Transactions on Ecology and the Environment vol 19, © 1998 WIT Press, www.witpress.com, ISSN 1743-3541

Combined use of genetic algorithms and boundary elements in aquifer restoration problems

K. L. Katsifarakis[1], N. Theodossiou[2] & P. Latinopoulos[3]
Division of Hydraulics and Environmental Engineering,
Dept. of Civil Engineering, A.U.Th.
GR-54006 Thessaloniki, Macedonia, GREECE
E-mail: klkats@civil.auth.gr[1], niktheo@vergina.eng.auth.gr[2], latin@civil.auth.gr[3].

Abstract

Restoration of contaminated aquifers is a very important issue, especially under conditions of water shortage. In this paper, a method is presented, which combines a boundary element technique, for the solution of the water flow problem, and genetic algorithms, as an optimization tool. The method is illustrated through an application to a contaminant plume stabilization problem, by means of pumping wells. Application examples show that the proposed combination forms a versatile tool, which may be useful in many problems of groundwater contamination control.

1. Introduction

Groundwater contamination is a very important environmental problem among those caused by the disposal of pollutants into natural water systems. This is not only because groundwater is in many cases the major source for water supply, but also because aquifer restoration requires time- and money- consuming procedures.

Several restoration techniques, which can be expressed as management problems, have been developed through the years involving hydrodynamic control and in-situ remediation. Greenwald and Gorelick

[1] provide a detailed literature review on the subject by classifying the available methodologies according to their general goals.

Minimization of groundwater pumping cost is an environmental problem per se, i.e. minimization of energy consumption and of the respective environmental impact. From the mathematical point of view, it is a typical optimization problem, and standard optimization techniques, e.g. linear and non-linear programming, have been used for its solution. The objective function, which should be minimized, is:

$$F = C \cdot \sum_{i=1}^{N} H_i Q_i \qquad (1)$$

where C is a constant and N is the number of the wells, while Q_i and H_i are the flow rate and the distance between ground level and water level at well i, respectively. Constraints are based on the concept of containment of the contamination in a certain (small) portion of the flow field, or on the concept of directing contamination away from production wells or sensitive areas. These constraints can be materialized mathematically as inequalities involving water velocity values V_j at the boundary of the contamination plume. Additional constraints may also be present, including upper bounds of water level drawdowns at the wells or at protected areas of the flow field.

The optimization tools should be combined with groundwater flow simulation codes, to calculate V_j and also φ_i, i.e. the hydraulic head at the wells. The latter enter the cost function through the respective H_I values. In most cases, finite elements or finite differences have been used as flow simulation tools.

In their works Theodossiou et al. [2] and Mylopoulos et al. [3] presented optimization techniques for aquifer cleanup, aiming at the minimization of treatment costs, expressed as linear or nonlinear objective functions. In the former work, a classical one-stage optimization approach was used, whereas in the latter a sequential optimization procedure was proposed. According to this procedure, at the end of each time period the effectiveness of the calculated optimum pumping strategy is checked, by updating the plume boundary through the application of a solute transport model. Finally Latinopoulos et al. [4], by implying a sensitivity analysis, examined the effects of the transmissivity changes upon the optimal solutions of the management problem.

In this paper, the optimization tool is based on genetic algorithms, while a boundary element code is used in groundwater flow calculations. Genetic algorithms have been recently applied to groundwater flow problems (e.g. McKinney and Min-Der Lin [5], Wagner [6], El Harrouni, Quazar & Cheng [7]). Boundary elements, on the other hand, have been used very efficiently for groundwater flow simulation in zoned aquifers (e.g. Latinopoulos and Katsifarakis [8], Katsifarakis, Andreatos and Vournelis [9]). The two techniques have been combined to calculate transmissivities in zoned aquifers, based on a restricted number of field measurements (Karpouzos & Katsifarakis [10]).

Genetic algorithms and boundary elements are briefly outlined in the following paragraphs. Then their use in minimizing total pumping costs is illustrated, through application to an aquifer, in which a contamination plume has been formed.

2. The optimization tool

Genetic algorithms are a mathematical tool, which can be used in many scientific fields (from biology to machine learning). They are particularly efficient in optimization problems, especially when the respective objective functions exhibit many local optima or discontinuous derivatives. There are already extensive textbooks, e.g. Goldberg [11] and Michalewicz [12], which deal with the theoretical background, the computational details and applications of genetic algorithms. Their main features are the following.

Genetic algorithms are essentially a mathematical imitation of a biological process, namely that of evolution of species. They start with a number of random, potential solutions of the problem. These solutions, which are called chromosomes, constitute the population of the first generation. In binary genetic algorithms, each chromosome is a binary string of predetermined length.

Each chromosome of the first generation undergoes evaluation, by means of a pertinent function or process. This process depends entirely on the specific application of genetic algorithms. Then, the second generation is produced, by means of certain operators, which imitate biological processes and apply to the chromosomes of the first generation. The main genetic operators are: a) selection b) crossover and c) mutation. Many other operators have been also proposed and used.

Selection is used first. It leads to an intermediate population, in which better chromosomes have, statistically, more copies. These copies substitute some of the worst chromosomes. Then, the other operators apply to a number of randomly selected members of this intermediate population. The result is an equal number of new chromosomes, i.e. new solutions, which replace the old ones. Thus, the next generation is formed. The whole process, i.e. evaluation-selection-crossover-mutation-other operators, is repeated for a predetermined number of generations. It is anticipated that, at least in the last generation, a chromosome will prevail, which represents the optimal (or at least a very good) solution of the examined problem.

Transactions on Ecology and the Environment vol 19, © 1998 WIT Press, www.witpress.com, ISSN 1743-3541

The genetic operators, which have been used in this problem, i.e. to accomplish contamination plume stabilization with minimal pumping cost, are outlined in the following paragraphs.

2.1 Selection

Selection can be accomplished in many ways. The most common processes are: a) The biased roulette wheel and b) The tournament method. The latter has been preferred, because it applies equally well to maximization and to minimization problems, while the former applies naturally to maximization problems only.

Selection through the tournament method starts with determination of the respective selection constant KK. Then it proceeds in the following way: KK chromosomes are randomly selected from the population of the current generation, and their fitness values are compared to each other. The chromosome with the best (largest or smallest) fitness value passes to the intermediate population. This process is repeated PS times, PS being the population size. In this way, the intermediate population is formed. Moreover, in our genetic code, the best chromosome of each generation is separately preserved through the selection process.

2.2 Crossover

Crossover applies to pairs of chromosomes, which are binary strings of length SL. Two chromosomes, which are named parents, are randomly selected from the intermediate population. An integer number XX, between 0 and (SL-1), is randomly selected, too. Then each parent binary string is cut to 2 pieces, immediately after position XX. The first piece of each parent is combined with the second piece of the other. In this way, two new chromosomes are formed, which are called offspring and substitute their parents in the next generation.

Crossover aims at combining the best features of both parents to one offspring. All chromosomes of the intermediate population have equal probability to undergo crossover. But this probability is actually larger for the better chromosomes of the parent generation, because they have got more copies in the intermediate population.

2.3 Mutation

Mutation applies to characters (genes), which form the chromosomes. In binary genetic algorithms, the gene, which is selected for mutation, is changed from 0 to 1 and vice versa. This process aims at: a) extending the search to more areas of the solution space (mainly in the first generations) and b) helping local refinement of good solutions (mainly in the last generations). The mutation probability is equal for all genes of all chromosomes. Its magnitude depends on the chromosome length SL, but generally it is much smaller than the respective crossover probability, because the latter refers to chromosomes, not to genes.

2.4 Antimetathesis

Many additional operators have been proposed in the literature, to further improve performance of genetic algorithms. A number of them are problem specific, while others are of general use. In this paper, one more operator, of general use, has been used [13]. The operator applies to pairs of successive positions (genes) of a chromosome. Any position (except for the last one) can be selected, with equal probability p_a. If the value of the selected gene equals to 1, it is set to 0, while that of the following gene is set to 1. The opposite happens if the value of the selected gene is 0. More explicitly, the following happen, with regard to gene pairs:

11 becomes 01 * 00 becomes 10 * 10 becomes 01 * 01 becomes 10

In the first 2 cases, the new operator is equivalent to mutation at the selected position. In the last 2 though, it is equivalent to mutation of both genes. Morover, it can be interpreted as a limiting case of the inversion operator. The new operator has been called antimetathesis. This name is in line with the tradition in genetic algorithm terminology, which calls for terms of greek origin. Antimetathesis and mutation are used interchangeably (in the even and odd generations respectively).

2.5 Handling constraints

The usual way to deal with constraints, is to include penalty functions in the evaluation process. Each penalty function affects the fitness value of chromosomes, which violate the respective constraint, increasing it in minimization problems and decreasing it in maximization ones. In this paper, such a penalty function is applied, in conjunction with velocity direction at the check points of the circumference of the contamination plume boundary.

3. Groundwater flow simulation

Groundwater flow simulation is necessary, in order to calculate the values of velocities and of hydraulic head drawdown. The boundary element method is used in this task. This method is very efficient in solving

Transactions on Ecology and the Environment vol 19, © 1998 WIT Press, www.witpress.com, ISSN 1743-3541

steady-state groundwater flow problems. It is based on the second Green's formula for transformation of surface to line integrals. Its main feature is that it does not require grid construction over the field area, since it is based on discretization of external and internal field boundaries. These boundaries are divided to pieces, which are called boundary elements. Calculations are performed in two stages. First, the values of hydraulic head φ and its vertical derivative $q = \partial\varphi/\partial n$ are calculated for each boundary element. Then φ or V is calculated separately for each internal point of the flow field. This is a very important advantage for our application, since φ and V values have to be calculated in very few internal points. Moreover, velocities are calculated directly (not through differences of adjacent φ values).

Finally, an equally important advantage is that areas of wells are described very accurately as concentrated «loads», i.e. without distributing well flow rates to grid elements.

A boundary element code, extensively tested in other applications (e.g. [9], [10]), has been used. It is based on constant boundary elements and its accuracy is satisfactory [8]. To be incorporated in the genetic algorithm, the code has been divided in two parts. The first, which includes data input and some preliminary calculations, is executed only once. The second, main, part has to be executed for every chromosome of every generation, since it is the main part of chromosome evaluation procedure.

4. Application example

The proposed method has been applied to the aquifer of figure 1. Its external boundary consists of 2 constant head parts (namely AB and ΓΔ) and two impermeable parts (namely BΓ and ΔA). Hydraulic head φ on AB equals 0, while on ΓΔ φ = 20m. The aquifer is confined, with thickness b=40m, porosity n=0.2 and transmissivity $T = 0.002m^2/s$.

As shown in fig.1, a contaminant plume has been formed, which moves towards boundary AB. Coordinates and water velocities at points 1 to 5, of the plume boundary, which are used as control points, are shown in table 1.

In order to prevent further contamination of the aquifer, two wells, located at the interior of the plume, are used. The respective coordinates appear in table 2, together with ground elevations (Elev) at the locations of the wells (with reference to φ=0 plane).

TABLE 1. Coordinates and velocities at the plume boundary.

point	x_i	y_i	$V_x \cdot 10^6$	$V_y \cdot 10^6$
1	1400	2100	-0.0013	-1.7986
2	1530	1800	-0.0222	-1.8036
3	1650	1700	-0.0319	-1.7966
4	1770	1800	-0.0256	-1.7852
5	1900	2100	-0.0130	-1.7817

TABLE 2. Coordinates and ground elevations at the locations of the wells

well	x_i	y_i	$Elev_i$
1	1700	1900	20
2	1600	2200	20

The optimization task is to minimize total pumping cost, which is required to prevent further aquifer contamination. It has been mentioned in the introduction, that the pumping cost can be expressed as:

$$F = C \cdot \sum_{i=1}^{N} H_i Q_i \tag{1}$$

where H_i can be written as

$$H_i = | Elev_i - \varphi_i | \tag{2}$$

In order to implement the genetic algorithm, which has been outlined in previous paragraphs, the following set of parameters has been selected: population size = 40, crossover probability = 0.40, mutation and anti-metathesis probability = 0.07, number of generations = 50, selection constant KK = 3, chromosome length SL= 12.

The chromosome length has been determined in the following way: Each chromosome represents a pair of well flow rate values Q_i, expressed as a binary number. To allow for a maximum flow rate of 50 *l*/s to be pumped from each well, 6 genes are needed for each Q_i. Thus, SL=12.

4.2 The evaluation procedure

The evaluation procedure for each chromosome includes the following steps:

a) Calculation of the hydraulic head φ_i at the location of each well, using the respective pair of well flow rate values, b) calculation of the velocities V_x and V_y at the five control points, and c) calculation of the fitness value VB.

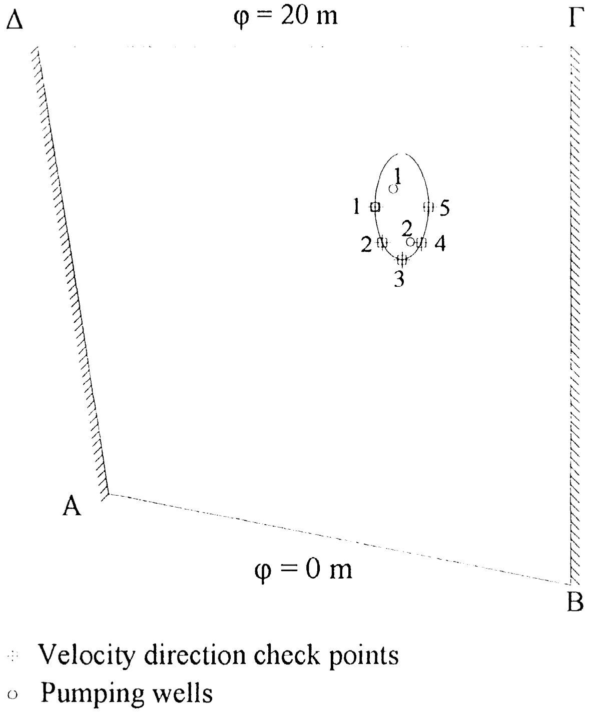

Figure 1. Laterally confined contaminated aquifer

In the first and second steps, the boundary element method is used, as explained in previous paragraphs. The boundary of the flow field has been divided in 44 elements.

In the third step, H_i for each well is calculated, by means of equation 2. Then, an initial fitness value VB is calculated from equation 1 (in which the constant C has been set to 1). After that, it is checked whether the constraints, regarding the contamination plume, are fullfilled. These constraints are expressed mathematically in the following way:

$$V_{x1} > 0 \; , \; V_{y2} > 0 \; , \; V_{y3} > 0 \; , \; V_{y4} > 0 \; , \; V_{x5} < 0 \qquad (3)$$

A penalty equal to 1.5 is added to VB, for each constraint violation. The number 1.5 is calculated considering one well pumping from an infinite field, with the same hydraulic features (b=40m, n=0.2, T = 0.002m^2/s) as the original one. To produce velocities of about $1.8 \cdot 10^{-6}$ m/s at a distance of 500 m, Q should equal 45 *l*/s. The resulting drawdown s_w at the well is close to 34 m. The penalty is set equal to the product $Q \cdot s_w \cong 1.5$. On the same line, the maximum well flow rate is set to $1.1 \cdot Q \cong 50$ *l*/s.

4.3 Typical results

The program has been run several times. Typical results appear in the first row of table 3. It includes the fitness value VB of the best chromosome, together with the respective Q_i, f_i and H_i values. Then, ground elevation at well 2 was set to 30m (instead of 20m) and the program was run again. Typical results appear in the second raw of the table 3. It can be seen that while both Q_1 and Q_2 change significantly, their sum remains practically constant. Moreover, VB is larger in the second case, as expected.

TABLE 3. Best chromosome's fitness value and respective Q_i, φ_i and H_i values

VB	Q_1 (*l*/s)	Q_2 (*l*/s)	φ_1 (m)	φ_2	H_1	H_2
0.9937	19	22	-4.74	-3.81	24.74	23.81
1.1718	28	12	-9.80	1.89	29.80	28.11

4.4 Computer time requirement

The time required to run the respective program is comparatively large (20 to 30 minutes on a Pentium at 133 MHz). This is due to the repetitive use of the boundary element code.

5. Final remarks

The combination of genetic algorithms and boundary elements, which has been described, offers an attractive alternative to classical optimization techniques in the field of groundwater hydraulics and contamination control. Genetic algorithms in particular, which are based on simple mathematics, can be easily adapted to each specific problem. Its relative drawback, i.e. comparatively large computer time requirement, is offset by simplicity of input data preparation, which saves a lot of time for the user.

References

[1] Greenwald, R. M. & Gorelick S.M., Particle travel times for contaminants incorporated into a planning model for groundwater plume capture, *Journal of Hydrology*, Vol 107, pp. 73-98, 1989.

[2] Theodossiou, N., Latinopoulos P. & Mylopoulos Y., Quality

management in groundwater systems using linear and nonlinear programming, *Proc. of the 2nd Conf. on Environmental Science and Technology*, pp. 482-490, Lesvos, Greece, 1991.

[3] Mylopoulos, Y., Latinopoulos, P. & Theodossiou, N., A combined use of simulation and optimization techniques in the solution of aquifer restoration problems, *Water Pollution: Modelling Measuring and Prediction*, eds. L.C. Wrobel and C.A. Brebbia, Computational Mechanics Publications, pp. 59-72, Southampton, 1991.

[4] Latinopoulos, P., Theodossiou, N., Mylopoulos, Y. & Mylopoulos, N., A sensitivity analysis and parametric study for the evaluation of the optimal management of a contaminated aquifer, *Water Resources Management*, Vol.8, pp. 11-31, 1994.

[5] McKinney, D.C. & Min-Der Lin, Genetic algorithm solution of groundwater management models, *Water Resources Research*, Vol. 30(6), pp. 1897-1906, 1994.

[6] Wagner, B. J., Sampling design methods for groundwater modeling under uncertainty, *Water Resources Research*, Vol. 31(10), pp. 2581-2591, 1995

[7] El Harrouni K., Quazar, D. & Cheng, A.H.-D, Boundary and parameter identification using genetic algorithms and boundary element method, *Proc. Int. Conf. Computer Methods and Water resources III*, eds. Y. Abousleiman, C.A. Brebbia, A.H.-D. Cheng & D. Quazar, pp. 487-495, Beirut, Lebanon, 1995.

[8] Latinopoulos P. & Katsifarakis, K.L., A boundary element and particle tracking model for advective transport in zoned aquifers, *J. of Hydrology*, Vol. 124(1-2), pp. 159-176, 1991.

[9] Katsifarakis K.L., Andreatos, N. & Vournelis, E., Application of boundary element techniques to flows through aquifers with zones of irregular shape, *Proc. Int. Conf. Computer Methods and Water resources III*, eds. Y. Abousleiman, C.A. Brebbia, A.H.-D. Cheng & D. Quazar, pp. 109-116, Beirut, Lebanon, 1995.

[10] Karpouzos, D. K. & Katsifarakis, K.L., Combined use of genetic algorithms and boundary elements to calculate zoned aquifer transmissivities, *Proc. 7th Panhellenic Conf. of the Greek Hydrotechnic Association*, pp. 245-252, Patras, Greece, 1997.

[11] Goldberg D. E., *Genetic algorithms in search, optimization and machine learning* Reading, Massachusetts: Addison-Wesley publishing company, 1989.

[12] Michalewicz Z., *Genetic algorithms + Data structures = Evolution programs* (2nd ed.), Springer-Verlag, 1994.

[13] Katsifarakis, K.L., & D.K. Karpouzos, Minimization of pumping cost in zoned aquifers by means of genetic algorithms, *Proc. of the IV Conf. Protection and Restoration of the Environment, Greece, 1998.*