:2 (54/27) topology (model 6) with no appreciable loss of accuracy. Finally, it is worth noting that the best results in the determination of *P* in the mixture were obtained with topologies including two neurons in the output layer, which allowed the CNNs to simultaneously deliver the concentrations of *P* and *GA* in the mixture. This can be ascribed to the above-mentioned synergistic effect present in the mixtures. For *GA*, the $F_{3,36}$ values obtained (211.77 and 642.84 for the training and test sets, respectively) also allow one to reject the hypothesis of identical means. Table 3 shows the subsets obtained by using the SNK test similarly as in the previous case. The size of the data set proved crucial in selecting the model to be used. In fact, the worst results were obtained when the set included mixtures containing increased amounts of *P* relative to *GA* (up to 6 μmol), which is consistent with the substantial synergistic effect of *P* on the determination of *GA*. Also, as in the previous case, once the size of the data set was established, the best results were obtained with topologies including two neurons in the output layer. Based on these results, the 3:3:2 (54/27) model was adopted with a view to resolving the *P*/*GA* mixture for the following reasons: (a) it afforded the resolution of mixtures spanning a broad concentration ratio; (b) the average SEP value obtained for *GA*, 8.98%, was analytically acceptable; and (c) the average SEP value for *P* was only slightly increased relative to the optimum model (see Table 3).

Generalization Ability of Pruned CNNs in the Simultaneous Determination of Pyrogallol and Gallic Acid. Figure 1B shows the architecture of the pruned CNN used to determine *P* and *GA* in mixtures. The simplicity of this architecture allows one to derive a straightforward equation system for the direct determination of the mixture components using (a) the parameters estimated by Gaussian regression of the response curve; (b) the optimized network weights; and (c) the sigmoidal transfer functions. The equations providing the best results with a 3:3:2 architecture (model 6) were as follows:

$$[\hat{P}]^* = 1.54\hat{h}_1 + 0.37\hat{h}_2 - 0.78\hat{h}_3 \quad (16)$$

$$[G\hat{A}]^* = -0.69 - 1.74\hat{h}_1 + 1.89\hat{h}_2 + 1.22\hat{h}_3 \quad (17)$$

$$\hat{h}_1 = \frac{1}{1 + exp(3.69 - 0.07\,\widehat{a}_m^* - 3.54\,\widehat{s}^*)} \quad (18)$$

$$\hat{h}_2 = \frac{1}{1 + exp(0.74 - 3.71\,\widehat{a}_m^* + 1.50\,\widehat{s}^* - 0.38\,\widehat{t}_m^*)} \quad (19)$$

$$h_3 = \frac{1}{1 + exp(0.80 - 1.93\,\widehat{a}_m^* + 7.05\,\widehat{s}^* + 2.30\,\widehat{t}_m^*)} \quad (20)$$

Note that $[\hat{P}]$ and $[G\hat{A}]$ are scaled values, so they must be rescaled to obtain the estimated concentrations of *P* and *GA*.

There is wide consensus that CNNs are effective algorithms for the nonlinear recognition and sorting of patterns but are like "black boxes", i.e. the models they use for these purposes are difficult to interpret. Because our models are very simple, they can be more easily approached and explained. Let us briefly discuss the most salient features of the proposed model in the light of eqs 16−20. From such equations it follows that (a) $\hat{h}_1$ is a sigmoidal function depending directly on the estimated dispersion, $\hat{s}$, of the Gaussian curve. It bears a direct relationship with *P* and an inverse one with *GA* and ranges from 0.035 to 0.4. (b) $\hat{h}_2$ is a sigmoidal function essentially dependent on the maximum height, $\hat{a}_m$. This is the function best accounting for the *P* and *GA* concentrations in the mixture. It is related in a linear, positive manner with both $[\hat{P}]$ and $[G\hat{A}]$ and ranges from 0.37 to 0.83. (c) $\hat{h}_3$ is a sigmoidal function inversely related with the dispersion, $\hat{s}$, and taking negligible values at *P* concentrations above 2 μmol. It bears an inverse linear relationship with $[\hat{P}]$ and a direct one with $[G\hat{A}]$ and ranges from 0.0005 to 0.13.

From the foregoing it follows that $\hat{s}$ and $\hat{a}_m$ are the two parameters of the Gaussian curve most strongly influencing the determination of the concentrations of *P* and *GA* in the mixtures; however, variable $\hat{t}_m$, which is not removed by the pruning process, provides better fitting in the resolution of such mixtures.

The magnitude of the relative errors in the concentrations estimated by using these equations can be clearly envisaged from columns 3 and 4 in Table 4, which shows the results for various synthetic binary mixtures containing variable amounts of *P* and *GA*. Mixtures containing the analytes in ratios from 1:1 to 4:4 can be accurately resolved with relative errors less than 10% in both components—mixtures with higher contents of *P* provide appreciable increased relative errors in *GA*, however.

Table 4. Comparison of the Quality Achieved in the Resolution of Pyrogallol/Gallic Acid Mixtures Using CNNs and SRA

[pyrogallol]/ [gallic acid]	3:3:2 CNNs error (%)		54/27 model SRA error (%)	
	pyrogallol	gallic acid	pyrogallol	gallic acid
1:1	−1.61	−5.99	−14.01	−13.39
1:2	1.82	−5.04	−13.84	−4.87
1:3	−4.88	5.73	−7.38	2.40
1:4	6.98	−5.06	9.57	−4.96
1:5	8.37	−3.64	10.35	−6.98
1:6	8.11	−1.18	16.80	−6.16
1:7	−5.89	−3.94	18.08	−11.42
2:1	3.78	−3.92	6.28	22.38
2:2	8.20	−9.25	7.59	−9.71
2:3	2.94	2.45	−3.18	13.34
2:4	5.27	−1.84	−2.58	6.25
2:5	4.30	2.72	−2.37	8.77
2:6	4.34	−3.03	−0.47	1.84
3:1	−5.21	9.17	1.31	12.58
3:2	−1.80	−8.50	0.65	−17.50
3:3	−2.33	5.77	0.33	7.43
3:4	−1.45	−2.62	−0.52	−0.16
3:5	−0.03	2.70	0.77	8.29
4:1	0.85	13.53	5.07	−12.68
4:2	−2.97	−12.49	1.62	−23.51
4:3	−2.10	−6.55	1.98	−13.61
4:4	1.73	−9.15	5.55	−13.31
5:1	−2.88	17.89	−2.33	16.27
5:2	6.47	−6.90	4.98	−9.94
5:3	1.65	−14.26	1.81	−17.04
6:1	−1.11	27.66	−5.34	38.39
6:2	1.06	−15.12	−3.18	−5.10

Response Surface Analysis versus Pruned CNNs for Multicomponent Determinations. Although the use of eqs 16−20 with pruned CNNs allows the accurate resolution of mixtures of *P* and *GA*, are CNNs the best choice for this purpose? Could a simpler alternative be used instead? These questions were answered with the aid of response surface

Table 5. Figures of Merit of Response Surface Analysis as Applied to the Resolution of Mixtures of Pyrogallol and Gallic Acid

analyte	term	coeff	SE	asymptotic 95% confidence interval
	Model 1. Data Set (Train/Test): 44/22			
pyrogallol	intercept	−1.460	0.060	−1.582 − −1.339
	$\hat{a}_m$	1.242	0.121	0.998 − 1.487
	$\hat{s}$	0.250	0.007	0.237 − 0.264
gallic acid	intercept	4.343	0.207	3.928 − 4.763
	$\hat{a}_m$	11.021	0.419	10.174 − 11.867
	$\hat{s}$	−0.635	0.023	−0.681 − −0.588
	Model 2. Data Set (Train/Test): 54/27			
pyrogallol	intercept	−1.660	0.177	−2.017 − −1.303
	$\hat{a}_m$	1.357	0.009	1.338 − 1.375
	$\hat{s}$	0.264	0.069	0.124 − 0.403
gallic acid	intercept	3.509	0.415	2.673 − 4.344
	$\hat{a}_m$	14.674	1.138	12.386 − 16.911
	$\hat{s}$	−0.642	0.025	−0.692 − −0.592
	$\hat{a}_m^2$	−3.428	0.862	−4.981 − −1.515

analysis (RSA). With this type of model,[34] we predicted the concentrations of *P* and *GA* by using a second-order polynomial including terms containing squares of the independent variables $\hat{a}_m$, $\hat{s}$, and $\hat{t}_m$ or their products. Thus, the polynomial response model adopted to predict the concentrations of *P* and *GA* in the mixture was represented by the following general expressions

$$[\hat{P}] = \beta_0 + \beta_1\hat{a}_m + \beta_2\,\hat{s} + \epsilon \tag{21}$$

$$[G\hat{A}] = \beta'_0 + \beta'_1\hat{a}_m + \beta'_2\,\hat{s} + \beta'_3\,\hat{a}_m^2 + \epsilon' \tag{22}$$

where β_0 and β'_0 are intercepts; β_1, β_2, β'_1, β'_2 and β'_3 are regression coefficients; and ϵ and ϵ' are random errors. The (54/27) and (44/22) data sets were also used here to estimate the parameters of the fitting functions, using the Levenberg-Marquard algorithm. Table 5 shows the regression coefficients—significantly different from zero—as well as the standard errors and the asymptotic 95% confidence interval for the proposed models according to the different data set sizes used. The equations derived for each model were as follows:

$$[\hat{P}]_{(44/22)} = -1.460 + 1.242\hat{a}_m + 0.250\,\hat{s} \quad (R^2 = 0.9908) \tag{23}$$

$$[G\hat{A}]_{(44/22)} = 4.343 + 11.021\hat{a}_m - 0.635\,\hat{s} \quad (R^2 = 0.9565) \tag{24}$$

$$[\hat{P}]_{(54/27)} = -1.660 + 1.357\hat{a}_m + 0.264\,\hat{s} \quad (R^2 = 0.9913) \tag{25}$$

$$[G\hat{A}]_{(54/27)} = 3.509 + 14.674\hat{a}_m - 0.642\,\hat{s} - 3.428\hat{a}_m^2 \quad (R^2 = 0.9483) \tag{26}$$

As can be seen, with a (44/22) data set eqs 23 and 24 exhibit a direct linear relationship of $\hat{a}_m$ with both $[\hat{P}]$ and $[G\hat{A}]$ as well as a direct linear relationship of $\hat{s}$ with $[\hat{P}]$ and an inverse one with $[G\hat{A}]$. With an expanded data set, (54/27), eq 25 exhibits the same behavior in $[\hat{P}]$; on the other hand, eq 26 exhibits not only a linear dependence of $[G\hat{A}]$ on $\hat{a}_m$ but also one with a_m^2. Therefore, increasing the *P* concentration in the mixture results in a more markedly quadratic dependence of *GA* on the $\hat{a}_m$ parameter of the Gaussian curves derived from eq 1. Equations 23−26 as applied to the generalization test provided the following SEP values: 4.37 and 4.35% for *P* with the (44/22) and (54/27) data set, respectively; and 10.05 and 11.49% for *GA* under the same conditions. These SEP values are very similar to those provided by pruned CNNs for *P* and higher than those for *GA*. In any case, it is analytically more interesting to compare the quality achieved in the resolution of *P*/*GA* mixtures with the two methodologies in terms of accuracy (expressed as relative errors). Columns 4 and 5 in Table 4 show the results obtained by using eqs 25 and 26, and the (54/27) data set, with those achieved by pruned CNNs. As can be seen, pruned CNNs provided lower relative errors (greater quality) for both components, particularly for *GA* and for mixtures with decreased amounts of *P* (compare the accuracy for 1:1 to 1:7 mixtures). The differences observed in the latter case—particularly—can be ascribed to the RSA models for *P* and *GA* being absolutely independent of $\hat{t}_m$, which is not the case with the pruning algorithm used with the proposed network topology. In summary, these results are quite favorable and testify to the good performance of pruned CNNs in multicomponent determinations based on the perturbation response of the analytes involved in a regular oscillating chemical reaction.

CONCLUSIONS

In this work we developed a novel approach for processing highly nonlinear data provided by oscillating chemical reactions as input data for pruned computational neural networks with a view to analyzing multicomponent mixtures.

The proposed CNN models are straightforward enough to represent the effect of each of the three parameters estimated, by NLR of the Gaussian curve fitted to the signal-time values, on the determination of the concentrations of the analytes in the mixture to be analyzed and interpreted. For this reason, the use of an estimating procedure in two steps (a nonlinear parametric procedure in the first step and a learning of a pattern set in the second step) opens up new avenues for characterizing the processes by which predictions in the kinetic resolution of mixtures are obtained according to the experimental design used; the ensuing models are quite simple, so they allow each of their independent variables to be readily explained and its sensitivity to be assessed. In addition, the poor results (relative errors were high), obtained when response surface analysis was used as an alternative to CNNs, confirm that CNNs are effective tools for solving the addressed analytical problem.

ACKNOWLEDGMENT

The authors wish to thank Spain's Dirección General Interministerial de Ciencia y Tecnología (DGICyT) for funding this research within the framework of Project ALI98-0676-CO2-02 (Department of Computer Science, University of Córdoba) and Project INTAS grant INTAS-OPEN-97-1094.

REFERENCES AND NOTES

(1) Pérez-Bendito, D.; Silva, M. Recent advances in kinetometrics. *Trends Anal. Chem.* **1996**, *15*, 232−240.
(2) Wentzell, P. D.; Karayannis, K. I.; Crouch, S. R. Simultaneous kinetic determination with the Kalman filter. *Anal. Chim. Acta* **1989**, *224*, 263−274.

(3) Rui, X.; Velasco, A.; Silva, M., Pérez-Bendito, D. Performance of the Kalman filter algorithm in differential reaction-rate methods. *Anal. Chim. Acta* **1991**, *251*, 313−319.
(4) Quencer, B. M.; Crouch, S. R. Multicomponent kinetic determination of lanthanides with stopped-flow, diode array spectrophotometry and the extended Kalman filter. *Anal. Chem.* **1994**, *66*, 458−463.
(5) Jiménez-Prieto, R.; Velasco, A.; Silva, M.; Pérez-Bendito, D. Kalman filtering of data from first- and second-order kinetics. *Talanta* **1993**, *40*, 1731−1739.
(6) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Riba, J.; Rovira, E. Kinetic spectrophotometric determination of Ga(III)-Al(III) mixtures by stopped-flow injection analysis using principal component regression. *Talanta* **1993**, *40*, 261−267.
(7) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Redon, M. Principal component regression for mixture resolution in control analysis by UV−visible spectrophotometry. *Appl. Spectrosc.* **1994**, *48*, 37−43.
(8) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Riba, J. Kinetic spectrophotometric method for analyzing mixtures of metal ions by stopped-flow injection analysis using partial least-squares regression. *Anal. Chem.* **1994**, *66*, 2905−2911.
(9) García Fraga, J. M.; Jiménez Abizanda, A. I.; Jiménez Moreno, F.; Arias León, J. J.; Havel, J. Kinetic spectrophotometric resolution of mixtures of europium, terbium, and lanthanum in a micellar medium using multivariate calibration. *Microchem. J.* **1996**, *54*, 32−40.
(10) Pettersson, A. K.; Karlberg, B. Simultaneous determination of orthophosphate and arsenate based on multi-way spectroscopic-kinetic data evaluation. *Anal. Chim. Acta* **1997**, *354*, 241−248.
(11) Kappes, T.; López-Cueto, G.; Rodríguez-Medina, J. M.; Ubide, C. Improved selectivity in multicomponent determinations through interference modelling by applying partial least squares regression to kinetic profiles. *Analyst* **1998**, *123*, 2071−2077.
(12) Ventura, S.; Silva, M.; Pérez-Bendito, D.; Hervás, C. Multicomponent kinetic determinations using artificial neural networks. *Anal. Chem.* **1995**, *67*, 4458−4461.
(13) Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; Redon, M. Artificial neural networks for multicomponent kinetic determinations. *Anal. Chem.* **1995**, *67*, 4477−4483.
(14) Ventura, S.; Silva, M.; Pérez-Bendito, D.; Hervás, C. Computational neural networks in conjunction with principal component analysis for resolving highly nonlinear kinetics. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 287−291.
(15) Ventura, S.; Silva, M.; Pérez-Bendito, D.; Hervás, C. Estimation of parameters of kinetic compartmental models by use of computational neural networks. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 517−521.
(16) Hervás, C.; Ventura, S.; Silva, M.; Pérez-Bendito, D. Computational neural networks for resolving nonlinear multicomponent systems based on chemiluminescence methods. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1119−1124.
(17) Blanco, M., Coello, J.; Iturriaga, H.; Maspoch, S.; Porcel, M. Simultaneous enzymatic spectrophotometric determination of ethanol and methanol by use of artificial neural networks for calibration. *Anal. Chim. Acta* **1999**, *398*, 83−92.
(18) López-Cueto, G.; Ostra, M.; Ubide, C. Linear and nonlinear multivariate calibration methods for multicomponent kinetic determinations in case of severe nonlinearity. *Anal. Chim. Acta* **2000**, *405*, 285−295.
(19) Scheeline, A.; Kirkor, E. S.; Kovacs-Boerger, A. E.; Olson, D. L. Analytical chemistry of nonlinear systems. *Mikrochim. Acta* **1995**, *118*, 1−42.
(20) Jiménez-Prieto, R.; Silva, M.; Pérez-Bendito, D. Approaching the use of oscillating reactions for analytical monitoring. *Analyst* **1998**, *123*, 1R-8R.
(21) Jiménez-Prieto, R.; Silva, M.; Pérez-Bendito, D. Simultaneous determination of gallic acid and resorcinol based on and oscillating chemical reaction by the analyte pulse perturbation technique. *Anal. Chim. Acta* **1996**, *334*, 323−320.
(22) Zupan, J.; Gasteiger, J. *Neural networks for chemists. An introduction*; VCH: Weinheim, 1993.
(23) Schulz, H. Neural networks in analytical chemistry-A new method or only a passing fashion? *GIT. Fachz. Lab*. **1995**, *39*, 1009−1010.
(24) Zupan, J.; Novic, M.; Ruisanchez, I. Kohonen and counterpropagation artificial neural networks in analytical chemistry. *Chemom. Intell. Lab. Syst.* **1997**, *38*, 1−23.
(25) Despagne, F.; Massart, D. L. Variable selection for neural networks in multivariate calibration. *Chemom. Intell. Lab. Syst*. **1998**, *40*, 145−163.
(26) Andersson, G. G.; Kaufmann, P. Development of a generalized neural network. *Chemom. Intell. Lab. Syst*. **2000**, *50*, 101−105.
(27) Weigend, A.; Rumelhart, D.; Huberman, B. Generalization by weight elimination with application to forecasting. *Adv. Neural Inform. Process. Syst.* **1991**, *3*, 875−882.
(28) Williams, P. M. Bayesian regularization and pruning using a Laplace prior. *Neural Comput.* **1995**, *7*, 117−143.
(29) Treadgold, N. K.; Gedeon, T. D. Exploring constructive cascade networks. *IEEE Trans. on Neural Networks* **1999**, *10*, 1335−1350.
(30) White, H. Learning in artificial neural networks: A statistical perspective. *Neural Comput.* **1989**, *1*, 425−464.
(31) Watanabe, S.; Fukumizu, K. Probabilistic design of layered neural networks based on their unified framework. *IEEE Trans. Neural Networks* **1995**, *6*, 691−702.
(32) Marquardt, D. W. An algorithm for least.squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Mathem*. **1963**, *11*, 431−441.
(33) Hartley, H. O. The modified Gauss−Newton method for the fitting of nonlinear regression functions by least-squares. *Technometrics* **1961**, *3*, 269−280.
(34) Rawlings, J. O.; Pantula, S. G.; Dickey, D. *Applied regression analysis*: *A research tool*; Springer-Verlag: New York, 1998.
(35) 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.
(36) Tikhonov, A. N. On solving incorrectly posed problems and method of regularization. *Doklady Akademii Nauk USSR* **1963**, *151*, 501−504.
(37) 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−Zhabotinskyi reaction. *Analyst* **2000**, *125*, 2118−2124.
(38) Minai, A. A.; 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.

CI010012J