International Journal of Applied Environmental Sciences
ISSN 0973-6077 Volume 2 Number 1 (2007) pp. 155–166

http://www.ripublication.com/ijaes.htm

Predicting Re-aeration Rates Using Artificial Neural Networks in Surface Aerators

[1]Achanta Ramakrishna Rao, PhD
and [2]Bimlesh Kumar, ME PhD

[1]Associate professor in the department of Civil Engineering,
Indian Institute of Science, Bangalore, India.
Email- ark@civil.iisc.ernet.in
[2]Research Student in Civil Engineering department at
Indian Institute of Science, Bangalore, India.
Email- bimk@civil.iisc.ernet.in

Abstract

The paper illustrates the application of a neural-network model to the modeling of mass transfer in unbaffled surface aeration tank fitted with six flat bladed rotors under geometrically similar conditions. Back-propagation with Levenberg-Marquadt algorithm is used for the modeling of neural-network. This paper discusses the ability of neural-network to model the mass transfer rate in unbaffled surface aeration tank. A thorough sensitive analysis has also been made to ascertain which variables are having maximum influence on re-aeration rates.

Keywords: Levenberg-Marquadt algorithm, Neural network, Sensitivity analysis, Surface aerator, Theoretical power per unit volume.

Introduction

Aeration is one of the important processes employed in water and wastewater treatment to reduce BOD. The basic phenomenon behind the process of aeration is a gas transfer, in which the gas molecules are exchanged between the liquid and the gas at the gas-liquid interface (Fair et al., 1971). The aeration process is also used to either removal of volatile substances and gases present in water and wastewater or improve the DO content in the water and wastewater or both at the same time.

Unfortunately, the phenomena underlying aeration process have not been fully understood, and a universal basis for design and operation does not exist. Furthermore, oxygen transfer rate from gas to liquid phase is dependent on several factors such as: method of aeration, power input, aerators' geometry, types of rotors and its speed, physicochemical properties of the liquid and mixing intensity or turbulence.

Artificial neural networks (ANNs) allows model of complex systems to be built without requiring the explicit formulation of the possible relationships that may exist between variables. Data-driven modeling employing artificial intelligence and machine learning methods are finding increasing relevance and importance in various engineering process. The goal of data-driven modeling is to build a system that can adapt and learn from practical data. In the past few years, artificial neural networks (ANNs) have been used in describing and modeling wastewater treatment processes. Artificial neural network models can be identified without a detailed knowledge of the kinetics of the system to be modeled. Also, ANN models can potentially contain a great deal of information about the system itself, including the same type of information contained in conventional deterministic models. The fact that these models can be continuously updated with minimal resource requirements makes them very attractive for application in a real-time control scenario (Ahmed et al., 2004). Artificial neural networks, coupled with an appropriate learning algorithm, can be used to learn complex relationships from a set of associated input-output vectors. In essence, a neural network can be viewed simply as a large dimensional regression model which can be used to model the aerators. Present work concentrates on surface aerators, as it was found that surface aerators (among the different types of aerators available) are more popular because of their better efficiency and ease in operation. Thus aim of the present study is to train a neural network to predict the oxygen transfer rates of the surface aerators and ascertain which parameters are having more influence on the process.

Neural Networks

ANNs are massively parallel, distributed and adaptive systems, modeled on the general features of biological networks with the potential for ever improving performance through a dynamical learning process (Bavarian, 1988). Neural networks are made up of a great number of individual processing elements, the neurons, which perform simple tasks. A neuron, schematically represented in Fig. 1, is the basic building block of neural network technology which performs a nonlinear transformation of the weighted sum of the incoming inputs to produce the output of the neuron. The input to a neuron can come from other neurons or from outside the network. The nonlinear transfer function can be a threshold, a sigmoid, a sine or a hyperbolic tangent function.

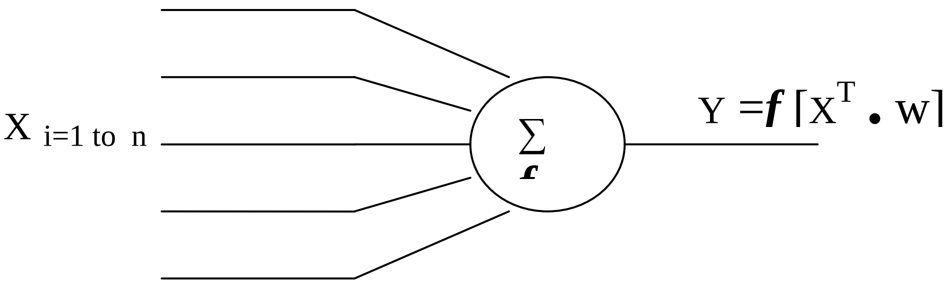

Figure 1 : A simple processing neuron

The most versatile learning algorithm for the feed forward layered network is back-propagation (Irie and Miyanki, 1988). The back-propagation learning law is a supervised error-correction rule in which the output error, that is, the difference between the desired and the actual output is propagated back to the hidden layers. Now, if the error at the output of each layer can be determined, it is possible to apply any method which minimizes the performance index to each layer sequentially.

Back-propagation algorithm with Levenberg-Marquardt algorithm

Multi-Layer Perceptrons (MLP) are perhaps the best-known type of feed forward networks. MLP has generally three layers: an input layer, an output layer and an intermediate or hidden layer. Neurons in the input layer only act as buffers for distributing the input signal x_i to neurons in the hidden layer. Each neuron j in the hidden layer sums up its input signals xi after weighting them with the strengths of the respective connections w_{ji} from the input layer and computes its outputs y_j as a function ***f*** of the sum, viz.

$$y_j = \boldsymbol{f}\,(\Sigma w_{ji}\, x_i) \tag{1}$$

where, ***f*** can be a simple threshold function or a sigmoid, hyperbolic tangent or radial basis function.

The output of neurons in the output layer is computed similarly. The back-propagation algorithm, a gradient descent algorithm, is the most commonly adopted MLP training algorithm. It gives the change Δw_{ji} in the weight of a connection between neurons j and i as follows:

$$\Delta w_{ji} = \eta\, \delta_j x_i \tag{2}$$

Where η is a parameter called the learning rate and δj is a factor depending on whether neuron j is an output neuron or a hidden neuron. For output neurons,

$$\delta_j = \left(\frac{\partial f}{\partial net_j}\right)\left(y_j^t - y_j\right) \tag{3}$$

and for hidden neurons,

$$\delta_j = \left(\frac{\partial f}{\partial net_j}\right)\sum_q \left(W_{qj}\ \delta_q\right) \tag{4}$$

In equation (3), net_j is the total weighted sum of input signals to neuron j and $y_j^{(t)}$ is the target output of neuron j. As there are no target outputs for hidden neurons, in equation (4), the difference between the target and actual output of a hidden neuron j is replaced by the weighted sum of the δ_q terms already obtained for neurons q connected to the output of j. Thus, iteratively, beginning with the output layer, the δ term is computed for neurons in all layers and weight updates determined for all connections.

Back-propagation searches on the error surface by means of the gradient descent technique in order to minimize the error. It is very likely to get stuck in local minima. Various other modifications to back-propagation to overcome this aspect of back-propagation have been proposed and the Levenberg-Marquardt modification (Hagan and Menhaj, 1994) has been found to be a very efficient algorithm in comparison with the others like Conjugate gradient algorithm or variable learning rate algorithm.

Levenberg-Marquardt works by making the assumption that the underlying function being modeled by the neural network is linear. Based on this calculation, the minimum can be determined exactly in a single step. The calculated minimum is tested, and if the error there is lower, the algorithm moves the weights to the new point. This process is repeated iteratively on each generation. Since the linear assumption is ill-founded, it can easily lead Levenberg-Marquardt to test a point that is inferior (perhaps even wildly inferior) to the current one. The clever aspect of Levenberg-Marquardt is that the determination of the new point is actually a compromise between a step in the direction of steepest descent and the above-mentioned leap. Successful steps are accepted and lead to a strengthening of the linearity assumption (which is approximately true near to a minimum). Unsuccessful steps are rejected and lead to a more cautious downhill step. Thus, Levenberg-Marquardt continuously switches its approach and can make very rapid progress.

The equations for changing the weights during training in Levenberg-Marquardt method are given as follows:

$$\Delta W = \left(J^T J + \mu I\right)^{-1} J^T e \tag{5}$$

Where J is the Jacobian matrix of the derivative of each error to each weight, μ is a scalar and e is an error vector. The Levenberg–Marquardt algorithm performs very well and its efficiency is found to be of several orders above the conventional back propagation with learning rate and momentum factor.

Surface Aerators

A typical surface aerator is shown in Fig. 2. The various geometric dimensions of the aerator are: A (cross-sectional area), H (water depth), h (distance between the horizontal bottom of the tank and the top of the blades) and D (diameter of the rotor). The rotor is fitted with six flat blades in symmetrical and evenly manner such that *b* and *l* are their linear dimensions. Rotor shaft is connected to a DC motor to rotate the rotor at desired speed.

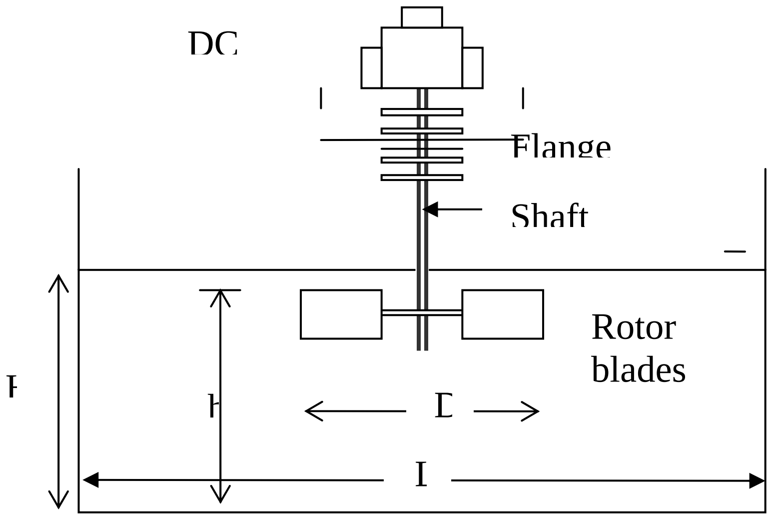

Figure 2 : Schematic diagram of a surface aerator

The rotor is rotated to create turbulence in the water body so that aeration takes place through the interface of atmospheric air and water surface. The rate of oxygen transfer depends on a number of factors like intensity of turbulence which in turn, depends on the speed of rotation, size, shape and number of blades, diameter and immersion depth of the rotor, and size and shape of aeration tank, as well as on the physical, chemical and biological characteristics of water (Rao 1999). The aeration process generally depends on three types of variables namely geometric, physical and dynamic variables.

Functions relating the oxygen transfer rates and these three types of the variables may be given by the following equations (Rao, 1999):

$k = f\,(\sqrt{A}/D,\ H/D,\ l/D,\ b/D,\ h/D,\ X$ (6)

Where $k = K_L a_{20}\,(\square/g^2)^{1/3}$ is the non-dimensional oxygen transfer parameter and $K_L a_{20}$ is the oxygen transfer coefficient at 20°C and the dynamic variables $X = F^{4/3}R^{1/3}$ is the parameter governing the theoretical power per unit volume in which $F = N^2D/g$ is the Froude number, and $R = ND^2/\square$ is the Reynolds number. The intensity of turbulence and wave action on the water are the major sources normally associated with surface aeration. Turbulence and viscous effects are generally described by the Reynolds number (R), where the surface wave action is described by the Froude number (F). The first six non-dimensional parameters represent the "geometric similarity" of the system and the last parameter represents the "dynamic similarity".

When the geometric similarity conditions are maintained, the functional relationship represented by Eq. 6 is reduced to a function of dynamic similarity (Rao, 1999) for any shape of aeration tank.

$k = f\,(X)$ (7)

As X is a combination of N, D and □and equals to $N^3D^2/g^{4/3}\ \square^{-1}$. Eq.7 can be expressed under geometrically similar conditions as follows:

$k = f\,(N, D, \square)$ (8)

Method of Analyzing

Experimental data obtained from our earlier studies (Rao, 1999 and Rao et. al., 2004) under geometrically similar conditions on six surface aerators (three sizes of square shaped and three sizes of circular shaped surface aerators) were used in the present analysis. In additions to this referenced data, some more experiments were conducted in both square and circular surface aerators to provide a sufficient number of data to the neural networks. The C/s area of a pair of square and circular surface aerators are the same as equals to A=1m^2, 0.5184m^2 and 0.1681m^2. Conditions of geometric similarity: $\sqrt{A}/D = 2.88$, $H/D = 1.0$, $l/D=0.3$, $b/D = 0.24$ and $h/H = 0.94$: as suggested by Udaya et al, 1991 were maintained in all the six surface aerators.

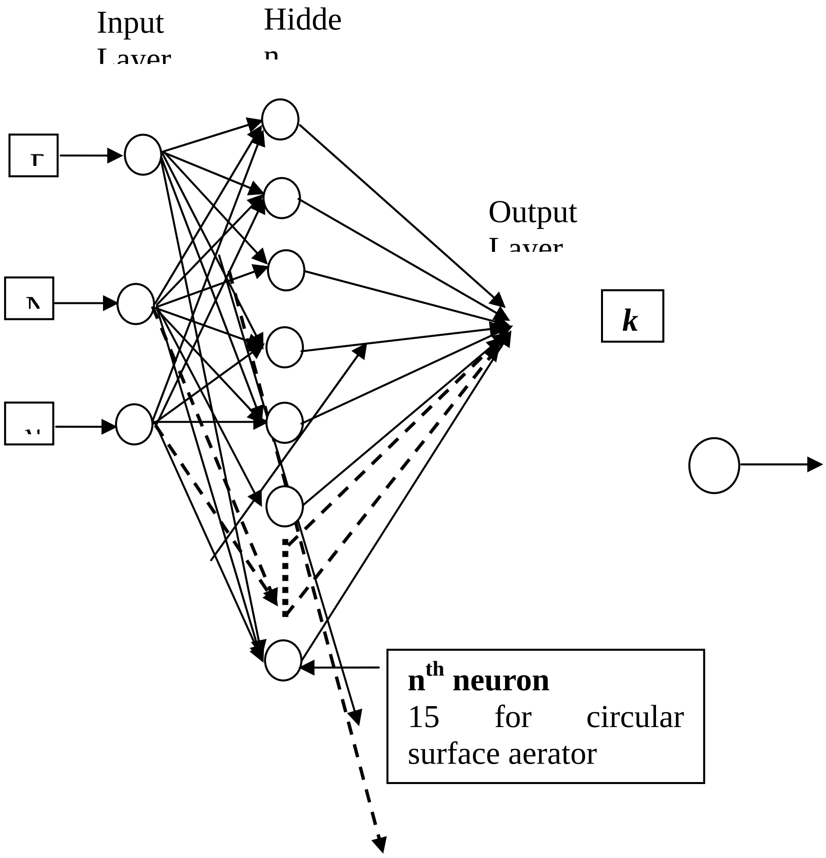

Figure 3 : Neural net architecture used in the analysis

The overall data points used in circular surface aerators are 210, out of which 150 has been assigned to training, 20 as cross-validation and 40 as test sets. Similarly in square surface aerators, the overall number of data points used is 230, out of which same number of data points 150 has been assigned for training, 20 as cross-validation as rest 60 for test evaluation. The aim is to produce and train a multilayer neural feed forward network in order to implement a function mapping the input vectors of real values (N, D, [illegible] to the real output values *k*.

The Transfer function used in the neural network is bipolar sigmoid. The idea behind choosing bipolar sigmoid functions as transfer functions is that it bears a greater resemblance to the biological neurons. In case of sigmoid functions, the output of the neurons varies continuously but not linearly with the input. The architecture of neural net used for analysis of surface aerators is shown in Fig.3. Before giving the input and output vectors to the neural net dataset have been randomized.

Results and Discussion

All the experimental data points collected were used to build and validate the ANN correlations. In the training algorithm of the back-propagation neural networks, the Mean Squared Error (MSE) was minimized for each epoch, i.e. iteration. The number of epochs was set 1000, and a termination rule was maintained as MSE of 0.0001 or maximum epoch. Weights are updated though batch learning. Batch learning computes the weight update for each input sample and store these values (without changing the weights) during one pass through the training set which is called an epoch. At the end of the epoch, all the weight updates are added together, and only then will the weights be updated with the composite value. This method adapts the weights with a cumulative weight update, so it will follow the gradient more closely.

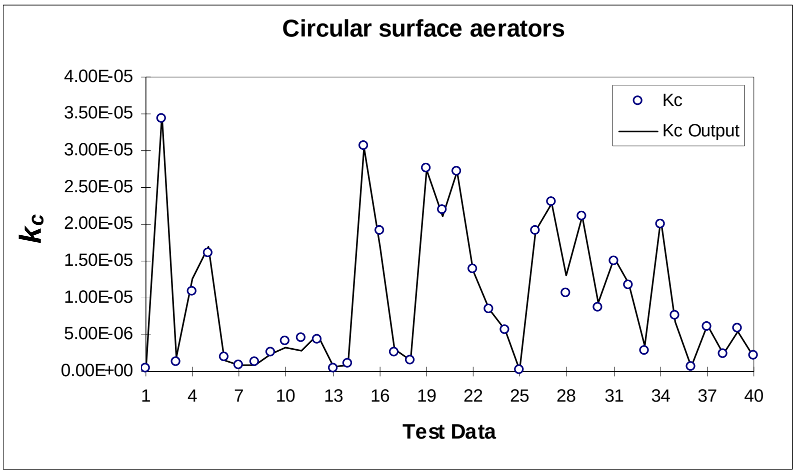

Figure 4 : Neural network modeling of circular surface aerators

It can be clearly seen from Fig. 4 and Fig.5 that the linear coefficient of correlation is very high between observed experimental data and values predicted through neural nets and it is 0.997 for circular surface aerators and 0.998 for square surface aerators. This shows the learning and generalization performance of the network is good. Thus it can be said that aeration phenomena can be modeled through neural networks. Therefore instead of proposing a compact correlation for modeling the re-aeration rates, a neural net with weights can be supplied for the modeling.

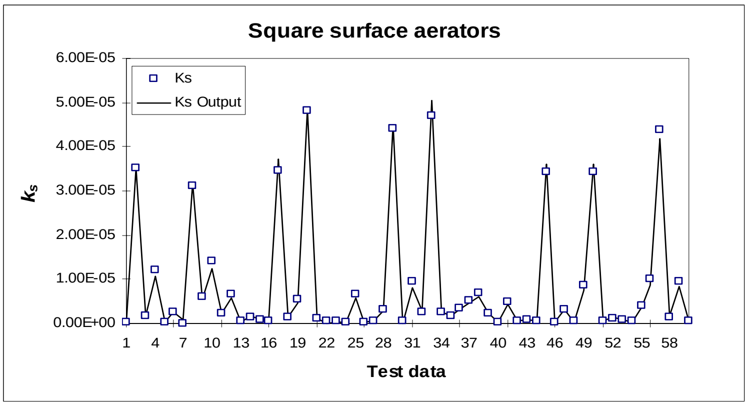

Figure 5 : Neural network modeling of square surface aerators

Sensitivity Analysis

Sensitivity analysis is a method for extracting the cause and effect relationship between the inputs and outputs of the network. The basic idea is that each input channel to the network is offset slightly and the corresponding change in the output is reported. To ascertain the influence of the input variables on output variables, sensitivity analysis is also carried out. This testing process provides a measure of the relative importance among the inputs of the neural model and illustrates how the model output varies in response to variation of an input. The first input is varied between its mean ± standard deviations while all other inputs are fixed at their respective means. Similar exercises have been made for all others input parameters.

Effect of rotor diameter

The re-aeration rates as a function of rotor diameter by keeping the other parameters as constant, as predicted by the ANN model is shown in Fig.6. Overall rotor diameter is having maximum influences on re-aeration rates. An increase in rotor diameter leads to an increase of the pumping capacity of the impeller, which results in an increase of re-aeration rates. It can be seen from the Fig. 10 that in the both cases of square and circular surface aerators, re-aeration rate increases as the rotor diameter increase.

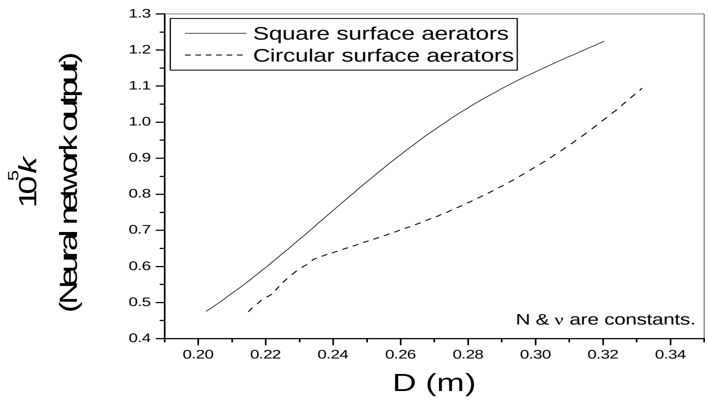

Figure 6 : Effect of rotor diameter on re-aeration rate

Effect of rotor speed

Figure 7 gives the re-aeration characteristics as a function of rotor speed. The oxygen transfer parameter k is found to increase with the rotor speed, which is in agreement with the literature (Rao, 1999). Increasing N intensifies the turbulences and thus the re-aeration rate. The nature of the curve shows that at lower speed increase in re-aeration rates is very slowly, and in the intermediate range the change in the slope is very steep.

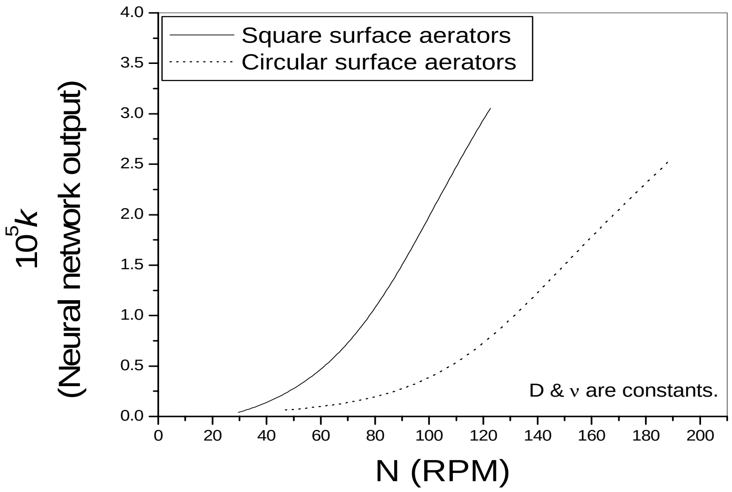

Figure 7 : Effect of rotor speed on re-aeration rate

Effect of viscosity

It is having least influence on the surface aeration process. Figure 8 illustrates the effect of viscosity on re-aeration rates. As it can be observed from Figs., *k* decreases with increasing viscosity. This means at lower viscosity re-aeration rates are high, which is in agreement with literature (Albal et al., 1983). In fact, an increase of the viscosity has been reported to decrease the turbulences leading to a decrease of *k*.

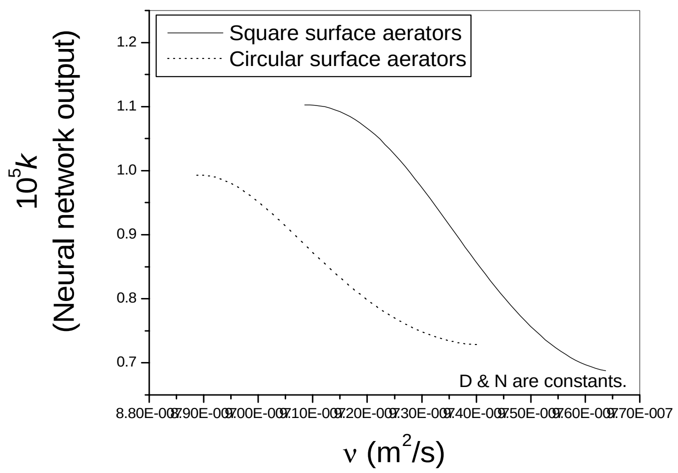

Figure 8: Effect of viscosity on re-aeration rate

Appendix

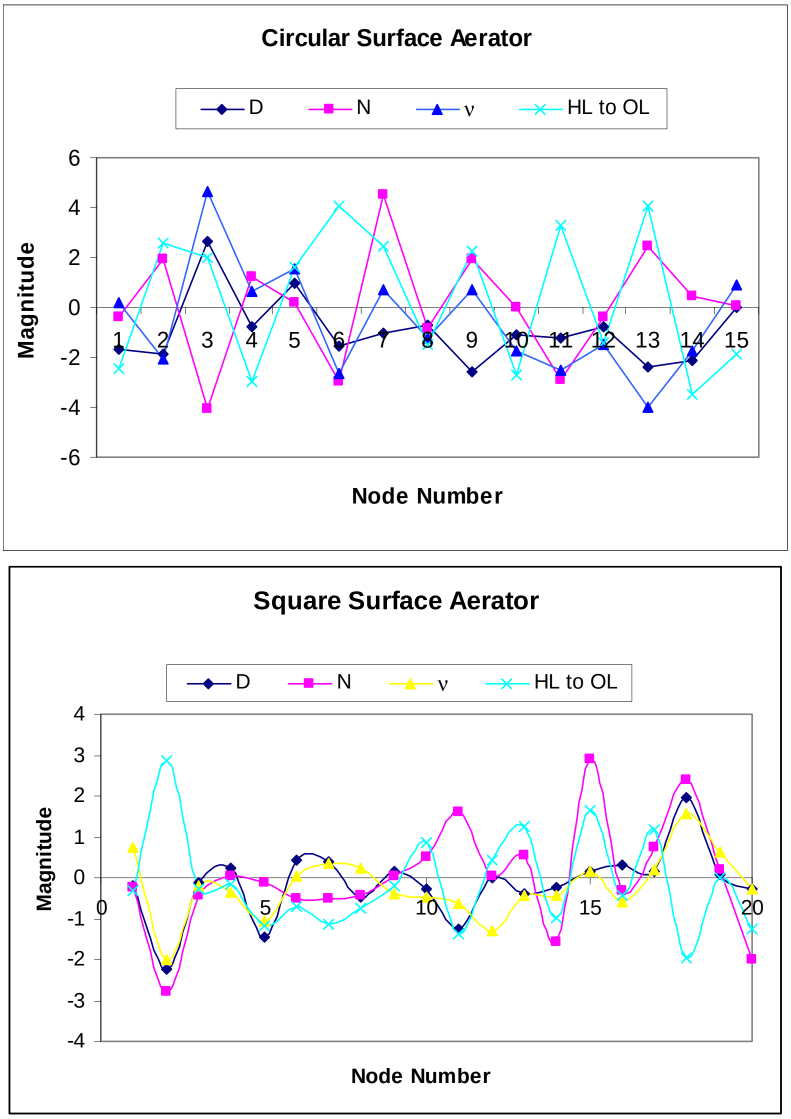

Figure 9 : Weight plots **a)** Circular Surface Aerators **b)** Square Surface Aerator [HL - Hidden Layer and OL – Output Layer]

Conclusion

- The results presented in this paper have clearly shown that the neural network methodology can be used efficiently to model the aeration phenomena in surface aerators. The main advantage of neural networks is to remove the burden of finding an appropriate model structure to fit experimental data or to find a useful regression equation. The network showed excellent learning performance and achieved good generalization. Thus this study provided an additional insight into abilities of neural networks to compete with empirical models. Neural network outputs well match the simulation equation given in the literatures.

- Sensitivity analyses with the trained neural net or during training could provide valuable additional information on the relative influence of various parameters on the aeration systems. The rotor diameter is having more influence on re-aeration than any other parameter. The re-aeration rate increases with increases in rotor diameter. The re-aeration rate will increase on increasing speed and it is found to be decreasing with increasing viscosity.

Notation

MSE = mean squared error;
NMSE = normalized mean squared error;
MAE = mean absolute error;
R^2 = linear correlation coefficient;
LM = Levenberg-Marquardt method;
A = cross-sectional area of an aeration tank (L^2);
b = width of the blade (L);
D = diameter of the rotor (L);
$\boldsymbol{f}$ = transfer functions;
$F = N^2D/g$, Froude number;
H = depth of water in an aeration tank (L);
h = distance between the top of the blades and the horizontal floor of the tank (L);
J = jacobian matrix;
K_La_{20} = overall oxygen transfer coefficient at 20 ^{o}C;
$k = K_La_{20}\ (\square/g^2)^{1/3}$, non-dimensional oxygen transfer coefficient;
k_c = non-dimensional oxygen transfer coefficient for circular tanks;
k_s = non-dimensional oxygen transfer coefficient for square tanks;
N = rotational speed of the rotor with blades (1/T);
$R = ND^2/\square$, Reynolds number;
x = input vectors;
$X = N^3D^2 / (g^{4/3}\square^{1/3}) = F^{4/3}R^{1/3}$ = theoretical power per unit volume parameter;
W=weights;
$\square$ = kinematic viscosity of water (M^2/T);
δ = error;
η = learning rate.

References

[1] Albal, R.S., Shah, Y.T., Shumpe, A. and Carr, N.L, 1983, “Mass transfer in multiphase agitated reactor”, Chem. Eng. J., 27, 67-80.

[2] Bavarian, B.,1998, “Introduction to neural networks for intelligence control”, IEEE controls Syst. Mag.3-7.

[3] Fair, G.M., Gayer, J.C. and Okum, D.A., 1971, “Elements of water supply and waste water disposal”, John Wiley and Sons, New York.

[4] Gamal El-Din, Ahmed, Smith, Daniel W., and Gamal El-Din, Mohamed, 2004, “Application of artificial neural networks in wastewater treatment”, J. Environ. Eng. & Sci, 3(S1): S81-S95.

[5] Hagan, M.T., and Menhaj, M.B., 1994 “Training Feedforward Networks with the Marquardt Algorithm”, IEEE Trans on Neural Networks vol. 5.

[6] Irie, B. and Miyanki, S., 1988, “Capabilities of three layer perceptrons”, In IEEE second Int. Conf. on Neural networks, San Diego, Vol.1, pp. 641-648.

[7] Rao A.R.K., 1999, “Prediction of reaeration rates in square, stirred tanks, Journal of Environmental Engineering”, ASCE, Vol. 125, No. 3, 215-233.

[8] Rao A.R.K., B.V.Bharathi Laxmi and K. Subba Narasiah, 2004, “Simulation of Oxygen Transfer Rates in Circular aeration Tanks", Water Qual. Res. J. Canada, Vol.39, No. 3,237-244.

[9] UdayaSimha L., Shrma K.V.N.S. and Rao A.R.K., 1991, “Effect of geometrical parameters for overall Oxygen transfer coefficient”, Proc. Symp. On Environmental Hydraulics, university of Honking.