ELSEVIER

Chemical Engineering and Processing 39 (2000) 323–334

Chemical Engineering and Processing

www.elsevier.com/locate/cep

Neural networks for modelling of chemical reaction systems with complex kinetics: oxidation of 2-octanol with nitric acid

E.J. Molga [a,*], B.A.A. van Woezik [b], K.R. Westerterp [b]

[a] *Chemical and Process Engineering Department, Warsaw University of Technology, ul. Warynskiego 1, 00-645 Warsaw, Poland*
[b] *Department of Chemical Engineering, Chemical Reaction Engineering Laboratories, Twente University, P.O. Box 2177, 7500 AE Enschede, The Netherlands*

Received 2 June 1999; received in revised form 13 September 1999; accepted 13 September 1999

Abstract

Application of neural networks to model the conversion rates of a heterogeneous oxidation reaction has been investigated — oxidation of 2-octanol with nitric acid has been considered as a case study. Due to a more complex and unknown kinetics of the investigated reaction the proposed approach based on application of neural networks is an efficient and accurate tool to solve modelling problems. The elaborated hybrid model as well as the modelling procedure have been described in detail. Learning data used to train the networks have been extracted from the experimental results obtained in an extensive investigation programme performed with a RC1 Mettler-Toledo reaction calorimeter. The influence of different operating conditions on the accuracy and flexibility of the obtained results has been investigated and discussed. It has been found that with the proposed approach the behaviour of a semi-batch reactor, i.e. its concentration and heat flow time profiles, can be predicted successfully within a singular series of experiments; however, the generalisation of the neural network approach to all series of experiments was impossible.

Keywords: Neural networks; Reaction kinetics; Liquid-liquid oxidation

1. Introduction

From performance and safety points of view an efficient modelling of chemical processes is needed for fast, reliable and accurate predictions of the reactor behaviour to optimise the process. But the development of a reactor model becomes often a laborious and expensive stage in the whole design or optimisation procedure. Usually the development of the reaction kinetics is a highly complicated task, particularly for complex reacting systems — as multiphase systems, polymerisation reactions or catalytic reactions. In such systems active, non-stable intermediates may influence significantly the reaction progress, but usually the exact reaction mechanism is not known as a result of the difficulties with the identification and quantitative determination of these intermediates.

The application of neural networks seems to be a promising tool to solve modelling problems for the cases where as a result of insufficient knowledge the governing mechanisms can not be formulated. Recently the application of neural networks has become more popular also in chemical engineering, particularly for control and fault diagnosis; the use of neural networks for modelling of chemical processes is relatively new — e.g. see Ref. [1].

Two main modelling strategies employing neural networks may be distinguished: the first one called 'the black-box approach', when the entire process is represented with the appropriate neural net, and 'the hybrid approach', which is a combination of both — traditional modelling of the process and a neural network representing the less known phenomena of the process. In the former case a generalisation of the obtained results to other systems, e.g. differing in size or operating conditions, is hardly possible, while the latter approach gives an exciting opportunity for knowledge generalisation provided the reaction kinetics are repre-

* Corresponding author. Tel.: +48-22-660-6293; fax: +48-22-825-1440.
E-mail address: molga@ichip.pw.edu.pl (E.J. Molga)

0255-2701/00/$ - see front matter
PII: S0255-2701(99)00093-8

sented with the trained neural network. This hybrid approach, since its introduction by Psichogios and Ungar [2], has been found to be smart and efficient to model complex reacting systems with unknown kinetics — e.g. see Ref. [3] where an application of this approach to model the pyrolysis of ethane is presented.

In our previous papers, the kinetics of a homogeneous catalytic esterification and heterogeneous aromatic nitration [4], a catalytic gas-liquid-solid hydrogenation [5] and a catalytic liquid-liquid hydrolysis [6] have been successfully modelled with neural networks. Neural networks have the ability to fit arbitrary complex non-linear relationships [7] and may thus be well suited for the approximation of kinetic expressions.

In this paper an application of neural networks to model the conversion rate of the oxidation of 2-octanol with nitric acid has been investigated. The flexibility and accuracy of the neural network approach, depending on the net architecture and treatment of available experimental data, have been studied. Some conclusions, useful for other studies, are formulated and discussed.

2. Experimental installation and procedure

The oxidation of 2-octanol with nitric acid has been studied in a reaction calorimeter (RC1). This apparatus provides an accurate measurement of the heat removal by cooling, which can be used to determine the rate of heat generation in the reaction mass, see e.g. Ref. [8].

The experimental set-up is shown in Fig. 1. The RC1 *(1)* forms the basis of the experimental set-up. It is equipped with a jacketed 1-l glass vessel of the type SV01. Baffles made of glass are placed in the reactor vessel. The reactor content is stirred by a propeller stirrer. The stirring speed is adjusted to 700 rpm. For further details and drawings of the RC1 see Refs [8,9].

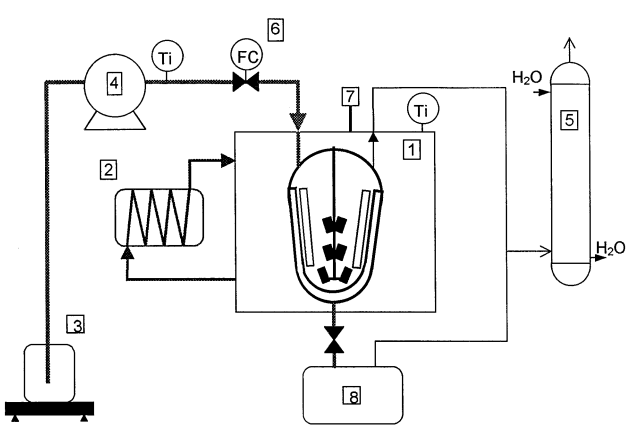

Fig. 1. Simplified flow-sheet of the experimental set-up. Ti, temperature indicator; FC, flow controller. See text for further details.

The reactor is operated in the semi-batch mode under isothermal conditions. To operate below room temperature an external cooling device *(2)* of the type Haake KT40 has been installed. The reactor is initially loaded with 0.4 kg of a 60 wt.% HNO_3-solution. Before the experiment is started, a small amount (0.1 g) of $NaNO_2$ is added as initiator. When the temperature of the reactor has reached a constant value the feeding system is started by activating the Mettler dosing controller RD10 *(6)*. The feed consists of pure 2-octanol in the supply vessel, which is located on a balance of the type Mettler pm3000 *(3)*. The organic compound is fed to the reactor by a Verder cogwheel pump *(4)* with a constant feed rate in the range of 0.05–0.4 kg/h. The nitric acid solution and the organic solution are immiscible and form a heterogeneous dispersion in the reactor. The nitric acid is used in excess and forms the continuous phase during the whole experiment. During the oxidation of 2-octanol, NO_X-gases are formed, which are washed in a scrubber *(5)* with water. After addition of 0.1 kg 2-octanol the dosing is automatically stopped by the RD10 *(6)*. The experiment is continued after dosing and stopped as soon as the maximum amount of 2-octanone has been formed. This maximum can be found by analysis of the organic phase, which is explained below.

During an experiment 4–10 samples of the dispersion are taken with a syringe as indicated by *(7)* in Fig. 1. The dispersion in the syringe separates directly in two phases. The concentration of acids in the aqueous phase is determined by titration with a 0.1 M NaOH-solution. The organic phase is first stabilized by washing it with water to remove small amounts of nitric acid and unstable nitro-compounds. The organic phase is then analyzed by GC to determine the concentration 2-octanol, 2-octanone, and carboxilic acids.

During all runs the temperatures are measured of the reactor content, the cooling oil, the feed and of the surroundings. The temperatures and the measured mass of the 2-octanol on the balance are monitored and stored by a computer. In case of an emergency, the computer opens the emergency cooling and the electric valve in the reactor bottom. The dumped reactor content is quenched in ice *(8)*.

The reaction system is investigated in the temperature range of 0–40°C for dosing times of 900–7200 s, see Table 1 where all performed experiments are listed.

3. Reaction pathways and observed phenomena

Oxidation reactions with nitric acid in general are very complex and usually several intermediates are formed, see Ref. [10]. The major uncertainty in describing this type of oxidation reaction is its mechanism, which is unknown. Different reacting species are pro-

Table 1
Experiments performed (data for learning have been extracted from runs written in **bold**)

Series	Dosing time, t_dos [s]	Reactor temperature, T_R [C]	Experiment name
I	7200	0	E_192
		0	**E_204**
		5	E_196
		10	**E_191**
		15	E_194
		20	**E_190**
		25	E_193
		40	**E_195**
II	3600	**0**	**E_178**
		10	E_169
		10	**E_229**
		15	E_186
		20	**E_181**
		25	E_189
		30	**E_170**
		40	**E_175**
III	1800	**0**	**E_179**
		10	E_173
		10	**E_228**
		15	E_185
		15	E_225
		20	**E_223**
		20	E_182
		25	E_188
		30	**E_171**
		40	**E_176**
IV	900	**0**	**E_180**
		10	**E_174**
		15	**E_184**
		15	E_224
		20	E_183
		20	**E_222**
		25	E_187
		25	E_221
		30	E_172
		30	**E_220**

posed like N_2O_4 by Ref. [11], NO^+ by Ref. [12] and NO_2^+ by Ref. [13]. This uncertainty makes elucidation of the real pathway extremely difficult or even impossible. The oxidation of cyclohexanol with nitric acid, which is very similar to the oxidation of 2-octanol, has been critically reviewed by Castellan et al. [14]. They concluded that at low temperatures — around 20°C — the oxidation proceeds mainly via an ionic-molecular mechanism, whereas it proceeds via a radical molecular mechanism at temperatures higher than 60°C. We are interested in obtaining 2-octanone, which can be produced at a temperature around 20°C. We now assume that nitrosonium (NO^+) is the reactive component, which reacts with the organic compound transferred from the organic phase into the aqueous phase. However, at high temperatures the radical mechanism probably may become important and care must be taken at this temperature level.

The oxidation of 2-octanol with nitric acid can be represented as a consecutive pathway as shown in Fig. 2. In this simplified reaction scheme the main reactions and components are depicted. First 2-octanol is partially oxidized to 2-octanone, which forms a stable product. This reaction does not start without adding an initiator like $NaNO_2$, which forms nitrous acid. The oxidation of 2-octanol to 2-octanone is an autocatalytic reaction whereby nitrous acid is formed and nitric acid is consumed.

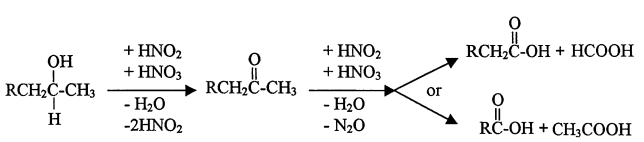

Fig. 2. Schematic reaction pathway for the oxidation of 2-octanol with nitric acid. Distribution of the obtained carboxylic acids is not shown in this scheme — see Longstaff and Singer [15] for more details.

The produced 2-octanone can be further oxidized to carboxylic acids. Depending on which carbon bond is broken hexanoic acid and acetic acid or heptanoic acid and formic acid are formed. The formic acid can again be oxidized to CO_2, see Ref. [15]. The amount of hexanoic acid as experimentally found is approximately two times larger than the amount of heptanoic acid. During the reaction nitrous acid and nitric acid are consumed. This reaction proceeds via some intermediate products, which are so unstable that they could not be found.

We have restricted ourselves to oxidation experiments of pure 2-octanol with 60 wt.% nitric acid. A small amount of nitrite is added to prevent the induction. The reaction system is still under investigation and it will the subject of another publication [16]. Although the complexity of the oxidation of 2-octanol (A) and 2-octanone (P) is high the stoichiometry can be simplified to only two equations as follows:

$$A + B \rightarrow P + 2B \quad (1)$$

$$P + B \rightarrow X \quad (2)$$

where B is the nitrosonium ion, which contributes for the autocatalytic effect, and X are carboxylic acids.

4. Neural network approach to model the conversion rates

The conventional kinetic model supplies the conversion rates as a function of the concentration of each relevant reactant as well as of the temperature. This kinetic modelling procedure consists of two essential steps: at first the reaction mechanism must be determined (assumed) and then the model parameters and their temperature dependencies are estimated. For some complex reacting systems non-stable or not measurable intermediate compounds must be introduced into the kinetic model, so the estimation of parameters and the verification of the model in this case is really difficult.

Using neural nets the modelled object or the phenomenon (here the reaction kinetics) is treated as a 'black-box', so merely the relevant input variables and their influence on the output signals are observed and mapped. From a mathematical point of view the neural network approximates a complex functional dependence and it is a mathematical superposition of all the expressions derived for a reliable and representative (but here unknown) kinetic model, without any explicit knowledge of these expressions.

The modelling procedure with the application of neural nets has been here executed in the following steps:

1. choice of input–output variables, relevant for the investigated system;
2. extraction of learning data from the experimental results;
3. choice of the network topology and learning method.

While point 3 is related directly to the existing calculation techniques, the first two points are integrally related to the system, so to execute them successfully quite a significant knowledge about the properties of the investigated system is required.

4.1. Choice of the input–output variables

A neural network, which at any operating conditions supplies the appropriate values of the conversion rates for the key reactants present in the system, works in fact as a kinetic model. Taking into account the properties of the investigated reacting system the input–output variables have been chosen as below.

From experimental data and the simplified reaction scheme (Eqs. (1) and (2)) it results that the state of the investigated reacting system can be uniquely described with the concentrations of the following reactants in the aqueous phase: c_A (2-octanol), c_P (2-octanone), c_X (total carbon acids) and c_N (nitric acid). The concentration of nitrosonium ion c_B, appearing in the stoichiometric equations, is a mathematical superposition of the previously listed variables (c_A, c_P, c_X, c_N), so it is expressed simultaneously with these variables. Including the reaction temperature, the set of the inlet signals finally consists of the following vector: $[T, c_A, c_P, c_X, c_N]^T$.

To estimate the listed concentrations, c_i, we have to know (determine, calculate) the solubility of each compound in the aqueous phase. But usually the distribution coefficients depend, for each reactant, on the temperature and also on the concentrations of other soluble compounds. Additionally, to study the investigated system accurately we probably need to use molar activities instead of simple molar concentrations. To avoid these problems the use of the global molar content of each compound in the entire reacting liquid-liquid system has been proposed. These global concentrations are defined for the investigated oxidation system as follows:

$$X_i = \frac{n_i}{\sum n_i} \quad (3)$$

where n_i is a global molar amount of the i-th compound in the reacting system. The molar contents n_A, n_P, n_X can be evaluated from, e.g. chromatographic analysis, while the values of n_N have to be estimated assuming that the initial concentration of this reactant equals to: $c_{N,in} = c_{HNO3,o} - c_{NaNO2,o} \approx c_{HNO3,o}$. The above defined global fractions have been introduced as an arbitrary, but a very convenient measure, for the

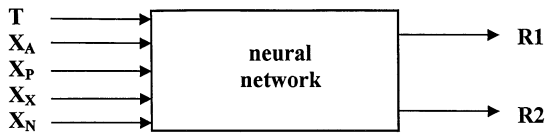

Fig. 3. Schematic diagram of the neural network representing the reaction kinetics.

purpose of the neural networks application. Using X_i, which expresses, as is shown in Eq. (3), the molar content of each reactant in the total heterogeneous liquid-liquid reaction mixtures, the input vector can be simplified. This is in contrast to the use of the well known concept of relative concentrations, which expresses separately the appropriate concentrations of any reactant in the organic and aqueous phases, respectively, one more input variable — being the volumetric fraction of the dispersed phase ε_d — has to be introduced into the input vector. Taking the above into account the concept of global concentrations has been chosen for the further analysis of the considered reacting system. The performed series of experiments differ mainly in the addition rate of the 2-octanol, so a possibility of incorporating of this feeding rate into the set of input variables has been also considered. But from the conceptual and practical reasons this possibility has been eliminated. The employed neural net is to represent the kinetics (in our case it incorporates also possible mass transfer resistances) so, it should work as kinetic expressions — i.e. for the set of intrinsically independent variables describing uniquely the state of the reacting system this net should supply the conversion rates for relevant reactants. The feeding rate of the 2-octanol is not related to the kinetics, it of course has an influence on the state of the system because at any moment of time the concentration of all reactants present in the reactor is a function of the feeding rate and the integrated conversion rate. This influence has been taken into account introducing the global concentrations as the input variables. These global concentrations uniquely define the state of the reaction mixture, independently on the history of the reaction mixture, so from practical point of view there is no need to introduce the feeding rate as an additional input variable. The feeding rate appears in Eq. (8) of the proposed below hybrid model to predict the behaviour of the reactor. Some comments on a possible inaccuracy introduced into the description as a result of use of the global concentrations are given below.

Using the simplified stoichiometric equations (Eqs. (1) and (2)) the behaviour of the reacting system, i.e. its rates of change, can be uniquely described with the overall conversion rate of the 2-octanol (A) into the 2-octanone (P) — r_1, and the overall conversion rate of the 2-octanone (P) into carboxylic acids (X) — r_2, where both r_1 and r_2 are expressed in moles per total volume of the reacting system and per second. To avoid a need for corrections of the reaction mixture volume as a result of possible changes of temperature and the mixture composition, it is more convenient to express each conversion rate per total initial number of moles of substrates (2-octanol and nitric acid) according to the relationship:

$$R_i = \frac{r_i V}{n_{Ao} + n_{No}} \tag{4}$$

The sum $(n_{Ao} + n_{No})$ used in the definition of Eq. (4) has been used to normalise the obtained numerical data of the conversion rates r_i.

The scheme of the proposed neural network representation of the reaction kinetics is shown in Fig. 3.

As a consequence of the used global concentrations and global conversion rates the proposed net represents not only the intrinsic reaction kinetics but also incorporates the simultaneous mass transfer phenomena, which play a role in the entire process. From the net performance point of view this fact is not relevant but it may be important for the knowledge generalisation. For the fast reaction regime the used overall reaction rates may incorporate the influence of the external mass transfer resistances, which are dependent on the interfacial area — so also on mixing and geometrical conditions. This problem will be discussed in detail later.

4.2. Extraction of the learning data from experimental results

To train a neural network successfully a representative set of learning data has to be prepared based on the available experimental results. This step has a decisive influence on the quality of the approximation of the reaction kinetics, and as such on the accuracy of process modelling as well as on the usability of the network for knowledge generalisation.

For the experimental investigations performed in this study two methods of extraction of learning data can be proposed:

- the learning data set is extracted only from the concentration measurements, so the available heat-flow results obtained in a RC1 reaction calorimeter are used as additional testing data — method M1;
- the learning data set is extracted on the basis of the concentration results obtained for the compound X and simultaneously of the heat-flow results, so the concentration results obtained for compound P are used only for testing — method M2.

The detailed algorithms for the preparation of the learning data are listed in Appendix A for both methods, respectively. Both algorithms result in an appropriate set of inlet-outlet learning data: $[T, X_A, X_P, X_X, X_N, R_1, R_2]^T$. Usually as many as 100 learning patterns have been extracted from each experimental run.

To compare and discuss properties of both methods used for the evaluation of the learning data the calculations described in the Appendix A have been performed with the use of the experimental results obtained for a singular run E_192. The obtained learning data are displayed in Figs. 4 and 5, respectively as a function of the dimensionless reaction time. Comparing the results plotted in these diagrams we can observe the following:

- as a result of the scarce sampling only a few concentration measurements during each oxidation run have been performed, so the interpolation polynomials for ζ_P and ζ_X versus θ are usually of an order as low as 2. In consequence, for method M1, exactly linear dependencies of R_i versus θ have been obtained — see Fig. 4;
- for method M2, the calculated dependence of R_1 versus θ strictly follows the shape of the curve Q_R versus θ obtained from the RC1 heat-flow measurement, although the absolute values of the conversion rate R_1 are not far from those obtained with method M1.

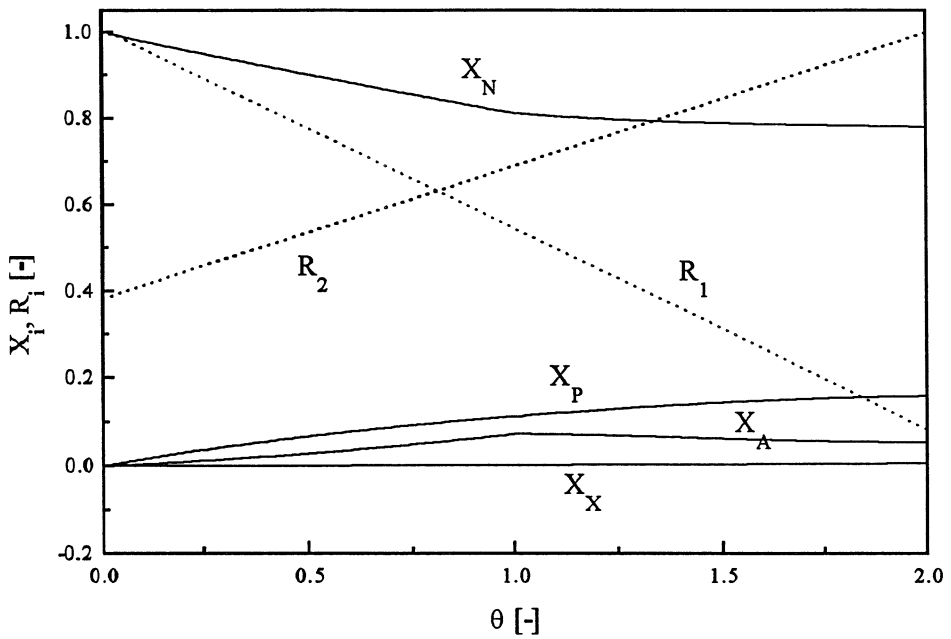

Fig. 4. Learning data extracted from run E_192 following the method M1 — see Appendix A.

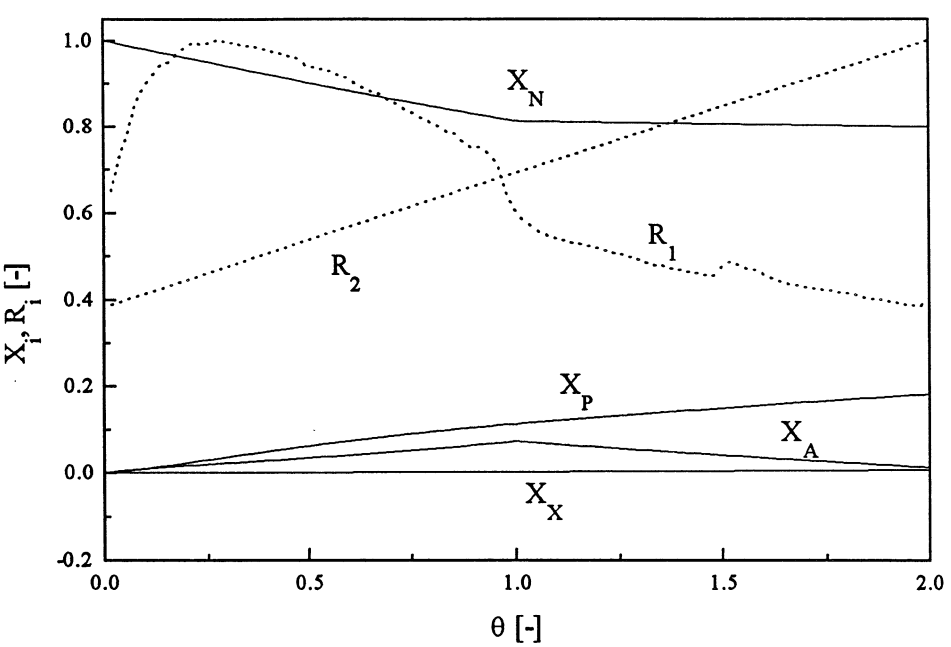

Fig. 5. Learning data extracted from run E_192 following the method M2 — see Appendix A.

Table 2
Accuracy of learning

Series	Net code	Net topology	Learning data set	Learning patterns P	rms error
I	N1	5-4-2	LS1: E_204, E_191, E_190, E_195	200	$6.06\ 10^{-4}$
II	N2	5-4-2	LS2: E_178, E_229, E_181, E_170, E_175	250	$2.00\ 10^{-3}$
III	N3	5-4-2	LS3: E_179, E_228, E_223, E_171, E_176	250	$2.85\ 10^{-3}$
IV	N4	5-4-2	LS4: E_180, E_174, E_184, E_222, E_220	250	$4.44\ 10^{-3}$
I–IV	N5	5-5-4-2	LS5: LS1+LS2+LS3+LS4	450	$8.41\ 10^{-2}$
I–III	N6	5-10-2	LS6: LS1+LS2+LS3	350	$3.51\ 10^{-2}$

The two neural networks trained with learning data obtained from both method M1 and M2, respectively predict experiment E_192 with a very good accuracy, although method M2 exhibits significant inaccuracies in the predicted ζ_A versus θ line at the beginning and at the end of the process. Taking this into account we have ultimately chosen the first method M1 to evaluate all data for learning.

Four sets of learning data, a separate set for each experimental series, have been produced with method M1 extracting learning data from approximately each second experiment performed — see Table 1, where experiments used for creating learning data are written in bold font. These runs taken for learning are listed also in Table 2.

Before presenting to the network, the learning data have been normalised and randomly mixed within each set.

4.3. Learning of the net to approximate the reaction rates

The net architecture has been optimised during the learning procedure to find the net topology as simple as possible but sufficiently complex to map accurately the chosen input and output variables. Usually only one hidden layer was enough to approximate the investigated reaction kinetics.

During the learning procedure a vector of the net parameters (weights) **W** has been modified to minimise differences between the outputs predicted with the net, y, and outputs used for learning, d. The following target function has been minimised:

$$E(\mathbf{W}) = \frac{1}{2}\sum_{j=1}^{P}\sum_{k=1}^{M}(y_k^{(j)} - d_k^{(j)})^2 \tag{5}$$

where the index j (in brackets) is a superscript not a power exponent. Basing on the methods described among others by White and Sofge [7] the optimisation of the net has been performed by pruning of each i-th neurone from the hidden layer(s) for which $\Sigma_{j=1}^{N}|W_{ij}| \approx 0$.

For the net with only one hidden layer, which consists of N neurones in the input layer, K neurones in the hidden layer and M neurones in the output layer, respectively-each k-th output can be calculated as follows:

$$y_k = f\left(\sum_{i=0}^{K} W_{ki}^{(o)} f\left(\sum_{j=0}^{N} W_{ij}^{(h)} x_j\right)\right) \tag{6}$$

where f is the so-called activation function.

In this study a Levenberg–Marquardt method [17] has been employed as a very efficient method for non-linear optimisation to find the optimal weight vector **W**. Not only the target function E but also the root mean square (rms) has been used as a criterion of the fitting quality, which is defined as:

$$\text{rms} = \sqrt{\frac{E}{P}} \tag{7}$$

where P is a number of learning patterns used for training.

The accuracy of learning obtained with different sets of learning data is shown in Table 2. A comparison of the output data (normalised values of the conversion rates R_1 and R_2) calculated by the trained network with the experimental data as used for learning is performed in Fig. 6 for the set LS2. As is visible in the diagram, a quite good accuracy of fitting has been obtained in case the net is trained with data extracted from singular series of experiments. Unfortunately, the accuracy of learning deteriorates significantly in case that the set of

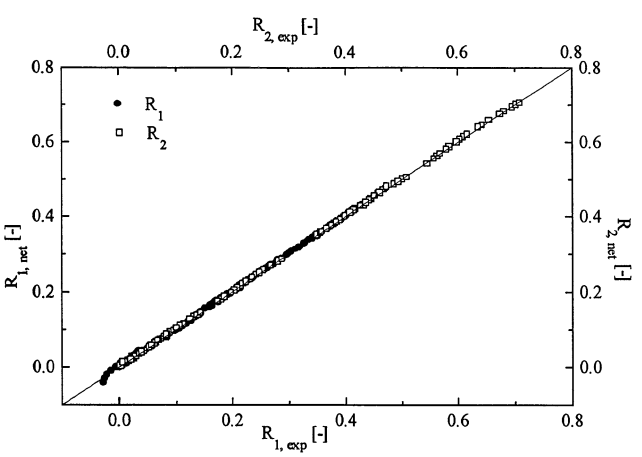

Fig. 6. Learning quality-comparison of the output data calculated with the trained network to the experimental ones; the learning data set — LS2.

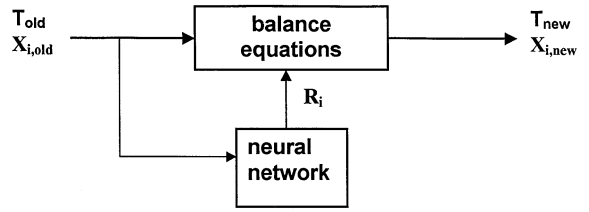

Fig. 7. Schematic diagram of a numerical representation of the hybrid first principle — neural network model.

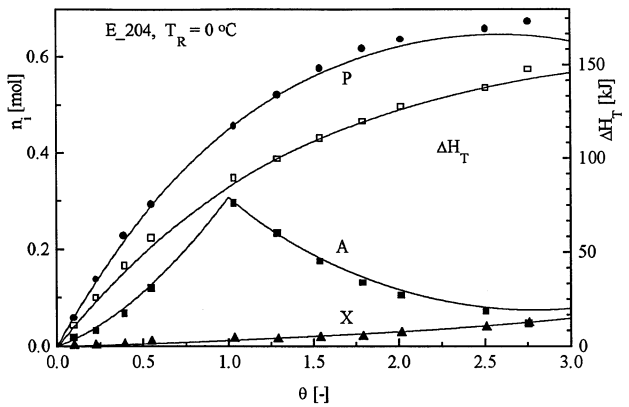

Fig. 8. Accuracy of predictions obtained with the hybrid model — comparison of the experimental and calculated molar content of the key reactants *A*, *P* and *X* in the reactor as well as the integrated heat effect ΔH_T; run E_204 used to create the learning data set.

learning patterns contains data taken from different experimental series: compare in Table 2 the values of rms errors obtained for LS1–LS4 with the LS5 and LS6 learning data sets, respectively. This observation is discussed in detail below.

5. Hybrid model of the reactor

The proposed neural network supplies after training the conversion rates at any operating conditions, so the mass and heat balances can be formulated and solved for different types of reactors and their operating modes (taking into account the restriction related to the mass transfer resistances mentioned above). For the semi-batch process performed and following the operating mode described in the experimental part of this paper — i.e. for the process where the 2-octanol is successively fed into the nitric acid being already in the reactor — the overall mass balances can be written as follows:

$$\frac{dn_A}{dt} = \varphi_{Ad} - R_1(n_{Ao} + n_{No}) \quad (8)$$

where φ_{Ad} [moles/s] is the molar feeding rate of compound A into the reactor.

$$\frac{dn_P}{dt} = (R_1 - R_2)(n_{Ao} + n_{No}) \quad (9)$$

$$\frac{dn_x}{dt} = R_2(n_{Ao} + n_{No}) \quad (10)$$

$$\frac{dn_N}{dt} = (-R_1 + R_2 - 2R_2)(n_{Ao} + n_{No}) = (-R_1 - R_2)(n_{Ao} + n_{No}) \quad (11)$$

The energy generated as a result of the reaction can be estimated according to the relationship:

$$Q_R = (R_1 \Delta H_P + R_2 \Delta H_x)(n_{Ao} + n_{No}) \quad (12)$$

The set of equations (Eqs. (8)–(12)), together with the heat balance equation which for a stirred tank reactor with a cooling jacket reads as follows:

$$n_R c_p \frac{dT_R}{dt} = Q_R - Q_C - \varphi_{Ad} c_{pA}(T_R - T_A) - Q_L, \quad (13)$$

makes a hybrid model for the considered semi-batch reactor, provided the conversion rates are supplied by the trained network. In Eq. (13) the cooling rate Q_C is equal to $UA(T_R - T_C)$, while Q_L is the rate of heat losses. For the isothermal experiments carried out in a RC1 reactor, we have immediately from the heat balance of Eq. (13), that the heat generation rate as a result of the reaction progress, Q_R — as it is calculated with Eq. (12) — is equal to that as determined directly from measurements.

A schematic diagram of the numerical representation of such model is shown in Fig. 7. A numerical version of the proposed hybrid model has been elaborated, where the trained neural network representing the kinetics has been implemented and a fourth order Runge-Kutta method has been applied to solve the system of balance equations of $8 \div 13$. Then for the demanded operating and initial conditions the hybrid model predicts the behaviour of the reactor.

6. Results and discussion

After the learning procedure each trained neural net has been tested with use of all experimental data. To this end the proposed hybrid model has been employed to predict the performance of the semi-batch stirred tank reactor operated at isothermal conditions in which the concentration profiles as well as the energy generated as a result of the reaction have been calculated as a function of time.

Comparisons of the experimental and calculated conversions of the key reactants *A*, *P* and *X* as well as the integrated heat effect ΔH_T are shown in Figs. 8 and 9, as a function of the reaction time for two runs of Series I. Data from run E_204 have been used to create the set of learning data LS1; it is visible in Fig. 8 that in

this case the hybrid model is able to reproduce the data extracted for learning with an excellent accuracy. This comparison is performed to test both, an accuracy of the implemented neural network as well as numerical procedures installed to integrate differential model equations. Data from run E_194 have not been used for learning — it can be seen in Fig. 9 that also in this case the neural net is able to predict the concentration and heat-flow results with a very good accuracy. Notice that, following method M1, the heat-flow results have not been used to train the net, so a comparison of calculated and experimental heat-flow results is an additional, independent test of the accuracy of the proposed approach.

Many testing calculations have been performed and their results are shown in Table 3, where the accuracy of the results obtained with four different hybrid models are compared. The used hybrid models differ in the type of the neural network implemented — see also Table 2. Notice that model HM1, employing net N1, has been used to model Series I, model HM2 for Series II, etc.

To display quantitatively the accuracy of predictions obtained with each type of the hybrid model the following measures have been used:

- for each run the relative error of predictions defined as:

$$RE_i = \frac{z_{i,\mathrm{exp}} - z_{i,\mathrm{cal}}}{z_{i,\mathrm{exp}}}, \tag{14}$$

where z_i is the conversion of the key organic compound — ζ_A, ζ_P, ζ_X or the integrated heat effect ΔH_T;

- for each series of measurements the root mean square (rms) error defined as:

$$\mathrm{rms} = \sqrt{\frac{1}{2}\frac{\sum_{i=1}^{L}(RE_i)^2}{L}} \tag{15}$$

where L is the number of experiments (experimental points available) taken for the comparison.

The root mean square values *rms* obtained for different categories of testing data are collected in Table 3. The

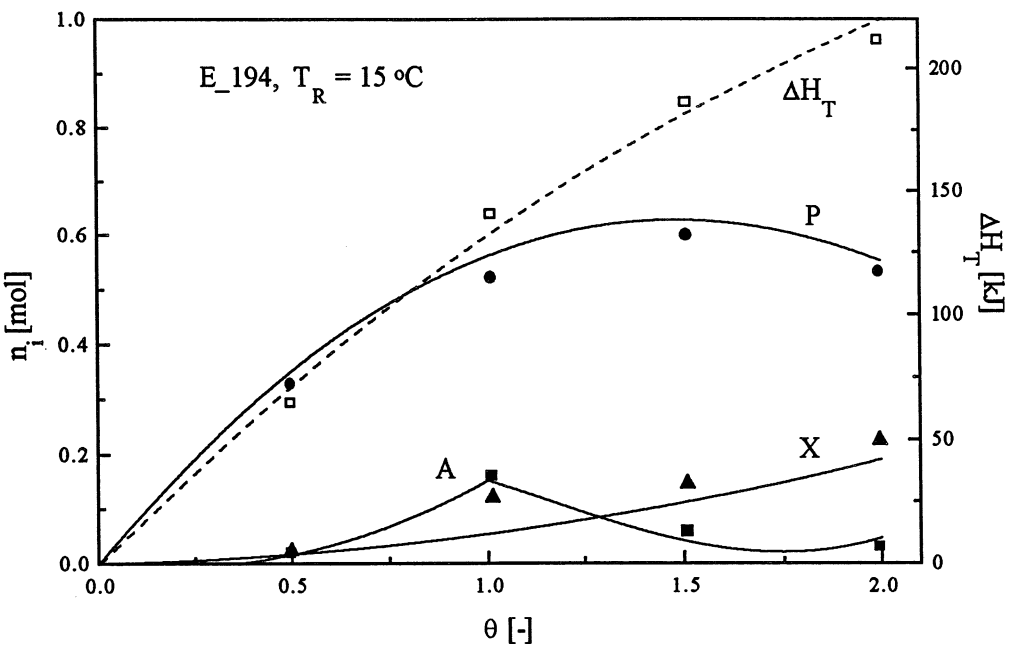

Fig. 9. Accuracy of predictions obtained with the hybrid model — comparison of the experimental and calculated molar content of the key reactants *A*, *P* and *X* in the reactor as well as the integrated heat effect ΔH_T; run E_194 not used to create the learning data set.

Table 3
Accuracy of the elaborated hybrid models

Net code	Hybrid model	Data used for learning		Data not used for learning	
		Concentration (rms error)	Heat-flow (rms error)	Concentration (rms error)	Heat-flow (rms error)
N1	HM1	0.047	0.107	0.049	0.101
N2	HM2	0.058	0.177	0.098	0.199
N3	HM3	0.054	0.194	0.066	0.139
N4	HM4	0.129	0.162	0.791	0.733

accuracy of the predictions of the concentration and heat-flow results for the runs used for learning and the runs not used for learning have been tested separately. The obtained results can be summarised as follows:

- the net trained with data taken from a singular series of experiments (the same dosing rate of 2-octanol) is able to represent accurately the reaction kinetics (together with simultaneous mass transfer phenomena), so the behaviour of the reactor can be accurately predicted in this case;
- an excellent accuracy of fitting has been obtained for series I, II and III, although some discrepancies for the integrated heat effect ΔH_T can be observed at higher reaction temperatures;
- for Series IV (experiments performed at the fastest dosing rate of 2-octanol) the neural network approach failed — only at the lowest temperatures of $T_R = 0$ and 10°C a sufficiently accurate approximation has been obtained;
- concentration data are reproduced much more accurately than integrated heat effect data;
- for Series I–III there is no significant difference between the prediction accuracies obtained for data used for learning and data not used for learning;
- the net trained with data extracted from one series of experiments can not be used to represent data from another series of experiments.

7. Summary

From the performed tests it is visible that the accuracy of the predictions is good as long as the learning of the net and reproducing of the reaction rates are restricted to a singular series of experiments. Each trial of knowledge generalisation and representation of runs of different series with only one neural net failed. This can be caused by the fact that in the proposed approach the neural net represents the intrinsic kinetics, but also the simultaneous mass transfer phenomena are incorporated. Droplet diameters in the dispersed system, so also the interfacial liquid-liquid contact area, depend on the Weber number, We, and on the volume fraction of the dispersed phase. The proposed neural nets, as a result of the used global concentrations, should be able to incorporate and represent also the influence of the interfacial contact area on the overall reaction rate, because the stirrer geometry and the stirrer speed have been kept the same from run to run. The observed discrepancies and poor abilities of the networks to generalise the knowledge within all experimental series can be caused by some phenomena which change the system properties from series to series, e.g. such as phase inversion or changes in the reaction mechanism caused by a temporary excess of some reactants in the reactor, etc. At certain conditions the inaccuracies in the estimation of the global concentrations X_i can play a significant role. These concentrations have been estimated on a basis of the composition of the organic phase: in view of the limited solubility of the organic compounds in the aqueous phase this method usually does not introduce significant errors in the values of X_i. But if the concentration of some organic compounds in the aqueous phase becomes significant, particularly for the formic and acetic acids produced during the reaction this is the case, the method used to estimate the global concentrations may introduce relevant discrepancies into the description. Based on the obtained results we can point out the limitations of the elaborated hybrid model with respect to the generalization are not related to the learning accuracy nor the optimization of the net architecture, so not to a network itself, but rather to the diversity of the modelled system. Because anticipated changes in the reaction mechanism as a result of a temperature rise and/or as a result of an excess of one reactant the approximation of the reaction kinetics with only one net becomes extremely difficult. To explain these poor generalisation abilities of the proposed approach further studies are needed.

At the current stage the reactor behaviour can be predicted within a singular series of experiments. The hybrid model employing the appropriate neural network can be used for this purposes as has been shown in Fig. 7. For each series of experiments using determined weights the overall reaction rates R_1 and R_2 can be estimated at any operating conditions with Eq. (6), after that the concentration profiles and energy generated as a result of the reaction can be predicted with Eqs. (8)–(13).

Because in the proposed approach the neural network incorporates and represents the intrinsic reaction kinetics coupled with external mass transfer resistances, special precautions have to be taken to model reactors much larger in volume than the RC1 reactor used in our experiments. Experimental scaling-up studies are needed to determine the limitations of the described method, but considering the general properties of liquid-liquid reacting dispersions some practical scaling-up rules can be formulated. These rules are based on the assumption that equal interfacial areas per unit volume of dispersion must be aimed at in laboratory and large reactor vessels, respectively. This implicates a complete geometric, kinematic, dynamic and thermal similarity of the equipment. Okufi et al. [18] state that usually geometric similarity and the rule of equal impeller tip speed provide the best scale-up criteria for equal interfacial areas. For the studied semi-batch process also the dosing rate is a significant scaling-up parameter in view of possible accumulation of unreacted 2-octanol in the reactor vessel. Then an equal ratio of the volumetric dosing rate of 2-octanol to the initial volume of the

nitric acid present in the reactor, is an additional scaling-up criterion for the system under consideration. In our investigations, for each experiment this ratio was equal to: (0.42/t_dos) [s^{-1}] where the dosing time t_dos [s] is as listed in Table 1.

8. Notation

c	molar concentration, mole m^{-3}
cp	molar specific heat capacity, J mol^{-1} K^{-1}
D_a	stirrer diameter, m
E	target function, -
f	activity function, -
ΔH	reaction enthalpy, J mol^{-1}
ΔH_T	integrated heat effect, J
K	number of neurons in the hidden layer
M	number of neurons in the output layer
n	number of moles, -
n_{Ao}	total amount of A added to the reactor, mole
N	number of neurons in the input layer
N	stirrer speed, s^{-1}
P	number of learning patterns, -
Q	power generated as a result of the reaction progress, W
r	conversion rate, mole m^{-3} s^{-1}
R	relative conversion rate (Eq. (4)), s^{-1}
RE	relative error (Eq. (14)), -
t	time, s
T	temperature
UA	product of the overall heat transfer coefficient and the heat exchange surface area for the reactor jacket, W K^{-1}
V	reactor volume, m^{-3}
W	weights vector,
$We=\frac{\rho_c N^2 D_a^3}{\sigma}$	the Weber number, -
X	global molar fraction, -
y	output of the net (Eq. (6)), -

Greek

$\theta = t/t_{dos}$	dimensionless time,
ζ	conversion (Eqs. (A2) and (A3)), -
φ	feeding rate, mole s^{-1}
ρ_c	density of the continuous phase, kg m^{-3}
σ	interfacial tension, N m^{-1}

Subscripts

A	2-octanol
P	2-octanone
X	carboxylic acids
B	nitrosonium ion
N	nitric acid
dos	dosing

Acknowledgements

These investigations were partially supported by the Netherlands Foundation for Chemical Research (SON) with the financial aid from the Netherlands Technology Foundation (STW). The authors wish to thank E.A.H. Ordelmans and S.P.W.M. Lemm for their contribution to the experimental work, and F. ter Borg, A.H. Pleiter and G.J.M. Monnink for technical support.

Appendix A. Extraction of learning data from experimental results

Experimental results available:

(a) Concentration measurements:	ζ_A, ζ_P, $\zeta_X = f(\theta)$ from chromatographic measurements, sampling frequency $\Delta\theta \cong 0.5$, usually 5 points for each run. In view of the limited number of samples, the concentration results have been interpolated with polynomial expressions.
(b) Heat-flow measurements:	T_R, $Q_R = f(\theta)$ from the RC1 reaction calorimeter, sampling frequency $\Delta\theta \cong 0.01$, usually more than 200 points for each run.

Method M1

For each run, at any moment of the reaction progress, the obtained experimental data have been prepared for the learning procedure as follows:

$$T = \frac{T_R - T_{R,min}}{T_{R,max} - T_{R,min}} \quad \text{(A1)}$$

$$n_P = \zeta_P n_{Ao} \quad \text{and} \quad n_X = \zeta_X n_{Ao}, \quad \text{(A2)}$$

where $\zeta_P = f(\theta)$ and $\zeta_X = f(\theta)$ are supplied by the interpolation polynominals

$$\zeta_A = \frac{\zeta_P + \zeta_X}{\theta} \quad (\text{if} \quad \theta > 1 \quad \text{then} \quad \theta = 1) \quad \text{(A3)}$$

Because the polynomial approximations $\zeta_P = f(\theta)$ and

$\zeta_X = f(\theta)$ are not sufficiently accurate at $\theta \cong 0$, so in the range $0 < \theta < 0.5$ a linear dependence of ζ_A on θ simply has been assumed

$$n_A = n_{Ao}\theta(1 - \zeta_A) \tag{A4}$$

$$n_N = n_{No} - n_{No2,o} - n_P - 2n_X \tag{A5}$$

$$\sum n = n_A + n_P + n_X + n_N \tag{A6}$$

$$X_i = \frac{n_i}{\sum n_i} \tag{A7}$$

where $i = A, P, X$ and N

$$\frac{dn_i}{dt} = \frac{d\zeta_i}{dt} n_{Ao} \tag{A8}$$

where $i = P$ and X

$$R_2 = \frac{1}{n_{Ao} + n_{No}} \frac{dn_X}{dt} \tag{A9}$$

$$R_1 = \frac{1}{n_{Ao} + n_{No}} \left(\frac{dn_P}{dt} + R_2 \right) \tag{A10}$$

Method M2

For each run, at any moment during the reaction, the obtained experimental data have been prepared for the learning procedure as follows:

$$T = \frac{T_R - T_{R,min}}{T_{R,max} - T_{R,min}} \tag{A11}$$

$$n_X = \zeta_X n_{Ao} \tag{A12}$$

where $\zeta_X = f(\theta)$ is supplied by the interpolation polynomial

$$R_2 = \frac{1}{n_{Ao} + n_{No}} \frac{dn_X}{dt} \tag{A13}$$

$$R_1 = \frac{\frac{Q_R}{n_{Ao} + n_{No}} - R_2 \Delta H_X}{\Delta H_P} \tag{A14}$$

$$\frac{dn_A}{dt} = \varphi_{Ad} - R_1(n_{Ao} + n_{No}) \tag{A15}$$

after integration and transformation we obtain $n_A = f(\theta)$

$$\frac{dn_P}{dt} = (R_1 - R_2)(n_{Ao} + n_{No}) \tag{A16}$$

idem $n_P = f(\theta)$

$$\frac{dn_N}{dt} = (-R_1 - R_2)(n_{Ao} + n_{No}) \tag{A17}$$

idem $n_N = f(\theta)$

$$\sum n = n_A + n_P + n_x + n_N \tag{A18}$$

$$X_i = \frac{n_i}{\sum n_i} \tag{A19}$$

where $i = A, P, X$ and N.

References

[1] A.B. Bulsari (Ed.), Neural Networks for Chemical Engineers, Computer-Aided Chemical Engineering 6, Elsevier, Amsterdam, 1995.

[2] D.C. Psichogios, L.H. Ungar, A hybrid neural network-first principles approach to process modelling, AIChE J. 38 (1992) 1499–1511.

[3] H.J. Zander, R. Dittmeyer, J. Wagenhuber, Dynamische modellierung chemischer reaktionssysteme mit neuronalen netzen und hybriden modellen, Chem. Ing. Technol. 71 (1999) 234–237.

[4] I.M. Galvan, J.M. Zaldivar, H. Hernandez, E. Molga, The use of neural networks for fitting complex kinetic data, Comput. Chem. Engng. 20 (1996) 1451–1465.

[5] E. Molga, K.R. Westerterp, Neural network based model of the kinetics of catalytic hydrogenation reactions, 1997 Studies in Surface Science and Catalysis, in: G.F. Froment, K.C. Waugh (Eds.), Proceedings of the International Symposium on Dynamics of Surfaces and Reaction Kinetics in Heterogeneous Catalysis, Antwerpen, Belgium, September 15–17, 1997, pp. 379–388.

[6] E. Molga, R. Cherbanski, Hybrid first principle-neural network approach to modelling of the liquid-liquid reacting system, Chem. Eng. Sci. 54 (1999) 2467–2473.

[7] D.A. White, D.A. Sofge (Eds.), Handbook of Intelligent Control, Neural, Fuzzy and Adaptive Approaches, van Nostrand Reinhold, New York, 1992, pp. 141–182.

[8] R. Reisen, B. Grob, Reaction calorimetry in chemical process development, Swiss Chem. 7 (1985) 39–43.

[9] A.G. Mettler-Toledo, Operating Instructions RC1 Reaction Calorimeter, Mettler-Toledo AG, Switzerland, 1993.

[10] Y. Ogata, Oxidations with nitric acid or nitrogen oxides, in: W.S. Trahanovsky (Ed.), Oxidation in Organic Chemistry, Part C, Academic Press, 1978, pp. 295–342.

[11] M. Horvath, I. Lengyel, G. Bazsa, Kinetics and mechanism of autocatalytic oxidation of formaldehyde by nitric acid, Int. J. Chem. Kinet. 20 (1988) 687–697.

[12] E.J. Strojny, R.T. Iwamasa, L.K. Frevel, Oxidation of 2-methoxyethanol to methoxyacetic acid by nitric acid solutions, J. Am. Chem. Soc. 9 (1971) 1171–1178.

[13] E. Camera, G. Modena, B. Zotti, On the behaviour of nitrate esters in acid solution. III. Oxidation of ethanol by nitric acid in sulphuric acid, Propell. Explos. Pyrotech. 8 (1983) 70–73.

[14] A. Castellan, J.C.J. Bart, S. Cavallaro, Nitric acid reaction of cyclohexanol to adipic acid, Cataly. Today 9 (1991) 255–283.

[15] J.V.L. Longstaff, K. Singer, The kinetics of oxidation by nitrous acid. Part II. Oxidation of formic acid in aqueous nitric acid, J. Chem. Soc. (1954) 2610–2617.

[16] B.A.A. van Woezik, K.R. Westerterp, The nitric acid oxidation of 2-octanol. A model reaction for multiple heterogeneous liquid–liquid reactions. Chem. Eng. Process. (1999) submitted for publication.

[17] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM J. Appl. Math. 11 (1963) 431–441.

[18] S. Okufi, E.S. Perez de Ortiz, H. Sawistowski, Scale-up of liquid-liquid dispersions in stirred tanks, Can. J. Chem. Eng. 68 (1990) 400–406.