1.0

1.1

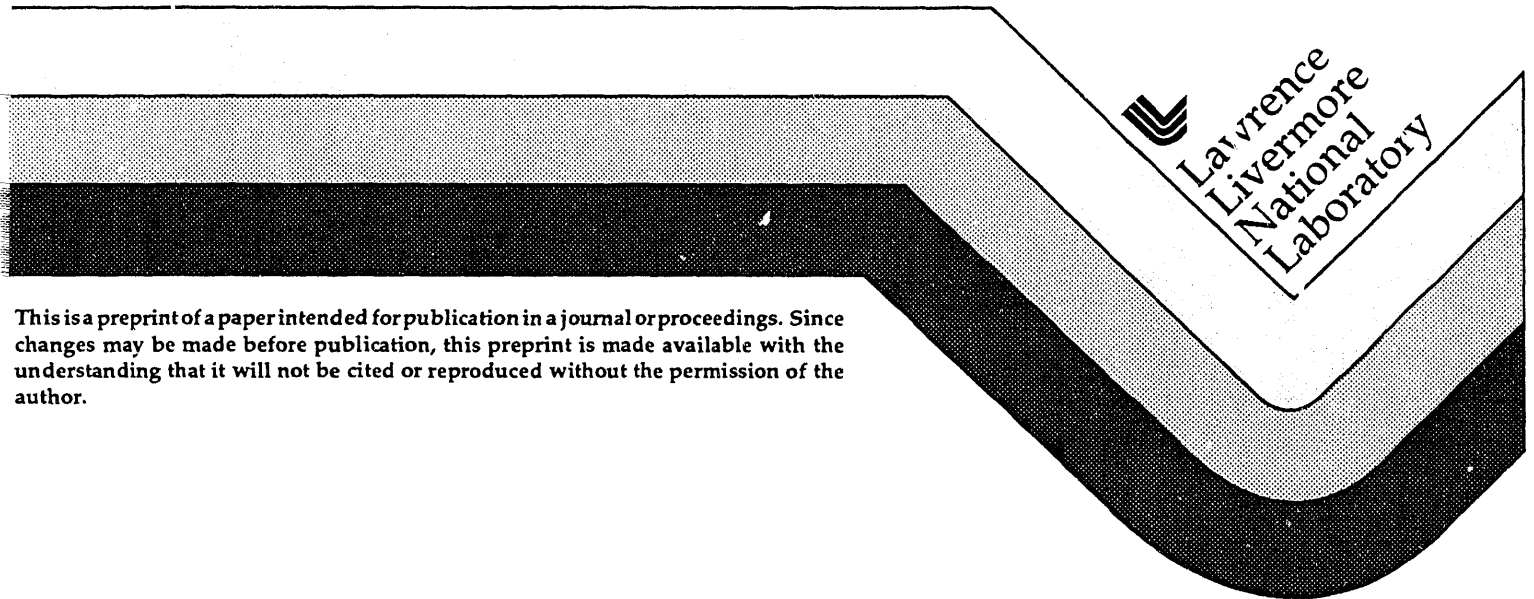1.25   1.4   1.6

2.8   2.5
3.2   2.2
3.6
4.0   2.0

1.8

1 of 1

Optimal Field-Scale Groundwater Remediation Using Neural Networks and the Genetic Algorithm

Leah L. Rogers
Farid U. Dowla
Virginia M. Johnson

DISCLAIMER

Optimal Field-Scale Groundwater Remediation Using Neural Networks and the Genetic Algorithm

Leah L. Rogers, Farid U. Dowla, and Virginia M. Johnson
Earth Sciences Department and Environmental Protection Department
Lawrence Livermore National Laboratory
Livermore, CA 94551 U.S.A.

We present a new approach for field-scale nonlinear management of groundwater remediation. First, an artificial neural network (ANN) is trained to predict the outcome of a groundwater transport simulation. Then a genetic algorithm (GA) searches through possible pumping realizations, evaluating the fitness of each with a prediction from the trained ANN. Traditional approaches rely on optimization algorithms requiring sequential calls of the groundwater transport simulation[1-6]. Our approach processes the transport simulations in parallel and "recycles" the knowledge base of these simulations[7,8], greatly reducing the computational and real-time burden, often the primary impediment to developing field-scale management models. We present results from a Superfund site suggesting that such management techniques can reduce cleanup costs by over a hundred million dollars.

Consider a hypothetical remediation example drawn from field data of volatile organic compounds (VOCs) at a well-characterized Superfund site (Figure 1a). A pump-and-treat remediation is optimized using a 2-D numerical groundwater flow and transport code (GFTC) to forecast the effectiveness of different pumping scenarios. The 28 potential well locations were chosen by trial and error as one pumping scenario that could contain the plume within its current boundaries. Let us now ask the management questions: 1) what subset of these 28 locations would minimize remediation costs and still contain the plume, and 2) what do we give up in terms of total contaminant mass extracted over the period of remediation?

The trained ANN can evaluate 2,000 pumping realizations per second, while the GFTC assesses only one in 2 hours on a SUN SPARCstation II. In this 28-well example, trained ANNs evaluated over 4 million pumping realizations. The top three realizations (Figure 1b) cost between $37 and $42 million over 50 years of remediation, as compared with $155 million for using all 28 wells.

In a somewhat different approach from the majority of ANN optimization applications[13,14], our approach combines ANN technology, inspired by neurobiological theories of massive

parallelism[15,16], with the GA, developed in analogy to natural selection and genetics [17,18] (Figure 2). An ANN is not programmed; rather, it learns from training examples developed by the GFTC. Once trained, the ANN predicts the success of new pumping realizations generated by the GA. The GA incorporates the predicted success into the search procedures, which create new generations of pumping realizations. This design strategy may also be useful with other optimization problems requiring a computationally intensive model for objective evaluation.

A set of 325 pumping realizations was run on the GFTC to develop training patterns to meet three objectives: 1) contain the plume within a fence of downgradient monitor wells (regulatory constraint), 2) maximize the total mass of solute extracted over the 50 years of remediation (mass), and 3) minimize cost (cost). The cost function included calculations of facilities, piping, operating, and water treatment costs. The costs of pumping and injection wells were differentiated. The cost of treated groundwater pumped to the surface was averaged between the different treatment protocols to be about 0.25 cent per liter.

An ANN was trained for each of the three objectives of remediation. The supervised learning

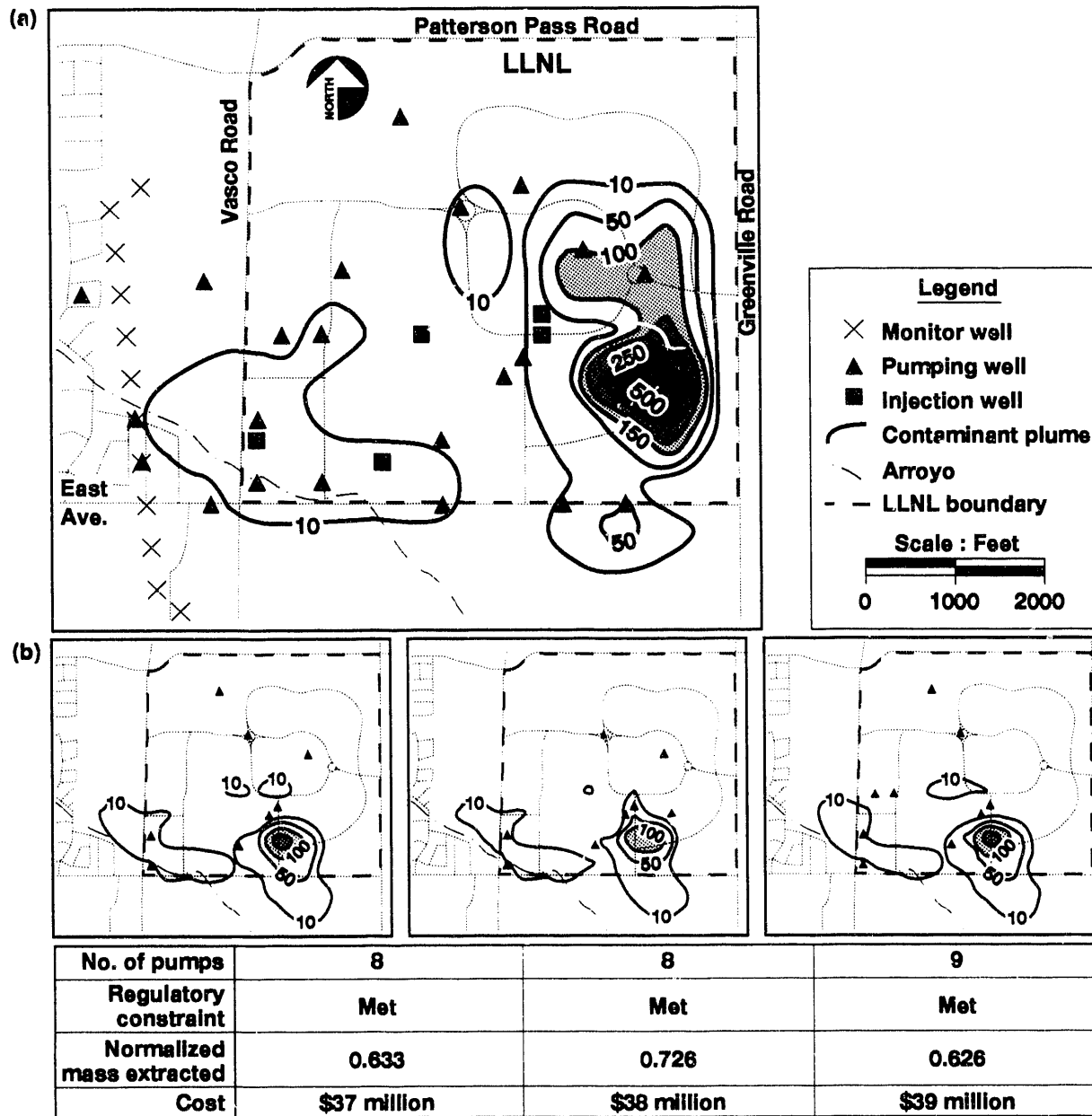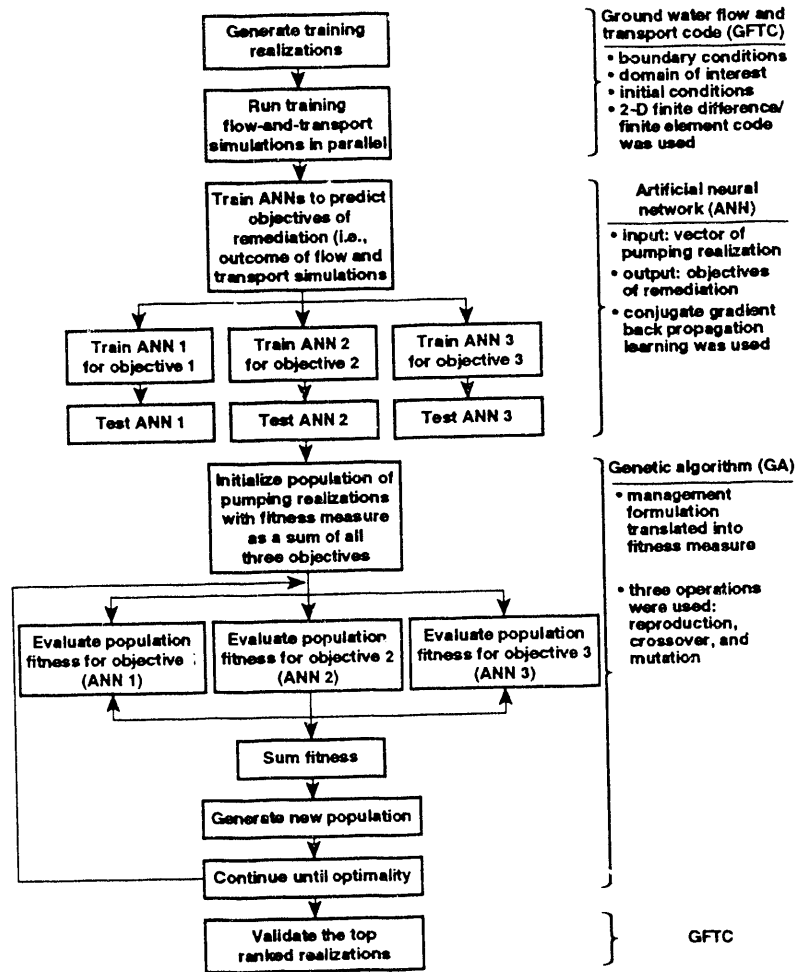| No. of pumps	8	8	9
Regulatory constraint	Met	Met	Met
Normalized mass extracted	0.633	0.726	0.626
Cost	$37 million	$38 million	$39 million

Figure 1: Volatile organic compound (VOC) distribution and GA-ANN optimization results.
(a) VOC concentration contours are shown for the approximately square-mile Lawrence Livermore National Laboratory (LLNL) site and hydraulically downgradient vicinity near Livermore, Calif.[9-11]. Also shown are 28 potential remediation well locations with 23 pumping (triangles), 5 injection (squares), and 12 monitoring (crosses) locations. The maximum well capacities were determined from lithology and long-term pumping tests in existing wells that were close to the well locations. Several locations had existing wells that had already been tested. The 2-D hybrid finite-element/finite-difference transport code, SUTRA[12], was used to solve the governing equations for confined areal groundwater flow and areal solute transport. In brief, the flow boundaries are no-flow fault zones to the northeast and southeast, with flux boundaries directly east and more distantly downgradient to the west. Several influent and effluent streams, several municipal and irrigation wells, and constant annual recharge from rainfall are included. Four hydraulic conductivity values represent four main hydrogeologic units, with a typical value being 3 m/day. A steady flow model was originally calibrated to the piezometric surface, which has had fairly constant topology over time. The time over which the remediation occurs is 50 years. (b) The most optimal pumping scenarios found after evaluation of over 4 million scenarios. Contours show VOC concentrations remaining after 50 years of remediation.

2

Figure 2. GA-ANN management flowchart. In our groundwater management approach, the input of the ANN characterizes the different realizations of pumping. The output characterizes the objectives and/or constraints of the optimization, such as whether regulatory goals have been met, value of cost functions or cleanup time, and mass of contaminant removed. The ANN was trained on patteᵣᵣₛ or realizations of pumping, with each pattern having 28 components representing the individual wells, each either 1 or 0 (i.e., discrete levels, 0 being off and 1 being on at the maximum well capacity). The 29th input parameter represents the percentage of nonpumping wells.

Generate training realizations

↓

Run training flow-and-transport simulations in parallel

↓

Train ANNs to predict objectives of remediation (i.e., outcome of flow and transport simulations)

↓

Train ANN 1 for objective 1 | Train ANN 2 for objective 2 | Train ANN 3 for objective 3

↓

Test ANN 1 | Test ANN 2 | Test ANN 3

↓

Initialize population of pumping realizations with fitness measure as a sum of all three objectives

↓

Evaluate population fitness for objective 1 (ANN 1) | Evaluate population fitness for objective 2 (ANN 2) | Evaluate population fitness for objective 3 (ANN 3)

↓

Sum fitness

↓

Generate new population

↓

Continue until optimality

↓

Validate the top ranked realizations

Ground water flow and transport code (GFTC)
• boundary conditions
• domain of interest
• initial conditions
• 2-D finite difference/ finite element code was used

Artificial neural network (ANN)
• input: vector of pumping realization
• output: objectives of remediation
• conjugate gradient back propagation learning was used

Genetic algorithm (GA)
• management formulation translated into fitness measure
• three operations were used: reproduction, crossover, and mutation

GFTC

algorithm of backpropagation was used to train the ANN. The conjugate gradient method[19] and weight-elimination procedures[20] were used to speed convergence and improve performance, respectively. The highest performing ANNs' ability to predict unseen patterns was 0.99 or better for all objectives (Table 1).

Once trained, the three ANNs were linked to a GA. A population of 239 pumping realizations was cycled through 20,000 generations of the GA using the three basic operators: 1) reproduction, 2) crossover, and 3) mutation (Table 2). The reproduction operator ranks parameter strings according to their objective function values. It is called the fitness function in analogy to the Darwinian survival of the fittest. Parameter strings are analogous to chromosomes or string creatures, with the objective function as the arbiter of survival[18]. The crossover operator determines how the string creatures pair for mating and where they crossover. The crossover site between two mated strings becomes the point at which information on the first string before the

site is paired with information on the second string after the site and vice versa. The somewhat controversial mutation operator is encouraged on the theory that reproduction and crossover may lose some valuable genetic material. In this simple GA, mutation was defined as a random alteration of a variable in a string and was used sparingly. The use of a GA had several advantages. First, the search began from a population of parameter realizations, not from a single realization. Second, GAs use probabilistic rules, not deterministic rules, to perturb the realizations. Third, the objective function information is used directly, rather than derivatives or other secondary descriptors.

Our results show that having all 28 wells on for 50 years will prevent plume migration to the west of LLNL and reduce on-site contamination at a cost of $155 million. However, the three most optimal scenarios of the 4 million examined would also contain the plume with little difference in mass removal (Figure 1b).

Table 1. ANN training and testing performance. Architecture represents the structure of the problem as presented to the ANN. An architecture of 29-7-1 allows the network to capture more nonlinearity, if any is present in the input data, than one of 29-2-1. When compared to the predictive accuracy of standard linear regression equations, the network performed equivalently to the regression technique on the most linear objective (#3) and outperformed it on the two less linear objectives (#1 and #2). Number of Function Evaluations indicates how long the network is allowed to refine its weights to reflect idiosyncrasies in the training input. Note the possibility of overfitting, i.e., training memorizes noise in the training set, resulting in lower testing performance. Training Accuracy is the correlation between the ANN's and the GFTC's prediction of each objective function for the 275 patterns used to train the ANNs. Generalization Accuracy is the same correlation for 50 test patterns new to the ANNs.

Architecture	Number of Function Evaluations	Training Accuracy	Generalization Accuracy
ANN #1: Regulatory Constraint			
29-7-1	1000	0.969	0.999[a]
29-7-1	5000	0.989	0.982
29-2-1	1000	0.905	0.859
29-2-1	5000	0.972	0.902
ANN #2: Total Mass ⌐ Solute Extracted			
29-7-1	1000	0.935	0.995[a]
29-7-1	5000	0.960	0.856
29-2-1	1000	0.861	0.939
29-2-1	5000	0.932	0.725
ANN #3: Cost of Remediation			
29-7-1	1000	0.997	0.999
29-7-1	5000	0.999	0.997
29-2-1	1000	0.999	0.999
29-2-1	5000	0.999	0.999[a]

[a]Indicates which ANN was linked to the GA.

Our earlier work on a 20-well, single-objective remediation example showed that the GA-ANN approach could solve the problem for an investment in GFTC calls comparable to one run of a traditional steepest-descent search[7,8]. The advantages were that, in the GA-ANN approach, the GFTC calls could be run in parallel on different machines, and results could be reused for different management formulations. For the 28-well multiple-objective remediation, it is impractical to duplicate this comparison with traditional methods. Approximately 300–400 sequential calls to the GFTC, totaling 1600–1800 CPU hours, would be needed for one run. Even if all convergence parameters had been perfectly set, several such runs would be required to avoid entrapment in local minima, thus resulting in several months of computational time. In contrast, the GA-ANN approach, with access to 10 machines, could complete the training-testing GFTC investment in three days. The actual

training of the ANNs took approximately five minutes, and the GA-ANN ran through four million scenarios in three days. Even budgeting for many training and testing experiments, the GA-ANN computational burden for this problem would be on the order of two weeks.

Although GA-ANN management methodology is in its infancy, it can be a robust and flexible tool to aid decision-making in groundwater remediation. The advantages of our approach are particularly promising for field-scale applications. This approach begins and ends with the classical basis of groundwater transport prediction, the GFTC. However, harnessing of the abilities of the GA and ANN allows several million pumping patterns to be examined with the same computational burden that traditional steepest-descent optimization algorithms would use to evaluate several hundred pumping patterns. This renders heretofore intractable field-scale applications practical.

Table 2. GA population generation. The table shows the transfer of information from one generation to the next according to the GA, with a mating example for pattern #1. The fitness parameter for the GA generations was a linear combination of three objectives: 1) whether the plume had been constrained behind selected monitor wells (value of 1 or 0), 2) the normalized and scaled (0.1 to 0.9) mass of solute that was removed over the 50 years of remediation, and 3) one minus the normalized and scaled (again 0.1 to 0.9) value of the cost of remediation over the 50 years. A mating example between generations 2 and 3 is shown for the first pattern of the population of 239. Pattern #1 made it into the mating pool and randomly drew pattern #110 as a mate to crossover after the fifth digit site, resulting in two new patterns in generation 3.

Current Generation

No.	Pattern of pumping (generation = 2)	Pumping param.	Prob. of selection	Regulatory constraint	Mass	Cost	Measure of fitness
#1	1 1 1 1 1 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0	0.29	0.024	1	0.296	0.646	1.943
			Total fitness of population		81.75		

Mating Pool After Reproduction

		Random mate	Random crossover site
#1	1 1 1 1 1 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0	110	5

Crossover Operator

Parent 1 #1	1 1 1 1 1	
	1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0	
Parent 2 #110	1 1 0 0 0	
	1 0 0 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 0 0	
Child 1	1 1 1 1 1 1 0 0 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 0 0	
Child 2	1 1 0 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0	

New Generation

No.	Pattern of pumping (generation = 3)	Pumping parameter
#1	1 1 1 1 1 1 0 0 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 0 0 0	0.21
#2	1 1 0 0 0 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 1 1 1 1 1 1 0 0	0.39

5

References

1. Ahlfeld, D.P., J.M. Mulvey, and G.F. Pinder, Contaminated groundwater remediation design using simulation, optimization, and sensitivity theory number 1: Model development, Water Resour. Res. 24(3):431–441, 1988a.

2. Ahlfeld, D.P., J.M. Mulvey, and G.F. Pinder, Contaminated groundwater remediation design using simulation, optimization, and sensitivity theory number 2: Analysis of a field site, Water Resour. Res. 24(3):442–452, 1988b.

3. Gorelick, S.M., C.I. Voss, P.E. Gill, W. Murray, M.A. Saunders, and M.H. Wright, Aquifer reclamation design: the use of contaminant transport simulation combined with nonlinear programming, Water Resour. Res. 20(4):415–427, 1984.

4. Gorelick, S.M., A review of distributed parameter groundwater management modeling methods, Water Resour. Res. 19(2):305-319, 1983.

5. Lefkoff, L.J., and S. Gorelick, Design and cost analysis of rapid aquifer restoration systems using flow simulation and quadratic programming, Groundwater 24(6):777-790, 1986.

6. Wagner, B.J., and S.M. Gorelick, Reliable aquifer remediation in the presence of spatial variable hydraulic conductivity: From data to design, Water Resour. Res. 25(10):2211-2225, 1989.

7. Rogers, L.L., and F. U. Dowla, Optimization of groundwater remediation using solute transport modeling and artificial neural networks, Lawrence Livermore National Laboratory, Livermore, CA, UCRL-JC-109016, 1992.

8. Rogers, L.L., Optimal groundwater remediation using artificial neural networks and the genetic algorithm, Ph.D. dissertation, Stanford University, Stanford, CA, 1992.

9. Dresen, M.D., E.M. Nichols, W.A. McConachie, K.S. Buchanan, and W.F. Isherwood, Remedial alternatives for VOCs in ground water west of LLNL, Lawrence Livermore National Laboratory, Livermore, CA, UCAR-10202, 1987.

10. Thorpe, R.K., W.F. Isherwood, M.D. Dresen, and C.P. Webster-Scholten (Eds.), CERCLA Remedial Investigations Report for the LLNL Livermore Site, Lawrence Livermore National Laboratory, Livermore, CA, UCAR-10299, 1990.

11. Tompson, A.R.B., E.M. Nichols, P.R. McKereghan, and M.C. Small, Summary of preliminary ground water simulations in the Livermore Regional Modeling Study: CFEST finite element code, Lawrence Livermore National Laboratory, Livermore, CA, UCRL-AR-107049, 1991.

12. Voss, C.I. , A finite-element simulation model for saturated-unsaturated, fluid-density-dependent groundwater flow with energy transport or chemically reactive single-species solute transport, U.S. Geological Survey, Water Resources Investigations Report 84-4369, 409p, 1984.

13. Hopfield, J.J., and D.W. Tank, Neural computation of decisions in optimization problems, Bio. Cybern. 52:141-152, 1985.

14. Hegde, S.U., J.L. Sweet, and W.B. Levy, Determination of parameters in a Hopfield/Tank computational network, Proc. International Conf. on Neural Networks II:291-298, July 1988.

15. Rumelhart, D.E., J.L. McClelland, and the PDP Research Group, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1, MIT Press: Boston, 1986.

16. McClelland, J.L., D.E. Rumelhart, and the PDP Research Group, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2, MIT Press: Cambridge, MA, 1986.

17. Holland, J., Adaptation in natural and artificial systems, MIT Press: Cambridge, MA, 1992.

18. Goldberg, D.E., Genetic algorithms in search, optimization, and machine learning, Addison-Wesley: Reading, MA, 1989.

19. Johansson, E.M., F.U. Dowla, and D.M. Goodman, Backpropagation learning for multi-layer feed-forward neural networks using the conjugate gradient method, International Journal of Neural Systems 2:291-301, 1992.

20. Weigend, A.S., D.E. Rumelhart, and B.A. Huberman, Generalization by weight-elimination with application to forecasting, Advances in Neural Information Processing 3 (NIPS'90), edited by R.P. Lippmann, J. Moody, and D.S. Touretzky, San Mateo, CA, Morgan Kaufmann, 1991.

Acknowledgments

DATE
FILMED
12 / 7 / 93

END