Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2017.1705

58 (2017) 202–213
January

Artificial neural networks for modeling and optimization of phenol and nitrophenols adsorption onto natural activated carbon

Y. El Hamzaoui[a], M. Abatal[b], A. Bassam[c,*], F. Anguebes-Franseschi[d], O. Oubram[e], I. Castaneda Robles[f], O. May Tzuc[c]

[a]*Instituto Tecnológico de Tijuana, Blvd. Industrial y Ave. ITR Tijuana S/N, Mesa de Otay, Tijuana B.C. 22500, Mexico, Email: youness@tectijuana.edu.mx*
[b]*Facultad de Ingeniería, Universidad Autónoma del Carmen, 24180, Ciudad del Carmen, Campeche, México, Email: moabatal@gmail.com*
[c]*Facultad de Ingeniera, Universidad Autónoma de Yucatán, Av. Industrias no Contaminantes por Periférico Norte, Apdo. Postal 150 Cordemex. Mérida, Yucatán, México, Tel. (+52) (999) 930-05-50. Ext. 1053; Fax: (+52) (999) 930-05-59; email: baali@correo.uady.mx, maytzuc@gmail.com*
[d]*Facultad de Química, Universidad Autónoma del Carmen. Calle 56 No. 4 Esq. Av. Concordia, Col. Benito Juárez CP 24180, Ciudad del Carmen, Campeche, México, Email: fanguebes@pampano.unacar.mx*
[e]*Facultad de Ciencias Químicas e Ingeniería, Universidad Autónoma del Estado de Morelos, Av.Universidad 1001, Col. Chamilpa, 62209, Cuernavaca, Morelos, México, Email: oumbra@uaem.mx*
[f]*Instituto Tecnológico Superior de Jerez, Libramiento Fresnillo-Tepetongo, Fracc. Los Cardos, CP 99863 Jerez de García Salinas, Zacatecas, México, Email: icastaneda@tecjerez.edu.mx*

Received 7 December 2015; Accepted 7 June 2016

ABSTRACT

An artificial neural network (ANN) approach was developed to predict the adsorption efficiency (*W*%) of phenol and nitrophenols onto activated carbon. We have studied the backpropagation of a three-layer feedforward network with Levenberg Marquardt, which describes the relationship between the adsorption efficiency as output and the operation conditions as contaminants (Phenol, Nitrophenols), initial contaminant concentration (C_i), pH and contact time. This model has been validated comparing it with both experimental measurement and simulated analysis and showed high agreement with very low percentage of error (0.5%) and high Pearson correlation (R^2 = 0.9868). The sensitivity analysis has also shown that the contact time was the most important influential parameter in this process. Based on the sensitivity analysis and neural networks model, we have developed an optimization algorithm (ANNi) for the calculation of the contact time into adsorption process when the initial conditions are well known and adsorption efficiency is required. ANNi could perform assessment with a minimal error. This technique is a very promising tool for modeling and optimization of the adsorption onto activated carbon process minimizing time and operation cost.

Keywords: Activated carbon; Phenols adsorption; Neural networks modeling; Sensitivity analysis; Water treatment

1. Introduction

Phenol and phenolic compounds are considered to have toxic effects on human health even in small concentrations. These compounds are common contaminants in the effluents of plastic, leather, paint, textile and petrochemical industries. These compounds are classified as high-priority pollutants due to their carcinogenic effects on humans, and it has been discovered that they are harmful to wildlife [1]. The US Environmental Protection Agency (EPA) has established a

* Corresponding author.

regulation to lower phenol content in wastewater to less than 1.0 mg/dm^3 [2]. Numerous researchers have reported the removal of organic compounds from wastewater by destructive processes including oxidation with ozone, manganese oxides [3] and recuperative processes such as biological treatment, membrane separation and solvent extraction. Among the proposed methods, the removal of phenols by adsorption technologies is regarded as one of the best methods because adsorption does not require high operation temperature which simplifies the operation procedure. Numerous adsorbent solids have been used to remove phenolic compounds from wastewater such as activated carbon (AC) [4], silica [5], polymeric resins [6], fly ash [7] and kaolinite [8].

This work studies the use of AC for the adsorption of phenols and nitrophenols. Specifically, our goal is to model and predict the behavior of the contaminant's (phenols and nitrophenols) adsorption efficiency onto AC, finding relations between the adsorption efficiency, operation conditions and contaminants. Such model could be used to take appropriate actions in order to remove these harmful poisons from the water.

However, modeling phenol and nitrophenols adsorption onto AC is a difficult process due to the complexity of the equations that involve the radiant energy balance, the spatial distribution of the adsorbed radiation, mass transfer and the mechanisms of adsorption transport involving attractive and repulsive superficial forces of molecules, diverse mobility forces between the molecules and shape molecular effects including dissymmetry on properties of matter as evaporation, condensation and reflection. The process depends on diverse factors and phenomena, exhibiting a nonlinear behavior which is difficult to describe by linear mathematical models, such as those derived from different variants of multivariate linear regression. Based on this reason, researchers have turned to nonlinear data-driven approaches for similar problems [9,10]. This paper proposes the use of direct and inverse artificial neural networks (ANNs). The developments in ANNs in the latest years make them able to describe the complex behavior of the system. ANNs have been broadly used as powerful tools to solve nonlinear multivariate mathematical problems about the adsorption process for water treatment [9–17].

The aim of this research is to develop a mathematical model using ANNs to learn and find a relation (transfer function and correlation) between experimental variables such as contaminants, initial concentration, pH, contact time and adsorption efficiency to predict the amount of phenol and nitrophenols adsorbed onto AC from aqueous solution. An empirical equation for adsorption efficiency of these contaminants was developed using the parameters of ANNs as the weights of networks. Furthermore, within the equilibrium field, the predicted results obtained from the optimized ANNs model were compared with the experimental data through the details of the computational approach, as well as the numerical validation with the statistical analysis was fully discussed.

Once ANNs have been established, we proceeded with inverting neural network into ANNi (artificial neural network inverse) using an optimization method to find the optimum parameter value (inputs); in this case, the contact time was the input parameter, for the required output (adsorption efficiency). In this research, we have found that the ANNi coupled with Nelder–Mead simplex method of optimization play an essential role to calculate the optimal operation conditions. Hernandez et al. [18,19] used this method to calculate the optimal input parameters for the required coefficient of performance (COP) for absorption heat transformers. In particular, this work is the first to use ANNi for estimating the optimal contact time of phenol and nitrophenols adsorption onto AC, when the adsorption efficiency is required. We have presented a comprehensive numerical and statistical comparison of the results, showing that ANNs and ANNi can be used to construct useful predictive models with optimization approach of the phenol and nitrophenols adsorption process onto AC.

2. Adsorption process

Adsorption process is the adhesion of the molecules from a mixture in a gaseous or liquid state to a solid surface. This process creates a film of the adsorbate on the surface of the adsorbent. In this work, the adsorbent used was a natural AC; it has had industrial interest for many years due to its uniform pore structure and appropriate selective adsorption [20,21]. The surface properties were calculated according to the t-plot analysis [22–24] showed in Table 1. An average diameter distribution of 0.3 nm was obtained according to the method described by Horváth and Kawazoe [25]. The morphology of the AC was examined using SEM technique (Fig. 1). This figure shows that the surface of AC presents

Table 1
Surface properties of activated carbon (AC) calculated by the t-plot analysis [25,27]

BET surface area	232.40 m^2/g
Average diameter	0.3000 nm
Adsorbent density	2.2460 g/cc
Surface atom density	38.4500 (molec/cm^2) × 1e14
Langmuir surface area	1.034E + 03 m^2/g
Surface tension	8.8500 erg/cm^2
Critical pressure	33.5000 atm

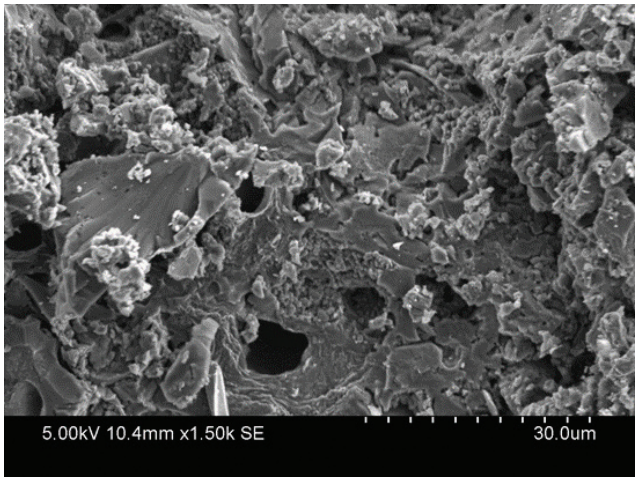

Fig. 1. SEM image of activated carbon (AC) at 1.50 K magnification.

different grain sizes and a particle diameter between 5 to 30 μm. It also shows that the external surface of AC is fairly heterogeneous with different size cavities.

We have used the breakthrough curves (*W*% vs. time) [26] in order to understand the different sides of the behavior presented by AC during adsorption process and determine the effects of experimental variables as pH and initial concentration of the contaminant upon the adsorption efficiency (*W*%).

The equation that describes the adsorption efficiency *W* (%) is given by:

$$W(\%) = \frac{C_i - C_f}{C_i} \times 100 \quad (1)$$

where C_i is the initial concentration of contaminant at the adsorbent, and C_f the final concentration of contaminant in the adsorbent at a determinate time.

2.1. Experimental setup

Batch mode experiments were conducted as follows: 0.1 g of AC was added to 10 mL of phenol (or 2-nitrophenol, 3-nitrophenol, 4-nitrophenol) at pH range 2.0–10.0. The mixtures were placed in centrifuge tubes and shaken in a rotary shaker for 5, 15, 30, 60, 120, 180, 240 and 360. After each specific contact time, the tubes were centrifuged at 3,500 rpm for 2 min to provide the separation between solid and liquid. The concentration of phenol, 2-nitrophenol, 3-ntrophenol and 4-nitrophenol was determined using ultraviolet-visible (UV-Vis) spectrophotometer (Evolution 220) at the maximum absorbance wavelengths 269 nm, 278 nm, 273 nm and 316 nm, respectively. In order to ensure the truthfulness of experiment results, all experiments were duplicated.

2.2. pH effect

The removal of pollutants from wastewater was affected significantly by the pH of the phase from which the removal occurs. Fig. 2 shows the effect of pH on the adsorption of phenol, 2-nitrophenol, 3-nitrophenol and 4-nitrophenol. The adsorption of phenolic compounds by AC diminishes with increasing pH values. The last correlation is explained knowing that at low pH ranges, chemisorption dominates in this range and chemisorption along with physisorption occurs at higher pH ranges. The pH of the solution affects the surface charge of the adsorbent degree of ionization and speciation of the adsorbate species, which might lead to change in kinetics and equilibrium characteristics of the adsorption process.

2.3. Contact time effect

The uptake values of phenol, 2-nitrophenol, 3-nitrophenol and 4-nitrophenol at pH = 6 from the solution as a function of contact time are presented in Fig. 3. The AC adsorption kinetics indicates that the time required to reach the adsorption equilibrium was approximately between 1 h and 3 h. This behavior could be explained by the properties of the phenolic compounds; the larger p-nitrophenol compound needed more time to be adsorbed onto the surface of the AC.

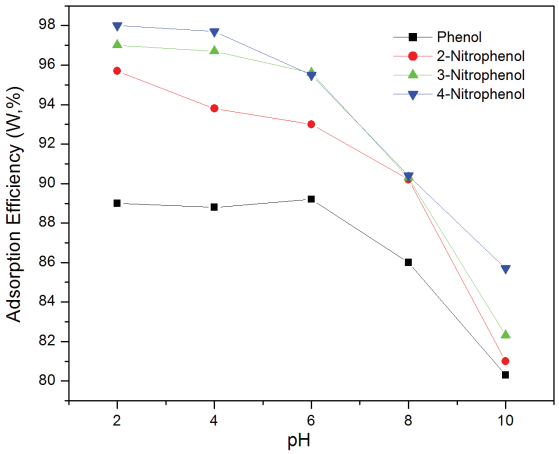

Fig. 2. Effect pH on Phenols uptake by AC.

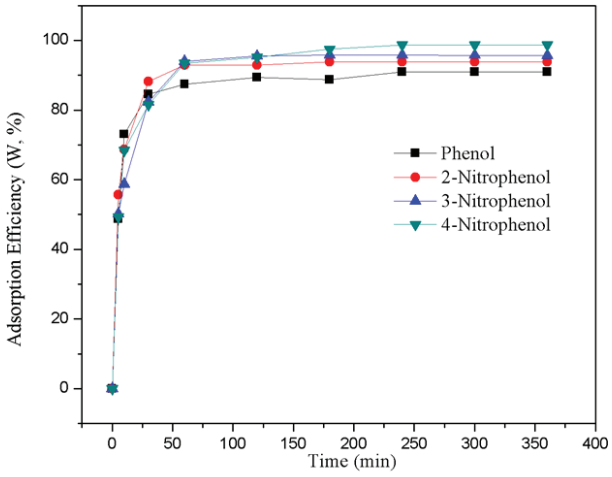

Fig. 3. Effect of contact time on the removal of phenol, 2-nitrophenol, 3-nitrophenol and p-nitrophenol on AC at pH = 6.

The difference between phenol and p-nitrophenol adsorption was most likely due to the lower solubility of 2-nitrophenol, 3-nitrophenol and 4-nitrophenol than phenol in aqueous solutions. A decrease in solubility was associated with an increase in adsorption capacity. Low solubility implies that there are weak forces between solvent and adsorbent molecules. As a result, a high amount of uptake occurred. Kumar and his colleagues have attributed the higher adsorption of p-nitrophenol than phenol onto AC to the difference in their chemical structures and to the positions of the functional groups [27].

3. ANN approach

An ANN is a structure confirmed by a number of interconnected unities, called neurons, which operate in parallel and present a natural tendency for learning from experimental data; for this reason, ANN can be used in several engineering applications [28,29]. The main objective of neural network architecture is to mimic the synapsis generated at biological nervous systems.

The ANN automates the process of model building and interpretation and enables us to obtain answers from a database. However, ANN shall be seen like a black box in which we introduce database as input variables. Each input is assigned with an appropriate weighting factor (w). The sum of the weighted inputs and the bias (b) produces the input for a transfer function which will generate an output value (Fig. 4). The main characteristic of this artificial intelligent model is that it does not require specific information about the physical behavior of the system or the way in which the data were obtained [30].

ANN can be trained to solve multivariable problems with nonlinear equations. The training process is accomplished by specific algorithms; the most broadly used of the algorithms is known as backpropagation. The architecture of an ANN is usually divided into three parts: an input layer, a hidden layer(s) and an output layer, where each one uses a transfer function [30]. At the training, the network learns from its errors until it obtains a model that describes the phenomenon the most accurately possible. During the training, weight and bias matrixes are generated; these are modified after each iteration until the ANN obtains the optimal values. In this work, the transfer functions used are the Tangent-Sigmoid function (Tansig, Eq. (2)) and the Linear function (Pureline, Eq. (3)):

$$\text{Tansig(n)} = \frac{2}{1+\exp(-2n)} - 1 \tag{2}$$

$$\text{Pureline}\ (n) = n \tag{3}$$

where n represents the weighted sum of the input values.

3.1. Numerical methodology

A numerical computational methodology was used to develop the sets of calculations about the adsorption efficiency process (Fig. 5). The methodology consists of three stages: (i) creation of a working database with experimental data collected during the adsorption process; (ii) development and evaluation of ANN models for the reliable estimation of adsorption efficiency W (%) on AC; and (iii) determination of W (%) using the optimized ANN architecture developed in this research and comparative statistical analysis between experimental data and estimated W (%) (inferred from the application of the ANNs tools).

(i) Experimental data set obtained during the phenols adsorption process: It is an experimental database, confirmed by 975 samples provided by Abatal et al. [31]; it consists of diverse contaminant adsorption values, calculated from the Eq. (1). The experimental data set is obtained considering diverse process parameters, including the type of contaminant present in the suspension (phenol, 2-nitrophenol, 3-nitrophenol and 4-nitrophenol), the initial contaminant concentration at the adsorption process (20, 40, 60, 80 and 100 ppm), the pH range (2, 4, 6, 8 and 10) and the contact time (0–360 min). Table 2 synthesizes the present data set, which is enough for training, testing and validation of the ANN model.

The data set created was randomly divided into two parts: 75% was destined to the ANN training process, and the remaining 25% was used for testing and validation phases, in order to obtain an accurate representation of the data distribution. The transfer function used in the hidden layer is sigmoid (Eq. 2); therefore, all samples should be normalized in the range of 0–1 [32,33]. Therefore, all the input data set Xi (from the training, validation and test sets) were scaled to a new normalized value $x_{i,N}$ as follows:

$$x_{i,N} = \frac{X_i}{1.1 \times X_{max}} \tag{4}$$

(ii) Development and evaluation of ANN models: Various ANN architectures were evaluated to obtain the model that provides a reliable estimate of W (%) of natural AC (Fig. 6). The ANN's model used was a network characterized by an input layer with four variables (contaminant, initial concentration [C_i], pH and adsorption time [t]), one hidden layer, and an output layer with a single variable W [%]). The assessment of multiple architectures of neurons with transfer functions is proposed as a suitable computational strategy to find out the relations between input and output variables [34].

(iii) Statistical analysis: The numerical data obtained by the several ANN models were statistically compared with experimental data to demonstrate the estimation accuracy. The analysis was performed by applying the

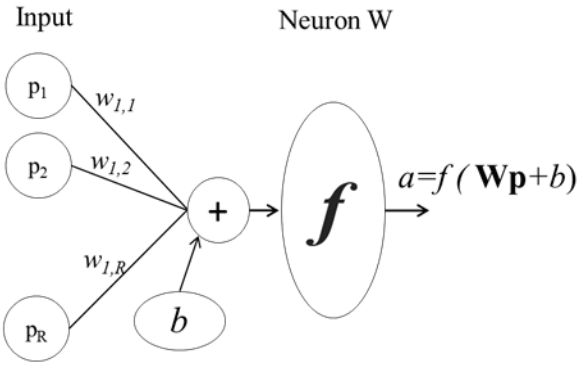

Fig. 4. A typical elementary network with R inputs.

Table 2
Characteristics of input and output variables about the ANN's model

Parameters	Min.	Max.	Units
Inputs parameters:			
Contaminant	1	4	–
Initial concentration (Ci)	20	100	ppm
pH (pH)	2	10	–
Time (t)	0	360	min
Outputs parameters:			
Adsorption efficiency (W)	0	100	%

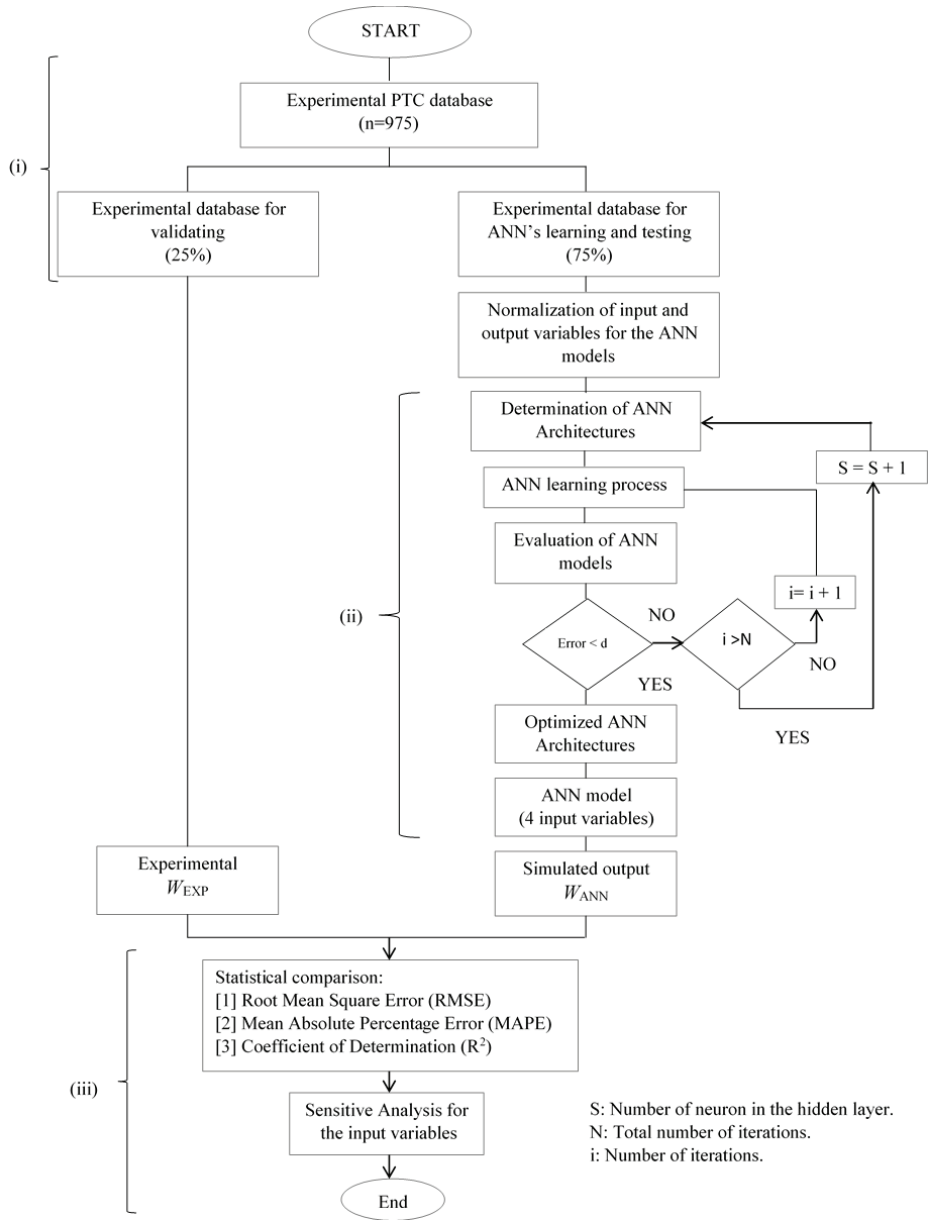

Fig. 5. Computational methodology.

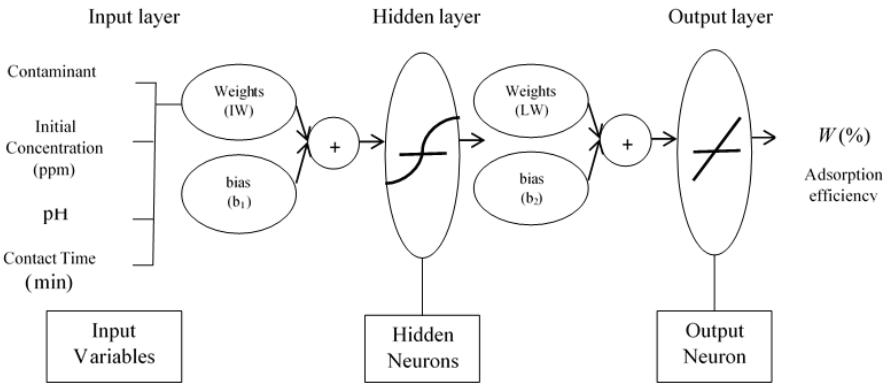

Fig. 6. Artificial neural network architecture for adsorption process.

statistical methods as root mean square error (RMSE, used to compute the differences between estimated and observed values [35]), mean absolute percentage error (MAPE, which is the computed average of errors (%) by estimating the predictions of a variable [26] and the correlation coefficient (R^2, which illustrates the strength of correlation of variability in a data set, generally between 0 and 1), given by the following equations:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(W_{\mathrm{Exp}(i)} - W_{\mathrm{Sim}(i)})^2}{n}} \quad (5)$$

$$\mathrm{MAPE} = \frac{\sum_{i=1}^{n}\left|\frac{W_{\mathrm{Exp}(i)} - W_{\mathrm{Sim}(i)}}{W_{\mathrm{Exp}(i)}}\right|}{n} \times 100(\%) \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(W_{\mathrm{Exp}(i)} - W_{\mathrm{Sim}(i)})^2}{\sum_{i=1}^{n}(W_{\mathrm{Exp}(i)} - \overline{W}_{\mathrm{Exp}})^2} \quad (7)$$

where $\overline{W} = \frac{1}{n}\sum_{i=1}^{n} W_i$; $W_{\mathrm{Sim}(i)}$ is the ANN estimated value, and $W_{\mathrm{Exp}(i)}$ is the experimental value of the variable W (%).

3.2. ANN model

The ANN is commonly trained in order to predict a specific output as a set of input values. This process, denominated learning, has as purpose adjusting the bias and connection weights among neurons to minimize the error expressed by the difference between the simulated output (generated by the weight adjusting process) and the output target (given by experimental data). The algorithm used at this work to achieve the optimization process of weights and bias is the algorithm known as Levenberg-Marquardt backpropagation, which is one of the most successful algorithms in increasing the convergence speed of the ANN architecture [36]. At the same way, the RMSE was the statistical criterion used to determine the estimation accuracy of the network according to the experimental data. Fig. 7 shows the algorithm used for the network training process, which was implemented in the ANN toolbox of the mathematical software MATLAB®.

To determine the optimal ANN architecture, we test a set of several network configurations with different number of neurons in the hidden layer (from 1 to 30 neurons). The transfer functions used in each set of the network training were the *Tansig function* (hidden layer) and *Pureline function* (output layer). The simulated data obtained after the training was compared statistically with experimental data through the Eqs. (5–7) to calculate the estimated accuracy of the model; the results of these comparisons are shown in Table 3. According to Table 3, we conclude that the best ANN model was accomplished with 20 neurons in the hidden layer. Therefore, the optimal configuration of the network was about 4-20-1, and we have found it convenient. This model presents smaller RMSE (5.4464%) and MAPE (0.5438%) values. Furthermore,

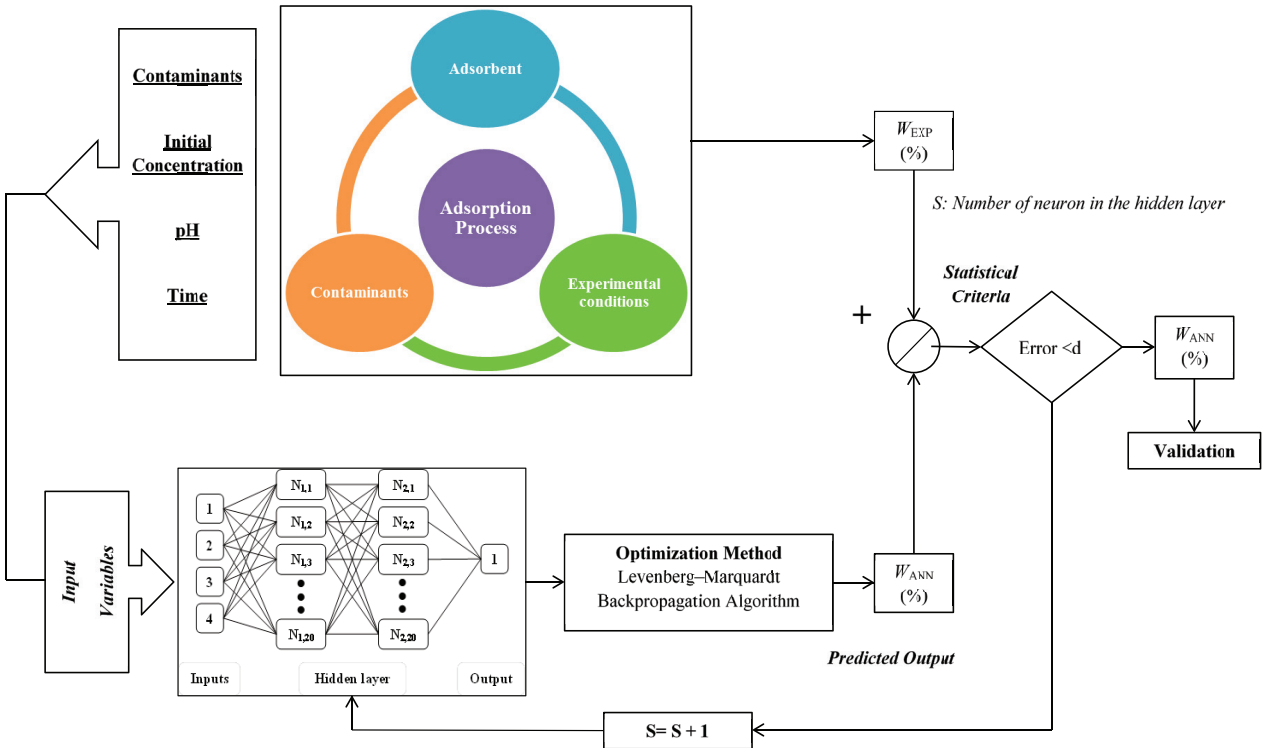

Fig. 7. Numerical procedure used in the ANN learning process, and the iterative architecture used by the ANN model to estimate adsorption percentage (S: number of the neurons in the hidden layer, and d: convergence criteria).

Table 3
Tests with different architectures of ANN

ANN architecture	Number of neurons	Epoch	RMSE (%)	MAPE (%)	R^2 (%)	Best linear equation
4-01-1	1	1,000	16.3303	4.6035	0.8739	y = 0.75x + 17.0
4-05-1	5	1,000	10.2760	1.9643	0.9521	y = 0.90x + 7.0
4-10-1	10	1,000	7.8692	0.9071	0.9722	y = 0.94x + 4.7
4-15-1	15	1,000	6.8930	0.8320	0.9787	y = 0.96x + 2.9
4-20-1	**20**	**1,000**	**5.4464**	**0.5438**	**0.9868**	**y = 0.98x + 1.8**
4-25-1	25	1,000	6.5692	0.7766	0.9807	y = 0.95x + 3.2
4-30-1	30	1,000	6.5307	0.8196	0.9809	y = 0.96x + 2.8

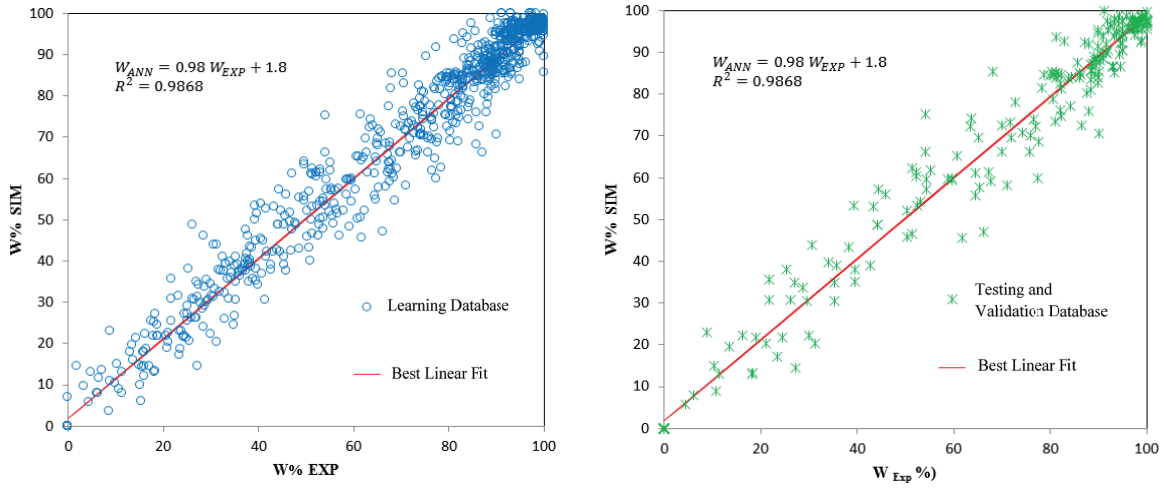

Fig. 8. Statistical comparison between simulated (ANN) and experimental *W* (%) data: (a) comparison at the learning process. (b) comparison at testing and evaluation process.

experimental (*WEXP*) and simulated data (*WANN*) were compared satisfactorily through a linear regression model, given by the following equation:

$$W_{ANN} = 0.98 \times W_{EXP} + 1.8 \quad (8)$$

with a regression coefficient $R^2 = 98.68\%$. It can be proved that the linear regression model given by ANN was subject to the hypothetical testing of confidence intervals using t-student within $\alpha = 5\%$ (level of the test). We thus see that finding the confidence intervals, respect to the slope is [–1.7523, 4.0354]. The same procedure was also held on the intercept we have [1.1470, 2.382].

Fig. 8 illustrates the comparative results between the experimental and simulated adsorption efficiency (*W* [%]) values used at the training (Fig. 8(a)) and testing (Fig. 8(b)) stages. In both figures, it can be seen that the calculated contamination adsorption efficiency showed the same behavior (R^2 = 0.9868) respect to the experimental contamination adsorption efficiency independently if it belongs to testing or training phase, indicating that samples selected for this process were representative of the phenomenon.

As pointed out before, the optimal weights and bias computed for the best ANN model are listed in Table 4, where IW represents the weights from input to hidden layer, LW the weights from hidden to output layer, *S* the total number of neurons in the hidden layer ($S = 20$), *K* the total number of neurons in the input layer ($K = 4$), and *b*1 and *b*2 the bias factors. On the basis of what has been just discussed, we shall be able to express the model with the help of the ANN's backpropagation algorithm (Fig. 6) considering the Tangent-Sigmoid transfer function (Eq. (2)), the linear transfer function (Eq. (4)) and the values in Table 4. The proposed model must take the form:

$$W_{ANN}(\%) = \sum_{j=1}^{S}\left[LW_{(1,j)}\left(\frac{2}{1+\exp\left(-2\left(\sum_{k=1}^{K}\left(IW_{(j,k)}In_{(k)}\right)+b1_{(j)}\right)\right)}-1\right)\right]+b2 \quad (9)$$

where *LW*, *IW*, *b*1, *b*2, *K* and *S* are described in Table 4, and *In* is the parameter value corresponding operation.

Table 4
Weights and bias parameters obtained for the ANN model developed

Number of neurons (s)*	Weights						
	Hidden layer (S = 20, K = 4) IW (s, k)				Output layer LW (s, l)	Bias	
	Contaminant (k = 1)	C_i (k = 2)	pH (k = 3)	t (k = 4)	W (%)	$b1$ (s)	$b2$
1	–6.99	6.17	–1.71	30.60	–12.55	19.00	–26.80
2	174.00	287.00	1.65	437	3.66	–654.0	
3	–8.35	19.90	0.74	–3.40	–15.62	7.21	
4	5.86	0.86	–2.58	–1.41	112.75	–1.31	
5	32.30	–78.50	0.85	–8.31	–14.05	–7.43	
6	–5.05	80.80	19.10	–3.50	56.76	–36.90	
7	–4.35	–3.98	–5.09	1.10	31.30	–5.94	
8	–7.83	–5.94	–3.18	1.26	40.18	–3.38	
9	–0.73	–0.25	0.66	–12.10	20.44	4.25	
10	0.29	–9.58	–0.28	22.40	–10.57	2.92	
11	32.80	–0.31	–1.33	256	17.32	–125.00	
12	–5.48	12.00	0.42	–41.90	–13.26	4.21	
13	7.07	–0.76	4.58	–1.32	–36.25	5.09	
14	–4.91	83.00	18.30	–3.25	–56.26	–38.10	
15	–77.10	–137.0	–137.0	87.40	–3.37	408.00	
16	30.90	–64.50	313.00	–17.10	–6.15	–321.00	
17	–5.00	–1.79	–2.15	6.17	24.33	19.40	
18	–0.13	–7.69	0.21	–28.00	9.62	7.06	
19	0.21	0.19	0.85	80.50	40.65	–8.80	
20	–0.29	0.03	–0.48	10.60	69.57	–0.20	

*s is the number of neurons in the hidden layer; k is the number of neurons in the input layer; l is the number of neurons in output layer (l = 1).

A validation of neural network was done to understand the behavior of the new ANN model development (Eq. (9)) in respect to experimental values. The validation of the ANN was accomplished through a comparison using data that weren't included in the training process [37]. Fig. 9 shows a comparison between the simulated and experimental breakthrough curves for the four contaminants (phenol, 2-nitrophinol, 3-nitrophenol and 4-nitrophenol) at different pH and concentration values. As can be seen, the eight graphics present a suitable reproduction of the curves demonstrating that the ANN model was able to adapt successfully to the different pH and to initial concentration values.

3.3. Sensitivity analysis

In order to determine the impact of each input variable on the W (%) simulated by the ANN model, a sensitivity analysis was developed using the Partial Derivatives (PD) method which has been described by Dimopoulos et al. [38]. The PD method depends on the IW and LW values (previously described in Table 4), the input value and the activation functions (Eqs. (2) and (3)). The relative contribution of the ANN inputs to the output data ($SSDi$) is given by:

$$\mathrm{SSD}_i = \sum_{i=1}^{N}\left[\left(f'(net_k)\sum_{j=1}^{S} LW_{1,j} f'(net_j) IW_{i,j}\right)^2\right] \quad (10)$$

where N is the total number of samples, and $f'(net_k)$ and $f'(net_j)$ are the derivatives of the activation function in the hidden layer and output layer, respectively. The SSD value allows us to classify the variables according to their increasing contribution to the output variable in the model.

The results of the sensitivity analysis done to the ANN model using the Eq. (10) are shown in Table 5. Whence, we found that the parameter that has the greatest impact in the adsorption process is the contact time (t), the second initial parameter is the pH following by the initial concentration (Ci) and finally the presence of contaminant.

4. ANNi

According to the sensitivity analysis, the contact time (t) is the most influential parameter in this process. On the other hand, in the experimental process, we need to know the contact time under different specified experimental conditions in

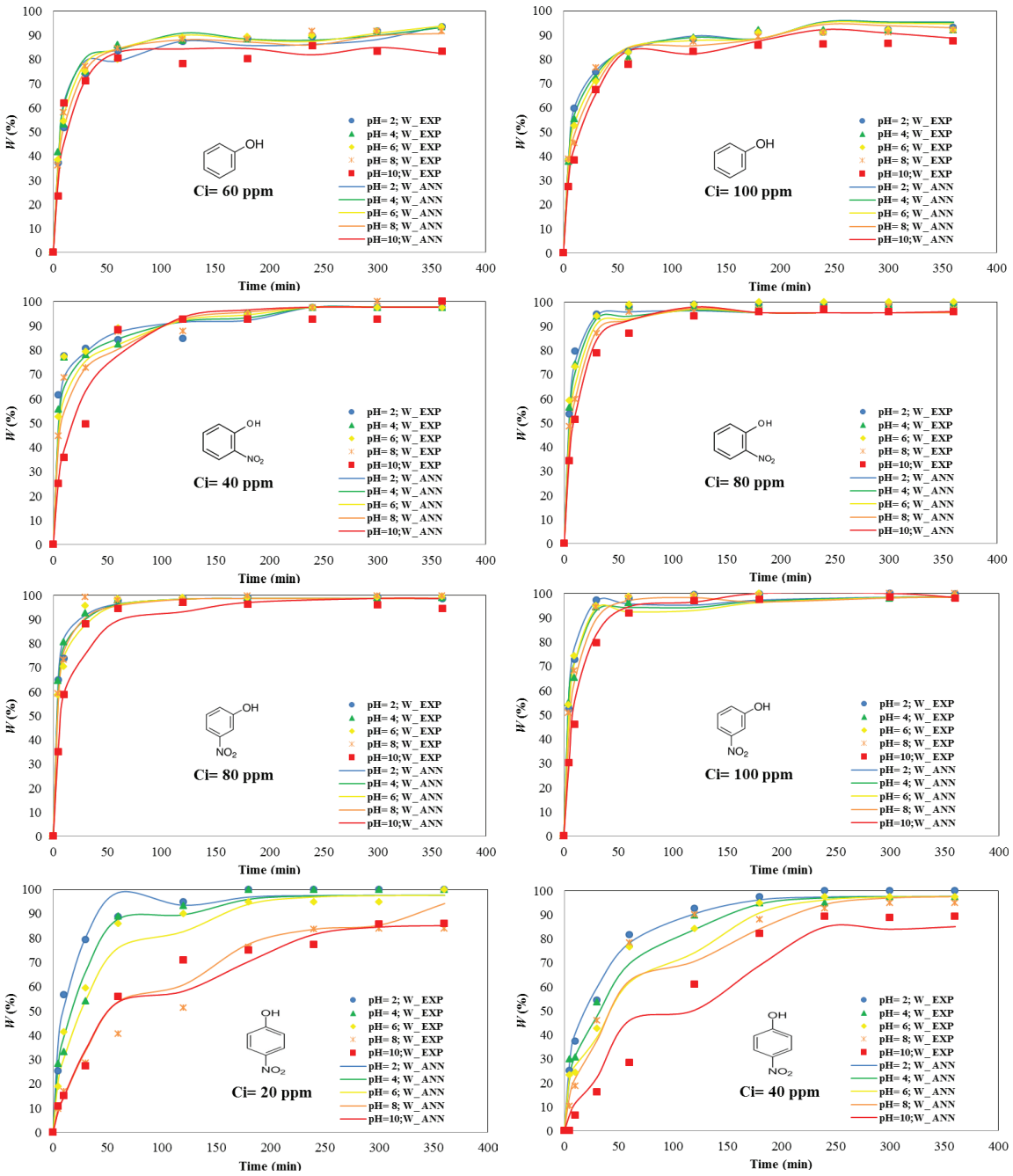

Fig. 9. Comparison between simulated and experimental curves of the adsorption process.

Table 5
Relative importance of input variables

Input variable	SSD
Contaminant (HC)	4.4445
Initial Concentration (*Ci*)	25.3926
Potential hydrogen (pH)	27.4573
Contact time (*t*)	675.0917

order to optimize the required experimental time and therefore minimize the operation cost. In this context, ANNi were applied to solve this optimization problem.

The problem of ANNi consists in estimating a value of the input variable from the required output parameter (Hernandez et al. [39]). ANNi works as follows:

$$W_{\mathrm{ANN}}(\%) = \mathrm{PURELIN}\left(\sum_{s}\left\{LW_{(1,s)}\cdot\left[\mathrm{TANSIG}\left(IW_{(s,k)}\cdot In_{(k)} + b1_{(s)}\right)\right]\right\} + b2_{(1)}\right) \tag{11}$$

$$W_{\mathrm{ANN}}(\%)=\sum_{s}\left\{LW_{(1,s)}\cdot\left[\frac{2}{1+e^{-2\left(b_{1(s)}+\sum_{k}IW_{(s,k)}\cdot In_{(k)}\right)}}-1\right]\right\}+b_{2(1)} \tag{12}$$

where IW_{i}, LW_{o}, b_1 and b_2 are the parameters of the neural networks as illustrated in Table 5. Eq. (12) can be expressed into Eq. (13). Then, we have:

$$W_{\mathrm{ANN}}(\%)=b_{2(1)}-\sum_{s}LW_{(1,s)}+\sum_{s}\left[\frac{2\cdot LW_{(1,s)}}{1+e^{-2\left(b_{1(s)}+\sum_{k}IW_{(s,k)}\cdot In_{(k)}\right)}}\right] \tag{13}$$

At this step, we have obtained the function which has to be optimized to get the optimal input parameter (s) $\ln_{(k=x)}$:

$$F\left(In_{(x)}\right)=b_{2(1)}-\sum_{s}LW_{(1,s)}+\sum_{s}\left[\frac{2\cdot LW_{(1,s)}}{1+e^{-2\left(IW_{(s,k)}\cdot In_{(x)}+\sum_{k\neq x}IW_{(s,k)}\cdot In_{(k)}+b_{1(s)}\right)}}\right] \tag{14}$$

It should be noticed that $\ln_{(x)} = t$ is the contact time for the adsorption efficiency (W%).

The Eq. (14) represents the objective function required to calculate the input value when the required output is well known. In this investigation, we are interested in the calculation of the process contact time for the required output value (W%). The general objective function highlighted in Eq. (14) was developed into the Eqs. (15–17) describing the relations between the contact time (t) that should be calculated and (W%). These relations could be illustrated into the following form:

$$\begin{aligned}F(t)=-E&+\frac{2LW_{(1,1)}}{1+e^{\left(X_1+2.53001t\right)}}+\frac{2LW_{(1,2)}}{1+e^{\left(X_2+12.29017t\right)}}+\frac{2LW_{(1,3)}}{1+e^{\left(X_3-2.8184t\right)}}\cdots\\&+\frac{2LW_{(1,4)}}{1+e^{\left(X_4+16.95003t\right)}}+\frac{2LW_{(1,5)}}{1+e^{\left(X_5-18.01032t\right)}}+\frac{2LW_{(1,6)}}{1+e^{\left(X_6+9.75632t\right)}}\cdots\\&+\frac{2LW_{(1,7)}}{1+e^{\left(X_7+19.0412t\right)}}+\frac{2LW_{(1,8)}}{1+e^{\left(X_8+17.8731t\right)}}+\frac{2LW_{(1,9)}}{1+e^{\left(X_9+15.0078t\right)}}+\cdots\\&+\frac{2LW_{(1,20)}}{1+e^{\left(X_{20}-0.74015t\right)}}\end{aligned} \tag{15}$$

$$\begin{aligned}E=W(\%)-b2_{(1)}&+LW_{(1,1)}+LW_{(1,2)}+LW_{(1,3)}+\\&LW_{(1,4)}+LW_{(1,5)}+\ldots+LW_{(1,20)}\end{aligned} \tag{16}$$

$$X_J=-2\left(IW_{(J,1)}\times In_{(1)}+IW_{(J,2)}\times In_{(2)}+IW_{(J,3)}\times In_{(3)}+b1_{(J)}\right);\quad J=1:20 \tag{17}$$

where $\ln_{(1)}$: contaminant; $\ln_{(2)}$: initial concentration (ppm); and $\ln_{(3)}$ = pH.

In order to solve these ANNi equations, we have suggested to combine the neural network model with the Nelder–Mead optimization method. We would like to stress, according to our knowledge that this approach is considered as a novelty for the adsorption process. The Nelder–Mead method is a nonlinear optimization algorithm which minimizes to zero an objective function in a multidimensional space. This algorithm is a direct search method that doesn't need using numerical or analytic gradient. This method has been described in detail by Nelder and Mead [40].

However, to validate the ANNi approach, we have randomly taken samples from the experimental database. This random sampling was accomplished matching 98% of adsorption efficiency; within this case, we shall consider the calculation under the experimental conditions (contaminant, initial concentration and pH); the contact time has been calculated and compared with experimental time reported in the laboratory.

The resulting validations and comparisons of this process are shown in Table 6, where there is an assessment of errors between experimental and estimated contact time, for phenol, 2-nitrophenol and 3-nitrophenol. The errors found were smaller (~1%) while those errors were relatively bigger according to the case of 4-nitrophenol, due to the failure of ANN model achieving enough accuracy about this case.

5. Conclusion

The adsorption efficiency of phenol and nitrophenols onto AC during the adsorption process has been modeled using ANNs at different operating conditions. A comparison of the effects of the contact time and pH for adsorption efficiency of phenol and nitrophenols was performed. ANNs were applied to express this process into the formula obtained, whence the operator could use such results without developing a real system. Using this method, it is possible to calculate the adsorption efficiency which is mathematically expressed as a function of pH, contact time and initial concentration of each contaminant. The developed ANNs model has also shown proper performance, and it converged with an accurate prediction of experimental data with RMSE, mean absolute error and a correlation coefficient of 5.4464, 0.015 and 0.9868, respectively; these results offer vital evidence for the advantages of this method. On the other hand, ANNi also succeeded to estimate the required contact time for adsorption efficiency using Nelder–Mead method with high accuracy (~5%). Thanks to this method (ANNi), it is possible to find and compute any unknown, whatever input variable in the adsorption process onto AC. It is important to point out that the elapsed time to calculate the optimum contact time is short. Therefore, it is feasible to accurately estimate more optimal parameters. This technique could be applied appropriately in control and automatization of industrial adsorption units in real time.

Therefore, it is believed that ANNs and ANNi could be used to handle many other kinds of problems related to adsorption process during the aqueous treatment. However, careful attention must be paid when using this technique because there are parameters that do not have physical

Table 6
Experimental and simulated random samples of the system

Component	Samples	Experimental conditions			Optimal time simulated by ANNi	Error $\lvert t_{exp} - t_{sim} \rvert / t_{exp}$
		C_i	pH	t_{exp} (min.)	t_{sim} (min.)	(%)
Phenol	57	40	2	180	181.5	0.83
	70	40	4	300	298.7	0.43
	130	60	6	360	361.9	0.53
	169	80	4	240	238.8	0.50
2-nitrophenol	288	20	8	240	238.4	0.67
	359	60	2	300	297.2	0.93
	390	60	8	360	361.4	0.39
	500	100	10	360	358.2	0.50
3-nitrophenol	506	20	2	120	121.1	0.92
	568	40	6	60	60.5	0.83
	622	60	8	60	59.3	1.16
	684	100	2	30	30.3	1.00
4-nitrophenol	805	40	6	360	371.7	3.25
	854	60	6	300	292.8	2.40
	893	80	4	240	247.1	2.96
	945	100	4	360	338.5	5.97

meaning and they could be attached and operated with ad hoc mathematical tricks that are not directly dictated by physical principles.

Acknowledgment

The authors express their gratitude, appreciation and acknowledgments to the financial support provided by CONACyT (CB-169133).

References

[1] N. Takahashi, T. Nakai, Y. Satoh, Y. Katoh, Variation of biodegradability of nitrogenous organic compounds by ozonation, Water Res., 28 (1994) 1563–1570.
[2] T. Min, B. Linda, C. Aicheng, Kinetics of the electrochemical oxidation of 2-nitrophenol and 4-nitrophenol studied by in situ UV spectroscopy and chemometrics, Electrochim Acta., 52 (2007) 6517–6524.
[3] L. Ukrainczyk, M.B. McBride, Oxidation of Phenol in acidic aqueous suspensions of manganese oxide, Clay Clay Miner., 40 (1992) 157.
[4] S. Nouri, F. Haghseresht, G.Q.M. Lu, Comparison of adsorption capacity of p-cresol and p-nitrophenol by activated carbon in single and double solute, Adsorption 8(3) (2003) 215–223.
[5] K. Hanna, I. Beurroies, R. Denoyerl, D. Desplantier-Giscard, A. Galarneu, F. Di Renzo, Sorption of hydrophobic molecules by organic/inorganic mesostructures, J Colloid Interface Sci., 252(2) (2002) 276–283.
[6] K. Abburi, Adsorption of phenol and p-chlorophenol from their single and bisolute aqueous solutions on Amberlite XAD-16 resin, J Hazard Mater., 105 (2003) 143–156.
[7] M. Sarkar, K.P. Acharya, B. Bhattacharya, Removal characteristics of some priority organic pollutants from water in a fixed bed fly ash column, J. Chem. Technol. Biot., 80 (2005) 1349–1355.
[8] M. Barhoumi, I. Beurroies, R. Denoyel, H. Said, K. Hanna, Co-adsorption of alkylphenols and nonionic surfactants onto kaolinite, Colloid. Surf. A., 219 (2003) 25–33.
[9] J. Díaz-Gómez, A. Parrales, A. Álvarez, S. Silva-Martínez, D. Colorado, J.A. Hérnandez, Prediction of global solar radiation by artificial neural network based on a meteorological environmental data, Desal. Wat. Treat., 12 (2015) 3210–3217.
[10] H. Heshmati, M. Torab-Mostaedi, H.G. Galini, A. Heydari, Kinetic, isotherm, and thermodynamic investigations of uranium (VI) adsorption on synthesized ion-exchange chelating resin and prediction with an artificial neural network, Desal. Wat. Treat., 4 (2015) 1076–1087.
[11] H.R. Vahidian, A.R. Soleymani, J.B. Parsa, Development of a four-layered ANN for simulation of an electrochemical water treatment process, Desal. Wat. Treat., 2 (2014) 388–398.
[12] R.H. Nia, M. Ghaedi, A.M. Ghaedi, Modeling of reactive orange 12 (RO 12) adsorption onto gold nanoparticle-activated carbon using artificial neural network optimization based on an imperialist competitive algorithm, J. Mol. Liq., 195 (2014) 219–229.
[13] H. Karim, M. Ghaedi, Application of artificial neural network and genetic algorithm to modeling and optimization of removal of methylene blue using activated carbon, J. Ind. Eng. Chem., 20 (2014) 2471–2476.
[14] E.A. Dil, M. Ghaedi, A. Ghaedi, A. Asfaram, M. Jamshidi, M.K. Purkait, Application of artificial neural network and response surface methodology for the removal of crystal violet by zinc oxide nanorods loaded on activated carbon: kinetics and equilibrium study, J. Taiwan Inst. Chem. E, 59 (2016) 210–220.
[15] M. Ghaedi, A. Daneshfar, A. Ahmadi, M.S. Momeni, Artificial neural network-genetic algorithm based optimization for the adsorption of phenol red (PR) onto gold and titanium dioxide Nanoparticles loaded on activated carbon, J. Ind. Eng. Chem., 21 (2015) 587–598.
[16] M. Maghsoudi, M. Ghaedi, A. Zinali, A.M. Ghaedi, M.H. Habibi, Artificial neural network (ANN) method for modeling of Sunset yellow dye adsorption using zinc oxide nanorods loaded on activated carbon: Kinetic and isotherm study, Spectrochim. Acta. A., 21 (2015) 587–598.

[17] M. Ghaedi, A. Ansari, F. Bahari, A.M. Ghaedi, A. Vafaei, A hybrid artificial neural network and particle swarm optimization for prediction of removal of hazardous dye Brilliant Green from aqueous solution using Zinc sulfide nanoparticle loaded on activated carbon, Spectrochim. Acta. A., 137 (2015) 1004–1015.
[18] J.A. Hernandez, A. Bassam, J. Siqueiros, D. Juarez-Romero, Optimum operating conditions for a water purification process integrated to a heat transformer with energy recycling using neural networks inverse, Renew. Energy, 34 (2009) 1084–1091.
[19] D. Colorado, J.A. Hernández, W. Rivera, H. Martínez, D. Juárez, Optimal operation conditions for a single-stage heat transformer by means of an artificial neural network inverse, Appl. Energy, 88 (2011) 1281–1290.
[20] D.M. Ruthven, Principles of Adsorption and Adsorption Processes, Wiley Interscience, 1984.
[21] J.G. Speight, The chemistry and Technology of Petroleum, 2nd ed., MarcelDekker Inc., 1991.
[22] D.D. Do, Adsorption Analysis: Equilibria and Kinetics, Imperial College Press, 1998.
[23] B.C. Lippens, J.H. de Boer, Studies on pore systems in catalysts: V. The *t* method, J. Catal., 4 (1965) 319–323.
[24] K. Nakai, J. Sonoda, S. Kondo, I. Abe, The analysis of surface and pores of activated carbons by the adsorption of various gases, Pure Appl. Chem., 65 (1993) 2181–2187.
[25] G. Horváth, K. Kawazoe, Method for the calculation of effective pore size distribution in molecular sieve carbon, J. Chem. Eng. Jpn., 16 (1983) 470–475.
[26] B. Singha, N. Bar, S.D. Das, The use of artificial neural networks (ANN) for modeling of adsorption of Cr(VI) ions, Desal. Wat. Treat., 3 (2014) 415–425.
[27] A. Kumar, S. Kumar, D.V. Gupta, Adsorption of phenol and 4-nitro phenol on granular activated carbon in basal salt medium: equilibrium and kinetics, J. Hazard. Mater., 147 (2007)155–166.
[28] S. Haykin, Neural Networks, 2nd ed., Prentice Hall, 1999.
[29] H. Demuth, M. Beale, Neural Network Toolbox for Use with MATLAB, User's Guide, version 4, The MathWorks, 2014.
[30] A. Bassam, R.A. Conde-Gutierrez, J. Castillo, G. Laredo, Direct neural network modelling for separation of linear and branched paraffins by adsorption process for gasoline octane number improvement, Fuel, 124 (2014) 158–167.
[31] M. Abatal, M.T. Olguin, Comparative adsorption behavior between phenol and p-nithophenol by Na- and HDTMA-clinoptilolite-rich tuff, Environ. Earth Sci., 69 (2013) 2691–2698.
[32] Y. Hamzaoui, J.A. Hernández, S. Silva-Martínez, A. Bassam, A. Álvarez, C. Lizama-Bahena, Optimal performance of COD removal during aqueous treatment of alazine and gesaprim commercial herbicides by direct and inverse neural network, Desalination, 277 (2011) 325–337.
[33] A. Bassam, I. Salgado-Tránsito, I. Oller, E. Santoyo, E.A. Jiménez, J.A. Hernandez, Optimal performance assessment for a photo-Fenton degradation pilot plant driven by solar energy using artificial neural networks, Int. J. Energy Res., 36 (2012) 1314–1324.
[34] R. Mohammadi, H. Eskandarloo, M. Mohammadi, Application of artificial neural network (ANN) for modelling of dyes decolorization by Sn/Zn-TiO_2 nanoparticles, Desal. Wat. Treat., 7 (2015) 1922–1933.
[35] D. Gnanasangeetha, D. SaralaThambavani, Modelling of As^{3+} adsorption from aqueous solution using *Azadirachta indica* by artificial neural network, Desal. Wat. Treat., 7 (2015) 1839–1854.
[36] M.T. Hagan, M. Menhaj, Training feedforward networks with the Marquardt algorithm, IEEE Trans. Neural Netw., 5 (1994) 989–993.
[37] P. Isasi Viñuela, I. Galván León, Redes Neuronales Artificiales: un Enfoque Práctico, Pearson Prentice Hall, 2003.
[38] I. Dimopoulos, P. Bourret, S. Lek, Use of some sensitivity criteria for choosing networks with good generalization ability, Neural Process. Lett., 2 (1995) 1–4.
[39] J.A. Hernández, D. Colorado, O. Cortés-Aburto, Y. El Hamzaoui, V. Velazquez, B. Alonso, Inverse neural network for optimal performance in polygeneration systems, Appl. Therm. Eng., 50 (2013) 1399–1406.
[40] J.A. Nelder, R.A. Mead, Simplex method for function minimization, Comput. J., 7 (1965) 308–313.