ANN-assisted forecasting of adsorption efficiency to remove heavy metals

Magdi BUAISHA[1] , **Şaziye BALKU**[2] , **Şeniz ÖZALP YAMAN**[1,*]

[1]Department of Chemical Engineering and Applied Chemistry, Faculty of Engineering, Atılım University, Ankara, Turkey

[2]Department of Energy Systems Engineering, Faculty of Engineering, Atılım University, Ankara, Turkey

Abstract: In wastewater treatment, scientific and practical models utilizing numerical computational techniques such as artificial neural networks (ANNs) can significantly help to improve the process as a whole through adsorption systems. In the modeling of the adsorption efficiency for heavy metals from wastewater, some kinetic models have been used such as pseudo first-order and second-order. The present work develops an ANN model to forecast the adsorption efficiency of heavy metals such as zinc, nickel, and copper by extracting experimental data from three case studies. To do this, we apply trial-and-error to find the most ideal ANN settings, the efficiency of which is determined by mean square error (MSE) and coefficient of determination (R^2). According to the results, the model can forecast adsorption efficiency percent (AE%) with a tangent sigmoid transfer function (tansig) in the hidden layer with 10 neurons and a linear transfer function (purelin) in the output layer. Furthermore, the Levenberg–Marquardt algorithm is seen to be most ideal for training the algorithm for the case studies, with the lowest MSE and high R^2. In addition, the experimental results and the results predicted by the model with the ANN were found to be highly compatible with each other.

Key words: Artificial neural network, adsorption, kinetic model, heavy metals

1. Introduction

Increasing awareness regarding the environment has led to the passing of strict regulations in reducing pollutants in treatment plants, namely of organic carbon, nutrients, and heavy metals. Wastewater contains numerous heavy metals, such as lead (Pb), cadmium (Cd), chromium (Cr), nickel (Ni), cobalt (Co), zinc (Zn), and copper (Cu), each with countless negative effects on both humans and the environment [1]. Some heavy metals are the result of industrial operations, such as mine drainage, metal processing, petroleum refining, tanning, photographic processing, or electroplating; others originate from the environment itself, such as from agricultural runoff and acid rain [2].

The presence of heavy metals in wastewater in high concentrations is not acceptable due to their toxic impacts on the environment and inadequacy of biological treatment. Metal removal plays an importance role in treatment plants; for this reason, it is crucial to be able to define precisely the quantity and quality of these materials when discharged into receiving waters in order to assess and improve the performance of wastewater treatment plants.

In particular, heavy metals also have an effect in terms of the biological treatment of wastewater. Recent studies showed that the heavy metal content in activated sludge systems has an inhibitory effect

---

*Correspondence: seniz.ozalpyaman@atilim.edu.tr

on the heterotrophic growth rate constant and on the lysis rate constant [3]. Another study examined the inhibitory effects of copper and zinc on the specific growth rate of autotrophic biomass in activated sludge, suggesting that copper reduces the nitrification process by up to 50% and zinc by 12% [4]. This makes the removal of such contents from wastewater a necessary task. To do this, heavy metal ions can be decreased by various conventional techniques, including chemical precipitation, reverse osmosis, electrochemical treatment, ion exchange, membrane filtration, coagulation, extraction, irradiation, and adsorption [5].

Adsorption technology has many advantages including low cost and high efficiency. For this reason, the technique has been considered among the best novel technologies available compared to others. In this respect, traditional sorbents such as activated carbons, clay minerals, chelating materials, and chitosan/natural zeolites have low sorption capacities. For this reason, nanomaterials are considered as the ideal materials to remove heavy metal ions in wastewater [6]. There are many factors, such as stability and the size of nanoparticles, which affect the applications of magnetic nanoparticles. For example, as the stability of these particles increases, the adsorption increases, whereas once the size of these nanoparticles increases, the adsorption level drops. This is because these particles perform better in case of nanoparticles being less than 10–20 nm [7]. Due to certain advantages, the use of nanomaterial to remove heavy metals is efficient, mainly due to increased adsorption [8].

1.1. Nanomaterials for adsorption

Heavy metals can be removed by inorganic material. For example, carbon nanotubes (CNTs) have high sorption efficiency [6]. There are other types of nanoparticles made of metal oxides, called "inorganic nanomaterials", which can reduce heavy metal ions in wastewater, such as silver nanoparticles, iron oxide, manganese oxide, titanium oxide, magnesium oxide, copper oxide, and cerium oxide. These oxides are characterized by large surface areas, low solubility, high separation efficiency in wastewater, and no environmental pollution effects. In particular, recent studies have shown that particle size is very important in order to obtain optimum separation of heavy metals during wastewater treatment processes [5].

1.2. Optimization of adsorption parameters

Each type of nanoparticle (NP) has its own adsorption capacities depending on the solution pH, temperature, initial heavy metal concentration, adsorbent dosage, and the presence of other coexisting ions. Therefore, both the adsorption performance and capacity are likely to increase if these parameters are optimized. According to the already existing literature on the topic, the adsorption capacity of Fe_3O_4 NPs increases with elevated solution pH [9]. Other experimental results revealed that considerable amounts of Cu(II) can be removed using chitosan-coated magnetic nanoparticles (CMNPs) at pH 6 [10]. Furthermore, the percent removal of Cr(VI) using Fe_3O_4@n-SiO_2 NPs increases to 2% by only increasing the temperature from 25 to 45 °C and keeping all the other experimental variables constant [11]. The dosage is another factor that influences the adsorption capacity of nanomaterial. According to previous studies [10], the ratio of Cu(II) ion removal from wastewater increases with elevating doses of CMNPs. The contact time, on the other hand, offers clues about the equilibrium time and the kinetics for the removal of heavy metals. For instance, extending the contact time increases the effectiveness of CMNPs in Cu(II) ion separation.

1.3. Adsorption isotherm

An adsorption isotherm explains how an adsorbent and an adsorbate interact with each other, making it an important element in process optimization. Numerous adsorption isotherm models are available that describe

the related equilibrium; some of them are the Langmuir, Freundlich, BET, Toth, Tempkin, Redlich–Peterson, Sips, Frumkin, Harkins–Jura, Halsey, Henderson, and Dubinin–Radushkevich models. According to the studies reviewed for the purpose of the present work, the Langmuir and Freundlich isotherm models are the best fitted models for Cr(VI) adsorption onto iron oxide nanosorbents [11].

1.4. Adsorption kinetics model

Adapting the experimental data to different kinetic models allows us to study the adsorption rate, the process type, and the product information on the adsorption, the kind of information that is commonly referred to as 'physisorption' or 'chemisorption'. The most popular kinetic models are Lagergren pseudo first-order and pseudo second-order kinetics. Among them, the pseudo second-order kinetic model has been regarded as the best fitted model with aqueous chromium adsorption onto different magnetic nanosorbents [11], aqueous Zn(II) ion adsorption onto Fe_3O_4 NPs [9], and aqueous Cu(II) ion adsorption onto CMNPs [10].

1.5. ANN model

In engineering, modeling is a sound and effective way to figure out removal operations [12]. However, doing so with ordinary mathematical or mechanistic models can be expensive and time-intensive given the vast number of tests needed. Apart from this, wastewater treatment is complicated and subject to many factors and variants. As a result, one cannot both model and improve such operations with traditional means [13]. Lately, empirical models have come to the forefront, such as least squares support vector machines, response surface methodology, and artificial neural networks (ANNs); the latter is based on biological neurons related to artificial intelligence and widely used in ventures like intelligent search, autonomous driving, big data, pattern recognition, and robotics [14].

For almost twenty years, ANNs have been chosen as an option to find out about intricate ties among factors based on the input and output data and no information as to the removal processes themselves [15]. Thanks to determining very nonlinear behaviors, ANNs are in use in environmental engineering fields such as wastewater treatment [16–20], membrane processes [21–25], and biosorption [26,27] to predict and model them. Hence, they are regarded as efficient in complicated removal procedures calling for numerous tests as well as chemical reagents [28]. However, limited data are available as to their use in forecasting the adsorption efficiency of heavy metal from aqueous solutions [29,30]. In this respect, ANN modeling is another one of these applications to determine the optimum conditions of the operating parameters for iron metal removal from other substances with the help of the sequential quadratic programming optimization algorithm [31]. As proposed by Fanaie et al. [32], ANNs have also been used in order to predict the efficiency of the biosorption process to remove 4-chlorophenol from aqueous solutions. The comparison between predicted and experimental results provided a high level of determination coefficient ($R^2 = 0.98$), which implies that the proposed model was able to predict the biosorption effectiveness with acceptable accuracy. Other examples of ANN modeling have been used to determine the optimum conditions for the operating parameters as a complex adsorption system model developed for metal removal using activated carbon [33], and another model to predict the adsorbed amount of Co(II) and Ni(II) ion metals from wastewater via carboxymethyl chitosan-bounded Fe_3O_4 nanoparticles [34].

In view of all the related work cited in the field, the present paper offers a discussion of different types of nanomaterials, adsorption isotherm models, and adsorption kinetics. As regards the control of the adsorption process, the present work is concerned with choosing the best ANN model and its related parameters to forecast the adsorption efficiency of zinc(II), nickel(II), and copper(II) in the selected cases. The model will later be

used as an error predictor to achieve our purpose; the operating variables mentioned in these studies are applied as inputs to set up a neural network in order to predict the removal efficacy of heavy metals from wastewater at any time as an output.

2. Results and discussion

The experimental results show that the NPs offer very high adsorption capacity in the case of heavy metals. The development of the ANN model may explain the performance of the decontamination process with less need for experimentation, thereby leading to less chemical substance consumption, less processing time, and a better understanding of the actual workings of the process.

The developed ANN model could describe the behavior of a complicated process within the extent of the experimental status. In the present study, an ANN is developed for the removal of heavy metals from wastewater using nanoabsorbents for three case studies as explained in the following sections.

2.1. Prediction of Zn(II) ion removal

To forecast Zn(II) removal, the best model contains three layers: an input with four operating parameters, a hidden layer with various neurons, and an output layer to consider removal efficiency. Maintaining all these parameters at the same level helps in determining ideal training algorithms. Ten, 20, and 30 neurons in the hidden layer are applied as well as two training algorithms as summarized in Table 1, leading to two different ANN models. Accordingly, the lowest MSE and highest R^2 values are found as 0.008 and 0.98, respectively, using the Levenberg–Marquardt (LM) algorithm (Trainlm), making it the most ideal algorithm to forecast Zn(II) removal with 10 neurons.

Table 1. Different training algorithms with different neurons in the hidden layer for predicting Zn(II) removal.

Algorithm	Training function	Number of neurons (H)	R^2	MSE (average performance)
Levenberg–Marquardt	Trainlm	10	0.98	0.008
		20	0.96	0.012
		30	0.95	0.015
Scaled conjugate gradient	Trainscg	10	0.92	0.700
		20	0.84	0.891
		30	0.82	0.924

With LM being the most ideal algorithm to forecast Zn(II) ion removal with 10 neurons, the results were compared with those of the scaled conjugate gradient (SCG) algorithm for the same number of neurons (10). For this purpose, Figure 1 represents the MSE for the two developed ANN models using the training, testing, and validation data results for the two algorithms. Based on this, the LM algorithm is shown to yield the lowest MSE for the training, testing, and validation sets, while the average performance of the LM algorithm is the most ideal in the case of the SCG algorithm at 0.008. For this reason, LM is regarded as the most ideal algorithm with the lowest MSE at epoch 8 and the best validation performance at 0.01041 with 10 neurons.

The LM and SCG algorithms were evaluated upon comparison of the forecast output results with experimental ones using separate training, validation, and testing datasets applying R^2. The plots for both experimental and predicted results appear in Figure 2 for the training, validation, and testing datasets.
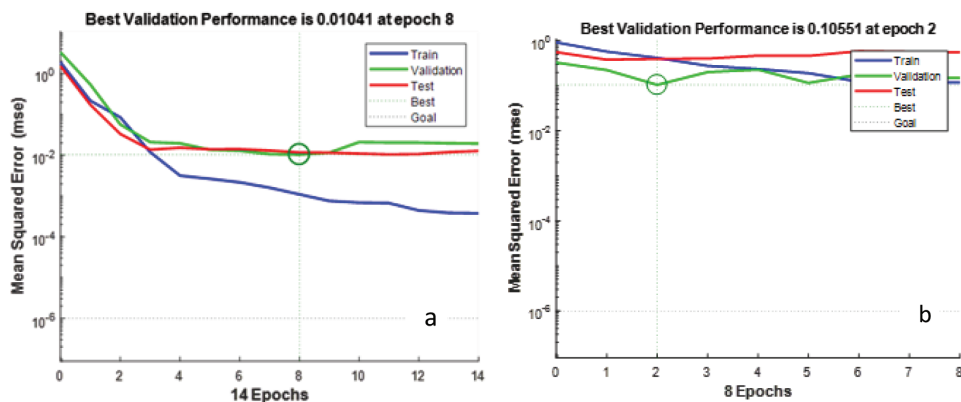
Figure 1. Performance graph between number of epochs and the mean squared error of prediction of Zn(II) ion removal for LM and SCG algorithm; number of neurons: 10. (a) LM algorithm; (b) SCG algorithm.
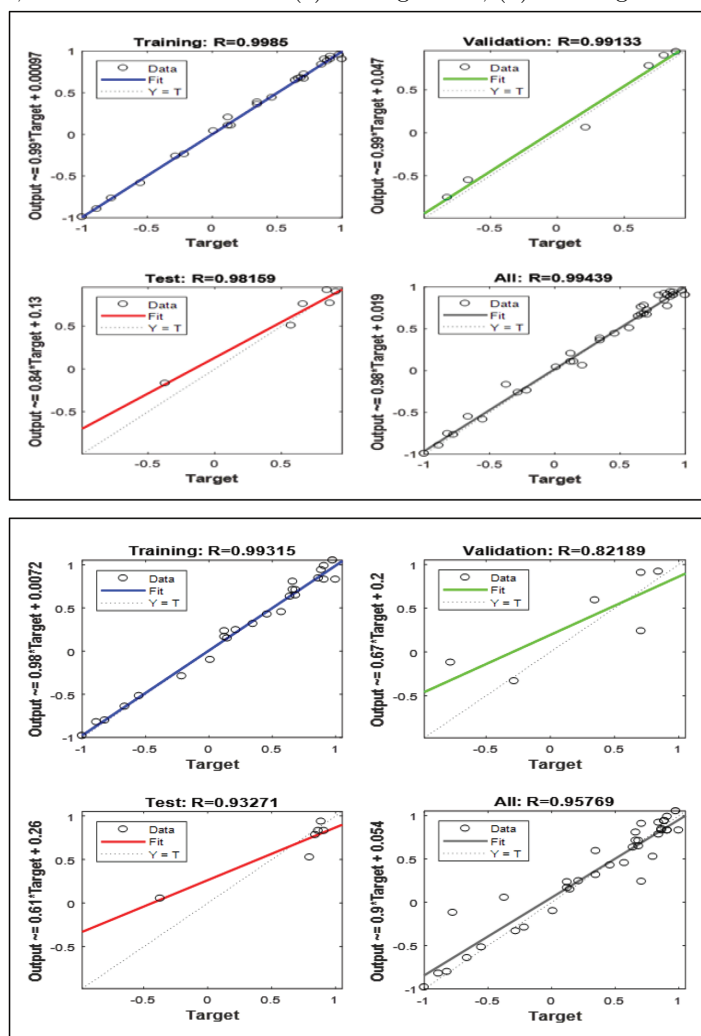

Figure 2. (a) Regression plot of Zn(II) ion removal using the correlation coefficient (R-value) for the Levenberg–Marquardt algorithm; number of neurons: 10. (b) Regression plot of Zn(II) ion removal using the correlation coefficient (R-value) for SCG algorithm; number of neurons: 10.

All these results show that the LM algorithm has better performance than the SCG algorithm with the R^2 value of 0.98. One can say that the ultimate ANN model chosen with the LM algorithm can forecast Zn(II) ions with the lowest MSE and high R^2 with 10 neurons. The level of agreement between the experimental data and ANN predictions justifies the reliability of the proposed ANN model, as shown in Figure 3.
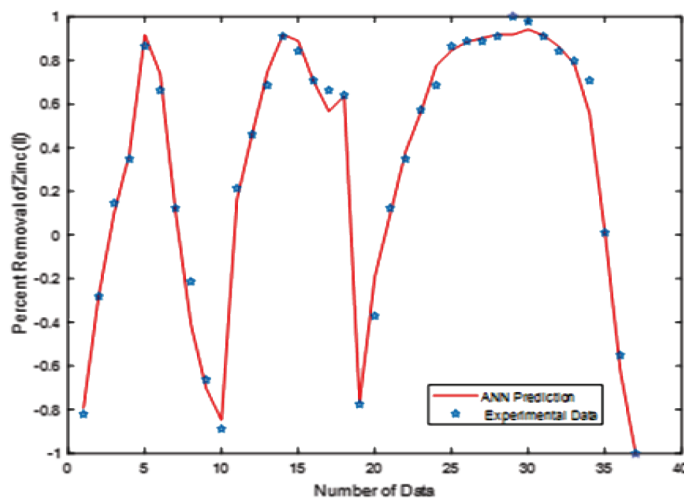

Figure 3. Comparison between experimental data of Zn(II) ions and ANN prediction as the normalized data using the LM algorithm.

2.2. Prediction of Ni(II) ion removal

In order to forecast Ni(II) ion removal, the model included three layers again: input with four operating parameters, hidden layer with various neurons, and output as removal efficiency. Maintaining the parameters at the same level, one can obtain the best training algorithm with 10, 20, and 30 neurons in the hidden layer. Two separate algorithms are used as shown in Table 2, offering two ANN models at the end. Accordingly, the lowest MSE and the highest R^2 values are obtained as about 0.075 and 0.84 using the LM (Trainlm), respectively. This provides the optimum training algorithm to estimate Ni(II) ion removal with 10 neurons.

Table 2. Different training algorithms with different neurons in the hidden layer for estimating Ni(II) removal.

Algorithm	Training function	Number of neurons (H)	R^2	MSE (the average performance)
Levenberg–Marquardt	Trainlm	10	0.84	0.075
		20	0.79	0.095
		30	0.75	0.123
Scaled conjugate gradient	trainscg	10	0.58	0.098
		20	0.44	0.128
		30	0.41	0.223

The LM algorithm being ideal to forecast Ni(II) ion removal with 10 neurons, the corresponding results for its comparison with the SCG algorithm were done with the same number neurons (10). Figure 4 contains the MSEs of the two newly developed ANN models with training, testing, and validation data results for the

algorithms. Based on this, for SCG the training error is far less than that for validation, implying that the SCG network overfits compared to LM; thus, the latter is chosen as the ideal algorithm with MSE of 0.075 for performance average. Additionally, the lowest MSE throughout validation occurs at epoch 5, which has the best validation performance, equal to 0.13819 using 10 neurons.
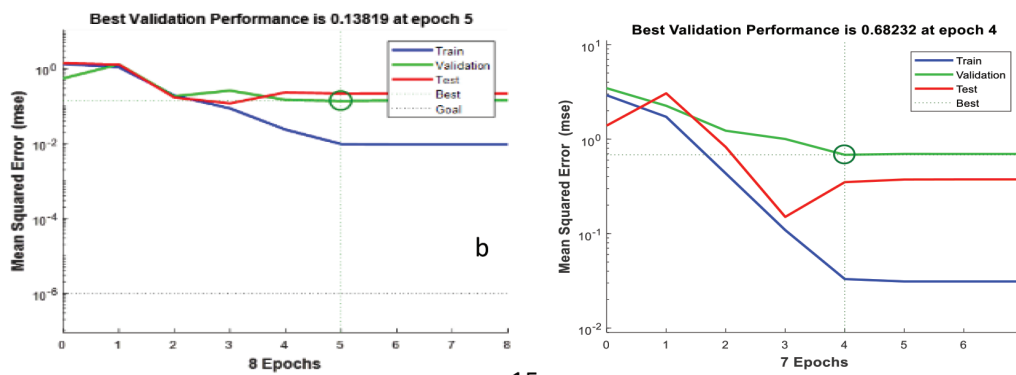

Figure 4. Performance graph between number of epochs and the mean squared error of prediction of Ni(II) ion removal for (a) LM and (b) SCG algorithm; number of neurons: 10.

The LM and SCG algorithms were compared in terms of output and experimental results using independent training, validation, testing, and all datasets with R^2. Once again, the plots for experimental and predicted results appear in Figure 5 for training, validation, and testing datasets.

The outcomes reveal that LM has better performance with better R^2 (0.84) than SCG. Therefore, the optimum ANN model with the LM algorithm can estimate Zn(II) ions with the lowest MSE and better R^2 with 10 neurons, and conformity between the experimental data and ANN predictions shows the accuracy of our ANN model (LM algorithm), as shown in Figure 6.

2.3. Prediction of Cu(II) ions removal

To forecast Cu(II) ion removal, a similar model is applied as in the previous cases for the extraction of Zn(II) and Ni(II) ions. Table 3 shows two different training algorithms. Between them, the LM (Trainlm) is the ideal algorithm to forecast Cu(II) ion removal with 10 neurons with the lowest MSE (0.002) and the highest R^2 values (0.99), and the related outcomes are compared with the SCG algorithm in terms of equal number of neurons (10).

Table 3. Different training algorithms and neurons in the hidden layer for Cu(II) ion removal.

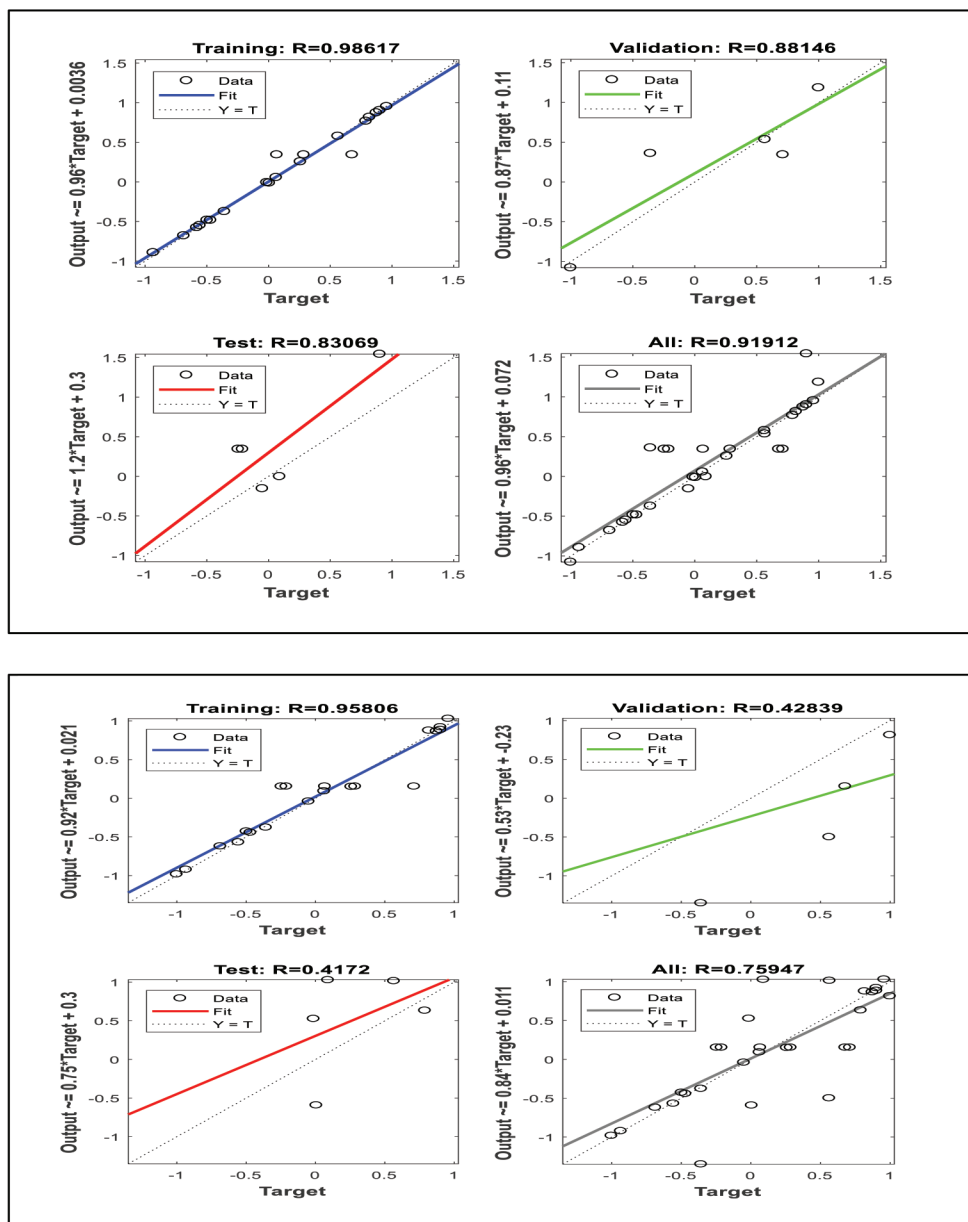| Algorithm	Training function	Number of neurons (H)	R^2	MSE (the average performance
Levenberg–Marquardt	Trainlm	10	0.99	0.002
		20	0.93	0.003
		30	0.89	0.006
Scaled conjugate gradient	trainscg	10	0.85	0.046
		20	0.60	0.069
		30	0.58	0.098

Figure 5. (a) Regression plot of Ni(II) ion removal using the correlation coefficient (R^2) for Levenberg–Marquardt algorithm; (b) regression plot of Ni(II) ion removal R^2 for SCG algorithm; number of neurons: 10.

Figure 7 highlights the MSE of the two developed ANN models using the training, testing, and validation results. Evidently, LM shows lower MSE for training, testing, and validation than SCG. Henceforth, LM is chosen as the ideal training algorithm with the average performance of 0.002. Furthermore, the lowest MSE in validation occurs at epoch 6, measured as 0.0034145 with 10 neurons as the best validation.

The LM and SCG algorithms are compared in terms of predicted output and experimental results with separate training, validation, and testing datasets and R^2. The corresponding plots appear in Figure 8.

Based on the outcomes, LM has obviously better performance in terms of R^2 (0.99) compared to that of SCG. Therefore, the chosen ANN model with the LM algorithm can forecast Cu(II) ions with the lowest MSE
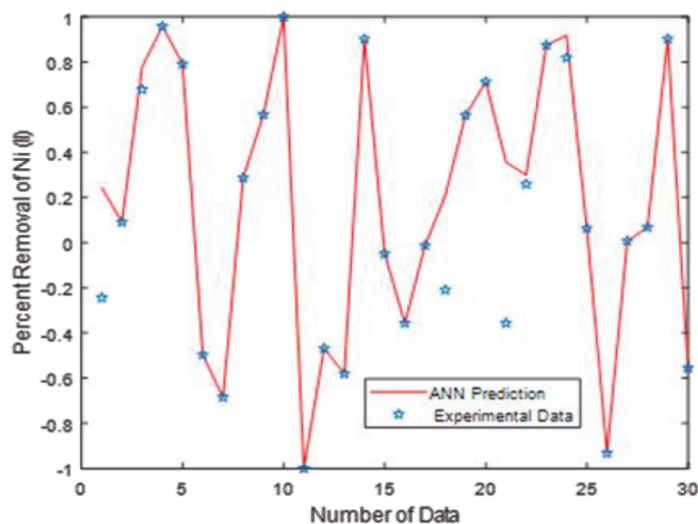
Figure 6. Comparison between experimental data of Ni(II) ions and ANN prediction as the normalized data.
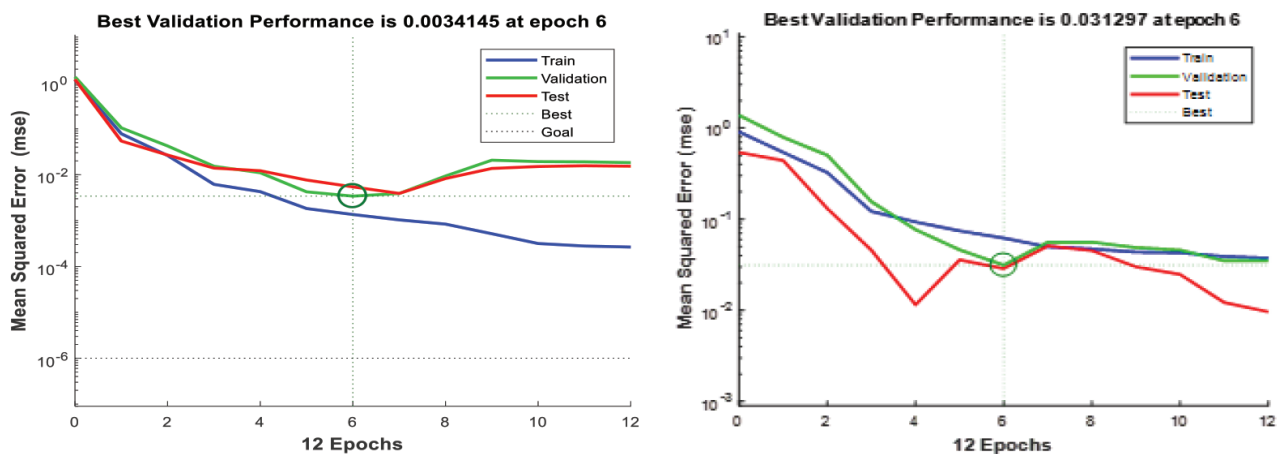

Figure 7. Performance graph between number of epochs and the mean squared error of prediction of Cu(II) ion removal for LM (a) and SCG algorithm (b); number of neurons: 10.

and the highest R^2 with 10 neurons. Based on the conformity of experimental data and ANN predictions, the accuracy of our ANN model (LM algorithm) can be depicted in Figure 9.

As a result, one can say that the efficiency of the ANN in the second case study was lower compared to the first and third case studies, which may be attributed to the small sample size of the experimental data in the second case study causing a lower learning performance for the ANN model [35]. However, the best correlation coefficients and MSE obtained from the ANN model are summarized in Table 4.

3. Materials and Methods

3.1. Data Collection

For the purpose of the present study, we refer to data formerly obtained in three other works. The extraction of zinc(II) ions via the bath adsorption process on magnetic Fe_3O_4 nanoparticles (case study 1) from wastewater has already been carried out and reported on by Shirsath and Shirivastava [9], who concluded that the process
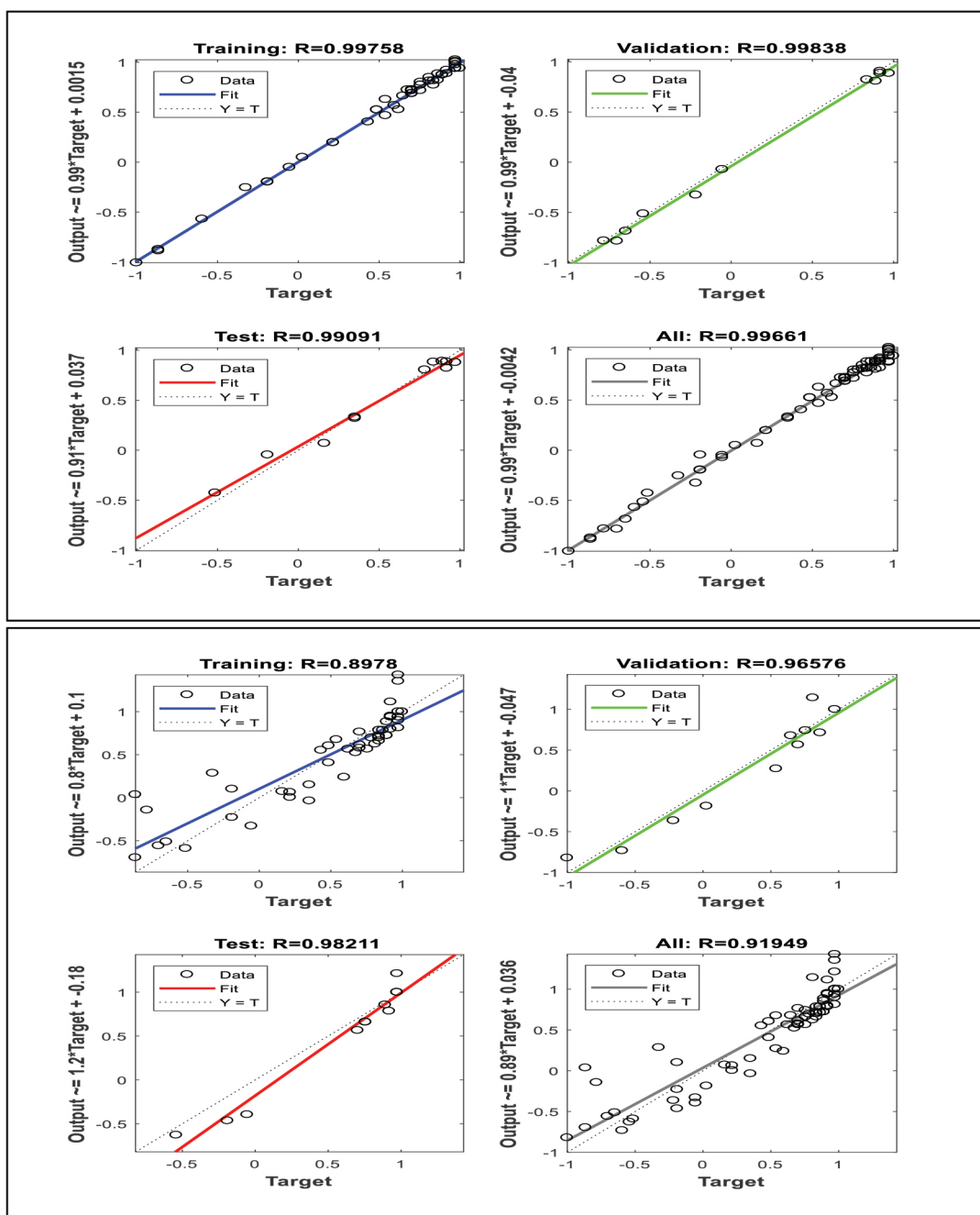
Figure 8. (a) Regression plot of Cu(II) ion removal using R^2 for LM Algorithm; (b) regression plot of Cu(II) ion removal using R^2 for SCG algorithm; number of neurons: 10.

could be influenced by different factors, such as the pH level, contact time (t), adsorbent dose, and initial concentration of Zn(II) ions. See Table 5 for a summary of the details. The experimental results showed that the pseudo second-order kinetic model adapts well to the equilibrium data to adsorb heavy metal by magnetic nanoparticles. In this case, the operating variables (pH, time, initial, Zn(II) ion, nanoadsorbent concentration) are applied as inputs while percent removal of Zn(II) is applied as output.

In the second case study, the optimization process parameters for the ablation of Ni(II), Cu(II), and Co(II)
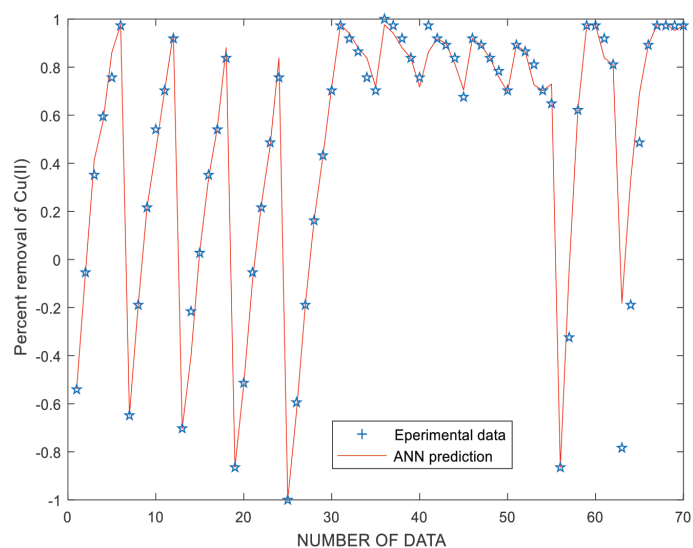
Figure 9. Comparison between experimental data Cu(II) ions and ANN prediction as the normalized data

Table 4. Comparison of ANN model in three cases.

Case study	MSE (the average performance)	Coefficient of determination (R^2)
Experimental data for removal of Zn(II) ions	0.008	0.98
Experimental data for removal of Ni(II)	0.075	0.84
Experimental data for removal of Cu(II)	0.002	0.99

Table 5. The experimental data for removal of Zn(II) ions.

pH	Time (s)	Initial Zn(II) ion concentration (mg L^{-1})	Nanoadsorbent concentration (mg L^{-1})	Percent removal of Zn(II) (% R)
1–10	0–225	20–100	1–2.5	10–99

doped alginate-coated chitosan nanoparticles (Alg-CS-NPs) from industrial effluents have also been shown to fit well into the pseudo second-order model proposed by Esmaeili and Khoshnevisan [36], with experimental data evidently establishing that the percent removal of Ni(II) ion rises up to 94%, as depicted in Table 6. In this case, the operating variables (pH, time, initial Ni(II) ion, biomass dose) are applied as inputs while percent removal of Ni(II) is applied as output.

Table 6. The experimental data for removal of nickel(II) ions.

pH	Contact time (s)	Initial concentration of Ni(II) (mg L^{-1})	Biomass dose (g)	Percent removal of Ni(II) (% R)
1–9	0–120	10–90	0.1–0.9	23–94.48

In the third case study, the extraction of Cu(II) ions using magnetic chitosan-coated magnetic nanoparticles from contaminated water has been already carried out and reported on by Neeraj et al. [10], who concluded

that the process can be influenced by different factors, such as the pH level, contact time (t), adsorbent dose, temperature, and initial concentration of Cu(II) ions. The experimental results show that the pseudo second-order kinetic model adapts well to the equilibrium data to adsorb heavy metal by magnetic nanoparticles. In this case study, the operating variables (initial Cu(II) concentration, contact time, adsorbent dose, temperature, and pH level) are applied as inputs while percent removal of Cu(II) is applied as output as depicted in Table 7.

Table 7. The experimental data for removal of Cu(II) ions.

Temperature (K)	pH	Contact time (min)	Initial concentration of Cu(II) (mg L^{-1})	Adsorbent dose (g L^{-1})	Percent removal of Cu(II) (% R)
303–335	2–8	0–30	100–500	0.4–3.2	25–99

3.2. ANN model development

ANN or linkage systems have been used in many fields, such as machine learning, computer science, and other research areas, focusing on a broad set of basic units known as neurons, with slight similarities to axons in a biological brain. In these networks, the bonds among the neurons transmit stimulating signals with different forces. In the same way, in wastewater treatment, scientific and practical models utilizing numerical computational techniques such as ANNs can significantly help to improve the process as a whole through adsorption systems.

The origin of ANNs goes back to the principle concept of artificial intelligence in an attempt to copy the workings of the human brain and the nervous system. For that purpose, learning and memorization processes were simulated using a set of mathematical correlations. As the outcome, the learning process by ANNs relies on a set of input variables and matching output applied to specify the relationship between the variables.

There are three different layers in ANNs; these are the input layer, hidden layer, and output layer. A multilayer ANN is used to generate models of a system using nonlinear combinations of input variables. The following measures must be taken for the conception and construction of a neural network model: the recording of the input, the output, and sample datasets, as well as the design, training, and verification of the neural network. Overall, pairs of input-output datasets are used as described in the three case studies selected (Tables 5–7).

After receiving the sample dataset, the operating parameters are chosen as input variables, and then heavy metal removal efficiency is chosen as an output variable. The dataset is randomly divided into test subsets, such as 70% for training, 15% for validation, and 15% to test the network. The ANN proposed in this study is a feedforward backpropagation network model with the function of a tangent sigmoid transfer (tansig) between the input and hidden layer with a range of neurons and function of a linear transfer (purelin) at the output node using an embedded MATLAB code (see supporting information). The training data form the biggest set and are used by the neural network to learn patterns presented in the data by updating the network weights. The testing data are used to evaluate the quality of the network. The final check of the performance and generalization ability of the trained network is made using validation data.

To compute the weight of a neuron in the hidden layer, Eq. (1) can be used:

$$W_b = \sum_{a=1}^{k} w_{ab} x_a \quad , \tag{1}$$

where k is the number of neurons in the input layer, w_{ab} is the connection weight between neuron a in the input layer and neuron b in the hidden layer, and x_a is the value of neuron a in the input layer. Similarly, the weight of a neuron in the output layer can be calculated as follows:

$$W_c = \sum\nolimits_{b=1}^{z} w_{bc} x_b \quad , \tag{2}$$

where z is the number of neurons in the hidden layer, w_{bc} represents the connection weight between neuron b in the hidden layer and neuron c in the output layer, and x_b stands for the value of neuron b in the hidden layer. The weight of the neuron in the hidden layer or output layer was used in the activation function, which produced a predicted output by Eq. (3):

$$y = f(W + B) \tag{3}$$

where y, f, W, and B are the output, activation function, weight, and bias in the hidden layer or output layer, respectively

Since, according to Bishop's study [37], more than one hidden layer is often not essential, our model has only one hidden layer. In this study, the training parameters are as specified in Table 8. Two training functions were examined to choose the one that best fit the collected data. All these parameters are fixed and checked for each training function with 10, 20, and 30 neurons in the hidden layer.

Table 8. Training parameters.

Training parameters	Range
Maximum number of validation checks	6
Maximum number of training epochs	1000
Performance goal	1e-6
Learning rate	0.01
Epochs between displays	25
Maximum time to train in seconds	infinity
Minimum performance gradient	1e-10

The final ANN structure has some of the operating variables as input with one hidden layer and the efficiency of heavy metal removal as the output of the output layer (Figure 10). Normalization is required so that all of the data are in a comparable range, thus normalizing the input and output datasets in the domain [–1, 1] prior to the training process. A trial-and-error method is applied to obtain the most favorable conditions, such as the number of hidden layers and the number of neurons in the hidden layer [38]. In general, different methods have been used and reported to obtain the optimum number of neutrons in the literature, such as the Taguchi method [39], genetic algorithm [40], k-fold validation [41], or design of experiment [42]. The trial-and-error approach is one of the most frequently applied methods [40–50] by researchers. One can use any of them. Herein, the trial-and-error method is applied to determine the optimal neuron numbers, where the lowest mean error is provided [46,48,49].

One hidden layer is used during the modeling of the process in order to simplify the system and consequently to reduce the time and cost of the simulation.

Researchers have proved that any continuous function would be convergent in the case of using one hidden layer in the network as long as the freedom level is enough [51]. In order to construct the appropriate
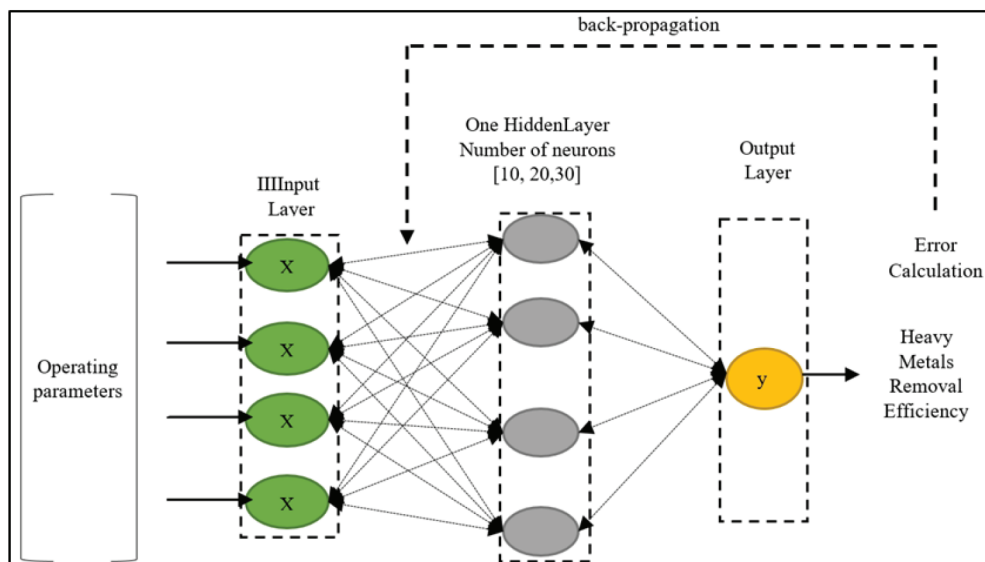
Figure 10. A single hidden layer neural network model with backpropagation for predicting adsorption efficiency.

network architecture, several attempts were made for each group until the most suitable learning rate; number of neurons and minimum power gradient were all optimized.

Each model is tested by measuring the coefficient of determination (R^2) and the MSE values between the predicted and actual figures in each attempt to determine the best structure with the lowest possible values. Such parameters show the most ideal training algorithm and the right number of neurons in the hidden layer. MSE is applied to show error related to the model as follows:

$$MSE == \frac{1}{N} \sum_{j=1}^{N} \left(y_{P,j} - y_{T,j} \right)^n \quad , \tag{4}$$

where Yp,j and yT,j are respectively the model prediction and the corresponding target of the variable and n is the number of the corresponding targets.

Accordingly, R^2 is determined as the square of the correlation (R-value) to reveal the degree of variability for the model between the input and the intended values for training, validation, and test datasets.

4. Conclusions

Recent studies showed that nanomaterials are suitable for the elimination of heavy metals in processes routinely carried out in wastewater treatment plants due to their unique properties and high efficacy. Although nanomaterials such as CNTs, nanometal oxides, and other organic sorbents have been successfully used for extracting heavy metal ions in wastewater, several problems still remain. In this respect, wastewater treatment on a large scale is the most important one.

ANNs provide a range of powerful new techniques for solving problems in sensor data analysis, fault detection, process identification, and control and have been used in a diverse range of chemical engineering applications.

The key component of the present paper is modeling based on an ANN, inspired by the neurological systems functioning in the human brain and attempting to imitate them for use in wastewater treatment.

Although these models are very different from those applied about the brain, they show an unanticipated degree of success in the adsorption process, which included the modeling of nonlinear systems.

Therefore, for the control of adsorption process, the ANN technique was applied to compare the predictions obtained with the experimental data extracted from three case studies in the field. The trial-and-error method was preferred to determine the best conditions and functions for artificial neural networks. Here, according to the adsorption tests carried out using different variables, the main purpose was to come up with a unique model to offer accurate forecasts of adsorption efficiency (AE %) related to zinc(II), nickel(II), and copper(II) ions. The ANN model suggested is based on feedforward backpropagation with the function of a tangent sigmoid transfer (tansig) between the input and the hidden layer with a range of neurons and function of a linear transfer (purelin) at the output node and applying an embedded MATLAB code. Two training functions are tested to select the best one fitting the data. All factors are kept constant and tested in each training function with 10, 20, and 30 neurons within the hidden layer.

Two functions were tested with the range of neurons upon the comparison of predicted results versus the experimental data based on the MSE and R^2. Accordingly, the best ANN model was chosen using the LM algorithm with ten neurons and was capable of forecasting AE % of zinc(II), nickel(II), and copper(II) ions with MSE of 0.008, 0.075, and 0.002 and with R^2 of 0.98, 0.84, and 0.99, respectively.

Thus, it is can be stated with relative certainty that the proposed ANN model can be used as a tool for the removal of heavy metals by magnetic nanoparticles in wastewater treatment plants. It is obvious that using nanotechnology is ideal for the removal of metal ions in sewage treatment. In this respect, theoretical architectures with ANNs are very important to determine the best model process for optimum results.

References

1. Fergusson JE. Heavy Elements: Chemistry, Environmental Impact and Health Effects. Oxford, UK: Pergamon, 1990.

2. Dean JG, Bosqui FL, Lannouette KL. Removing heavy metals from waste water. Environmental Science Technology 1977; 6: 518-524.

3. Pai TY, Wang SC, Lo HM, Chiang CF, Liu MH et al. Novel modeling concept for evaluating the effects of cadmium and copper on heterotrophic growth and lysis rates in activated sludge process. Hazardous Materials 2009; 166: 200-206.

4. Juliastuti S, Baeyens J, Creemers C, Bixio D, Lodewyckx E. The inhibitory effects of heavy metals and organic compounds on the net maximum specific growth rate of the autotrophic biomass in activated sludge. Hazardous Materials 2003; 100: 271-283.

5. Wang X, Guo Y, Yang L, Han M, Zhao J et al. Nanomaterials as sorbents to remove heavy metal ions in wastewater treatment. Journal of Environmental Analysis and Toxicology 2012; 2: 100-154.

6. Tyagi I, Gupta VK, Sadegh H, Ghoshekandi RS, Makhlouf ASH. Nanoparticles as adsorbent; a positive approach for removal of noxious metal ions. Science, Technology and Development 2017; 34: 195-214.

7. Rosenberg E. Heavy metals in water: presence, removal and safety. Johnson Matthey Technology 2015; 59: 293-297.

8. Gautam RK, Chattopadhyaya MC. Advanced Nanomaterials for Wastewater Remediation. New York, NY, USA: CRC Press, 2016.

9. Shirsath DS, Shirivastava VS. Adsorptive removal of heavy metals by magnetic nanoadsorbent: an equilibrium and thermodynamic study. Applied Nanoscience 2015; 5: 927-935.

10. Neeraj G, Krishnan S, Kumar PS, Shriaishvarya KR, Kumar VV. Performance study on sequestration of copper ions from contaminated water using newly synthesized high effective chitosan coated magnetic nanoparticles. Journal of Molecular Liquids 2016; 214: 335-346.

11. Ilankoon N. Use of iron oxide magnetic nanosorbents for Cr (VI) removal from aqueous solutions: a review. Journal of Engineering Research and Applications 2014; 4: 55-63.

12. Ghaedi AM, Ghaedi M, Pouranfard AR, Ansari A, Avazzadeh Z et al. Adsorption of Triamterene on multi-walled and single-walled carbon nanotubes: artificial neural network modeling and genetic algorithm optimization. Journal of Molecular Liquids 2016; 216: 654-665.

13. Elmolla ES, Chaudhuri M, Eltoukhy MM. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. Journal of Hazardous Materials 2010; 179 (1-3): 127-134.

14. Zhang Y, Pan B. Modeling batch and column phosphate removal by hydrated ferric oxide-based nanocomposite using response surface methodology and artificial neural network. Chemical Engineering Journal 2014; 249: 111-120.

15. Assefi P, Ghaedi M, Ansari A, Habibi MH, Momeni MS. Artificial neural network optimization for removal of hazardous dye Eosin Y from aqueous solution using Co_2O_3-NP-AC: Isotherm and kinetics study. Journal of Industrial and Engineering Chemistry 2014; 20 (5): 2905-2913.

16. Zhu J, Zurcher J, Rao M, Meng MQH. An online wastewater quality predication system based on a time-delay neural network. Engineering Applications of Artificial Intelligence 1998; 11 (6): 747-758.

17. Gontarski CA, Rodrigues PR, Mori M, Prenem LF. Simulation of an industrial wastewater treatment plant using artificial neural networks. Computers & Chemical Engineering 2000; 24 (2-7): 1719-1723.

18. Pai TY, Tsai YP, Lo HM, Tsai C, Lin C. Grey and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent. Computers & Chemical Engineering 2007; 31 (10): 1272-1281.

19. Sahoo GB, Ray C. Predicting flux decline in crossflow membranes using artificial neural networks and genetic algorithms. Journal of Membrane Science 2006; 283 (1-2): 147- 157.

20. Chen H, Kim AS. Prediction of permeate flux decline in crossflow membrane filtration of colloidal suspension: a radial basis function neural network approach. Desalination 2006; 192 (1-3): 415-428.

21. Shetty GR, Shankararaman C. Predicting membrane fouling during municipal drinking water nanofiltration using artificial neural networks. Journal of Membrane Science 2003; 217 (1-2): 69-86.

22. Guadix A, Zapata JE, Almecija MC, Guadix EM. Predicting the flux decline in milk cross-flow ceramic ultrafiltration by artificial neural networks. Desalination 2010; 250 (3): 1118-1120.

23. Libotean D, Giralt J, Giralt F, Rallo R, Wolfe T et al. Neural network approach for modeling the performance of reverse osmosis membrane desalting. Journal of Membrane Science 2009; 326 (2): 408-419.

24. Prakash N, Manikandan SA, Govindarajan L, Vijayagopal V. Prediction of biosorption efficiency for the removal of copper (II) using artificial neural networks. Journal of Hazardous Materials 2008; 152 (3): 1268-1275.

25. Fagundes-Klen MR, Ferri P, Martins TD, Tavares CRG, Silva EA. Equilibrium study of the binary mixture of cadmium–zinc ions biosorption by the *Sargassum filipendula* species using adsorption isotherms models and neural network. Biochemical Engineering Journal 2007; 34 (2): 136-146.

26. Yetilmezsoy K, Demirel S. Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (*Pistacia vera* L.) shells. Journal of Hazardous Materials 2008; 153 (3): 1288-1300.

27. Sadrzadeh M, Mohammadi T, Ivakpour J, Kasiri N. Neural network modeling of Pb^{2+} removal from wastewater using electrodialysis. Chemical Engineering and Processing Process Intensification 2009; 48 (8): 1371-1381.

28. Mandal S, Mahapatra SS, Sahu MK, Patel RK. Artificial neural network modelling of As(III) removal from water by novel hybrid material. Process Safety and Environmental Protection 2015; 93: 249-64.

29. Kashaninejad M, Dehghani AA, Kashiri M. Modeling of wheat soaking using two artificial neural networks (MLP and RBF). Journal of Food Engineering 2009; 91 (3): 602-607.

30. Singh KP, Basant A, Malik A, Jain G. Artificial neural network modeling of the river water quality-A case study. Ecological Modelling 2009; 220 (6): 888-895.

31. Balku S. Magnetic removal of iron from fireclays: optimum conditions. Industrial Ceramics 2009; 29: 83-89.

32. Fanaie VR, Karrabi M, Amin MM. Biosorption of 4-chlorophenol by dried anaerobic digested sludge: artificial neural network modeling, equilibrium isotherm, and kinetic study. International Journal of Environmental Science and Technology 2017; 14: 37-48.

33. Khataee A, Khani A. Modeling of nitrate adsorption on granular activated carbon (GAC) using artificial neural network (ANN). International Journal of Chemical Reactor Engineering 2009; 7: 1-16.

34. Allahkarami E, Igder A, Fazlavi A, Rezai B. Prediction of Co (II) and Ni (II) ions removal from wastewater using artificial neural network and multiple regression models. Physicochemical Problems of Mineral Processing 2017; 53 (2):1105-1118.

35. Esmaeili A, Khoshnevisan N. Optimization of process parameters for removal of heavy metals by biomass of Cu and Co-doped alginate-coated chitosan nanoparticles. Bioresource Technology 2016; 218: 650-658.

36. Bishop CM. Neural Networks for Pattern Recognition. New York, NY, USA: Oxford University Press, 1995.

37. Maier HR, Morgan N, Chow CW. Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters. Environmental Modelling & Software 2004; 19: 485-494.

38. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Networks 1989; 2: 359-366.

39. Madić MJ, Radovanović MR. Optimal selection of ANN training and architectural parameters using Taguchi method: a case study. FME Transactions 2011; 39 (2): 79-86.

40. Panchal FS, Panchal M. Review on methods of selecting number of hidden nodes in artificial neural network. International Journal of Computer Science and Mobile Computing 2014; 3 (11): 455-464.

41. Abbasi H, Seyedain ASM, Mohammadifar MA, Emam-Djomeh Z. Comparison of trial and error and genetic algorithm in neural network development for estimating farinograph properties of wheat-flour dough. Nutrition and Food Sciences Research 2015; 2 (3): 29-38.

42. Laosiritaworn W, Chotchaithanakorn N. Artificial neural networks parameters optimization with design of experiments: an application in ferromagnetic materials modeling. Chiang Mai Journal of Science 2009; 36 (1): 83-91.

43. Gazi, M, Oladipo AA, Ojoro ZE, Gulcan HO. High-performance nanocatalyst for adsorptive and photo-assisted Fenton-like degradation of phenol: modeling using artificial neural networks. Chemical Engineering Communications 2017; 204 (7): 729-738.

44. Oladipo AA, Vaziri R, Abureesh MA. Highly robust $AgIO_3$/MIL-53 (Fe) nanohybrid composites for degradation of organophosphorus pesticides in single and binary systems: application of artificial neural networks modelling. Journal of the Taiwan Institute of Chemical Engineers 2018; 83: 133-142.

45. Muhsin WA, Zhang J, Lee J. Modelling and optimization of a crude oil hydrotreating process using neural networks. Chemical Engineering Transactions 2016; 52: 211-216.

46. Cenk N, Budak G, Dayanik S, Sabuncuoglu I. Artificial neural network modeling and simulation of in-vitro nanoparticle-cell interactions. Journal of Computational and Theoretical Nanoscience 2014; 11 (1): 272-282.

47. Xu S, Chen L. A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining. In: ICITA 2008 5th International Conference on Information Technology and Applications; 2008. pp. 683-686.

48. Pourzangbar A, Saber A, Yeganeh-Bakhtiary A, Ahari LR. Predicting scour depth at seawalls using GP and ANNs. Journal of Hydroinformatics 2017; 19 (3): 349-363.

49. Ahmed FE. Artificial neural networks for diagnosis and survival prediction in colon cancer. Molecular Cancer 2005; 4 (1): 29-40.

50. Shojaee SA, Hezave AZ, Lashkarbolooki M, Shafipour ZS. Prediction of the binary density of the ionic liquids+ water using back-propagated feed forward artificial neural network. Chemical Industry & Chemical Engineering Quarterly 2014; 20 (3): 325-338.

51. Mao R, Zhu H, Zhang L, Chen A. A new method to assist small data set neural network learning. In: ISDA'06 Sixth International Conference on Intelligent Systems Design and Application; 2006. pp. 17-22.

Supporting information. MATLAB codes.

```
%Solve an Input-Output Fitting problem with a Neural Network
% the code generated by MATLAB
% Created 10-Jun-2018 11:38:54
%
% variables are defined:
%
% p - input data.
% T - target data.
% first case
x=load('first.dat');
P=[x(1:37,1:4)];
T=[x(1:37,5)];
% secod case
%x=load('second.dat');
%P=[x(1:30,1:4)];
%T=[x(1:30,5)];
%Third case
%x=load('third .dat');
%P=[x(1:70,1:5)];
%T=[x(1:70,6)];
% normaliztion data
[pn,minp,maxp,tn,mint,maxt]=premnmx(P',T');
s = pn;
t = tn;
% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
% trainFcn ='trainscg' ; % Scaled Conjugate Gradient
trainFcn = 'trainlm'; % Levenberg-Marquardt backpropagation.
% Create a Fitting Network
% hiddenLayerSize = 20
% hiddenLayerSize = 30
hiddenLayerSize = 10;
net = fitnet(hiddenLayerSize,trainFcn);
% Selection of internal transfer functions
net.layers{1}.transferFcn = 'tansig';
net.layers{2}.transferFcn = 'purelin';
```

```
% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
net.trainParam.epochs=1000;
net.trainParam.lr=0.01;
net.trainParam.goal =1e-6
net.trainParam.min_grad=1e-10
net.trainParam.show=25
net.trainParam.time =inf
% Train the Network
[net,tr] = train(net,s,t);
% Test the Network
y = net(s);
e = gsubtract(t,y);
performance = perform(net,t,y)
% View the Network
view(net)
% Plots
figure, plotperform(tr)
figure, plottrainstate(tr)
figure, ploterrhist(e)
figure, plotregression(t,y)
figure, plotfit(net,s,t)
%weights
IW = net.IW{1,1} %weights for the connection from the first input to the first layer
b1 = net.b{1}%the bias values for the first layer
b2 = net.b{2}%the bias values for the second layer
LW = net.LW{2,1}%weights for the connection from the first layer to the second layer
plot ( 1:number of data , y','r',1: number of data ,tn','p')
```