Advances in Colloid and Interface Science 245 (2017) 20–39

Contents lists available at ScienceDirect

Advances in Colloid and Interface Science

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

Historical Perspective

Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review

Abdol Mohammad Ghaedi*, Azam Vafaei

Department of Chemistry, Gachsaran Branch, Islamic Azad University, P.O. Box 75818-63876, Gachsaran, Iran

ARTICLE INFO

Keywords:
Adsorption
Removal
Dye
ANN
ANFIS
SVM

ABSTRACT

Artificial neural networks (ANNs) have been widely applied for the prediction of dye adsorption during the last decade. In this paper, the applications of ANN methods, namely multilayer feedforward neural networks (MLFNN), support vector machine (SVM), and adaptive neuro fuzzy inference system (ANFIS) for adsorption of dyes are reviewed. The reported researches on adsorption of dyes are classified into four major categories, such as (i) MLFNN, (ii) ANFIS, (iii) SVM and (iv) hybrid with genetic algorithm (GA) and particle swarm optimization (PSO). Most of these papers are discussed. The further research needs in this field are suggested. These ANNs models are obtaining popularity as approaches, which can be successfully employed for the adsorption of dyes with acceptable accuracy.

1. Introduction

In recent years, discharge of waste products (mainly the synthetic dyes) from different industries leading to aquatic and environmental pollution is a serious global problem of great concern [1–4]. Due to toxicity of dyes, it is necessary to remove them from wastewater before them discharge to the natural environment. Various physico-chemical and biological wastewater treatment technologies such as aerobic or anaerobic bioremediation systems, oxidative process, Fenton's reagent, ozonation, electrochemical destruction, electrokinetic coagulation, ion exchange, irradiation, membrane filtration, decolorization by white-rot fungi and adsorption have been used for the removal of dyes from industrials effluents [5], as each method has its own advantages and disadvantage. Among all of these technologies, the adsorption is one of the most effective methods for wastewater treatment [5–7].

Modeling and simulation of a process is obtaining data about how a process will behave without doing practical experiments. The application of modeling for many processes in science and engineering is well accepted. The removal percentage is a key parameter for adsorption processes but is not available for most of the operating conditions due to time-consuming and expensive experimental test. Therefore, it is essential to predict removal (%) using several operating variables. Adsorption process is complex process; therefore, due to complexity of the relationship between input parameters and output, it is difficult to be modeled using statistical models. Computational intelligence models are often more flexible compared to statistical models when modeling complex datasets with possible nonlinearities or missing data [8]. Recently, powerful artificial intelligence (AI) prediction method, such as adaptive neuro fuzzy inference system (ANFIS), random forest (RF), least square support vector regression (LSSVR), and artificial neural network (ANN) in modeling adsorption process have been successfully used [9–15]. ANN inspired by the biological nervous processing, it can be employed to solve and model a many of complex engineering systems due to its simplicity, robustness, reliability and nonlinearity. ANN model can be learned from experimental data to solve of the complex non-linear, multi-dimensional functional relationships without any prior assumptions about their nature.

The previously cited reviews of adsorption removal and engineering applications of ANN confirmed that, there is no specific review reported on the usage of ANN for adsorption removal process. The two main aims of the this review are as follows: (i) to summarize the investigations related to adsorption of dyes done by ANNs models, (ii) to present the further research needs of ANNs for the adsorption of dyes.

2. ANN models

The ANN models can create the nonlinear relationship between independent and dependent variables based on a set of experimental data. The time distribution of the number of researches reported using the ANN models removal of dyes using adsorption is indicated in Fig. 1. From Fig. 1, it is clear that the number of studies reported in this area is gradually enhancing during last decade. Modeling of dyes adsorption using ANN models is becoming common with researchers working in this area. The reported studies for adsorption of dyes are categorized

* Corresponding author.
E-mail address: abm_ghaedi@yahoo.com (A.M. Ghaedi).

http://dx.doi.org/10.1016/j.cis.2017.04.015

Available online 26 April 2017

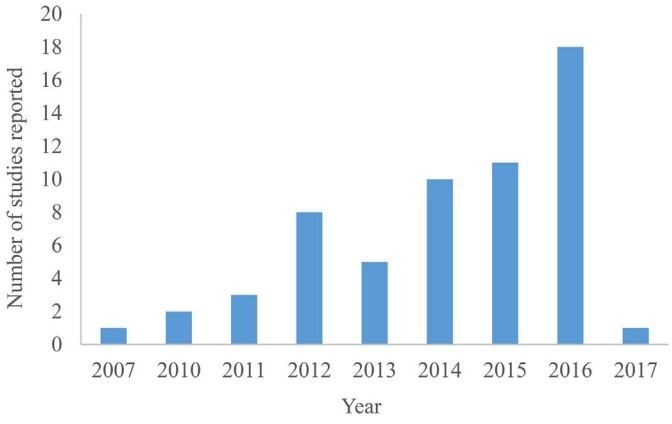

Fig. 1. Applications of ANN models for dye adsorption reported in this review.

into four subsections based on the type of ANN model as follows: (i) MLFNN, (ii) ANFIS, (iii) SVR, and (iv) hybrid models.

2.1. Multilayer feedforward neural networks

Artificial neural network was introduced by McCulloch and Pitts [16] act like a "black box" model derived from a simplified concept of the human brain. It has been widely accepted as an approach for prediction, control systems, classification, optimization and decision-making in various fields. ANNs are a parallel distribution processing method combined of neurons and weights, and are based on the principle that a highly connected system of simple processing elements that can train complex interrelationships between input(s) and output (s) variables. Neurons are single processing elements, which linked to neurons in the next layer and therefore forming different types of ANNs. Multi-layer perceptron (MLP) is one of the most well-known types of ANNs, of which consists of an input layer, an output layer and one or more hidden or intermediate layers. The commonly employed ANN structures for adsorption are MLFFN. The basic information on ANN structures is very much available in detail in the many literatures [17–19].

2.1.1. Important parameters of the ANN model

The ANN parameters including, the number of hidden layers, the number of neurons in each layer, network structure, transfer function, training algorithm, momentum factor and training rate are to be optimized for constructing an ANN model. The number of neurons in the input layer is usually equal to the number of parameters such as temperature, contact time, amount of adsorbent, initial concentration of feed, pH, and particle size of adsorbent that affect the performance of adsorption process. The number of neurons in output layer is usually equal to the number of variables considered for forecasting the adsorption performance (usually removal percentage). A layer(s) of processing neurons connects the input and the output layers are named as the hidden layer(s). The number of hidden layers and the number of hidden neurons, momentum factor and training rate values can be changed to obtain the results with required accuracy.

2.1.2. Training process of the ANN model

Training the networks is the process of adjusting the weights and biases between the processing neurons using a suitable learning method and a set of matched input-output data. The learning methods are divided into unsupervised, evolutionary, reinforcement and supervised learning. The outputs are the dependent parameters that the network generates for the matched inputs, which are then compared to the desired or correct outputs. The training process progresses until the network outputs correspond to the desired outputs. Modifying the weights and biases shall decrease the error between the network output and the correct output. The training process is stopped automatically when all the errors are within the required tolerance or the maximum epochs is reached [20,21].

The multilayer feed forward neural network structure is the most widely used neural network. A back-propagation (BP) algorithm as one of the strongest learning algorithms is a gradient descent algorithm that can be employed to learn these multilayer feed forward networks with differentiable transfer functions. As mentioned above, the ANNs are built with three layers: input, output and hidden layers. Each layer has different numbers of neurons. The network performance is characterized by the connections between these neurons (elements). The relationship between the input vector X_i^n and output vector Y_k^{n+1} of this element can be defined as follows:

$$Y_k^{n+1} = f\left(\sum_i W_{ki}^n X_i^n\right) \tag{1}$$

where f(x) is the linear or nonlinear transfer function, and Y_k^{n+1} is output of unit k in the nth layer, W_{ki}^n is a weight from unit i in nth layer to unit k in (n + 1)th layer. An input vector is indicated to the units of the input layer. Elements in the next layer calculate a weighted sum of the input, and output the result of a nonlinear function to the sum. The learning method is based on a gradient search, with a criterion of errors between the values of network output and desired output:

$$E = \sum_{n=1}^{N} (O_n - O_d)^2 \tag{2}$$

which, E is the total sum squared error of all data in the training set, in which O_n is the network output for the nth data and O_d is the desired output. In the training process, the weights of all the connecting nodes are modified until the required error level is obtained or the maximum number of iteration is reached. The learning algorithm employed here for the weights is

$$\Delta W_{ki}^n(m+1) = -\eta \frac{\partial E}{\partial W_{ki}^n} + \mu \Delta W_{ki}^n(m) \tag{3}$$

where $\Delta W_{ki}^n(m)$ is the correction of the weight at the mth learning step, η is the training rate (a small parameter to alter the correction each time), and μ is the momentum factor (decrease an oscillation and helps quick convergence). Network learning adjusts using suitable values of these parameters [22].

As described before, each input is weighted with a suitable value, and then the sum of the weighted inputs and the bias constructs the input to the transfer (activation) function f. Nodes (neurons) can employ any differentiable transfer function f to produce their output [23]. Three transfer functions that are the most commonly utilized transfer function for multilayer networks are described below. Multi-layer networks often apply the log-sigmoid (logsig) transfer function.

$$y = logsig(x) = \frac{1}{(1 + exp(-x))} \tag{4}$$

The function logsig produces outputs between 0 and 1 as the node's net input goes from negative to positive infinity. Alternatively, the tan-sigmoid as transfer function can be used.

$$y = tansig(x) = \frac{2}{(1 + exp(-2^*x))} - 1 \tag{5}$$

Sigmoid outputs nodes are often employed for pattern recognition problems, while linear (purelin) transfer function is applied for function fitting problems shows the purelin transfer function.

$$y = purelin(x) = x \tag{6}$$

The steps of ANN training processes used for the adsorption shows in Fig. 2. The steps involved in this processes are summarized below.

1. The data set have to be collected from experiments. Before this, the

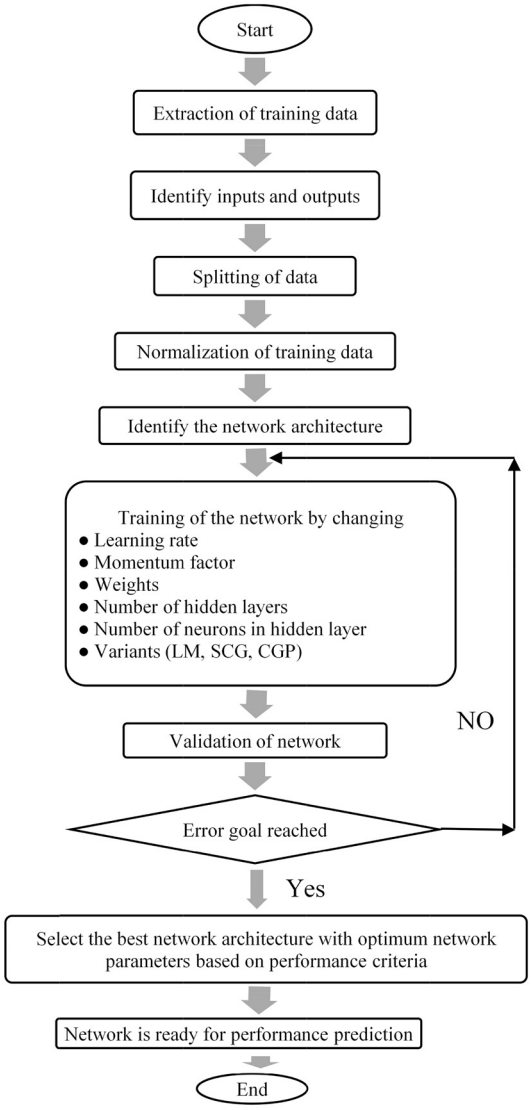

Fig. 2. Flow chart of ANN training processes [17].

variables that affect the performance of adsorption process have to be recognized.
2. The inputs and output variables have to be identified.
3. Normalization of inputs and outputs data is performed in the range between 0 and 1; between −1 and 1 or 0.1 and 0.9 (Eqs. 7–9). Normalization of inputs and outputs increases the training rate of the network.

$$y = x_i - x_{min}/x_{max} - x_{min} \quad (7)$$

$$y = x_i - x_{min}/x_{max} - x_{min} \quad (8)$$

$$y = x_i - x_{min}/x_{max} - x_{min} \quad (9)$$

where y is the normalized value of x_i. The x_{max} and the x_{min} are the maximum and minimum value of x_i, respectively.

4. The collected data is randomly divided into two data sets: training data set and testing data set. The training data set is employed to learn in the ANN to generate the model output. The testing data set is utilized to examine the parameters of the trained ANN. About 70–80% of the randomly selected data sets are appointed as training data sets and remaining data can be applied for testing the model.
5. Constructing and training of the ANN model, then the ANN model parameters are to be optimized to achieve results with good accuracy.
6. Selection of the best ANN model based on performance criteria and extraction of results from the optimum trained model.

The model performance is considered using statistical parameters including the coefficient of determination (R^2), average relative error (ARE), correlation coefficient (R), root mean square values (RMSE), coefficient of variance (COV), sums of squares error (SSE), mean absolute error (MAE), Chi-square statistic (χ^2), sum of the absolute error (SAE), and mean squared error (MSE) by altering the model parameters. The model parameters were optimized by trial and error method or by novel optimization algorithm to obtain better results. Some of the statistical parameters have been presented to evaluate ANN model, defined by

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(|y_{prd,i} - y_{exp,i}|)^2 \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_{prd,i} - y_{exp,i})^2}{\sum_{i=1}^{N}(y_{prd,i} - y_m)^2} \quad (11)$$

$$SSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{exp,i} - y_{prd,i})^2}{N}} \quad (12)$$

$$ARE = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{y_{exp,i} - y_{prd,i}}{y_{exp,i}}\right| \quad (13)$$

$$\chi^2 = \sum_{i=1}^{N}\left[\frac{(y_{prd,i} - y_{exp,i})^2}{y_{prd,i}}\right] \quad (14)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{prd,i} - y_{exp,i})^2}{N}} \quad (15)$$

$$SAE = \sum_{i=1}^{N}|y_{prd,i} - y_{exp,i}| \quad (16)$$

where $y_{prd,i}$ is the ith predicted property characteristic, $y_{exp,i}$ is the ith measured value, y_m is the mean value of $y_{exp,i}$, and N is the number of data. The R^2 depicts the fit of the predicted output variable approximation curve to the experimental data output variable curve. Higher R^2 and lower MSE show a model with better output prediction.

Reviewing applications of ANN model in removal of synthetic dyes from wastewaters by adsorption and other physicochemical methods.

2.1.3. Modeling of dye adsorption

The applications of ANN model for estimation of dye adsorption are reviewed and presented in Table 1. In related work, Aber et al. [24], in 2007, applied a three-layer feed forward backpropagation network (FFBPN) for predicting of acid orange 7 (AO7) removal from aqueous solutions by powdered activated carbon. They developed a FFBPN with three neurons in the input layer (representing initial dye concentration, initial pH and contact time) and one neuron in the output layer (representing dye concentration in solution after t min). They considered tan-sigmoid as transfer function of neurons at first and second layers, and log-sigmoid at output layer. The input values were normalized between −1 and 1 and output values were normalized between 0.2 and 0.8. The network with trainscg as training function and a 3-2-1 configuration was identified as the optimal structure. The performance of model in this study was measured by using mean relative error (MRE). The MRE result for ANN model used was found to be 5.81%. This value indicated that the ANN model had very good predictive performance. They concluded that ANN model could be used as a

Table 1
A summary of investigations on modeling of dye adsorption.

Authors [references]	Dye	Adsorbent	Optimal number of neurons	Evaluation indices
Aber et al. [24]	Acid orange 7	Powdered activated carbon	3-2-1	MRE = 5.81%
Xu and Hu [25]	Methylene blue	Activated carbon	–	AE = 0.0014
Dutta et al. [26]	Reactive black 5	TiO_2	3-10-1	R = 0.964; R = 0.993
Balci et al. [27]	Basic Blue 41 Reactive Black 5	Tree barks (*Eucalyptus camaldulensis*)	3-5-1	MSE = 0.00620594 MSE = 0.00119229
Cavas et al. [28]	Methylene blue	Dead leaves of *Posidonia oceanica* (L.)	3-15-1	R = 0.998
Celekli and Geyik [29]	Lanaset Red G	*Chara contraria*	4-23-1	R^2 = 0.999
Khonde and Pandharipande [30]	Bromocresol red Alizarin red Malachite green Methylene blue	Rice husk carbon	3-5-5-2	RMSE(train) = 0.025 RMSE (test) = 0.024
Saha and Mishra [31]	Safranin onto	Modified rice husk	3-10-1	R^2 = 0.988
Chowdhury and Saha [32]	Methylene blue	NaOH-modified rice husk	3-13-1	R^2 = 0.995
Chowdhury et al. [33]	Crystal violet	Eggshells	4-8-1	R^2 = 0.978
Çelekli et al. [34]	Lanaset red G	Walnut husk	4-20-1	R^2 = 0.995 MSE = 0.4993–5.0057
Elemen et al. [35]	Reactive red 141	HDTMA-bentonite	4-5-1	R^2 = 0.978 MSE = 0.027364
Deshmukh [36]	Indigo dye	Active carbon obtained from Coconut Shell	6-5-5-1	RMSE(train) = 0.072831 RMSE (test) = 0.106985
Dutta et al. [37]	Direct blue 86	Active carbon	5-8-1	R^2 = 0.982
Maleki et al. [38]	Reactive Red 198	Potato peel powder	5-7-1	R^2 = 0.98, RMSE = 4.3
Chakraborty et al. [39]	Crystal violet	Modified rice straw	3-17-1	R^2 = 0.997
Yang et al. [40]	Acid black 172	Bamboo biochar	5-5-1	MSE = 1.048×10^{-4} R^2 = 0.996
Hosseini Nia et al. [41]	Reactive orange 12	Gold nanoparticle-activated carbon	3-9-1	R^2 = 0.9720,MSE = 0.0007
Ghaedi et al. [42]	Methylene blue Brilliant green	Graphite oxide nanoparticle	3-10-1 3-11-1	MSE = 0.0012, R^2 = 0.982 MSE = 0.0001, R^2 = 0.981
Ghaedi et al. [43]	Sunset yellow	Zn (OH)2 nanoparticles-activated carbon	3-9-1	R^2 = 0.9782 MSE = 0.0013
Ghaedi et al. [44]	Reactive orange 12	Copper sulfide nanoparticles-activated carbon	3-12-1	R^2 = 0.9959, MSE = 0.000223
Ghaedi et al. [12]	Methyl orange	Gold nanoparticles-activated carbon Tamarisk	3-11-1 3-20-1	R^2 = 0.9580, MSE = 0.00082 R^2 = 0.9886, MSE = 0.0006
Çoruh et al. [45]	Malachite green Acid blue 161	Waste marble dust	4-12-1	R > 0.89, MSE < 0.01
Karimi and Ghaedi [46]	Methylene blue	Activated carbon	5-5-1	R^2 = 0.998, AAD% = 1.65
Zeinali et al. [47]	Methylene blue Brilliant green (BG)	Graphite oxide nanoparticle	3-10-1	R^2 = 0.9980, MSE = 0.0513 R^2 = 0.9944, MSE = 0.0674
Assefi et al. [48]	Eosin Y	Co_2O_3-NP-AC	3-23-1	R^2 = 0.9991, MSE = 1.49e-04
Coruh et al. [49]	Congo red	Seafood shell	3-10-1	MSE = 3.8127e-005
Ghaedi et al. [9]	Eosin B	Cobalt oxide nanoparticle-activated carbon	3-10-1	R^2 = 0.9965, MSE = 0.00016 R^2 = 0.9936, MSE = 0.00029
Ghaedi et al. [50]	Sunset yellow	Nickel sulfide nanoparticle loaded on activated carbon	3-5-1	R^2 = 0.99, MSE = 0.0003
Ghaedi et al. [51]	Malachite green	Copper nanowires loaded on activated carbon	3-11-1	R^2 = 0.9658, MSE = 0.0017
Ghaedi et al. [52]	Sunset yellow	Activated carbon prepared from wood of orange tree	3-13-1 3-15-1	R^2 = 0.9966, MSE = 0.0001 R^2 = 0.9961, MSE = 0.0001
Ghaedi et al. [53]	Phenol red	Au-NP-AC TiO_2-NP-AC	4-15-1 4-19-1	R^2 = 0.9994, MSE = 5.66e-05 R^2 = 0.9729, MSE = 0.0022
Hajati et al. [54]	Acid yellow 41 unset yellow	SnO_2 nanoparticle loaded activated carbon	5-10-2	R^2 = 0.99, MSE = 0.53 R^2 = 0.98, MSE = 0.79
Maghsoudi et al. [55]	Sunset yellow	Zinc oxide nanorods loaded on activated carbon	3-6-1	R^2 = 0.998, MSE = 0.0008
Malekbala et al. [56]	Acid red 57 (AR57)	Mesoporous carbon-coated monolith	3-21-1	R^2 = 0.997 MSE = 0.9365–6.6529
Mahmoodi et al. [57]	BB41 BR46 BR18 BB41 + BR18 BB41 + BR46	NiO-MnO_2 Nanocomposite	3-7-1 3-7-1 3-8-1 3-4-2 3-6-2	R^2 = 0.9977 R^2 = 0.9955 R^2 = 0.9989 R^2 = 0.9950 R^2 = 0.9995
Azad et al. [58]	Chrysoidine G Rhodamine B Disulfine blue	Ni doped FeO(OH)-NWs-AC	6-4-1 6-5-1 6-6-1	R^2 = 0.9997, MSE = 0.0055 R^2 = 0.9999, MSE = 0.0033 R^2 = 0. 9996, MSE = 0.0046
Babaei et al. [59]	Methylene blue	Activated spent tea (AST).	5-10-1	R^2 = 0.999
Asfaram et al. [60]	Methylene blue	Zinc sulfide nanoparticles with activated carbon (ZnS-NPs-AC)	4-8-1	R^2 = 0.9984, RMSE = 0.00065 R^2 = 0.9983, RMSE = 0.00071
Debnath et al. [61]	Congo Red	Fe_2O_3 nanoparticles	4-5-1	R^2 = 0.991, MSE = 0.00235.
Dil et al. [62]	Methylene blue Auramine O Crystal Violet Eosin Yellow	ZnO–NR–AC	6-3-1 6-5-1 6-2-1 6-4-1	R^2 = 0.9853, MSE = 0.000683 R^2 = 0.999730,MSE = 0.000014 R^2 = 0.987920, MSE = 0.000656 R^2 = 0.997093, MSE = 0.00011

(continued on next page)

Table 1 *(continued)*

Authors [references]	Dye	Adsorbent	Optimal number of neurons	Evaluation indices
Dil et al. [63]	Crystal Violet	ZnO-NR-AC	4-4-1	R^2 = 0.9998, MSE = 0.000959 R^2 = 0.9815,MSE = 0.000014
Celekli et al. [64]	Basic Red (BR) 46	Walnut husk (WH).	5-25-1	R^2 = 0.9988, SSE = 0.5848
Heibati et al. [65]	Ethidium bromide	Natural pumice and iron-coated pumice	4-2-2-1 4-1-2-1	R^2 = 0.9998, MSE = 0.005
Jamshidi et al. [66]	Eosin B Brilliant green	ZnS nanoparticles loaded activated carbon	4-7-1 4-7-1	R^2 = 0.9589, MSE = 0.0021 R^2 = 0.9455, MSE = 0.0022
Kooh et al. [67]	Methyl violet 2B	Soya bean waste	6-4-1	R^2 = 0.9946
Mahmoodi et al. [68]	Basic Red 18 Basic Blue 41	CuO–NiO nanocomposite	3-11-1 3-10-1	R^2 = 0.9904 R^2 = 0.9964
Salehi et al. [69]	Crystal violet	Magnetic activated carbon	5-8-1	R^2 = 0.9980, MAPE = 0.38%
Tanhaei et al. [70]	Methyl orange	Chitosan/Al_2O_3/Fe_3O_4 core-shell composite microsphere	2-8-1	R^2 = 0.998, MSE = 101.67

predictive model and as a replacement of kinetic studies in adsorption field. In similar work in 2010, Xu and Hu [25] used FFBPN to model the complex nonlinear relationship of the temperature, dosage of activated carbon, time and solution concentration as inputs with methylene blue (MB) adsorption onto activated carbon as output. It was reported that, the neural network model predicts the measured values with absolute error (AE) of 0.0014.

Dutta et al. [26] proposed an ANN model for predicting the adsorption efficiency (percentage of dye removal), and adsorption capacity (loading) of reactive black 5 dye on TiO_2 surface. They trained ANN model with three input parameters, such as, pH, TiO_2 dose and initial dye concentration. They found the optimum number of neurons in the hidden layer by trial and error method. In their work, an ANN with 3-10-1 configuration predicts both adsorption efficiency and adsorption capacity. They tested different algorithms and transfer functions, as in Tables 2 and 3. They reported that among the algorithms used "Levenberg-Marquardt backpropagation" algorithm obtains most satisfactory results and among the transfer function appled "poslin" is the most suitable for adsorption efficiency. For adsorption efficiency ANN model, the value of correlation coefficient (R) with most suitable combination "Resilient backpropagation" algorithm and "poslin" transfer function is 0.964. They also reported that for adsorption capacity model, "poslin" transfer function indicates satisfactory result, and among algorithms, Levenberg-Marquardt backpropagation and resilient backpropagation algorithm are the most suitable. In their work, the most suitable combination model of resilient backpropagation algorithm with "poslin" transfer function forecasted adsorption capacity with R value of 0.993.

ANN model was used by Balci et al. in 2011 to estimate of textile dyes (Basic Blue 41–BB41 and Reactive Black 5–RB5) adsorption in glass columns by using tree barks (*Eucalyptus camaldulensis*) [27]. In their network, three input neurons represent the volume of water (L), bed depth (cm), and effluent dye concentration (mg/L). The output neurons represent the dye concentration of the treated water. The network was optimized to a classic 3-5-1 configuration. They reported that model predictions were closer to the experimental data with MSE values of 0.00620594 and 0.00119229, respectively, for BB41 and, RB5 test data. The results indicated that ANN model could describe present

Table 2
Summary of trial and error method used for adsorption efficiency ANN model development [26].

Algorithm	Function	Transfer function for hidden layer	Transfer function for output layer	Correlation coefficient ®	Comments
Conjugate gradient backpropagation with Polak-Ribiere updates	traincgp	tansig	purelin	–	Gives arbitrary results
		satlin		–	Gives arbitrary results
		poslin		0.929	Gives satisfactory results
Levenberg-Marquardt backpropagation	trainlm	tansig	purelin	0.962	Gives satisfactory results
		satlin		0.943	Gives satisfactory results
		poslin		0.964	Gives satisfactory results
Gradient descent with momentum and adaptive learning rate backpropagation	traingdx	tansig	purelin	–	Gives arbitrary results
		satlin		–	Gives arbitrary results
		poslin		0.754	Gives satisfactory results
Resilient backpropagation	trainrp	tansig	purelin	0.969	Gives satisfactory results
		satlin		–	Gives arbitrary results
		poslin		0.96	Gives satisfactory results
Scaled conjugate gradient backpropagation	trainscg	tansig	purelin	0.985	Gives satisfactory results
		satlin		–	Gives arbitrary results
		poslin		0.93	Gives satisfactory results

Table 3
Summary of trial and error method used for adsorption capacity ANN model development [26].

Algorithm	Function	Transfer function for hidden layer	Transfer function for output layer	Correlation coefficient (R)	Comments
Conjugate gradient backpropagation with Polak-Ribiere updates	traincgp	tansig	purelin	–	Gives arbitrary results
		satlin		–	Gives arbitrary results
		poslin		0.973	Results are satisfactory
Levenberg-Marquardt backpropagation	trainlm	tansig	purelin	0.996	Results are satisfactory
		satlin		0.838	Gives satisfactory results
		poslin		0.979	Gives satisfactory results
Gradient descent with momentum and adaptive learning rate backpropagation	traingdx	tansig	purelin	–	Gives arbitrary results
		satlin		–	Gives arbitrary results
		poslin		0.97	Gives satisfactory results
Resilient backpropagation	trainrp	tansig	purelin	0.989	Gives arbitrary results
		satlin		–	Gives arbitrary results
		poslin		0.993	Gives satisfactory results
Scaled conjugate gradient backpropagation	trainscg	tansig	purelin	0.818	Gives arbitrary results
		satlin		–	Gives arbitrary results
		poslin		0.973	Gives satisfactory results

system.

Cavas et al. [28] used Thomas and ANN models for prediction of methylene blue removal by dead leaves of *Posidonia oceanica* (L.). They performed dynamic removal of MB from aqueous solution in a fixed-bed column. In their work, a total of 1215 data points (50% for training, 25% for validation, and 25% for testing) were employed to train and test the performance of ANN. Three variables including bed height, flow rate and time were considered for predicting the effluent MB concentration. They simulated single and double hidden layer networks with different hidden neuron sizes to obtain optimal structure of the ANN model. They showed that MSE was decreased when hidden neuron sizes were increased up to 20. Increase of hidden neurons beyond this value is not significant impact on the MSE, and it remains almost constant. For ANN model with one hidden layer, they set the optimal value of the number of hidden neurons as 15 where their training error and testing error start to diverge because overfitting happens when the difference between training error and test error increases with further increment of the number of hidden neuron. The ANN was optimized to 3-15-1 configuration for predicting the effluent MB concentration with good R value of 0.998. They have also carried out numerical experiments with two hidden layer ANN structure by changing the number of hidden neurons of the second hidden layer from 3 to 33 while varying the number of hidden neurons of first hidden layer from 3 to 33. They indicated that the ANN with two hidden layers is somewhat better to the one with single layer in terms of MSE values. However, they did not prefer to apply this type of ANN configuration not to fit exactly to noisy training data. They compared the experimentally measured effluent MB concentration with forecasted results achieved from neural network modeling. The results obtained shows ANN modeling performance is good enough to estimate experimental results, and ANN model better fitted with experimental data compared to Thomas model (Fig. 3 and Table 4).

Celekli and Geyik [29] in 2011 constructed a three layer ANN to forecast the removal efficiency of Lanaset Red (LR) G on *Chara contraria* based on 2304 experimental data sets. The input layer had four neurons as initial dye concentration, particle size, contact time, and initial pH regime. They considered amount of adsorbed LR G on the adsorbent as one neuron at the output layer. They selected the number of neurons in the hidden layer as 23 neurons based on minimum error values (%) for

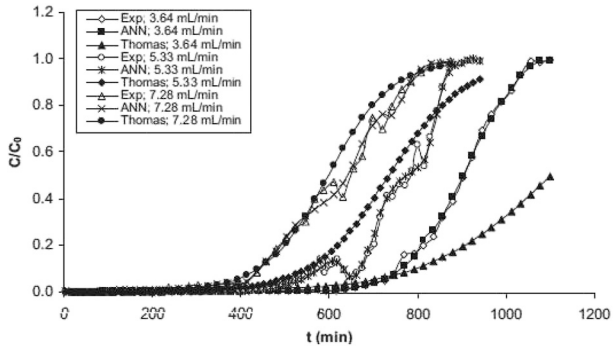

Fig. 3. Experimental data, Thomas model outputs and ANN model outputs for removal of MB in fixed-bed column at different flow rates (3.64, 5.33 and 7.28 mL/min) and bed heights: 9 cm [28].

Table 4
Mean Square Errors for comparison of Thomas and ANN models [28].

Z (cm)	F (mL/min)	ANN vs. experimental data	Thomas vs. experimental data
3	3.64	0.001453	0.013090
3	5.33	0.000631	0.008983
3	7.28	0.000549	0.008168
6	3.64	0.000386	0.005136
6	5.33	0.000553	0.003072
6	7.28	0.000426	0.002134
9	3.64	0.000121	0.049122
9	5.33	0.000347	0.006852
9	7.28	0.000603	0.003444

training and testing data sets. The network with a 4-23-1 configuration predicts the dye uptake at any time as the output with R2 value of 0.999. They also fitted pseudo second-order model to the experimental data. The comparison analysis, pseudo second-order kinetic and ANN models using determination coefficient (R^2), and error analyses, showed that both models had well fitted to experimental data ($R^2 > 0.973$). Nevertheless, the ANN was more appropriate model to describe the adsorption of LR G on *Chara contraria*, as shown in Table 5.

Table 5
Values of kinetic parameters, correlation coefficient, and error analyses for the adsorption of LR G on Chara contraria (pHi 1, particle size = 45–125 μm, C_o = 50–800 mg/L, M = 1 g/L, and t = 360 min) [29].

	Q_{exp}	40.12	70.25	130.21	230.66	307.48	372.39
	$q_{predict}$	38.74	73.62	129.48	229.5	314.39	373.52
	R2	0.989	0.99	0.989	0.996	0.986	0.994
	SSE	1.946	1.623	1.717	1.486	2.528	1.964
	AE	1.647	1.348	1.386	1.113	1.697	1.505
	ARE	4.513	2.168	1.269	0.54	0.603	0.447
	MSE	3.789	2.635	2.948	2.207	6.391	3.858
	χ^2	1.586	0.658	0.47	0.178	0.365	0.19
Pseudo second-order	q_{exp}	39.6	69.26	129.27	227.12	301.08	364.03
	$q_{predict}$	0.959	0.974	0.985	0.979	0.988	0.987
	R^2	2.149	2.86	3.982	8.28	7.965	10.221
	SSE	1.694	2.258	3.133	7.126	6.949	8.951
	AE	5.761	4.052	2.954	3.533	2.491	2.642
	ARE	4.62	8.179	15.86	68.55	63.438	104.477
	MSE	2.842	2.579	2.603	5.727	3.738	5.056
	χ^2	39.6	69.26	129.27	227.12	301.08	364.03

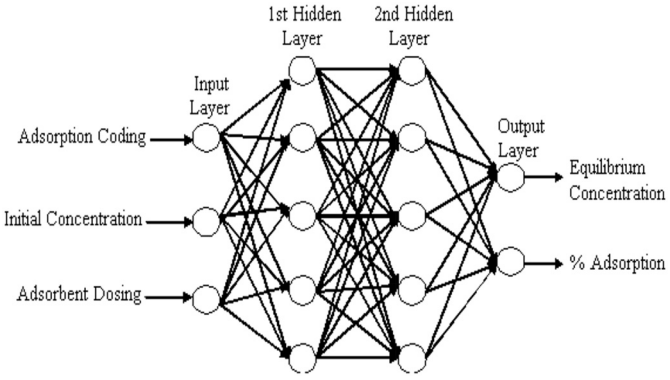

Fig. 4. Artificial neural network architecture [30].

Khonde and Pandharipande [30] in 2012 used the elite-ANN approach for predicting the adsorption of bromocresol red, alizarin red, malachite green and methylene blue dyes from aqueous solution using rice husk carbon. They developed ANN model by using BP network with three inputs namely, adsorbent dosing, adsorbate coding, and initial concentration of adsorbate; two hidden layers with five neurons each and two outputs such as equilibrium concentration and adsorption percentage, as shown in Fig. 4. The ANN was optimized to a 3-5-5-2 configuration. It was concluded that the proposed model has excellent accuracy and the ANN model can be used as an effective tool in modeling equilibrium concentration and % adsorption of dyes on rice husk carbon adsorbent. The ANN predictions are closer to the actual results with good MRE values of 0.025 and 0.024 for training and testing data set, respectively. In another work, Saha and Mishra [31] in 2012, modeled adsorption of safranin onto chemically modified rice husk in an upward flow packed bed reactor by using a three-layer network with a linear transfer function with backpropagation neural network. They considered three parameters namely time, pH and flow rate as inputs for predicting the % dye removal. The ANN was optimized to 3-10-1 configuration for predicting output with good R^2 value of 0.988.

Chowdhury and Saha [32] in 2012 employed an ANN model to forecast of dynamic methylene blue adsorption from aqueous solutions by using NaOH-modified rice husk (NMRH). In their work, the tansig function with BP algorithm at hidden layer and a purelin function at output layer were employed. A total 946 experimental data points were divided into three sets (training (50%), validation (25%), and testing sets (25%)). To avoid the scaling effect of parameter values, the input and output values were normalized in the 0.1–0.9 range. The input layer of the network consists of three neurons namely flow rate, time and bed height while the output layer includes the effluent MB concentration as the only neuron. To determine the optimum number of hidden neuron, they varied the number of neurons at hidden layer from 1 to 20. They repeated each topology three times to avoid random correlation because of random initialization of the weights. It was reported that the optimal topology with a 3-13-1 configuration predicts the R^2 values of 0.995. A R^2 of this high displays the reliability of the constructed ANN model. They also compared ANN model with the bed depth service time (BDST) and the Thomas models to determine the most suitable model for describing the dynamic dye adsorption process using the experimental data. They used various error functions such as R^2, ARE, SAE and $\chi 2$ to perform error analysis between actual data and data estimated by the models. The comparative error analysis indicated that due to the values of R^2, the ANN was better at the predicting effluent MB concentration compared to the Thomas and BDST models. The results of the error analyses are shown in Table 6. It was concluded that ANN model could be applied in designing an automated effluent treatment plant for adsorption MB by NMRH. In similar work, Chowdhury et al. [33] in 2012 forecasted removal of crystal violet (CV) dye from aqueous solution by adsorption onto Eggshells using ANN model. They used adsorbent dose (0.5–5 g), pH (3 − 10), temperature (293–303 K) and initial dye concentration (20–100 mg/L) as inputs and selected adsorption efficiency as the desired output. A total 210 experimental mean data points were used to construct the ANN structure with a tan-sigmoid transfer function at hidden layer and a linear transfer function at output layer. They normalized and then divided the all data into three sets with the same proportion similar to their previous work [32]. It was reported that a 4-8-1 network configuration forecasts the dye removal efficiency with R^2 value of 0.978 when compared to the experimental data (Fig. 5). It was concluded that forecasting based on the constructed ANN model could predict the behavior of the dye adsorption process under different experimental conditions.

Çelekli et al. [34] in 2012 presented artificial neural networks and pseudo second-order kinetic models to estimate the sorption of Lanaset red (LR) G on walnut husk (WH). Operating variables namely adsorbent dose, initial pH value, contact time, dye concentration, and particle size were employed as the inputs to predict the dye removal efficiency at any time as an output. They employed the tan-sigmoid transfer function with back-propagation algorithm at hidden layer and a linear transfer

Table 6
Error function assessment for comparison of BDST, Thomas, and ANN models [32].

F (mL/min)	Z (cm)	BDST vs. experimental data				Thomas vs. experimental data				ANN vs. experimental data			
		R^2	ARE	SAE	χ^2	R^2	ARE	SAE	χ^2	R^2	ARE	SAE	χ^2
10	5	0.9461	3.8961	4.4425	1.2467	0.9673	3.1532	3.5693	0.5729	0.9935	1.0368	1.4276	0.0192
15	5	0.9488	3.9134	4.5387	1.3018	0.9641	3.1987	3.3918	0.6547	0.9871	1.0234	1.5232	0.0176
20	5	0.9457	4.0137	4.2372	1.3577	0.9616	3.2383	3.3637	0.4312	0.9959	1.0621	1.5198	0.0207
5	3	0.9539	3.7845	4.1362	1.1852	0.973	3.3211	3.6733	0.5621	0.9897	1.1087	1.3317	0.0117
5	6	0.9444	3.8342	4.2781	1.0314	0.9652	3.2716	3.5478	0.7287	0.9928	1.1053	1.2339	0.0185
5	9	0.9462	3.8893	4.1834	1.2731	0.9611	3.3747	3.7662	0.9164	0.9906	1.1011	1.3565	0.0126

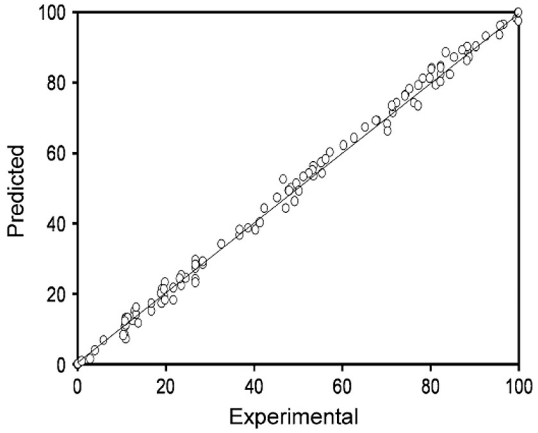

Fig. 5. Comparison of the experimental data with those predicted by the ANN model [33].

Table 7
Method comparison between pseudo second and ANN [34].

	C_o (mg/L)	50	100	200	400	600	800
ANN	q_{exp}	36.76	67.35	100.89	129.58	151.38	187.74
	$q_{predict}$	38.48	67.17	100.01	129.31	150.98	186.4
	R^2	0.995	0.996	0.998	0.998	0.999	0.998
	MSE	0.4993	1.2504	1.3067	2.0387	1.807	5.0057
Pseudo sec-ond	$q_{predict}$	36.08	65.95	102.77	129.59	150.72	185.07
	R^2	0.984	0.97	0.974	0.99	0.974	0.963
	MSE	1.2956	8.0492	17.0451	9.1594	34.3554	75.1481

function at output layer. In their study, 408 experimental mean data points were applied to construct the ANN model. All data were normalized in the range between 0.1 and 0.9. Samples were divided into training (278 samples), validation (65 samples), and test (65 samples) sets. To obtain the optimum number of hidden nodes, they changed the number of nodes from 2 to 25. Due to minimum error values (%) for training and testing sets, the number of neurons in the hidden layer was found as 20 neurons. Therefore, the network with a 12-5-1 configuration was selected as optimal ANN model. The ANN model was optimized to 4-20-1 configuration for predicting the sorption of LR G on the WH. Error analyses and regression analyses were performed to compare experimental and predicted data obtained from pseudo second kinetic and ANN models. R^2 values were obtained to be 0.992 and 0.998 for pseudo second and ANN models, respectively. These values of R^2 indicated that both models had well-fitting to experimental data. However, due to values of R^2 (0.995) and MSE (0.4993–5.0057), the ANN was found to be more suitable model to describe this sorption process (Table 7). In order to obtain the relative importance of the various input variables on the output variable, neural net weight matrix was applied. Their results showed that pH was the most efficient input (43%), followed by initial dye concentration (40%) for the uptake of LR G on WH. According to the results obtained, it was reported that ANN model could be applied in design and scale-up for adsorption of LR G on the WH.

Elemen et al. [35] in 2012 have developed a three-layered ANN model to predict the adsorption of the reactive red (RR) 141 on organoclay (HDTMA-bentonite) based on experimental data achieved in a batch system. The input parameters such as contact time (0–1440 min), initial dye concentration (20–00 mg/L), temperature (30–40 °C), and adsorbent dosage (0.05–0.1 g/L) were studied to predict the decolorization efficiency of the dye. A total 100 experimental data points were applied to construct the ANN model. The samples were divided to training and testing sets that included 85 and 15 samples, respectively. They also tested different transfer functions and algorithm, to obtain the most suitable network model based on the minimum value of MSE. They reported that the "logsig" transfer function is the most appropriate for adsorption efficiency calculation. Among algorithms used, "scaled conjugate gradient backpropagation" algorithm obtains the most satisfactory results (Table 8). The network using the most proper combination of the "scaled conjugate gradient backpropagation" algorithm and the "logsig" transfer function with 4-5-1 configuration predicts values of the output variable closer to experimental values with a correlation coefficient value of 0.978 and MSE value of 0.027364. It was concluded that artificial neural network provided reasonable predictive performance.

Deshmukh [36] in 2012 used a four-layer artificial neural network for prediction of indigo dye removal by the active carbon obtained from coconut shell. In this study, the six variables such as pH, particle size, adsorbent dose, type of activation, initial concentration of dye, and type of adsorbent were considered as inputs while pollutant removal efficiency (PRE) was the output. From the collected database (173 data points), 153 datasets were randomly selected as training sets (approximately 88% each of total). The rest, 20 datasets were applied for ANN testing. It has been reported that the ANN model with a 6-5-5-1 configuration, a learning rate of 0.3 and momentum of 1532 predicts the PRE with RMSE values of 0.072831 and 0.106985, respectively for training and testing data sets. In order to evaluate the prediction performance using the developed ANN model, they have given random variations in two or more input variables. In this case, they got forecasted output values close to actual values. Therefore, it was concluded that ANN can be successfully used as a potent model for PRE estimation in a fast and reliable ways, compared to the ordinary

Table 8
Summary of trial and error method used for adsorption efficiency ANN model development [35].

Algorithm	Function	Transfer function for hidden layer	Transfer function for output layer	Correlation coefficient (R^2)	Mean square error (MSE)
Scaled conjugate gradient backpropagation	trainscg	logsig	purelin	0.97	0.027364
		tansig		0.87	0.025364
		poslin		0.83	0.11193
Levenberg Marquardt backpropagation	trainlm	logsig	purelin	0.91	0.039921
		tansig		0.92	0.11825
		poslin		0.82	0.11357
Gradient descent with momentum backpropagation	traingdm	logsig	purelin	0.58	0.15655
		tansig		0.51	0.18299
		poslin		0.67	0.17413
Conjugate gradient backpropagation with Powelle Beale restarts	traincgp	logsig	purelin	0.91	0.082202
		tansig		0.92	0.11452
		poslin		0.84	0.096518
Resilient backpropagation	trainrp	logsig	purelin	0.88	0.12522
		tansig		0.9	0.12868
		poslin		0.84	0.16308

mathematical modeling. In another work, Dutta et al. [37] in 2012 have modeled the rate of adsorption of direct blue 86 using microwave assisted activated carbon by using a three-layer feed forward neural network. In their work, the Levenberg-Marquardt back-propagation algorithm, the tan sigmoid transfer function at hidden layer and the linear transfer function were used to train the ANN model. Five inputs namely pH, initial dye concentration, contact time, temperature and adsorbent dose were selected as the input variables whereas, the dye adsorption capacity was chosen as the output variable. The ANN with a 5-8-1 configuration forecasts the output values in good agreement with agreement with experimental values with correlation coefficient value of 0.982.

Maleki et al. [38] in 2013 employed the artificial neural network and multiple linear regression (MLR) models to anticipate the reactive red 198 (RR198) removal efficiency from aqueous solutions using potato peel powder based sorbent (PP). The five operational parameters such as initial pH, sorbent particle average size, dose of sorbent, and contact time and initial dye concentration were chosen as inputs of models while the dye removal efficiency (DR%) was considered as output variable. The 90 data points of DR% as a data set was randomly divided into three sets; 54 data as a training set, 18 data as a validation set, and 18 data as testing set. In their study, the ANN model with back propagation algorithm was used. A trial and error procedure was applied to determine the number of hidden layers and nodes. The MLR model was developed based on the same datasets for ANN model as the most popular linear model. The results show that the MLR model does not have good predictability for DR% because sorption process have a very complex mechanism with nonlinear relationships. The optimum architectural of ANN with 5–7-1 configuration estimate the DR% with the R^2 and RMSE values of 0.98 and 4.3, respectively. It was reported that the ANN as a powerful modeling method was successfully used to model the adsorption process.

Chakraborty et al. [39] in 2013 applied ANN method to model of dynamic adsorption of crystal violet from aqueous solution onto citric-acid-modified rice (*Oryza sativa*) straw. In order to compare among different models to obtain the most appropriate model for describing the adsorption kinetics of crystal violet in the fixed-bed column system, three models namely Thomas, BDST and ANN were fitted to the actual data and an extensive error analysis was performed between experimental data and data forecasted by the models. A dataset with 831 data points were used to build the ANN structure. The method used for constructing the ANN structure was much like in their previous work [32,33]. In order to obtain the optimum number of hidden nodes, a trial and error method was used, in which the number of nodes were changed from 1 to 20. Fig. 6 indicates the relation between network error and number of neurons in the hidden layer. As can be seen, the MSE is minimum just about 17 neurons. Therefore, the ANN with 3-17-1 configuration was selected as optimum model. Their results show that ANN calculated results are closer to experimental results with good R^2 value of 0.997. All the three models, ANN, Thomas, and BDST had good fitting to the actual data. However, according to the R^2 and the various error functions values, the ANN model was the best-fitting model compared to the other models.

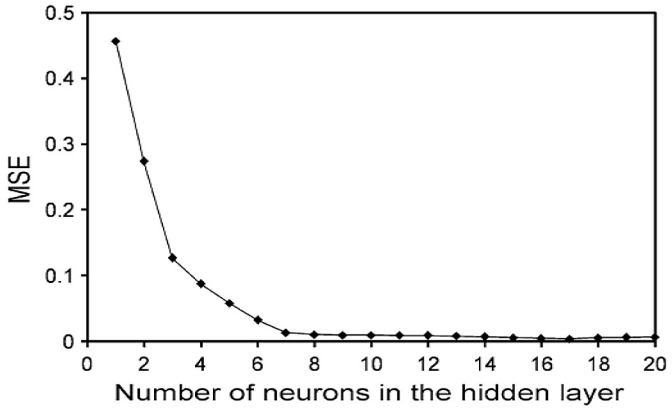

Fig. 6. The impact of the neuron numbers in the hidden layer on the performance of the neural network [39].

Yang et al. [40] in 2013 have studied the removal of color from solutions containing metal-complex dye acid black 172 using the adsorption using bamboo biochar. They have also constructed a back propagation ANN model to investigate the impacts of the operational parameters such as temperature, pH, ionic strength, initial dye concentration, and contact time on the adsorption capacity of bamboo biochar. The data sets with 180 samples were randomized and divided into three subsets; training (180 samples), validation (36 samples) and testing (36 samples). All samples were normalized in the range 0.1–0.9. The ANN with 5-5-1 configuration simulates the output with MSE and R^2 values of 1.048×10^{-4} and 0.996, respectively. The relative importance of input parameters was calculated using the network weights as follows:

$$Ij = \frac{\sum_{m=1}^{m=Nh}\left(\left(\frac{|W_{jm}^{ih}|}{\sum_{k=1}^{Ni}|W_{km}^{ih}|}\right)\times|W_{mn}^{ho}|\right)}{\sum_{k=1}^{k=Ni}\left\{\sum_{m=1}^{m=Nh}\left(\frac{|W_{jm}^{ih}|}{\sum_{k=1}^{Ni}|W_{km}^{ih}|}\right)\times|W_{mn}^{ho}|\right\}} \quad (17)$$

where *Ij* is the relative importance of the jth input parameter on the output parameter; *Ni* and *Nh* refer to the input and hidden neuron numbers, respectively; W shows connection weight; the superscripts 'i,' '*h*' and '*o*' represent input, hidden and output layers, respectively; and subscripts '*k*,' '*m*' and '*n*' are input, hidden and output the number of neurons, respectively. Their results using analysis based on the ANN illustrated that the temperature with a relative importance of 28.67% appeared to be the most effective variable in the adsorption process for dye removal (Fig. 7), followed by time (21.05%), ionic strength (18.25%), pH (17.77%) and dye concentration (14.23%). Since all the values of relative importance for variables were above 10%, it can be concluded that all of the parameters had strong influences on the adsorption capacity of metal-complex dye acid black 172 on bamboo biochar.

Hosseini Nia et al. [41] in 2014 have constructed a three layer ANN,with Levenberg–Marquardt back-propagation algorithm with 1000 iterations, a tangent sigmoid transfer function (tansig) at the hidden layer, and a linear transfer function (purelin) at the output layer, to predict the adsorption of reactive orange 12 (RO 12) onto gold nanoparticle-activated carbon. The data were normalized between 0.1

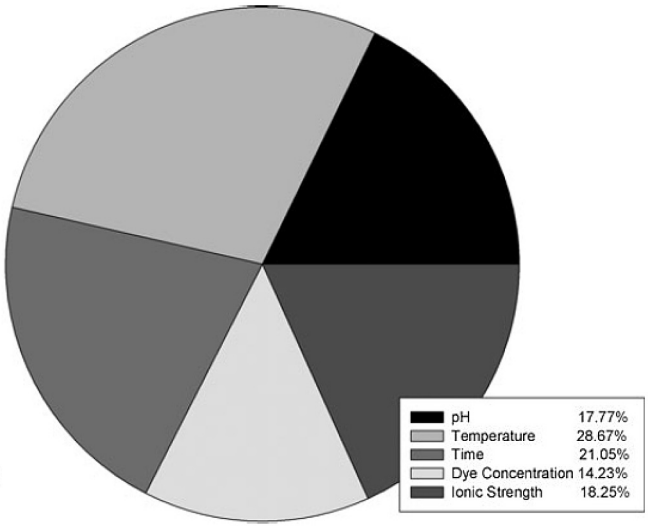

Fig. 7. Sensitivity analysis using artificial neural network [40].

Table 9
Comparison of MSE and R2 obtained using the MLR and ANN models [41].

Model	Training set		Testing set	
	MSE	R^2	MSE	R^2
ANN	0.0007	0.9720	0.0012	0.9596
MLR	0.0108	0.6050	0.0133	0.5420

and 0.9 and then randomly were divided into training (168 data) and testing (72 data). In their study, three parameters such as concentration of dye, adsorbent dosage and contact time were considered as the network input parameters, while the only output parameter is removal percentage (R%). The network was optimized to a 3-9-1 configuration. This network is capable of estimating the R% with the R2 and MSE values of 0.9720 and 0.0007, respectively. The comparison of the results obtained by the developed models indicated that the ANN model is better than the MLR model for the estimation of reactive orange 12 adsorption onto gold nanoparticles loaded on activated carbon (Table 9). Similarly, Ghaedi et al. [42–44] constructed the ANN models to predict the removal of dye by adsorption onto adsorbent. Their results indicated that the ANN model is better than MLR model for prediction of dye removal using adsorbent. It was reported that ANN predicted removal percentage were closer to the experimental results with the high values of R^2 and low values MSE. In another work, Ghaedi et al. [12] in 2014 have studied the adsorption of methyl orange dye by using Gold nanoparticles loaded on activated carbon (Au-NP-AC) and Tamarisk. They also developed an ANN model with four neurons in the input layer (include adsorbent dosage, initial dye concentration, contact time and stirrer speed) and one neuron in the output layer (representing removal percentage). After normalization, 182 data points were randomly divided into 2 groups; 182 and 78 points as training and testing subsets, respectively. The numbers of neurons in the hidden layer were varied between 1 and 30 to optimize the model structure. For Au-NP-AC adsorbent, the network was optimized with a 4-11-1 configuration. It was reported that ANN predicted results are closer to the experimental results with R^2 and MSE values of 0.9580 and 0.00082, respectively. For Tamarisk adsorbent, the ANN with 4-20-1 configuration predicts the output with R^2 and MSE values of 0.9886 and 0.0006, respectively. Fig. 8 displays the experimental data versus the predicted values of normalized removal dye for training, and testing sets using the ANN and MLR models. Table 10 also compares the results achieved by using the MLR and ANN models. Their results indicated that the ANN model was better than the MLR model for estimation of dye removal using both adsorbents.

Çoruh et al. [45] in 2014 have applied the experimental data to construct an ANN model to forecast acid blue 161 and malachite green dyes adsorption from aqueous solutions using waste marble dust. In their work, a three-layer ANN, an input layer with four neurons, namely initial dye concentration, temperature, contact time, and waste marble dust amount and one output layer with one neuron includes the percentage of the malachite green or acid blue 161 dyes removal was employed. All data were scaled between 0 and 1. Then, data were divided into three subsets. The training data set includes of 40 data points and the validation set contains of 6 data points and the testing set consists of 8 data points. They also used a logarithmic sigmoid transfer function (logsig) at hidden layer and output layer and the gradient descent backpropagation with adaptive learning rate algorithm as the best BP algorithms. Different neuron numbers in the hidden layer were tested to obtain the best model (Fig. 9). It was chosen as 12 neurons based on the minimum value of the mean squared error of the training and testing sets. The prediction model was indicated that the ANN model with a 4-12-1 configuration is able to forecast the efficiency of adsorption for the removal acid blue 161 and malachite green dyes by marble sludge dust with $R > 0.89$ and $MSE < 0.01$. They concluded that their results are quite reasonable and accepted, and ANN model is able to model the batch experimental system for the removal of malachite green and acid blue 161 dyes by waste marble dust.

Karimi and Ghaedi [46] in 2014 applied the ANN to model of methylene blue removal by using activated carbon. They developed a three-layer MLP with a tangent sigmoid transfer function at hidden layer and a linear transfer function at output layer. The network has five neurons in input layers including initial adsorbent dosage, contact time, initial pH, initial dye concentration, and agitated rate. The output is the percentage value of MB removal (R%). All data points were divided into three data sets (70 for training data set, 11 for validation data set, and 27 for testing data set). Scaling of inputs and outputs is done in the range between -1 and 1. By trial and error approach for primarily designed model, five neurons were selected for hidden layer. The ANN with a 5-5-1 configuration predicts dye removal percentage with R^2 and absolute average deviation percent (ADD%) values of 0.998 and 1.65, respectively. It was reported that the lowest ADD% and the high R2 between network forecasting and corresponding experimental data display goodness of the ANN model performance to predict the adsorption process under different conditions.

Zeinali et al. [47] in 2014 have applied an ANN model to simultaneous adsorption of methylene blue (MB) and brilliant green (BG) onto graphite oxide nanoparticle. A three-layer feed-forward artificial neural network with Levenberg–Marquardt back-propagation algorithm with 1000 iterations, a tangent sigmoid transfer function at hidden layer and a linear transfer function at output layer was developed. All experimental data set were randomly divided into two subsets, 100 and 44 data points for training and testing. Principal component analysis (PCA) was used to reduce the inputs. Thus, three PC was considered as neurons at input layer. The network was optimized by a 3-10-1 configuration for both dyes. It was reported that, ANN models predicted removal percent are closer to the experimental results with R^2 value of 0.9980 and MSE value of 0.0513 for MB dye, and R^2 value of 0.9944 and MSE value of 0.0674 for BG dye. They concluded that there is a good agreement between actual data and forecasted data using ANN model. In another work, Assefi et al. [48] in 2014 developed an ANN model for forecasting the adsorption of hazardous dye Eosin Y from aqueous solution onto cobalt (III) oxide nanoparticle loaded on activated carbon (Co_2O_3-NP-AC). A total of 360 data points were randomly divided into three subsets (70% for training, 15% for validation and 15% for testing subsets). Three variables including adsorbent dosage, contact time and initial eosin Y concentration were considered as inputs. It was reported that an ANN with 3-23-1 configuration simulate output with MSE and R^2 values of 1.49e-04 and 0.9991, respectively.

Coruh et al. [49] in 2015 have applied the ANN model for estimation adsorption of Congo red onto seafood shell in a batch system. In their work, a three-layer NARX network with tangent sigmoid transfer function at hidden layer, an input layer with three neurons (including time, adsorbent dosage and concentration), one hidden layer with 10 neurons, and an output layer with 1 neuron (adsorption efficiency) were constructed. To achieve the suitable weights and bias values for ANN model, a backpropagation (BP) algorithm applied combined gradient decent momentum (GDM) optimization. The training parameters such as delta max = 50, number of epochs = 100, error goal = 0.001, and Mu = 0.00001 were used. The network having 3-10-1 configuration predicted the removal percentage with MSE value of 3.8127e-005 at epoch 14. It was reported that ANN is applied some experimental data to forecast the response of the experiments at new similar conditions for the problem of the removal of dye from aqueous solution.

Ghaedi et al. [9] in 2015 used an artificial neural network for prediction of eosin B removal using cobalt oxide nanoparticle-activated carbon. A three-layer ANN with a tansig transfer function at hidden layer, a purelin transfer function at output layer, and Levenberg–Marquardt backpropagation algorithm with 1000 iterations were employed.

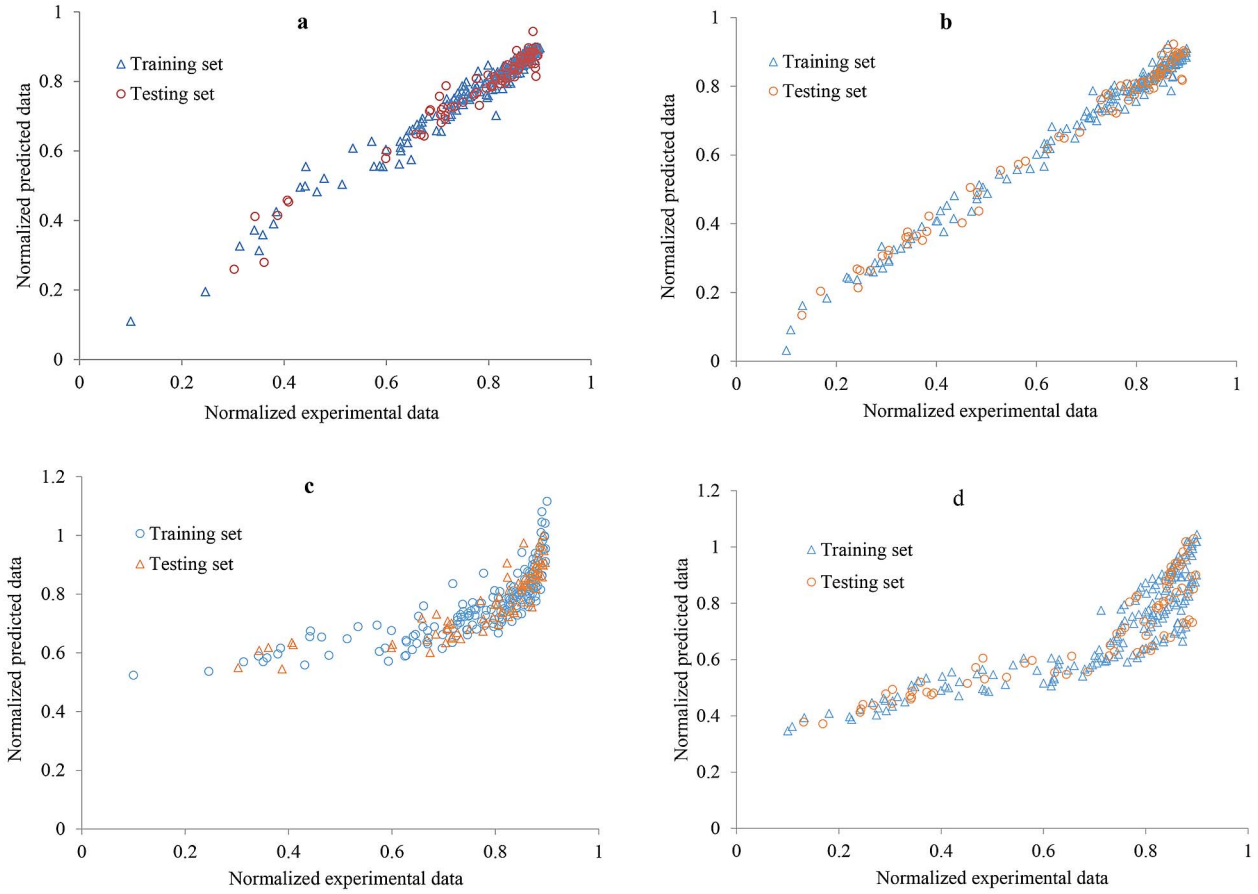

Fig. 8. The predicted values of normalized removal versus experimental data for training, and testing sets using the ANN model (a) Au-NP-AC, (b) Tamarisk, and using the MLR models, (c) Au-NP-AC, (d) Tamarisk [12].

A total 360 data point after scaling between 0.0 and 1.0 were randomly divided into two subsets (70% for training and 30% for testing). Three variables including concentration of dye, adsorbent dosage and contact time were considered as neurons at input layer, and removal percentage was used neuron at hidden layer. The network was optimized to a classic 3-10-1 configuration. It was reported that network estimations were closer to the experimental data with R^2 values of 0.9965 and 0.9936, and MSE values of 0.00016 and 0.00029, respectively for training and testing dataset. In another work, Ghaedi et al. [50] in 2015 applied the ANN method for modeling of sunset yellow dye removal using nickel sulfide nanoparticle loaded on activated carbon. The network with a 3-5-1 configuration predicts output with R^2 and MSE values of 0.99 and 0.0003, respectively. In related work, Ghaedi et al. [51] in 2015 employed an artificial neural network modeling for prediction of malachite green adsorption onto copper nanowires loaded on activated carbon. Total data points were normalized between −1 and 1. Out of 248 data points, 186 data points were applied for training, while the remaining data points were used for testing. The network with a 3-11-1 architecture predicts the removal percent with MRE of 0.0017 and R^2 of 0.9658 for testing data set. In a further work, Ghaedi et al. [52] in 2015 have compared ultrasonic with stirrer performance for removal of sunset yellow by using activated carbon prepared from

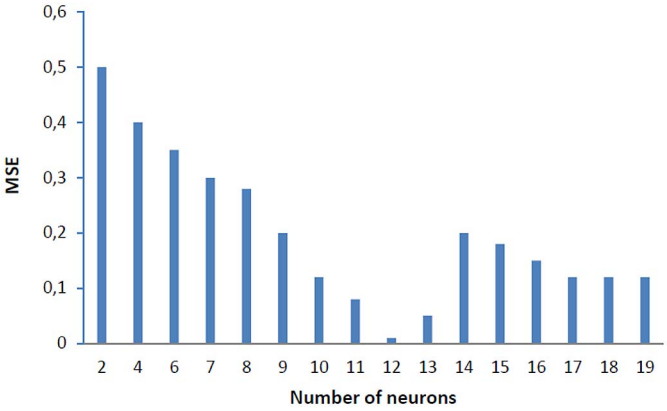

Fig. 9. The number of hidden layer neurons in the ANN structure [45].

Table 10
Comparison of MSE and R^2 obtained using the MLR and ANN models [12].

Model	Gold nanoparticles				Tamarisk			
	Training set		Testing set		Training set		Testing set	
	MSE	R^2	MSE	R^2	MSE	R^2	MSE	R^2
ANN	0.0012	0.9420	0.00082	0.9580	0.00044	0.9893	0.00060	0.9886
MLR	0.0082	0.5973	0.0071	0.6379	0.011	0.732	0.0123	0.782

wood of orange tree. They also developed an artificial neural network model to simulate removal (%) of sunset yellow dye based on experimental data. In their work, 168, 72 and 72 points were used for training, testing and validation data sets. The ANN was optimized to 3-13-1, and 3-15-1 configurations respectively for ultrasonic and stirrer. The R^2 and MSE values of testing data set were 0.9966, 0.0001 and 0.9961, 0.0001 for ultrasonic and stirrer, respectively. In another work, Ghaedi et al. [53] in 2015 used an ANN model for predicting the adsorption of phenol red onto gold (Au-NP-AC) and titanium dioxide (TiO2-NP-AC) nanoparticles loaded on activated carbon. In their work, the initial concentration, pH, contact time, and adsorbent mass were considered as input variables, and total data points were scaled between 0 and 1. It was reported that an ANN with 4-15-1 configuration forecast removal efficiency R^2 and MSE values of 0.9994 and 5.66e-05 for Au-NP-AC adsorbent. Similarly, an ANN with 4-19-1 configuration predicts output with R^2 and MSE values of 0.9729 and 0.0022 for TiO_2-NP-AC adsorbent. It was concluded that there is a good agreement between experimental data and predicted data by ANN model.

Hajati et al. [54] in 2015 used a three- layer ANN to predict the simultaneous removal of acid yellow 41 (AY41) and sunset yellow (SY) using SnO_2 nanoparticle-loaded activated carbon. Parameters namely contact time, pH, amount of adsorbent, AY41 concentration and SY concentration were considered as inputs. It was reported that ANN with 5-10-2 configuration forecast removal of AY41 dye with R^2 value of 0.99 and MSE value of 0.53 for AY41 dye and R^2 value of 0.98 and MSE value of 0.79 for SY dye. In another work, Maghsoudi et al. [55] in 2015 predicted sunset yellow dye adsorption onto zinc oxide nanorods loaded on activated carbon using an artificial neural network method. A dataset with 270 experimental data was divided into 70% for training and 30% for testing data sets. It was reported that the network having 3-6-1 configuration forecasted the removal percentage with MSE value of 0.0008 and R2 value of 0.998.

Malekbala et al. [56] in 2015 developed a powerful ANN model to predict the removal efficiency of acid red 57 (AR57) on mesoporous carbon-coated monolith (MCCM). The three-layer ANN with tangent sigmoid transfer function at hidden layer and a linear transfer function at output layer was constructed in their work consists of three neurons in input layer including initial pH, initial dye concentration, and contact time with one neuron in the output layer represents the amount of adsorbed AR57 on the MCCM. The data set of 504 points were divided into three subsets training (70%), validation (15%), and testing (15%). All data points were scaled between 0 and 1. A series of structure was used to evaluate the optimum number of hidden neurons, in which the number of neurons changed from 2 to 25. Based on the MSE values, a number of hidden neurons equal to 21 were accepted. The optimized ANN with 3-21-1 architecture predicts the output follow out the targets very well with good R^2 value over 0.997 and MSE values in the range of 0.9365–6.6529. The comparison between ANNs and pseudo-second-order kinetic models was drawn to figure out the best model for the adsorption of AR57 on MCCM. It was reported that both models had satisfactory fittings to experimental data. However, according to the values of R^2 and MSE, the ANN was specified to be the more suitable model to describe the adsorption of AR57 on the MCCM. They also used the neural net weight matrix to characterize the relative importance of the different input parameters on the output variables. Their results showed that the initial dye concentration was the most efficient variable (48%), followed by contact time (39%) and pH (13%) for the adsorption of AR57 on MCCM.

Mahmoodi et al. [57] 2016 used an artificial neural network as an intelligent system to model the Basic Blue 41 (BB41), Basic Red 18 (BR18), and Basic Red 46 (BR46) dyes removal from single and binary using NiO-MnO_2 nanocomposite. Their network constants of three neurons in input layer such as adsorbent dosage, adsorption time and initial dye concentration with one neuron in the output layer dye removal. A three layer FFBP with LM training algorithm was used using sigmoid and linear transfer functions at hidden and output layers, respectively. All data were normalized between -1 and $+1$. There were 70 and 40 actual data points for each dye in single and binary systems, respectively. Seventy percent of all data sets were employed for training, 15% for validation and 15% were applied as testing data. The results indicated that the predictions of the ANN models were in close agreement with experimental data with R^2 values in the range of 0.9950–0.9995 for single and binary systems. In another work, Azad et al. [58] in 2016 employed a three-layer ANN model to predict simultaneous adsorption of ternary dyes (Chrysoidine G (CG), Rhodamine B (RB) and Disulfine blue (DB)) by by Ni doped Ferric Oxy Hydroxide FeO(OH) nanowires on activated carbon (Ni doped FeO (OH)-NWs-AC). In their work, the network inputs are initial pH, initial CG, RB, DB concentration, adsorbent mass, and sonication time and the output of the network is removal percentage. The ANN was optimized to 6-4-1, 6-5-1 and 6-6-1 configurations for predicting the removal percentage of DB, RB and CG dyes, respectively. The ANN estimations are closer to the experimental results with good MSE and R^2 values (0.0055 and 0.9997), (0.0033 and 0.9999) and (0.0046 and 0.9996) for removal of CG, RB and DB dyes, respectively. The results obtained using ANN were compared with response surface methodology (RSM) results based on statistical parameters such as RMSE, AAD, MAE, and R^2. Their results showed that ANN had better prediction performance as compared to RSM.

Babaei et al. [59] in 2016 have developed an ANN model to predict the methylene blue (MB) dye removal from aqueous solution in a batch system onto activated spent tea (AST). The input parameters such as time, temperature, pH, adsorbent dose and dye concentration were studied to predict the dye removal efficiency (%). Eighty on experimental sets were applied to feed the ANN architecture. The data sets were divided into training, validation and test subsets that included 61, 10, and 10 points, respectively. It was reported that ANN with 5-10-1 configuration forecast dye removal efficiency (%) with R^2 value of 0.999 for testing set. Their results confirmed that the ANN model is a good way of forecasting the experimental data within the adopted ranges. They also calculated the relative importance of input parameters on the value of dye removal efficiency (%). It was reported that the relative importance of time, adsorbent dose, dye concentration, pH and temperature on dye removal efficiency was 35.44, 30.03, 14.10, 10.96, and 9.47%, respectively.

Asfaram et al. [60] in 2016 utilized an ANN method for modeling the adsorption of methylene blue from aqueous solutions on a composite of zinc sulfide nanoparticles with activated carbon (ZnS-NPs-AC). Four inputs of the network representing, sonication time, MB concentration, pH, and adsorbent amount. The output of the network was the efficiency of MB adsorption. The network was optimized by a 4-8-1 configuration. The R^2, and RMSE for training were 0.9984 and 0.00065, while for validation, these values were 0.9983 and 0.00071, respectively.

Debnath et al. [61] in 2016 applied an artificial neural network for predicting removal of Congo Red (CR) from aqueous media using Fe_2O_3 nanoparticles. Their network consists of four neurons in input layer representing solution pH, initial CR concentration, adsorbent dose and contact time with one neurons in the output layer representing removal (%) of dye. Experimental data set were normalized between 0 and 1, then were divided randomly into three subsets (91 data for training, 20 data for testing, and 20 data for validation set). A trial and error method was utilized for finalization of optimal number of nodes in hidden layer. It was reported that optimal ANN with 4-5-1 configuration predicts the removal (%) of dye with good R^2 value of 0.991 and MSE value of 0.00235.

Dil et al. [62] in 2016 utilized an ANN model to anticipate simultaneous adsorption of Methylene blue (MB), Auramine-O (AO), Crystal Violet (C 27 V) and Eosin Yellow (EY) dyes from aqueous solution onto ZnO–NR–AC. Six experimental parameters including the initial MB, EY, CV and AO concentration, amount of sorbent and sonication time were considered as inputs. The output was removal

percentage of dyes. The ANN models were optimized with 6-3-1, 6-5-1, 6-2-1 and 6-4-1 configurations. It was reported that the optimum topology predict removal (%) of training set with R^2 values of 0.9853, 0.999730, 0.987920, 0.997093 and MSE values of 0.000683, 0.000014, 0.000656, 0.00011, respectively for MB, EY, CV and AO dyes. While for testing set, R^2 values were 0.9008, 0.817730, 0.889242, 0.92535 and MSE values were 0.004747, 0.006726, 0.00422, 0.00406, respectively for MB, EY, CV and AO dyes. In further work, Dil et al. [63] in 2016 have studied he removal of crystal violet dye from aqueous solution by ultrasound assisted adsorption using zinc oxide nanorods loaded on activate carbon (ZnO-NR-AC) as an adsorbent. They also developed an ANN model for estimating removal (%) of crystal violet dye based on experimental data. The network inputs were sonication time, initial dye concentration, pH, and adsorbent dosage. The outputs of the network consist of the removal percentage of crystal violet. A total of 31 data points was divided into two subsets (75% training data and 25% testing data). It was reported ANN with 4-4-1 configuration predict the output of training data set with R^2 and MSE values of 0.9998 and 0.000959, while for testing data set, these values were 0.9815 and 0.0011, respectively. Their results indicate that there is a low divergence between forecasted and experimental data.

Celekli et al. [64] in 2016 developed the artificial neural network, gene expression programming (GEP), logistic, and pseudo-second-order kinetic models to forecast the removal efficiency of Basic Red (BR) 46 by walnut husk (WH). In their work, a three-layer ANN with a tangent sigmoid and linear transfer functions, respectively at hidden and output layers was applied. Particle size, initial dye concentration, initial pH regime, temperature and contact time were considered as inputs. The amount of adsorbed BR 46 on the adsorbent was considered as one neuron at output layer. About 2160 experimental mean sets were separated into three sets, such as training (1468 samples), validation (346 samples) and testing (346 samples). All inputs were normalized between -1 and $+1$. The output was normalized between 0 and 1. The network configuration was optimized by a 5-25-1 configuration. For training of the GEP models, maximum numbers of generations were between 5000 and 10,000. It was reported that results of kinetic models fittings to the experimental data showed that ANN was found to be more appropriate model to describe the sorption of BR 46 on WH. Moreover, Fig. 10(a)–(d) clearly displays that ANN had very well agreement with experimental q_t values at all dye concentrations with higher R^2 and lower SSE values than those of other applied models. They applied neural net weight matrix to assess the relative importance of operating factors on output parameters. Their results indicated that contact time with a relative importance of 48% appeared to be the most efficient variable in the adsorption process of BR 46 on WH, followed by initial concentration of dye (40%), particle size (5%), pH (4%), and temperature (3%). Regression analyses were carried out in order to compare experimental (q_{exp}) and forecasted ($q_{predict}$) data from the kinetic models. It was reported that values of R^2 were found to be 0.9572, 0.9880, 0.9973, 0.9988 for GEP, pseudo-second-order kinetic, logistic, and ANN models, respectively.

Heibati et al. [65] in 2016 employed an artificial neural network for modeling ethidium bromide (EtBr) adsorption on natural pumice and iron-coated pumice. It was reported that a simple back-propagation feed-forward network with three hidden layers including of 2–2-1 and 1–2-1 neurons in each layer was proposed as optimum topology with MSE value of 0.005 and R^2 value of 0.9998. In another work, Jamshidi et al. [66] in 2016 used an ANN model for predicting simultaneous ultrasonic assisted adsorption of eosin B (EB) and brilliant green (BG) on ZnS nanoparticles loaded activated carbon. In their work, the input layer consists of four neurons representing the amount of adsorbent, concentration of EB and BG dyes and contact time. The network with 4-7-1 configuration was identified as an optimum topology for both dyes. It was reported that ANN predicted dye removal (%) were closer to the experimental results with R^2 of 0.9589 and MSE of 0.0021 for BG dye, while for EB dye, these results were 0.9455 and 0.0022, respectively. It was reported that the results reveal acceptable agreement between experimental and forecasted data.

An artificial neural network model were utilized by Kooh et al. [67] in 2016 to predict the adsorption of methyl violet 2B onto soya bean waste in batch system under different conditions. All input data set were scaled to values between 0 and 1.0, then these data were randomly divided into training set (90%) and testing set (10%). Training and testing of ANN model was performed using a ten-fold cross validation approach. Each network for different number of hidden layers and hidden layer neurons had been tested, in which the network with 6-4-1 configuration produced best result with R^2 value of 0.9946. Due to the high value of R^2, it was concluded that the ANN model presented a very good predictive performance for the adsorption process. In another work, Mahmoodi et al. [68] in 2016 used an ANN model to predict adsorption Basic Red 18 (BR18) and Basic Blue 41 (BB41) dyes on CuO–NiO nanocomposite. In their work, a three-layer ANN model with a tangent sigmoid transfer function at hidden layer, a linear transfer function at output layer and backpropagation with Levenberg–Marquardt training algorithm was developed. The ANN models with 3-10-1 and 3-11-1 configurations predict efficiency of dye removal with R^2 values of 0.9964 and 0.9904 respectively for BB41 and BR 18 dyes. It was reported that the ANN modeling could effectively forecast the behavior of the adsorption process. Salehi et al. [69] in 2016 have predicted adsorption of crystal violet on magnetic activated carbon (MAC) using ANN and compared with response surface methodology. The same design was applied for both models. In their work, the input parameters of both models included pH, initial dye concentration, temperature, amount of magnetic activated carbon and time. The output was the percentage of dye adsorption. Various topologies with different number of neurons in hidden layer were considered, in which the ANN with eight neurons in hidden layer had the minimum value of MSE. The ANN with 5-8-1 configuration forecasted adsorption (%) with R^2 value of 0.9980 which indicated acceptable and suitable agreements between predicted and experimental data. In order to evaluate the prediction potential of both models, mean absolute percentage error (MAPE) and R^2 values were calculated. Fig. 11 and Table 11 display comparative results of the experiment and predicted values by both models. Their results indicate that the APE values of ANN model are less than RSM model for each experimental. The MAPE and R^2 using ANN were 0.38% and 0.9980, respectively. On the other hand, RSM has MAPE of 0.59% and R^2 of 0.9953. According to their results, it was reported that the ANN model had higher accuracy and ability in estimation of dye adsorption than RSM model.

A multilayer perceptron (MLP) neural network was constructed for forecasting the adsorption capacity of methyl orange (MO) onto Chitosan/Al_2O_3/Fe_3O_4 core-shell composite microsphere (CAMF) surface by Tanhaei et al. [70] in 2016. In their study, initial concentration of MO and time were considered as input parameters and adsorption capacity was output parameter. The number of neurons in hidden layer was obtained by trial and error method. The fifty-four experimental data were applied to build MLP model. All experimental data were randomly partitioned into three subsets (60% training, 20% validation, and 20% testing). The entire set of input parameters were normalized between 0 and 1. The ANN with 2-8-1 configuration predicts the adsorption capacity with R^2 value of 0.998 and MSE value of 101.67 for all training experimental data, in which the estimations of MLP network had very good agreement with actual data. It was reported that the MLP model was superiorly able to anticipate the adsorption capacity using contact time and different initial dye concentrations.

2.2. Adaptive neuro fuzzy inference system (ANFIS)

Jang [71] in 1993 suggested a novel architecture named adaptive neuro-fuzzy inference system or adaptive network-based fuzzy inference system (ANFIS). It is a kind of feed-forward artificial neural network that each layer is a neuro-fuzzy system component based on

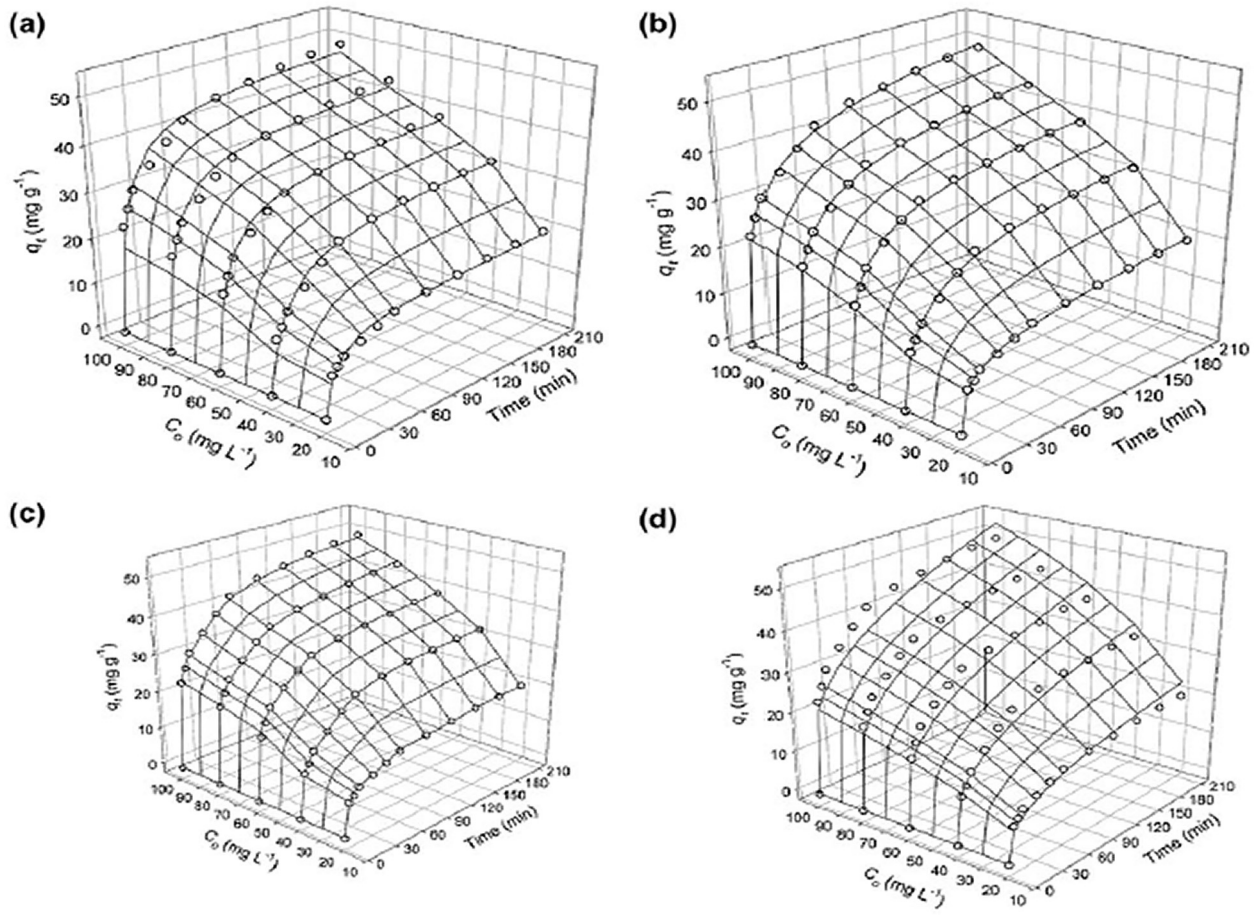

Fig. 10. Comparison of the experimental data with predicted data from (a) pseudo-second-order kinetic, (b) logistic, (c) ANN, and (d) GEP models (pH 10, particle size = 125 μm, Co = 20–100 mg/L, t = 0–210 min, and temperature = 313 K) [64].

Takagi–Sugeno fuzzy inference system. Since it combines both of the learning abilities of a neural network and reasoning abilities of fuzzy logic, it has potential to receive the benefits of both in a single methodology. By employing a hybrid-learning method, the ANFIS architecture can build an input-output model based on constructing a set of fuzzy if-then rules with appropriate membership functions (according to human knowledge) to produce the conditional input-output pairs. Hence, the ANFIS architecture can be used to model nonlinear systems. The details of ANFIS can be found elsewhere [71–73]. Ghaedi et al. [11] in 2013 used a principle component analysis-adaptive neuro-fuzzy inference system (PCA-ANFIS) model for prediction of methylene blue adsorption onto activated carbon obtained from *Pistacia khinjuk*. In their work, 262 experimental data points was split into training set (185 samples) and testing set (79

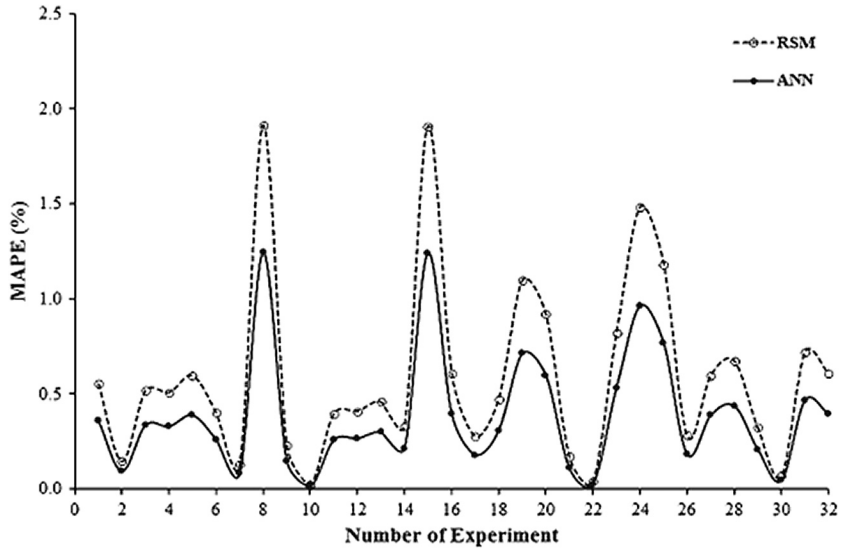

Fig. 11. MAPE (%) of RSM and ANN versus number of experiments [69].

Table 11
Comparative results of the experiment and predicted values by ANN and CCD for test data set [69].

No.	Factors					Adsorption (%)			APE (%)	
	pH	Time (min)	Temp. (°C)	Initial dye con. (mg/L)	MAC amount (mg)	Exp.	ANN	RSM	ANN	RSM
1	6	35	35	15	30	98.33	98	97.82	0.34	0.52
2	9	35	35	35	30	77.65	77.64	77.63	0.02	0.03
3	7.5	5	45	25	45	94.78	95.07	95.23	0.31	0.47
4	10.5	25	45	25	45	87.01	87.68	88.04	0.77	1.18
5	9	15	55	15	60	100.33	99.94	99.73	0.39	0.6
6	6	15	55	35	60	100.1	100.57	100.82	0.47	0.72
MAPE (%)									0.38	0.59

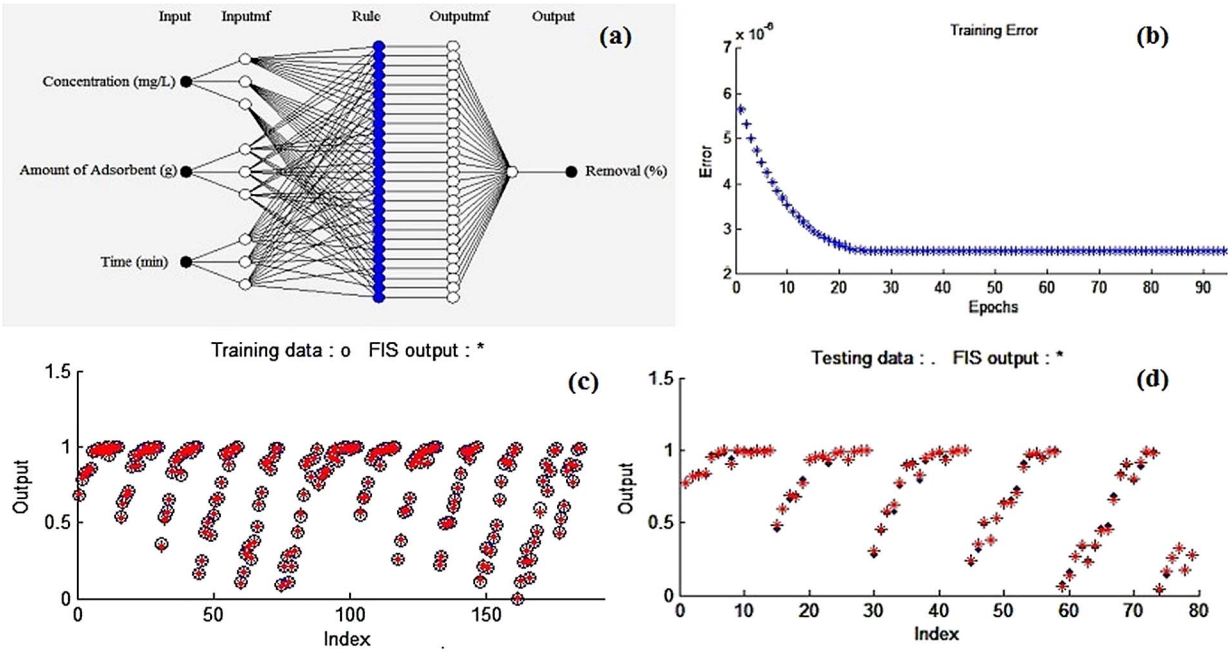

Fig. 12. (a) The architecture of the ANFIS model; (b) error versus the number of epochs; (c and d) distribution of predicted and experimental data of removal [11].

samples). PCA was applied for preprocessing of input variables such as concentration, contact time, and amount of adsorbent. The output variable (removal percentage) was scaled between 0 and 1 to avoid numerical overfitting. Their results showed that the minimum error values of 2.516×10^{-6} and 0.0312 were achieved, respectively for training and testing subsets. The optimal ANFIS model included the output layer with linear membership function type, input layer with 6,6,6 membership function numbers, hybrid method as optimum method and iteration values of 100 was chosen. Fig. 12(a–d) displays the structure of ANFIS, error versus the number of epochs, distribution of predicted and experimental data of removal (%) for training and testing sets, respectively. The optimal ANFIS model predicts the output with R2 value of 0.9968 and MSE value of 0.000271 for testing data. They also employed PCA-MLR and PCA-PLS as linear regressions. It was concluded that the forecasted data by using PCA-ANFIS model are within good agreement with the actual data and the PCA-ANFIS results are better than the PCA-MLR and PLS models in their study.

Aghajani and Tayebi [74] applied a model based on ANFIS for anticipating the adsorption of reactive red 198 from aqueous solution by composite of mesoporous material SBA-15 modified by the surfactant Cetyltrimethylammonium bromide (CTAB). An architecture of ANFIS model contains of five layers, such as fuzzy layer, product layer, normalized layer, defuzzy layer and total output layer was used (Fig. 13). In their study, a fuzzy clustering-based method with 17 number of clusters was applied for clustering. The variables, namely, temperature, pH, dosage, concentration and time were used as inputs. The amount of adsorption was considered as the output. Actual data were randomly divided into two groups: training (70%) and testing (30%) groups. Mean RMSE values of 1.91 and 2.81acheived after 50 runs in training and testing data, respectively. Based on the error between experimental and predicted output, it was reported that ANFIS is a model with high accuracy and efficiency in the prediction of adsorption process.

2.3. Support vector machine (SVM)

Support vector machine as a new machine learning approach introduced by Vapnik [75,76]. Based on structural risk minimization (SRM) trains learning machine instead of the empirical risk minimization, it has a good generalization performance in classification and regression. The combination of three variables, for example, the kernel type with its corresponding kernel function variable, ε of the ε-insensitive loss function and the penalty factor C control the regression performance of SVM [77]. The penalty factor C as a regularization variable is used to control the trade-off between maximizing the margin and to minimize the training error. Too small and too large values of C will cause insufficient stress on fitting the training data and over-fitting the training data, respectively. Generalization ability of SVM is controlled using the kernel type and its parameters. The type of data noise specifies the optimal value of ε [78].

Least squares-support vector machine (LS-SVM) developed by Suykens and Vandewalle in 1999 [79]. It has been known as an emerging and attractive semisupervised statistical learning approach [80,81]. It can be used to resolve linear and nonlinear multivariate

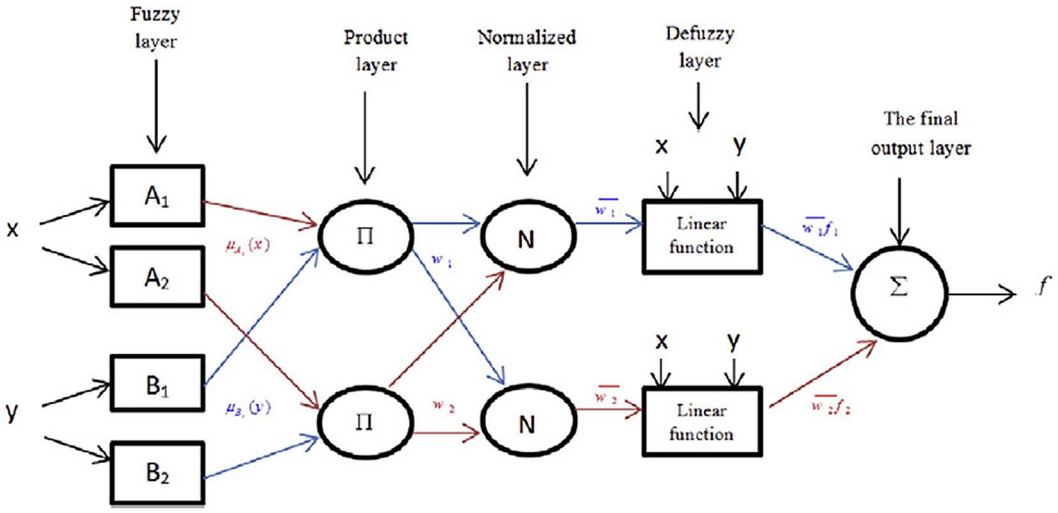

Fig. 13. ANFIS architecture [74].

calibration problems in a relatively quick way. LS-SVM contains similar advantages of SVM, also causes it easy for realization by computer software. There are only two variables requiring optimization, containing the regularization parameter (γ) and the kernel parameter (σ^2). The theory of LS-SVM is introduced based on the traditional support vector machine, the least squares linear systems is presented into SVM, rather than the traditional support vector machine directly using quadratic programming to resolve classification and regression problems. More details of the LSSVM algorithm can be obtained elsewhere [79,80,82]. Briefly, LS-SVM determines a linear relation ($y = w^T x + b$) between the output variable (y) and the input variables (x). The best fitting line is supposed to be the one that minimizes the lose function (L):

$$L = \frac{1}{2} \| w \|^2 + \gamma \sum_{i=1}^{n} e_i^2 \tag{18}$$

subject to

$$e_i = y_i - w^T x - b \tag{19}$$

where w presents an adjustable weight vector; b shows the scalar threshold; γ displays the regularization parameter, which specifies the adjustment between minimizing the training error (e_i) and minimizing model complexity. The final LS-SVM regression model can be represented as a nonlinear regression function:

$$y(x_i) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b \tag{20}$$

where $K(x, x_i)$ shows a kernel function, y presents the forecasted value and α is the Lagrange multiplier. LS-SVM regression has attracted wide attention and has been employed in adsorption field recently [60].

Asfaram et al. [60] in 2016 applied linear models (central composite design (CCD) and MLR) and nonlinear methods (LS-SVM and ANN) for prediction of the methylene blue (MB) adsorption using composite of zinc sulfide nanoparticles with activated carbon. In their work, four variables such as sonication time, initial MB concentration, pH, and adsorbent Mass were considered as inputs, while the output was dye removal efficiency. RMSE values of 0.00013, 0.00071 and 0.00117, and R^2 values of 0.9996, 0.9983 and 0.9978 were obtained respectively for the LS-SVM, ANN and CCD models using the validation set. It was concluded that nonlinear models (LS-SVM and ANN) indicate better performance than the linear models for the forecasting of MB dye adsorption. In similar works, Ghaedi et al. [83] in 2014 have used the LS-SVM model for the prediction of the MB dye adsorption onto copper oxide nanoparticles loaded on activated carbon, and Mahmoodi et al. [84] in 2015 have employed this method for modeling of Basic Blue 41 (BB41), Basic Red 46 (BR46), and Basic Red 18 (BR18) dyes removal using MnO2 nanoparticles. In both works, based on values of R^2, it was reported that the LSSVM approach is a proper and powerful method for modeling of dye adsorption.

2.4. Hybrid models (PSO-ANN and GA-ANN application in dye adsorption process)

Backpropagation uses gradient descent learning as one of the most popular algorithms for training neural networks. One of the most inherent limitations of using gradient descent learning in training artificial neural networks is that the ANN trapped in the local minima and may not finding a global minimum. However, a momentum term can be employed to solve this learning difficulty. Metaheuristic optimization algorithms including particle swarm optimization (PSO) and genetic algorithms (GA) have been applied for training ANNs rather than gradient decent learning. It has been seen that these algorithms overcome the local minima problem.

The PSO has originally developed by Kennedy based on simulating social behavior [85]. Shi and Eberhart introduced the inertia weight parameter into the original particle swarm optimizer to obtain a better performance [86]. The PSO technique has been successfully used in many areas and many applications [87–89]. Khajeh et al. [90] in 2013 used hybrid PSO with ANN model and response surface methodology for the forecasting of MB adsorption onto silver nanoparticles from water samples. In their work, six input parameters such as pH, sample and eluent flow rates, sample and eluent volume, and amount of adsorbent on extraction percent of dye as output were considered. Tansig and purelin were applied as transfer functions for hidden and output layers. The values of 10, 2, 2, 6, 0.05, 0.1 and 10 were used for particle in swarm, cognitive component, social component, maximum velocity, minimum inertia weight, maximum inertia weight and iterations for optimization simulation to obtain good extraction percent of MB. Their results indicated that the ANN with 0.78% prediction error has better accuracy of prediction than the RSM model with 1.68% prediction error. Fig. 14 shows the experimental versus predicted extraction percent using ANN and RSM models for training and testing (unseen data). The R^2 values for testing data by ANN and RSM were 0.95 and 0.89, respectively. Therefore, it was concluded that the ANN has higher generalization ability than RSM due to its universal capability to predict non-linearity of the system.

Ghaedi et al. [91] in 2015 used a hybrid approach of a ANN combined with PSO for prediction of brilliant green dye adsorption from aqueous solution onto zinc sulfide nanoparticle loaded on activated carbon. Fig. 15 shows the flowchart of the PSO. As can be

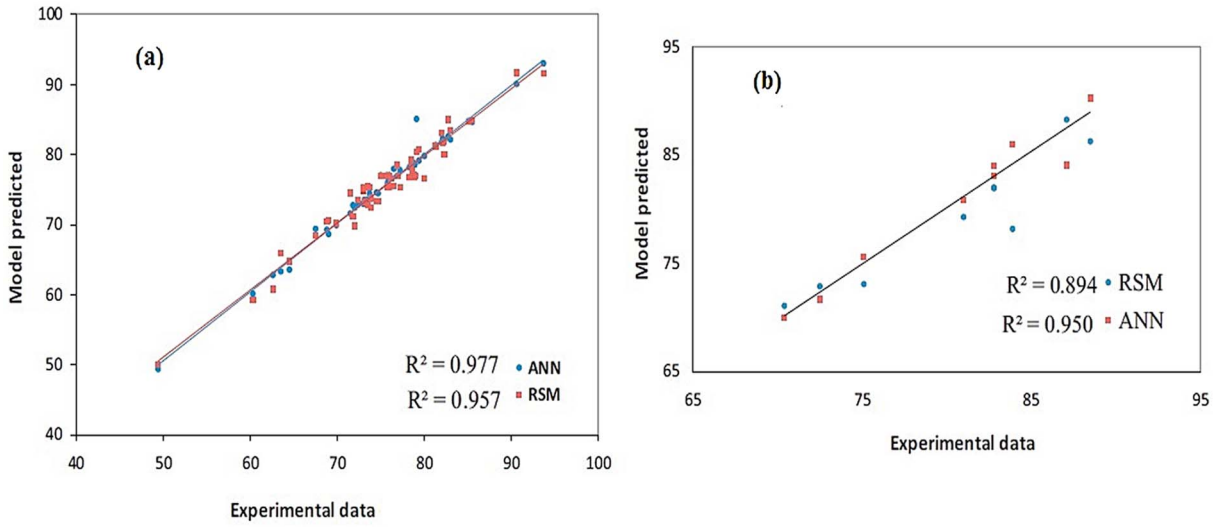

Fig. 14. RSM and ANN predicted vs. experimental data for MB; (a) training and (b) testing data [90].

seen from this figure the steps to implement the PSO are as follows:

1. Generation and initialization of particle array
2. Evaluation of objective function until obtaining better position
3. Determines new g best value
4. Calculation of particles new velocity
5. Update particle's position and repeating the steps

In their study, three variables such as contact time, adsorbent dosage and concentration were considered as inputs, while the output

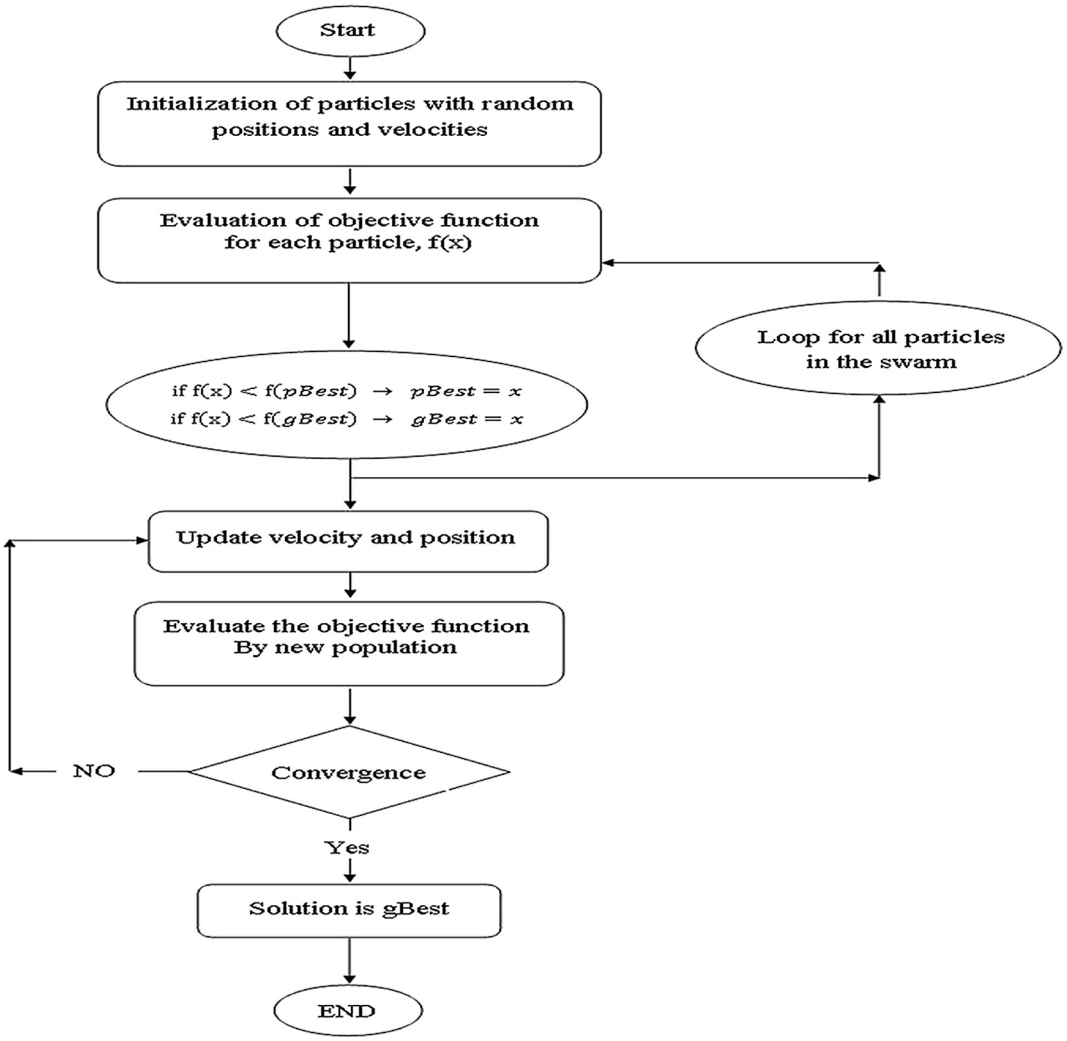

Fig. 15. Flow chart of the PSO [91,92].

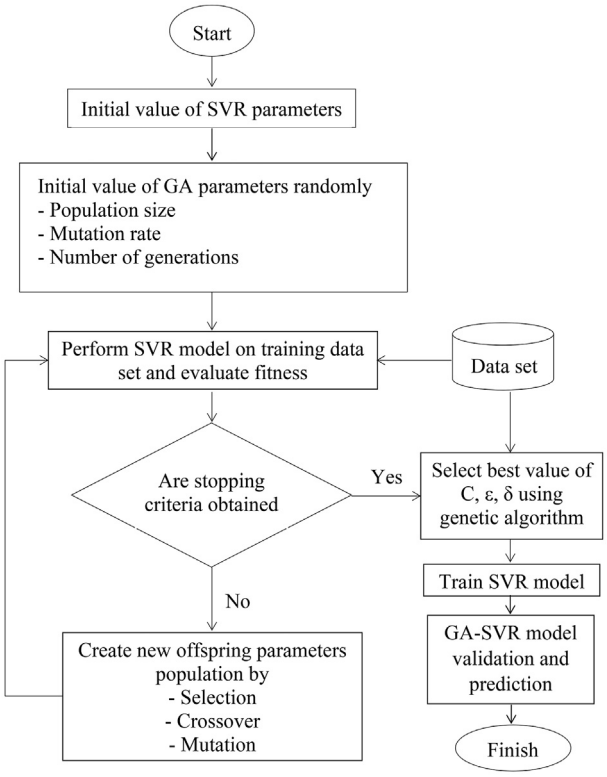

Fig. 16. Flow chart of the combined GA with SVR model [100].

was removal percentage of MO. The optimal ANN-PSO model predicts the output with the R^2 values of 0.9610 and 0.9506; MSE values of 0.0020 and 0.0022 for the training and testing data points, respectively. Comparison of the results obtained using MLR and ANN models showed that a three layer ANN model with 13 neurons in the hidden layer was a more proper approach to simulate the adsorption of BG onto ZnS-NP-AC.

In similar work, Agarwal et al. [92] in 2016 applied this hydride model for forecasting of methyl orange adsorption from aqueous solutions using lead oxide nanoparticles loaded activated carbon. The optimal ANN-PSO model with three layers and six neurons in the hidden layer forecasts the removal percentage of dye with R^2 and MSE values of 0.97 and 0.00093, for testing data set, respectively. It was concluded that ANN-PSO model is a suitable approach for prediction of the MO dye adsorption onto PbO-NP-AC.

GA is random search algorithm that is proper for complex nonlinear problems based on the natural genetics and natural processes of evolution [93,94]. GA has been widely applied as a search technique in large-scale and complex nonlinear optimization problems in many fields of science and engineering [95–98]. Khajeh et al. [99] in 2016 developed a combination of ANN and GA approach for predicting of MB adsorption onto wood sawdust from water samples. The pH, temperature, amount of sawdust and time were used as the input variables, while the percentage of extraction of MB was considered as the output. In another work, Ghaedi et al. [100] applied a hybrid model of SVR with GA for malachite green adsorption onto multi-walled carbon. Fig. 16 shows the GA procedure that can be applied in the SVR model to find the optimal parameters. Firstly, choose random algorithm parameters such as population size, generations, mutation rate and crossover rate; secondly, calculate the fitness of each individual in the population; thirdly, generate the offspring through selection, crossover and mutation; fourthly, selection of optimal parameters (consider the termination condition. If the offspring satisfies the termination condition, then stop the algorithm; otherwise, continue to optimize); and lastly, output the optimal parameters. In their work, data were randomly split into two subsets; training (176 data points), and testing (75 data points). The inputs were concentration, amount of adsorbent, and contact time, while the output was the removal (%). The optimal GA-SVR model predicts the output with MSE value of 0.0034 and R2 value of 0.9195 for testing data set.

3. Conclusions and future perspectives

Researchers from around the world have published many research studies on the prediction of dye adsorption using ANNs. This paper reviews the important research studies on dye adsorption forecasting using ANN methods. The literature reviewed in this paper showed that, the ANN approaches can be successfully applied for the modeling and forecasting of dye adsorption process with acceptable accuracy compared to conventional linear models such as MLR and PLS. The ANN model, despite major limitations such as extrapolation errors, over training errors, and optimization of network configuration, was widely employed for forecasting the performance of dye adsorption because of its simplicity. The optimization methods namely GA, and PSO were used for optimizing the ANN configuration. It seems that the hybrid networks with optimization approaches are more efficient to the performance of dye adsorption. ANFIS, LSSVR, and hybrid network with PSO and GA methods for modeling and predicting of dye adsorption with acceptable accuracy were discusses. The future research suggestions of ANNs in the field of dye adsorption for carrying out extensive studies are as follows:

(i) The prediction capability of other models such as radial basis function network (RBFN), neural gas network, group method of data handling (GMDH), and random forest (RF) for dye adsorption needs further research studies.
(ii) The hybridization of the GMDH and LSSVM (GLSSVM) forecasting methods, regarding the potential of predict dye adsorption with more accurate results.
(iii) Limited studies have been reported with the optimization algorithms combined ANN approaches. Hence, it is necessary to extend the optimization of network configuration for modeling adsorption process using an evolutionary computation method such as GA, PSO, ant colony algorithm, simulated annealing (SA), tabu search, firefly algorithm, artificial bee colony, teaching-learning-based optimization (TLBO), harmony search, shuffled frog-leaping algorithm (SFLA), and invasive weed optimization (IWO).

Based on the details and discussions presented in this paper, it can be concluded that, ANN methods are an excellent models for the adsorption of dyes. The information presented here would be highly useful to the researchers working in the field of dye adsorption and also those using the ANN models in their investigations.

Acknowledgment

This work was supported by the Gachsaran Branch, Islamic Azad University, Iran.

References

[1] Gupta VK, Gupta B, Rastogi A, Agarwal S, Nayak A. A comparative investigation on adsorption performances of mesoporous activated carbon prepared from waste rubber tire and activated carbon for a hazardous azo dye–Acid Blue 113. J Hazard Mater 2011;186:891–901.
[2] Uluozlu OD, Sari A, Tuzen M, Soylak M. Biosorption of Pb(II) and Cr(III) from aqueous solution by lichen (Parmelina tiliaceae) biomass. Bioresour Technol 2008;99:2972–80.
[3] Arunarani A, Chandran P, Ranganathan BV, Vasanthi NS, Sudheer Khan S. Bioremoval of Basic Violet 3 and Acid Blue 93 by Pseudomonas putida and its adsorption isotherms and kinetics. Colloids Surf B Biointerfaces 2013;102:379–84.
[4] Sari A, Mendil D, Tuzen M, Soylak M. Biosorption of Cd(II) and Cr(III) from aqueous solution by moss (*Hylocomium splendens*) biomass: equilibrium, kinetic and ther-

modynamic studies. Chem Eng J 2008;144:1–9.
[5] Salleh MAM, Mahmoud DK, Karim WAWA, Idris A. Cationic and anionic dye adsorption by agricultural solid wastes: a comprehensive review. Desalination 2011;280:1–13.
[6] Crini G. Non-conventional low-cost adsorbents for dye removal: a review. Bioresour Technol 2006;97:1061–85.
[7] Demirbas A. Agricultural based activated carbons for the removal of dyes from aqueous solutions: a review. J Hazard Mater 2009;167:1–9.
[8] Karlaftis M, Vlahogianni E. Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transportation Research Part C: Emerging Technologies 2011;19:387–99.
[9] Ghaedi M, Ansari A, Assefi Nejad P, Ghaedi A, Vafaei A, Habibi MH. Artificial neural network and Bees Algorithm for removal of Eosin B using Cobalt Oxide Nanoparticle-activated carbon: isotherm and kinetics study. Environ Prog Sustain Energy 2015;34:155–68.
[10] Dehghanian N, Ghaedi M, Ansari A, Ghaedi AM, Vafaei A, Asif M, Agarwal S, Tyagi I, Gupta VK. A random forest approach for predicting the removal of Congo red from aqueous solutions by adsorption onto tin sulfide nanoparticles loaded on activated carbon. Desalination and Water Treatment 2016;57:9272–85.
[11] Ghaedi M, Ghaedi AM, Abdi F, Roosta M, Vafaei A, Asghari A. Principal component analysis- adaptive neuro-fuzzy inference system modeling and genetic algorithm optimization of adsorption of methylene blue by activated carbon derived from Pistacia khinjuk. Ecotoxicol Environ Saf 2013;96:110–7.
[12] Ghaedi M, Ghaedi AM, Ansari A, Mohammadi F, Vafaei A. Artificial neural network and particle swarm optimization for removal of methyl orange by gold nanoparticles loaded on activated carbon and Tamarisk. Spectrochim Acta A Mol Biomol Spectrosc 2014;132:639–54.
[13] Ghaedi M, Ghaedi AM, Negintaji E, Ansari A, Vafaei A, Rajabi M. Random forest model for removal of bromophenol blue using activated carbon obtained from Astragalus bisulcatus tree. Journal of Industrial and Engineering Chemistry 2014;20:1793–803.
[14] Ghaedi M, Hosaininia R, Ghaedi AM, Vafaei A, Taghizadeh F. Adaptive neuro-fuzzy inference system model for adsorption of 1,3,4-thiadiazole-2,5-dithiol onto gold nanoparticales-activated carbon. Spectrochim Acta A Mol Biomol Spectrosc 2014;131:606–14.
[15] Ghaedi A, Vafaei A, Mohagheghian M, Afshar N, Hafezi S. Fuzzy modelling of concentration in chamomile solution using reverse osmosis. Fresen Environ Bull 2012;21:634–43.
[16] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;5:115–33.
[17] Mohanraj M, Jayaraj S, Muraleedharan C. Applications of artificial neural networks for thermal analysis of heat exchangers–a review. International Journal of Thermal Sciences 2015;90:150–72.
[18] Ata R. Artificial neural networks applications in wind energy systems: a review. Renew Sustain Energy Rev 2015;49:534–62.
[19] Pérez-Sánchez B, Fontenla-Romero O, Guijarro-Berdiñas B. A review of adaptive online learning for artificial neural networks. Artificial Intelligence Review 2016:1–19.
[20] Kalogirou S. Applications of artificial neural networks in energy systems. Energ Conver Manage 1999;40:1073–87.
[21] Yadav M, Sehrawat N, Sangwan A, Kumar S, Beniwal V, Singh AK. Artificial Neural Network (ANN): application in media optimization for industrial microbiology and comparison with response surface methodology (RSM). Adv Appl Sci Res 2013;4:457–60.
[22] Zhang Z, Barkoula N-M, Karger-Kocsis J, Friedrich K. Artificial neural network predictions on erosive wear of polymers. Wear 2003;255:708–13.
[23] Gardner J, Craven M, Dow C, Hines E. The prediction of bacteria type and culture growth phase by an electronic nose with a multi-layer perceptron network. Measurement Science and Technology 1998;9:120.
[24] Aber S, Daneshvar N, Soroureddin SM, Chabok A, Asadpour-Zeynali K. Study of acid orange 7 removal from aqueous solutions by powdered activated carbon and modeling of experimental results by artificial neural network. Desalination 2007;211:87–95.
[25] Xu Y, Hu X. Activated carbon adsorption of methylene blue based on artificial neural network [J]. Materials Review 2010;24:023.
[26] Dutta S, Parsons SA, Bhattacharjee C, Bandhyopadhyay S, Datta S. Development of an artificial neural network model for adsorption and photocatalysis of reactive dye on TiO_2 surface. Expert Systems with Applications 2010;37:8634–8.
[27] Balci B, Keskinkan O, Avci M. Use of BDST and an ANN model for prediction of dye adsorption efficiency of *Eucalyptus camaldulensis* barks in fixed-bed system. Expert Systems with Applications 2011;38:949–56.
[28] Cavas L, Karabay Z, Alyuruk H, Doğan H, Demir GK. Thomas and artificial neural network models for the fixed-bed adsorption of methylene blue by a beach waste *Posidonia oceanica* (L.) dead leaves. Chem Eng J 2011(171):557–62.
[29] Çelekli A, Geyik F. Artificial neural networks (ANN) approach for modeling of removal of Lanaset Red G on *Chara contraria*. Bioresour Technol 2011;102:5634–8.
[30] Khonde R, Pandharipande S. Artificial Neural Network modeling for adsorption of dyes from aqueous solution using rice husk carbon. International Journal of Computer Applications 2012;41.
[31] Saha P, Das Mishra R. Adsorption of safranin onto chemically modified rice husk in a upward flow packed bed reactor: artificial neural network modeling. Biotechnol Adv 2012;44:7579–83.
[32] Chowdhury S, Saha PD. Artificial neural network (ANN) modeling of adsorption of methylene blue by NaOH-modified rice husk in a fixed-bed column system. Environ Sci Pollut Res 2013;20:1050–8.
[33] Chowdhury S, Chakraborty S, Saha PD. Removal of crystal violet from aqueous solution by adsorption onto eggshells: equilibrium, kinetics, thermodynamics and artificial neural network modeling. Waste and Biomass Valorization 2013;4:655–64.
[34] Çelekli A, Birecikligil SS, Geyik F, Bozkurt H. Prediction of removal efficiency of Lanaset Red G on walnut husk using artificial neural network model. Bioresour Technol 2012;103:64–70.
[35] Elemen S, Kumbasar EPA, Yapar S. Modeling the adsorption of textile dye on organoclay using an artificial neural network. Dyes and Pigments 2012;95:102–11.
[36] Deshmukh S. Indigo dye removal by using coconut shell adsorbent and performance evaluation by artificial neural network. Emerging trends in science, engineering and technology. Springer; 2012. p. 655–63.
[37] Dutta M, Ghosh P, Basu I. Prediction of Adsorption Capacity of Microwave Assisted Activated Carbon for the decolorization of Direct Blue 86 by using Artificial Neural Network. Applied Sci 2012;5:414–22.
[38] Maleki A, Daraei H, Khodaei F, Bayazid-Aghdam K, Rezaee R, Naghizadeh A. Investigation of potato peel-based bio-sorbent efficiency in reactive dye removal: artificial neural network modeling and genetic algorithms optimization. Journal of Advances in Environmental Health Research 2013;1:21–8.
[39] Chakraborty S, Chowdhury S, Saha PD. Artificial neural network (ANN) modeling of dynamic adsorption of crystal violet from aqueous solution using citric-acid-modified rice (*Oryza sativa*) straw as adsorbent. Clean Technologies and Environmental Policy 2013;15:255–64.
[40] Yang Y, Lin X, Wei B, Zhao Y, Wang J. Evaluation of adsorption potential of bamboo biochar for metal-complex dye: equilibrium, kinetics and artificial neural network modeling. International Journal of Environmental Science and Technology 2014;11:1093–100.
[41] Hosseini Nia R, Ghaedi M, Ghaedi AM. Modeling of reactive orange 12 (RO 12) adsorption onto gold nanoparticle-activated carbon using artificial neural network optimization based on an imperialist competitive algorithm. J Mol Liq 2014;195:219–29.
[42] Ghaedi M, Zeinali N, Ghaedi A, Teimuori M, Tashkhourian J. Artificial neural network-genetic algorithm based optimization for the adsorption of methylene blue and brilliant green from aqueous solution by graphite oxide nanoparticle. Spectrochim Acta A Mol Biomol Spectrosc 2014;125:264–77.
[43] Ghaedi M, Ghaedi A, Negintaji E, Ansari A, Mohammadi F. Artificial neural network–imperialist competitive algorithm based optimization for removal of sunset yellow using $Zn(OH)_2$ nanoparticles-activated carbon. Journal of Industrial and Engineering Chemistry 2014;20:4332–43.
[44] Ghaedi M, Ghaedi A, Abdi F, Roosta M, Sahraei R, Daneshfar A. Principal component analysis-artificial neural network and genetic algorithm optimization for removal of reactive orange 12 by copper sulfide nanoparticles-activated carbon. Journal of Industrial and Engineering Chemistry 2014;20:787–95.
[45] Coruh S, Gürkan H, Kilic E, Geyikci F. Prediction of adsorption efficiency for the removal malachite green and acid blue 161 dyes by waste marble dust using ANN. Global NEST J 2014;16:676–89.
[46] Karimi H, Ghaedi M. Application of artificial neural network and genetic algorithm to modeling and optimization of removal of methylene blue using activated carbon. Journal of Industrial and Engineering Chemistry 2014;20:2471–6.
[47] Zeinali N, Ghaedi M, Shafie G. Competitive adsorption of methylene blue and brilliant green onto graphite oxide nano particle following: derivative spectrophotometric and principal component-artificial neural network model methods for their simultaneous determination. Journal of Industrial and Engineering Chemistry 2014;20:3550–8.
[48] 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;20:2905–13.
[49] Coruh S, Dogan G, Coruh U. Adsorption efficiency for the removal congo red by seafood shell using ANN. Proceedings of the 14th International Conference on Enronmental Science and Technology. Greece: Rhodes; 3-5 September 2015.
[50] Ghaedi M, Zeinali N, Maghsoudi M, Purkait M. Artificial neural network (ANN) method for modeling of sunset yellow dye adsorption using nickel sulfide nanoparticle loaded on activated carbon: kinetic and isotherm study. J Dispers Sci Technol 2015;36:1339–48.
[51] Ghaedi M, Shojaeipour E, Ghaedi AM, Sahraei R. Isotherm and kinetics study of malachite green adsorption onto copper nanowires loaded on activated carbon: artificial neural network modeling and genetic algorithm optimization. Spectrochim Acta A Mol Biomol Spectrosc 2015;142:135–49.
[52] Ghaedi AM, Ghaedi M, Karami P. Comparison of ultrasonic with stirrer performance for removal of sunset yellow (SY) by activated carbon prepared from wood of orange tree: artificial neural network modeling. Spectrochim Acta A Mol Biomol Spectrosc 2015;138:789–99.
[53] Ghaedi M, Daneshfar A, Ahmadi A, Momeni MS. Artificial neural network-genetic algorithm based optimization for the adsorption of phenol red (PR) onto gold and titanium dioxide nanoparticles loaded on activated carbon. Journal of Industrial and Engineering Chemistry 2015;21:587–98.
[54] Hajati S, Ghaedi M, Mahmoudi Z, Sahraei R. SnO_2 nanoparticle-loaded activated carbon for simultaneous removal of Acid Yellow 41 and Sunset Yellow; derivative spectrophotometric, artificial neural network and optimization approach. Spectrochim Acta A Mol Biomol Spectrosc 2015;150:1002–12.
[55] Maghsoudi M, Ghaedi M, Zinali A, Ghaedi AM, Habibi MH. Artificial neural network (ANN) method for modeling of sunset yellow dye adsorption using zinc oxide nanorods loaded on activated carbon: kinetic and isotherm study. Spectrochim Acta A Mol Biomol Spectrosc 2015;134:1–9.
[56] Malekbala MR, Hosseini S, Masoudi Soltani S, Malekbala R, Choong TS, Eghbali Babadi F. Development, application, and evaluation of artificial neural network in

investigating the removal efficiency of Acid Red 57 by synthesized mesoporous carbon-coated monoliths. Desalination and Water Treatment 2015;56:2246–57.
[57] Mahmoodi NM, Hosseinabadi-Farahani Z, Chamani H. Dye adsorption from single and binary systems using NiO-MnO_2 nanocomposite and artificial neural network modeling. Environ Prog Sustain Energy 2016.
[58] Azad FN, Ghaedi M, Asfaram A, Jamshidi A, Hassani G, Goudarzi A, et al. Optimization of the process parameters for the adsorption of ternary dyes by Ni doped FeO(OH)-NWs–AC using response surface methodology and an artificial neural network. RSC Adv 2016;6:19768–79.
[59] Babaei AA, Khataee A, Ahmadpour E, Sheydaei M, Kakavandi B, Alaee Z. Optimization of cationic dye adsorption on activated spent tea: equilibrium, kinetics, thermodynamic and artificial neural network modeling. Korean Journal of Chemical Engineering 2016;33:1352–61.
[60] Asfaram A, Ghaedi M, Azqhandi MA, Goudarzi A, Dastkhoon M. Statistical experimental design, least squares-support vector machine (LS-SVM) and artificial neural network (ANN) methods for modeling the facilitated adsorption of methylene blue dye. RSC Adv 2016;6:40502–16.
[61] Debnath A, Deb K, Das NS, Chattopadhyay KK, Saha B. Simple chemical route synthesis of Fe_2O_3 nanoparticles and its application for adsorptive removal of Congo red from aqueous media: artificial neural network modeling. J Dispers Sci Technol 2016;37:775–85.
[62] Dil EA, Ghaedi M, Ghaedi A, Asfaram A, Goudarzi A, Hajati S, et al. Modeling of quaternary dyes adsorption onto ZnO–NR–AC artificial neural network: analysis by derivative spectrophotometry. Journal of Industrial and Engineering Chemistry 2016;34:186–97.
[63] Dil EA, Ghaedi M, Ghaedi A, Asfaram A, Jamshidi M, Purkait MK. Application of artificial neural network and response surface methodology for the removal of crystal violet by zinc oxide nanorods loaded on activate carbon: kinetics and equilibrium study. J Taiwan Inst Chem Eng 2016;59:210–20.
[64] Çelekli A, Bozkurt H, Geyik F. Artificial neural network and genetic algorithms for modeling of removal of an azo dye on walnut husk. Desalination and Water Treatment 2016;57:15580–91.
[65] Heibati B, Rodriguez-Couto S, Ozgonenel O, Turan NG, Aluigi A, Zazouli MA, et al. A modeling study by artificial neural network on ethidium bromide adsorption optimization using natural pumice and iron-coated pumice. Desalination and Water Treatment 2016;57:13472–83.
[66] Jamshidi M, Ghaedi M, Dashtian K, Ghaedi AM, Hajati S, Goudarzi A, et al. Highly efficient simultaneous ultrasonic assisted adsorption of brilliant green and eosin B onto ZnS nanoparticles loaded activated carbon: artificial neural network modeling and central composite design optimization. Spectrochim Acta A Mol Biomol Spectrosc 2016;153:257–67.
[67] Kooh MRR, Dahri MK, Lim LB, Lim LH, Malik OA. Batch adsorption studies of the removal of methyl violet 2B by soya bean waste: isotherm, kinetics and artificial neural network modelling. Environ Earth Sci 2016;75:1–14.
[68] Mahmoodi NM, Hosseinabadi-Farahani Z, Bagherpour F, Khoshrou MR, Chamani H, Forouzeshfar F. Synthesis of CuO–NiO nanocomposite and dye adsorption modeling using artificial neural network. Desalination and Water Treatment 2016;57:17220–9.
[69] Salehi I, Shirani M, Semnani A, Hassani M, Habibollahi S. Comparative study between response surface methodology and artificial neural network for adsorption of crystal violet on magnetic activated carbon. Arabian Journal for Science and Engineering 2016:1–11.
[70] Tanhaei B, Ayati A, Lahtinen M, Mahmoodzadeh Vaziri B, Sillanpää M. A magnetic mesoporous chitosan based core-shells biopolymer for anionic dye adsorption: kinetic and isothermal study and application of ANN. J Appl Polym Sci 2016.
[71] Jang J-S. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 1993;23:665–85.
[72] Yilmaz I, Kaynar O. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. Expert Systems with Applications 2011;38:5958–66.
[73] Buyukbingol E, Sisman A, Akyildiz M, Alparslan FN, Adejare A. Adaptive neuro-fuzzy inference system (ANFIS): a new approach to predictive modeling in QSAR applications: a study of neuro-fuzzy modeling of PCP-based NMDA receptor antagonists. Bioorg Med Chem 2007;15:4265–82.
[74] Aghajani K, Tayebi H-A. Adaptive Neuro-Fuzzy Inference system analysis on adsorption studies of Reactive Red 198 from aqueous solution by SBA-15/CTAB composite. Spectrochim Acta A Mol Biomol Spectrosc 2017;171:439–48.
[75] Vladimir VN, Vapnik V. The nature of statistical learning theory. Heidelberg: Springer; 1995.
[76] Vapnik VN, Vapnik V. Statistical learning theory. New York: Wiley; 1998.
[77] Liu H, Xue C, Zhang R, Yao X, Liu M, Hu Z, et al. Quantitative prediction of log k of peptides in high-performance liquid chromatography based on molecular descriptors by using the heuristic method and support vector machine. J Chem Inf Comput Sci 2004;44:1979–86.
[78] Wang B, Chen J, Li X, Wang YN, Chen L, Zhu M, et al. Estimation of soil organic carbon normalized sorption coefficient (Koc) using least squares-support vector machine. QSAR & Combinatorial Science 2009;28:561–7.
[79] Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural Processing Letters 1999;9:293–300.
[80] Suykens JA, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J, Suykens J, et al. Least squares support vector machines. World Scientific; 2002.
[81] Adankon MM, Cheriet M, Biem A. Semisupervised least squares support vector machine. IEEE Trans Neural Netw 2009;20:1858–70.
[82] Ming Z. Research on least squares support vector machines algorithm. 2015; 2015.
[83] Ghaedi M, Ghaedi AM, Hossainpour M, Ansari A, Habibi MH, Asghari AR. Least square-support vector (LS-SVM) method for modeling of methylene blue dye adsorption using copper oxide loaded on activated carbon: kinetic and isotherm study. Journal of Industrial and Engineering Chemistry 2014;20:1641–9.
[84] Mahmoodi NM, Hosseinabadi-Farahani Z, Chamani H. Nanostructured adsorbent (MnO_2): synthesis and least square support vector machine modeling of dye removal. Desalination and Water Treatment 2015:1–10.
[85] Kennedy J. Particle swarm optimization. Encyclopedia of machine learning. Springer; 2011. p. 760–6.
[86] Shi Y, Eberhart R. A modified particle swarm optimizer. Evolutionary Computation Proceedings, 1998 IEEE World Congress on Computational Intelligence, The 1998 IEEE International Conference on: IEEE. 1998; 1998. p. 69–73.
[87] Khajeh M, Sarafraz-Yazdi A, Moghadam AF. Modeling of solid-phase tea waste extraction for the removal of manganese and cobalt from water samples by using PSO-artificial neural network and response surface methodology. Arabian Journal of Chemistry 2013;141:712–5.
[88] Liu S, Xu L, Li D, Li Q, Jiang Y, Tai H, et al. Prediction of dissolved oxygen content in river crab culture based on least squares support vector regression optimized by improved particle swarm optimization. Computers and Electronics in Agriculture 2013;95:82–91.
[89] Xing J-J, Liu Y-F, Li Y-Q, Gong H, Zhou Y-P. QSAR classification model for diverse series of antimicrobial agents using classification tree configured by modified particle swarm optimization. Chemom Intel Lab Syst 2014;137:82–90.
[90] Khajeh M, Kaykhaii M, Sharafi A. Application of PSO-artificial neural network and response surface methodology for removal of methylene blue using silver nanoparticles from water samples. Journal of Industrial and Engineering Chemistry 2013;19:1624–30.
[91] Ghaedi M, Ansari A, Bahari F, Ghaedi AM, Vafaei A. A hybrid artificial neural network and particle swarm optimization for prediction of removal of hazardous dye brilliant green from aqueous solution using zinc sulfide nanoparticle loaded on activated carbon. Spectrochim Acta A Mol Biomol Spectrosc 2015;137:1004–15.
[92] Agarwal S, Tyagi I, Gupta VK, Ghaedi M, Masoomzade M, Ghaedi A, et al. Kinetics and thermodynamics of methyl orange adsorption from aqueous solutions—artificial neural network-particle swarm optimization modeling. J Mol Liq 2016;218:354–62.
[93] Goldberg DE. Genetic algorithms in search, optimization and machine learning. Reading, MA: Addison-Wesley; 1989.
[94] Holland JH. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. U Michigan Press; 1975.
[95] Ziver AK, Pain CC, Carter JN, de Oliveira CRE, Goddard AJH, Overton RS. Genetic algorithms and artificial neural networks for loading pattern optimisation of advanced gas-cooled reactors. Annals of Nuclear Energy 2004;31:431–57.
[96] Morbiducci U, Tura A, Grigioni M. Genetic algorithms for parameter estimation in mathematical modeling of glucose metabolism. Comput Biol Med 2005;35:862–74.
[97] Abu Qdais H, Bani Hani K, Shatnawi N. Modeling and optimization of biogas production from a waste digester using artificial neural network and genetic algorithm. Resources, Conservation and Recycling 2010;54:359–63.
[98] Liu S, Tai H, Ding Q, Li D, Xu L, Wei Y. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. Mathematical and Computer Modelling 2013;58:458–65.
[99] Khajeh M, Sarafraz-Yazdi A, Natavan ZB. Combination of artificial neural network and genetic algorithm method for modeling of methylene blue adsorption onto wood sawdust from water samples. Toxicol Ind Health 2016;32:437–46.
[100] Ghaedi M, Dashtian K, Ghaedi A, Dehghanian N. A hybrid model of support vector regression with genetic algorithm for forecasting adsorption of malachite green onto multi-walled carbon nanotubes: central composite design optimization. Phys Chem Chem Phys 2016;18:13310–21.