

GA APPLICATION TO DETERMINE OPTIMAL PUMPING POLICY IN HETEROGENEOUS UNCONFINED AQUIFER

B. APARNA* and A. K. RASTOGI†

Department of Civil Engineering

Indian Institute of Technology Bombay, India-400076

**baparna@iitb.ac.in*

†akr@civil.iitb.ac.in

In this paper, an irregular heterogeneous unconfined aquifer is simulated by finite element method. Groundwater head distribution is obtained at various nodes of the flow domain involving a set of boundary conditions, source, and sink terms. Finite element model is coupled with genetic algorithm techniques to minimize the cumulative drawdown in the 11 wells subject to constraints on aquifer heads and pumping rates. The drawdowns and the corresponding optimal pumping rates in each well are determined by real-coded GA, and the results are compared with binary-coded genetic algorithm. The values of the GA parameters were chosen based on the sensitivity analysis of crossover, mutation, and number of generations. The results showed that real-coded GA yielded marginally lower values of cumulative drawdowns in comparison to binary-coded GA. The results also suggest near-global optimality of the pumping rates, which are further examined by statistical reliability analysis.

1. Introduction

Groundwater is an important source of water supply in many parts of the world today. Its use in agriculture, industries, municipalities, and rural homes continues to increase. The ever-increasing demand on groundwater has resulted in many environmental imbalances. Excessive extraction rates have caused undesirable effects of critical lowering of water table, increased pumping costs, land subsidence, salination, desertification, changes in hydraulic pressure, and underground flow directions. When several pumping wells are involved particularly in a heterogeneous anisotropic aquifer, a careful decision is required by water managers to withdraw water from these wells. Optimal assessment of groundwater withdrawal rates and the corresponding drawdowns are of utmost importance for adequate management of groundwater resources.

In the present work, a simulation model based on Galerkin's finite element approach has been developed that predicts the head in the flow domain. Aquifer head distribution and its sensitivity to pumping and recharge rates are also discussed. Real-coded genetic algorithm is coupled to the flow simulation model to assess the optimal drawdowns for this combinatorial optimization problem involving 11 pumping wells in a heterogeneous unconfined aquifer. Composite scaled sensitivity of parameters is evaluated to assess the reliability of estimated results.

2. Problem Description

The unconfined aquifer flow domain is divided into nine zones based on hydraulic conductivity and specific yield values. The flow region is discretized into 90 linear triangular elements involving 56 nodes (Fig. 1). The head at constant boundary is 100m and the aquifer consists of 11 pumping wells (Fig. 2). Initially, the pumping rate is assumed to be $1500\text{m}^3/\text{d}$ at each well. An average annual recharge of $0.002\text{m}/\text{d}$ is considered applicable in the region.

The zones of the flow domain are classified based upon the hydraulic conductivity and specific yield values as given in Table 1. Coefficient of anisotropy is equal to 2.

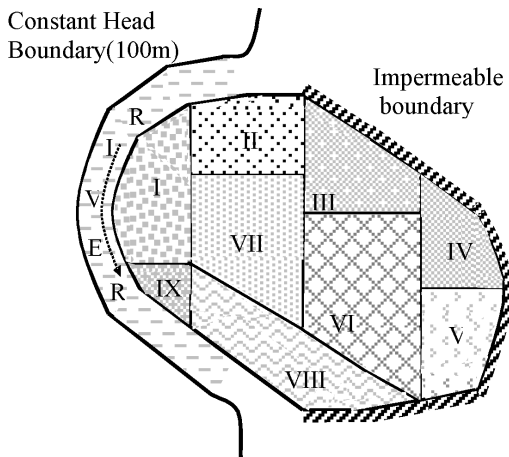


Fig. 1. Flow domain of heterogeneous isotropic unconfined aquifer.

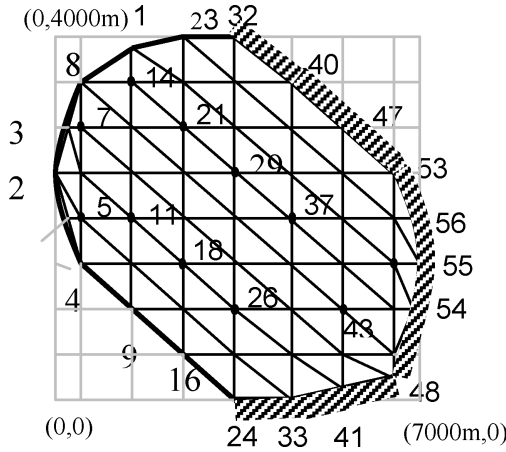


Fig. 2. Finite element discretization. No. of nodes: 56; No. of elements: 90.

Table 1. Aquifer properties in different zones.

Zone	I	II	III	IV	V	VI	VII	VIII	IX
K_x (m/d)	200	50	100	200	150	50	120	150	150
K_y (m/d)	100	25	50	100	75	25	60	75	75
S_y	0.25	0.05	0.12	0.25	0.18	0.05	0.15	0.18	0.18

The governing equation describing the flow in a two-dimensional heterogeneous anisotropic unconfined aquifer is given as¹

$$\frac{\partial}{\partial x} \left[K_x h \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_y h \frac{\partial h}{\partial y} \right] = S_y \frac{\partial h}{\partial t} + Q_w \delta(x - x_i)(y - y_i) - q \quad (1)$$

subject to the following initial and boundary conditions:

$$\begin{aligned} h(x, y, 0) &= h_0(x, y) & x, y &\in \Omega, \\ h(x, y, t) &= h_1(x, y, t) & x, y &\in \partial\Omega_1, \\ K h \frac{\partial h}{\partial n} &= q(x, y, t) & x, y &\in \partial\Omega_2, \end{aligned}$$

where $h(x, y, t)$ is the piezometric head (m), K_x and K_y are hydraulic conductivities in x - and y -directions (m/d), S_y is specific yield, x and y are space variables (m), Q_w is the source or sink term ($-Q_w =$ source, $Q_w =$ sink in $\text{m}^3/\text{d}/\text{m}^2$), t is time in days, Ω is the flow region, $\partial\Omega$ is the boundary region ($\partial\Omega_1 \cup \partial\Omega_2 = \partial\Omega$), $\frac{\partial}{\partial n}$ is the normal derivative, $h_0(x, y)$ is

the initial head in the flow domain (m), $h_1(x, y, t)$ is the known head value of the boundary head (m), $q(x, y, t)$ is the known inflow rate ($\text{m}^3/\text{d}/\text{m}$), and δ is the Dirac delta function.

After Galerkin's finite element formulation, the above partial differential equation is transformed to

$$[G]\{h_I^{t+\Delta t}\} + [P]\{h_I^t\} = \{F\}, \quad (2)$$

where $[G]$ is the conductance matrix containing hydraulic conductivity terms; $[P]$ is the storage matrix containing storativity terms; Δt is the time step size; $\{F\}$ is the net flux vector; $h_I^{t+\Delta t}$ is the unknown head vector; and h_I^t is the known head vector at time t .

The linear simultaneous equations derived above are solved to obtain the head distribution at nodal points using Gauss-Seidel iterative method for the given initial and boundary conditions, recharge, pumping, hydraulic conductivity, and the specific yield values. Iterations are terminated when the difference in heads between two successive iterations is below a set of tolerance level ε , which is presently considered as 0.0001.

3. Groundwater Head Distribution

The groundwater levels may not vary uniformly throughout the flow domain due to pumping by various wells from a heterogeneous anisotropic unconfined aquifer. These are influenced by the pumping conditions, boundary conditions, properties of the porous media (hydraulic conductivity, specific yield, saturated aquifer thickness), aquifer recharge, and also by the distribution of wells in the aquifer domain. Flow pattern in the aquifer and the groundwater head distributions obtained by FEM simulation at a section in the flow domain is shown in Fig. 3. Drawdown in well at nodes 7 and 21 is different due to their location. Adequacy of the model is established by examining its sensitivity to pumping and recharge rates by satisfying a daily mass balance criterion. Head variation near the river (node 5) is shown in Fig. 4 which approaches nearly a steady state due to river recharge. The optimization problem is aimed at minimizing the collective drawdown from the wells, and is represented as

$$\text{Min } F = \sum_{t=1}^{t \max} \sum_{i=1}^n H_I(i) - H_F(i, t) \quad (3)$$

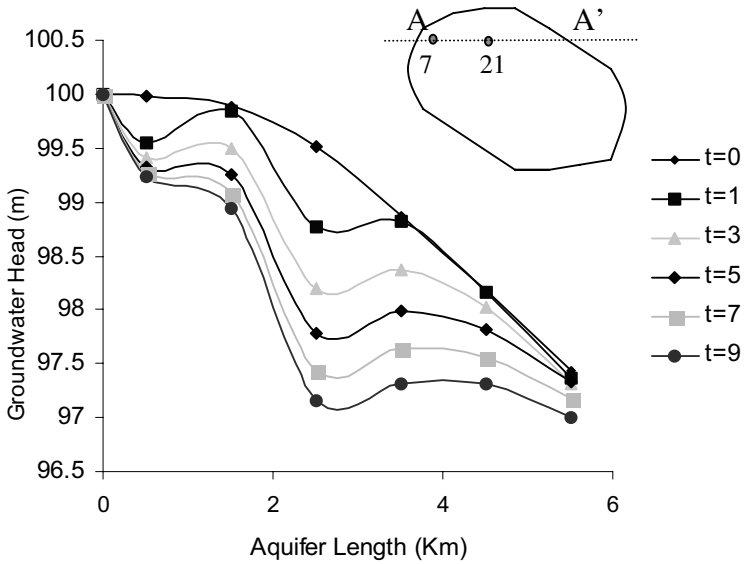


Fig. 3. Groundwater head distribution along section A-A'.

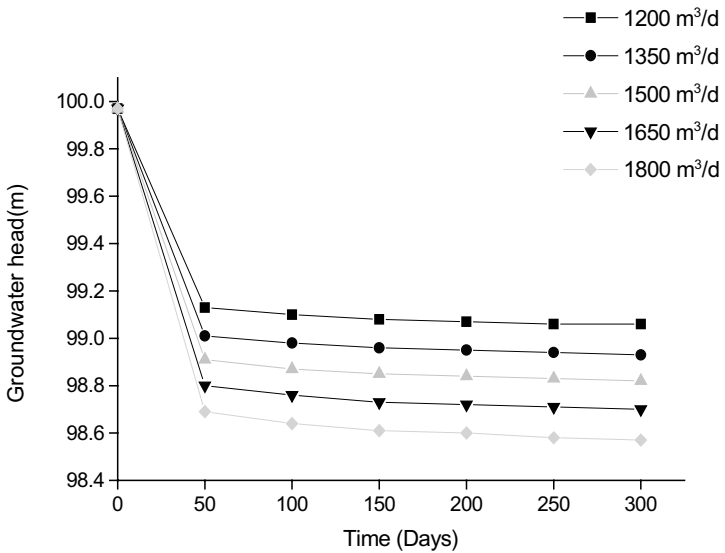


Fig. 4. Head variation in well at node 5 for different pumping rates.

subject to the following set of constraints:

$$Q_L \leq \sum Q \leq Q_U,$$

$$h_L \leq h(i, t) \leq h_U,$$

and

$$q_L \leq q_i \leq q_U.$$

The objective function F represents the cumulative drawdown due to continuous pumping from the 11 wells; t is the time period and t_{\max} is the maximum time for which the pumping is done (1 year presently); n represents the number of wells; Q is the total pumping rate during operational period t ; $H_I(i)$ represents the initial head in well i while $H_F(i, t)$ is the final head in well i at the end of the operational period; Q_L and Q_U represent the lower and upper bounds of total pumping rate which are 5000 m³/d and 10,000 m³/d, whereas h_L and h_U are the lower and upper bounds of the groundwater heads in the wells which are 85 m and 100 m, respectively depending upon the availability of the pump; q_i is the pumping rate in well i ; and q_L and q_U are the lower and upper bounds on individual pumping rates of each well.

4. Genetic Algorithm

Genetic algorithm (GA) is a heuristic method to find approximate solutions to combinatorial optimization problems. Genetic algorithms are a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, natural selection, and recombination (or crossover). Important steps involved in GA methodology can be summarized as follows:

- Generation of initial population of strings;
- Selection of coding scheme;
- Running the flow simulation finite element model for all the sets (strings) of variables;
- Evaluation of each string in a population (computing fitness or objective function);
- Performing the computation with genetic operators;
- Selection of the best strings and formulation of mating pool;
- Crossover of the selected strings;
- Mutation of the strings;
- Terminating condition.

In this study, real-coded GA is applied, and later on the results are compared with binary-coded GAs which are discussed in the following sections.

4.1. Real-coded GA

The above optimization problem is solved by real-coded GA in which the variables are used directly without adopting any coding scheme. In order to estimate the cumulative drawdown, an initial population size which linearly varies with the number of variables ($N = 10n$, where n is the number of variables) is chosen. For this case, the number of variables is 11, i.e. the pumping rates at each of the 11 wells. GA starts with an initial population of a set of individual pumping strategies, which are randomly generated and meet the constraint on the total withdrawal rate. For each particular solution, the finite element model is run to evaluate the groundwater heads. Fitness of each individual is evaluated subsequently based upon the cumulative drawdown function. Individuals are selected by tournament selection based on their fitness values. The winners of the tournament selection are further processed by crossover and mutation operations (directly applied on real parameter values) to produce new individuals, but keeping the population size fixed. One generation refers to one complete cycle of selection, crossover, and mutation. If the optimization criteria are not met, the creation of a new generation starts. This cycle of generations is performed until the optimization criterion is reached.

4.1.1. Simulated binary crossover for real-coded GA

The crossover that has been applied for the present case of real-coded GA is a simulated binary crossover (SBX) operator, which is particularly suitable here, because the spread of children solutions around parent solutions can be controlled using a distribution index η_c . A large value of η_c allows only near-parent solutions to be created, whereas a small value of η_c allows distant solutions to be created. For that reason, η_c value of 8 is considered for the present problem. Another aspect of this crossover operator is that it is adaptive, allowing any solution to be created in the beginning to have a more focused search when the population is converging. The procedure for calculating the children solutions (c_1 and c_2) from parent solutions (p_1 and p_2) is as follows²: A uniform random number (u) between 0 and 1

is generated, and the spread factor β is computed as

$$\begin{aligned}\beta &= (2u)^{\frac{1}{\eta_c+1}} && \text{if } u \leq 0.5 \\ &= \left(\frac{1}{2(1-u)}\right)^{\frac{1}{\eta_c+1}}, && \text{otherwise,}\end{aligned}\tag{4}$$

where η_c is a distribution index of SBX and a non-negative real number. The children solutions are subsequently calculated as follows:

$$\begin{aligned}c_1 &= 0.5[(1 + \beta)p_1 + (1 - \beta)p_2] \\ c_2 &= 0.5[(1 - \beta)p_1 + (1 + \beta)p_2]\end{aligned}\tag{5}$$

These two children solutions are symmetric about the parent solutions. A value of the distribution index (η_c) allows that the children solutions are closer to the parent solutions. A smaller value of η_c makes a more uniform distribution in the range, and if η_c equals to 0, then it makes a nonuniform distribution in the range.

4.1.2. Polynomial mutation for real-coded GA

A polynomial probability distribution is used to create a solution c in the vicinity of a parent solution p under the mutation operator. The following procedure is used for a parameter $p \in [p_l, p_u]$.

1. Create a random number u between 0 and 1.
2. Calculate the parameter δ_q as follows:

$$\begin{aligned}\delta_q &= [2u + (1 - 2u)(1 - \delta)^{n_m+1}]^{\frac{1}{n_m+1}} - 1 && \text{if } u \leq 0.5 \\ &= 1 - [2(1 - u) + 2(u - 0.5)(1 - \delta)^{n_m+1}]^{\frac{1}{n_m+1}} && \text{otherwise,}\end{aligned}\tag{6}$$

where $\delta = \min[(p - p_u), (p - p_l)] / ((p_u - p_l))$.

3. Then, the mutated child c can be calculated as follows:

$$c = p + \delta_q(p_u - p_l).\tag{7}$$

From the above equations, the normalized perturbation = $((c - p) / (p_u - p_l))$ of the mutated solutions in both positive and negative sides separately. We observe that this value is $O(\frac{1}{n_m})$. In order to get a mutation effect of 1% perturbation in solutions, we set $n_m = 100 + t$, and the probability of

mutation is changed as follows:

$$p_m = \frac{1}{n} + \frac{t}{t_{\max}} \left(1 - \frac{1}{n} \right), \quad (8)$$

where t and t_{\max} are the current generation number and the maximum number of generations allowed, respectively. Thus, in the initial generation, we mutate on an average, one variable $p_m = \frac{1}{n}$ with an expected 1% perturbation and as generations proceed, we mutate more variables with lesser expected perturbation. In the real-coded mutation, the amount of perturbation in a variable can be controlled by fixing the parameter η_m (distribution index for mutation). An η_m value of 30 is taken after some careful initial experimentation.

4.2. Binary-coded GA

For the binary coding of GA, the string length for each parameter is taken as 7. The string length is reasonable in keeping view of the accuracy of the solution. The selection operator considered is the same as that taken for real-coded GA. Single point crossover and bitwise mutation operators are applied to evaluate the cumulative drawdown.

5. Sensitivity of GA

The success and performance of GA is dependent on several factors like number of generations, probabilities of crossover, mutation rate, and population size. An acceptable balance between exploration and exploitation can also be accomplished through appropriate selection of these factors. Figure 5 shows the convergence of GA process while optimizing the population size. It is evident from the figure that the optimum value of population for which the objective function is minimum is 110. Similar tuning of GA parameters was carried out for a number of generations (50), crossover, and mutation probabilities, to select the most applicable values for the present problem. Goldberg³ suggested that good performance might be achieved from GA using high crossover probability and low mutation probability. Sensitivity to crossover probability is carried out using a population size of 110 and mutation probability of 0.1. The results show (Fig. 6) that the objective function is minimum for a crossover probability of 0.8. For the determination of sensitivity to mutation probability, a

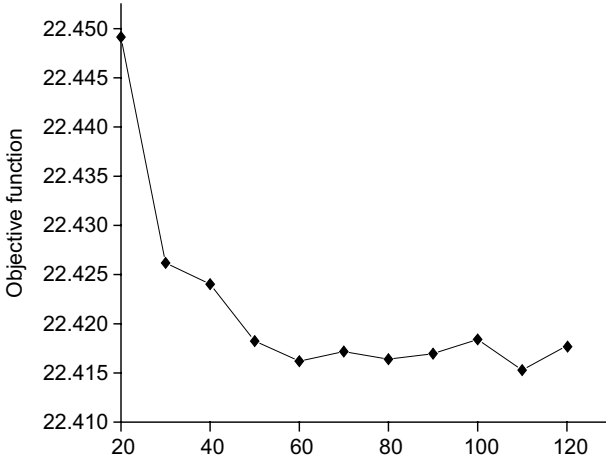


Fig. 5. Effect of population size on objective function.

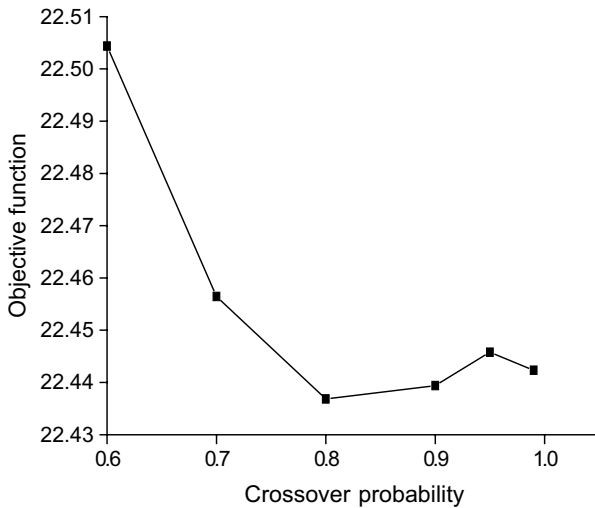


Fig. 6. Effect of crossover probability on objective function.

crossover probability of 0.8 is used. With the initial population taken as 110, the objective function attains a minimum at mutation probability 0.1, and for this value it was found that the solution converged faster in comparison to any other probability value.

6. Results and Conclusions

Individual and cumulative drawdowns in 11 wells by real and binary GAs are shown in Tables 2 and 3, respectively. Optimal pumping rates for the drawdown are summarized in Table 4 for binary GA. However, real GA results are very similar and not reproduced here for brevity. Cumulative optimal drawdown in 11 wells by real and binary GAs is shown in Fig. 7, where a close match can be observed. For the reliability analysis, the composite scale sensitivity parameter⁴ was worked out which suggested that the drawdown near the river wells was more reliable compared to the wells near the impervious boundary of the flow region. FEM aquifer analysis

Table 2. Optimal drawdown (m) in 11 wells by real-coded GA.

Well No.	$T = 100$ days	$T = 200$ days	$T = 365$ days
5	0.484767	0.497149	0.504237
7	0.423515	0.433589	0.439417
11	1.029083	1.095222	1.13293
14	0.572364	0.59372	0.60597
18	1.545848	1.725679	1.82821
21	2.310775	2.554207	2.692876
26	2.256996	2.696994	2.948904
29	3.087477	3.617759	3.921039
37	4.146849	5.09629	5.64315
43	3.368417	4.761373	5.56794
51	3.211266	4.727302	5.606377
Cumulative	22.43736	27.79928	30.89105

Table 3. Optimal drawdown (m) in 11 wells by binary-coded GA.

Well No.	$T = 100$ days	$T = 200$ days	$T = 365$ days
5	0.516797	0.529983	0.537549
7	0.456912	0.467675	0.473912
11	1.079607	1.149942	1.190144
14	0.585693	0.60844	0.621516
18	1.648056	1.839187	1.948465
21	2.368883	2.627478	2.775208
26	2.346852	2.814479	3.083002
29	3.174618	3.738085	4.061333
37	4.277928	5.28749	5.87092
43	3.591697	5.073979	5.935438
51	3.436777	5.050179	5.989259
Cumulative	23.48382	29.18692	32.48675

Table 4. Optimal pumping rates (m^3/d) in pumping wells by binary-coded GA.

Well No.	$T = 100$ days	$T = 200$ days	$T = 365$ days
5	704.7232	703.5173	666.3831
7	721.44	715.8291	683.5269
11	683.2956	673.7941	619.8667
14	720.1143	713.0156	670.7007
18	706.3097	689.6951	619.2156
21	719.2965	701.2842	629.7749
26	675.6772	656.5659	575.1869
29	692.1406	669.9077	586.8978
37	692.9689	662.5927	571.4718
43	609.4508	593.2166	506.4367
51	588.4892	578.2648	493.6202
Cumulative	7513.906	7357.683	6623.081

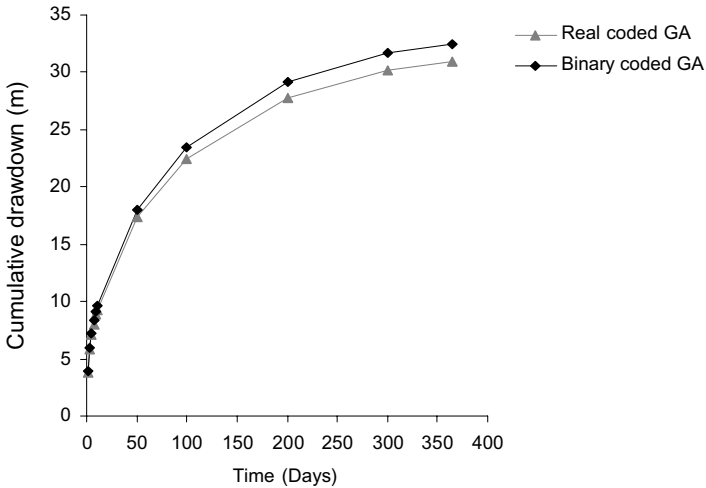


Fig. 7. Cumulative drawdown at various times by real and binary GA.

suggested that there is a rapid decline in groundwater heads for a period of almost 50 days. After 50 days, the groundwater levels tend to recover in some wells located near the river boundary due to a dynamic balance developed between the groundwater withdrawn with the aquifer boundary recharge. However, near the impervious boundary it took nearly 300 days to achieve constant drawdown conditions in the wells. This indicates that the water levels, in general, would recover at a faster rate if the wells

are appropriately located near the river due to the contribution of the boundary flux to the aquifer, but at a slower rate elsewhere, depending on the distance from the boundary. Drawdown increase near impermeable boundary is attributed to the fact that the cone of depression cannot expand across the boundary due to the absence of lateral flow conditions. Coupling of real-coded GA with the simulation model assessed optimal cumulative drawdown for the 11 wells as 30.9m with the accompanying total withdrawals of 6298.18 m³/d. However the optimal results varied only marginally when compared with the binary-coded GA which estimated 5% higher total drawdown followed by a matching increase in cumulative pumping. Both the results satisfied the demands and constraints on the pumping and aquifer head throughout the year. Therefore, the present study concludes that the nonconventional techniques of real- and binary-coded GA can give meaningful results for groundwater withdrawal in heterogeneous anisotropic aquifers from practical consideration.

References

1. J. Bear, *Hydraulics of Groundwater* (McGraw-Hill, New York, 1979).
2. K. Deb and M. Goyal, A combined genetic adaptive search GeneAS engineering design, *Comp. Sci. Informatics* **26**(4) (1996) 30–45.
3. D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning* (Addison-Wesley, Reading, MA, 1989).
4. M. C. Hill, R. L. Cooley and D. W. Pollock, A controlled experiment in groundwater flow model calibration, *Groundwater* **36**(3) (1998) 520–535.
5. M. P. Samuel and M. K. Jha, Estimation of aquifer parameters from pumping test data by genetic algorithm optimization technique, *J. Irrigation Drain. Eng.* **129**(5) (2003) 348–35.