Journal of Urban and Environmental Engineering, v.2, n.1, p.1-7

ISSN 1982-3932
doi: 10.4090/juee.2008.v2n1.001007

Journal of Urban and Environmental Engineering

www.journal-uee.org

PREDICTION OF BOD AND COD OF A REFINERY WASTEWATER USING MULTILAYER ARTIFICIAL NEURAL NETWORKS

Eldon R. Rene[1*] and M. B. Saidutta[2]

[1]*Department of Chemical Engineering, University of La Coruna, Spain*
[2]*Department of Chemical Engineering, National Institute of Technology Karnataka, Surathkal, India*

Received 23 October 2007; received in revised form 19 March 2008; accepted 02 April 2008

Abstract: In the recent past, artificial neural networks (ANNs) have shown the ability to learn and capture non-linear static or dynamic behaviour among variables based on the given set of data. Since the knowledge of internal procedure is not necessary, the modelling can take place with minimum previous knowledge about the process through proper training of the network. In the present study, 12 ANN based models were proposed to predict the Biochemical Oxygen Demand (BOD_5) and Chemical Oxygen Demand (COD) concentrations of wastewater generated from the effluent treatment plant of a petrochemical industry. By employing the standard back error propagation (BEP) algorithm, the network was trained with 103 data points for water quality indices such as Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Phenol concentration, Ammoniacal Nitrogen (AMN), Total Organic Carbon (TOC) and Kjeldahl's Nitrogen (KJN) to predict BOD and COD. After appropriate training, the network was tested with a separate test data and the best model was chosen based on the sum square error (training) and percentage average relative error (% ARE for testing). The results from this study reveal that ANNs can be accurate and efficacious in predicting unknown concentrations of water quality parameters through its versatile training process.

Keywords: Artificial neural networks; COD; BOD; sum square error; percentage average relative error; predictions

[*] Correspondence to: Eldon R. Rene. E-mail: eldonrene@yahoo.com

INTRODUCTION

Industrial activities consume a huge amount of natural water, utilizable resources and energy thereby discharging enormous wastewater to the natural environment. It is therefore necessary to analyse any industrial wastewater to determine its reuse potential and the degree of treatment required prior to its ultimate disposal or to device suitable measures for the recovery of useful products. It is of great importance in water quality control that the amount of organic matter present in the system be known and that the quantity of oxygen required for its stabilisation be determined. Over the years, different physico-chemical tests have been developed to determine the organic and inorganic content of wastewater (Metcalf & Eddy, 1995). In general, these tests may be divided into those used to measure gross concentrations of organic matter greater than 1 mg/L and those used to measure trace concentrations in the range of 10^{-6} to 10^{-3} gm/L.

Laboratory methods commonly used today to measure the gross amount of organic matter (greater than 1 mg/L) in wastewater includes the following: (a) Biochemical Oxygen Demand (BOD_5), (b) Chemical Oxygen Demand (COD) and (c) Total Organic Carbon (TOC). These three parameters are used in wastewater treatment operations to estimate the influent and effluent characteristics and treatment efficiency. The use of TOC as an analytical parameter has become more common in recent years especially for the treatment of industrial wastewater. Partly, this is due to the fact that the TOC determinations can be carried out in triplicate within minutes compared with the five days required for the BOD_5 test (Sawyer *et al.*, 1994). Apart from these, the easily measurable parameter for any industrial wastewater includes indices like Total Suspended Solids (TSS), Total Dissolved Solids (TDS), Phenol concentration, Ammoniacal Nitrogen (AMN) and Kjeldahl's Nitrogen (KJN) (Metcalf & Eddy, 1995).

A review of the existing literature in this field reveals that correlation among these parameters seldom exists. It could be difficult to understand the dynamics of relationship between these parameters because they primarily depend on the process of the target industry, raw material/by-product composition, composition of chemicals discharged in wastewater and thus their non-linear relationship makes universal generalization difficult.

The main objective of this paper is to predict the BOD and COD concentrations of a refinery wastewater using different combinations of easily measurable water quality indices like TOC, TSS, TDS, Phenol, AMN and KJN using back error propagation (BEP) neural network. The best network architecture was determined by selecting the appropriate network topology.

ARTIFICIAL NEURAL NETWORKS

The three-layer back propagation network has been proved to be universal function approximations in the field of environmental prediction (Poggio & Girosi, 1990). Neural networks has been applied to solve and predict problems related to the following; biodegradation kinetics of organic compounds (Shuurmann & Muller, 1994), estimating optimum alum doses in water treatment (Maier *et al.*, 2004) and long term tidal waves (Lee, 2004).

The ANN theory

Neural networks are powerful data driven modelling tools that has the ability to capture and represent complex input/output relationships. The development of neural computational techniques emerged from the desire to develop an artificial system that could perform multiple, complex and intelligent tasks similar to those performed by the human brain. ANNs consists of a system of simple interconnected processing element called neurons.

This gives the ability to model any non-linear process through a set of unidirectional weighted connections (Haykin, 1999). The neuron accepts input from single or multiple sources and produces output by a simple calculating process guarded by a non-linear transfer function. A three-layered network (Bandyopadhyay & Chattopadhyay, 2007) with an input layer, hidden layer and output layer is shown in **Fig. 1**.

The input layer consists of a set of neurons, each representing an input parameter and propagates the raw information to the neuron in the hidden layer, which in turn transmits them to the neurons in the output layer. Each layer consists of several neurons and the layers are connected by the connection weights (*W*). The most commonly used transfer function is the sigmoid function as described by:

$$f(x) = \frac{1}{1+e^{-x}} \quad (1)$$

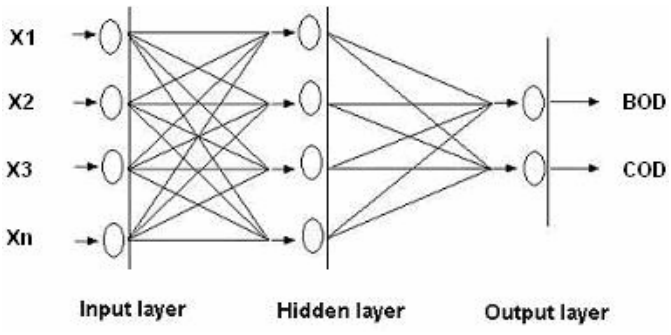


Fig. 1 Schematic of a three layer neural network.

This produces output in the range of 0–1 and introduces non-linearity into the network, which gives the power to capture nonlinear relationships. The back propagation network is the most prevalent supervised ANN learning model (Rummelhart *et al.*, 1986). It uses the gradient descent algorithm to correct the weights between interconnected neurons (Maier & Dandy, 1998).

During the learning process of the network, the algorithm computes the error between the predicted and specified target values at the output layer. The error function at the output layer can be defined by:

$$E = \frac{1}{2}\sum\left(O_d - O_P\right)^2 \qquad (2)$$

where, E is the error function, O_d is the desired output and O_p is the output predicted by the network.

Important Network Parameters

A good network architecture requires selecting the most dependable values of network parameters like: number of hidden layers, the number of neurons in the hidden layer N_H, the activation function $f(x)$, the learning rate of the network η, epoch size ε, momentum term α and training cycles T_C. The best values for parameters: η, α, N_H, and T_C are normally estimated by a trial and error approach. The learning rate η and momentum α can play an important role in the convergence of the network. The η value of a network affects the size of steps taken in weight space (Maier & Dandy, 1998). If η is too small, the algorithm would take more time to converge. The momentum term α accelerates the convergence of the error during the learning process by adding a fraction to the precious weight update. The values of η and α varies between 0–1 and is normally estimated by trial and error (Hamed *et al.*, 2004).

MATERIALS AND METHODS

Data handling procedure

The various wastewater parameters such as TSS, BOD, COD, TOC, phenol concentration, AMN, KJN and TDS were obtained from the quality control laboratory of a refinery located in Mangalore, India. Water samples collected from the effluent treatment plant after tertiary treatment were analyzed for the above mentioned parameters, which were later divided into training set (103) and test set (40). The ranges of various values of different parameters used for training and testing are shown in **Table 1**.

Software's used

Neural network based predictions were simulated using the software NNMODEL. Their performance was evaluated by the Sum square error (SSE) values for training obtained directly from the software, while the test data was evaluated using percentage average relative error, % ARE. Low SSE and low % ARE values theoretically mean that the predictions are precise and accurate.

Table 1. Range of water quality parameters used for training and testing

Sl. No.	Parameters (mg/L)	Training data		Testing data	
1	BOD	2–34	13.52	6.1–34	15.61
2	COD	12–160	61.64	38–114	72.076
3	TOC	3.1–18.5	8.21	4–18.5	9.67
4	TSS	4–71	18.60	6–41	18.13
5	TDS	343–1851	858.62	480–1720	973.73
6	AMN	1.4–92	19.04	9.5–94	31.90
7	KJN	1.8–93.4	20.83	10.3–96.8	34.48
8	Phenol	0.08–0.8	0.29	0.1–0.8	0.31

The percentage Average Relative Error (% ARE) was estimated from this relation,

$$\%ARE = \frac{1}{N}\sum\frac{\left|A_{Expt} - A_{Pred}\right|}{A_{Expt}} \times 100 \qquad (3)$$

ANN Based Models – inputs and outputs

A total of 12 ANN based models were evaluated in this study for predicting the BOD and COD of refinery wastewater. These models are shown in **Table 2**.

RESULTS

Prediction of BOD

The training of these models were started with the default values of NN model with a training count of 1000 and 4 hidden neurons in the hidden layer. From the next trail, the optimum training count for the network was decided. This was done by trial and error by checking the SSE and the % ARE after each cycle of training. The optimum training count was the one which gave a minimum SSE and lower % ARE for the test data. After deciding the maximum training count for these models the number of hidden neurons in the hidden layer were varied by small increments by maintaining constant training count until the desired SSE and % ARE for the test data was obtained. The training was done for these models by varying the learning rates of the network (0.35 to 0.75) and it was observed that there was no significant change in the SSE values after training. However, by varying the training count and the number of neurons in the hidden layer, the performance of the network greatly improved. The variation of SSE with different training count and hidden layers for model A3 is shown in **Table 3**. From these values it was observed that the SSE tends to cease after a particular time of training and almost remains constant throughout the training period.

Table 2. Various models developed using neural networks and their best SSE values

Model No.	Input Parameters	Output	Best SSE value
A1	TOC, phenol, TSS, TDS.	BOD	0.003 822
A2	TOC, phenol, TSS, TDS.	COD	0.006 053
A3	TOC, Phenol, TSS, AMN	BOD	0.003 403
A4	TOC, Phenol, TSS, AMN	COD	0.005 531
A5	TOC, Phenol, TSS, TDS, KJN	BOD and COD	0.003 585
A6	TOC, Phenol, TSS	BOD	0.003 725
A7	TOC, Phenol,TSS	COD	0.006 914
A8	TOC, Phenol, TSS, TDS,	COD and BOD	0.004 547
A9	TOC, Phenol, TDS	BOD	0.004 705
A10	TOC, Phenol,TDS	COD	0.007 055
A11	TOC	BOD	0.005 317
A12	TOC	COD	0.007 651

Prediction of COD

The training was initially carried out with the default values of the software NNModel. Later, the optimum training count for the network was determined. The same procedure that was applied for BOD was followed, thereby varying the number of hidden neurons in the hidden layer in small increments and by maintaining constant training count till the desired SSE and % ARE for the test data was obtained. It was observed that the SSE tends to slow down without showing any decrement in its value and then tends to increase to a certain extent before again decreasing and then remaining constant throughout the remaining period of training. This kind of behavior was noticed in model A7 at a training count of 6000 and 9 hidden neurons in the hidden layer which produced a SSE of 0.008 413 which was quite high compared to the other training cycles.

These models were also trained with different learning rates (0.5 to 0.75), but the network showed no positive improvement in reducing the SSE and the % ARE. Therefore, all these models were trained with the default values of NNModel for learning rates. The variation of SSE with the training count and hidden neurons in the hidden layer for the best model developed to predict COD is shown from **Table 4**.

Prediction of BOD and COD in a combined model

Two models were developed to predict BOD and COD simultaneously. The variation of SSE with different training count and different hidden layers for model A8 is shown in **Table 5**.

Table 3. Variation of SSE with different training count and hidden neurons for Model A3

Training count	Hidden neurons	Sum square error
default	4	0.004 491
5000	4	0.003 947
5000	5	0.003 983
5000	6	0.003 942
5000	**7**	**0.003 403**
7500	8	0.003 481
10 000	8	0.003 946

The same procedure as followed earlier to determine the optimum training count and good SSE was followed for these models.

DISCUSSION AND CONCLUSIONS

The measured and predicted BOD and COD values from different models are shown in **Figs 2–13** respectively. After each set of training, % ARE for the test data was calculated. The various % ARE values obtained for the test data using these models are shown in **Table 6**.

Table 4. Variation of SSE with different training count and hidden neurons for Model A10

Training count	Hidden neurons	Sum square error
default	4	0.007 400
2000	4	0.007 324
2500	5	0.007 463
1000	6	0.007 345
1500	**8**	**0.007 055**
2500	8	0.007 155

Table 5. Variation of SSE with different training count and hidden neurons for Model A8

Training count	Hidden neurons	Sum square error
default	4	0.005 532
2000	5	0.004 588
5000	**6**	**0.004 547**
5000	7	0.004 891
5000	8	0.005 096
7500	8	0.004 888
10 000	8	0.004 634

Table 6. %ARE for the BOD and COD test data

Model No (BOD)	% ARE	Model No (COD)	% ARE
A1	14.7479	A2	13.4163
A3	11.6614	A4	13.5600
A6	12.8236	A7	15.9200
A9	15.0126	A10	6.9729
A11	12.8982	A12	10.0821

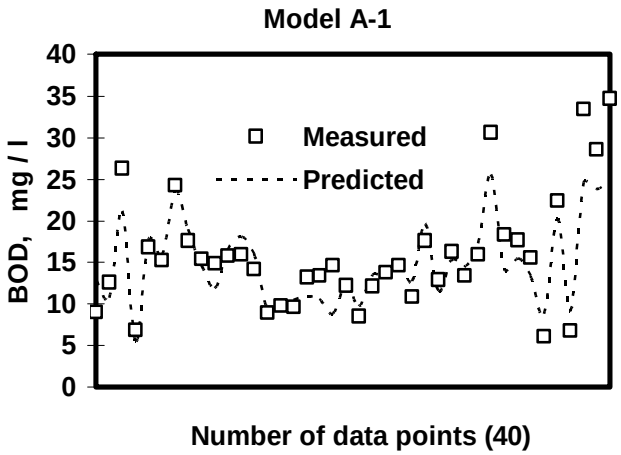


Fig. 2 Measured and predicted test data for BOD concentration from Model A-1.

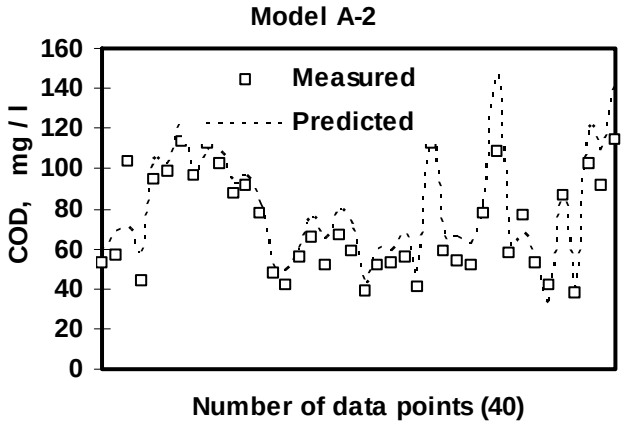


Fig. 3 Measured and predicted test data for COD concentration from Model A-2.

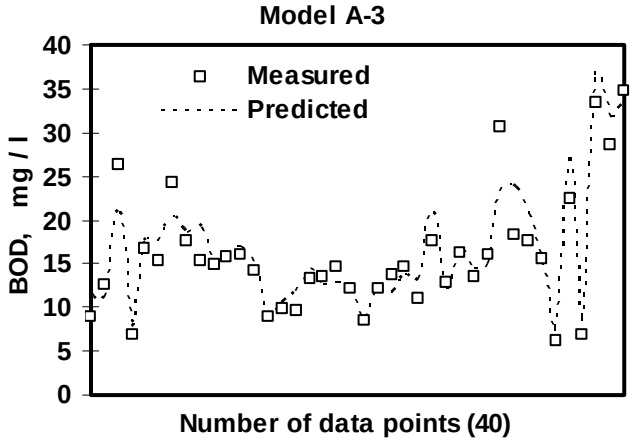


Fig. 4 Measured and predicted test data for BOD concentration from Model A-3.

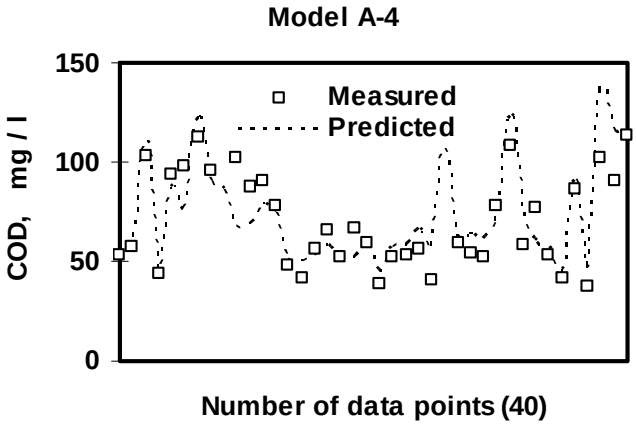
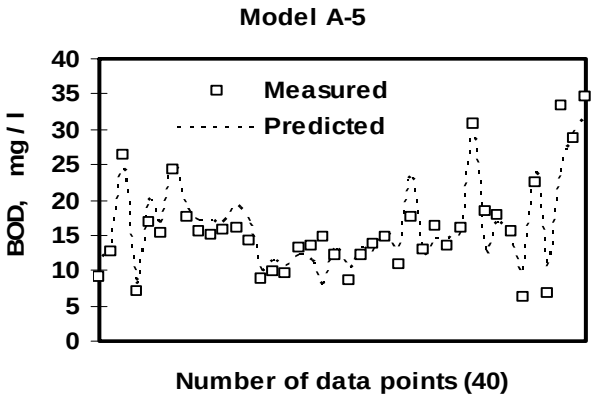
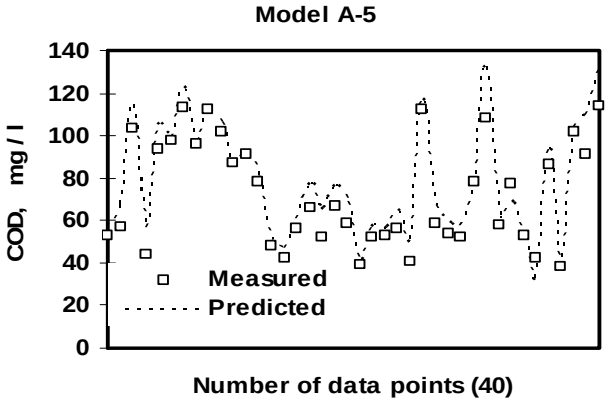


Fig. 5 Measured and predicted test data for COD concentration from Model A-4.



(a)



(b)

Fig. 6 Measured and predicted test data for (a) BOD and (b) COD concentration from Model A-5.

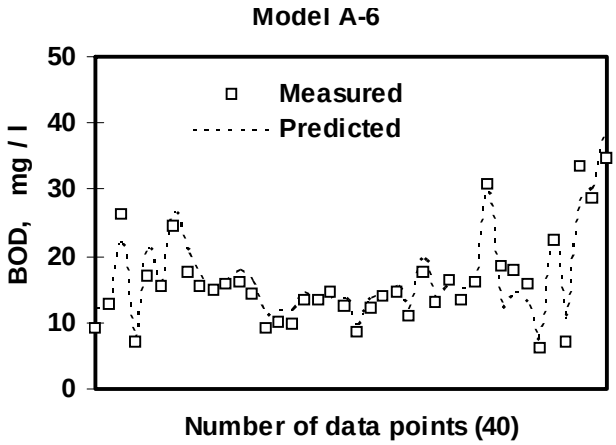


Fig. 7 Measured and predicted test data for BOD concentration from Model A-6.

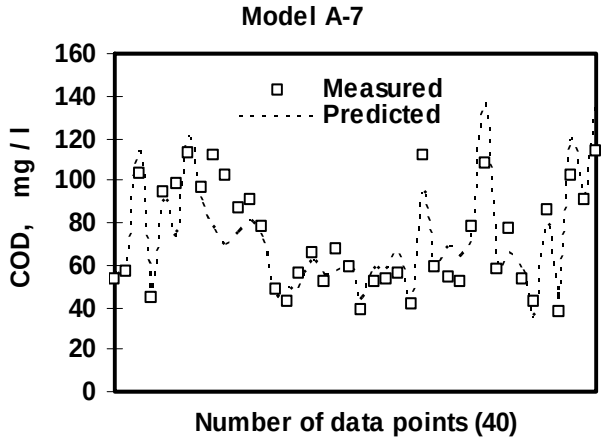
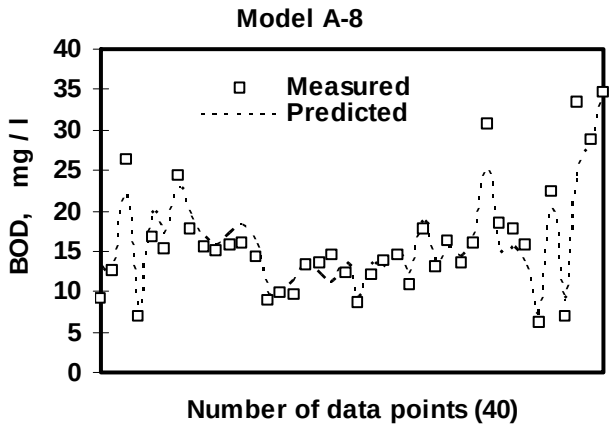
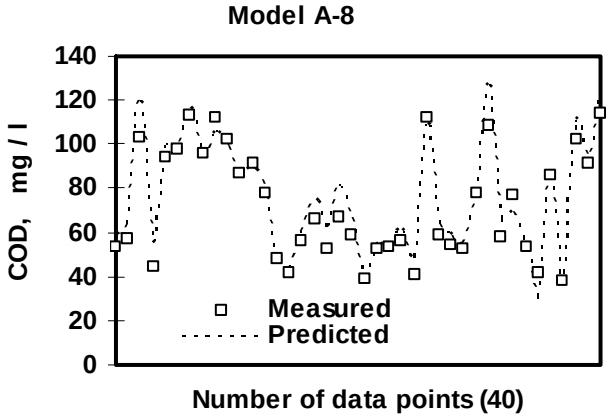


Fig. 8 Measured and predicted test data for COD concentration from Model A-7.



(a)



(b)

Fig. 9 Measured and predicted test data for (a) BOD and (b) COD concentration from Model A-8.

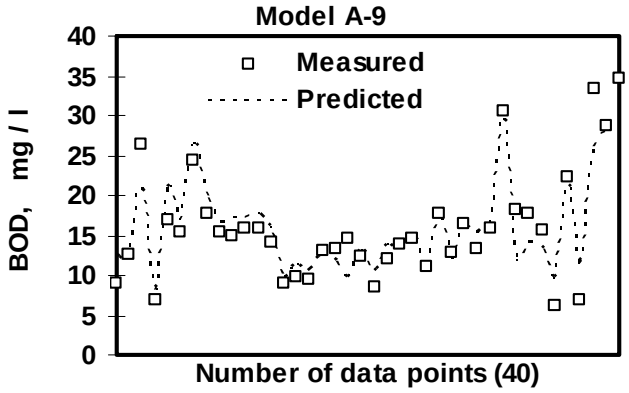


Fig. 10 Measured and predicted test data for BOD concentration from Model A-9.

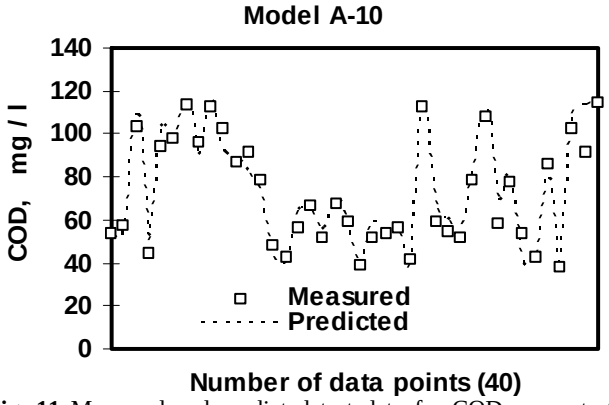


Fig. 11 Measured and predicted test data for COD concentration from Model A-10.

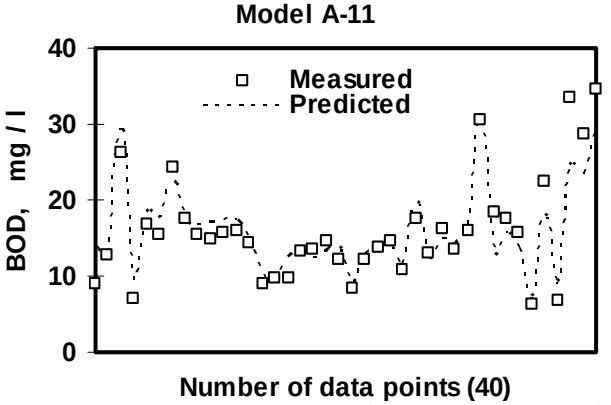


Fig. 12 Measured and predicted test data for BOD concentration from Model A-11.

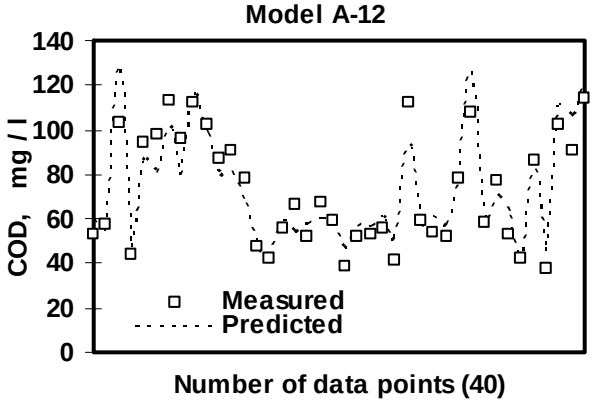


Fig. 13 Measured and predicted test data for COD concentration from Model A-12.

From this table, it is evident that the Model A3 with TOC, Phenol, TSS and AMN as the input parameters was the best model for predicting BOD with a SSE of 0.003 403 in the training data and % ARE of 11.6614 when tested with the test data. Model A3 gave good results at a training count of 5000 and 7 hidden neurons in the hidden layer. All the other models showed comparatively poorer results than model A3 while both training and testing. While testing model A3 with the 40 test data's, 21 (52%) data points were found to be within the 10% limit. Similarly, for the different models developed to predict COD, it was inferred that model A10 with TOC, Phenol and TDS as the input parameters produced better results for predicting COD. This model was formulated with a training count of 1500 and 8 hidden neurons in the hidden layer indicating the training capability of the network. This model gave a SSE of 0.007 055 and when tested with the test data yielded % ARE of 6.9729, which was remarkably good compared to the other models. It is noteworthy to mention that, out of the 40 data points used for testing the network, 30 (75%) data points were found to be within the 10% level of significance. On the other hand, from the results obtained for models developed to predict both BOD and COD simultaneously, it was clearly evident that model A8 with TOC, Phenol, TSS and TDS as the input parameters was able to predict good results for both BOD and COD compared to model A5. Model A8 produced showed better results at a training count of 5000 and 6 hidden neurons in the hidden layer. This model gave a SSE of 0.004547 for the training data and when tested with an external test data gave % ARE of 8.201 for BOD and 11.0835 for COD.

This model gave commendable results when compared with the previous best model for BOD (A3) that produced a SSE of 0.003 403 and % ARE of 11.6614, however for COD it was able to produce satisfactory results compared to the best model for COD (A10). During BOD predictions, 57% (23/40) of the error residuals were found to be below 10% of the measured value, while for COD it was 67% (27/40).

The results of models obtained from NN Model collectively show good statistical significance at the 10% level for the test data. Model A3, was able to predict BOD using TOC, Phenol, TSS, AMN as the model inputs, while Model A10 at a training count of 5000 and 7 hidden neurons in the hidden layer, while Model A10 gave good results for COD using TOC, Phenol and TDS as the inputs at a training count of 1500 and 8 hidden neurons. Interestingly, the combined model A8 developed to predict both BOD and COD was found more effective using TOC, Phenol, TSS and TDS as the inputs. The results from this neural prediction showed very less % ARE values, indicating that the predictions are highly acceptable. Similar data driven modelling approaches can be developed to suit any industrial situation to predict fluctuating effluent concentrations well in advance.

Acknowledgement The authors would like to thank the refinery Senior Manager, Mr. S. Ramesh and the Lab Supervisor Mr. H. Shreekrishna Sharma, for their help during the course of this research work.

REFERENCES

Bandyopadhyay, G. & Chattopadhyay, S. (2007) Single hidden layer artificial neural network models versusmultiple linear regression model in forecasting the time series of total ozone. *Int. J. Environ. Sci. Tech.* **4**(1), 141–149.

Hamed, M.M., Khalafallah, M.G., Hassanien, E.A. (2004) Prediction of wastewater treatment plant performance using artificial neural networks. *Environ. Mod. Soft.* **19**, 919–928.

Haykin, S. (1999) *Neural Networks – A comprehensive foundation*. 6th Indian reprint, Pearson Education, Inc. Singapore.

Hornik, K., Stinchcombe, M., White, H. (1989) Multilayer feed forward networks as universal approximators, *Neural Networks* **2**, 359–356.

Lee, T-L. (2004) Back-propagation neural network for long-term tidal predictions. *Ocean Eng.* **21**, 225–238.

Maier, H.R., Morgan, N., Chow, C.W.K. (2004) Use of artificial neural networks for predicting optimal alum doses and treated water quality parameter. *Environ. Mod. Soft.* **19**, 485–494.

Maier, H.R. & Dandy, G.C. (1998) The effects of internal parameters and geometry on the performance of back propagation neural networks: an empirical study. *Environ. Mod. Soft.* 13, 193–209.

Metcalf, E. (1995) *Wastewater Engineering, Treatment, Disposal and Reuse*. 5th Edition, McGraw Hill, NY.

Poggio, T., Girosi, F., (1990) Networks for approximation and learning. *Proc. IEEE*, **78**(9), 1481–1497.

Rummelhart, D.E., Hinton, G.E., Williams, R.J. (1986). Learning representations by back-propagation errors. *Nature* **323**, 533–536.

Sawyer C.N., McCarty P.L., Parkin G.F. (1994) *Chemistry for Environmental Engineering*. 4th Edition, McGraw-Hill International Editions.

Schuurmann, G., Muller, G. (1994) Back-propagation neural networks recognition vs. prediction capability. *Environ. Toxicol. Chem.* **13**, 743–747.