Hydrology Journal, 29(1-2) *March-June 2006*

Neural network based solutions for locating groundwater pollution sources

JITENDRA KUMAR, ASHU JAIN & RAJESH SRIVASTAVA
Department of Civil Engineering, Indian Institute of Technology, Kanpur-208 016

Abstract The identification of pollution sources in aquifers is an important area of research for hydrologists and governmental agencies. It may be possible to locate the polluting industry, given the data in terms of pollutant concentration measurements at observation wells and the aquifer parameters. Traditionally, hydrologists have relied on the conceptual methods for the identification of groundwater pollution sources. Recently, artificial neural networks (ANNs) have emerged as an attractive and easy to implement alternative to solve complex problems efficiently. Some researchers have used ANNs for the identification of pollution sources in aquifers. A major problem with most previous studies using ANNs has been the large size of the neural networks that are needed to model the inverse problem. The breakthrough curves at an observation well may consist of hundreds of concentration measurements, and presenting all of them to the input layer of an ANN not only results in huge networks but also requires large amount of training and testing data sets to develop the ANN models. In this paper, we present the results of a study aimed at estimating groundwater pollution source location using ANNs through the use of two different methods of presenting the breakthrough curves data as inputs to the ANN models. To simplify the ANN architectures, these methods do not employ the whole breakthrough curves as the inputs to the ANNs. The feed-forward multi-layer perceptron type of ANN architecture was employed to develop various ANN models which were trained using the back-propagation method. The results show that the ANNs can be very efficient tools for locating pollution sources and that it is possible to obtain good ANN model performance even with extremely simplified architectures involving a very few input variables.

Key words: pollution identification; aquifers;ANN.

INTRODUCTION

Groundwater is a major source of water in India for agriculture, municipal, and industrial sectors. The quality of groundwater has traditionally been very good requiring no or minimal treatment in most cases. However, over the last couple of decades, the groundwater has been at a high risk of being contaminated by the harmful chemicals due to many reasons such as rapid industrialization, increased use of pesticides, and increase in the number of underground petrol storage tanks. Once an aquifer has been contaminated, it may take a very long time and considerable expenditure to restore it to an usable state. Due to the large costs of cleaning operations of contaminated aquifers, it is necessary to identify the source of the pollution so that suitable punitive measures could be imposed on the polluting industry/individual/agency to recover some of the costs. Complete identification of pollution source involves determination of source concentrations,

duration, and location. The physical processes involved in the movement of water and contaminants in the aquifers are highly complex, non-linear, and dynamic processes affected by a wide range of physical variables. The identification of the pollution sources is much more complex in the sense that it requires an inverse modeling of the flow and contaminant transport. Over the past few decades, various investigators have looked at the problem of identification of groundwater pollution sources using a wide variety of techniques.

The simplest approach is to use forward simulations with assumed source location and release history and compare the results with the observed data. However, it is not very efficient due to the infinite number of possible combinations and some type of optimization method has to be used to obtain the best solution. Probably the earliest such study was that of Gorelick et al. (1983) who used forward-time simulations with an optimization model based on linear programming and multiple regressions. They incorporated the transport model as constraints in the form of a response matrix. Wagner (1992) considered an inverse model as a non-linear maximum likelihood estimation problem for simultaneous model parameter estimation and source characterization. Mahar & Datta (1997, 2000) combined the identification of a pollutant source with the optimal design of a monitoring network for an efficient identification process. Mahar & Datta (2000) used a classical nonlinear optimization technique to estimate the magnitude, location and duration of groundwater pollution sources under transient conditions. A different approach was proposed by Skaggs & Kabala (1994). They attempted to reconstruct the history of the plume using Tikhonov Regularization (TR). 1-D solute transport through a saturated homogeneous medium was studied with a complex contaminant release history and assuming no prior knowledge of the release function. Samarskaia (1995) applied the TR with fast Fourier transforms to a groundwater contamination source reconstruction problem. Liu & Ball (1999) used modified TR technique to study a contaminant release at Dover Air Force Base, Delaware, USA. They used field measured concentration profiles in lowpermeability porous media that underlie the contaminated aquifer. Singh et al. (2004) used ANNs for identification of unknown groundwater pollution sources. The ANN was trained to identify source characteristics based on simulated contaminant concentration measurement data at specified observation locations in the aquifer. The performance of ANN models was found to be very effective for source identification. Singh & Datta (2004) utilized a trained ANN to simultaneously solve the problems of estimating unknown groundwater pollution sources and estimating unknown hydro-geologic parameters.

It is clear that a variety of techniques have been investigated by researchers for pollution source identification. Recently, ANNs have also been employed for this purpose; however, most of the studies reported earlier have focused on identifying the source release history at potential locations. In real aquifers, identifying the location of the pollution source is extremely important for taking punitive measures. Further, a major problem in many of the earlier studies attempting to use ANNs for pollution source identification has been the very large and complex ANN architectures needed to present the whole breakthrough curves to the ANNs. It may be possible to obtain good performance in identification of pollution source

location by presenting an ANN with a smaller number of inputs characterizing the breakthrough curves in some manner. The authors have not encountered any study attempting to reduce the dimensionality of ANN architectures for the purpose of identification of pollution source locations.

The objectives of this study are to (a) investigate the use of ANN methodology to estimate the distance of the pollution source from an observation well where the measured breakthrough curve is available, and (b) investigate the use of two methodologies using simplified representation of the breakthrough curves as inputs to the ANNs to estimate the distance of the pollution source from an observed well. The paper begins with a brief overview of the ANN technique followed by the model development before presenting the results and making concluding remarks.

ARTIFICIAL NEURAL NETWORKS

An ANN is a highly inter-connected network of many simple processing units called "neurons" or "neurodes". Neurons having similar characteristics are grouped in a single layer. For example, the neurons in an input layer receive input from an external source, and transmit the same to a neuron in an adjacent layer, which could either be a hidden layer or an output layer. Each neuron in an ANN is also capable of comparing an input to a threshold value. The input output data presented to an ANN are normally scaled between 0 and 1. The ANN stores the information captured from the input vector as the 'strengths of the connections' between the neurons. The most commonly used ANN in engineering applications is a feed-forward ANN (see Figure 1). In this figure, each neuron is represented by a circle and each connection by a line. The ANN shown in Figure 1 consists of three layers: an input layer consisting of three neurons, a hidden layer also consisting of three neurons, and an output layer consisting of one neuron. In a feed-forward ANN, the inputs presented to the neurons in an input layer are propagated in a forward direction and the output vector is calculated through the use of a non-linear function called activation function. The activation function should be continuous, differentiable, and bounded from above and below. Then, knowing the output, the error at the output layer from the ANN model can be computed. The computed error is then back propagated through the network and the 'connection strengths' are updated using some training mechanism such as 'generalized delta rule' (Rumelhart et al., 1986). This process of feed-forward calculations and back-propagation of error is repeated until an acceptable level of convergence is reached. This whole process is known as training of the ANN. Once the network has been trained, it can be used for prediction. A feed-forward ANN with generalized delta rule as its training mechanism was employed in the present study to develop all ANN models. More details of an ANN can be found in ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000).

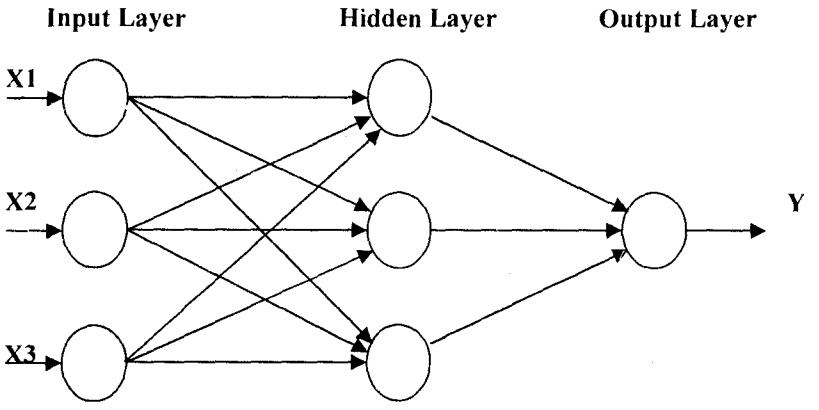

Fig. 1. Structure of a Feed-forward ANN

MODEL DEVELOPMENT

Data Generation

The data for the problem of pollution source identification consist of observed breakthrough curve at observation wells and the source strengths, duration, and location. Since in the present study, the objective is to estimate the distance of the pollution source given a breakthrough curve, the pollution source is assumed to be active for constant duration injecting a conservative pollutant at a constant rate. The breakthrough curves at different distances were then calculated using the governing equations for the pollutant transport as follows:

The one-dimensional transport of conservative solutes through a homogeneous saturated semi-infinite porous media is represented by the advection-dispersion equation:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - v\frac{\partial C}{\partial x} \tag{1}$$

in which C is the concentration, t is time, D is the dispersion coefficient, x is the distance and v is the groundwater velocity. Its solution requires one initial and two boundary conditions which are dictated by the type of problem considered. In this study, we consider an initially uncontaminated aquifer with a pollutant source of constant strength releasing the pollutant till certain time and stopping after that. The initial and boundary conditions are then given by;

$$C(x,0) = 0 \tag{2}$$

$C(0,t) = C_0$ for $t \le T_0$ and 0 otherwise

$C(\infty,t) = 0$

in which T_0 is the duration of release and C_0 is concentration at source. The solution of the above equations is obtained by utilizing the solution to a step input according to Ogata & Banks (1961) as:

$$C = \frac{C_0}{2}\left[erfc\left(\frac{x-vt}{\sqrt{4Dt}}\right) + e^{\frac{vx}{D}} erfc\left(\frac{x+vt}{\sqrt{4Dt}}\right)\right] \quad \text{for } t < T_0 \tag{3}$$

$$C = \frac{C_0}{2}\left[erfc\left(\frac{x-vt}{\sqrt{4Dt}}\right) + e^{\frac{vx}{D}} erfc\left(\frac{x+vt}{\sqrt{4Dt}}\right)\right] - \frac{C_0}{2}\left[erfc\left(\frac{x-v(t-T_0)}{\sqrt{4D(t-T_0)}}\right) + e^{\frac{vx}{D}} erfc\left(\frac{x+v(t-T_0)}{\sqrt{4D(t-T_0)}}\right)\right]$$

for $t > T_0$ (4)

Equations (3) and (4) were employed to generate the breakthrough curves using the following data: v = 0.1 m/day; D = 0.1 m^2/day; C_0 = 1000 mg/l; and T_0 = 120 days. The time interval of the breakthrough curves was taken as 30-days. A total of 2,000 input-output patterns were generated, of which 1,500 were used for training, 250 were used for validation, and 250 were used for testing. The data were scaled in the range of 0.1 and 0.9 to avoid saturation during training. A brief description of various error measured employed in this study for model development is provided next.

Performance Statistics

Five different standard performance statistics were employed for model development. These are normalized root mean square error (NRMSE), Nash-Sutcliffe efficiency (E), coefficient of correlation (R), average absolute relative error (AARE), and threshold statistics (TS). These can be calculated using the following equations:

$$NRMSE = \frac{\sqrt{\frac{1}{N}\sum(XO-XE)^2}}{\overline{XO}} \tag{5}$$

$$E = 1 - \frac{\sum(XE-XO)^2}{\sum(XO-\overline{XO})^2} \tag{6}$$

$$R = \frac{\sum (XO - \overline{XO})(XE - \overline{XE})}{\sqrt{\sum (XO - \overline{XO})^2 \sum (XE - \overline{XE})^2}} \quad (7)$$

$$AARE = \frac{1}{N} \sum \left| \frac{XE - XO}{XO} \right| \times 100\% \quad (8)$$

$$TS_x = \frac{N_x}{N} \quad (9)$$

Where XO is the observed value of the variable, XE is the estimated value of the variable from a model, $\overline{XO}$ is the average observed value of the variable, $\overline{XE}$ is the average estimated value of the variable, N_x is the number of data points estimated for which the absolute relative error (ARE) is less than x%, N is the total number of data points predicted, and all the summations run from 1 to N. The value of x of 1%, 10%, and, 50% were considered in this study to compute threshold statistics.

The normalized root mean square error (NRMSE) is a relative measure of the residual variance from the model. The value of NRMSE close to 0.0 indicates better performance. The correlation coefficient (R) measures the degree of linear dependence between two series. It ranges between -1.0 and 1.0 with a higher value indicating good linear dependence between observed and estimated values. The coefficient of efficiency (E), proposed by Nash & Sutcliffe (1970), is also one of the widely employed relative statistic in hydrologic literature. A value of 1.0 represents a perfect prediction while a model with E = 0.0 is no more accurate than predicting the mean observed value. The ANN models are trained by minimizing the sum squared error (SSE) at the output layer, which is similar to the global error statistics such as NRMSE, E, and R. In order to test the robustness of the ANN model developed, it is important to test the model using some other performance statistics such as average absolute relative error (AARE) and threshold statistics (Jain et al., 2001; Jain & Ormsbee, 2002, 2004; Jain & Srinivasulu, 2004; and Jain & Srinivasulu, 2006). The AARE provides an overall average error in estimating the variable being modeled, clearly, lower the AARE better is the model performance. The threshold statistics provides information on the distribution of the prediction errors, and higher TS values indicate better model performance.

ANN Model Development

The ANN models developed in this study consisted of three layers: an input layer, a hidden layer, and an output layer. Two different ANN models were developed that differed in the manner of presenting the breakthrough curve to the input layer of the ANN. In the first method (called ANN-1 Model), the whole breakthrough curve was divided into ten parts, and the pollution concentrations at the eleven end points were computed. This time distribution of the pollutant concentrations was then modeled as 22-N-1 in ANN-1 model. The second method (called ANN-2

Model) employed the first, second and third temporal moments of the breakthrough curves to capture the essential features of the curves using only a few characteristics. The n^{th} temporal moment at a given sampling point is defined as

$$M_n = \int_0^\infty C(t) t^n dt \tag{10}$$

The zeroth moment represents the area under the breakthrough curve, which is indicative of the mass passing through the sampling point for a constant groundwater velocity. The first moment, normalized with the zeroth moment, represents the mean arrival time of the plume:

$$\bar{t} = \frac{\int_0^\infty C(t)\, t\, dt}{\int_0^\infty C(t)\, dt} \tag{11}$$

The higher order moments are typically computed about the mean arrival time and are written as

$$\mu_n = \frac{\int_0^\infty C(t)(t-\bar{t})^n\, dt}{\int_0^\infty C(t)\, dt} \qquad n = 2,3,... \tag{12}$$

The second moment therefore represents the spread about the mean and the third provides an indication of the skewness of the breakthrough curve. Thus, the structure of the ANN-2 model (3-N-1) was considerably simplified. The output neuron in both the ANN models represented the distance of the pollution source from the location where the breakthrough curve was observed.

The main task in the development of the ANN model is the determination of the optimum number of neurons (N) in the hidden layer. The number of neurons in the hidden layer is, in fact, responsible for capturing (or mapping) the dynamic and complex relationship among various input and output variables considered. The sigmoid activation function was used as the transfer function at both hidden and output layers. This study employed the popular back-propagation training algorithm using step-wise learning with momentum factor. The value of learning coefficient of 0.075 and momentum correction factor of 0.075 was used while training. The value of N was varied from 1 to 20, and for each N, the back-propagation algorithm was used to minimize SSE at the output layer. Each of the ANN architectures was trained for a maximum of 50,000 iterations or when the SSE reached 0.0005. Three error statistics, namely R, SSE, and AARE were used to determine the best ANN architecture (or optimal N). Figure 2 and Figure 3 show the graphs between the number of hidden neurons and different error statistics during training from ANN-1 and ANN-2 models, respectively. It can be noted from

the two figures that the correlation coefficient R is almost constant for all the hidden neurons for both ANN models. The SSE first decreases, and then almost remains constant when the number of hidden neurons is increased. These two statistics (R & SSE) therefore do not help much in determining the best number of hidden neurons to capture the complex relationships inherent in the input and output data. Therefore, in order to select the best architectures for each ANN model, ANN architectures having number of hidden neurons equal to 3, 6, 8, and 10 were selected for the ANN-1 model for further analysis. Similarly, ANN architectures having number of hidden neurons equal to 5, 10, and 14 were selected for the ANN-2 model for further analysis. These selections were made based on the AARE values being minimum for the number of hidden neurons (see Figure 2 and Figure 3) for the two models, respectively. The analyses for the selection of the best ANN architectures for each of the models were based on the training results only. The results in terms of various performance statistics from the selected models during training, validation, and testing are presented in Table 1, and Table 2 from ANN-1 and ANN-2 models, respectively.

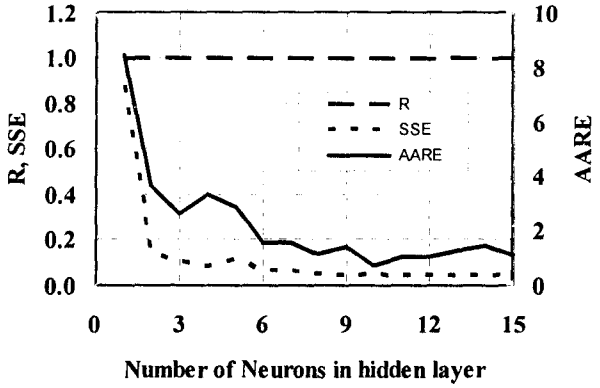

Fig. 2. Error Statistics v/s number of hidden neurons for ANN-1Model

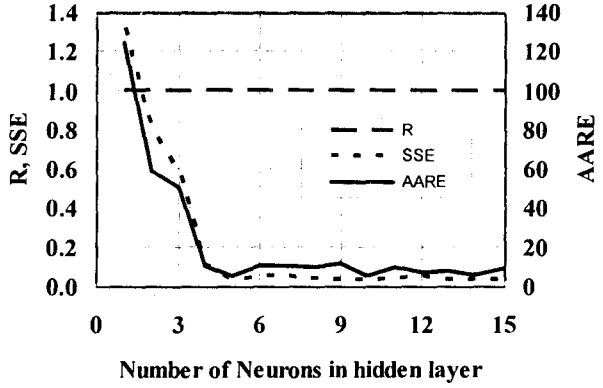

Fig. 3. Error Statistics v/s number of hidden neurons for ANN-2 Model

RESULTS & DISCUSSIONS

Analyzing the results for ANN-1 models from Table 1, it can be noted that the values of correlation coefficient and Nash-Sutcliffe efficiency in excess of 0.95 were obtained from all the models, which represents excellent performance. The values of AARE of less than 3% from all models during training, validation, and testing shows that the ANN-1 model is very efficient in predicting the distance of the pollution source from the sampling well very accurately. It is to be noted that the 22-10-1 architecture of the ANN-1 model is deemed to be the best since its performance in terms of AARE during training and validations data sets is the best (0.71% and 2.28%, respectively), and its performance during testing is also very good. The 22-10-1 ANN model performed comparable to the other models in terms of NRMSE, E, and R statistics. This model also had the best TS statistic for most ARE levels. The results in the form of a scatter plot from the 22-10-1 ANN model is shown in Figure 4. The narrow and uniform spread around the ideal line indicates that it was able to predict the distance of the pollution source very accurately for all magnitudes.

Table 1. Performance Statistics from ANN-1 Models

Model	NRMSE	E	R	AARE	TS1	TS10	TS50
During Training							
22-03-1	0.0052	0.999919	0.999960	2.58	86.27	97.87	99.53
22-06-1	0.0042	0.999946	0.999974	1.56	89.27	99.07	99.67
22-08-1	0.0036	0.999962	0.999982	1.13	90.93	99.33	99.73
22-10-1	0.0036	0.999962	0.999982	0.71	91.00	99.07	99.87
During Validation							
22-03-1	0.1266	0.951940	0.976004	2.90	60.80	92.80	99.20
22-06-1	0.1264	0.952036	0.976091	2.71	67.60	97.20	99.20
22-08-1	0.1265	0.952010	0.976074	2.98	65.60	97.20	99.60
22-10-1	0.1265	0.951993	0.976074	2.28	67.60	94.80	99.60
During Testing							
22-03-1	0.1263	0.951756	0.975920	2.75	61.60	92.80	99.20
22-06-1	0.1262	0.951837	0.975972	2.81	66.80	97.20	98.40
22-08-1	0.1262	0.951821	0.975957	2.47	66.40	96.00	99.20
22-10-1	0.1263	0.951790	0.975963	2.95	67.60	94.40	98.80

Analyzing the performance of ANN-2 model from Table 2, it can be noted that the performance of all the models is very good in terms of most of the error statistics and deteriorated slightly (as compared to ANN-1 model) in terms of the other statistics. All the selected ANN-2 models obtained R & E values in excess of 0.99 during training and in excess of 0.90 during validation and testing data sets,

which can be characterized as very good. However, the AARE values are larger compared to the ANN-1 models indicating marginally poor performance. The AARE value close to 5% during training and about 20% during validation and testing is considered reasonable for practical applications. There was significant deterioration in TS1 statistic meaning that the number of data points estimated from ANN-2 models having errors less than 1% have gone down considerably. This is due to the loss of information, as expected, when only three inputs are provided to the ANN. Based on all the results in Table 2, the 3-5-1 model is found to be the best among this category. It obtained AARE values of 5.08%, 18.0%, and 15.4% during training, validation, and testing, respectively. The performance of the 3-5-1 model was comparable to the other models in terms of E, NRMSE, and R statistics also, therefore, it was selected as the best following the principle of parsimony.

Table 2. Performance Statistics from ANN-2 Models

Model	NRMSE	E	R	AARE		TS1
TS10	TS50					
During Training						
03-05-1	0.0096	0.999723	0.999871	5.08	76.73	95.73	99.47
03-10-1	0.0096	0.999722	0.999872	5.10	77.13	97.13	99.07
03-14-1	0.0091	0.999750	0.999884	5.54	79.87	97.07	99.00
During Validation							
03-05-1	0.1685	0.914798	0.961237	18.0	05.60	50.00	98.40
03-10-1	0.1644	0.918902	0.964239	25.9	05.20	54.00	93.20
03-14-1	0.1774	0.905566	0.960817	18.8	05.60	45.20	95.20
During Testing							
03-05-1	0.1670	0.915639	0.961648	15.4	05.60	50.80	98.40
03-10-1	0.1631	0.919600	0.964529	21.8	05.20	54.80	93.20
03-14-1	0.1753	0.907121	0.961131	18.7	04.80	46.00	96.00

Although the performance of the first method was certainly better than the second one, as expected, it must be emphasized that the second method involves far simpler ANN architectures involving only three neurons in the input layer as compared to 22 in the first method. Comparing the performances of the best models based on the two methodologies (22-10-01 & 3-5-1), it can be observed that the performance of 22-10-1 ANN model is marginally better than that of the 3-5-1 ANN model in terms of the most of the error statistics considered in this study. Therefore, the 3-5-1 ANN model can be used for the purpose of groundwater pollution source location estimation.

SUMMARY & CONCLUSION

This paper presents the results of a study aimed at estimating groundwater pollution source location using ANNs through the use of two different methods of presenting the breakthrough curves as inputs to the ANN models. The first method divides the whole breakthrough curve in 10 equal parts, and the second methods uses first three temporal moments of the breakthrough curves, to be presented to the ANNs. The feed-forward multi-layer perceptron type of ANN architecture was employed to develop various ANN models trained using back-propagation method. The data for ANN model development were generated using the analytical solution of the problem of one-dimensional steady flow and transient contaminant transport in homogeneous aquifer. The training data set was divided into training and validation to prevent over-training and/or under-training. A wide variety of standard performance statistics were used to evaluate the performance of various ANN models.

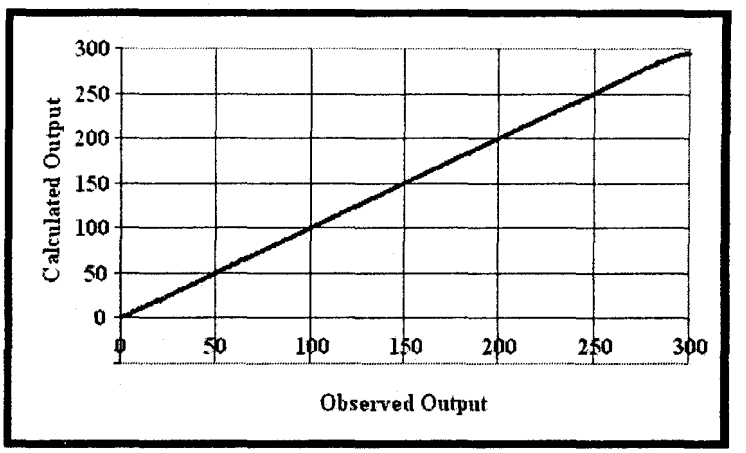

Fig. 4. Scatter Plot from 22-10-1 ANN Model

The results obtained in this study indicate the suitability of the ANNs in solving the complex problem of inverse modeling for groundwater pollution source identification. The first method that involves presenting the breakthrough curve using 10 divisions was found to be superior to the second methodology. This is expected because presenting the breakthrough curve to an ANN using only three input neurons results is certain amount of loss of important information contained in the data. However, the performance of the second type of ANN models was found to be comparable to that of the first type of ANN models in terms of NRMSE. E, & R but the major differences were found in terms of TS & AARE statistics. This demonstrates the necessity of using a wide variety of statistical parameters to evaluate the performance of various ANN models developed on the same data set.

A limitation of the study presented has been that perfect data obtained from the analytical solution of the groundwater flow and transport problem were employed

for ANN model development. In reality, the breakthrough curves obtained from aquifers contain many type of errors e.g. measurement errors etc. How the ANN models will be able to perform when presented with the noisy data remains to be investigated. Also, the data employed in the current study come from a homogeneous aquifer, and it would be interesting to develop ANN models using actual data for existing aquifers that involve heterogeneity in terms of various flow and contaminant transport parameters. It is hoped that further research efforts will focus some of these directions.

REFERENCES

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology (2000), Artificial neural networks in hydrology I: Preliminary concepts, *J. Hydrol. Engg.*, ASCE, **5 (2)**, 115-123.

Gorelick, S.M., Evans, B.E., & Remson, I. (1983). Identifying sources of groundwater pollution: an optimization approach, *Water Resour. Res.* , **19(3)**, 779–790.

Jain, A. & Ormsbee, L.E. (2002). Evaluation of Short-Term Water Demand Forecast Modeling Techniques- Conventional Methods versus AI, *J. American Water Works Assoc.*, **94 (7)**, 64-72.

Jain, A. & Ormsbee, L.E. (2004). An Evaluation of the Available Techniques for Estimating Missing Fecal Coliform Data, *J. Amer. Wat. Resour. Assoc.*, **40(6)**, 1617-1630.

Jain, A. & Srinivasulu, S. (2004). Development of effective and efficient rainfall-runoff models using integration of deterministic, real-coded genetic algorithms, and artificial neural network techniques, *Water Resources Research*, **40(4)**, W04302, doi: 10.1029/ 2003WR002355.

Jain, A. & Srinivasulu, S. (2006). Integrated approach to modelling decomposed flow hydrograph using artificial neural network and conceptual techniques, *J. Hydrol.*, **317(3-4)**, 291-306.

Jain, A., Varshney, A.K., & Joshi, U.C. (2001). Short-Term Water Demand Forecast Modeling at IIT Kanpur using Artificial Neural Networks, *Water Resour. Mgmt.*, **15(5)**, 299-321.

Liu, C. & Ball, W.P. (1999). Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware, *Water Resour. Res.*, **35(7)**, 1975–1985.

Mahar, P.S. & Datta, B. (1997). Optimal monitoring network and ground-water pollution source identification, *J. Water Resour. Plng. & Mgmt.*, ASCE **123**, 199–207.

Mahar, P.S. & Datta, B. (2000). Identification of pollution sources in transient groundwater systems, *Water Resour. Mgmt.* **14(3)**, 209–227.

Nash, J. E. & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models: I. A discussion of principles, *J. Hydrol.*, **10**, 282-290.

Ogata, A., & Banks, R.B. (1961). A solution of the differential equation of longitudinal dispersion in porous media, *U.S. Geol. Surv.*, Prof. Pap. No. **411-A**.

Rumelhart, D.E., Hinton, G.E. & Williams, R. J. (1986). Learning representations by back-propagating errors, *Nature*, **323**, 533-536.

Samarskaia, E. (1995). Groundwater contamination modeling and inverse problems of source reconstruction, *SAMS*, **18–19**, 143–147.

Singh, R. M., & Datta, B. (2004). Groundwater pollution source identification and simultaneous parameter estimation using pattern matching by artificial neural network. *Env. Forensics*, **5 (3)**, 143-153.

Singh, R.M., Datta, B., & Jain, A. (2004). Identification of unknown groundwater pollution sources using artificial neural networks, *J. Wat. Resour. Plng. & Mgmt.*, ASCE **130(6)**, 506-514.

Skaggs, T.H. & Kabala, Z.J. (1994). Recovering the release history of a groundwater contaminant, *Water Resour. Res.* **30**, 71–79. No. 1.

Wagner, B.J. (1992). Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modeling, *J. Hydrol.*, **135**, 275–303.