

Genetic Algorithms and Neural Networks Based Optimization Applied to the Wastewater Decolorization by Photocatalytic Reaction

GABRIEL DAN SUDITU¹, MARIUS SECULA¹, CIPRIAN GEORGE PIULEAC¹, SILVIA CURTEANU^{1*}, IOANNIS POULIOS²

¹"Gh. Asachi" Technical University Iasi, Department of Chemical Engineering, 71A D. Mangeron, Blvd., 700050, Iasi, Romania

²Aristotle University of Thessaloniki, Department of Chemistry, Lab. Physical Chemistry, 54124, Thessaloniki, Greece

This paper proposes a genetic algorithm and a neural network based procedure to estimate the optimal conditions for a dyestuff wastewater treatment process consisting of a heterogeneous photocatalytic oxidation. A simulated dyestuff effluent containing the azo dye Reactive Black 5 is decolorized by a photocatalytic reaction using TiO₂ P-25 as catalyst in the presence of Fe⁺³ and H₂O₂. A simple feed forward neural network with one hidden layer was projected and used to predict the evolution in time of the decolorization of this type of wastewater. The neural model was included in the optimization procedure solved with a simple genetic algorithm. The goal of the optimization is to calculate the optimal reaction conditions (illumination time and amounts of reagents) which assure an imposed value for the transmittance.

Keywords: wastewater, photocatalysis, Reactive Black 5 dye, neural network modeling, genetic algorithm optimization.

Heterogeneous and homogeneous solar photocatalytic detoxification methods (TiO₂/H₂O₂, Fe⁺³/H₂O₂) have shown recently great promise for the treatment of industrial wastewater, groundwater and contaminated air. Additionally, the semiconductor mediated photocatalytic process has shown also great potential for disinfection of air and water, thus making possible a number of applications [1-3]. General description of heterogeneous and homogeneous photocatalysis under artificial or solar irradiation is presented in several excellent review articles [4-7].

The photocatalytic degradation of azo dyes containing different functionalities using TiO₂ or Photo-Fenton reagent as photocatalysts in aqueous solutions under solar and UV-A irradiation has been described in a series of researches [8-10]. The combination of both processes, as has been reported by other authors [11, 12], leads to an enhancement of the removal rate of the pollutants due to the fact that Fe⁺³ ions and H₂O₂ act as scavengers of the electrons, which are photogenerated in the conduction band of TiO₂, while the produced Fe⁺² ions can participate again in the Fenton reaction.

The phenomenological treatment of such photochemical systems are very complex. In general, the rate of reaction in heterogeneous photocatalytic systems is a complex nonlinear function of catalyst loading, light intensity, initial solution pH, reactant and oxidants concentration etc. Due to these reasons, the ability of systems such as neural networks (NN) to recognize and reproduce cause-effect relationships through training, for multiple input-output systems, has gained popularity recently, in various areas of chemical engineering, also in the field of photocatalytic treatment of wastewater [13-16].

In recent years, neural networks have been much studied because of their capability to approximate any continuous nonlinear functions. Neural networks possess the ability to learn what happens in a process without actually modeling the physical and chemical laws that govern the system [17]. The use of neural networks has become increasingly recommended for applications

where the mechanistic description of the interdependence between variables is either unknown or very complex. They are now the most popular artificial tool with applications in areas such as pattern recognition, classification, process control, optimization [18-20]. Different types of neural network applications are reviewed in our precedent work [21].

In correlation with the aspects considered in the present paper, several applications of neural networks can be mentioned.

D. Salari et al. [22] used a three-layer neural network to predict the methyl tert-butyl ether concentration in a photooxidative degradation process in the presence of H₂O₂ under UV light illumination. S. Göb et al. [23] has been developed an empirical model based on artificial neural networks for fitting the experimental data obtained in a laboratory batch reactor for the degradation of 2,4-dimethyl aniline (2,4-xyldine), chosen as a model pollutant. A. Duran et al. [24] applied neural network modeling to the degradation of Reactive Blue 4 dye solutions in order to evaluate the use of the Fenton reagent under UV irradiation conditions. In a recent paper, A. Duran and J.M. Monteagudo [25] determine the influence of four parameters (pH and initial concentrations of TiO₂, Fe(II) and H₂O₂) on the value of the decoloration kinetic rate constant using neural networks.

The solution of an optimisation problem can be found through, among others, deterministic or stochastic approaches. The former composes the traditional optimisation methods (direct and gradient-based methods) and have the disadvantages of requiring the first and/or second-order derivatives of the objective function and/or constraints or of being not efficient in non-differentiable or discontinuous problems. Furthermore, the deterministic methods are dependent on the chosen initial solution and can tend to converge towards local extrema of the fitness function, which is clearly unsatisfactory for problems where the fitness varies non-monotonously with the parameters. The stochastic methods, such as genetic algorithms (GA), do not possess these drawbacks. GAs are

* email: silvia_curteanu@yahoo.com

part of the so-called evolutionary algorithms and compose a search and optimisation tool with increasing application in scientific problems. They do not need to have any information about the search space, just needing an objective/fitness function that assigns a value to any solution [26].

Because of their flexibility, ease of operation, minimal requirements and global perspective, GAs have been successfully used in a wide variety of problems [27]. GA has found various applications in chemical engineering including process control, gas pipeline design, pattern recognition of multivariate chemical data, optimization of reaction rate parameters, multipurpose chemical batch plant design, and scheduling [28]. Details about different types of genetic algorithms - simple genetic algorithm or different adaptations for problems with multiple constraints - and a series of their applications in optimization of the chemical processes has been described in the reviews of Bhaskar [29] and Nandasana [30].

The main goal of this paper is to develop a general procedure based on neural networks and genetic algorithms which could be applied to complex optimization problems.

A series of experiments have been carried out to study the decolorization of a synthetic dye stuff effluent containing the Reactive Black 5 (RB5) as a model azo dye, by photocatalytic degradation using TiO_2 P-25 as catalyst in the presence of Fe^{+3} and H_2O_2 . Neural networks are used as efficient modeling tool and genetic algorithms as solving method of optimization; a neural model of the process, trained with experimental data, is included in the optimization procedure. The case study is the estimation of the evolution in time of a dyestuff wastewater treatment process consisting of a photocatalytic oxidative reaction. It is followed a certain decolorization degree related to the optimal working conditions: illumination time, amounts of catalyst TiO_2 P-25, H_2O_2 and Fe^{+3} . The GA optimization procedure has proved easily to apply with useful and accurate results.

Experimental part

Reagents

The synthetic wastewater (CSW) was made according to the recipe used for dyeing of cotton fabrics and was of the following composition: 0.07 g L^{-1} Reactive Black 5, 0.1 g L^{-1} HCOOH , 0.250 g L^{-1} Perydrol FHB 15 %, 0.375 g L^{-1} Sequon, 0.5 g L^{-1} Na_2CO_3 , 0.5 g L^{-1} NaOH , 3 g L^{-1} NaCl . The CSW has an initial pH value of 12 and a DOC of 0.15 g L^{-1} (0.5 g L^{-1} COD and 0.12 g L^{-1} BOD₅).

The chemical structure of Reactive Black 5 ($\text{C}_{26}\text{H}_{21}\text{N}_5\text{O}_{19}\text{S}_6\text{Na}_4$) is given in figure 1.

TiO_2 P-25 of Degussa (anatase/rutile = 3.6/1, surface area $56 \text{ m}^2 \text{ g}^{-1}$ nonporous) was used for all photocatalytic experiments. All other reagent-grade chemicals, such as $\text{FeCl}_3 \cdot 6 \text{ H}_2\text{O}$, H_2O_2 , etc., were purchased through Merck and were used without further purification.

Procedures and analysis

Experiments were carried-out in a closed Pyrex cell of 500 ml capacity, provided with ports, at the top, for bubbling

air necessary for the reaction to take place. The reaction mixture was maintained as suspension by magnetic stirring. Previous to irradiation, the reaction mixture was left 15 minutes in the dark with the aim at achieving the maximum adsorption of the dye onto the semiconductor catalyst surface. The irradiation was performed with a 9 tral lamp. The spectral response of the irradiation source (Osram Dulux S 9W/78 UV-A) according to the producer is ranged between 340 and 400 nm, with a maximum at 366 nm and two additional weak lines in the visible region. The photon flow per unit volume of the incident light was determined by chemical actinometry using potassium ferrioxalate. The initial light flux, under exactly the same conditions as in the photocatalytic experiments, was evaluated to be $7.16 \cdot 10^{-4} \text{ Einstein min}^{-1}$.

In all cases, during the experiments, 500 ml acidified ($\text{pH}_0 = 3.2$) dye solution containing the appropriate amounts of semiconducting powder, H_2O_2 and Fe^{+3} was magnetically stirred before and during irradiation. Specific quantities of samples were withdrawn at periodic intervals and filtered through a $0.45 \mu\text{m}$ filter (Schleicher and Schuell) in order to remove the catalyst particles. With the aim at assessing the extent of color removal, changes in the concentration of the dye were observed from its characteristic absorption band using a UV-VIS spectrophotometer Shimadzu UV-160 A. Since a linear dependence between the initial concentration of the dye solution and his optical density at 585 nm was observed, the photodecomposition was monitored spectrophotometrically at this wavelength. The pH values of the solution were monitored with a Metrohm pH-meter, while the reaction temperature was kept constant at $25 \pm 0.1 \text{ }^\circ\text{C}$.

Results and discussions

Processing of experimental data

Table 1 presents the experimental results obtained in different conditions of initial RB5, TiO_2 P-25, H_2O_2 and Fe^{+3} concentrations.

In the first column of the table 1 is given the time intervals for collecting experimental data. For the first ten minutes, the samples were collected from two to two minutes, then the time interval increases to five minutes. Table 1 contains, in the last column, the RB5 concentration from the sample, calculated as function of transmittance and using an initial reference curve.

For the 16 series of experimental data in table 1, the influence of reaction conditions on the rate of dye elimination was studied using different graphical representations.

Figures 2 and 3 render the influence of the ratio between H_2O_2 and Fe^{+3} in the presence of 0.25 g L^{-1} TiO_2 P-25 (fig. 2) and 1 g L^{-1} TiO_2 P-25 (fig. 3), on the decolorization process. The best results in both cases – shortest reaction time - correspond to maximum amounts of the components, respectively 800 mg L^{-1} H_2O_2 and 56 mg L^{-1} Fe^{+3} .

The influence of TiO_2 P-25 amount on the dye elimination rate is illustrated in figure 4 and the best result corresponds to 1.08 g L^{-1} .

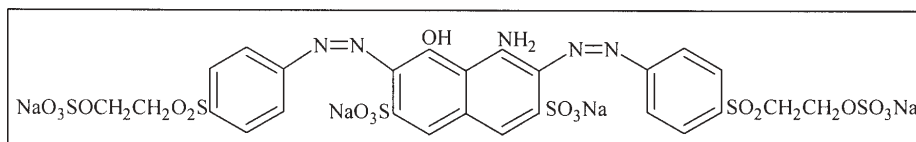


Fig. 1. The structure of Reactive Black 5

Table 1
RESULTS OF THE PHOTOCATALYTIC DECOLORIZATION OF A SIMULATED
DYE STUFF EFFLUENT UNDER VARIOUS EXPERIMENTAL CONDITIONS ($\text{pH}_0 = 3.2$)

Time [min]	TiO ₂ P-25 [g L ⁻¹]	H ₂ O ₂ [mg L ⁻¹]	Fe ⁺³ [mg L ⁻¹]	Transmittance [%]	RB5 concentration [mg L ⁻¹]
Series 1. 70 mg L ⁻¹ RB5 + 1 g L ⁻¹ TiO ₂ P-25 + 800 mg L ⁻¹ H ₂ O ₂ + 56 mg L ⁻¹ Fe ⁺³					
0	1	800	56	6	54.94
2	1	800	56	84	3.4
4	1	800	56	97.1	0.59
6	1	800	56	97.2	0.59
8	1	800	56	97.3	0.59
10	1	800	56	100	0
Series 2. 70 mg L ⁻¹ RB5 + 0.25 g L ⁻¹ TiO ₂ P-25 + 800 mg L ⁻¹ H ₂ O ₂ + 56 mg L ⁻¹ Fe ⁺³					
0	0.25	800	56	5	58.5
2	0.25	800	56	66	8.11
4	0.25	800	56	92	1.63
6	0.25	800	56	97.1	0.59
8	0.25	800	56	97.2	0.59
10	0.25	800	56	97.3	0.59
15	0.25	800	56	98	0.39
20	0.25	800	56	98.1	0.39
25	0.25	800	56	100	0
Series 3. 70 mg L ⁻¹ RB5 + 1 g L ⁻¹ TiO ₂ P-25 + 800 mg L ⁻¹ H ₂ O ₂ + 7 mg L ⁻¹ Fe ⁺³					
0	1	800	7	6	54.94
2	1	800	7	56	11.32
4	1	800	7	90	2.06
6	1	800	7	97	0.59
8	1	800	7	99	0.2
10	1	800	7	100	0
Series 4. 70 mg L ⁻¹ RB5 + 0.25 g L ⁻¹ TiO ₂ P-25 + 800 mg L ⁻¹ H ₂ O ₂ + 7 mg L ⁻¹ Fe ⁺³					
0	0.25	800	7	5	58.5
2	0.25	800	7	32	22.5
4	0.25	800	7	73	6.15
6	0.25	800	7	92	1.63
8	0.25	800	7	96	0.8
10	0.25	800	7	99	0.2
15	0.25	800	7	100	0
Series 5. 70 mg L ⁻¹ RB5 + 1 g L ⁻¹ TiO ₂ P-25 + 200 mg L ⁻¹ H ₂ O ₂ + 56 mg L ⁻¹ Fe ⁺³					
0	1	200	56	6	54.94
2	1	200	56	80	4.36
4	1	200	56	96	0.8
6	1	200	56	97	0.59
8	1	200	56	98	0.39
10	1	200	56	99	0.2
15	1	200	56	100	0
Series 6. 70 mg L ⁻¹ RB5 + 0.25 g L ⁻¹ TiO ₂ P-25 + 200 mg L ⁻¹ H ₂ O ₂ + 56 mg L ⁻¹ Fe ⁺³					
0	0.25	200	56	5	58.5
2	0.25	200	56	26	26.31
4	0.25	200	56	65	8.41
6	0.25	200	56	86	2.95
8	0.25	200	56	90	2.06
10	0.25	200	56	95	1
15	0.25	200	56	96	1
20	0.25	200	56	97	0.59
30	0.25	200	56	99	0.2
40	0.25	200	56	100	0
Series 7. 70 mg L ⁻¹ RB5 + 1 g L ⁻¹ TiO ₂ P-25 + 200 mg L ⁻¹ H ₂ O ₂ + 7 mg L ⁻¹ Fe ⁺³					
0	1	200	7	6	54.94
2	1	200	7	36	19.95
4	1	200	7	79	4.6
6	1	200	7	90	2.06
8	1	200	7	94	1.21
10	1	200	7	96	0.8
15	1	200	7	98	0.39
20	1	200	7	100	0

Series 8. 70 mg L ⁻¹ RB5 + 0.25 g L ⁻¹ TiO ₂ P-25 + 200 mg L ⁻¹ H ₂ O ₂ + 7 mg L ⁻¹ Fe ⁺³					
0	0.25	200	7	5	58.5
2	0.25	200	7	26	26.31
4	0.25	200	7	70	6.97
6	0.25	200	7	88	2.5
8	0.25	200	7	95	1
10	0.25	200	7	97	0.59
15	0.25	200	7	100	0
Series 9. 70 mg L ⁻¹ RB5 + 1.0806 g L ⁻¹ TiO ₂ P-25 + 500 mg L ⁻¹ H ₂ O ₂ + 31.5 mg L ⁻¹ Fe ⁺³					
0	1.0806	500	31.5	7	51.93
2	1.0806	500	31.5	86	2.95
4	1.0806	500	31.5	97	0.59
6	1.0806	500	31.5	98	0.39
8	1.0806	500	31.5	99	0.2
10	1.0806	500	31.5	100	0
Series 10. 70 mg L ⁻¹ RB5 + 0.169375 g L ⁻¹ TiO ₂ P-25 + 500 mg L ⁻¹ H ₂ O ₂ + 31.5 mg L ⁻¹ Fe ⁺³					
0	0.169375	500	31.5	5	58.5
2	0.169375	500	31.5	39	18.39
4	0.169375	500	31.5	85	3.17
6	0.169375	500	31.5	96	0.8
8	0.169375	500	31.5	98	0.39
10	0.169375	500	31.5	99	0.2
15	0.169375	500	31.5	99.1	0.2
20	0.169375	500	31.5	100	0
Series 11. 70 mg L ⁻¹ RB5 + 0.625 g L ⁻¹ TiO ₂ P-25 + 500 mg L ⁻¹ H ₂ O ₂ + 61.2675 mg L ⁻¹ Fe ⁺³					
0	0.625	500	61.2675	5	58.5
2	0.625	500	61.2675	84	3.4
4	0.625	500	61.2675	96	0.8
6	0.625	500	61.2675	98	0.39
8	0.625	500	61.2675	100	0
Series 12. 70 mg L ⁻¹ RB5 + 0.625 g L ⁻¹ TiO ₂ P-25 + 500 mg L ⁻¹ H ₂ O ₂ + 1.7325 mg L ⁻¹ Fe ⁺³					
0	0.625	500	1.7325	5	58.5
2	0.625	500	1.7325	14	38.39
4	0.625	500	1.7325	37	19.42
6	0.625	500	1.7325	60	9.98
8	0.625	500	1.7325	80	4.36
10	0.625	500	1.7325	90	2.06
15	0.625	500	1.7325	94	1.21
20	0.625	500	1.7325	98	0.39
30	0.625	500	1.7325	100	0
Series 13. 70 mg L ⁻¹ RB5 + 0.625 g/L TiO ₂ P-25 + 864.5 mg L ⁻¹ H ₂ O ₂ + 31.5 mg L ⁻¹ Fe ⁺³					
0	0.625	864.5	31.5	6	54.94
2	0.625	864.5	31.5	84	3.4
4	0.625	864.5	31.5	98	0.39
6	0.625	864.5	31.5	99	0.2
8	0.625	864.5	31.5	100	0
Series 14. 70 mg L ⁻¹ RB5 + 0.625 g L ⁻¹ TiO ₂ P-25 + 135.5 mg L ⁻¹ H ₂ O ₂ + 31.5 mg L ⁻¹ Fe ⁺³					
0	0.625	135.5	31.5	5	58.5
2	0.625	135.5	31.5	57	10.98
4	0.625	135.5	31.5	85	3.17
6	0.625	135.5	31.5	95	1
8	0.625	135.5	31.5	96	0.8
10	0.625	135.5	31.5	98	0.39
15	0.625	135.5	31.5	98.1	0.39
20	0.625	135.5	31.5	100	0
Series 15. 70 mg L ⁻¹ RB5 + 0.625 g L ⁻¹ TiO ₂ P-25 + 500 mg L ⁻¹ H ₂ O ₂ + 31.5 mg L ⁻¹ Fe ⁺³					
0	0.625	500	31.5	7	51.93
2	0.625	500	31.5	67.5	7.68
4	0.625	500	31.5	95	1
6	0.625	500	31.5	97	0.59
8	0.625	500	31.5	98	0.39
10	0.625	500	31.5	100	0
Series 16. 70 mg L ⁻¹ RB5 + 0.625 g L ⁻¹ TiO ₂ P-25 + 500 mg L ⁻¹ H ₂ O ₂ + 31.5 mg L ⁻¹ Fe ⁺³					
0	0.625	500	31.5	7	51.93

Time [min]	TiO ₂ P-25 [g L ⁻¹]	H ₂ O ₂ [mg L ⁻¹]	Fe ³⁺ [mg L ⁻¹]	Transmittance [%]	RB5 concentration [mg L ⁻¹]
2	0.625	500	31.5	77	5.1
4	0.625	500	31.5	97	0.59
6	0.625	500	31.5	98	0.39
8	0.625	500	31.5	98.1	0.39
10	0.625	500	31.5	100	0

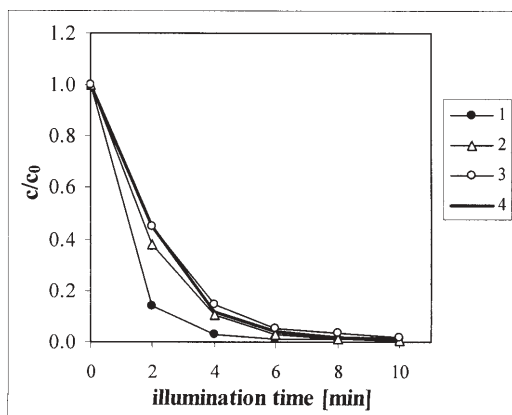


Fig. 2. The influence of the Fe³⁺ and H₂O₂ concentrations on the decolorization of the simulated wastewater in the presence of 0.25 g L⁻¹ TiO₂ P-25: 1 - 800 mg L⁻¹ H₂O₂ + 56 mg L⁻¹ Fe³⁺; 2 - 800 mg L⁻¹ H₂O₂ + 7 mg L⁻¹ Fe³⁺; 3 - 200 mg L⁻¹ H₂O₂ + 56 mg L⁻¹ Fe³⁺; 4 - 200 mg L⁻¹ H₂O₂ + 7 mg L⁻¹ Fe³⁺

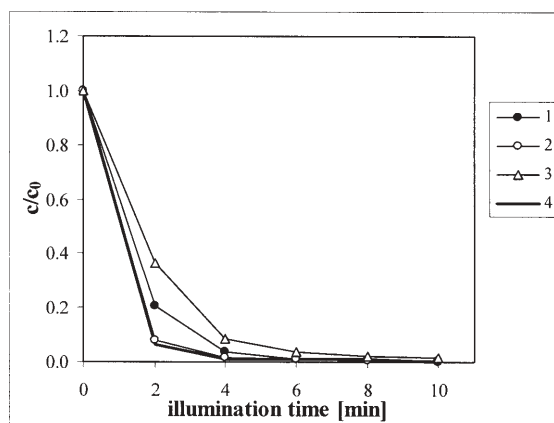


Fig. 3. The influence of the Fe³⁺ and H₂O₂ concentrations on the decolorization of the simulated wastewater in the presence of 1 g L⁻¹ TiO₂ P-25: 1 - 800 mg L⁻¹ H₂O₂ + 7 mg L⁻¹ Fe³⁺; 2 - 200 mg L⁻¹ H₂O₂ + 56 mg L⁻¹ Fe³⁺; 3 - 200 mg L⁻¹ H₂O₂ + 7 mg L⁻¹ Fe³⁺; 4 - 800 mg L⁻¹ H₂O₂ + 56 mg L⁻¹ Fe³⁺

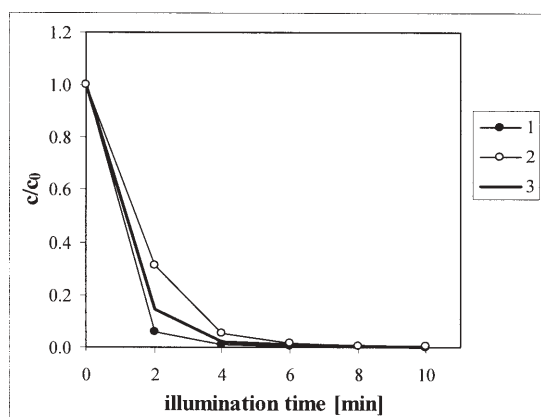


Fig. 4. The influence of TiO₂ P-25 amount on the decolorization of the simulated wastewater in the presence of 500 mg L⁻¹ H₂O₂ and 31.5 mg L⁻¹ Fe³⁺: 1 - 1.08 g L⁻¹ TiO₂ P-25; 2 - 0.17 g L⁻¹ TiO₂ P-25; 3 - 0.625 g L⁻¹ TiO₂ P-25

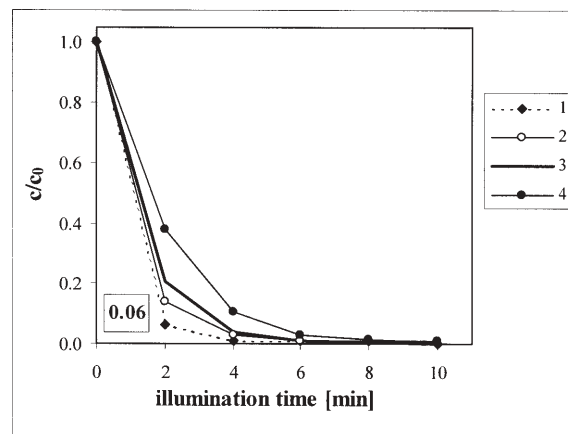


Fig. 5. The influence of the amounts of TiO₂ P-25 and Fe³⁺ on the decolorization of the simulated wastewater in the presence of 800 mg L⁻¹ H₂O₂: 1 - 1 g L⁻¹ TiO₂ P-25 + 56 mg L⁻¹ Fe³⁺; 2 - 0.25 g L⁻¹ TiO₂ P-25 + 56 mg L⁻¹ Fe³⁺; 3 - 1 g L⁻¹ TiO₂ P-25 + 7 mg L⁻¹ Fe³⁺; 4 - 0.25 g L⁻¹ TiO₂ P-25 + 7 mg L⁻¹ Fe³⁺

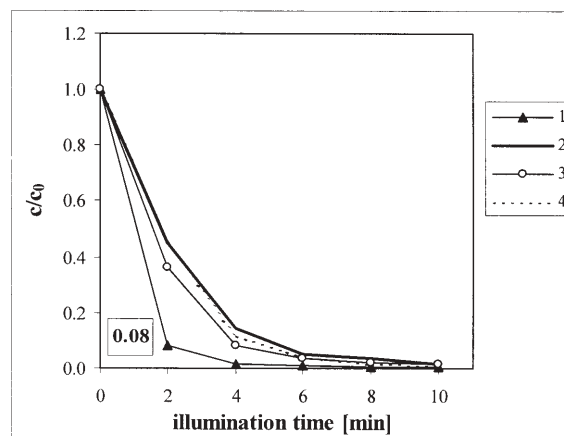


Fig. 6. The influence of the amounts of TiO₂ P-25 and Fe³⁺ on the decolorization of the simulated wastewater in the presence of 200 mg L⁻¹ H₂O₂: 1 - 1 g L⁻¹ TiO₂ P-25 + 56 mg L⁻¹ Fe³⁺; 2 - 0.25 g L⁻¹ TiO₂ P-25 + 56 mg L⁻¹ Fe³⁺; 3 - 1 g L⁻¹ TiO₂ P-25 + 7 mg L⁻¹ Fe³⁺; 4 - 0.25 g L⁻¹ TiO₂ P-25 + 7 mg L⁻¹ Fe³⁺

Figure 5 presents the influence of Fe³⁺ and catalyst amounts on the wastewater process working with maximum amount of H₂O₂ (800 mg L⁻¹). The minimum time corresponds to the 1 g L⁻¹ TiO₂ P-25 and maximum amount of Fe³⁺ (56 mg L⁻¹). The same influence is shown in figure 6, but for minimum amount of H₂O₂ (200 mg L⁻¹). The best result is obtained in the same conditions as in the previous case, respectively 1 g L⁻¹ TiO₂ P-25 and 56 mg L⁻¹ Fe³⁺. Further, comparing the figures 5 and 6, for 1 g L⁻¹ TiO₂ P-25 and 56 mg L⁻¹ Fe³⁺, 94 % of the dye is eliminated after 2 minutes (fig. 5) where the amount of H₂O₂ was 800 mg L⁻¹ compared to 92 % for 200 mg L⁻¹ H₂O₂ in figure 6.

Neural network modeling

Neural network based modeling was applied to predict the rate of the process as function of the reaction conditions.

Experimental data from table 1 were used to train different neural networks which model the transmittance as function of reaction conditions. 10 % of these data represent the validation data set and the remaining data is the training data set.

One major problem in the development of neural network model is determining the network architecture, *i.e.* the number of hidden layers and the number of neurons in each hidden layer. Firstly, potentially good topologies must be identified. Nevertheless, no good theory or rule accompanies the neural network topology that should be used and trial-and-error is still required. This is done by testing several topologies and comparing the prediction errors. Smaller errors indicate potentially good architectures, *i.e.* neural network topologies with chances to train well and to output good results [31].

In this study, the four inputs of neural networks are: time (min), amount of catalyst (TiO_2 P-25, g L^{-1}), amounts of H_2O_2 and Fe^{+3} (mg L^{-1}) and the network output is the transmittance (%).

Table 2 contains different feed forward topologies tested with selected training data and the main performances for these networks: MSE (Mean Squared Error), r (correlation between experimental and neural network outputs) and E_p (percent error). The structure of a network of MLP type (multilayer perceptron) is given by the number of neurons in the input layer, corresponding to the four input variables, then the number of hidden neurons (in one or two layers) and, finally, the number of neurons in output layer for the output variable.

Hidden neurons, as well as output layer neuron, use hyperbolic tangent as nonlinear activation function. This type of function, compared to linear or logistic activation function, produces better learning. All the network weights were initialized as random numbers in the interval (-0.5, 0.5). The network was trained using back-propagation algorithm. For better performance, momentum is used to allows the network to respond not only to the local gradient, but also to recent trends in the error surface. Without momentum, the neural network may get stuck in a shallow local minimum; with momentum, it can slide through such a minimum. The training is considered finished at the point where the network error (MSE) becomes sufficiently small (less than 0.001). Consequently, a configuration of 4 input neurons, one hidden layer with 10 hidden neurons and an output layer of 1 neuron is selected, having $\text{MSE} = 0.000678$, $r = 0.9992$ and percentage error, $E_p = 5.2863\%$.

Table 2

DIFFERENT TOPOLOGIES TESTED FOR THE FEED FORWARD NEURAL NETWORKS

Network topology	MSE	r	E_p
MLP(4:5:1)	0.001603	0.9984	10.3904
MLP(4:10:1)	0.000678	0.9992	5.2863
MLP(4:20:1)	0.000420	0.9995	3.35617
MLP(4:30:1)	0.000347	0.9996	2.3251
MLP(4:50:1)	0.000443	0.9994	3.5525
MLP(4:100:1)	0.000325	0.9996	2.4893
MLP(4:9:3:1)	0.000262	0.9997	1.2661
MLP(4:12:4:1)	0.000253	0.9997	1.3532
MLP(4:24:8:1)	0.000229	0.9997	0.5297
MLP(4:33:11:1)	0.000214	0.9997	0.5122
MLP(4:42:14:1)	0.000202	0.9997	0.4222

We select a neural network from table 2 that best balanced generalization performance against network size and complexity, that means MLP(4:10:1) (marked in grey in table 2). This neural model does not have the best performances, but against other networks with two hidden layers, the differences are not significant. So we prefer a simple architecture with performances good enough.

Good predictions are obtained with the neural model MLP (4:10:1), on training data: average relative errors of 1.2660 % for transmittance and correlation between experimental data and network prediction were 0.9997. Relative errors were calculated using the following formula:

$$E_r = p_{exp} - p_{net} / p_{exp} \cdot 100 \quad (1)$$

where p represents the parameter under study (transmittance), indexes *exp* and *net* denote experimental and network values.

Several examples are presented in figures 7 and 8 which show a comparison between the two sets of data, experimental and network outputs. In these figures, series noted 7 and 8 in table 1 are chosen.

White bars represent the experimental values, and grey bars – neural network results. Good agreement between the two data sets proves that the neural model has learned well the behavior of the process – the increase of transmittance with time.

Zero value in the figures 7 and 8 represents a relative value, that means the moment of turning on the UV lamp. In fact, until zero moment on the graphics, the catalyst was mixed with wastewater for 15 min. During this time, a certain amount of dye is absorbed on the catalyst TiO_2 P-25 (5 % approximately).

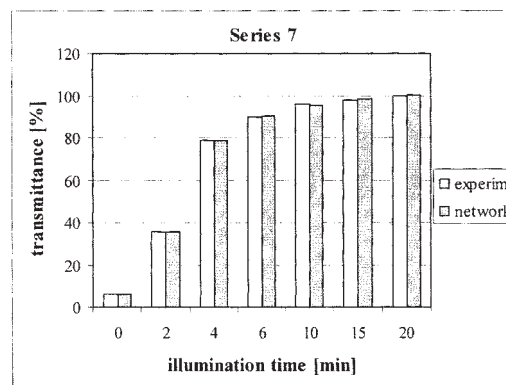


Fig. 7. Experimental data and neural network predictions on training data for a series of photocatalytic experiments obtained under the following conditions: 70 mg L^{-1} RB5 + 1 g L^{-1} TiO_2 P-25 + 200 mg L^{-1} H_2O_2 + 7 mg L^{-1} Fe^{+3}

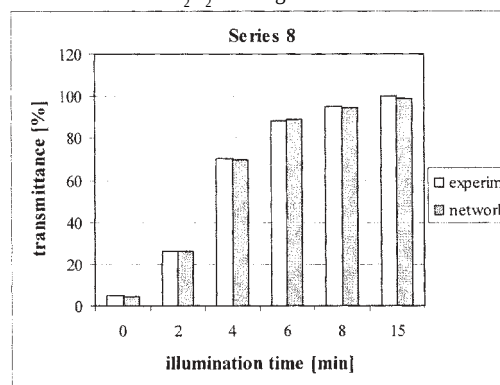


Fig. 8. Experimental data and neural network predictions on training data for a series of photocatalytic experiments obtained under the following conditions: 70 mg L^{-1} RB5 + 0.25 g L^{-1} TiO_2 P-25 + 200 mg L^{-1} H_2O_2 + 7 mg L^{-1} Fe^{+3}

Table 3
PREDICTIONS OF MLP(4:10:1) ON VALIDATION DATA

Time [min]	TiO ₂ [g L ⁻¹]	H ₂ O ₂ [mg L ⁻¹]	Fe ⁺³ [mg L ⁻¹]	Experimental transmittance [%]	Network transmittance [%]	E _r
6	1	800	56	97.2	98.05	0.8760
2	0.25	800	56	66	66.07	0.1106
4	1	800	7	90	92.35	2.6111
2	1	200	56	80	79.61	0.4886
15	0.25	200	56	96	96.02	0.0208
8	1	200	7	94	93.85	0.1595
10	0.25	200	7	97	96.8	0.2061
4	0.625	500	1.7325	37	35.90	2.9729
4	0.625	135.5	31.5	85	88.10	3.6470
8	0.625	500	31.5	98.1	98.70	0.6079
average relative error						1.3083
correlation						0.9982

A key issue in neural network based process modeling is the robustness or generalization capability of the developed models, i.e. how well the model performs on unseen data. Thus, a serious examination of the accuracy of the neural network results requires the comparison with experimental data, which were not used in the training phase (previously unseen data). That is why the validation data set is considered as part of experimental data from table 1. The predictions of the network on validation data are given in table 3.

It can be noticed a satisfactory agreement between the two categories of data, experimental and neural network predictions, with an average relative error of 1.3083, maximum relative error less than 4% and a correlation of 0.9982. For this reason, the projected neural model MLP (4:10:1) can be used to make predictions under different reaction conditions, substituting the experiments that are time and material consuming.

Genetic algorithm optimization

The general solution of an optimization problem can be obtained in terms of the following four elements: an accurate model of the process, a selected number of control variables, an objective function and a suitable numerical method for solving the specified optimization problem.

In this study, the optimization problem includes the neural model, which is represented as:

$$\text{NN [Inputs: } t, \text{TiO}_2, \text{H}_2\text{O}_2, \text{Fe}^{+3} \text{; Output: } D] \quad (2)$$

The vector of control variables, u , has the components:

$$u = [t, \text{TiO}_2, \text{H}_2\text{O}_2, \text{Fe}^{+3}] \quad (3)$$

An admissible control input u^* should be formed in such a way that the performance index, J , defined by the following equation, are minimized:

$$J = \left(1 - \frac{D_f}{D_d} \right)^2 \quad (4)$$

subject to:

$$\begin{aligned} u_{\min} \leq u \leq u_{\max} ; \quad 0 \leq t \leq 60 \text{ min} ; \quad 0.1 \leq \text{TiO}_2 \leq 1 \text{ g L}^{-1} \\ 100 \leq \text{H}_2\text{O}_2 \leq 800 \text{ mg L}^{-1} ; \quad 1 \leq \text{Fe}^{+3} \leq 65 \text{ mg L}^{-1} \end{aligned} \quad (5)$$

with t representing illumination time, D_f – transmittance at the end of the process and D_d – an imposed value for this parameter.

The transmittance, D , is a measure of the dye elimination from wastewater, with values between 0 and 100 %. High values for this parameter mean high degree of dye elimination.

The constraints are very important to define the range of variation of parameters and to disregard possible solutions that could be interesting in a theoretical approach to the problem.

In other words, the optimization problem to be solved can be formulated as follows: which are the optimal working conditions (time, amount of catalyst TiO₂ P-25, amounts of H₂O₂ and Fe⁺³) necessary to obtain an imposed transmittance value under the given experimental conditions?

The optimization procedure includes a neural network (NN) model and it is solved with a simple genetic algorithm. The fitness function of the GA is the scalar objective function (4). Figure 9 illustrates this optimization procedure. Genetic algorithm provides, after an iterative calculus, the optimal values for decision variables (time, TiO₂, H₂O₂, Fe⁺³), which are the inputs for the neural network model. With these inputs, the neural network computes the final value of transmittance, D_f which will be compared with the desired value, D_d . If the two values are identical or there is a very tight difference between them, we can conclude that the task of the optimization, represented by the minimum of the objective function, J , is achieved.

Genetic algorithms are intelligent stochastic optimization techniques based on the mechanism of natural selection and genetics. GAs start with an initial set of solutions, called population. Each solution in the

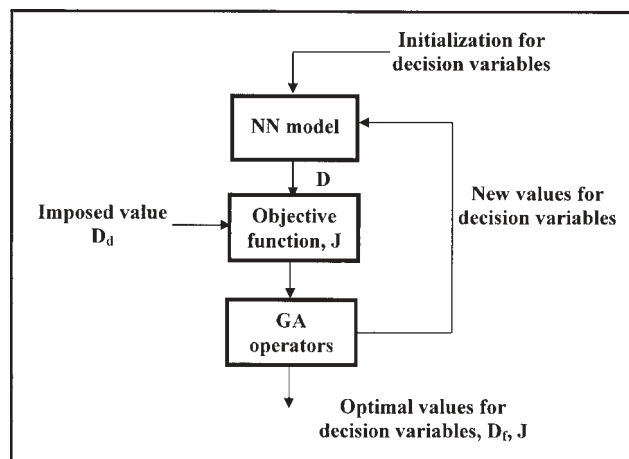


Fig. 9. The optimization method based on NN and GA.

population is called a chromosome (or individual), which represents a point in the search space. The chromosomes are evolved through successive iterations, called generations, by genetic operators (selection, crossover and mutation) that mimic the principle of natural evolution. A set of solutions are analyzed and modified by genetic operations simultaneously, where selection operator can select some "good" solutions as seeds, crossover operator can generate new solutions hopefully retaining good features from parents, and mutation operator can enhance diversity and provide a chance to escape from local optima [32]. In a GA, a fitness value is assigned to each individual according to a problem-specific objective function. Generation by generation, the new individuals, called offspring, are created and survive with chromosome in the current population, called parents, to form a new population.

In our GA model, we used real values encoding for the chromosomes. There are other approaches for genetic algorithm based optimization which use binary solution representation, as it is the simplest type of encoding, in which chromosomes are composed only of 1's and 0's. Even the number of alleles is thus rather small (two), this encoding is very common, because it is very easy to use. However, value encoding is more general, because genes are real numbers. Some experiments [33] have shown that real value encoding is more efficient, with better precision of the solutions.

The initial population is generated randomly. Offspring are created by genetic operators and it is stored in a population pool that is a collection of offspring and their parents.

Selection compares the chromosomes in the population aiming to choose those which will take part in the reproduction process. There are different methods for the selection phase; our paper uses rank selection which first

rank the population and then every chromosome receives fitness from this ranking. The recombination (crossover) has as main purpose the recombination the features of two randomly selected parents from the mating pool with the aim of producing better offspring. The variant of crossover used in this study supposes different points for all genes, that means the new individual will no longer be on the line segment that links its parents. After recombination, offspring undergoes to mutation. Generally, the mutation refers to the creation of a new chromosome from one and only one individual with predefined probability. Mutation is used to produce small perturbations on chromosomes to promote diversity of the population. Our GA includes a variant of mutation named resetting. A gene value is reset to a random value in its search interval. The purpose is to refresh the search process, in case when the genetic diversity of the population decreases (so no longer converges to the solution) or the algorithm has converged into a local optimum.

After the three operators are carried, the offspring is inserted into the population, replacing the parent chromosomes in which they were derived from, producing a new generation. The best individual is copied directly into the new population (the elitism technique) and the rest of the individuals are replaced by the new generations.

The termination criterion determines when GA will stop. In other words, the genetic operations are repeated until a termination condition is met. In our implementation, we stopped GA if the maximum number of generations has been executed or the pre-set number of generations without improvement in the last best solution has been reached.

Population size, number of generations, crossover probability and mutation probability are known as the control parameters of genetic algorithms. The values of these parameters must be specified before the execution

Table 4
OPTIMAL DECISION VARIABLES OBTAINED FOR DIFFERENT VALUES OF THE IMPOSED TRANSMITTANCE

D_d [%]	Time [min]	TiO₂ [g L ⁻¹]	H₂O₂ [mg L ⁻¹]	Fe⁺³ [mg L ⁻¹]	D_f [%]	J
75	4.16	0.51	434.18	18.86	75	0.000020
76	5.10	0.62	495.93	13.72	76	0.000000
77	4.98	0.57	549.10	13.67	77	0.000048
78	4.65	0.32	317.52	47.05	78	0.000002
79	6.61	0.69	408.97	7.23	79	0.000078
80	3.83	0.49	282.23	40.28	80	0.000155
81	4.57	0.58	301.21	20.12	81	0.000000
82	5.12	0.75	275.49	16.72	82	0.000000
83	3.99	0.71	267.29	25.51	83	0.000002
84	6.6	0.63	299.41	10.04	84	0.000003
85	6.37	0.39	222.07	43.08	85	0.000015
86	7.04	0.19	306.15	41.37	86	0.000007
87	6.82	0.55	474.79	12.13	87	0.000000
88	5.31	0.48	250.71	41.69	88	0.000041
89	9.64	0.65	505.92	2.73	89	0.000227
90	7.82	0.51	397.44	10.36	90	0.000006
91	8.03	0.43	469.03	9.14	91	0.000000
92	4.73	0.61	297.71	30.68	92	0.000002
93	5.86	0.37	398.20	45.51	93	0.000000
94	8.04	0.35	299.41	39.99	94	0.000000
95	10.68	0.21	273.60	34.94	95	0.000000
96	11.73	0.69	321.60	19.81	96	0.000000
97	14.10	0.59	411.16	17.21	97	0.000000
98	14.94	0.32	363.10	52.65	98	0.000000
99	20.04	0.30	453.32	50.08	99	0.000000
100	19.23	0.51	517.31	50.87	100	0.000000

of GA and they depend on the nature of the objective function. One important thing in using GA as solving method is to adjust its parameters to the particular problem approached so as to obtain good solutions and to preserve the algorithm from a preliminary convergence.

Several tests lead to the following values: pop_size = 50 (size of initial population), max_gen = 100 (maximum number of generations), cross_rate = 0.9 (rate of crossover) and mut_rate = 0.03 (rate of mutation).

The fitness function in GA is the objective function of the optimization procedure. The results of the optimization are represented by the values of the decision variables (illumination time, amounts of catalyst, H_2O_2 and Fe^{+3}) that lead to a minimum value of the objective function, which means the achievement of the imposed value for the transmittance.

The optimization procedure is implemented in *Matlab* 7.5 with original software, as specific functions were programmed for each phase of the genetic algorithm.

The variation domains for decision variables are represented by the constraints (5). Imposing different values for transmittance, in the interval 75 % ÷ 100 %, the optimal working conditions obtained by GA optimization procedure are given in table 4.

In all cases in table 4, $D_t = D_d$, that means the goal of the optimization has been accomplished.

Therefore, the results in table 4 represent the optimal working conditions (illumination time, amounts of catalyst,

H_2O_2 and Fe^{+3}) necessary for obtaining a pre-established value for transmittance. It was expected that high values for transmittance involve higher amounts of reagents and longer illumination time.

Supplementary calculations were added to the data in table 4 concerning the energy necessary for the decolorization process composed from the energy consumed by the light source and for the mixture stirring. Table 5 is the correspondent for table 4 and the results are in decreased order by the energy. In this table, two values were marked with grey bands, having small necessary energy and relatively high values for transmittance (small ratio c/c_0 , respectively, where c is the concentration of the dye and c_0 – initial concentration).

Several remarks should be made concerning the optimization statistics.

Searching the optimum, GA starts from some initial points because the first population is generated randomly. But, during the evolution, new potential solutions are introduced through mutation process.

GA is a heuristic search method, therefore the global optimal solution can not be guaranteed. This is why if a solution is not always satisfactory, the algorithm runs many times and each of run could give a different solution. In other words, GA can provide local optimum where a better solution than all the other ones assigns its genes with priority and “wins” the population. This is the best solution in a particular case, but possibly not a good solution in

Table 5
THE ENERGY NECESSARY FOR DECOLORIZATION PROCESS ATTACHED
TO THE DATA FROM TABLE 4

D_d [%]	Time [min]	TiO_2 [g L ⁻¹]	H_2O_2 [mg L ⁻¹]	Fe^{+3} [mg L ⁻¹]	c/c_0	Energy, [kJ]
99	20.04	0.30	453.32	50.08	0.00	10.82
100	19.23	0.51	517.31	50.87	0.00	10.38
98	14.94	0.32	363.10	52.65	0.01	8.07
97	14.1	0.59	411.16	17.21	0.01	7.61
96	11.73	0.69	321.60	19.81	0.01	6.33
95	10.68	0.21	273.60	34.94	0.02	5.77
89	9.64	0.65	505.92	2.73	0.04	5.21
94	8.04	0.35	299.41	39.99	0.02	4.34
91	8.03	0.43	469.03	9.14	0.03	4.34
90	7.82	0.51	397.44	10.36	0.04	4.22
86	7.04	0.19	306.15	41.37	0.05	3.80
87	6.82	0.55	474.79	12.13	0.05	3.68
79	6.61	0.69	408.97	7.23	0.08	3.57
84	6.6	0.63	299.41	10.04	0.06	3.56
85	6.37	0.39	222.07	43.08	0.06	3.44
93	5.86	0.37	398.20	45.51	0.03	3.16
88	5.31	0.48	250.71	41.69	0.05	2.87
82	5.12	0.75	275.49	16.72	0.07	2.76
76	5.1	0.62	495.93	13.72	0.10	2.75
77	4.98	0.57	549.10	13.67	0.09	2.69
92	4.73	0.61	297.71	30.68	0.03	2.55
78	4.65	0.32	317.52	47.05	0.09	2.51
81	4.57	0.58	301.21	20.12	0.07	2.47
75	4.16	0.51	434.18	18.86	0.10	2.25
83	3.99	0.71	267.29	25.51	0.07	2.15
80	3.83	0.49	282.23	40.28	0.08	2.07

general. Then, GA should be run many times, choose and keep the best solution obtained along the runs.

For a binary GA, the “scheme theorem” proves that there is a nonzero probability to obtain the solution of a problem. But for GA with real encoding, this theorem is not always valid and it is recommended the multiple execution of the optimization algorithm, followed by the choose of the best solution among those obtained.

The optimization procedure based on a simple genetic algorithm and a neural network model applied in this paper is easy to manipulate and provides accurate results. In this way, a theoretical complete analysis of the decolorization process approached here is performed, with useful information for the practical applications.

Conclusions

This paper provides a general and simple optimization methodology, based on genetic algorithms and neural networks, applied to a wastewater decolorization process. The genetic algorithm solves the optimization problem (minimum transmittance at the end of the process) and the neural network constitutes the model included in the optimization procedure which estimates the evolution in time of a dyestuff wastewater treatment process consisting of a photocatalytic oxidative reaction..

Simple architecture of the neural network is proposed for process modeling: feed forward neural network with one hidden layer. The training and testing phases of the modeling procedure is conducted with experimental data performed in different reaction conditions (catalyst and Fenton reagent amounts). Good predictions are obtained with the neural model in validation phase, so this neural network gives a very good representation for the wastewater treatment process analysis.

The simple genetic algorithm proves to be a good tool for solving the optimization problem, providing important information useful in experimental practice.

The method can be easily extended and adapted to other environmental oriented processes, with high chances of providing accurate results by simple handling.

References

1. PERAL, J., DOMENECH, X., OLLIS, D., 1997, *J. Chem. Technol. Biotechnol.*, 70, 117:140
2. ALFANO, O., BAHNEMANN, D., CASSANO, A., DILLERT, R., GOSLICH, R., 2000, *Catal. Today*, 58, 199:230
3. MALATO, S., BLANCO, J., VIDAL, V., ALARCON, D., MALDONADO, M., CACERES, J., GERNJAK, W., 2003, *Solar Energy*, 75, 329:336
4. SERPONE, N., EMELINE, A.V., 2002, *Int. J. Photoenergy*, 4, 91:131
5. HOFFMAN, M., MARTIN, S., CHOI, W., BAHNEMANN, D., 1995, *Chem. Rev.*, 95, 69:96

6. FUJISHIMA, A., RAO, T., TRYK, D., 2000, *J. Photochem. Photob. C: Photochem. Reviews*, 1, 1:21
7. HERRMANN, J.M., 1999, *Catal. Today*, 53, 115:29
8. MURUGANANDHAM, M., SOBANA, N., SWAMINATHAN, M., 2006, *J. Hazard. Mat.*, 137, 1371:1376
9. AUGUGLIARO, V., BAIOCCHI, C., PREVOT, A., LÓPEZ, E., LODDO, V., MALATO, M., MARCÍ, G., PALMISANO, L., PAZZI, M., PRAMAURO, E., 2002, *Chemosphere*, 49, 1223:1230.
10. NEPPOLIAN, B., CHOI, H.C., SAKTHIVEL, S., ARABINDOO, B., MURUGESAN, V., 2002, *J. Hazard. Mat.*, 89, 303:317
11. MARUGAN, J., LOPEZ-MUNOZ, M.J., GERNJAK, W., MALATO, S., 2006, *Ind. Eng. Chem. Res.*, 45, 8900:8908
12. MROWETZ, M., SELLI, E., 2004, *J. Photochem. Photob. A: Chemistry*, 162, 89:95
13. PAREEK, V.K., BRUNGS, M.P., ADESINA, A.A., SHARMA, R., 2002, *J. Photochem. Photob. A: Chem.*, 149, 139:146
14. DURAN, A., MONTEAGUDO, J.M., MOHEDANO, M., 2006, *Applied Catalysis B: Environmental*, 65, 127:134
15. TOMA, F.L., GUESSASMA, S., KLEIN, D., MONTAVON, G., BERTRAND, G., CODDET, C., 2004, *J. Photochem. Photob. A: Chem.*, 165, 91:96
16. SALARI, D., DANESHVAR, N., AGHAZADEH, F., KHATAEE, A.R., 2005, *J. Haz. Mat.*, 125, 205:210
17. XIONG, Z., ZHANG, J., 2005, *Chem. Eng. Proc.*, 44: 477
18. XIONG, Z., ZHANG, J., 2005, *J. Proc. Control*, 15: 11
19. NG, C.W., HUSSAIN, M.A., 2004, *Chem. Eng. Proc.*, 43: 559
20. ZHANG, J., 2004, *Ind. Eng. Chem. Res.*, 43: 1030
21. CURTEANU, S., POPA, A.M., 2004, *Rev. Roum. Chim.*, 49: 3
22. SALARI, D., DANESHVAR, AGHAZADEH, F., KHATAEE, A.R., 2005, *J. Hazard. Mat.*, B125: 205
23. GÖB, S., OLIVEROS, E., BOSSMANN, S.H., BRAUN, A.M., NASCIMENTO, C.A.O., GUARDAN, R., 2001, *Water Science & Technology*, 44: 339
24. DURAN, A., MONTEAGUDO, J.M., MOHEDANO, M., 2006, *Applied Catalysis B: Environmental*, 65: 127–134.
25. DURAN, A., MONTEAGUDO, J.M., 2007, *Water Research*, 41: 690
26. DEB, K., 1999, in *Sadhana-Academy Proceedings in Engineering Sciences*, 24, 293:315
27. DEB, K., 2001, *Multiobjective optimization using evolutionary algorithms*, Chichester, UK: Wiley
28. YAN, X.F., CHEN, D.Z., HU, S.X., 2003, *Comp. Chem. Eng.*, 27, 1393:1404
29. BHASKAR, V., GUPTA, S.K., RAY, A.K., 2000, *Reviews in Chemical Engineering*, 16, 1:54
30. NANDASANA, A.D., RAY, A.K., GUPTA, S.K., 2003, *Int. J. Chem. Reactor Eng.*, 1, 1:15
31. FERNANDES, F.A.N., LONA, L.M.F., 2005, *Brazilian J. Chem. Eng.*, 22: 323-330
32. WANG, L., ZHANG, L., ZHENG D.Z., 2006, *Comput. & Operat. Res.*, 33, 2969:2971.
33. MICHALEWICZ, Z., 1996, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer Verlag, Berlin

Manuscript received: 11.12.2007