Journal of Molecular Catalysis A: Chemical 331 (2010) 86–100

Contents lists available at ScienceDirect

Journal of Molecular Catalysis A: Chemical

journal homepage: www.elsevier.com/locate/molcata

Review

Artificial neural networks modeling of contaminated water treatment processes by homogeneous and heterogeneous nanocatalysis

A.R. Khataee[a,*], M.B. Kasiri[b]

[a] Department of Applied Chemistry, Faculty of Chemistry, University of Tabriz, Tabriz, Iran
[b] Faculty of Applied Arts, Tabriz Islamic Art University, Tabriz, Iran

ARTICLE INFO

Article history:
Received 30 June 2010
Received in revised form 24 July 2010
Accepted 29 July 2010
Available online 11 August 2010

Keywords:
TiO_2 nanoparticles
ZnO nanoparticles
Neural network
Photocatalysis
Fenton
Photo-Fenton
Electro-oxidation

ABSTRACT

Artificial neural networks (ANNs) are computer based systems that are designed to simulate the learning process of neurons in the human brain. ANNs have been attracting great interest during the last decade as predictive models and pattern recognition. Artificial neural networks possess the ability to "learn" from a set of experimental data (e.g. processing conditions and corresponding responses) without actual knowledge of the physical and chemical laws that govern the system. Therefore, ANNs application in data treatment is especially important where systems present nonlinearities and complex behavior. In recent years "advanced oxidation processes" (AOPs), including homogeneous and heterogeneous nanocatalytic processes, have been proposed to oxidize quickly and non-selectively a broad range of water pollutants. Due to the complexity of reactions in AOPs, the effect of different operational parameters involved are very difficult to determine, leading to uncertainties in the design and scale-up of chemical reactors of industrial interest. It is evident that this problem can not be solved by simple linear multivariate correlation. Artificial neural networks are a promising alternative modeling technique. This paper briefly describes the application of artificial neural networks for modeling of water and wastewater treatment using various homogeneous and heterogeneous nanocatalytic processes. Examples of early applications of ANNs in modeling and simulation of photocatalytic, photooxidative and electrochemical treatment processes are reviewed.

Contents

* Corresponding author. Tel.: +98 411 3393165; fax: +98 411 3340191.
E-mail addresses: a_khataee@tabrizu.ac.ir (A.R. Khataee), m_kasiri@tabriziau.ac.ir, masud.kasiri@gmail.com (M.B. Kasiri).

1381-1169/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.molcata.2010.07.016

1. Introduction

In the last decades great strides have been made in neural network technology. This breakthrough has led to increasing research on a wide variety of scientific applications [1,2]. The interest to ANNs is reflected in the number of scientists, the amounts of funding, the number of large conferences, and the number of journals associated with neural networks.

Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision, and control systems. In this document, we have briefly described the application of artificial neural networks for modeling of different water and wastewater advanced treatment processes.

Conventional water and wastewater treatment processes have been long established in removing many chemical and microbial contaminants of concern to public health and the environment. However, the effectiveness of these processes has become limited over the last two decades because of the following three new challenges [3,4]:

- Increased knowledge about the consequences of water pollution and the public desire for better quality water;
- Diminishing water resources and rapid population growth and industrial development. The reuse of municipal and industrial wastewaters and the recovery of potential pollutants used in industrial processes become crucially important;
- Advances in the manufacturing industry and the growing market associated with advanced treatment processes have resulted in substantial improvements to the versatility and costs of these processes at the industrial scale.

To resolve these new challenges and better use of economical resources, various AOPs have been proposed, tested, and applied to meet both current and anticipated treatment requirements. Advanced treatment technologies have been demonstrated to remove various potentially harmful compounds that could not be effectively removed by conventional treatment processes. Among them, homogeneous and heterogeneous nanocatalytic processes have been proven to successfully remove a wide range of challenging contaminants and hold great promise in water and wastewater treatment.

This review starts with a chapter in which the topology of artificial neural networks is discussed. In chapter 3, training, validation and test of the neural networks are described. Chapters 4 to 7 discuss the early applications of ANNs in modeling and simulation of various homogeneous and heterogeneous nanocatalytic processes including photocatalysis, Fenton, photo-Fenton and Electro-oxidation processes.

2. Artificial neural networks structure

The success in obtaining a reliable and robust network depends strongly on the choice of process variables involved as well as the available set of data and the domain used for training purposes [5]. In this chapter, we are to describe the topology of artificial neural networks including transfer functions, learning processes and training algorithms.

Each network consists of artificial neurons grouped into layers and put in relation to each other by parallel connections. The strength of these interconnections is determined by the weight associated with them. For every ANN, the first layer constitutes the input layer (independent variables) and last one forms the output layer (dependent variables). One or more neuron layers called hidden layers can be located between them.

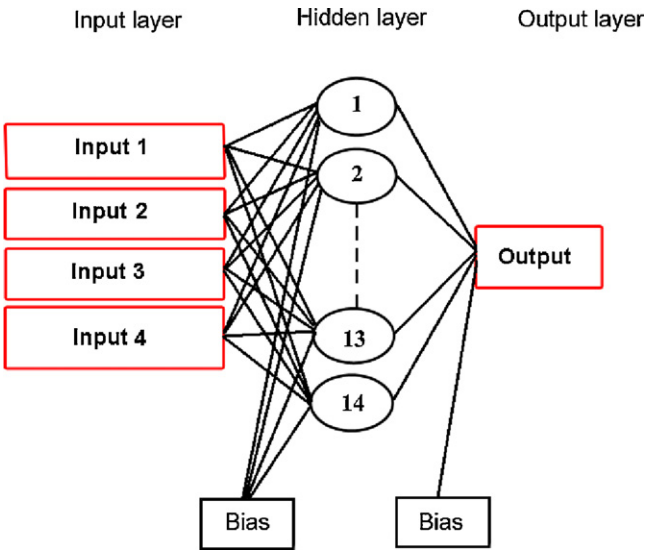

Fig. 1. A typical feed-forward neural network with a single hidden layer containing 14 neurons.

The hidden layers act like feature detectors and in theory, there can be more than one hidden layer. Universal approximation theory, however, suggests that a network with a single hidden layer with a sufficiently large number of neurons can interpret any input–output structure [5–8]. The number of neurons in the hidden layer is determined by the desired accuracy in the neural predictions. Hence, it may be considered as a parameter for the neural net design. In the feed-forward neural net, all the neurons of a particular layer are connected to all the neurons of the layer next to it. The input layer of neurons acts like a distributor and the input to this layer is directly transmitted to the hidden layer. The inputs to hidden and output layers are calculated by performing a weighted summation of all the inputs received from the preceding layer. The weighted sum of the inputs are transferred to the hidden neurons, where it is transformed using an activation function. The output of hidden neurons in turn, acts as inputs to output neurons where it undergoes another transformation. Fig. 1 shows a typical feed-forward neural network with a single hidden layer containing 14 neurons [9].

The topology of an artificial neural network is then determined by number of its layers, number of nodes in each layer and the nature of transfer functions. Optimization of ANN topology is probably the most important step in the development of a model.

2.1. Transfer functions

Different transfer functions can be used as the neuron activation function. Table 1 shows the characteristics of some of these functions. The graph in the square to the right of each transfer function represents the symbol of each transfer function.

Table 1
Characteristics of different transfer functions.

Function name	Input/output	Icon	Acronym (*Matlab Name*)
Hard-limit	$a=0$ if $n<0$ $a=1$ if $n\geq 0$		*hardlim*
Symmetric hard-limit	$a=-1$ if $n<0$ $a=1$ if $n\geq 0$		*hardlims*
Linear	$a=n$		*purelim*
Saturated linear	$a=0$ if $n<0$ $a=n$ if $0\leq n\leq 1$ $a=1$ if $n>1$		*satlin*
Symmetric saturated linear	$a=-1$ if $n<-1$ $a=n$ if $-1\leq n\leq 1$ $a=1$ if $n>1$		*satlins*
Positive linear	$a=0$ if $n<0$ $a=n$ if $n\geq 0$		*poslin*
Sigmoid	$a=\frac{1}{1+\exp^{-n}}$		*logsig*
Tangent hyperbolic	$a=\frac{e^n-e^{-n}}{e^n+e^{-n}}$		*tansig*
Competitive	$a=1$ if n is maximum $a=0$ otherwise	C	*compet*

Three commonly used functions are "hard limit", "linear" and "sigmoid" [10]. The hard-limit transfer function limits the output of the neuron to either 0, if the net input argument n is less than 0, or 1, if n is greater than or equal to 0. More specifically, a negative input does not pass the threshold, then the function returns the value 0 (0 can be interpreted as false), then a non-negative input exceeds the threshold, and the function returns 1 (true).

The other important transfer function is called linear. This function is very simple; it directly transfers its input to its output:

$$a = n \tag{1}$$

In this case, the output of the neuron corresponds to its level of activation and the zero output occurs when:

$$w^{\mathrm{T}}p = b \tag{2}$$

The sigmoid transfer function takes the input, which can have any value between plus and minus infinity, and squashes the output into the range 0 to 1.

Sigmoid transfer function is the most widely used transfer function for the input and hidden layers and given by Eq. (3).

$$a = \frac{1}{1+e^{-n}} \tag{3}$$

The sigmoid function is very non-linear because there is a discontinuity when $w^{\mathrm{T}}p=b$. This transfer function is commonly used in back propagation networks, because it is differentiable. Finally, it must be noted that the "hyperbolic tangent" is a symmetrical version of the sigmoid function [10].

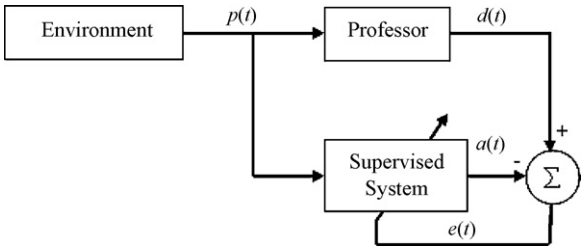

Fig. 2. Supervised learning scheme.

2.2. Learning processes

Among the desirable properties of a neural network, the ability of learning from its environment to improve its performance is surely the most important. But what then is learning? Unfortunately, there is no general, universally accepted definition, as this term is associated with many different concepts that depend on each person's perspective. Learning is a dynamic and iterative process that modifies the parameters of a network. This process is a respond to the signals that the network receives from its environment [11,12].

In most topologies, learning results to a change of synaptic efficiency, or in other words, to a change in the amount of the weights connecting the neurons of one layer to another. There are three basic types of learning including supervised, reinforcement and unsupervised learning. Supervised learning is characterized by the presence of a "professor" who has a deep knowledge of the environment in which the neural network evolves. Supervised learning is illustrated schematically in Fig. 2.

The environment is unknown to the network. This produces an input $p(t)$ that is routed to both the professor and the network. Thanks to its intrinsic knowledge, the professor produces a desired output $d(t)$ for this input. It is assumed that this response is optimal. It is then compared (by subtraction) with the output of the network to produce an error signal $e(t)$ which will be re-injected into the network to change its behavior through an iterative process. This iteration eventually allows the network to simulate the response of the professor [13].

The reinforcement learning overcomes some limitations of supervised learning. It is a kind of supervised learning, but with a hint of satisfaction instead of a scalar error signal vector. In practice, the use of reinforcement learning is complex to implement. But, it is important to understand the difference between this type of learning and supervised one. Supervised learning has an error signal that not only calculates an index of satisfaction, but also estimates the local gradient indicating a direction for the adaptation of synaptic weights. This information is provided by the professor who makes all the difference. In reinforcement learning, the absence of error signal makes the calculation of this gradient impossible. To estimate the gradient, the network is forced to attempt actions and observe the outcome, eventually infer a direction of change for the synaptic weights. To do this, it implements a process of trial/error while delaying the reward offered by the satisfaction index [14].

Unsupervised or self-organized learning is characterized by the complete absence of professor, neither a signal error, as in the case of supervised, nor an index of satisfaction, as in the case for reinforcement. Therefore, there is an environment that provides inputs, and a network must learn without external intervention. By assimilating the input from the environment to a description of its internal state, the major task of a network is then to model this state as best as possible. At the end of learning process, the network has developed an ability to form internal representations of environmental variations permitting to encode their characteristics and, therefore, automatically create similar classes of responses [15].

2.3. Training algorithms

2.3.1. Back propagation algorithm

Back propagation was created by generalizing the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions [16]. Standard back propagation is a gradient descent algorithm in which the network weights are moved along the negative of the gradient of the performance function. The term back propagation refers to the manner in which the gradient is computed for nonlinear multilayer networks. There are a number of variations on the basic algorithm that are based on other standard optimization techniques, such as conjugate gradient and Newton methods.

There are many variations of the back propagation algorithm, several of which are described in this chapter. The simplest implementation of back propagation learning updates the network weights and biases in the direction in which the performance function decreases most rapidly, the negative of the gradient. An iteration of this algorithm can be written as follows:

$$x_{k+1} = x_k - \alpha_k g_k \tag{4}$$

where x_k is a vector of current weights and biases, g_k is the current gradient, and α_k is the learning rate [16].

There are two different ways in which gradient descent algorithm can be implemented: incremental mode and batch mode. In incremental mode, the gradient is computed and the weights are updated after each input is applied to the network. In other words the network is presented with cases from the training data one at a time and the weights are revised after each case in an attempt to minimize the error function. In batch mode, all the inputs are applied to the network before the weights are updated [17]. Some important types of back propagation algorithms are described below.

The basic back propagation algorithm adjusts the weights in the steepest descent direction (negative of the gradient); the direction in which the performance functions is decreasing most rapidly. It turns out that, although the function decreases most rapidly along the negative of the gradient, this does not necessarily produce the fastest convergence. In the conjugate gradient algorithms a search is performed along conjugate directions, which produces generally faster convergence than steepest descent directions [18].

The gradient method has most of the benefits of Newton's method but without the inconvenience of having to calculate and invert the Hessian matrix. The Hessian matrix (or simply the Hessian) is the square matrix of second-order partial derivatives of a function; it describes the local curvature of a function of many variables [19].

Conjugate gradient algorithm discussed above requires a line search at each iteration. This line search is computationally expensive, because it requires that the network response to all training inputs be computed several times for each search. The scaled conjugate gradient algorithm (*scg*), developed by Moller [11], was designed to avoid the time-consuming line search. This algorithm combines the model-trust region approach, used in the Levenberg-Marquardt algorithm, described in Section 2.3.3.

2.3.2. Quasi-Newton algorithm

Newton's method is an alternative to the conjugate gradient methods for fast optimization. The basic step of Newton's method is:

$$x_{k+1} = x_k - A_k^{-1} g_k \tag{5}$$

where A_k^{-1} is the Hessian matrix (second derivatives) of the performance index at the current values of the weights and biases. Newton's method often converges faster than conjugate gradient methods. Unfortunately, it is complex and expensive to compute the Hessian matrix for feed-forward neural networks.

There is a class of algorithms that is based on Newton's method, but which does not require calculation of second derivatives. These are called quasi-Newton (or secant) methods. They update an approximate Hessian matrix at each iteration of the algorithm. The update is computed as a function of the gradient [10,20].

2.3.3. Levenberg-Marquardt algorithm

Like the quasi-Newton methods, the Levenberg-Marquardt algorithm was designed to approach second-order training speed without having to compute the Hessian matrix. When the performance function has the form of a sum of squares (as is typical in training feed-forward networks), then the Hessian matrix can be approximated as:

$$H = J^{\mathrm{T}} J \tag{6}$$

and the gradient can be computed as:

$$g = J^{\mathrm{T}} e \tag{7}$$

where J is the Jacobian matrix that contains first derivatives of the network errors with respect to the weights and biases, and e is a vector of network errors. The Jacobian matrix can be computed through a standard back propagation technique that is much less complex than computing the Hessian matrix.

The Levenberg-Marquardt algorithm uses this approximation to the Hessian matrix in the following Newton-like update:

$$x_{k+1} = x_k - [J^{\mathrm{T}} J + \mu I]^{-1} J^{\mathrm{T}} e \tag{8}$$

When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible [20].

Thus, μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm.

3. Optimization of a neural network

The primary objective of this chapter is to explain how to use the training functions to train neural networks to solve specific problems. There are generally four steps in the training process:

- Assemble the training data,
- Create the network object,
- Train the network, and
- Simulate the network response to new inputs.

This chapter discusses a general optimization process of a neural network topology. It describes the architecture of a typical multilayer feed-forward network. Then the simulation and training of the network objects will be presented.

Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities.

Feed-forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range −1 to +1.

In order to determine the optimum number of hidden nodes, a series of topologies are used, in which the number of nodes was

varied from for example 2 to 20. A mathematic function such as the mean square error (MSE) is normally used as the error function. MSE measures the performance of the network according to the following equation:

$$\mathrm{MSE} = \frac{\sum_{i=1}^{i=N}(y_{i,\mathrm{pred}} - y_{i,\mathrm{exp}})^2}{N} \quad (9)$$

where N is the number of data point, $y_{i,\mathrm{pred}}$ is the network prediction, $y_{i,\mathrm{exp}}$ is the experimental response and i is an index of data.

All ANNs are trained using one appropriate algorithm such as scaled conjugate gradient, quasi-Newton, Levenberg-Marquardt algorithms and so on. The primary goal of training is minimizing the error function by searching for a set of connection weights and biases that causes the ANN to produce outputs that are equal or close to target values. For instance, the back propagation algorithm minimizes the error function between the observed and the predicted output in the output layer, through two phases. In the forward phase, the external input information signals at the input neurons are propagated forward to compute the output information signal at the output neuron. In the backward phase, modifications to the connection strengths are made based on the basis of the difference in the predicted and observed information signals at the output neuron [21].

If the used transfer function in the hidden layer is sigmoid, all samples must be normalized in the range of 0.1–0.9 [5]. So all of the data sets (X_i) were scaled to a new value x_i as follows:

$$x_i = 0.8\left(\frac{X_i - X_{\min}}{X_{\max} - X_{\min}}\right) + 0.1 \quad (10)$$

Each topology would be repeated several times to avoid random correlation due to the random initialization of the weights. The network performance could be stabilized after inclusion of defined number of nodes on the hidden layer. For instance, Salari et al. [22] have shown that based on the approximation of error function (MSE), a number of hidden neurons equal to sixteen was adopted in a feed-forward back propagation neural network and was used for modeling of the process studied.

3.1. Test of the fitted model

In order to calculate modeling errors, all of the outputs would performed an inverse range scaling to return the predicted responses to their original scale and compared them with experimental responses. Usually two lines are used to show the success of the prediction [23,24]. The one is the perfect fit (predicted data = experimental data), on which all the data of an ideal model should lay. The other line is the line that best fits on the data of the scatter plot with equation $Y = ax + b$ and it is obtained with regression analysis based on the minimization of the squared errors. If there were a perfect fit (outputs exactly equal to targets), the slope would be 1, and the y-intercept would be 0. The third variable is the correlation coefficient (R-value) between the outputs and targets. It is a measure of how well the variation in the output is explained by the targets. If this number is equal to 1, then there is perfect correlation between targets and outputs.

3.2. Relative importance of input variables

The ANN used provides the weights that are coefficients between the artificial neurons. These weights 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 [25].

The neural net weight matrix can be used to assess the relative importance of the various input variables on the output variables. Garson [26] has proposed an equation based on partitioning of connection weights:

$$I_j = \frac{\sum_{m=1}^{m=N_h}\left(\left(\left|W_{jm}^{ih}\right| / \sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|\right) \times \left|W_{mn}^{ho}\right|\right)}{\sum_{k=1}^{k=N_i}\left\{\sum_{m=1}^{m=N_h}\left(\left|W_{km}^{ih}\right| / \sum_{k=1}^{N_i}\left|W_{km}^{ih}\right|\right) \times \left|W_{mn}^{ho}\right|\right\}} \quad (11)$$

where I_j is the relative importance of the jth input variable on output variable, N_i and N_h are the number of input and hidden-neurons, respectively, W's are connection weights, the superscripts 'i', 'h' and 'o' refer to input, hidden and output layers, respectively, and subscripts 'k', 'm' and 'n' refer to input, hidden and output neurons, respectively.

Kasiri et al. [24] have used this equation to evaluate the relative importance of the input variables on color removal of the dye solution by UV/H_2O_2 process. The researchers have found that among the input variables studied, initial concentration of H_2O_2 with a relative importance of 48.89% appeared to be the most influential parameter in UV/H_2O_2 decolorization process.

3.3. Improving generalization

One of the problems that occur during neural network training is called "overfitting". The error on the training set is driven to a very small value, but when new data is presented to the network the error is large. The network has memorized the training examples, but it has not learned to generalize to new situations.

One method for improving network generalization is to use a network that is just large enough to provide an adequate fit. The larger network you use, the more complex the functions the network can create. If you use a small enough network, it will not have enough power to overfit the data. Unfortunately, it is difficult to know beforehand how large a network should be for a specific application. There are two other methods for improving called "regularization" and "early stopping" [12].

The first method for improving generalization is called regularization. This involves modifying the performance function, which is normally chosen to be the sum of squares of the network errors on the training set.

Another method for improving generalization is called early stopping. In this technique the available data is divided into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the validation set. The error on the validation set is monitored during the training process. The validation error normally decreases during the initial phase of training, as does the training set error. However, when the network begins to overfit the data, the error on the validation set typically begins to rise. When the validation error increases for a specified number of iterations, the training is stopped, and the weights and biases at the minimum of the validation error are returned.

Khataee and Mirzanjani [27] have described that a good method for preventing the overfitting is to use the validation data set periodically to compute the error rate for it while the network is being trained. The validation error decreases in the early epochs of back propagation but after a while, it begins to increase. The point of minimum validation error is a good indicator of the best number of epochs for training and the weight at that stage are likely to provide the best error rate in new data [5]. Their results indicated that the minimum error of the validation set could be achieved in the epochs just about 5000. After 5000 epochs the mean square error slightly increased; therefore, 5000 has been selected as the optimum epoch number.

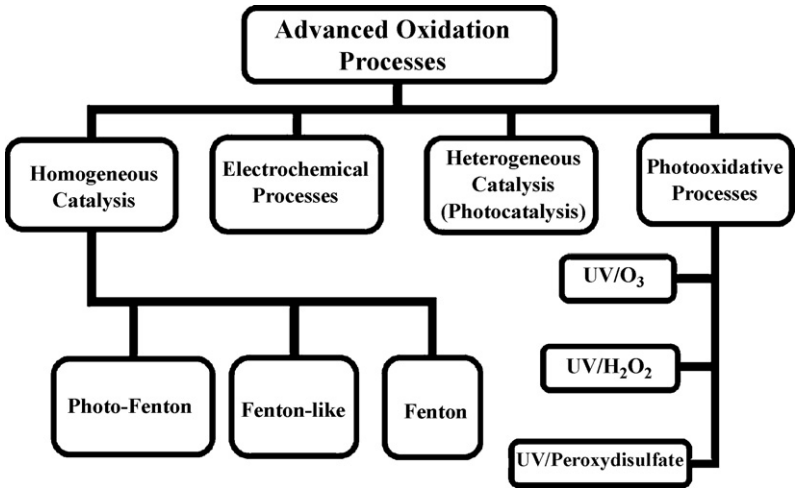

Fig. 3. Advanced oxidation processes (AOPs).

4. ANN modeling of heterogeneous nanocatalytic processes

In recent years, advanced oxidation processes (AOPs) have been developed as an alternative to the conventional water and wastewater treatment methods. These processes are based on the generation of very reactive species such as hydroxyl radicals that have been proposed to oxidize quickly and nonselectively a broad range of organic pollutants [28–32]. Different AOPs have been schematically indicated in Fig. 3. AOPs have been developed to degrade the nonbiodegradable contaminants of water into harmless species (e.g. CO_2, H_2O, etc.).

Heterogeneous photocatalysis via combination of a nanocatalyst (e.g. TiO_2 and ZnO) and UV light is considered one of the promising AOPs for destruction of soluble organic pollutants found in water and wastewater [30,32]. Some of the beneficial characteristics of TiO_2 nanoparticles in comparison to other photocatalysts are high photocatalytic efficiency, physical and chemical stability, low cost and low toxicity [33]. Interaction of TiO_2 nanoparticles with the photons that possess energy equal or higher to the band gap may cause separation of conduction and valence bands (see Fig. 3). This event is known as electron–hole pair generation. For TiO_2 nanoparticles, this energy can be supplied by photons with energy in the near ultraviolet range. This property promotes TiO_2 nanoparticles as a promising candidate in photocatalysis where solar light can be used as the energy source. Indeed, when TiO_2 nanoparticles is illuminated with $\lambda < 390$ nm light, an electron excites out of its energy level and consequently leaves a hole in the valence band. As electrons are promoted from the valence band to the conduction band, they generate electron–hole pairs (Eq. (12)) [33–35]:

$$TiO_2 + h\nu(\lambda < 390\ \text{nm}) \rightarrow e^- + h^+ \quad (12)$$

Valence band (h^+) potential is positive enough to generate hydroxyl radicals ($OH^{\bullet}$) at TiO_2 surface and the conduction band (e^-) potential is negative enough to reduce molecular oxygen as shown in Fig. 4.

The hydroxyl radical is a powerful oxidizing agent which may attack the organic matters present at or near the surface of TiO_2 nanoparticles. It is capable to degrade toxic and bioresistant compounds into harmless species (e.g. CO_2, H_2O, etc.). TiO_2 nanomaterials are successfully used for the photocatalytic remediation of a variety of organic pollutants such as hydrocarbons and chlorinated hydrocarbons (e.g. CCl_4, $CHCl_3$, C_2HCl_3, phenols, chlorinated phenols, surfactants, pesticides, and dyes) as well as reduction deposition of heavy metals such as Pt^{4+}, Pd^{2+}, Au^{3+}, Rh^{3+} and Cr^{3+} from aqueous solutions. TiO_2 nanomaterials have also been effective in the destruction of biological organisms such as bacteria, viruses, and molds [36–40].Due to the complexity of the reactions in the photocatalytic processes, the kinetic parameters of the various steps involved are very difficult to determine, leading to uncertainties in the design and scale-up of chemical reactors of industrial interest. This is caused by the complexity of solving the equations that involve the radiant energy balance, the spatial distribution of the absorbed radiation, mass transfer, and mechanisms of a photocatalytic degradation process involving radical species. Because of these reasons, the modeling of the photocatalytic processed via artificial neural networks is quite appropriate. As it was mentioned, one of the characteristics of modeling based on ANNs is that it does not require the mathematical description of the phenomena involved in the process. The examples of ANN modeling of photocatalytic water and wastewater treatment processes are summarized in Table 2.Artificial neural networks have been used for modeling of TiO_2 photocatalytic degradation of 2,4-dihydroxybenzoic acid, chosen as a model water contaminant, as a function of the concentrations of substrate and catalyst. The experimental design methodology has been applied to the choice of an appropriate set of experiments well distributed in the experimental region (Doehlert uniform array). Contrary to a classical treatment of the data, based on apparent rate constants modeled by a quadratic polynomial function, neural network analysis of the

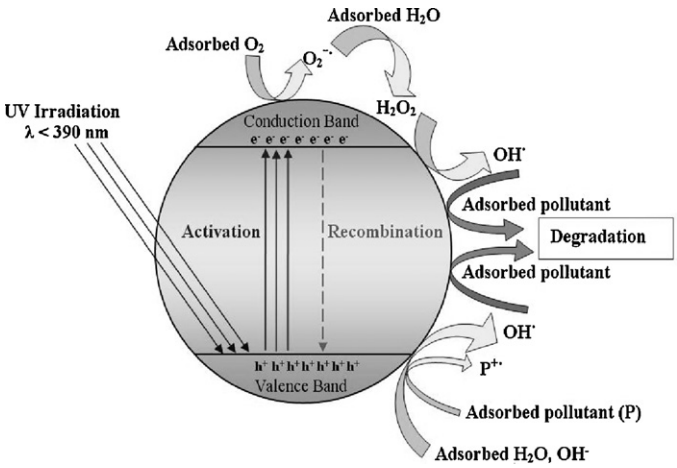

Fig. 4. Generation of photocatalytic active species at the surface of TiO_2 nanoparticles.

Table 2
ANN modeling of heterogeneous photocatalytic water and wastewater treatment processes.

Treatment Process	Treatment target	ANN architecture	Training function	Layers No.	ANN topology	Data no.	Input	Output	Epochs no.	Refs.
UV/TiO_2	Nitrilotriacetic acid (NTA)	Feed-forward back propagation	–	3	3:5:1	58	$[TiO_2]_0$, $[NTA]_0$, Time	NTA degradation (%)	5000	[41]
Solar/Fenton/TiO_2	Imipramine	Feed-forward back propagation	Conjugate gradient descent	3	3:3:1	16	$[H_2O_2]_0$, $[Fe(II)]_0$, $[TiO_2]_0$	Imipramine degradation (%)	95	[42]
UV/TiO_2	Nitrogen oxides (NO and NO_x)	Feed-forward	Quick propagation	5	3:7:4:3:2	488	$[TiO_2]_0$, Exposed surface, Time	NO and NO_x degradation (%)	2000	[43]
Solar/Fenton/TiO_2	Reactive blue 4	Feed-forward back propagation	Marquardt non-linear	–	–	19	pH, $[TiO_2]_0$, $[Fe(II)]_0$, $[H_2O_2]_0$	Decolorization kinetic constant (min^{-1})	–	[44]
Solar/H_2O_2/TiO_2	IGCC power station effluents	Feed-forward back propagation	Marquardt non-linear	–	–	11	$[TiO_2]_0$, $[H_2O_2]_0$	Cyanide and Formates degradation constant (min^{-1})	–	[45]
Solar/H_2O_2/ZnO	IGCC power station effluents	Feed-forward back propagation	Marquardt non-linear	–	–	11	$[ZnO]_0$, $[H_2O_2]_0$	Cyanide and Formates degradation constant (min^{-1})	–	[46]
UV/Immobilized–TiO_2	C. I. Basic red 46	Feed-forward back propagation	Train scaled conjugate gradient	3	4:8:1	140	pH, $[Dye]_0$, Time, UV light intensity	Decolorization (%)	5000	[47]
UV/TiO_2	2,4-dihydroxybenzoic acid (DHBA)	Feed-forward back propagation	–	3	3:4:1	66	$[DHBA]_0$, $[TiO_2]_0$, Time	$[DHBA]_t$ (mg/L)	20000	[48]
UV/TiO_2	Ethylene diamine tetraacetic acid (EDTA)	Feed-forward back propagation	–	3	3:4:4	26	$[TiO_2]_0$, $[EDTA]_0$, pH	EDTA degradation at 30, 60, 90, and 120 min	50	[49]

same experimental data does not require the use of any kinetic or phenomenological equations and allows the simulation and the prediction of the pollutant degradation as a function of irradiation time, as well as prediction of reaction rates, under varying conditions within the experimental region [48].An artificial neural network model has been developed to predict the photocatalytic decolorization of C.I. Basic Red 46 solution. This dye, commonly used as a textile dye, has been photocatalytically removed using supported TiO_2 nanoparticles irradiated by a 30 W UV-C lamp in a batch reactor. The photocatalyst has been industrial Degussa P25 (crystallite mean size 21 nm) immobilized on glass beads by a heat attachment method. The process of the dye decolorization in the presence of TiO_2 nanoparticles has been experimentally studied through changing the initial dye concentration, UV light intensity and initial pH. The influence of inorganic anions such as chloride, sulfate, bicarbonate, carbonate and phosphate on the photocatalytic decolorization of the dye has been investigated. The findings indicated that the proposed ANN model provides reasonable predictive performance ($R^2 = 0.96$). The influence of each parameter on the variable studied has been assessed: initial concentration of the dye being the most significant factor, followed by the initial pH and reaction time [47].A chemometric study on the TiO_2 photocatalytic degradation of nitrilotriacetic acid (NTA) in aqueous media under UV radiation has been carried out taking into account the multiple variables that take part in the system. To save redundant number of experiments, the system has been managed under chemometric techniques for several variables as NTA and TiO_2 concentrations, pH and irradiation time. Multiple-way analysis of the variance (MANOVA) has been applied to find the statistically significant variables. An artificial neural network has been used to build an empirical model of the system. All measurements have been driven under experimental designs: a full-factorial design has been used to analyze significant factors through MANOVA, and a Doehlert design, which was modified by spatial rotation, was applied in order to have a satisfactory number of levels for the factor time to be able to train the ANN. The study allows the knowledge and prediction of the behavior of the system as well as to work out kinetic parameters and to optimize their variables. The results of kinetic parameters obtained with the neural network agree with independent experimental results [41].Toma and co-workers [43] have been proposed an artificial neural network model to analyze the photocatalytic removal of nitrogen oxides over TiO_2 powders. It should be mentioned that the nitrogen oxides NO_x (especially NO and NO_2) forming by cars traffic, combustion of coals and thermal power plants participate in the formation of acid rain, greenhouse effect, photochemical pollution and major problems on the human health. 488 experimental sets have been used to feed the ANN model. The network input contains three neurons representing TiO_2 powder quantity (g), reaction time (min) and exposed surface (cm^2). The output pattern comprise two neurons representing the photocatalytic response, namely NO and NO_x efficiencies. Experimental sets are organized in training and test samples. The first category is used to tune neuron network weights and the second category to test the network configuration. A stopping criterion is applied corresponding to a fixed number of training and test cycles: 2000 cycles are achieved for each network configuration after approximately 1 h of computation time. The structure of the optimized ANN model characterize by three hidden layers containing seven, four and three neurons, respectively (see Table 2). Predicted results are in good agreement with experimental ones ($R^2 = 0.9857$). With the ANN optimized structure, it is possible to quantify the effect of each experiment variable by varying independently each of them and collecting NO and NO_x efficiencies [43].Emilio and co-workers [49] have been also used the chemometric techniques including full factorial and Doehlert experimental designs, multivariate analysis by MANOVA and artificial neural networks for the photocatalytic

reaction of ethylenediaminetetraacetic acid (EDTA) over TiO_2 in aqueous solution. EDTA concentration, TiO_2 amount, pH of the solution and irradiation time have been chosen to build a set of experiments for the analysis. Correlation plots among variables have been built a model for prediction the behavior of the photocatalytic system and optimizing parameters.The heterogeneous assisted photocatalytic degradation processes (i.e. Solar/H_2O_2/TiO_2 and Solar/H_2O_2/ZnO) of wastewater from a thermoelectric power station under concentrated solar light irradiation using a Fresnel lens has been also reported [45,46]. The efficiency of photocatalytic degradation processes has been determined from the analysis of cyanide and formate removal. The experimental kinetic constants have been fitted using neural networks. The ANN applied has been solved with two neurons and using a simple exponential activation function and the strategy is based on a back propagation calculation. Input variables have been initial concentration of hydrogen peroxide and photocatalyst while output data have been cyanides and formates degradation constant. The analysis of the relevance of each variable with respect to the others has been reported. It has been found that the initial concentration of hydrogen peroxide is the most significant factor affecting the degradation kinetic rate constants of cyanide and formate [45].

From the above studies, it can be concluded that the artificial neural networks can describe the behavior of the complex reaction system such as photocatalytic processes in the range of experimental conditions adopted. Simulation based on the ANN models can estimate the behavior of the photocatalytic processes under different conditions.

5. ANN modeling of homogeneous catalytic processes

As it was mentioned, AOPs are the alternative to the conventional water and wastewater treatment methods. Hydroxyl radicals ($^{\bullet}OH$), highly reactive species generated in sufficient quantities by these systems, have the ability to oxidize the majority of the organics in the polluted waters [50]. As shown in chapter 4, common AOPs involve heterogeneous and homogeneous photocatalysis [51,52].

Homogeneous photocatalytic processes including Fenton and photo-Fenton methods have been widely used for destruction of organics in the polluted waters. In this section, a brief mechanistic description of these techniques as well as the results of ANN modeling will be presented.

In 1876, Fenton's pioneering work pointed out the possible use of a mixture of H_2O_2 and Fe^{2+} to destroy tartaric acid. Most people, however, consider that Fenton's chemistry began in 1894 when he published a deeper study on the strong promotion of the oxidation of this acid with such a reagent. During the period 1901–1928, the stoichiometry of the reaction between H_2O_2 and Fe^{2+} has been studied. The extraordinary practical usefulness of Fenton's reagent for the oxidation of organic compounds was first assumed in the 1930s as a radical mechanism for the catalytic decomposition of H_2O_2 by iron salts [53].

Fenton's reagent (Fe^{2+}/H_2O_2) is known to have different treatment functions, as mentioned earlier, depending on the $H_2O_2/FeSO_4$ ratio. When the amount of Fe^{2+} employed exceeds that of H_2O_2, the treatment tends to have the effect of chemical coagulation. When the two amounts are reversed, the treatment tends to have the effect of chemical oxidation.

The wastewater treatment applying Fenton and photo-Fenton processes is, in general, quite complex. This is caused by the complexity of solving the equations that involve the radiant energy balance, the spatial distribution of the absorbed radiation, mass transfer, and the mechanisms of a photochemical or photocatalytic degradation involving radical species. Since the process depends on several factors, modeling of these processes involves many problems. It is evident that these problems cannot be solved by simple linear multivariate correlation. Artificial neural networks are now commonly used in many areas of chemistry and they represent a set of methods that may be useful in solving such problems (see Table 3) [7,8,54–56].

Elmolla et al. [57] have used neural network modeling to predict the performance of Fenton process for removal of antibiotics (amoxicillin, ampicillin and cloxacillin). The configuration of the back propagation neural network giving the smallest MSE was three layers ANN with tangent sigmoid transfer function (tansig) at hidden layer with 14 neurons, linear transfer function (purelin) at output layer and Levenberg–Marquardt back propagation training algorithm (LMA). ANN predicted results were very close to the experimental results with correlation coefficient (R^2) of 0.997 and MSE 0.000376.

Yu et al. [58] have also developed a novel Fenton process control strategy using ANN models and oxygen-reduction potential (ORP) monitoring to treat two synthetic textile wastewaters containing two common dyes namely Reactive Blue 49 (RB49) and Reactive Brilliant Blue (RBB). Experimental results indicate that the ANN models can predict precisely the color and chemical oxygen demand (COD) removal efficiencies for synthetic textile wastewaters with correlation coefficients of 0.91–0.99.

When UV irradiation is combined with some powerful oxidant, such as H_2O_2, organic dye degradation efficiency can be significantly enhanced due to hydroxyl radical generation caused by the photolysis of H_2O_2 (Eq. (16)), and these highly reactive non–selective radicals may further react with the organic substrate [59]. This process demonstrated high efficiency in the treatment of different types of organic dyes [60–63]. The efficiency of Fenton process could be significantly increased under light irradiation, where Fe^{3+} ions are constantly reduced to the Fe^{2+}, (Eq. (13)), [64] and the Fenton process is improved by the participation of photogenerated Fe^{2+}:

$$Fe^{3+} + H_2O + h\upsilon \rightarrow Fe^{2+} + OH^{\bullet} + H^{+} \quad (13)$$

ANN modeling of photo-Fenton processes is the subject of numerous papers (see Table 3). Durán et al. [65] have developed a three layers (4:4:2) feed-forward network to predict degradation rate constant of cyanides and formates under UV/Fe(II)/H_2O_2 process in a integrated gasification combined cycle (IGCC) power station effluent. In a similar work, they have used a two layers network to evaluate the efficiency of photo-Fenton process [66]. The network has been trained by a marquardt non-linear fitting algorithm to simulate the output parameters: decolorization and mineralization rate constants. Simulation from ANNs equations has proved that the initial concentration of hydrogen peroxide in aqueous dye solutions is the main parameter affecting the photo-decolorization kinetics (see Table 3).

Gob et al. [54] have studied photo-Fenton removal of 2,4-dimethyl aniline (2,4-xylidine) from contaminated water. A three layers (3:8:1) feed-forward back propagation network has been developed and trained using 50,000 data sets. Comparison made between predicted and experimental output values ($R^2 = 0.995$) show that ANN is a successful technique to predict 2,4-xylidine concentration in the treated solution (see Table 3).

Treatment of saline wastewater contaminated with hydrocarbons by the photo-Fenton process has been the subject of another ANN modeling. In this work, Moraes et al. [8] have followed TOC content of the treated wastewater using a three layer (5:2:1) feed-forward back propagation network. Totally 1000 data sets have been used for training the network. There has been a good agreement between experimental and predicted output values with high correlation coefficients of 0.950 and 0.965 for learning and test sets, respectively.

Table 3
ANN modeling of Fenton and photo–Fenton water and wastewater treatment processes.

Treatment Process	Treatment target	ANN architecture	Training function	Layers no.	ANN topology	Data no.	Input	Output	Epochs no.	Refs.
Photo–Fenton	Saline solutions containing raw gasoline	Feed–forward back propagation	–	3	5:2:1	–	Reaction time, $[TOC]_0$, $[Fe(II)]_0$, $[H_2O_2]_0$, $[NaCl]_0$	*[TOC]	1000	[8]
UV/Fe–ZSM5/H_2O_2	Acid Red 14	Feed–forward back propagation	Scaled conjugate gradient algorithm	3	4:10:1	25	$[Dye]_0$, $[H_2O_2]_0$, Initial pH, $[Catalyst]_0$,	Degradation (%)	–	[24]
Photo–Fenton	Reactive Blue 4	Feed–forward back propagation	Marquardt non–linear fitting algorithm	3	4:2:1	19	$[H_2O_2]_0$, pH, $[TiO_2]_0$, $[Fe(II)]_0$	Decolorization kinetic constant	–	[74]
Photo–Fenton	Imipramine	Feed–forward back propagation	Conjugate gradient descent	3	3:3:1	–	$[H_2O_2]_0$, $[Fe(II)]_0$, $[TiO_2]_0$	Imipramine degradation (%)	95	[42]
Photo–Fenton	2,4–dimethyl aniline	Feed–forward back propagation	–	3	3:8:1	–	$[Fe(II)]_0$, $[H_2O_2]$, Reaction time	*[Xyl]	50000	[54]
Fenton	Antibiotics (amoxicillin, ampicillin and cloxacillin)	Feed–forward back propagation	Levenberg–Marquardt	3	5:14:1	120	Reaction time, H_2O_2/COD molar ratio, H_2O_2/Fe(II) molar ratio, pH, $[Antibiotics]_0$	*COD removal (%)	–	[57]
Fenton	Reactive Brilliant Blue, Reactive Blue 49	–	–	–	–	–	$[Fe(II)]_0$, $[H_2O_2]_0$	COD removal, Effluent color removal	–	[58]
UV/Fe(II)/H_2O_2	Integrated Gasification Combined Cycle (IGCC) power station effluent	Feed–forward back propagation	–	3	4:2:2	–	$[Fe(II)]_0$, $[H_2O_2]_0$, pH, Temperature	Degradation rate constant of cyanides and formates	–	[65]
Photo–Fenton	Reactive Blue 4 (RB4)	Marquardt non–linear fitting algorithm	–	2	5:2	–	$[RB4]_0$, $[Fe(II)]_0$, $[H_2O_2]_0$, pH, Temperature	Decolorization and mineralization rate constant	–	[66]
Ferrioxalate–assisted photo–Fenton	Orange II	Feed–forward back propagation	Marquardt non–linear fitting algorithm	3	7:2:2	–	H_2O_2 Flow, $[Fe(II)]_0$, pH, $[H_2C_2O_4]_0$, Temperature Solar power, UV Dosage	Decolorization and mineralization kinetic rate constants	–	[67]
Photo–Fenton	Polyvinyl alcohol (PVA)	Feed–forward back propagation	–	3	4:8:1	432	Reaction time, $[DOC]_0$, $[Fe(II)]_0$, $[H_2O_2]$	[PVA]	10000	[68]
Solar driven photo–Fenton	Phenol	Feed–forward back propagation	–	3	5:6:1	–	$[DOC]_0$,$[Fe(II)]_0$, H_2O_2 feed rate, Accumulated radiant energy, Irradiation time	*[DOC]	10000	[70]
Solar photo–Fenton	Reactive Blue 4 (RB4)	Feed–forward back propagation	Marquardt non–linear fitting algorithm	3	5:2:1	–	pH, $[Fe(II)]_0$, $[H_2O_2]_0$, $[RB4]_0$, Temperature	Decolorization rate pseudo constant	–	[75]
Solar photo–Fenton–ferrioxalate					6:2:1		$[Oxalic\ acid]_0$			
Solar photo–Fenton	Orange II	–	–	3	5:2:1	–	pH, $[FeSO_4]_0$, $[H_2O_2]_0$, $[Orange\ II]_0$, Temperature	Decolorization kinetic rate constant	–	[76]
Solar photo–Fenton–ferrioxalate					4:2:1		$[FeSO_4]_0$, $H_2O_2]_0$, Temperature, $[Oxalic\ acid]_0$			
Ferrioxalate–assisted solar photo–Fenton	Orange II	–	–	3	7:2:2	50	pH, $[Fe(II)]_0$, H_2O_2 flow, Temperature, $[H_2C_2O_4]_0$, Solar power, Air flow	Decolorization and mineralization kinetic rate constants	–	[77]

TOC: Total Organic Carbon, Xyl: 2,4- dimethyl aniline, COD: Chemical Oxygen Demand. DOC: Dissolved Organic Carbon.

Table 4
ANN modeling of photooxidative processes for water and wastewater treatment.

Treatment Process	Treatment target	ANN Architecture	Training function	Layers No.	ANN topology	Data no.	Input	Output	Epochs no.	Refs.
UV/H_2O_2	Acid Orange 7	Feed–forward back propagation	Scaled conjugate gradient algorithm	3	4:8:1	228	$[Dye]_0$, $[H_2O_2]_0$, Initial pH, Reaction time	Color removal (%)	–	[23]
UV/H_2O_2	Reactive Red 120	Counter–propagation learning strategy	Kohonen algorithm	4	7:8:8:4	–	UV light intensity, $[H_2O_2]_0$, $[NaCl]_0$, Reaction time, $[NaOH]_0$, [Dye], [*Oxidant]$_0$	Absorbance, *COD, TOC, TIC	200	[25]
UV/peroxydisulfate	Basic Blue 3	Feed–forward back propagation	Scaled conjugate gradient algorithm	3	4:8:1	177	Reaction time, $[Dye]_0$, $[Peroxydisulfate]_0$, UV light intensity,	Color removal (%)	5000	[27]
UV/H_2O_2	Methyl tert-butyl ether (MTBE)	Feed–forward back propagation	Scaled conjugate gradient algorithm	3	4:8:1	64	$[MTBE]_0$, $[H_2O_2]_0$, pH, Reaction time	$[MTBE]_t$	–	[89]
UV/H_2O_2	Acid Brown 75, Acid Orange 52, Acid Orange 10, Direct Red 28	Feed–forward back propagation	Levenberg–Marquardt	3	7:18:1	757	Azo bond number, Sulphonate group number, $[Dye]_0$, pH, H_2O_2 volume, Temperature, Time of operation	Decolorization degree	34	[90]
UV/H_2O_2	Acid Orange 52	Feed–forward back propagation	Levenberg–Marquardt	3	5:16:1	218	H_2O_2 volume, pH, Temperature, $[Dye]_0$, Time of operation	Decolorization (%)	22	[91]
UV/H_2O_2	Acid Brown 75	Feed–forward back propagation	Levenberg–Marquardt	3	5:116:1	528	$[Dye]_0$, pH, H_2O_2 volume, Temperature, Time of operation	Decolorization degree	49	[92]

COD: Chemical Oxygen Demand, TOC: Total Organic Carbon, TIC: Total Inorganic Carbon, Oxidant: *m*-nitrobenzene sulphonate.

Monteagudo et al. [67] have also developed a neural network for modeling of ferrioxalate-assisted solar photo-Fenton degradation of Orange II aqueous solutions. The three layers (7:2:2) feed-forward back propagation network has been trained using Marquardt non-linear fitting algorithm which can predict the decolorization kinetic rate constant chosen as the output variable. They have reported that experimental results and ANNs fittings of the process are in good agreement with an average error lower than 16% for dye decolorization.

Calza et al. [42] have developed a neural network to predict the performance of a photo-Fenton process for removal of imipramine from the contaminated water. The three layers network (3:3:1) has been trained with conjugate gradient descent algorithm during 95 epochs. The linear regression between the network prediction and the corresponding experimental data prove that modeling of photo-Fenton removal of imipramine using artificial neuron network is a satisfactory method. Modeling photo-Fenton removal of Reactive Blue 4 is another example of the application of neural network technique. Duran et al. [44] have developed a three layers (4:2:1) feed-forward neural network using back propagation algorithm. Totally 19 data sets and marquardt non-linear fitting algorithm have been used for training the network that is successfully enabled to predict the output variable, i.e. decolorization kinetic constant.

Giroto and co-workers [68] have also developed a neural network to predict polyvinyl alcohol abatement in aqueous solution by photo-Fenton process. A three layers (4:8:1) feed-forward back propagation network has been trained using 432 data sets and during 10,000 epochs to predict polyvinyl alcohol concentration at the end of photo-Fenton process. High correlation coefficient (R^2 = 0.996) between experimental and predicted values of the output variable shows the success of the modeling.

Nogueira et al. [69] have also developed a neural network for modeling the performance of a solar driven photo-Fenton process used for removal of phenol from the effluents. In this work, dissolved organic carbon (DOC) content of the treated solution has been chosen as the output variable and a three layers (5:6:1) feed-forward back propagation neural network has been trained during 10,000 epochs.

Homogeneous photo-Fenton has some disadvantages such as (i) the tight range of pH in which the reaction proceeds, (ii) the need for recovering the precipitated catalyst after the treatment and (iii) deactivation by some ion complexing agents like phosphate anions [70]. An alternative method could be the use of heterogeneous solid Fenton nanocatalysts, such as transition metals containing zeolites, clays, bentonits and so on [198–200]. The use of synthetic zeolites is very promising due to their unique properties such as micro-porous structure, high surface area and ion exchange capacity, which could give them advantage over other carriers [71–73].

Kasiri et al. [24] have studied photo-Fenton process using Fe-ZSM5 zeolite as heterogeneous nanocatalyst for removal of Acid Red 14 from the contaminated water. They have used an artificial neural network for modeling of the process. The three layers (4:10:1) network has been trained using scaled conjugate gradient algorithm and the degradation efficiency of the process has been chosen as the output variable. Modeling results show a good agreement between experimental and predicted results with a high correlation coefficient (R^2 = 0.996). The results of modeling confirm that neural network modeling could effectively reproduce experimental data and predict the behavior of the process.

6. ANN modeling of UV/peroxide processes

Hydrogen peroxide (H_2O_2) is a strong oxidant (standard potential 1.80 and 0.87 V at pH 0 and 14, respectively) [78] and its application in the treatment of various inorganic and organic pol-

Table 5
ANN modeling of electrochemical water and wastewater treatment processes.

Treatment Process	Treatment target	ANN architecture	Training function	Layers No.	ANN topology	Data no.	Input	Output	Epochs no.	Refs.
Peroxi-coagulation	C.I. Basic blue 3, Malachite green, C.I. Basic red 46, C.I. Basic yellow 2	Feed-forward back propagation	trainscg	3	4:14:1	60	Electrolysis time, Initial pH, Applied current, $[Dye]_0$	Decolorization (%)	–	[9]
Peroxi-coagulation	C.I. Basic yellow 2	Feed-forward back propagation	trainscg	3	4:16:1	117	Electrolysis Time, Initial pH, Applied current, $[Dye]_0$	Decolorization (%)	–	[22]
Electrochemical oxidation	Real wastewater (Wastewater of Specialty Chemical Industry)	Feed-forward back propagation	Levenberg–Marquart	3	3:5:1	124	Current density, Electrolysis duration, Supporting electrolyte concentration	COD removal (%)	–	[94]
					3:7:1					
					3:9:1					
				4	3:3:3:1					
					3:3:5:1					
					3:3:7:1					
				3	7:10:1	49	Current density, Electrolysis Time, Initial pH, $[Dye]_0$, Conductivity, Retention time, Inter-electrodes distance	Decolorization (%)	–	[95]
Electrocoagulation	Cr(IV)	Feed-forward back propagation	–	3	4:10:1	212	Current density, Time of electrolysis, $[Cr(VI)]_0$, Concentration of electrolyte	$[Cr(VI)]_t$	–	[99]
Electrocoagulation	Acid blue 113	Feed-forward back propagation	trainlm	5	4:10:10:10:1	144	Effluent concentration, Electrolyte pH, Current density, Electrolysis time	COD removal (%)	–	[100]
Electrolysis	Phenolic compounds (phenol, 4-chlorophenol, 2,4-dichlorophenol, 2,4,6-trichlorophenol, 4-nitrophenol and 2,4-dinitrophenol)	Feed-forward back propagation	Different functions	4	7:30:25:1	420	Temperature, $[COD]_0$, pH, Current density, Current charge passed, Types of chlorine phenol compound, Type of nitrophenols compounds	COD removal (%)	–	[111]
					7:25:20:1					
Electrodialysis	Pb^{2+}	Multilayer perceptron	Levenberg–Marquardt	4	4:5:4:1	81	$[Pb^{2+}]_0$ Temperature, Flow rate, Voltage	$[Pb^{2+}]$	–	[112]
Coagulation	Drinking water	Self organizing map	–	4	7:8:0:3	202	Turbidity, Color, Absorption 254, Residual aluminum, pH, Alum dose, DOC	Turbidity, Color, Absorption 254 nm, Residual aluminum	–	[113]
					10:8:0:2					
					9:8:8:1					

lutants is well established. Numerous applications of H_2O_2 in the removal of pollutants from wastewater, such as sulfites, hypochlorites, nitrites, cyanides, and chlorine have been reported [79].

Oxidation by H_2O_2 alone is not effective for high concentrations of certain refractory contaminants, such as highly chlorinated aromatic compounds and inorganic compounds, because of low rates of reaction at reasonable H_2O_2 concentrations. Transition metal salts (e.g. iron salts), ozone and UV–light can activate H_2O_2 to form hydroxyl radicals (Eqs. (14)–(16)) which are strong oxidants:

- Ozone and hydrogen peroxide

$$O_3 + H_2O_2 \rightarrow OH^{\bullet} + O_2 + HO_2 \quad (14)$$

- Iron salts and hydrogen peroxide [80]

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^{\bullet} + OH^- \quad k_1 \approx 70\,M^{-1}\,s^{-1} \quad (15)$$

- UV-light and hydrogen peroxide

$$H_2O_2 + UVC \rightarrow 2OH^{\bullet} \quad (16)$$

The principle behind the beneficial effects observed using ultraviolet light in combination with hydrogen peroxide or ozone as compared to the individual application, lies in the fact that the rate of generation of free radicals is significantly enhanced in the case of combination technique, which is very similar to ultrasoundy H_2O_2 or ultrasoundy O_3 processes. Only difference being the energy required for the generation of free radicals from dissociation of ozone or hydrogen peroxide is given by the UV light as against cavitating bubbles in the case of ultrasound [81,82].

It is widely accepted that the first step in the UV/H_2O_2 process is the attack of the photon against the hydrogen peroxide molecule and the subsequent formation of hydroxyl radical ($OH^{\bullet}$) (Eq. (16)) [83].

High concentrations of H_2O_2 do not necessarily favor the kinetics of the reaction, for after the reaction starts, the steps of propagation can be prevented by the excess of hydrogen peroxide. This excess can act as a hydroxyl radical self-consumer (Eq. (17)) [84].

$$OH^{\bullet} + H_2O_2 \rightarrow H_2O + HO_2^{\bullet} \quad (17)$$

This scavenging effect of hydroxyl radical produces the hydroperoxy radical, less reactive than the hydroxyl radical. Thus, hydrogen peroxide in excess may react with the hydroxyl radical and compete with the attack of this radical to the organic matters in the solution during the photolysis [85].

The kinetics of UV/H_2O_2 process is favored up to H_2O_2 addition critical point. The critical point is related to various factors such as the amount of hydrogen peroxide added, reaction media pH, UV radiation wavelength, concentration of organic matters and structural characteristics, besides other specific factors like the presence of inorganic salts, which affect the reaction performance of the hydroxyl radical [86].

Wastewater treatment by applying photooxidative processes is, in general, quite complex. The mathematical equations describing the performance of theses processes involve the radiant energy balance, the spatial distribution of the absorbed radiation, mass transfer, and the mechanisms involving radical species. It is clear that such a complex problem cannot be solved by simple linear multivariate correlation. Artificial neural networks are very useful to solve this problem as they do not require the mathematical description of the phenomena involved in the process [23,87,88]. So, there are numerous papers regarding ANN modeling of these processes (see Table 4).

Aleboyeh et al. [23] have developed an artificial neural network to predict the performance of a UV/H_2O_2 removal of Acid Orange 7 from the aqueous solution. The network has been trained using totally 228 data sets divided into training, validation and test subsets, each of them containing 114, 57 and 57 data sets, respectively. A three layers network (4:8:1) has been optimized to predict the decolorization efficiency of UV/H_2O_2 process. The comparison between experimental values and predicted output variables using the adopted neural network model show that this network predicts the output variable with a high correlation coefficient ($R^2 = 0.996$). The results of modeling confirm that neural network modeling could effectively reproduce experimental data and predict the behavior of the process.

Salari and co-workers [89] have used ANN technique for modeling of MTBE removal by UV/H_2O_2 process. Totally 64 data sets have been used for training, validation and test of the model. The configuration of the back propagation neural network giving the smallest MSE has been a three layers ANN with tangent sigmoid transfer function (tansig) at hidden layer with 8 neurons, linear transfer function (purelin) at output layer trained with scaled conjugate gradient algorithm. ANN predicted results are very close to the experimental ones with correlation coefficient of 0.998. The ANN model can then describe the behavior of the complex reaction system with the range of experimental conditions adopted.

Modeling of UV/H_2O_2 removal of Reactive Red 120 has been also studied by Slokar et al. [25]. A four layers (7:8:8:4) neural network has been developed during 200 epochs on the basis of counter-propagation learning strategy and Kohonen algorithm. This optimized network can effectively predict the output variables including absorbance, chemical oxygen demand, total organic carbon and total inorganic carbon of the dye solution (see Table 4).

Guimaraes and Silva have established a hybrid neural model for decolorization of azo dyes by UV/H_2O_2 involving the study of process variables and structural parameters [90]. Decolorization degree of the studied dyes including Direct Red 28, Acid Brown 75, Acid Orange 52 and Orange 10 has been chosen as the output variable. The network has been trained by Levenberg–Marquardt algorithm during 34 epochs and consequently, a three layers (7:18:1) neural network is optimized for modeling of UV/H_2O_2 process. The neural model provided optimum estimates for the decolorization based on the absorbance measurement as an output variable, with correlation coefficients above 0.96 for the training, validation and test sets, indicating the optimum model generalization capacity.

As a conclusion, artificial neural networks can successfully describe the performance of the photooxidative processes in the range of the variables studied, in spite of the complexity involved such processes.

7. ANN modeling of electrochemical treatment processes

Electrochemical technologies can be applied for the treatment of effluents released from a wide range of industries or processes. These techniques have been receiving greater attention in recent years due to their distinctive advantages such as environmental compatibility (the main reactant is the electron which is a clean reagent), and versatility (a plethora of reactors and electrode materials, shapes, and configurations can be utilized). It is noteworthy that the same reactor can be used frequently for different electrochemical reactions with only minor changes and also the electrolytic processes can be scaled easily from the laboratory to the plant, allowing treatment volumes ranging from milliliters to millions of liters. Electrochemical methods are generally safe because of the mild conditions and innocuous nature of the chemicals usually employed. Electrodes and cells can be designed to minimize power losses due to poor current distribution and voltage drops and making the processes more competitive in energy consumption than the conventional techniques [93].

Artificial neural networks have been widely used for modeling of electrochemical methods of water and wastewater treatment (see Table 5).

For example, Basha et al. [94] have applied ANN method for modeling the electrochemical degradation in batch, batch recirculation and continuous recycle modes in managing the wastewater discharged by a typical medium-scale specialty chemical industry. They have developed three and four layers back propagation feed-forward nets trained with Levenberg-Marquardt algorithm. It has been reported that single hidden layer feed-forward back propagation neural network is adequate enough to predict the performance of the process.

Daneshvar et al. [95] have used a three layer feed-forward back propagation neural network for modeling the electrocoagulation removal of C.I. Basic Yellow 28.

Electrocoagulation (EC) as an electrochemical method has been developed to overcome the drawbacks of conventional water and wastewater treatment technologies. EC process provides a simple, reliable and cost-effective method for the treatment of wastewater without any need for additional chemicals, and thus the secondary pollution. It also reduces the amount of sludge, which needs to be disposed [96–98].

Daneshvar and co-workers have reported that ANN modeling could be successfully used to investigate the cause effect relationship in electrocoagulation process [95]. The ANN model could describe the behavior of the complex reaction system with the range of experimental conditions adopted. Simulation based on the ANN model can estimate the behavior of the system under different conditions.

Aber et al. [99] have also developed a three layer feed-forward back propagation neural network for modeling of the electrocoagulation removal of Cr(VI) from the polluted solutions. The network was developed with sigmoidal transfer function as a transfer function in the hidden and output layers. 212 experimental sets have been used to develop the ANN model. The linear regression between the network prediction and the corresponding experimental data (R^2 = 0.976) proves that modeling of the electrocoagulation removal of Cr(VI) using artificial neural network is a good and precise method to predict the residual concentration of Cr(VI) under different conditions.

A five layer (4:10:10:10:1) back propagation feed-forward network has been also developed to predict the electrocoagulation removal of Acid Blue 113. The training function was *trainlm* and totally 144 experimental sets have been used to develop the ANN model. The authors have reported that ANN model predictions satisfactorily match with experimental observation [100].

One of the most popular electrochemical advanced oxidation process (EAOP) is anodic oxidation (AO) consisting in the destruction of organics in an electrolytic cell under the action of hydroxyl radical formed as intermediate from water oxidation to O_2 at the surface of a high O_2-overvoltage anode:

$$M + H_2O \rightarrow M(OH^{\bullet}) + H^+ + e^- \quad (18)$$

where M($OH^{\bullet}$) denotes the hydroxyl radical adsorbed on the anode M or remaining near its surface [101].

More potent indirect electro-oxidation methods with hydrogen peroxide electrogeneration are being also developed for wastewater remediation. In these techniques, H_2O_2 is continuously supplied to the contaminated solution from the two-electron reduction of O_2 usually at carbon-felt [102–107] and carbon-polytetrafluoroethylene O_2-diffusion [74,108] cathodes:

$$O_2(g) + 2H^+ + 2e^- \rightarrow H_2O_2 \quad (19)$$

The electro-Fenton process performs when Fe^{2+} is added to the solution.

$$Fe^{2+} + H_2O_2 + H^+ \rightarrow Fe^{3+} + OH^{\bullet} + H_2O \quad (20)$$

The peroxi-coagulation process is performed with a sacrificial Fe anode, which continuously supplies soluble Fe^{2+} to the solution from the following anodic oxidation reaction [109]:

$$Fe \rightarrow Fe^{2+} + 2e^- \quad (21)$$

Fe^{2+} thus produced is quickly oxidized by electrogenerated H_2O_2 from reaction (21) yielding a solution saturated with Fe^{3+}, whereas the excess of this ion precipitates as hydrated Fe(III) oxide ($Fe(OH)_3$).

$$Fe^{2+} + H_2O_2 \rightarrow Fe(OH)^{2+} + OH^{\bullet} \quad (22)$$

Then, pollutants can be removed by the combined action of their degradation with $OH^{\bullet}$ generated from reaction (22) and their coagulation with the $Fe(OH)_3$ precipitate formed.

Peroxi-coagulation differs from classical electrocoagulation with a Fe anode and a graphite cathode, since electrocoagulation does not degrade soluble organics because no significant H_2O_2 is produced in the medium, such as previously found for aniline and 4-chlorophenol [109,110].

Salari et al. [22] have developed a feed-forward back propagation neural network for modeling of peroxi-coagulation removal of C.I. Basic Yellow 2. The network is developed with sigmoidal transfer function as a transfer function in the hidden and output layers. The three layer network was trained with scaled conjugate gradient algorithm using totally 117 data sets. The linear regression between the network prediction and the corresponding experimental data (R^2 = 0.9713) proves that modeling the peroxi-coagulation removal of BY2 using artificial neuron network is a satisfactory method. The authors have concluded that artificial neural network modeling has been successfully used to investigate the cause effect relationship in peroxi-coagulation process.

In a similar work, Zarei et al. [9] have used neural network modeling to predict the performance of a peroxi-coagulation removal of four dyes namely Basic Blue 3, Malachite Green, Basic Red 46 and Basic Yellow 2 using carbon nanotube-PTFE cathode. They have used 60 data sets to develop a three layer (4:14:1) feed-forward network to predict the decolorization efficiency of the process as the output of the neural network. The linear regression between the network prediction and the corresponding experimental data (R^2 = 0.989) show that ANN modeling the peroxi-coagulation removal of Basic Yellow 2 as a model dye is a precious method.

Neural networks have also been developed to model the electrolysis of wastes polluted with phenolic compounds, including phenol, 4-chlorophenol, 2,4-dichlorophenol, 2,4,6-trichlorophenol, 4-nitrophenol and 2,4-dinitrophenol [111]. In this work, Piuleac and co-workers have proposed feed-forward four layers networks (7:30:25:1 and 7:25:20:1) using different training algorithms to predict COD content of the treated solutions.

Sadrzadeh et al. [112] have developed a neural network to predict the separation of lead ions from wastewater using electrodialysis method. In this work, 81 data sets were used to train a multilayer perceptron network using Levenberg-Marquardt algorithm. The developed four layers network (4:5:4:1) can effectively predict lead ions concentration in the treated solution. The authors have reported that ANN modeling technique have many favorable features such as efficiency, generalization and simplicity, which make it an attractive choice for modeling of complex systems, such as wastewater treatment processes.

It can be concluded that artificial neural networks are capable to simulate the complex relationships existing between input and

output process variables in electrochemical methods. ANNs can overcome the difficulty of modeling such processes where different phenomena such as mass and heat transfers and mechanic fluids are involved to run the overall process.

8. Conclusions

Artificial neural networks have been widely used in modeling and simulation of homogeneous and heterogeneous water and wastewater treatment processes. Catalytic treatment processes are, in general, quite complex. This is caused by the complexity of solving the equations that involve the radiant energy balance, the spatial distribution of the absorbed radiation, mass transfer, and the mechanisms of a photochemical or photocatalytic degradation involving radical species. This paper confirms that ANN modeling is an effective and simple approach to successfully describe the behavior of these complex processes, in which manipulated operational variables show a combined effect, within the range of experimental conditions investigated. ANNs are capable to simulate the complex relationships existing between input and output process variables in different homogeneous and heterogeneous water and wastewater treatment processes.

As an effective tool to evaluate the success of ANN modeling, a relationship between the predicted results of the designed ANN model and experimental data has been normally conducted. The high correlation coefficient (R^2) between the network prediction and the corresponding experimental data proves that modeling these nanocatalytic processes using artificial neuron network is a satisfactory method. This technique might therefore be useful in homogeneous and heterogeneous processes modeling, as well as in the design, scale-up and industrial application of water and wastewater treatment processes.

Acknowledgements

We are grateful to the University of Tabriz and Tabriz Islamic Art University, Iran for the all supports.

References

[1] W.S. McCulloch, W. Pitts, Bull. Mafh. Biophys. 5 (1943) 113–133.
[2] B. Krose, P.V. Smagt, An Introduction to Neural Networks, eighth ed., The University of Amsterdam, 1996.
[3] H. Zhou, D.W. Smith, J. Environ. Eng. Sci 1 (2002) 247–264.
[4] J. Mallevialle, P.E. Odendall, M.R. Wiesner, Water Treatment Membrane Processes, McGraw–Hill, New York, 1996.
[5] F. Despange, D.L. Massart, Analyst 123 (1998) 157–178.
[6] S. Lek, J.F. Guegan, Ecol. Model. 120 (1999) 65–73.
[7] V.K. Pareek, M.P. Brungs, A.A. Adesina, R. Sharma, J. Photochem. Photobiol. A 149 (2002) 139–146.
[8] J.E.F. Moraes, F.H. Quina, C.A.O. Nascimeto, D.N. Silva, O. Chiavone-Filho, Environ. Sci. Technol. 38 (2004) 1183–1187.
[9] M. Zarei, A. Niaei, D. Salari, A.R. Khataee, J. Electroanal. Chem. 639 (2010) 167–174.
[10] D. Howard, M. Beale, M. Hagan, Neural Network Toolbox For Use with MATLAB® User's Guide, Version 5, Mathworks Inc, 2006.
[11] M.F. Moller, Neural Networks 6 (1993) 525–533.
[12] Z. Zhang, K. Friedrich, Comp. Sci. Technol. 63 (2003) 2029–2044.
[13] M. Parizeau, Réseaux de Neurones, GIF–21140 et GIF–64326, Université Laval, 2006.
[14] J. Si, A. Barato, W. Powell, D. Wunsch, Handbook of Learning and Approximate Dynamic Programming, IEEE Press, John Wiley & Sons, 2004.
[15] C. Robert, C. Guilpin, A. Limoge, J. Neurosci. Meth. 79 (1998) 187–193.
[16] S. Gob, E. Oliveros, S.H. Bossmann, A.M. Braun, R. Guardani, C.A.O. Nascimento, Comput. Ind. Eng. 53 (2007) 95–122.
[17] C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, USA, 1995.
[18] A. Alvarez, Neural Process. Lett. 16 (2002) 43–52.
[19] H. Neudecker, J.R. Magnus, Matrix Differential Calculus with Applications in Statistics and Econometrics, John Wiley & Sons, New York, 1998.
[20] M.T. Hagan, M. Menhaj, IEEE Trans. Neural Netw. 5 (1999) 989–993.
[21] M.M. Hamed, M.G. Khalafallah, E.A. Hassanien, Environ. Model. Soft. 19 (2004) 919–928.
[22] D. Salari, A. Niaei, A.R. Khataee, M. Zarei, J. Electroanal. Chem. 629 (2009) 117–125.
[23] A. Aleboyeh, M.B. Kasiri, M.E. Olya, H. Aleboyeh, Dyes Pigments 77 (2008) 288–294.
[24] M.B. Kasiri, H. Aleboyeh, A. Aleboyeh, Environ. Sci. Technol. 42 (2008) 7970–7975.
[25] Y.M. Slokar, J. Zupan, A.M.L. Marechal, Dyes Pigments 42 (1999) 123–135.
[26] G.D. Garson, AI Expert 6 (1991) 47–51.
[27] A.R. Khataee, O. Mirzajani, Desalination 251 (2010) 64–66.
[28] A.R. Khataee, V. Vatanpour, A.R. Amani, J. Hazard. Mater. 161 (2009) 1225–1233.
[29] N. Daneshvar, D. Salari, A.R. Khataee, J. Photochem. Photobiol. A 157 (2003) 111–116.
[30] N. Daneshvar, D. Salari, A.R. Khataee, J. Photochem. Photobiol. A 162 (2004) 317–322.
[31] N. Daneshvar, A. Aleboyeh, A.R. Khataee, Chemosphere 59 (2005) 761–767.
[32] A.R. Khataee, M.B. Kasiri, J. Mol. Catal. A: Chem. 328 (2010) 8–26.
[33] U. Diebold, Surf. Sci. Rep. 48 (2003) 53–229.
[34] A.R. Khataee, H. Aleboyeh, A. Aleboyeh, J. Exp. Nanosci. 4 (2009) 121–137.
[35] A.R. Khataee, M.N. Pons, O. Zahraa, J. Hazard. Mater. 168 (2009) 451–457.
[36] T. Yates Jr., T.L. Thompson, Chem. Rev. 106 (2006) 4428–4453.
[37] M.R. Hoffmann, S.T. Martin, W. Choi, D.W. Bahnemannt, Chem. Rev. 95 (1995) 69–96.
[38] N. Daneshvar, D. Salari, A. Niaei, A.R. Khataee, J. Environ. Sci. Heal. B 41 (2006) 1273–1290.
[39] D.M. Blake, P.C. Maness, Z. Huang, E.J. Wolfrum, J. Huang, Sep. Purif. Meth. 28 (1999) 1–50.
[40] K. Tennakone, K.G.U. Wijayantha, J. Photochem. Photobiol. A 113 (1998) 89–92.
[41] C.A. Emilio, J.F. Magallanes, M.I. Litter, Anal. Chim. Acta 595 (2007) 89–97.
[42] P. Calza, V.A. Sakkas, A. Villioti, C. Massolino, V. Boti, E. Pelizzetti, T. Albanis, Appl. Catal. B-Environ. 84 (2008) 379–388.
[43] F.L. Toma, S. Guessasma, D. Klein, G. Montavon, G. Bertrand, C. Coddet, J. Photochem. Photobiol. A 165 (2004) 91–96.
[44] A. Duran, J.M. Monteagudo, Water Res. 4 (2007) 690–698.
[45] J.M. Monteagudo, A. Duran, J. Guerra, F. Garcia-Pena, P. Coca, Chemosphere 71 (2008) 161–167.
[46] A. Duran, J.M. Monteagudo, I.S. Martin, F. Garcia-Pena, P. Coca, J. Hazard. Mater. 144 (2007) 132–139.
[47] A.R. Khataee, Environ. Technol. 30 (2009) 1155–1168.
[48] E. Oliveros, F. Benoit-Marquie, E. Puech-Costes, M.T. Maurette, C.A.O. Nascimento, Analusis 26 (1998) 326–332.
[49] C.A. Emilio, M.I. Litter, J.F. Magallanes, Helvetica Chim. Acta 84 (2001) 799–813.
[50] A. Aleboyeh, H. Aleboyeh, Y. Moussa, Dyes Pigments 57 (2003) 67–75.
[51] P.R. Gogate, A.B. Pandit, Adv. Environ. Res. 8 (2004) 553–597.
[52] F. Ibney Hai, K. Yamamoto, K. Fukushi, Crit. Rev. Environ. Sci. Technol. 37 (2007) 315–377.
[53] E. Brillas, I. Sires, M.A. Oturan, Chem. Rev. 19 (2009) 6570–6631.
[54] S. Gob, E. Oliveros, S.H. Bossmann, A.M. Braun, R. Guardani, C.A.O. Nascimento, Chem. Eng. Process 38 (1999) 373–382.
[55] A.R. Khataee, V. Vatanpour, M. Rastgar Farajzadeh, Turkish J. Eng. Environ. Sci. 32 (2008) 367–376.
[56] J.A. Stegemann, N.R. Buenfeld, J. Hazard. Mater. 90 (2002) 169–188.
[57] E.S. Elmolla, M. Chaudhuri, M.M. Eltoukhy, J. Hazard. Mater. 179 (2010) 127–134.
[58] R.F. Yu, H.W. Chen, K.Y. Liu, W.P. Cheng, P.H. Hsieh, J. Chem. Technol. Biotechnol. 85 (2010) 267–278.
[59] O. Legrini, E. Oliveros, A.M. Braun, Chem. Rev. 93 (1993) 671–698.
[60] F.A.P. Costa, E.M. dos Reis, J.C.R. Azevedo, J. Nozaki. Sol. Energy 77 (2004) 29–35.
[61] G.M. Collona, T. Caronna, B. Marcandalli, Dyes Pigments 41 (1999) 211–220.
[62] N.H. Ince, Water Res. 33 (1999) 1080–1084.
[63] Y.S. Shen, D.K. Wang, J. Hazard. Mater. 89 (2002) 267–277.
[64] N. Daneshvar, A.R. Khataee, J. Environ. Sci. Health. A 41 (2006) 1–14.
[65] A. Durán, J.M. Monteagudo, I. San Martin, R. Sánchez–Romero, J. Hazard. Mater. 167 (2009) 885–891.
[66] A. Durán, J.M. Monteagudo, M. Mohedano, Appl. Catal. B-Environ. 65 (2006) 127–134.
[67] J.M. Monteagudo, A. Durán, I. San Martin, M. Aguirre, Appl. Catal. B-Environ. 95 (2010) 120–129.
[68] J.A. Giroto, R. Guardani, A.C.S.C. Teixeira, C.A.O. Nascimento, Chem. Eng. Process 45 (2006) 523–532.
[69] K.R.B. Nogueira, C.S.C. Teixeira, C.A.O. Nascimento, R. Guardani, Braz. J. Chem. Eng. 25 (2008) 671–682.
[70] M.B. Kasiri, H. Aleboyeh, A. Aleboyeh, Appl. Catal. B-Environ. 84 (2008) 9–15.
[71] H. Kusic, N. Koprivanac, L. Selanec, Chemosphere 65 (2006) 65–73.
[72] T. Kawai, K. Tsutsumi, Colloid. Polym. Sci. 273 (1995) 787–792.
[73] H. Kusic, A.L. Bozic, M. Koprivanac, S. Papic, Dyes Pigments 74 (2007) 388–395.
[74] E. Brillas, J.C. Calpe, J. Casado, Water Res. 34 (2000) 2253–2262.
[75] A. Durán, J.M. Monteagudo, E. Amores, Appl. Catal. B-Environ. 80 (2008) 42–50.
[76] J.M. Monteagudo, A. Durán, C. Lopez-Almodovar, Appl. Catal. B-Environ. 83 (2008) 46–55.
[77] J.M. Monteagudo, A. Durán, I. San Martin, M. Aguirre, Appl. Catal. B-Environ. 89 (2009) 510–518.

[78] Degussa Corporation, Environmental Uses of Hydrogen peroxide (H_2O_2), Allendale, NJ, 2001.
[79] R. Venkatadri, R.W. Peters, Hazard. Waste Hazard. Mater. 10 (1993) 107–149.
[80] T. Rigg, W. Taylor, J. Weiss, J. Chem. Phys. 22 (1954) 575–577.
[81] P.R. Gogate, A.B. Pandit, Adv. Environ. Res. 8 (2004) 553–597.
[82] R. Munter, Proc. Estonian Acad. Sci. Chem. 50 (2001) 59–80.
[83] M.B. Kasiri, H. Aleboyeh, A. Aleboyeh, Environ. Technol. 31 (2010) 165–173.
[84] A. Mohey, J.A. Libra, U. Wiesmann, Chemosphere 52 (2003) 1069–1077.
[85] H.Y. Shu, M.C. Chang, H.J. Fan, J. Hazard. Mater. 113 (2004) 201–208.
[86] P.R. Gogate, A.B. Pandit, Adv. Environ. Res. 8 (2004) 501–551.
[87] A. Niaie, J. Towfighi, A. Khataee, K. Rostamizadeh, Pet. Sci. Technol. 25 (2007) 967–982.
[88] A.R. Khataee, Pol. Chem. Technol. 11 (2009) 38–45.
[89] D. Salari, N. Daneshvar, F. Aghazadeh, A.R. Khataee, J. Hazard. Mater. 125 (2005) 205–210.
[90] O.L.C. Guimaraes, M.B. Silva, Chem. Eng. Process 46 (2007) 45–51.
[91] O.L.C. Guimaraes, M.H. dos Reis Chagas, D.N.V. Filho, A.F. Siqueira, H.J.I. Filho, H.O.Q. de, M.B. Aquino, Silva, Chem. Eng. J. 140 (2008) 71–76.
[92] O.L.C. Guimaraes, D.N.V. Filho, A.F. Siqueira, H.J.I. Filho, M.B. Silva, Chem. Eng. J. 141 (2008) 35–41.
[93] K. Rajeshwar, J.G. Ibanez, Environmental Electrochemistry Fundamentals and Applications in Pollution Abatement, Academic Press, San Diego, CA, 1997, pp. 362–363.
[94] A.C. Basha, P.A. Soloman, M. Velan, L.R. Miranda, N. Balasubramanian, R. Siva, J. Hazard. Mater. 176 (2010) 154–164.
[95] N. Daneshvar, A.R. Khataee, N. Djafarzadeh, J. Hazard. Mater. 137 (2006) 1788–1795.
[96] N. Daneshvar, H. Ashassi Sorkhabi, M.B. Kasiri, J. Hazard. Mater. 112 (2004) 55–62.
[97] M. Yousuf, A. Mollah, R. Schennach, J.R. Parga, D.L. Cocke, J. Hazard. Mater. 84 (2001) 29–41.
[98] G. Chen, Sep. Purif. Technol. 38 (2004) 11–41.
[99] S. Aber, A.R. Amani–Ghadim, V. Mirzajani, J. Hazard. Mater. 171 (2009) 484–490.
[100] A. Arul Murugan, T. Ramamurthy, B. Subramanian, C.S. Kannan, M. Ganesan, Int. J. Chem. Reactor Eng. 7 (2009) A83.
[101] B. Marselli, J. Garcia–Gomez, P.A. Michaud, M.A. Rodrigo, C. Comninellis, J. Electrochem. Soc. 50 (2003) D79–D83.
[102] P. Drogui, S. Elmaleh, M. Rumeau, C. Bernard, A. Rambaud, Water Res. 35 (2001) 3235–3241.
[103] M.A. Oturan, N. Oturan, C. Lahitte, S. Trevin, J. Electroanal. Chem. 507 (2001) 96–102.
[104] B. G()zmen, M.A. Oturan, N. Oturan, O. Erbatur, Environ. Sci. Technol. 37 (2003) 3716–3723.
[105] K. Hanna, S. Chiron, M.A. Oturan, Water Res. 39 (2005) 2763–2773.
[106] S. Irmak, S.I. Yavuz, O. Erbatur, Appl. Catal. B-Environ. 63 (2006) 243–248.
[107] M. Diagne, N. Oturan, M.A. Oturan, Chemosphere 66 (2007) 841–848.
[108] E. Brillas, B. Boye, I. Sires, J.A. Garrido, R.M. Rodriguez, C. Arias, P.L. Cabot, C. Comninellis, Electrochim. Acta 49 (2004) 4487–4496.
[109] E. Brillas, R. Sauleda, J. Casado, J. Electrochem. Soc. 145 (1998) 759–765.
[110] C.A. Martínez-Huitle, E. Brillas, Appl. Catal. B-Environ. 87 (2009) 105–145.
[111] C.G. Piuleac, M.A. Rodrigo, P. Canizares, S. Curteanu, C. Saez, Environ. Model. Soft. 25 (2010) 74–81.
[112] M. Sadrzadeh, T. Mohammadi, J. Ivakpour, N. Kasiri, Chem. Eng. J. 144 (2008) 431–441.
[113] H.R. Maier, N. Morgan, C.W.K. Chow, Environ. Model. Soft. 19 (2004) 485–494.