Prediction of Phenolic Compound Migration Process through Soil Media using Artificial Neural Network Approach

Supriya Pal, Kalyan Adhikari, Somnath Mukherjee, Sudipta Ghosh

Abstract—This study presents the application of artificial neural network for modeling the phenolic compound migration through vertical soil column. A three layered feed forward neural network with back propagation training algorithm was developed using forty eight experimental data sets obtained from laboratory fixed bed vertical column tests. The input parameters used in the model were the influent concentration of phenol(mg/L) on the top end of the soil column, depth of the soil column (cm), elapsed time after phenol injection (hr), percentage of clay (%), percentage of silt (%) in soils. The output of the ANN was the effluent phenol concentration (mg/L) from the bottom end of the soil columns. The ANN predicted results were compared with the experimental results of the laboratory tests and the accuracy of the ANN model was evaluated.

Keywords—Modeling, Neural Networks, Phenol, Soil media

I. INTRODUCTION

ENVIRONMENTAL pollution is considered as a negative effect of industrial activities. Lithosphric and aquatic environment (Soil, lakes, rivers etc.) are highly contaminated due to emanating of various toxic compounds. An example of such compound is phenol. The phenolic compounds pollute the environment through on route discharge of industrial wastewater from coal conversion, coke preparation, petroleum refineries, and pesticide, insecticide, pulp and paper, plastic, textile, dye, polymeric resin, and pharmaceutical industries [1]. Phenolic compounds are very harmful to organisms even at low concentrations due to its toxicity and carcinogenicity properties [2]. Several methods have so far been proposed for removal of phenol, e.g., biological oxidation, thermal liquid phase oxidation, photochemical conversion, catalytic oxidation, physical adsorption, and solvent extraction [3]. Among these methods, adsorption is a well-established and powerful technique for treating domestic and industrial effluents. Viraraghavan and Alfaro investigated the adsorptive capacity of phenol by peat, flyash and bentonite with an initial phenol concentration of 1000 mg/L. They found that the adsorption capacity of these

adsorbents was much lower. Kinetic study results revealed that a long equilibrium time (15 h) was needed for the adsorption of phenol by the materials [4]. Taha et al. studied the adsorption behavior of phenol in granite residual soil and kaolinite. Batch adsorption tests were conducted at both low phenol concentration and higher phenol concentration. Three adsorption isotherms, linear, Freundlich and Langmuir were used in the analysis. It was found that residual soil possesses a greater adsorption capacity compared with kaolinite and granite residual soil has a greater potential for use as a soil liner material [5]. Phenol adsorption by natural soil has already been done experimentally by researchers using laboratory batch adsorption test and fixed bed vertical soil column tests. But the modeling of the complex processes of phenol migration through soil media using soft computing tools is lacking. Artificial neural network are capable to capture complex and non linear relationship between input and output patterns. Annadurai and Lee used artificial neural network for the modeling of adsorption of Phenol using Pseudomonas Pictorum (NICM-2074). The input parameters of ANN model were Maltose dose (g/L), Phosphate (g/L), pH and Temperature. The predictive results of the ANN model and Multiple Regression Analysis (MRA) were also compared. The values of coefficient of correlation for ANN model and MRA model were 0.97 and 0.90.This indicates that the results of ANN model are better than that of MRA model [6]. Yetilmezsoy and Demirel used back propagation neural network for the modeling of Pb (II) adsorption from aqueous solution using Antep Pistachio Shells by using MATLAB software. They had done modeling with ten types of back propagation training algorithms available in MATLAB software. They found that results of Levenberg Marquardt back propagation algorithm (Coefficient of Correlation = 0.93) are better than other training algorithms. The efficiency of adsorption of Pb (II) using Pistachio Shells is @ 99%. The maximum adsorption capacity of Pistachio Shell for Pb (II) was observed 27.1mg/g [7]. Bashseer et al. investigated the feasibility of utilizing the concept of neural nets in developing networks for predicting the breakthrough curves of fixed-bed absorbers. Close agreement was observed between the breakthrough curves predicted by the developed neural network and those obtained from the mathematically based adsorption model(HSDM)[8].The objective of this study was to model the phenolic compound migration through soil media based on the experimental result on fixed bed vertical soil column tests. The artificial neural network approach for modeling the phenol migration was also attempted and the accuracy of the model was judged using some statistical parameters such as correlation coefficient, mean square error.

F. A. Supriya Pal is with the Civil Engineering Department, National Institute of Technology, Durgapur 713209, West Bengal, India (phone: +91-9434713292; e-mail: supriya_pal@ rediffmail.com).

S. B. Kalyan Adhikari is with the Geology Department, National Institute of Technology, Durgapur 713209, West Bengal, India.

T. C. Somnath Mukherjee is with the Civil Engineering Department, Jadavpur University, Kolkata 700 032, West Bengal, India

F.D. Sudipta Ghosh is with the Civil Engineering Department, Jadavpur University, Kolkata 700 032, West Bengal, India

World Academy of Science, Engineering and Technology
International Journal of Geological and Environmental Engineering
Vol:5, No:3, 2011

II. MATERIALS AND METHODS

A. Collection of Soil Samples

Soil samples were collected from different neighboring sites in Durgapur, Dist. - Burdwan, West Bengal, India. Soil samples were collected from 50 – 100 cm depth below the ground surface by auger boaring. The collected samples were stored in polythene bags and taken to the laboratory. Samples were oven dried for 24 hours and the physical properties such as grain size distribution, specific gravity, organic carbon content (%), porosity, compaction characteristics, liquid limit (%), plastic limit (%) etc. were analyzed as per standard methods.

B. Fixed Bed Vertical Soil Column Tests

The fixed bed vertical soil column test was performed to obtain effluent phenol concentration with time. A series of columns of circular cross section (inner diameter of 3 cm) and 30 cm high made of perspex glass were used in the studies. Homogeneous soil from the sites were placed in the columns and compacted in different heights (5 cm, 10 cm and 15 cm) and then saturated with water. Glass wool and glass balls were placed at the top and bottom end of the soil samples. The synthetic stock solutions of phenol of 1000 mg/L were prepared by dissolving a known amount of analytical grade of phenol in distilled water. The stock solutions were diluted to make 10, 20, 30, 40 mg/L concentration of phenol for tests. Diluted phenol solutions (10 – 40 mg/L) were used as influent phenol concentration. The solutions were allowed to pass through the soil columns and the effluent phenol solution from the bottom end of the column was collected in 100 ml capacity glass bottles in different time interval. The collected phenol solution was then filtered through Whitman No. 42 filter paper and the phenol concentration were analyze using a Techcomp UV 2300 Spectro photometer with a wavelength of 510 nm after developing color with 4-aminoantipyrine and potassium ferricyanide in an alkaline medium. Thus the effluent phenol concentration with respect to time was obtained.

C. Artificial Neural Network (ANN)

Neural networks are simplified models of the biological structure of human brains. A neural network model consists of an interconnected assembly of simple processing elements, neurons, which are organized in layered fashion [9]. In general, neural network model architecture, as shown in Fig.1 consists of three main layers: an input layer (independent variables), an output layer (dependent variables) and one or more intermediate hidden layers.
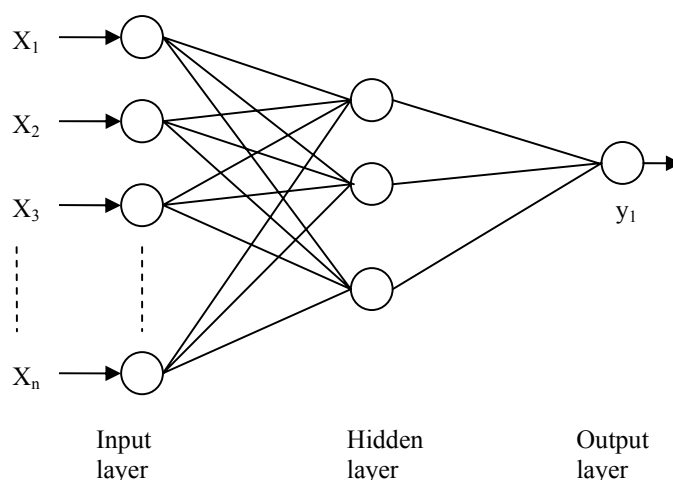

Fig.1 Artificial Neural Network Architecture

In this present work, Neural Network Toolbox V 8.0 of MATLAB$^{(R)}$ mathematical software was used to predict the effluent phenol concentration. A three layered feed forward neural network with back propagation training algorithm was used to develop the ANN model. Tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer were used. The neurons in the hidden layer were varied to develop most efficient model. The five input parameters used in the model were the influent concentration of phenol(mg/L) on the top end of the soil column, depth of the soil column (cm), elapsed time after phenol injection (hr), percentage of clay (%), percentage of silt (%) in soils. The range of input data used in the model is shown in Table 1.

TABLE I
RANGE OF INPUT VARIABLES USED IN THE MODEL

Sr.No.	Input variables	Range
1.	Influent concentration of phenol(mg/L)	10 - 40
2.	Depth of the soil column (cm)	5 - 15
3.	Elapsed time after phenol injection (hr)	0 - 490
4.	Percentage of clay (%)in soils	18 – 45
5.	Percentage of silt (%) in soils	42 - 61

Forty-eight experimental data sets were used to develop the model architecture. The data sets were divided into three groups. The first group consisting of one half of data sets (24 data sets) were used to train the network, the second group consisting of one fourth of data sets (12 data sets) were used for testing the model and third group consisting of remaining one fourth data sets (12 data sets) were used for validating the developed model. In the ANN analysis, it is very much

World Academy of Science, Engineering and Technology
International Journal of Geological and Environmental Engineering
Vol:5, No:3, 2011

necessary to preprocess the data. Data preprocessing helps in ensuring equal attention to all variables [10]. Preprocessing can be in the form of data sealing, normalization and transformation [11]. So, just before training stage, the data sets obtained in the fixed bed vertical column test were scaled in the range of 0 and 1 using equation no.1.

$$P_N = \frac{p_{act} - p_{min}}{p_{max} - p_{min}} \qquad (1)$$

Where, p_N, p_{act}, p_{min} and p_{max} are normalized, observed, minimum and maximum values of the data series respectively. The neural network output data was again post processed by converting it into denormalizing unit before comparison is made.

III. RESULT AND DISCUSSIONS

The Levenberg- Marquardt (LM) back-propagation algorithm was used for the training of ANN model. The trained ANN model was then tested and validated with the experimental results to estimate the effluent phenol concentration with time. Initially the training of the ANN model was started with 2 neurons in the hidden layer. It was then increased upto 20 neurons. Different MSE (mean squared error) values were found with different number of neurons in the hidden layer. Fig.2 shows the relation between the MSE values and number of neurons in the hidden layer.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (T_i - A_i) \qquad (2)$$

N = Number of data point
T_i = Network predicted value at the ith data.
A_i = Experimental value at the ith data.
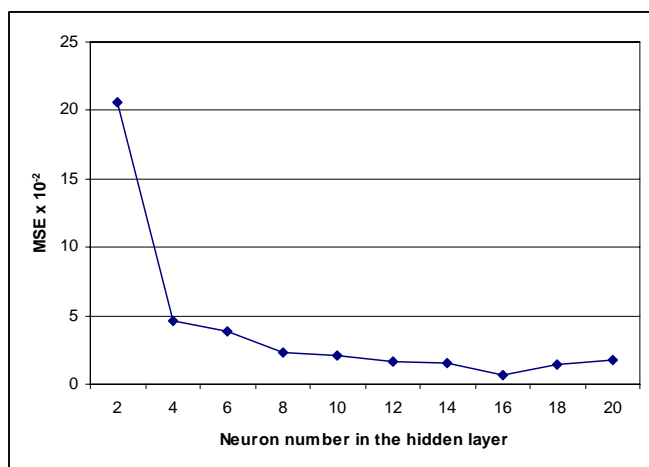i = an index of the data.

Fig. 2 Relation between the MSE and number of neurons in the hidden layer

It was observed from Fig.2 that MSE value was 0.206253 for 2 numbers of neuron in the hidden layer. It was then decreasing with increase number of neuron and found minimum (0.006921) at 16 numbers of neurons in the hidden

layer. The MSE value again shows increasing trend with the increase number of neuron in the hidden layer. So, the 16 numbers of neuron in the hidden layer may be considered as optimum for this ANN model.

Fig.3 shows the training, validation and test mean squared error for the Levenberg-Marquardt algorithm. The training was stopped after 35 epochs.
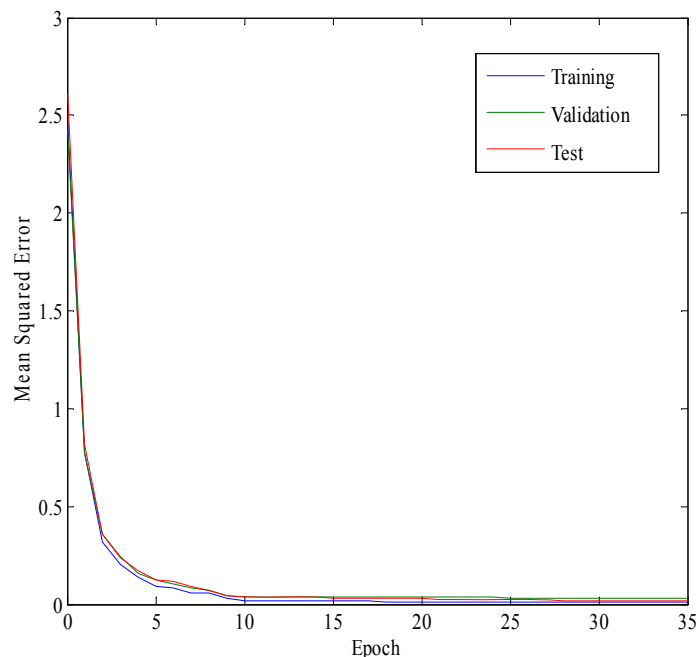

Fig. 3 Training, validation and test mean squared errors for the Levenberg-Marquardt algorithm
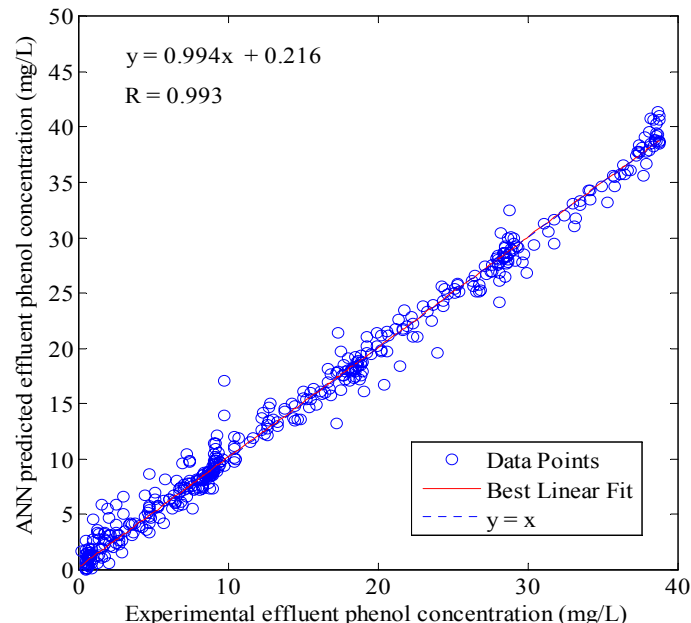

Fig.4 Comparison of the experimental and ANN predicted values of effluent phenol concentration.

Fig.4 shows the experimental and ANN predicted values of effluent phenol concentration from the bottom end of the soil

World Academy of Science, Engineering and Technology
International Journal of Geological and Environmental Engineering
Vol:5, No:3, 2011

column. The coefficient of correlation of 0.993 was found from the plot in this figure. The coefficient of correlation, R, is used to assess the accuracy of the ANN model. Smith suggested the following guide for values of IRI between 0.0 to 1.0 [12].

IRI ≥ 0.8 strong correlation exists between two sets of variables.

$0.2 \leq IRI \leq 0.8$ correlation exists between the two sets of variables.

IRI ≤ 0.2 weak correlation exists between the two sets of variables.

This result indicates that the ANN model was able to predict the effluent phenol concentration quite reasonably. Fig.5 shows the experimental result and ANN predicted values of effluent phenol concentration with time. A three layered ANN with tangent sigmoid transfer function (tansig) at hidden layer with 16 neurons and linear transfer function (purelin) at output layer with Levenberg- Marquardt (LM) training algorithm best predicts the phenol migration through vertical soil column.
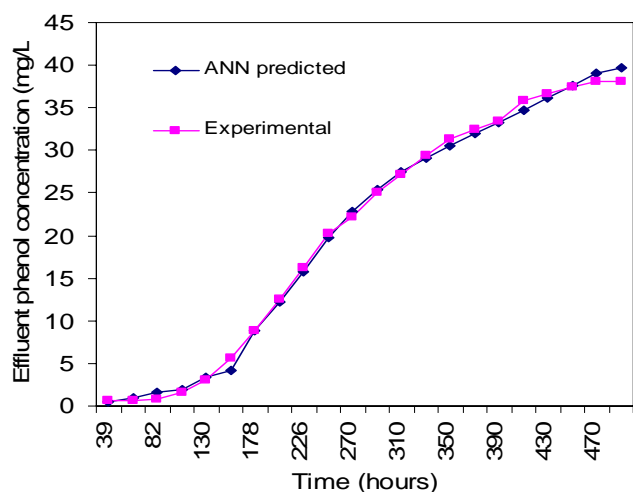

Fig.5 Comparison between ANN predicted and Experimental values of effluent phenol concentration with contact time (Silt=42%, Clay=45%, Influent concentration of phenol=40 mg/l, Depth of soil column=15 cm)

IV. CONCLUSION

In the present study, a three layer ANN with a tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer were used to predict the effluent phenol concentration with time passed through vertical soil column of different depths. All the ANN modeling and analysis were carried out using Neural Network Toolbox V 8.0 of MATLAB$^{(R)}$ mathematical software. The architecture of the ANN model had a 5-16-1 network trained with Levenberg-Marquardt (LM) back propagation algorithm. The MSE value was found to be 0.006921. The predicted results of ANN model were compared with laboratory fixed bed vertical soil column test results. Close agreement was found between ANN predicted values and experimental values with IRI value as 0.993. The ANN model is found to estimate

the breakthrough pattern satisfactorily. Therefore, it can be concluded that ANN has a good potential to model the complex phenol migration process through soil media.

REFERENCES

[1] F.Gonen and Z.Aksu,"A comparative adsorption/biosorption of phenol to granular activated carbon and immobilized activated sludge in a continuous packed bed reactor"; Chemical Engineering Communication; 190, (2003), 763-778.
[2] A.T.M.Din, B.H. Hameed and A.L. Ahmad, "Batch adsorption of phenol onto physiochemical- activated coconut shell", Journal of Hazardous Materials; 161(2-3) (2009), 1522-1529.
[3] P.P. Mitra, and T.K. Pal, "Treatment of effluent containing phenol and catalytical conversion", Int. Chem. Eng., 1994, 41 (1), 26-30.
[4] T. Viraraghavan and J. Alfaro,"Removal of phenol from wastewater by adsorption on peat, fly ash and bentonite", J. Environ. Eng, 1992, 5: 311– 317.
[5] M.R. Taha, T.O. Leng, A.B. Mohamad and A.A.H. Kadhum, "Batch adsorption tests of phenol in soils", Bull Eng Geol Env, 2003, 62:251–257
[6] G. Annadurai and J.F. Lee,"Application of artificial neural network model for the development of optimized complex medium for phenol degradation using seudomonas pictorum (NICM 2074)", Biodegradation, 2007, 18:383–392.
[7] K.Yetilmezsoy and S. Demirel, "Artificial neural network (ANN) approach for modeling of Pb (II) adsorption from aqueous solution by Antep pistachio (Pistacia Vera L.) shells", Journal of Hazardous Materials, 2008, 153:1288-1300.
[8] I.A. Basheer   and Y.M. Najjar,"Predicting dynamic response of adsorption columns with neural nets", Journal of Computing in Civil Engineering, ASCE, Vol.10, No.1, January,1996.
[9] O. Gunadym, "Estimation of soil compaction parameters by using statistical analyses and artificial neural networks", Environ Geol, 2009, 57:203-215.
[10] H.R. Maier and G.C. Dandy, "Neural Networks for the prediction and forecasting of water resources variables. A review of modeling issues and applications", Environmental Modeling & Software, 2000, 15, 101-124.
[11] T. Masters,"Practical neural network recipes in C++", Academic Pres, 1993, San Diego, California.
[12] G. N. Smith,"Probability and statistics in civil engineering: An introduction", Collins,1986, London.