Practical Remedial Design Optimization for Large Complex Plumes

Richard C. Peralta, M.ASCE[1]; Ineke M. Kalwij, M.ASCE[2]; and Shengjun Wu[3]

Abstract: Powerful simulation/optimization (S/O) models exist for designing groundwater well systems and pumping strategies. However, it can be challenging to use S/O modeling effectively for large, complex, and computationally intensive problems within project time and cost constraints. Here, we present a generic two-stage optimization procedure for making S/O modeling more practical. Application is illustrated for developing optimal transient 30-year pump-and-treat designs for Blaine Naval Ammunition Depot (NAD), Nebraska, and using an innovative hybrid advanced genetic algorithm with tabu search features (AGT). AGT includes standard genetic algorithm and tabu search features plus healing, elitism, threshold acceptance, and a new subset/subspace decomposition optimization. The screening stage simplifies the optimization problem, and selects desirable remediation wells from among many candidates. During this stage, computational effort is lessened by reducing the number of state variables needing evaluation, and the solution space dimensionality (including temporal dimensions). Subset/subspace decomposition optimization of steady flow rates is used to identify desirable sets of candidate wells. The transient optimization stage develops mathematically optimal time-varying pumping rates for well subsets identified by the screening stage. It also includes reoptimization using the original objective function plus goal programming to increase strategy robustness. Initializing the AGT with feasible solutions reduces computational effort. Within a short period the procedure developed optimal pump and treat system designs for NAD. The procedure yields better objective function values than trial and error. Because optimization causes tight constraints, the computed strategy is sensitive to changes in model parameters. Increasing strategy robustness via AGT and goal programming degrades the value of the initial objective function.

DOI: 10.1061/(ASCE)0733-9496(2008)134:5(422)

CE Database subject headings: Groundwater management; Optimization; Groundwater pollution; Remedial action; Plumes; Wells; Pumps.

Introduction

S/O modeling is an effectual tool for preparing mathematically cost-effective and environmentally sustainable groundwater quality management solutions. For example, Aly and Peralta (1999a,b), Johnson and Rogers (1995), McKinney and Lin (1995), and Peralta et al. (2003) develop optimal pump-and-treat designs. Minsker and Shoemaker (1998), Liu and Minsker (2004), Shieh and Peralta (2005), Smalley et al. (2000), and Yoon and Shoemaker (2001) optimize in situ bioremediation. Cieniawski et al. (1995), Reed and Minsker (2004), and Wagner (1999) optimize the groundwater monitoring design.

S/O models use simulation modules to predict system response to stimuli, and use optimization algorithms to obtain an optimal solution. Developing optimal remediation pumping strategies requires identifying the best well locations (possibly including injection and/or extraction wells) and pumping rates. Groundwater remediation problems can be highly nonlinear and complex, and the solution can be computationally intensive. Often, they are best solved using heuristic optimization (HO) and hybrid modeling techniques, because these more easily avoid entrapment in locally optimal solutions than traditional nonlinear programming (Marryott et al. 1993; McKinney and Lin 1994; Johnson and Rogers 1995; Aly and Peralta 1999a; Yoon and Shoemaker 1999). HO includes genetic algorithms (GAs), simulated annealing (SA), and tabu search (TS). Monographs by Goldberg (1989), Laarhoven and Aarts (1987), and Glover and Laguna (1997) provide details on GA, SA, and TS, respectively. Hybrid models generally improve optimization efficiency by combining two analytical and/or HO techniques (Zheng and Wang 1999; Hsiao and Chang 2002; Shieh and Peralta 2005; Espinoza et al. 2005).

Optimizing complex remediation problems involves selecting candidate remediation well locations from hundreds or thousands of possible well locations, and pumping rates, yielding an infinite number of solutions. For simple sites and problems, such as presented by Umatilla Army Ammunition Depot (site is described in Becker et al. 2006), well location and pumping rates can be optimized simultaneously. However, sometimes groundwater flow and contaminant transport simulation algorithms require much time, and the problem might require solution within a predefined period. In such cases, considering all possible candidate well locations during optimization might not be temporally efficient, because it slows optimizer convergence to the best subset of wells.

[1]Professor, Biological and Irrigation Engineering Dept., Utah State Univ., 4105 Old Main Hill, Logan, UT 84321-4105 (corresponding author). E-mail: richard.peralta@usu.edu

[2]Simulation/Optimization Laboratory, Biological and Irrigation Engineering Dept., Utah State Univ., 4105 Old Main Hill, Logan, UT 84321-4105. E-mail: imk@cc.usu.edu

[3]Simulation/Optimization Laboratory, Biological and Irrigation Engineering Dept., Utah State Univ., 4105 Old Main Hill, Logan, UT 84321-4105. E-mail: swu@frontiernet.net

Note. Discussion open until February 1, 2009. Separate discussions must be submitted for individual papers. The manuscript for this paper was submitted for review and possible publication on May 26, 2006; approved on August 20, 2007. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 134, No. 5, September 1, 2008. ©ASCE, ISSN 0733-9496/2008/5-422–431/$25.00.

Here we demonstrate a generic procedure, helpful for solving complex remediation problems within a fixed time period. The procedure entails: (1) a screening stage that uses subset/subspace decomposition optimization and steady stimuli (pumping) simulations to identify candidate decision variables (remediation wells); and (2) a transient optimization stage that optimizes time-varying pumping rates for subsets of candidate decision variables (remediation wells). Here the optimizer applied in both stages is an innovative advanced genetic algorithm with tabu search features (AGT).

The procedure is applied to develop mathematically optimal well systems and pumping strategies for managing trichloroethylene (TCE) and trinitrotoluene (TNT) plumes at NAD. It is used to solve three optimization problem formulations. The first two formulations minimize cleanup and containment cost. The third formulation minimizes the maximum total pumping of any period needed for containment. The procedure presented herein was applied for solving these formulations between June and September 2002, as part of a project for the Environmental Security Technology Certification Program (ESTCP). The goal of the project was to demonstrate the benefit of applying S/O modeling for solving transport problems compared to a trial-and-error approach that uses simulation models only. Blaine optimization problem formulations were prepared by Blaine Naval Ammunition Depot (NAD) personnel working with the ESTCP team. Contract deliverables were feasible optimal strategies having the best objective function possible. To avoid qualitative comparison, the contract specified that sensitivity analysis would not need to be performed. In other words, an actual implementable design was not to be submitted. Instead, the task was to use optimization to prepare a preliminary design that subsequently could be made more robust based upon professional experience. For this paper, we performed sensitivity analyses for computed optimal strategies. For one formulation we use goal programming with AGT to modify least-cost pumping rates to increase strategy robustness, and show the attendant cost increase.

The ESTCP project completion report states that ". . . applying the transport optimization algorithms to these complex real-world sites . . . required expertise to limit the potential solution space to be searched . . . These approaches require substantial expertise and professional insight" (ESTCP 2004b). The fact sheet describing project results concludes with "Computational complexity still poses challenges for transport simulation/optimization, and expertise is required in the posing and solving of the problems" (ESTCP 2004a). Becker et al. (2006) summarize project results, without discussing details.

This paper responds to the above needs and similar project participant requests. It tells how to simplify the application of S/O models to optimization problems. It is especially useful when designing under temporally stressful circumstances. The project and this paper assume all desirable characteristics of optimal solutions are represented in the optimization problem being solved. Although the paper does not emphasize other characteristics sometimes important in selecting a pumping strategy or design, the same simplifications can apply for stochastic optimization.

Optimization Design Procedure

Overview

The two-stage activities discussed are most beneficial when addressing large complex, and computationally intensive optimization problems that require time-varying decision variable strategy solutions, especially when parallel processing is not an option. For such problems, forcing a transient optimizer to consider all possible decision variables is inefficient—the number of possible permutations becomes prohibitively large. The proposed procedure employs modeler experience and tailored algorithms to avoid permutations not needing simulation—greatly reducing computation time. Table 1 lists the generic processes used in the screening and transient optimization stages. Both stages employ an innovative advanced GA having some TS capabilities (AGT), discussed in the next subsection. The screening stage also uses a new AGT subset/subspace decomposition optimization feature. The transient optimization stage uses more AGT capabilities to compute time-varying pumping rates and incorporates goal programming to enhance design robustness.

AGT and Subset/Subspace Decomposition Optimization Algorithms

A new AGT algorithm (Fig. 1) develops continuous domain time-varying solutions for multiple stress periods and management periods simultaneously. AGT employs standard GA operations such as parent selection, crossover, mutation, and advanced features including elitism and healing. With elitism only the M best strategies to date are used in the parent selection. Healing ensures that a new pumping strategy created via crossover and mutation satisfies decision variable-based constraints. AGT handles violation of state variables constraints by adding penalties to the objective function value proportional to the degree of constraint violation a pumping strategy causes. AGT includes a threshold acceptance features that forces the algorithm to only simulate a strategy having an unpenalized objective function value that is at least a predefined value better than the best objective function value to date. This feature applies when the objective function is based solely on decision variables.

AGT incorporates some TS capabilities to intensify search in the solution space region that potentially yields superior strategies, and to avoid regions that yield inferior results. TS remembers the best strategies (or elite strategies) to date and develops strategies in the neighborhood of the elite strategies by allowing only elite strategies as parents and setting an upper limit on the solution space size. Further, TS maintains a list of tabu (inferior) strategies, and ensures that there is a minimum acceptable distance (search coarseness) between a newly developed strategy and the tabu strategies. The AGT with subset/subspace decomposition optimization is well suited for the two-stage procedure, although other optimizers might be used.

The new subset/subspace decomposition optimization algorithm is hereafter referred to as subset optimization (Fig. 2). It prepares and contrasts optimal steady pumping strategies to select desirable subsets of candidate wells for subsequent transient optimization. Subset optimization initially considers all defined candidate well locations in the entire study area or in selected subregions. To best explore the potentially very nonlinear decision space, subset optimization employs AGT. After a specified number of generations, the algorithm ranks subsets based on either the number of feasible strategies or best objective function (OF) values. Then it develops preliminary feasible and optimal strategies for the first N-ranked subsets sequentially (where N is a prespecified number). Subset optimization outputs are N optimal strategies for N subsets of remediation well combinations.

Table 1. Overview of Screening and Transient Optimization Stages

	Process	Process action	Output
Screening (A–D)	A. Select preliminary sets of decision spaces	A1. Analyze optimization problem	
		A1.1. Analyze modeled system (simulation model)	
		A1.2. Analyze objective function and its components	
		A1.3. Identify constraints/constraint locations, degree of complexity, and contaminant hotspots	
		A1.4. Determine how decision spaces and variables affect modeled system	
		A1.5. Identify areas to which stimuli should not be applied	
	B. Simplify optimization problem	B1. Reduce effort needed to identify satisfactory system stales (feasible solutions)	
		B1.1. Combine simulated system components	
		B1.2. Reduce number of state variables needing evaluation and comparison with bounds or constraints	
		B1.3. Apply decomposition	
		B2. Reduce solution space (decision variable) dimensionality	
		B2.1. Identify optimal stress periods of goal-achievements	
		B2.2. Use same stimulus value for a decision variable in different stress periods	
	C. Perform screening simulations	C1. Simulate within multiple subsets of decision dimensions and steady stimuli strategies	
		C2. Evaluate each simulation based upon of value and constraint violations	
		C3. Rank simulations based on penalized of value	
	D. Perform subset/subspace decomposition optimization	D1. Select potential candidate decision variables locations and combinations	
		D2. Initialize optimizer with feasible strategies	
		D3. (a) Optimize decision variable values for multiple subsets of decision dimensions. (b) Sequentially develop preliminary feasible and optimal strategies for the first N-ranked subsets	Multiple subsets of decision spaces and decision variable values
Transient optimization (E)	E. Employ optimizer	E1. Initialize optimizer with feasible strategies	Optimal pumping strategy
		E2. Optimize one subset of decision variables at a time	
		E3. Perform a postoptimization sensitivity analysis	
		E4. Increase pumping strategy robustness (reoptimization):	
		E4.1. Tighten the optimization problem formulation	More robust optimal pumping strategy
		E4.2. Invoke goal programming in combination with the primary objective function (multiobjective optimization)	

Screening Stage

In this stage the modeler: (1) selects sets of decision space dimensions or (in the context of this research) candidate well locations; (2) assesses how to simplify the optimization; (3) employs screening simulation model(s) (Table 1); and (4) employs AGT subset optimization to finalize subsets of candidate wells. To select decision variable dimensions (candidate locations) one uses study area and optimization problem formulation information (Table 1, A1.1–A1.5). To reduce computational effort, and increase likelihood of developing an optimal strategy, one uses simplification concepts (Table 1, B1–B2). Simplification reduces the numbers of individual constituent transport simulations performed, postsimulation processing, and decision space dimensionality. Simplification makes it easier for subsequent Processes C and D (Table 1) to rule out inferior well locations and determine potentially good candidate well locations.

Process Action B1 aims to make it easier to identify strategies that yield solutions feasible with respect to state variables. Via action B1.1, one considers simulating only indicator or composite contaminants. This might be a lesser number than that for which the transport simulator was calibrated.

For example, reasonably similar contaminants can be combined into a composite plume that used weighting based on cleanup standards. This requires normalizing the concentration of each constituent to a representative concentration level of the surrogate contaminant according to the ratio of the cleanup levels (Geotrans 2002)

$$C_\alpha = C_\alpha \frac{\mathrm{CL}_\beta^s}{\mathrm{CL}_\alpha} \tag{1}$$

where C_α=concentration of contaminant α; CL_β^s=cleanup level for surrogate contaminant β, and CL_α=cleanup level for contaminant α.

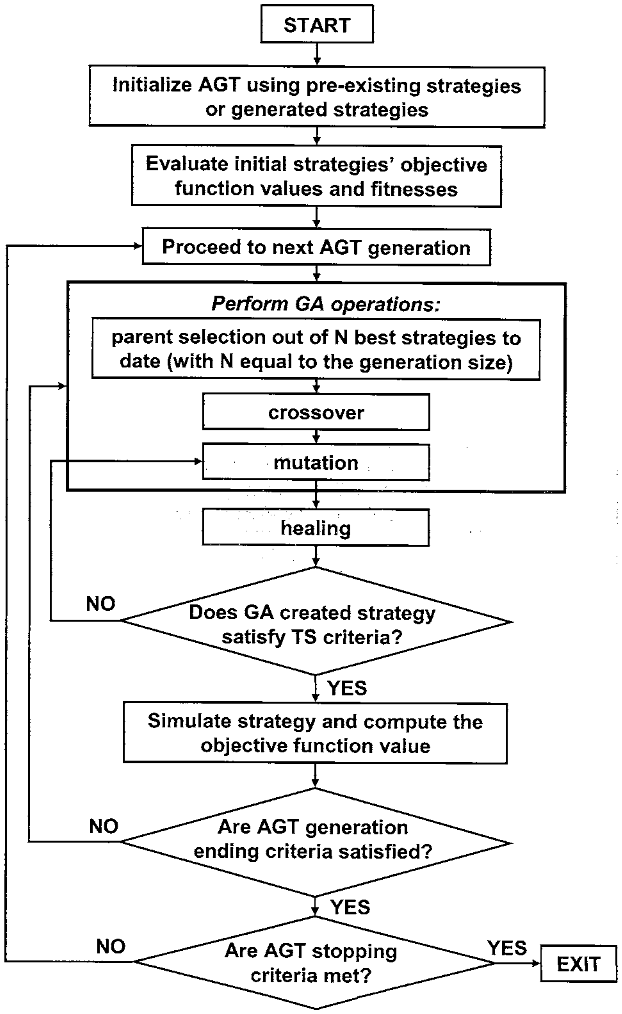

Fig. 1. AGT flow chart

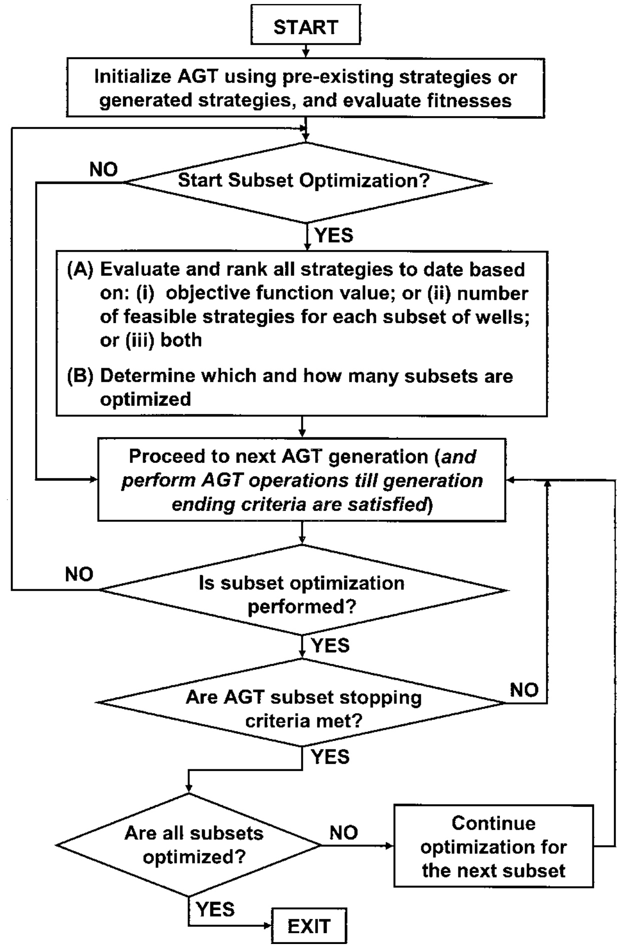

Fig. 2. Subset optimization flow chart

Action B1.2 urges the reader not to automatically use extraneous variables, especially global concentration variables, to satisfy curiosity. A global max concentration variable is the maximum concentration that exists anywhere (in any and all cells) within a specified zone (group of cells) at a particular time. Thus, although only one state variable might be needed for a cleanup constraint, determining that value requires much processing. On the other hand, in hydraulic optimization, computing or constraining a head value is trivial and very fast if superposition is used.

Action B.1.3 suggests reducing the number of needed transport simulations by decomposition. Decomposition can be illustrated by a multiplume situation, in which the plumes can be treated somewhat separately although by the same pump-and-treat system. For example, one can begin by developing a reasonably optimal strategy to address a Contaminant Number 1, and using that strategy near Contaminant 1 while then addressing Contaminant Number 2. One does not simulate Contaminant 1 transport while initially addressing Contaminant 2. Only after a good solution is obtained for Contaminant 2 might one also simulate transport of both contaminants during final optimization.

Process Action B2 aims to eliminate regions of the solution (decision) space that will need to be searched during the transient optimization stage. This reduces the number of strategies under consideration. Action B2.1 suggests using experience and preliminary analyses to identify the particular times (stress periods) at which optimization goals are best achieved. For example, assume a situation in which one seeks to minimize the present value cost of achieving cleanup by the end of period four. Assume one can determine, from input cost and other data, that the minimal cleanup cost will occur by achieving cleanup during period three. In that case, Action B2.1 suggests that one allow the optimization algorithm to consider only such strategies by imposing a cleanup constraint on period three, and omitting the cleanup constraint for other stress periods. One will not allow the optimizer algorithm to consider strategies that achieve cleanup in preceding stress periods, and one would simulate only three periods, not four. In effect, one performs preliminary simulations and evaluations to determine the time at which one a constraint should be optimally satisfied, and then adds or changes constraint(s) to the initially posed optimization problem to assure that occurs. The intent is that adding such constraints reduces optimizer search efforts in a potentially very nonlinear solution space.

Action B.2.2 urges reducing decision space dimensionality by forcing strategies from multiple periods to have the same values. For a six period problem, for example, one can force the first four periods to employ the same decision variable values, and the last two periods to use different values. Usually one allows rates to change when system state changes occur. For example, in a multiwell distributed plume cleanup problem, one might employ steady pumping rates in all periods until cleanup is achieved in part of the plume, but one would allow rates to change in subsequent periods. Such constraints might be different during the screening and transport optimization stages.

Process C (Table 1) employs a screening simulation module that: (1) simulates multiple subsets of predefined well locations and steady pumping strategies; (2) evaluates each strategy/simulation based on the OF value and constraint violations; and (3) ranks simulations based on penalized OF value.

Process D (Table 1) emphasizes subset optimization. Based on previous analysis and simulations, AGT receives feasible solutions for selected candidate well location subsets. AGT optimization yields the *N* best subsets of remediation wells.

Because Table 1 activities and processes can be interrelated, for a particular optimization problem formulation, they might be performed simultaneously and/or sequentially. Recommendations are adaptable to parallel processing situations.

Transient Optimization Stage

The transient optimization stage (process E in Table 1) follows the screening stage. Transient optimization's main output is the optimal pumping strategy (resulting from the best subset of candidate well locations). This stage begins by using the candidate wells and feasible pumping strategies developed during the screening stage. It can use different subsets of candidate wells in different optimization runs. If time is available, one can perform AGT optimization for each of the *N* best subsets of wells resulting from subset optimization in the screening stage. If time is limited, one might perform AGT only for the best subset of wells.

Beginning with good feasible solutions greatly reduces the number of simulations needed to obtain refined optimal strategies. In essence, during the screening stage one employs experience and practical knowledge to simplify the problem that the optimizer must solve, and speed solution during the transient optimization stage. A variety of optimization algorithms can be used to develop optimal transient solutions after candidate wells are selected and feasible strategies are created. Guidance concerning transient optimizer selection is beyond the scope of this paper.

Ideally a developed optimal strategy is also robust—i.e., it will satisfy constraints and achieve specified goals in the field, even if the field physical system differs somewhat from the assumed model system. However, a computed optimal strategy might not be robust if the stochastic nature of aquifer parameters is not incorporated within the optimization. Stochastic optimization is infrequently used because data are usually not available to develop the reasonable probability density functions or realizations that stochastic optimization needs.

One way to increase strategy robustness is to tighten the optimization problem formulation—causing the optimizer to try to achieve lower concentrations in model cleanup and exclusion zones than are actually required in the field. This approach, used here, employs multiobjective optimization by coupling concentration goal programming with the original objective function.

Application and Results

NAD Study Area and Optimization Problem Formulations

Blaine NAD is located east of Hastings, Neb. NAD was placed on the Environmental Protection Agency's national priority list in 1986 due to groundwater contamination. NAD has significant groundwater contamination by volatile hydrocarbons and explosives from solid waste and explosives disposal and wastewater discharge (Geotrans 2002). There is currently no pump-and-treat system installed at NAD. Selected NAD characteristics and model features are presented in Table 2. Groundwater flow and solute transport are simulated employing a modular finite-difference groundwater flow model, MODFLOW (McDonald and Harbaugh 1988) and a modular three-dimensional transport model, MT3DMS (Zheng and Wang 1998). The MT3DMS transport model is designed for simulating TNT and TCE transport in layers 1–5. Layer 6 has no contamination. During the nonirrigation season flow is predominantly to the east and southeast with an average hydraulic gradient of 0.001. Irrigation season pumping alters groundwater flow directions significantly horizontally and vertically. Individual cones of depression around irrigation wells caused many small capture zones to be formed. Initial TCE and TNT plumes are presented in Figs. 3 and 4, respectively.

The Formulation 1 goal is to minimize cost of containing and remediating TCE and TNT plumes within 30 years. The cost OF is (Geotrans 2002) Minimize

$$Z = (\mathrm{CCE} + \mathrm{CCT} + \mathrm{CCD} + \mathrm{FCM} + \mathrm{FCS} + \mathrm{VCE} + \mathrm{VCT} + \mathrm{VCD}) \tag{2}$$

Subject to

Table 2. NAD Characteristics and Model Features

NAD characteristics/model features	Value
NAD area (km^2)	197.5
Model area (km^2)	357
Hydrogeological units' thickness[a,b]	
Unconfined aquifer (m)	3–5
Upper confining layer (m)	0.3–0.9
Semiconfined aquifer (m)	30–46
Model discretization	
• Number of layers	6
• Number of rows	82
• Number of columns	136
• Minimum cell size (m × m)	122 × 122
• Maximum cell size (m × m)	610 × 610
Hydraulic conductivity (*K*) range (m/day)	
• Unconfined Aq. (Layer 1)	3–24
• Upper confining layer (Layer 2)	0.0006–0.2
• Semiconfined Aq. (Layers 3–6)	46–76
Vertical K range (m/day)	
• From Layer 1 to Layer 2	0.0001–0.0015
• From Layer 2 to Layer 3	0.0001–0.059
• Between semiconfined aquifers	4.6–7.8
Number of stress periods (SPs)	60
Duration odd numbered SPs[c]	76
Duration even numbered SPs[d]	289
Planning horizon (years)	30
Number of management periods (MPs)	6
Duration one MP (years)	5 (10 SPs)
Number of irrigation wells in model	951
Combined TCE and TNT plume length (km)	12.2

[a]From the ground surface downwards.

[b]The unconfined and upper confined are discontinuous in some portions of the study area.

[c]Corresponds with the irrigation season.

[d]Coincides with the nonirrigation season.

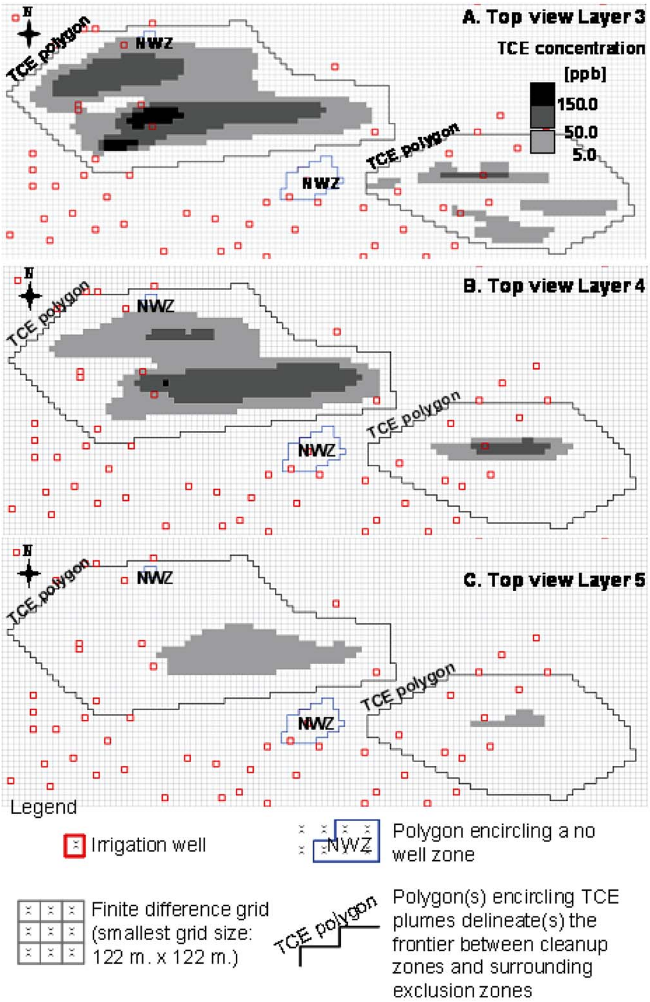

Fig. 3. Top view of partial study area showing TCE initial concentrations in: (A) Layer 3; (B) Layer 4; and (C) Layer 6

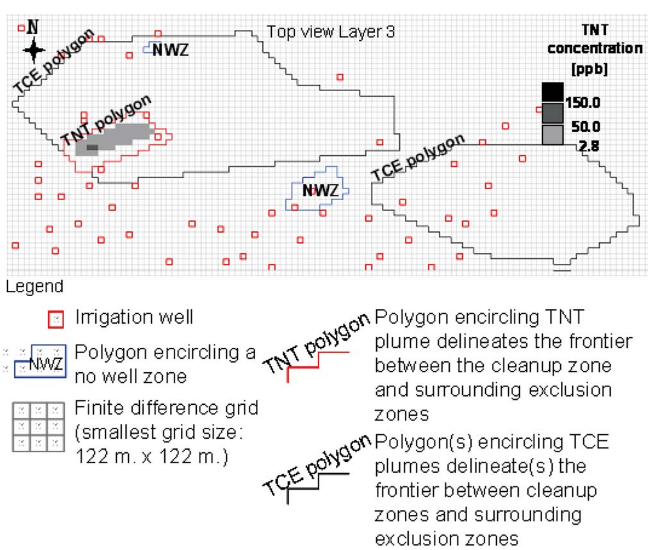

Fig. 4. Top view of partial study area showing TNT initial concentrations in Layer 3

Table 3. Cost OF Components and Values

Cost components	Description	Value (×1000)
CCE	Capital cost of new extraction wells ($)	400
CCT	Capital cost of contaminated water treatment facility ($/m^3/day)	5.72
CCD	Capital cost of discharge piping ($/m^3/day)	8.18
FCM	Fixed operation and management cost ($/year)	115
FCS	Fixed sampling and analysis cost ($/year)	300
VCE	Variable cost of well operations ($/m^3/day)	0.25
VCT	Variable cost of treatment ($/m^3/day)	1.54
VCD	Variable cost of discharge ($/m^3/day)	0.36

Note: A 3.5% discounting rate is used for each component to compute the present value of the OF value.

$$\mathrm{Conc}_{s,z,t} \leqslant \mathrm{Conc}^{U}_{s,z,t} \tag{3}$$

where cost components of Eq. (2) are detailed in Table 3; $\mathrm{Conc}_{s,z,t}$=maximum concentration for species s in zone z at time t. U=upper bound on the state variable. The cleanup goal for TCE and TNT is 5 parts per billion (ppb) and 2.8 ppb, respectively. This must be achieved within 30 years in all cells within the cleanup zones (Figs. 3 and 4). In this example the contamination must be kept within the containment zone, so cleanup zone and containment zone are synonymous. Concentration cannot exceed 5 and 2.8 ppb in the exclusion zones for TCE and TNT (Figs. 3 and 4). Both constraints are evaluated at the end of every management period (MP). Table 4 lists additional constraints and characteristics of the NAD strategy development process.

Although NAD personnel participated in posing all three optimization problem formulations, environmental regulators might or might not agree with the formulations. Some regulators might oppose including economic discounting that can favor slowing remediation. Because public funds would be used for

Table 4. Additional NAD Optimization Constraints and Characteristics

Number	Constraints/characteristics
1	Wells can only be added and pumping rates can only be changed at the beginning of modeling Years 1, 6, 11, 6, 21, and 26.
2	Upper bounds on pumping from wells screened in one, two, or three layers are 22.1 L/s (350 gpm), 44.2 L/s (700 gpm), and 66.2 L/s (1,050 gpm), respectively.
3	No remediation wells are allowed in specified restricted areas of Layer 6, and in cells with irrigation wells (Figs. 3 and 4).
4	No cell should go dry[a] (i.e., have zero saturated thickness).
5	Upper limit on number of remediation wells is 25 (only for Formulation 3).
6	The 30-year planning period is discretized into six 5-year management periods (MPs), and 60 simulation model stress periods (SPs).
7	Input data includes 60 SPs of time-varying background irrigation pumping rates, that are not subject to optimization.
8	To be optimized are timing and installation of extraction wells and pumping rates for each 5-year MP.
9	Layers 1 and 2 are excluded from optimization due to high uncertainty in contaminant concentrations.

[a]This constraint is added because the MT3DMS transport code could not numerically handle dry cells.

Table 5. Optimization Results for Formulations 1–3

Formulation	OF (million \$)	OF (L/s)	Number of remediation wells	TCE cleanup (years)	TNT cleanup (years)	Total pumping range (L/s)	Total pumping range (gpm)	Improvement from trial-and-error approach (%)
1	40.8	—	10	30	29	156.8–213.1	2,486–3,378	19
2	18.9	—	10	30	29	156.8–213.1	2,486–3,378	33
3	—	134.9	25	n.a.	n.a.	134.3–134.9	2,129–2,139	26

[a]n.a.=not available.

remediation, other regulators might consider it desirable to expend as few funds as needed to appropriately satisfy environmental regulations.

Formulation 2 has the same OF and constraints as Formulation 1 except that it assumes diversion of 151.4 L/s (2,400 gal/min) of extracted water. This rate is therefore omitted from the treatment or discharge cost:

$$\text{If } (Q_{\text{MAX}} \leqslant 157.7 \text{ L/s}) \text{ then CCT} = 0$$

$$\text{If } (Q_{\text{MAX}} > 157.7 \text{ L/s}) \text{ then CCT} = 1.0^{*}[Q_{\text{MAX}} - 151.4 \text{ L/s}] \tag{4}$$

where Q_{MAX}=maximum total flow of water extracted by remediation wells in any MP.

The Formulation 3 goal is to minimize the maximum total remediation pumping rate in any management period of a 30-year simulation (min-max OF)

$$\min(Q_{\max}) \tag{5}$$

Formulation 3 includes the same constraints as Formulation 1 (Table 4), except that the cleanup constraint does not apply here, and limits the maximum number of new remediation wells to 25.

Formulation 1

Screening Stage

To attempt to reduce temporal dimensionality, we reviewed the optimization problem formulation and effect of discounting. This indicated that one should delay well installation and pumping to the extent practical, i.e., achieve cleanup as close to Year 30 as possible. Hence we only made 30-year simulations.

Reducing the number of state variables needing evaluation when each strategy is tested involved two steps. First, Geotrans (2002) used the previously mentioned weighting procedure to combine TCE, 1,1-dichloroethene (DCE), and Royal Demolition Explosive (RDX) into composite TCE-dominated plumes (Fig. 3). Then, noting that the TNT plume (Fig. 4) was much smaller than the TCE plumes, our initial strategy design simulations only addressed the TNT plume. These determined that the TNT plume could be contained and remediated by installing one or two wells. Thereafter, during screening, one well pumped at a fixed rate within the TNT plume, and only TCE transport was simulated and addressed. This reduced the time required for a single 30-year flow and transport simulation by about 28%, (from 50 to 36 min on a 1.8 GHz hertz Pentium 4 central processor unit).

To attempt to reduce decision variable dimensionality, we noted that the cost of installing one well was relatively large compared with pumping cost. Thus, one would want to install as few wells as possible, even if adding a well could reduce total pumping slightly.

We took several approaches to develop sets of good candidate well locations for both the western main TCE plume, and the eastern small TCE plume. First we placed candidate wells at leading edges of the TCE plumes to achieve plume containment. Then we added candidate wells in high concentration areas to satisfy cleanup constraints, and modified both sets of positions to reduce the number of wells needed. This included installing a well in Year 20 in the northeastern part of the main TCE plume to remediate contamination entering Layer 3 from Layer 2.

Because most western main TCE plume contamination was in Layers 3 and 4, candidate wells would extract from those layers. Wells were not needed in Layer 5, because Layer 5 contamination was removed by nearby wells drawing from Layers 3 and 4. Most eastern small TCE plume contamination is in Layer 3. Candidate wells in most of this plume only extract from Layer 3. Only in the northern part of the eastern small TCE plume does a candidate well need to extract from both Layers 3 and 4.

En toto, screening involved over 200 30-year simulations using steady pumping. Subset optimization helped finalize remediation well locations in the small TCE plume. Time was not available to perform subset optimization for candidate wells in the main TCE plume.

Transient Optimization Stage

Transient optimization was performed for one subset of candidate remediation well locations. The AGT optimizer was set to develop up to 800 strategies (100 generations with eight simulations per generation), and to perform flow and transport simulation for each. One can assume that optimizing for only one subset requires only half as many simulations as optimizing for two subsets.

AGT transient optimization was initialized with a feasible steady pumping strategy having an OF value of \$48.7 million. During this stage, TCE transport was always simulated, but TNT was generally not. The candidate well within the TNT plume pumped at rates within bounds that assured satisfying TNT cleanup and containment constraints.

AGT gradually converges to a least-cost transient pumping strategy OF value of \$40.8 million. The AGT optimizer was not able to improve the OF value further because of the tightness of the TCE maximum concentrations to their upper bounds. For example, maximum TCE concentrations are 4.98 and 4.97 ppb within exclusion zones at the end of 5 and 25 years, respectively. The maximum TCE cleanup zone concentration is 4.99 ppb at the end of Year 30. Table 5 has additional strategy details. Because of economic discounting, optimization tends to increase total period pumping rate with time. For the optimal strategy, Fig. 5 shows the candidate well locations and TCE concentrations greater than 5 ppb after 25 years for Layers 3–5. For the optimal strategy, Fig. 6 shows the well locations and TNT concentrations greater than 2.8 ppb after 25 years for Layer 3.

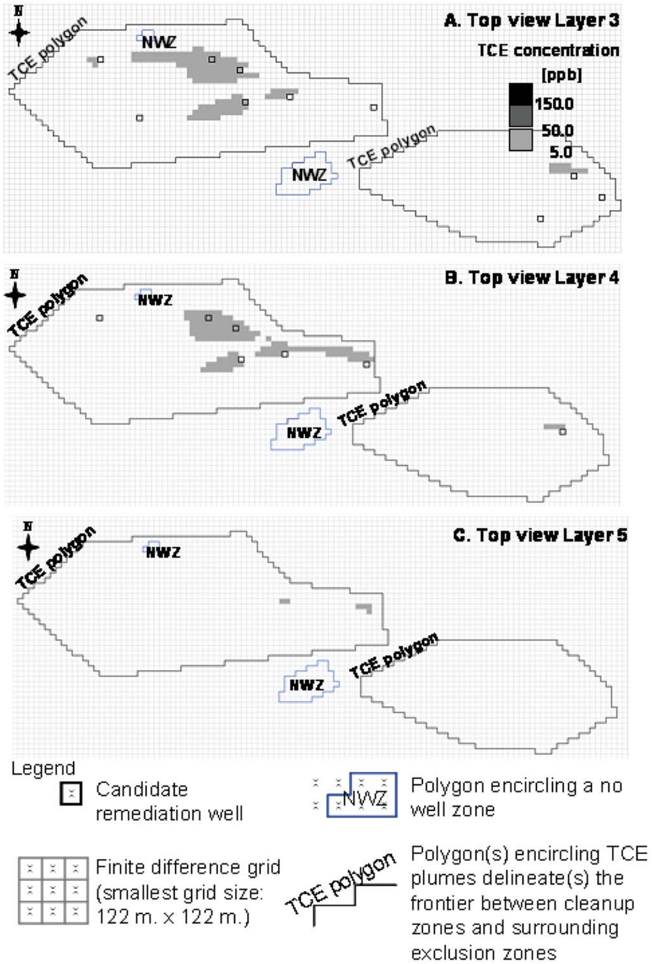

Fig. 5. Top view of partial study area showing Formulation 1 TCE concentrations >5 ppb after 25 years in: (A) Layer 3; (B) Layer 4; and (C) Layer 5

The optimal Formulation 1 strategy is 19% less costly than the strategy simultaneously developed by an experienced consultant using the traditional trial-and-error approach (Becker et al. 2006). This computed optimal pumping strategy might be globally optimal for the employed wells, but it is not robust. Sensitivity analysis showed that increasing or decreasing hydraulic conductivity (K) array values (using global multipliers as small as $\pm$ 1.00%) causes infeasible solutions.

To create a strategy more robust for conductivity, we employed an alternative multiobjective optimization problem (Formulation 1b). The Formulation 1b objective function seeks to bring TCE maximum concentration values below 4.5 ppb (instead of 5 ppb) while minimizing the remediation cost (multiobjective optimization). The optimal result of this exercise yields a cost of \$46.5 million (a 13.8% increase) and a robustness range of −5–9% (K-array multiplication factors of 0.95 and 1.09, respectively). Fig. 7 contrasts robustness of strategies developed using Formulations 1 and 1b, and illustrates the robustness of intermediate Formulation 1b optimization run strategies.

Fig. 7 demonstrates a common phenomenon—increasing robustness degrades achievement of the primary objective function goal. Here, as the maximum concentration goal is lowered, robustness increases, but total remediation cost increases. More pumping is needed to reduce the maximum concentration values. It costs \$5.7 million to obtain a 14% robustness range.

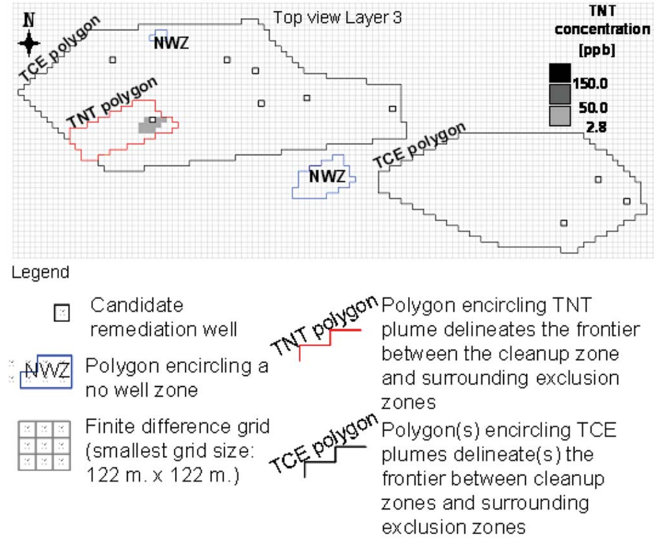

Fig. 6. Top view of partial study area showing Formulation 1 TNT concentrations >2.8 ppb after 25 years in Layer 3

Both the flow and transport simulation models contain many parameters that can be quite stochastic. For all formulations, mathematical optimization will cause cleanup and containment constraints to be tight. Changes in assumed model parameters that cause a contaminant to move more quickly or more slowly or affect concentrations within the model, can cause such constraints to be violated when a strategy developed using unaltered parameters is tested for altered parameters. Thus, in addition to changes in hydraulic conductivity, computed optimal strategies can be sensitive with respect to changes in assumed effective porosity, dispersivity, partitioning coefficient, half life, and other parameters. In design practice, sensitivity analysis includes as many as appropriate, but usually, only one parameter is changed at a time.

Formulation 2

Analyzing the OF components for Formulations 1 and 2 revealed that the optimal strategy for Formulation 1 would also be optimal for Formulation 2—diverting up to 151.4 L/s of extracted water

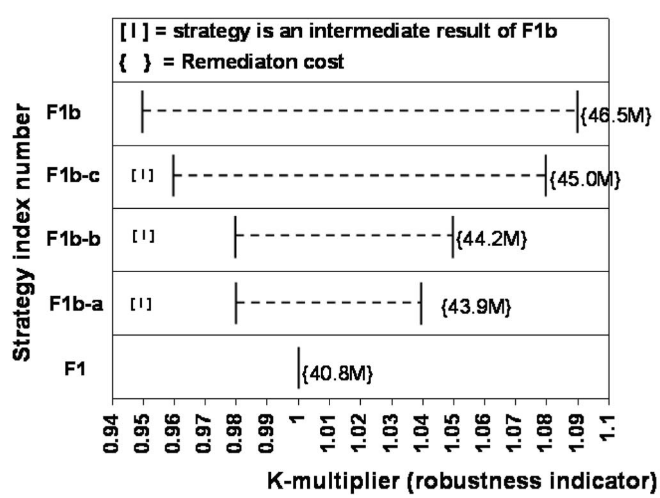

Fig. 7. Robustness of optimal strategies for Formulations 1 and 1b and of intermediate strategies

from the treatment train would significantly reduce cost but would not affect well locations and pumping rates of an optimal strategy. Therefore, we assumed that the optimal strategy for Formulation 1 is also optimal for Formulation 2 and did not perform additional screening, optimization, or sensitivity analysis for Formulation 2. Table 5 summarizes the results.

Formulation 3

Screening Stage

The screening stage for Formulation 3 was performed simultaneously with that for Formulation 1. Formulation 3 differs in: using a mini-max flow objective function, omitting cleanup constraints; and allowing up to 25 remediation wells. Because using 25 wells is a more relaxed optimization problem than using fewer wells, and cost was not to be considered, we automatically allowed 25 wells. Screening for Formulation 3 involves placing candidate wells close enough to the leading edge of the TCE plumes that they can capture it. Therefore many of the wells were positioned in lines nearly parallel to containment zone boundaries. No effort was made to reduce the number of wells because that could potentially adversely affect the objective function value (by potentially causing pumping to increase). For this problem posed in collaboration with the Blaine facility for ESTCP demonstration, obtaining the best objective function value we could for a feasible solution within time constraints was the goal. Because some employed candidate pumping wells were in cells adjacent to each other, probably fewer pumping wells could be used without significantly increasing (harming) the objective function value. There was no time or need to test that view.

Transient Optimization Stage

During transient optimization, AGT is employed for 1,200 simulations (150 generations with eight simulations per generation). During optimization only TCE transport was simulated. As in Formulation 1, the candidate well within the TNT plume pumped at rates within bounds that assured satisfying TNT cleanup and containment constraints.

Transient optimization is performed for one subset of candidate remediation well locations. It was simplified by optimizing steady pumping rates for the first five management periods, and optimizing different rates for the final period. Thus transient optimization only needed to compute two optimal rates for each well instead of six (six are needed if one allows a different rate for each MP). This reduced processing time of an optimization run by about 80%.

The AGT converges to a Formulation 3 OF value of 134.95 L/s (2,139 gal/min). Table 5 summarizes the results. The exclusion zone constraints are tight to the upper bounds at the end of Years 5, 20, and 25, with TCE maximum concentration values of 4.99, 4.99, and 4.98 ppb, respectively.

Sensitivity analysis revealed that reducing K-array values by as much as 35% (a K-array multiplication factor of 0.65) yields feasible strategies. Increasing K-array values always caused infeasible solutions—concentrations exceeding maximum concentration level occurred outside the containment zone (i.e., within the exclusion zone).

Summary and Conclusions

In real-world practice, time and money constraints restrict how much effort one can expend on a design. Designing well systems and pumping strategies for managing large complex contaminant plumes can be aided by transport optimization software. However, to obtain reasonably optimal solutions for large complicated design problems, one should use the software carefully.

One wants to intelligently reduce the size of the solution space being explored by the heuristic optimizer. Such a reduction increases the chances that the optimizer will obtain close to a globally optimal solution. Without such a reduction, the optimizer will unnecessarily explore much solution space. Reducing the size of the solution space can be accomplished in several ways.

Here we propose a two-stage process for designing optimal pump-and-treat systems and pumping strategies. The screening stage involves steady pumping simulations and optimizations. It selects candidate pumping wells and develops feasible pumping strategies to initiate the subsequent transient optimization stage. Within the screening stage one employs professional experience and judgment to reduce problem solution space and computer processing time. Basically one tries to minimize state variables considered and decision dimensions and variables. Stated differently, in the screening stage, one tries to:

1. Reduce the temporal dimensionality of the solution space. In our application during the optimization stage we only need to optimize strategies that will pump for all six management periods (MPs)—we will not optimize strategies that achieve cleanup and containment in fewer than six MPs. Several trial simulations were performed to conclude that increasing well numbers and pumping rates early in the management period would not speed cleanup sufficiently to offset the increased cost;
2. Reduce the decision variable dimensionality of the solution space. Experience identifies subsets of wells that should yield reasonably good feasible solutions. From those, subset optimization selects the best subsets of candidate wells, and develops feasible solutions (pumping strategies);
3. Reduce the number of state variables needing evaluation and comparison with constraints. We identify a well location and minimum pumping rate that is adequate for managing one contaminant species (TNT). We subsequently generally only simulate the other species (TCE) during the transient optimization stage, reducing the simulation time needed to evaluate a single strategy by about 30%. This is a decomposition technique; and
4. Reduce the number of simulations needed to get obtain a reasonably optimal strategy. We develop an initial feasible solution for each subset of candidate wells that we provide to the optimizer. During the transient optimization stage the optimizer does not waste time in trying to develop a (first) strategy that satisfies all constraints.

In essence, we use experience and practical knowledge to simplify and focus the problem that we pass to the transient optimization stage.

The transient optimization stage begins by developing optimal strategies for the subsets of candidate wells selected in the screening stage. It employs the AGT, a new optimizer that couples advanced GA with some TS capabilities and employs subset/subspace decomposition optimization, to significantly reduce solution space size and simulation numbers. The AGT computes a mathematically optimal time-varying pumping strategy for the best subset of wells. Of course, one can further change candidate well locations if indicated.

Postoptimization sensitivity analysis for the hydraulic conductivity parameter shows that Formulation 1–3 optimal results are not robust, because of the tightness of final predicted TCE concentrations to their upper bounds. An alternative optimization problem formulation (Formulation 1 plus goal programming) sought to bring maximum TCE concentrations below 4.5 ppb while minimizing remediation cost. The resulting optimal strategy is more robust (robustness range of −5%–9%), but the primary objective function value increases 13.8%.

In conclusion, heuristic optimization is a powerful tool for plume remediation design, and can yield significant benefits (>19% improvement compared to a normal trial-and-error approach). However, if deadlines are important and tight, and for complicated problems that have long simulation run time, one should try to reduce the computational effort of formal optimization. The modeler's experience and logic is invaluable for doing this, especially in the screening stage that precedes optimization.

References

Aly, A. H., and Peralta, R. C. (1999a). "Comparison of a genetic algorithm and mathematical programming to the design of groundwater cleanup systems." *Water Resour. Res.*, 35(8), 2415–2425.

Aly, A. H., and Peralta, R. C. (1999b). "Optimal design of aquifer cleanup systems under uncertainty using a neural network and a genetic algorithm." *Water Resour. Res.*, 35(8), 2523–2532.

Becker, D., Minsker, B., Greenwald, R., Zhang, Y., Harre, K., Yager, K., Zheng, C., and Peralta, R. C. (2006). "Reducing long-term remedial costs by transport modeling optimization." *Ground Water*, 44(6), 864–875.

Cieniawski, S. E., Etheart, J. W., and Ranjithan, S. (1995). "Using genetic algorithms to solve a multiobjective groundwater monitoring problem." *Water Resour. Res.*, 31(2), 399–409.

Environmental Security Technology Certification Program (ESTCP). (2004a). "Application of flow and transport optimization codes to groundwater pump-and-treat systems." *Factsheet ER-0010*, Dept. of Defense, Washington, D.C.

Environmental Security Technology Certification Program (ESTCP). (2004b). "Cost and performance report, application of flow and transport optimization codes to groundwater P&T systems." *Rep. No. CU-0010*, Dept. of Defense, Washington, D.C.

Espinoza, F. P., Minsker, B. S., and Goldberg, D. E. (2005). "Adaptive hybrid genetic algorithm for groundwater remediation design." *J. Water Resour. Plann. Manage.*, 131(1), 14–24.

Geotrans (2002). "Transport optimization." Hastings Naval Ammunition Depot, Hastings formulations documentation.

Glover, F., and Laguna, M. (1997). *Tabu search*, Kluwer Academic, Norwell, Mass.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, Reading, Mass.

Hsiao, C., and Chang, L. (2002). "Dynamic optimal groundwater management with inclusion of fixed costs." *J. Water Resour. Plann. Manage.*, 128(1), 57–65.

Johnson, V. M., and Rogers, L. L. (1995). "Location analysis in groundwater remediation using neural networks." *Ground Water*, 33(5), 749–758.

Liu, Y., and Minsker, B. S. (2004). "Full multiscale approach for optimal control of in-situ bioremediation." *J. Water Resour. Plann. Manage.*, 130(1), 26–32.

Marryott, R. A., Dougherty, D. E., and Stollar, R. L. (1993). "Optimal groundwater management. 2: Application of simulated annealing to a field-scale contamination site." *Water Resour. Res.*, 29(4), 847–860.

McDonald, M. D., and Harbaugh, A. W. (1988). "A modular three-dimensional finite-difference groundwater flow model." *U.S. Geological Survey techniques of water-resources investigations*, Book 6, Chap. A1, USGS, Denver.

McKinney, D. C., and Lin, M.-D. (1994). "Genetic algorithm solution of groundwater management models." *Water Resour. Res.*, 30(6), 1897–1906.

McKinney, D. C., and Lin, M.-D. (1995). "Approximate mixed-integer nonlinear programming methods for optimal aquifer remediation design." *Water Resour. Res.*, 31(3), 731–740.

Minsker, B. S., and Shoemaker, C. A. (1998). "Dynamic optimal control of in situ bioremediation of ground water." *J. Water Resour. Plann. Manage.*, 124(3), 149–161.

Peralta, R. C., Kalwij, I., and Wu, S. (2003). "Practical simulation/optimization modeling for groundwater quality and quantity management." *Proc., MODFLOW and More: Understanding through Modeling*, International Groundwater Modeling Center, Golden, Colo., 784–788.

Reed, P., and Minsker, B. S. (2004). "Striking the balance: Long term groundwater monitoring design for conflicting objectives." *J. Water Resour. Plann. Manage.*, 130(2), 140–149.

Shieh, H. J., and Peralta, R. C. (2005). "Optimal in-situ bioremediation design by hybrid genetic algorithm-simulated annealing." *J. Water Resour. Plann. Manage.*, 131(1), 67–78.

Smalley, J. B., Minsker, B. S., and Goldberg, D. E. (2000). "Risk-based in situ bioremediation design using a noisy genetic algorithm." *Water Resour. Res.*, 36(10), 3043–3052.

van Laarhoven, P. J. M., and Aarts, E. H. L. (1987). *Simulated annealing: Theory and applications*, D. Reidel, Dordrecht, The Netherlands.

Wagner, B. J. (1999). "Evaluating data worth for groundwater management under uncertainty." *J. Water Resour. Plann. Manage.*, 125(5), 281–288.

Yoon, J.-H., and Shoemaker, C. A. (1999). "Comparison of optimization methods for groundwater bioremediation." *J. Water Resour. Plann. Manage.*, 125(1), 54–63.

Yoon, J.-H., and Shoemaker, C. A. (2001). "An improved real-coded GA for groundwater bioremediation." *J. Comput. Civ. Eng.*, 15(3), 224–231.

Zheng, C., and Wang, P. P. (1998). "MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems." *Technical Rep. Prepared for Waterways Experiment Station. US Army Corps of Engineers*, Dept. of Geology and Mathematics, Univ. of Alabama, Tuscaloosa, Ala.

Zheng, C., and Wang, P. P. (1999). "An integrated global and local optimization approach for remediation system design." *Water Resour. Res.*, 35(1), 137–148.