Combination of Anns and Heuristic Algorithms in Modelling and Optimizing of Fenton Processes for Industrial Wastewater Treatment

Hüseyin Cüce
Giresun University: Giresun Universitesi
Ozge Cagcag Yolcu
Giresun University: Giresun Universitesi
Fulya AYDIN TEMEL (fulya.temel@giresun.edu.tr)
Giresun University: Giresun Universitesi https://orcid.org/0000-0001-8042-9998

Research Article

Keywords: Fenton process, Multilayer Perceptron, Single Multiplicative Neuron Model, Particle Swarm Optimization, Genetic Algorithm, RSM

Posted Date: May 7th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-382899/v1

Combination of ANNs and heuristic algorithms in modelling and optimizing of Fenton processes for industrial wastewater treatment

Hüseyin CÜCE[1], Ozge Cagcag Yolcu[2], Fulya AYDIN TEMEL[3,*]

[1]Department of Geomatic Engineering, Faculty of Engineering, Giresun University, Giresun, Turkey
[2]Department of Industrial Engineering, Faculty of Engineering, Giresun University, Giresun, 28200, Turkey
[3]Department of Environmental Engineering, Faculty of Engineering, Giresun University, Giresun, Turkey

*Corresponding author:
E-Mail: fulya.temel@giresun.edu.tr

ABSTRACT

In this study, it is aimed to evaluate COD removal performance of Classical-Fenton and Photo-Fenton Processes from cosmetic wastewater by different prediction models. Besides Response Surface Methodology (RSM), three neural networks were used to more reliably and effectively predict the behavior of dependent variable at different values of relevant parameters. These neural networks; multi-layer perceptron trained by Levenberg-Marquardt (MLP-LM); multi-layer perceptron and single multiplicative neuron model trained by particle swarm optimization algorithm (MLP-PSO; SMN-PSO). H_2O_2 doses, Fe(II) doses, and H_2O_2/Fe(II) rates were independent variables of prediction models to optimize both processes in batch reactors. The generated predictions for whole data set were compared with each other. The prediction performances of models were evaluated by RMSE and MAPE error criteria. Regression analysis was also applied to determine the performance of the best model. The results obtained from all prediction tools showed that the model produces the best predictive results in almost all cases is SMN-PSO model in terms of both criteria. In addition, the genetic algorithm was utilized for SMN-PSO model results to find the optimum values of the study. Thus, without the need to perform many different experiments, the optimum parameter values can be determined to get maximum removal ratios.

Keywords: Fenton process, Multilayer Perceptron, Single Multiplicative Neuron Model, Particle Swarm Optimization, Genetic Algorithm, RSM.

1. Introduction

One of the significant sources of environmental pollution is industrial wastewaters. The consumption of freshwater used during production is also increasing with industrial developments (Friha et al. 2014). The production of wastewater in the cosmetic industry is excessively important such as in other industries. In addition to the wastewater generated during the production of products, large amounts of wastewater are also released during the cleaning process of reactors, pipes, filling lines, and other mechanical equipment used in the

production of products (de Andrade et al. 2020). The cosmetic industry can be divided into five groups: haircare, make-up, skincare, fragrances, and others (Melo et al. 2013). The wastewaters from cosmetic industries contain a high concentration of low biodegradability organic compounds, suspended solids, fats and oils, natural oils, surfactants, solvents, cosmetic ingredients, fragrances, colorants, dyes, bleaches, sunscreen agents, polymers, exfoliators, emollients, chelating agents, UV filters, antioxidants, pH adjusters, and several chemical compounds (Bautista et al. 2007; Banerjee et al. 2016; Bom et al. 2019). The most frequently detected compounds are the recalcitrant organic matter among these pollutants in water sources, and other major pollutants, non-ionic and anionic surfactants have been determined in river water, groundwater, etc, and in their sediments and biota samples in the world (Banerjee et al. 2016). For this reason, the toxicity of cosmetic wastewater should be eliminated or reduced to acceptable limits require before discharging to receiving environments.

Chemical and biological treatment technologies have been applied to treat cosmetic wastewater. The most investigated technologies have also been coagulation/flocculation (El-Gohary et al. 2010), dissolved ozone flotation (Wiliński et al. 2017), electro-coagulation, membrane systems (Monsalvo et al. 2014), submerged membrane bioreactor (Friha et al. 2014), up-flow anaerobic sludge blanket reactor (Puyol et al. 2011), advanced oxidation processes (AOPs) such as Fenton (Naumczyk et al. 2014) or photo-Fenton processes (Muszyński et al. 2019), catalytic wet peroxide oxidation (Bautista et al. 2010), etc. Compared to other treatment methods, AOPs are known as expensive methods due to their high energy and chemical requirements (Oller et al. 2011). The common feature of AOPs applied with different operating conditions is the formation of hydroxyl radicals at normal pressure and room temperature. The hydroxyl radical is a non-selective and strong oxidant that reacts with three different mechanisms. These mechanisms are hydrogen abstraction, radical addition or electron transfer. Moreover, AOPs are always improving with new equipment as well as the application of the most efficient methods (Paździor et al. 2019). AOPs can be grouped as Fenton-based, Ozone-based, Photo-catalytic, EAOPs, and others. Fenton-based processes among AOPs are Classical Fenton, Electro-Fenton, Photo-Fenton, Sono-Fenton, Photo-Fenton/TiO_2, Photo-Sono-Fenton, Photo-Electro-Fenton, Sono-Electro-Fenton. Fenton processes are preferred in the treatment of various wastewaters because of their flexible operation, easy system, the ability to react at wide temperature ranges and under atmospheric pressure (Fernandes et al. 2018).

The efficient use of treatment methods depends on determining the optimum operating conditions. A limited number of trials are carried out with the experiment sets created for this purpose and the optimum conditions are tried to be determined according to the data obtained here. Recently, Artificial Neural Networks (ANNs) and Response Surface Methodology (RSM) have come forward as effective experimental modelling and optimization methods, especially for nonlinear systems. In literature, some approaches have been asserted to model Fenton-based processes and to make predictions about these processes under certain operational parameters. While most of these approaches include statistical-based models like RSM, machine learning-based models such as ANNs have been started to apply as an alternative modelling and prediction tool. Although ANNs are the subject of many studies for time series prediction (Egrioglu et al. 2013; Cagcag Yolcu et al. 2018; Yolcu et al. 2019),

there are limited studies in the literature that use different ANN types to model Fenton-based processes (Elmolla et al. 2010; Zarei et al. 2010; Jaafarzadeh et al. 2012; Sabour and Amiri 2017; Radwan et al. 2018; Baştürk and Alver 2019; Talwar et al. 2019; Tolba et al. 2019; Gholizadeh et al. 2021). Modelling a dependent variable through certain independent variables is essentially done to predict this dependent variable at different and especially non-existent independent variable values. RSM is known as a traditional modelling and prediction tool in literature. But, this traditional methodology has some limitations such as linear model and distribution assumptions. The fundamental principle of linear modelling is that if the relationships among variables are not linear, it gives a lower performance. This situation is generally encountered in environmental sciences and ecological researches. Therefore, some variables need to be converted (Lek et al. 1996). As the computer systems, ANNs are improved to learn the data and to generate new information in an analogy to the human brain. By means of its hidden layer, ANN learns the data structure and presents the suitable models. ANNs are preferred to traditional methodology when programming is impossible or exceedingly difficult.

The main goal of modelling in the present study is to predict the behavior of the dependent variable in different values of the relevant parameters more reliably and effectively. For this purpose, it is aimed to investigate and predict COD removal performance of Classical and Photo-Fenton Processes from cosmetic wastewater by four different prediction models. The first of the applied models is the traditional model, RSM. The second is multi-layer perception (MLP) trained with the Levenberg-Marquardt training algorithm (MLP-LM). And, MLP and Single Multiplicative Neuron Network trained by particle swarm optimization algorithm (MLP-PSO; SMN-PSO) are also the distinctive models and innovative aspect of this study.

In this study, the effects of H_2O_2 doses, Fe(II) doses, and H_2O_2/Fe(II) rates were determined as independent variables to optimize both Fenton processes in a batch reactor. The scopes of the study were as follows: (*i*) to demonstrate the generalized abilities of models via training, validation, and test; (*ii*) to evaluate the model performances by RMSE and MAPE criteria; (*iii*) to do a comprehensive comparison of model results, (*iv*) to determine the performance of the best model by regression analysis; (*v*) to obtain optimal values of the best model by genetic algorithm; (*vi*) to compare the removal performances of Fenton processes.

2. Materials and Methods

2.1. Materials

The cosmetic wastewater from the production of automobile care products was supplied from a company in Nevşehir city of Turkey. The automobile care products produced in this company are car washing shampoos, multi-purpose cleaning products, tire care/cleaning products, car waxes, lubricants for car care, air conditioning care/cleaning products, special cleaning products for rim care, brake pad cleaners. The real wastewater samples were taken as 2-hour composite samples into 5 L of bottles, brought to the laboratory by a storage box, and stored in the refrigerator at 4˚C. The main properties of this wastewater were determined and were given in Table 1.

Table 1. The main properties of wastewater

Parameters	Values	Parameters	Values
Temperature (°C)	20.5	Color (420 nm)	1371
pH	7.48	Color (485 nm)	1262
E. Conductivity (µs/cm)	1782	Color (508 nm)	1235
Dissolved oxygen (mg/L)	4.3	COD (mg/L)	1128.9
Orthophosphate (mg/L)	0.63	Sulfate (mg/L)	544.6
Anionic surfactant (mg/L)	1.09		

2.2.Experimental procedure

The Photo-Fenton system has four main parts. As it is seen in Figure 1, the system consists of UV-C radiation lamps of 8 watts mounted parallely, two magnetic stirrers (Mtops MS200), 500 mL of reactors, and a light-proof wooden cabin. The dimensions of the cabin are 50 cm x 50 cm x 42 cm (L x W x H). The Classical-Fenton system consists of a 500 mL of reactor and a Jar Test Flocculator (Velp JLT6). Both processes were conducted in a batch system with 200 mL of wastewater samples. The influences of Fe(II) dose, H_2O_2 dose, and H_2O_2/Fe(II) rates on COD removal were investigated by both Fenton processes. The wastewater sample and reagents were mixed firstly at 300 rpm for two minutes. Then, mixing speed was decreased at 90 rpm and mixed for 45 minutes in Classical-Fenton reactors and 20 min in Photo-Fenton reactors. All experiments were performed in the room temperature. After the precipitation process, the mixture was filtered by using 0.45 µm of membrane filters and COD concentration was analyzed by Closed Reflux Method by using a thermoreactor-Hach LT200 and a spectrophotometer-Hach DR3900 according to the standard method (Baird et al. 2017). E. conductivity and pH were measured by a multi-meter during experiments (Hach HQ40d). COD removal performances are calculated as follows:

$$Removal\ efficiency(\%) = \frac{C_i - C_f}{C_i} x100 \qquad (1)$$

where C_f and C_i are the final and inlet COD concentrations (mg/L), respectively.

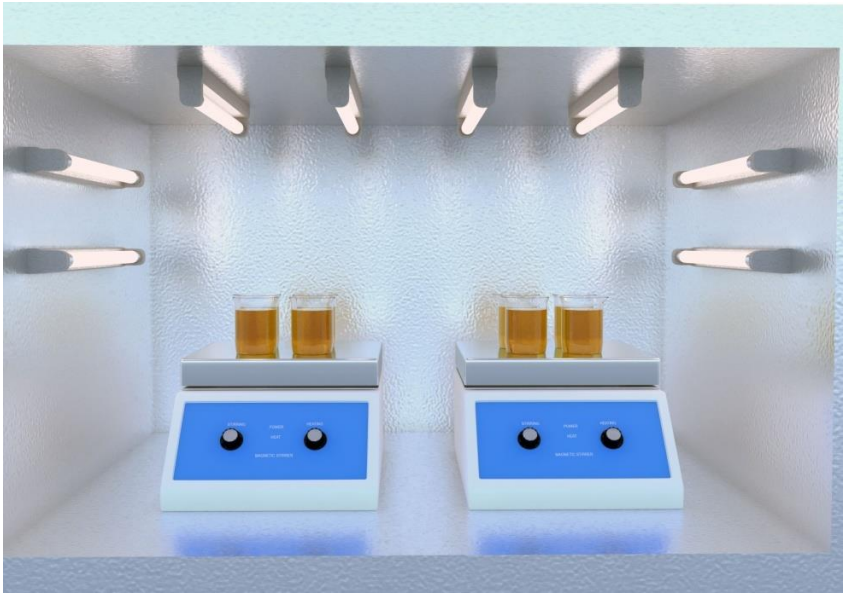


Fig. 1. A schematic representation of Photo-Fenton Process

2.3. Multi-Layer Percepton Neural Networks

Feed-forward neural networks are one of the most popular architectures owing to their structural flexibility, capabilities of well-representational, and a large number of training algorithms available as well as well-known machine learning (Haykin 1999). They are basically designed to replicate the ability for creating and designing new information of the human brain. MLP was firstly proposed by Werbos with the intention of solving the nonlinear problems due to its architectural structure including hidden layer(s) (Werbos 1974). Then, Rumelhart et al. were improved MLP methods (Rumelhart et al. 1986). MLP methods have been widely used for so many areas such as prediction, classifications, modelling etc. This network comprises neurons regulated in layers as input, outputs, and one or more hidden layers. Every neuron is attached to all neurons of the next layer. A number of neurons in the hidden layer have a crucial effect on the performance of the network (Li et al. 2017). Having a data-driven feature that comes from including hidden layers in its structure enables these kinds of neural networks to have flexible and adaptable models for nonlinear problems. The neurons are attached by weights and output signals that are a function of the sum inputs to the neurons modified by a simple nonlinear transfer, or activation function.

2.4. Single Multiplicative Neuron Model

Single multiplicative neuron model (SMN) was firstly introduced to the literature by Yadav. et al. (Yadav et al. 2007). SMN structure has just one neuron as a hidden layer, unlike the MLP. Having this feature makes SMN more advantageous, especially in solving the determination of appropriate structure problems for MLP. In SMN structure, there is only one neuron in the model and instead of an addition operator; the process of multiplication is applied to the signal accruing to the neuron. In Eq.(1), $\Omega(x, \Theta)$ function comprises of the product of weighted inputs. Θ is a vector which include the weights (w_j), X_{ij} is i_{th} sample for j_{th} input, and the biases (b_j) of the model and can be shown with $\Theta = (w_1, w_2, \ldots, w_m, b_1, b_2, \ldots, b_m)$. There are m inputs which are showed with $(X_1, X_2, \ldots X_m)$ and just one output given by y and also f shows the activation function which is the function that specifies the nonlinear relationships between inputs and output. The net value of the neuron is calculated as:

$$net_i = \Omega(x, \Theta) = \prod_{j=1}^{m} \left(w_j X_{ij} + b_j\right), \quad i = 1,2, \ldots n \tag{1}$$

$$y_i = f(net_i) \quad i = 1,2, \ldots n \tag{2}$$

2.5. Particle Swarm Optimization

PSO is a kind of heuristic optimization method, proposed firstly by Kennedy and Eberhart (Kennedy and Eberhart 1995). PSO was improved by adding some coefficients to the optimization process (Shi and Eberhart 1999; Ma et al. 2006). The most significant feature of this algorithm is the ability to reach the optimum point from several points at the same time. So, having this feature gives the opportunity to PSO algorithm to reach global optimum by escaping local optimum. Because of its high solution quality, simplicity, and good convergence properties, recently, the PSO algorithm has been widely applied to the data. In the PSO method, each particle has a position and speed that represents the solution to the optimization problem and the search direction in the search space. While the best positions of

particles are stored in *Pbest* vectors, the best state of all particles is stored in *Gbest* vectors representing the global optimum.

In this study to be able to train the SMN, modified PSO is utilized. The process of modified PSO analysis has some steps which differentiate this model from the traditional one. These are cognitive (c_1) and social (c_2) coefficients, the inertia parameter (w). These parameters are calculated for each iteration by using the following equations.

$$c_1 = (c_{1f} - c_{1i})\frac{t}{maxt} + c_{1i} \tag{3}$$

$$c_2 = (c_{2f} - c_{2i})\frac{t}{maxt} + c_{2i} \tag{4}$$

$$w = (w_2 - w_1)\frac{maxt - t}{maxt} + w_1 \tag{5}$$

Here, (c_{1i}, c_{1f}) are possible intervals for cognitive coefficients, (c_{2i}, c_{2f}) are ranges for social coefficients and (w_1, w_2) are inertia parameters. *maxt* gives a maximal number of iterations, and *t* is a valid iteration number. And finally, new values of positions and velocities are calculated with equations given below;

$$V_{id}^{k+1} = [w \times V_{id}^{k} + c_1 \times rand_1 \times (Pbest_{id} - X_{id}) + c_2 \times rand_2 \times (Gbest - X_{id}) \tag{6}$$

$$X_{id}^{k+1} = X_{id} + V_{id}^{k+1} \tag{7}$$

where $rand_1$ and $rand_2$ are random numbers between 0 and 1. After reaching the predetermined iteration number, $Gbest$'s results are taken optimal parameters of the system.

2.6. Genetic Algorithm

Genetic algorithm (GA) was presented by Holland (Holland 1992) and improved by Goldberg (Goldberg 1989). GA is one of the heuristic optimization methods used to find benefit solutions to complicated problems. It contains important parts as the population for selection, crossover, and mutation. Firstly, some random solutions (individuals) that are each containing several features (chromosomes) are created in the algorithm. According to the laws of genetics, cross-over and mutations happen in chromosomes to create the second generation of individuals with more different properties. The calculations for GA function were performed through MATLAB 2018b.

3. Results and discussions

In this study, it is basically aimed to predict COD removal performance of Classical and Photo-Fenton Processes from cosmetic wastewater by using a traditional method (RSM) and three state-of-the-art models (MLP-LM, MLP-PSO, and SMN-PSO). The experiment design properties of Classical and Photo-Fenton Processes are summarized in Table 2. A scheme of the combination of ANNs and heuristic algorithms is presented in Fig.2.

227 **Table 2.** The experimental conditions and architectures of neural networks

Expr. No	Fenton process	Independent Variables	Fixed Variables	MLP Architecture	SMN Architecture
1	Classic	A. Fe(II) dose (50-400 mg/L) B. H_2O_2 dose (600 mg/L, 900 mg/L)	1. pH/3 2.Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90 rpm	from 2-1-1 to 2-4-1	2-1-1
2	Photo	A. Fe(II) dose (50-400 mg/L) B. H_2O_2 dose (600 mg/L, 900 mg/L)	1. pH/3 2.Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90 rpm	from 2-1-1 to 2-4-1	2-1-1
3	Classic	A. H_2O_2 dose (200-1050 mg/L) B. Fe(II) dose (150 mg/L, 300 mg/L, 400 mg/L)	1. pH/3 2. Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90 rpm	from 2-1-1 to 2-4-1	2-1-1
4	Photo	A. H_2O_2 dose (200-1050 mg/L) B. Fe(II) dose (150 mg/L, 300 mg/L, 400 mg/L)	1. pH/3 2.Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90rpm	from 2-1-1 to 2-4-1	2-1-1
5	Classic	A. H_2O_2 dose (200-1050 mg/L) B. Fe(II) dose (50-400 mg/L)	1. pH/3 2. Temperature/ 23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90rpm	from 2-1-1 to 2-4-1	2-1-1
6	Photo	A. H_2O_2 dose (200-1050 mg/L) B. Fe(II) dose (50-400 mg/L)	1. pH/3 2. Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90 rpm	from 2-1-1 to 2-4-1	2-1-1
7	Classic	A. Contact time (0-60 min) B. Fe(II) dose (150 mg/L, 400 mg/L)	1. pH/3 2. Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90 rpm 5.H_2O_2 dose/900 mg/L	from 2-1-1 to 2-4-1	2-1-1
8	Photo	A. Contact time (0-60 min) B. Fe(II) dose (300 mg/L)	1. pH/3 2. Temperature/23±2°C 3. Fast mixing speed/300 rpm 4. Slow mixing speed/90 rpm 5.H_2O_2 dose/600 mg/L	from 2-1-1 to 2-4-1	2-1-1

228
229
230
231
232
233
234
235

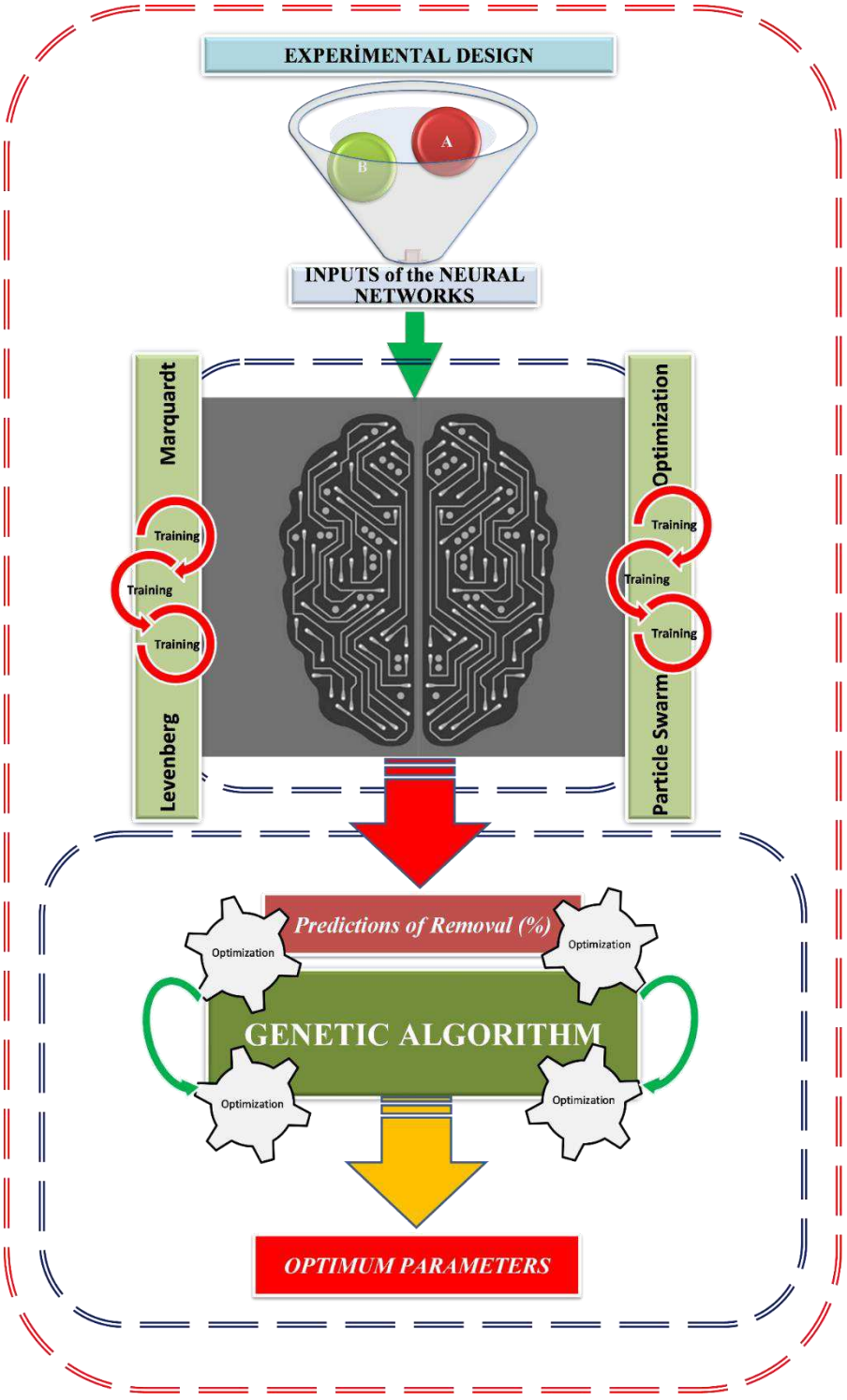


Fig.2. A scheme of the combination of ANNs and heuristic algorithms

3.1. Performance measures

The prediction results produced from RSM, MLP-LM, MLP-PSO, and SMN-PSO models were evaluated from different perspectives. Firstly, RMSE and MAPE that is revealed the predictive performance of models, the basic statistical criteria widely applied in prediction literature were discussed.

$$RMSE = \sqrt{\frac{1}{n}\sum_{p=1}^{n}(Target_p - Output_p)^2} \quad (8)$$

$$MAPE = \text{mean}\left(\left|\frac{Target_p - Output_p}{Target_p}\right|\right) \quad , p = 1,2,\cdots,n \quad (9)$$

The other perspective one is the analysis of the regression model to be created for the predictions and target values, and some properties of this regression model which is given by Eq. (10). For a successful prediction tool, the regression and determination coefficients of the model are expected to be equal to 1 or quite close to 1. In addition, scatter diagrams showing the symphony between predictions and actual values were illustrated.

$$Y_t = \beta\hat{Y}_t + \varepsilon_t \quad (10)$$

3.2. RSM modelling and optimization

This study contains basic information rather than detailed information about RSM since it is a well-known statistical method. The second-order polynomial model structure for the case where there are k independent variables is written as:

$$RE(\%) = \beta_0 + \sum_{i=1}^{k}\beta_i X_i + \sum_{i=1}^{k}\beta_{ii}X_i^2 + \sum_{i=1}^{k}\sum_{i\neq j=1}^{k}\beta_{ii}X_iX_j + \varepsilon \quad (11)$$

For all experiment designs including Classical and Photo-Fenton Processes, the results of the determined data in terms of un-coded factors using RSM are given in Table 3. Considering the results given in Table 1, it is seen from the MAPE values that the predictions obtained by RSM in 5 of 8 experiments contain approximately 5% or more error. On the other hand, the percentage error was around 4% for 3rd experiment, while it was 2.5% for the 7th experiment and 1.33% for the 8th experiment. Considering R^2-value, which is another measure of the success of the model, it is seen that relatively high R^2 values (higher than 90%) are obtained for 2nd, 7th and 8th experiments in parallel with the MAPE values.

Table 3. The RSM results for the experiments

Exp. No	Optimized Model (by omitting insignificant terms)	Optimal Values	y	d	$R^2(\%)$	RMSE	MAPE
1	$\widehat{RE} = 6.1 + 0.4036A - 0.000605A^2$	A. 329.293 B. 900	84.4790	0.7782	92.98	4.4125	6.4823%
2	$\widehat{RE} = 25.04 + 0.3072A - 0.000382A^2$	A. 400 B. 900	88.8368	0.8161	96.74	2.8959	3.9062%
3	$\widehat{RE} = -14.4 + 0.1396A + 0.0250B - 0.000071A^2$	A. 1027.96 B. 400	74.9213	0.6828	91.56	4.9924	8.3217%
4	$\widehat{RE} = 10.4 + 0.0628A + 0.1243B$	A. 1050 B. 400	87.4843	0.8063	89.46	4.4770	6.1295%
5	$\widehat{RE} = -71.9 + 0.2434A + 0.3373B - 0.000156A^2 - 0.000588B^2$	A. 826.768 B. 322.222	82.9769	0.7847	86.31	7.1229	11.5837%
6	$\widehat{RE} = -6.9 + 0.1027A + 0.1947B - 0.000068A^2 - 0.000384B^2 + 0.000107AB$	A. 1050 B. 399.015	86.9010	07973	84.10	5.9786	7.1423%
7	$\widehat{RE} = 28.54 + 1.669A + 0.03699B - 0.01608A^2$	A. 52.222 B. 400	87.1502	0.7868	98.15	1.8424	2.4908%
8	$\widehat{RE} = 39264 + 1.479A - 0.01518A^2$	A. 48.729 B. 300	90.4605	0.7439	97.82	1.4658	1.3340%

3.3. Neural networks-based modelling

Artificial neural networks have been widely used in many scientific areas. In particular, thanks to the rapid development of computer technology in recent years, neural network-based prediction models have started to be used frequently. One of these models is the MLP-LM. While derivative-based training algorithms such as Levenberg-Marquardt learning algorithm could sometimes get stuck in local optimum, particle swarm optimization carries no such risk.

From this point of view, unlike existing studies in the literature that use MLP for similar purposes, MLP and SMN trained by PSO were used as prediction tools in this study. Unlike MLP, there is no architectural selection problem since SMN has only one neuron, making it more applicable. Moreover, using a multiplicative multiplication aggregation function instead of an additive aggregation function makes SMN more flexible and successful, especially in solving nonlinear problems. The use of SMN, which has these features, as a predictive tool in this field is another pioneering and distinguishing feature of this study compared to other studies in the literature. In this respect, this study is the first study that takes into account all the above mentioned issues in its literature. Also, analysis and modelling using MLP-LM, MLP-PSO, and SMN-PSO were performed with MATLAB program codes created by researchers of this study. MLP and SMN structures with the two inputs are given in Figure 3.

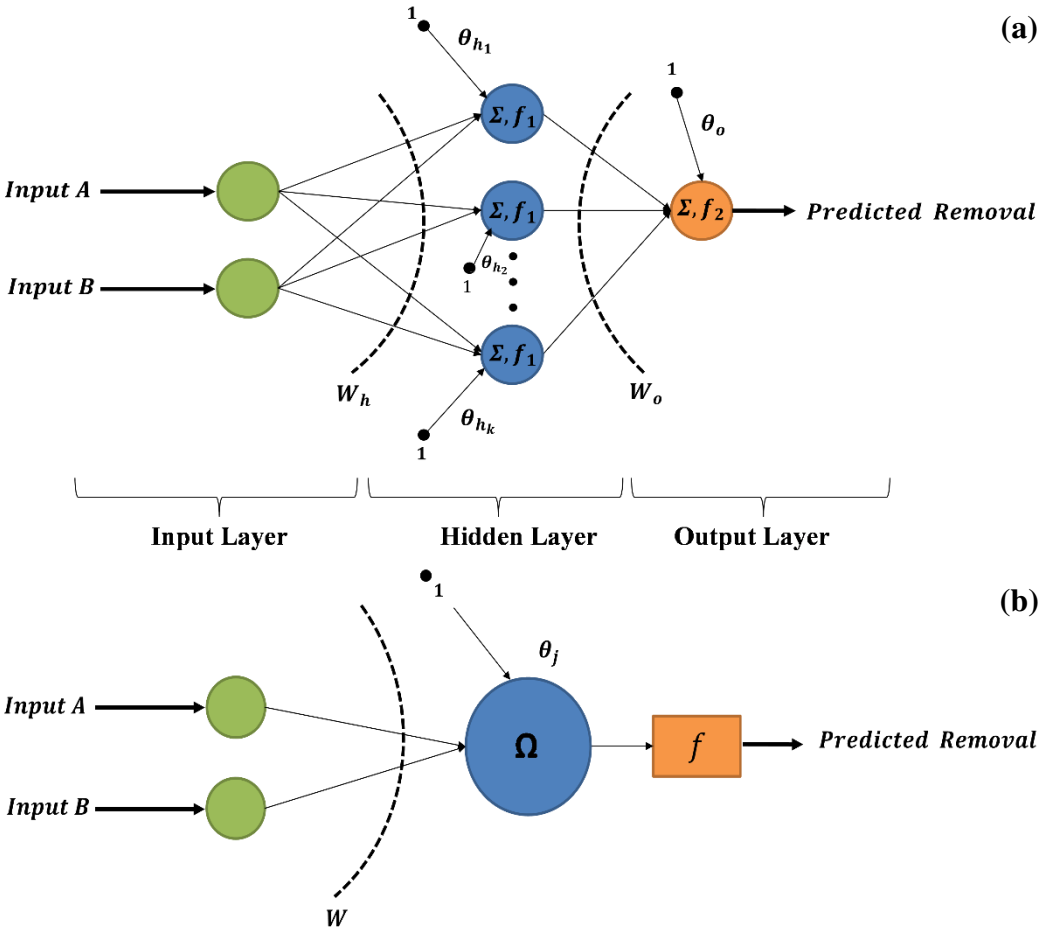


Fig. 3. An illustration of MLP structure / 2-k-1 Architecture (a) and SMN structure (b)

For all experiment designs of Classical and Photo-Fenton Processes, the prediction results produced by MLP-LM, MLP-PSO, and SMN-PSO models are given in Table 4 and Table 5, in terms of RMSE and MAPE criteria, respectively. These tables also give the success rankings of the prediction models according to the corresponding criteria. When the findings given in Tables 4 and 5 were evaluated, it was seen that all three NN-based prediction models produced better predictive results than the RSM in all cases except for one exception, in terms of both criteria. Among these three NN-based models, it is seen that SMN-PSO was the model that produced the best prediction results, as expected, according to the average success rankings created for both criteria. The main reason for this situation is that SMN-PSO trained by PSO has less risk of getting stuck in local optimums than MLP-LM trained with Levenberg-Marquardt algorithm.

Table 4. The prediction results in terms of RMSE

Case	Processes	# Samples		SMN-PSO		MLP-PSO		MLP-LM	
				RMSE	Rank	RMSE	Rank	RMSE	Rank
Effect of Fe (II) doses	Classical-Fenton Process	10	Training	1.7930	**1**	2.4818	2	2.9371	3
		3	Validation	1.5488	**1**	1.6673	2	4.3519	3
		3	Test	0.6645	**1**	0.7423	2	0.8924	3
		16	ALL	1.5943	**1**	2.1152	2	3.0153	3
	Photo-Fenton Process	10	Training	1.7108	**1**	2.2669	3	2.0894	2
		3	Validation	0.5633	**1**	1.5524	3	1.0260	2
		3	Test	0.2093	**1**	0.3939	2	0.5513	3
		16	ALL	1.3773	**1**	1.9217	3	1.7271	2
Effect of H_2O_2 doses	Classical-Fenton Process	11	Training	1.7904	**1**	2.0809	2	2.6345	3
		4	Validation	1.2094	2	1.7601	3	1.1319	**1**
		4	Test	1.0630	**1**	1.5196	2	2.0760	3
		19	ALL	1.5497	**1**	1.9093	2	2.2793	3
	Photo-Fenton Process	15	Training	1.1129	**1**	1.7680	2	2.4760	3
		5	Validation	1.2857	3	1.2258	**1**	1.2715	2
		5	Test	0.7732	**1**	0.9021	2	1.0530	3
		25	ALL	1.0924	**1**	1.5293	2	2.0551	3
Effect of H_2O_2/Fe(II) rates	Classical-Fenton Process	19	Training	1.5455	**1**	1.7818	2	4.0121	3
		5	Validation	1.5539	**1**	1.7180	2	1.8323	3
		5	Test	0.9173	**1**	1.0169	2	1.8114	3
		29	ALL	1.4581	**1**	1.5617	2	3.4192	3
	Photo-Fenton Process	25	Training	1.0027	**1**	1.5506	2	3.3785	3
		5	Validation	0.9724	**1**	1.2980	2	3.7574	3
		5	Test	0.6980	**1**	1.0178	2	2.8843	3
		35	ALL	0.9606	**1**	1.4512	2	3.3702	3
Effect of contact time	Classical-Fenton Process	10	Training	1.1510	2	1.6974	3	0.7898	**1**
		2	Validation	0.5010	2	0.3977	**1**	1.1039	3
		2	Test	0.4060	2	0.2439	**1**	0.4436	3
		14	ALL	1.0028	2	1.4454	3	0.8049	**1**
	Photo-Fenton Process	3	Training	0.8431	**1**	0.8520	2	1.8142	3
		2	Validation	0.8725	**1**	1.2502	2	2.1566	3
		2	Test	0.4968	**1**	1.0160	2	2.3276	3
		7	ALL	0.7699	**1**	1.0260	2	2.0706	4
	Average		Training	**1.1250**		2.2500		2.6250	
			Validation	**1.5000**		2.0000		2.5000	
			Test	**1.1250**		1.8750		3.0000	
			ALL	**1.1250**		2.2500		2.7500	

Table 5. The prediction results in terms of MAPE

Case	Processes	# Samples		SMN-PSO		MLP-PSO		MLP-LM	
				MAPE	Rank	MAPE	Rank	MAPE	Rank
Effect of Fe (II) doses	Classical-Fenton Process	10	Training	3.1235%	**1**	4.2101%	2	5.4812%	3
		3	Validation	2.5373%	**1**	2.9869%	2	4.6784%	3
		3	Test	1.0634%	**1**	1.1402%	3	1.0784%	2
		16	ALL	2.6273%	**1**	3.4052%	2	4.5051%	3
	Photo-Fenton Process	10	Training	2.2558%	**1**	3.4481%	3	2.3766%	2
		3	Validation	0.9992%	**1**	2.2900%	3	1.4986%	2
		3	Test	0.2051%	**1**	0.5819%	2	0.7329%	3
		16	ALL	1.6357%	**1**	2.6935%	3	1.9038%	2
Effect of H_2O_2 doses	Classical-Fenton Process	11	Training	2.7672%	**1**	4.6369%	3	2.9823%	2
		4	Validation	2.1089%	**1**	2.5006%	3	1.7866%	2
		4	Test	1.5827%	**1**	1.6765%	2	3.3245%	3
		19	ALL	2.3793%	**1**	3.5639%	3	2.8027%	2
	Photo-Fenton Process	15	Training	1.6213%	**1**	2.5519%	3	2.3834%	2
		5	Validation	1.1778%	**1**	1.6106%	2	1.1670%	3
		5	Test	1.1082%	**1**	1.3129%	2	1.3742%	3
		25	ALL	1.4300%	**1**	2.1159%	3	1.9383%	2
Effect of H_2O_2/Fe (II)	Classical-Fenton Process	19	Training	2.2681%	**1**	3.0225%	2	4.5416%	3
		5	Validation	2.0428%	**1**	2.7955%	2	3.8585%	3
		5	Test	1.3458%	2	1.1457%	**1**	2.7890%	3
		29	ALL	2.0702%	**1**	2.5643%	2	4.1217%	3
	Photo-Fenton Process	25	Training	1.4591%	**1**	2.2576%	2	4.7201%	3
		5	Validation	1.3227%	**1**	1.8462%	2	4.9370%	3
		5	Test	1.0807%	**1**	1.4088%	2	4.3294%	3
		35	ALL	1.3855%	**1**	2.0776%	2	4.6953%	3
Effect of contact time	Classical-Fenton Process	10	Training	1.7168%	2	2.1805%	3	0.9750%	**1**
		2	Validation	0.7531%	2	0.5466%	**1**	1.6131%	3
		2	Test	0.5132%	2	0.2744%	**1**	0.5944%	3
		14	ALL	1.4072%	2	1.6748%	3	1.0118%	**1**
	Photo-Fenton Process	3	Training	0.9751%	**1**	1.0655%	2	1.6829%	3
		2	Validation	1.0624%	**1**	1.5248%	2	2.1396%	3
		2	Test	0.4825%	**1**	1.1284%	2	2.4598%	3
		7	ALL	0.8593%	**1**	1.2147%	2	2.0353%	4
	Average		Training	**1.1250**		2.5000		2.3750	
			Validation	**1.1250**		2.1250		2.7500	
			Test	**1.2500**		1.8750		2.8750	
			ALL	**1.1250**		2.5000		2.5000	

In addition, SMN does not have architectural selection problems, and thanks to the multiplicative aggregation function it uses, it is more flexible and successful than MLP in these types of problems with non-linearity dominance. Thus, it is seen that the SMN-PSO produced predictive results with errors of around 2% and 3%, and in many cases much lower than these rates, up to values less than 1%, considering the training, validity, test sets and even all data sets for each experiment. Another important reason for preference to use NN-based prediction tools is that unlike RSM, producing in stronger results in terms of reliability and consistency by analyzing the data in different parts such as training, validation, and test

sets. Especially when considering estimates from out-of-sample test sets, the MAPE values were about 1% for almost all experiments. In 2nd, 7th and 8th experiments, the percentage errors even below 1% were observed in the out-of-sample data set for SMN-PSO. These findings indicate that SMN-PSO produces highly satisfactory and consistent predictions even for the out-of-sample data sets. One of the most important aspects of these results is that in cases that experimental designs are difficult or costly, satisfactory predictive results for out-of-sample data sets without the need for any additional experiments are proof that they can be produced with SMN-PSO.

Another way to reveal the superior predictive ability of a prediction tool is to examine some properties of a linear regression model to be established between predictions and target values. Such a strategy was also followed for SMN-PSO, which showed superior performance among the three NN-based prediction models. For the estimate of the regression coefficient ($\hat{\beta}$), a satisfactory and applicable prediction tool and also the coefficient of determination (R^2) of the model $Y_t = \beta \hat{Y}_t + \varepsilon_t$ are desired to be 1 or quite close to 1. Table 6 presents the findings for this regression analysis.

Table 6. The results of the regression analysis for SMN-PSO predictions

Exp. No	Process	# Samples	$Y = \hat{\beta} Y_{pre}$	95% Confidence Interval of β		R^2(%)
				Lower Bound	Upper Bound	
1	Classical-Fenton Process	16	$Y = 1.006850 Y_{pre}$	0.993262	1.020438	99.9399
2	Photo-Fenton Process	16	$Y = 1.0002654 Y_{pre}$	0.988809	1.011720	99.9567
3	Classical-Fenton Process	19	$Y = 1.008765 Y_{pre}$	0.995706	1.021824	99.9317
4	Photo-Fenton Process	25	$Y = 1.001056 Y_{pre}$	0.994011	1.008101	99.9721
5	Classical-Fenton Process	29	$Y = 0.994154 Y_{pre}$	0.985077	1.003231	99.9444
6	Photo-Fenton Process	35	$Y = 0.999413 Y_{pre}$	0.994233	1.004593	99.9779
7	Classical-Fenton Process	14	$Y = 1.005520 Y_{pre}$	0.997223	1.013818	99.9810
8	Photo-Fenton Process	7	$Y = 0.997960 Y_{pre}$	0.988405	1.007515	99.9908

The findings given in Table 4 can be investigated from three different angles. For all experiments, the beta coefficient estimates obtained in the regression estimation equation are pretty close to 1 as expected for a successful prediction tool, a sign that the predictions produced by SMN-PSO are very close to the actual observations. Moreover, the confidence intervals of β (with 95% probability) coefficients cover 1 and also had a very narrow frame. In other words, it can be said that beta coefficients are equal to 1 with a probability of 95%. In addition, the fact that the determination coefficient, R^2, is very close to 1 for each experiment can be seen as proof of the existence of a very high linear relationship between the predictions of SMN-PSO and the actual removal values. This is another feature that a superior prediction tool should have, just like the SMN-PSO.

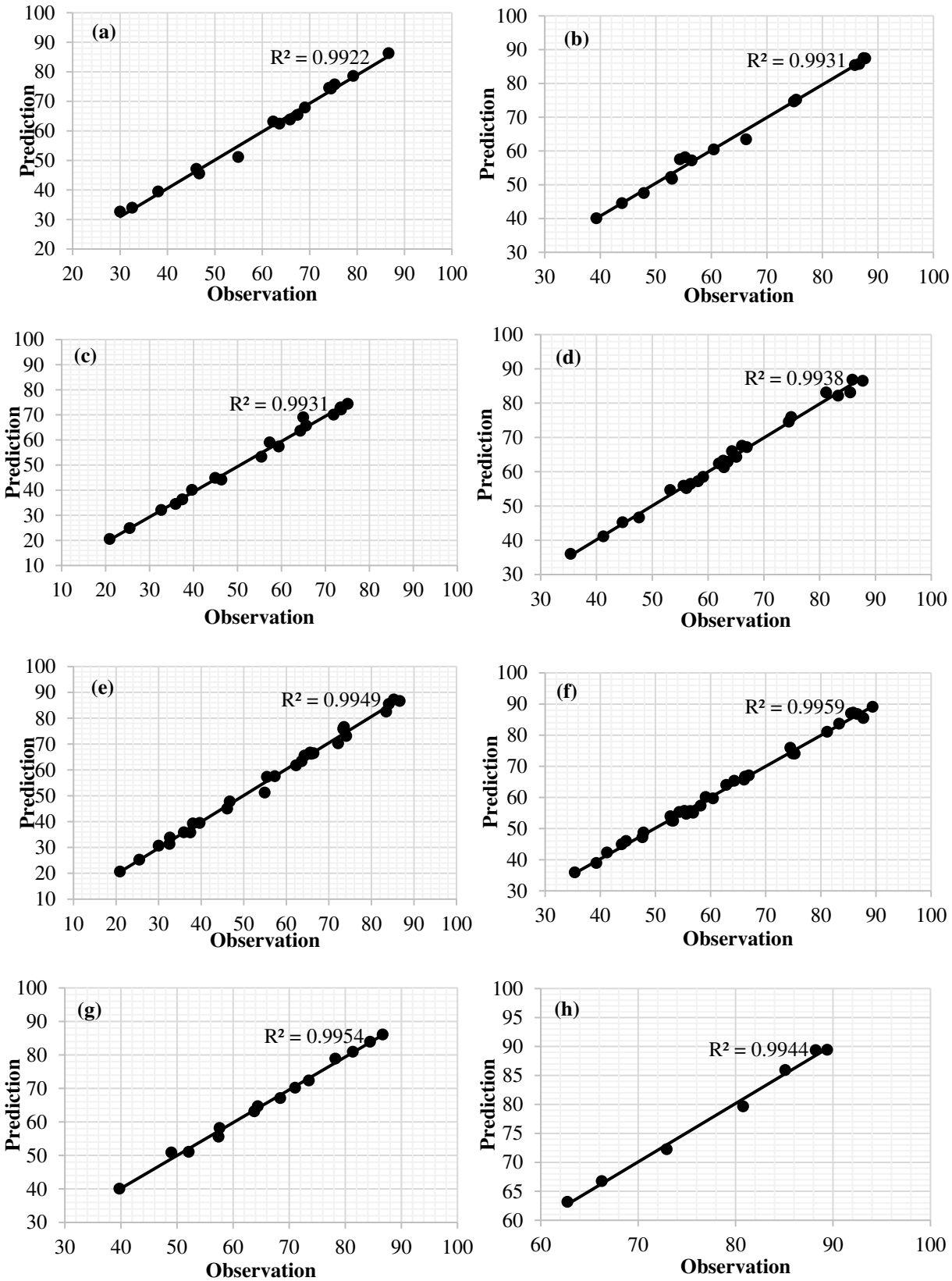


Fig. 4. The scatter gram of observed and predicted removals (a: Exp.1; b: Exp. 2; c: Exp. 3; d: Exp. 4; e: Exp. 5; f: Exp. 6; g: Exp. 7; h: Exp. 8)

Moreover, besides all these statistical evaluations, scatter diagrams displaying the observed and predicted removal efficiency values were used to visualize the superior

prediction performance of the SMN-PSO as the best of the NN-based models used in this study. In a scatter plot, for a prediction tool which is produced satisfactory predictions, it is expected that most of the points on the scatter plots are in proximity to the line segment. The scatter plots, given in Figure 4, also contain examples of exactly this situation. In other words, in the scatter plots, the points were spread very close to the line.

3.4. Optimization via genetic algorithm

After the modelling process, as the final goal of the study, the various operating conditions used for the analysis process were optimized. For this purpose, GA was used to optimize the independent variables to maximize the removal efficiency rate. Here, the basic principle is to determine the independent variable values that will maximize the efficiency of the removal. Generally, the objective function intended to be maximized is given as in Eq. (12).

$$y = f(X_1, X_2) \tag{12}$$

where y is the removal performance of Classical and Photo-Fenton Processes from cosmetic wastewater, X_1 and X_2 represent the independent variables such as H_2O_2 doses, Fe(II) doses, and H_2O_2/Fe(II) rates. The optimization process was carried out for only the SMN-PSO, which showed the highest prediction performance among the three NN-based prediction models used in this study. In this case, the objective function can be given as follows:

$$y = f\left(\frac{1}{1 + e^{-((X_1 \times w_1 + b_1) \times (X_2 \times w_2 + b_2))}}\right) \tag{12}$$

The basic framework of the optimization process for each experiment is summarized in Table 7. An important advantage in modelling this type of data by using NN is that yields of removal corresponding to non-existent parameter values in performed experiments can also be revealed. Moreover, by using the fitness function of a trained neural network, an optimization process performed through GA can determine the optimum values of the experimental parameters to maximize efficiency. Furthermore, it is possible that these values differ from the parameter values used in the experiments already performed. From this perspective, using the fitness function from the SMN-PSO, the parameters of the experiments were optimized via GA. Thus, without the need for many different experiments, the optimum parameter values can be determined to get the maximum removal ratios. When the results given in Table 5 are exemplified; for Experiment 1, with Fe(II) dosage 399.999 mg/L and H_2O_2 dosage 726.1816 mg/L, the optimized condition leads to maximum removal performance (86.4965%) with 99.70% desirability.

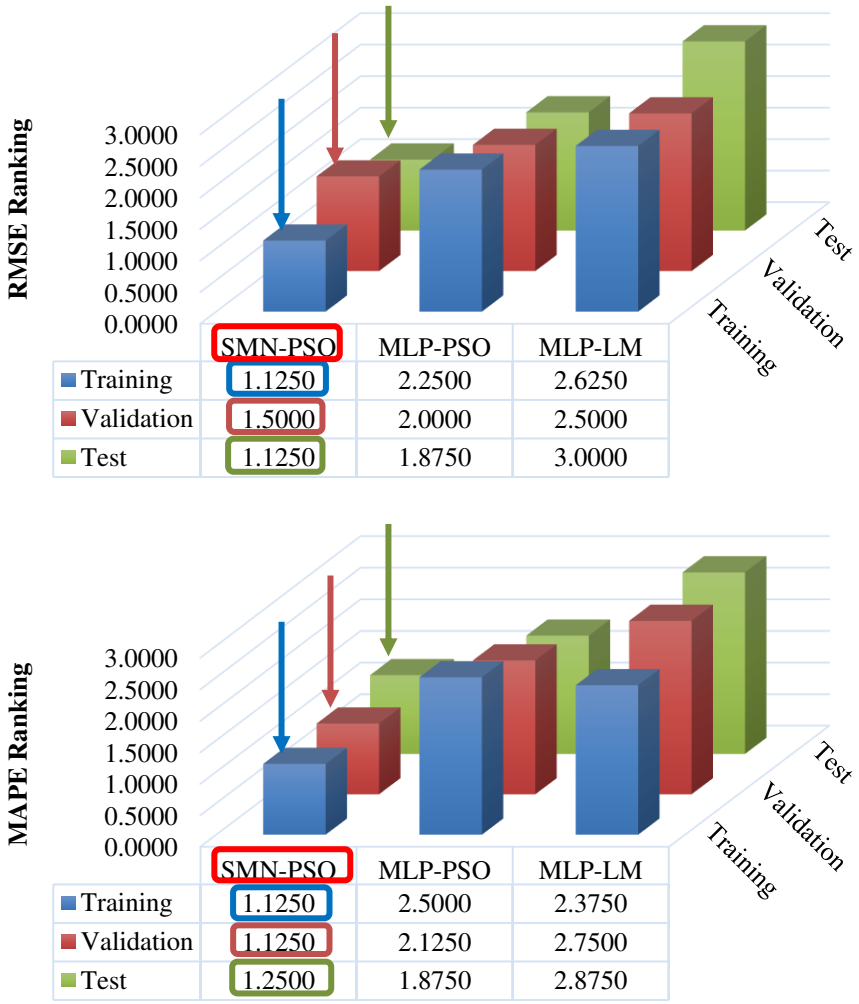
The results of the optimization transaction can be also used to comparatively evaluate the performance of Classical and Photo-Fenton processes. From these results, under the optimum conditions, it is clearly seen that Photo-Fenton processes have higher COD removal performance from cosmetic wastewater in each experiment.

Table 7. The framework of the optimization process

Exp. No	Process	Constraints	Optimal Values	Objective Function Values	Desirability
1	Classical-Fenton	$50 \leq Fe(II)\ (mg/L) \leq 400$ $600 \leq H2O2(mg/L) \leq 900$	399.9999 726.1686	86.4965	99.70%
2	Photo-Fenton	$50 \leq Fe(II)\ mg/L \leq 400$ $600 \leq H2O2(mg/L) \leq 900$	399.9998 894.0719	87.4862	99.48%
3	Classical-Fenton	$200 \leq H2O2(mg/L) \leq 1050$ $150 \leq Fe(II)(mg/L) \leq 400$	992.0797 350.3543	75.0394	99.90%
4	Photo-Fenton	$200 \leq H2O2(mg/L) \leq 1050$ $150 \leq Fe(II)(mg/L) \leq 400$	1029.1429 399.9869	87.4848	99.52%
5	Classical-Fenton	$200 \leq H2O2(mg/L) \leq 1050$ $50 \leq Fe(II)\ (mg/L) \leq 400$	951.9432 400	86.6637	100.00%
6	Photo-Fenton	$200 \leq H2O2(mg/L) \leq 1050$ $50 \leq Fe(II)\ (mg/L) \leq 400$	992.0543 399.9994	89.2568	99.71%
7	Classical-Fenton	$0 \leq Contact\ time\ (min) \leq 60$ $150 \leq Fe(II)(mg/L) \leq 400$	60 335.3312	86.6868	99.99%
8	Photo-Fenton	$0 \leq Contact\ time\ (min) \leq 60$ $Fe(II)(mg/L) = 300$	60 300	89.4118	100.00%

3.5. Comparison of the results

When the results obtained from each model were examined in detail, it was observed that NN-models produced much better prediction results than RSM. The averages of success rankings created by taking into account the performance of all models for eight different experiments, according to RMSE and MAPE criteria, also support this situation. Considering the averages of success rankings, as can be clearly seen in Figures 5-6, SMN-PSO had much better performance than other NN-based models for training, validation, and test sets. Moreover, when all data sets were considered, it was also observed that SMN-PSO had a superior prediction performance than both other NN-based models and RSM.



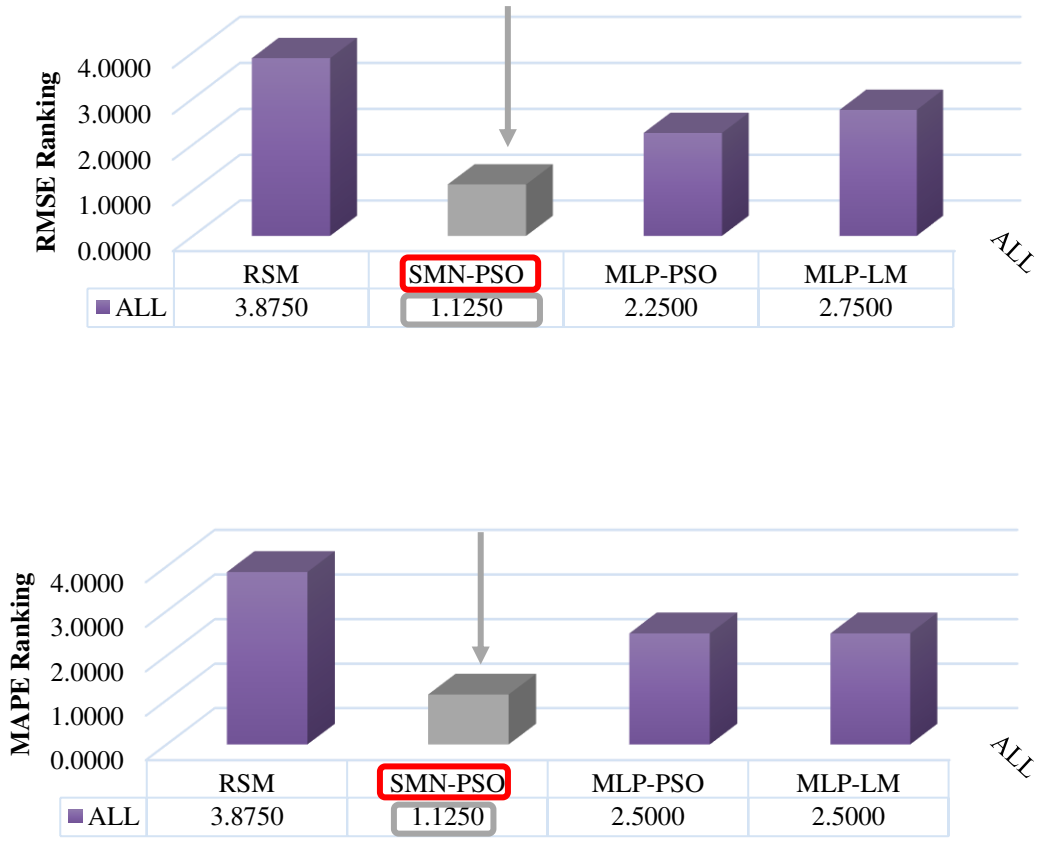
420

421 **Fig. 5.** Comparing NN-based models in terms of RMSE and MAPE

422

423

424



425 **Fig. 6.** Comparing all models in terms of RMSE and MAPE

426

427 ## 4. Conclusion

428 In this study, the prediction of COD removal performance of Classical and Photo-Fenton
429 Processes from cosmetic wastewater was performed by using statistics-based RSM and NN-
430 based three machine learning models. Among all methods, SMN-PSO was observed to be the
431 model that produces the best predictive results in almost all cases.

432 • While SMN-PSO mostly produced predictions with an RMSE value of around 1%, in
433 many cases, it was observed that this value was even below 1%.
434 • In terms of the MAPE criterion, which is a measure of the percentage error, SMN-
435 PSO revealed errors of less than 1% and 2% in most cases and even less than 1% in
436 some cases.
437 • Considering the average of the success rankings, it can be clearly seen that the SMN-
438 PSO produces far superior prediction results compared to other models in terms of
439 both criteria.
440 • The SMN-PSO, which has the best predictive performance in almost all situations,
441 also performed satisfactorily and competitively in other situations.

442 There are several main reasons why SMN-PSO exhibits superior prediction success;

- Since the SMN-PSO is trained by PSO, it does not get caught in local optimum traps unlike derivative-based training algorithms such as LM.
- With the multiplicative aggregation function, SMN-PSO has a higher ability to the adaptation to nonlinear problems.
- Moreover, SMN-PSO, unlike MLPs, does not include a problem such as determining the architecture.

Another unique aspect of this study is that the independent variables are optimized by a genetic algorithm for SMN-PSO, which is used as a function approximation and produces the best prediction results. Here, using the genetic algorithm, the independent values to maximize COD removal performance are obtained in a particular search field. COD removal performances corresponding to the obtained optimum independent variable values provided high desirability values. Moreover, by this means, it is not necessary to perform an additional experiment to achieve all these.

With this study, the optimum conditions can be determined more accurately by modeling the findings obtained from the limited number of experiments conducted to determine the performance of the treatment processes. In this way, technical difficulties, costs, manpower, and time are no longer a problem, and the estimates of the models trained with the available data allow the process to work efficiently. In future studies, statistical-based prediction models and NN-based machine learning models can be combined to get better predictions of removal performance.

References

Baird RB, Eaton AD, Rice EW (2017) Standard Methods for the Examination of Water and Wastewater. 23rd edition

Banerjee P, Dey T kumar, Sarkar S, et al (2016) Treatment of cosmetic effluent in different configurations of ceramic UF membrane based bioreactor: Toxicity evaluation of the untreated and treated wastewater using catfish (Heteropneustes fossilis). Chemosphere 146:133–144. https://doi.org/10.1016/j.chemosphere.2015.12.004

Baştürk E, Alver A (2019) Modeling azo dye removal by sono-fenton processes using response surface methodology and artificial neural network approaches. J Environ Manage 248:109300

Bautista P, Mohedano AF, Gilarranz MA, et al (2007) Application of Fenton oxidation to cosmetic wastewaters treatment. J Hazard Mater 143:128–134. https://doi.org/10.1016/j.jhazmat.2006.09.004

Bautista P, Mohedano AF, Menéndez N, et al (2010) Catalytic wet peroxide oxidation of cosmetic wastewaters with Fe-bearing catalysts. Catal Today 151:148–152. https://doi.org/10.1016/j.cattod.2010.01.023

Bom S, Jorge J, Ribeiro HM, Marto J (2019) A step forward on sustainability in the cosmetics industry: A review. J Clean Prod 225:270–290. https://doi.org/10.1016/j.jclepro.2019.03.255

Cagcag Yolcu O, Bas E, Egrioglu E, Yolcu U (2018) Single Multiplicative Neuron Model

485 Artificial Neural Network with Autoregressive Coefficient for Time Series Modelling.
486 Neural Process Lett 47:1133–1147. https://doi.org/10.1007/s11063-017-9686-3

487 de Andrade PM, Dufrayer CR, Ionashiro EY, de Brito NN (2020) The use of metallurgical
488 waste for heterogeneous photo Fenton-Like treatment of cosmetic effluent. J Environ
489 Chem Eng 8:104148. https://doi.org/10.1016/j.jece.2020.104148

490 Egrioglu E, Aladag C, Yolcu U, et al (2013) Fuzzy Time Series Method Based on
491 Multiplicative Neuron Model and Membership Values. Am J Intell Syst 3:33–39.
492 https://doi.org/10.5923/j.ajis.20130301.05

493 El-Gohary F, Tawfik A, Mahmoud U (2010) Comparative study between chemical
494 coagulation/precipitation (C/P) versus coagulation/dissolved air flotation (C/DAF) for
495 pre-treatment of personal care products (PCPs) wastewater. Desalination 252:106–112.
496 https://doi.org/10.1016/j.desal.2009.10.016

497 Elmolla ES, Chaudhuri M, Eltoukhy MM (2010) The use of artificial neural network (ANN)
498 for modeling of COD removal from antibiotic aqueous solution by the Fenton process. J
499 Hazard Mater 179:127–134. https://doi.org/10.1016/j.jhazmat.2010.02.068

500 Fernandes NC, Brito LB, Costa GG, et al (2018) Removal of azo dye using Fenton and
501 Fenton-like processes: Evaluation of process factors by Box–Behnken design and
502 ecotoxicity tests. Chem Biol Interact 291:47–54

503 Friha I, Karray F, Feki F, et al (2014) Treatment of cosmetic industry wastewater by
504 submerged membrane bioreactor with consideration of microbial community dynamics.
505 Int Biodeterior Biodegrad 88:125–133. https://doi.org/10.1016/j.ibiod.2013.12.015

506 Gholizadeh AM, Zarei M, Ebratkhahan M, Hasanzadeh A (2021) Phenazopyridine
507 degradation by electro-Fenton process with magnetite nanoparticles-activated carbon
508 cathode, artificial neural networks modeling. J Environ Chem Eng 9:104999.
509 https://doi.org/10.1016/j.jece.2020.104999

510 Goldberg DE (1989) Genetic Algorithms in Search, Optimization and Machine Learning 13th
511 ed. Edition. Addison-Wesley Publishing Company, Boston, United States

512 Haykin S (1999) Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice-Hall,
513 New York

514 Holland JH (1992) Adaptation in Natural and Artificial Systems: An Introductory Analysis
515 with Applications to Biology, Control, and Artificial Intelligence. MIT Press

516 Jaafarzadeh N, Ahmadi M, Amiri H, et al (2012) Predicting Fenton modification of solid
517 waste vegetable oil industry for arsenic removal using artificial neural networks. J
518 Taiwan Inst Chem Eng 43:873–878. https://doi.org/10.1016/j.jtice.2012.05.008

519 Kennedy J, Eberhart R (1995) Particle Swarm Optimisation. In: Proceedings of IEEE
520 international conference on neural networks. Piscataway, NJ: IEEE Service Center,
521 Perth, Australia, pp 1942–1948

522 Lek S, Delacoste M, Baran P, et al (1996) Application of neural networks to modelling
523 nonlinear relationships in ecology. Ecol Modell 90:39–52. https://doi.org/10.1016/0304-
524 3800(95)00142-5

525 Li D, Lv R, Si G, You Y (2017) Hybrid neural network-based prediction model for
526 tribological properties of polyamide6-based friction materials. Polym Compos 38:1705–
527 1711

528 Ma Y, Jiang C, Hou Z, Wang C (2006) The formulation of the optimal strategies for the
529 electricity producers based on the particle swarm optimization algorithm. IEEE Trans
530 Power Syst 21:1663–1671. https://doi.org/10.1109/TPWRS.2006.883676

531 Melo ED de, Mounteer AH, Leão LH de S, et al (2013) Toxicity identification evaluation of
532 cosmetics industry wastewater. J Hazard Mater 244–245:329–334.
533 https://doi.org/10.1016/j.jhazmat.2012.11.051

534 Monsalvo VM, Lopez J, Mohedano AF, Rodriguez JJ (2014) Treatment of cosmetic
535 wastewater by a full-scale membrane bioreactor (MBR). Environ Sci Pollut Res
536 21:12662–12670. https://doi.org/10.1007/s11356-014-3208-x

537 Muszyński A, Marcinowski P, Maksymiec J, et al (2019) Cosmetic wastewater treatment with
538 combined light/Fe0/H2O2 process coupled with activated sludge. J Hazard Mater
539 378:120732. https://doi.org/10.1016/j.jhazmat.2019.06.009

540 Naumczyk J, Bogacki J, Marcinowski P, Kowalik P (2014) Cosmetic wastewater treatment by
541 coagulation and advanced oxidation processes. Environ Technol (United Kingdom)
542 35:541–548. https://doi.org/10.1080/09593330.2013.808245

543 Oller I, Malato S, Sánchez-Pérez JA (2011) Combination of Advanced Oxidation Processes
544 and biological treatments for wastewater decontamination-A review. Sci Total Environ
545 409:4141–4166. https://doi.org/10.1016/j.scitotenv.2010.08.061

546 Paździor K, Bilińska L, Ledakowicz S (2019) A review of the existing and emerging
547 technologies in the combination of AOPs and biological processes in industrial textile
548 wastewater treatment. Chem Eng J 376:120597.
549 https://doi.org/10.1016/j.cej.2018.12.057

550 Puyol D, Monsalvo VM, Mohedano AF, et al (2011) Cosmetic wastewater treatment by
551 upflow anaerobic sludge blanket reactor. J Hazard Mater 185:1059–1065.
552 https://doi.org/10.1016/j.jhazmat.2010.10.014

553 Radwan M, Gar Alalm M, Eletriby H (2018) Optimization and modeling of electro-Fenton
554 process for treatment of phenolic wastewater using nickel and sacrificial stainless steel
555 anodes. J Water Process Eng 22:155–162. https://doi.org/10.1016/j.jwpe.2018.02.003

556 Rumelhart E, Hinton G, Williams R (1986) Learning internal representations by error
557 propagation. The M.I.T. Press, Cambridge, pp 318–362

558 Sabour MR, Amiri A (2017) Comparative study of ANN and RSM for simultaneous
559 optimization of multiple targets in Fenton treatment of landfill leachate. Waste Manag
560 65:54–62. https://doi.org/10.1016/j.wasman.2017.03.048

561 Shi Y, Eberhart RC (1999) Empirical study of particle swarm optimization. Proc 1999 Congr
562 Evol Comput CEC 1999 3:1945–1950. https://doi.org/10.1109/CEC.1999.785511

563 Talwar S, Verma AK, Sangal VK (2019) Modeling and optimization of fixed mode dual
564 effect (photocatalysis and photo-Fenton) assisted Metronidazole degradation using ANN
565 coupled with genetic algorithm. J Environ Manage 250:.

https://doi.org/10.1016/j.jenvman.2019.109428

Tolba A, Gar Alalm M, Elsamadony M, et al (2019) Modeling and optimization of heterogeneous Fenton-like and photo-Fenton processes using reusable Fe3O4-MWCNTs. Process Saf Environ Prot 128:273–283. https://doi.org/10.1016/j.psep.2019.06.011

Werbos P (1974) Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Harvard University

Wiliński PR, Marcinowski PP, Naumczyk J, Bogacki J (2017) Pretreatment of cosmetic wastewater by dissolved ozone flotation (DOF). Desalin Water Treat 71:95–106. https://doi.org/10.5004/dwt.2017.20552

Yadav RN, Kalra PK, John J (2007) Time series prediction with single multiplicative neuron model. Appl Soft Comput J 7:1157–1163. https://doi.org/10.1016/j.asoc.2006.01.003

Yolcu U, Egrioglu E, Bas E, et al (2019) Probabilistic forecasting, linearity and nonlinearity hypothesis tests with bootstrapped linear and nonlinear artificial neural network. J Exp Theor Artif Intell 00:1–22. https://doi.org/10.1080/0952813X.2019.1595167

Zarei M, Khataee AR, Ordikhani-Seyedlar R, Fathinia M (2010) Photoelectro-Fenton combined with photocatalytic process for degradation of an azo dye using supported TiO2 nanoparticles and carbon nanotube cathode: Neural network modeling. Electrochim Acta 55:7259–7265. https://doi.org/10.1016/j.electacta.2010.07.050

Declarations:

Ethics approval: Not applicable

Consent to participate: Not applicable

Consent to publish: Not applicable

Authors' contributions:

HC: Investigation, Experiment, Reviewing

FAT: Supervision, Data curation, Writing- Original draft preparation, Visualization, Investigation, Reviewing and Editing

OCY: Software, Validation, Writing- Original draft preparation, Investigation, Reviewing and Editing

Funding: No Funding

Competing interests: The authors declare that they have no competing interests.

Availability of data and materials: Not applicable

Figures

Figure 1

A schematic representation of Photo-Fenton Process

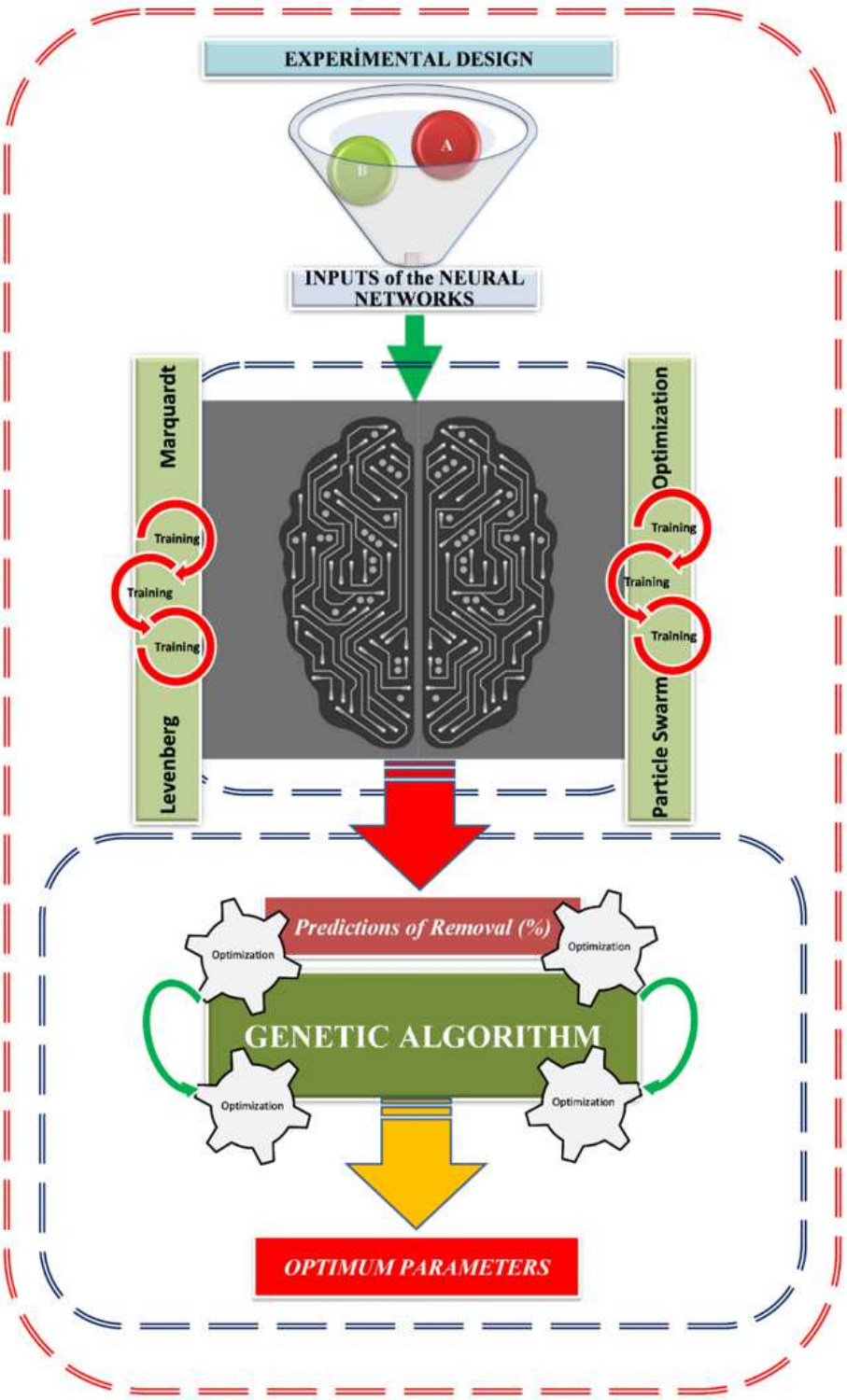


Figure 2

A scheme of the combination of ANNs and heuristic algorithms

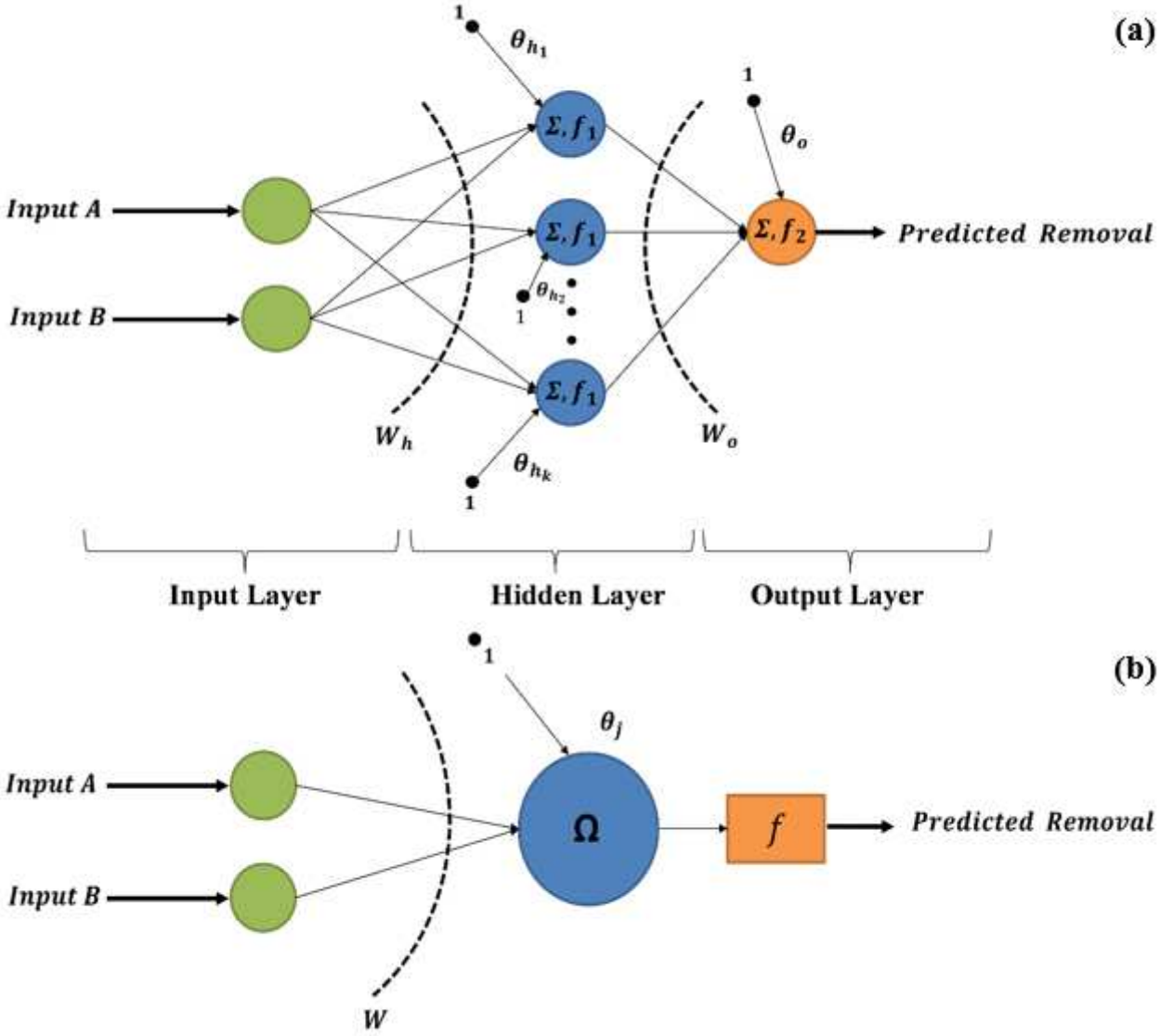


Figure 3

An illustration of MLP structure / 2-k-1 Architecture (a) and SMN structure (b)

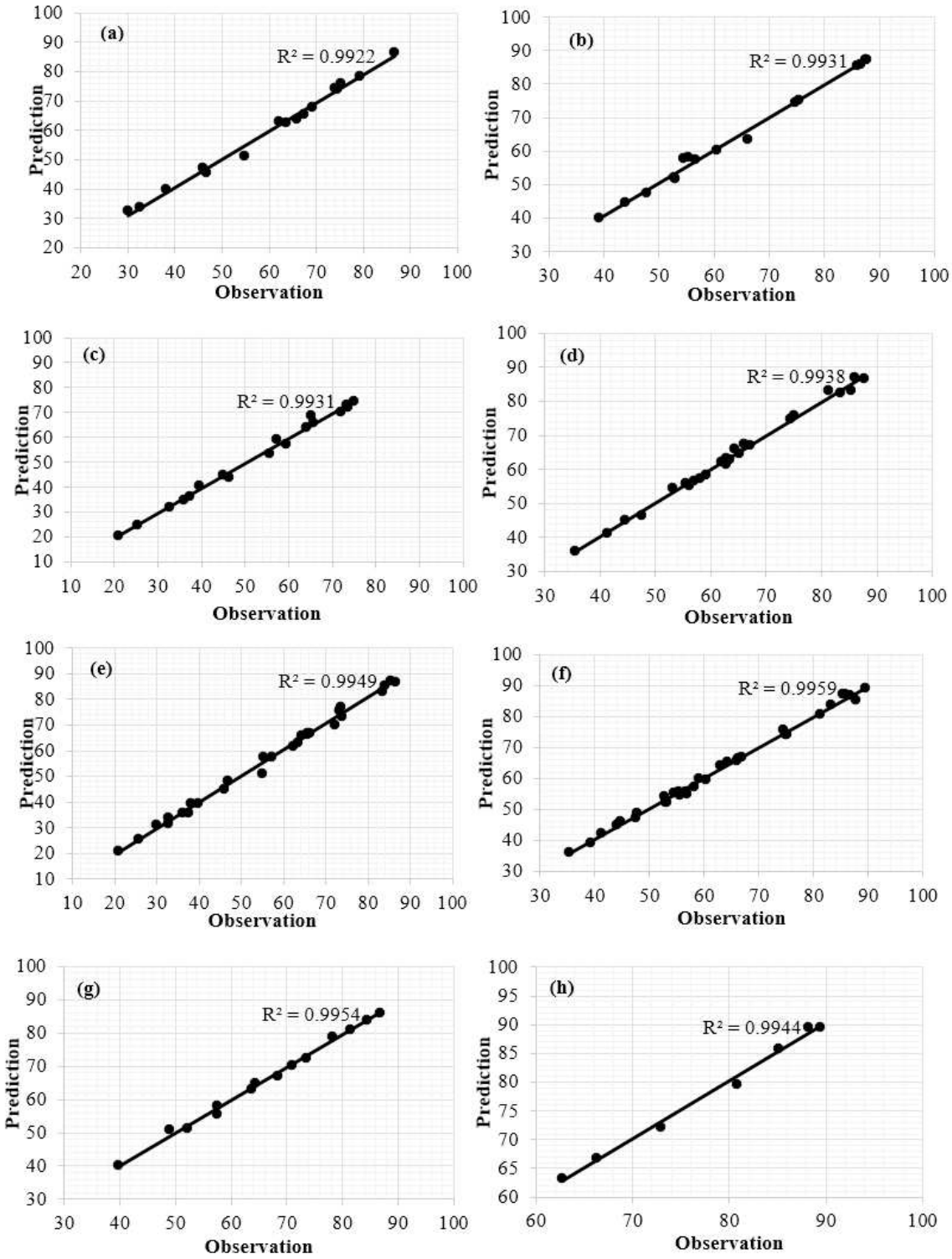
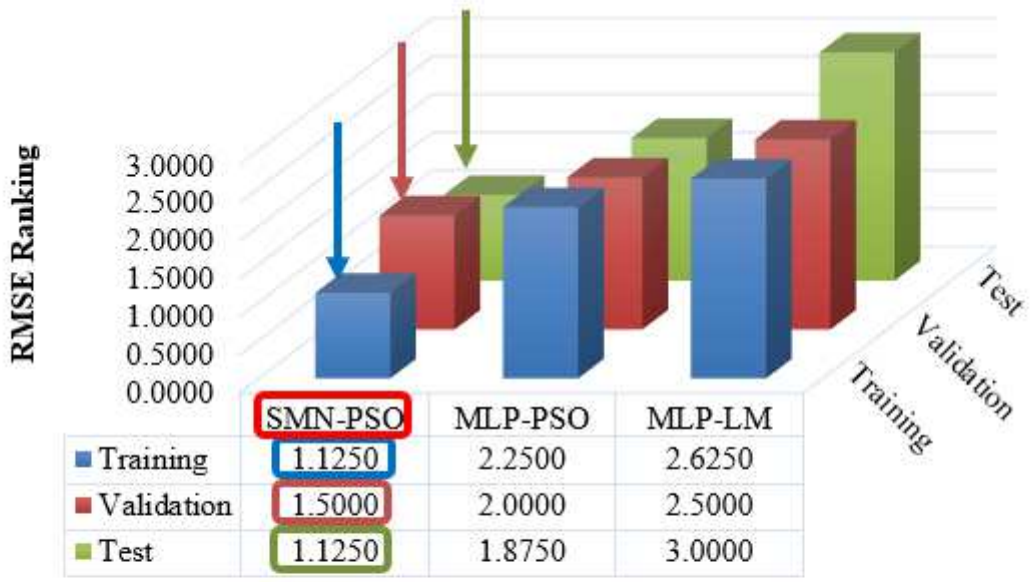
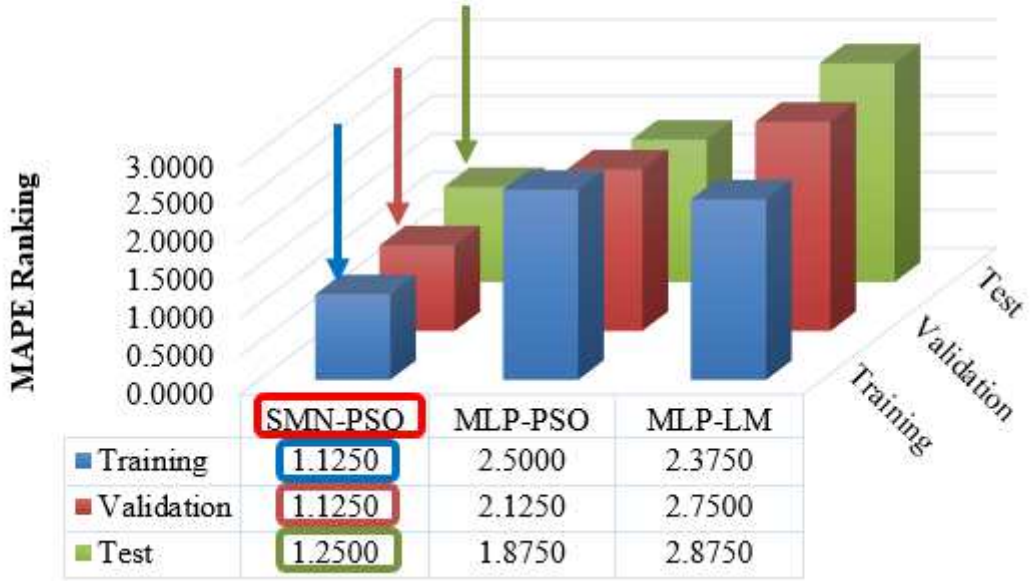


Figure 4

The scatter gram of observed and predicted removals (a: Exp.1; b: Exp. 2; c: Exp. 3; d: Exp. 4; e: Exp. 5; f: Exp. 6; g: Exp. 7; h: Exp. 8)



	SMN-PSO	MLP-PSO	MLP-LM
Training	1.1250	2.2500	2.6250
Validation	1.5000	2.0000	2.5000
Test	1.1250	1.8750	3.0000



	SMN-PSO	MLP-PSO	MLP-LM
Training	1.1250	2.5000	2.3750
Validation	1.1250	2.1250	2.7500
Test	1.2500	1.8750	2.8750

Figure 5

Comparing NN-based models in terms of RMSE and MAPE

Figure 6

Comparing all models in terms of RMSE and MAPE