



THE USE OF NEURAL NETWORKS FOR FITTING COMPLEX KINETIC DATA

I. M. GALVÁN,^a J. M. ZALDÍVAR,^a H. HERNÁNDEZ,^b E. MOLGA^c

^a Joint Research Centre of the Commission of the European Communities, Institute for Safety Technology, Process Engineering Division TP 680, 21020-Ispra (VA) Italy

^b Joint Research Centre of the Commission of the European Communities, Institute for Prospective Technological Studies Edificio World Trade Center, Isla de la Cartuja s/n, 41092-Sevilla Spain

^c Chemical and Process Engineering Department, Warsaw Technical University, ul. Warynskiego 1, 00-645 Warszawa Poland

(Received 25 April 1995)

Abstract—In this paper the use of neural networks for fitting complex kinetic data is discussed. To assess the validity of the approach two different neural network architectures are compared with the traditional kinetic identification methods for two cases: the homogeneous esterification reaction between propionic anhydride and 2-butanol, catalysed by sulphuric acid, and the heterogeneous liquid–liquid toluene mononitration by mixed acid. A large set of experimental data obtained by adiabatic and heat flux calorimetry and by gas chromatography is used for the training of the neural networks. The results indicate that the neural network approach can be used to deal with the fitting of complex kinetic data to obtain an approximate reaction rate function in a limited amount of time, which can be used for design improvement or optimisation when, owing to small production levels or time constraints, it is not possible to develop a detailed kinetic analysis.

Copyright © 1996 Elsevier Science Ltd

Keywords: Kinetic identification; neural networks; tendency modelling

INTRODUCTION

In the last years batch and semibatch operations have become more popular owing to their versatility which allows the production of special chemicals — with very good yields — in small amounts (when compared with those of continuous processes), and permits a rapid change from one process to another with minor modifications.

Despite the fact that significant economical benefits could be derived from optimisation of such processes, the complex reaction network that characterises them and the shorter market window make, in many instances, the development of a detailed understanding of the reaction kinetics impractical, which is the first prerequisite for the optimisation procedure.

It is therefore desirable to develop approximate kinetic models which can be used to maximise the yield and selectivity of the desired product, and improve the safety of its production.

The kinetic analysis of complex systems of chemical reactions is a difficult problem. The major complication results from the complex stoichiometry and thermodynamics and an enormous number of possible kinetic models — combinations of reaction schemes and kinetic equations for each reaction — from which to choose. Furthermore, heterogeneous reactions involve mass transfer and hence the overall conversion rate is also

influenced by factors such as stirred speed, interfacial area, diffusion coefficients, etc.

Various routes have been developed to perform this kinetic analysis. Fundamentally, these are:

(a) Traditional approach: it consists in developing a detailed kinetic model based on molecular kinetics and studies only some of the reactions in the scheme, or at least, some which reduce the complexity of the problem to a set of simple reactions. This alternative is not always possible since even for simple reactions the complete scheme is still not understood.

(b) Tendency modelling: to overcome the limitations of the white box modelling, a considerable interest has appeared, during the last years, in the development of alternative kinetic identification techniques, i.e. the so called tendency modelling (Filippi *et al.*, 1986), which is based on a phenomenological approach and on the estimation of a small number of kinetic parameters by matching the model predictions to experimental data of previous batches. This is called grey model approach because it does not require a detailed kinetic description (white model) but takes into account general knowledge about the process such as mass and energy balances. This option reduces the kinetic scheme to a very simplified one, often lumping chemical species with simple kinetics, normally power functions, in the stages of the scheme considered.

The methods applied for calculating the parameters in

such a system are normally based upon regression, and have many variations, see Himmelblau *et al.* (1967) for a review. The most utilised methods have been the differential method (Kennard and Melsen, 1985) and the integral method (Himmelblau *et al.*, 1967).

(c) Approximation methods: the development of mathematical analysis has led to the discovery and study of important classes of approximating functions, which, under certain conditions, have proved to be the natural means of approximating other, more or less arbitrary functions. These approximation functions include polynomials, trigonometric series, orthogonal functions, splines, etc. In this case, the adjusting parameters are basically viewed as vehicles for fitting the data and, in principle, do not reflect physical considerations about the reacting system. For these reasons, this approach is called black model.

In this paper, the use of neural network techniques as an alternative — black model — approach to solving the problem of fitting experimental data to obtain an approximate reaction rate function is illustrated and the results are compared with some existing methods. In this context, the multilayer neural network (MNN) and the radial basis function network (RBFN) can be considered as merely two additional (parameterized) families of functions which have evolved in recent years. The Stone–Weierstrass theorem provides a sound basis to determine whether a particular class of neural networks can approximate arbitrary continuous functions. Based on this theorem, different authors have proved that the MNNs, with as few as a hidden layer, are capable of approximating any continuous functions (Cybenko, 1989; Funahashi, 1989; Gallant and White, 1988; Hornik *et al.*, 1989). In the case of RBFNs the studies are quite straightforward and there are some results about its universal approximative properties (Poggio and Girosi, 1990).

ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (NN) are parallel interconnected networks of simple computational elements that are intended to interact with the objects of the real world in a similar way to the biological nervous systems (Kohonen, 1988). Formally, an NN is an oriented graph in which the nodes represent a set of processing units, called neurones, and the arches (or connections) represent the information flow channels. Each connection between two neurones has an associated value called weight (w_{ij}) which specifies the strength of the connection from unit j to unit i . Positive and negative values determine excitatory and inhibitory connections, respectively.

The choice of a specific class of networks for the approximation of a nonlinear map depends on a variety

of factors dictated by the context and is related to the accuracy desired and the prior information available concerning the input–output pairs (patterns).

In this work we shall distinguish two classes of networks which will be described briefly: (1) Multilayer Neural Networks (MNN), and (2) Radial Basis Networks (RBFN).

Multilayer neural networks

The MNN have their nodes divided into disjoint subsets, called layers. Each node forms a weighted sum of the inputs from previous layers to which it is connected, adds a threshold value and produces its activation or output value,

$$\text{out}_i = f\left(\sum_{j=1}^n w_{ij}x_j + \theta_i\right) \quad (1)$$

where x_j ($j=1, \dots, n$), w_{ij} ($j=1, \dots, n$) and θ_i are the inputs, weights and threshold associated to the unit i , respectively. f is called activation function and the most often used are the smooth sigmoidal functions:

$$f(x) = \frac{1 - e^x}{1 + e^x} \text{ or } f(x) = \frac{e^x}{1 + e^x} \quad (2)$$

This output value serves as input to the next layer to which the node is connected, and the process is repeated until output values are obtained in the output layer.

The number of layers in the network, and the number of nodes in each layer are important parameters of the network. Once these have been selected, by trial and error in each specific context, only the adjustable weights have to be determined to specify the network completely. At this point the so-called learning rule comes into operation, which determines the modification of the weights and the threshold values. Most learning rules are formulated with a specific goal, e.g. move the weight vector to a position which minimises or maximises some particular cost or performance function. The weights in the MNN are generally adjusted to minimise the error between the output (y) of the network and some desired output (o),

$$\text{Error} = \frac{1}{N_p} \sum_{i=1}^{N_p} (y_i - o_i)^2 \quad (3)$$

where N_p is the number of pattern presented to the NN.

Owing to the presence of nonlinear functions in the network, it also follows that the output of the network depends non linearly on the weights. Many useful optimisation techniques exist, but the most commonly used method makes use of gradient information. The algorithm used for gradient calculation is the classic backpropagation algorithm (Rumelhart *et al.*, 1986).

Radial basis networks

An alternative to MNN are the radial basis networks (Moody and Darken, 1989). The RBFN is a three layer network in which the first layer only transmits the external input to the hidden layer with no connections. The hidden layer performs a fixed nonlinear transformation with no adjustable parameters. The output layer combines linearly the outputs of the nodes in the hidden layer. The equation that describes the output of the RBFN is the following:

$$y = \sum_{i=1}^N w_i \phi_i(u) + \theta \quad (4)$$

where w_i ($i=1, \dots, N$) are the weights in the network, u is the input vector to the network, ϕ_i are functions that have an appreciable value only in the neighbourhood of a vector c_i which is called the center of ϕ_i and θ is the threshold associated to the output node.

These types of function are called radial basis functions and the most general class contains the Gaussian functions that have the form:

$$\phi_i(u) = \exp\left(-\sum_{j=1}^n \frac{(u_j - c_{ij})^2}{2\sigma_i^2}\right) \quad (5)$$

where $c_i = (c_{i1}, \dots, c_{in})$ and σ_i are the center and width of the Gaussians, respectively.

Designing an RBFN involves determination of the parameters w_i , c_i and σ_i . This problem may be solved by breaking it into two phases. The first phase determines the centers c_i and the standard deviation σ_i in an unsupervised manner, while the second phase performs the optimisation via supervised training to determine w_i .

Phase 1: During this phase the centers and the widths of the n radial basis functions are computed. The K-means clustering (Duda and Hart, 1973) or Kohonen algorithm (Kohonen, 1982; Kohonen, 1984) can be used to find the n clusters in the input training examples. The centers of the clusters determine values of the c_i and the width σ_i is calculated as the distance from the center c_i to its two nearest neighbours.

Phase 2: The weights w_i in the RBFN are found from minimisation of the error function, equation 3. Since these parameters are related linearly to the output and hence to the output error, they can be adjusted using the least squares method.

The number of the radial basis units or hidden units needed to approximate a given function is a critical factor. Generally this number tends to increase exponentially with the dimension of the input space, so that the approach becomes practically unfeasible when the dimension of the input space is high. In such cases it is possible to use a different learning performance in order to reduce the number of the radial basis units required. We will call it supervised method (Moody and Darken,

1989). The difference of this model with respect to the RBFN model is in the learning algorithm. A number of Gaussians units, whose centers are initialised using the Kohonen algorithm, is fixed. Afterwards, the centers and widths are also adapted as the weights, to minimise the error function (equation 3) in a supervised manner. Since the output is a nonlinear function of those parameters the rule used to adapt them is based on the gradient method. The derivatives of the output with respect to the adjustable parameters w_i , c_i and σ_i can be easily calculated.

CASE STUDIES: ESTERIFICATION AND NITRATION REACTIONS

To assess the validity of the NN approach for kinetic identification, two different reactions have been chosen: acid catalysed esterification reaction between propionic anhydride and 2-butanol (homogeneous) and toluene nitration by mixed acid (heterogeneous).

Esterification reaction

The system 2-butanol/propionic anhydride (acid) which has been studied by Snee *et al.* (1992) in safety assessment was selected as a specific case. This homogeneous reaction has some characteristics that make it very interesting for testing studies: moderately exothermic with no danger of decomposition reactions; reaction rate variable as a function of catalyst (strong acid, i.e. sulphuric acid); the reaction exhibits a second-order kinetics when no strong acid is present and a kind of autocatalytic behaviour when sulphuric acid is introduced.

Experimental part. An extensive experimental programme was carried out using adiabatic and heat flow calorimetry supported by chemical analysis. The adiabatic experiments were performed using the PHI-TEC calorimeter (Singh, 1989), while the isothermal and isoperibolic — constant jacket temperature — experiments were made using an RC1 reaction calorimeter (Riesen and Grob, 1985) and a small jacketed vessel (approximately 200 ml). Gas Chromatography (GC) was used for the concentration analysis of different species.

The adiabatic experiments are summarised in Table 1. A series of measurements in the PHI-TEC were made using a different molar ratio between propionic anhydride (PAN), 2-butanol (2-B) and propionic acid (PAC) catalysed by various concentrations of sulphuric acid (Sul) with a total sample mass of approximately 60 g. Reagents were added with the sample container positioned inside the calorimeter and adiabatic conditions were established from ambient temperature without using the 'heat-wait-search' mode of operation.

The isothermal variation of the rate of heat generation with time was measured using the Mettler RC1 calorim-

Table 1. Adiabatic experiments: initial conditions

Ex.	$n_{PAH}:n_{2-B}:n_{PAC}$	$x_{Sul} \cdot 10^3$
1	1:1:2	0.0
2	1:1:0	0.0
3	1:2:0	0.0
4	1:1:0	2.2
5	1:1:0	2.9
6	1:1:0	3.6
7	1:1:0	4.3
8	1:1:0	5.9
9	1:1:2	1.5
10	1:2:0	1.9
11	1:3:0	1.5
12	1:2:1	1.4

eter. Table 2 summarises the isothermal experiments. During these experiments, the concentration of 2-butanol and 2-butyl propionate was measured by GC. The alcohol was added to the reactor and time was allowed for the control system to achieve the temperature selected for isothermal measurement. Propionic anhydride (or propionic acid, or a mixture) at ambient temperature was then added quickly to the reactor and the reaction was allowed to proceed until completion.

The isoperibolic experiments were carried out in the same way as the isothermal tests in the bench-scale reactor by configuring it to operate at constant jacket temperature. Experiments were performed at a series of fixed jacket temperatures with an equimolar mixture of propionic anhydride/2-butanol catalysed by the addition of the same amount of sulphuric acid (see Table 3).

The data from adiabatic (temperature versus time) and isothermal (concentrations versus time) experiments were used to fit the kinetic expressions whereas the

Table 2. Isothermal experiments: initial conditions

Ex.	T_{sp} (K)	$n_{PAH}:n_{2-B}:n_{PAC}$	$x_{Sul} \cdot 10^3$
13	343.2	0:1:1	0.0
14	343.2	0:1:1	3.0
15	343.2	0:1:1	6.0
16	343.2	1:1:0	0.0
17	331.6	1:1:0	0.0
18	303.2	1:1:0	3.0
19	303.2	1:1:1	2.0
20	303.2	1:1:2	1.5
21	303.2	2:2:1	2.4

Table 3. Experimental isoperibolic test matrix: initial conditions

Ex.	T_e (K)	$n_{PAH}:n_{2-B}:n_{PAC}$	$x_{Sul} \cdot 10^3$
22	293.9	1:1:0	3.0
23	295.7	1:1:0	3.0
24	298.2	1:1:0	3.0
25	300.7	1:1:0	3.0
26	303.2	1:1:0	3.0
27	308.2	1:1:0	3.0

isoperibolic data were used to test the accuracy of such kinetics through the comparison of experimental versus the theoretical by simulation.

Identification of kinetic scheme. The esterification mechanism involving acidic catalysis has been studied in detail and a number of general possibilities have been recognised (March, 1992). There are four possible mechanisms depending on the following criteria: unimolecular or bimolecular, and acyl or alkyl cleavage. All four of these are SN1, SN2 or tetrahedral mechanisms. This applies to the reaction between 2-butanol and propionic anhydride as well as propionic acid. Moreover, in this relatively simple reaction system another group of reactions have been recognised:

– The reaction between the sulphuric acid and the 2-butanol to form the mono-alkyl sulphuric acid which acts as a different catalyst of the reaction (Dhanuka *et al.*, 1977) and can favour the ionisation of the alcohol to carbonium and, in consequence, the alkyl cleavage (Deno and Newman, 1951).

– The reaction between the sulphuric acid and the propionic anhydride to form the propyl sulphate which acts as a different catalyst of the reaction (Yvernault, 1955).

In addition, different authors (Haldar and Rao, 1992a, b) have shown the influence of the Hammett's acidity function on the reaction rate which depend on the sulphuric and propionic acid concentrations, and temperature.

For all these reasons, only empirical kinetic models can be found in literature, in which the reaction rate constant depends on the sulphuric acid concentration, the initial ratio between reagent concentrations, the organic acid concentration, etc. (Sreeramulu and Rao, 1973; Troupe and Kobe, 1950; amongst others).

The reaction rate without sulphuric acid

The analysis of experiments without sulphuric acid (Ex. 1–3, 16 and 17) showed that in this particular case the reaction rate followed a second-order expression (see Fig. 1 (a)),

$$r = k_0 \cdot C_{PAH} \cdot C_{2-B} \quad (6)$$

where $k_0 = 5.36178 \cdot 10^7 \cdot \exp(-80.47864/R \cdot T_m)$ l/mol·s. Furthermore, the propionic acid seemed not to influence the reaction mechanism.

The reaction rate between propionic acid and 2-butanol

In a separate set of experiments (Ex. 13–15), it was demonstrated that the reaction rate between the propionic acid and 2-butanol was negligible in presence of propionic anhydride (the constant rate is approximately 200 slower), and the reaction could be fitted using an elementary reversible second-order kinetics,

$$r = k_1 \cdot C_{2-B} \cdot C_{PA} - k_{-1} \cdot C_{BPi} \cdot C_W \quad (7)$$

in which the sulphuric acid influence could be correlated by the following expressions (Dhanuka *et al.*, 1977),

$$k_1 = \alpha_1 + \beta_1 \cdot x_{SuI} \tag{8}$$

$$k_{-1} = \alpha_{-1} + \beta_{-1} \cdot X_{SuI} \tag{9}$$

where α was the specific reaction rate constant for non-catalytic reaction and β was a constant ($\alpha_1 = 1.4813 \cdot 10^{-7}$, $\alpha_{-1} = 3.8072 \cdot 10^{-8}$, $\beta_1 = 3.746 \cdot 10^{-3}$, $\beta_{-1} = 1.4863 \cdot 10^{-3}$ at $T_m = 343$ K).

The sulphuric acid as catalyst

The following experimental observations were found during the reaction between 2-butanol and propionic anhydride catalysed by sulphuric acid:

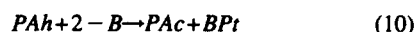
- The initial reaction rate from adiabatic experiments seems proportional to the concentration of sulphuric acid. In addition, the heat of reaction does not change, between the experimental error, with the addition of sulphuric acid (Snee and Hare, 1992).

- The propionic acid concentration increases the reaction rate, producing a sort of autocatalytic behaviour only if there is sulphuric acid.

- After some concentration level the propionic acid seems not to increase the reaction rate anymore.

Since the elucidation of the real pathway was so complex, it was abandoned for another more empirical approach. The existence of two catalysts was postulated (Dhanuka *et al.*, 1977). The former accelerating the

second order reaction and the latter producing a first order reaction expression. Furthermore, the transformation of the catalysts was correlated with the acidity function (Haldar and Rao, 1992a) and the concentration of 2-butanol. This empirical model can be written as,



where the main reaction can be expressed as,

$$r_2 = \left(k_0 + k_2 \cdot C_{Cat_1} \right) C_{PAh} \cdot C_{2-B} + k_3 \cdot C_{Cat_2} \cdot C_{PAh} \tag{12}$$

$$r_3 + k_4 \cdot 10^{-H_R} \cdot C_{Cat_1} \cdot C_{2-B} \tag{13}$$

and H_R is related to the acidity function (Rochester, 1970) and is correlated as,

$$H_R = - \left(P_1 \cdot C_{Cat_1} + P_2 \cdot C_{PAc} \right) \left(P_3 + \frac{P_4}{T_m} \right) \tag{14}$$

The parameters for these expressions were obtained by minimisation through the complete set of experimental data from Table 1 and Table 2 (see Fig. 2). Afterwards, the isoperibolic experiments (Table 3) were used to check their validity. Kinetic parameters are shown in Table 4.

Table 8. Supervised RBFN performances for reaction rate prediction.

Esterification reaction			Nitration reaction		
R. B. units	Error	parameters	R. B. units	Error	parameters
5	$2.9 \cdot 10^{-5}$	36	6	$1.5 \cdot 10^{-3}$	49
10	$1.2 \cdot 10^{-5}$	71	10	$2.9 \cdot 10^{-4}$	81
15	$8.1 \cdot 10^{-6}$	106	15	$7.9 \cdot 10^{-5}$	121

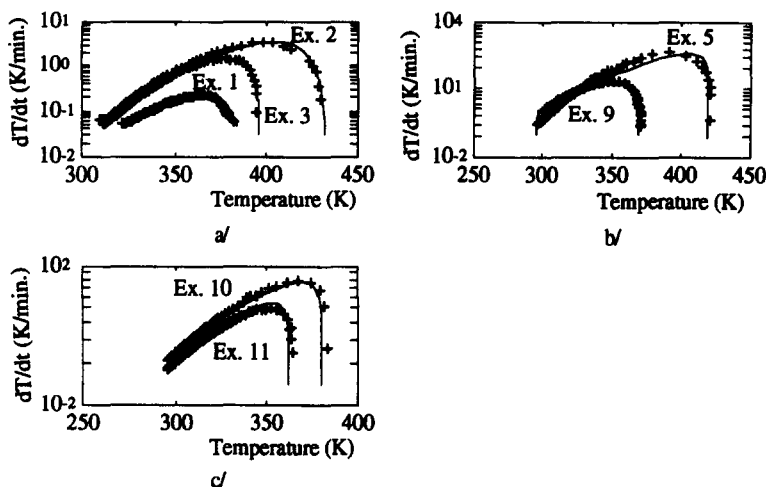


Fig. 1. Experimental (+) and simulated (continuous lines) rates of self-heating for the esterification in the PHI-TEC calorimeter. (a) Without sulphuric acid (see equation 6); (b) and (c) with sulphuric acid (see equations 12-14).

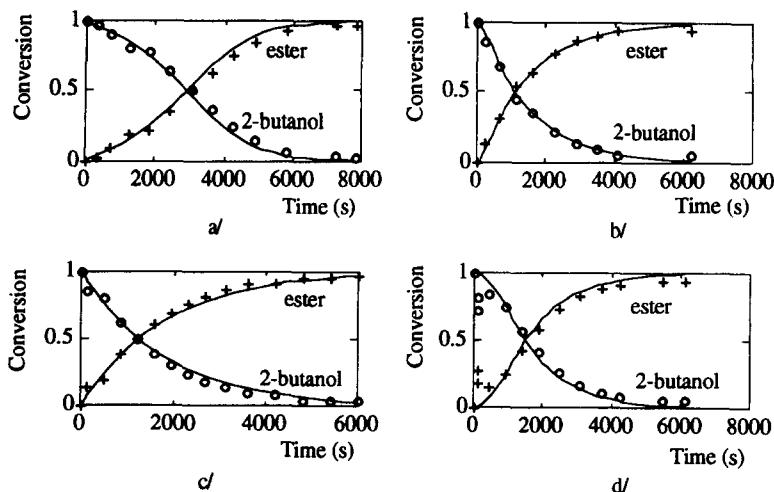


Fig. 2. Experimental (+ and o) and simulated (continuous lines) conversion profiles for the esterification with sulphuric acid. (a) Ex. 18; (b) Ex. 19; (c) Ex. 20; (d) Ex. 21.

In Fig. 3 the temperature profiles of six isoperibolic experiments at different jacket temperatures are plotted. The initial drop in reactor temperature is due to the endothermic mixing of the reagents. As can be seen, the

Table 4. Kinetic parameters of equations (12)–(14)

A_2	2.8074×10^{10}
Ea_2	79.1595
A_3	3.9480×10^{10}
Ea_3	69.9746
A_4	1.4031×10^8
Ea_4	76.6172
p_1	2.002×10^{-1}
p_2	3.205×10^{-2}
p_3	- 21.3754
p_4	12706.0

behaviour of the reaction mass shows that in these conditions we are working in a parametric sensitivity region, defining sensitivity as the variation of the maximum of reactor temperature with respect to that of the jacket, i.e. dT_m/dT_c . In Ex. 22 at jacket temperature of 293.9 K the maximum temperature reached by the reaction mass is 310.5 K whereas in Ex. 27 with a jacket temperature of 308.2 K the maximum temperature is 407.5 K, i.e. a change of 14.3 degrees in jacket temperature produces a change of 97 K in the maximum temperature reached by the reaction mass. This shows how an exothermic reaction can proceed under subcritical (controlled) and supercritical (runaway) conditions. In the latter, the reaction stops when all reagents have been consumed.

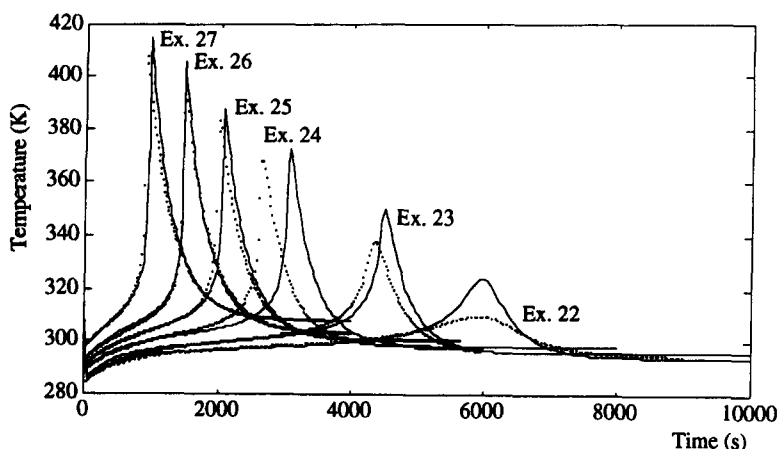


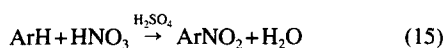
Fig. 3. Experimental (dotted lines) and simulated (continuous lines) reactor temperatures of isoperibolic batch experiments of Table 3 using the empirical model (equations 12–14). 884.0 g. of propionic anhydride were added to 503.2 g of 2-butanol. The molar fraction of sulphuric acid was 3.0×10^{-3} .

Table 5. Test matrix of isothermal experiments in the RC1 calorimeter

Temperatures (K)	288, 298, 308, 318
Sulphuric acid strength (wt.%)	78.9, 74.8, 72.4, 70.6, 68.0, 66.1, 63.9, 61.4
Stirring speeds (s^{-1})	6.67, 10, 13.33
Feeding times (h)	batch, 2, 4

Toluene nitration by mixed acid

The estimation of conversion rate in aromatic nitrations by mixed acid ($H_2SO_4-HNO_3-H_2O$) in discontinuous reactors involves a considerable number of problems owing to the fact that, in these heterogeneous liquid-liquid reactions, chemical reaction and mass transfer phenomena occur simultaneously. Furthermore, as the nitration proceeds water is formed, equation (15), and the acid phase composition varies. This is the reason for which the observed second-order rate constant decreases appreciably in the course of the reaction (Schofield, 1980).



A series of toluene nitration experiments (see Table 5) were carried out in an RC1 reaction calorimeter to study the dynamic behaviour and the influence of the different initial and operating conditions (temperature, sulphuric acid strength, stirrer speed and feeding rate) in discontinuous toluene mononitration.

From the analysis of experimental and simulated results it was concluded (Zaldívar *et al.*, 1992a) that toluene mononitration can follow different reaction regimes; at low temperature or low sulphuric acid strength, the overall conversion rate is in the kinetic regime (slow liquid-liquid reaction), in this case, the rate of mass transfer is not enhanced by the occurring reaction and the reaction proceeds in the bulk of the acid phase (Zaldívar *et al.*, 1995a)

$$R = \epsilon_{aq} \cdot K_2 \cdot C_{ArH}^{aq} \cdot C_{HNO_3}^{aq} \quad (16)$$

However, as we increase temperature or sulphuric acid strength the nitration rate increases and the heterogeneous system enters the fast regime. In this situation the conversion rate is affected not only by the physical and chemical characteristics of the system but also by the mechanical features of the equipment. The former includes viscosities and densities of the phases, interfacial surface properties, diffusion coefficients, distribution coefficients of reagents and products between the phases and chemical reaction constants. The latter includes, for example, the type and diameter of the impeller, vessel geometry, the flow rate of each phase and the rotational speed of the impeller. Furthermore, dispersion phenomena such as coalescence and breakage of droplets, drop size distribution and phase inversion phenomena will affect the extent of conversion and the selectivity of the reactions (Villiermaux, 1981). Nor-

mally, under the conditions employed during the experiments, the reaction rate could be fitted using the fast liquid-liquid reaction expression (Zaldívar *et al.*, 1995b):

$$R = a \cdot \sqrt{k_2 \cdot D_{ArH} \cdot C_{ArH}^{aq} \cdot C_{HNO_3}^{aq}} \quad (17)$$

where a is the effective interfacial area which is normally evaluated by the expression:

$$a = \frac{6 \cdot \epsilon_d}{d_{32}} \quad (18)$$

ϵ_d is the fraction of dispersed phase and d_{32} is the Sauter mean diameter which can be computed as:

$$d_{32} = \frac{\sum_i n_i d_i^3}{\sum_i n_i d_i^2} \quad (19)$$

where n_i is the number of droplets with a diameter d_i . The Sauter mean diameter cannot be evaluated through a simple analysis. Therefore, it is customary to evaluate the maximum stable diameter d_{max} encountered in the impeller zone of the vessel and multiply it by an empirical factor to obtain d_{32} , see Sprow (1967). The general form of the correlations is (Tavlarides and Stamatoudis, 1981):

$$\frac{d_{32}}{D_a} = A \cdot f(\epsilon_d) \cdot We^{-0.6} \quad (20)$$

where $f(\epsilon_d)$ normally represents a linear correlation of the volume fraction of dispersed phase:

$$f(\epsilon_d) = 1 + B \cdot \epsilon_d \quad (21)$$

and We is the Weber number of the main flow, defined as:

$$We = \frac{\rho_c \cdot n_a^2 \cdot D_a^3}{\sigma} \quad (22)$$

The mathematical model as well as the experimental determination of its parameters and constants is described in detail elsewhere (Zaldívar *et al.*, 1992a, 1995a, b) and will therefore not be discussed here. The model employed, based on the effective interfacial area between the two liquids and the description of mass transfer with reaction using the film model is the most common approach. Figure 4 shows the calculated and simulated, equation (17), reaction rate as a function of sulphuric acid strength during semibatch toluene mono-

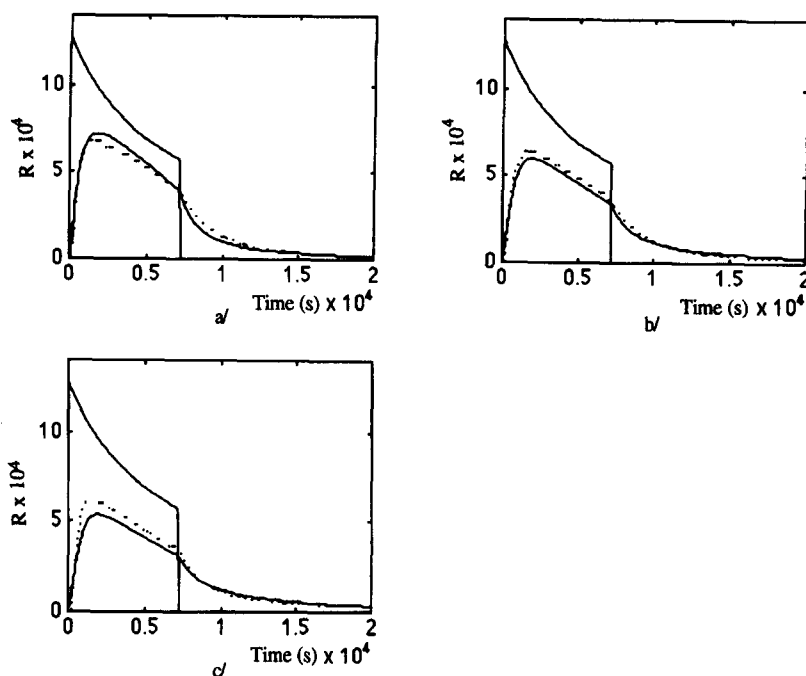


Fig. 4. Experimental (dotted lines) versus simulated, equation (16), reaction rate (mol/l.s) for semibatch toluene mononitration during the fast reaction regime experiments (x , sulphuric acid strength): (a) $x=74.8$ wt.%; (b) $x=72.4$ wt.%; (c) $x=70.6$ wt.%. $T_{op}=308.16$ K, and $Na=6.67$ s $^{-1}$.

nitration experiments. The instantaneous reaction rate — no accumulation of nitric acid in the reactor — is also plotted. As can be seen the agreement between experimental and simulated results is satisfactory. However, the effort required to carry out such analysis was not negligible.

NEURAL NETWORK APPROACH

Many studies have proved the ability of NNs to approximate any nonlinear relationship between a set of inputs and outputs. In our case we are interested in finding the reaction rate value (output of the network) as a function of state variables (inputs to the network). In our system the inputs will be temperatures and concentrations plus some parameters related to the equipment for the heterogeneous system.

The implementation of the neural network approach consists of two steps: the learning procedure and the evaluation procedure.

Learning procedure: training data set

The performance of the NNs in estimating the reaction rate is directly related to the training data. To achieve robust performance, data must be distributed across all the regions of the input space that are of interest. It has been demonstrated that performance of NNs often deteriorates when extrapolating into regions of the input

space for which the network has not been trained (Venkatasubramanian and Chan, 1989).

Two different training data sets for esterification and nitration were generated. The reaction rate was estimated from calorimetric reconstruction using a special version of the reaction calorimeter simulator (Hernández *et al.*, 1993). In this version, the mathematical model is reduced by introducing the measured variables, i.e. reactor, jacket, ambient and feeding temperatures, and replacing their values in the respective differential and algebraic equations (Hernández *et al.*, 1993). The estimation step is carried out by using a recursive least square algorithm with constant forgetting factor (Ljung, 1987). Our simulator produces similar results to the RC1 evaluation package (see RC1 Mettler operating instructions manual) for the calculation of the heat flow owing to the chemical or physical process in the reactor. In the particular case of esterification and nitration the heat of dilution was subtracted from the total heat produced during the process using a model developed in a previous work (Zaldívar *et al.*, 1992b). Once the rate of heat generated is estimated the rate of reaction can be obtained assuming known enthalpy of reaction, in our case equation (10) or equation (15). To validate this calorimetric reconstruction, independent GC measurements (Molga *et al.*, 1993) were used.

In the case of esterification reaction the variables used as inputs to the NN for fitting the reaction rate are: reactor temperature (T_m), propionic anhydride concentra-

tion (C_{PAH}), 2-butanol concentration (C_{2-B}), propionic acid concentration (C_{PAc}), and sulphuric acid concentration (C_{Sul}). The training data set was the same as those used for the determination of kinetic function, i.e. the data from experiments listed in Table 1 and Table 2.

In the case of nitration the reactor temperature, and the overall concentrations of water, nitric and sulphuric acids and toluene were chosen as the inputs to the NN. As the training data set contained experiments carried out in fast nitration regime the interfacial area, which is proportional to the conversion rate (see equation (17)) should be introduced. The interfacial area, equations (18)–(22) is a function of dispersed phase volume and the Weber number, which turns out to be a function of continuous phase density, interfacial tension, stirrer diameter and stirrer speed. It would be interesting to introduce into the network the characteristics of the equipment to see if it is possible to consider scaling-up effects. However, experimental data from different equipment would be necessary to perform such a study. As these data were not available the stirred speed was considered as the only input to the NN, since the stirrer diameter and the other shape factors are constant during all the experiments.

In the esterification case the training set had 2000 patterns while 10,000 patterns were used in the nitration case. The patterns were randomly sampled before being presented to the network. All data were normalised in the interval [0, 1].

Different MNN architectures with three layers have been trained using the backpropagation procedure. The mean-square error (see equation 3) obtained for each architecture in both cases: esterification and nitration is shown in Table 6. The number of adjusted parameters (connection weights) in the network is also specified. As can be seen, 36 parameters are enough to obtain the desired error in the esterification case, while in the nitration case it is necessary to add more hidden nodes to obtain the same order of magnitude for the error. The

results obtained with 10 or 15 hidden nodes are acceptable.

The RBFN model was also used to approximate the reaction rate. The centers c_i and the widths σ_i (equation 5) were determined in an unsupervised manner using the Kohonen algorithm. Owing to the fact that the number of patterns is considerable, instead of using the least squares method, the gradient method was employed to calculate the weights. The calculation of the derivatives of the output with respect to the weight vector is immediate using equation (4). The number of radial basis units was determined by trial and error. Table 7 shows that in the nitration case 350 units are necessary to cover the input space to obtain an acceptable approximation, while 200 hidden units are required in the esterification case to obtain error of similar magnitude to those found using the MNN.

The results obtained with the supervised method aiming to decrease the number of hidden units are shown in Table 8. As can be seen, in both cases: esterification and nitration, the number of radial basis units required to obtain similar errors using this type of learning is considerably reduced in comparison with the unsupervised learning method (RBFN).

Validation procedure: test data set

The performance of trained neural network was tested using different experiments to those used for the training procedure. In the case of the esterification, the test set consisted of the isoperibolic experiments described in Table 3. In this case, the different trained NN were introduced in the numerical simulator of the reaction calorimeter (Zaldivar *et al.*, 1990) replacing the empirical equation of the rate of reaction. As can be seen in Fig. 5 the results using the MNN are similar to those obtained using the empirical model. However, in the case of RBFN (Fig. 6) the agreement is less satisfactory. Owing to the high dimension of the input space it is difficult to find the location of the centres of the radial basis

Table 6. MNN performances for reaction rate prediction

Esterification reaction			Nitration reaction		
Architec.	Error	parameters	Architec.	Error	parameters
5-5-1	$5.1 \cdot 10^{-5}$	36	6-6-1	$9.5 \cdot 10^{-4}$	49
5-10-1	$4.3 \cdot 10^{-5}$	71	6-10-1	$6.8 \cdot 10^{-5}$	81
5-15-1	$4.7 \cdot 10^{-5}$	106	6-15-1	$8.8 \cdot 10^{-5}$	121

Table 7. RBFN performances for reaction rate prediction

Esterification reaction			Nitration reaction		
R. B. units	Error	parameters	R. B. units	Error	parameters
60	$9.8 \cdot 10^{-4}$	61	60	$2.4 \cdot 10^{-3}$	61
100	$3.9 \cdot 10^{-4}$	101	100	$4.1 \cdot 10^{-4}$	101
200	$2.7 \cdot 10^{-5}$	201	350	$9.8 \cdot 10^{-5}$	401

Table 8. Supervised RBFN performances for reaction rate prediction

Esterification reaction			Nitration reaction		
R. B. units	Error	parameters	R. B. units	Error	parameters
5	$2.9 \cdot 10^{-5}$	36	6	$1.5 \cdot 10^{-3}$	49
10	$1.2 \cdot 10^{-5}$	71	10	$2.9 \cdot 10^{-4}$	81
15	$8.1 \cdot 10^{-6}$	10^6	15	$7.9 \cdot 10^{-5}$	121

function in order to obtain positive generalisation properties. Hence patterns that should appertain to different classes fall within the same radial unit in the network and as the units have local character the outputs of the NN are close.

When the supervised method is used to obtain the reactor temperature profiles this effect is partially corrected. Because the widths are adapted to minimise the error, the local representation capacities of the units in the network are lost, i.e. one pattern could activate

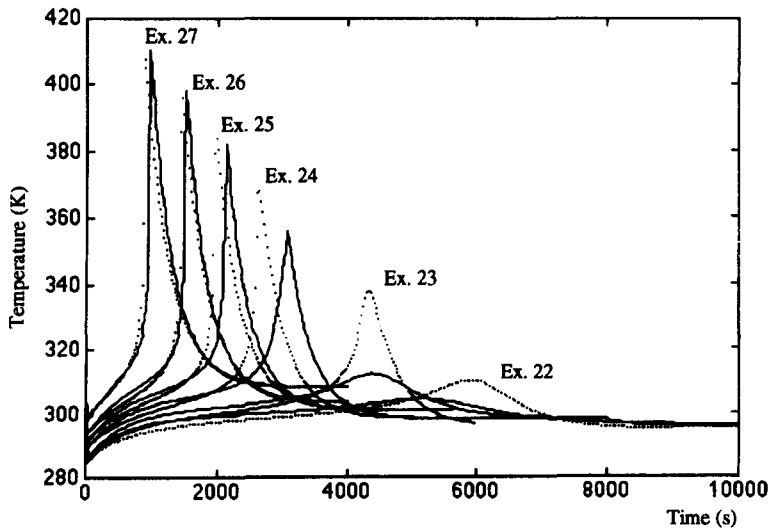


Fig. 5. Experimental (dotted lines) and simulated (continuous lines) reactor temperature profiles using the MNN model (5-10-1) for the isoperibolic experiments described in Fig. 3.

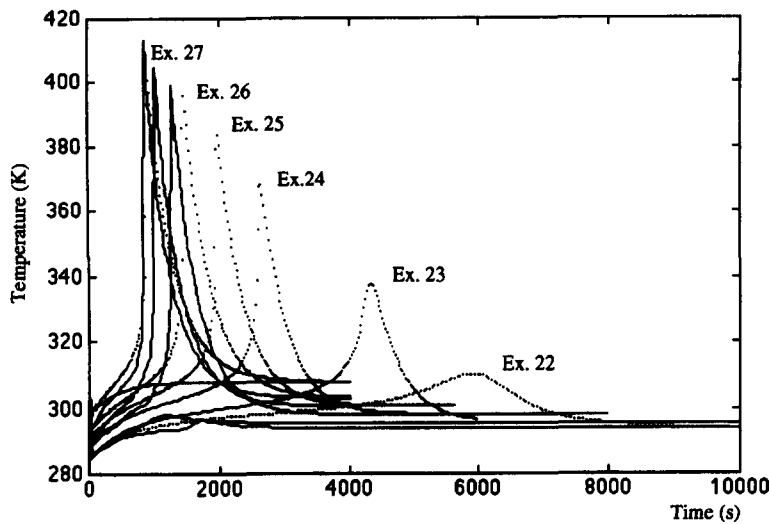


Fig. 6. Experimental (dotted lines) and simulated (continuous lines) reactor temperature profiles using the RBFN model (200 units) for the isoperibolic experiments described in Fig. 3.

different units, even units whose centre is far away. As can be seen in Fig. 7 the results are better than the approximation obtained with the RBFN model.

In the case of the nitration reaction some of the experiments that were not included in the training data set have been used to test the NNs reaction rate estimation. In Figs 8 and 9 the experimental and

estimated reaction rate results for the different networks are compared for two experiments.

Figure 8 (b) shows the reaction rate estimated with the RBFN model. In this case, the results are very similar to the results obtained with the MNN model because the input patterns for this experiment are well represented in the training data set and hence these patterns fall within

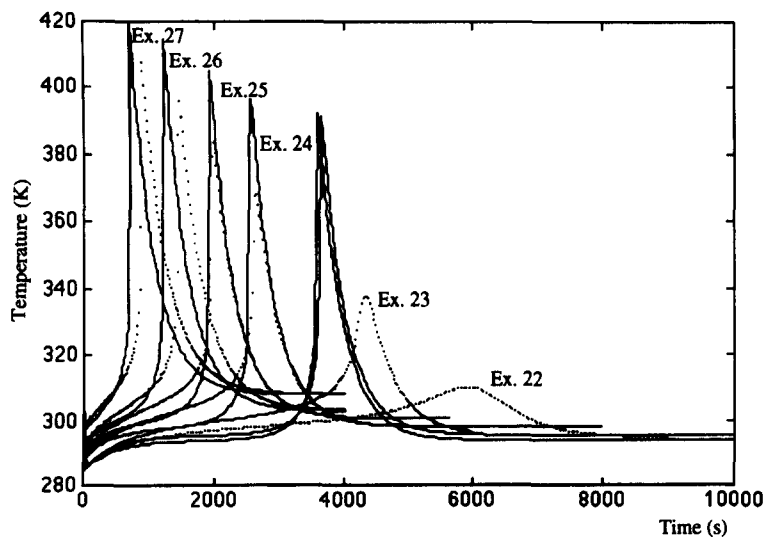


Fig. 7. Experimental (dotted lines) and simulated (continuous lines) reactor temperature profiles using the supervised method (10 units) for the isoperibolic experiments described in Fig. 3.

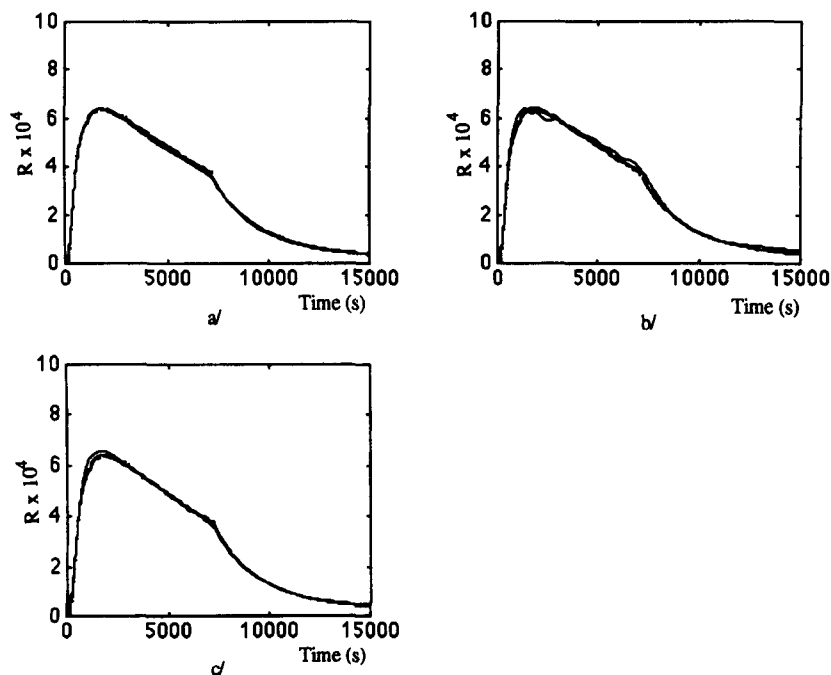


Fig. 8. Experimental (dotted lines) versus estimated reaction rate (mol/l.s) for semibatch toluene mononitration during the fast reaction regime experiments: (a) MNN (6-10-1), (b) RBFN (350 units), (c) Supervised method (15 units). x (sulphuric acid strength)=72.4 wt.%, $T_p=308.16$ K, feeding time=2 h and $Na=6.67$ s⁻¹.

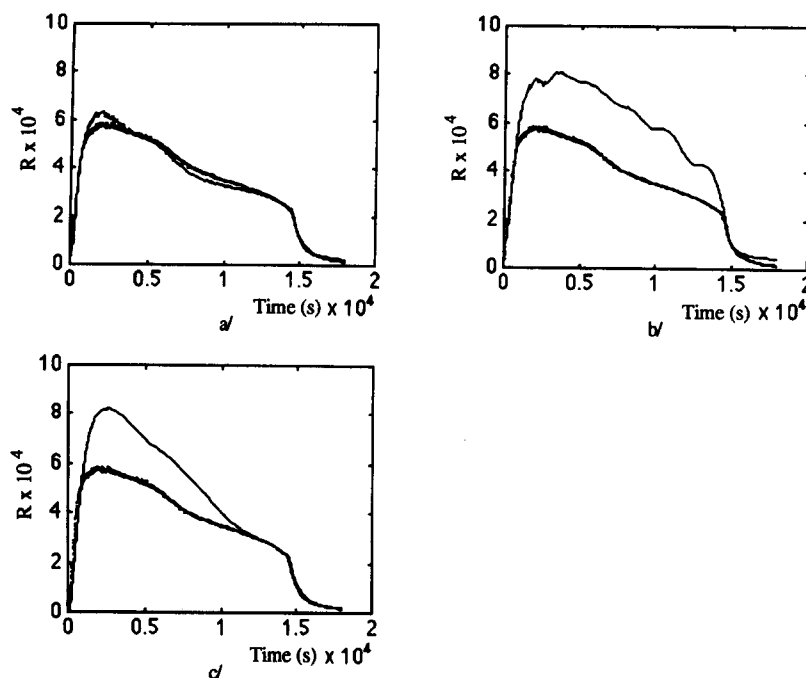


Fig. 9. Experimental (dotted lines) versus estimated reaction rate (mol/l.s) for semibatch toluene mononitration during the fast reaction regime experiments: (a) MNN (6-10-1), (b) RBFN (350 units), (c) Supervised method (15 units). x (sulphuric acid strength)=78.9 wt.%. $T_p=308.16$ K, feeding time=4 h and $N_a=10.0$ s $^{-1}$.

the right radial unit. However, the results for the other experiment (Fig. 9(b)) are less satisfactory. The situation is the same as that in the esterification case.

DISCUSSION

All the approximation methods have their own characteristics. Conditions under which one method is preferable to another are also known but no one method is absolutely the best. In this work, two types of neural network architectures have been implemented and tested to predict the reaction rate of two different systems. However, since polynomials, orthogonal functions and splines can also be used to approximate such reaction rates, it is not readily apparent as to why neural networks should be preferred.

During the past few years extensive studies have revealed that both MNN and RBFN have some advantages over conventional methods for approximation. For example, the polynomial or orthogonal series approximation are difficult to implement in hardware. In contrast, neural networks are readily implementable in hardware. Another important feature of the neural networks is that the networks contain only the linear operations of multiplication by a scalar constant and summation in addition to the nonlinear function activation which is known. This makes it mathematically attractive for the approximations of functions.

Assuming that neural networks are to be used for the

approximation of reaction rate, the immediate question that arises concerns the choice of the neural networks to be used, i.e., an MNN or an RBFN. In this section we will discuss the advantages and disadvantages for each model.

MNN have been applied with some success as regression models in a wide variety of chemical engineering problems. From fault classifiers (Hoskins and Himmelblau, 1988; Watanabe *et al.*, 1989) to model and control chemical reactors (Bhat and McAvoy, 1990; Zaldívar *et al.*, 1991). However, difficulties were encountered. The first limitation is in the determination of the required number of hidden nodes and the resulting problem in pruning redundant nodes or weights. Another disadvantage of MNN is that the parameters (weights) appear in a non-linear form, so the gradient method has to be used for their adjustment. This implies that the parameters can converge to local minima. Furthermore, the adjustment of a single parameter of the network affects the output globally. Hence, all the weights have to be adjusted simultaneously for each pattern. This reduces the effect of previous learning or in other words causes a rapid forgetfulness owing to seldom-seen input-output pairs. Both the above factors contribute to the slow convergence rate or slow learning. About 15 or 20 thousand iterations over the training set are necessary to stabilise the weights in the trained architectures.

Broomhead and Lowe (1988) and Moody and Darken (1989) considered, to improve network training time and

generalisation characteristics, the use of RBFN. The learning of an RBFN is faster than that of MNN. There are two factors that contribute to this phenomenon. The first is that owing to the fact that the parameters are linearly related to the error, convergence is guaranteed. The second factor is the local character of the radial basis functions in contrast to the global character of the sigmoidal functions. When an input pattern is presented to the network, only the hidden nodes whose centres are close to the input will be active. The magnitude of this activation will depend on the width of each Gaussian. Thus for each pattern only some weights will be adjusted and the convergence will be faster. Furthermore, local representation should provide better generalisation characteristics (Leonard and Kramer, 1990). However, owing to the high input space (the dimension is 5 in the esterification and 6 in the nitration case) to guarantee a good reaction rate approximation and generalisation properties with RBFN, the number of fitted parameters has to be much larger than in the approximation with multilayer neural network.

As we have said, to reduce the number of radial basis units it is possible to use the supervised method. This method yields high precision results, but places no architectural restrictions on parameters of the network. For instance, the widths are not restricted to remain small, so the local representation is not guaranteed. Furthermore, the learning in this method is a non-linear optimisation problem because the dependence of the output of the network with respect to these parameters is non-linear. Those characteristics are responsible for the loss of the fast convergence and local character proprieties of the RBFN model.

Owing to the similar performance learning between the supervised method and the MNN model it is interesting to compare them. In general the supervised network does not offer advantages over the standard multilayer network because the convergence is also slow and the number of parameters to adjust is not reduced. In some cases the results obtained with this kind of network, of comparable size with the multilayer network, are better. Note that in the esterification case the supervised method achieved better performance than the more traditional sigmoidal units (Table 5 and Table 7). However, it is not possible to conclude a general result because a good approximation depends on the Gaussian shape and we cannot guarantee how it will be using the gradient method to adapt its centers and its width.

CONCLUSIONS

The results obtained indicate that the MNN model is more adequate for determining kinetic functions of complex processes in contrast to the RBFN model.

Simulations carried out using both MNN and RBFN for the approximating functions $C(R^+, R)$ for small values of n ($n \leq 3$) clearly show the superiority of RBNN (White and Sofge, 1992). For large values of n the radial basis functions needed become excessive, and the location of the basis functions should be chosen with considerable care. Hence some prior information is needed. The MNN does not require such information. Consequently it is preferred when the dimensionality of the input space is high.

Furthermore, the comparison of the neural network approach with the traditional identification methods white or grey models showed that when the kinetic scheme is simple or when the amount of data available is limited other methods are suitable, since the knowledge about the exponential dependence on temperature and the function of concentration as a product of power of different compounds can be included. However, if a considerable amount of thermo-kinetic data is available, there is little knowledge about the kinetic pathway and/or the amount of time available to determine a reaction rate expression is limited, the neural network approach can be used at least as a first approximation for process design and optimisation.

NOTATION

a =Interfacial area (m^2/m^3)
 A =Frequency factor
 C =Concentration (mol/l)
 D =Diffusion coefficient (m^2/s)
 E_a =Activation Energy
 f =Sigmoidal activation function
 H_R =Acidity function
 k =Rate constant
 n =Molar hold-up
 p =Parameter of kinetic expression
 r =Reaction rate ($mol.l^{-1}.s^{-1}$)
 R =Overall conversion rate in heterogeneous systems ($mol.l^{-1}.s^{-1}$)
 R =Gas constant
 T =Temperature (K)
 u =Input vector to the network
 w =Weight vector
 x =Input vector to the neuron
 x =Molar fraction
 y =Output vector
 e =Volumetric fraction
 q =Threshold value
 j =Gaussian activation function
Subscripts and superscripts
 aq =Aqueous/acid phase
 PAh =Propionic Anhydride
 $2-B$ =2-butanol
 PAC =Propionic Acid
 BPr =2-butyl propionate
 i =Interphase
 Sul =Sulphuric Acid
 W =Water
 m =Reaction mixture
 e =Heat transfer fluid
 sp =Set point

Acknowledgements—The authors would like to express their appreciation to J. Lighthart, T. J. Snee, and C. Barcons for their contribution during the experimental part; and F. Panetsos and V. Calenbuhr for their valuable discussions on the subject.

REFERENCES

- Bhat N. and T. J. McAvoy, Use of neural nets for dynamic modelling and control of chemical process systems. *Comput. chem. Engng.* **14**, 573–583 (1990).
- Broomhead D. S. and D. Lowe, Multivariate functional interpolation and adaptive networks. *Complex Syst.* **2**, 321–355 (1988).
- Cybenko G., Approximation by superposition of sigmoidal functions. *Mathematics of control, Signals and Systems* **2**(4), 303–314 (1989).
- Deno N. C. and M. S. Newman, The racemization of alkyl hydrogen sulfates in sulfuric acid. *J.A.C.S.* **73**, 1920–1923 (1951).
- Dhanuka V. R., V. C. Malshe and S. B. Chandalia, Kinetics of the liquid phase esterification of carboxylic acids with alcohols in the presence of acid catalysts: Re-interpretation of published data. *Chem. Eng. Sci.* **32**, 551–556 (1977).
- Duda R. O. and P. E. Hart, *Pattern classification and scene analysis*. Wiley, New York (1973).
- Filippi C., J. L. Greffe, J. Bordet, J. Villermaux, J. L. Barnay, P. Bonte and C. Georgakakis, Tendency modelling of semibatch reactors for optimization and control. *Chem. Eng. Sci.* **41**, 913–920 (1986).
- Funahashi K., On the approximate realization of continuous mappings by neural networks. *Neural Networks* **2**, 183–192 (1989).
- Gallant A. R. and H. White, *There exists a neural network that does not make avoidable mistakes*. *IEEE Second International Conference on Neural Networks*, 657–664. SOS Printing, San Diego, CA, (1988).
- Haldar R. and D. P. Rao, Experimental studies on parametric sensitivity of a batch reactor. *Chem. Eng. Technol.* **15**, 34–38 (1992a).
- Haldar R. and D. P. Rao, Experimental studies on semibatch reactor parametric sensitivity. *Chem. Eng. Technol.* **15**, 39–43 (1992b).
- Hernández H., J. M. Zaldívar and C. Barcons, Development of a mathematical model and a numerical simulator for the analysis and optimization of batch reactors. *Comput. Chem. Engng.* **17S**, 45–50 (1993).
- Himmelblau D. M., C. R. Jones and K. R. Bischoff, Determination of rate constants for complex kinetics models. *Ind. Eng. Chem., Process Design and Development* **6**, 539–543 (1967).
- Hornik K., M. Stinchcombe and H. White, Multilayer feedforward networks are universal approximators. *Neural Networks* **2**, 359–366 (1989).
- Hoskins J. C. and D. M. Himmelblau, Artificial neural networks models of knowledge representation in chemical engineering. *Computers chem. Engng* **12**, 881–890 (1988).
- Kennard M. L. and A. Melsen, Mechanisms and kinetics of diethanolamine degradation. *Ind. Eng. Chem., Fundamentals* **24**, 129–140 (1985).
- Kohonen T., Self-organizing formation of topologically correct feature maps. *Biol. Cyber.* **43**, 59–69 (1982).
- Kohonen T., *Self-Organisation and associative memory*. Springer-Verlag, Berlin (1984).
- Kohonen T., An introduction to neural computing. *Neural Networks* **1**, 3 (1988).
- Leonard J. A. and M. A. Kramer, Limitations of the back-propagation approach to fault diagnosis and improvement with radial basis functions, presented at the AIChE Annual Meeting, Chicago (1990).
- Ljung L., *System identification- theory for the user*. Prentice Hall, Eng Cliffs, NJ, 303–337 (1987).
- March J., *Advanced organic chemistry: Reactions, mechanisms, and structure*, 4th Edition, Wiley and Sons, New York (1992).
- Molga E., C. Barcons and J. M. Zaldívar, Mononitration of toluene and quantitative determination of the isomer distribution by Gas Chromatography. *Afinidad* **50**, 15–20 (1993).
- Moody J. and C. J. Darken, Fast learning in networks of locally-tuned processing units. *Neural Computation* **1**, 281–294 (1989).
- Poggio T. and F. Girosi, Networks for approximation and learning. *Proc. IEEE* **78**(9), 1481–1486 (1990).
- Riesen R. and B. Grob, Reaction calorimetry in chemical process development. *Swiss Chem.* **7**, 39–43 (1985).
- Rochester C. H., *Acidity Functions*, Academic Press, London (1970).
- Rumelhart D., G. Hinton, and J. McClelland, Learning internal representations by error propagation, in D. Rumelhart, and J. McClelland (Ed.) *Parallel Distributed Processing*, MIT Press (1986).
- Schofield K., *Aromatic nitration*, Cambridge University Press, Cambridge (1980).
- Singh J., PHI-TEC: Enhanced vent sizing calorimeter, in *International Symposium on Runaway Reactions*, AIChE, 313–330 (1989).
- Snee T. J., C. Barcons, H. Hernández and J. M. Zaldívar, Characterisation of an exothermic reaction using adiabatic and isothermal calorimetry. *Journal of Thermal Analysis* **38**, 2729–2747 (1992).
- Snee T. J. and J. A. Hare, Development and application of a pilot scale facility for studying runaway exothermic reactions. *J. Loss Prev. Process Ind.* **5**, 46–54 (1992).
- Sprow F. B., Distribution of drop sizes produced in turbulent liquid–liquid dispersions. *Chem. Engng. Sci.* **22**, 435–442 (1967).
- Sreeramulu V. and P. B. Rao, Kinetics of esterification of isobutyl alcohol with palmitic acid using sulfuric acid catalyst. *Ind. Eng. Chem. Process Des. Develop.* **12**, 483–485 (1973).
- Tavlarides L. L., and M. Stamatoudis, The analysis of interphase reactions and mass transfer in liquid–liquid dispersions, in *Advances in Chemical Engineering*, vol 11, (T. B. Drew, G. R. Cokelet, J. W. Hoopes Jr. and T. Vermeulen, Eds), Academic Press, New York (1981).
- Troupe R. A. and K. Kobe, Kinetics of methanol–lactic acid reaction. *Ind. Eng. Chem.* **42**, 801–810 (1950).
- Venkatasubramanian V. and K. Chan, A neural network methodology for process fault diagnosis. *AIChE J.* **35**, 1993–2003 (1989).
- Villermaux J., Drop break-up and coalescence. Micromixing effect in liquid–liquid reactors, in *Multiphase chemical reactors VI Fundamentals*, (Rodriguez, A. E., J. M. Calo, J. M. and Sweed, N. H., Eds), Sijthoff et Noordhoff, Alphen aan den Rijn, 285–362 (1981).
- Watanabe K., I. Matsuura, M. Abe, M. Kubota and D. M. Himmelblau, Incipient fault diagnosis of chemical processes via artificial neural networks. *AIChE JI* **35**, 1803–1812 (1989).
- White D. A. and D. A. Sofge (Eds), *Handbook of Intelligent Control*, Multiscience Press, New York (1992).
- Yvernault T., Acid catalysis of the hydrolysis of acetic anhydride in acetic acid. *Compt. rend.* **241**, 485–487 (1955).
- Zaldívar J. M., H. Hernández and C. Barcons, Development of a mathematical model and numerical simulator for a reaction calorimeter. FISIM, RC1 version. Joint Research Center of C.E.C., *TN nú 1.90.109* Ispra, Italy (1990).
- Zaldívar J. M., F. Panetsos and H. Hernández, Control of batch reactors using neural networks. *Chemical Engineering and Processing* **31**, 173–180 (1991).
- Zaldívar J. M., C. Barcons H. Hernández, E. Molga and T. J. Snee, Modelling and optimization of semibatch toluene mononitration with mixed acid from performance and safety viewpoints. *Chem. Eng. Sci.* **47**, 2517–2522 (1992a).

- Zaldívar J. M., H. Hernández, C. Barcons and R. Nomen, Heat effects due to dilution during aromatic nitrations by mixed acid in batch conditions. *J. Thermal Anal.* **38**, 2575–2582 (1992b).
- Zaldívar J. M., M. A. Al-s, E. Molga, H. Hernández and K. R. Westerterp, Aromatic nitrations by mixed acid. Slow liquid-liquid reactions. *Chem. Engng. and Processing* **34**, 543–559 (1995).
- Zaldívar J. M., E. Molga, M. A. Al-s, H. Hernández and K. R. Westerterp, Aromatic nitrations by mixed acid. Fast liquid-liquid reactions. *Chem. Engng. and Processing* **35**, 91–105 (1995).