Artificial Neural Network Approach with Back Propagations for Modelling SS-removal in GMF

SA'AD ABU-ALHAIL AL-KHALIL
Department of Civil Engineering, College of Engineering,
University of Basrah Basra city, South of Iraq, Iraq
saadaboalheel@yahoo.com

Abstract—**In this paper, removal efficiency of suspended solid SS from water was investigated in glasses media filter called GMF. Removal efficiency of SS is obtained by using laboratory glasses media filter "GMF" where this efficiency is used as target function in Artificial neural networks .The remain characteristics are used as input parameters for ANNs where these parameters include raw water quality, operation conditions and glasses media characteristics. The model result showed that optimal number of neurons is nine neurons. As a final observations, the study shows that Artificial neural networks with back propagation algorithm is a good tool that can be used in Prediction Removal efficiency of GMF whereas the results was indicated that the BP model has good convergence performance during training, and the predictions of outflow suspended solid removal efficiency coincided well with the measured values.**

Keyword- Artificial Neural Network, Back propagation, GMF, Water Turbidity

I. Introduction

Deep bed filtration is an effective process in removing particles of various nature and sizes that are present in water and wastewater. Removal of these particles by deep bed filtration involves complex mechanisms. First, particles in suspension are transported near filter grains by mechanisms such as sedimentation, interception, diffusion, inertia and hydrodynamic effect. However, the effective removal of these particles depends on the attachment mechanism, which depends on the surface forces acting between particles and filter grains when their separation distance becomes in the order of nano-meters. The factors, which affect these forces eventually, affect the performance of deep bed filtration. Generally, several mathematical filtration models have been proposed to describe Suspended solid removal by filter whereas macroscopic and microscopic theories are widely accepted. In the macroscopic approach, first order kinetics in the removal of particulate is assumed, whereas the microscopic approach takes into account single collector efficiencies. Filtration equations describing the deep bed filtration was proposed by Yao et al. (1971). Similarly, there are many other models available in the literature including Iwaski (1937), Brinkman, (1947). Happel (1958), Kuwabare (1959), Mints(1966), Ives (1969), Payatakes (1973), Yao (1971), Payatakes and Tien (1976). O'Melia and Ali (1978), Tien and payatakes (1979),Wiesner(1987)Ives(1989),Tien,(1989) [1,2,3,4,5,6, and 7]. In recent years, Artificial neural networks (ANN) have been successfully applied to a wide variety of domains; in the field of water treatment, they have been used, e.g. for predicting the behavior of Wastewater treatment plants [8]. For simulating a combined humic Substances coagulation and membrane filtration process [9] or for estimating H_2O_2 addition critical point in a decoloration process [10]. ANN has been used to predict the breakthrough of activated carbon filters by pesticides in surface waters. [11]. ANN has used also to modeling the oil sands extraction processes [12]. The objective of the neural network is to transform the input into meaningful outputs with minimum error which requires: (i) select suitable variables, whose influence on the output is larger than the influence. (ii) training models of increasing complexity, (iii) selecting the model that generalizes best, provides the most accurate predictions in situations that are not present in the training set, (iv) estimating the performance of the selected model on fresh data, which was used neither for training nor for models election[13]. ANN is a system that is built in accordance with the human brain. Therefore, ANN consists of a few types of many, simple, nonlinear functional blocks, which are called neurons. Neurons are organized into layers, which are mutually connected by highly parallel synaptic weights. It is composed of many artificial neurons that are linked together According to specific network architecture. The Mechanism of Artificial Neural network is the program exhibits a learning ability synaptic weights can be strengthened or weakened during the learning process, and in this way, information can be stored in the neural network where exhibits a very high operational speed[14],[15]. Due to the learning ability, the ANN can behave as an adaptive system, which automatically reacts to changes in its surroundings and it can be to solve even such types of problems as are unsolvable by linear systems. Neural networks are parameterized non- linear models whose parameters are estimated by training from examples. So this advantageously can be used when no satisfactory prior knowledge is available about the physical process, and about the mechanisms that it involves. Furthermore, they require a smaller number of parameters than alternative parameterized non-linear regression models such as polynomials: the number of parameters of neural networks varies linearly with the number of variables of the process, whereas it varies exponentially for polynomials [16]. In this study, artificial neuron network back propagation algorithm is used to predict the performance of glasses media filter.

II. MATERIALS AND METHODS

A. Experimental set up

Filtration Process was carried out by a pilot laboratory filter. The laboratory scale GMF column has been designed to operate identically to full scale granular filters. Using deferent filter media with different bed depth tests on this unit to provide operational data which was scaled up to full size. The media in each filter was supported on a PVC orifice plate drilled with 5 mm holes and covered by a wire mesh to prevent glasses media passing through the orifices. A scale was made in front of tube column to estimate bed height. The laboratory filter was used to measure the experimental removal efficiency of suspension clay in GMF and concentration profiles through the filter bed. The Deep Bed Filter Column is a clear acrylic unit mounted in a floor standing framework approximately 2 m high with flanged end pieces to allow easy access. The medium is supported on a corrosion resistant gauze mesh below is packed 1kg of 0.01m Ballotini to ensure good wash water distribution . Slotted sampling tubes inserted through the wall penetrate into the media, and are fitted with control valves so that suspension samples can be taken are kinetically. Plain tubes also penetrate the medium through the wall to transmit pressure to a manometer system. These sampling and manometer probes are located at 0.02 m depth intervals, but staggered in position, over 0.8 m depth. Consequently as shown in figure (1) where Deep bed filter consist of pump, sump tank , flow controller, rotameter, control valves, tubing, sampling tubes and bank of water differential manometers as shown in figure below. An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

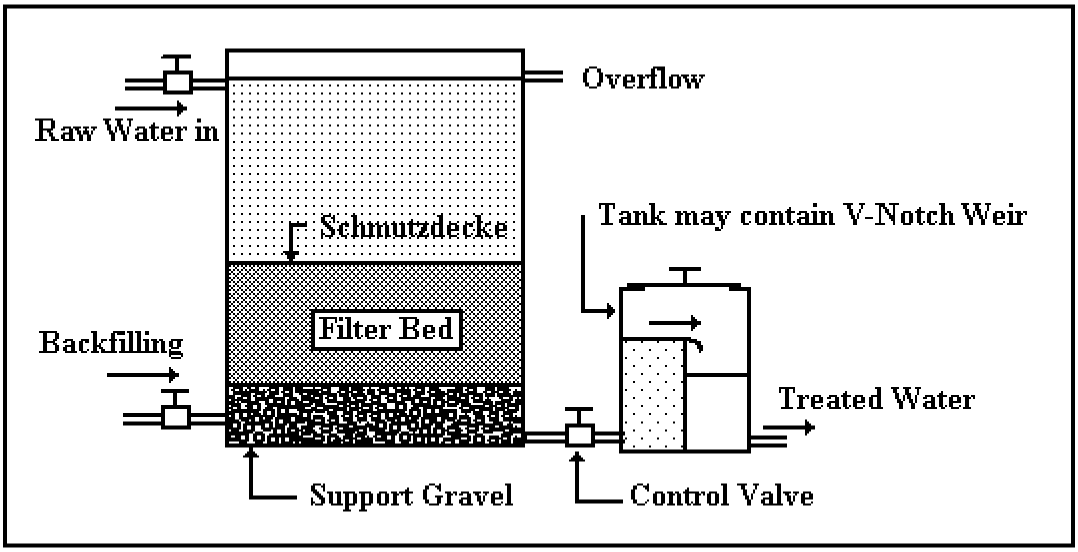


Fig.1. GMF process

B. ANN with BP Model : Derivation the Output Function by General Optimization Method

Artificial Neural network (ANN) with back –propagation algorithm was used for predicting the performance of deep bed .where it is a suitable model for predicting the experimental data. Back propagation was created by generalizing the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities. Feed forward Network often has one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range –1 to +1.

1) Model Formulation

A universal algorithm for the feed ward Neural Network was expressed as follows:

For an input layer neuron is equal to its input(x), $x0j=xj$ for $(j=1, 2\ldots n)$.The subscript j indicates the jth input layer neuron, and n is the number of input layer neurons. The output from the input neuron where weighted and fed to the neuron in the hidden layer .the net input of the ith hidden layer neuron is:

$$y_j = \sum_{i=1}^{n} x_j^{\,0}\ w(i, j) \qquad ...(1)$$

Where w (i, j) indicate the weight of the connection between the jth input layer neuron and the ith hidden layer neuron. In this study, we want to optimize the neurons hidden weight, so the multilayer percepton (MLP) Network with the general Optimization method could be used. A n-ary Gabor polynomial for approximation can be written as,

$$f(x) = w_0 + w_1 x_1 + w_2 x_2 + ... + w_{11} x_1^2 + w_{12} x_1 x_2 + ... w_{12} .... x_2$$

$$y_{out} = w^T . x \qquad ...(2)$$

For the sigmoid function: $y_j^1 = \dfrac{1}{1+\exp^{-w_j^T (x-a_j)}}$ ...(3)

Where j= (1, 2,… m) that is explained in figure (2).

$$y_j^1 = \frac{1}{1+\exp^{-w_j^T (x-a_j^{\,1})}}, \quad y^1 = (y_1^1, y_2^1, y_j^1) \qquad y_j^2 = \frac{1}{1+\exp^{-w_j^T (y-a_j^{\,2})}}, \quad y^2 = (y_1^2, y_2^2)$$

$$y_{out} = \sum_{i=1}^{2} (w^{3T}, y^2) \qquad ...(4) \qquad y_{out} = w_j^3 \frac{1}{1+\exp^{-w_j^T (x-a^2{}_j)}} \qquad ....(5)$$

Equations (4) and (5) were derived for evaluation the performance of deep bed filter.

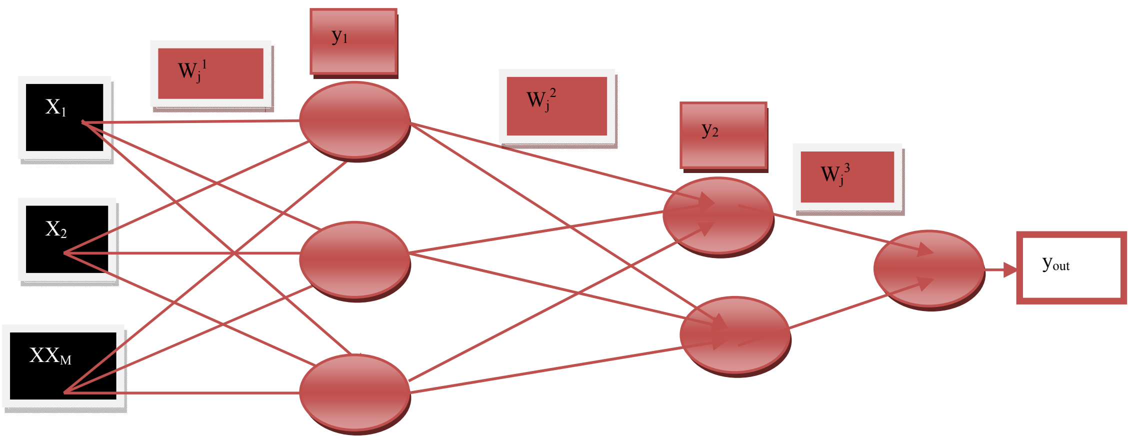


Fig.2. Network Architecture

2) Notation And Error Function

The notation and error functions of a MLP network can be presented by the output weight optimization - back propagation (OWO-BP) algorithm [17]. The OWO-BP technique iteratively solves linear equations for output weights and uses back propagation with full batching to change hidden weights. A set of N is training patterns (Xp, Tp) where the pth input vector xp and the pth desired output vector Tp have dimensions N and Nout respectively. The activation yp(n) of the nth input unit for training pattern p is:

$$y_p(n) = x_p(n) \qquad ...(6) \qquad x_j(p) = w(i, j).y_i(p) \qquad ...(7) \qquad y_i(p) = f(x_p(j))$$

Where, Xp (n) denotes the nth element of Xp . j, is the hidden unit of neurons, ith unit is in any previous layer and w(j,i) denotes the weight connecting the ith unit to the jth unit. Xj(p), is the net input. yj(p) , is the output activation for the pth training pattern. If the activation function f is sigmoidal then,

$$f(x_p(j)) = \frac{1}{1+\exp^{-x_p(j)}}$$

For the k is number of target unit, the net input Xk (p) and the output activation yj(op) for the pth training patterns are:

$$X_k(op) = \sum_i w(k,i).y_i(p) \qquad ...(8) \qquad nX_k(op) = y_k(op)$$

The error function for pth pattern is

$$E_p = \sum_{k=1}^{Nout} (y_k(p) - y_k(op))^2 \quad ..(9)$$

Where yk(p) denote to the measured value of the process output for observation p in kth element. yk(op), is the prediction of the model for observation p. In order to train a neural network in batch mode, the error function for the kth output unit is defined as

$$E(k) = \frac{1}{N}\sum_{k=1}^{Nout} (y_k(p) - y_k(op))^2 \quad ..(10)$$

The overall performance of a MLP Neural Network, measured as Mean Square Error (MSE), can be written

$$E = \sum_{k=1}^{Nout} E(k) = \frac{1}{N}\sum_{p=1}^{N} E_p \quad ...(11)$$

3) Output Weight Change by General Optimization Method

In this section , we are described the model that can be predicted the output weight change in MLP neural Network .The calculation are presented by the general optimization method[18],[19] .This model have been prospered with a single layer of nonlinear neurons. Also, synaptic weight change rules for the neurons in the hidden layer are measured in this section by the same as a method .the weight change rules for the output neurons can be expressed by.

$$\Delta w_j^i = -c\,\frac{\partial E_p}{\partial w^i{}_{j,k}} \quad ...(12) \qquad w_j^{i,new} = \Delta w_j^i + w_j^i \quad ...(13)$$

$$w_{k,j}(p+1) = w^i{}_{j,k}(p) - -c\,\frac{\partial E_p}{\partial w^i{}_{j,k}} \quad ...(14)$$

In this work, the output data consist from one target, which represent the removal efficiency of the filter, therefore k value was substituted by the unity .The error function becomes:

$$E_p = \Sigma_{k=1}^{Nout} (y_1(p) - y_1(op))^2 \quad ..(15)$$

By the derivative of equation (9) and substituted equation (6), get

$$\frac{\partial E_p}{\partial w_j^i} = 2(y_k(p) - y_k(op)).\frac{\partial y(op)}{\partial w_j^i} \quad ...(16)$$

By deriving equation (5) and substituting in equation (16), gets (17).

$$\frac{\partial E_p}{\partial w_j^3} = 2(w_j^3 \frac{1}{1+\exp^{-w_j^T(x-a^2{}_j)}} - y_k(op)).\frac{1}{1+\exp^{-w_j^T(x-a^2{}_j)}} \quad ...(17)$$

Substituting equation (17) into equation (14), we can get the weight change rules for the output neuron of Network Architecture.

$$w_{1,j}^i(p+1) = w_{1,j}^i(p) - c.2(w_j^3 \frac{1}{1+\exp^{-w_j^T(x-a^2{}_j)}} - y_k(op)).\frac{1}{1+\exp^{-w_j^T(x-a^2{}_j)}} ...(18)$$

Synaptic weight change rules for the neurons of the hidden layer is given by the derivation,

$$w_{1,j}^i(p+1) = w_{1,j}^i(p) - c\,\frac{\partial E_p}{\partial w_{i,1}^j} \quad ...(19) \qquad \frac{\partial E_p}{\partial w_i^j} = 2(y_k(p) - y_k(op)).\frac{\partial y(op)}{\partial w_i^j}$$

$$\frac{\partial E_p}{\partial w_i^j} = (y_k(p) - y_k(op)).\frac{\partial}{\partial w_i^j}\frac{1}{1+\exp^{-w_i^T(xp-a^2{}_i)}} \quad ...(20)$$

$$\frac{\partial}{\partial w_i^j}\frac{1}{1+\exp^{-w_i^T(x-a^2{}_i)}} = \frac{\exp^{-w_i^T(x-a^2{}_i)}}{(1+\exp^{-w_i^T(xp-a_i)})^2}.\frac{\partial}{\partial w_i^j}(-w_i^T.(xp-a_i))$$

$$\frac{\partial}{\partial w_i^j}(-w_i^T.(xp-a_i)) = X_i^p \quad ...(21)$$

By Substituting equation (21) into equation (20) and then in equation (19), we get the weight change rules for the neurons of the hidden layer,

$$w_{1,i}^{j}(p+1)=w_{1,i}^{j}(p)-c.2(y_k(p)-y_k(op)).\frac{\exp^{-w_i^T(x-a^2{}_i)}}{(1+\exp^{-w_i^T(xp-a\ i)})^2}.(-X_i^{\ p})...(22)$$

4) Learning Factor Calculation

One problem with using BP in this manner is that the proper value of the learning factor is difficult to determine. If the gradient vector has large energy, we may need to use a small learning factor to prevent the error function E from blowing up. The Learning Factor can be determined as follows. Assume that a learning factor Z is small so that the error surface is well approximated by a frenzied plane. Then the change in E due to the change in w(j,i) in equation (19) is approximately

$$\Delta \mathrm{E}=\frac{\partial \mathrm{E}}{\partial \mathrm{w(j,i)}}.\mathrm{w(j,i)}=-\mathrm{C}\left(\frac{\partial \mathrm{E}}{\partial \mathrm{w(j,i)}}\right)^2 \quad ...(23)$$

Assume that we want to calculate C , so that the error function E is reduced by a factor ζ which is close to, but less than, 1. We then get,

$$\Delta E=\zeta E-E=-C.\sum_j\left(\sum_i\left(\frac{\partial E}{\partial w(j,i)}\right)^2\right) \quad ....(24) \quad \textit{And}$$

$$C=\frac{C'E}{\sum_j\left(\sum_i\left(\frac{\partial E}{\partial w(j,i)}\right)^2\right)} \quad ...(25)$$

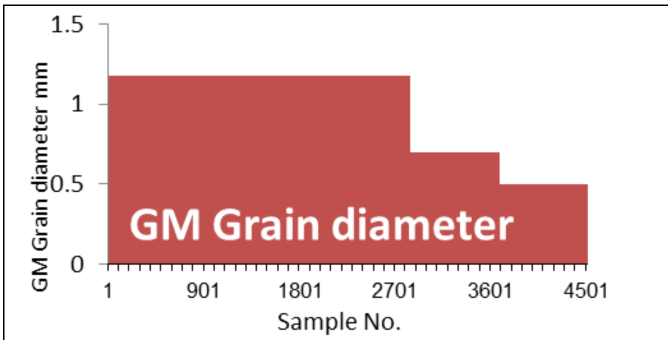
Where,

$$C^{'}=(1-\zeta)$$

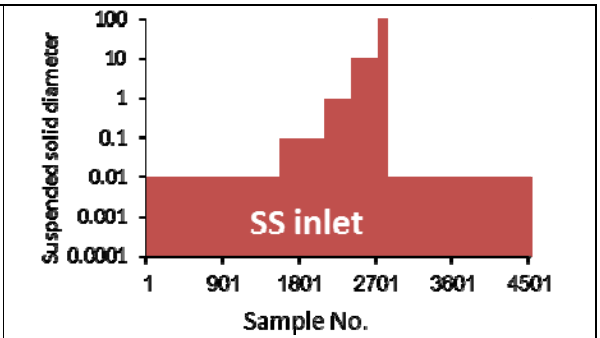
Using these equations, the learning factor Z is automatically determined from the gradient andC' , where C' is a number between 0 and 0.1. The learning factor C in equation (25) is then used in (19).

C. Data set for training and validating the FNN

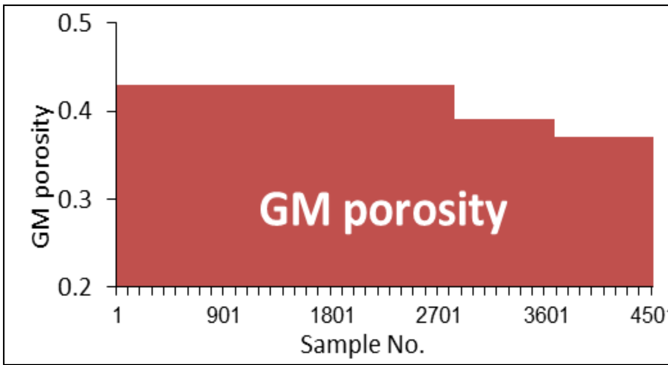
The Input parameters of ANN model with back- propagation algorithm are parameters which are media depths , media grain size , filtration rate , particle diameter , media porosity , feed concentration , operation time. and output parameter is removal efficiency of GMF ,where these data are collocated from our experimental work with different operation condition .this mean that ,we will change one of the input factor and remaining the other constant ,such as the procedure in my experimental work .Three values was used for media grain size , feed concentration and porosity .The data of media depths , filtration rate , particle diameter, and operation time expressed as 7,5,5, and 15 respectively . These data are showed in figure (3 (a – g)).



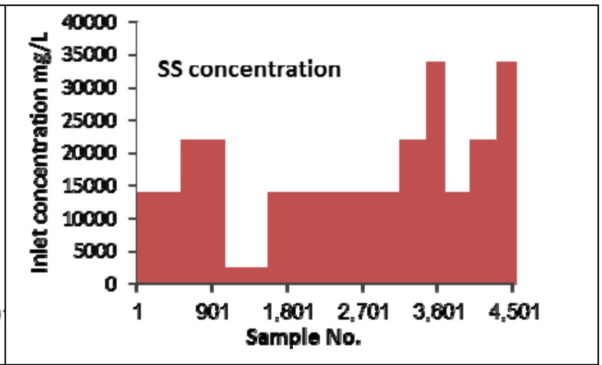
A. Grain diameter



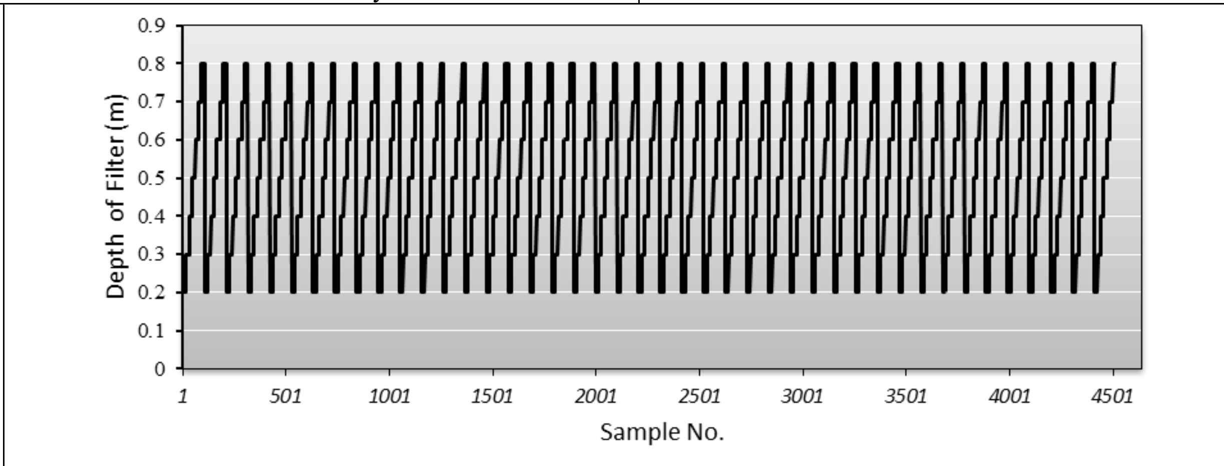
B. Inlet SS diameter um



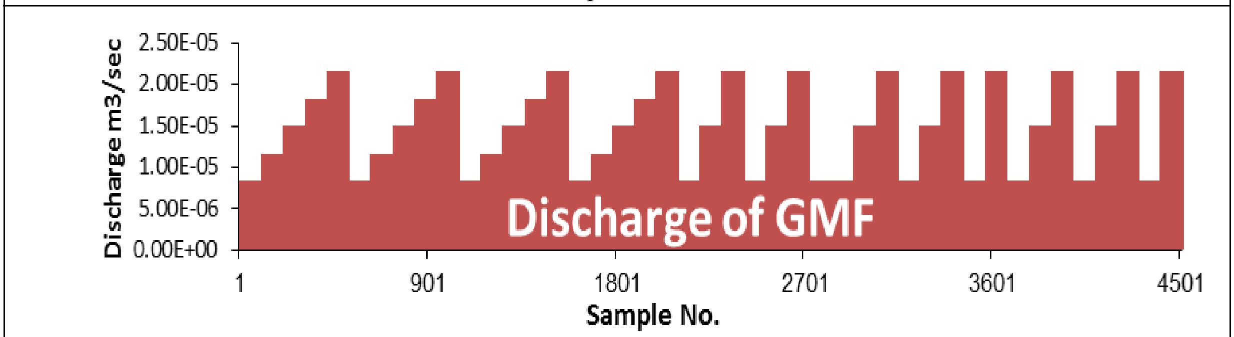
C. Porosity



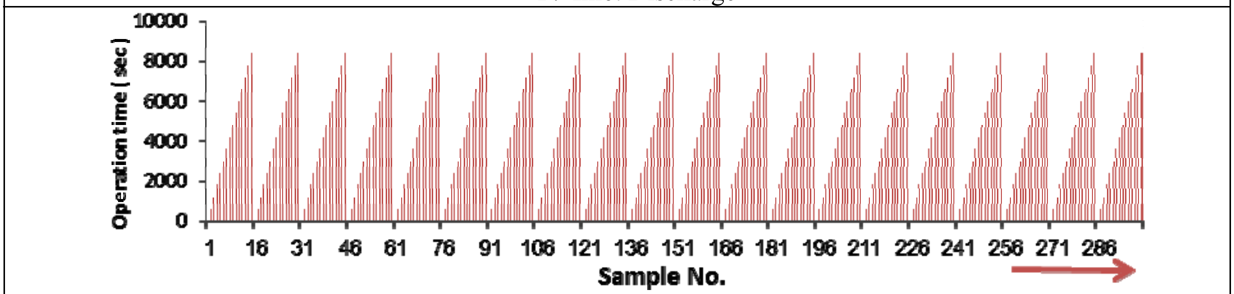
D. Inlet SS Concentration



E. Depth of GMF



F. Inlet Discharge



G. Operation time

Fig.3. Input Parameters in ANN model of GMF

D. ANN- Architectures Investigation in GMF

To build ANN network, it is very important to select the number of hidden neurons that gave best training, validation and test results where our study will investigate the number of hidden neurons in hidden layer of ANN network and the root mean square generalization error. In this present work , the 4515 measurements were randomly divided into three statistically parts: 60% for training , 20% for validation , and 20 % for testing .The data are organized in sequences, each sequence corresponding to one experimental of removal efficiency curve. The performance of ANN was evaluated in terms of root mean square (RMSE) criterion to minimize the error between observed data and the predicted data which is derived in 2.2.1.3 section. Another widely used criterion is the coefficient of determination (R2) or regression analysis, between predicated and observed date, it is provides a measure of a strength of the correlation .The best formula known to calculate this index , is the Pearson product –moment correlation coefficient as follows :

$$R^2 = \left[\frac{\sum_{i=1}^{N} (y_i - y_{i,ave})\left(y_i(p) - y_{i,ave}(p)\right)}{\sqrt{\sum_{i=1}^{N} (y_i - y_{i,ave}) \sum_{i=1}^{N} \left(y_i(p) - y_{i,ave}(p)\right)^2}} \right]^2 \quad ...(26)$$

Where:

y_i is measured value ; $y_{i,ave}$ is average of measured value; $y_{i,ave(p)}$ is simulated value; $y_{i,ave\,(p)}$ is average of simulated value

Levenberg- Marquardt training method was applied in order to improve convergence speed and performance of the network [20], [21],[22]. Another widely used criterion is the coefficient of efficiency which can be determined as follows

$$CE = 1 - \frac{\sum_{i=1}^{N} (yi - yi(p))^2}{\sum_{i=1}^{N} (yi - yi, ave)^2} \quad ...(27)$$

Different number of neurons 2 to 45 were investigated depending on RMSE, CE, R^2. The main objective of regression analysis is to determine the suitable locations for the data required for the modelling activities .The scatter diagrams were constricted for several cases of several inputs and outputs. For each number of neurons, a linear regression analysis were conducted and the coefficient of determination (R2) was computed to determine how a variation in the input parameter (s) could explain variations in the corresponding output . In order to minimize the number of parameters, seven hidden neurons were selected for the FNN model where it presented greater value of (R^2) and lower value of RMSE than other neurons as shown in figures (4&5).

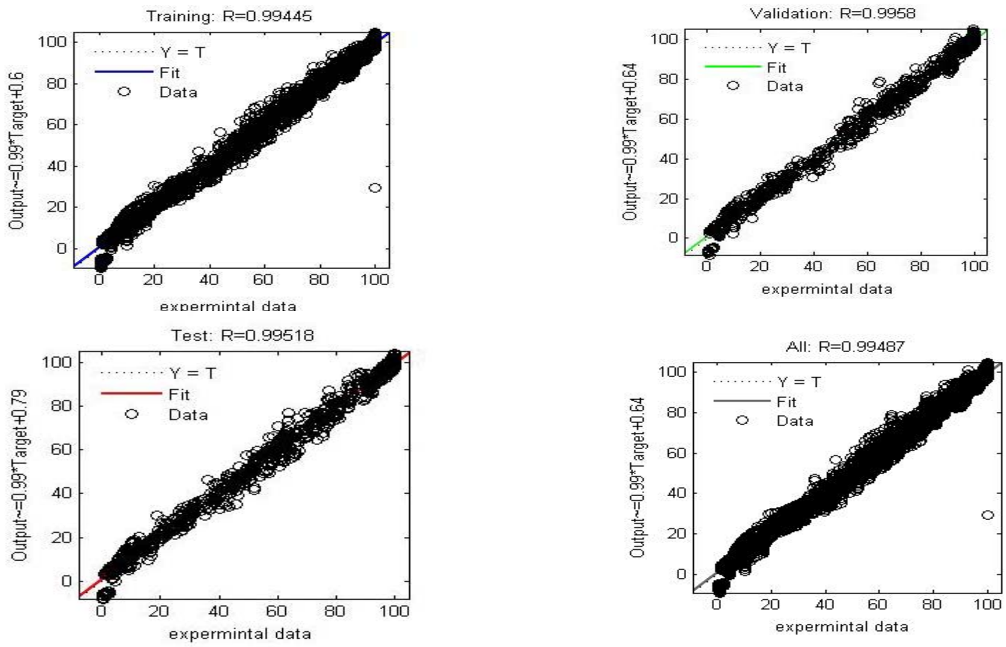


Fig.4. Regression Analysis at Seven Neurons Hidden Layer Using ANN

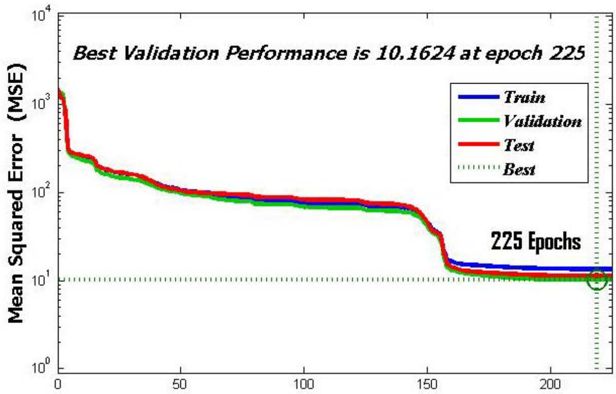


Fig.5. Early Stopping Criteria for Seven Neurons

The optimum number of neuron in ANN model is depend on The RMSE for each number of neuron whereas the optimum number of neuron is shown in figure (6).

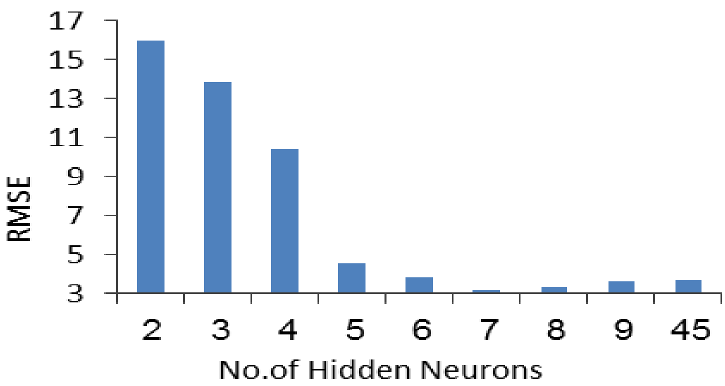


Fig.6. Neurons in ANN Hidden Layer

The final optimized architecture of FNN is shown in Fig. (7) Corresponding to 7 neurons and 72 rules in the ANN architecture.

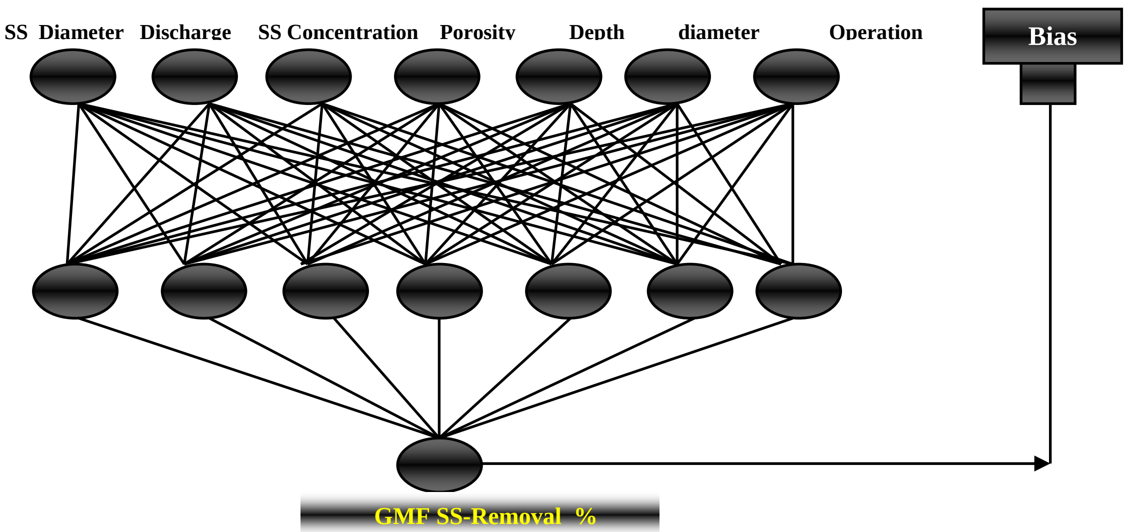


Fig.7. Optimum Architectures of the feed-forward ANN

III. RESULTS AND DISCUSION

The simulation result of SS- removal efficiency with respect to sample number for each neuron can be obtained as shown in figures (8)&(9) below where as mentioned previously the ANN with seven neurons gave best simulation result. Figures (8) and (9) represents a review for the fast prediction of removal efficiency in GMF where we can design and analysis any filter unit directly by using this figure and all figures of input data (Figure (3)). In design, we can selected the desirable SS-removal efficiency of GMF and project this value on (9) we get the sample number and then we will use the figure (3a-g) of input parameters to Known the specification of this sample which include (water quality , media properties and operation time) . In the analysis of constructed filter unit, we can take the specification of filter which includes (water quality, media properties and operation time). Then we can get on sample number from input parameter figures and by project the sample number on figure (9), we get the prediction value of removal efficiency for this unit.

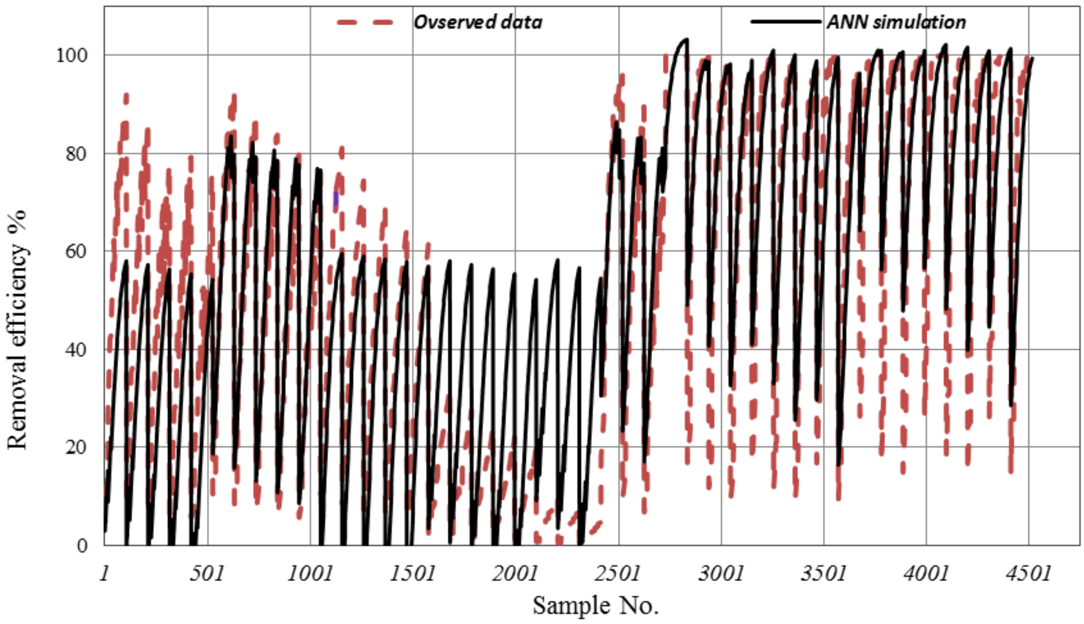


Fig.8. Worst Simulation Results of GWF at only two neuron hidden layer using Back Propagation ANN Model

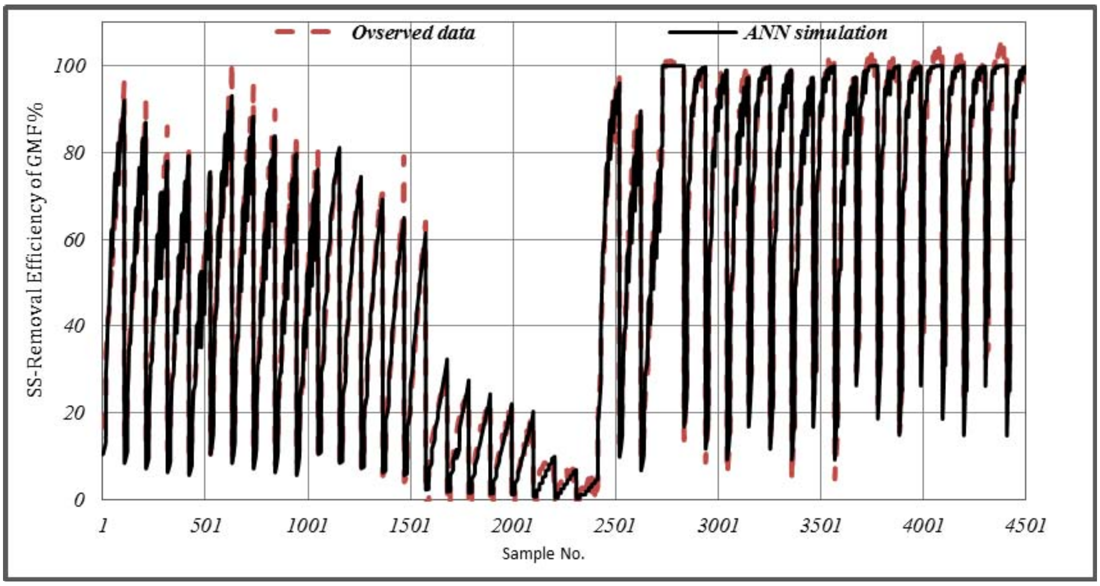


Fig.9.Optimum Simulation Results of GWF at seven neuron hidden layer using Back Propagation ANN Model

IV. CONCULSION

In this paper, the model and mechanism of ANN reviewed and the formulas for each layer given. ABP neural network model was developed for the SS- removal efficiency in GMF. The suspended solid removal efficiency that obtained from laboratory filter used as target function in ABP neural network while the other properties of filter such as raw water quality, operation conditions and media characteristics used as input parameters. As a result, ANN with back propagation algorithm is a good tool that can be used in Prediction Removal efficiency of GMF where Results indicated that the BP model has a good convergence performance during training, and the predictions of outflow suspended solid removal efficiency coincided well with the measured values.

REFERENCES

[1] Iwasaki, T., 1937. Some notes on sand filtration. J. Am. Water Works Assoc. 29: 1591.C.
[2] Tien,B.V. Remoarao, 2007,(Granuler Filtration of Aersols and Hydrosols).2nd.Ch.2: p.29. ISBN:978-1-85617-458-9.
[3] Brinkman, H.C., 1947. A calculation of the viscous force exerted by a flowing fluid on a dense swarm of particles, J. Appl. Sci. Res. A1,p. 27.
[4] Payatakes, A.C., Tien, C., Turian, R.M., 1973a. A new model for granular porous media: Part I. Model formulation. AIChE J. 19, 58.
[5] Payatakes, A.C., Tien, C., Turian, R.M., 1973b. A new model for granular porous medium- Part II, Numerical solution of steady state incompressible Newtonian flow through periodically constricted tubes. AIChE J. 19 (1), 67.
[6] O'melia and Ali,W,(1978).The role of retained particles in deep bed filtration ,Prog. Wat.Techno. 10: p. (123-137).
[7] H. Stoppiglia, G. Dreyfus, R. Dubois, Y. Oussar, 2003. Ranking a random feature for variable and feature selection . J. Mach. Learn. Res. 1399–1414.
[8] Belanche, J.J. Valdes, J. Comas, I.R. Roda, M. Poch, Towards 1999.a model of input–output behaviour of wastewater treatment plants using soft computing techniques, Environ. Model. Software 14 (5) 409–419.
[9] M. Al-Abri, N. Hilal, 2008.Artificial neural network simulation of combined humic Substance coagulation and membrane filtration, Chem. Eng. J. 141 [27–34.].
[10] O.L.C. Guimaraes, D.N.Va. Filho, A.F. Siquerra, H.J.L. Filho, M.B. Silva, 2008.Optimization of the azo dyes decoloration process through neural networks: determination of the H2O2 critical point, Chem. Eng. J. (141) 35–41.
[11] C. Faur , Modeling the breakthrough of activated carbon filters by pesticides surface waters with static and recurrent neural networks. Chemical Engineering Journal 145 (2008) 7–15.
[12] Qing J. Zhang, Ronald P. Sawatzky, E. Dean Wallace, Michael J. London, and Stephen J. Stanley 2004. Artificial neural network modeling of oil sands extraction processes J. Environ. Eng. Sci. 3(S1): S99–S110.
[13] Rumelhart, D. E. & McClelland, J. L., 1986. Parallel Distributed Processing Explorations in the Microstructure of Cognition, Vols. 1 and 2. MIT Press, Cambridge, MA.
[14] Lippmann, R. P., 1987.An introduction to computing with neural nets. IEEE Acoust. Speech Signal Process, 4(2), 4-22.
[15] Fitpatrick,j.A., and Spillman ,L.A,1973.filtration of aqueous Latex suspension Through beds of glass speres .J.Collid Inter sci.,43(2) , 350-369.
[16] Marren, A., Harston, C., Pap, R., 1990. Hand book of Neural Computing Applications. Academic Press Inc., San Diego, CA.
[17] Howard Demuth, Mark Beale, Martin Hagan, 2006.Neural Network Toolbox User's Guide version (5). The Mathworks, Inc.
[18] M.T. Manry, et al, 1994. Fast Training of Neural Networks for Remote Sensing, Remote Sensing Reviews, Vol. 9, 77-96.
[19] K. Rohani, M.S. Chen and M.T. Manry, 1992.Neural subnet design by direct polynomial mapping, IEEE Transactions on Neural Networks, Vol. 3, No. 6, 1024-1026.
[20] D.B. Parker, 1982. "Learning Logic," Invention Report S81-64, File 1, Office of Technology Licensing, Stanford University.
[21] S.Haykin, 1999 .Neural Networks: A comprehensive Foundation. 2nd edition . Prentice Hall Co., NJ, USA.
[22] D.R Baughman ,Y,A.Liu , ,1995.Neural network in Bioprocessing and chemical Engineering , Acadmic press , San Diego , CA,USA. L.A.