ELSEVIER

Dyes and Pigments 42 (1999) 123–135

DYES and PIGMENTS

The use of artificial neural network (ANN) for modeling of the H_2O_2/UV decoloration process: part I

Yness March Slokar[a,*], Jure Zupan[b], Alenka Majcen Le Marechal[a]

[a]Faculty for Mechanical Engineering, Smetanova 17, 2000 Maribor, Slovenia
[b]National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia

Received 5 November 1998; accepted 18 January 1999

Abstract

A brief introduction into artificial neural networks (ANNs) is given, with emphasis on counter-propagation learning strategy, as well as their use for the purpose of modeling and optimization of H_2O_2/UV decoloration process. The use of Plackett–Burman partial factorial design for seven variables on three different levels, for the selection of experiments, needed to calculate the significance of variables, is described. Results of learning with Kohonen ANN are described, and the best prediction assembly suggested. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Artificial neural network; Decoloration; Variables; Partial factorial design; Ecological parameters; Modeling

1. Introduction

Effluents of dye manufacturers and dye user companies are usually highly polluted with different types of dyes. Even though there are not many dyes that are proven to be carcinogenic for humans, the extent of dyes in surface waters is rising, and methods for their removal need to be evaluated. Due to the large number of different dyes that are available on the market at present (over 3000) [1], it is almost impossible to find a perfect method which would satisfactorily purify the waste-waters, regardless of the chemical nature of the pollutant. Many possibilities have been reported [2], but these are more or less selective.

Even for the same dye, the decoloration process may depend on many different factors. For a decoloration process, one usually has to consider the time needed for the decoloration to be completed, i.e. the time needed for the dye to be either completely removed or to be removed up to a reasonable amount. It is desired that this time is as short as possible. For these reasons our research group wanted to optimize the decoloration process, which implies that a model to predict the time needed for decoloration to conclude had to be obtained.

Modeling of the decoloration process involves many problems, since the process depends on many factors, i.e. we are dealing with a multivariate system. Furthermore, the concentration of the dye is not the only parameter of interest; ecological parameters such as COD, BOD and TOC, are also important. This means our system is also a multi-response one. It is also evident that these problems cannot be solved by simple linear multivariate correlation [3]. Recently, artificial neural

* Corresponding author c/o Tiziana D'Adda, Strada D'Azeglio 55, 43100 Parma, Italy.

PII: S0143-7208(99)00022-4

networks (ANNs) represent a set of methods that may be useful in solving such problems [4–9].

Therefore, our modeling was done by means of ANNs; the procedures used and results obtained from this approach are shown in this present paper.

2. Theory

2.1. Artificial neural network

In standard terms, a model is a set of analytical functions, describing relationships between input and output variables, fitting experimental data. In other words, with inputs X^S and outputs T^S, we need to obtain f_i (X^S) functions, which will for each input vector X^S, represented with m variables ($x_1^S, x_2^S, ...x_i^S, ...x_m^S$), give values y_i^S, that will differentiate from the real outputs t_i^S as little as possible (Fig. 1). Models may be mechanical, empirical or based on artificial neural network.

It is difficult to describe, in a simple way, what ANNs are. The basic concept of ANN is an ensemble of artificial neurons, simulated on a computer. The function representing an artificial neuron in ANN acts similarly to the action of a biological neuron. It accepts signals (X^S) from neighbouring neurons, processes them in a predefined way, and, depending on the outcome of this processing, neuron j decides whether to 'fire' an output signal (y_j) or not (Fig. 2). The output signal can be either 0 or 1 when dealing with binary neurons, or any real value between 0 and 1 when dealing with 'real value' neurons, and is calculated as follows:

$$y_j^S = \text{out}_j = \frac{1}{1 + \exp\left(-\sum_{i=1}^{m} w_{ji} x_i^S\right)} \tag{1}$$

where

j = neuron j;

m = number of variables in input vectorX^S;

w_{ji} = weight in neuron j for i-th variable in input vector;

$x_i{}^S$ = ith variable in input vector X^S.

Weights w_{ji} are coefficients in the jth artificial neuron, which are analogous to synapse strengths between the axons and dendrites in real biological neurons. Therefore, each weight decides what proportion of the incoming signal will be transmitted into the neuron's body [10]. However, it should be pointed out, that all artificial neurons are simulated by the *same* function.

The essence of modeling with ANN is the 'supervised learning' of neurons. 'Supervised learning' defines the training procedure in which input and output (requested response) values are known in advance, and are both used during the learning phase. It basically means adjusting the

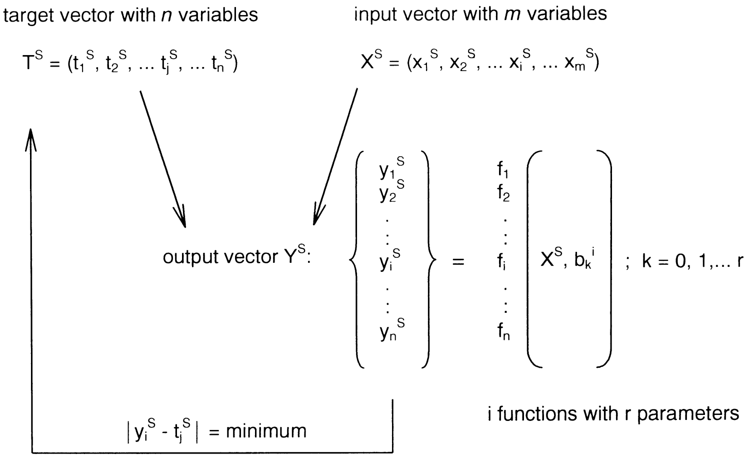

target vector with n variables
$T^S = (t_1^S, t_2^S, ... t_j^S, ... t_n^S)$
input vector with m variables
$X^S = (x_1^S, x_2^S, ... x_i^S, ... x_m^S)$
output vector Y^S:
y_1^S y_2^S ... y_i^S ... y_n^S = f_1 f_2 ... f_i ... f_n (X^S, b_k^i) ; k = 0, 1,... r
i functions with r parameters
$|y_i^S - t_j^S|$ = minimum

Fig. 1. The relation between input and output variables at modeling.

j..... neuron j

m .. number of variables in input vector X^S

w_{ji}.. weight in neuron j for i-th variable in input vector

x_i^S . i-th variable in input vector X^S

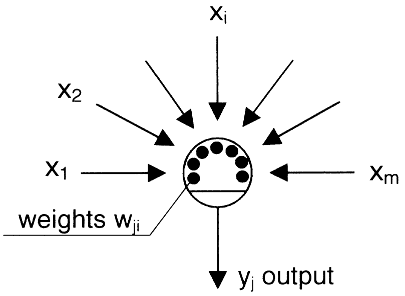


Fig. 2. Schematic representation of an artificial neuron *j*.

weights in neurons with regard to the difference between the outputs, predicted by the model (system), and the actual responses. The model corrects itself in such a way that in the next run the outputs will be predicted better, i.e. the difference between predicted outputs and actual responses will be smaller. This means that for each step, the performance of the model is improved with regard to the previous step. These steps are continuously repeated until

a. the difference for all input vectors X^S between the model and actual responses are within certain tolerance limits, or
b. the iteration quota (the number of learning attempts) is exceeded [11].

For the ANNs there are two supervised learning strategies, viz. (a) error back-propagation and (b) counter-propagation.

Although they have similar names, they are quite different when understanding the procedures for adapting weights in neurons. The former has a disadvantage in that the trained weights in neurons, representing the final synapse strengths, do not have any specific meaning or relevance to the problem the ANN is supposed to solve. This deficiency is not the case with the latter learning strategy, which is why it was used for the modeling used in our waste-water decoloration process. The disadvantage of the counter-propagation ANN is that, on average, it gives slightly worse responses than error back-propagation ANN.

The counter-propagation network is a multilayer one, namely, it has two layers, viz. (a) the input layer (also called the Kohonen layer) and (b) the output layer (Fig. 3). Each single neuron in the input layer is connected to the neuron in the output layer, at exactly the same topological location.

With counter-propagation learning strategy, both the input and target values are input into the network. The available data need to be divided into more sets [4]. These sets are called training, control (or fine tuning) and test set. As evident from the name, the first set is used for training the net. The weights in the input layer neurons are adapted according to the input vector X^S, while weights in the output layer neurons are adapted according to the target vector T^S. Each step of the counter-propagation learning strategy consists of four acts, viz.

i. input vector X^S enters the Kohonen layer and locates the central neuron with weights that are the most similar to the input variables [Eq. (2)];
ii. all the weights in the central neuron, and in a certain neighbourhood of a central neuron, are corrected according to Eq. (3);
iii. the position of the central neuron 'points out' the neuron in the output layer which the target vector T should enter;
iv. all weights in the central neuron and in a certain neighborhood of a central neuron in the output layer are corrected according to Eq. (4).

These four acts are repeated for each object–target pair until all the pairs are sent through the net—at this point, one *epoch* of learning is accomplished.

$$\text{neuron } c \leftarrow \min\left\{\sum_{i=1}^{m}\left(x_i^S - w_{ji}\right)^2\right\} \tag{2}$$

$$w_{ij}^{K(\text{new})} = w_{ji}^{K(\text{old})} + \eta(n_t)a\left(r_c - r_j, n_t\right)x\left(x_j^S - w_{ji}^{K(\text{old})}\right) \tag{3}$$

$$w_{ij}^{o(\text{new})} = w_{ji}^{o(\text{old})} + \eta(n_t)a\left(r_c - r_j, n_t\right)x\left(t_j^S - w_{ji}^{o(\text{old})}\right) \tag{4}$$

where

K = weights in the Kohonen layer;

o = weights in the output layer;

c = central neuron;

n_t = number of iteration steps;

$\eta(n_t)$ = function monotonically decreasing between the two pre-defined values $a_{\max}$ and $a_{\min}$ with increasing number of iteration steps;

$r_c - r_j$ = topological distance of neuron j from the central neuron c;

$a(r_c{-}r_j, n_t)$ = neighborhood function defining the percentage of actual correction (between 0 and 1) of the weight in neuron.

Eqs. (3) and (4) are practically the same, with the exception of the component used for correction calculation [x_j^s in Eq. (3) and t_j^s in Eq. (4) for correction of the weights in the Kohonen and output layer, respectively].

The larger the distance $r_c - r_j$, the smaller is the correction of weights. At the beginning of the learning phase, the correction encompasses the weights of all neurons in the network. During the learning process, the size of the neighborhood in which weights in neurons are corrected, is shrinking, until in the last epoch only the weights of the central neuron are corrected. Speed of shrinking depends on the number of iteration steps [11]. Additionally, the corrections are *decreasing* through the learning process.

The aim of this learning strategy is to make 'topologically close neurons', i.e. neurons that are

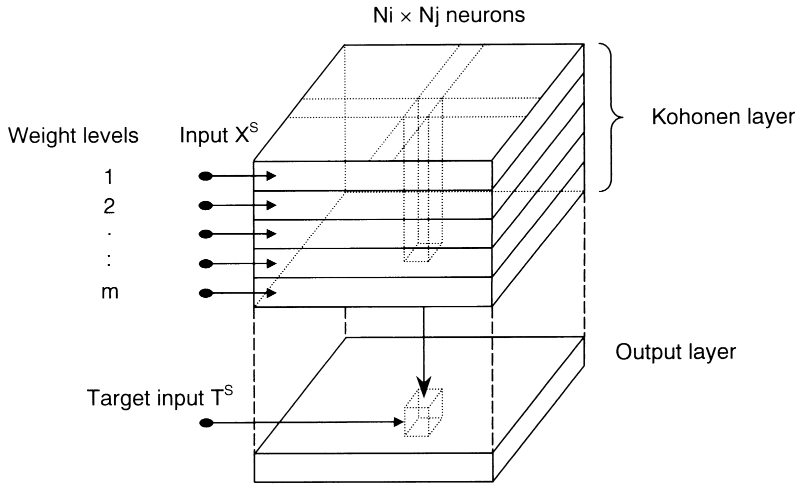


Fig. 3. Counter-propagation neural network architecture.

located physically closely respond to similar inputs, while neurons located far apart should respond to different inputs. Since, during the learning phase, the answers are distributed over the entire assembly of weights in the output layer neurons, this means that *each* output neuron 'learns' an answer, even if its 'pointer' in the Kohonen layer has never been excited during the learning. Each plain level of network therefore represents a two-dimensional map of a certain variable, and with overlapping of these maps, correlation among the variables can be obtained [10].

2.2. Experimental design

It is desired that the training set consists of such input data so that the individual objects are distributed over the entire measurement space as evenly as possible. This means that all the variables should be represented evenly between the possible maximum and minimum extreme values. Dyeing charts always tolerate the addition of individual chemicals added to the dye in a dyebath. If one would want to consider all the possible combinations of both dye concentration and concentrations of all the chemicals, thousands of experiments would need to be carried out to determine the significance of influences and to optimize the decoloration process.

In our previous work [12] we concluded there were seven principal variables influencing the decoloration process. In general, the more variables, the more complicated is the model, therefore, we tried to reduce the number of the variables. For this purpose we had to determine the significance of the influence on each of these seven variables. The influence of a variable is significant if the difference D_i [Eqs. (5) and (6)] is greater than experimental error. Since the number of experiments needed to be carried out increases exponentially with the number of variables, already, from four variables on, it is almost impossible to perform all the experiments that would cover the whole experimental area. Therefore an experimental design has to be made.

$$Di_{\max} = \sum_{j=1}^{4} \frac{yj^{(\text{at} \quad oi)}}{4} - \sum_{j=1}^{4} \frac{y_j^{(\text{at} \quad +i)}}{4} \tag{5}$$

and

$$Di_{\min} = \sum_{j=1}^{4} \frac{yj^{(\text{at} \quad oi)}}{4} - \sum_{j=1}^{4} \frac{y_j^{(\text{at} \quad -i)}}{4} \tag{6}$$

The experiments needed to determine the significance of the variables' influences were determined by Plackett–Burman partial factorial design (Fig. 4) [3].

In this experimental design, each of the variables is four times at nominal and four times at extreme level, of the latter once at maximum level and at another time at minimum level. This implies that for each variable the experiments are divided into two groups: (i) in the first group the variable of interest has nominal level and (ii) in the second group it has extreme level.

In each of the groups, all the other variables are twice at nominal and twice at extreme level. By comparing the averages the influences of all the rest of the variables are, therefore, zero. Influences are calculated as shown in Eqs. (5) and (6).

3. Experimental

As has already been pointed out, the decoloration of C.I. Reactive Red 120 [11] with the H_2O_2/UV process depends on seven variables, namely: (i) the intensity of UV lamp radiation; (ii) the addition of hydrogen peroxide; (iii) decoloration time; (iv) concentration of dye; (v) concentration of salt; (vi) concentration of an alkali; (vii) concentration of an oxidizing agent.[1]

The intensity of UV radiation (i), the addition of H_2O_2 (ii) and decoloration time (iii) depend on dyebath composition (iv)–(vii), and the latter depends on dye concentration in the dyebath, i.e. the desired depth of color. With decoloration processes, where named variables have different values, the responses (i)–(iv) also differ a great deal: (i) absorbance, (ii) COD (chemical oxygen demand), (iii) TOC (total organic carbon) and (iv) TIC (total inorganic carbon).

[1] Lamberti Redoks L2C (LR L2C), which needs to the added to the dyebath in order to prevent possible reduction of the dye during the dyeing process.

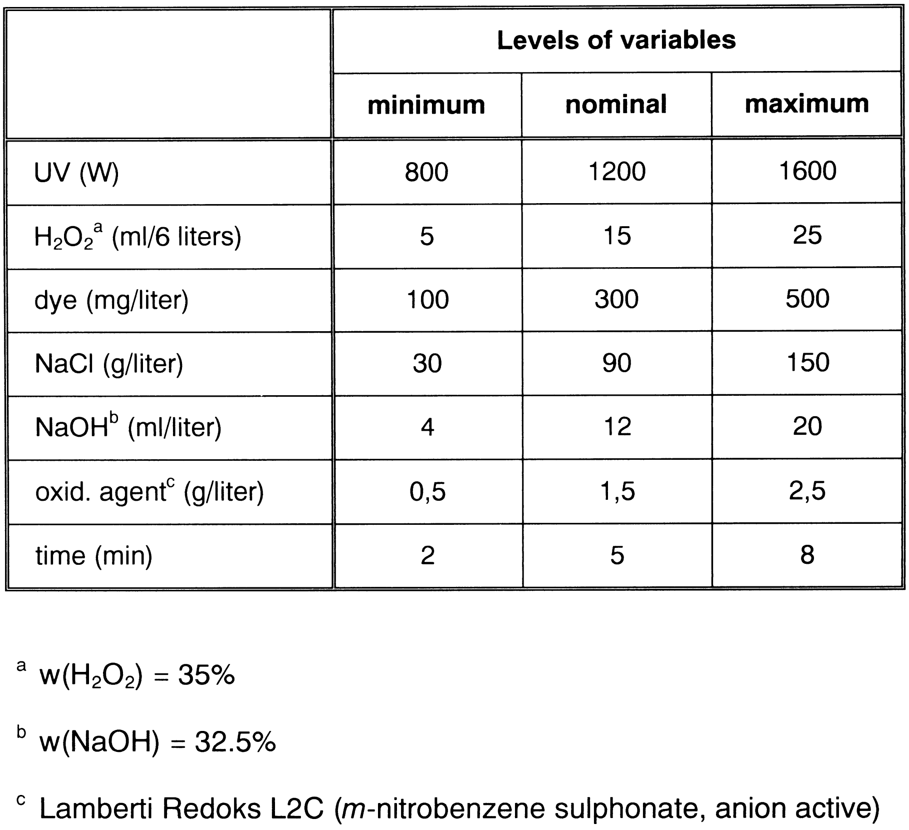
Experiment	Variables (factors)							Measurement
	A	B	C	D	E	F	G	(result)
1	o	o	o	o	o	o	o	y_1
2	o	o	+	o	+	+	+	y_2
3	o	+	o	+	o	+	+	y_3
4	o	+	+	+	+	o	o	y_4
5	+	o	o	+	+	o	+	y_5
6	+	o	+	+	o	+	o	y_6
7	+	+	o	o	+	+	o	y_7
8	+	+	+	o	o	o	+	y_8
9	o	o	−	o	−	−	−	y_9
10	o	−	o	−	o	−	−	y_{10}
11	o	−	−	−	−	o	o	y_{11}
12	−	o	o	−	−	o	−	y_{12}
13	−	o	−	−	o	−	o	y_{13}
14	−	−	o	o	−	−	o	y_{14}
15	−	−	−	o	o	o	−	y_{15}

o nominal level + maximum level − minimum level

Fig. 4. Adopted Plackett–Burman partial factorial design for seven variables, used for the determination of variables' influences in upper and lower intervals.

The purpose of modeling was, therefore, to optimize the procedure in such a way that it would be possible to 'prescribe' the appropriate intensity of UV-light and the addition of peroxide according to the dyebath composition that is actually used in the textile industry. With appropriate decoloration conditions one has in mind the *optimum* conditions in which the decoloration time would be as short as possible, and also the pertinent ecological parameters.

The levels of input variables are given in Fig. 5. The absorbance was measured spectrophotometrically on an HP Diode Array Spectrophotometer 8452A. The absorbance was measured



	Levels of variables		
	minimum	nominal	maximum
UV (W)	800	1200	1600
H_2O_2[a] (ml/6 liters)	5	15	25
dye (mg/liter)	100	300	500
NaCl (g/liter)	30	90	150
NaOH[b] (ml/liter)	4	12	20
oxid. agent[c] (g/liter)	0,5	1,5	2,5
time (min)	2	5	8

[a] $w(H_2O_2) = 35\%$

[b] $w(NaOH) = 32.5\%$

[c] Lamberti Redoks L2C (*m*-nitrobenzene sulphonate, anion active)

Fig. 5. The levels of input variables for the decoloration with H_2O_2/UV.

at the wavelength of maximum absorbance of the dye, i.e. $\lambda = 508$ nm. Ecological parameters COD, TOC and TIC were also determined.

Decolorations were performed on a Solvay Interox pilot plant [12].

4. Results

Values of absorbance and ecological parameters for the experiments carried out following Plackett–Burman partial factorial design are given in Fig. 6.

In Fig. 7 the influences of individual variables obtained from the difference between average values at nominal and maximum levels ($D_{i,\max}$) and at nominal and minimum levels ($D_{i},_{\min}$) are given.

For modeling, neural assemblies with different numbers of neurons were used. Also, the value of $a_{\max}$, as well as the number of iteration steps, has been changed (Fig. 8). All other conditions were the same all the time. As a result, we chose the assembly which with cross-validation showed the smallest differences between the predicted and the real target values.

The rest of the conditions, that were constant all the time were: number of weights in each neuron: *11*; toroid boundary conditions: *no*; minimal correction factor ($a_{\min}$) = 0.01; type of neighborhood correction: *triangular*; furthest neuron for corrections: *linearly decreasing with increasing number of epochs*; type of the type of the best match: *neuron with weights most similar to the input*.

It transpired that the results were best predicted with an ANN architecture of 8 × 8 neurons, the value of maximal correction factor $a_{\max}$ being 0.5 and at 200 epochs. The results of cross-validation are given in Fig. 9.

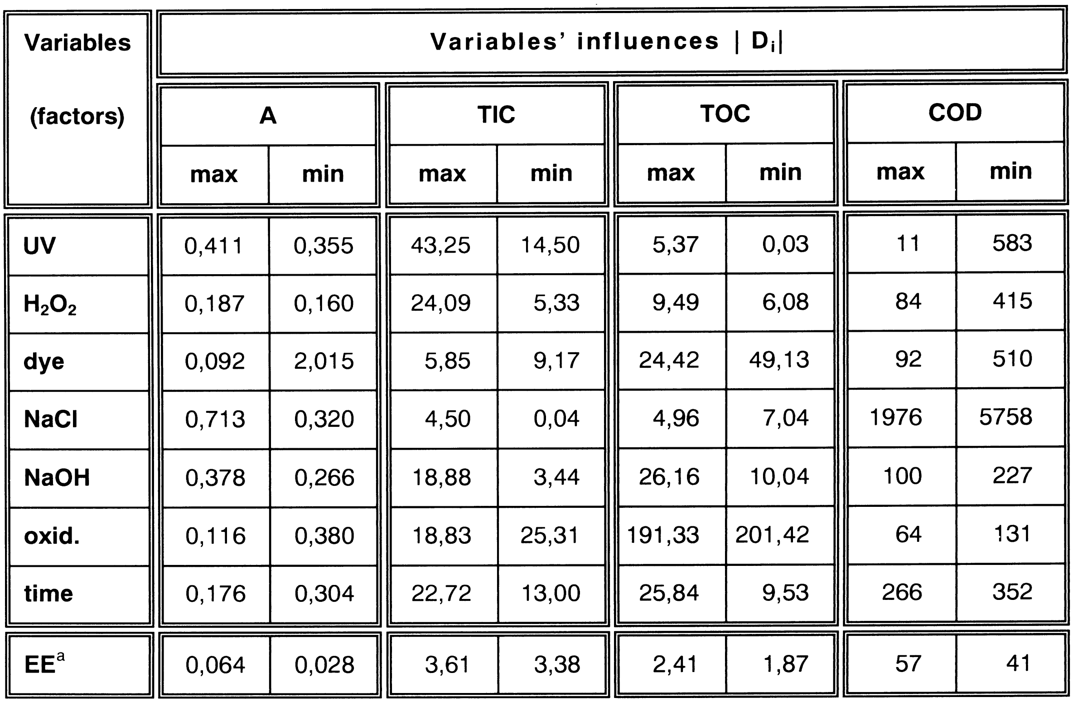
	Variables (factors)							Results			
Experiment	UV (W)	H_2O_2 (ml/6l)	dye (mg/l)	NaCl (g/l)	NaOH (ml/l)	oxid.a. (g/l)	time (min)	A[a] (-)	TIC (mg/l)	TOC (mg/l)	COD (mg/l)
1	1200	15	300	90	12	1,5	5	2,502	124,7	344,0	8800
2	1200	15	500	90	20	2,5	8	2,295	54,7	583,6	9190
3	1200	25	300	150	12	2,5	8	1,855	45,1	570,9	11090
4	1200	25	500	150	20	1,5	5	1,330	36,0	352,0	10880
5	1600	15	300	150	20	1,5	8	1,745	47,8	361,7	10920
6	1600	15	500	150	12	2,5	5	2,190	141,0	573,6	11010
7	1600	25	300	90	20	2,5	5	2,427	87,7	518,5	8760
8	1600	25	500	90	12	1,5	8	2,641	97,2	395,8	8910
9	1200	15	100	90	4	0,5	2	0,878	17,6	137,2	9250
10	1200	5	300	30	12	0,5	2	2,784	25,0	189,4	2340
11	1200	5	100	30	4	1,5	5	1,110	50,7	342,2	3560
12	800	15	300	30	4	1,5	2	2,840	46,9	390,2	2970
13	800	15	100	30	12	0,5	5	0,980	19,6	124,7	4200
14	800	5	300	90	4	0,5	5	2,781	20,0	182,9	9260
15	800	5	100	90	12	1,5	2	1,168	19,1	329,5	9240

[a] A - absorbance

Fig. 6. The input variables and corresponding output values obtained with H_2O_2/UV decoloration.

5. Discussion

It is evident from Fig. 7 that variables with influences on individual answers smaller than the experimental error are rare. None of the variables has a significant influence on all the answers. Dye has no influence on absorbance, but only for experiments with variables at maximum level. For the same experiments, UV irradiation has no influence on COD values. The latter have also no influence on COD values, and salt has no influence on TIC values, both valid for experiments with variables at minimum level. In other words, all of the variables have a significant influence on at least



Variables (factors)	Variables' influences | D_i |							
	A		TIC		TOC		COD	
	max	min	max	min	max	min	max	min
UV	0,411	0,355	43,25	14,50	5,37	0,03	11	583
H_2O_2	0,187	0,160	24,09	5,33	9,49	6,08	84	415
dye	0,092	2,015	5,85	9,17	24,42	49,13	92	510
NaCl	0,713	0,320	4,50	0,04	4,96	7,04	1976	5758
NaOH	0,378	0,266	18,88	3,44	26,16	10,04	100	227
oxid.	0,116	0,380	18,83	25,31	191,33	201,42	64	131
time	0,176	0,304	22,72	13,00	25,84	9,53	266	352
EE[a]	0,064	0,028	3,61	3,38	2,41	1,87	57	41

[a] EE – experimental error

Fig. 7. Influences of individual variables on H_2O_2/UV decoloration.

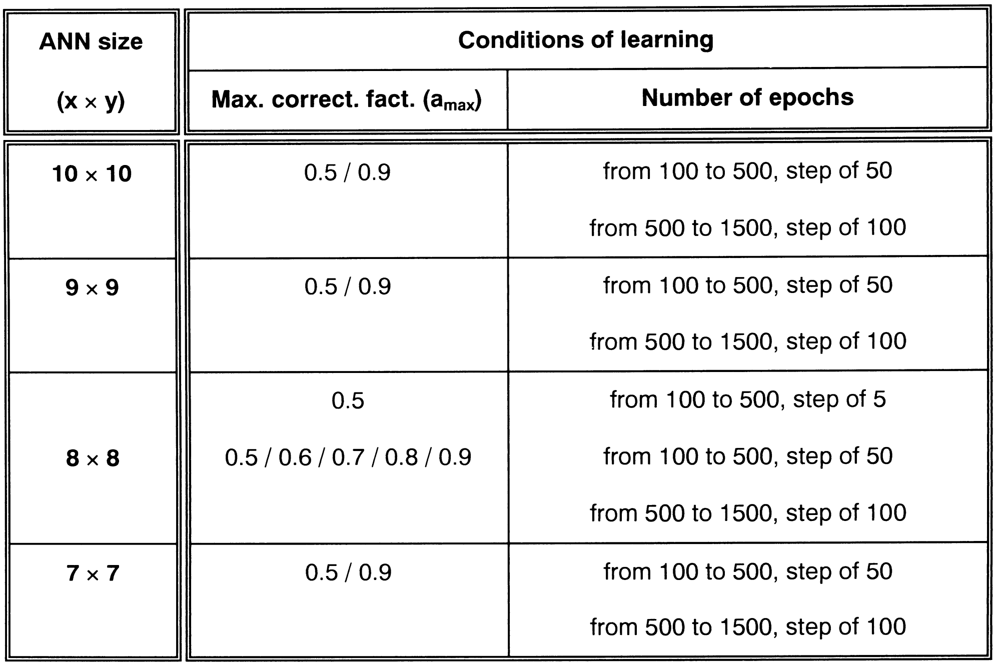
one of the answers. As a consequence, all seven variables had to be taken into account for modeling.

Although there are 49 neurons in the 7 × 7 assembly, it transpired that they were not enough for 15 experiments to be uniformly distributed over it. Regardless of the number of epochs and the value of maximal correction factor, in most of the cases, two or even three experiments fell on the same central neuron. Since the choice of a central neuron depends on weights most similar to the input, it is evident that some of the input–target pairs were simply too alike for the neurons in the 7 × 7 assembly to differentiate between them; this is not acceptable for modeling purposes.

With the rest of the assemblies, it turned out the outputs (Y_{model}) are predicted well, after a certain number of epochs is exceeded. In general the predictions of outputs are better with a higher value of maximal correction factor. With 'better predictions of outputs' one has in mind the prediction error, which should be as small as possible, and the 'stabilizing' of the network (predicting outputs that are totally equal to actual target values), which should be achieved as soon as possible. With $a_{\max} = 0.9$ this already happens at 250 epochs, while with $a_{\max} = 0.5$ the assembly needs twice as many iteration steps. This also means that the conditions with lower values of maximal correction factor are more strict.

Nevertheless, also below the stated number of epochs, the errors between predicted outputs and actual targets are very small. With $a_{\max} = 0.9$ the largest error is 6.1% (8 × 8 neurons), and regardless of the size of the architecture, all the errors are smaller than 0.6% at 150 epochs. The largest prediction error, with a value of maximal correction factor $a_{\max} = 0.5$, was made by 10 × 10 neurons network (approximately 26%). At 200 epochs the errors for all ANN fall below 4%.

Part of the KCTRCVF [12] program is also a cross-validation of the network. Cross-validation is a standard procedure for error estimation in a



ANN size	Conditions of learning	
(x × y)	Max. correct. fact. (a_{max})	Number of epochs

10 × 10	0.5 / 0.9	from 100 to 500, step of 50
		from 500 to 1500, step of 100

9 × 9	0.5 / 0.9	from 100 to 500, step of 50
		from 500 to 1500, step of 100

8 × 8	0.5
	0.5 / 0.6 / 0.7 / 0.8 / 0.9
		from 100 to 500, step of 5
		from 100 to 500, step of 50
		from 500 to 1500, step of 100

7 × 7	0.5 / 0.9	from 100 to 500, step of 50
		from 500 to 1500, step of 100

Fig. 8. The conditions of ANN learning.

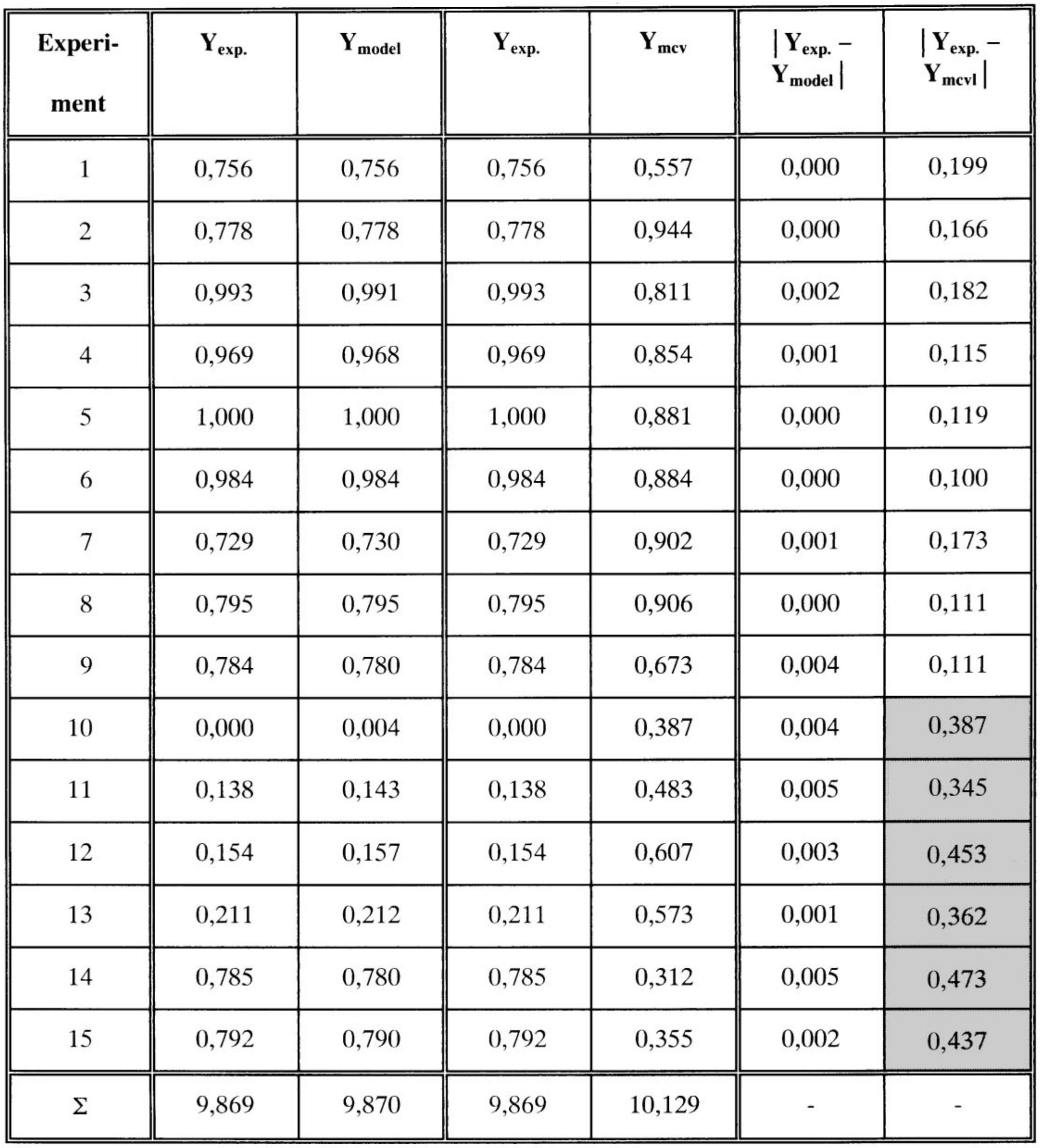
system with a small number of objects. Therefore, the KCTRCVF program enables the network to establish by self-checking, i.e. how well is it trained and how capable it is in predicting accurate outputs (Y_{mcv}). For all the assemblies it established, the output results are much better predicted for the experiments with variables at maximum level, as opposed to the predicted outputs for the experiments with variables at minimum level, i.e. in the latter case, the average errors were 23.2% larger.

Such results are probably the consequence of differences in the experimentally determined outputs (Y_{exp}). For experiments with variables on maximum level, these differences are much smaller than those for the experiments with variables on minimum level. In the former case, the difference between the largest and the smallest answer is 30%, while in latter case this difference is 80%. If one has in mind, that with cross-validation, the network predicts the outputs for the experiments that were left *out while learning*, then differences in output prediction errors for maximum and minimum level are easy explainable. If differences within outputs for variables with similar inputs are small, differences between Y_{exp} and Y_{mcv} will be in the same size class (maximum level), and if differences within outputs for variables with similar inputs are large, differences between Y_{exp} and Y_{mcv} will also be larger, respectively (minimum level).

Maximum correction factor a_{max} had no significant influence on learning of the 10 × 10 assembly. Regardless of its value, all the experiments are uniformly distributed over the assembly. In general, prediction errors for the experiments with variables in maximum level are smaller, with $a_{max} = 0.5$ compared to those with $a_{max} = 0.9$.

At 9 × 9 assembly, the influence of maximum correction factor is more perceivable. With $a_{max} = 0.5$, two experiments fall on the same central neuron more often than in the case with $a_{max} = 0.9$. Obviously corrections with lower a_{max} are too small for the neurons to differentiate between two different experiments with similar variables.

Learning of 8 × 8 assembly proved to be large enough for the experiments to be uniformly distributed over it, as long as the maximum correction factor and the number of epochs are adequate. As a consequence, it was necessary to determine the



Experiment	$Y_{exp.}$	Y_{model}	$Y_{exp.}$	Y_{mcv}	$\lvert Y_{exp.} - Y_{model} \rvert$	$\lvert Y_{exp.} - Y_{mcvl} \rvert$
1	0,756	0,756	0,756	0,557	0,000	0,199
2	0,778	0,778	0,778	0,944	0,000	0,166
3	0,993	0,991	0,993	0,811	0,002	0,182
4	0,969	0,968	0,969	0,854	0,001	0,115
5	1,000	1,000	1,000	0,881	0,000	0,119
6	0,984	0,984	0,984	0,884	0,000	0,100
7	0,729	0,730	0,729	0,902	0,001	0,173
8	0,795	0,795	0,795	0,906	0,000	0,111
9	0,784	0,780	0,784	0,673	0,004	0,111
10	0,000	0,004	0,000	0,387	0,004	0,387
11	0,138	0,143	0,138	0,483	0,005	0,345
12	0,154	0,157	0,154	0,607	0,003	0,453
13	0,211	0,212	0,211	0,573	0,001	0,362
14	0,785	0,780	0,785	0,312	0,005	0,473
15	0,792	0,790	0,792	0,355	0,002	0,437
Σ	9,869	9,870	9,869	10,129	-	-

Fig. 9. Results of cross-validation for the specified ANN and conditions used.

most appropriate value of the maximum correction factor. Learning of 8 × 8 assembly was therefore carried out under the same conditions as before, while $a_{\max}$ was changed from 0.5 to 0.9, with the step of 0.1.

The results showed that the amplitude of the prediction error oscillation increases with increasing maximum correction factor value, which confirmed earlier findings. With increasing number of epochs, the oscillation amplitude becomes even more evident. In view of these observations, $a_{\max} = 0.5$ was chosen as the most appropriate value of the correction factor, while the number of epochs should be less than 500. Nevertheless, 'less than 500' is a very wide range, so the 8 × 8 assembly with $a_{\max} = 0.5$ was run with 100–500 epochs, step of 5.

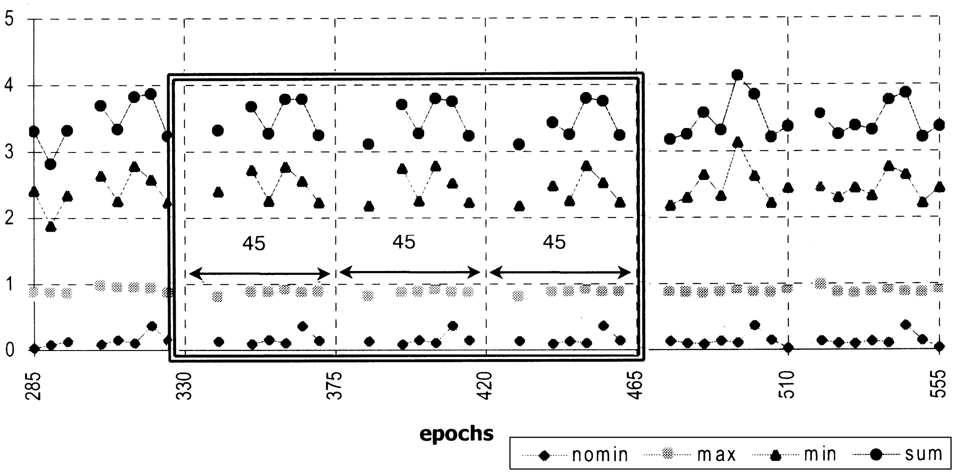


Fig. 10. Cross-validation errors for Kohonen network (8 × 8 assembly, $a_{\max} = 0.5$, epochs in step of 5).

After exceeding 300 epochs, the magnitude of errors starts to follow a certain pattern, which is observable at all the levels (nominal, maximal and minimal). The pattern is very apparent in Fig. 10.

Experiments 9 and 12 fall on the same central neuron at 330, 375, 420 and 465 epochs, i.e. every 45 epochs. Experiments 13 and 15 fall on the same central neuron at 335, 345, 380, 390, 425, 435 and 470 epochs, i.e. every 10 + 35 = 45 epochs. Moreover, the prediction error for experiments with variables on nominal level reaches relative extreme value every 45 epochs. The pattern is therefore repeated every 45 epochs. If the pattern starts at 330 epochs, it can be concluded that from this number of repetitions on prediction, errors are more or less indifferent to the number of iteration steps. Therefore it was reasonable to look for optimal solution below 300 epochs.

Consequently it was established that the sum of errors at predicting outputs after cross-validation is the least at 200 epochs, which is why the 8 × 8 assembly, with $a_{\max} = 0.5$ and 200 epochs, was chosen as the best one for the H_2O_2/UV decoloration model.

6. Conclusions

Decoloration of C.I. Reactive Red 120 with hydrogen peroxide, activated with UV beams was evaluated with respect to seven factors. Mostly, these factors have a negative influence, which means that the decoloration process is slowed down. It is desirable for the industrial treatment of waste waters to be reasonable from both an ecological and an economical point of view. Consequently the decoloration process needs to be optimized, i.e. we have attempted to prescribe the best conditions for the existing treatment. The model that needs to be prescribed prior to optimization was to be effected by means of the artificial neural network. Since modeling becomes more complicated with increasing numbers of variables, the reduction of the latter was desirable, which is why the significance of the variables needed to be determined. Variables, that turned out as significant ones, were introduced to a Kohonen network.

As responses, absorbance was chosen, because it gives information on the dye still remaining in the treated water, as well as ecological parameters (TIC, TOC and COD), that are considered to give information on the toxicity of the waste-water. On the basis of the results obtained, the following conclusions can be made:

- After 8 min, the absorbance for all the experiments diminishes. Reduction of the absorbance is larger for experiments with variables on the minimal level, while for all the experiments the reduction is larger with more hydrogen peroxide, less alkali and less oxidizing agent (LR L2C) present in the bath. Alkali and oxidizing agent have the greatest influence on the absorbance.

- As expected, TOC values are greater in baths with more dye and more oxidizing agent present; both substances cause higher values of TIC.
- After 8 min, COD values remain practically the same, which probably means, that the time is not sufficient for the dye to degrade into ecologically unobjectionable products, but only up to triazine derivatives and/or other aromatic amines, which are ecologically objectionable. The greatest influence on COD is salt.
- Influences of all the variables are greater than the experimental bias, which means that all of them are significant, and need to be taken into consideration for optimization.
- Kohonen assembly with 7 × 7 neurons turned out to be too small for all the 15 experiments to evenly distribute over it; 9 × 9 and 10 × 10 assemblies are suitable, but not optimal.
- The best prediction results for H_2O_2/UV decoloration were obtained with a network of 64 (8 × 8) neurons, with 200 epochs and a value of the maximal correction factor 0.5.
- The estimated error, using cross-validation, is 20%.

Acknowledgements

Financial support of the Ministry of Science and Technology, Slovenia, through grants J1-8900-104 and R9-2481-0795, is gratefully acknowledged.

References

[1] Richardson ML. Risk assessment of chemicals in the environment. Royal Society of Chemistry 1988:398–413.

[2] Slokary M, Majcen Le Marechal A. Methods of decoloration of textile wastewaters. Dyes and Pigments 1998;37(4):335–56.

[3] Massart DL, Vandeginste BGM, Deming SN, Michotte Y, Kaufmann L. Chemometrics: a textbook. Amsterdam: Elsevier, 1988.

[4] Zupan J, Gasteiger J. Neural networks: a new method for solving chemical problems or just a passing phase? (a review). Analytica Chimica Acta 1991;248:1–30.

[5] Gasteiger J, Zupan J. Neuronale Netze in der Chemie. Angewandte Chemie 1993;105:510–36.

[6] Gasteiger J, Zupan J. Neural networks in chemistry. Angewandte Chemie International Edition in English 1993;32:503–27.

[7] Kateman G. Neural networks in analytical chemistry. Chemometrics and Intelligent Laboratory Systems 1993;19:135–42.

[8] de B Harrington P. Minimal neural networks: differentiation of classification entropy. Chemometrics and Intelligent Laboratory Systems 1993;19:142–54.

[9] Zupan J, Gasteiger J. Neural networks for chemists: an introduction. Weinheim VCH, 1993.

[10] Zupan J. Introduction to artificial neural network (ANN) methods: what they are and how to use them. Acta Chimica Slovenica 1994;41:327–52.

[11] Zupan J, Novič M, Gasteiger J. Neural networks with counter-propagation learning strategy used for modeling. Chemometrics and Intelligent Laboratory Systems 1995;27:175–87.

[12] Majcen Le Marechal A, Slokar YM, Taufer T. Decoloration of chlorotriazine reactive azo dyes with H_2O_2/UV. Dyes and Pigments 1997;33(4):281–98.

[13] Zupan J. Program K-CTR for Kohonen and counter-propagation ANNs. Report KI DN-1452, 1994.