ELSEVIER

Advances in Water Resources 25 (2002) 51–65

Advances in Water Resources

www.elsevier.com/locate/advwatres

Multi-objective optimal design of groundwater remediation systems: application of the niched Pareto genetic algorithm (NPGA)

Mark Erickson [a], Alex Mayer [a,*], Jeffrey Horn [b]

[a] Department of Geological Engineering and Sciences, Michigan Technological University, 1400 Townsend Drive, Houghton, Michigan, USA
[b] Department of Mathematics and Computer Science, Northern Michigan University, Marquette, Michigan, USA

Received 19 October 2000; received in revised form 3 May 2001; accepted 19 May 2001

Abstract

A multiobjective optimization algorithm is applied to a groundwater quality management problem involving remediation by pump-and-treat (PAT). The multiobjective optimization framework uses the niched Pareto genetic algorithm (NPGA) and is applied to simultaneously minimize the (1) remedial design cost and (2) contaminant mass remaining at the end of the remediation horizon. Three test scenarios consider pumping rates for two-, five-, and 15 fixed-location wells as the decision variables. A single objective genetic algorithm (SGA) formulation and a random search (RS) are also applied to the three scenarios to compare performances with NPGA. With 15 decision variables, the NPGA is demonstrated to outperform both the SGA algorithm and the RS by generating a better tradeoff curve. For example, for a given cost of $100,000, the NPGA solution found a design with 75% less mass remaining than the corresponding RS solution. In the 15-well scenario, the NPGA generated the full span of the Pareto optimal designs, but with 30% less computational effort than that required by the SGA. The RS failed to find any Pareto optimal solutions. The optimal population size for the NPGA was found by sensitivity analysis to be approximately 100, when the total computational cost was limited to 2000 function evaluations. The NPGA was found to be robust with respect to the other algorithm parameters (tournament size and niche radius) when using an optimal population size. The inclusion of niching produced better results in terms of covering the span of the tradeoff curve. As long as some niching was included, the results were insensitive to the value of the parameter that controls niching ($\sigma_{\text{share}} > 0$).

1. Introduction

When faced with subsurface remediation management problems, decision makers must frequently weigh multiple objectives such as minimizing cost, minimizing health risk, minimizing cleanup time, and maximizing reliability. In these cases, it may be of value to the decision makers to view the tradeoffs between the conflicting objectives, providing a more effective means of selecting and implementing the best-suited remedial alternative for a given site.

The majority of applications of optimization tools to subsurface remediation problems have been based on single objective optimization methods. Single objective methods can accommodate multiobjective problems in several ways, such as minimizing a weighted, linear combination of the objective functions or minimizing a single objective while transforming the remaining objectives into constraints. However, these methods rely on a priori knowledge of the appropriate weights or constraint values. Furthermore, they are only capable of finding individual points on the tradeoff curve (or surface) for each problem solution.

True multiobjective methods have the potential to simultaneously generate all possible optimal combinations of objectives, with less effort than other approaches. Multiobjective problems involve several objective functions, each of which is a function of decision (d) and state variables (s). A generic multiobjective problem can be stated as:

$$O_1 = \max / \min[f_1(d_1, d_2, \ldots d_{n_1}; s_1, s_2, \ldots s_{n_2})]$$
$$O_2 = \max / \min[f_2(d_1, d_2, \ldots d_{n_1}; s_1, s_2, \ldots s_{n_2})]$$
$$\vdots$$
$$O_m = \max / \min[f_m(d_1, d_2, \ldots d_{n_1}; s_1, s_2, \ldots s_{n_2})]$$

* Corresponding author. Tel.: +906-487-3372; fax: +906-487-3371.
E-mail address: asmayer@mtu.edu (A. Mayer).

0309-1708/02/$ - see front matter
PII: S0309-1708(01)00020-3

subject to the following set of constraints:

$$d_1 \leqslant d_1^*, \quad d_2 \leqslant d_2^*, \ldots, d_{n_1} \leqslant d_{n_1}^*,$$
$$s_1 \leqslant s_1^*, \quad s_2 \leqslant s_2^*, \ldots, s_{n_2} \leqslant s_{n_2}^*,$$

where there are m objective functions, n_1 decision variables, and n_2 state variables.

As already mentioned, previous approaches for solving the multiobjective problem have involved reducing the problem dimension, either by combining all objectives into a single objective (e.g. [27]) or optimizing one while the rest are constrained (e.g., [3]). Once this reduction has been made, single-objective optimization methods (e.g. linear programming, [10]; non-linear programming, [11]; mixed-integer programming, [25], simulated annealing, [4]; genetic algorithms, [21]) can be applied to the optimal remediation design problem. Wagner [26] and more recently, Freeze and Gorelick [7], provide extensive reviews on the applications of optimization to groundwater remediation design.

Alternative multiobjective methods optimize all objectives simultaneously, eliminating the need for determining appropriate weights or formulating constraints. Multiobjective approaches in this category operate on the concept of "Pareto domination", which states that one candidate dominates another only if it is at least equal in all objectives and superior in at least one. The "degree of domination" for a design is proportional to the number of designs it is dominated by. For example, in Fig. 1, where the objective is to minimize both objective functions, designs 1 and 2 dominate design 3 because they are superior in both objectives. Moreover, designs 1 and 2 are said to be "non-dominated" because there are no existing designs that dominate them. This concept is utilized by evolutionary algorithms such as the multiobjective genetic algorithm (MOGA, e.g. [6]) and the niched Pareto GA used in this work.

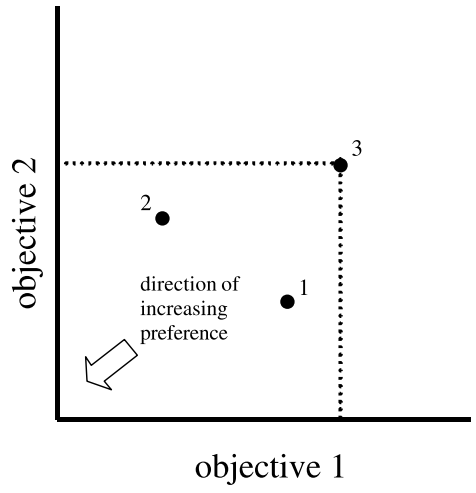

Fig. 1. Ranks of candidate designs based on the concept of Pareto domination.

Ritzel et al. [24] applied two variations of the genetic algorithm (GA), a Pareto GA and a vector-evaluated genetic algorithm (VEGA), to a multiobjective, groundwater pollution containment problem. The multiobjective problem was formulated to minimize the containment design cost while maximizing the design's reliability. The Pareto GA relied on a ranking scheme that ordered the population according to each containment design's degree of domination. The VEGA searches for multiple solutions to multiobjective problems simultaneously by selecting a fraction of the next population, based on the associated values of each objective function. Although the VEGA is considered a multiobjective optimization method, Richardson et al. [22] reported that VEGA tended to favor the extrema of the objective functions, such that only the endpoints of the tradeoff curve were found. Ritzel et al. [24] concluded that the Pareto GA was superior to the VEGA in finding the largest portion of the Pareto optimal solutions.

Cieniawski et al. [2] investigated the performance of four GA formulations in solving a multiobjective groundwater monitoring problem where they simultaneously maximized reliability of a monitoring system and minimized the contaminant plume size at time of first detection. They implemented a weighted GA, VEGA, Pareto GA and a VEGA/Pareto GA combination and compared them to results generated by simulated annealing. The VEGA/Pareto GA method was shown to be more computationally efficient and more successful at generating the greatest portion of the tradeoff curve than the other GA formulations. They recommended that a form of fitness sharing [9] be used to enhance the Pareto GA in this area, where crowding in the Pareto optimal solutions is alleviated by decreasing the fitness of crowded individuals.

Previous approaches for optimal groundwater remediation design have largely focused on single-objective optimization. Although some groundwater quality management efforts have considered multiobjective optimization, these approaches did not succeed in generating a sufficient representation of full range of Pareto optimal designs. In this work, we present an improved version of the niched Pareto genetic algorithm (NPGA), a multiobjective technique originally developed by Horn et al. [13], and apply it to a hypothetical contaminated groundwater remediation scenario.

There are several parameters (i.e., population, niche radius, tournament size, crossover rate, and mutation rate) that control the performance of this version of the NPGA. Although theoretical guidelines have been suggested for choosing optimal values for some of the algorithm parameters, these guidelines apply only to idealized, abstract problems [14–16]. The objectives of this work are to (a) develop a multiobjective approach for optimal groundwater remediation design using the NPGA, (b) explore the sensitivity of the NPGA to the

parameters that control the behavior of the algorithm in terms of computational performance, and (c) compare the NPGA to two other optimization approaches: a single-objective genetic algorithm (SGA) and a random search (RS).

2. The niched pareto genetic algorithm

Horn et al. [13] developed an evolutionary multiobjective optimization algorithm based on a suggestion by Goldberg [8] that introduced speciation along with the theory of a spatially ordered search space. This method, known as NPGA, extends the traditional GA to multiple objectives through the use of Pareto domination ranking and fitness sharing (or niching). Exploitation of the entire set of Pareto optimal designs is maximized by the selection pressure induced by the Pareto ranking and tournament competitions, and the diversity is maintained by fitness sharing. The addition of the sharing function is expected to overcome the challenge of finding and maintaining the entire tradeoff curve during the optimization process, as noted by Ritzel et al. [24]. This method is a promising approach for solving multiobjective optimization problems because of its adaptability to a wide variety of problems and its ability to search non-linear and discontinuous search spaces without relying on the need for continuous first and second derivatives.

McKinney and Lin [21], Ritzel et al. [24], and Huang and Mayer [16] give detailed descriptions of the traditional GA selection, reproduction, and mutation operators and a general overview of the GA as applied to single-objective, groundwater quality management problems. The heuristic parameters that are common between single and multiobjective GAs are population size, crossover probability, and mutation probability. Both sets of methods also can rely on tournament competition for deciding which candidates should go forward into the next generation. The extension of the traditional GA to the NPGA involves the addition of two specialized genetic operators: (1) Pareto domination ranking and (2) continuously updated fitness sharing. These operators alter the traditional mechanism of selection by partial ordering of the population and by maintaining diversity in the population through successive generations.

Tournament competition and fitness sharing create two principal genetic pressures that control the evolutionary process in the optimization algorithm. Selection pressure is controlled by the tournament size and propagates the designs towards the optimal frontier. Larger tournament sizes induce greater selection pressures. Fitness sharing promotes diversity by dispersing the designs over the limits of the tradeoff curves. The amount of searching performed by the NPGA is contingent on the size of the population and the level of selection pressure applied.

Fig. 2 summarizes the basic steps followed by the NPGA. To initiate the selection process in the NPGA, each individual in the population of designs is assigned a rank equal to the degree of Pareto domination experienced by that design. The degree of domination, or rank, of an individual design is the total number of designs in the population that dominate that design. A design is said to dominate another individual in the population if it is at least equal in all objectives to that individual and better in at least one. Non-dominated designs, or those that are not dominated by any individuals in the population, are assigned a rank of zero. In Fig. 3, where the objectives are to minimize both cost and mass remaining, an example of a Pareto domination ranked population of 10 designs is shown.

Once the entire population has been ranked according to the Pareto domination rank, candidate designs are chosen for reproduction. The mechanism of selection used here is similar to the tournament selection process described by Ritzel et al. [24] and Cieniawski et al. [2]. The controlling variable in tournament selection competitions is the tournament size. Tournament selection begins by randomly selecting a group of candidates from the population of ranked designs. The candidates in the tournament selection are then pitted against each other by comparing their respective ranks. If there is a single candidate with the lowest rank (i.e. less dominated), this candidate is the "clear winner" of the tournament and is selected for reproduction. If all the lowest ranked candidates in the tournament are non-dominated or otherwise equal in rank (i.e. no clear winner), none of the candidates are preferred and the tournament selection ends in a tie. Fig. 4 demonstrates each of these outcomes for a tournament size of two (the minimum tournament size). If the two candidates designated by triangles were selected in the tournament, the design designated by the open triangle would be the clear winner, because it is dominated by fewer designs than the solid triangle. If the candidates designated by the squares in Fig. 4 were chosen for the tournament, the tournament would end in a tie ("no clear winner").

In the case of a tie, an additional process is needed to select a tournament winner. Fitness sharing is a method for selecting the winning candidate that promotes the dispersal of candidate designs along the Pareto front. Our application of fitness sharing involves assessing the degree of crowding, or population density, experienced by each candidate. The population density around each candidate is calculated within a specified Cartesian distance (in objective function-space), called the niche radius (see Fig. 5). The niche count is calculated by summing the number of designs within the niche radius of each candidate, weighted by the radial distance between the candidate and the other designs, or

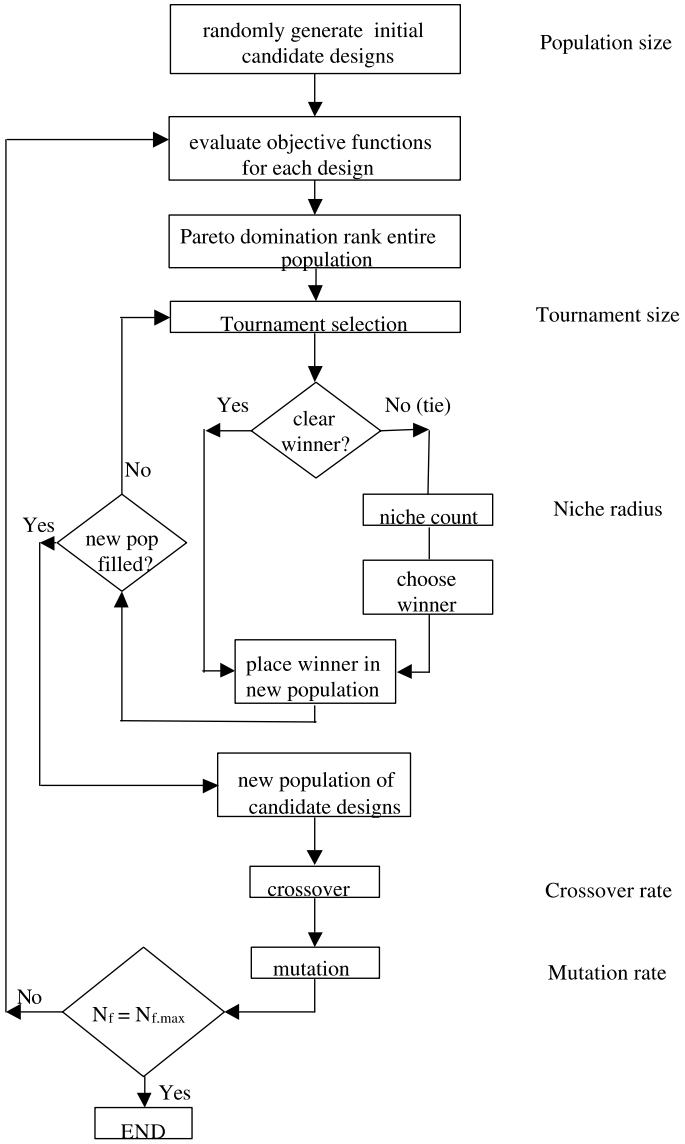

Fig. 2. Process flowchart for the NPGA. The parameters that control specific algorithm steps are listed in italics to the right of the flow chart.

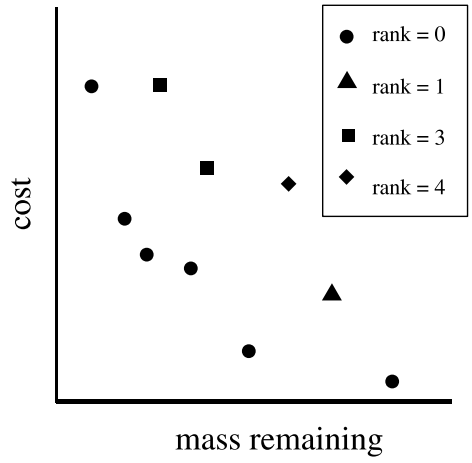

Fig. 3. Pareto domination ranking for a population of ten designs. Designs of equal rank are designated by the same symbol.

$$m_i = \sum_{j \in \text{pop}} \left(1 - \frac{d'_{ij}}{\sigma_{\text{share}}}\right), \tag{1}$$

where $d'_{i,j}$ is the scaled, radial distance between candidate i and candidate j and σ_{share} is the niche radius. Both $d_{i,j}$ and σ_{share} are measured in scaled, objective function space, such that $0 < d'_{i,j} < \sqrt{2}$. The values of the objective functions are scaled as in

$$O'_i = \frac{O_i - O_{i,\min}}{O_{i,\max} - O_{i,\min}}, \tag{2}$$

where O'_i, $O_{i,\min}$, and $O_{i,\max}$ are the scaled, minimum, and maximum values of objective O_i, respectively.

Thus, in the no clear winner case, preference is given to the candidate with the lowest niche count. In Fig. 5, the design shown as a solid square has two candidate designs within the niche radius, whereas the design

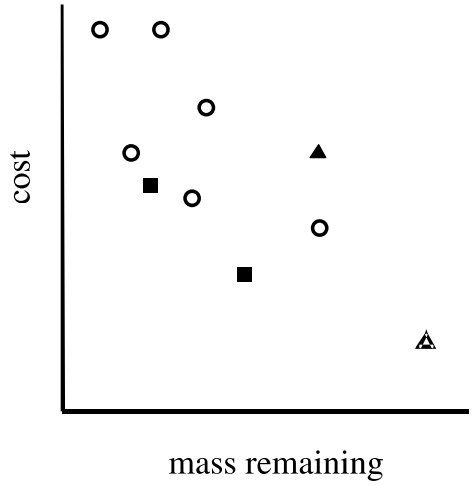

Fig. 4. Two possible tournament selection outcomes, where the candidates have equal ranks (designated by a ■) or unequal ranks (designated by either an △ or ▲). The open triangle is the winner in the case of unequal ranks.

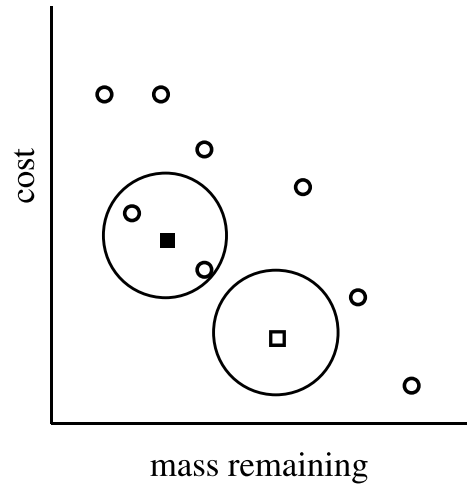

Fig. 5. Performing fitness sharing on two tournament selection candidates ranks (designated by either an □ or ■) with equal ranks. The winner here is the less crowded design represented by the open square.

shown as an open square has no candidate designs within this distance. The open square would be given preference and would be the winner of the tournament selection process.

3. Numerical experiments

The objectives of the numerical experiments are to (1) investigate the performance of the NPGA as a function of the size of the search space, (2) compare the NPGA performance with two other methods for generating the tradeoff curve, and (3) investigate the NPGA performance as a function of algorithm parameters.

3.1. Multiobjective management model

The application of the multiobjective problem focuses on the active remediation of a hypothetical, contaminated groundwater site by pump-and-treat (PAT) technology. The decision variables are well extraction rates as a function of location. It was assumed that the ex situ treatment technology for the contaminated groundwater is granular activated carbon (GAC). The two objectives are to minimize cost and maximize cleanup performance. The cost of a PAT remediation system includes capital costs incurred from installation of recovery wells and operational costs from pumping and groundwater treatment as in

$$\min J = a_1 N_{\mathrm{ew}} + \sum_{k=1}^{N_{\mathrm{ew}}} \left[(a_2 Q_k H_k T) + \sum_{l=1}^{NTSP} \left(a_3 t_l Q_k \frac{C_{k,l}}{K_{AB} C_{k,l}^{1/n}} \right) \right], \tag{3}$$

where J is the total cost of the remedial design, N_{ew} is the number of active wells, $NSTP$ is the number of time steps within the remediation horizon $T = \Sigma t_l$, Q_k is the extraction rate of well k, H_k is the total drawdown at well k, $C_{k,l}$ is the average flow-weighted concentration removed by well k in time step l, t_l is the length of time step l, K_{AB} and $1/n$ are the Freundlich GAC adsorption parameters for a given contaminant and carbon adsorbent, and a_1, a_2, and a_3 are the coefficients for capital, pumping, and treatment costs, respectively. The treatment cost term is linearly related to the GAC utilization rate, $\dot{m}_{\mathrm{GAC}}$, which is based on a steady-state mass balance on the GAC reactor

$$q\dot{m}_{\mathrm{GAC}} = QC_{k,l} - QC^*,$$

where q is the concentration of the contaminant on the adsorbent and is given by a Freundlich isotherm equation, $q = K_{AB}(C^*)^{1/n}$ [23], and C^* is the target effluent concentration from the GAC reactor. The second term on the right-hand side is ignored because $C_{k,l} \gg C^*$ for most of the simulation period.

The second objective function is based on a measure of cleanup performance. The usual approach is to require that point concentrations at monitoring or pumped wells do not exceed a target concentration. The approach used here is different in that cleanup performance is treated as a variable to be maximized, rather than as a fixed constraint. Furthermore, the total contaminant mass remaining in the aquifer (MR) is used as a measure of cleanup performance, rather than point concentrations. In the simulations, the use of MR avoids the arbitrariness of selecting well locations for monitoring performance. In the field, point concentrations are inherently uncertain, due to variability in aquifer properties and measurement errors. Since the MR is

determined by integrating over point concentration values, it tends to average out the uncertainty.

The mass remaining objective function is formulated as:

$$\min MR' = 100\frac{MR}{MI}, \tag{4}$$

where MR' is the percent mass remaining after the remedial horizon is complete; MI and MR are the initial mass present in the aquifer at the beginning of the remediation horizon and the mass remaining at the end of the remediation horizon, respectively. In the simulations of the aquifer-contaminant system, MI is a known. However, in the field, the value of MI usually is not known, unless the initial contaminant release has been documented or there is a conservative constituent present and the composition of the contaminant source is known.

In groundwater management and remediation problems, drawdown constraints are usually enforced to protect against aquifer dewatering. Constraints can be applied explicitly in optimization problems by formulating penalty functions that decrease the values of the objective functions proportionally to the magnitude of constraint violations. Although the algorithm performance is sensitive to the form of the penalty function, there are only qualitative guidelines for choosing the optimal formulation. In an effort to avoid the use of penalty functions, the drawdown constraint is enforced by limiting the maximum extraction rate per well $Q_k^{\max}$, as in the following:

$$0 \leqslant Q_k \leqslant Q_k^{\max} \quad \text{for } Q_k^{\max} = Q_{\mathrm{T}}^{\max}/N_{\mathrm{ew}}, \quad yk = 1, \ldots, N_{\mathrm{ew}}, \tag{5}$$

where $Q_{\mathrm{T}}^{\max}$ is the maximum total extraction rate, $Q_{\mathrm{T}}^{\max}$ is determined by running a series of single-well flow simulations, where the extraction rate is varied until the maximum allowable drawdown (15% of the aquifer thickness) is reached. The constraint is applied through the encoding/decoding of the decision variables, such that the value of the individual well extraction rates cannot exceed $Q_k^{\max}$. Although this approach for enforcing the drawdown constraint is straightforward for the cases presented here, it may not be appropriate for heterogeneous cases or cases with a large number of decision variables.

The state variables, hydraulic head and contaminant concentration are determined with a groundwater flow and contaminant transport simulator. The steady-state confined groundwater flow equation for a non-deforming, saturated, aquifer system is

$$\nabla(K \cdot \nabla h) = \sum_k Q'_k \delta(x - x_k, y - y_k), \tag{6}$$

where K is the hydraulic conductivity, Q'_k is the extraction rate per unit aquifer volume from well k located at x_k and y_k, and δ is the delta Dirac function. The hydraulic head, h, is related to the total drawdown, H, by $H = z_{gs} - h + h_l$, where z_{gs} is the ground surface elevation and h_l is the estimated head loss due to piping in treatment train.

Contaminant concentrations are determined by solving the mass balance equation for a neutrally buoyant, conservative aqueous chemical constituent, given by

$$\frac{\partial C}{\partial t} + \nabla(\mathbf{v}C) - \nabla(\boldsymbol{D} \cdot \nabla C) = -\sum_j \frac{C_k}{n} Q'_k \delta(x - x_k, y - y_k), \tag{7}$$

where C is the aqueous concentration in the aquifer, C_k is the aqueous concentration removed by well k, and n is the effective porosity. The hydrodynamic dispersion tensor, $\mathbf{D}$, is defined as

$$\mathbf{D} = (\alpha_{\mathrm{T}}|v| + D^*)\mathbf{I} + (\alpha_{\mathrm{L}} - \alpha_{\mathrm{T}})\frac{vv}{|v|}, \tag{8}$$

where α_{L} and α_{T} are the effective longitudinal and transverses dispersivity coefficients, $\mathbf{I}$ is the unit tensor; and D^* is the molecular diffusivity. The pore velocity, $\mathbf{v}$, is given by Darcy's relationship as

$$n\mathbf{v} = -K\nabla h. \tag{9}$$

We employ a two-dimensional finite difference approximation to solve the groundwater flow Eq. (6) and a particle-tracking method to solve the mass transport Eq. (7). The numerical codes have been validated by Maxwell [18,19]. Additional background information pertaining to the development of this numerical simulator can be found in [17].

3.2. Description of groundwater system

To explore the performance and efficiency of the NPGA, the simulation/optimization algorithm was applied to a simple, hypothetical contaminated groundwater site. The site is modeled as a confined homogeneous aquifer 1000 m long by 1010 m wide and 30 m thick (see Fig. 6). The two-dimensional finite-difference grid system consists of 10-m square grid blocks. Constant-head boundaries are imposed on the east and west sides of the model domain and no-flow boundaries are imposed on the north and south sides. The aquifer is modeled as having a homogeneous, isotropic hydraulic conductivity. Groundwater extraction wells are modeled as being open over the entire thickness of the confined aquifer. Removal of contaminant by the treatment system is simulated as equilibrium, non-linear adsorption onto the GAC.

Trichloroethlyene (TCE), a commonly observed and studied groundwater contaminant, is used as the hypo-

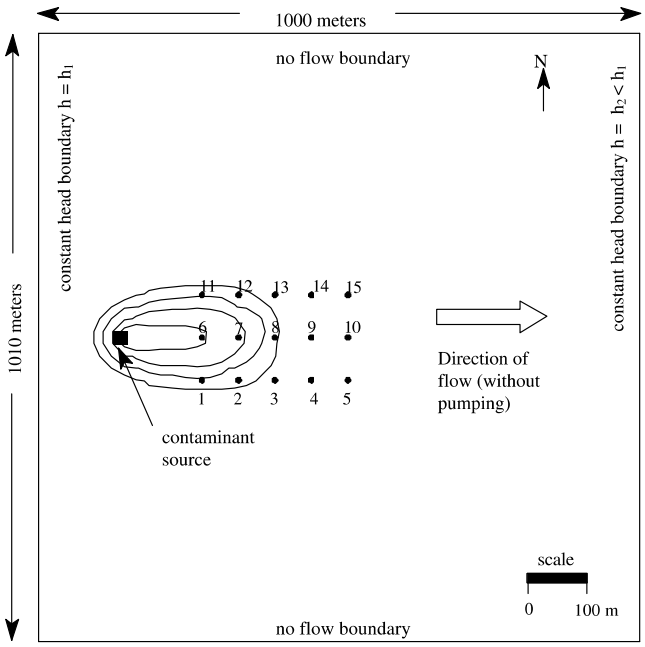

Fig. 6. Plan view of the hypothetical aquifer system showing the initial contaminant plume and fixed-location extraction wells used in the numerical experiments. The contaminant plume is contoured at 100, 10, 1, and 0.1 ppm levels.

thetical contaminant and is treated as a conservative, dissolved species. A constant source of approximately 750 ppm is used to generate the initial concentration plume shown in Fig. 6. The contaminant plume evolves until approximately 1000 kg of TCE is released into the confined aquifer system as a dissolved species. The source is then removed and active remediation begins. The length of the remediation horizon is ten years. Table 1 summarizes the aquifer, contaminant, and treatment system properties. The aquifer properties (porosity, hydraulic conductivity, and background pore velocity) are similar to those found for a non-uniform, fine to medium sand. The longitudinal dispersivity was chosen to be equivalent to the gridblock size so that numerical errors would be minimized and the transverse dispersivity is set as a typical fraction of the longitudinal dispersivity ($\alpha_T = \alpha_L/5$).

3.3. Structure of numerical experiments

We chose to examine the performance of the NPGA with two, five, and 15 decision variables, where each decision variable corresponds to the pumping rate at a fixed-location well. The NGPA parameters we investigate are population size, tournament selection size, and niche radius. Table 2 lists the cost coefficients used in the numerical experiments. Table 3 describes the problem characteristics for each scenario.

SGA and RS methods. We compare the performance of the NPGA with two methods for generating the tradeoff curve: an SGA and RS. With the SGA method, minimizing cost is the objective and the percent mass remaining is held as a constraint. The tradeoff curve is constructed by executing separate SGA runs, each with a different value of the mass remaining constraint. The percent mass remaining constraint was imposed via the multiplicative penalty approach described by Chan Hilton and Culver [1]. The constrained optimization problem can be formulated as

$$\min\left[J\left(1+wMR'\right)\right], \tag{10}$$

where the cost function, J, is identical to the cost function in Eq. (3), w is the constraint violation weight, and MR' is defined as Eq. (4). We assumed that three separate SGA runs with the constraint values equally distributed (on a log scale) over the possible range of mass remaining would be sufficient to delineate the tradeoff curve.

The RS is a simplistic approach to finding optimal solutions, where values of the decision variables are randomly generated from a uniform distribution. The tradeoff curve is formed by the Pareto optimal subset of the randomly generated set of decision variables. We compare the NGPA, SGA, and RS in terms of (a) the Pareto optimality of the tradeoff curves and (b) the span of the tradeoff curves. Table 4 lists the values of the algorithm parameters used in the comparisons. We use the same random seed in each numerical experiment.

Coding of decision variables, size of search space, and scaling of objective functions. For all three methods, the decision variables are coded in binary form. The binary

Table 1
Parameters for flow, transport and treatment simulations

Parameter	Value	Source
Porosity	0.25 (dimensionless)	Assumed
Hydraulic conductivity	3.82×10^{-5} m/s	Assumed
Background pore velocity	2.7×10^{-2} m/d (west to east)	Calculated from constant head boundary conditions and above parameters
Longitudinal dispersivity	10 m	Assumed
Transverse dispersivity	2 m	Assumed
GAC adsorption coefficient, K_{AB}	28.4 (mg/gm) (l/mg)$^{1/n}$	Hand [12]
GAC adsorption coefficient, $1/n$	0.48	Hand [12]
GAC effluent concentration	5 ppm	Fetter [5]

Table 2
Cost coefficients used in numerical experiments

Parameter	Value	Source
Capital cost coefficient, a_1		
8 in. (two-well scenario)	10,800 \$/well	Means [20]
6 in. (five-well scenario)	8500 \$/well	Means [20]
4 in. (15-well scenario)	5800 \$/well	Means [20]
Power cost coefficient, a_2	1.05	Means [20]
Treatment cost coefficient, a_3	2.14	Means [20]

representation is transformed to a real form by scaling with maximum ($Q_k^{\max}$) and minimum ($Q_k^{\min} = 0$) allowable pumping rates. The size of the search space, or the number of possible solutions, N_p, is dependent on the number of bits, N_b, used to represent the decision variables in the search methods, which corresponds to the precision of the decision variables, dQ_k. The number of bits is calculated as follows:

$$2^{N_b} - 1 = \frac{Q_k^{\max} - Q_k^{\min}}{dQ_k} \tag{11}$$

and the number of possible solutions is given by

$$N_p = \left(2^{N_b}\right)^{N_{ew}}. \tag{12}$$

The objective functions are scaled as described in Eq. (2), so that the scaled distances in Eq. (1) can be calculated. The maximum and minimum objective function values do not change from generation to generation; they are fixed at the beginning of a run. The maximum cost value is estimated using results from initial sets of optimization runs. The NPGA results were insensitive to changes in the maximum objective function values. A 100% increase and 50% decrease in the maximum cost were found to have no impact on final results. The minimum value for cost assuming no action is zero; similarly, the minimum mass removed is 0%. The maximum mass removal is 100%.

Performance measures. The algorithms (NPGA, SGA, or RS) proceed until a fixed amount of computational effort is consumed. One unit of computational effort is equivalent to one set of objective function evaluations, which involves one execution each of the flow and transport simulators and the corresponding calculation of the cost and mass remaining objective function values. The computational time required to run the flow and transport simulators is typically more than 90% of the total time needed to evaluate the objective functions. Thus, one unit of computational effort is approximately equal to the effort required for one run of the flow and transport simulators. One execution of the flow and transport simulators consumes approximately 2.5 min of CPU time on a Sun Ultra 80 workstation with a 450 MHz processor.

Since the computational effort is fixed for the simulation/optimization experiments, we use as performance measures: (a) the percentage of Pareto optimal solutions found by each method and (b) a qualitative evaluation of the span of the tradeoff curve covered by these solutions. We determine the percentage of Pareto optimal solutions found by a method m by aggregating all of the Pareto optimal solutions found by the methods and calculating P_m/P_M , where P_M is the total number of Pareto optimal solutions found by all methods and P_m is the number of Pareto optimal solutions found by method m. The endpoints of the tradeoff curve are defined as the maximum MR' achievable ($\sim$100%) and the minimum MR' achievable, given the maximum pumping rate constraint.

Table 3
Problem characteristics for two-, five-, and 15-well scenario scenarios

Scenario	Candidate well locations[a]	Maximum number of objective function evaluations allowed	Decimal precision in flow rate (m^3/d)	Maximum flow rate per well (m^3/d)	Number of bits per decision variable	Number of possible designs
2 wells	7, 8	500	2.5	250	7	$\sim 10^4$
5 wells	2, 6, 7, 8, 12	1000	2.5	100	6	$\sim 10^9$
15 wells	All wells	2000	2.5	33	4	$\sim 10^{18}$

[a] Note that these locations refer only to the possible locations for the given scenario, and that not all of these wells will be active for a given design. Refer to Fig. 6 for locations.

Table 4
Optimization algorithm parameters for the multiobjective (NPGA) and single-objective (SGA) solution methods applied to two-, five-, and 15-well scenarios

Parameter	NPGA 2 wells	NPGA 5 wells	NPGA 15 wells	SGA all scenarios
Population size	50	50	100	50
Tournament selection size	2	2	10	2
Niche radius	0.5	0.05	0.5	–
Probability of crossover	0.9	0.9	0.9	0.9
Probability of mutation	0.001	0.001	0.001	0.001
MR' constraint violation weight	–	–	–	150

Although the SGA was nominally subjected to a stopping criterion based on a fixed amount of computational effort, each SGA run converged before the stopping criterion was reached. We report the number of flow and transport simulations expended to reach convergence, which is defined as the point where the entire population converges to a single solution. The best solution (lowest cost) found during the run was used as the optimal value.

Archiving. In the NPGA, SGA and RS methods, we implement design archiving. Archiving eliminates redundant objective function evaluations by saving the objective function values for every design, beginning with the initial population. If, in a subsequent generation, a design is found to have been evaluated previously, the previous evaluation is used, and only the remaining, new designs are evaluated. The archives are searched by comparing the active set of decision variables to the sets stored in the archives, using a hashing algorithm.

Archiving typically resulted in reducing the number of objective function evaluations by 50%, which is approximately the same as the reduction observed for CPU time. As the number of digits used to code the decision variables or the number of decision variables increases, the archive search requires more CPU time and storage space. However, only the binary decision variables and the associated, floating point values of the objective function are stored, so that only of the order of 100 bytes are required for each design. For the simulations reported in this work, the archiving effort did not exceed 0.01% of the total CPU time for a simulation and the maximum archive file size was less than 10 kB.

4. Results and discussion

We present the results in terms of the distribution of the Pareto offline data. The Pareto offline data consist of the Pareto optimal designs found in each generation, beginning with the initial population. In subsequent generations, Pareto optimal designs are copied into the Pareto offline data set, and any designs from previous generations that are now dominated are removed.

Fig. 7 shows the results from a typical NPGA run, where the cost is plotted against the mass remaining on a log scale. In this computational experiment, the NPGA ran for 200 generations with a population size of 50, tournament selection size of 2, and niche radius of 0.005. The PAT remedial design involved estimating extraction rates for fifteen fixed-location wells (see Fig. 6 for locations). In Fig. 7, the Pareto offline designs are shown for every 50th generation. In this case, the minimum MR' possible was 2%, due to the maximum extraction rate constraint. The improvements seen in the final set of optimal solutions are significant when compared to the initial population. As the algorithm proceeded, the span of the tradeoff curve increased until Pareto optimal designs were found over the entire possible range of tradeoffs. While designs in the higher cost region of the tradeoff curve (greater than about $140,000) appear to have converged quickly (within 50 generations), the remaining portion of the tradeoff curve was not generated until at least 150–200 generations elapsed.

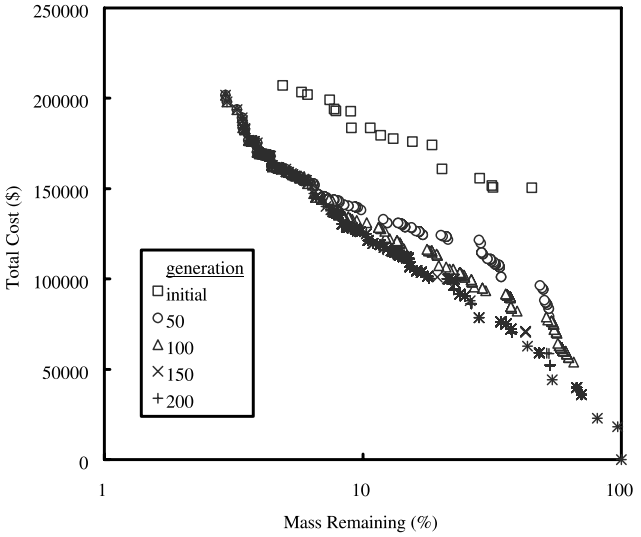

Fig. 7. NPGA results for the 15-well scenario: Pareto optimal designs at intervals of 50 generations with a population size of 50 designs.

The designs also improved considerably in terms of Pareto optimality as the algorithm proceeded. For example, the cost of designs with about 5% mass remaining improved by nearly 20% from the first to the 50th generation. The cost of the designs with 50% mass remaining improved by about 50% from the first to the 100th generation. Although the position of the starting set of solutions varies from random population to random population, the convergence behavior shown in Fig. 7 is typical of all of the runs conducted with the NPGA.

In Fig. 8, the extraction rate distribution and final contaminant plume contours are given for 60%, 15%, and 4% MR' designs, respectively. These figures demonstrate that the NPGA is finding reasonable results, since as MR' decreases, more extraction wells are required. In addition, the distribution of extraction rates is symmetric about the mean direction of flow, for the most part, as is the final contaminant plume. The extraction rate distribution shown in Fig. 8(c) is not an entirely intuitive result, however, since the two wells not on the plume centerline (wells #3 and #15, see Fig. 6 for well numbering system) are not opposite each other. The optimization algorithm cannot be expected to recognize the physical nature of the optimal solution, i.e. symmetry, but it should yield optimal results. To assess the optimality of the asymmetric solution, a new simulation

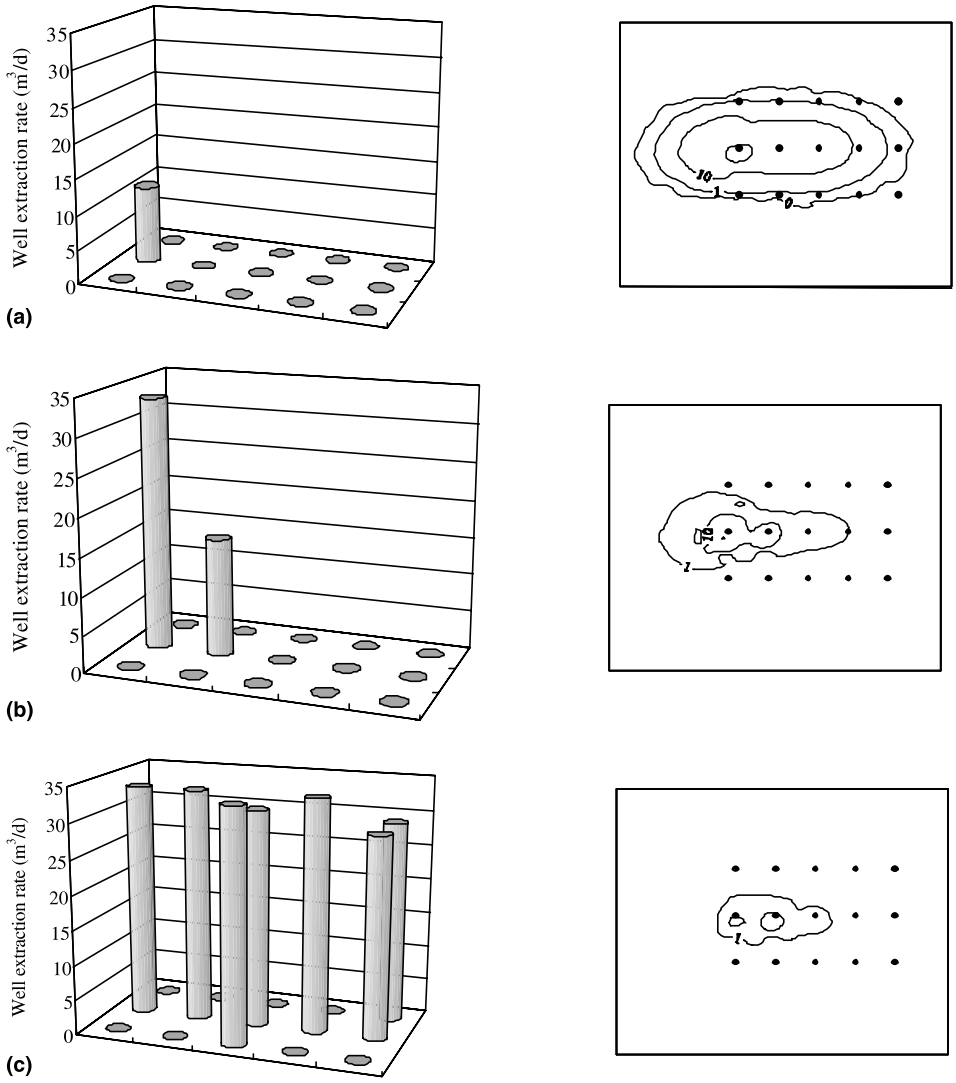

Fig. 8. NPGA results for the 15-well scenario: distribution of extraction rates (left-hand side) and contaminant plume at the end of remediation (right-hand side) for: (a) a 60% *MR* design, (b) a 15% *MR* design, and (c) a 4% *MR* design. Plume contours are spaced by a factor of 10 ppm. Well locations are indicated by dots.

was conducted where the extraction rate at well #15 was exchanged with that of #13 to produce an almost symmetric design. A simulation using the new, almost symmetric design resulted in costs and mass removals that were only 2.5% and 0.5% lower, respectively, than the cost associated with the asymmetric design. These small differences imply that, since the solution is insensitive to the pumping rate distribution in the off-centerline wells, the optimization algorithm will not necessarily find a symmetric solution.

The relationships of the total extraction rates, treatment costs, and total costs to percent mass remaining are illustrated in Fig. 9. For higher values of MR', treatment costs dominate the overall cost of a design. However, as MR' decreases below 10%, installation costs are increasingly significant, since the number of wells required to meet the given value of MR' increases sharply below this level. The total extraction rate also increases sharply for $MR' < 10\%$, since below this level, the volume of water that must be extracted to meet the given level of MR' increases exponentially.

As mentioned in the Section 3, it is likely that the initial mass (MI) will be uncertain in a field situation. If the transport equation is linear in concentration, then the concentrations are linearly related to the source concentration and thus the initial mass of contaminant. In this case, the effects of uncertainty in MI can be analyzed by shifting the tradeoff curves along the MI axis. If the processes active in the contaminant-aquifer system result in non-linear transport, a more involved analysis is required and is the subject of the ongoing work.

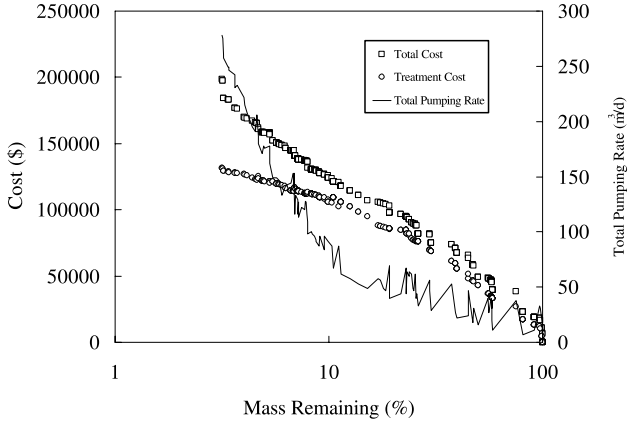

Fig. 9. NPGA results for the 15-well scenario: total pumping rates, treatment costs, and total costs for final Pareto optimal designs.

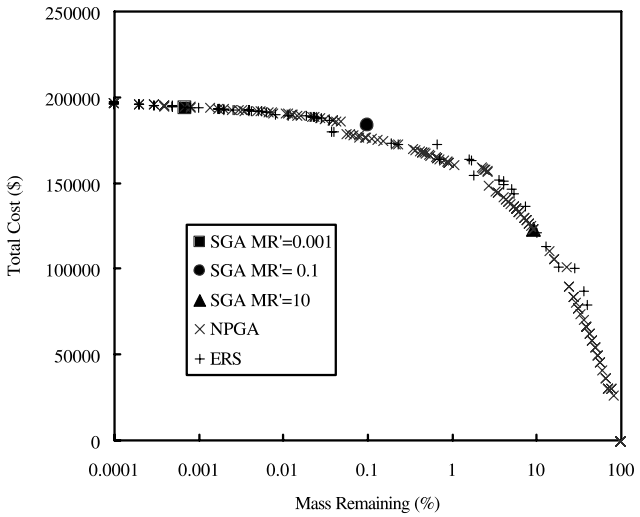

Fig. 10. Final Pareto offline results for two-well scenario found by the NPGA, SGA and RS.

4.1. Sensitivity to number of decision variables and comparison of NGPA performance with SGA and RS

We applied the NPGA, SGA, and RS to hypothetical remediation scenarios using two, five, and 15 as the number of wells for which optimal extraction rates must be found. Table 5 shows a component cost breakdown for $MR' = 5\%$ designs for these scenarios. The cost breakdown shows that, in general, the treatment costs account for the majority of the total costs. The results in Table 5 also show that, for designs that can achieve $MR' = 5\%$, not all of the possible wells are used in each scenario.

For the two-well tradeoff curve (Fig. 10), the cost increases from $MR' = 100\%$ until a level of approximately $MR' = 0.1\%$ is achieved, below which the cost begins to approach an asymptotic value. The asymptotic behavior is primarily due to the fact that, for $MR' < 0.1\%$, the concentration of water extracted from the aquifer is below the desired effluent concentration for the treatment system. Thus, as MR' decreases further below 0.1%, the only additional increase in cost is due to the energy requirements for pumping.

The results in Fig. 10 show that each method is equally effective at generating the trend of a tradeoff curve, for the case where the search space is relatively small. If the NPGA, SGA, and RS results are aggregated, we find that the NPGA generated 70% of the aggregated Pareto optimal designs, spanning the entire range of MR'. The RS found the remainder of the aggregate Pareto optimal designs, spanning $MR' = 0.0001\%$ to $MR' = 40\%$. The NPGA and RS runs each consumed the limit of 500 objective function evaluations. The three SGA runs, using mass remaining constraints of 0.001%, 0.1% and 10%, consumed a total of 641 evaluations. All of the SGA runs converged to a single solution before the limit of 500 objective function evaluations was reached.

The final Pareto offline results for the five-well scenario are shown in Fig. 11. In this scenario, the costs increase sharply as a function as MR' decreases, but approach an asymptotic value as MR' drops below about 1%. This behavior is similar to that of the two-well case, which is explained by the fact that the treatment costs approach zero for $MR' < 0.1\%$. The NPGA generated 95% of the aggregate Pareto optimal designs, compared to 4% for the RS. The minimum cost designs found by SGA for the three constraint values of $MR' = 0.05\%$, 5%, and 50% are nearly Pareto optimal solutions. The SGA designs were found at the expense of 1661 evaluations. The NPGA and RS runs each consumed the limit of 1000 objective function evaluations. Although the NGPA clearly found more Pareto optimal solutions than the other methods, the tradeoff curves generated by each method are of a similar shape and position.

The Pareto offline results for the fifteen-well scenario are shown in Fig. 12. These results exhibit approximately a linear (cost)-log (MR') relationship over

Table 5
Comparison of Pareto optimal 5% MR' designs found by the NPGA for the two-, five-, and 15-well scenarios

Test scenario	Number of active wells	Total flow rate (m^3/d)	Total capital cost ($)	Total pumping cost ($)	Total treatment cost ($)	Total cost ($)
2 wells	1	81	10,800	1100	126,400	138,300
5 wells	2	76	17,000	1000	123,500	141,400
15 wells	6	173	34,000	2400	122,500	159,100

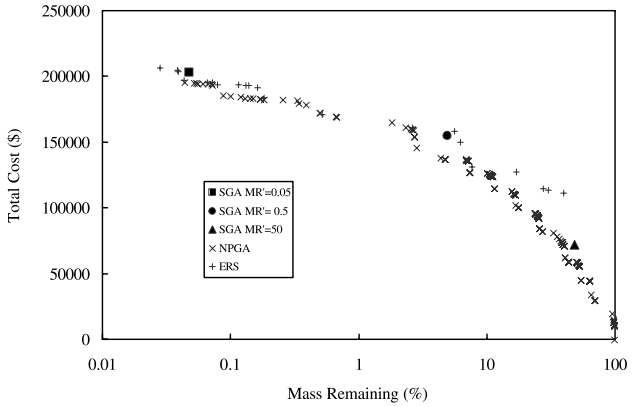

Fig. 11. Final Pareto offline results for five-well scenario found by the NPGA, SGA and RS.

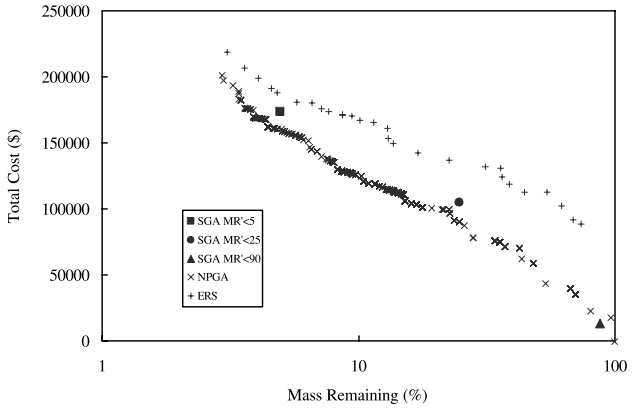

Fig. 12. Final Pareto offline results for 15-well scenario found by the NPGA, SGA and RS.

the entire range of MR'. This relationship is due to the fact the number of wells required to achieve a given value of MR' increases linearly with a log change in MR'. All three methods were capable of finding designs that spanned the full range of mass remaining values. However, the results show that the NPGA found solutions that would achieve 25–250% better performance than the RS, measured in terms of the MR' achieved for given values of cost. For example, a \$100,000 design generated by the NPGA achieved 20% mass remaining, whereas an RS design with a cost of \$100,000 achieved only 65% mass remaining. In addition, the NPGA produced 352 Pareto optimal solutions that span the entire range of tradeoffs, while the RS failed to find any designs on the Pareto front. Although the three SGA runs found designs that are near optimal in cost with respect to the NPGA, the total computational expense was 2713 objective function evaluations, as compared to 2000 for the NPGA.

The inability of the five- and 15-well cases to identify designs with mass removals as low as the two-well case is caused by the drawdown constraint. Recalling Eq. (1), the drawdown constraint is enforced by setting a maximum extraction rate per well, which is equal to the total maximum extraction rate divided by the number of wells. Since the total maximum extraction rate is constant, the total extraction is applied over a successively broader area as the number of wells increases. The result is that the efficiency of the remediation effort declines as the number of wells increases.

4.2. NPGA parameter sensitivity

Population size and tournament size. In general, we observed that there is an optimal population size that gives the best tradeoff between finding designs that are Pareto optimal and designs that span the tradeoff curve. This observation is illustrated in Fig. 13 where populations of 100 and 150 designs covered a greater span of the tradeoff curve than the population of 50 designs. The results from the run with a population size of 50 indicate that designs for larger values of MR' ($MR' > 40\%$) were not found. Although the NPGA run with a population of 150 found Pareto optimal designs that spanned the possible range of MR' values, it performed worst in terms of the number of Pareto optimal designs found, as shown in Table 6. This result is due to the fixed limit on the number of objective functions. Under these circumstances, runs with smaller population sizes can propagate through more generations, thus advancing the Pareto front farther than runs with larger population sizes. Thus, a population size of 100 seems a good balance of search breadth versus search duration. We have also found that when the population size was at an optimal value, the NPGA was less sensitive to selection pressure, crossover rates, and decision variable precision.

Selection pressure also can be increased by enlarging the tournament size. Fig. 14 shows a comparison of results obtained with tournament sizes of 2, 4 and 10 for

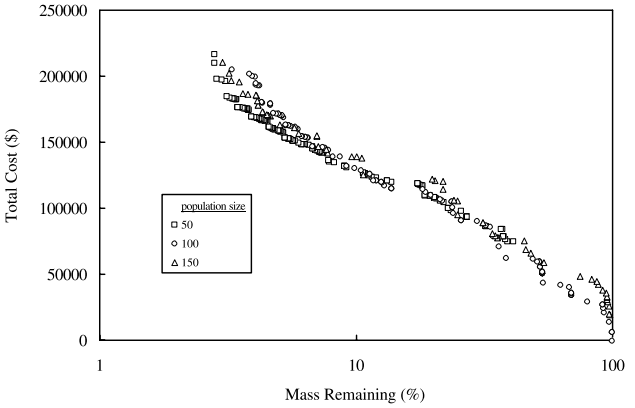

Fig. 13. Final Pareto offline results for 15-well scenario using NPGA: population sizes of 50, 100, and 150 designs and a tournament size of 2 designs.

Table 6
Performance results for NPGA Pareto optimal solutions for three population sizes

Population size	Number of generations	Number of objective function evaluations allowed	% of population that is Pareto optimal
50	89	2000	95
100	26	2000	50
150	15	2000	30

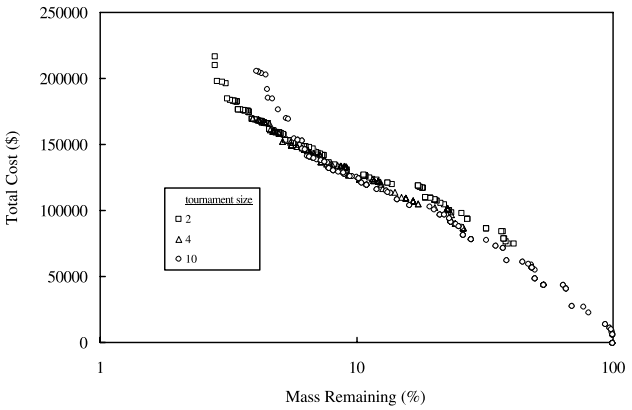

Fig. 14. Final Pareto offline results for 15-well scenario using NPGA: tournament sizes of 2, 4 and 10 and a population of 50.

a population size of 50. The increase in tournament size to 10 improved the span of the tradeoff curve significantly, although there was a corresponding decrease in Pareto optimality for the highest cost designs.

Niching. In cases with low tournament size, we observed that niching produces a greater span of the tradeoff curve. As shown in Fig. 15(a), where the tournament size is 2 and the population is 50, the entire tradeoff curve is spanned for $\sigma_{\text{share}} > 0$. Even without explicit niching, the algorithm is able to maintain some diversity on the Pareto optimal front. This performance is most likely due to the Pareto domination ranking tournaments, since individuals in more crowded areas will tend to have more individuals dominating them.

When the tournament size is increased to 10, the effect of niching on the span of the tradeoff curve is almost

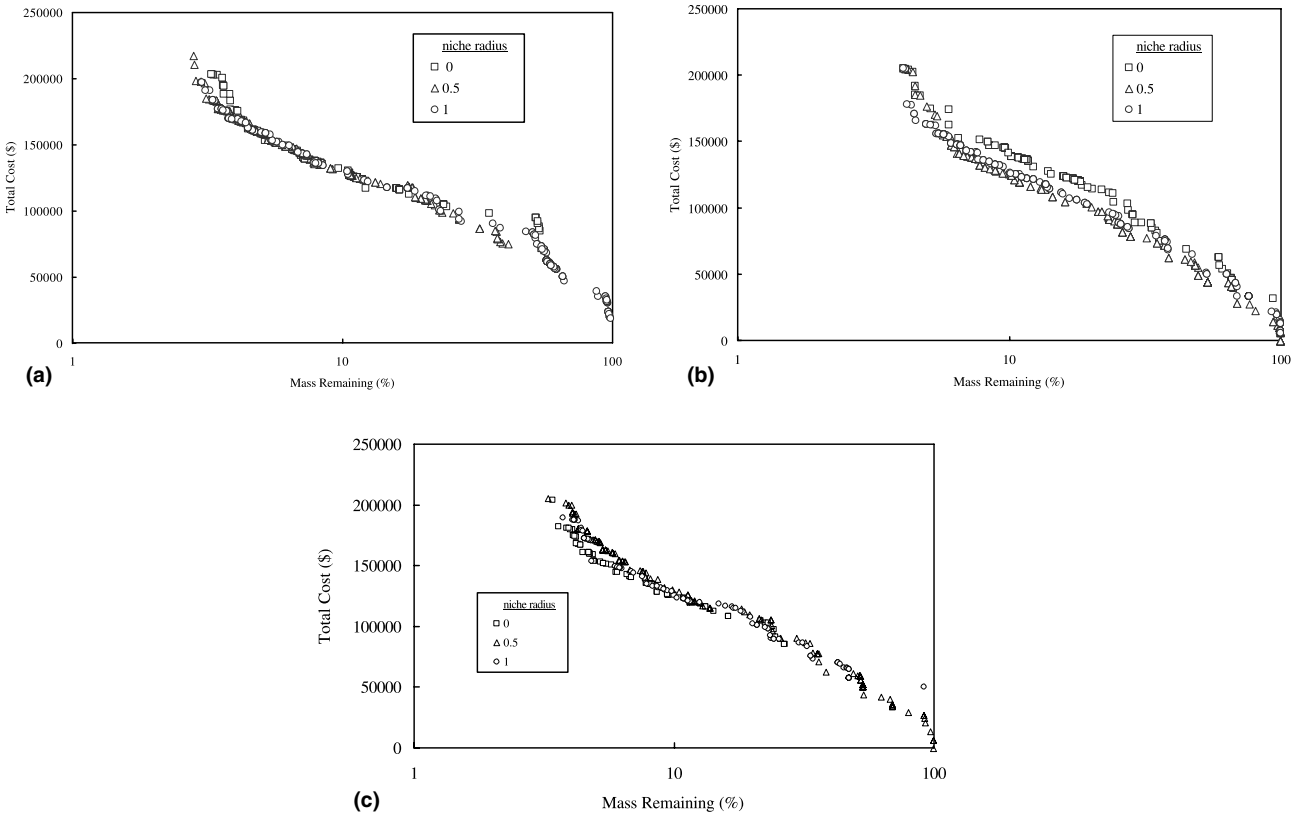

Fig. 15. Final Pareto offline results for 15-well scenario using NPGA: niche radii of 0, 0.5, and 1.0 with: (a) a population of 50 designs and a tournament size of 2, (b) a population of 50 designs and a tournament size of 10, and (c) a population of 100 designs and a tournament size of 2.

negligible, as shown in Fig. 15(b). This result is due to the high selection pressure induced by large tournament sizes, creating populations that are more consistently diverse over the evolutionary process. The results in Fig. 15(b) also indicate that the increase in tournament size produces better results in terms of Pareto optimality for the cases where niching is applied. When the population size is increased from 50 to 100 (compare Fig. 15(a) and (c)), the cases with niching perform better in terms of Pareto optimality. It is likely that with the smaller population (50), there is limited diversity within a single generation, such that the addition of niching improves the search by maintaining greater diversity. Conversely, for the larger population size (100), there is sufficient diversity, and so niching does not help. From our limited results, there appears to be no optimal setting of σ_{share}, other than $\sigma_{\text{share}} > 0$. Indeed, the choice of niche radius seems itself to be a problem with two conflicting objectives: spanning the tradeoff curve versus Pareto optimality.

5. Conclusions

In summary, the NPGA has been applied to a groundwater quality management problem consisting of active remediation by pump-and-treat. The sensitivity of the algorithm to the parameters that control the behavior of the algorithm was assessed, namely population size, tournament size, and niche radius. A population of 100 designs gave the best performance in terms of distribution of designs along the tradeoff curve. For the two- and five-well cases, smaller populations (50 designs) lacked sufficient diversity to adequately explore the range of tradeoff curves. While larger population sizes (150 designs) offered more diversity, a larger fraction of the generated designs was inferior, for all of the cases. When the population size is at an optimal value, the NPGA is less sensitive to selection pressure, values of crossover and mutation probabilities, and decision variable precision.

There is an optimal degree of selection pressure, with respect to the span of the tradeoff curve and the number of Pareto optimal points. Increasing the selection pressure via the tournament size produced solutions with a more complete span of the tradeoff curve. However, too much selection pressure resulted in more inferior solutions. Increasing the amount of selection pressure via increasing the tournament size (from 2 to 10) produced solutions that spanned wider tradeoff curves, but were inferior in some portions of the curve. Niching appears to increase the span of the tradeoff curve for various population sizes and selective pressures (i.e., tournament sizes), but its effect on Pareto optimality is unclear. It appears that the results are insensitive to the value of the parameter that controls niching (σ_{share}), as long as some niching is allowed.

A series of test problems was conducted where the NPGA was compared to two other methods, SGA and RS. As the problems increased in complexity, by considering additional decision variables, the NPGA was more effective and efficient than either the SGA or the RS in finding more Pareto optimal designs that span the entire tradeoff curve. When applied to a 15-well scenario, the niched Pareto genetic algorithm outperformed the SGA by finding 352 Pareto optimal designs with 30% less effort than the three single-objective runs. The NPGA produced solutions that were 25–250% better performance than the RS, measured in terms of the mass remaining for a given cost.

The NPGA can be an effective method for producing tradeoff curves for subsurface remediation problems. Tradeoff curves such as those presented here may give decision makers the capability of making better informed decisions. The NPGA can accommodate additional objective functions, such as maximizing reliability or minimizing remediation time. The algorithm is flexible with respect to the number and types of decision variables that can be considered. For example, future studies may include extraction rates as a function of location and time. Further tests of the applicability of this approach should consider realistic, contaminated sites, especially those with a significant degree of heterogeneity.

Acknowledgements

The authors wish to thank Reed M. Maxwell, from the Lawrence Livermore National Laboratory, for the gracious use of his numerical groundwater flow and transport simulators. The majority of this work has been made possible by an EPA Office of Research and Development grant under the contract number CR826614-01-0. A portion of the computational work was conducted on the MTU Computational Science and Engineering program's High Performance Computing Platform, funded by NSF. Any views presented herein by the authors are not necessarily reflective of the views of the EPA or NSF.

References

[1] Chan Hilton AB, Culver TB. Constraint handling for genetic algorithms in optimal remediation design. J Water Resour Plan Manage 2000;126(3):128–37.

[2] Cieniawski SE, Eheart JW, Ranjithan S. Using genetic algorithms to solve a multiobjective groundwater monitoring problem. Water Resour Res 1995;31(2):399–409.

[3] Das D, Datta B. Development of multiobjective management models for coastal aquifers. J Water Resour Plan Manage 1999;125(2):76–87.

[4] Dougherty DE, Marryott RA. Optimal groundwater management, 1, simulated annealing. Water Resour Res 1991;27(10): 2493–508.
[5] Fetter CW. Contaminant hydrogeology. 2nd. ed. Upper Saddle River, NJ: Prentice Hall; 1999.
[6] Fonseca CM, Fleming PJ. Multiobjective optimization and multiple constraint handling with evolutionary algorithms I: a unified formulation. IEEE Trans Systems, Man, Cybernetics, Part A: Systems Humans 1998;28(1):26–37.
[7] Freeze RA, Gorelick SM. Convergence of stochastic optimization and decision analysis in the engineering design of aquifer remediation. Ground Water 1999;37(6):934–54.
[8] Goldberg DE. Genetic algorithms in search, optimization, and machine learning. Reading, MA: Addison-Wesley; 1989.
[9] Goldberg DE, Richardson J. Genetic algorithms with sharing for multimodal function optimization. In: Grefenstette JJ, editor. Genetic algorithms and their applications: proceedings of the second international conference on genetic algorithms. San Mateo, CA: Morgan Kaufmann; 1987. p. 41–9.
[10] Gorelick SM. A model for managing sources of groundwater pollution. Water Resour Res 1982;18(4):773–81.
[11] Gorelick SM, Voss CI, Gill PE, Murray W, Saunders MA, Wright MH. Aquifer reclamation design: the use of contaminant transport coupled with non-linear programming. Water Resour Res 1984;20:415–27.
[12] Hand DW. Personal communication, Department of Civil and Environmental Engineering. Houghton, MI: Michigan Technological University; 2000.
[13] Horn J, Nafpliotis N, Goldberg DE. A niched Pareto genetic algorithm for multiobjective optimization. In: Proceedings of the First IEEE Conference on Evolutionary Computation (ICEC '94), Piscataway, NJ: IEEE Service Center; 1994. p. 82–7.
[14] Horn J. The nature of niching: genetic algorithms and the evolution of optimal, cooperative populations. PhD thesis, University of Illinois at Urbana Champaign, Urbana, Illinois, 1997.
[15] Horn J, Nafpliotis N. Multiobjective optimization using the niched Pareto genetic algorithm. Technical Report IlliGAl Report 93005, University of Illinois at Urbana-Champaign, Urbana, IL, 1993.
[16] Huang C, Mayer AS. Pump-and-treat optimization using well locations and pumping rates as decision variables. Water Resour Res 1997;33(5):1001–12.
[17] LaBolle EM, Fogg GE, Tompson AFB. Random-walk simulation of transport in heterogeneous porous media: local mass-conservation problem and implementation methods. Water Resour Res 1996;32(3):583–93.
[18] Maxwell RM. Understanding the effects of uncertainty and variability on groundwater-driven health risk assessment. PhD dissertation, University of California, Berkeley, CA, 1998.
[19] Maxwell RM, Pelmulder SD, Tompson AFB, Kastenburg WE. On the development of a new methodology for groundwater driven health risk assessment. Water Resour Res 1998;34(4):833–47.
[20] Means RS. Environmental remediation cost data – assemblies. 6th annual ed. R.S. Means Co. and Talisman Partners, Englewood, CO, 2000.
[21] McKinney DC, Lin M. Genetic algorithm solutions of groundwater management models. Water Resour Res 1994;30(6):1897–906.
[22] Richardson JT, Palmer MR, Liepins G. Hilliard M. Some guidelines for genetic algorithms with penalty functions. In: Proceedings of the Third International Conference on Genetic Algorithms. San Mateo, CA: Morgan Kaufman; 1989. p. 191–7.
[23] Randke S, Snoeyink J. Evaluating GAC adsorptive capacity. J Amer Water Works Assoc 1983;75(8):406–18.
[24] Ritzel BJ, Eheart JW, Ranjithan S. Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. Water Resour Res 1994;30(5):1589–603.
[25] Sawyer CS, Ahlfeld DP, King AJ. Groundwater remediation design using a three-dimensional simulation model and mixed-integer programming. Water Resour Res 1995;31(5):1373–85.
[26] Wagner BJ. Recent advances in simulation-optimization groundwater management modeling. Rev. Geophys. Supplement, US Natl. Rep. Int. Union Geod. Geophys., 1995;33:1021–8.
[27] Xiang Y, Sykes JF, Thomson NR. Alternative formulations for optimal groundwater remediation design. J Water Resour Plan Manage 1995;121(2):171–81.