ELSEVIER

Thermochimica Acta 257 (1995) 127–137

thermochimica acta

# Hybrid connectionist modelling of the kinetics of thermal decomposition processes

C. Aldrich*, J.S.J. van Deventer

*Department of Chemical Engineering, University of Stellenbosch, Private Bag X5018, Stellenbosch, 7599, South Africa*

Received 31 May 1994; accepted 9 November 1994

**Abstract**

The kinetics of thermogravimetric decomposition reactions are often difficult to model explicitly, because model parameters can depend on process conditions in an ill-defined way. Implicit kinetic models, such as those based on neural nets, often require extensive data and are not usually suitable for extrapolation of experimental data. In this paper it is shown that by combining explicit phenomenological models with artificial neural nets, more accurate modelling and extrapolation of these types of processes can be achieved.

*Keywords:* Experimental; Hybrid connectionist modelling; Kinetics; TGA

## 1. Introduction

Numerous mathematical methods have been proposed for the determination of the kinetics of solid state reactions from isothermal or non-isothermal thermogravimetric analyses. These methods are typically based on the use of a general rate equation of the form

$$d\alpha/dt = kf(\alpha) = k_0 e^{-E/RT} f(\alpha) \quad (1)$$

or in terms of temperature

$$d\alpha/dT = kf(\alpha)/\beta = k_0 e^{-E/RT} f(\alpha)/\beta \quad (2)$$

where $\alpha$ denotes some measure of the mass or concentration of the solid at time $t$, $\beta$ the rate of heating ($\beta = dT/dt$) and $k$ a rate constant. This rate constant $k$ is typically related

* Corresponding author.

0040-6031/95/$09.50 © 1995 – Elsevier Science B.V. All rights reserved
*SSDI* 0040-6031(94)02201-I

to the temperature via an Arrhenius function $k = k_0 e^{-E/RT}$, with a pre-exponential factor $k_0$ and an activation energy $E$.

Eqs. (1) and (2) can subsequently be used in either a differential or an integral form (Eqs. (3) and (4) to describe the mechanism and kinetics of reactions

$$g(\alpha) = kt = k_0 e^{-E/RT} t \tag{3}$$

$$g(\alpha) = k_0 RT^2/\beta E[1 - 2RT/E]e^{-E/RT} t \tag{4}$$

This is usually accomplished by plotting the logarithms of either of these equations (Eqs. (1) or (3), or in terms of temperature, Eqs. (2) or (4)) in order to evaluate the kinetic parameters $k$, $k_0$ and $E$. The mechanism of the reaction is indicated by the functional form of $\alpha$, i.e. $f(\alpha)$ or its integral $g(\alpha)$.

Unforturnately these methods are often prone to significant shortcomings, in that by plotting the differential or integral forms $f(\alpha)$ or $g(\alpha)$ against the reciprocal temperature $1/T$, the functional forms are linearised to such an extent that the evaluation of a particular reaction mechanism become seriously compromised.

Some strategies have been suggested [1] to overcome these problems by determining analytical curves for all possible forms of $f(\alpha)$ and $g(\alpha)$ and comparing them with the experimental data. This approach can constitute a very elaborate exercise, apart from difficulties in obtaining sufficient experimental data for such a comprehensive analysis. Even when it is possible to determine the kinetics and mechanism of a solid state reaction without undue difficulty [1], these techniques do not lend themselves to convenient analysis of the multiple or complicated reactions often found in industry.

Hashimoto et al. [2], for example, showed that the kinetics of the thermal regeneration of activated carbon loaded with *p*-nitrophenol and a surface-active agent depended on multiple first-order reactions, each with a different activation energy and frequency factor. In complex multi-mechanism reactions, where mechanisms are superimposed, it is difficult to distinguish between these individual mechanisms [3]. These processes can be modelled by techniques such as multiple adaptive regression splines (MARS) [4] or artificial neural nets, which do not require explicit specification of a process model. Unfortunately these implicit methods often require extensive data to construct an accurate model of the process and are typically not suitable for extrapolation.

In this paper a hybrid approach to the modelling of the kinetics of thermal decomposition processes is discussed. The method is based on the use of simple back-propagation neural nets in conjunction with available fundamental knowledge of the reaction kinetics. As a consequence the technique requires relatively few data and is capable of significantly better extrapolation of the experimental data.

## 2. Connectionist systems

A number of books and a considerable volume of literature on artificial neural networks have appeared in recent years and only a very brief overview is provided in this paper [5–7].

Connectionist systems or artificial neural nets consist of a large number of fundamentally primitive computational elements connected to each other on a massive scale.

These computational elements are usually arranged in a layered structure which consists of an input and an output layer, and also possibly one or more hidden layers, as shown in Fig. 1. Unlike input and output layers, hidden layers are not connected to the information environment in which the net is operating. In feedforward nets, such as those used in this investigation, information is passed from the input layer to successive hidden layers (if present) and the output layer in the net. During the process the computational elements or nodes in the layers of the net transform the weighted sums of all their inputs (the potentials of the computational elements) by means of transfer functions which typically map the potentials to smaller domains than that of the inputs, as indicated in Fig. 2

$$z_i(t+1) = \phi[\textstyle\sum_j w_{ij} z_j(t) - \Theta_i] \tag{5}$$

where $z_i(t)$ is the output of the $i$th process node at time $t$ (or the $t$th iteration), $w_{ij}$ is the weight or connection between nodes $i$ and $j$, and $\Theta_i$ is the bias of the $i$th node.

The form of the transfer function $\phi$ may vary, but it could be a linear, step or sigmoidal transfer function, among others, with a domain typically much smaller than that of the potential, i.e. the sum of the weighted inputs, of the process unit.

The weights which characterise the connections between process elements can be modified by various suitable training algorithms [8]

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij} \tag{6}$$

where

$$w_{ij} = -\tau(\partial\varepsilon/\partial w_{ij}) \tag{7}$$

and $\tau$ is the learning rate of the net, and $\varepsilon$ an error criterion. These training algorithms are designed to minimise the mean square error between the desired and the actual

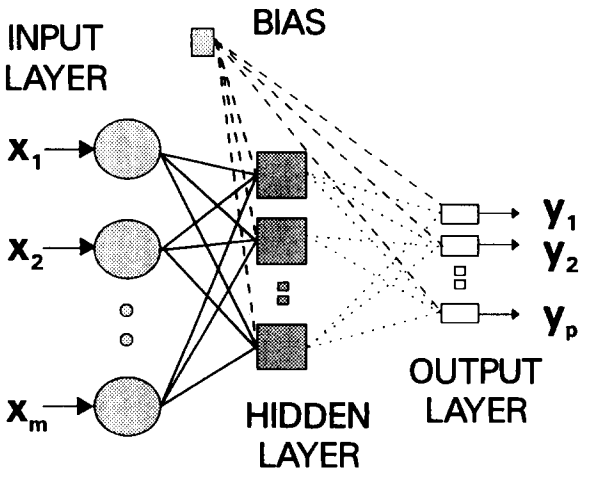


Fig. 1. General structure of a back-propagation neural net with one hidden layer.

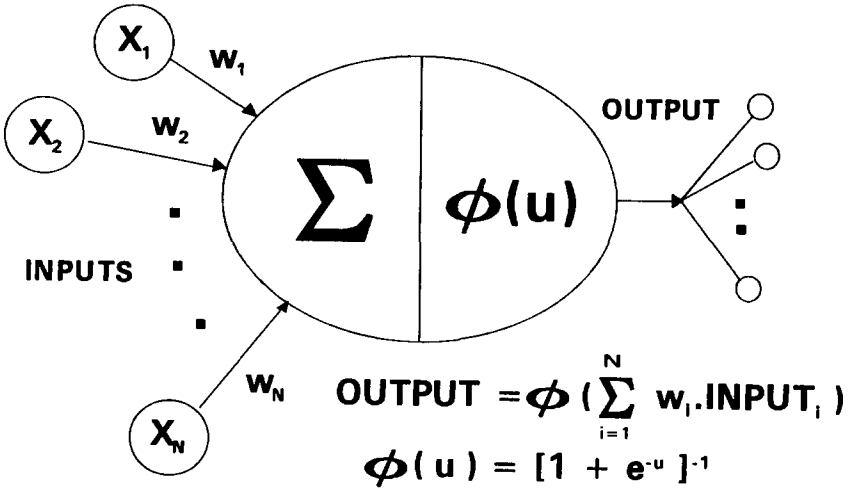


Fig. 2. Sigmoidal process unit of a back-propagation neural net.

output of the net [6]

$$\varepsilon = \frac{1}{2}\sum(d_j - z_j)^2 \tag{8}$$

where $d_j$ is the desired output of the net and $z_j$ the actual output.

By presenting a neural net with a representative range of data exemplifying a functional relationship, the net is thus able to form a generalised internal representation of this functional relationship. The trained net can consequently be used to predict mapped output from previously unseen inputs, regardless of the complexity of the mapping.

Artificial neural nets have been applied with notable success to, among others, the modelling of mineral and chemical processes, process control, fault diagnosis, as well as to classification problems of various types [9–11].

## 3. Modelling methodology

The modelling methodology is similar to the approaches followed by Psichogios and Ungar [12] and Reuter and Bernhard [3]. Consider the dynamic system represented by Eqs. (9) and (10)

$$\mathrm{d}\boldsymbol{x}/\mathrm{d}t = \boldsymbol{\Phi}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{p}) \tag{9}$$

$$\boldsymbol{p} = \boldsymbol{\phi}(\boldsymbol{x}, \boldsymbol{u}) \tag{10}$$

where $\boldsymbol{x}$ denotes the state vector of the system, $\boldsymbol{u}$ the control and $\boldsymbol{p}$ a vector of process parameters. The functional relationship between $\boldsymbol{p}$ and the state and control variables $\boldsymbol{x}$ and $\boldsymbol{u}$ is often difficult to derive from first principles, especially with regard to complex

processes such as the kinetics of thermal decomposition. Despite the difficulty in relating the process parameters to the state and control variables, knowledge of these parameters under a wide range of operating conditions is crucial to the efficient reactor design and operation. The neural net can either be used to represent the functional relationship $\boldsymbol{\phi}(\boldsymbol{x},\boldsymbol{u},\boldsymbol{p})$ or else available knowledge of this relationship can be retained by using a neural net to represent $\boldsymbol{p} = \boldsymbol{\phi}(\boldsymbol{x},\boldsymbol{u})$. This approach allows the modelling of complex processes without having to identify constituent or multiple reaction mechanisms. The technique is discussed in more detail by means of examples below.

### 3.1. *Example 1: Decomposition of phenol on activated carbon [13]*

In the first example, the use of a neural net to represent the non-isothermal decomposition of phenol adsorbed on activated carbon is illustrated. This process has previously been modelled by van Deventer and Camby [13] by means of a two-stage kinetic model. The phenol (A) first decomposes irreversibly into a non-volatile fragment (B) and a volatile fragment (C). This decompositon step can be written as

$$\mathrm{A(ads)} \rightarrow b\mathrm{B(ads)} + c\mathrm{C(ads)} \tag{11}$$

$$\mathrm{C(ads)} \rightarrow \mathrm{C(g)} \tag{12}$$

The total amount of phenol ($q$) adsorbed at any time is $q = q_A + q_B$, provided that the second reaction (Eq. (12)) is rapid compared to the first (Eq. (11)).

The decomposition reaction is assumed to be first order

$$-\mathrm{d}(q - q_B^0)/\mathrm{d}t = k_A(q - q_B^0) \tag{13}$$

with $q_B^0$ being the initial mass of adsorbed intermediate product per unit mass of initial virgin carbon. The reaction constant $k_A$ is, moreover, assumed to follow an Arrhenius relationship, i.e.

$$k_A = k_A^0 \exp(-E_A/R_0 T) \tag{14}$$

At a higher temperature, the intermediate product (B) decomposes into cracked products C′ that desorb easily, as well as an adsorbed char residue (R)

$$b\mathrm{B(ads)} \rightarrow r\mathrm{R(ads)} + c'\mathrm{C'(ads)} \tag{15}$$

$$\mathrm{C'(ads)} \rightarrow \mathrm{C'(g)} \tag{16}$$

The kinetics of the decomposition of (B) can similarly be expressed as

$$-\mathrm{d}(q - q_R^\infty)/\mathrm{d}t = k_B(q - q_R^\infty) \tag{17}$$

with $k_B$ also following an Arrhenius relationship of the form

$$k_B = k_B^0 \exp(-E_B/R_0 T) \tag{18}$$

Instead of further refinement of the model as discussed by van Deventer and Camby [13], a neural net can be used to model the kinetics of the decomposition process by considering the ratio of the total mass of adsorbate per unit mass of initial virgin carbon

to the initial mass of adsorbed phenol per unit mass of initial virgin carbon ($q/q_A^0$) as a function of the initial loading ($q_A^0$) and the time $t$, i.e. $q/q_A^0 = f_{NN}(q_A^0, t)$.

*3.1.1. Neural net model of decomposition kinetics*

By using a simple back-propagation neural net, the relationship $q/q_A^0 = f(q_A^0, T)$ could be represented without the need for explicit knowledge of the function $f_{NN}$. The net consisted of an input layer with two nodes (one for each input variable), a hidden layer with three sigmoidal nodes and a single node sigmoidal output layer (for the response variable ($q/q_A^0$), similar to the net shown in Fig. 1.

The experimental data were divided into a training and a test set, each comprised of exemplars of the form $\{q_A^0, t | q/q_A^0\}$. The net was subsequently trained by repeatedly presenting it with the data in the training set, until the root-mean-square (RMS) error between the predicted $q/q_A^0$ values and the experimental $q/q_A^0$ values was minimised, as shown in Fig. 3 and explained previously.

After convergence, the performance of the net was evaluated against the exemplars in the test set to ensure that the net had generalised the underlying trends in the data (instead of learning the data themselves). The results for activated carbon with an initial loading of 82.5 mg (g phenol)$^{-1}$ are shown in Fig. 4. These results are typical for other phenol loadings as well, and as can be seen from this figure, the net was able to form an accurate representation of the kinetics of the process, with an average absolute error of less than 0.3%.

*3.2. Example 2: Calcination of limestone*

The calcination of limestone has been studied extensively, due to its importance in industrial processes and pollution control. The reaction (Eq. (19)) involves the en-

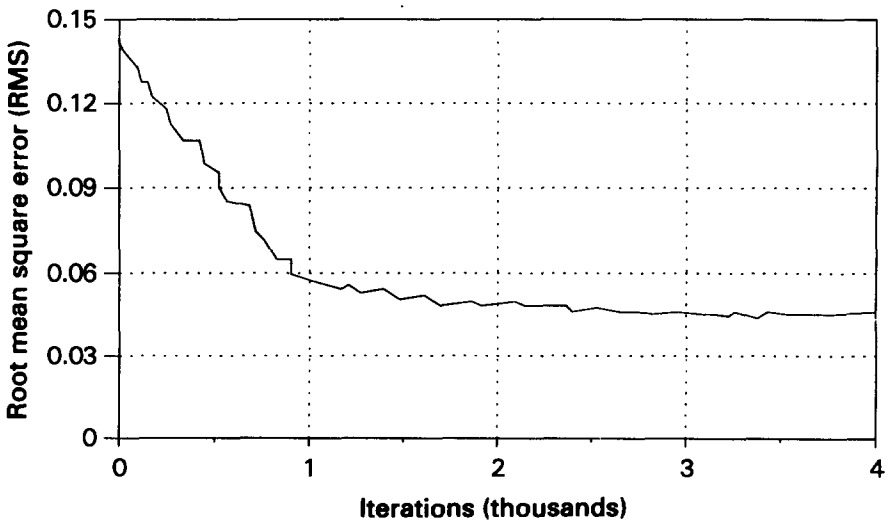


Fig. 3. Root mean square (RMS) error in output of neural net during training (Example 1).

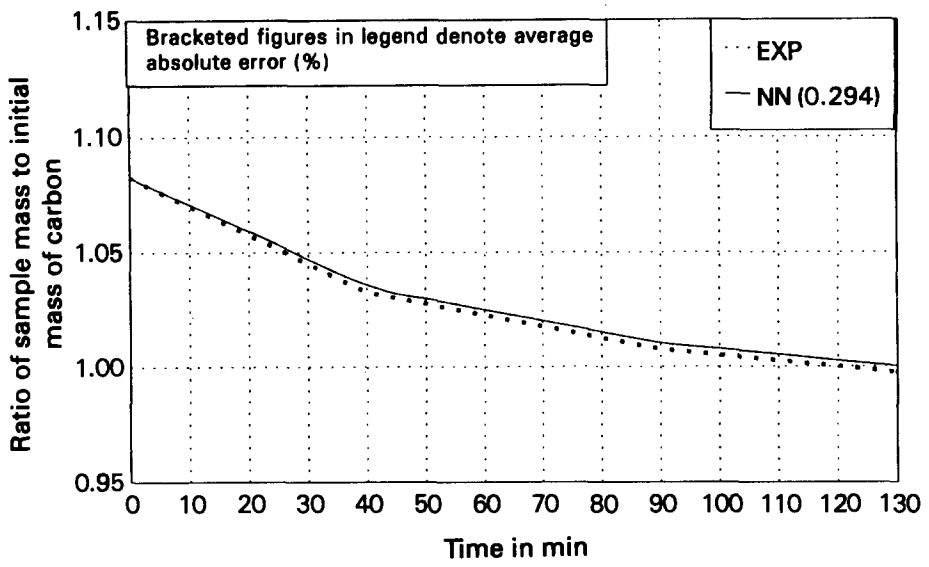


Fig. 4. Neural net model of decomposition of phenol on activated carbon.

dothermic decomposition of solid calcium carbonate into solid calcium oxide and carbon dioxide gas [14]

$$CaCO_3(s) \rightarrow CaO(s) + CO_2(g) \tag{19}$$

The calcination reaction depends on the reaction temperature, the size, shape and pore structure of the calcium carbonate particles, the background pressure of the carbon dioxide, the presence of impurities in the calcium carbonate, etc. [15].

The process kinetics are represented by

$$-d[CaCO_3]/dt = f([CaCO_3], T, d) \tag{20}$$

where $[CaCO_3]$ denotes the concentration of the calcium carbonate at time $t$, $T$ the temperature at which the reaction is taking place, and $d$ the average particle size of the calcium carbonate. The functional relationship between the rate of decomposition of the calcium carbonate and the process conditions was modelled by a sigmoidal back-propagation neural net with one hidden layer. The input layer consisted of three nodes (one each for the calcium carbonate concentration $[CaCO_3]$, the temperature $T$ and the average particle size $d$, while the output layer of the net had a single node corresponding to the decomposition rate of the solid ($-d[CaCO_3]/dt$).

The training and test sets of exemplars were constructed from laboratory data obtained in experiments with six different particle sizes, namely 3.5–4.0, 4.0–5.6, 5.6–6.7, 6.7–11.2, 11.2–16.0 and 16.0–25.0 mm, at three different temperatures, 1000, 1100 and 1200°C. The training set consisted of 400 exemplars, while the test set consisted of 100 exemplars. The net converged rapidly (within 120 000–15 000 iterations, as shown by the solid lines in Fig. 5) and predicted the test data with an average absolute error of less than 8%, as shown by the solid lines in Fig. 6.

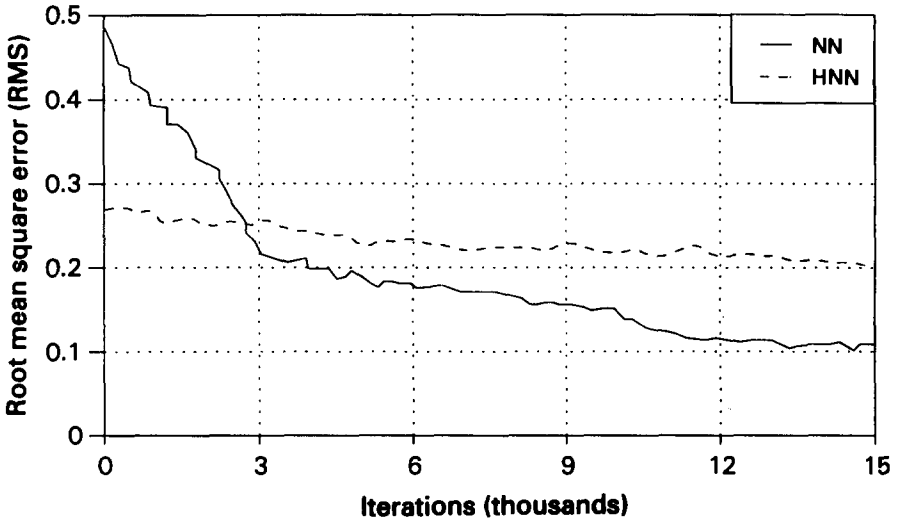


Fig. 5. Root mean square (RMS) error in output of neural net (NN) and hybrid neural net (HNN) model of the calcination of calcium carbonate during training (Example 2).

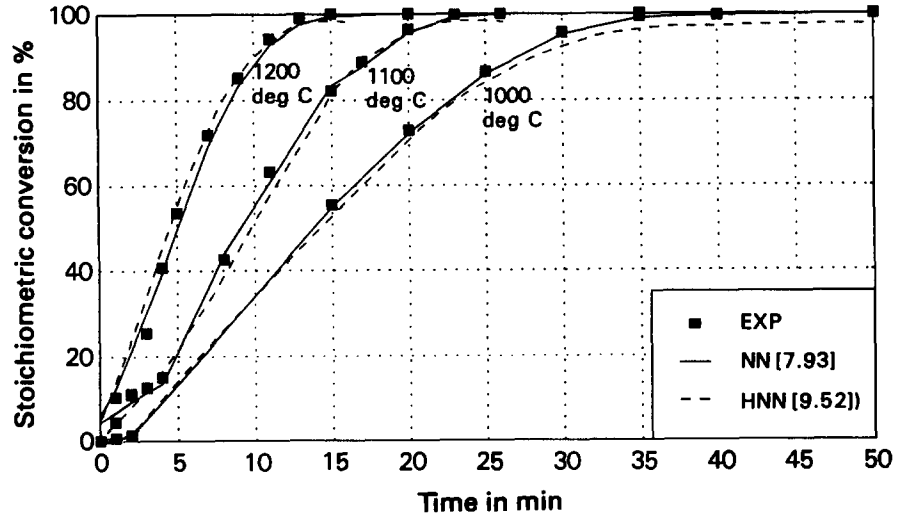


Fig. 6. Neural net (NN) and hybrid neural net (HNN) modelling of the calcination of calcium carbonate (figures in square brackets indicate average absolute percentage errors).

Instead of using a neural net to model the calcination process *in toto*, the net can be used to model the dependency of the kinetics on the calcium carbonate concentration and particle size, while an Arrhenius relationship can be used to model the effect of temperature, that is

$$-\mathrm{d}[\mathrm{CaCO_3}]/\mathrm{d}t = f_{\mathrm{NN}}([\mathrm{CaCO_3}], d)\mathrm{e}^{-E/RT} \tag{21}$$

The net had the same structure as those used previously, except that it had an input layer with two nodes (one each for the calcium carbonate concentration [$CaCO_3$] and the average particle size $d$). During training the net converged fairly rapidly (within 15 000 iterations), as indicated by the broken lines in Fig. 5 and was capable of forming an accurate (average absolute error of less than 10%) internal representation of the relationship $f_{NN}([CaCO_3], d)$, as shown by the broken lines in Fig. 6.

### 3.3. Example 3: Kinetics of the thermal regeneration of spent activated carbon [16]

Van Deventer and Camby studied the regeneration of activated carbon from a gold adsorption plant, as well as carbon loaded with phenol in a steam–nitrogen atmosphere in a fluidised bed. They derived a kinetic model for the mass loss of the loaded carbon $M$, as a function of time $t$ and temperature $T$. As before, this relationship can be modelled in full by a neural net, that is $M = f_{NN}(t, T)$. A simple back-propagation neural net with a hidden layer containing three sigmoidal process units was trained on experimental data representing the process at temperatures below 800°C. These results are shown as solid lines in Fig. 7. The net was subsequently used to extrapolate the experimental data at 900°C, the results of which are shown as dotted lines in Fig. 7. As can be seen from this figure, the net was not able to extrapolate the process kinetics at 900°C very accurately (the average absolute error was 18.09%, compared to an average absolute error of 5.93% for the solid-line curves).

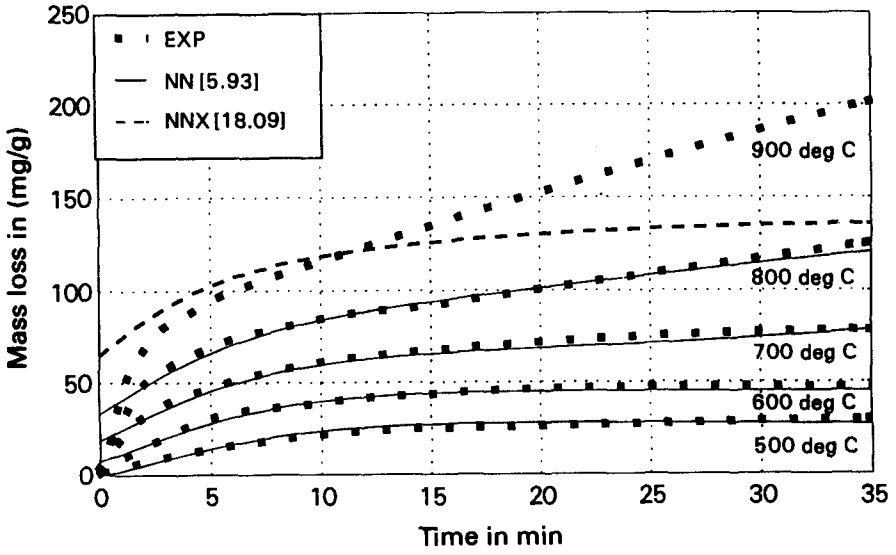


Fig. 7. Neural net (NN) model of the regeneration of activated carbon loaded with phenol. Broken lines indicate experimental data (EXP), solid lines indicate the prediction of the neural net (NN) in process regions represented by its training data base, and dotted lines (NNX) indicate prediction in process regions not represented by its training data base. Figures in square brackets denote average absolute percentage errors.

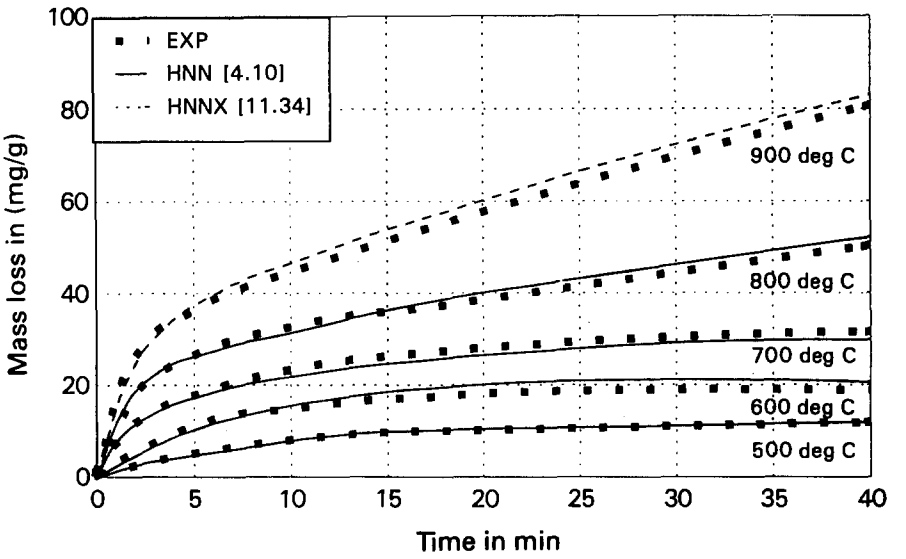


Fig. 8. Hybrid neural net (HNN) model of the regeneration of activated carbon loaded with phenol. Broken lines indicate experimental data (EXP), solid lines indicate the prediction of the hybrid neural net (HNN) in process regions represented by its training data base, and dotted lines indicate prediction of the net (HNNX) in process regions not represented by its training data base. Figures in square brackets denote average absolute percentage errors.

A hybrid model was subsequently constructed. This model consisted of a simple net with one input node, a three-node sigmoidal hidden layer and a single output node, combined with an Arrhenius-type equation to model the effect of temperature. The net was trained to represent the relationship between the mass loss $M$ and time $t$ at a temperature of 500°C only, while the temperature was modelled explicitly. The results of this model are shown in Fig. 8. The net was able to predict the experimental data at 500°C accurately (with an average absolute error of 4.10%). The ability of the net to predict the data at other temperatures was hampered somewhat by inaccuracies in the explicit temperature relationship, but it was nonetheless capable of good all-round performance, as shown in Fig. 8. The net could predict the data at 900°C with an average absolute error of less than 11.4% compared to the 18.09% of the previous neural net model. While the performance of the hybrid net should not deteriorate significantly at higher temperatures, that of the larger net $M = f_{NN}(t, T)$ can be expected to deteriorate dramatically.

## 4. Conclusions

The examples discussed in this paper merely serve as case studies to elucidate the use of neural nets to model the kinetics of decomposition and other similar processes. In

principle these techniques can be applied to numerous other types of mass transfer processes, as long as sufficient data are available to describe these phenomena. The sufficiency of the data depends on the complexity of the process, i.e. the dimensionality of the input space, as well as on the degree of interaction between the input variables.

The hybrid neural net method is a general approach which eliminates the need for the development of complicated fundamental models. An advantage of this approach is that the net already contains a partial model of the process, so that relatively few data are required to model the process parameters of the system. The hybrid modelling technique is based on the assumption that the partial model is reasonably accurate. If that is not the case, or if the partial model contains a large number of complex parameters, the hybrid net may not afford a significant advantage over the use of standard neural net methods.

## References

[1] A.B. Phadnis and V.V. Deshpande, Thermochim. Acta, 62 (1983) 361.
[2] K. Hashimoto, K. Miura and T. Watanabe, AIChE J., 28 (1982) 737.
[3] M.A. Reuter and D. Bernhard, Miner. Eng., 7(1) (1994) 1–19.
[4] J.H. Friedman, Ann. Statistics, 19 (1991) 1.
[5] R. Hecht-Nielsen, Neurocomputing, Addison-Wesley Publishing Company, 1990.
[6] R.P. Lippmann, IEEE ASP Mag., 35 (1987) 4.
[7] P.D. Wasserman, Neural Computing: Theory and Practice, Von Nostrand-Reinhold, 1989.
[8] D.E. Rumelhart, G.E. Hinton and R.J. Williams, Parallel Distributed Processing, 1 (1986) 318.
[9] N.V. Bhat and T.J. McAvoy, Comput. Chem. Eng., 14(4/5) (1990) 573.
[10] D.J. Cooper, L. Megan and R. F. Hinde, Jr., AIChE J., 38(1) (1992) 41.
[11] J.C. Hoskins, K.N. Kaliyur and D.M. Himmelblau, AIChE J., 37(1) (1991) 137.
[12] D.C. Psichogios and L.H. Ungar, AIChE J.,, 38(10) (1992) 1499.
[13] J.S.J. van Deventer and B.S. Camby, Thermochim. Acta, 104 (1986) 61–70.
[14] T.R. Rao, Can. J. Chem. Eng., 71 (1993) 481.
[15] S. Keener and S.-J. Khang, Chem. Eng. Commun., 117 (1992) 279.
[16] J.S.J. van Deventer and B.S. Camby, Thermochim. Acta, 136 (1988) 179.