Chemosphere 200 (2018) 330–343

Contents lists available at ScienceDirect

Chemosphere

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

Review

A review on experimental design for pollutants removal in water treatment with the aid of artificial intelligence

Mingyi Fan [a], Jiwei Hu [a, b, *], Rensheng Cao [a], Wenqian Ruan [a], Xionghui Wei [c]

[a] Guizhou Provincial Key Laboratory for Information Systems of Mountainous Areas and Protection of Ecological Environment, Guizhou Normal University, Guiyang 550001, Guizhou, China
[b] Cultivation Base of Guizhou National Key Laboratory of Mountainous Karst Eco-environment, Guizhou Normal University, Guiyang 550001, Guizhou, China
[c] Department of Applied Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China

HIGHLIGHTS

- Fundamentals, advantages and limitations were discussed for ANNs, GA and PSO.
- Studies were summarized on modeling of removal processes using ANNs, GA and PSO.
- Predicting performances for removal processes were compared between different ANNs.
- Developments of AI tools were described for the optimization of removal processes.

ARTICLE INFO

Article history:
Received 31 August 2017
Received in revised form
27 January 2018
Accepted 19 February 2018
Available online 20 February 2018

Handling Editor: Xiangru Zhang

Keywords:
Water treatment
Environmental pollutants
Experimental design
Artificial intelligence
Artificial neural networks
Genetic algorithm

ABSTRACT

Water pollution occurs mainly due to inorganic and organic pollutants, such as nutrients, heavy metals and persistent organic pollutants. For the modeling and optimization of pollutants removal, artificial intelligence (AI) has been used as a major tool in the experimental design that can generate the optimal operational variables, since AI has recently gained a tremendous advance. The present review describes the fundamentals, advantages and limitations of AI tools. Artificial neural networks (ANNs) are the AI tools frequently adopted to predict the pollutants removal processes because of their capabilities of self-learning and self-adapting, while genetic algorithm (GA) and particle swarm optimization (PSO) are also useful AI methodologies in efficient search for the global optima. This article summarizes the modeling and optimization of pollutants removal processes in water treatment by using multilayer perception, fuzzy neural, radial basis function and self-organizing map networks. Furthermore, the results conclude that the hybrid models of ANNs with GA and PSO can be successfully applied in water treatment with satisfactory accuracies. Finally, the limitations of current AI tools and their new developments are also highlighted for prospective applications in the environmental protection.

Contents

* Corresponding author. Guizhou Provincial Key Laboratory for Information Systems of Mountainous Areas and Protection of Ecological Environment, Guizhou Normal University, Guiyang 550001, Guizhou, China.
E-mail address: jwhu@gznu.edu.cn (J. Hu).

https://doi.org/10.1016/j.chemosphere.2018.02.111
0045-6535/

1. Introduction

Clean water is not easily available to all since increasing human activities and industrialization have led to the release of contaminated effluents into freshwater systems (Huang et al., 2009c; Dubey et al., 2015). These effluents cause the water contamination due to inorganic and organic pollutants, e.g., nutrients, heavy metals and persistent organic pollutants (POPs) (Cheung et al., 2003; Hoh and Hites, 2005; Luo et al., 2013; Li et al., 2016b). Problems with water pollution have received a significantly more attention than ever before from both the public and environmentalists (Ma et al., 2011; Liu et al., 2014). The pollutants removal in water treatment has been extensively studied with different methods, such as adsorption, solvent extraction, reduction, flocculation, coagulation, chemical or biological oxidation and membrane filtration (Kumar et al., 2007, 2008). The modeling and optimization are important procedures in pollutants removal processes to increase their efficiency without increasing the costs (Shojaeimehr et al., 2013). The conventional methods for the modeling and optimization of pollutants removal processes require the determination of a dependent variable for every combination of independent variables, just varying only one at a time while keeping all others as constants in batch studies (Singh et al., 2010). Such methods obviously require a broad range of experiments to be performed, which would be expensive and time consuming (Tak et al., 2015). Moreover, these methods would not be able to reveal the influence of the interactions between the independent variables (Sahu et al., 2009). The complex mechanisms (such as adsorption, reduction and oxidation) for water treatment may also hinder the modeling by using traditional methods (Nandi et al., 2010).

Artificial intelligence (AI) has recently gained a tremendous advance in various applications e.g., autonomous driving, big data, pattern recognition, intelligent search, image understanding, automatic programming, robotics and human-computer game, thus it will greatly impact the human society. AI tools mainly encompass artificial neural networks (ANNs), genetic algorithm (GA), support vector machine (SVM), random forest (RF), boosted regression tree (BRT), simulated annealing (SA), Monte Carlo simulation (MCS), particle swarm optimization (PSO), immune algorithm (IA), ant colony algorithm (ACA), imperialist competitive algorithm (ICA) and decision tree (DT). AI tools have also been combined with experimental design (e.g., response surface methodology and uniform design) in order to improve the precision of optimal solution prediction. Both ANN and RSM models take full advantage of the experimental data due to their connatural capability to extract information from operation parameters, which can describe the complex relationships in environmental engineering without understanding the in-depth mechanisms of removal processes (Bagheri et al., 2015; Fan et al., 2016). With the deepening of sustainable development concept, this technique has been used for pollutants removal because of their obvious advantages, such as reduction in reagent consumption and experimental work (Acharya et al., 2006). However, the applications of AI tools have still been limited in environmental protection as a whole due to their focus in water treatment (Mandal et al., 2015; Mendozacastillo et al., 2015; Reynelavila et al., 2015).

The objective of this review is to critically discuss the fundamentals and advantages of AI tools (e.g., ANNs, GA and PSO) as well as combined approaches (e.g., orthogonal design, response surface methodology and uniform design). This article also summarizes up-to-date studies concerning the modeling and optimization of the pollutants removal processes in water treatment by using multilayer perception network, fuzzy neural network, radial basis function network, self-organizing map network. Furthermore, this study demonstrates that hybrid models of ANN coupled with GA or PSO (ANN-GA or ANN-PSO) can be successfully applied in water treatment with satisfactory accuracies. Finally, we briefly describe the limitations of current AI tools and their new developments (e.g., deep learning ANNs, SVM, SA and MCS) for the modeling and optimization of pollutants removal processes.

2. Fundamentals, advantages and limitations of experimental designs and AI tools

2.1. Artificial neural networks

ANNs as one of the major AI approaches were derived from biological neurons (Zhang and Pan, 2014; López et al., 2016), which can solve multivariate non-linear problems with a suitable amount of data and an appropriate training algorithm (Wang and Deng, 2015; Li et al., 2016a). Each ANN consists of artificial neurons regrouped into layers and put in relation to each other by connections. ANNs are popular tools in machine learning due to their ability to learn rather complicated functions, which are non-linear statistical data modeling and data mining tools (Hu et al., 2010). ANNs provide a platform for mapping relationships between input and output parameters in removal processes (Chu, 2003; Aleboyeh et al., 2008). The capability of self-learning and self-adapting of ANNs can be successfully exploited for the prediction of output under the influence of various operational parameters. The

advantage of ANNs is that the mathematical description is not required for the phenomena involved in the process, therefore less time is needed for the model development than that for the traditional mathematical modeling (Assefi et al., 2014). Therefore, ANNs are a good alternative for conventional empirical modeling based on polynomial and linear regressions (Kose, 2008; Jin et al., 2015; Luo et al., 2015a,b).

Simulating water treatment processes and establishing the models can be applied with different types of ANNs, including multilayer perception (MLP) network, radial basis function (RBF) network, fuzzy neural network (FNN), self-organizing map (SOM) network, recurrent neural network (RNN), chaotic neural network (CNN), convolutional neural network and deep belief network (DBN) (Ouarda and Shu, 2009).

2.1.1. The category of artificial neural networks

MLP-ANN is composed of an input layer, an output layer and one or more hidden layers. The MLP-ANN and RBF-ANN are the feed-forward ANNs in which the neurons of each layer are interconnected from each preceding layer to the following layer, without lateral or feedback connections. The structure of RBF-ANN is similar to that of MLP-ANN, while the activation functions of RBF-ANN are different from those of MLP-ANN (Fig. S1). The general activation function of RBF-ANN is the Gaussian density function, which is defined by a "center" position and a "width" parameter (Deshmukh et al., 2012). The FNN model is a hybrid system combining the theories of fuzzy logic with neural network and thus can make effective use of easy interpretability of fuzzy logic as well as superior learning ability and adaptive capability of neural network (Fig. S2) (Huang et al., 2009a). The SOM networks learn to cluster groups of similar input patterns from a high dimensional input space in a non-linear fashion onto a low dimensional (most commonly two-dimensional) discrete lattice of neurons in a competitive layer (Fig. S3) (Kalteh et al., 2008).

Unlike the feed-forward neural networks, the RNN can handle one example at a time, retaining a state and memory since other networks training has long been difficult with a number of parameters (Fig. S4). Recently, deep learning is derived from the research on ANNs of multiple hidden layers with supervised learning, unsupervised learning and reinforcement learning. For example, convolutional neural network is a supervised learning model, which reduces the number of parameters to improve the training performance by using the relative relation of space. On the contrary, DBN is an unsupervised learning model that can solve the optimization problems of the deep learning structure.

The CNN and ESN models can be used for the prediction of pollutants removal in water treatment. Huang and Fang (2010) used the CNN model to conduct comprehensive evaluation of water resources allocation effects, which improves the efficiency of algorithm, strengthens objectivity of the evaluated results and also has a strong adaptability and replicability. This CNN model takes full advantage of ergodicity of chaotic motion and the ability of neural networks to handle complex non-linear mapping. Moreover, Sacchi et al. (2007) developed the echo-state network (ESN) that is one of RNN model for forecasting hydropower plant reservoir water inflow. The performance of the ESN model is fairly good for water inflow forecast in comparison with that of the SOM, RBF and FNN models. In addition, as a novel method, the application of SVM is still limited in the environmental protection field. The basic idea of SVM is to map the input variables into a higher dimensional feature space by a kernel function, *K(xi,xj)*, and then to carry out a linear regression in this space. SVM has shown the advantages in the problem of small samples, non-linear relationship and high dimensional pattern recognition. Besides, SVM can find the global optimal solution because it is a convex optimization problem. Ren et al. (2006) indicated that the SVM model performed well both in fitness and in prediction capacity.

2.1.2. Back-propagation (BP) artificial neural network

It has been well recognized that a neural network with one hidden layer is capable of approximating any finite nonlinear function with high accuracy, whereas an over-three-hidden-layer system is known to cause an unnecessary computational overload (Kim and Gilley, 2008). Hence, a MLP network with one hidden layer trained by the back-propagation (BP) algorithm is frequently used to build the ANN model (Fig. S5). BP is by far the most commonly employed training algorithm for MLP network (Khayet and Cojocaru, 2011; Wen et al., 2013). It is an iterative optimization process where the mean square error is minimized by adjusting the values of the weight and bias between the neurons (Fig. S6) (Yasin et al., 2014). The derivatives of the error function with respect to the weights are estimated using the error back propagation technique, in which the error in the final layer is propagated backwards to estimate the derivatives in the preceding layer. The optimization is carried out using the gradient descent method in which the weights are moved in the direction of negative gradient. The output from a neuron in the hidden layer is the transformation of the weighted sum of output from the input layers, which is given as follows:

$$c_j = k\left(\sum_{i=1}^{d} w_{ij}a_i + b_1\right) \tag{1}$$

where b_1 is the bias of hidden layer, a_i is the ith output from the input layer, w_{ij} is the weight connecting neuron i in the input layer to neuron j in the hidden layer, d is the number of input variables, k is the transfer function, and c_j is the jth output in the hidden layer.

The output from the neuron in the output layer is the transformation of the weighted sum of output in the hidden layer, which is given as follows:

$$e_k = k'\left(\sum_{j=1}^{m} w_{jk}c_j + b_2\right) \tag{2}$$

where b_2 is the bias of output layer, w_{jk} is the weight connecting neuron j in the hidden layer to the neuron k in the output layer, m is the number of neuron in output layer, k' is the transfer function, and e_k is the kth output from the output layer.

2.1.3. Training and testing of artificial neural networks

There are various transfer functions used in the ANN model, such as hard linear (hardlim), symmetrical hard linear (hardlim), linear (purelin), saturately linear (satlin), log-sigmoid (logsig) and hyperbolic tangent sigmoid (tansig). The transfer function commonly used is the tansig function and purelin function. The tansig transfer function can be used in input layer and hidden layer, while a purelin transfer function can be used in the output layer. The tansig transfer function can be expressed as:

$$f(x) = \left(1 - e^{-2x}\right) \Big/ \left(1 + e^{-2x}\right) \tag{3}$$

The purelin transfer function is defined as:

$$f(x) = x \tag{4}$$

The primary goal of training is to minimize the error function by searching for a set of connection weights and biases that cause the ANNs to produce the outputs equal to (or close to) targets (Fig. S7). The process of optimizing the connection weights and biases can be

carried out using local algorithms, which fall into two categories: the first-order methods (e.g., gradient descent) and the second-order methods (e.g., the Newton's method). Gradient descent only evaluates the first-order derivatives of objective function with gradient vector, whereas the Newton's method can estimate the second-order derivatives of objective function with the Hessian matrix. In addition to the two algorithms (gradient descent and Newton's method), other algorithms, e.g. scaled conjugate gradient, Levenberg-Marquardt, Powell-beale restarts, Fletcher-Reeves update and quasi-Newton, can also be used as the learning algorithms for ANNs in order to produce a network with the best R^2 value but a small MSE.

Determining the number of neurons in the hidden layer is an important task when an ANN is designed (Ghanbary et al., 2012). On the one hand, the convergence accuracy of the network may be lowered by a limited number of neurons in the hidden layer; on the other hand, a large number of neurons in the layer may result in the model over-fitting and generalization reduction. Therefore, the number of neurons in the hidden layer is obtained by the trial and error method so that the error between the experimental values and predicted values is minimized.

To ensure that all variables are important in the input data, principal component analysis can be performed as an effective procedure for the determination of input parameters. The input data used for the ANNs are generated from the experimental design, such as OD and RSM (the related description can be found in supplementary material). It is thus ensured that the data are uniformly distributed in the domain with high representativeness. The available data are generally divided into training, validation and testing subsets to develop an ANN model. The training set is used to estimate the weights and biases to fit the input-output relation according to its error. The validation set is used to decide when to stop training in order to avoid over-fitting and/or which network structure is optimal, whereas the test set is used to assess the generalization ability of the trained model. The data of input and output for an efficient ANN are normalized between 0.1 and 0.9 to avoid a numerical overflow due to large or small weights. The normalization equation used is as follows:

$$y = 0.8 \times (x_i - x_{\min})/(x_{\max} - x_{\min}) + 0.1 \tag{5}$$

where y is the normalized value of xi, and xmax and xmin are the maximum and minimum values of xi, respectively.

The Garson's equation partitions the hidden and output connection weights, which can be utilized to evaluate the relative importance of each independent variable in the ANN models. This equation is given as follows:

$$R_{in} = \sum_{j=1}^{A}\left(\left(|w_{ij}|\Big/\sum_{c=1}^{B}|w_{cj}|\right)|w_{jn}|\right)\Big/\sum_{i=1}^{B}\left(\sum_{j=1}^{A}\left(\times\left(|w_{ij}|\Big/\sum_{c=1}^{B}|w_{cj}|\right)|w_{jn}|\right)\right) \tag{6}$$

where R stands for the percentage of influence for the input neurons, w is the connection weight between the input, hidden and output neurons, B represents the number of neurons in input layer, A is the number of neurons in hidden layer and n is the number of output neurons.

The experimental value is compared with the predicted value by calculating mean square error (MSE). If the MSE is higher than a suggested value, it is propagated back from the output layer to the input layer, and the connection weights are modified accordingly until the number of iterations is achieved or the MSE is obtained under the suggested value. To determine the performance of the ANN model, the MSE and the coefficient of determination (R2) can be used. The MSE represents the errors associated with the model and should be calculated with the following equation:

$$MSE = (1/N)\sum_{i=1}^{N}(Y_p - Y_e)^2 \tag{7}$$

where N is the total number of model output, Yp is the predicted output, and Ye is the experimental value. And R2 represents the degree of correlation between the experimental and predicted values, which can be computed with the following equation:

$$R^2 = 1 - \sum_{i=1}^{N}(Y_p - Y_e)^2 \Big/ \sum_{i=1}^{N}\left(Y_p - \overline{Y_e}\right)^2 \tag{8}$$

where $\overline{Y_e}$ is the average of the experimental values.

ANNs have been applied to model the removal processes in environmental engineering, including adsorption (Kumar et al., 2007; Oguz and Ersoy, 2010), advanced oxidation processes (Kasiri et al., 2008), membrane systems (Guria et al., 2005), bioreaction (Kundu et al., 2014) and electrocoagulation (Mirsoleimani-Azizi et al., 2015). Compared with RSM, the disadvantage of ANNs is that, because of their structured nature (black box), it is difficult for ANNs to reveal the system information, such as the interactive effect of two variables on the removal processes (Jaiswal et al., 2005). The reproducibility of ANNs is also poor since the weight and bias between the neurons are given randomly. Furthermore, there is a high computational burden and proneness to over-fitting for training ANN with a large dataset (Elmolla et al., 2010). Finally, the conventional gradient-based optimization methods cannot be used efficiently for the global optimum search with complex nonlinear optimization problems because these methods require the objective function to be continuous, differentiable and more importantly smooth.

2.2. Different heuristic algorithms

It is well known that a plain gradient descent algorithm can become trapped at shallow local optima (Mohanraj et al., 2012). Another similar approach is simplex method, which is characterized by a derivative-free line search method. This is based on the idea of comparing the values of the objective functions at the N+1 vertices of a polytope in N-dimensional space and moving the vertex towards the minimum point as an optimization is in progress (El-Wakeel, 2014). This limitation of ANN has motivated researchers to generate ideas of merging or hybridizing ANN with other heuristic approaches in the search for a better performance. Therefore, AI tools (e.g., GA, PSO, ACA, IA, ICA, SA and MCS) inspired from social behavior or natural phenomena were found to be good alternatives for the conventional algorithms (e.g. gradient search techniques) to search for the global optima quickly and efficiently (Mohebbi et al., 2008; Arhami et al., 2013; Dehghani et al., 2014).

2.2.1. Genetic algorithms

GA is a numerical search tool operating according to procedures that resemble the principles of natural selection and genetics (Marseguerra et al., 2005). After an ANN-based process model is developed with a good prediction accuracy and generalization ability, GA can be applied to adjust the weights of hidden and output layers of a fixed set of connections (Chang and Hou, 2006; Ghaedi et al., 2015b). The optimization strategy based on GA can be described as a global optimization procedure with the advantage of being independent on the initial value to achieve the convergence. GA method is developed according to a series of steps, including solutions encoding, fitness computation based on the objective function, selection of the best chromosomes and the genetic

propagation of chosen parent chromosomes by genetic operators, like crossover and mutation (Fig. S8) (Jiang et al., 2014). Both crossover and mutation are implemented to produce the new and better populations of chromosomes (Yetilmezsoy and Demirel, 2008; Amirov et al., 2014). The process repeats several iterations until one chromosome has the best fitness, which is considered to be the optimal solution. Besides, IA introduces the immune operator based on the original GA, which includes two steps (vaccination and immune selection). The former is to improve the fitness for the solution and the latter is to prevent the population degradation. IA is good at the multi-target search because its search target has a certain dispersion and independence, while GA is used to obtain the global optimum solution since its search target is single and exclusive (Luo et al., 2015a,b; Wu et al., 2015).

2.2.2. Particle swarm optimization (PSO)

PSO as a well-known heuristic approach is inspired by the behavior of a bird flock. It is an evolutionary algorithm proposed by Kennedy and Eberhart, which can avoid trapping in a local minimum because it is not based on gradient descent algorithm (Ghaedi et al., 2015a). It starts with the following steps: (1) generation of initial population with random positions and velocities; (2) evaluation of fitness function for each particle (the former value will be replaced when a new position with better fitness value is obtained); (3) calculation of new velocity for the particles; (4) update of the particle position by moving toward maximal objective function; (5) completion of this algorithm when the iteration number reaches the maximum (Khajeh et al., 2013a). The advantage of PSO is that it depends only on the particle velocity to complete the search process (no crossover and mutation operations), therefore the convergence rate of PSO is higher than that of GA. PSO also has a memory ability, which can remember the best location of a particle and pass it to other particles. Moreover, the real number coding is used for PSO, which is directly determined by the solution of a problem. The disadvantage of PSO is that it lacks the dynamic adjustment of velocity and is easy to fall into a local optimum, resulting in difficult convergence and low convergence accuracy.

2.2.3. Other heuristic algorithms

ACA is superior to the above mentioned algorithms because of its strong robustness, however it needs a long search time, has a slow convergence rate and is easy to fall into a local optimum. ICA, unlike GA, PSO and ACA, was inspired from social behavior simulating the colonial assimilation mechanism and competition mechanism of the empire. The convergence rate and accuracy of ICA is higher than those of PSO and GA. Besides, SA is based on the ideas from statistical mechanics that is a probabilistic optimization technique with the potential of finding a global optimal solution for practical large-scale problems. SA method differs from the traditional descent algorithms in that this search algorithm for a neighborhood solution search allows not only the downhill moves but also the uphill moves.

Uncertainty in input parameters is related to the accuracy and representativeness of the input parameters used for predictions. MCS is a statistical sampling technique in obtaining a probabilistic approximation to the solution of a mathematical model. The uncertainty in model predictions due to uncertainties in input parameters is a function of the magnitudes and shapes of the probability density functions (PDFs) of the uncertainties in individual inputs (Hanna et al., 1998). The mechanism of this method is to repeat the generation of random parameters from their probability distribution, and then compute the statistics of the output. MCS technique is widely employed for the analysis of uncertainty in modeling, which enables the quantification of the model output uncertainty resulting from uncertain model parameters, input data or model structure (Shrestha et al., 2009). Therefore, MCS can be adopted to deal with input parameters uncertainties in order to use ANNs correctly. In general, MCS relies on the two random sampling processes, such as simple random sampling (SRS) and Latin hypercube sampling (LHS). Since MCS is easy to apply, its robustness and accuracy are not dependent on the problem dimension or complexity except it requires a large number of simulation runs to obtain a good estimate. In order to deal with this limitation, the Latin hypercube sampling (LHS) technique can be employed since it converges with small sample size.

2.2.4. Comparison of ANN-GA and ANN-PSO with RSM

The ANN-GA and ANN-PSO models with the higher R^2 value and the smaller average error offered more accurate predictions than the RSM models for the pollutants removal as demonstrated in Table S1 (Singh et al., 2012; Khajeh et al., 2013a, 2013b; Zhang and Pan, 2014; Mohan et al., 2015; Dil et al., 2016; Fatemeh Mehrabia et al., 2016). Using the ANN-GA and ANN-PSO based tool, the removal efficiencies of various pollutants have been improved, as compared to that of RSM. Both RSM and ANN-GA were applied to model and optimize the processes for the Cd(II) and Rhodamine B removal. The results indicated that the predicted removal efficiency of the ANN-GA model was more accurate than that of the RSM model (Fan et al., 2017a; Shi et al., 2017). Furthermore, the results from ANN-GA and ANN-PSO models showed a high accuracy for predicting the Cu(II) removal efficiency as compared to RSM model (Fan et al., 2017b).

Although ANN is superior to RSM, these models complement each other in interpreting the experimental removal efficiency. ANN is more reliable in capturing the nonlinear relationship between the removal efficiency and process variables, while RSM notes the statistical importance of the individual process variables and their interactions via ANOVA. However, the main limitation of RSM is that it assumes only quadratic non-linear correlation. Therefore, RSM can be effectively used when the search window is appropriately narrowed down, which makes the search process highly dependent upon the searched space. It will require either extra experiments or good priori knowledge of the system to fix the search window. ANN can easily overcome the above discussed limitation of RSM because it can inherently capture almost any form of non-linearity. Thus, the liberal search space can be chosen by using ANN although the correlation in the search space is more complex than the quadratic.

3. Applications of AI tools in the modeling and optimization of pollutant removal processes

A classification tree of AI techniques for pollutants removal in water treatment is presented in Fig. 1, which shows that two categories of ANN approaches (single methods and hybrid methods) are employed in water treatment for modeling, predicting and optimizing the pollutants removal.

In Table 1, we collected 44 ANN models reported for modeling and optimization of pollutant removal processes, including 7 models for the nutrients removal, 11 models for the heavy metals removal, 6 models for the POPs removal, and 20 models for the other pollutants removal. The majority of the ANN models present a good modeling and optimization ability (indicated by MSE and R^2).

3.1. Datasets used for artificial intelligence approaches in water treatment

In order to evaluate the performance of artificial intelligence approaches, it is required that the related datasets define their level

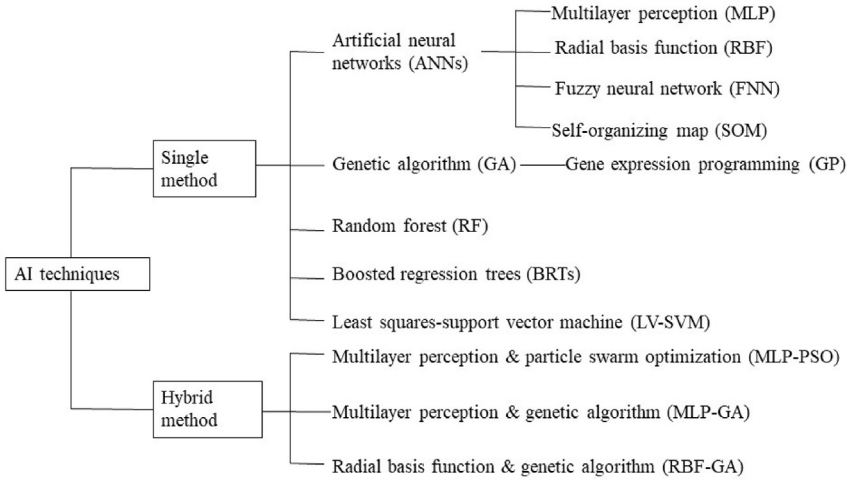

Fig. 1. A classification tree of AI techniques for pollutants removal in water treatment.

of accuracy. (Kalantari et al., 2017). In Table 2, the classification of public datasets is shown, which were employed by researchers for modeling the removal processes of nutrients, heavy metals, POPs and other pollutants in water treatment.

3.2. Nutrients removal

A number of researchers have studied the BP-ANN model and recommended it for the modeling and optimization of the water treatment (Akratos et al., 2009; Kundu et al., 2013; Santin et al., 2016). Kundu et al. (2013) employed the BP-ANN model to predict the performance of a laboratory scale sequencing batch reactor. This model possesses a distinctive ability of learning nonlinear functional relationship without requiring the sophisticated underlying mechanism of the complex process. This study indicated that the ANN model has a great potential for the prediction of nutrients removal efficiency in biological systems. Akratos et al. (2009) also developed the BP-ANN model to predict the phosphorous removal in horizontal subsurface flow constructed wetlands (HSFCWs). The modeling results showed a satisfactory correlation between the experimental and predicted data, providing a theoretical guidance for the removal of nutrients in wastewater by using a physical method. Moreover, Santin et al. (2016) proposed a BP-ANN model to support the design of riparian buffer strips (RBS) as a function of the desired nitrogen filtering effectiveness. Based on this result, the BP-ANN model can effectively capture the relations among the variables related to the mitigation potential of nutrient pollution.

For BP-ANN, the individual relations between the input and output variables cannot be developed by engineering judgment so that the model tends to be a black box, while FNN is transparent and hence its if-then rules are ready to understand and interpret. A FNN model and a BP-ANN model were proposed by Huang et al. (2009b) for estimating the nutrient concentrations and overcoming the problem of delayed measurements. The FNN model can be effectively applied to a bench-scale anoxic/oxic process to cope with influent variations, which are typical of municipal wastewater. Furthermore, this result showed that the outcome of the FNN model is better than that of the BP-ANN model due to the less error. The FNN modeling technique has many favorable features (e.g., efficiency, generalization and simplicity), which make it an attractive choice for modeling the complex systems.

The RBF networks have also been successfully applied for solving the dynamic system problems, because they can predict the behavior directly from the input and output data. Bagheri et al. (2014) modeled a sequencing batch reactor using the MLP-ANN and the RBF-ANN to predict the effluent concentrations of total phosphorus and ammonia. The results indicated that output data of the MLP-ANN model with higher R^2 value and lower root mean square error are more precise than those of the RBF-ANN model. The MLP-ANN model has a stronger approximation and generalization ability than the RBF-ANN model. The precision of MLP-ANN and RBF-ANN models is determined not only by the amount of input data to train the networks but also by the correlation of these data. The high correlation of the input data affected the MLP-ANN model more than the RBF-ANN model, on the contrary, the amount of input data affected the RBF-ANN model more than the MLP-ANN model.

The SOM model is based on an unsupervised neural network algorithm, which has been used to analyze, cluster and model various types of large databases (Zhang et al., 2008). However, the SOM model has unfrequently been implemented in water treatment processes in comparison to traditional neural networks. The SOM model was applied to predict the outflow nutrient concentrations for integrated constructed wetlands (ICWs) treating the farmyard runoff (Zhang et al., 2008). This model performed well in predicting the nutrient concentrations, which was measured cost-effectively. It is demonstrated that the SOM model is an appropriate approach for monitoring wastewater treatment processes in ICWs.

Several studies took advantage of the GA coupled with ANN to generate optimum operational variables for the studied process because GA does not easily get trapped in a local minimum. Zhang and Pan (2014) explored the potential of ANN and RSM for modeling phosphate removal from water by the nanocomposite absorbent HFO-201. In combination with GA, the optimized ANN model exhibited an acceptable predictive strength for the experimental runs, indicating its potential for predicting the behavior of complex water treatment processes. Moreover, ANOVA tests and sensitivity analysis were employed to find the relative importance of independent variables. The order for these variables influencing the removal process was: sulfate concentration > adsorbent dosage > pH > temperature.

3.3. Heavy metals removal

It is apparent that the numerous applications of BP-ANN have been successfully applied in the removal of heavy metals because of their reliable, robust and salient characteristics for establishing the non-linear relationships between variables in water treatment

Table 1
Different ANN models and heuristic algorithms for applications in the pollutants removal from wastewater.

No.	Input and output variables		Types of the ANN model	Sizes of training, validation and testing datasets			MSE	R^2	References
	Input variables	Output variables		Training data sets	Validation data sets	Testing data sets			
1	Initial COD, NH_4^+-N, MLVSS and DO concentrations, contact time and pH	COD and NH_4^+-N removal efficiencies	BP-ANN	15	7	7	1.32 (COD) 1.63 (NH_4^+-N)	0.96 (COD) 0.958 (NH_4^+-N)	Kundu et al. (2013)
2	Inverse specific, wastewater temperature and hydraulic residence times	*Ortho*-phosphate and total P removal efficiencies	MLP-ANN	50%	25%	25%	0.17	0.58	Akratos et al. (2009)
3	Vegetation cover type, soil type, mean nitrogen influent and removal effectiveness	Buffer width	BP-ANN	70	15	15	–	–	Santin et al. (2016)
4	Oxidation-reduction potential, pH and dissolved oxygen	Nitrogen and COD removal efficiencies	BP-ANN FNN	29	3	–	0.88 (FNN) 2.60 (BP-ANN)	–	Huang et al. (2009b)
5	Influent concentration, filling time, reaction time, aeration intensity, SRT and MLVSS	Effluent concentrations of TSS, TP, COD and NH_4^+-N	MLP-ANN RBF-ANN	70%	15% (MLP-ANN)	15% (MLP-ANN) 30% (RBF-ANN)	1.06 (MLP-ANN) 0.24 (RBF-ANN)	0.98 (MLP-ANN) 0.99 (RBF-ANN)	Bagheri et al. (2014)
6	DO, temperature, pH, chloride and conductivity	Concentrations of ammonia-nitrogen and soluble reactive phosphorus	SOM	245	–	79	0.03	0.943	Zhang et al. (2008)
7	Initial pH, sulfate concentration, operation temperature and adsorbent dosage	Phosphate removal efficiency	BP-ANN-GA	25	8	4	3.4	0.99	Zhang and Pan (2014)
8	Initial pH, As(III) concentration, contact time, temperature, material dose and agitation speed	As(III) removal efficiency	BP-ANN	63	–	42	0.29	0.98	Mandal et al. (2015)
9	Initial concentrations of Cd(II), Ni(II), Pb(II) and pH	Cd(II), Ni(II) and Pb(II) removal efficiencies	BP-ANN	158	34	33	5	0.97	Reynelavila et al. (2015)
10	Sorbent characteristics and metal ion properties	Cd(II), Ni(II), Zn(II) and Pb(II) removal efficiencies	BP-ANN	183	39	39	0.06	0.99	Mendozacastillo et al. (2015)
11	Initial concentrations of Pb(II) and MG, materials dosage, pH and ultrasonication time	Pb(II) and MG removal efficiencies	MLP-ANN	20	6	6	0.0010 (Pb(II)) 0.0040 (MG)	0.9997 (Pb(II)) 0.9999 (MG)	Dil et al. (2017)
12	Initial concentrations of Cd(II) and MB, pH, adsorbent mass and contact time	Cd(II) and MB removal efficiencies	BP-ANN BRT	36	8	8	$2.27*10^{-5}$ (BP-ANN) $1.33*10^{-5}$ (BRT)	0.9896 (BP-ANN) 0.9912 (BRT)	Mazaheri et al. (2017)
13	Initial Cr(VI) concentration, pH, temperature and CuONPs dosage	Cr(VI) removal efficiency	BP-ANN-GA	24	–	7	0.06	0.995	Mohan et al. (2015)
14	Contact time, pH, adsorbent dosage and initial Pb(II) concentration	Pb(II) removal efficiency	BP-ANN-GA	19	3	3	0.14	0.999	Yasin et al. (2014)
15	Amount of tea waste, pH, concentrations of complexing agent and eluent, eluent volume and eluent flow rates	Mn(II) and Co(II) removal efficiencies	BP-ANN-PSO	76%	12%	12%	0.01 (Mn(II)) 0.0026 (Co(II))	0.9807 (Mn(II)) 0.9838 (Co(II))	Khajeh et al. (2013b)
16	Temperature, pH, Cd(II) concentration, contact time and agitation rate	Cd(II) removal efficiency	BP-ANN FNN	15	–	–	8.41(BP-ANN) 4.22 (FNN)	0.888 (BP-ANN) 0.961 (FNN)	Singh et al. (2006)
17	Emulsification time, ultrasonic power, stirring speed, sulfuricacid concentration, etc	Cu(II) removal efficiency	RBF-ANN	–	–	–	0.0016	0.997	Messikh et al. (2014)
18	Turbidity, conductivity, redox potential, outflow water temperature, DO and pH	Cu(II) removal efficiency	SOM	182	–	66	0.5	0.97	Lee and Scholz (2006)
19	Initial carbaryl concentration, pH, biomass dose and contact time	Carbaryl removal efficiency	MLP-ANN	22	–	7	0.523	0.921	Chattoraj et al. (2014)
20	Contact time, pH, temperature and initial trichlorophenol concentration	Trichlorophenol removal efficiency	MLP-ANN	–	–	–	0.3489	0.999	Dlamini et al. (2014)
21	Retention time, influent feed and COD	Aniline removal efficiency	BP-ANN	300	–	120	0.0018	0.99	Delnavaz et al. (2010)
22	Initial naphthalene concentration, fluence rate, salinity, temperature and contact time	Naphthalene removal efficiency	BP-ANN	116	38	38	0.0018	0.943	Jing et al. (2014)
23	Contact time, pH, temperature, adsorbent concentration and initial triamterene concentration	Triamterene removal efficiency	BP-ANN-GA	45	–	19	0.0005	0.9856	Ghaedi et al. (2016)
24	Toluene inlet concentration, retention time, temperature, moisture content of the biofilter bed and pH	Toluene removal efficiency	RBF-ANN	40	–	27	3.03	0.9755	Deshmukh et al. (2012)
25	Current density, electrolysis time, initial pH and dye concentration, conductivity, retention time of sludge and distance between electrodes	Color removal efficiency	BP-ANN	25	12	12	–	0.974	Daneshvar et al. (2006)

Table 1 (*continued*)

No.	Input and output variables		Types of the ANN model	Sizes of training, validation and testing datasets			MSE	R^2	References
	Input variables	Output variables		Training data sets	Validation data sets	Testing data sets			
26	Initial concentrations of MB, BG and CV, pH, adsorbent dosage, sonication time,	Removal efficiencies of MB, BG and CV	BP-ANN	70%	15%	15%	0.0004 −0.0011	0.9976 −0.9986	Asfaram et al. (2015)
27	Initial concentrations of MB and MG, pH, adsorbent mass and ultrasonication time	Removal efficiencies of MB and MG	BP-ANN	70%	15%	15%	0.0006 (MB) 0.0012 (MG)	0.9997 (MB) 0.9990 (MG)	Asfaram et al. (2016a)
28	Initial MB concentration, pH, adsorbent mass and sonication time	MB removal efficiency	BP-ANN LS-SVM	75%	–	25%	$9.79*10^{-7}$ (BP-ANN) $2.64*10^{-8}$ (LS-SVM)	0.9984 (BP-ANN) 0.9995 (LS-SVM)	Asfaram et al. (2016b)
29	Initial MG, DB and MB concentrations, adsorbent mass and sonication time	Removal efficiencies of MG, DB and MB	BP-ANN	70%	15%	15%	0.00001 −0.00006	0.9989 −0.9993	Bagheri et al. (2016)
30	Initial concentrations of IC and SO, adsorbent mass and sonication time	Removal efficiencies of IC and SO	BP-ANN	70%	15%	15%	0.00006 (IC) 0.00005 (SO)	0.9991 (IC) 0.9997 (SO)	Dastkhoon et al. (2017)
31	Adsorbent size, initial pH, initial dye concentration and contact time	Lanaset Red G removal efficiency	BP-ANN-GA	784	184	184	1.3314	0.999	Çelekli et al. (2012)
32	Adsorbent dosage, pH, contact time and initial MG concentration	Malachite green removal efficiency	BP-ANN-GA	186	–	62	0.0017	0.9658	Ghaedi et al. (2015b)
33	Initial CG concentration, adsorbent mass and sonication time	CG removal efficiency	Random forest	–	–	–	0.0045	0.9315	Bagheri et al. (2015)
34	Influent flow rate, return and excess sludge flow rate, influent COD concentration, total kjeldahl nitrogen, influent concentration of suspended solids. dissolved oxygen concentration and chemical oxygen load	COD removal efficiency	BP-ANN	240	–	50	2.64	0.91	Güçlü and Sükrü Dursun (2008)
35	Amount of silver nanoparticles, pH, volume and flow of sample and eluent	MB removal efficiency	BP-ANN-PSO	76%	12%	12%	1.1881	0.98	Khajeh et al. (2013a)
36	Initial dye concentration, amount of ZnS-NP-AC and contact time	Brilliant green removal efficiency	BP-ANN-PSO	252	–	108	0.0021	0.9558	Ghaedi et al. (2015a)
37	Initial methyl orange (MO) concentration, adsorbent dosage and contact time	MO removal efficiency	BP-ANN-PSO	270	–	90	0.0009	0.97	Agarwal et al. (2016)
38	Reaction time, H_2O_2/COD and H_2O_2/Fe^{2+} molar ratio, pH and antibiotic concentration	COD removal efficiency	BP-ANN	60	30	30	0.0004	0.997	Elmolla et al. (2010)
39	Porosity, wastewater temperature and hydraulic residence times	BOD removal efficiency	BP-ANN	406	203	203	–	0.68	Akratos et al. (2008)
40	Reaction time, pH, herbicide concentration, contaminant, US ultrasound, UV light intensity, $[TiO_2]_o$, $[K_2S_2O_8]_o$, and SR solar radiation	COD removal efficiency	BP-ANN	138	69	69	0.0003	0.9913	Hamzaoui et al. (2011)
41	MLSS, HRT and contact time	COD removal efficiency	BP-ANN	70%	15%	15%	0.1486	0.9999	Hazrati et al. (2017) d
42	Initial pH, $[H_2O_2]/[Fe^{2+}]$ mole ratio and Fe(II) dosage	Mass content ratio and mass removal efficiency of COD	MLP-ANN	11	3	4	1.62	0.9767	Sabour and Amiri (2017)
43	Adsorbent mass, pH, sonication time, initial MB and MG concentrations	MB and MG removal efficiencies	BP-ANN RBF-ANN	70%	15%	15%	$1.06*10^{-5}$ (MPL-ANN) $2.88*10^{-6}$ (RBF-ANN)	0.9878 (MLP-ANN) 0.9991 (RBF-ANN)	Asfaram et al. (2017)
44	Influent water quality and chemical dosages	COD removal efficiency	FNN	32	–	32	0.001	–	Huang et al. (2009a)

(Mandal et al., 2015; Mendozacastillo et al., 2015; Reynelavila et al., 2015). To reduce the cost of experimental runs, Mandal et al. (2015) made the prediction for the adsorption of the As(III) onto cerium hydroxylamine hydrochloride using the BP-ANN model. Comparison of the experimental and predicted results showed that this model can predict the As(III) adsorption efficiency with acceptable accuracy. Reynelavila et al. (2015) also used a BP-ANN approach for the analysis and modeling of multi-component sorption data. The results indicated that the application of ANN offers several advantages in the analysis and modeling of the multi-component sorption of heavy metal ions from an aqueous solution. In addition, the BP-ANN model was utilized by Mendozacastillo et al. (2015) for predicting the sorption of heavy metals onto lignocellulosic biomasses using relevant characteristics and properties of both the adsorbent and the metallic pollutant. They demonstrated that the BP-ANN model can be used for predicting the adsorption performance of ligoncellulosic biomasses in other operational conditions. Furthermore, a feed-forward neural network was successfully applied by Dil et al. (2017) for the ultrasound-assisted simultaneous removal of Pb(II) from aqueous solutions. The values of R^2 and MSE were 0.9997 and 0.9999 from the ANN model with 14 neurons in the hidden layer for the Pb(II) removal. Mazaheri et al. (2017) also employed the boosted regression tree (BRT), BP-ANN and RSM for modeling the adsorption of Cd(II) onto walnut carbon and computationally investigating the influence of important variables. It can be concluded that though RSM is the widely used approach

Table 2
Datasets classifications for the various pollutants removal in the water treatment.

Dataset type	Description	Size	Reference
NH_4^+-N removal	Total datasets were obtained from experimental results in the laboratory.	29 samples	Kundu et al. (2013)
Phosphorus removal	The used datasets were generated from the RSM-CCD design.	74 samples	Zhang and Pan (2014)
As(III) removal	Total datasets were obtained from experimental results in the laboratory.	105 samples	Mandal et al. (2015)
Pb(II), Cd(II), Ni(II) and Zn(II) removal	The datasets were gathered from kinetic and isotherm experiments.	261 samples	Mendozacastillo et al. (2015)
Cr(VI) removal	The used datasets were generated from the RSM-CCD design.	31 samples	Mohan et al. (2015)
Cu(II) removal	The used datasets were collected by monitoring the effluent concentration of all 12 filter for more than 2 years.	248 samples	Lee and Scholz (2006)
Carbaryl removal	The used datasets were generated from the RSM-BBD design.	29 samples	Chattoraj et al. (2014)
Aniline removal	The used datasets were collected from the moving bed biofilm reactor in 3 months continuous period.	420 samples	Delnavaz et al. (2010)
Naphthalene removal	The used datasets were obtained from the previous study (Jing et al., 2013).	192 samples	Jing et al. (2014)
Toluene removal	The used datasets were collected from the biofilter system.	67 samples	Deshmukh et al. (2012)
MB removal	The used datasets were generated by the RSM-CCD design.	30 samples	Asfaram et al. (2016b)
COD removal	The used datasets were obtained from the previous study (Elmolla and Chaudhuri, 2009).	120 samples	Elmolla et al. (2010)
BOD removal	The used data came from five pilot-scale horizontal subsurface flow constructed wetlands for two years	812 samples	Akratos et al. (2008)

for optimization of the dye removal, the BRT and ANN approaches may present a good alternative for a limited dataset.

GA and PSO are expected to avoid the local optima frequently by promoting the exploration of the search space, which are different from the local search algorithms (e.g., the gradient descent). The adsorption parameters for CuO nanoparticles were manually optimized by Mohan et al. (2015) in batch mode through one-factor-at-a-time approach followed by reoptimization of the process with statistical and computational methods (RSM and BP-ANN-GA). Using the BP-ANN-GA model, the Cr(VI) removal efficiency from water solutions was improved by 8.1%, as compared to that of RSM. Besides, Yasin et al. (2014) developed a BP-ANN-GA model for the simulation and optimization of lead ions removal from aqueous solutions. The results of sensitivity analysis showed that the solution pH is the most effective parameter for the lead ions removal. The residual error between the predicted and experimental values was 2.5%, which confirmed the validity of the established BP-ANN-GA model. Furthermore, the RSM and BP-ANN-PSO models were developed by Khajeh et al. (2013b) to model and optimize the tea waste extraction process for the Mn(II) and Co(II) removal. The generalization and predictive capabilities of both RSM and BP-ANN-PSO have demonstrated the superiority of BP-ANN-PSO in capturing the non-linear relationship of the removal process.

Singh et al. (2006) developed a FNN model and a BP-ANN model for the prediction of adsorption of cadmium on hematite. The FNN model was proven to be more efficient in predicting cadmium adsorption than the BP-ANN model. Besides, the RBF-ANN model was applied by Messikh et al. (2014) in the prediction for the copper removal efficiency using the emulsion liquid membrane process. The results obtained from this model were close to the experimental data, hence the performance of the RBF-ANN model was impressive for predicting the copper removal efficiency. On the one hand, the RBF-ANN is a feed-forward and local adjustment network, which makes its training rate faster than other neural networks; on the other hand, the units typically required in the hidden layer of RBF-ANN model are more than those in the hidden layer of MLP-ANN model. Such a large number of units offer a great flexibility to RBF networks, but exposes their analysis to the danger of over-fitting. Furthermore, Lee and Scholz (2006) developed the SOM model to predict the copper concentration in constructed wetlands for treating the urban runoff. The accuracy of prediction with SOM was satisfactory based on monitoring in real time. Moreover, the SOM model provides an outstanding visual information to identify the relationships between the variables describing complex removal processes.

3.4. POPs removal

The applications of BP-ANN to simulate wastewater treatment processes, especially POPs removal, have recently gained an increasing attention (Delnavaz et al., 2010; Chattoraj et al., 2014; Dlamini et al., 2014; Jing et al., 2014). The *Lemna major* biomass was used by Chattoraj et al. (2014) for the removal of carbaryl from aqueous solutions. Both the RSM model and the BP-ANN model were applied to simulate and optimize the adsorption process, and the results showed that the predictive capability of the RSM model is higher than that of the ANN model. In addition, Dlamini et al. (2014) developed a BP-ANN model for predicting the trichlorophenol (TCP) adsorption by odorata stem. It was demonstrated that the ANN model can be used to simulate the removal of TCP from aqueous solutions using the Chromolaena odorate stem. Moreover, Delnavaz et al. (2010) also used the BP-ANN to establish the models in predicting aniline treatment efficiency using a moving bed biofilm reactor (MBBR) in various conditions, such as retention time and influent COD. This BP-ANN model was found to provide an efficient and robust tool in predicting MBBR performance for treating aromatic amine compounds. Furthermore, Jing et al. (2014) were focused on developing a simulation model for the photodegradation of naphthalene in seawater by using the ANN model. The findings of this study showed that the ANN model can effectively predict the process of the photo-induced PAH degradation. The results of ANOVA analysis demonstrated that the order for the independent variables influencing the removal process was: fluence rate > temperature > reaction time > salinity > initial concentration. This was in good agreement with the sensitivity analysis results by using the Garson approach.

ANNs may not solve all problems in the real world because they have some limitations, i.e. it does not guarantee the global optimal solutions. Hence, GA has been the most well-known AI methodology for optimization problems since it requires no smoothness, derivability or continuity of an objective function and is able to reach the global optimum in a random but systematical search. Ghaedi et al. (2016) modeled the effect of influential variables by using the multiple linear regression (MLR) and the ANN-GA for the rapid adsorption of triamterene with the single-walled carbon nanotubes and multiwalled carbon nanotubes. They concluded that the performance of the ANN-GA model is better than that of the MLR model.

A comparative modeling study has been carried out by Deshmukh et al. (2012) using RSM and RBF-ANN to predict and optimize the performance of a biofilter system for treating toluene. The prediction errors for the RSM model and the RBF-ANN model were 7.76% and 3.03%, respectively, and the R^2 values obtained were 0.8826 and 0.9755, respectively. The results indicated the superiority of the RBF-ANN model in the prediction capability due to its ability to approximate the high degree of non-linearity between the input and output variables.

3.5. Other pollutants removal

Daneshvar et al. (2006) investigated the decolorization of the textile methine dye solution containing C. I. Basic Yellow 28 by using a BP-ANN model. This ANN model successfully described the behavior of the electrocoagulation process, thus the model could estimate the behavior of the system under different conditions. RSM and BP-ANN were also employed by Asfaram et al. (2015, 2016a,b) to model and optimize the simultaneous dyes adsorption onto various materials. The BP-ANN model demonstrated a better predictability than the RSM model based on the values of R^2, root mean square error (RMSE), mean absolute error (MAE) and absolute average deviation (AAD). Bagheri et al. (2016) focused on the combination of RSM and BP-ANN to optimize and model the entire adsorption following ternary dyes concentration evaluation by derivative spectrophotometry. A satisfactory agreement was obtained between the values predicted by using the ANN model and the experimental data. RSM and BP-ANN were used by Dastkhoon et al. (2017) to model the removal process and predict the dyes (Safranin-O (SO) and Indigo Carmine (IC)) removal capacity of nanowires under different conditions. The sensitivity analysis of BP-ANN model demonstrated that sonication time was the essential and inessential factor for the SO and IC removal with the relative importance of 36.62% and 12.60%, respectively. In addition, Çelekli et al. (2012) selected the lentil straw to remove an azo metal complex dye by using an ANN-GA model. The results indicated that the initial dye concentration has the strongest effect on dye uptake, followed by pH. Ghaedi et al. (2015b) used the copper nanowires loaded on activated carbon for the removal of malachite green (MG) by using MLR and BP-ANN-GA. The low value of MSE and high value of R^2 demonstrated the superiority of the BP-ANN-GA model. Furthermore, the random forest (RF) model and the RSM model were developed by Bagheri et al. (2016) for the ultrasonic-assisted chrysoidine G removal by copper sulfide nanoparticles loaded on activated carbon. The results showed that the experimental data were in a good agreement with the data predicted by the RF model.

Güçlü and Sükrü Dursun (2008) developed an ANN model that successfully predicted the effluent COD concentrations and monitored the large-scale wastewater treatment plant performance. The results showed that this ANN model can be used as a general modeling tool in a wide variety of other treatment systems. Elmolla et al. (2010) investigated the implementation of BP-ANN for the prediction of antibiotic degradation in terms of COD removal by the Fenton process. The results demonstrated that neural network modeling can effectively simulate and predict the behavior of the process. Akratos et al. (2008) have successfully used BP-ANN to optimize the removal of biochemical oxygen demand and COD in horizontal subsurface flow constructed wetland. In addition, a direct and inverse ANN approach was employed by Hamzaoui et al. (2011) to predict the COD removal during the degradation of alazine and gesaprim commercial herbicides. The good correlation and small error obtained from the ANN model indicated a high agreement between the predicted data and experimental data. Hazrati et al. (2017) developed a BP-ANN for predicting the COD concentration to meet the effluent discharge standards in a membrane bioreactor treating petrochemical wastewater at various hydraulic retention times. This BP-ANN model was proven to be a good approach with 0.9999 of R^2 and 0.1486 of MSE. Furthermore, RSM and BP-ANN were used by Sabour and Amiri (2017) to discuss the Fenton process performance (COD removal) in landfill leachate treatment. The better results of R^2 (0.97–0.98), RMSE (1.45–1.86) and error (2–4%) demonstrated the relative superiority of BP-ANN model compared to that of the RSM model.

Other neural networks (e.g., RBF-ANN, FNN and SOM) have also been used for other pollutants removal in wastewater treatment. Asfaram et al. (2017) employed the BP-ANN and RBF-ANN approaches to model and optimize the removal efficiency of methylene blue and MG. The values of RMSE, MAE and AAD obtained from BP-ANN and RBF-ANN models show a good predictability of the RBF-ANN model in comparison with that of the BP-ANN model. In addition, Huang et al. (2009a) developed a FNN model predictive control scheme for studying the coagulation process in a paper mill. The results indicated that the FNN model effectively tackled both the environmental and economic objectives of the system in a real-time basis, although a disadvantage of this model is its high dependence on the quality of training data. It is common for a raw database to contain some redundant and conflicting data. Thus, it is sometimes necessary for the raw training database to be pre-treated to remove redundancies and resolve conflicts in the data. Furthermore, excitation-emission matrix (EEM) spectra coupled with SOM was employed by Du et al. (2012) for the investigation of organic matter removal in a water treatment plant. The results indicated that the SOM technique was used as an effective tool for EEM spectra analysis. This is helpful for the modeling and optimization of water treatment process parameters and the improvement of its performance.

3.6. Comparison between the performances of single methods and hybrid methods

In order to get an overall comprehension, the obtained values of MSE and R^2 for single methods and hybrid methods are compared in Figs. 2 and 3. The majority of MSE values obtained from the BP-ANN, BRT, LS-SVM and RF models were lower than those from MLP-ANN, RBF-ANN and FNN (Fig. 2). The R^2 values obtained from the BP-ANN, RBF-ANN, BRT and LS-SVM models were higher than those from MLP-ANN, FNN and RF models. In summary, the BP-ANN, BRT and LS-SVM are the most successful single methods among AI techniques for the pollutants removal in water treatment. In addition, the performances of hybrid methods for the pollutants removal generally demonstrated their accuracy and robustness with the low MSE values and high R^2 values (Fig. 3).

4. Conclusions and prospects

AI techniques (e.g., ANNs, GA and PSO) are presently considered as attractive alternative tools in the environmental engineering due

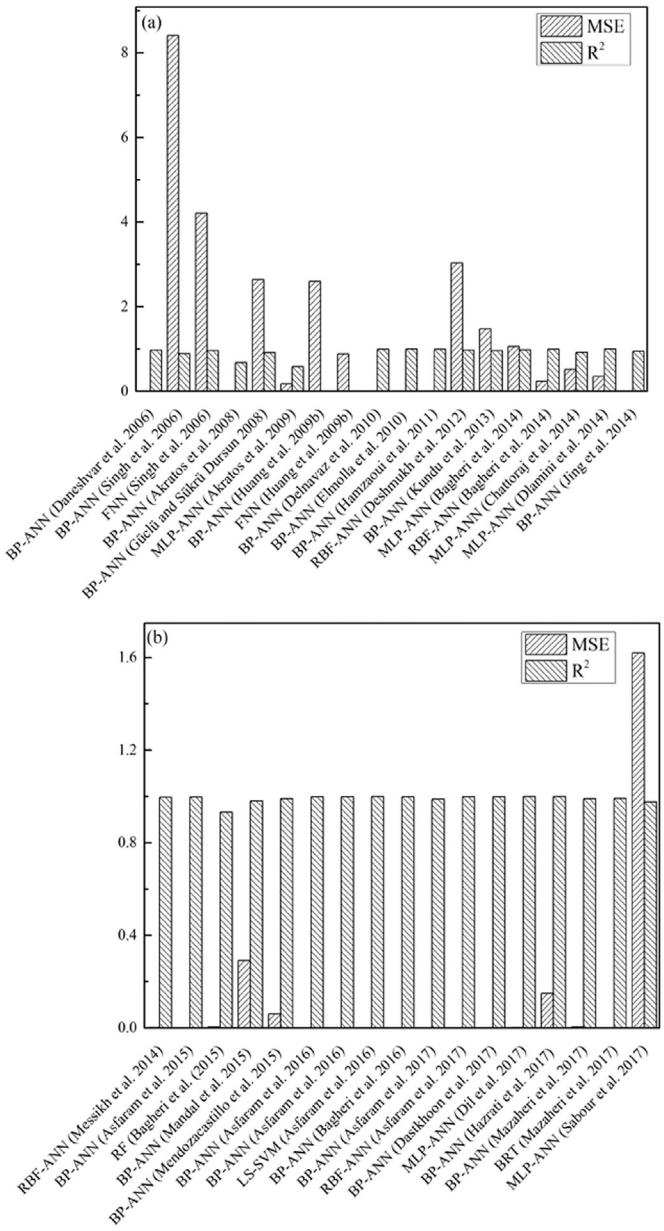

Fig. 2. Comparison of the MSE and R^2 value obtained from single methods for the pollutants removal in water treatment.

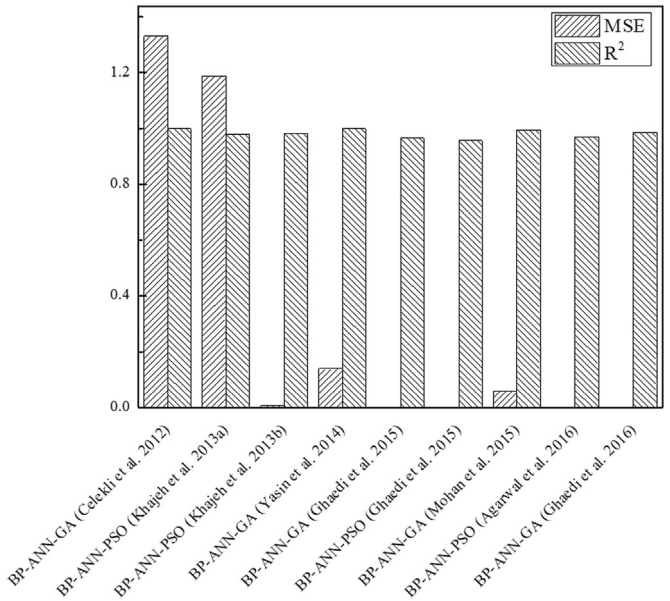

Fig. 3. Comparison of hybrid methods in terms of the MSE and R^2 values.

to their specific features, such as non-linearity, adaptivity and generalization (Mohebbi et al., 2008). ANNs are good AI tools for optimizing the removal processes because of their following advantages: (i) in comparison to other approaches, ANNs can model a multivariable system to extract complex nonlinear relationships between the variables by means of the data training; (ii) ANNs can overcome the limitations of conventional mathematical models by extracting the required information using the training data, which have not required a prior specification of suitable fitting functions. The disadvantages of ANNs are as follows: (i) ANNs have a poor reproducibility because the weight and bias between neurons are given randomly; (ii) ANNs are easy to fall into a local optimal solution, thus they cannot provide an optimal condition for the engineering application in water treatment. Therefore, hybridization of ANNs with other AI tools (such as SA, MCS, PSO, IA, ACA, ICA and DT) have a potential to be used for generating optimal operational variables for the removal processes due to the ability to search for the global optima.

ANN-GA and ANN-PSO models have recently been developed for modeling and optimization of the pollutant removal processes. The training data for ANN are usually selected using either a random sampling method or an experimental design (e.g., RSM and OD). However, the random sampling method does not ensure that the training data uniformly cover the domain, especially when the number of input variables is large while training dataset is relatively small. The main drawback of RSM is that a relatively large input dataset is required to build a sufficiently accurate response surface. Therefore, uniform design (UD) in combination with ANNs can be used for predicting the efficiency of the pollutants removal processes in laboratories especially when the experiment has many variables and multiple levels or the experimental cost is highly expensive. The reason for this is that the experimental trials for UD are slightly fewer than those for OD and significantly fewer than those for RSM. Finally, if we need to model the removal processes of operational plants with real-time sensors, we can utilize ANNs to treat useful data available directly from the actual operation of the plants rather than from experimental design.

Deep learning models (such as convolutional neural network and DBN) can be used to learn large volumes of data because their learning capacities can be controlled by varying their depth and breadth. Thus, convolutional neural network and DBN have a potential use for the modeling and optimization of the pollutants removal processes that have large and complex datasets (i.e. big data).

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant No. 21667012 and the National 111 Project of China under Grand No. D17016.

List of abbreviations

AAD	Absolute average deviation
ACA	Ant colony algorithm
AI	Artificial intelligence

ANN	Artificial neural network
ANOVA	Analysis of variance
BBD	Box-Behnken design
BOD	Biochemical oxygen demand
BP-ANN	Back propagation artificial neural network
BRT	Boosted regression tree
CCD	Central composite designs
CNN	Chaotic neural network
COD	Chemical oxygen demand
DBN	Deep belief network
DM	Doehlert matrix
DO	Dissolved oxygen
DT	Decision tree
FNN	Fuzzy neural network
GA	Genetic algorithm
IA	Immune algorithm
ICA	Imperialist competitive algorithm
MAE	Mean absolute error
MCS	Monte Carlo simulation
MLP	Multilayer perception
MSE	Mean square error
OD	Orthogonal design
POPs	Persistent organic pollutants
PSO	Particle swarm optimization
RBF	Radial basis function
RF	random forest
RMSE	Root mean square error
RNN	Recurrent neural network
RSM	Response surface methodology
SA	Simulated annealing
SM	Simplex method
SOM	Self-organizing map
SVM	Support vector machine
UD	Uniform design

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.chemosphere.2018.02.111.

References

Acharya, C., Mohanty, S., Sukla, L.B., Misra, V.N., 2006. Prediction of sulphur removal with Acidithiobacillus sp. using artificial neural networks. Ecol. Model. 190, 223–230.

Agarwal, S., Tyagi, I., Gupta, V.K., Ghaedi, M., Masoomzade, M., Ghaedi, A.M., Mirtamizdoust, B., 2016. Kinetics and thermodynamics of methyl orange adsorption from aqueous solutions-artificial neural network-particle swarm optimization modeling. J. Mol. Liq. 218, 354–362.

Akratos, C.S., Papaspyros, J.N.E., Tsihrintzis, V.A., 2008. An artificial neural network model and design equations for BOD and COD removal prediction in horizontal subsurface flow constructed wetlands. Chem. Eng. J. 143, 96–110.

Akratos, C.S., Papaspyros, J., Tsihrintzis, V.A., 2009. Artificial neural network use in ortho-phosphate and total phosphorus removal prediction in horizontal subsurface flow constructed wetlands. Biosyst. Eng. 102, 190–201.

Aleboyeh, A., Kasiri, M.B., Olya, M.E., Aleboyeh, H., 2008. Prediction of azo dye decolorization by UV/H_2O_2 using artificial neural networks. Dyes Pigments 77, 288–294.

Amirov, A., Gerget, O., Devjatyh, D., Gazaliev, A., 2014. Medical data processing system based on neural network and genetic algorithm. Procedia Soc. Behav. Sci. 131, 149–155.

Arhami, M., Kamali, N., Rajabi, M.M., 2013. Predicting hourly air pollutant levels using artificial neural networks coupled with uncertainty analysis by Monte Carlo simulations. Environ. Sci. Pollut. R 20, 4777–4789.

Asfaram, A., Ghaedi, M., Hajati, S., Goudarzi, A., 2015. Ternary dye adsorption onto MnO_2 nanoparticle-loaded activated carbon: derivative spectrophotometry and modeling. Rsc Adv 5, 72300–72320.

Asfaram, A., Ghaedi, M., Hajati, S., Goudarzi, A., 2016a. Synthesis of magnetic γ-Fe_2O_3-based nanomaterial for ultrasonic assisted dyes adsorption: modeling and optimization. Ultrason. Sonochem. 32, 418–431.

Asfaram, A., Ghaedi, M., Azqhandi, M.H.A., Goudarzi, A., Dastkhoon, M., 2016b. Statistical experimental design, least square-support vector machine (LS-SVM) and artificial neural network (ANN) methods for modeling of facilitated adsorption of methylene blue dye. Rsc Adv 6, 40502–40516.

Asfaram, A., Ghaedi, M., Azqhandi, M.H.A., Goudarzi, A., Hajati, S., 2017. Ultrasound-assisted binary adsorption of dyes onto Mn@CuS/ZnS-NC-AC as a novel adsorbent: application of chemometrics for optimization and modeling. J. Ind. Eng. Chem. 54, 377–388.

Assefi, P., Ghaedi, M., Ansari, A., Habibi, M.H., Momeni, M.S., 2014. Artificial neural network optimization for removal of hazardous dye Eosin Y from aqueous solution using Co_2O_3-NP-AC: isotherm and kinetics study. J. Ind. Eng. Chem. 20, 2905–2913.

Bagheri, A.R., Ghaedi, M., Hajati, S., Gahedi, A.M., Goudarzi, A., Asfaram, A., 2015. Random forest model for the ultrasonic-assisted removal of chrysoidine G by copper sulfide nanoparticles loaded on activated carbon; response surface methodology approach. Rsc Adv 5, 59335–59343.

Bagheri, A.R., Ghaedi, M., Asfaram, A., Hajati, S., Ghaedi, A.M., Bazrafshan, A., Rahimi, M.R., 2016. Modeling and optimization of simultaneous removal of ternary dyes onto copper sulfide nanoparticles loaded on activated carbon using second-derivative spectrophotometry. J. Taiwan Inst. Chem. E 65, 212–224.

Bagheri, M., Mirbagheri, S.A., Ehteshami, M., Bagheri, Z., 2014. Modeling of a sequencing batch reactor treating municipal wastewater using multi-layer perceptron and radial basis function artificial neural networks. Process Saf. Environ 93, 111–123.

Çelekli, A., Bozkurt, H., Geyik, F., 2012. Use of artificial neural networks and genetic algorithms for prediction of sorption of an azo-metal complex dye onto lentil straw. Bioresour. Technol. 129C, 396–401.

Chang, H., Hou, W.C., 2006. Optimization of membrane gas separation systems using genetic algorithm. Chem. Eng. Sci. 61, 5355–5368.

Chattoraj, S., Mondal, N.K., Das, B., Roy, P., Sadhukhan, B., 2014. Carbaryl removal from aqueous solution by Lemna major biomass using response surface methodology and artificial neural network. J. Environ. Chem. Eng 2, 1920–1928.

Cheung, K.C., Poon, B.H.T., Lan, C.Y., Wong, M.H., 2003. Assessment of metal and nutrient concentrations in river water and sediment collected from the cities in the Pearl River Delta, South China. Chemosphere 52, 1431–1440.

Chu, K.H., 2003. Prediction of two-metal biosorption equilibria using a neural network. Eur. J. Miner. Process. Environ. Protect. 3, 119.

Daneshvar, N., Khataee, A.R., Djafarzadeh, N., 2006. The use of artificial neural networks (ANN) for modeling of decolorization of textile dye solution containing C. I. Basic Yellow 28 by electrocoagulation process. J. Hazard Mater. 137, 1788–1795.

Dastkhoon, M., Ghaedi, M., Asfaram, A., Azqhandi, M.H.A., Purkait, M.K., 2017. Simultaneous removal of dyes onto nanowires adsorbent use of ultrasound assisted adsorption to clean waste water: chemometrics for modeling and optimization, multicomponent adsorption and kinetic study. Chem. Eng. Res. Des. 124, 222–237.

Dehghani, M., Saghafian, B., Nasiri Saleh, F., Farokhnia, A., Noori, R., 2014. Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte Carlo simulation. Int. J. Climatol. 34, 1169–1180.

Delnavaz, M., Ayati, B., Ganjidoust, H., 2010. Prediction of moving bed biofilm reactor (MBBR) performance for the treatment of aniline using artificial neural networks (ANN). J. Hazard Mater. 179, 769–775.

Deshmukh, S.C., Senthilnath, J., Dixit, R.M., Malik, S.N., Pandey, R.A., Vaidya, A.N., Omkar, S.N., Mudliar, S.N., 2012. Comparison of radial basis function neural network and response surface methodology for predicting performance of bofilter treating toluene. J. Softw. Eng. Appl. 5, 595–603.

Dil, E.A., Ghaedi, M., Asfaram, A., Mehrabi, F., Bazrafshan, A.A., Ghaedi, A.M., 2016. Trace determination of safranin O dye using ultrasound assisted dispersive solid-phase micro extraction: artificial neural network-genetic algorithm and response surface methodology. Ultrason. Sonochem. 33, 129–140.

Dil, E.A., Ghaedi, M., Asfaram, A., Hajati, S., Mehrabi, F., Goudarzi, A., 2017. Preparation of nanomaterials for the ultrasound-enhanced removal of Pb^{2+} ions and malachite green dye: chemometric optimization and modeling. Ultrason. Sonochem. 34, 677–691.

Dlamini, D.S., Mishra, A.K., Mamba, B.B., 2014. Artificial neural network simulations and experimental results: removal of trichlorophenol from water using Chromolaena odorata stem. Water S A (Pretoria) 40.

Du, E.D., Guo, Y.Q., Sun, Y., Gao, N.Y., 2012. The investigation of organic matter removal in water treatment plant by EEM spectra coupled with self-organizing map. Spectrosc. Spectr. Anal. 32, 1846–1851.

Dubey, R., Bajpai, J., Bajpai, A.K., 2015. Green synthesis of graphene sand composite (GSC) as novel adsorbent for efficient removal of Cr(VI) ions from aqueous solution. J. Water Process Eng 5, 83–94.

El-Wakeel, A.S., 2014. Design optimization of PM couplings using hybrid particle swarm optimization-simplex method (PSO-SM) algorithm. Elec. Power Syst. Res. 116, 29–35.

Elmolla, E.S., Chaudhuri, M., 2009. Optimization of Fenton process for treatment of amoxicillin, ampicillin and cloxacillin antibiotics in aqueous solution. J. Hazard Mater. 170, 666–672.

Elmolla, E.S., Chaudhuri, M., Eltoukhy, M.M., 2010. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. J. Hazard Mater. 179, 127–134.

Fan, M.Y., Li, T.J., Hu, J.W., Cao, R.S., Wu, Q., Wei, X.H., Li, L.Y., Shi, X.D., Ruan, W.Q., 2016. Synthesis and characterization of reduced graphene oxide-supported nanoscale zero-valent iron (nZVI/rGO) composites used for Pb(II) removal. Materials 9, 687.

Fan, M.Y., Li, T.J., Hu, J.W., Cao, R.S., Wei, X.H., Shi, X.D., Ruan, W.Q., 2017a. Artificial neural network modeling and genetic algorithm optimization for cadmium removal from aqueous solutions by reduced graphene oxide-supported nanoscale zero-valent iron (nZVI/rGO) composites. Materials 10, 544.

Fan, M.Y., Hu, J.W., Cao, R.S., Xiong, K.N., Wei, X.H., 2017b. Modeling and prediction of copper removal from aqueous solutions by nZVI/rGO magnetic nanocomposites using ANN-GA and ANN-PSO. Sci. Rep. 7, 18040.

Fatemeh Mehrabia, A.V., Ghaedib, Mehrorang, Ghaedia, Abdol Mohammad, Alipanahpour Dilb, Ebrahim, Asfaramb, Arash, 2016. Ultrasound assisted extraction of Maxilon Red GRL dye from water samples using cobalt ferrite nanoparticles loaded on activated carbon as sorbent: optimization and modeling. Ultrason. Sonochem. 1350–4177, 30282–30286.

Güçlü, D., Dursun, S., 2008. Amelioration of carbon removal prediction for an activated sludge process using an artificial neural network (ANN). Clean. - Soil, Air, Water 36, 781–787.

Ghaedi, M., Ansari, A., Bahari, F., Ghaedi, A.M., Vafaei, A., 2015a. 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 137, 1004–1015.

Ghaedi, M., Shojaeipour, E., Ghaedi, A.M., Sahraei, R., 2015b. 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 142, 135–149.

Ghaedi, A.M., Ghaedi, M., Pouranfard, A.R., Ansari, A., Avazzadeh, Z., Vafaei, A., Tyagi, I., Agarwal, S., Gupta, V.K., 2016. Adsorption of triamterene on multi-walled and single-walled carbon nanotubes: artificial neural network modeling and genetic algorithm optimization. J. Mol. Liq. 216, 654–665.

Ghanbary, F., Modirshahla, N., Khosravi, M., Behnajady, M.A., 2012. Synthesis of TiO_2 nanoparticles in different thermal conditions and modeling its photocatalytic activity with artificial neural network. J. Environ. Sci. 24, 750–756.

Guria, C., Bhattacharya, P.K., Gupta, S.K., 2005. Multi-objective optimization of reverse osmosis desalination units using different adaptations of the non-dominated sorting genetic algorithm (NSGA). Comput. Chem. Eng. 29, 1977–1995.

Hamzaoui, Y.E., Hernández, J.A., Silva-Martínez, S., Bassam, A., Álvarez, A., Lizama-Bahena, C., 2011. Optimal performance of COD removal during aqueous treatment of alazine and gesaprim commercial herbicides by direct and inverse neural network. Desalination 277, 325–337.

Hanna, S.R., Chang, J.C., Fernau, M.E., 1998. Monte Carlo estimates of uncertainties in prediction by a photochemical grid model (UAM-IV) due to uncertainties in input variables. Atmos. Environ. 32, 3619–3628.

Hazrati, H., Moghaddam, A.H., Rostamizadeh, M., 2017. The influence of hydraulic retention time on cake layer specifications in the membrane bioreactor: experimental and artificial neural network modeling. J. Environ. Chem. Eng 5, 3005–3013.

Hoh, E., Hites, R.A., 2005. Brominated flame retardants in the atmosphere of the East-Central United States. Environ. Sci. Technol. 39, 7794–7802.

Hu, J.W., Zhang, X., Wang, Z., 2010. A review on progress in QSPR studies for surfactants. Int. J. Mol. Sci. 11, 1020–1047.

Huang, MaY., Wan, J., Yan, W., 2009a. Simulation of a paper mill wastewater treatment using a fuzzy neural network. Expert Syst. Appl. 36, 5064–5070.

Huang, M., Wan, J., Ma, Y., 2009b. Monitoring of anoxic/oxic process for nitrogen and chemical oxygen demand removal using fuzzy neural networks. Water Environ. Res. 81, 654–663.

Huang, X., Fang, G., 2010. Water resources allocation effect evaluation based on chaotic neural network model. J. Comput. 5, 1169–1176.

Huang, X., Hu, J., Li, C., Deng, J., Long, J., Qin, F., 2009c. Heavy-metal pollution and potential ecological risk assessment of sediments from Baihua Lake, Guizhou, P.R. China. Int. J. Environ. Health Res. 19, 405–419.

Jaiswal, S., Benson, E.R., Bernard, J.C., Wicklen, G.L.V., 2005. Neural network modelling and sensitivity analysis of a mechanical poultry catching system. Biosyst. Eng. 92, 59–68.

Jiang, B., Zhang, F., Sun, Y., Zhou, X., Dong, J., Zhang, L., 2014. Modeling and optimization for curing of polymer flooding using an artificial neural network and a genetic algorithm. J. Taiwan Inst. Chem. E 45, 2217–2224.

Jin, L., Hu, J., Wei, X., Fu, L., Li, L., 2015. Dehalogenation of persistent halogenated organic compounds: a review of computational studies and quantitative structure-property relationships. Chemosphere 131, 17–33.

Jing, L., Chen, B., Zhang, B.Y., Zheng, J.S., 2013. The effects of salinity and temperature on the photolysis of naphthaleneusing UVC irradiation. In: Proceedings of the 36th Arctic and Marine Oil Spill Program (AMOP) Technical Seminar on Environmental Contamination and Response, Halifax, Canada, pp. 591–600.

Jing, L., Chen, B., Zhang, B., 2014. Modeling of UV-induced photodegradation of naphthalene in marine oily wastewater by artificial neural networks. Water, Air, Soil Pollut. 225, 1–14.

Kalantari, A., Kamsin, A., Shamshirband, S., Gani, A., Alinejad-Rokny, H., Chronopoulos, A.T., 2017. Computational intelligence approaches for classification of medical data: state-of-the-art, future challenges and research directions. Neurocomputing. https://doi.org/10.1016/j.neucom.2017.01.126.

Kalteh, A.M., Hjorth, P., Berndtsson, R., 2008. Review of the self-organizing map (SOM) approach in water resources: analysis, modelling and application. Environ. Model. Software 23, 835–845.

Kasiri, M.B., Aleboyeh, H., Aleboyeh, A., 2008. Modeling and optimization of heterogeneous photo-Fenton process with response surface methodology and artificial neural networks. Environ. Sci. Technol. 42, 7970–7975.

Khajeh, M., Kaykhaii, M., Sharafi, A., 2013a. Application of PSO-artificial neural network and response surface methodology for removal of methylene blue using silver nanoparticles from water samples. J. Ind. Eng. Chem. 19, 1624–1630.

Khajeh, M., Sarafraz-Yazdi, A., Moghadam, A.F., 2013b. 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. Arab. J. Chem. 10, 1663–1673.

Khayet, M., Cojocaru, C., 2011. Artificial neural network modeling and optimization of desalination by air gap membrane distillation. Separ. Purif. Technol. 86, 171–182.

Kim, M., Gilley, J.E., 2008. Artificial Neural Network estimation of soil erosion and nutrient concentrations in runoff from land application areas. Comput. Electron. Agric. 64, 268–275.

Kose, E., 2008. Modelling of colour perception of different age groups using artificial neural networks. Expert Syst. Appl. 34, 2129–2139.

Kumar, J.R., Lee, H.I., Lee, J.Y., Kim, J.S., Sohn, J.S., 2008. Comparison of liquid–liquid extraction studies on platinum(IV) from acidic solutions using bis(2, 4, 4-trimethylpentyl) monothiophosphinic acid. Separ. Purif. Technol. 63, 184–190.

Kumar, K.V., Porkodi, K., Rondon, R.L.A., Rocha, F., 2007. Neural network modeling and simulation of the solid/liquid activated carbon adsorption process. Ind. Eng. Chem. Res. 47, 486–490.

Kundu, P., Debsarkar, A., Mukherjee, S., 2013. Artificial neural network modeling for biological removal of organic carbon and nitrogen from slaughterhouse wastewater in a sequencing Batch reactor. Adv. Artif. Neural Syst 2013, 1–15.

Kundu, P., Debsarkar, A., Mukherjee, S., Kumar, S., 2014. Artificial neural network modelling in biological removal of organic carbon and nitrogen for the treatment of slaughterhouse wastewater in a batch reactor. Environ. Technol. 35, 1296–1306.

López, M.E., Rene, E.R., Boger, Z., Veiga, M.C., Kennes, C., 2016. Modelling the removal of volatile pollutants under transient conditions in a two-stage bioreactor using artificial neural networks. J. Hazard Mater. 324, 100–109.

Lee, B.H., Scholz, M., 2006. Application of the self-organizing map (SOM) to assess the heavy metal removal performance in experimental constructed wetlands. Water Res. 40, 3367–3374.

Li, L.Y., Hu, J.W., Shi, X.D., Fan, M.Y., Luo, J., Wei, X.H., 2016a. Nanoscale zero-valent metals: a review of synthesis, characterization, and applications to environmental remediation. Environ. Sci. Pollut. Res. 23, 1–21.

Li, L.Y., Hu, J.W., Shi, X.D., Ruan, W.Q., Luo, J., Wei, X.H., 2016b. Theoretical studies on structures, properties and dominant debromination pathways for selected polybrominated diphenyl ethers. Int. J. Mol. Sci. 17, 927.

Liu, F., Yang, J.H., Zuo, J., Ding, M., Gan, L., Xie, B., Pei, W., Bo, Y., 2014. Graphene-supported nanoscale zero-valent iron: removal of phosphorus from aqueous solution and mechanistic study. J. Environ. Sci. 26, 1751–1762.

Luo, J., Chen, C., Xie, J., 2015a. Multi-objective immune algorithm with preference-based selection for reservoir flood control operation. Water Resour. Manag. 29, 1447–1466.

Luo, J., Hu, J.W., Wei, X.H., Fu, L.Y., Li, L.Y., 2015b. Dehalogenation of persistent halogenated organic compounds: a review of computational studies and quantitative structure-property relationships. Chemosphere 131, 17–33.

Luo, J., Hu, J.W., Zhuang, Y., Wei, X.H., Huang, X.F., 2013. Electron-induced reductive debromination of 2, 3, 4-tribromodiphenyl ether: a computational study. J. Mol. Model. 19, 3333–3338.

Ma, X., Li, Y., Li, X., Yang, L., Wang, X., 2011. Preparation of novel polysulfone capsules containing zirconium phosphate and their properties for Pb^{2+} removal from aqueous solution. J. Hazard Mater. 188, 296–303.

Mandal, S., Mahapatra, S.S., Sahu, M.K., Patel, R.K., 2015. Artificial neural network modeling of As(III) removal from water by novel hybrid material. Process Saf. Environ 93, 249–264.

Marseguerra, M., Zio, E., Podofillini, L., 2005. Multiobjective spare part allocation by means of genetic algorithms and Monte Carlo simulation. Reliab. Eng. Syst. Saf. 87, 325–335.

Mazaheri, H., Ghaedi, M., Ahmadi Azqhandi, M.H., Asfaram, A., 2017. Application of machine/statistical learning, artificial intelligence and statistical experimental design for the modeling and optimization of methylene blue and Cd(II) removal from a binary aqueous solution by natural walnut carbon. Phys. Chem. Chem. Phys. 19, 11299–11317.

Mendozacastillo, D.I., Villalobosortega, N., Bonillapetriciolet, A., Tapiapicazo, J.C., 2015. Neural network modeling of heavy metal sorption on lignocellulosic biomasses: effect of metallic ion properties and sorbent characteristics. Ind. Eng. Chem. Res. 54, 443–453.

Messikh, N., Chiha, M., Ahmedchekkat, F., Bsoul, A.A., 2014. Application of radial basis function neural network for removal of copper using an emulsion liquid membrane process assisted by ultrasound. Desalin. Water Treat 56, 1–10.

Mirsoleimani-Azizi, S.M., Amooey, A.A., Ghasemi, S., Salkhordeh-Panbechouleh, S., 2015. Modeling the removal of endosulfan from aqueous solution by electrocoagulation process using artificial neural network (ANN). Ind. Eng. Chem. Res. 54, 9844–9849.

Mohan, S., Singh, Y., Verma, D.K., Hasan, S.H., 2015. Synthesis of CuO nanoparticles through green route using Citrus limon juice and its application as nanosorbent for Cr(VI) remediation: process optimization with RSM and ANN-GA based model. Process Saf. Environ 96, 156–166.

Mohanraj, M., Jayaraj, S., Muraleedharan, C., 2012. Applications of artificial neural networks for refrigeration, air-conditioning and heat pump systems-A review. Renew. Sust. Energ. Rev. 16, 1340–1358.

Mohebbi, A., Taheri, M., Soltani, A., 2008. A neural network for predicting saturated

liquid density using genetic algorithm for pure and mixed refrigerants. Int. J. Refrig. 31, 1317–1327.
Nandi, B.K., Moparthi, A., Uppaluri, R., Purkait, M.K., 2010. Treatment of oily wastewater using low cost ceramic membrane: comparative assessment of pore blockingandartificial neural network models. Chem. Eng. Res. Des. 88, 881–892.
Oguz, E., Ersoy, M., 2010. Removal of Cu^{2+} from aqueous solution by adsorption in a fixed bed column and neural network modelling. Chem. Eng. J. 164, 56–62.
Ouarda, T.B.M.J., Shu, C., 2009. Regional low-flow frequency analysis using single and ensemble artificial neural networks. Water Resour. Res. 45, 114–122.
Ren, Y., Liu, H., Yao, X., Liu, M., Hu, Z., Fan, B., 2006. The accurate QSPR models for the prediction of nonionic surfactant cloud point. J. Colloid Interface Sci. 302, 669–672.
Reynelavila, H.E., Bonillapetriciolet, A., Rosa, G.D.L., 2015. Analysis and modeling of multicomponent sorption of heavy metals on chicken feathers using Taguchi's experimental designs and artificial neural networks. Desalin. Water Treat 55, 86.
Sabour, M.R., Amiri, A., 2017. Comparative study of ANN and RSM for simultaneous optimization of multiple targets in Fenton treatment of landfill leachate. Waste Manag. 65, 54–62.
Sacchi, R., Ozturk, M.C., Principe, J.C., Carneiro, A.A.F.M., 2007. Water inflow forecasting using the echo state network: a Brazilian case study. Cad. Saúde Pública 30, 2403–2408.
Sahu, J.N., Acharya, J., Meikap, B.C., 2009. Response surface modeling and optimization of chromium(VI) removal from aqueous solution using Tamarind wood activated carbon in batch process. J. Hazard Mater. 172, 818–825.
Santin, F.M., Silva, R.V.D., Grzybowski, J.M.V., 2016. Artificial neural network ensembles and the design of performance-oriented riparian buffer strips for the filtering of nitrogen in agricultural catchments. Ecol. Eng. 94, 493–502.
Shi, X., Ruan, W., Hu, J., Fan, M., Cao, R., Wei, X., 2017. Optimizing the removal of Rhodamine B in aqueous solutions by reduced graphene oxide-supported nanoscale zerovalent iron (nZVI/rGO) using an artificial neural network-genetic algorithm (ANN-GA). Nanomaterials 7, 134.
Shojaeimehr, T., Rahimpour, F., Khadivi, M.A., Sadeghi, M., 2013. A modeling study by response surface methodology (RSM) and artificial neural network (ANN) on Cu^{2+} adsorption optimization using light expended clay aggregate (LECA). J. Ind. Eng. Chem. 20, 870–880.
Shrestha, D.L., Kayastha, N., Solomatine, D.P., 2009. A novel approach to parameter uncertainty analysis of hydrological moedels using neural networks. Hydrol. Earth Syst. Sci. 13, 1235–1248.
Singh, K.P., Gupta, S., Singh, A.K., Sinha, S., 2010. Experimental design and response surface modeling for optimization of Rhodamine B removal from water by magnetic nanocomposite. Chem. Eng. J. 165, 151–160.
Singh, K.P., Singh, A.K., Gupta, S., Rai, P., 2012. Modeling and optimization of reductive degradation of chloramphenicol in aqueous solution by zero-valent bimetallic nanoparticles. Environ. Sci. Pollut. Res. 19, 2063–2078.
Singh, T., Singh, V., Sinha, S., 2006. Prediction of cadmium removal using an artificial neural network and a neuro-fuzzy technique. Mine Water Environ. 25, 214–219.
Tak, B.Y., Tak, B.S., Kim, Y.J., Park, Y.J., Yoon, Y.H., Min, G.H., 2015. Optimization of color and COD removal from livestock wastewater by electrocoagulation process: application of Box-Behnken design (BBD). J. Ind. Eng. Chem. 38, 307–315.
Wang, J., Deng, Z., 2015. Modeling and prediction of oyster norovirus outbreaks along gulf of Mexico coast. Environ. Health Perspect. 124, 627–633.
Wen, X., Fang, J., Diao, M., Zhang, C., 2013. Artificial neural network modeling of dissolved oxygen in the Heihe River, Northwestern China. Environ. Monit. Assess. 185, 4361–4371.
Wu, J., Peng, D., Li, Z., Zhao, L., Ling, H., 2015. Network intrusion detection based on a general regression neural network optimized by an improved artificial immune algorithm. PLoS One 10, e0120976.
Yasin, Y., Ahmad, F.B.H., Ghaffari-Moghaddam, M., Khajeh, M., 2014. Application of a hybrid artificial neural network-genetic algorithm approach to optimize the lead ions removal from aqueous solutions using intercalated tartrate-Mg-Al layered double hydroxides. Environ. Nanotech. Monit. Manage 1–2, 2–7.
Yetilmezsoy, K., Demirel, S., 2008. Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (Pistacia Vera L.) shells. J. Hazard Mater. 153, 1288–1300.
Zhang, L., Scholz, M., Mustafa, A., Harrington, R., 2008. Assessment of the nutrient removal performance in integrated constructed wetlands with the self-organizing map. Water Res. 42, 3519–3527.
Zhang, Y., Pan, B., 2014. Modeling batch and column phosphate removal by hydrated ferric oxide-based nanocomposite using response surface methodology and artificial neural network. Chem. Eng. J. 249, 111–120.