The modeling of lead removal from water by deep eutectic solvents functionalized CNTs: Artificial neural network (ANN) approach

Seef Saadi Fiyadh[a] , Mohammed Abdulhakim AlSaadi[a,b*], Ahmed El-Shafie[c**], Mohamed Khalid AlOmar[b,c], Sabah Saadi Fayaed[d], Ako R. Hama[d] , Sharifah Bee[a]

[a]Nanotechnology & Catalysis Research Centre (NANOCAT), IPS Building, University of Malaya, 50603 Kuala Lumpur, Malaysia

[b]University of Malaya Centre for Ionic Liquids, University Malaya, Kuala Lumpur 50603, Malaysia

[c]Department of Civil Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia

[d]Civil Engineering Department, Faculty of Engineering, Komar University of Science and Technology, Sulaymaniyah, Iraq

*E-mail: mdsd68j@gmail.com, Tel: +60163630693, Fax: +60 3 7967 5311

**E-mail: elshafie@um.edu.my, +60122003857

Abstract

The main challenge in the lead removal simulation is the behavior of non-linearity relationships between the process parameters. The conventional modeling technique usually deals with problem by linear method. The substitute modelling technique is ANN system, and it is selected to reflect the non-linearity in the interaction among the variables in the function. Herein, the synthesized deep eutectic solvents (DESs) were used as functionalized agent with carbon nanotubes (CNTs) as adsorbents of (Pb^{2+}). Different parameters were used in the adsorption study including, pH (2.7 to 7), adsorbent dosage (5 to 20 mg), contact time (3 to 900 min) and Pb^{2+} initial concentration (3 to 60 mg/ L). The number of experimental trials to feed and train the system was 158 runs conveyed in lab scale. Two ANN types were designed in this work, the FFBP and LR, both methods are compared based on their predictive proficiency in terms of the (MSE), (RMSE), (RRMSE), (MAPE) and (R^2) based on the testing dataset. The ANN model of lead removal was subjected to accuracy determination and the results showed determination coefficient (R^2) of 0.9956 with MSE 1.66 10^{-4}. The maximum relative error is 14.93% for the FBNN model.

Keywords: carbon nanotubes; deep eutectic solvents; water treatment; lead ions; neural network; feed forward

1. Introduction

The removal of heavy metal ions from water has been a crucial step to curb the resulting environmental problem. Any presence of heavy metals in water is recognized as a threat to both the human health and aquatic organisms (Wang & Chen 2006). In addition, the properties of heavy metals to be non-biodegradable and has the tendency to build up in living organism may lead to various disease. Heavy metals may be present in the solution as free ions or in the form of molecules, and chelate metal ligands in any water streams (Salisu 2016). Lead is known to be one of the primary toxins of the heavy metals, discharge into the environment by battery manufacturing, metal electroplating, pigment and dye industries (Majumdar 2010). The consumption of each such contaminate water may affect kidney, brain, liver and central nervous system which subsequently will lead to irreversible brain damage, weakness of muscles and nervous disorders (Geetha 2015). Research has been done to prove that the adsorption technique is one of the effective method to extract metal ions from water solution (Pimentel 2007). The effectiveness of adsorption is majorly dependent on the selection of appropriate process condition, including the mass of sorbent, duration of the process, pH and temperature of the system (Lourie & Gjengedal 2011). Many studies have been done on different materials to be used as an adsorbent to extract metal cations from water, for example, activated carbon (Chen & Wu 2004), clay minerals (Oubagaranadin & Murthy 2010), biomaterials (Gupta 2006) and pistachio vera shells (Yetilmezsoy & Demirel 2008). However, these adsorbents have not proven satisfying results.

Researchers have suggested carbon nanotubes(CNT) to be used as most effective adsorbents to extract numerous pollutants such as dyes, metal ions phenols, aniline, drugs and other contaminants (Ibrahim 2016). The qualities of CNTs such as large surface area,

diameter and the shorter equilibrium time than other materials contribute to its applications. CNTs have the potential to be used in variety of applications due to its distinctive electrical, physical and chemical properties.

CNTs have been successfully used to extract various heavy metals, for instance chromium, copper, zinc, lead, cadmium, arsenic and mercury (Ihsanullah 2016). Researchers have reported that carbon nanotubes are efficient and is significantly adsorb lead more than copper and cadmium in the suitable pH value. Besides, the presence of different ions, the strength of ions and the pH value are the major criteria that influences the adsorption of Pb^{2+} (Kabbashi 2009b).

Moreover, CNTs proved a great affinity for interaction with different compounds (Ihsanullah 2016). Therefore, functionalization is the key to improve the activity of CNTs. The conventional functionalization usually involves hash acids and non-environmentally friendly chemicals with complicated processes. Consonantly, the need for environmentally friendly functionalization agents with simple chemical processes is crucial (Hayyan 2015).

Recently, deep eutectic solvents (DESs) have gain an enormous amount of interest due to its involvement in many applications. DESs was first introduced as a low-cost development or replacement of ionic liquids (ILs). DESs have many advantages over ILs in team of availability of the row martials and easy to synthesis with minimum environmentally harmful waste (AlOmar 2016c). Therefore, DESs have conquer many fields of science. Lately, DESs involvement in many nanotechnologies related fields including media for synthesis of nanoparticles (Chen 2014; Chakrabarti 2015; Jia 2015; Xu 2016), electrolyte in nanostructure sensors (Zheng 2014), and electrolyte in nanoparticle deposition (Abbott 2009; Gu & Tu 2011; Renjith 2014). Functionalization agent of CNTs (AlOmar 2016a; AlOmar 2017). AlOmar et,al. 2016 have used choline

chloride based DESs as a functionalization agent of CNTs to prepare a sufficient adsorbent of Pb^{2+} ions. consequently, the dataset prepared from that work has been implemented for the modelling in this work (AlOmar 2016b).

New techniques such as artificial neural network (ANN) has been considered as a less complicated model in the sophisticated biological network. The substitute technique of modelling, artificial neural network system (ANN), is been selected in order to represent the non-linear function relationship among variables. The artificial neural network (ANN) techniques do not require any mathematical induction since the ANN analyses examples and recognizes the patterns in a series of inputs and outputs of dataset without any prior assumptions about their characteristics and interrelations (Mandal 2009). The speciality of the ANN to generalize and identify the pattern of any non-linear and complex development makes it an influential modelling means. Neural network has the ability to extract complicated data that is beyond the capability to be observed by a human or any computer technique. Experiments have been successfully performed to use ANN to model the adsorption of lead ions by pistachio Vera L. shells (Yetilmezsoy & Demirel 2008), the removal of Laneset Red G on Chara contraria (Mjalli 2007), Laneset Red G on walnut husk removal efficiency (Çelekli 2012), and the intercalated tartrate-Mg-Al layered double hydroxides as an adsorbent (Yasin 2014). Several studies have recently been conducted on water quality prediction models (Wu & Xu 2011; Chibole 2013). Moreover, there are some research have been applied on different areas for example, modeling the fermentation media optimization (Desai 2008b), modeling of a microe-wave-assisted extraction method (Khajeh 2011).

1.1 Problem statement

The Artificial neural network (ANN) is used to predict the adsorbent capacity of Pb^{2+} from water by using a set of experimental data that have been prepared in advance. The main advantages of using the ANN are their precision, salient and efficacy in apprehending the non-

linear relationship current between the variables of multi-input or output in complicated system. Moreover, the availability in abundance and easy handing, the ANN is economically and the best option to predict Pb^{2+} adsorption. From an economical perspective, ANN can be utilised as a substitute of CNTs that is relatively higher in cost to study the adsorption process. Furthermore, the adsorption prediction model can play a key role in providing relevant information related to the input variables. In addition to that, the development of such models can be considered as low-cost, and reduce the engineering effort.

1.2 Objective

The aims of this study are: (1) to create an ANN model to establish the relationship that exists between the adsorbent dosage, concentration of Pb^{2+}, pH and contact time to predict the DES-CNTs adsorption capacity of Pb^{2+} from water solution based on the experimental data set prepared in lab scale(AlOmar 2016b). (2) The adsorption Capacity of the DES-CNTs adsorbent for Pb^{2+} will be predicted by using ANN model and compare it with the experimentally measured values. (3) Two neural network types will be designed and compared based on the performance of the network.

2. Materials and methods

2.1 Experimental

In a previous study, a novel Pb^{2+} adsorbent was prepared based on pristine CNTs oxidized with $KMnO_4$, and then functionalized by choline chloride: triethylene glycol (salt:HBD) 1:2 DES (TEG). The preparation of the adsorbent was in two stages, the primary oxidation involved sonication of P-CNTs with $KMnO_4$ for 2 h @ 65 °C, later the resulted oxidized CNTs (K-CNTs) was sonicated with DES for 3 h @ 65 °C to produce KTEG-CNTs. The adsorbent was comprehensively characterized by indicting the RAMAN shift using Raman spectroscopy. The functional groups associated with the functionalization process was analysed using FT-IR. The surface charge, surface area and surface morphology was investigated by Zeta potential,

IBT surface area FESEM and TEM respectively. The structural phase was also investigated by conducting the XRD profile. Moreover, batch adsorption study was performed at ambient condition. Adsorbent dosage, initial concentration, pH value and contact time were taken as variables to the response of adsorption capacity of KTEG-CNTs. 158 points were taken to study the influence of each parameter and the interaction among them on the adsorption capacity. The restriction taken for each parameter is listed in Table 1. The work flow is demonstrated in scheme 1.

2.2 Design of Artificial neural networks (ANN) structure

The NN Toolbox R2014a of MATLAB is a mathematical software that was used to predict the adsorption capacity of functionalized carbon nanotube to adsorb Pb^{2+} from water solution. A total of 158 experimental datasets were prepared and used to create the ANN model. The experimental variables are initial concentration of Pb^{2-}, adsorbent dosage, pH and contact time.

Artificial neural network (ANNs) is a sophisticated statistical approach that created to behaves similar to the nervous system of human by developing a logical model containing of interconnective neurons system in a computing network (Kurt & Kayfeci 2009; Hemmat Esfe 2015). The neural network is used to resolve complicated test models such as pattern recognition, classification and estimation.

The supervised and the unsupervised are the two major types of ANNs that can be used in classification or regression. At the supervised model, the network is trained in order to adjust the optimum weight values between neurons that makes it able to produce the desired output value(s) after taking different number of instructing data from the previous experimental examples. Whereby, for the unsupervised model, there is no preferred design value during the introducing of the input to the structure. The supervised method was applied in this work.

The 158-present data were allocated into training and testing sets, where it comprises of four (4) input and one (1) output, and the testing files contained only the output parameter that

were not operated for the training processes. The data were subdivided as defined percentages to prepare separate data sets for training and testing processes of the ANN model. Nevertheless, the division is organised on the basis that the training data forms the major share of the latter. Subsequently, the data were switched within the spreadsheet and analysis was done to invalidate the presence of existing combination of trend and the inherent characteristics within the data (GK ; Zhang & Govindaraju 2003; Sarangi & Bhattacharya 2005).

A total of two (2) types of neural network were designed to analyse the feed-forward back-propagation (BP) and the layer recurrent, as to develop a NN. The number of neuron(s), layer(s), training and testing sets and the type of transfer function need to be determine carefully.

The suitable training algorithm can only be determined upon the identification of the complexity of the problem, the number of data point in the training set, the value of biases and weights in the network and the maximum error target. Six training function as presented in table 2, were used and compared based on their performance to select the best suitable training function in both the feed-forward back-propagation (BPNN) and the layer recurrent. For all training function, three hidden layers were selected for the feedforward Backpropagation (BP) and five hidden layers for the layer recurrent (LRNN), the number of hidden layers were selected by try and error to design the best NN structure.

Similarly, to selection the optimum number of hidden neurons to be used in the network is one of the major challenges for neural networks, using imprudent hidden neurons will lead to overfitting problem, this will cause to the network an over-estimate the complexity of the goal problem. It significantly effects the generalization performance, which cause a significant deviation in predictions. In the network optimization, 2 hidden neurons were used as the first hypothesis up to fifteen for both the feedforward backpropagation (BP) and the layer

recurrent(LR). In this perception, the determining of the optimum number of hidden neurons to avoid over-fitting problem is critical in function estimation using NN.

The transfer function is one of the most important factors in the model creation, in this study a three-different transfer function are used (TANSIG, PURELIN and LOGSIG) to choose the optimum one for the model.

2.3 Feed-forward Back-propagation (BP)

The feed-forward Back-propagation is usually used learning algorithm in ANN application, which used the back-propagation system as the gradient decent technique to minimize network error. Each layer in the BPANN has several neurons and each neuron transmits input values and processes to the next layer. As shown in figure1, the value of the input variable is multiplied by the connection weights w_{ii} which connects the input to the hidden layer.

The FFBP models consist of input layers, hidden layers and output layer in a multilayer neural network. The input layer consists of nodes, the hidden layer contains a nodes and the output layer consists of the K nodes. Consequently, the can be written as:

$$= \left(+ \sum_{=1} \cdot \left(+ \sum_{=1} \cdot \right) \right) \quad (1)$$

The is the transfer function in the equation 1 or its the activation function, b_{jk} and a_{ij} (= 1,2,3, ; =1,2, 3, ; =1,2, 3,) are the weight values, is the input number, and and are the deviation. The function in equation (1) is a type of mapping rule to transfer the neurons from the weighted input to output, also it is a strategy type to introduce the nonlinear into the FBNN network (Kothari & Agyepong 1996).

There are plenty of Feed-forward Back-propagation Neural Network (FBNN) transfer function in the backpropagation unit. The following transfer function selected principles used

as a monotonous non-decreasing, differentiable and continuous function. In this work the most universal binary logistic sigmoid transfer function is used and it is written as following:

$$( ) = \frac{1}{1+ \ ^{-}} \tag{2}$$

The optimal parameters can be selected by adjusting the weight values of the network as the FBNN owned by a supervisory learning algorithm technique (El-Shafie & El-Manadely 2011), and optimum means the different between the target values or actual values and the network output achieved the minimum or the target that is:

$$= \frac{1}{2} \sum_{=1} (\quad - \quad)^2 \tag{3}$$

To produce an output vector ..., for the ANN which is close as possible to target vector ,...*t* a learning, also named as training process, is employed to find the optimum bias vectors and the optimum weight matrices , that reduce the error which been established in advance that typically has the from:

$$= \sum \sum (\quad - \quad)^2 \tag{4}$$

Here, = ANN output; is the target output ; = the output value of nodes and = training patterns number. The training is a process which the connection weights of the ANN are adapted by a continuous process of stimulation through the situation in which the network is embedded.

The input data is normalized in the range of (0 to 1) form to avoid the overfitting, the adopted FBNN model structure, it is realized that all the unite at the same layer does not connect to each another, and the connections between the developed layers can be expressed based on the weighted coefficient (El-Shafie 2007).

The weighted signals and bias from the input neurons are summed by the hidden neurons and then projected through the transfer function. In the (FFBP) algorithm, the inputs forwarded into the network until the end of the network, output are initiated and compared to the target value and the error is calculated (Kurt & Kayfeci 2009; Hemmat Esfe 2015).

The back-propagation learning is to establish the relationship between the target data and the input data that is usually assigned with a random initial weight and later updating them by comparing the results between the actual and target values. In diversity of research using neural computations, consist of different transfer functions were possible to use depending upon problem nonlinearity and the complexity of data, in order to design the proper network.

2.4 Layer recurrent

The artificial recurrent neural network (RNNs) preform a great and different classes of computational modeling which is usually created by more or less detailed analogy with the biological brain module. In the layer-recurrent, many abstract neurons and also may called processing elements or units which are interconnected by likewise distracted synaptic links or connection, which enable activation to pass through the network.

The (LRNN) is almost the same as the (FNN) except that both the hidden and output layer of the (LRNN) has recurrent connection associated with a tab delay which is different with the (FNN). In addition to the input space, the RNN works on an internal state space, which is a trace of what has already been processed by the network. Neurons in RNN can be connected to any other neurons in the same or a previous layer. the recurrent neural network (RNNs) consist of input layer, hidden layer, and output layer, with activation function, feedback connection weights, and interconnection weights. Figure 2 illustrates the flow of the input samples in the LRNN architecture.

The first hidden layer of the LRNN are connected to the inputs and the following layer which assemble the networks output and has a connection from the previous layer. The input

weights of the hidden layer come from the input samples, and the following layer has a weight comes from the preceding hidden layer. Though, the hidden layer does not directly affect with the external environment, they have great influence on the following layer or the output layer of the ANN.

The LRNN is categorized by the appearance of a backward connection in the hidden and output layers providing backward connection initiated from each output of hidden layer connected to one of the weights of the input layer by the context unit. Moreover, the backward connection from the output of the hidden layer as presented in figure 2, is occupied from the real output during the training prosses of the LRNN. The selection of proper network parameters such as the hidden layers number, the neurons number in each layer, and the function of transfer types which is one of the most important network parameters considered for the architecture.

2.5 **Evaluation indicators for simulation models**

Two competing neural networks have been developed, the feedforward back-propagation and layer recurrent for modeling was used in this study. The total of 158 experimental data were divided into two subsets of testing and training (Desai 2008a; Lee 2011) for developing ANN model. The experimental variables are Pb^{2+} concentration, adsorbent dosage, pH, and contact time. The assessment of multicriteria was carried out. Therefore, the (FNN) and (RNN) models performance was determined by comparing the actual data and the simulated data. The simulation behaviour of each model was evaluated by using the root mean square error (RMSE), relative error (RE), mean square error (MSE), relative root mean square error (RRMSE) and the mean absolute percentage error (MAPE). Formulas for calculation MSE, RMSE, RE, RRMSE, and MAE were given below and as follow:

$$= \frac{1}{} \sum_{=1} \left(\quad ( ) - \quad ( ) \right)^{2} \qquad (5)$$

$$= \left[\frac{1}{} \sum_{=1} \left(\quad () - \quad () \right)^2 \right]^{\frac{1}{2}} \tag{6}$$

$$= \left[\frac{1}{} \sum_{=1} \left(\frac{() - ()}{()} \right)^2 \right]^{\frac{1}{2}} \tag{7}$$

$$= \frac{1}{} \sum_{=1} \left| \frac{\left(\quad () - \quad () \right)}{()} \right| \times 100 \tag{8}$$

$$^2 = 1 - \frac{}{} \tag{9}$$

$$= \frac{() - ()}{()} \times 100 \tag{10}$$

Where

= the actual value.

= the simulated value.

= the regression sum of squares.

= the sum of squares of residuals.

Generally, MSE, RMSE, RE, MAPE and RRMSE equations were selected to indicate the models performance, were based on the obtained result by comparing the evaluated error of the actual and simulated model. The best model the model with the smallest error is considered.

3. Results and Discussion

In this section, the used methods to select the optimal neuron number, training function, performance of the selected model, model performance evaluation and the relative error between the predicted and the actual results are discussed.

3.1 Model performance evaluation

The functioning of every model was presented by using the MSE, RMSE, RRMSE and MAPE the values of each simulated method is tabulated in table 3.

With reference to the MSE values of every model listed in Table 3, it is observed that the MSE of FBNN model is$1.66\ 10^{-4}$ which is a practical value and reflects the efficacy and higher accuracy in comparison to the LRNN model to simulate the relationship between initial concertation of Pb^{2+}, adsorbent dosage, pH and contact time to analyse the adsorbent capacity of KTEG-CNTs.

A simulation model is categorised to be reasonable and accurate if the MAPE value is below 30% and 5% respectively. Based on this acceptance criteria, the MAPE value for FBNN that is 4.10% is considered to as accurate, whereas the MAPE value of 5.60% for LRNN does not fall in the range of accuracy.

It is evident that the FBNN provides better result in comparison to LRNN. Furthermore, the FBNN model had resulted in RMSE and RRMSE value of $1.28\ 10^{-2}$ and $5.76\ 10^{-2}$ respectively which is lesser than the results obtained in LRNN model. The probability of error in simulated value is low with the result of RMSE and RRMSE getting closer to zero. Therefore, these criteria further confirm that the FBNN method provide higher degree of accuracy in comparison to the LRNN method.

3.2 Training and testing dataset

This section discusses the part of a typical multilayer network workflow. In common practice, the data is first partitioned into two sets. The first set is termed as training set, that is used to record the gradient and modify the network weights and biases. Testing set is the second subset, which is not applied during the training but is functional to compare various models and plot the errors of test sets. Generally, each backpropagation training group starts with several initial weights and biases, and various division of data into training and testing sets. These different conditions can lead to varied solutions for the same problem. In this study,

different training and testing percentage used to find the optimal set with minimum error, the used sets are presented in Table 4:

The used data in this study is 158 were separated into training and testing sets, different sets of data were used as presented in Table 4 to find the optimal set for model creation.

At 70% training and 30 % testing the MSE for the FBNN is $2.44\ 10^{-2}$ and for the LRNN is $1.08\ 10^{-1}$ whereby, when increasing the percentage of the training set to 75% with 25% for the test set, the MSE decreased to $9.73\ 10^{-3}$ for the FBNN and $9.41\ 10^{-2}$ for the LRNN. Furthermore, the percentage of training increased to 80% with decreased the percentage of the test set to 20% there was an improvement in the result of the MSE it shows $5.87\ 10^{-3}$ for FBNN and $2.96\ 10^{-2}$ for the LRNN. The minimum MSE for the FBNN and LRNN was at 85% at the training and 15% at the test set with $1.66\ 10^{-4}$ for the FBNN and $7.22\ 10^{-4}$ for the LRNN. While, by increasing the training set to 90% for the training and 10% for testing the MSE showed a higher value with $6.67\ 10^{-2}$ for the FBNN and $8.45\ 10^{-2}$ for LRNN. Hence, the optimal split is 85% for the training and 15% for the test. The results demonstrate that FBNN model is more accurate than LRNN as the values of the MSE for all the sets for the FBNN lower than the MSE for LRNN.

3.3 Neurons number optimization

The best structure of the ANN model and its specification difference are determined with reference to the smallest value of MSE of the test dataset. With an increasing the number of the neurons, the network generates different MSE values for the testing dataset. Figure3 illustrate the relationship between the number of neurons in each hidden layer and MSE obtained.

For the feedforward backpropagation, 3 hidden layers were used. The MSE value of the network result is higher for the 2 neurons (MSE $7.43\ 10^{-2}$) and 3 ($7.60\ 10^{-2}$) hidden layer neurons, than those with 4 shows a higher drop to (MSE $6.51\ 10^{-2}$) While, for 5 (MSE $5.03\ 10^{-}$

$^{-2}$), 6 (MSE 4.25 10^{-2}) and 7 (MSE 3.54 10^{-2}). The value of the MSE reduced significantly from 3.54 10^{-2} to 1.48 10^{-2} with the application of 8 hidden neurons and continue to decrease with subsequent rise in the neurons number from 8 to 10. Therefore, the neural network consisting of 10 hidden neurons with MSE value of 1.66 10^{-4} was selected as the best case based on the MSE value.

Furthermore, when the neurons number increase to 13, the MSE value displayed a slight increment from 1.66 10^{-4} to 4.98 10^{-3}. Subsequently, a further addition in the neurons number from 13 to 15 the result of the MSE is sharply increased. The increment might be assign to the MSE characteristics input vector used and performance index used in this work.

While for layer recurrent as shown in figure3 the number of hidden neurons was tested on 5 hidden layers as the best for the structure of the network. 2 hidden neurons were used at the first try and showed (9.01 10^{-2} MSE), while for 3 neurons the MSE decreased to (8.04 10^{-2}) and, for 4 the MSE increased (8.27 10^{-2}). While when using 5 neurons there is a decline in the MSE with 7.03 10^{-2}), however, increasing the neurons to 6 the MSE showed a slightly decrease to (3.83 10^{-2}), the number of hidden neurons were later increased to 7 and 8 hidden neurons to improve the stabilization of the network, the MSE decreased to (2.27 10^{-2} and 2.06 10^{-2}) respectively, with the increasing the neurons to 9 the MSE reached to (1.94 10^{-3}). furthermore, with 10 neurons the MSE decreased to 7.22 10^{-4}which shows the best stabilization of the network, with 11 and 12 neurons the MSE resulted displayed a higher value which are 1.96 10^{-3} and 4.68 10^{-3} respectively. Finally, with 13,14 and 15 neurons the MSE rises to 1.04 10^{-2}, 9.84 10^{-3} and 1.86 10^{-2}. This confirm that using 10 neurons at each hidden layer shows the best performance of the network.

3.4 Selection of the training function for FBNN and LRNN.

It is laborious to find the fastest training algorithm for a given problem due to the complexity of the problem that depends on various factors. This section discusses on the

comparison of the various training algorithms. Networks are trained on six different training functions and identified based on the R^2value and the MSE. Six training functions shown in Table5 were used and compared to select the best suitable training function in both the feedforward and backpropagation (BP) and the layer recurrent (LR).

For the feed-forward back-propagation the comparison study resulted that the Bayesian regularization backpropagation (trainrb) had resulted in smallest value of MSE in comparison to different sets of algorithms such as the Levenberg Marquardt bbackpropagation (trainlm) algorithm. As shown in Table 5, the smallest MSE was obtained about $1.66\ 10^{-4}$, and 0.9956 of R^2 for trainbr function presented in figure 4 which reflects a great performance of the network. This was followed by the trainlm with a MSE of $6.79\ 10^{-4}$. However, both trainrb and trainlm shows a great behaviour than the other algorithms such as trainbgf, traincgb, traincgf and traincgp.

The structure and the combinatorial characteristic of the test data influences the results optimality initiated by some BP algorithms. Therefore, the problem complexity was solved by the results of several analysis of training algorithms used for the comparison.

Whereby, for the layer recurrent the first benchmark is using trainbfg as training function with tansig transfer function and gave a result of the MSE is $1.28\ 10^{-2}$. While for the trainbr training function the MSE is decreased to $7.22\ 10^{-4}$, which shows a great performance of the network. However, by using a different training function such as traincgb and traincgf the MSE is $4.44\ 10^{-3}$ and $2.89\ 10^{-2}$which are greater than the MSE of trainbr training function. While, the MSE for the traincgp is $1.32\ 10^{-2}$, The performance of the trainlm is also showed a great value of the MSE $2.82\ 10^{-3}$ which is also can be considered as one of the best suitable training function for the network.

It can be realized that the trainbr training function is the best suitable for both feed-forward backpropagation and layer recurrent network.

3.5 Relative error indication

Relative error is one of the indications of error in the model prediction values comparing the predicted values to the actual values, measurements and calculation can be characterized with regard to their precision and accuracy. The term accuracy can be defined as how closely the predicted value matches the actual value, whereas precision is referred to how closely values matches with each other.

The highest relative error value is found to be 14.93 % for the FBNN model and 18.67% for the LRNN model calculated by equation (10) which is considered as an acceptable value.

Based on the results shown in figure5, it is observed that occurred error for all the testing dataset is less than 14.93% for the FBNN whereby, the maximum error for the LRNN is less than 18.67% for the LRNN. Which indicate that the FBNN model is more accurate than the LRNN model. This proves the effectiveness and reliability of the proposed approach to extract features from input data. The hybrid FBNN algorithm network model is able to provide a perfect prediction of KTEG-CNTs as Pb^{2+} absorber from water. The uncertainty in this work might come from the accuracy of the initial concentration of Pb^{2+}, amount of adsorbent dosage, and the initial pH adjustment, as the amount of the materials used is very small amount. Also, the humidity and the temperature of the room is not considered in this study which might affect the accuracy of the results.

3.6 The effect of pH on the adsorption capacity

The pH is one of the very important factors which can affect the quantity and the form of Pb^{2+} in water, and the interactions of the mineral and Pb^{2+}, and the quantity and the form of the minerals surface sites (Chen & Wang 2007).

The pH effect on adsorption capacity studied by mixing 12.5 mg adsorbent dosage with 5 mg/l concertation at 15 min contact time with range pH values from 1 to 10. The experiment results present that the pH of the solution was found as an important factor effacting the adsorption efficiency. The pH increment led to significant increase in adsorption capacity until pH 5.0,

then the adsorption capacity became steady with increasing pH. It is well known that at pH greater than 7.0, the dominant species of Pb^{2+} are $Pb(OH)^{+}$ and $Pb(OH)_2$. This complexation may occur due to the extensive presence of OH^- at this pH level which resulted in a precipitation form (Gupta 2011). In addition, the decreasing of H^+ plays a significant role in the mechanism of Pb^{2+} adsorption due to the decreasing of competition of the active sides of the adsorbent. The agreement of the ANN model predictions as a pH function is presented in Fig. 6. From fig. 6, it can be noticed that the ANN model outputs showed almost the same behaviour as the experimental data, this can prove that the ANN model can to predict the adsorption capacity of Pb^{2+} removal from water satisfactorily.

3.7 The effect of adsorbent dosage on the adsorption capacity

The Adsorbent dosage is one of the important parameters involved in the adsorption process, adsorbent dosage effect on the Pb^{2+}removal is examined by keeping the involved factors as constant, at time 10 minutes, pH 5.0, and 5 mg/L of Pb^{2+} initial concentration. The Pb^{2+} removal capacity is decreased from 47.46 mg/m to 19.704 mg/g by increasing the adsorbent dosage from 5 mg to 12.5 mg and 19.704mg/g to 12.392 mg/g by increasing the adsorbent dosage from 12.5 mg to 20 mg. The decreasing in the uptake capacity with increasing in the adsorbent dosage might be attributed with increasing the adsorbent surface area following in an increase of more active sites (Kumar & Phanikumar 2013; Das 2014). The ANN technique is used for the modeling and prediction of the obtained data from the experimental work, the prediction results shows a good agreement with the experimental result trend. The ANN outputs and the experimental results as the function of dosage versus the uptake capacity are presented in fig.7.

3.8 The effect of initial concentration

The initial concentration is one of the factors involved is this work, the effect of initial concentration of Pb^{2+} ions is studied by changing the initial concentration from 5 mg/l to 60

mg/l. The other factors were fixed at, time 60 min, pH 2.7 and adsorbent dosage 5 mg. From the presented results in fig. 8, it can be seen that the uptake capacity of Pb^{2+} ions at 5 mg/l concentration was 47.7 mg/g whereby increasing the Pb^{2+} concentration to 60 mg/l the uptake capacity increased to 225.05 mg/g. This might be attributed due to the increase in the driving force of the mass transfer which lead to an increase in the uptake capacity of Pb^{2+} ions from water solution. At low concentration, the Pb^{2+} ions interact at the adsorbent active site whereas at higher Pb^{2+} concentration, the adsorbent active site will be saturated and the removal percentage will be lower (Hamza 2013). The obtained data from the experimental work are trained and predicted by using the ANN modeling techniques. The ANN model prediction found satisfactory for the experimental data observation. The experimental and predicted output of the ANN are presented in fig. 8.

3.9 The effect of contact time

The contact time is one of the involved parameter in the experimental work, the contact time effect is studied with varying the contact time from 5 min to 120 min. The other involved parameters are kept as constant, initial concentration 5 mg/l, adsorbent dosage 5 mg and pH 5. The uptake capacity at 5 min time is 31.98 mg/g whereby, at 80 min the uptake capacity reached to 48.1mg/g, the maximum uptake capacity at the equilibrium of time is 49.3 mg/g. It is clearly from the results presented in fig.9 that 90% removal occurred at the 80 min (Kabbashi 2009a). This due to the availability of vacant sites at the adsorbent particles hence, the adsorption rate will be higher at the beginning of the reaction. The ANN model was used for the modeling and prediction of the obtained results, it can be seen from fig.9 that the ANN model predicted the experimental data satisfactorily.

4. Conclusion

The artificial neural network (ANN) has been successfully used to predict the removal of the Pb^{2+} from aqueous solution by using DES functionalized CNTs. The (tansig) transfer

function was used in this study for modeling. Two different neural network types were developed in this work the (BP-ANN) and (LR-ANN), Both the models are created with same aim function and restriction with the same structure of dataset. The optimal topology of ANN was obtained during training phase using (trainbr) algorithm. The results showed that the network with 10 neurons in each hidden layer with three hidden layers, showed the best performance. Moreover, the supervised (multi-layer feed-forward neural network) used in this study.

The (MSE) of the (BP-ANN) model prediction is $1.66\ 10^{-4}$ with the (R^2) of 0.9956. The favourable features of the ANN modeling technique was originate to have many criteria such as generalization, efficiency and simplicity, which make it a preferable choice for the modeling of complex systems, such as removal of Pb^{2+} ions from water processes.

5. Acknowledgment

The authors express their thanks to the University of Malaya for funding this research. UMRG (RP017B-13AET)

6. References

Abbott A. P., El Ttaib K., Frisch G., McKenzie K. J. and Ryder K. S. (2009). Electrodeposition of copper composites from deep eutectic solvents based on choline chloride. *Physical Chemistry Chemical Physics* **11**(21), 4269-77.

AlOmar M. K., Alsaadi M. A., Hayyan M., Akib S. and Hashim M. A. (2016a). Functionalization of CNTs surface with phosphonuim based deep eutectic solvents for arsenic removal from water. *Applied Surface Science* **389**, 216-26.

AlOmar M. K., Alsaadi M. A., Hayyan M., Akib S., Ibrahim M. and Hashim M. A. (2017). Allyl triphenyl phosphonium bromide based DES-functionalized carbon nanotubes for the removal of mercury from water. *Chemosphere* **167**, 44-52.

AlOmar M. K., Alsaadi M. A., Hayyan M., Akib S., Ibrahim R. K. and Hashim M. A. (2016b). Lead removal from water by choline chloride based deep eutectic solvents functionalized carbon nanotubes. *Journal of Molecular Liquids* **222**, 883-94.

AlOmar M. K., Hayyan M., Alsaadi M. A., Akib S., Hayyan A. and Hashim M. A. (2016c). Glycerol-based deep eutectic solvents: Physical properties. *Journal of Molecular Liquids* **215**, 98-103.

Çelekli A., Birecikligil S. S., Geyik F. and Bozkurt H. (2012). Prediction of removal efficiency of Lanaset Red G on walnut husk using artificial neural network model. *Bioresource Technology* **103**(1), 64-70.

Chakrabarti M. H., Manan N. S. A., Brandon N. P., Maher R. C., Mjalli F. S., AlNashef I. M., Hajimolana S. A., Hashim M. A., Hussain M. A. and Nir D. (2015). One-pot electrochemical gram-scale synthesis of graphene using deep eutectic solvents and acetonitrile. *Chemical Engineering Journal* **274**, 213-23.

Chen H. and Wang A. (2007). Kinetic and isothermal studies of lead ion adsorption onto palygorskite clay. *Journal of Colloid and Interface Science* **307**(2), 309-16.

Chen J. P. and Wu (2004). Acid/Base-Treated Activated Carbons: Characterization of Functional Groups and Metal Adsorptive Properties. *Langmuir* **20**(6), 2233-42.

Chen P. H., Hsu C.-F., Tsai D. D.-W., Lu Y.-M. and Huang W.-J. (2014). Adsorption of mercury from water by modified multi-walled carbon nanotubes: adsorption behaviour and interference resistance by coexisting anions. *Environmental Technology* **35**(15), 1935-44.

Chibole O. K. (2013). Modeling River Sosiani's water quality to assess human impact on water resources at the catchment scale. *Ecohydrology & Hydrobiology* **13**(4), 241-5.

Das B., Mondal N., Bhaumik R. and Roy P. (2014). Insight into adsorption equilibrium, kinetics and thermodynamics of lead onto alluvial soil. *International Journal of Environmental Science and Technology* **11**(4), 1101-14.

Desai K. M., Survase S. A., Saudagar P. S., Lele S. and Singhal R. S. (2008a). Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan. *Biochemical Engineering Journal* **41**(3), 266-73.

Desai K. M., Survase S. A., Saudagar P. S., Lele S. S. and Singhal R. S. (2008b). Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan. *Biochemical Engineering Journal* **41**(3), 266-73.

El-Shafie A., Taha M. R. and Noureldin A. (2007). A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. *Water Resources Management* **21**(3), 533-56.

El-Shafie A. H. and El-Manadely M. S. (2011). An integrated neural network stochastic dynamic programming model for optimizing the operation policy of Aswan High Dam. *Hydrology Research* **42**(1), 50.

Geetha P., Latha M. S., Pillai S. S. and Koshy M. (2015). Nanoalginate based biosorbent for the removal of lead ions from aqueous solutions: Equilibrium and kinetic studies. *Ecotoxicology and Environmental Safety* **122**, 17-23.

GK P., Nale J. and Muluneh W. Modelling Reference Evapotranspiration Using Artificial Neural Network: A Case Study of Ameleke watershed, Ethiopia.

Gu C. and Tu J. (2011). One-Step Fabrication of Nanostructured Ni Film with Lotus Effect from Deep Eutectic Solvent. *Langmuir* **27**(16), 10132-40.

Gupta V. K., Agarwal S. and Saleh T. A. (2011). Synthesis and characterization of alumina-coated carbon nanotubes and their application for lead removal. *Journal of Hazardous Materials* **185**(1), 17-23.

Gupta V. K., Rastogi A., Saini V. K. and Jain N. (2006). Biosorption of copper(II) from aqueous solutions by Spirogyra species. *Journal of Colloid and Interface Science* **296**(1), 59-63.

Hamza I. A., Martincigh B. S., Ngila J. C. and Nyamori V. O. (2013). Adsorption studies of aqueous Pb (II) onto a sugarcane bagasse/multi-walled carbon nanotube composite. *Physics and Chemistry of the Earth, Parts A/B/C* **66**, 157-66.

Hayyan M., Abo-Hamad A., AlSaadi M. A. and Hashim M. A. (2015). Functionalization of graphene using deep eutectic solvents. *Nanoscale Research Letters* **10**(1), 324.

Hemmat Esfe M., Afrand M., Yan W.-M. and Akbari M. (2015). Applicability of artificial neural network and nonlinear regression to predict thermal conductivity modeling of Al2O3–water nanofluids using experimental data. *International Communications in Heat and Mass Transfer* **66**, 246-9.

Ibrahim R. K., Hayyan M., AlSaadi M. A., Hayyan A. and Ibrahim S. (2016). Environmental application of nanotechnology: air, soil, and water. *Environmental Science and Pollution Research* **23**(14), 13754-88.

Ihsanullah, Abbas A., Al-Amer A. M., Laoui T., Al-Marri M. J., Nasser M. S., Khraisheh M. and Atieh M. A. (2016). Heavy metal removal from aqueous solution by advanced carbon nanotubes:

Critical review of adsorption applications. *Separation and Purification Technology* **157**, 141-61.

Jia H., An J., Guo X., Su C., Zhang L., Zhou H. and Xie C. (2015). Deep eutectic solvent-assisted growth of gold nanofoams and their excellent catalytic properties. *Journal of Molecular Liquids* **212**, 763-6.

Kabbashi N. A., Atieh M. A., Al-Mamun A., Mirghami M. E., Alam M. and Yahya N. (2009a). Kinetic adsorption of application of carbon nanotubes for Pb (II) removal from aqueous solution. *Journal of Environmental Sciences* **21**(4), 539-44.

Kabbashi N. A., Atieh M. A., Al-Mamun A., Mirghami M. E. S., Alam M. D. Z. and Yahya N. (2009b). Kinetic adsorption of application of carbon nanotubes for Pb(II) removal from aqueous solution. *Journal of Environmental Sciences* **21**(4), 539-44.

Khajeh M. M. a. M. (2011). "Comparison of Response Surface Methodology and Artificial Neural Network in Predicting the Microwave-Assisted Extraction Procedure to Determine Zinc in Fish Muscles" *Food and Nutrition Sciences* **Vol. 2**(No. 8,), pp. 803-8.

Kothari R. and Agyepong K. (1996). On lateral connections in feed-forward neural networks. In: *Neural Networks, 1996., IEEE International Conference on*, pp. 13-8 vol.1.

Kumar M. S. and Phanikumar B. (2013). Response surface modelling of Cr6+ adsorption from aqueous solution by neem bark powder: Box–Behnken experimental approach. *Environmental Science and Pollution Research* **20**(3), 1327-43.

Kurt H. and Kayfeci M. (2009). Prediction of thermal conductivity of ethylene glycol–water solutions by using artificial neural networks. *Applied Energy* **86**(10), 2244-8.

Lee J.-W., Suh C., Hong Y.-S. T. and Shin H.-S. (2011). Sequential modelling of a full-scale wastewater treatment plant using an artificial neural network. *Bioprocess and biosystems engineering* **34**(8), 963.

Lourie E. and Gjengedal E. (2011). Metal sorption by peat and algae treated peat: Kinetics and factors affecting the process. *Chemosphere* **85**(5), 759-64.

Majumdar S. S., Das S. K., Chakravarty R., Saha T., Bandyopadhyay T. S. and Guha A. K. (2010). A study on lead adsorption by Mucor rouxii biomass. *Desalination* **251**(1–3), 96-102.

Mandal S., Sivaprasad P. V., Venugopal S. and Murthy K. P. N. (2009). Artificial neural network modeling to evaluate and predict the deformation behavior of stainless steel type AISI 304L during hot torsion. *Applied Soft Computing* **9**(1), 237-44.

Mjalli F. S., Al-Asheh S. and Alfadala H. E. (2007). Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance. *Journal of Environmental Management* **83**(3), 329-38.

Oubagaranadin J. U. K. and Murthy Z. V. P. (2010). Isotherm modeling and batch adsorber design for the adsorption of Cu(II) on a clay containing montmorillonite. *Applied Clay Science* **50**(3), 409-13.

Pimentel P. M., González G., Melo M. F. A., Melo D. M. A., Silva Jr C. N. and Assunção A. L. C. (2007). Removal of lead ions from aqueous solution by retorted shale. *Separation and Purification Technology* **56**(3), 348-53.

Renjith A., Roy A. and Lakshminarayanan V. (2014). In situ fabrication of electrochemically grown mesoporous metallic thin films by anodic dissolution in deep eutectic solvents. *Journal of Colloid and Interface Science* **426**, 270-9.

Salisu A., Sanagi M. M., Abu Naim A., Wan Ibrahim W. A. and Abd Karim K. J. (2016). Removal of lead ions from aqueous solutions using sodium alginate-graft-poly(methyl methacrylate) beads. *Desalination and Water Treatment* **57**(33), 15353-61.

Sarangi A. and Bhattacharya A. K. (2005). Comparison of Artificial Neural Network and regression models for sediment loss prediction from Banha watershed in India. *Agricultural Water Management* **78**(3), 195-208.

Wang J. and Chen C. (2006). Biosorption of heavy metals by Saccharomyces cerevisiae: A review. *Biotechnology Advances* **24**(5), 427-51.

Wu G. and Xu Z. (2011). Prediction of algal blooming using EFDC model: Case study in the Daoxiang Lake. *Ecological Modelling* **222**(6), 1245-52.

Xu K., Wang Y., Ding X., Huang Y., Li N. and Wen Q. (2016). Magnetic solid-phase extraction of protein with deep eutectic solvent immobilized magnetic graphene oxide nanoparticles. *Talanta* **148**, 153-62.

Yasin Y., Ahmad F. B. H., Ghaffari-Moghaddam M. and 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. *Environmental Nanotechnology, Monitoring & Management* **1–2**, 2-7.

Yetilmezsoy K. and Demirel S. (2008). Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (Pistacia Vera L.) shells. *Journal of Hazardous Materials* **153**(3), 1288-300.

Zhang B. and Govindaraju R. S. (2003). Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. *Journal of Hydrology* **273**(1–4), 18-34.

Zheng Y., Ye L., Yan L. and Gao Y. (2014). The electrochemical behavior and determination of quercetin in choline chloride/urea deep eutectic solvent electrolyte based on abrasively immobilized multi-wall carbon nanotubes modified electrode. *Int. J. Electrochem. Sci* **9**, 238-48.

Table 1 The range of input and output parameters

Parameters	Minimum	Maximum
Adsorbent Dosage (g)	5	20
Initial Concentration of Pb^{2+} (ppm)	3	60
PH	2.7	7
Contact Time (min)	3	900
Uptake Capacity (mg/g) (output)	7.12	294.5

Table 2 The selected training functions.

Name of training function	Training function
Quasi-Newton back-propagation	Trainbfg
Bayesian regularization back-propagation	Trainbr
Powell Beale conjugate gradient back-propagation	Traincgb
Polak Ribi-ere conjugate gradient back-propagation	Traincgp
Fletcher Reeves conjugate gradient back-propagation	Traincgf
Levenberg Marquardt back-propagation	Trainlm

Table 3 Evaluation indicators

	FBNN	LRNN
MSE	0.000166	0.00072
RMSE	0.012871	0.02687
RRMSE	0.057678	0.074638
MAPE	4.101775	5.605057

Table 4 The training and testing sets.

		MSE	
Training %	Testing %	FBNN	LRNN
70	30	0.024487	0.108435
75	25	0.009733	0.094116
80	20	0.005876	0.029661
85	15	0.000166	0.000722
90	10	0.066796	0.084524

Table 5 The training function, R^2 and MSE

	Feedforward backpropagation		layer recurrent	
Training function	**MSE**	**R^2**	**MSE**	**R^2**
trainbfg	0.00903	0.7465	0.012892	0.856
trainbr	0.00016	0.9956	0.000722	0.990
traincgb	0.00297	0.9021	0.004446	0.956
traincgf	0.04631	0.4931	0.028921	0.663
traincgp	0.02303	0.6915	0.013207	0.891
trainlm	0.00068	0.9766	0.002822	0.968

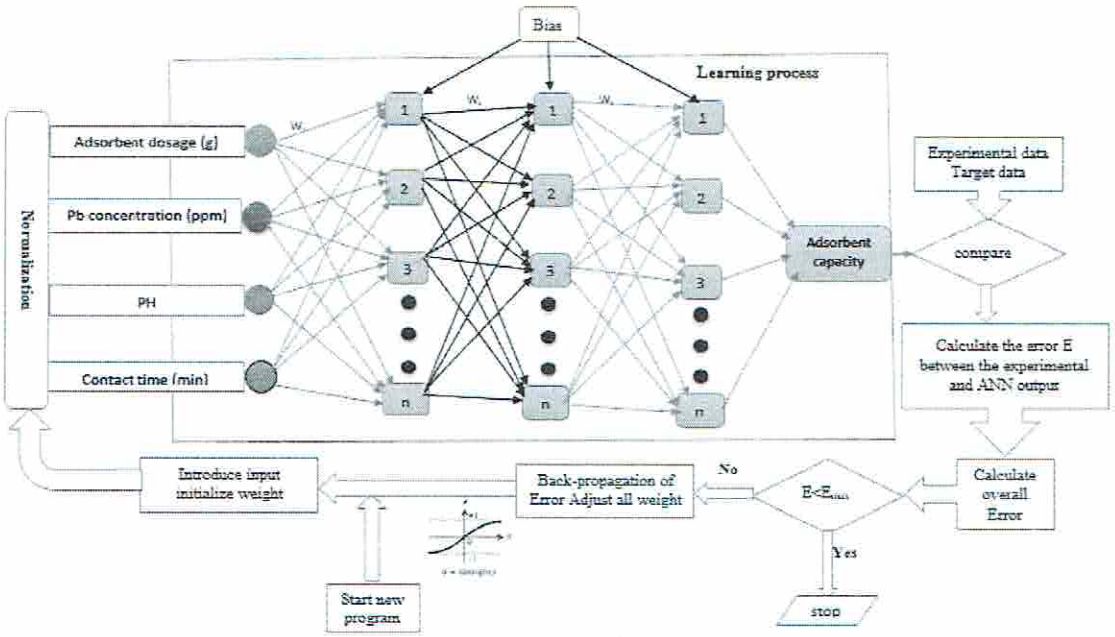


Fig. 1 Feed-forward back-propagation neural network structure.

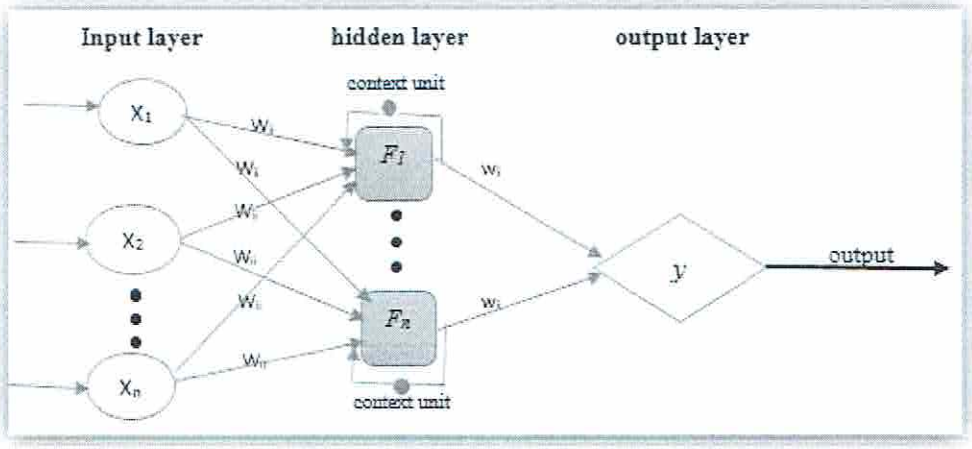


Fig. 2 The architecture system of the LRNN

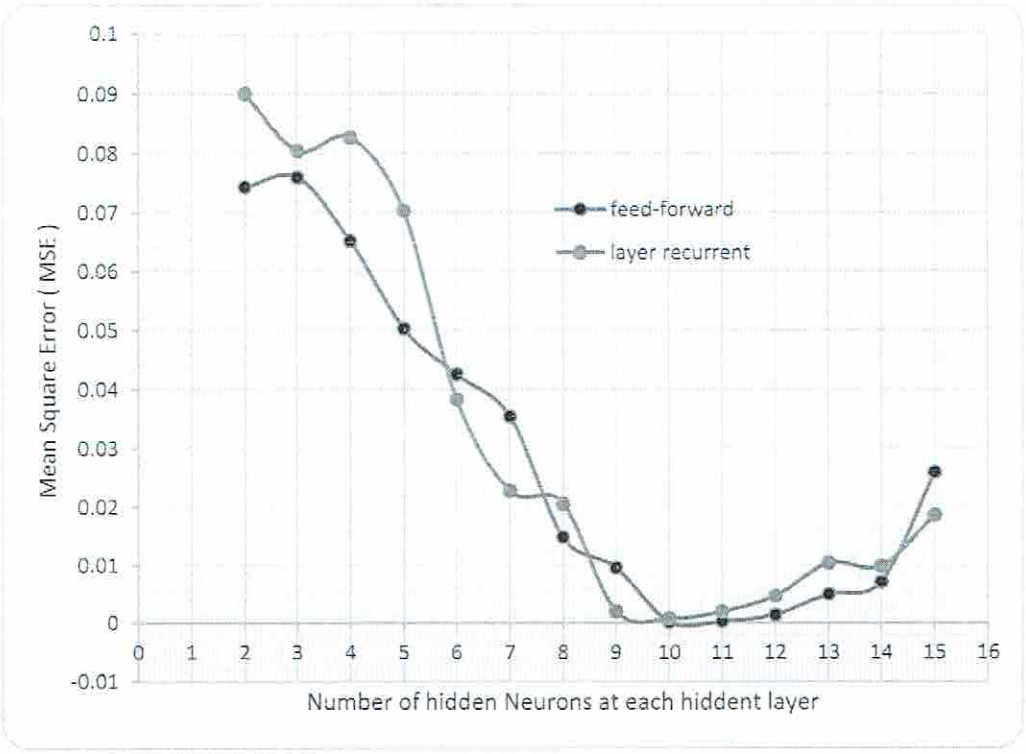


Fig. 3 The neurons number at each hidden layer with the MSE value.

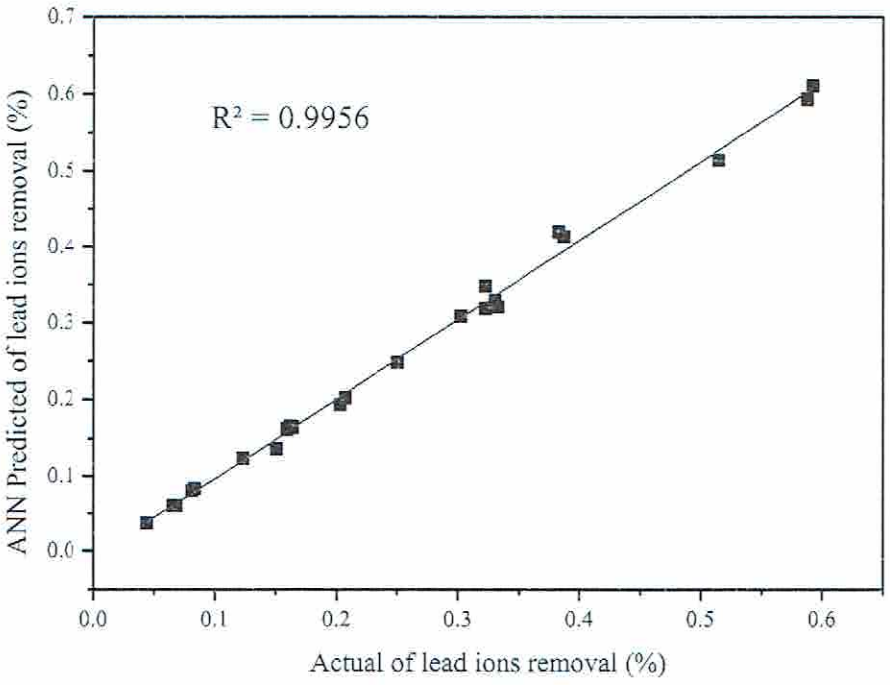


Fig. 4 The R^2 of feedforward neural network

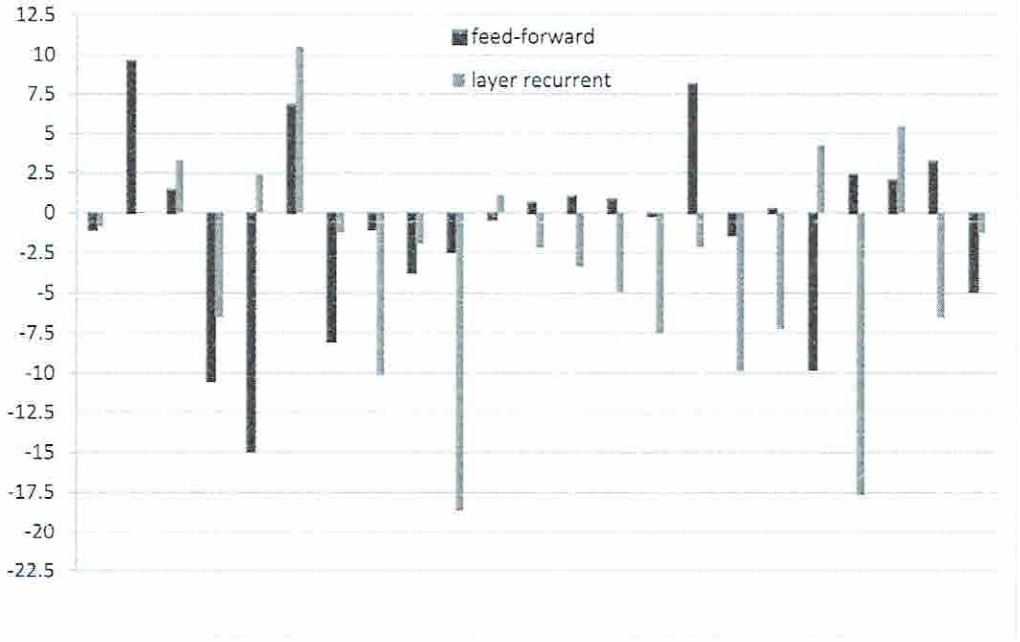


Fig. 5 Illustrates the accuracy of the hybrid model.

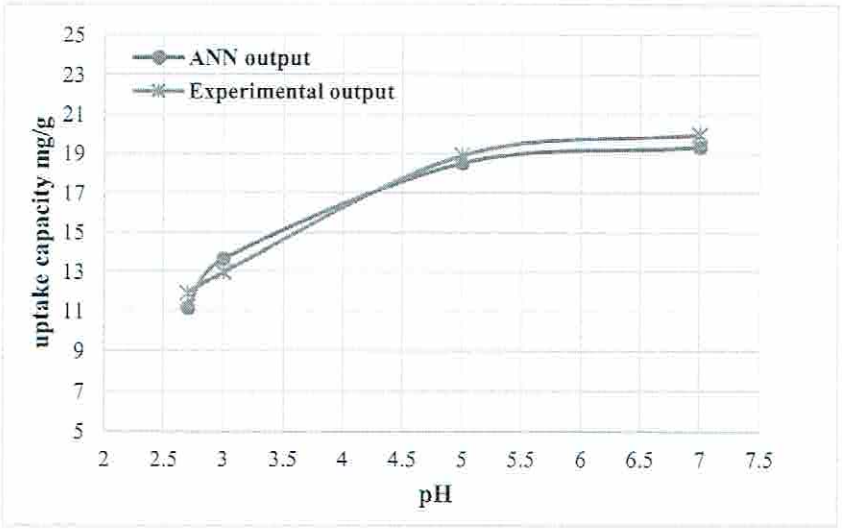


Fig.6 Agreement between ANN and experimental outputs with various pH values.

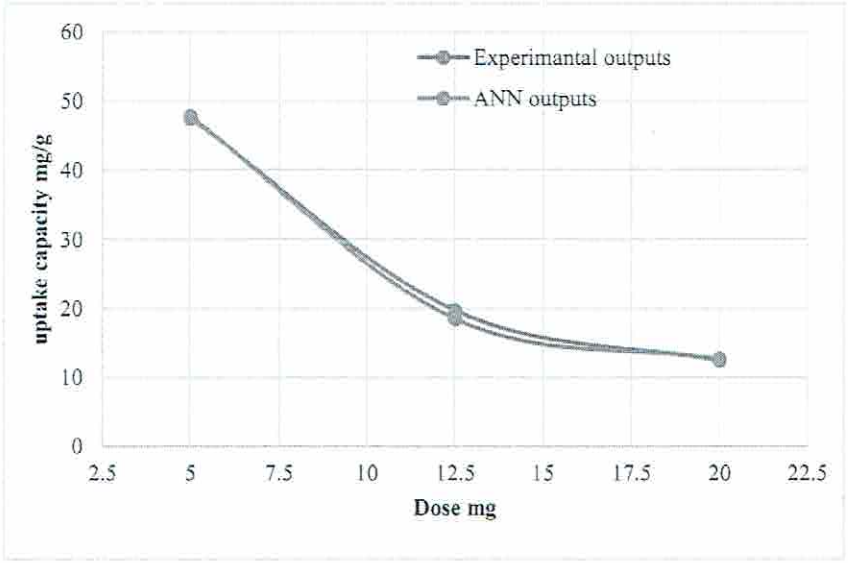


Fig 7. Experimental and ANN output as the function of adsorbent dosage

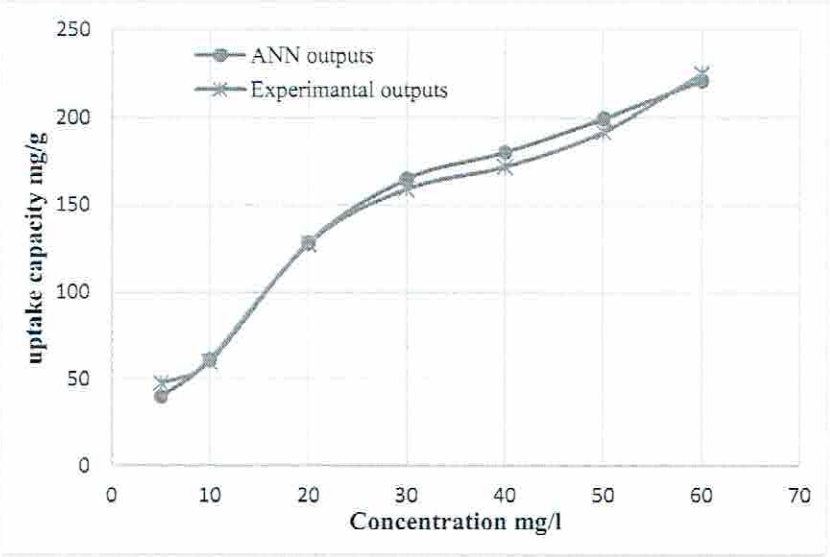


Fig. 8 Experimental and ANN output as the function of initial concentration.

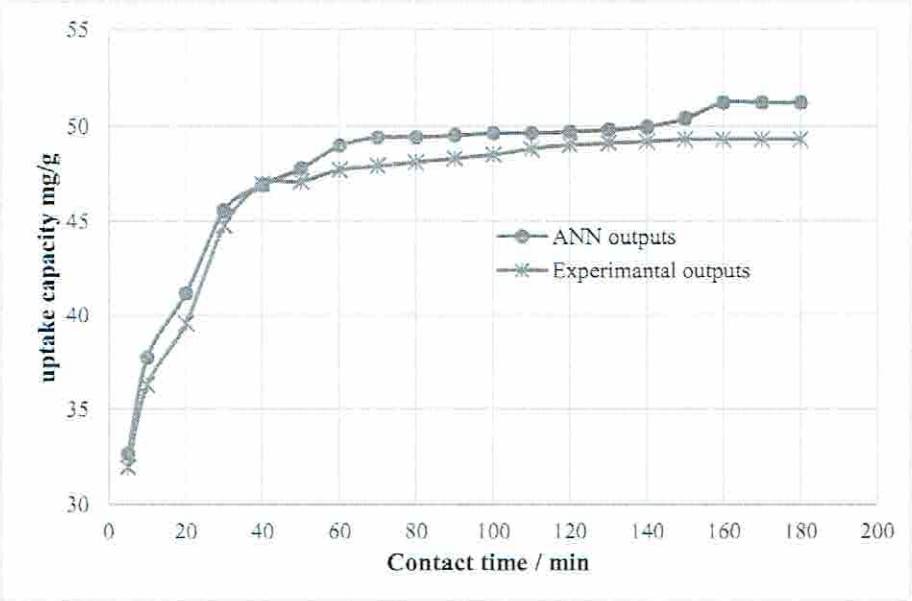
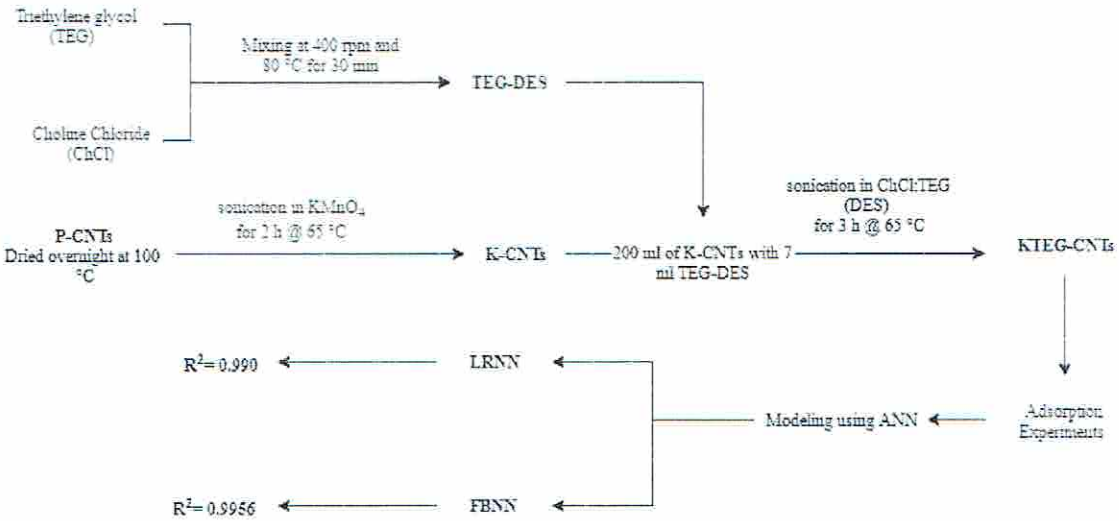


Fig. 9 Experimental and ANN output as the function of contact time.



Scheme 1 The work flow