International Journal of Civil and Structural Engineering– IJCSE
Volume 3 : Issue 1 [ISSN : 2372-3971]
Publication Date: 18 April, 2016

Determination of illegal pumping and monitoring network using genetic algorithm based simulation-optimization model

[1]Arnesh Das, [2]Rajeev Gandhi BG,[2] Rajib Kumar Bhattacharjya[2]

[1]Delhi Technological University, [2] Indian Institute of Technology Guwahati

ABSTRACT

Groundwater is an important source of freshwater, more so in areas away from the surface water sources. Due to the substantial growth in industry and agriculture as well as the increased use of municipal water, the demand of groundwater has been increasing continuously in many parts of the world. This has depleted the groundwater table in many parts of the country as well as in other parts of the world. As a result of over exploitation of groundwater, the quality of groundwater is also deteriorating rapidly. Hence there is a need to monitor the exploitation of groundwater in an aquifer vulnerable from quality and quantity aspects. The overexploitation or the illegal pumping of groundwater can be assessed by using inverse optimization techniques. In this study, an inverse optimization model is proposed to identify the illegal pumping locations and pumping rates. The performance of the model is highly related to the location and number of monitoring wells used in the model. As such, a modified formulation is also used to design an optimal monitoring network. We used Genetic algorithms to solve the inverse optimization model. For obtaining physically meaningful solution, the groundwater simulation model needs to incorporate with the optimization model. The simulation model simulates the physical processes in the groundwater aquifer by solving the governing groundwater flow equation. We solve the groundwater flow equation using finite difference approach. The model is then linked with the optimization model to determine the location and pumping schedule of the wells. The applicability of the proposed methodology is evaluated using a hypothetical study area involving a two dimensional aquifer. The evaluation shows that this methodology can be used for solving practical problems in the real world.

Keywords: aquifer, genetic algorithm, groundwater, inverse optimization, simulation

I.INTRODUCTION

Groundwater is an important source of water, more so in areas where there are no freshwater sources nearby. However, due to rapid population growth, industrialization and urbanization, groundwater has been overexploited. This has led to the depletion of the groundwater table in many parts of the world. Besides the quantity, the quality of groundwater has also been deteriorating rapidly. Often it happens that due to excess exploitation of groundwater in a particular area, there is not much usable water left for people in nearby areas. It is necessary that such a problem is addressed at the earliest so that water can be equally used by all sections of people in the locality. Hence, it is necessary that there are proper sustainable planning and management strategies in place for the optimal and efficient operation of groundwater systems. This can be done by using inverse optimization model. In this approach, the groundwater simulation model is combined with the optimization model for detecting locations of unknown well sources and their pumping rates. The simulation model simulates the physical processes in a groundwater aquifer by solving the governing groundwater flow equation. The model is then linked with the optimization model to determine the location and pumping schedule of illegal wells.

Similar works have been done in the past in different parts of the world. Saffi and Cheddadi, 2010 [1] identified illegal groundwater pumping in semi confined aquifers by minimising the error between the observed and simulated head values. Tung and Chou, 2004 [2] tried to identify the spatial distribution of groundwater pumping using optimization techniques.. There are numerous studies associated with the solution of well locations and pumping rates identification using genetic algorithm based simulation/ optimization models. Ayvaz and Karahan, 2008 [3] studied the identification of unknown well locations and their pumping rates in two dimensional aquifers. They proposed a simulation model consisting of finite difference solution of governing groundwater flow equation that was then combined with a genetic algorithm based optimization model. Huang and Mayer, 1997 [5] used genetic algorithm for obtaining optimal solution of remediation system design by selecting well locations as discrete decision variables. Mahinthakumar and Sayeed, 2005 [6] solved the source identification problem by combining genetic algorithm with a local search algorithm. In their research, genetic algorithm simultaneously identified the location and concentration of single source and then a local search algorithm was used to fine-tune the genetic algorithm solution. A point to be noted is that genetic algorithm finds the optimal solution by considering both well locations and pumping. Datta, Chakrabarty and Dhar, 2010 [7] proposed a methodology that involved using a classical non linear optimization model that was linked to a flow and transport simulation model for identifying unknown groundwater pollution sources. This proposed methodology could overcome many of the limitations of other methods given by early researchers. Aral, Guan and Maslia, 2001 [9] used an optimization model for identification of contaminant source locations. It was attempted to minimise the difference between the simulated and observed concentrations at the observation sites. Borah and Bhattacharjya, 2013 [10] used Groundwater Modelling System (GMS) in combination with MATLAB based optimization method for solving groundwater source identification problem. The simulation of flow and transport processes was done in GMS and the optimization model minimised the difference between the observed and simulated concentrations. Singh, Datta and Jain, 2004 [11] made use of artificial neural networks for identification of unknown groundwater pollution sources. The training of artificial neural network involved simulated concentration measurement data at specified observation sites in the aquifer for identification of sources.

In the present study, an attempt has been made to determine the locations of wells and their pumping rate values using a simulation and optimization model. A two dimensional aquifer is considered. The aquifer is considered as homogeneous and isotropic. Initially, with the locations and pumping values of the wells known, a simulation model is run. This would give us the head values (observed head values) at the different points of the grid considered. Next, another simulation model is run but this time the locations as well as the pumping values of the wells are not known. This would give us the simulated head values at different points of the grid. These simulation models are then combined with an optimization model where the difference between the observed head values and the simulated head values is minimised. Basically, an objective function is considered and it is attempted to minimise the value of this function.

II. METHODOLOGY

A. Problem Definition

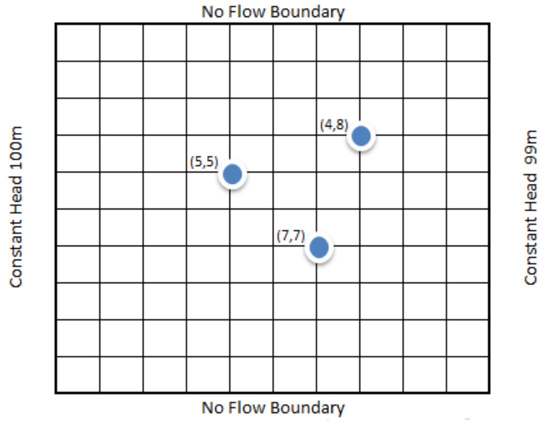

Fig 1: Illustrative study area showing the location of pumping wells and boundary conditions

Fig 1 shows a two dimensional aquifer of length and breadth each equal to 500 metres. The aquifer considered is homogeneous and isotropic. The hydraulic conductivity and storativity value of the aquifer are 15 m/day and 0.02 respectively. The locations of the wells are specified. The coordinates of the locations of wells are (4,8), (5,5) and (7,7). The simulation is run for five days and is being run in MATLAB.The pumping values of the wells on different days are as specified in Table I. These values are the actual pumping values and we have assumed them to generate observation data.

Table I: Day by day pumping schedules at all the three pumping well locations

DAY	PUMPING VALUE (m^3/day) WELL 1	PUMPING VALUE (m^3/day) WELL 2	PUMPING VALUE (m^3/day) WELL 3
1	500	800	700
2	550	850	0
3	650	0	0
4	0	0	800
5	300	750	850

B. Method Used

As has been mentioned before, an objective function is considered and it is attempted to minimise this function. The objective function (F) is given by

$$F = \sum_{k=1}^{k=121} \sum_{t=1}^{t=5} w_k \left(H_{ok}^{t} - H_{sk}^{t} \right)^2 \qquad (1)$$

In the above equation, summation is run from k=1 to k=121 i.e for all the 121 points of the 11X11 line grid. Also, since the head values were computed at five time steps, therefore summation is run from t=1 to t=5. In this equation, ‘ w_k ’ represents a binary function which gives the location of the wells. Its value can be either 0 or 1 depending on whether a well is present at that point. The w_k value is one at the point where a well is present and at other points it would be zero. It has also been assumed that there are at least 30 wells in the aquifer and the maximum number of wells does not exceed 70. H_{ok}^{t} And H_{sk}^{t} represent the observed head and simulated head values of the different points of the 2D aquifer grid and at different timesteps. These are obtained by solving the groundwater flow equation. The groundwater flow equation is solved using fully implicit finite difference procedure and coding done in MATLAB.

The groundwater flow equation for a homogeneous isotropic two dimensional aquifer is given by

$$K \frac{\partial^2 \phi}{\partial x^2} + K \frac{\partial^2 \phi}{\partial y^2} + q_s = S_s \frac{\partial \phi}{\partial t} \qquad (2)$$

Where K is the hydraulic conductivity of the aquifer (LT^{-1}); q_s is the source or sink value in the aquifer which includes the pumping or injection well’s pumping rate ($L^3T^{-1}L^{-3}$); S_s is the specific storativity which is the ratio of aquifer storativity and thickness of aquifer (L^{-1}); ϕ is the piezometric head in the aquifer; x, y and t are the coordinates of space and time.

Using Taylor’s series for solving this equation and after simplification and rearrangement, the finite difference approximation of the 2D flow equation at any point (i,j) can be expressed as

$$A_1 \phi^{n+1}{}_{i,j} + A_2 \phi^{n+1}{}_{i+1,j} + A_3 \phi^{n+1}{}_{i-1,j} + A_4 \phi^{n+1}{}_{i,j+1} + A_5 \phi^{n+1}{}_{i,j-1} = B \phi^{n}{}_{i,j} \qquad (3)$$

Here

$A_1 = -1-2K_1-2K_2$; $A_2 = K_1$; $A_3 = K_1$; $A_4 = K_2$; $A_5 = K_2$; $B = -1$

Where $K_1 = \dfrac{K(\Delta t)}{S_S (\Delta x)^2}$ and $K_2 = \dfrac{K(\Delta t)}{S_S (\Delta y)^2}$

The values at n^{th} time step are considered to be known values and those at $n+1^{th}$ time step are unknown. There are a set of boundary conditions, namely constant head boundary and no flow boundary. By assuming the initial conditions in the aquifer the iterations are started by applying the equation at each space point i,j. After eliminating 22 points from the total 121 points in the aquifer due to the boundary conditions, there are 99 unknown grid points and the final set of simultaneous equations to be solved are obtained. The simultaneous equations are solved by solving matrix equation $A\phi^{n+1}=B\phi^{n}$ by matrix inversion method giving the values of piezometric head towards the next time step. These are to be solved simultaneously by taking the updated values of piezometric heads as inputs for successive time steps basing on the number of days the simulation need to run.

III. RESULTS AND DISCUSSIONS

An optimization problem usually involves the maximisation or minimisation of a real function with the help of chosen input values and then computing the value of the function. The present study involves the minimisation of the objective function considered. It was carried out using ‘optimtool’ in MATLAB. In the problem setup of optimization tool, genetic algorithm was selected as the solver. In the present problem, there are a total of 117 variables. For the first 18 variables which represent the locations and pumping values of the well, the lower bounds for the first, seventh and thirteenth variables are 1 while for the rest of the fifteen variables, it is 0. The first, seventh and thirteenth variables represent the locations while the rest represent the pumping values of the wells. The location points are generally given by ‘i’ th row and ‘j’ th column. It is converted into a single integer value for input as a variable with i running from 2 to10 and j running from 2 to 10, considering that the boundary rows and columns in the grid are not having any pumping location by using k = (i-1) + (j-2)*9. Thus the upper bounds are 81 for the locations of the wells while for the pumping values, it is taken as 10000. For the remaining 99 variables, the lower bound is 0 while the upper bound is 1. For these 99 variables, the value can be either 0 or 1 as explained earlier. The population size was 1170 (which is arrived at by multiplying 10 with number of variables).The population type was selected as double vector. There exists a default constraint that if a well is located at a point then there has to be a well at that particular point, Rank was selected as the scaling function and stochastic function as the selection function. In reproduction, the default values of 2 for elite count and 0.8 for crossover function were used. The mutation function was constraint dependent. The crossover function was scattered and migration direction was taken as forward. The default values of 0.2 and 20 were used for fraction and interval respectively. In constraint parameters, for initial penalty and penalty factor, the default values of 10 and 100 were used. In stopping criteria, the default values of infinity, minus infinity, 50, infinity and $1x10^{-3}$ were used for time limit, fitness limit, stall generations, stall time limit and non-linear constraint tolerance respectively. The generations were altered from 100 to 10000 and the function tolerance was altered from 10^{-6} to 10^{-9}.

The maximum absolute error obtained was 0.00042. It took around 10 hours to run the optimization. After running the optimization model the flow contours are drawn in the area for the five days in which the simulations are run which are shown in the figures 2 to 6. The observation well locations are shown in the figure 9. A total of 41 observation well locations were detected by the end of optimization model which satisfies the constraint that the number of wells should be in between 30 to 70 in the model. Later, a check has been performed to test the accuracy of the model by taking the observations at only these 41 observation well locations, and the same optimization was performed with only the pumping locations and the pumping values at respective locations as variables summing them up to 18 variables, by eliminating the 99 variables which were taken for the observation well locations from 117 variables. This test has given satisfactory results with the objective function minimized close to tolerance in only 10 minutes. The results show the exact simulated pumping values as the earlier one and the objective function was found out to be 2.4255×10^{-7}. The objective obtained by the end of this test simulation is quite satisfactory and thus these obtained 41 observation well locations can be considered for the identification of illegal pumping from the well locations in the aquifer domain.

Table II: Observed and simulated values of locations and pumping obtained by optimization model

Pumping Location 1			Pumping Location 2			Pumping Location 3		
Variable (L,T)	Observed Value	Simulated Value	Variable (L,T)	Observed Value	Simulated Value	Variable (L,T)	Observed Value	Simulated Value
57,-	57	57	51,-	51	51	31,-	31	31
57,1	500	500.00	51,1	700	699.9998	31,1	800	800.00
57,2	550	550.00	51,2	0	$4.202x10^{-4}$	31,2	850	849.9999
57,3	650	649.9999	51,3	0	$6.542x10^{-4}$	31,3	0	$6.013x10^{-6}$
57,4	0	$3.965x10^{-4}$	51,4	800	799.9999	31,4	0	$1.787x10^{-5}$
57,5	300	299.9999	51,5	850	850.00	31,5	750	750.00

The results are described in the "table II". The variables in the table are described as (L,T) in which L means Location variable and T means day on which pumping is done. The observed and simulated values are the ones generated in the problem and the values that are obtained through optimization model respectively.

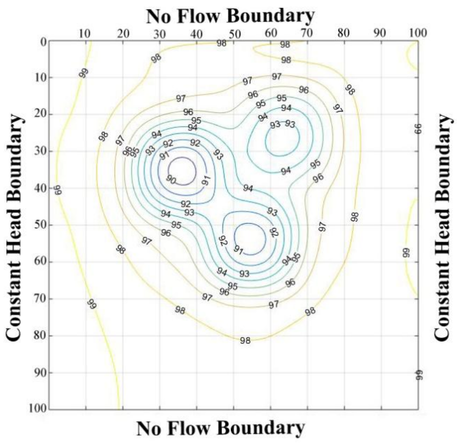

Fig.2 Contour Diagram for Day 1

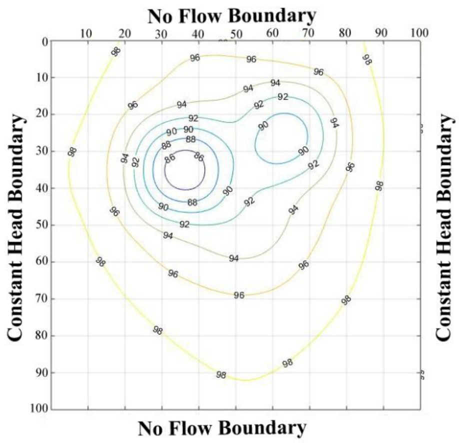

Fig.3 Contour Diagram for Day 2

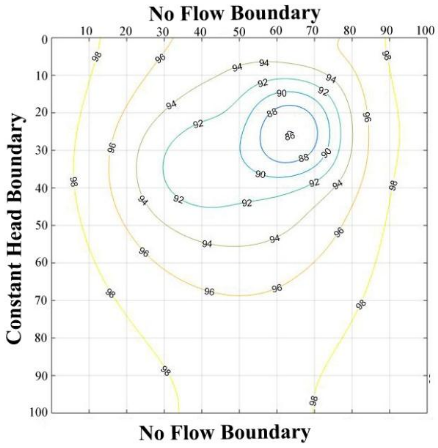

Fig.4 Contour Diagram for Day 3

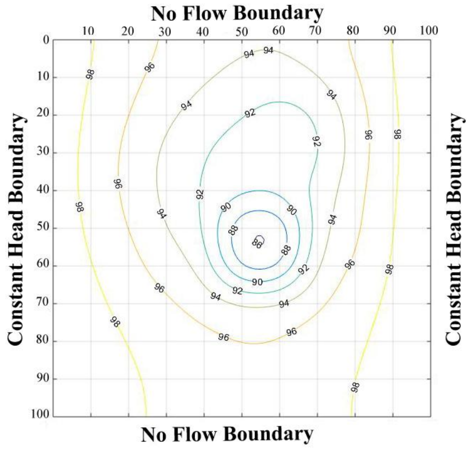

Fig.5 Contour Diagram for Day 4

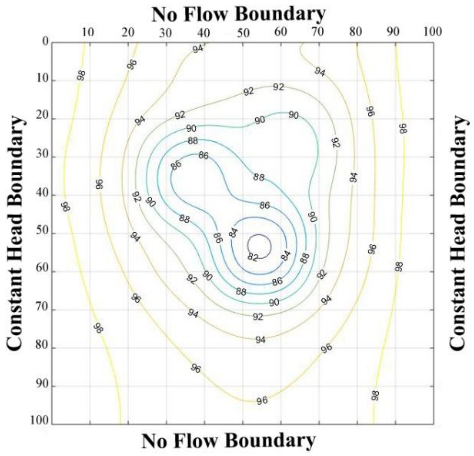

Fig.6 Contour Diagram for Day 5

By the end of optimization with genetic algorithm, the penalty values i.e, the objective function values and the best optimal solution are shown below in the figures 7 and 8 respectively.

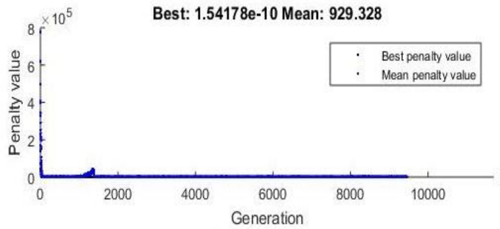

Fig 7: Objective function vs generation in optimization model

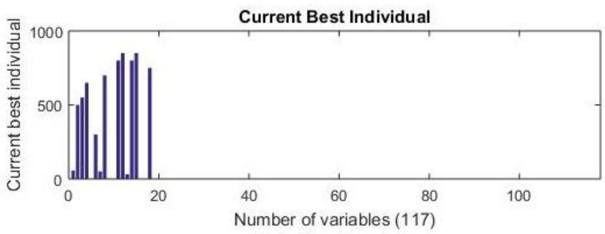

Fig 8: Best individual (optimal solution) obtained by the end of optimization

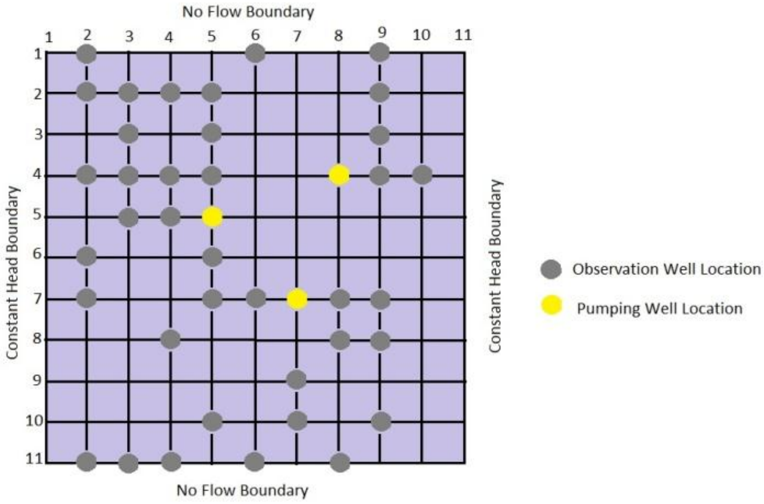

Fig 9: Observation well locations obtained by the end of optimization model

IV. REFERENCES

1.Saffi,M and Cheddadi,A.(2010) “Identification of illegal groundwater pumping in semi confined aquifers”, Hydrological Sciences Journal-Journal Des Sciences 55(8), 1348-1356.

2. Tung,C. and Chou,C. (2004) “Pattern classification using tabu search to identify spatial distribution of groundwater pumping”, Hydrogeology Journal, Vol.12 (5),488-496.

3.Tamer Ayvaz, M and Karahan,H (2008) “A simulation/optimization model for the identification of unknown groundwater well locations and pumping rates”, Journal of Hydrology,357,76-92.

4. Ayvaz,M.T (2008) “Identification of pumping well characteristics in a heterogeneous aquifer system using a genetic algorithm approach, PhD dissertation, Department of Civil Engineering, Pamukkale University, Denizli, Turkey.

5. Huang,C, Mayer,A.S (1997) “Pump and treat optimization using well locations and pumping rates as decision variables”, Water Resources Research 33(5), 1001-1012.

6. Mahinthakumar,G and Sayeed,M.(2005) “Hybrid genetic algorithm local search methods for solving groundwater source identification inverse problems”, Journal of Water Resources Planning and Management-ASCE 131(1), 45-57.

7. Datta,B., Chakrabarty,D., Dhar, A., (2010) “ Identification of unknown groundwater pollution sources using classical optimization with linked simulation”, Journal of Hydro-environment Research, 5(1), 25-36.

8. Datta,B., Chakrabarty,D., Dhar, A., (2009), “Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters”, Journal of Hydrology, 376(1-2), 48-57.

9. Aral,M.M, Guan,J.B, Maslia,M.L (2001), “Identification of contaminant source location and release history in aquifers”, Journal of Hydrologic Engineering 6(3), 225-234.

10. Borah, T and Bhattacharjya,R.K (2013) “Solution of source identification problem using GMS and MATLAB” ISH Journal of Hydraulic Engineering, Vol 19, No.3, 297-304.

11. Singh, R.M, Datta, B, Jain, A.(2004) “Identification of unknown groundwater pollution sources using artificial neural networks” Journal of Water Resources Planning and Management 130(6), 506-514.