

Square Surface Aerator: Process Modeling and Parameter Optimization

A. R. K. Rao* and B. Kumar

Department of Civil Engineering, Indian Institute of Science, Bangalore 560012, India

ABSTRACT. The performance of an aeration system generally depends upon the geometric and dynamic parameters. The main purpose of doing experiment in the area of surface aeration system is to interpret the laboratory results for the field application that is to scale-up the results. This requires a geometrical similarity conditions. Finding geometrical optimal conditions of a surface aeration system through experiments involves physical constraints and classically parameters can be optimized by varying one variable at a time and keeping others as constants. In the real experimental process, it is not possible to vary all others geometric parameters simultaneously. In such a case, the model of the system is built through computer simulation, assuming that the model will result in adequate determination of the optimum conditions for the real system. In this paper, two approaches have been used to model the phenomena: i) Multiple regression and ii) Neural network. It has been found that neural network approach is showing better predictability compared to the multiple regression approach. In process of optimization, the pertinent dynamic parameter is divided into a finite number of segments over the entire range of observations. For each segment of the dynamic parameter, the neural network model is optimized for the geometrical parameters spanning over the entire range of observations. Thus each segment of the dynamic parameter has its set of optimal geometrical conditions. Results obtained are having less variation among them and they are very nearer to the experimental optimal conditions. Input parameter significance test of neural network model reveals that, in general the blade width of rotor is a major geometric parameter to enhance the aeration process.

Keywords: Dynamic similarity, geometric similarity, multiple regressions, neural network, prediction error surface, surface aerator

1. 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 remove the 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. The important functions of the surface aerators are to enhance oxygen transfer process and liquid phase mixing for ensuring the oxygen availability in all parts of the contactor and suspension of microorganisms. There are many types of aerators used in practice, such as cascade, spray nozzles, diffused or bubble aerators and surface aerators. Among the several types of aerators, surface aerators are more popular because of their better efficiency and ease in operation.

Usually the sizes of treatment plants are much larger than the reactors used in the manufacturing processes. A large number of impeller designs are used in surface aeration applications but most of the information (White et al., 1934; Nagata et al., 1956; Sano and Usui, 1985; Wu, 1995) is for standard configuration of impeller clearance and submergence. Further, the mass transfer coefficients have been reported over a nar-

row range of design and operating variables (Zlokarnik, 1979; Backhurst et al., 1988; Ognean, 1993). Patil et al. (2004) have optimized the conditions for the operation of surface aerators and put forward the optimized values of geometric parameters. But they have not extended their study for developing a mathematical correlation for prediction of mass transfer coefficient and power number, which are expected to be useful for the design engineers.

Aeration process is represented by three different groups of parameters, namely geometric, dynamic and physical. Out of these physical groups of parameter are invariant in the analysis, as they are more or less constant. The objective of doing studies on aeration process is to interpret the laboratory result into the field installation. This means scaling up of laboratory geometric dimensions for the field installation. It requires a geometrical similarity condition that is to say that the field installation should be built on a definite geometric ratio of the laboratory setup.

The physical and chemical processes taking place in the aeration tank are complex and closely coupled to the underlying transport processes, in particular the flow field. Therefore, a detailed understanding of the hydrodynamics of aeration tank (velocity field, turbulence, stress field etc.) is useful for optimum design. The geometry is the key to understanding mixing. In fact, the geometry is so important that processes can be considered "geometry specific". Optimality of the geometric parameters can be obtained by keeping the one para-

* Corresponding author: ark@civil.iisc.ernet.in

meter as a variable at a time and keeping the others constant. This approach has some limitations and it is difficult to keep the one parameter constant at every possible limits. However based on certain intuitions and guess, this approach generally gives a working result. It can also be said from the earlier studies (Nagata, 1975; Rao, 1999) that aeration process is uniquely described by the dynamic parameters on optimal geometric conditions.

Numerical process does not have such limitations. The idea of this paper is first to formulate a general model which describe the best approximation of the aeration process, then by keeping the dynamic parameter constant at one value, optimize the general model with all the geometric parameter varying within the experimental observation limit. The objective of this paper is to find an optimal geometric similarity condition computationally and compare the result to the experimental findings.

2. Surface Aeration Process

A typical surface aerator with six flat blades, used in this study, is shown in Figure 1.

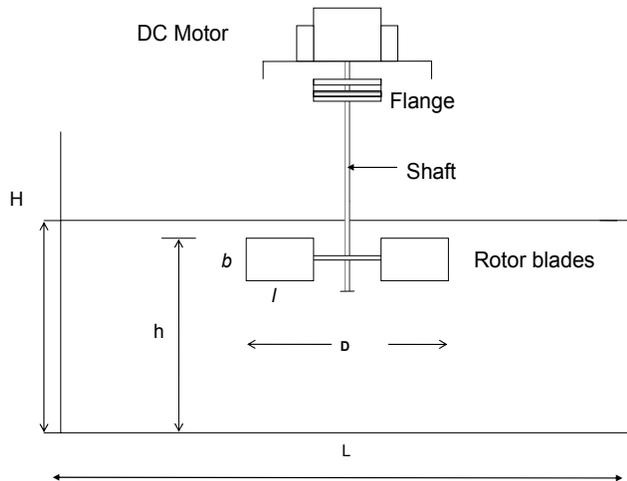


Figure 1. Schematic diagram of a surface aerator.

The main component of these surface aerators is an impeller or rotor, to which the blades are fitted. The rotor is rotated to create turbulence in the water body so that aeration takes place through the interface of atmospheric oxygen and the water surface.

The aeration process generally depends on three types of variables namely geometric, physical and dynamic variables and they are all explained as follows:

Geometric variables: Cross-sectional area of the tank (A); depth of water in the tank (H); diameter of the rotor (D); length of the blades (l); width of the blades (b); the distance between the top of the blades and the horizontal floor of the tank (h); and the number of blades (n).

Physical variables: density of air (ρ_a); density of water

(ρ_w); and kinematic viscosity of water (ν).

Dynamic variables: rotational speed of the rotor with blades (N).

The variables, which can influence the oxygen, transfer coefficient at 20°C (i.e. $K_L a_{20}$) for a given shape of an aeration tank are given by

$$K_L a_{20} = f(A, H, D, l, b, h, n, N, g, \rho_a, \rho_w, \nu) \quad (1)$$

Equation 1 may be expressed in terms of non-dimensional parameters as follows:

$$K = f(\sqrt{A/D}, H/D, l/D, b/D, h/D, n, \rho_a/\rho_w, R, F) \quad (2)$$

or it can be expressed as (Rao, 1999; Rao et. al., 2004):

$$K = f(\sqrt{A/D}, H/D, l/D, b/D, h/D, n, \rho_a/\rho_w, X) \quad (3)$$

where $K = K_L a_{20} / N$ is the non-dimensional oxygen transfer parameter and $X = F^{4/3} R^{1/3}$ is the parameter governing the theoretical power per unit volume. The first six non-dimensional parameters represent the "geometric similarity" of the system and the last parameter represents the "dynamic similarity".

As $A = L^2$, $\sqrt{A/D}$ can be written as L/D . The number of blades, n in the present experiments is constant. Also, the parameter ρ_a/ρ_w is considered as invariant. Thus, these two parameters are omitted in the analysis. Therefore, the functional relationship of Equation 3 can now be expressed as

$$K = f(L/D, H/D, l/D, b/D, h/D, X) \quad (4)$$

2.1. Problem Formulation

The problem formulation analyzed in this paper is at best described by the following steps:

- 1) Generalize model for the Equation 4 by different approach such as Multiple regression, Neural network etc.
- 2) Select the model which gives less prediction error variance from the model.
- 3) Find the most important or sensitive parameters of the model.
- 4) Now optimize the K , by changing the geometric parameters and keeping the X at certain values.
- 5) Repeat the process '4' to the entire experimental range of X .
- 6) Compare with the experimental results.

The schematic diagram on whole modeling and optimization process is shown in Figure 2.

3. Modeling Process

A commercial Microsoft Windows based Mathematics software Matalab® v.7 was used throughout the studies. The mathematical representation of the input/output function of a

simulation model will be represented (Kleijnen and Sargent, 2000) as:

$$Y = g(X) \tag{5}$$

The **X** vector for a simulation includes the input variables and **Y** is the output variables. General models are typically developed separately for each component of **Y**, that is, for each coordinate function of *g*. The major issues in modeling include: i) the choice of a functional form for *g*, ii) the selection of a set of **X** points at which to observe *y* (run the full model) to adjust the fit of *g* and iii) the assessment of the adequacy of the fitted model. To model any process, there are two basic approaches; first, extrapolation of model experiments based on the principles of similitude (soft modeling or empirical modeling) and, second, mathematical analysis of the complete (or controlling) mechanism. While the second of these has unlimited potential value, it also has serious limitations in practice. Often the relationships are too involved to permit rigorous definition or the resultant mathematical expressions are too complex for economical solution, even with computing equipment. This leaves us with soft modeling/ empirical modeling. There are various ways to do empirical modeling of a systems viz., multiple regression analysis, neural network, genetic algorithm etc. Present paper uses the two most common approaches to model the process such as multiple regression analysis and neural network.

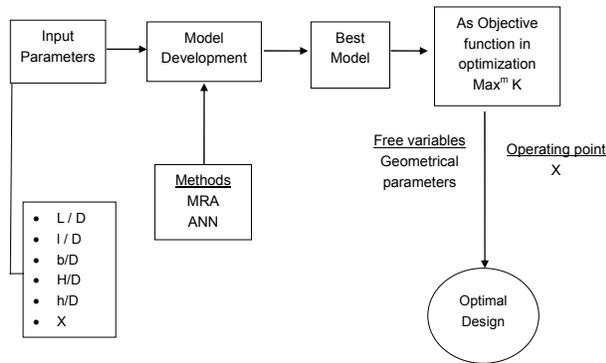


Figure 2. Flow diagram of the modeling process.

To judge the quality of the results obtained in multiple regression analysis and neural network, two indicators were used:

- The correlation coefficient *R* between observed and estimated values, or the determination coefficient R^2 ;
- Prediction error variance (PEV): PEV is a very useful way to investigate the predictive capability of the model. It gives a measure of the precision of a model's predictions. A low PEV (close to zero) means that good predictions of the model.

3.1. Data Source: Experimental Work

Experimental data obtained from our earlier studies (Rao,

1999) on square surface aerators were used in additions to the present the experimental data in order to provide a sufficient number of data for the modeling the process. Experimental optimal points for different parameters have been published in Rao (1999). Three sizes of square surface aeration tanks of c/s area (1, 0.5184 and 0.1684 m²) were tested under laboratory conditions. Determination of *K* has been made as follows:

According to two-film theory (McWhirter et al., 1995), the oxygen transfer coefficient at T °C, $K_L a_T$ may be expressed as follows:

$$K_L a_T = [\ln(C_s - C_0) - \ln(C_s - C_t)]/t \tag{6}$$

where, *ln* represents natural logarithm and the concentrations C_s , C_0 and C_t are dissolved oxygen (DO) concentrations in parts per million (ppm), C_s = the saturation DO concentration at time tending to very large values, C_0 is at $t = 0$ and C_t is at time $t = t$. The value of $K_L a_T$ can be obtained as slope of the linear plot between $\ln(C_s - C_t)$ and time *t*. The value of $K_L a_T$ can be corrected for a temperature other than the standard temperature of 20°C as $K_L a_{20}$, using the Vant- Hoff Arrhenins equation (WEF and ASCE, 1988):

$$K_L a_T = K_L a_{20} \theta^{(T-20)} \tag{7}$$

where θ is the temperature coefficient 1.024 for tap water.

At first, water in the tank was deoxygenated by adding the required amount of cobaltous chloride (CoCl₂) and sodium sulphite (Na₂SO₃) (Metcalf & Eddy Inc., 2004) and thoroughly mixing water. The deoxygenated water was re-aerated by rotating the rotor at desired speeds and maintaining the variables as per data presented in Table 1. When the DO concentration began to rise, readings were taken at regular intervals till DO increased up to about 80% of the DO saturation value. Lutron Dissolved Oxygen meter was used to measure the DO concentration in water. The DO meter was calibrated with the modified Winkler's method (AWWA, 1985).

Table 1. Descriptive Statistics

Parameters	Min ^m	Max ^m	Mean	Std. Dev.
L/D	1.61	6	3.21	0.752
I/D	0.048	0.38	0.275	0.0449
b/D	0.038	0.304	0.22	0.036
H/D	0.516	2.53	1.08	0.311
h/D	0.584	2.38	1.02	0.293
X	0.003137	7.57	0.819	1.16

The known values of DO measurements in terms of C_t at regular intervals of time *t* (including the known value of C_0 at $t = 0$) a line is fitted, by linear regression analysis of Equation 6, between the logarithm of $(C_s - C_t)$ and *t*, by assuming different but appropriate values of C_s such that the regression that gives the minimum "standard error of estimate" is taken and thus the values of $K_L a_T$ and C_s were obtained simultaneously. The values $K_L a_{20}$ are computed using Equation 7 with θ

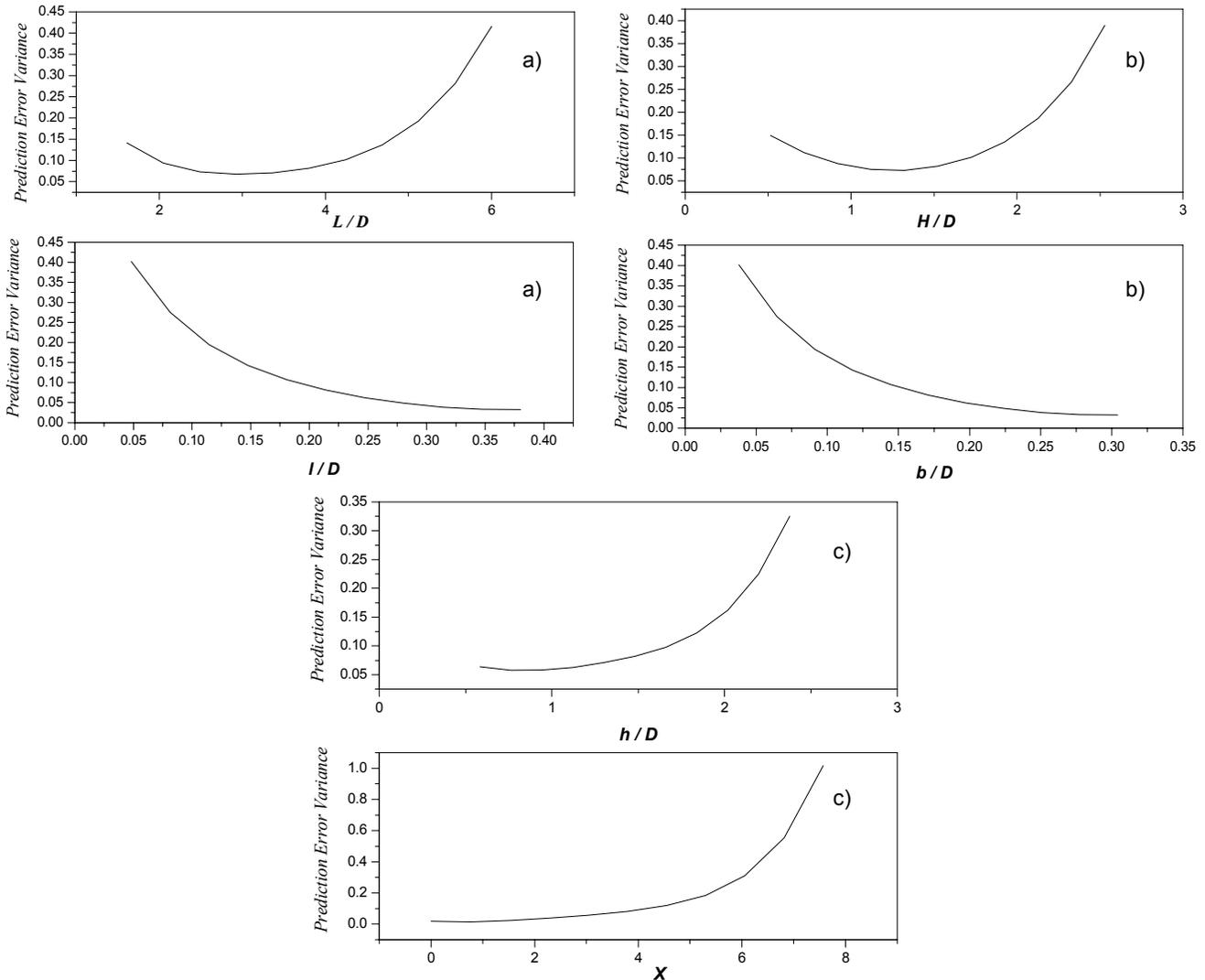


Figure 3. PEVs of multiple regression models a), b) and c).

= 1.024 as per the standards for clean water (WEF and ASCE, 1988). Thus the values of $K_{La_{20}}$ were determined for different rotational rotor speeds N in all of the geometrically similar tanks. Thus K has been determined for different values of parameters listed in Equation 4.

3.2. Multiple Regression Analysis: MRA

Multiple regression analysis (MRA) is probably one of the most commonly used methods for multivariate data analysis. It can be used to test hypotheses regarding the relationship of one or more predictors to a dependent variable. It also lends itself easily to predicting or forecasting the values of the dependent variable.

The concept of regression analysis lies in the idea to predict the scores of one dependent variable Y from the scores of one or several independent variables X_1, X_2, \dots, X_m in an optimal way. Standard multiple regression can only accurately es-

timate the relationship between dependent and independent variables if the relationships are linear in nature. If the relationship between independent variables and the dependent variable is not linear, the results of the regression analysis will underestimate the true relationship. The descriptive statistics of the parameters are given in the Table1. Total experimental runs are 400. The regression equation of the process is given as:

$$K = -0.7538 \frac{b}{D} + 0.0016 \frac{H}{D} + 0.0006 \frac{L}{D} + 0.5903 \frac{l}{D} + 0.0017X + 0.255 \frac{b}{D} \frac{L}{D} - 0.008 \frac{L}{D} X + 0.0118 \frac{l}{D} X + 0.004 \frac{H}{D} \frac{L}{D} - 0.0087 \frac{H}{D} X - 0.0014 \frac{b}{D} \frac{l}{D} + 0.0093 \frac{b}{D} X - 0.004 \left(\frac{L}{D}\right)^2 - 0.1987 \frac{l}{D} \frac{L}{D} - 0.003X^2 \quad (8)$$

Equation 8 gives the correlation coefficient $R^2 = 0.9262$. Although the correlation coefficient of the model is reasonab-

ly good, but the model prediction error is very high compared to the next model as shown in the following Figures 3a to 3c.

3.3. Artificial Neural Network: ANN

Artificial neural networks (ANNs) provide a useful and effective tool for modeling the complex and poorly understood processes that occur in nature, as they are able to extract functional relationships between model inputs and outputs from data without requiring explicit consideration of the actual data generating process (Hayken, 1994). ANNs are non-linear mapping structures based on the function of the human brain. They are considered universal and highly flexible approximators for any data.

In ANN, the computational or processing elements are called neurons. Like a natural neuron, they have many inputs but only a single output, which can stimulate other neurons in the network. Neurons from one layer are connected to all neurons in the adjacent layer(s). The number of input and output units depends on the representations of the input and the output objects, respectively. Training the network consists of using a training data set to adjust the connection weights in order to minimize the error between observed and predicted values. The units in neighboring layers are fully interconnected with links corresponding to synapses. The strengths of connections between two units are called "weights". In each hidden layer and output layer the processing unit sums its input from the previous layer and then applies the "tansigmoid" function to compute its output to the following layer according to the following equations:

$$y_j = \sum W_{ij} X_i \tag{9}$$

$$f(y_j) = \frac{e^{y_j} - e^{-y_j}}{e^{y_j} + e^{-y_j}}$$

where W_{ij} is the weight of the connection between unit j in the current layer to unit i in the previous layer and X_i is the output value from the previous layer. $f(Y_j)$ is conducted to the following layer as an output value. ANN learns an approximate non-linear relationship by a procedure called "training", which involves varying weight values. Training is defined as a search process for the optimized set of weight values (Weight plot is given in the appendix, Figure 9) which can minimize the squared error between the estimation and experimental data of units in the output layer. A back-propagation method with the Levenberg-Marquadt (LM) algorithm has been widely applied for training ANN. The LM algorithm is widely applied to many different domains and is faster and produces better results than other training methods (Hagan and Menhaj, 1994; Tan and van Cauwenberghe, 1999). To update weights and biases, the LM algorithm uses an approximation to the Hessian matrix. Learning rate has been set to 0.01 and learning cycle to 5000.

The optimal number of neurons in the hidden layer of the network has been determined by running the program code

(MATLAB® Platform V. 7) several times with different networks and chooses the number that produced the greatest network performance. The best architecture of the neural model has been shown in Figure 4 and results of the training and testing for the best model in Figure 5.

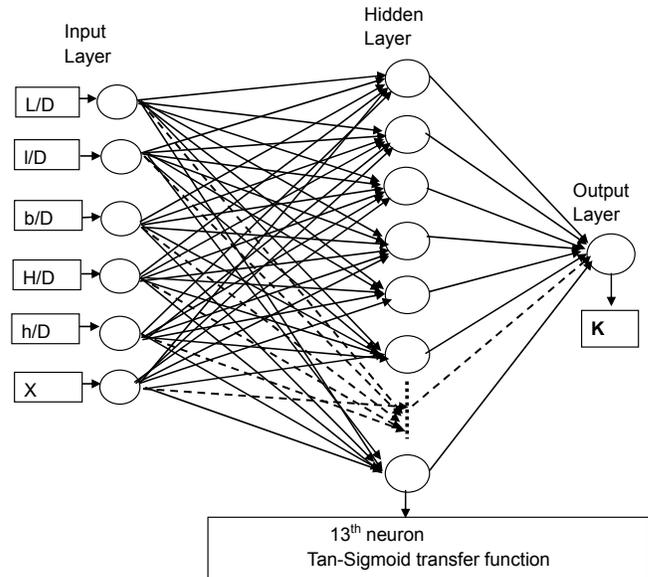


Figure 4. Best configurations of artificial neural network.

It can be clearly seen from Figure 5 that the linear coefficient of correlation is very high between observed experimental data and values predicted through neural nets and it is 0.964 in training and 0.958 in testing for square surface aerators. This shows the learning and generalization performance of the network is good. Prediction Error Variance (PEV) is a very useful way to investigate the predictive capability of model. It gives a measure of the precision of a model's predictions. Figures 6a to 6c show how well the present model predicts over the design region, when compared to the PEV of multiple regression model of Figures 3a to 3c.

3.4. Discussion between MRA and ANN Model

In order to assess the usefulness and the main limitations of ANNs, it is interesting to have an overview of alternative tools often applied in multivariate calibration. The most popular method is MRA. The attraction of MRA lies in the ease of model interpretation since the estimated parameters relate the parameter of interest to a set of original variables. MRA may be viewed as a special case ANN model that uses linear transfer functions and no hidden layers. If the linear model performs as well as a more complex ANN, then using the non-linear ANN may not be justified; thus, linear models are useful as a basis for comparison. In our analysis, both the methods show a high R^2 (0.93 for MRA and 0.96 for ANN), but the predictability of the MRA method is very poor as shown by its high value of prediction error variance. Their ability to learn and derive X - Y relationships from the presentation of a set of training samples, ANN avoids the time-consuming and

possibly expensive task of hard model identification. In addition, the fundamental principle of distributing information among several weights and nodes renders the ANN model robust with respect to random noise in the input data.

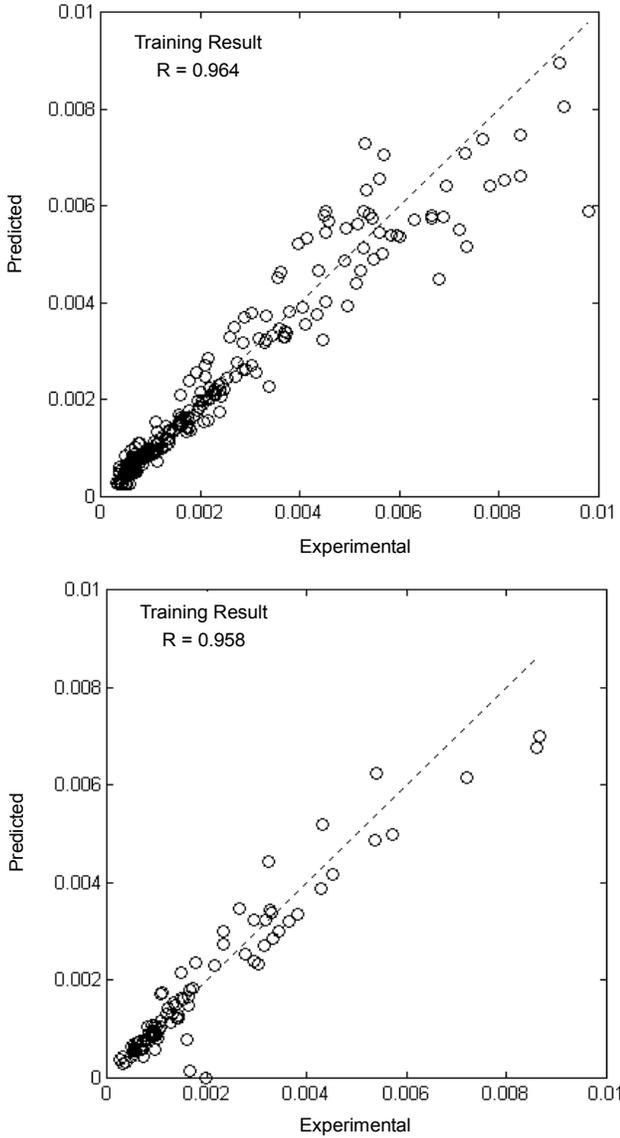


Figure 5. Neural network result.

If one is not careful, however, a drawback of the flexibility of ANNs is their tendency to overfit calibration data and the resulting lack of generalization ability, that is, the capability of a model to produce a valid estimate of the correct output when a new input is presented to the NN. Also, the flexibility of ANNs can lead to unreliable results in situations of extrapolation.

3.5. Input Significance Testing

Geometric parameters do have effects on surface aeration system; it is a well known fact. The purpose of doing in-

put significant analysis is to find out that which parameter is having more influence on mass transfer process. As, it is necessary to know the impact of the input variables, some authors have proposed methods allowing determination of the impact of the input variables on outputs such as connection weight, Garson's algorithm etc. In the present paper, four such methods have been applied to know the importance of the input variables on output.

Connection weights: Calculates the product of the raw input-hidden and hidden-output connection weights between each input neuron and output neuron and sums the products across all hidden neurons (Olden and Jackson, 2002). In the neural network, the connection weights between neurons are the links between the inputs and the outputs, and therefore are the links between the problem and the solution. The relative contributions of the independent variables to the predictive output of the neural network depend primarily on the magnitude and direction of the connection weights. Input variables with larger connection weights represent greater intensities of signal transfer, and therefore are more important in the prediction process compared to variables with smaller weights.

Garson's algorithm: Garson (1991) proposed a method for partitioning the neural network connection weights in order to determine the relative importance of each input variable in the network. The methodology for this algorithm is as follows:

(a) For each hidden neuron h , divide the absolute value of the input-hidden layer connection weight by the sum of the absolute value of the input-hidden layer connection weight of all input neurons, i.e.

For $h = 1$ to nh , For $i = 1$ to ni ,

$$O_{ih} = \frac{|W_{ih}|}{\sum_{i=1}^{ni} |W_{ih}|} \quad (10)$$

(b) For each input neuron i , divide the sum of the Q_{ih} for each hidden neuron by the sum for each hidden neuron of the sum for each input neuron of Q_{ih} , multiply by 100. The relative importance of all output weights attributable to the given input variable is then obtained.

For $i: 1$ to ni

$$RI (\%)_i = \frac{\sum_{h=1}^{nh} Q_{ih}}{\sum_{n=1}^{nh} \sum_{i=1}^{ni} Q_{ih}} \times 100 \quad (11)$$

Sensitivity analysis: A number of investigators have used sensitivity analysis to determine the spectrum of input variable contributions in neural networks. For example, the Senso-nets approach includes an additional weight in the network for each input variable representing the variable's sensitivity (Schleiter et al., 1999). Scardi and Harding (1999) added a

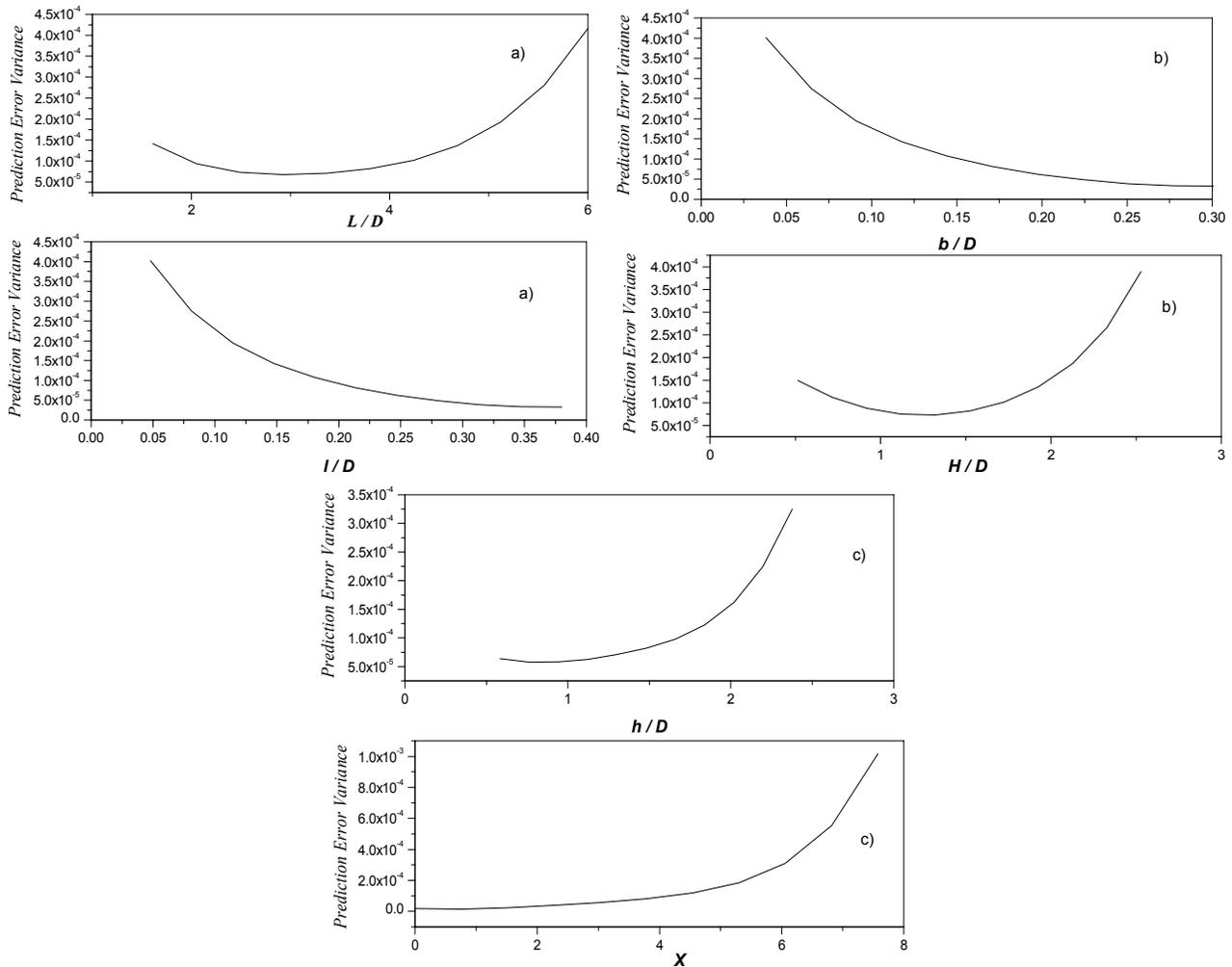


Figure 6. PEVs of ANN models a), b) and c).

white noise to each input variable and examined the resulting changes in the mean square error of the output. Traditional sensitivity analysis involves varying each input variable across its entire range while holding all other input variables constant; so that the individual contributions of each variable are assessed. This approach is somewhat cumbersome, however, because there may be an overwhelming number of variable combinations to examine (Lek et al., 1996). As a result, it is common first to calculate a series of summary measures for each of the input variables (e.g. minimum, maximum, quartiles, percentiles), and then vary each input variable from its minimum to maximum value, in turn, while all other variables are held constant at each of these measures (e.g. Ozesmi and Ozesmi, 1999). Relationships between each input variable and the response can be examined for each summary measure, or the calculated response can be averaged across the summary measures.

Forward stepwise addition: This method consists of adding or rejecting step by step one input variable and noting the effect on the output result. Based on the changes in MSE, the

input variables can be ranked according to their importance in several different ways depending on different arguments. For instance the largest changes in MSE due to input deletions can allow these inputs to be classified by order of significance. In another approach the largest decrease in MSE can identify the most important variables, i.e. the most relevant to the construction of a network with a small MSE (Gevrey et al., 2003). First, ten models were generated, each using only one of the available variables input. Then, nine models were generated, combining the variable that resulted in the smallest error (for a single input variable) with each of the remaining variables, this procedure was repeated using models with three input variables, and more till all the variables were added (Maier et al., 1998). The order of integration of the input variables in the network is the order of the importance of their contributions.

Results of the above approaches have been shown in Figure 7 and calculations of all approaches are tabulated in Table 2. The mass transfer phenomena are strongly affected by the geometrical parameters; it is ascertained from Figure 7. Ex-

cept one approach, all approaches have shown the most significant parameter is b/D .

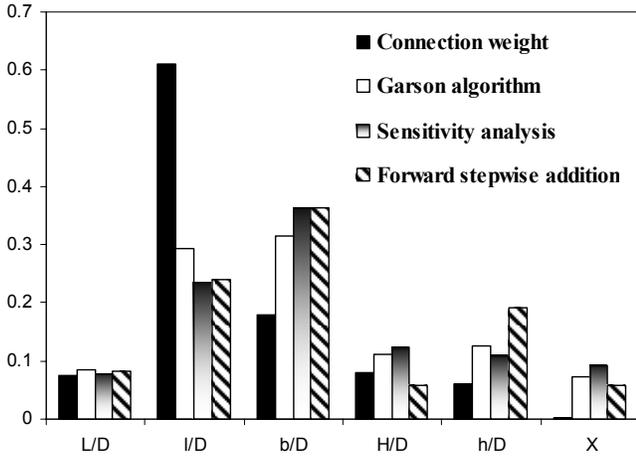


Figure 7. Importance of the input variables.

Table 2. Input Significance Test

Variables	1	2	3	4
L/D	0.074144	0.083591	0.076639	0.081647
I/D	0.611251	0.291974	0.234083	0.24071
b/D	0.178425	0.314447	0.363145	0.36329
H/D	0.081081	0.112561	0.124286	0.058283
h/D	0.061244	0.12515	0.109836	0.1922
X	0.003355	0.072278	0.091976	0.058283

*1-Connection weight, 2-Garson algorithm, 3-Sensitivity analysis and 4-Forward stepwise addition.

4. Optimal Conditions of Geometrical Parameters

In general, the optimization problems can be viewed in terms of minimization (or maximization) of the objective function, $F(X)$, under the following inequality and/or equality constraints:

$$\left. \begin{aligned} G_i(X) &\geq 0 \\ H_i(X) &\geq 0 \end{aligned} \right\} \quad (12)$$

where $G_i(X)$ is the inequality constraint and $H_i(X)$ is the equality constraint. In the case of a fully-trained ANN, $F(X)$ corresponds to the predicted value of response variable adopted as the unit in the output layer and X is a set of causal factors used as the units in the input layer.

Optimization problem has been dealt as a Point by Point problem. In a point-by-point problem, a single optimization run can determine optimal control parameter values at a single operating point. To optimize control parameters over a set of operating points, an optimization can be run for each point. Calibration generation and optimization toolboxes provided in the MATLAB® have been used in the analysis. The steps in-

involved in this procedure are as follows:

(1) At particular value of X , Maximize $K = K(L/D, b/D, H/D, h/D, I/D, X)$

(2) Subjected to constraints:

Min value $< L/D < Max$ value

Min value $< b/D < Max$ value

Min value $< H/D < Max$ value

Min value $< h/D < Max$ value

Min value $< I/D < Max$ value

At first, a deterministic optimization of the best general model has been done by keeping the X at lower range. After one set of analysis, the value of X has been altered to another one. One set of analysis at $X = 4.1$ is shown in Figure 8. The entire range of X has been divided into 10 parts. “foptcon” function provided by the MATLAB® environment has been used to optimize the neural model. “foptcon” is a single-objective optimization subject to the constraints - “foptcon” attempts to find a constrained minimum of a scalar function of several variables starting at an initial estimate. Table 3 shows the optimal value of geometrical parameters at every X . The variation in optimal points is about 1%, when compared to extreme lower and upper ranges.

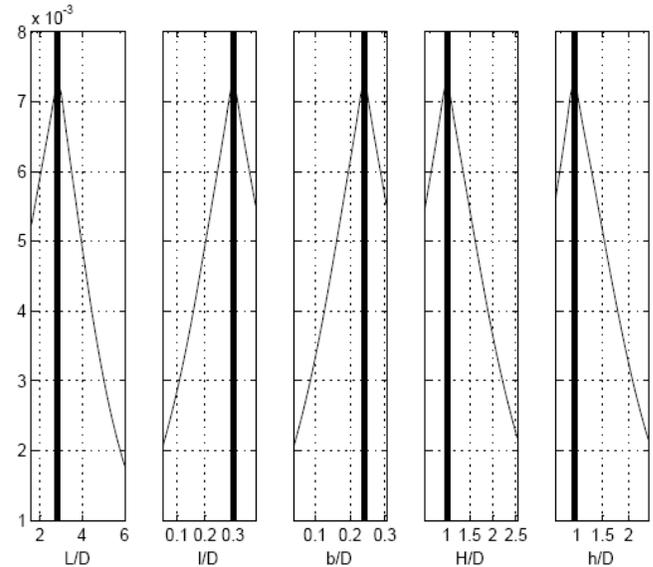


Figure 8. Optimal conditions of geometrical parameters.

5. Conclusions

1) Neural network model gives a best prediction of the aeration phenomena comprising of geometrical and dynamic variables.

2) Blade width is the most significant geometrical parameter affecting the re-aeration process.

3) Experimental optimization has physical constraints

Table 3. Optimal Points

Parameter	Optimal Points at X										Optimal Points (Experimental) of Rao (1999)
	0.0032	0.822	1.64	2.46	3.28	4.1	5	5.9	6.6	7.5	
L/D	2.85	2.879	2.879	2.879	2.879	2.868	2.88	2.88	2.88	2.88	2.88
l/D	0.33	0.32	0.305	0.31	0.305	0.305	0.305	0.30	0.30	0.30	0.3
b/D	0.28	0.27	0.25	0.26	0.25	0.245	0.245	0.24	0.24	0.24	0.24
H/D	1.1	1.005	1.004	1.001	0.999	1.006	0.997	0.999	0.999	0.999	1
h/D	0.95	0.97	0.94	0.945	0.945	0.955	0.9465	0.94	0.94	0.94	0.94

while optimizing the geometric parameters of a surface aeration system, whereas numerically the parameters can be varied without such limitations.

4) The results obtained by the numerical optimization are very nearer to the optimal points obtained by the experiments.

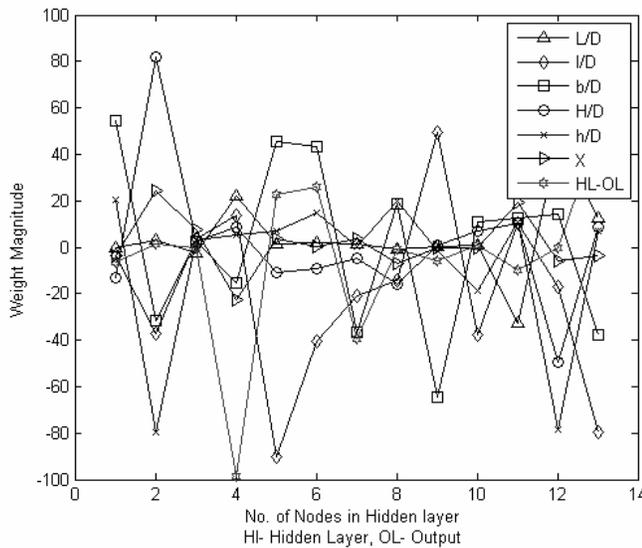


Figure 9. Weight plot.

References

Backhurst, J.R., Harker, J.H. and Kaul, S.N. (1988). The performance of pilot and full-scale vertical shaft aerators, *Water Res.*, 22, 1239.

Fair, G.M., Gayer, J.C. and Okum, D.A. (1971). *Elements of Water Supply and Waste Water Disposal*, John Wiley & Sons, New York.

Garson, G.D. (1991). Interpreting neural-network connection weights. *Artif. Intell. Expert*, 6(4), 47-51.

Gevrey, M., Dimopoulos, I. and Lek, S. (2003). Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecol. Model.*, 160, 249-264.

Hagan, M.T. and Menhaj, M.B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.*, 5, 989-993.

Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*, Macmillan, New York.

Kleijnen, J.P.C. and Sargent, R.G. (2000). A methodology for fitting and validating metamodels in simulation. *Eur. J. Oper. Res.*, 120, 14-29.

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J. and Aulagnier, S. (1996). Application of neural networks to modeling nonlinear relationships in ecology. *Ecol. Model.*, 90, 39-52.

McWhirter, J.R., Chern, J. and Hutter, J.C. (1995). Oxygen mass transfer fundamentals of surface aerators. *Ind. Eng. Chem. Res.*, 34, 2644-2654.

MATLAB® Version 7, The mathworks.com.

Maier, H.R., Dandy, G.C. and Burch, M.D. (1998). Use of artificial neural networks for modelling Cyanobacteria anabaena spp. in the river Murray, South Australia. *Ecol. Model.*, 105, 257-272.

Metcalf & Eddy Inc. (2004). *Waste Water Engineering: Treatment Disposal and Reuse*, Tata McGraw-Hill, New Delhi.

Nagata, S., Yokoyama, T. and Maeda, H. (1956). Memoires Fac. Eng., in G.B. Tatterson (1991), *Fluid Mixing and Gas Dispersion in Agitated Tanks*, McGraw-Hill, USA.

Nagata, S. (1975). *Mixing Principles and Applications*, John Wiley & sons, New York.

Olden, J.D. and Jackson, D.A. (2002). Illuminating the “blackbox”: understanding variable contributions in artificial neural networks. *Ecol. Model.*, 154, 135-150.

Ognean, T. (1993). Dimensionless criteria for estimating oxygen transfer in aeration system. *Biotechnol. Bioeng.*, 41, 1014.

Ozesmi, S.L. and Ozesmi, U. (1999). An artificial neural network approach to spatial habitat modeling with inter specific interaction. *Ecol. Model.*, 116, 15-31.

Patil, S.S., Deshmukh, N.A. and Joshi, J.B. (2004). Mass transfer characteristics of surface aerators and gas-inducing impellers. *Ind. Eng. Chem. Res.*, 43, 2765.

Rao, A.R.K (1999). Prediction of reaeration rates in square, stirred tanks. *J. Environ. Eng., ASCE*, 125(3), 215-233.

Rao, A.R.K., Bharathi Laxmi, B.V. and Subba Narasiah, K. (2004). Simulation of oxygen transfer rates in circular aeration tanks. *Water Qual. Res. J. Can.*, 39(3), 237-244.

Sano, Y. and Usui, H. (1985). Interrelations among mixing time, power number and discharge flow rate number in baffled mixing vessel. *J. Chem. Eng. Japan*, 18, 47.

Scardi, M. and Harding, L.W. (1999). Developing an empirical model of phytoplankton primary production: A neural network case study. *Ecol. Model.*, 120, 213-223.

Schleiter, I.M., Borchardt, D., Wagner, R., Dapper, T., Schmidt, K.D., Schmidt, H.H. and Werner, H. (1999). Modeling water quality, bioindication and population dynamics in lotic ecosystems using neural networks. *Ecol. Model.*, 120, 271-286.

AWWA (American Public Health Association) (1985). Standard methods for the examination of water and wastewater, 16th Edition, WCPA, Washington, DC.

Tan, Y. and Van Cauwenbergh, A. (1999). Neural-network-based d-step-ahead predictors for nonlinear systems with time delay. *Eng. Appl. Artif. Intell.*, 12, 21-25.

WEF and ASCE Manual of practice for water pollution control (1988). *Aeration a Waste Water Treatment Process*, Water Environment Federation, Alexandria, VA, and ASCE, New York.

White, A.M. and Brenner, E. (1934). Studies in agitation. *Trans. Am. Inst. Chem. Eng.*, 30, 585.

Wu, H. (1995). An issue on application of a disk turbine for gas-

liquid mass transfer. *Chem. Eng. Sci.*, 50(17), 2801.

Zlokarnik, M. (1979). Scale-up of surface aerators for waste water treatment. *Adv. Biochem. Eng.*, 11. 157.