Kuwait J. Sci. Eng. **36** (2A) pp. 165-178, 2009

Back propagation neural networks to predict the performance of anoxic sulfide oxidizing reactor

QAISAR MAHMOOD[*], YOUSAF HAYAT[**], GHULAM JILANI[***], ZAHID HUSSAIN[**], MUHAMMAD RASHID AZIM[****] AND MOHAMMAD WASEEM[*****]

[*]*Department of Environmental Sciences, COMSATS Institute of Information Technology, Abbottabad 22060, Pakistan*
[**]*Department of Mathematics, Statistics and Computer Science NWFP Agricultural University Peshawar, Pakistan*
[***] *Department of Soil Sciences and Water Conservation, PMAS Arid Agriculture University Rawalpindi Pakistan*
[****] *Department of Botany, Federal Government Post Graduate College H-8, Islamabad, Pakistan*
[*****] *Department of Biology, Allama Iqbal Open University, Islamabad, Pakistan*

ABSTRACT

During the present investigation the data collected from a lab-scale Anoxic Sulfide Oxidizing (ASO) reactor was used in a neural network system to predict performance. Five uncorrelated components of the influent wastewater were used as inputs to the artificial neural network model to predict the final effluent concentrations using back-propagation and general regression algorithms. The best prediction performance was achieved when the data was fed to a back propagated neural network. Within the range of experimental conditions tested, it was concluded that the ANN model gave predictable results for sulfide and nitrite removal from wastewater through the ASO process. The model did not predict the formation of sulfate in an acceptable manner.

Keywords: ASO reactor; Back propagation neural network analysis; effluent sulfide prediction; effluent nitrite prediction; Wastewater treatment.

INTRODUCTION

Preliminary lab scale studies have demonstrated that Anoxic Sulfide Oxidizing (ASO) reactor has been very efficient in the biooxidation of sulfide utilizing nitrite as an electron acceptor in synthetic wastewater (Mahmood et al., 2007). An ASO reactor is an upflow bioreactor containing biomass fed with synthetic wastewater containing sulfide and nitrite. After a specified time period sulfide is converted to sulfate/elemental sulfur while nitrite is oxidized to gaseous nitrogen. Microbial biomass contained in bioreactor is a mixture of various microbial populations collected from anaerobic digestion plant of Sibao wastewater treatment plant (WWTP) located in Hangzhou City, China. Being an anoxic process, potential aeration cost savings are possible using in ASO technology for simultaneous removal of both sulfide and nitrite. The overall reactions occurring in an ASO reactor are as under;

$$3HS^- + 8NO_2^- + 5H^+ \rightarrow 3SO_4^- + 4N_2 + 4H_2O \ \Delta G_m^\theta = -2944KJ/mol...Eq. \quad (1)$$

$$3HS^- + 2NO_2^- + 5H^+ \rightarrow 3S^\circ + N_2 + 4H_2O \ \Delta G_m^\theta = -917KJ/mol.......Eq. \quad (2)$$

Problems related to biological treatment involve dynamical processes, which are very difficult to be described by deterministic modeling techniques. Besides, the process control shows a few difficulties, for instance, the presence of non-linear systems with too many degrees of freedom and uncertainties. Modeling a wastewater treatment process can certainly be put in this category. The biological processes involved in wastewater treatment are highly complex, constituted by different flow phases and biological stages. Applications using neural networks have rapidly increased recently. It has been shown that this technology is appropriate to control problems, making it possible to achieve a better performance than those obtained by conventional models.

In 1943, the original concept of artificial neurons was first proposed by McCulloch and Pitts. Since 1980, there have been a lot of investigations on application of neural networks in control of continuous wastewater treatment processes (Syu and Chen, 1998). Recently, artificial neural networks have also been applied to wastewater treatment processes. The major work being done is the neural network on-line control of treatment processes (Krovvidy and Wee, 1990). Krovvidy and Wee (1993) have reported different AI approaches for studying the sequence of the wastewater treatment systems. The Hopfield network, a network structure proposed by Hopfield (1982, 1984) for solving optimization problems from the concept of energy function, was applied to obtain the optimum sequence of the treatment units in a wastewater plant (Krovvidy et al., 1994). With its precise identification ability, the back-propagation neural network proposed by Werbos (1974) has interested many researches (Khalid et al., 1994, Sastry et al., 1994). Its applications have been distributed in many different types of systems (Bhat and McAvoy, 1990; Boskovic and Narendra, 1995; Glassey et al., 1994; Oin et al., 1994; Su and McAvoy, 1993; Tusar et al., 1992; Yang and Linkens, 1994). It has been successfully applied across a large range of domains such as image recognition, medicine, and molecular biology and, more recently, ecological and environmental sciences (Iglesias et al., 2004; Tutu et al., 2005).

Improper operation of a wastewater treatment plant wastewater treatment plant (WTP) brings serious environmental and social problems, as its effluents can cause or spread various diseases to human beings, as well as to destabilize the natural environment where these wastes are disposed (Belanche et al., 1992). A great contribution to a more efficient operation of wastewater treatment process is attained when predictions of the plant parameters can be issued from observations of past parameters. This will allow better process control.

The main purpose of this work was to predict the effluents of an ASO process using neural networks in a wastewater treatment plant using historical information.

EXPERIMENTAL PROCEDURES

Data collection

Data used for present analysis was collected from a lab scale Anoxic Sulfide Oxidizing (ASO) reactor operating for more than 130 days at Department of Environmental Engineering, Zhejiang University Hangzhou China. Data were collected for various parameters like influent and Effluent sulfide, nitrite, nitrogen, sulfate and pH values. For the detailed experimental procedures Mahmood et al., (2007) may be consulted.

The artificial neural network (ANN)

In the early 1940s, McCulloch and Pitts (1943) explored the competitive abilities of networks made up of theoretical mathematical models when applied to the operation of simple artificial neurons. The structure of an ANN defines the overall architecture of the network, including one input layer, one output layer, and usually one or two hidden layers (Fig. 1).

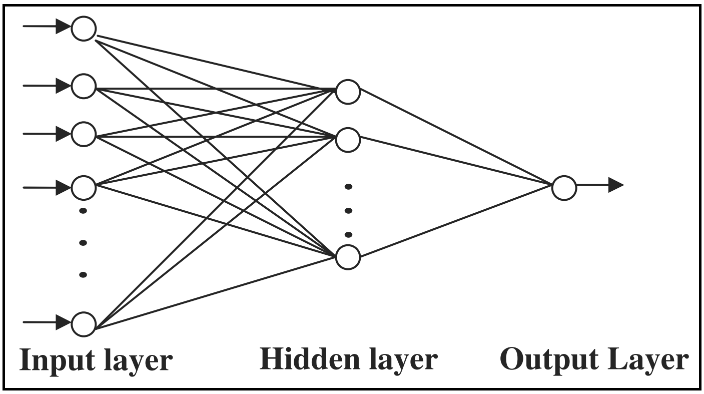

Fig. 1

Each neuron receives a weighted sum from each neuron in the preceding layer and provides an input to each neuron of the next layer. Thus,

$$net = \sum_{i=1}^{n} W_i X_i \text{.....................................} \quad Eq. \quad (3)$$

Where *net* is the summation of the input signal, and W_i denotes an element of the weight vector $\boldsymbol{W}$, and X_i is an element of the input vector $\boldsymbol{X}$. For a given network and input vector, the output vector is totally determined by the weights. The process of finding optimal weights is to find optimal weight, called "training". The training algorithms used in this study were back-propagation. The ANN parameters were: learning rate 0.05, the momentum constant 0.9, the values of MSE were 10^{-10} and the activation function is Tan-sigmoid (Table 1). In this process, the input units and their desired output value are set for the network. The activations of the units are then calculated, feeding forward layer-by-layer from the inputs to the output. A logistic threshold function was used,

$$O = \frac{1}{1 + e^{-net}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad Eq. \qquad (4)$$

Where O is the output of the network. Once the network output value has been produced, it is compared with the target output specified in the training data set. Following this comparison, a backwards adjustment of the weights is performed and the training is stopped when the minimum error for test data found.

$$MSE = \frac{1}{n}\sum_{1}^{n}(P - O)^2 \ldots\ldots\ldots\ldots\ldots \quad Eq. \qquad (5)$$

Where MSE is the mean square error considering prediction (P) and observed values (O) for n testing data vectors. The correlation coefficient (r) or determination coefficient (R^2) was used to evaluate the prediction. When $r = 1$, there is a perfect positive linear correlation between P and O. When $r = -1$, there is a perfectly linear negative correlation between P and O. When $r = 0$, there is no correlation between P and O. Intermediate values present partial correlation.

Table 1

Layer of hidden neurons	Activation function	Learning rate	Momentum constant	Convergence criterion	Epochs
1	Tan-sigmoid	0.05	0.9	1e-10	2000

In our study, the input vector consists of 4 values (Influent pH, sulfide, Nitrite, and Hydraulic Residence Time). The output value is one variable like

Sulfide, N_2 etc. The input data and output data were divided into training and testing data sets. The success of training was determined with the average sum square value between desired output vector and the predicted value, and the final error goal was set to 10^{-10}.

Statistical and graphical work

Statistical and graphical work was carried out using Sigma Plot^{TM} v.10.

RESULTS AND DISCUSSIONS

Influent Concentrations selection

The data to perform artificial neural network analysis was obtained from ASO process utilizing sulfide as electron donor and nitrite as electron acceptor. Various inputs in ASO reactor like sulfide and nitrite can result in the production of sulfate or sulfur, gaseous nitrogen along with formation of some quantities of ammonium in the reactor. Hydraulic residence time (HRT) and pH can also be regarded as input variables in ANN analysis to predict the output. Thus four input variables (sulfide, nitrite, pH and HRT) can result in the production of three outputs (sulfate, nitrogen and ammonium). The range of influent sulfide and nitrite used was 32~1920 mgS/L and 37.75~2265.25 mgN/L, respectively.

Network selection

A neural network that uses gradient-descent error learning is designed and used in our prediction. The neural network has one input layer and one output layer. In the training of a BP neural network, 4 inputs (sulfide, nitrite, pH and HRT) and one output vector sets are generated from each data set. Experimental data from reactor operation of about 120 days were used as the learning set. The number of neurons in the hidden layer is generally selected from the different levels, such as one hidden layer (i.e. 4, 8, 12 layers) and two hidden layers (i.e. 25-12, 25-9, 12-8, 12-4), and 8 neurons in the hidden layer were found successful in the training process for experimental data. The training process has been completed approximately in 2,000 iterations. When the training is completed, a neural network is designed using the obtained weights. In two layers network, a 4-8-1 network is chosen and it can be successfully modeled for the current data.

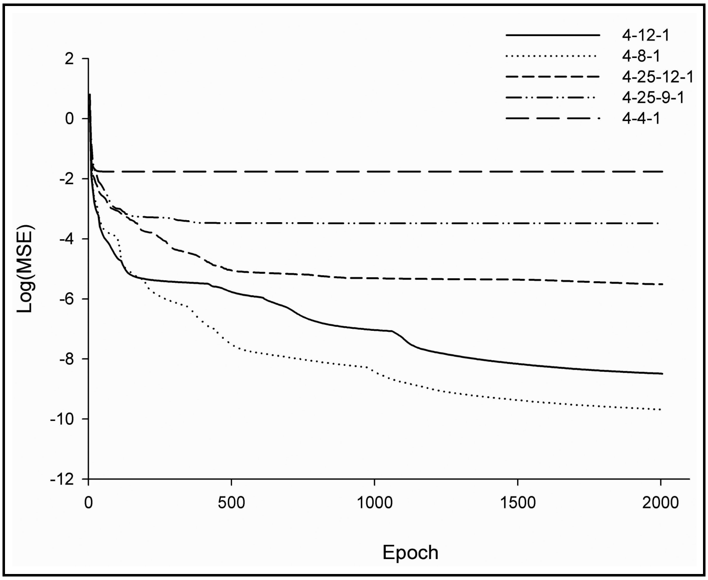

Fig.2

The optimum performance of the ASO reactor ranged for sulfide and nitrite influent concentrations of 32~1664 mgS/L and 37.75~1963.25 mgN/L, respectively; where sulfide and nitrite removal percentages remained 99%~99.8% and 80%~94.47% respectively. An optimum operation of a biological wastewater treatment plant (WWTP) requires the effluent concentrations below the limits set by National Environmental Quality Standards (NEQS) or United States Environmental Protection Agency (USEPA). Though ASO process does not meet the demands of NEQS, however it is very efficient in removing high sulfur and nitrogen loads. Moreover it can polish the effluents from secondary treatment. The results showed a build up of higher amounts of ammonia in the reactor. Higher influent sulfide and nitrite concentrations (> 2000 mg/L) accompanied by ammonia caused overall inhibition of the process. Hence, the ultimate selection of the range of the influent data was used for ANN analysis.

ANN prediction

The prediction of experimental data was conducted using the 8-1 ANN. Training data setting was based on the 50% - 50% sampling technique. Initially, one-half of the cases were randomly selected and used with the BP network. Subsequently, the BP networks were trained for 2000 iterations. The root mean squared training error achieve by the networks is approximately 10^{-10}. Determining the training ending point for the BP network is a delicate task. After training, the correlation coefficient between training and predictive values reaches 0.98 for field data and 0.99 for the training data, respectively.

The relation between observed and predicted parameters during the performance of ASO reactor has been presented in Figures 3-7. Table 1 shows regression analysis for prediction of various parameters through ANN. As shown in Fig. 3-7, when the predicted values calculated by ANN were compared with the observed values, strong and positive correlation coefficient reached 0.805, 0.829, 0.182, 0.986, and 0.926 for pH, sulfide, sulfate, nitrite and nitrogen respectively. Model is very suitable to predict the effluent pH, sulfide, nitrite and nitrogen (Fig.3-6). However, this model is weaker in predicting the effluent sulfate (r = 0.18) as shown in Table 2 and Fig.7.

Table 2

Statistical parameters	pH	Sulfide	Sulfate	Nitrite	Nitrogen
R	0.8051	0.8291	0.182	0.9686	0.926
RSqr	0.6481	0.6874	0.331	0.9383	0.8575

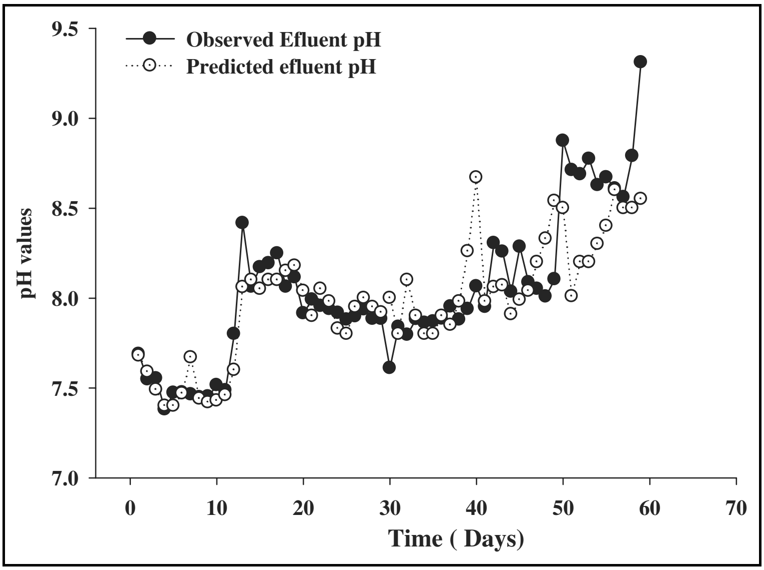

Fig. 3

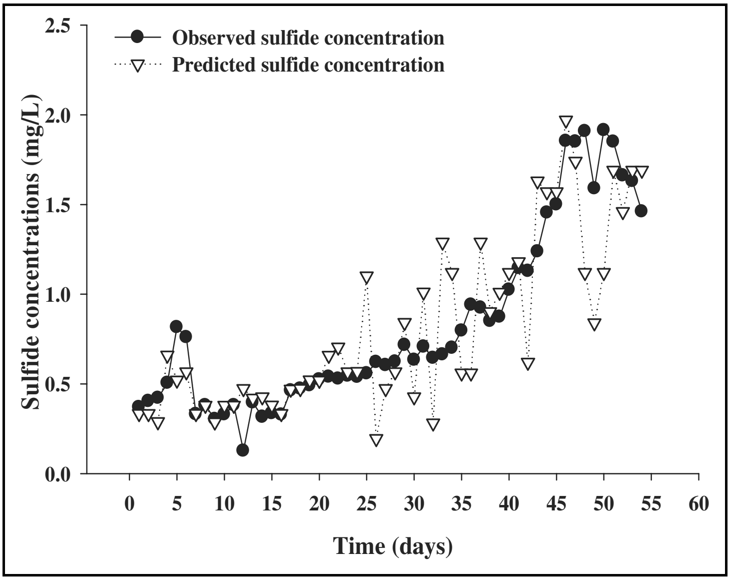
Observed sulfide concentration
Predicted sulfide concentration
Sulfide concentrations (mg/L)
Time (days)
2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
25
30
35
40
45
50
55
60

Fig. 4

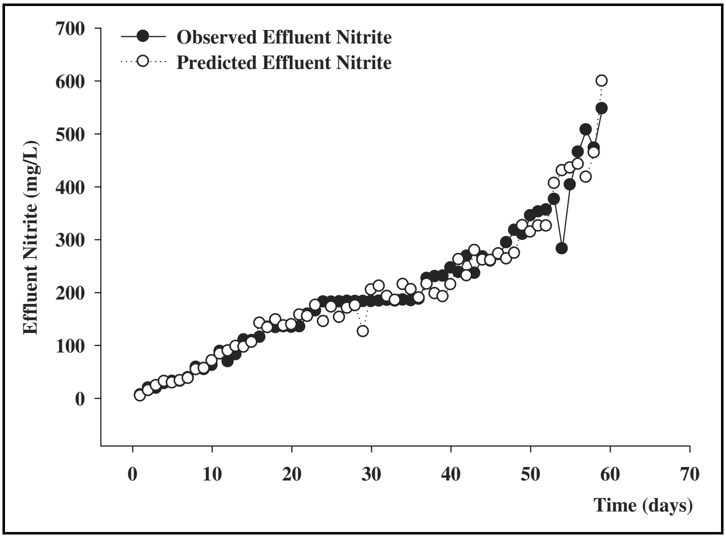
Observed Effluent Nitrite
Predicted Effluent Nitrite
Effluent Nitrite (mg/L)
Time (days)
700
600
500
400
300
200
100
0
0
10
20
30
40
50
60
70

Fig. 5

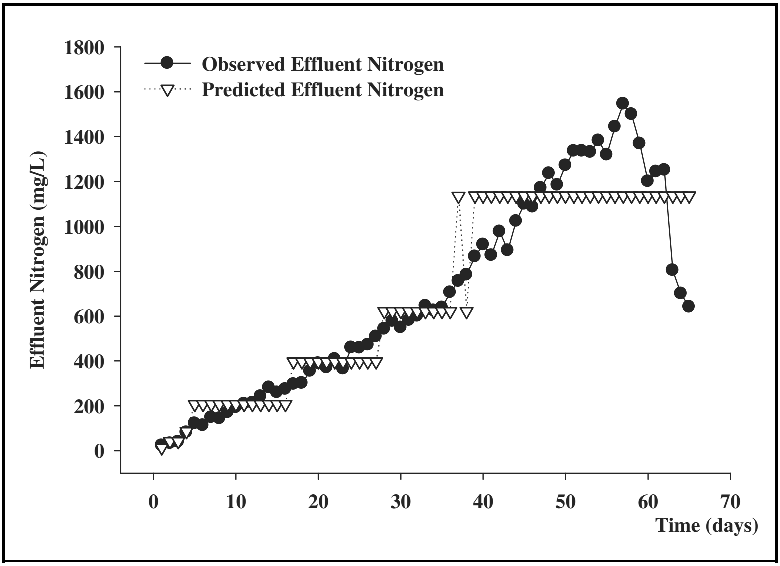
Observed Effluent Nitrogen
Predicted Effluent Nitrogen
Effluent Nitrogen (mg/L)
1800
1600
1400
1200
1000
800
600
400
200
0
0
10
20
30
40
50
60
70
Time (days)

Fig. 6

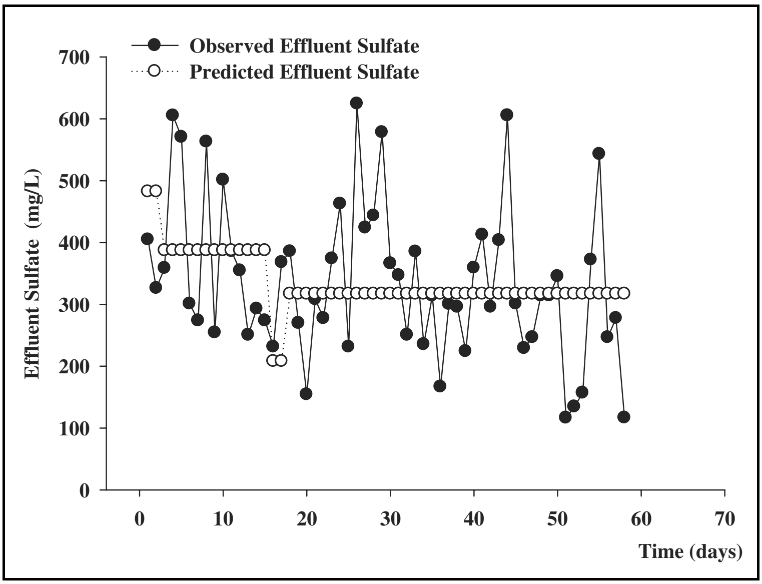
Observed Effluent Sulfate
Predicted Effluent Sulfate
Effluent Sulfate (mg/L)
700
600
500
400
300
200
100
0
0
10
20
30
40
50
60
70
Time (days)

Fig. 7

DISCUSSIONS

In our study, ANNs have been shown to be able to model and predict effluents of Anoxic Sulfide Oxidizing process based on the influent data. Present approach provides quite satisfactory prediction of effluent pH, sulfide, nitrite and nitrogen for ASO reactor. However, the model can not predict the formation of sulfate and ammonium for ASO process. Such unpredictable behavior of ANN model for sulfate is not surprising as there were many fluctuations in sulfate formation during actual experiment (Fig. 7). So model can be accepted based on the sulfide and nitrite removal prediction. The rates of biochemical processes are governed by rules formally analogous to the rules valid for pure chemical processes. However, reaction rates of individual biochemical sub processes are nonlinearly dependent on the plant state and environmental variables such as temperature, pH etc. Frequently the dependence is nonmonotonic; some processes take place only in a narrow band of concentrations (Hrycej, 1997).

The water resources management is a highly complex issue covering a wide spectrum of activities in the field of assessment, planning, designing, operation and maintenance (Zhao and McAvoy, 1996). From more general point of view, all techniques can be applied for prediction, simulation, identification, classification and optimization. Deterministic models are often used for simulation of various processes related to the management of water such as hydrodynamic, morphological, ecological, water quality, groundwater flow etc. All these models use detailed description and fine quantization of the undergoing processes. On the contrary, neural networks do not require the explicit knowledge of physical processes and the relations can be fitted on the basis of measured data. In many or most occasions it was shown that the neural networks tend to give better result than the deterministic models, provided that the process under consideration is not changed in time. If significant variables are known, without knowing the exact relationships, ANN is suitable to perform a kind of function fitting by using multiple parameters on the existing information and predict the possible relationships in the coming future. This sort of problem includes rainfall-runoff prediction, water level and discharge relations, drinking water demand, flow and sediment transport, water quality prediction etc. In order to represent data more efficiently, it is needed to extract the most important features in the data set. Unsupervised neural networks often incorporate self-organizing features, enabling them to find unknown regularities, meaningful categorization and patterns in the presented input data (Zhao and McAvoy, 1996).

Neural networks have frequently been used for environmental process modeling and control (Melas et al., 2000). For NN part of the scheme, an error back propagation network and a recurrent neural network are used. The scheme was tested with both simulated and measured plant data. Melas et al., (2000) presented a BPNN model developed for 24-h prediction of photochemical pollutant levels. The model relates peak pollutant concentrations to meteorological and emission variables and indices, as well as maximum hourly pollutant concentrations during the three previous days of the week. An experiment was used to investigate sensitivity of model predictions to uncertainty in the input data. The experiment shows an agreement between observed and predicted concentrations. However, the results of the BPNN model were only marginally better than those obtained by regression models in the literature, possibly due to inadequacy of input parameters (i.e. some important predictors were either not included in the analysis because they were unavailable or under-represented) and inadequate representation of the data used (i.e. only local measurements of some data were obtained and they may not be representative of a larger area).

Reich et al., (1999) proposed a BPNN with momentum to identify the apportionment of a small number of sources from a data set of ambient concentrations of a given pollutant. The trained network was able to identify the most likely emission parameters of an unknown source, and hence, to determine the relative allocation of the point sources involved. The real case is a noisy generation with a higher uncertainty than the given examples. The neural network was able to solve the inverse source-receptor problem in the presence of noisy or ambiguous data. However, it was unable to adapt to data outside the validation region of the selected set of examples.

CONCLUSIONS

Within the range of experimental conditions tested, it was concluded that the ANN model gave predictable results for sulfide and nitrite removal from wastewater through ASO process. The model did not predict the formation of sulfate and ammonium to an acceptable manner. Apart from experimentation, ANN model can help to simulate the results of such experiments in finding the best optimal choice for ASO based denitrification.

REFERENCES

Belanche, L. Sànchez, M., Cortés, U. & Serra, P. 1992. A knowledge-based system for the diagnosis of waste-water treatment plants. Proc. Of 5th IEA/AIE-92. Lecture Notes in Artificial Intelligence, 604:324-336.

Bhat, N. McAvoy, & T. J. 1990. Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems. Comput Chem Eng, 14:573-582.

Boskovic, J. D. & Narendra, K. S. 1995. Comparison of Linear, Nonlinear and Neural Network Based Adaptive Controllers for A Class of Fed-batch Fermentation Processes. Automatica, 31:817-823.

Chan, C. W. & Huang, G. H. 2003. Artificial intelligence for management and control of pollution minimization and mitigation processes. Engineering Applications of Artificial Intelligence,16:75-90.

Glassey, J., Montague, G. A., Ward, A. C. & Kara, B. V. 1994. Artificial Neural Network Based Experimental Design Procedures for Enhancing Fermentation Development. Biotechnol. Bioeng, 44:397-403.

Hopfied, J. J. 1982. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proc Natl Acad Sci USA,79:2554-2558.

Hopfield, J. J. 1984. Neurons with Graded Response Have Collective Computational Properties such as those of Two-state Neurons. Proc Natl Acad Sci USA, 81:3088-3092.

Hrycej, T.1997. Neurocontrol Towards an Industrial Control Methodology in Biological Wastewater Treatment Control.. John Wiley and Sons Inc. USA, p331-362.

Iglesias, A., Varela, B. A., Cotos, J. M., Taboada, & J. A., Dafonte C. 2004. A comparison between functional networks and artificial neural networks for the prediction of fishing catches. Neural Computing and Applications, 13:24-31.

Khalid, M., Omatu, S. & Yusof, R. 1995. Temperature regulation with Neural Networks and Alternative Control Schemes. IEEE Trans Neural Networks, 6:572-582.

Krovvidy, S. & Wee, W. G. 1990. A Knowledge Based Neural Network Approach for Wastewater Treatment System. Int Joint Conf Neural Networks., 1:327-332.

Krovvidy, S. Wee, W. G. 1993. Wastewater Treatment Systems from Case-based Reasoning, Mach Learning.,10: 341-363.

Krovvidy, S., Wee, W. G., Suidan, M., Summers, R. S., Coleman & J. & J. Rossman, L. 1994. Intelligent Sequence Planning for Wastewater Treatment Systems. IEEE Expert., 9:15-20.

Mahmood, Q., Ping, Z., Jing, C., Donglei, W., Baolan & H., Jinye, L. 2007. Anoxic sulfide biooxidation using nitrite as electron acceptor. J. Hazard. Mat., 147: 249-256.

McCulloch, W.S. & Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. Bull math Biophys, 5:115-133.

Melas, D., Kioutsiouskis, L. & Ziomas, L.C. 2000. Neural network model for predicting peak photochemical pollutant levels. Journal of the Air & Waste Management Association., 50:495-501.

Oin, Z., Reid, J. E., Litchfield, B. J., Ren, J. L. &Chang, S.W. 1994. A Prototype Neural Network Supervised Control System for *Bacillus thuringiensis* Fermentations. Biotechnol Bioeng., 43:483-489.

Reich, S.L., Gomez, D.R. & Dawidowski, L.E. 1999. Artificial neural network for the identification of unknown air pollution sources. Atmospheric Environment, 33:3045-3052.

Sastry, P. S., Santharam, G. & Unikrishnan, K. P. 1994. Memory Neuron Networks for Identification and Control of Dynamical Systems. IEEE Trans. Neural Networks., 5:306-319.

Su, H. T., and McAvoy, T. J. 1993. Neural Model Predictive Control of Nonlinear Chemical Processes. Proc. 1993 Int Symp Intell Control. 358-365.

Syu, Mei-J. & Chen Bow-C. 1998. Back-propagation Neural Network Adaptive Control of a Continuous Wastewater Treatment Process. Ind Eng Chem Res., 37:3625-3630.

Tusar, M., Zupan, J., & Gasteiger, J. 1992. Neural Networks and Modeling in Chemistry. J. Chem. Phys., 89, 1517-1529.

Tutu, H., Cukrowska, E. M., Dohnal, V. & Havel, J. 2005. Application of artificial neural networks for classification of uranium distribution in the Central Rand goldfield, South Africa, Environmental Modeling & Assessment., 10:143-152.

Werbos, P. 1974. Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. Dissertation, Harvard University, Cambridge, MA,

Yang, Y. Y., & Linkens, D. A. 1994. Adaptive Neural Network Based Approach for the Control of Continuously Stirred Tank reactor. *IEE* Proc Control Theory Appl., 141:341-349.

Zhao, H., & McAvoy, T.J. 1996. Application of ANN & Fuzzy Logic in IWM. www.stowa-nn.ihe.nl/Applications_ANN_Fuzzy_Logic.htm.

Submitted : *5/5/2008*
Accepted : *24/5/2009*

إعادة نشر الشبكات العصبية للتنبؤ بأداء مفاعل الكبريت الأنوكسي المؤكسد

[*] قيصر محمود ، [**] يوسف حياة ، [***] غلام جيلاني ، [**] زاهد حسين ، [*] محمدى راشد عظيم و محمد وسيم

[*] قسم النبات – كلية الدراسات العليا – إسلام آباد – باكستان
[**] قسم الرياضيات – الإحصاء وعلوم الكمبيوتر – جامعة بشاور الزراعية – باكستان
[***] قسم علوم التربية – جامعة رو البندي للزراعة القاحلة – باكستان

خلاصة

إن المعلومات التي تم الحصول عليها من مختبر Anoxic sulfide oxidizing (ASO) تم استخدامها في نظام الشبكات العصبية للتنبؤ بالأداء. تم استخدام خمس مكونات غير مترابطة لمياه الصرف الصحي الداخلة كعناصر داخلة في نموذج الشبكة العصبية الصناعية للتنبؤ بالتركيز النهائي لمياه الصرف الصحي الخارجة وذلك باستخدام إعادة النشر و اللورغاريتمات العامة التراجعية. لقد تم الحصول على تنبؤ أحسن أداء عندما تم استخدام المعلومات لإمداد إعادة نشر الشبكة العصبية

ومن خلال مدى ظروف التجارب التي تم اختبارها، تم التوصل إلى استنتاج أن نموذج ANN قد تنبأ بنتائج لإزاحة الكبريت و النترات من مياه الصرف الصحي خلال عملية (ASO).

كما أن النموذج لم يتنبأ في تكوين الكبريت بطريقة مقبولة.