

## **Determination of the hydraulic properties of a multilayered aquifer from pumping test data using genetic algorithms**

**Y. TAKESHITA, I. KOHNO**

*Department of Environmental and Civil Engineering, Okayama University,  
Okayama 700-8530, Japan*

e-mail: [yujitake@cc.okayama-u.ac.jp](mailto:yujitake@cc.okayama-u.ac.jp)

**K. YASUI**

*Shimizu Corporation, Minato-ku, Tokyo 105-0023, Japan*

**Abstract** A new approach to parameter determination methodology is presented which estimates soil hydraulic properties and aquifer thickness from pumping test data with the aid of genetic algorithms (GA) incorporating finite element analysis. GA are search algorithms based on the mechanics of natural selection and natural genetics, which combine an artificial survival of the fittest with genetic operators abstracted from nature. The advantages of our proposed method are the possibility of identifying the optimal hydraulic properties and geological formations of aquifers. Pumping test data measured under multilayered conditions were used to evaluate our proposed method.

### **INTRODUCTION**

Pumping tests are widely used to evaluate the hydraulic characteristics of aquifers. These pumping tests are usually performed in a multilayered aquifer system. The exact determination of hydraulic properties of each aquifer layer is very important for the correct groundwater flow prediction. It is, however, difficult to analyse the data obtained from the pumping test under these conditions "analytically".

In this paper, a new approach to determining soil hydraulic properties from drawdown data, which are obtained by pumping tests in a multilayered aquifer system, has been developed. In this method axisymmetric seepage flow finite element analysis and genetic algorithms (GA) are used. GA provide the possibility of finding the optimal hydraulic conductivity, specific storage and thickness of each aquifer layer. Time-drawdown data, which are observed in the two layered confined diluvial aquifers, are used to evaluate the applicability of our proposed method.

### **METHOD**

#### **Study site**

The study was conducted in the confined diluvial aquifer at Okayama City in Japan. The geological conditions of the study site are shown in Fig. 1(a). The diluvial sand-gravel layers (Dg.) are revealed as a confined aquifer existing in this region. The

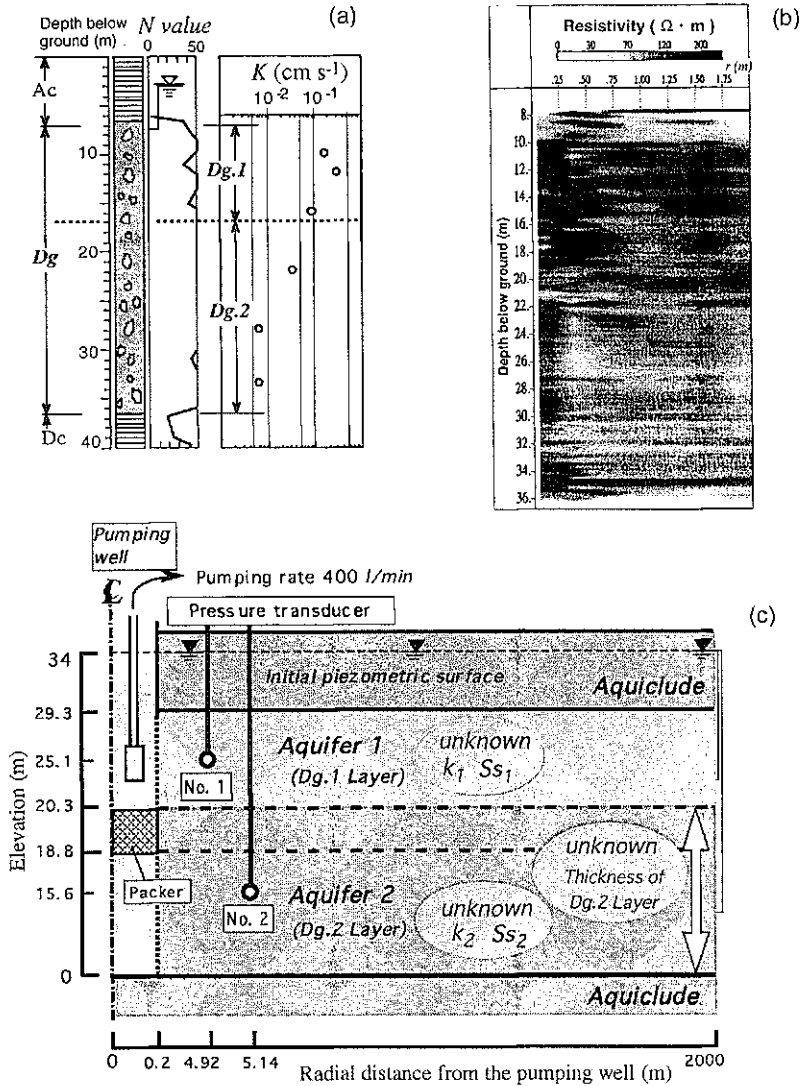


Fig. 1 (a) Geological conditions of the study site. (b) Resistivity distribution by BRP. (c) Schematic diagram of well construction and geological conditions of pumping test.

thickness of the Dg. layer is about 30 m. Borehole resistivity profiler (BRP) and single piezometer tests were performed to estimate the vertical profile of hydraulic conductivity in the Dg. layer. Figure 1(b) shows the resistivity profile measured by BRP. The resistivity distribution is divided into two parts. Its boundary is located at a depth of about -17 m from the ground surface. The resistivity value of the lower layer is lower than that of the upper layer. These resistivity measurements suggest that the hydraulic conductivity of the upper layer is greater than that of the lower layer.

The single piezometer tests were performed at different depths in layer Dg. The single piezometer test involves the injection of water into a borehole, followed by measuring the volume of injected water in steady state. The vertical profile of

measured hydraulic conductivity shows values from  $8.3 \times 10^{-2}$  to  $2.4 \times 10^{-1}$  cm s<sup>-1</sup> in the upper layer and from  $5.7 \times 10^{-3}$  to  $3.0 \times 10^{-2}$  cm s<sup>-1</sup> in the lower layer (Fig. 1(a)). On the basis of these measurements, the confined diluvial aquifer was divided into two layers, Dg.1 and Dg.2. The boundary between them is assumed to be at a depth of 17 m. This aquifer system, therefore, consists of two aquifers, each with its own hydraulic characteristics, and separated by an interface that allow unrestricted crossflow.

The well screen of the pumping well was designed through the entire aquifer except for the depth range 17–18.5 m (elevation 18.8–20.3 m in Fig. 1(c)). Air packers were installed at that depth to test each aquifer independently. Different types of pumping tests were carried out by changing the pumped layer. In this study, pumping test data obtained during pumping from layer Dg.1 were used to determine the hydraulic properties of layers Dg.1 and Dg.2. Drawdowns were measured by pressure transducers which were installed in each aquifer layer by using a bentonite seal. The pressure transducers and pumping system were automatically controlled by a personal computer. The pumping rate was kept constant. The minimum time interval for measurements of drawdowns was one second.

## **Genetic algorithms**

GA are search algorithms based on the mechanics of natural selection and natural genetics, which combine an artificial survival of the fittest with genetic operators abstracted from nature (Goldberg, 1989). Using GA, the search begins from a population of parameter realizations, not from a single realization as in more conventional optimization procedures. The GA use probabilistic, not deterministic rules for perturbation of the realizations, and the objective function information is used directly rather than derivatives or the secondary descriptors.

The basic idea in using GA as an optimization method is to represent a population of possible solutions, referred to as “strings”, in a binary-type encoding, and to manipulate these encoded possible solutions through simulated reproduction, crossover, and mutation. The patterns of (1,0) in the individual binary strings represent characteristics of the corresponding model parameters. The length of the strings depends on the magnitudes of the model parameters, and the accuracy requirement of the final optimal solution. The strings in the initial population are generated by a random number generator. Consequently, GA have no need to determine the initial estimates of unknown parameters.

The strings are evaluated using an objective function. Once the strings are evaluated, the next step is to select a temporary population of the possible solutions by using a stochastic selection processes. As the aim is to propagate the better solutions, those with better fitness values should have a higher probability of proceeding to the next cycle, referred to as “generation”. The selection is repeated until a mating pool as large as the original population is selected. The mating pool represents the population of parent strings that survive to create child strings. The crossover is an operator to determine how the string creatures are paired up for mating and where strings actually cross over. The crossover is performed on each mated pair with a certain probability, referred to as crossover probability.

The last operator of GA is mutation. In the binary coded strings, mutation is typically performed by flipping the values of the 0/1 bit of each string that is selected. The probability of selecting a string for mutation is controlled by the mutation probability. The purpose of mutation is to prevent the GA from prematurely converging to a local minimum. After the mutation, a new population of strings is created. The new population is evaluated again according to the objective function. The next step is to check the convergence criterion, which is based on the change of either the objective function or the optimized parameters. If it is met, the procedure stops. Otherwise, it goes back to the reproduction stage to start another generation until the convergence criterion is met. The number of generations needed depends on the number of unknown parameters to be estimated and the convergence criterion used.

### Parameter estimation using genetic algorithms

Although Javadel & Witherspoon (1983) developed analytical solutions for the drawdown in both layers of a confined two-layered aquifer system, this method is not applicable for our pumping test data because the assumptions needed in their method are not satisfied in our pumping test site conditions. In this study, soil hydraulic properties in layers Dg.1 and Dg.2 and the thickness of Dg.2 layer are estimated using GA.

Drawdown data obtained from pumping test are simulated as axisymmetric transient flow through the rigid porous medium. The axisymmetric flow equation in a confined aquifer can be written as follows:

$$S_s \frac{\partial h}{\partial t} = \frac{\partial}{\partial r} \left( K \frac{\partial h}{\partial r} \right) + \frac{K}{r} \left( \frac{\partial h}{\partial r} \right) + \frac{\partial}{\partial z} \left( K \frac{\partial h}{\partial z} \right) \quad (1)$$

where  $z$  is the vertical coordinate (positive upwards),  $h$  is the pressure head at distance  $r$  from the pumping well at the time  $t$  since pumping started,  $K$  is the hydraulic conductivity, and  $S_s$  is specified storage. The solution of equation (1) is obtained by a finite element modelling of non-steady seepage flow.

Various gradient-based inverse methods have been used for parameter estimation of soil hydraulic properties. Generally the sums of weighted differences between observed data and computed prediction are evaluated as the objective function. A commonly used expression for the objective function  $\mathbf{O}(\mathbf{b})$  is the weighted least squares, as defined by equation (2):

$$\mathbf{O}(\mathbf{b}) = \sum_{i=1}^N w_i (h_m(t_i) - h_c(t_i, \mathbf{b}))^2 \quad (2)$$

where  $w_i$  is weighting factor taken as unity in the simulations considered here because measured pressure head errors were assumed to be constant.  $N$  is the number of measurements of the pressure head  $h_m(t_i)$ . Pressure head data  $h_c(t_i, \mathbf{b})$  are computed at prescribed times  $t_i$  and as a function of the set of optimized parameters in the vector  $\mathbf{b}$  which consists of the five unknown parameters,  $K_1$ ,  $K_2$ ,  $S_{s1}$ ,  $S_{s2}$  and the thickness of layer Dg.2.

Using GA, five unknown parameters are encoded as binary strings. The data structure of encoded parameters are shown in Table 1. The ranges of  $K_1$  and  $K_2$  are

**Table 1** Data structure of encoded parameters.

String no.	$K_1$ or $K_2$ (cm s <sup>-1</sup> )	Binary code (7 bit)	String no.	$S_{s1}$ or $S_{s2}$ (cm <sup>-1</sup> )	Binary code (6 bit)	String no.	Thickness of Dg.2 layer (m)	Binary code (3 bit)
0	$5.0 \times 10^{-4}$	0000000	0	$9.0 \times 10^{-8}$	000000	0	20.3	000
1	$5.2 \times 10^{-4}$	0000001	1	$9.5 \times 10^{-8}$	000001	1	20.1	001
:	:	:	:	:	:	2	19.9	010
63	$8.6 \times 10^{-3}$	0111111	31	$6.5 \times 10^{-6}$	011111	3	19.7	011
64	$8.8 \times 10^{-3}$	1000000	32	$7.0 \times 10^{-6}$	100000	4	19.5	100
:	:	:	:	:	:	5	19.3	101
126	$3.2 \times 10^{-1}$	1111110	62	$4.0 \times 10^{-4}$	111110	6	19.1	110
127	$3.4 \times 10^{-1}$	1111111	63	$4.5 \times 10^{-4}$	111111	7	18.8	111

from  $5.0 \times 10^{-4}$  to  $3.4 \times 10^{-1}$  cm s<sup>-1</sup>. The ranges of  $S_{s1}$  and  $S_{s2}$  are from  $9.0 \times 10^{-8}$  to  $4.5 \times 10^{-4}$  cm<sup>-1</sup>. The range of thickness of layer Dg.2 is from 18.8–20.3 m.  $K_1$  and  $K_2$  are divided into 128 possibilities which can be represented by a 7-bit binary code.  $S_{s1}$  and  $S_{s2}$  are divided into 64 possibilities, which can be represented by a 6-bit binary code. The thickness of layer Dg.2 is divided into eight possibilities, which can be represented by a 3-bit binary code. There are about  $2^{29}$  ( $= 2^7 \times 2^6 \times 2^7 \times 2^6 \times 2^3$ ) different combinations of unknown parameters. This results in a 29-bit long binary code for each individual string composed of all five unknown parameters. One evaluation of the string corresponds to each axisymmetric seepage flow simulation run.

## RESULTS AND DISCUSSION

In this paper, we employ transient drawdown data which were measured in layers Dg.1 and Dg.2 to perform our GA-based parameter estimation procedure. In our GA operations, a population size of 60, crossover probability of 0.6 and mutation probability of 0.1, are used as the empirical parameters. These parameters may need frequent adjustment to ensure successful and efficient application of the GA method. Estimated hydraulic conductivity, specific storage in each aquifer and thickness of Dg.2 are listed in Table 2.

**Table 2** Aquifer parameters estimated by GA.

GA Run	RMSE (cm)	$K_1$ (cm s <sup>-1</sup> )	$K_2$ (cm s <sup>-1</sup> )	$S_{s1}$ (cm <sup>-1</sup> )	$S_{s2}$ (cm <sup>-1</sup> )	Thickness of Dg.2 (m)
Trial 1	1.39	$6.6 \times 10^{-2}$	$7.6 \times 10^{-3}$	$2.5 \times 10^{-7}$	$9.5 \times 10^{-5}$	20.3
Trial 2	1.38	$6.6 \times 10^{-2}$	$7.6 \times 10^{-3}$	$8.5 \times 10^{-7}$	$1.0 \times 10^{-4}$	20.3
Trial 3	1.37	$8.2 \times 10^{-2}$	$8.2 \times 10^{-3}$	$7.0 \times 10^{-7}$	$1.0 \times 10^{-4}$	20.3

Figure 2 shows the measured data from the pumping test. The results for pressure head are presented in these figure, where the solid and broken lines refer to the computed data using parameters estimated by GA. The symbols denote the measured data. As can be seen in this figure, the behaviour of computed drawdown was in good agreement with the measured data.

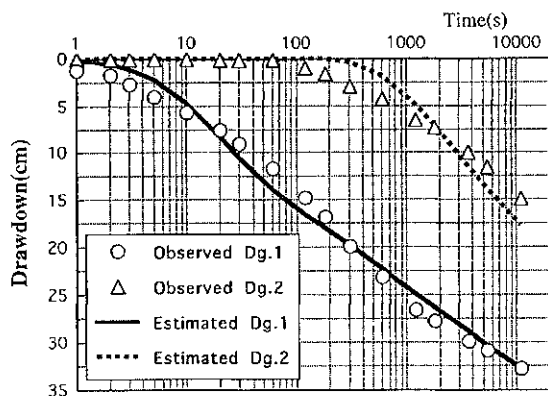


Fig. 2 Estimated and measured pressure head with time.

Differences between the optimizations were evaluated using the root mean squared error (*RMSE*) between optimized and measured transient data defined by the following equation:

$$RMSE = \left\{ \sum_{j=1}^N (y_j - Y(\mathbf{b}))^2 / N \right\}^{1/2} \tag{3}$$

where  $Y(\mathbf{b})$  is the estimated drawdown data, and  $y$  the observed drawdown data.

Figure 3 shows the evolution of the relationship between *RMSE* and number of generations in the search process of the GA. The GA solution converged to a certain condition after about 10 generation runs.

The parameter estimation from our pumping test indicates that the GA-based parameter estimation method works well. The GA method has some advantages over traditional gradient-based methods. The method is very stable and robust because it does not need to evaluate derivatives. This advantage will become more obvious when dealing with a highly nonlinear problem. The most obvious limitation of the GA method, on the other hand, is the large number of forward model simulation runs required. For our pumping test problem with five unknown parameters, it typically

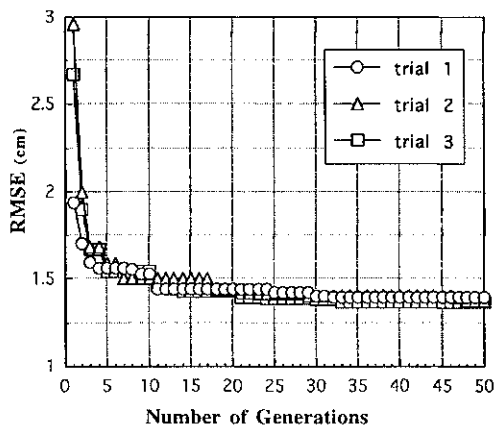


Fig. 3 Evolution of the *RMSE* in the search process of the GA.

requires about 20 or 30 generations for the GA solution to converge. Each generation, in turn, requires many forward runs depending on the population size. The GA, however, can search for the reasonable solution which is difficult for conventional gradient-based procedures to estimate from immensity solution sets.

## **SUMMARY**

We have developed a numerical code for a GA-based parameter estimation procedure to estimate soil hydraulic parameters from transient outflow experiments. Measured drawdown data were used to evaluate the objective function in our GA-based method. The results presented in this study indicate that the GA-based parameter estimation method works well. Hydraulic conductivity, specific storage and thickness of aquifer were estimated simultaneously. Estimated parameters are accurate enough for practical reasons. The GA-based parameter estimation method is a more stable and robust approach than traditional gradient-based methods. This advantage will become more obvious when dealing with a highly nonlinear problem such as inverse problems of estimating unsaturated soil hydraulic parameters.

**Acknowledgements** The authors are grateful to H. Uekuma and A. Nishimura (OYO Corporation, Japan) for providing us the geotechnical instruments and the field data.

## **REFERENCES**

- Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Massachusetts, USA.
- Javandel, I. & Witherspoon, P. A. (1983) Analytical solution of a partially penetrating well in a two-layer aquifer. *Wat. Resour. Res.* 19, 567–578.