Proceedings of the 14th International Conference on Environmental Science and Technology
Rhodes, Greece, 3-5 September 2015

GROUNDWATER PARAMETER DETERMINATION BY MEANS OF A GENETIC ALGORITHM WITH CHROMOSOMES OF VARIABLE LENGTH

CHAIROPOULOS K.[1] and SIDIROPOULOS E.[1]

[1] Aristotle University, Faculty of Engineering.
E-mail: nontas@topo.a.0uth.gr

ABSTRACT

The present paper deals with the problem of determining the hydraulic conductivity distribution over the extent of an aquifer on the basis of hydraulic head measurements at several different locations. The objective function of this inverse problem consists of the sum of square differences between measured and calculated hydraulic heads. For that purpose, a steady-state groundwater model is adapted and solved by means of finite differences for various possible hydraulic conductivity distributions, thus yielding calculated hydraulic head values. The conductivity is assumed to follow a zoning pattern. In contrast to current literature, the number of zones is not taken as known in the present treatment. This poses a problem which is solved by means of a genetic algorithm with chromosomes of variable length, in order to deal with the determination of a number of zones. The typical chromosome consists of genes expressing the coordinates of each zone's center and the hydraulic conductivity value assigned to that zone. Thus, the length of the chromosome is connected to the number of zones. Special genetic operations are defined in order to recombine and mutate such chromosomes.

The overall calibration problem is characterized as a mixed-integer nonlinear optimization problem. Numerical results are obtained for a hypothetical rectangular aquifer and for various possible zoning patterns. A discussion relating to the pertinent literature concludes the paper.

Keywords: genetic algorithms, variable length chromosomes, finite differences method, clustering, groundwater parameters

1. Introduction

The direct problem of simulating the function of an aquifer lies in calculating the hydraulic heads over the extent of an aquifer. This calculation is based on the knowledge of an aquifer's geometric features, boundary conditions and value of the hydraulic conductivity over the extent of the aquifer. Knowledge of the hydraulic conductivity form field measurements is impractical and it can only be supplemented by estimates. For that reason, determination of hydraulic conductivity by means of inverse problems methodologies is well established (Tsai *et al*., 2003, Ayvaz, 2007). The complexity of the problems that arise has led to the application of evolutionary methods. Typically, the aquifer's region is assumed to be divided into zones of constant conductivity and the values of the conductivities are determined so that the resulting hydraulic heads are calculated as close as possible to the corresponding already measured values. The number of zones is important as it determines the dimension of the problem and it is generally assumed to be known in advance. In a more general setting, a procedure introduced by Sun *et al* (1998) starts from a single zone and solves the problem for a successively larger number of zones, while the increase in the number of zones is performed in a systematic way. This process was applied by Ayvaz (2007) for the aquifer parameter determination by fuzzy clustering. The present paper presents a method that does not require prior knowledge of the number of zones. The genetic algorithm of this paper determines the number as well as the extent of the zones. This is achieved by introducing chromosomes of a varying length accompanied by special genetic operators (Sidiropoulos, 2013).

2. Genetic algorithm with chromosomes of variable length

The objective function of the problem has the form

$$f = \sum_{i=1}^{n} (h_{mi} - h_{ci})^2 \tag{1}$$

where h_{mi} denotes measured and h_{ci} calculated hydraulic head values and n is the number of points where measurements take place. These calculated values are obtained by solving a two-dimensional steady state groundwater flow problem with the following governing equation:

$$\frac{\partial}{\partial x}\left(K\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K\frac{\partial h}{\partial y}\right) + w = 0 \tag{2}$$

where K is the hydraulic conductivity as a function of position and w a source-sink term.

The differential equation (2) is solved by means of block-centered finite differences after the area of the aquifer is divided into cells, according to the scheme employed by Wang and Zheng (1998). The cells are grouped into zones or clusters of equal hydraulic conductivity. According to the clustering procedure used, the positions of the cluster centers are defined first while each cell is subsequently assigned to the zone with the nearest center. Thus, the independent variables of the problem are the coordinates of the zone centers and the constant values of the hydraulic conductivity in each zone. These will also be the independent variables of the objective function (Equation 1). Accordingly, the typical chromosome of the genetic algorithm will consist of a concatenation of as many similar parts as the number of zones. Each of these parts will include three collections of genes respectively representing the x and y coordinates of the zone center and the hydraulic conductivity value assigned to that zone. The decimal representations of the variables will be converted to binary ones, the length of which will be determined on the basis of the pre-defined respective precisions (Michalewicz, 1996). Since the number of zones is not known a priori, the chromosomes are allowed to have varying lengths corresponding to various possible zonations.

Following these problem definitions, parameters of the algorithm that have to be pre-defined are the following: range of the hydraulic conductivity values and of the zone center coordinates; minimum and maximum of possible zone numbers; precision required of the continuous variables; size of the population; probabilities of crossover and mutation; and, finally, the number of generations. In order to form the initial population, an algorithm that produces pseudorandom integers is employed to obtain numbers of zones that will be used for the composition of the various chromosomes. The construction of the starting population is followed by its evaluation. The chromosomes provide the data for every zone and, through the clustering method, the distribution of the hydraulic conductivity over the aquifer emerges. The direct problem is solved using the FDM and the objective function gets evaluated.

The mutation operator takes the standard form of the binary genetic algorithm (Michalewicz, 1996), but crossover between chromosomes of unequal length will require a new, specially designed operator, generalizing the standard crossover operator (Sidiropoulos, 2013). This new operator is demonstrated schematically in the Appendix, where a different random separator is placed on each one of the chromosomes C_1 and C_2, resulting in offspring O_1 and O_2.

3. Determination of the hydraulic conductivity distribution

The standard aquifer is rectangular with 1500 meters in length and 1500 meters in width and is divided into 225 square cells. Its boundaries consist of constant-head and impermeable layers. The groundwater flow is considered steady and isotropic. The example is studied as two-dimensional. The numerical example is based on the premise of the extraction of the measured hydraulic heads via a known hydraulic conductivity distribution pattern. Subsequently, the algorithm attempts to reconstruct the same pattern using as data only the array of the control points measurements.

The parameters of the algorithm are the following: K_{min}=80m/d, K_{max}=170m/d; initial population N_{pop}=40; number of generations N_{gen}=500; crossover probability 20%; mutation probability 2%. The maximum error was E=0.94. In the examples that follow, the direct problem was solved first with an assumed conductivity distribution. The hydraulic heads that resulted were then taken as measured values on a selected number of cells, as shown in Figure 1. These values were subsequently used for the inverse problem which attempts to reconstruct the initial distribution.

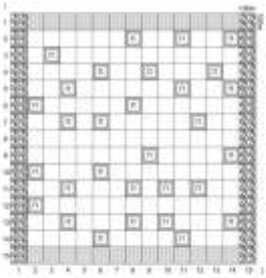

Figure1: Hypothetical aquifer

In Figures 2, 3, 4 and 5 that follow, the initial and the reconstructed distribution are juxtaposed for comparison purposes.

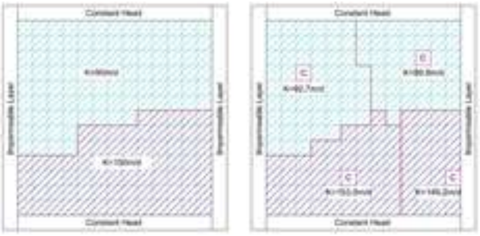

Figure 2: Two-zone problem

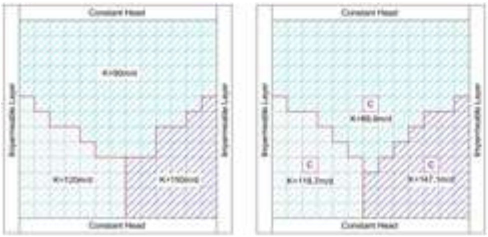

Figure 3: Three-zone problem

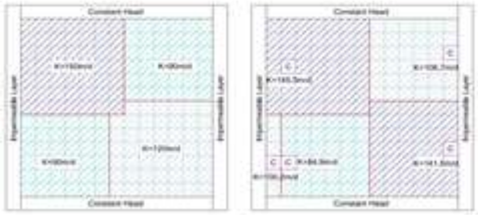

Figure 4: Four-zone problem

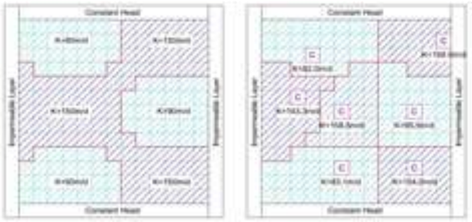

Figure 5: Six-zone problem

4. Conclusions

The present paper demonstrates the structure and the function of a genetic algorithm constructed to calculate the distribution of the hydraulic conductivity in homogeneous zones over the extent of an aquifer without prior knowledge of the number of zones, in contrast to the treatments presented in the current literature. Thus, finding the dimension of the inverse problem was included in the same unified algorithm. The numerical results that were obtained for the various hypothetical zoning patterns of the model aquifer show the algorithm’s capacity to reproduce diverse hydraulic conductivity distributions. Even in cases that the algorithm did not retrieve the exact number of zones, the pattern produced was close to the actual one, thus permitting the user to easily discern the true number of zones. Also, regarding the values of the hydraulic heads that result from the reconstructed hydraulic conductivity, they were very close to the hypothetical values. Investigation of more complex and more realistic inverse problems will contribute to a further development of the present method.

REFERENCES

1. Ayvaz, T.M. (2007). Simultaneous determination of aquifer parameters and zone structures with fuzzy c-means clustering and meta-heuristic harmony search algorithm. Advances in Water Resources, 30, 2326-2338.
2. Michalewicz Z., (1996). «Genetic Algorithms + Data Structures = Evolution Programms». Springer
3. Sidiropoulos, E. 2013. A genetic algorithm with variable chromosome length (in Greek). Honorary Volume dedicated to Prof. D. Tolikas, Aristotle University of Thessaloniki, pp.47-56.
4. Sun, N.Z., Yang, S.L. and Yeh, W.W.G., 1998. A proposed stepwise regression method for model structure identification. Water Resources Research, 34(10), 2561-2572.
5. Tsai, F. T-C, Sun, N. and Yeh, W.W.G. 2003. A combinatorial optimization scheme for parameter structure identification in groundwater modeling. Ground Water, 41(2), 156-169.
6. Wang, P.P.and Zheng, C., 1998. An efficient approach for successively perturbed groundwater models. Advances in Water Resources, 21, 499-508.

APPENDIX. CROSSOVER OPERATOR FOR CHROMOSOMES OF UNEQUAL LENGTHS.