

Modelling of the p,p'-Dinitrodibenzyl Electroreduction by Using an Artificial Neural Network

Maria Olea^a

Department of Chemical Engineering, Babes-Bolyai University of Cluj-Napoca, Romania

(Received November 14, 2006)

Abstract

An artificial neural network-based model was developed to predict the performances of the p,p'-dinitrodibenzyl (DNDB) electroreduction under different operating conditions. Six feed-forward networks with a three-layer structure were trained and then optimized through the validation process. The performance of the neural network was examined with respect to the learning coefficient, number of hidden nodes and number of iterations required to reduce the total error to an acceptable level. The best predictions were obtained with a neural network having an 8-12-2 architecture and "tansig", "tansig" as threshold functions. The 8 nodes of the input layer correspond to two input variables, current density and quantity of electricity passed, and six state variable, temperature, initial concentration of DNDB in the electrolyte, mass transfer coefficient, electrode surface area, number of electrons involved, and volume of the catholyte. The 2 nodes of the output layer correspond to current efficiency and yield of the main reaction product, the p,p'-Diaminodibenzyl (DADB).

^aPresent address:

Laboratorium voor Petrochemische Techniek, Department of Chemical Engineering, Ghent University, Krijgslaan 281- S5, B-9000 Ghent, Belgium
Tel.+ 32 9 264 4538, Fax +32 9 264 4999, E-mail: Maria.Olea@UGent.be

1. Introduction

p,p'-Dinitrodibenzyl electroreduction to p,p'-Diaminodibenzyl is a complex electrochemical reaction (Olea, 1996; Danciu, 1998). This reaction was carried out at constant current density in a divided batch electrolyzer, with a rapid external recirculation of the hydro-alcoholic electrolyte at 333K (Oniciu, 1990). During the electroreduction of DNDB, two side chemical reactions, namely a condensation reaction and a reaction that lead to the formation of quinoine-imine intermediates, occur as well. These intermediates can afterwards be electroreduced, even easier than DNDB. Thus, the analytical (phenomenological) modelling of the DNDB electroreduction, i.e., via mass, energy, momentum and potential balance equations, becomes difficult. Also, the exhaustive experimentation required to collect the kinetic data needed to the analytical modelling of the process turns out to be cumbersome and prohibitively costly.

In recent years, the artificial neural networks (ANNs) based modelling approach has opened a new avenue for developing empirical models useful to predicting the performance of the process outside the experimental domain. Especially, neural networks proved to be a powerful tool for catalyst development and optimization. Many applications evidenced the strong capability of the neural networks to provide relations between the catalyst components and its activity. As for example, an artificial neural-network analysis has been applied to the prediction of conversions and yields in the catalytic reaction of decomposition of NO into N₂ and O₂ over Cu/ZSM-5. With a proper network structure and training sets, the evolved neural-network system provides plausible activities of catalysts from a discrete and finite number of experimental data. A two- or three-dimensional graphical expression of the calculated data seems to be useful to reveal the relationship between the catalytic activity

and reaction conditions. Therefore, they can be utilized in finding optimum reaction conditions or in examining effects of various factors on catalytic reaction (Sasaki, 1995).

Based on the properties of neural network, improved back-propagation networks were developed and used for the design of the best catalyst for propane ammoxidation (Hou, 1997), methane oxidative coupling (Huang, 2001), and preferential CO oxidation in excess hydrogen (Omata, 2005).

Besides of the prediction of catalytic activity, neural networks were successfully used for modelling the kinetics of catalytic reactions, models applicable to catalyst scale up and process control and optimization in the frame of combinatorial catalysis (Serra, 2003). Moreover, in the frame of high-throughput synthesis, the ANN kinetic models for a series of zeolites aided at the rapid determination of the occurrence and crystallinity of zeolite beta and competing phases (Moliner, 2005).

On the other hand, a neural network has been found relevant for the estimation of acid strengths of mixed oxides, and it has been shown that not only the interpolation but also the extrapolation of given acid strength data is possible within reasonable experimental error (Hattori, 1995).

The ANNs being nonparametric models have the advantage that the model coefficients can be estimated directly from the samples of process input-output data.

The present paper reports the artificial neural network model developed for predicting the current efficiency and DADB yield as a function of the electrochemical reactor operating conditions, such as current density, quantity of electricity passing through the reactor, temperature, initial concentration of DNDB, electrode surface area.

As the current efficiency is a measure of the electrochemical reactor's ability to utilize current in favor of one specific component, it could also be a measure of the electrochemical process selectivity.

At our best knowledge, it is for the first time when the ANN approach is used for modeling electrocatalytic data.

2. Artificial neural network-based modelling

Artificial neural network is a system loosely modeled on the human brain. It consists of many processing elements, neurons or nodes that work in parallel. Nodes are connected to each other in layers (input, hidden and output), and the layers are interconnected. A "neuron" is essentially a regression equation with a non-linear output, each neuron (except those of the input layer) having the task to process the signals received from its dendrites to its own threshold function and to output the answer through its axon to the rest of the neurons. Therefore, as it is shown in Fig. 1, each node has two features: (1) the activity a (or state of the node) resulting from the inputs and (2) the output, which is a function of the activity, $f(a)$. The activity of node j is simply the weighted sum of the inputs given by:

$$a_j(\mathbf{X}) = \mathbf{W}_j^T \mathbf{X}$$

where $\mathbf{X} = (X_1, X_2, \dots, X_n, 1)^T$ is the measurement or input pattern vector, which reflects the state of the process and $\mathbf{W}_j^T = (W_{1j}, W_{2j}, \dots, W_{n+1j})$ is the weight or gain vector. The output function $f(a_j)$ maps the current state of activation to an output signal (Hoskins, 1998).

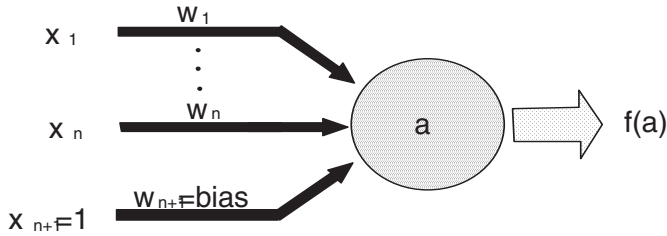


Fig. 1. Features of the artificial neuron.

ANNs have been applied in solving a wide variety of problems. But, in order to “fit” a particular ANN to a particular problem it must be trained (or learned) to generate a correct response for a given set of inputs (Mitchell, 1997). Learning (or training) an ANN means that if one gives a *training* set of known pairs of input data and output data, the ANN calculates a set of output data from the given set of input data, and it iteratively adjusts its own network patterns by the training algorithm so as to make the calculated output data as close to the given output data as possible (Kito, 1994). The performance of the network is then compared by evaluating the error function using an independent *validation* set, and the network having the smallest error with respect to the validation set is selected. This approach is called the *holdout method*. Since this procedure can itself lead to some over fitting to the validation set, the performance of the selected network should be confirmed by measuring its performance on a third independent set of data called a *test* set. Learning in the artificial neural networks can be supervised or unsupervised. Supervised learning means that the network has some information present during learning (training) to tell what the correct answer should be. In contrast, unsupervised learning means that the network has not such knowledge of the current answer and thus cannot know exactly what the correct output should be (Bhagat, 1990).

Two supervised learning algorithms were tested for the off-line training of our six multi-layer feed-forward proposed networks: the error-back propagation algorithm (EBP) (Rumelhalt, 1986) and the Levenberg-Marquardt algorithm (Hagan, 1994). The EBP algorithm is based on a gradient descent technique known as the generalized delta rule (GDR), while the Levenberg-Marquardt algorithm is based on a non-linear optimization technique (Freeman & Skapura, 1991). Because the Levenberg-Marquardt algorithm had the fastest convergence, we used this algorithm only in our further attempts to design the best ANN for our purpose.

Designing an ANN consists of arranging neurons in various layers, deciding the type of connection among neurons for different layers, as well as among the neurons within a layer, deciding the way a neuron receives input and produces output, and determining the strength of connection within the network by allowing the network to learn the appropriate values of connection weights by using a training data set.

3. ANNs approach for the DNDB electroreduction

For the ANN-based modelling of the DNDB electroreduction, six feed-forward networks with eight input - two output neuron structure were designed. The reactor operating conditions represented by the eight inputs are: input I-current density, i , (A/m^2), input II-the quantity of electricity passed, Q_p , (C), input III-temperature, T , (K), input IV-initial concentration of DNDB in the electrolyte, C_{DNDB}^0 , (mol/m^3), input V-mass transfer coefficient, K_m , (m/s), input VI-surface electrode area, A , (m^2), input VII-number of electrons involved in the DNDB electroreduction, n , and input VIII-volume of catholyte, V , (m^3). The two output neurons for each of the six networks represented the calculated current efficiency, CE_c , and calculated yield of DADB, Y_{cDADB} . Measured current efficiency, CE_m

and measured DADB yield, Y_{mDADB} , were obtained through the quantitative analysis of the experimental data of product distribution.

The number of nodes in the hidden layer is not fixed and requires to be optimized.

The six neural networks were trained and optimized using the experimental data (inputs) given in Table 1. The training set consisted of sixteen patterns. Initial weights assigned randomly between ± 0.1 worked best for this system. The validation set consisted of three patterns. The performance of the neural network was examined with respect to the learning coefficient, number of hidden nodes and number of iterations required to reduce the total error to an acceptable level. Along with the initial process conditions used for the training of the networks, Table 1 presents the measured current efficiency, CE_m , and yield, Y_{mDADB} , as well as the calculated, CE_c , and Y_{cDADB} , respectively, and the total error.

Pattern No.	i [A/m^2]	$Q_p \cdot 10^{-5}$ [C]	CE_m [%]	CE_c [%]	Y_{mDADB} [%]	Y_{cDADB} [%]	Total Error
1	500	1.30	80.70	80.64	81.03	81.37	-0.28
2	1000	0.90	74.50	74.72	71.90	72.05	-0.37
3	1000	1.45	82.29	82.09	70.34	70.63	-0.09
4	1000	1.50	90.27	91.07	88.23	87.55	-0.12
5	1000	1.80	83.78	83.62	85.52	85.65	-0.03
6	1000	2.30	56.82	56.89	77.61	77.75	-0.21
7	1500	1.50	85.68	85.69	76.25	76.18	0.06
8	1500	1.90	80.26	80.24	89.11	89.09	0.04
9	2000	0.63	59.92	59.99	90.55	90.75	-0.27
10	2000	0.90	48.12	48.57	46.05	46.59	-0.85
11	2000	0.92	71.90	71.48	68.73	68.35	0.80
12	2000	1.25	91.80	91.10	67.17	67.95	-0.01
13	2000	1.40	72.27	72.39	79.09	79.88	-0.90
14	2000	1.80	69.84	69.79	74.42	74.53	-0.06
15	2000	1.90	65.50	65.43	76.19	76.98	-0.72
16	2000	2.00	62.04	62.07	71.60	72.58	-0.83

Table 1. Measured and calculated current efficiency and DADB yields for different current densities and quantities of electricity passed

We did not include in Table 1 the initial values for the inputs III-VIII because their values were not changed from experiment to experiment. They were: $T = 333 \text{ K}$, $C_{\text{DNDB}}^{\circ} = 229 \text{ mol/m}^3$, $K_m = 1.37 \cdot 10^{-5} \text{ m/s}$, $A = 1.76 \cdot 10^{-2} \text{ m}^2$, $n = 12$, and $V = 1.5 \cdot 10^{-3} \text{ m}^3$.

The network having 12 neurons in the hidden layer and as threshold functions the sigmoid activation functions "tansig", "tansig", gave the best performance (smallest error with respect to the validation set), as it shown in Table 2.

$CE_m[\%]$	$CE_c[\%]$	$Y_{m\text{DADB}}[\%]$	$Y_{c\text{DADB}}[\%]$
87.28	87.85	75.29	75.32
94.38	93.66	85.78	85.63
90.19	91.28	76.25	75.99

Table 2. Results obtained during the validation procedure for the ANN which gave the best results (8-12-2 architecture, "tansig", "tansig").

It should be added that ANNs are not unique, that is many nets can produce identical outputs from prespecified inputs (Himmelblau, 2000).

4. Simulation results

Simulation studies were then performed in order to check if the ANN developed here could be used for the prediction of the current efficiency and product yield under other reactor operating conditions, outside its training regime.

Figures 2 to 5 present the results of our simulation studies. We studied the influence of current density at constant quantity of electricity passed as well as the influence of the electricity passed at constant current density on the current efficiency and DADB yield. The simulation results were consistent with the experimental results.

By varying the current density between 500 A/m^2 and 2000 A/m^2 , the response of the electrochemical process in respect with the current efficiency strongly depends on the quantity of electricity (Fig. 2).

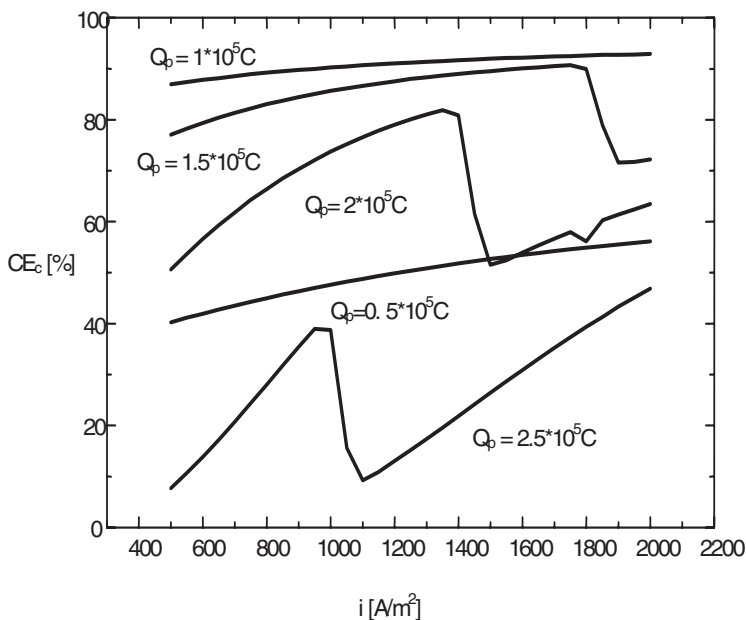


Fig. 2. Variation of the ANN predicted current efficiency with current density at constant amount of electricity passed.

By passing through the reactor of some relatively small quantities of electricity ($Q_p = 0.5 \cdot 10^5$ and $1 \cdot 10^5 \text{ C}$, respectively), an almost linear increase of the current efficiency with current density was observed. In this region the side reactions are less probable. By increasing the amount of electricity passed through the reactor (that means the increasing of the reaction time, too) the side reactions are more likely to happen. If the reaction time is long enough,

the quinoine-imine intermediates formed through the side reaction accumulated, their further electroreduction to the cathode was then possible.

As for the situation presented in Fig. 3, the conclusion was evident: ANN could detect the existence of an optimal reaction time. Moreover, as the maximum of the CE_c is obtained for current densities about 1600 A/m^2 , the simulation results are very consistent with the experimental ones. The experimental value for the optimum current density was 1550 A/m^2 . The same value was obtained using an optimization procedure (Oniciu, 1990).

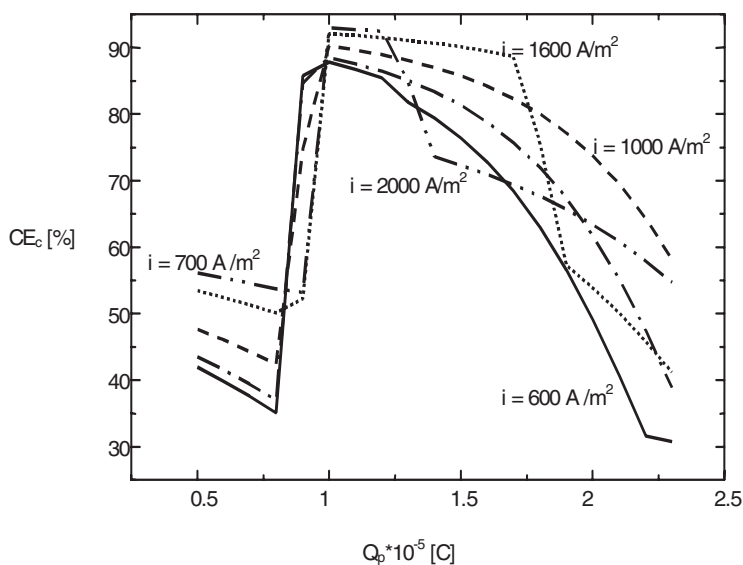


Fig. 3. Variation of the ANN predicted current efficiency with electricity passed at constant current density.

As regarding the influences on the DADB yield, the conclusion was: the maximum yield could be obtained in the conditions of a current density of about 1500 A/m^2 (the same value

like in the case of maximum current efficiency) but for $Q_p = 2 \cdot 10^5$ C, so the reaction time should be two times longer (Fig. 4 and 5). Of course, if the time is longer we could obtain a bigger amount of DADB, but this affects the selectivity (once the reaction time is longer the side reactions are more favorable).

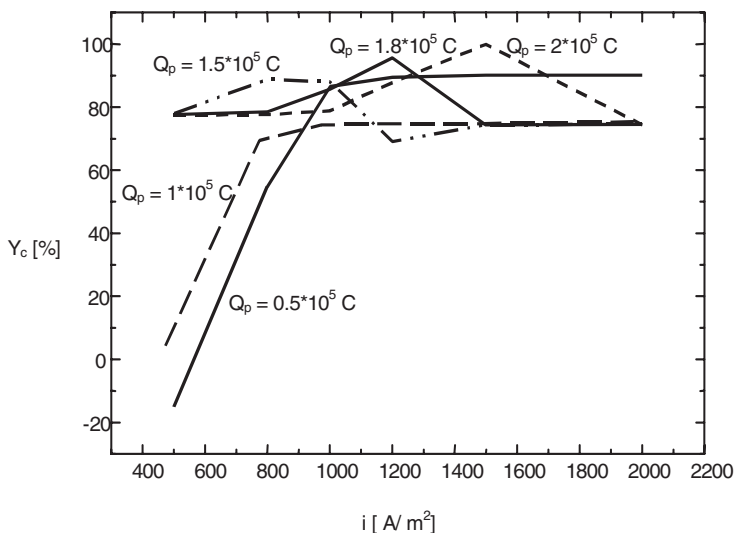


Fig. 4. Variation of the ANN predicted DADB yield with current density at constant amount of electricity passed.

We have found interesting features by analyzing the shape of the curves presented in Fig. 5. As it can be seen in Fig. 6, all these curves seem to have the same shape (they are almost parallel). We can say that in this region of current densities the reaction mechanism was the same. We are only talking about the direct electroreduction of the DNDB without considering the DNDB electrocatalytical hydrogenation. But, at higher current densities and

higher quantities of electricity passed, this process occurs and it is competitive with the H_2 formation process.

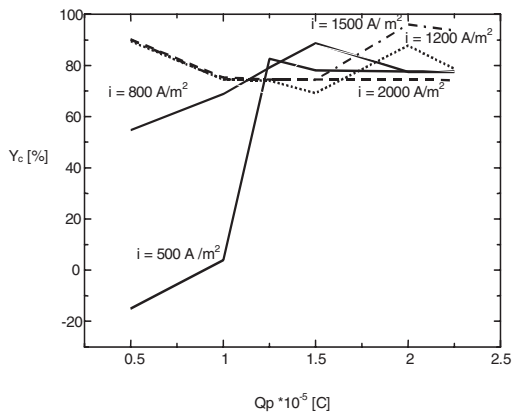


Fig. 5. Variation of the ANN predicted DADB yield with electricity passed at constant current density.

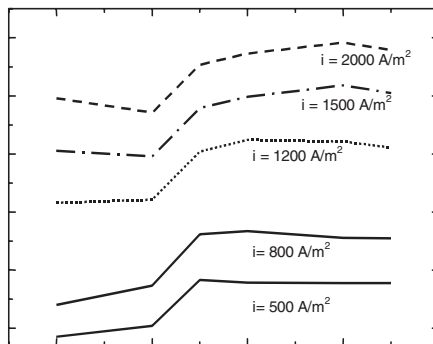


Fig. 6. The shape of the curves showed in Figure 5.

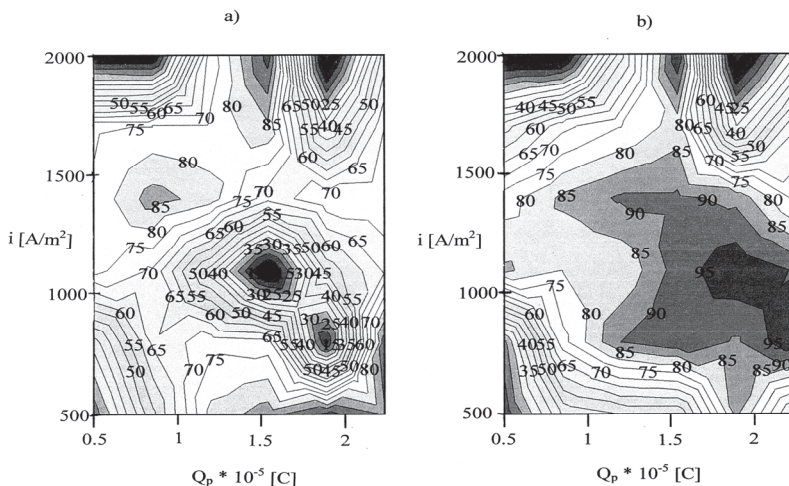


Fig. 7. a) Contour plots for $CE_c = f(i, Q_p)$; b) Contour plots for $Y_c = f(i, Q_p)$.

Next simulation step was to vary both the current density and the quantity of electricity passed.

Fig. 7 presents as contour plots, the 3D representations of CE_c and Y_c , respectively, versus Q_p and i . Simultaneously, a good current efficiency and a good yield could be obtained at Q_p about $0.8 \cdot 10^5$ C and i about 1480 A/m².

Experimental measurements of current efficiency and DADB yield performed using these values for input parameters revealed a good agreement between experimental and ANN predicted output data.

Therefore, the developed ANN proved to be able to predict the behavior of our complex reaction system.

5. Conclusion

This study proves that ANN can be successfully used for the simulation of the behavior of a complex electrochemical process, allowing in the same time the determination of the

optimal operation conditions. We have carried out simulations varying one by one the variables and the two variables at once, current density and quantity of electricity, in fact imposing different reaction rates and different reaction times. We observed that the predictions, which are in fact related to the conversion and the selectivity of the electrochemical process, improve significantly when the two variables were used for simulation.

References

- P. Bhagat: "An introduction to neural nets", *Chem. Eng. Progress*, 86 (1990) 55-58.
- V. Danciu: Ph.D. Thesis, "Babes-Bolyai" University, Cluj-Napoca, Romania, (1998).
- J. A. Freeman, D. M. Skapura: *Neural Networks: Algorithms, Applications and Programming Techniques*, Addison-Wesley, Reading, MA, 1991.
- M.T. Hagan, M. Menkaj: "Training feedforward networks with the Marquardt algorithm", *IEEE Transactions on Neural Networks*, 5 (1994) 989-993.
- D.M. Himmelblau: "Applications of artificial neural networks in chemical engineering", *Korean J. Chem. Eng.*, 17 (2000) 373-392.
- J.C. Hoskins, D.M. Himmelblau: "Artificial neural networks models of knowledge representation in chemical engineering", *Comp. Chem. Eng.*, 12 (1988) 881-885.
- Z-Y. Hou, Q. Dai, X-Q. Wu, G-T. Chen, *Appl. Catal. A.*, 161 (1997) 183-190.
- K. Huang, F-Q. Chen, D-W. Lü, *Appl. Catal. A.*, 219 (2001) 61-68.
- S. Kito, T. Hattori, Y. Murakani: "Estimation of catalytic performance by neural-network-product distribution in oxidative dehydrogenation of ethylbenzene", *Appl. Catal. A*, 114 (1994) 173-178.
- T.M. Mitchell: *Machine Learning*, WCB/McGraw Hill, Boston, 1997.
- M. Moliner, J.M. Serra, A. Corma, E. Argente, S. Valero, V.Botti, *Microporous and Mesoporous Materials*, 78 (2005) 73-81.
- M. Olea, Ph.D. Thesis, "Babes-Bolyai" University, Cluj-Napoca, Romania, 1996.
- K. Omata, Y. Kobayashi, M. Yamada, *Catal. Commun.* 6 (2005) 563-567.
- L. Oniciu, M. Olea, M. Schinteie: "Electrochemical reactor for p,p'-diaminodibenzyl electrosynthesis", *Revista de Chimie*, 41 (1990) 345-347.
- L. Oniciu, M. Olea, V. Danciu: "Determination of optimal parameters of p,p'-dinitrodibenzyl reduction to p,p'-diaminodibenzyl", *Revista de Chimie*, 41 (1990) 799-801.
- D.E. Rumelhart, G.E. Hinton, R.J. Williams: *Parallel Distributed Processing*, vol. 1, MIT Press, Cambridge, MA, 1986.
- M. Sasaki, H. Hamada, Y. Kintaichi, T. Ito, *Appl. Catal. A*, 132 (1995) 261-270.
- J.M. Serra, A. Corma, E. Argente, S. Valero, V. Botti, *Appl. Catal. A*, 254 (2003) 133-145.
- T. Hattori, S. Kito, *Catal. Today*, 23 (1995) 347-355.