Optimal design of groundwater remediation systems using a multi-objective fast harmony search algorithm

Qiankun Luo · Jianfeng Wu · Xiaomin Sun ·
Yun Yang · Jichun Wu

Abstract A new multi-objective optimization methodology is developed, whereby a multi-objective fast harmony search (MOFHS) is coupled with a groundwater flow and transport model to search for optimal design of groundwater remediation systems under general hydrogeological conditions. The MOFHS incorporates the niche technique into the previously improved fast harmony search and is enhanced by adding the Pareto solution set filter and an elite individual preservation strategy to guarantee uniformity and integrity of the Pareto front of multi-objective optimization problems. Also, the operation library of individual fitness is introduced to improve calculation speed. Moreover, the MOFHS is coupled with the commonly used flow and transport codes MODFLOW and MT3DMS, to search for optimal design of pump-and-treat systems, aiming at minimization of the remediation cost and minimization of the mass remaining in aquifers. Compared with three existing multi-objective optimization methods, including the improved niched Pareto genetic algorithm (INPGA), the non-dominated sorting genetic algorithm II (NSGAII), and the multi-objective harmony search (MOHS), the proposed methodology then demonstrated its applicability and efficiency through a two-dimensional hypothetical test problem and a three-dimensional field problem in Indiana (USA).

Q. Luo · J. Wu (✉) · X. Sun · Y. Yang · J. Wu
Key Laboratory of Surficial Geochemistry, Ministry of Education,
Department of Hydrosciences,
School of Earth Sciences and Engineering,
Nanjing University, Nanjing, 210093, China
e-mail: jfwu@nju.edu.cn

Introduction

Groundwater contamination, as one of the most important health-related environmental problems, has brought serious adverse effects on the environment and human health and attracted more and more attention around the world. Groundwater remediation is one of the major technical and environmental challenges in the field of water resources because completion of groundwater remediation often needs to undergo a relatively long time horizon of up to several decades or more. Over the past three decades, the coupled simulation-optimization (S/O) models were often used for optimal design of groundwater remediation systems and have been successfully applied to a variety of groundwater management problems (Gorelick 1983; Wagner and Gorelick 1989; Culver and Shoemaker 1992; Tiedeman and Gorelick 1993; Wagner 1995; Rizzo and Dougherty 1996; Minsker and Shoemaker 1998; Zheng and Wang 1999a; Mayer et al. 2002; Cai et al. 2003; Wu et al. 2005). Until now, most studies have dealt with the groundwater remediation design as a single objective optimization problem involving minimizing cleanup time, minimizing health risks, minimizing remediation cost, and minimizing contaminant mass remaining in the aquifer. The methods used for solving these single-objective optimization problems include linear programming (Ahlfeld et al. 1988; Ahlfeld and Mulligan 2000), nonlinear programming (Haggerty and Gorelick 1994; Karatzas and Pinder 1996), dynamic programming (Culver and Shoemaker 1992; Hsiao and Chang 2002), simulated annealing (Dougherty and Marryott 1991), genetic algorithm (McKinney and Lin 1994, 1996; Ritzel et al. 1994; Huang and Mayer 1997; Wang and Zheng 1997; Zheng and Wang 2002; Guan and Aral 2004; Ko et al. 2005), and some other evolutionary strategies (Bayer and Finkel 2004, 2007). For real-world groundwater remediation problems, decision makers often need to consider these competing objectives simultaneously. These multiple competing objectives will lead to a series of compromised solutions, known as non-dominated solutions or Pareto-optimal solutions (Deb 2001; Deb et al. 2002), instead of the single optimal solution to groundwater remediation problems.

Recently, multi-objective evolutionary algorithms (MOEAs) have been reported to solve groundwater remediation problems. Many evolutionary algorithms have proven to outperform traditional methods because of their ability to obtain a set of Pareto optimal solutions with different target units of measurement in a single optimization run. Evolutionary algorithms use a population of solutions and can be easily extended to maintain a diverse set of solutions in a single run. For instance, Ritzel et al. (1994) compared a Pareto genetic algorithm (Pareto GA) with a vector evaluated genetic algorithm (VEGA) in a multi-objective groundwater pollution containment application by the pump-and-treat (PAT) method, and the Pareto GA was shown to be superior to the VEGA. Cieniawski et al. (1995) applied four different GA-based codes to a multi-objective groundwater-monitoring problem, and compared their results with those based on simulated annealing in terms of both performance and computational burden. The multi-objective GA-based formulations were able to find a large number of convex and nonconvex points of trade off curve in a single iteration. More recently, Erickson et al. (2002) used a niched Pareto genetic algorithm (NPGA) to solve the PAT groundwater remediation problem, where two objectives were used: minimization of the remediation cost and minimization of contaminant mass remaining at the end of the remediation horizon. The NPGA was demonstrated to outperform both the single genetic algorithm (SGA) and the random search (RS) by generating a better tradeoff curve. Singh and Minsker (2008) developed a probabilistic multi-objective genetic algorithm (PMOGA) and applied it to a field-scale PAT design problem at the Umatilla Chemical Depot site at Hermiston (Oregon, USA). The results showed that the uncertainty-based PMOGA can give valuable information about remediation options, resulting in both cost-effectiveness and lower uncertainty. Singh and Chakrabarty (2010) coupled the non-dominated sorting genetic algorithm II (NSGAII) coded in C with FORTRAN programs (MODFLOW and MT3DMS) and used this methodology to obtain a tradeoff between remediation cost and clean water extraction rate. The results indicated that the proposed method is promising for wide applicability in the field of groundwater remediation. Wu et al. (2011a, b) developed an improved niched Pareto genetic algorithm (INPGA) to solve the multi-objective PAT groundwater remediation problem in the Massachusetts Military Reservation (MMR, USA) site, and added the message passing interface (MPI) for parallel computing and the operation library of individual fitness to improve calculation speed. The comparative results proved that the INPGA was superior to the NPGA in finding the tradeoff curve with a range of applicable Pareto optimal solutions.

The aim of this study is to present a new multi-objective programming algorithm, named multi-objective fast harmony search (MOFHS) algorithm, and then the proposed algorithm is coupled with the flow and transport simulators to search for Pareto-optimal solutions of the multi-objective optimization problem associated with groundwater-remediation problems.

Simulation and optimization model

Flow and transport simulation model
In this study, the flow and transport simulation model is based on the three-dimensional (3D) finite-difference groundwater flow simulator MODFLOW (McDonald and Harbaugh 1988; Harbaugh and McDonald 1996) and the 3D contaminant fate and transport simulator MT3DMS (Zheng and Wang 1999b). Under the S/O framework used in this study, the flow and transport equations have to be solved repeatedly, that is to say, the MODFLOW and MT3DMS simulators have to be called by the main optimization program repeatedly.

Multi-objective optimization model
A general multi-objective optimization problem can be stated as (Rao 1991):

$$\text{Minimize } \mathbf{y} = F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots, f_k(\mathbf{x})) \tag{1}$$

subject to

$$g_i(\mathbf{x}) \leq 0, i = 1, 2, \cdots, M \tag{2}$$

$$h_j(\mathbf{x}) = 0, j = 1, 2, \cdots, P \tag{3}$$

$$l_i \leq x_i \leq u_i, i = 1, 2, ..., N \tag{4}$$

where $\mathbf{y} = (y_1, y_2, \cdots, y_k) \in Y$, $y_i = f_i(\mathbf{x})$ is the i-th objective function among k objectives, Y is the objective function space; $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in X$, $\mathbf{x}$ is a n-dimensional decision variable vector that represents a solution, X is the set of feasible solutions, restricted by M inequality (Eq. 2) and P equality (Eq. 3) constraints; l_i and u_i are the lower and upper bounds of the i-th decision variable (x_i), respectively.

Usually, the term "Pareto-optimal solutions" or "non-dominated solutions" is used to characterize the optimal solutions to the multi-objective optimization model as given by Eqs. 1–4 in that the multiple objectives are usually conflicting with each other. Mathematically, a Pareto-optimal solution or non-dominated solution can be defined as follows: if $x^* \in X$, and if-and-only-if there is no existence of $x \in X$ such that $F(x) < F(x^*)$, then $\mathbf{x}^*$ is one of the Pareto-optimal solutions to a multiple objective optimization model in the decision variable space, X.

In this study, the generic groundwater remediation system aims to minimize the remediation cost and the contaminant mass remaining in the aquifer at the end of the remediation horizon, while satisfying some specific constraints. The first

objective function is defined as the total remediation cost through the engineering planning horizon, including the fixed cost associated with well drilling, capital cost associated with well installation, and operation cost associated with pumping and/or treatment over the full duration of the project. The second objective function is measured by the percentage of mass remaining in the aquifer at the end of the operational period of the PAT system.

This problem can be mathematically expressed as follows (Zheng and Wang 2003):

$$\text{Minimize } J_1 = RC = \alpha_1 \sum_{i=1}^{N_w} w_i + \alpha_2 \sum_{i=1}^{N_w} w_i d_i$$
$$+ \alpha_3 \sum_{i=1}^{N_w} \sum_{t=1}^{N_t} w_i |Q_i^t| \Delta t_t$$
$$+ \alpha_4 \sum_{i=1}^{N_w} \sum_{t=1}^{N_t} w_i M_i^t \quad (5)$$

$$\text{Minimize } J_2 = MR = \left(\frac{mass_{end}}{mass_0} \right) \times 100\% \quad (6)$$

where J_1 is the total remediation cost (RC) in terms of a currency unit, N_w is the number of potential pumping wells to be optimized, w_i is a binary variable indicating whether well i is drilled (yes if $w_i=1$; no if $w_i=0$), d_i is the depth of well bore associated with well i, Q_i^t is the extraction/injection rate associated with well i during the 't'th management period, N_t is the total number of management periods, and Δt_t is the duration of the 't'th management period, M_i^t is the amount of solute mass removed by well i during the 't'th management period, and α_i (i=1, 2, 3, 4) is the cost coefficient associated with well drilling, well installation, water extraction/injection, or water treatment. J_2 is the percentage of mass remaining (MR) in the aquifer at the end of the remediation horizon, $mass_0$ and $mass_{end}$ are the total solute mass in the aquifer at the beginning and end of the remediation horizon, respectively. The number of potential pumping wells (N_w) and the extraction/injection rate associated with well i (Q_i^t) are the decision variables.

The constraints for the remediation design problems can be expressed as (Zheng and Wang 2003):

$$\sum_{i=1}^{N_w} w_i \leq N_W, \quad (7)$$

$$h_j^{\min} \leq h_j \leq h_j^{\max}, j = 1, 2, \cdots, N_h \quad (8)$$

$$h_k^{out} - h_k^{in} \geq \Delta h_k^{\min}, k = 1, 2, \cdots, N_g \quad (9)$$

$$C_l^{\min} \leq C_l \leq C_l^{\max}, l = 1, 2, \cdots, N_C \quad (10)$$

$$Q_i^{\min} \leq Q_i \leq Q_i^{\max}, i = 1, 2, \cdots, N_w \quad (11)$$

Equation 7 is a constraint indicating that the total number of actual wells must not exceed a prescribed number, N_W. Equation 8 gives a set of constraints indicating that the hydraulic head at location j, h_j, must be within the specified lower and upper bounds (h_j^{min} and h_j^{max}) during any specific management period, N_h is the total number of head constraints. Equation 9 gives a set of constraints indicating that the head difference between a pair of nodes at up-gradient and down-gradient locations (h_k^{out} and h_k^{in}) must be greater than a minimum value (Δh_k^{min}) during any specific management period to ensure the efficiency of the PAT system; N_g is the total number of head pairs. Equation 10 gives a set of constraints indicating that the solute concentration at location l, C_l, must be within the specified lower and upper limits (C_l^{min} and C_l^{max}) during any specific management period; N_c is the total number of concentration constraints. Equation 11 gives a set of constraints indicating that the pumping capacity of well i at any specific management period must be within the specified minimum and maximum values (Q_i^{min} and Q_i^{max}). In most cases, only one is compulsory among these two constraints given by Eqs. 9 and 10, respectively.

Solution by optimization algorithms

To find the optimal solutions to the two-objective optimization problem by different evolutionary algorithms, the constrained model must be modified to an unconstrained fitness measure by adding the constraint violations to the objective function as penalties. The constrained optimization model is then transformed to an unconstrained one given by

$$\text{Minimize } F_1 = J_1 + V_1 + V_2 + V_3 + V_4 \quad (12)$$

$$\text{Minimize } F_2 = J_2 \quad (13)$$

and

$$V_1 = \beta_1 \cdot \max \left(0, \sum_{i=1}^{N_w} w_i - N_W \right) \quad (14)$$

$$V_2 = \beta_2 \cdot \sum_{j=1}^{N_h} \max \left(0, \frac{h_j^{\min} - h_j}{h_j^{\max} - h_j^{\min}}, \frac{h_j - h_j^{\max}}{h_j^{\max} - h_j^{\min}} \right) \quad (15)$$

$$V_3 = \beta_3 \cdot \sum_{k=1}^{N_g} \max \left(0, \frac{\Delta h_k^{\min} - (h_k^{out} - h_k^{in})}{\Delta h_k^{\min}} \right) \quad (16)$$

$$V_4 = \beta_4 \cdot \sum_{l=1}^{N_c} \max\left(0, \frac{C_l^{min} - C_l}{C_l^{max} - C_l^{min}}, \frac{C_l - C_l^{max}}{C_l^{max} - C_l^{min}}\right) \quad (17)$$

where F_i (i=1, 2) are the fitness values; V_i (i=1, 2, 3, 4) are the penalty amount of constraint violations with respect to the constraints on the total number of wells (Eq. 7), hydraulic head (Eq. 8), hydraulic gradient direction (Eq. 9), and water quality (Eq. 10); and β_i (i= 1, 2, 3, 4) are penalty coefficients. The constraint on the pumping capacity of wells (Eq. 11) is implicitly represented by the search coding.

Multi-objective fast harmony search algorithm

Improved fast harmony search (IFHS) algorithm

The harmony search (HS) proposed by Geem et al. (2001) is a nature-inspired algorithm that mimics the improvisation of music players. The HS algorithm uses the concept that musical performances seek a perfect state of harmony determined by aesthetic estimation, to optimize the objective function. The harmony in music is analogous to the optimization solution vector, and the musician's improvisations are analogous to the local and global search schemes in optimization techniques (Geem et al. 2001). The HS algorithm generates a new solution vector by considering all of the existing solution vectors and uses harmony memory considering rate (HMCR) and pitch adjustment rate (PAR) for finding the solution vector in the search space, unlike the GA (only considering two parent vectors; Geem et al. 2001; Geem 2010). These features increase the flexibility of the HS algorithm and then produce better solutions. Thus, the HS algorithm is of good ergodicity and has been successfully applied in many areas such as ecological optimization (Geem and Williams 2007), water distribution network design (Geem 2006), structural design (Degertekin 2008), groundwater management (Ayvaz 2009), and inverse pollution source identification (Ayvaz 2010).

The iteration process of the HS algorithm is based on the parameters HMCR, PAR and bandwidth (BW). The parameter HMCR, which varies between 0 and 1, controls the balance between exploration and exploitation. For example, HMCR=0 behaves like a purely random search. While the entire solution space is explored by HMCR, PAR aims to satisfy the diversity of the harmony memory. The parameter PAR combined with BW potentially controls the convergence rate of the algorithm and fine-tunes the optimized solution vectors. When PAR is small, and BW is large, it will increase the iterations needed to find the optimal solution, showing the poor performance of the algorithm. A small BW increases the fine-tuning of solution vectors in final generations, however, in early generations BW should be set a fairly big value to increase the diversity of solution vectors. Moreover, large PAR with small BW usually leads to the improvement of solution quality in final generations and make the algorithm converge to optimal solution vectors (Mahdavi et al. 2005).

The IFHS algorithm proposed by Luo et al. (2011) incorporates two methods to enhance the accuracy and convergence rate of HS. The first method is based on the improved harmony search (IHS) algorithm proposed by Mahdavi et al. (2005). The IHS-based IFHS uses the previous iterations (see Eqs. 20 and 21 in the following) to generate the new solution vectors and no longer needs to pre-set the value of BW. This can help it quickly find the range where the optimal solution is located. The second method relies on the basic idea of GA that produces a new population in any iteration. The IFHS generates a new population of solution vectors instead of only one new single solution vector. In this way, it would not only increase the diversity of the solutions, but also accelerate the convergence speed of the algorithm. Luo et al. (2011) has shown the powerful global search ability and fast convergence rate of IFHS for single objective optimization problems associated with hydrogeological parameter identification of groundwater systems. For the sake of completeness, the main elements of IFHS are briefly recapitulated as follows:

Step I. Initialize the problem and algorithm parameters

Initialize the optimization problem and algorithm parameters: the harmony memory size (HMS), HMCR, PAR and the number of improvisations (NI).

Step II. Initialize the harmony memory

The harmony memory (HM or the HM matrix) is a memory location where all the solution vectors are stored. The HM matrix is filled with as many randomly generated solution vectors as the HMS:

$$\begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_{N-1}^1 & x_N^1 & f_i(\mathbf{x}^1) \\ x_1^2 & x_2^2 & \cdots & x_{N-1}^2 & x_N^2 & f_i(\mathbf{x}^2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{HMS-1} & x_2^{HMS-1} & \cdots & x_{N-1}^{HMS-1} & x_N^{HMS-1} & f_i(\mathbf{x}^{HMS-1}) \\ x_1^{HMS} & x_2^{HMS} & \cdots & x_{N-1}^{HMS} & x_N^{HMS} & f_i(\mathbf{x}^{HMS}) \end{bmatrix}$$
$$(18)$$

where $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ represents a set of decision variables, N is the number of the decision variables, and $f_i(\mathbf{x})$ (in this study, i=1, 2) are the objective function values corresponding to the decision variable set $\mathbf{x}$. HMS is the harmony memory scale that shows the number of the solution vectors.

Step III. Improvise a new harmony memory

A new harmony vector $\mathbf{x}' = (x_1', x_2', x_3', ..., x_N')$ is generated based on three rules: (1) memory consideration, (2) pitch adjustment, and (3) random selection.

The memory consideration procedure is intuitively given as follows:

$$If\ rand(0,1) < HMCR$$
$$\quad x'_i \leftarrow x'_i \in \left\{ x_i^1, x_i^2, x_i^3, ..., x_i^{HMS} \right\}$$
$$Else \tag{19}$$
$$\quad x'_i \leftarrow x'_i \in rand(l_i, u_i)$$
$$End\ If$$

where rand(0, 1) is the random number generator for generating a uniformly distributed random number between 0 and 1, and rand(l_i, u_i) is a uniformly distributed random number between the minimum and maximum values of the decision variable x_i.

Then the pitch adjustment procedure evolves according to the parameter PAR:

$$If\ rand(0,1) < PAR(gn)$$
$$\quad x'_i \leftarrow x_i^{best} + r_1 \times (x_i^{best} - x_i)$$
$$Else \tag{20}$$
$$\quad x'_i \leftarrow x'_i$$
$$End\ If$$

and

$$PAR(gn) = PAR_{\min} + \frac{(PAR_{\max} - PAR_{\min})}{NI} \times gn \tag{21}$$

where gn is the current iteration step number, x_i^{best} representatives the current best solution, x_i is a random solution vector, r_1 is a random number between 0 and 1, $PAR_{\max}$ and $PAR_{\min}$ are the maximum and minimum value of the PAR (Mahdavi et al. 2005).

Step IV. Update the harmony memory

To update the harmony memory, the solution vectors in the old and new harmony memory matrices are ranked by their objective function values. Then the top 50 % of solution vectors are added to the new harmony memory matrix for the next iteration.

Step V. Check the stopping criterion

If the maximum number of generations is achieved, computing is terminated. Otherwise, go to step III and continue running the program.

MOFHS algorithm

Horn et al. (1994) developed an evolutionary multi-objective optimization algorithm, known as the niched Pareto genetic algorithm (NPGA), based on a suggestion by Goldberg (1989) that introduced speciation along with the theory of a spatially ordered search space. Erickson et al. (2002) used the NPGA to solve a multi-objective groundwater quality management problem involving remediation by the PAT system. Wu et al. (2011a) proposed the INPGA to promote the solving ability of

the NPGA algorithm. The improvement of the INPGA algorithm lies in the use of the Pareto solution set filter, the elite individual preservation strategy and the neighborhood space Mühlenbein mutation (Deb et al. 2002). The INPGA is comparable to the NSGA-II in finding the Pareto-optimal or near Pareto-optimal solutions to the multi-objective optimization problems (Wu et al. 2011a). Moreover, the message passing interface (MPI) for parallel computing and the operation library of individual fitness are introduced in the INPGA to improve calculation speed. Sivasubramani and Swarup (2011) proposed a multi-objective harmony search (MOHS) algorithm for an optimal power flow problem and compared its performance with that of the NSGAII method. The comparison showed that the MOHS was able to generate true and well distributed Pareto optimal solutions for multi-objective optimization problems (Sivasubramani and Swarup 2011).

In this study, the MOFHS algorithm inherits the structural framework of NPGA and improvements are made by adding the Pareto solution set filter and the elite individual preservation strategy to guarantee the diversity and accelerate the convergence of the Pareto solutions to the true Pareto fronts of multi-objective optimization problems (Wu et al. 2011a). The Pareto solution set filter is introduced to avoid the situation whereby some individuals appearing in the Pareto optimal solution may disappear because the population size of the evolution is limited. The Pareto solution set filter is able to accommodate a certain number of Pareto solutions. In each generation, the optimal individuals are stored in the Pareto solution set filter, and when a new individual is added, the filter operation is performed to ensure the solutions in the Pareto solution set filter are the "true" Pareto optimal. When the filter capacity is saturated, the individuals with the greatest population density will be removed to ensure the uniformity of the Pareto optimal solutions. The capacity of the Pareto solution set filter can be set equal to the size of population or any other reasonable value (100 is given in this study). Simultaneously, the elite individual preservation strategy is applied to preserve the best individual that has appeared during the evolution. The parent population is mixed with the offspring population into a new population, and each of them is assigned a rank according to the degree of Pareto domination ranking and crowding. The smaller the rank value is, the better the individual stands for the solution quality. The first half of the mixed population with smaller rank value will be considered as the new parent population to the next round of evolution. With this strategy, it can accelerate the convergence speed to the Pareto front, and ensure that the Pareto optimal solutions are distributed more uniformly (Deb et al. 2002).

In order to further improve the calculation speed, the operation library of individual fitness is also introduced. The operation library of individual fitness is a library that can accommodate all individuals and their fitness values. In the simulation and optimization process, the individuals and their fitness values of the initial measurement are stored in the operation library of individual fitness. In the

following iterations, each individual must be first compared with the individual library to check if the fitness value evaluation is needed. If there is a matching individual in the individual library, the flow and solute transport models (MODFLOW/MT3DMS) no longer need to run and directly read the corresponding individual fitness value. Otherwise, if the individual is not consistent with an existing library, the flow and transport simulators need to run to calculate the individual's fitness value and save the individual as well as its fitness value into the individual library. This step can improve the computational efficiency and reduce the running time of the program (Wu et al. 2011a). Figure 1 shows the flowchart of MOFHS linked with MODFLOW and MT3DMS.

Performance measures

There are two goals in a multi-objective optimization: (1) convergence to the Pareto-optimal set, and (2) maintenance of diversity in solutions of the Pareto-optimal set. Thus efforts should be made in devising two metrics to evaluate the performance of a multi-objective optimization algorithm: one for measuring the convergence of solutions to the Pareto-optimal front and the other for measuring the diversity of solutions. Many performance metrics such as the convergence metrics, the diversity metrics, and so on, have been suggested (Deb 2001; Deb et al. 2002; Zitzler

1999). In this study, two performance metrics have been used for directly evaluating the quality of solutions obtained by the MOFHS algorithm. The first metric γ (the convergence metric) measures the extent of convergence to a known set of the true Pareto-optimal front (Deb et al. 2002). For each solution among the obtained Pareto-optimal solutions, its minimum Euclidean distance apart from those solutions on the true Pareto-optimal front can be calculated. Then the average of these distances is considered as the convergence metric γ which is mathematically given by

$$d_i = \min_{j=1}^{P^*} \sqrt{\sum_{m=1}^{k} \left(\frac{f_m(h_i) - f_m(p_j)}{f_m^{\max} - f_m^{\min}} \right)} \tag{22}$$

$$\gamma = \frac{\sum_{i=1}^{H} d_i}{H} \tag{23}$$

where the parameter d_i is the minimum Euclidean distance of the i-th obtained solution apart from those solutions on the true Pareto-optimal front, P^* is the number of the true Pareto optimal solutions, H is the number of the obtained Pareto optimal solutions, and f_m^{max} and f_m^{min} are the maximum and minimum value of the m-th objective function among the entire set of solutions per iteration. The smaller the value of γ, the better the convergence toward the Pareto-optimal front. When all obtained solutions lie exactly on the true Pareto-optimal front, this metric takes a value of zero.

Although the first metric γ can provide some information about the spread in obtained solutions, this method uses a different metric to measure the spread in solutions obtained by an algorithm directly. The second metric Δ (the diversity metric) measures the extent of spread achieved among the obtained solutions defined as follows (Deb et al. 2002):

$$\Delta = \frac{d_f + d_l + \sum_{i=1}^{N-1} |d_i - \overline{d}|}{d_f + d_l + (N-1)\overline{d}} \tag{24}$$

where the parameters d_f and d_l are the Euclidean distances between the extreme solutions and the boundary solutions of the obtained nondominated set. The parameter $\overline{d}$ is the average of all distances d_i, assuming that there are N solutions on the best nondominated front. The maximum value of the above metric can be greater than one. However, an idealized distribution would make all distances d_i equal to $\overline{d}$ and $d_f = d_l = 0$. Thus, for the most widely and uniformly spread out set of nondominated solutions, the value of Δ would be zero or extremely close to zero. For any other distribution, the value of the metric
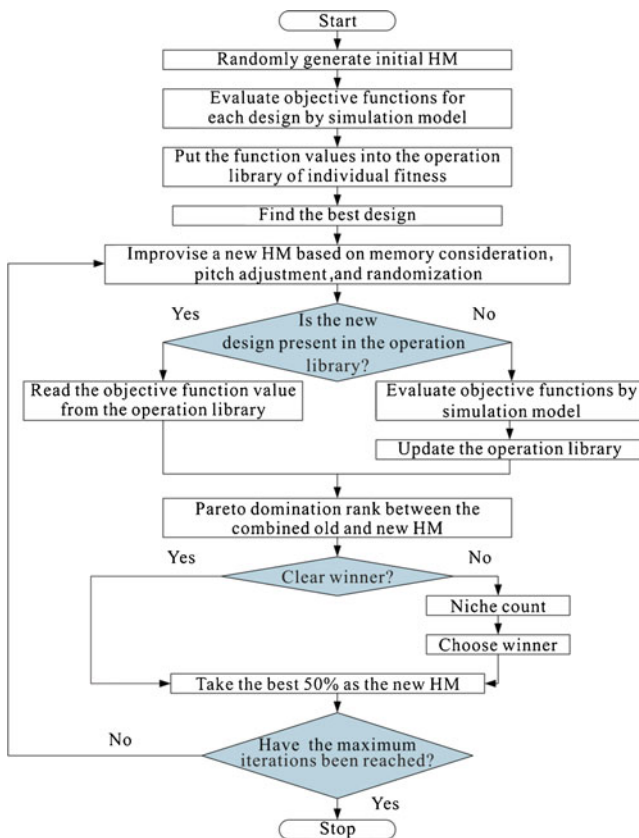

Fig. 1 Flowchart of the MOFHS algorithm for multi-objective optimal design of groundwater remediation systems. *HM* harmony memory

would be greater than zero. For two distributions having identical values of d_f and d_l, the metric Δ takes a higher value with worse distributions of solutions within the extreme solutions.

Note that in this study the true Pareto-optimal front were obtained from the mixed ranking of the optimal results from INPGA, MOFHS, NSGAII and MOHS algorithms. And all calculations are completed on a laptop computer with a 1.80 GHz CPU.

Application to a two-dimensional (2D) hypothetical problem

Description of the application

In this section, the MOFHS algorithm is applied to optimal design of a PAT system subject to the flow and transport constraints adopted from Zheng and Wang (2003). The aquifer system used for this example is a 2D model that consists of 17 rows and 23 columns (Fig. 2). The model grid spacing is uniformly 150 m along rows and columns. It is assumed that the aquifer has already been contaminated by an existing contaminant plume shown in Fig. 2. The flow domain is bounded by the constant-head boundaries on the east and west sides, and no-flow boundaries on the north and south sides. The head values for the two constant-head boundaries are set equal to 35 m on the west side and 25 m on the east side. For the transport model, boundary conditions are zero-mass-flux on the west, north and south sides, and specified advective flux on the east side. Other input parameters of the flow and transport model are given in Table 1. Four candidate pumping wells shown in Fig. 2 are pre-selected to contain and cleanup the contaminated groundwater. It is assumed that the pumping rate is within the specified minimum and maximum values (0 and 10,000 m³/d) and the remediation time horizon is fixed at 5 years.
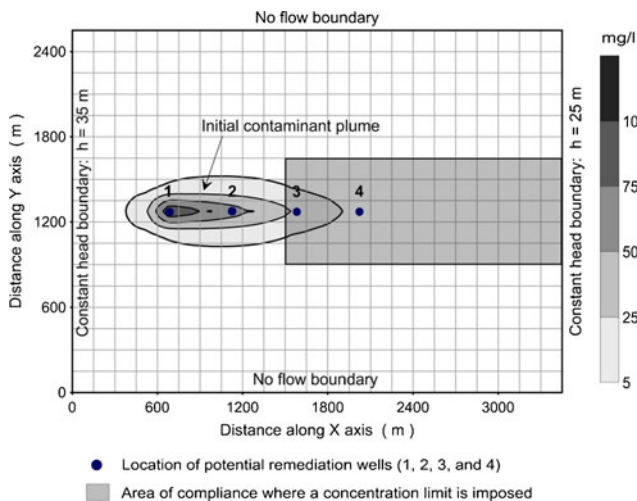

Fig. 2 Diagram showing the configuration of the pump-and-treat (PAT) system design for the hypothetical aquifer (modified from Zheng and Wang 2003)

Table 1 Input parameters of the flow and transport models for the hypothetical test problem (after Zheng and Wang 2003)

Parameter	Value
Ambient hydraulic gradient	0.003
Porosity	0.20
Transmissivity (m²/d)	500
Saturated thickness (m)	30.0
Storage coefficient	0.0001
Longitudinal dispersivity (m)	50.0
Transverse dispersivity (m)	10.0

Optimization model

The first objective considered for this application is to minimize the total cost needed to install and operate the PAT system while satisfying the constraint that the maximum concentration level (MCL) is below 20 mg/l within the area of compliance (see Fig. 2) at the end of the 5-year project horizon. For this simple application, it is supposed that the first objective function only includes two terms of the right-hand side of Eq. 5: the fixed cost associated with well drilling, and operation cost associated with pumping during the entire 5-year management period. Accordingly, the first objective function can be mathematically expressed as

$$\text{Minimize } J_1 = RC = \alpha_1 \sum_{i=1}^{Nw} w_i + \alpha_3 \sum_{i=1}^{Nw} w_i |Q_i| \Delta t \quad (25)$$

where α_1 is set to 10,000 USD (US dollars) and α_3 is set to 0.4 USD/m³ in this study.

The second objective is to minimize the total mass remaining in the aquifer at the end of the project duration, which is still given by Eq. 6. And the MCL constraint can be stated as

$$C_m \leq C* \quad (26)$$

where C_m is the calculated concentration at any monitoring location, and C^* is the MCL constraint within the area of compliance (set to 20 mg/l in this study).

Optimization results

The parameters of MOFHS used for this problem are set as follows: HMCR, 0.95; PAR_{min}, 0.1; PAR_{max}, 0.95 and the parameters of INPGA, NSGAII and MOFHS are set according to the actual situation of the applied example and references (Deb et al. 2002, Mahdavi et al. 2005, Sivasubramani and Swarup 2011, Wu et al. 2011a). In order to facilitate comparison, the number of iterations and size of population are set identically to 100 for the four optimization algorithms.

Figure 3 shows the comparison of the Pareto solutions obtained from the INPGA-based, MOFHS-based, NSGAII-based and MOHS-based optimization model runs for 100 generations. Compared with those based on NSGAII and MOHS the Pareto solutions along the tradeoff curve based on INPGA and MOFHS are more
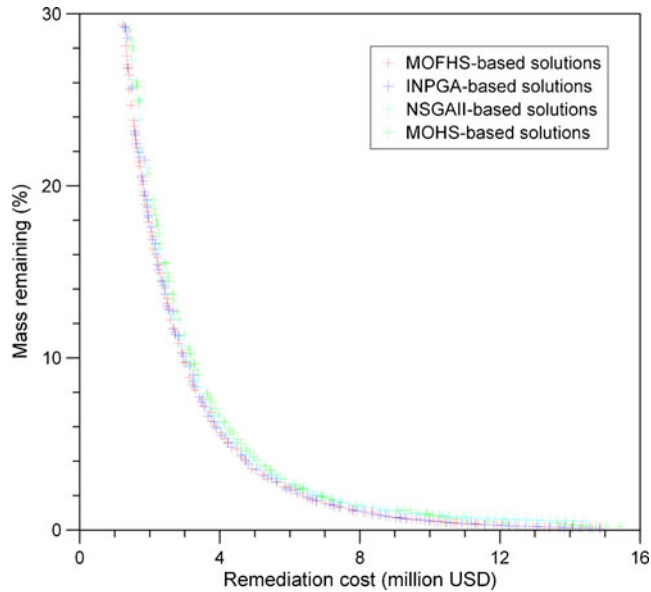
1504

Fig. 3 Comparison of the final Pareto-optimal solutions based on INPGA, MOFHS, NSGAII and MOHS algorithms

uniformly distributed and more fully reflect the tradeoff between the two conflicting objective functions. With the decrease of the remaining pollutants in the aquifer, the total remediation cost is clearly nonlinear growth. Further, with careful comparison between the results of the INPGA and MOFHS algorithm, the tradeoff curve based on MOFHS provides a more perfect range of options for decision making under different circumstances. Table 2 further shows that the quality of Pareto solutions based on MOFHS is better than that based on the other three optimization algorithms. The convergence and diversity metrics of MOFHS, γ and Δ, are 0.009968 and 0.2945, respectively, slightly smaller than those of INPGA, NSGAII and MOHS since all of these are small. Moreover, to obtain the Pareto solutions as shown in Fig. 3, a MOFHS run needs to repetitively implement 8,662 times of the flow and transport model and the optimization runtime is 609 s, while an INPGA run needs to execute 9,038 times of the flow and transport model and the runtime is about 677 s, a NSGAII run needs to calculate 10,000 times of the flow and transport model and the runtime is about 1055 s, and an MOHS run needs to call 10,000 times of the flow and transport model too and the run time is about 1,094 s (shown in Table 2).

Table 2 Comparison of computational load of the INPGA, MOF-HS, NSGAII and MOHS algorithms

Algorithm	Computing time (sec)	Convergence metric (γ)	Diversity metric (Δ)	Repetitive runs of the simulation model
INPGA	677	0.010675	0.4396	9,038
MOFHS	609	0.009968	0.2945	8,662
NSGAII	1,055	0.017302	0.5798	10,000
MOHS	1,094	0.024015	0.5994	10,000

Comparatively, the MOFHS has approximately 10.4 % more runtime saving than INPGA and approximately 72.3 and 79.6 % saving than NSGAII and MOHS, respectively. Therefore, it is suggested that the MOFHS be a little bit appealing because of its higher computational efficiency and stronger searching ability.

MOFHS parameter sensitivity
In this study, sensitivity analyses of the main MOFHS parameters including HMCR, PAR_{min}, PAR_{max} and niche radii are carried out to identify the importance of input parameters and quantify the effects of parameter variations on the optimization results.

Figure 4, showing the tradeoff curves of the remediation cost versus the percentage of mass remaining, is obtained from four MOFHS runs for different HMCR values of 0.35, 0.5, 0.75 and 0.95. Note that the values of PAR_{min}, PAR_{max} and niche radii are set to 0.1, 0.95 and 0.05 respectively for this case. It is found that, for the case of fixed PAR_{min}, PAR_{max} and niche radii, the greater the value of HMCR is, the more uniformly the Pareto solutions by MOFHS are distributed along the Pareto front. This is consistent with those shown in Table 3 where the value of convergence metric γ decreases from 0.068011, 0.011062, 0.0108062 to 0.009968 and the value of diversity metric Δ decreases from 0.8763, 0.6848, 0.4907 to 0.2945 as the HMCR increases from 0.35, 0.5, 0.75 to 0.95. Moreover, the comparison of computational time required by the MOFHS for different HMCR values is also shown in Table 3. When the HMCR is set to a small value, the MOFHS runs are relatively computationally inefficient, requiring runtime up to 876 s for HMCR= 0.35 and 864 s for HMCR=0.5. However, with the increase of the HMCR value from 0.5, 0.75 to 0.95, the computational time of the MOFHS run decreases from 864, 787 to 609 s, respectively. This shows that an adequately large HMCR could help MOFHS produce better optimization results in terms of both tradeoff curve and computational effort.

Figure 5 shows how well the Pareto-optimal solutions depend on different combinations of PAR_{min} and PAR_{max}, supposing that HMCR and niche radii are set to 0.95 and 0.05 respectively. In the cases of the combinations of the fixed PAR_{min} (PAR_{min}=0.1) and fluctuating PAR_{max}, the Pareto solutions along the tradeoff curve are distributed more uniformly as the value of PAR_{max} increases from 0.35, 0.65 to 0.95 as shown in Fig. 5a–c. Comparison of the values of convergence metric γ and diversity metric Δ in Table 4 shows that when the combination of (PAR_{min}, PAR_{max}) is set equal to (0.1, 0.95), the algorithm produces the best result, where the values of γ and Δ are 0.009968 and 0.2945, respectively. In the cases of the combinations of the fixed larger PAR_{min} (PAR_{min}=0.35) and varying PAR_{max} (PAR_{max}=0.65 and 0.95), the Pareto solutions along the tradeoff curves shown in Fig. 5d and e are distributed less uniformly than those shown in Fig. 5c. As shown in Table 4, the values of convergence metric γ are 0.011381 and 0.011072 corresponding to Fig. 5d and e,
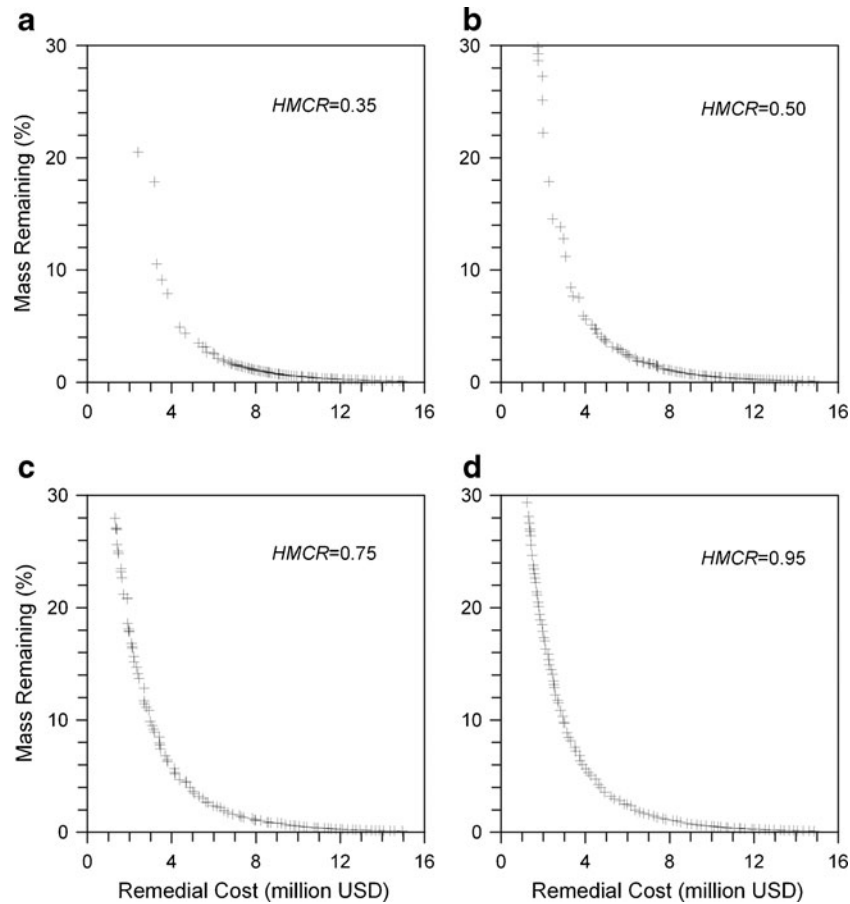
Fig. 4 **a–d** Final Pareto offline results obtained from the MOFHS for different *HMCR*. The plus symbol (+) denotes the Pareto-optimal solutions

respectively. These values are approximately 9.3 times of that of 0.009968 corresponding to Fig. 5c. Meanwhile, the values of diversity metric Δ are 0.4123 and 0.4987 corresponding to Fig. 5d and e, which are also much greater than that of 0.2945 corresponding to Fig. 5c. Even more unfortunate is that the MOFHS algorithm was trapped into premature convergence in the case of Fig. 5d (PAR$_{min}$=0.35 and PAR$_{max}$=0.65). Similarly, in the case of PAR$_{min}$=0.65 and PAR$_{max}$=0.95, the algorithm was still trapped into premature convergence even though the Pareto solutions along the tradeoff curve shown in Fig. 5f are distributed quite uniformly, where the diversity metric of Δ=0.2949 is very close to that of Δ=0.2945 based on Fig. 5c. On the other hand, there is no substantial difference in the computational efficiency between the cases as shown in Table 4 except for the two afore-mentioned premature cases. Thus, the combination of an adequately small PAR$_{min}$ and an adequately large PAR$_{max}$ could be helpful for the MOFHS to generate high-quality Pareto solutions, which corresponds with the description of Mahdavi et al. (2005).

To examine the effects of the niche radii on the algorithmic performance, several MOFHS runs were implemented with fixed *HMCR*, *PAR$_{min}$*, *PAR$_{max}$* (*HMCR*=0.95, *PAR$_{min}$*=0.1, *PAR$_{max}$*=0.95) and different niche radii set to 0.01, 0.05, 0.1 and 0.5, respectively. Shown in Fig. 6 is the comparison of the Pareto solutions obtained from the MOFHS runs with different niche radii. At first view, it is found that the Pareto solutions shown in Fig. 6a–d spread over quite a wide and uniform span of the tradeoff curve. It is very difficult to distinguish the quality of Pareto solutions by the MOFHS with different niche radii. However, as shown in Table 5, when the niche radii is set to 0.05, the MOFHS generates the smallest values of convergence metric (γ) and diversity metric (Δ). Moreover, when the niche radius is set to 0.05, the MOFHS is the least time-consuming. Unfortunately, when niche radius is set to 0.5 (shown in Fig. 6d), the algorithm was trapped into premature convergence. Therefore, selection of a suitable niche radius can help to improve the stability and robustness of the MOFHS in finding the Pareto solutions to multi-objective

Table 3 Computational load of the MOFHS for different *HMCR*

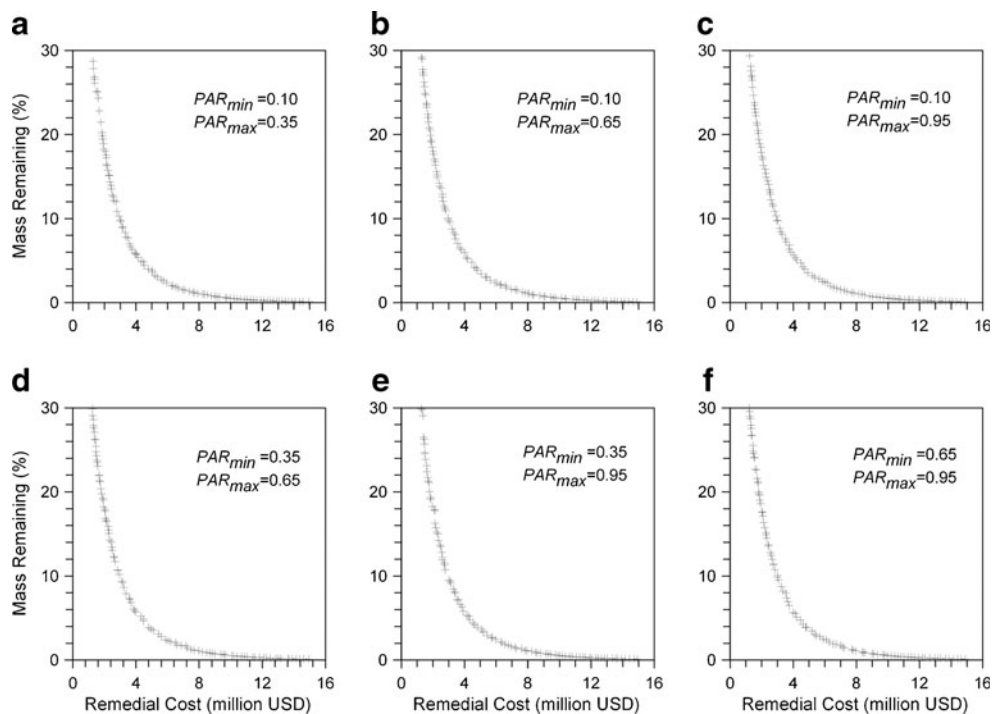| HMCR	Computing time (s)	Convergence metric (γ)	Diversity metric (Δ)
0.35	876	0.068011	0.8763
0.50	863	0.011062	0.6848
0.75	787	0.010806	0.4907
0.95	609	0.009968	0.2945

Fig. 5 **a–f** Final Pareto offline results obtained from the MOFHS for different combinations of PAR_{min}, PAR_{max}. The plus symbol (+) denotes the Pareto-optimal solutions

optimization problems. Comparison of the performance of the MOFHS shown in Table 5 shows that 0.05 is the best choice of the niche radius for this synthetic application.

Application to a 3D field problem in Indiana

The field application considered in this study is the optimization of an existing PAT system at a gasoline terminal site at Granger, Indiana (Fig. 7a). Extensive field investigations showed that groundwater beneath and down-gradient of the site was contaminated by dissolved compounds associated with petroleum hydro-carbons (Hathaway and Andrews 1990). The study aquifer of interest consists of an upper deposit unit of fine and medium grain sands with occasional discontinuous lenses of silty sand, and a lower deposit unit of medium to coarse sands with some gravels (Fig. 7c). Wang and Zheng (1997) developed a GA-based single

Table 4 Computational load of the MOFHS for different PAR (PAR_{min}, PAR_{max})

(PAR_{min}, PAR_{max})	Computing time (s)	Convergence metric (γ)	Diversity metric (Δ)
(0.10, 0.35)	642	0.010841	0.3963
(0.10, 0.65)	614	0.011674	0.4011
(0.10, 0.95)	609	0.009968	0.2945
(0.35, 0.65)	554	0.011381	0.4123
(0.35, 0.95)	630	0.011072	0.4987
(0.65, 0.95)	566	0.011611	0.2949

objective optimization model for optimal design of groundwater remediation system for this site. Figure 7b shows the configuration of the flow and transport models developed by Wang and Zheng (1997). The finite-difference mesh for the 3D flow and transport models consists of 61 rows, 40 columns and 4 layers, covering an area of approximately 3,536 m (11,600 ft) in the X-direction by 6,233 m (20,450 ft) in the Y-direction. The vertical four modeled layers are uniformly 7.6 m (25 ft) thick. The upper two layers represent the fine-to-medium sand unit, and the lower two layers represent the coarse sand unit. The horizontal mesh spacing is uniformly 15.2 m by 15.2 m (50 ft by 50 ft) in the detailed study area (labeled as zone ABCD in Fig. 7b) and increases gradually toward the model boundaries. Figure 7b shows the eight potential pumping wells of the PAT system and the initial concentrations of 1,2-DCA (1,2-dichloroethane) in groundwater in the detailed study area of layer 3. The boundary conditions for the flow model are specified-head for the four sides and no flow at the bottom. The boundary conditions for the transport model are no mass flux on the four sides. Primary input parameters for the flow and transport models for this site are listed in Table 6. Details of the flow and transport model can also be found in previous work (Wang and Zheng 1997; Zheng and Wang 2003).

In the S/O approach, the flow model is assumed to be steady-state. The two conflicting objectives considered in the current study are similar to the previous 2D hypothetical example except that the fixed capital cost is negligible compared with the pumping and treatment cost. Thus, the
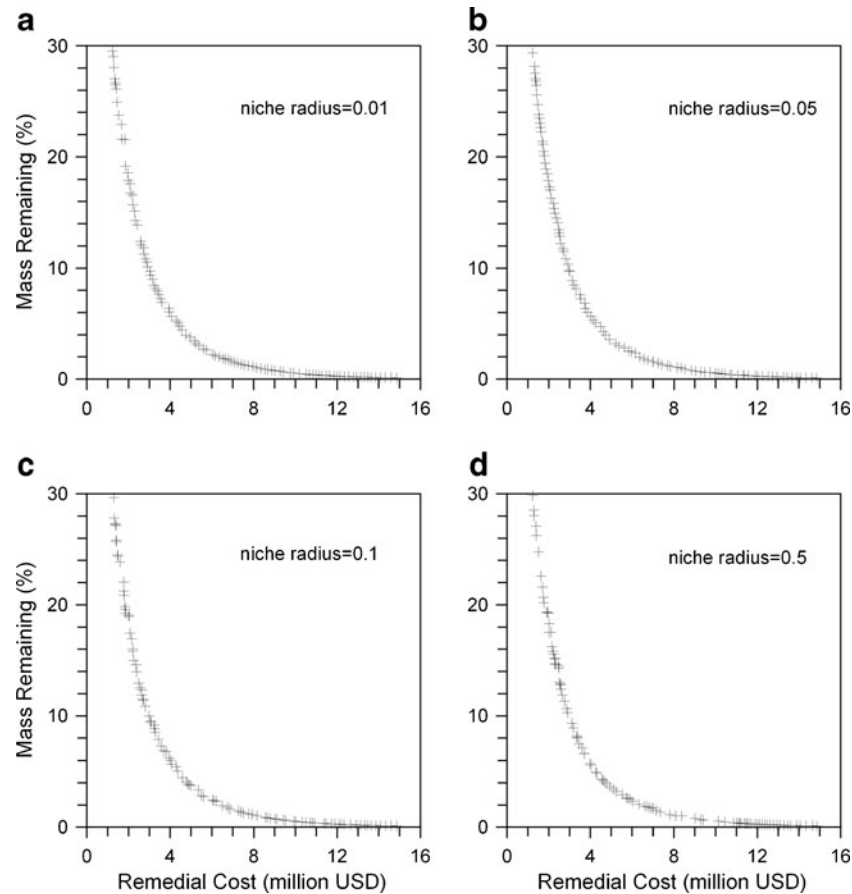
Fig. 6 **a–d** Final Pareto offline results obtained from the MOFHS for different niche radii. The plus symbol (+) denotes the Pareto-optimal solutions

first objective function for this problem can be simplified as the total pumping volume given in the following.

$$\text{Minimize} \, J_1 = \Delta t \sum_{i=1}^{8} |Q_i| \qquad (27)$$

where Q_i is the pumping rate of well i, Δt_i is the duration of pumping associated with well i.

The second objective function and the MCL constraint are also given by Eqs. 6 and 26, respectively. Note that, for this field application, the MCL is set to 42 mg/l within the entire model by the end of 500 days (Wang and Zheng 1997). The pumping capacity for each well is specified to be 850 m³/d (30,000 ft³/d).

In order to further demonstrate the applicability and efficiency of the MOFHS algorithm, comparisons are made between INPGA, MOFHS, NSGAII and MOHS

algorithms. The parameters of algorithms are the same as those set for the 2D hypothetical problem.

Figure 8 shows the tradeoff curve obtained from the four optimization algorithms run for the Indiana site. It is obvious that the tradeoff curve of the MOFHS method is the best of all, not only for its diversity but also for its convergence to the true Pareto front. For the computational efficiency, the MOFHS needs about 216 min to complete the optimization run for this application, while the INPGA needs 223 min, the NSGAII needs 364 min, and the MOHS needs 362 min.

In Fig. 8, any point on the tradeoff curve based on MOFHS corresponds to an optimal pumping strategy that satisfies all the constraints and achieves the clean-up targets by the end of 500 days. There is a nonlinear relationship between the total pumping volume and the mass remaining in the aquifer. With the decrease of the total pumping volume, the contaminant mass remaining in the aquifer will increase polynomially, approximately given by a quadratic equation with a coefficient of determination of 0.9987:

Table 5 Computational load of the MOFHS for different niche radii

Niche radii	Computing time (sec)	Convergence metric (γ)	Diversity metric (Δ)
0.01	731	0.010173	0.4009
0.05	609	0.009968	0.2945
0.10	717	0.011576	0.5193
0.50	681	0.010507	0.6078

$$J_2 = 6.57 \times J_1^2 - 37.18 \times J_1 + 94.47 \qquad (28)$$

where J_1 (total pumping volume) and J_2 (contaminant mass remaining) are measured in the units of million cubic meters and dimensionless percentages, respectively.
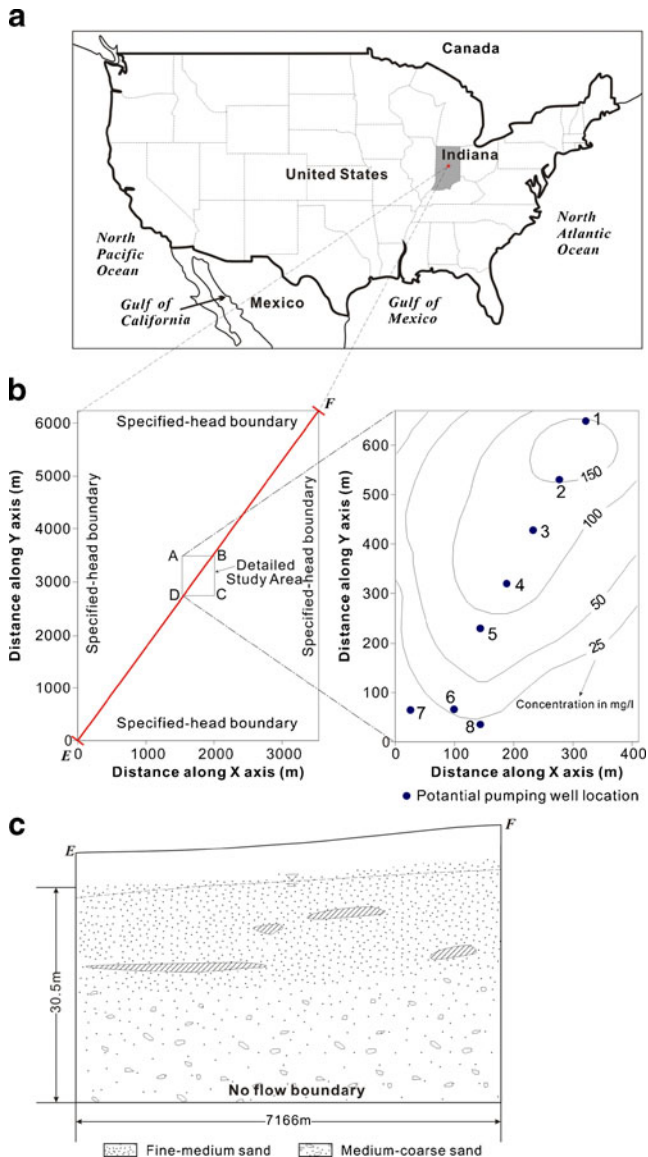
Fig. 7 **a** The field application site at Granger, Indiana (USA). **b** The configuration of the entire model domain and the initial plume in the detailed study area labeled as *ABCD* (concentration of 1,2-DCA, in model layer 3) and potential pumping well locations. **c** The cross section map showing the alluvial deposits of the aquifer along the section plane *E–F* (modified from Wang and Zheng 1997)

Table 6 Aquifer parameters of the flow and transport models for the 3D field problem (after Wang and Zheng 1997)

| Parameter	Value
Hydraulic conductivity of fine to medium sand (m/d)	18.29 (60 ft/d)
Hydraulic conductivity of coarse sand (m/d)	158.50 (520 ft/d)
Ratio of vertical to horizontal hydraulic conductivity	0.1
Recharge rate (cm/a)	12.7 (5 in/a)
Porosity	0.3
Retardation factor	2.0
Longitudinal dispersivity (m)	3.05 (10.0 ft)
Transverse dispersivity (m)	0.61 (2.0 ft)
Saturated thickness (m)	30.48 (100 ft)
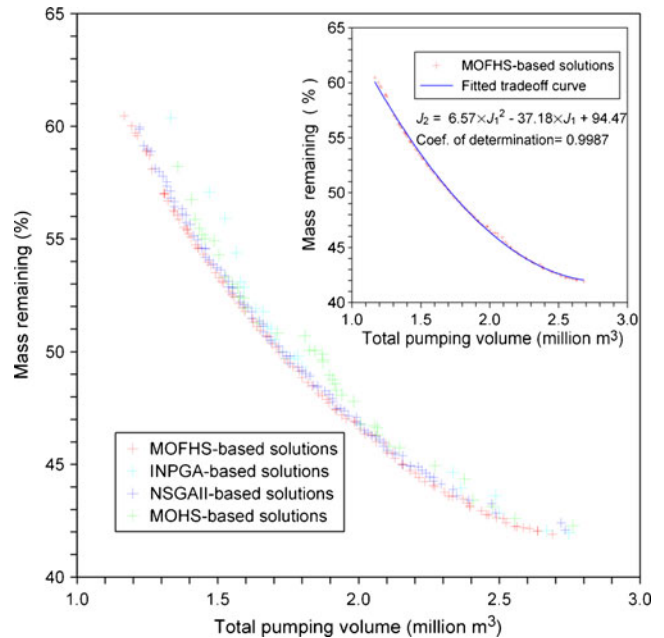
Fig. 8 Tradeoff curve for total amount of pumping rate vs. mass remaining (%) in model layer 3 based on different algorithms for the gasoline terminal site at Granger, Indiana

Note that the mass remaining in the aquifer for this field application consists of the immobile mass adsorbed by solid matrix and the mobile mass dissolved in groundwater. According to the transport model for this field application, the initial contaminant mass in the aquifer is 4,722 kg, about half of which is immobile. However, only mobile mass in an aquifer can be removed by the PAT system so that the value of J_2 (contaminant mass remaining) as shown in Fig. 8 is relatively high, approximately ranging from 42 to 60 %. This indicates that groundwater remediation by the PAT technique is not very efficient for this field application even though a little bit of immobile mass can transform into mobile with the decrease of mobile mass in an aquifer during the remediation period. As for the detailed cost-effectiveness analysis of groundwater cleanup, it is beyond the scope of this work.
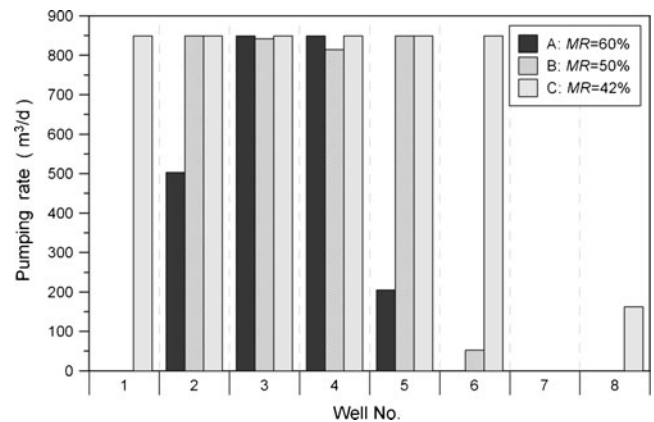

Fig. 9 Distribution of pumping volume under three different optimal strategies under specific mass remaining. constraints: *A*-MR=60 %, *B*-MR=50 %, and *C*-MR=42 %. Well locations are shown in Fig. 7b

Shown in Fig. 9 are the distributions of three optimal pumping strategies based on the MOFHS corresponding to three specific constraints on the percentages of contaminant mass remaining: 42, 50 and 60 %. Figure 9 clearly shows how the pumping volume changes with the contaminant mass remaining. When the percentage of contaminant mass remaining is set to 60 %, the contaminants in groundwater are pumped from wells 2–5, and the total pumping volume from these four wells is 2,407 m^3/d (85,000 ft^3/d). When the percentage of mass remaining is decreased to 50 %, the total pumping volume increases to 3,408 m^3/d (120,368 ft^3/d). At this time, except for wells 2–5, well 6 is also selected to be an operational well for pumping. When the percentage of mass remaining is 42 %, the total pumping volume achieves 5,260 m^3/d (185,756 ft^3/d). Similarly, except for well 7, all of the other potential wells (wells 1–6 and well 8) are selected by the optimization model to be operational pumping wells, whereas the pumping rate from well 8 is rather small, merely 163 m^3/d (5,756 ft^3/d). Thus, it can be concluded that wells 2–5 make an important contribution to removal of the contaminant mass in the aquifer. Especially, wells 3 and 4 are always at their maximum capacity while wells 2 and 5 are not. As shown in Fig. 7c (cross section along the line EF), the direction of groundwater flow is from northeast to southwest, so the upstream pollutant in the aquifer near well 2 will move downstream and could be captured by wells 3 and 4 even though well 2 is close to the centroid of the plume. Application of the MOFHS to the Indiana site demonstrates that the newly developed model can be used to solve large-scale field problems involving multiple objective groundwater contaminant optimizations under complex hydrogeologic conditions.

Conclusions

In this study, a new multi-objective optimization technique, namely MOFHS, is developed for optimal design of nonlinear water-management problems under general hydrogeological conditions. As an extension of the IFHS, this new algorithm makes improvements by adding the Pareto solution set filter and the elite individual preservation strategy so that it can guarantee the uniformity and integrity of the Pareto front of multi-objective optimization problems. Furthermore, the operation library of individual fitness is used to reduce the computational load of the algorithm. The MOFHS algorithm was then coupled with the commonly used groundwater flow and transport simulators, MODFLOW and MT3DMS, to solve multi-objective groundwater-quality management problems consisting of active remediation by the PAT system.

The coupled MOFHS-based simulation-optimization model was first used to solve a 2D hypothetical problem to demonstrate the efficacy of a PAT remediation system under the dual objectives of cost-effectiveness and mass removal. Comparison of the results based on the INPGA, NSGAII, MOHS and MOFHS show that the proposed MOFHS is an effective method for producing perfect tradeoff curves for optimal groundwater remediation design. Furthermore, the sensitivities of MOFHS parameters including HMCR, PAR_{min}, PAR_{max} and niche radii were assessed for the hypothetical application. The results show that the best parameter combination for the hypothetical application is: HMCR, 0.95; PAR_{min}, 0.1; PAR_{max}, 0.95; niche radii, 0.05.

The MOFHS was then applied to optimize an existing PAT system involving 1,2-DCA and complex hydrogeologic conditions at a gasoline terminal site at Granger, Indiana. Comparison of the results based on the INPGA, NSGAII, MOHS and MOFHS demonstrated that the proposed MOFHS is flexible in coping with different number of decision variables, and the coupled MOFHS-based model can successfully yield Pareto-optimal solutions (tradeoff curves) for both hypothetical and practical applications.

Further work should focus on investigating the applicability of the MOFHS in solving remediation system-design problems under consideration of uncertainty of contaminant source and aquifer properties. Another work is to improve the computational efficiency of the MOFHS in searching for Pareto-optimal solutions to address real-world multi-objective optimization problems.

References

Ahlfeld DP, Mulligan AE (2000) Optimal management of flow in groundwater systems. Academic, San Diego, CA, 185 pp

Ahlfeld DP, Mulvey JM, Pinder GF, Wood EF (1988) Contaminated groundwater remediation design using simulation, optimization, and sensitivity theory: 1. model development. Water Resour Res 24(3):431–441

Ayvaz MT (2009) Application of harmony search algorithm to the solution of groundwater management models. Adv Water Resour 32(6):916–924

Ayvaz MT (2010) A linked simulation-optimization model for solving the unknown groundwater pollution source identification problems. J Cont Hydrol 117:46–59

Bayer P, Finkel M (2004) Evolutionary algorithms for the optimization of advective control of contaminated aquifer zones. Water Resour Res 40:W06506. doi:10.1029/2003WR002675

Bayer P, Finkel M (2007) Optimization of concentration control by evolution strategies: formulation, application, and assessment of remediation solutions. Water Resour Res 43:W02410. doi:10.1029/2005WR004753

Cai X, Rosegrant MW, Ringler C (2003) Physical and economic efficiency of water use in the river basin: implications for efficient water management. Water Resour Res 39:1013. doi:10.1029/2001WR000748

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

Culver TB, Shoemaker CA (1992) Dynamic optimal control for groundwater remediation with flexible management periods. Water Resour Res 28(3):629–641

Deb K (2001) Multi-objective optimization using evolutionary algorithms. Wiley, Chichester, UK, 245 pp

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Trans Evolut Comput 6(2):182–197

Degertekin SO (2008) Optimum design of steel frames using harmony search algorithm. Struct Multidiscip Optim 36(4):393–401

Dougherty DE, Marryott RA (1991) Optimal groundwater management: 1. simulated annealing. Water Resour Res 27 (10):2497–2508

Erickson M, Mayer A, Horn J (2002) Multi-objective optimal design of groundwater remediation systems: application of the niched Pareto genetic algorithm (NPGA). Adv Water Resour 25 (1):51–65

Geem ZW (2006) Optimal cost design of water distribution networks using harmony search. Eng Optim 38(3):259–280

Geem ZW (2010) Multi-objective optimization of time-cost trade-off using harmony search. J Constr Eng Manage 136(6):711–716

Geem ZW, Williams JC (2007) Harmony search and ecological optimization. Int J Energy Environ 1(2):150–153

Geem ZW, Kim JH, Loganathan GW (2001) A new heuristic optimization algorithm: harmony search. Simulation 76(2):60–68

Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Boston, MA, 412 pp

Gorelick SM (1983) A review of distributed parameter groundwater management modeling method. Water Resour Res 19(2):305–319

Guan J, Aral MM (2004) Optimal design of groundwater remediation systems using fuzzy set theory. Water Resour Res 40: W01518. doi:10.1029/2003WR002121

Haggerty R, Gorelick SM (1994) Design of multiple contaminant remediation: sensitivity to rate-limited mass transfer. Water Resour Res 30(2):435–446

Harbaugh AW, McDonald MG (1996) Programmer's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model. US Geol Surv Open-File Rep 96-486, 220

Hathaway DL, Andrews CB (1990) Fate and transport modeling of organic compounds from a gasoline spill. In: Proc. Petroleum Hydrocarbons and Organic Chemicals in Groundwater: Prevention, Detection and Restoration. National Groundwater Association, Houston, TX, pp 563–576

Horn J, Nafpliotis N, Goldberg DE (1994) Multi-objective optimization using the niched Pareto genetic algorithms. Proceedings of the First IEEE Conference on Evolutionary Computation (ICEC'94), IEEE Service Center, Piscataway, NJ, pp 82–87

Hsiao CT, Chang LC (2002) Dynamic optimal groundwater management with inclusion of fixed costs. J Water Resour Plann Manage ASCE 128:57–65

Huang C, Mayer AS (1997) Pump-and-treat optimization using well locations and pumping rates as decision variables. Water Resour Res 33(5):1001–1012

Karatzas GP, Pinder GF (1996) The solution of groundwater quality management problems with nonconvex feasible region using a cutting plane optimization technique. Water Resour Res 32 (4):1091–1100

Ko NY, Lee KK, Hyun Y (2005) Optimal groundwater remediation design of a pump and treat system considering clean-up time. Geosci J 9(1):23–31

Luo QK, Wu JF and Yang Y (2011) An improved fast harmony search algorithm for identification of hydrogeological parameters. In: 2011 7th Int. Conf. on Natural Computation. Shanghai, China, July 2011, pp 1986–1989

Mahdavi M, Fesanghary M, Damangir E (2005) An improved harmony search algorithm for solving optimization problems. Appl Math Comput 188:1567–1579

Mayer AS, Kelley CT, Miller CT (2002) Optimal design for problems involving flow and transport phenomena in saturated subsurface systems. Adv Water Resour 25:1233–1256

McDonald MG, Harbaugh AW (1988) A modular three-dimensional finite-difference ground-water flow model. USGS Techniques of Water Resources Investigations, Book 6, US Geological Survey, Reston, VA

McKinney DC, Lin MD (1994) Genetic algorithms solution of groundwater management models. Water Resour Res 30 (6):1897–1906

McKinney DC, Lin MD (1996) Pump-and-treat ground-water remediation system optimization. J Water Resour Plann Manag ASCE 122(2):128–136

Minsker BS, Shoemaker CA (1998) Dynamic optimal control of in situ bioremediation of ground water. J Water Resour Plann Manage ASCE 124(3):149–161

Rao SS (1991) Optimization theory and application. Wiley, New Delhi

Ritzel BJ, Eheart JW, Ranjithan S (1994) Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. Water Resour Res 30(5):1589–1604

Rizzo DM, Dougherty DE (1996) Design optimization for multiple management period groundwater remediation. Water Resour Res 32(8):2549–2561

Singh TS, Chakrabarty D (2010) Multi-objective optimization for optimal groundwater remediation design and management systems. Geosci J 14(1):87–97

Singh A, Minsker B (2008) Uncertainty-based multi-objective optimization of groundwater remediation design. Water Resour Res 44:W02404. doi:10.1029/2005WR004436

Sivasubramani S, Swarup KS (2011) Multi-objective harmony search algorithm for optimal power flow problem. Electr Power Energy Syst 33:745–752

Tiedeman C, Gorelick SM (1993) Analysis of uncertainty in optimal groundwater contaminant capture design. Water Resour Res 29 (7):2139–2154

Wagner BJ (1995) Recent advances in simulation-optimization groundwater management modeling. Rev Geophys (S1):1021–1028

Wagner BJ, Gorelick SM (1989) Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: from data to design. Water Resour Res 25(10):2211–2225

Wang M, Zheng CM (1997) Optimal remediation policy selection under general condition. Ground Water 35(5):757–764

Wu JF, Zheng CM, Chien CC (2005) Cost effective sampling network design for contaminant plume monitoring under general hydrogeological conditions. J Contam Hydrol 77:41–65

Wu JF, Peng W, Qian JZ, Wu JC, Zheng CM (2011a) INPGA-based multi-objective management model for optimal design of groundwater remediation system: I. methodology and its experimental validation (in Chinese with English abstract). Geol Rev 57(2):277–284

Wu JF, Peng W, Qian JZ, Wu JC, Zheng CM (2011b) INPGA-based multi-objective management model for optimal design of groundwater remediation system: II. application to the MMR site (in Chinese with English abstract). Geol Rev 57(3):437–443

Zheng CM, Wang PP (1999a) An integrated global and local optimization approach for remediation system design. Water Resour Res 35(1):137–148

Zheng CM, Wang PP (1999b) MT3DMS: a modular three-dimensional multi-species transport model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems. Documentation and User's Guide, Contract Report SERDP-99-1. U S Army Engineer Research and Development Center, Vicksburg, MS

Zheng CM, Wang PP (2002) A field demonstration of the simulation optimization approach for remediation system design. Ground Water 40(3):258–265

Zheng CM, Wang PP (2003) MGO: a modular groundwater optimizer incorporating MODFLOW/MT3DMS, documentation and user's guide. University of Alabama and Groundwater Systems Research Ltd., Tuscaloosa, AL

Zitzler E (1999) Evolutionary algorithms for multiobjective optimization: methods and applications.PhD Thesis, Swiss Federal Institute of Technology (ETH), Switzerland