

# Identification of Unknown Groundwater Pollution Sources and Determination of Optimal Well Locations Using ANN-GA Based Simulation-Optimization Model

Sophia Leichombam, Rajib Kumar Bhattacharjya

Department of Civil Engineering, Indian Institute of Technology Guwahati, Guwahati, India  
Email: l.sophia@iitg.ernet.in, rkbc@iitg.ernet.in

Received 14 February 2016; accepted 28 March 2016; published 31 March 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

The linked simulation-optimization model can be used for solving a complex groundwater pollution source identification problem. Advanced simulators have been developed and successfully linked with numerous optimization algorithms for identification of groundwater pollution sources. However, the identification of pollution sources in a groundwater aquifer using linked simulation-optimization model has proven to be computationally expensive. To overcome this computational burden, an approximate simulator, the artificial neural network (ANN) model can be used as a surrogate model to replace the complex time-consuming numerical simulation model. However, for large-scale aquifer system, the performance of the ANN-based surrogate model is not satisfactory when a single ANN model is used to predict the concentration at different observation locations. In such a situation, the model efficiency can be enhanced by developing separate ANN model for each of the observation locations. The number of ANN models is equal to the number of observation wells in the aquifer. As a result, the complexity of the ANN-based simulation-optimization model will be related to the number of observation wells. Thus, this study used a modified formulation to find out the optimal numbers of observation wells which will eventually reduce the computational time of the model. The performance of the ANN-based simulation-optimization model is evaluated by identifying the groundwater pollutant sources of a hypothetical study area. The limited evaluation shows that the model has the potential for field application.

## Keywords

Linked Simulation-Optimization, Groundwater Pollutant Source, Inverse Optimization, Artificial Neural Networks, Genetic Algorithm

---

## 1. Introduction

Identifying the groundwater pollutant source is a very difficult and time-consuming process due to the involvement of aquifer simulation model with the optimization model. As simulation model is linked with the optimization model, the model is known as the simulation-optimization model. The source identification model can be regarded as an inverse problem and can be solved using inverse optimization technique. The complexity of the aquifer simulation model has led the solution of the source identification model highly computationally expensive. The focus of every source identification model is not only to solve the problem but also to make it efficient with respect to computational time. Numerous attempts have been made by adopting various simulation models but the real challenge lies on the degree of effectiveness of the simulator. There are various techniques for solving the groundwater source identification problems *viz.* response matrix approach, embedded technique and linked simulation-optimization methods. The study on unknown groundwater pollution source identification using response matrix method was first adopted by Ref. [1]. They used the least square method and linear programming for solving the optimization model. This approach was adopted by many researchers in the earlier stage of identification of unknown groundwater pollutant sources [2]-[7]. Although the evolvement gave various platforms for further research but one of the major disadvantages was the non-satisfactory performance while solving the non-linear system. Later, Ref. [8] applied embedded technique for optimal source identification, but this technique is limited to only small-scale problems of groundwater management. In spite of the limitations, the embedded approach has been adopted by various researchers [9]-[13]. Taking into account the lacunae faced by the researchers in the field of groundwater management modeling, the linked simulation-optimization model originated. Ref. [14] formulated a non-linear linked simulation-optimization model and successfully identified the groundwater pollution source locations and release history using a new approach known as Progressive Genetic Algorithm (PGA). The results obtained were found to be encouraging even with large measurement errors. Ref. [15] used Genetic Algorithm for solving the multiple unknown pollutant sources. They linked the optimization model to an external flow and transport simulation model and eliminated the computational complexity in solving large embedded non-linear optimization models.

One of the major advantages of using linked simulation-optimization is that the simulator is externally linked to the optimization model. As such, any type of complex simulator also can be easily incorporated into the optimization model. Numerous researchers adopted various groundwater simulation models by linking with the optimization model. Ref. [16] adopted linked simulation-optimization approach for solving the groundwater pollutant source identification using the groundwater simulation model SUTRA. They have used the optimization solver MINOS for solving the problem. With the advancement of technologies, the solution of the complex physical processes of groundwater flow and transport can be performed by some of the standard simulation models like MODFLOW, MT3DMS, FEMWATER, HST3D, SEAWAT etc. Although the linked simulation-optimization model has the capability of successfully solving the groundwater problems but their performances are computationally expensive as there requires a large number of iterations to reach an optimal solution. One of the best techniques to deal with computational efficiency is the use of approximate simulation model. Artificial neural network (ANN) model is one of the most effective as well as popular models used for replacing the numerical aquifer simulation model. This has been confirmed by various groundwater management studies that introduction of ANN enhances the computational efficiency of the linked simulation-optimization model. The large scale complex problems such as pattern recognition, nonlinear modeling, classification, association, and control can be solved using ANN from given patterns [17]. Ref. [18] applied the ANN model on three types of groundwater and management problems and affirmed the efficiency after using ANN. Ref. [19] studied on the identification of groundwater pollution sources using ANN model and found that the performance of the ANN model was highly resistant even with large concentration measurement error. Later [20] solved ANN-GA based saltwater intrusion problems in a coastal aquifer and reported that the proposed model requires less computational time compared to embedded approach. Ref. [21] used ANN model to study the missing data scenario for pollutant source identification problem and showed that the model is capable of extracting the relation between the pollutant sources and their corresponding contaminant concentrations. Ref. [22] successfully solved the groundwater management model by developing ANN-PSO model. The objective of the model was to minimize the pumping cost of the well and the cost of the pipeline of Dore river basin in France. Ref. [23] adopted hybrid optimization technique for identification of groundwater pollution sources and found that the computational time can be reduced to a good extent when compared to an ANN based and GMS based optimization model.



**Table 1.** Hydrological parameters used in the study area.

Parameters	Values
Hydraulic conductivity in x direction, $K_x$ (m/s)	0.0002
Hydraulic conductivity in y direction, $K_y$ (m/s)	0.0002
Effective porosity, $\epsilon$	0.25
Time steps, $\Delta t$ (months)	3
Longitudinal dispersivity, $\alpha_L$ (m)	40
Transverse dispersivity, $\alpha_T$ (m)	9.6

**Table 2.** Source fluxes for different time steps (g/s).

Sources	Time Step 1	Time Step 2	Time Step 3	Time Step 4	Time Step 5
S1	908.42	1130.50	653.35	902.13	721.25
S2	644.02	1023.87	1139.88	781.09	889.77
S3	0	0	0	0	0
S4	0	1024.16	652.05	1117.45	889.77
S5	987.08	0	0	1104.82	639.93

**Table 3.** Pumping rates of the wells at the pumping well location of the aquifer.

Time Step	Pumping Rates (m <sup>3</sup> /day)	Time Step	Pumping Rates (m <sup>3</sup> /day)
1	327.024	11	272.52
2	163.512	12	218.016
3	218.016	13	327.024
4	318.528	14	163.512
5	109.008	15	381.528
6	327.024	16	217.72
7	272.520	17	272.520
8	163.512	18	218.010
9	381.528	19	327.024
10	109.008	20	272.520

### 3. Groundwater Source Identification Model

The groundwater pollution sources can be identified using inverse optimization technique. There are mainly two objective functions used in the optimization model. The first objective function minimizes the difference between the observed and the simulated concentrations at the observation locations. Whereas the second objective function will allow the model to select those observation wells whose concentrations are large and monitors effectively throughout the stress period. It can be further added that the second objective function will not select that observation wells which are located very far from the pollutant sources as the contaminant concentration observed in those wells will be negligible at different time steps. So, the combination of these two objectives will allow the model to identify the pollutant source effectively and the select only those optimal well locations which will monitor large contaminant concentration. The objective function can be written as:

$$\text{Minimize } f = w_1 \sum_{m=1}^P \sum_{n=1}^K |CO_m^n - CS_m^n|^2 \cdot \frac{1}{\left(1000 + \sum_{m=1}^P \sum_{n=1}^K CO_m^n\right)^2} \cdot z_m + w_2 \frac{D}{1 + \sum_{m=1}^P \sum_{n=1}^K CO_m^n} \tag{1}$$

Subject to the constraints

$$\sum_{m=1}^P z_m \leq M_{\max} \quad (2)$$

$$\sum_{m=1}^P Z_m \geq M_{\min} \quad (3)$$

where:  $CO_m^n$  and  $CS_m^n$  represent the observed and simulated concentrations at  $m^{\text{th}}$  observation well location at time period  $n$ ;  $z_m$  is the binary decision variable which indicates whether an observation well will be selected or not at the location “ $m$ ”.  $z_m$  equals to 1 if an observation well is installed and 0 otherwise.  $M_{\max}$  and  $M_{\min}$  represents the maximum and minimum permissible wells that are to be installed for the given period.  $P$  and  $K$  are the total no of well locations and the time period that are considered, respectively.  $w_1$  and  $w_2$  are weights for the first and second objective functions.  $D$  is a constant to be assigned by the user.

For obtaining the unknown contaminant sources in an aquifer, an aquifer simulator is required to be linked with the optimization model. The aquifer process can be simulated by solving the flow and transport equation using a numerical technique. As mentioned earlier, the numerical simulation model is computationally expensive. The computational time can be reduced using ANN based surrogate model. But developing an ANN model for each of the observation well will also be an issue of increasing computational time for a large study area. Therefore, the modified formulation presented above searches for an optimal number of observation wells which will reduce the computational burden of the model and eventually identify the pollutant sources efficiently. The ANN simulation will be performed only for those locations where the  $z$  value is equal to one. As mentioned earlier, the simulation has been performed for five years and the search was performed to find the optimal number of wells at the end of 1st year, 3rd year and 5th year. A total number of 30 potential observation wells were used for all the three time periods. From the initially placed potential observation wells, the objective function directs the optimization model to select the best match concentrations between the observed and simulated from all the considered locations and time steps. The constraint sets a boundary condition for the objective function to choose an optimal number of wells between 10 and 20 numbers of monitoring wells for the three time periods.

The additional binary decision variable introduced in the constraint function decides whether an observation well will be selected or not given by 0 or 1, respectively. It is to be noted that the placing of a large number of observation wells far from the pollutant sources might not identify the pollutant source efficiently because the location of the wells being far from the source will become redundant. Hence, the optimal number of well locations around the pollutant sources will efficiently identify the pollutant sources. The optimization model is solved using genetic algorithms as it has a binary variable.

In this evaluation process, the ANN models developed for each observation location will be repeatedly called by the optimization model. During each successive generation, the existing population undergoes the subsequent steps of selection, crossover and mutation to produce an advanced breed. This continues until the termination criteria of the simulation-optimization are not satisfied. The best offspring tends to give the best optimal solution of the inverse optimization problem. The genetic algorithm parameters used are shown in [Table 4](#). The schematic representation of the ANN-GA methodology is shown in [Figure 2](#).

## 4. Simulation Model

MODFLOW and MT3DMS models are used for simulating the groundwater flow and transport processes. In the present methodology, the data pattern required for training the ANN model is generated using the MT3DMS model.

### 4.1. Simulation of Groundwater Flow and Transport

The simulation of groundwater flow and transport in an aquifer can be solved using the two partial differential

**Table 4.** Genetic algorithm parameters used in the present methodology.

Parameter	Adopted Value	Function Parameter	Adopted
Population size	200	Scaling function	Rank
Generations	2000	Selection function	Stochastic uniform
Crossover fraction	0.8	Mutation function	Constraint dependent
Elite count	0.5	Crossover function	Scattered

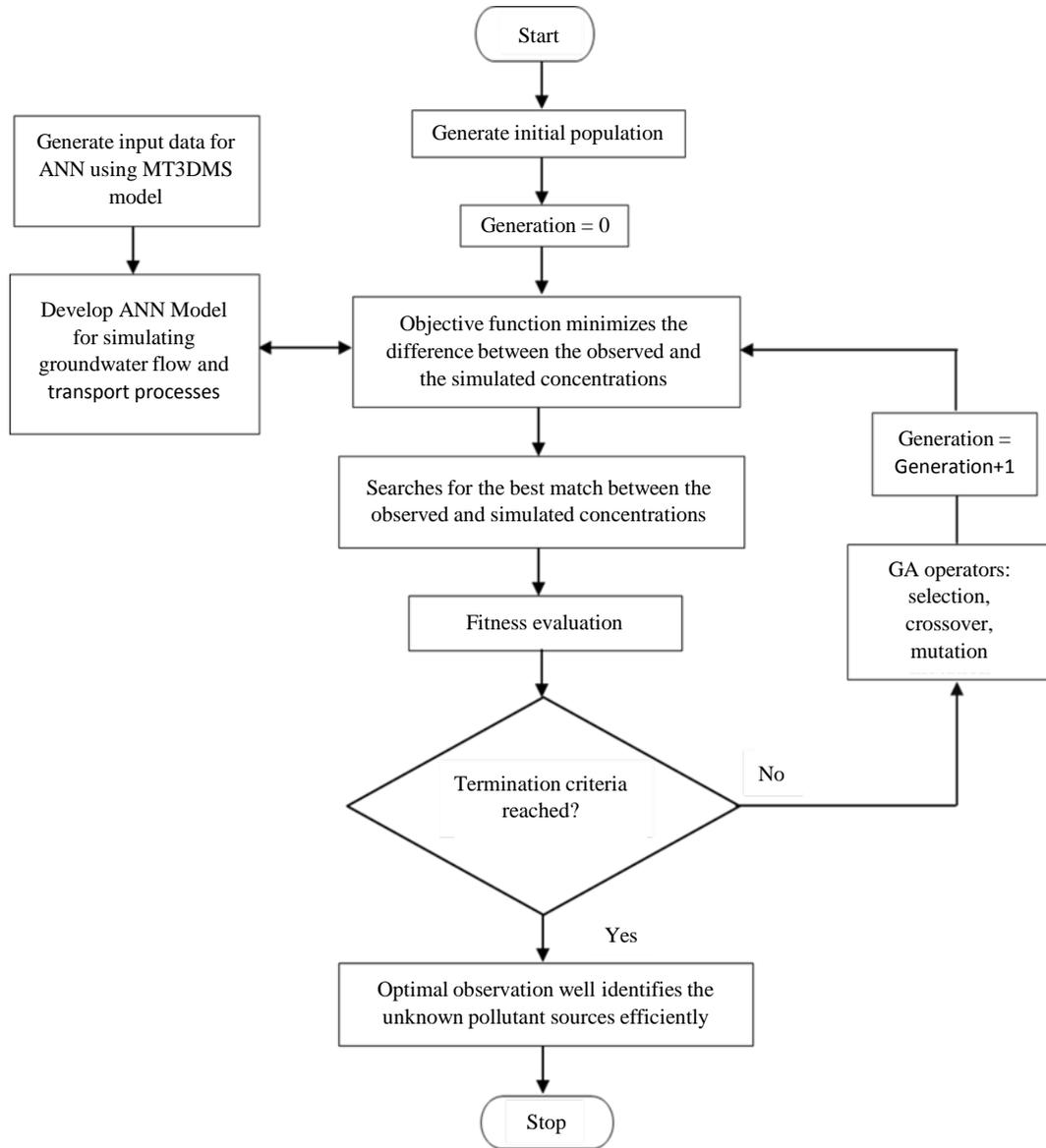


Figure 2. Schematic representation of the ANN-GA simulation-optimization model.

equations. The groundwater flow equation for two-dimensional flow in a confined homogeneous aquifer can be written according to [24] as:

$$\frac{\partial}{\partial x_i} \left( T_{ij} \frac{\partial h}{\partial x_j} \right) = S_s \frac{\partial h}{\partial t} + Q - W \quad i, j = 1, 2 \quad (4)$$

where:  $T_{ij} = K_{ij}b$  is the transmissivity tensor ( $L^2T^{-1}$ );  $K_{ij}$  is the hydraulic conductivity tensor ( $LT^{-1}$ );  $b$  is the saturated thickness of the aquifer (L);  $S_s$  is the storage coefficient;  $t$  is the time (T);  $h$  is the hydraulic head (L);  $Q$  is the pumping rate per unit area ( $LT^{-1}$ );  $W$  is the recharge flux per unit area ( $LT^{-1}$ ); and  $x_i$  and  $x_j$  are Cartesian co-ordinates.

The governing equation for contaminant mass transport in groundwater is given by [25] and can be written as:

$$\frac{\partial(cb)}{\partial t} = \frac{\partial}{\partial x_i} \left( bD_{ij} \frac{\partial c}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (bcv_i) - \frac{c'W}{\varepsilon} - \frac{c'Q}{\varepsilon} \quad i, j = 1, 2 \quad (5)$$

where:  $c$  is the concentration of dissolved chemical species ( $\text{ML}^{-3}$ );  $b$  is the saturated thickness of the aquifer;  $D_{ij}$  is the coefficient of hydrodynamic dispersion ( $\text{L}^2\text{T}^{-1}$ );  $c'$  is the concentration of the dissolved chemical in source or sink fluid ( $\text{ML}^{-3}$ );  $v_i$  is the velocity component ( $\text{LT}^{-1}$ );  $W$  is the recharge flux per unit area ( $\text{LT}^{-1}$ );  $Q$  is the pumping rate per unit area ( $\text{LT}^{-1}$ );  $\varepsilon$  is the effective porosity of the aquifer.

## 4.2. Development of ANN Model as Approximate Simulator

Ref. [26] developed the first simple neural network and now ANN is considered as the best approximate model due to its ability to solve a problem by storing the information of large input and output pattern. The basic processing unit of an ANN model is the neuron which resembles the biological nervous system of the human brain. The ANN model is developed with three neuron layers. These are input, output, and hidden layers. In this whole process of learning, the input layer receives the input data and the synaptic weight present between the neurons stores the knowledge [27] by following the pattern of the data set. Once the ANN model is trained according to the data pattern, subsequently it can be used to predict the output.

## 4.3. Architecture and Generation of ANN Pattern

For the present methodology, single hidden layer architecture is adopted. There are no definite rules for selecting the number of hidden layers and the number of neurons in the architecture, it is determined on the basis of trial and error. The ANN model is trained using Levenberg-Marquardt algorithm. A unipolar sigmoidal transfer function and a purely linear transfer function are used for the neurons in the hidden layer and in the output layer of the network respectively. There are number of important ANN parameters which must be selected through comparative study. The ANN parameters are the learning rate and the momentum rate. Once the combination of the ANN parameters are decided for designing an ANN model it can be further used in the training of ANN model. The learning rate signifies the rate at which the weight is changes in the training phase. A fast learning rate will require less computational time to train the network. Whereas a slow learning rate slows down the training procedure and would require more computational time [28]. To overcome this, the momentum rate is introduced which will improve the learning rate by further adjusting the weight vector and would ultimately give an upgraded solution. The input data to the ANN model are the groundwater pollutant source fluxes which are active for five-time steps at an interval of three months. The corresponding contaminant concentrations at the specified observation location will constitute the output pattern for the ANN model. A total number of 2000 input-output patterns were generated using the MT3DMS model. The input and output files for the MODFLOW and MT3DMS model has been developed using Groundwater Modeling System (GMS) platform. Once the files are generated, the models are run using Matlab [23]. From the total number of generated data, 60% is used for training the ANN model which will subsequently learn the input pattern producing the desired outputs. The remaining 40% patterns are further used for testing and validating the ANN model.

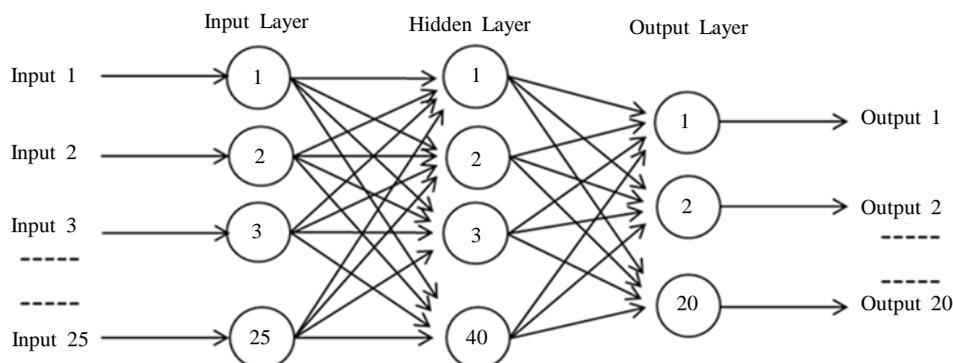
## 4.4. Development of the Artificial Neural Network

There are five pollutant sources in the study area and the simulation is performed for a period of five years at an interval of ninety days. As such a total number of 25 pollutant source fluxes will be used as the input to the ANN model. The pollutant sources are active for five-time steps. The output from the ANN model is the concentration at the observation locations for all the five years which will be equal to 20. There are 30 observation wells in the aquifer. As such, 30 models have been developed. Then each of these models will be used in predicting the contaminant concentrations at the well locations. A total number of three layers are used with one input layer, one hidden layer, and one output layer. The developed architecture for the ANN model can be represented as 25-40-20 which has been shown in Figure 3. The self-teaching ability of the ANN model depicts to be the highly effective approach for the inverse optimization problem. If the training algorithms are selected effectively, then the developed model tends to be a robust one.

## 5. Results and Discussion

### 5.1. Performance of the ANN Model

The ANN models developed for the thirty observation wells of the illustrative study area are incorporated with the optimization model for identification of pollutant sources and the optimal location of the monitoring wells.



**Figure 3.** A three layer neural network architecture.

**Figure 4** shows the performance of the ANN model as an approximated simulator of the groundwater transport processes. The scatter plot shows that there is a very high correlation between the actual and the predicted normalized contaminant concentrations. The actual concentrations are simulated using the groundwater transport model MT3DMS whereas the predicted concentrations are obtained using the ANN model. From the value of the  $R^2$ , it can be observed that the performance of the ANN model as an approximate simulator is quite good.

## 5.2. Identification of the Contaminant Sources

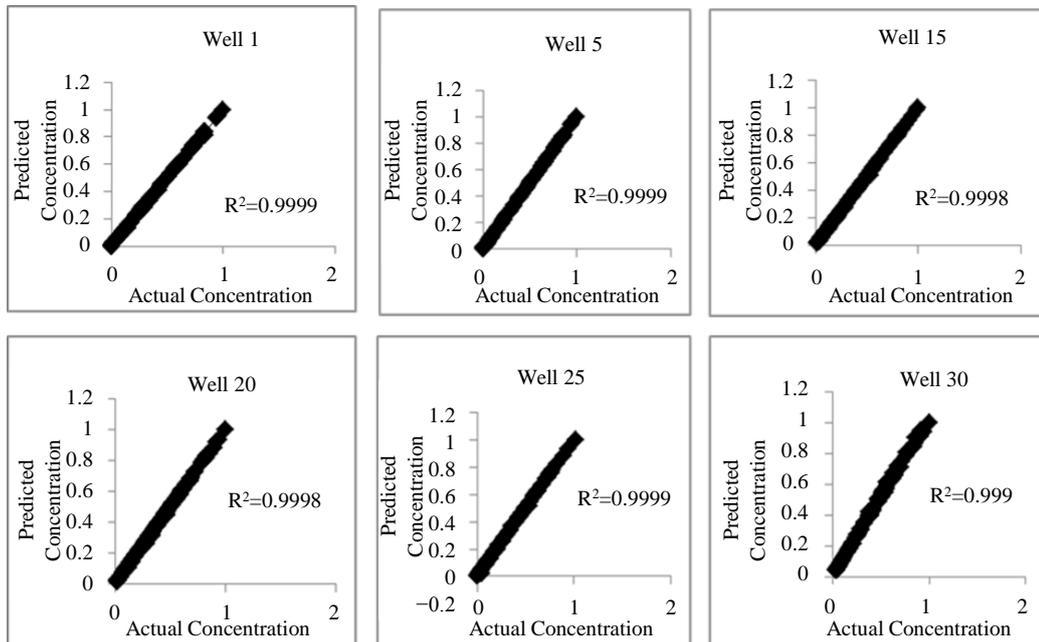
As discussed above, the developed ANN model is linked with the optimization model. The optimization problem is solved using Genetic Algorithms. In every generation of Genetic Algorithms, the ANN models have to be run to calculate the simulated concentration at observation locations. Although there are thirty ANN models need to be run in every generation of Genetic Algorithms, it will actually run the ANN model of only those locations where the  $z$  value is one. Thus, the number of ANN model run in every generation is always less than the maximum number of observation wells. The results obtained using the proposed model is compared between the actual and the predicted pollutant sources as shown in **Figure 5**.

**Table 5** shows the estimated pollutant sources and is compared with the actual source values of the study area by calculating the absolute relative error. The pollutant sources are active for five-time steps *i.e.* 15 months only. There are altogether four active sources and one dummy source. It is observed that the sources (S1, S2, S3, S4 and S5) are predicted very efficiently. Even though the source pollutant S3 is a dummy source, some source concentration has been identified due to the effect from other pollutant sources. Therefore, with the increase of number of sources and the overlapping of the pollutant plume, there will be less probability of the identification of the sources.

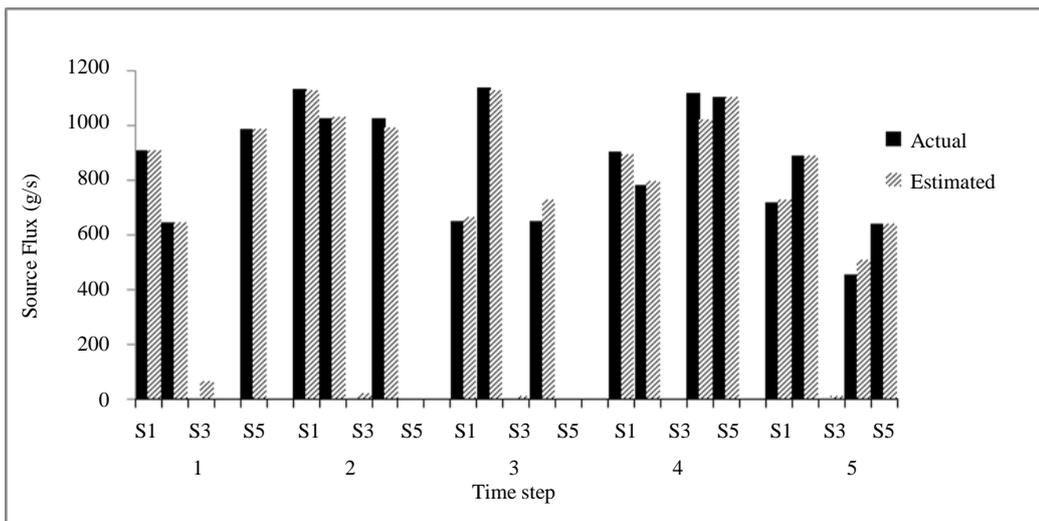
## 5.3. Optimal Location of Observation Wells for Different Time Period

Even though the movement of the pollutant sources in the groundwater is considered to be a slow process but considering the dynamic nature of the pollutant sources, the location for optimal number of observation wells will be selected at the end of first year, third year and fifth year. The model selects optimal number of wells for each time period from a total number of 30 observation wells locations. In the first year, the optimal well locations selected by the model are close to the source locations. It can be seen that the location of the optimal well follows the path of the plume and a total number of 15 wells are selected ranging between 10 and 20 which are the maximum and the minimum number of optimal wells allowed by the model (**Table 6**). A total number of five optimal wells are detected close to the source S5 (with  $I = 28, j = 53, k = 1$ ). The model selects a sum of six optimal wells around the sources S4 ( $i = 60, j = 48, k = 1$ ) and S3 ( $i = 71, j = 59, k = 1$ ). However, it is observed that S2 ( $i = 73, j = 48, k = 1$ ) being a dummy source no optimal wells were selected around it by the model. The model manages to select four optimal wells for the pollutant source S1 ( $i = 8, j = 30, k = 1$ ). The source locations along with the optimal observation well locations at the end of the first time period are shown in **Figure 6**.

The observation well locations for the second and third time periods can be seen in **Figure 7** and **Figure 8** respectively. The grid location of the optimal wells for the second and third time period is given in **Table 7** and **Table 8** respectively. It is seen that some of the wells fall far away from the outer contour of the contaminant



**Figure 4.** Scatter plot of actual and predicted contaminant concentration for five years.



**Figure 5.** Comparison between the actual fluxes and the estimated source fluxes for five stress period.

plume. This is because the optimal wells detected will monitor for two long years and the concentration of the wells being dynamic in nature compels the wells to change their location with time. Hence, with the passage of time, the optimal wells are moving along the direction of the plume. It is also observed that the total number of optimal wells selected by the model is 10 for both the time periods.

Most of the selected optimal observation wells for all the three time period are found to be very close to the pollutant sources. It is due to the fact that the pollutant sources are active for five-time steps only. In the remaining time steps, the pollutant source becomes inactive resulting in the decrease of the pollutant concentration for the remaining time steps. It can also be noted that the observation well (W 30) located very far from the pollutant sources is not selected by the model as the well is not within the reach of the contaminant concentration. As mentioned earlier, ANN model is developed for all the thirty observation wells and each of these selected optimal wells represents those ANN models which perform as the best approximate simulator. The ANN-GA model took only 6 hours of CPU time to compute the simulation-optimization problem on a 2.27 GHz processor.

**Table 5.** Absolute relative error between actual and estimated sources.

Time Steps	Source Locations	Actual Sources (g/s)	Estimated Sources (g/s)	Absolute Relative Error (%)
1	S1	908.42	909.54	-0.12
	S2	644.02	643.20	0.12
	S3	0	66.6	-
	S4	0	7.73	-
	S5	987.08	986.203	0.08
2	S1	1130.5	1125.20	0.46
	S2	1023.87	1029.73	-0.57
	S3	0	22.16	-
	S4	1024.16	988.78	3.45
	S5	0	0.38	-
3	S1	653.35	663.18	-1.50
	S2	1139.88	1125.71	1.24
	S3	0	12.34	-
	S4	652.05	728.73	-11.76
	S5	0	0.48	-
4	S1	902.15	892.55	1.06
	S2	781.09	796.56	-1.98
	S3	0	4.73	-
	S4	1117.45	1017.66	8.92
	S5	1104.82	1101.45	0.30
5	S1	721.24	726.87	0.77
	S2	889.77	886.64	0.35
	S3	0	10.23	-
	S4	457.91	509.22	-11.20
	S5	639.93	642.24	-0.36

**Table 6.** Results for optimal observation well locations at the end of first year.

Time period	Number of selected optimal wells	Location of the optimal observation wells (i, j, k)
1	15	W1 (30, 56, 1) W3 (32, 53, 1) W4 (63, 52, 1) W5 (66, 49, 1) W6 (70, 61, 1) W9 (72, 49, 1) W10 (75, 50, 1) W14 (82, 53, 1) W16 (86, 32, 1) W19 (87, 32, 1) W21 (89, 31, 1) W22 (86, 32, 1) W24 (87, 32, 1) W25 (89, 31, 1) W26 (92, 32, 1)

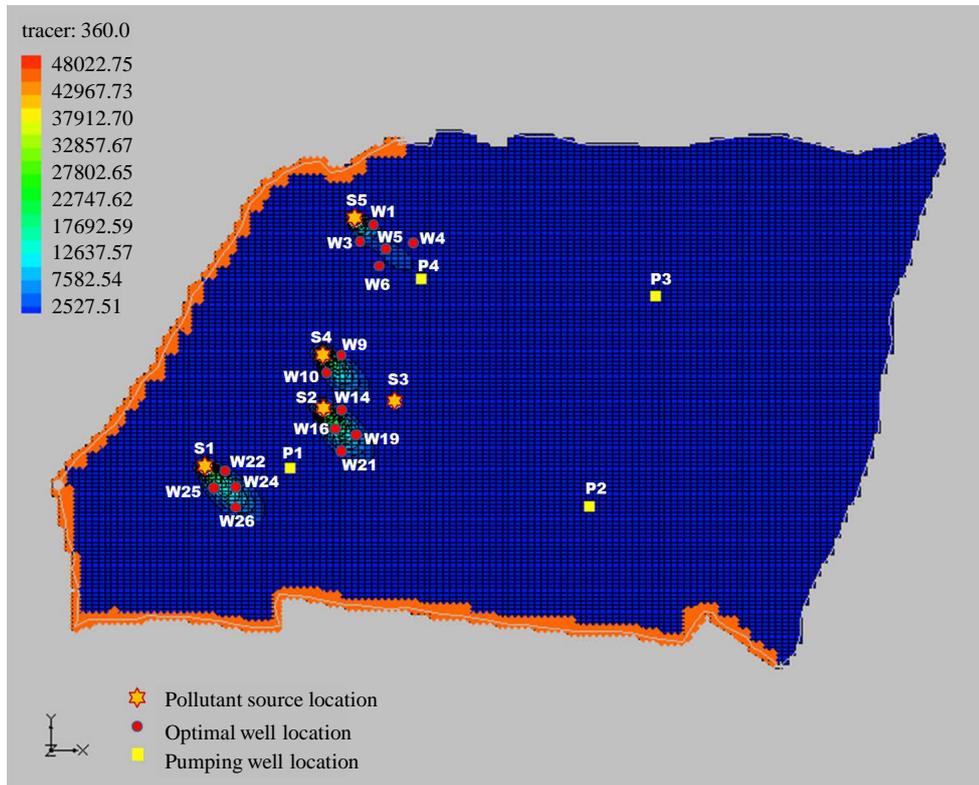


Figure 6. Study area showing optimal observation well location at the end of first year.

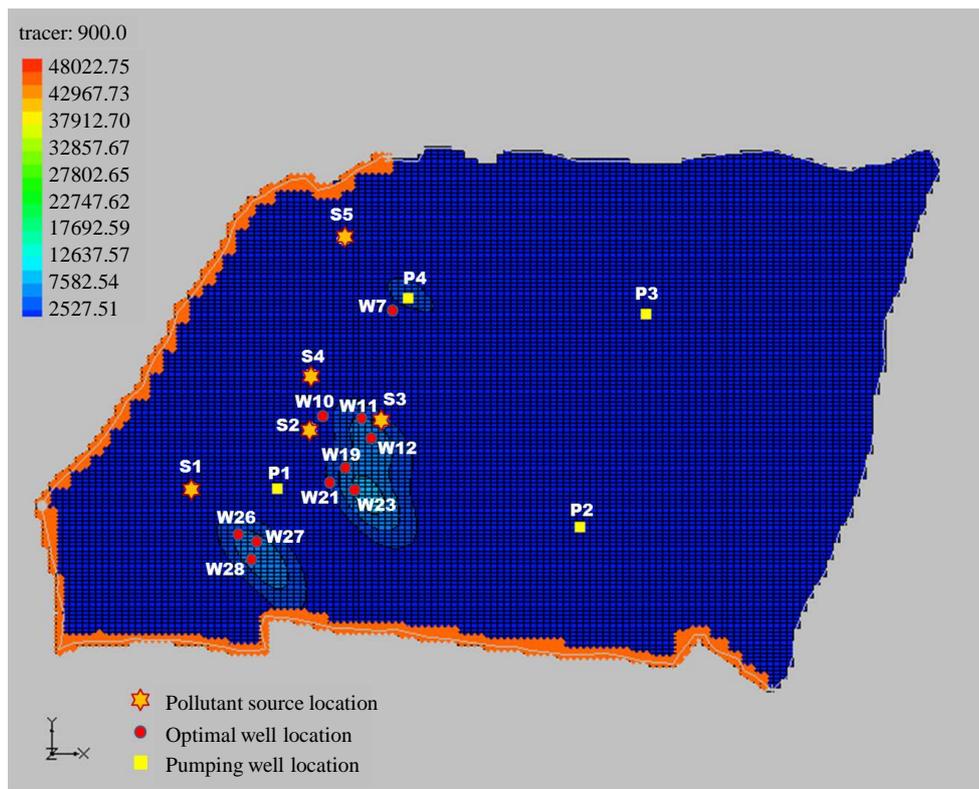
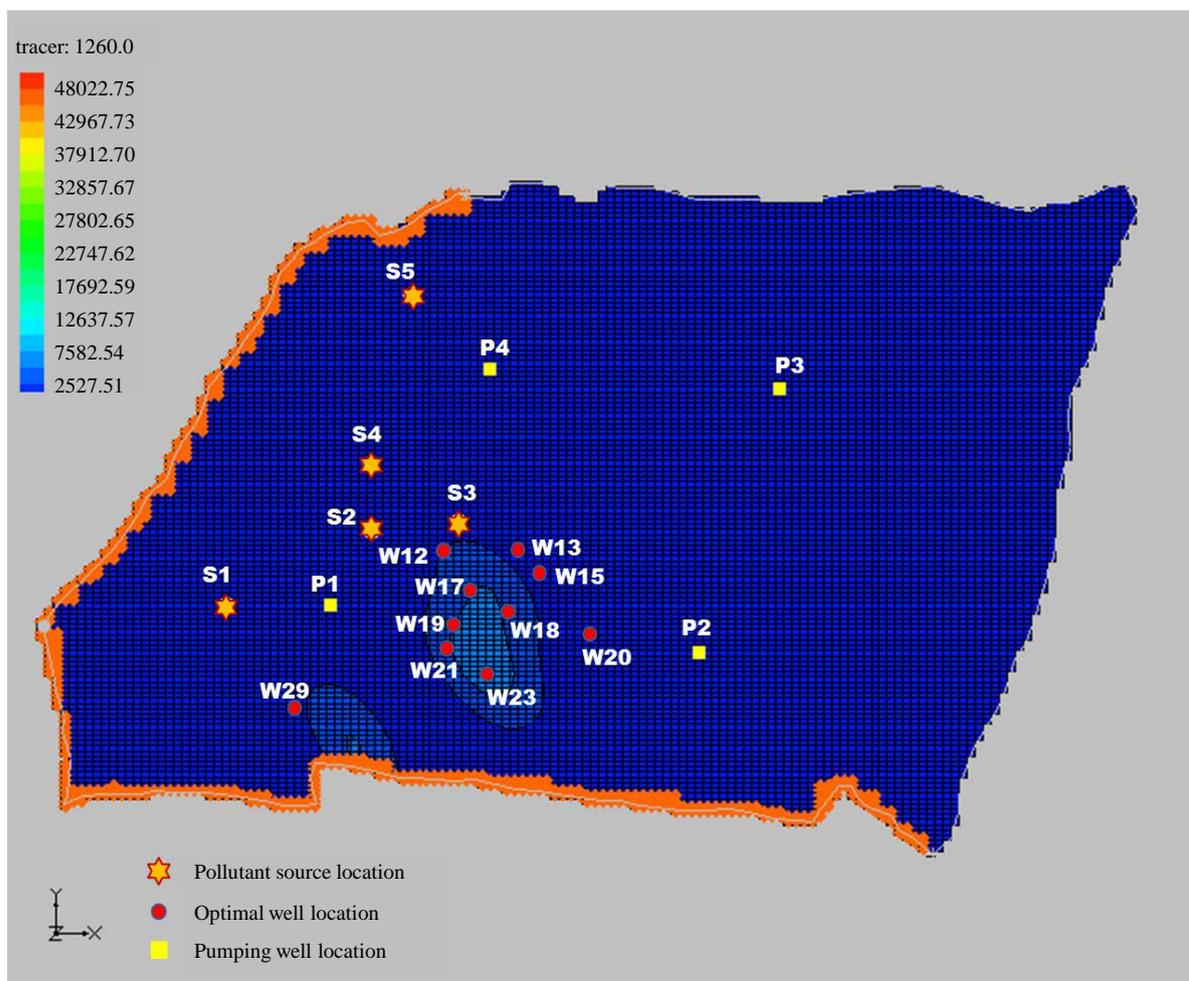


Figure 7. Study area showing optimal observation well location at the end of third year.

**Table 7.** Results for optimal observation well locations at the end of third year.

Time period	Number of selected optimal wells	Location of the optimal observation wells $(i, j, k)$
2	10	W7 (38, 62, 1) W10 (66, 49, 1) W11 (68, 53, 1) W12 (69, 57, 1) W19 (82, 53, 1) W21 (82, 50, 1) W23 (86, 61, 1) W26 (92, 32, 1) W27 (93, 36, 1) W28 (99, 36, 1)



**Figure 8.** Study area showing optimal observation well locations at the end of fifth year.

## 6. Conclusion

The present methodology has adopted linked simulation-optimization approach for efficient identification of pollutant sources in an aquifer. For reducing the computational time of the groundwater flow and transport processes, ANN model has been used as an approximate simulator. The developed methodology has been

**Table 8.** Results for optimal observation well locations at the end of fifth year.

Time period	Number of selected optimal wells	Location of the optimal observation wells ( <i>i, j, k</i> )
3	10	W12 (69, 57, 1) W13 (70, 61, 1) W15 (74, 63, 1) W17 (75, 59, 1) W18 (80, 61, 1) W19 (82, 53, 1) W20 (82, 66, 1) W21 (83, 50, 1) W23 (86, 61, 1) W29 (103, 36, 1)

applied in a large study area of approximately 17 km<sup>2</sup>. A total number of 30 observation wells are used for the present study area. Thirty ANN models were developed for each of the thirty observation wells as performance of the simulation model is not satisfactory when a single ANN model is used to predict the concentration of all the observation wells. However, simulations performed by all these ANN models are also taking considerable CPU time. Therefore, the present study has focused on the selection of optimal observation wells which will ultimately reduce the computational time. The solution results show that the location of the optimal wells changes from one time period to another and shows the dynamic state of the network. The wells were selected by the model considering the budgetary constraints. These optimal wells detect the pollutant concentration very efficiently, matching the estimated fluxes with the actual fluxes. The relative error for the estimated source fluxes is found to be very negligible when compared with actual sources. From the performance evaluation of the results, it is seen that the ANN model can efficiently use as an approximate simulator for a large study area. However, detailed and hydrological parameters of a real large aquifer are required for evaluating the performance of the present methodology.

## References

- [1] Gorelick, S.M., Evans, B. and Remson, I. (1983) Identifying Sources of Groundwater Pollution: An Optimization Approach. *Water Resources Research*, **19**, 779-790. <http://dx.doi.org/10.1029/WR019i003p00779>
- [2] Heidari, M. (1982) Application of Linear Systems Theory and Linear Programming to Groundwater Management in Kansas. *Water Resources Research*, **18**, 1003-1012.
- [3] Willis, R. and Liu, P. (1984) Optimization Model for Groundwater Planning. *Water Resources Planning and Management*, **110**, 333-347. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1984\)110:3\(333\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1984)110:3(333))
- [4] Yazicigil, H., Al-Layla, R. and De Jong, R.L. (1987) Optimal Management of a Regional Aquifer in Eastern Saudi Arabia. *Water Resources Bulletin*, **23**, 423-434. <http://dx.doi.org/10.1111/j.1752-1688.1987.tb00820.x>
- [5] Skaggs, T.H. and Kabala, Z.J. (1995) Recovering the Release History of a Groundwater Contaminant Plume: Method of Quasi-Reversibility. *Water Resources Research*, **3**, 2669-2673. <http://dx.doi.org/10.1029/95WR02383>
- [6] Datta, B. and Dhiman, S.D. (1996) Chance-Constrained Optimal Monitoring Network Design for Pollutants in Groundwater. *Journal of Water Resources Planning Management*, **122**, 180-188. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1996\)122:3\(180\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1996)122:3(180))
- [7] Aral, M.M. and Guan, J. (1996) Genetic Algorithms in Search of Groundwater and Pollutant Sources. *Advances in Groundwater Pollution Control and Remediation*, **9**, 347-369. [http://dx.doi.org/10.1007/978-94-009-0205-3\\_17](http://dx.doi.org/10.1007/978-94-009-0205-3_17)
- [8] Mahar, P.S. and Datta, B. (1997) Optimal Monitoring Network and Groundwater Pollution Source Identification. *Journal of Water Resources Planning and Management*, **123**, 199-207. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1997\)123:4\(199\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1997)123:4(199))
- [9] Aguado, E. and Remson, I. (1974) Groundwater Hydraulics in Aquifer Management. *Journal of the Hydraulics Division*, **100**, 103-118.
- [10] Willis, R. and Neuman, B.A. (1977) Management Model for Groundwater Development. *Journal of Water Resources Planning and Management*, **103**, 159-171.

- [11] Peralta, R.C. and Datta, B. (1990) Reconnaissance-Level Alternative Optimal Groundwater Use Strategies. *Water Resources Planning and Management*, **116**, 676-692. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(1990\)116:5\(676\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(1990)116:5(676))
- [12] Mahar, P.S. and Datta, B. (2001) Optimal Identification of Ground-Water Pollution Sources and Parameter Estimation. *Journal of Water Resources Planning and Management*, **127**, 20-29. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(2001\)127:1\(20\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(2001)127:1(20))
- [13] McPhee, J. and Yeh, W.W.G. (2008) Groundwater Management Using Model Reduction via Empirical Orthogonal Functions. *Water Resources Planning and Management*, **134**, 161-170. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(2008\)134:2\(161\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(2008)134:2(161))
- [14] Aral, M.M., Guan, J. and Maslia, M.L. (2001) Identification of Contaminant Source Location and Release History in Aquifers. *Journal of Hydrologic Engineering*, **6**, 225-234. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2001\)6:3\(225\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2001)6:3(225))
- [15] Singh, R. and Datta, B. (2006) Identification of Groundwater Pollution Sources Using GA-Based Linked Simulation Optimization Model. *Journal of Hydrological Engineering*, **11**, 101-109. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2006\)11:2\(101\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2006)11:2(101))
- [16] Datta, B., Chakrabarty, D. and Dhar, A. (2009) Simultaneous Identification of Unknown Groundwater Pollution Sources and Estimation of Aquifer Parameters. *Journal of Hydrology*, **376**, 48-57. <http://dx.doi.org/10.1016/j.jhydrol.2009.07.014>
- [17] American Society of Civil Engineers Task Committee on Application of Artificial Neural Networks in Hydrology (2000) Artificial Neural Networks in Hydrology. I: Preliminary Concepts. *Journal of Hydrologic Engineering*, **5**, 115-123. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2000\)5:2\(115\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2000)5:2(115))
- [18] Copola, E., Poulton, M., Charles, E., Dustman, J. and Szidarovszky, F. (2003) Application of Artificial Neural Networks to Complex Groundwater Management Problems. *Natural Resources Researches*, **12**, 303-320. <http://dx.doi.org/10.1023/B:NARR.0000007808.11860.7e>
- [19] Singh, R.M., Datta, B. and Jain, A. (2004) Identification of Unknown Groundwater Pollution Sources Using Artificial Neural Networks. *Water Resources Planning and Management*, **130**, 506-514. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(2004\)130:6\(506\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(2004)130:6(506))
- [20] Bhattacharjya, R.K. and Datta, B. (2009) ANN-GA-Based Model for Multiple Objective Management of Coastal Aquifers. *Journal of Water Resources Planning and Management*, **135**, 314-322. [http://dx.doi.org/10.1061/\(ASCE\)0733-9496\(2009\)135:5\(314\)](http://dx.doi.org/10.1061/(ASCE)0733-9496(2009)135:5(314))
- [21] Singh, R.M. and Datta, B. (2007) Artificial Neural Network Modelling for Identification of Unknown Pollution Sources in Groundwater with Partially Missing Concentration Observation Data. *Water Resources Management*, **21**, 557-572. <http://dx.doi.org/10.1007/s11269-006-9029-z>
- [22] Gaur, S., Sudheer, C., Graillet, D., Chahar, B.R. and Kumar, D.N. (2013) Application of Artificial Neural Network and Particle Swarm Optimization for the Management of Groundwater Resources. *Water Resources Management*, **27**, 927-941. <http://dx.doi.org/10.1007/s11269-012-0226-7>
- [23] Borah, T. and Bhattacharjya, R.K. (2014) Development of Unknown Pollution Source Identification Models Using GMS ANN-Based Simulation Optimization Methodology. *Journal of Hazardous, Toxic and Radioactive Waste*, **19**, 1-12. [http://dx.doi.org/10.1061/\(ASCE\)HZ.2153-5515.0000242](http://dx.doi.org/10.1061/(ASCE)HZ.2153-5515.0000242)
- [24] Bear, J. (1972) Dynamics of Fluid in Porous Media. American Elsevier Publishing Company, New York.
- [25] Bredehoeft, J.D. and Pinder, J.F. (1973) Mass Transport in Flowing Groundwater. *Water Resources Research*, **12**, 194-210. <http://dx.doi.org/10.1029/WR009i001p00194>
- [26] McCulloch, W.S. and Pitts, W. (1943) A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, **5**, 115-133. <http://dx.doi.org/10.1007/BF02478259>
- [27] Haykin, S. (1994) Neural Networks: A Comprehensive Foundation. Prentice Hall PTR, Upper Saddle River.
- [28] Bhattacharjya, R., Datta, B. and Satish, M. (2007) Artificial Neural Networks Approximation of Density Dependent Saltwater Intrusion Process in Coastal Aquifers. *Journal of Hydrologic Engineering*, **12**, 273-282. [http://dx.doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:3\(273\)](http://dx.doi.org/10.1061/(ASCE)1084-0699(2007)12:3(273))