We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Application of Simulated Annealing in Water Resources Management: Optimal Solution of Groundwater Contamination Source Characterization Problem and Monitoring Network Design Problems

Manish Jha and Bithin Datta

1. Introduction

Estimating various characteristics of an unknown groundwater pollutant source can be formulated as an optimization problem using linked simulation-optimization. Meta-heuristics based optimization algorithms such as Simulated Annealing (SA), Genetic Algorithm (GA), Tabu Search etc. are now being accepted as reliable, faster and simpler ways to solve this optimization problem. In this chapter we discuss the suitability of a variant of traditional Simulated Annealing (SA) known as the Adaptive Simulated Annealing (ASA) in solving unknown groundwater pollutant source characterization problem. Growing anthropogenic activities and improper management of their impacts on groundwater quality has resulted in widespread contamination of groundwater worldwide. Coupled with ever increasing water demand leading to increased reliance on groundwater, it has resulted in a widespread recognition of public health risk posed by contaminated groundwater. This has triggered massive efforts for better management of groundwater quality in general and remediation of contaminated aquifers in particular. The sources of contamination in groundwater are often hidden and inaccessible. Characteristics of these pollutant sources such as their location, periods of activity and contaminant release history are often unknown. Groundwater contaminant source identification problem aims at estimating various characteristics of an unknown groundwater pollutant source using measured contaminant concentrations at a number of monitoring locations over a period of time. It has been widely accepted that for any remediation strategy to work efficiently, it is very important to know the pollutant source characteristic. A detailed account of different categories of source identification problems and various approaches to solve them has been presented in Pinder [26].

Typically, groundwater contamination is first detected by one or more arbitrarily located well or monitoring location. Unknown Groundwater source identification problem specifically attempts to ascertain the following source characteristics:

1. Source Type (point, areal etc.)
2. Spatial location and extent of the source
3. Release pattern (slug, continuous, intermittent etc.)
4. Point of time when the source first became active (start time)
5. Contaminant flux released as a function of time elapsed since start time (release history)

Source type is often obvious. In some cases, information on groundwater contaminant source location may be available from preliminary investigations. If an exhaustive record of pollutant inventory and industrial activities of the area is available, it may be possible to infer start time. Release history of the source, however, is difficult to ascertain as the source is not physically accessible for measurements and hence it is unlikely that any accurate temporal record of contaminant fluxes released from the source exists.

Release history reconstruction problem is one of the most widely studied groundwater source identification problems. Ascertaining release history of the contaminant sources from available contaminant concentration measurements is an inverse problem as it requires solving groundwater flow and transport equations backwards in time and space. The process of solving this inverse problem is essentially the process of finding various unknown characteristics of source using observed information about the transport media and the effects caused by the source. In such circumstances, a solution cannot be guaranteed, especially when observed information is sparse. Even if the solution exists, it may not be unique. This is because different combinations of various source characteristics can produce the same effect at a monitoring location. Moreover, the solution of this problem is highly sensitive to measurement errors either in the observation data or model parameters and hence this problem has been classified as an ill-posed inverse problem[33] . When this inverse problem has to be solved by using inaccurate values of media parameters such as hydraulic conductivity and porosity and contaminant concentration observed at arbitrarily placed monitoring wells, it becomes even more challenging to obtain a reliable solution.

Methods proposed in the past to solve this ill-posed inverse problem can be broadly classified as optimization approaches, analytical solutions, deterministic direct methods and probabilistic and geo-statistical simulation approaches. A detailed review of these methodologies can be found in Atmadja & Bagtzoglou [2]; Michalak & Kitanidis [25]; Bagtzoglou & Atmadja [3] and Sun et al. [31, 32]. The most effective of all suggested methods seems to be those based on optimization or probabilistic and geo-statistical simulation. Of the optimization methods, linked simulation-optimization approaches have been established as one of the most efficient methods. In this approach, a numerical groundwater flow and transport simulation model is linked to the optimization model. All the linked simulation-optimization approaches aim at solving a minimization problem with an objective function representing the difference in measured concentration and simulated concentration at various monitoring locations. The optimization model generates candidate solutions for various source characteristics. This is used as an input for the simulation model to generate estimated contaminant concentration observations at designated monitoring locations. The

optimization algorithm then calculates the value of objective function by calculating the difference in contaminant concentrations estimated by simulation model and actual observed values at the same monitoring locations over a period of time. Over a number of iterations, the optimization algorithm minimizes the objective function value. Most prominent approaches in this category are linear programming with response matrix approach [9], nonlinear optimization with embedding technique [21–23], artificial neural network approach [28–30], constrained robust least square approach [31, 32], classical optimization based approach [5–7], genetic algorithm based approach [1, 24, 27] etc.

In recent past, heuristic global search approaches such as Genetic Algorithm [12], Harmony Search, Tabu Search, Ant-Colony Optimization, Simulated Annealing[19] have developed rapidly and have been applied to a wide range of optimization problems. One of the major reasons for their popularity is the fact that these optimization methods do not easily get trapped in the local optima, thereby maximizing the probability of achieving a global optimal solution. Genetic algorithm (GA) and its variants, in particular, have been widely applied for solving unknown pollutant source identification. [11, 24, 27]. Genetic Algorithms are computational optimization algorithms that simulate the laws of natural genetics and natural selection and use it to search for the optimal solution.

Apart from GA or its variants, Simulated Annealing has also been used in solving inverse problems in groundwater management. Simulated annealing is inspired by the physical process of annealing in metallurgy which involves heating and controlled cooling of a material to reduce defects in crystal structure. The atoms are excited by heat and they become agitated while getting into higher energy states. The slow cooling allows a better chance for these atoms to achieve lower energy states than the ones they started with. In simulated annealing, a current solution may be replaced by a random "neighborhood" solution chosen with a probability that depends on the difference between corresponding function values and on a global parameter T (called temperature) that is gradually decreased in the process. Implementations of simulated annealing has been relatively limited because the traditional simulated annealing algorithm is reported to converge slower compared with GA or its variants. However, faster variants of simulated annealing have been developed and one of the most promising variants in terms of convergence speed is Adaptive Simulated Annealing (ASA) [14, 15]. The ASA code was first developed in 1987 as Very Fast Simulated Re-annealing (VFSR) [13]. Ingber & Rosen [16] showed that VFSR is at least an order of magnitude superior to Genetic Algorithms in convergence speed and is more likely to find the global optima during a time limited search.

Linked simulation-optimization based approaches are computationally intensive as the simulation model has to be run many thousands of times before an acceptable solution is produced. This has been a deterrent to any desktop based implementation of the simulation-optimization approach. Faster convergence can reduce the computational burden significantly and thereby enhance the possibility of a desktop based implementation of linked simulation-optimization approach. This paper investigates the applicability of ASA to unknown groundwater contaminant source release history reconstruction problem and compares its performance to genetic algorithm based solution. The performance evaluation of competing simulation-optimization approaches are based on a realistic scenario of missing measurement data, where contaminant concentration measurements are available a few years after the sources have ceased to exist. Apart from the convergence speed, the two algorithms are compared for their ability to produce accurate source release histories with moderately

erroneous data and with uncertainty in estimation of hydro-geological parameters.

One of the most important factors that affects the execution time and accuracy of solutions generated by linked simulation-optimization approaches is the choice of observation locations. Poorly chosen contaminant observation locations often produce misleading results and hence it becomes important that after the initial estimation of the contaminant sources, a monitoring network is designed and implemented. In this study we use a monitoring network designed specifically to enhance the efficiency of source identification. However, a detailed discussion of the methodology used for monitoring network design is beyond the scope of this book.

2. Methodology

The linked simulation-optimization approach consists of two parts. An optimization algorithm generates the candidate solutions corresponding to various unknown groundwater source characteristics. The candidate solutions are used as input in the numerical groundwater transport simulation model to generate the concentration of contaminant in the study area. The generated concentration at designated monitoring locations is matched to the observed values of contaminant concentrations at various time intervals at the same locations. The difference between simulated and observed concentration is used to calculate the objective function value which is utilized by the optimization algorithm to improve the candidate solution. The process continues until an optimal solution is obtained. A detailed schematic representation of this process of using SA as the optimization algorithm in a linked simulation-optimization model is presented in Figure 1. The classical simulated annealing (SA) algorithm has many associated guiding parameters such as the initial parameter temperature, annealing schedule, acceptance probability function, goal function etc. Effective application of the classical simulated annealing to a particular optimization problem normally involves a lot of trials and adjustments to achieve ideal values for all or most of these parameters. ASA, which is a variant of classical SA, helps overcome this difficulty to a certain extent by automating the adjustments of parameters controlling temperature schedule and random step selection thereby making the algorithm less sensitive to user defined parameters compared with classical SA. This additional ability of ASA combined with inherent ability of classical SA to find the global optimal solution even when multiple local optimums exists, makes it a natural choice for solving the groundwater pollutant source identification problem.

2.1. Governing equations

The three-dimensional transport of contaminants in groundwater can be represented by the following partial differential equation [17]

$$\frac{\partial C}{\partial t} = D_{ij} \frac{\partial}{\partial x_i} \left(\frac{\partial C}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (\vartheta_i C) + \frac{q_s}{\theta} C_s + \sum_{k=1}^{N} R_k \tag{1}$$

Where

C is the concentration of contaminants dissolved in groundwater,ML^{-3};
t is time, T;
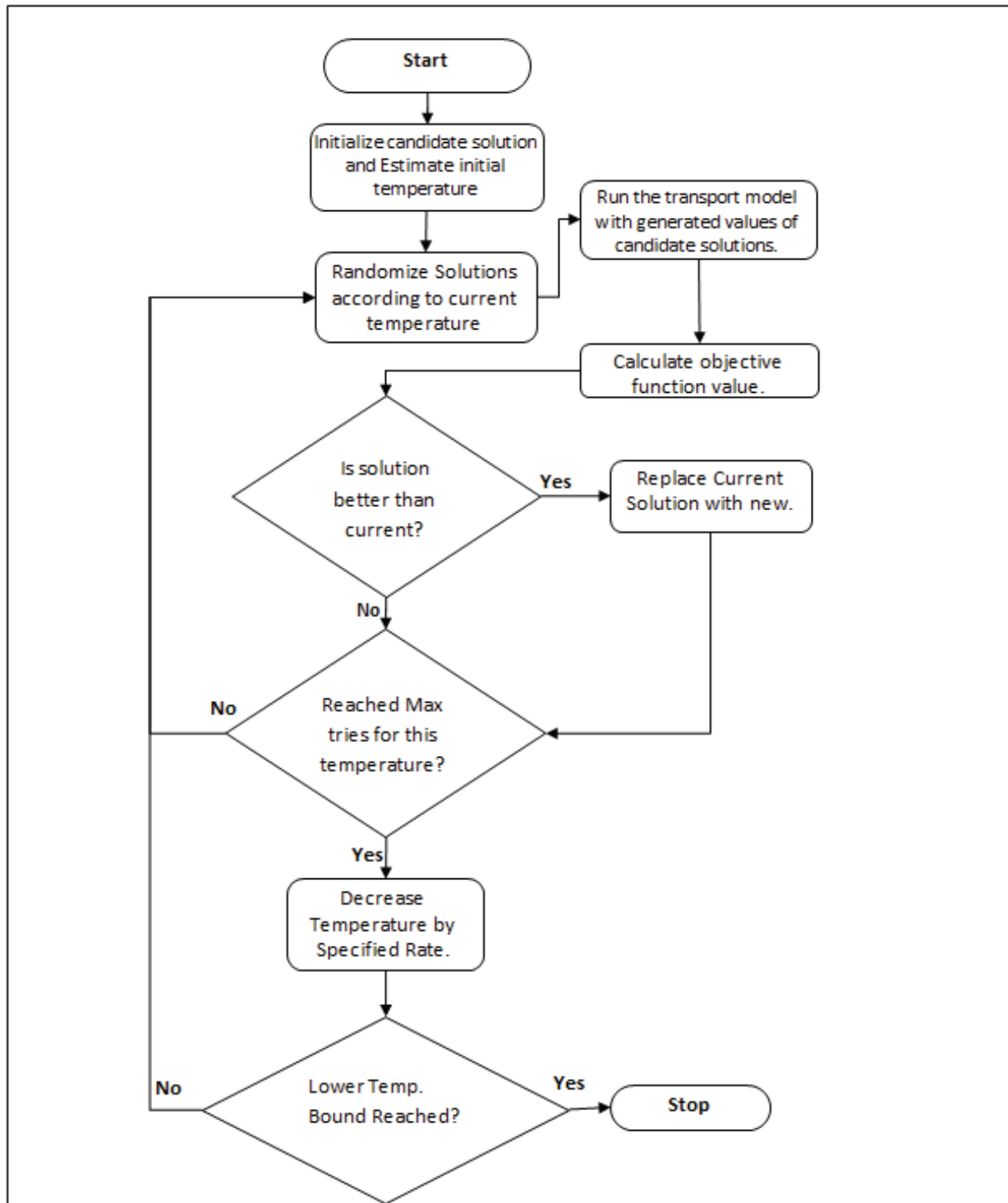x_i, x_j is the distance along the respective Cartesian coordinate axis, L;

Figure 1. Schematic Representation of Linked Simulation-Optimization Model using SA

D_{ij} is the hydrodynamic dispersion coefficient, L^2T^{-1};

ϑ_i is the seepage or linear pore water velocity, LT^{-1};

q_s is the volumetric flux of water per unit volume of aquifer representing sources (positive) and sinks (negative), T^{-1};

C_s is the concentration of the sources or sinks, ML^{-3};

θ is the porosity of the porous medium, dimensionless;

N is the number of chemical species considered;

$\sum_{k=1}^{N} R_k$ is the chemical reaction term for each of the N species considered, $ML^{-3}T^{-1}$.

In order to solve this transport equation, linear pore water velocity needs to be known for the study area. Hence, it becomes necessary to first calculate the hydraulic head distribution using a groundwater flow simulation model. The partial differential equation for groundwater flow is given by the following equation:

$$\frac{\partial}{\partial x}\left(K_{xx}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial x}\left(K_{yy}\frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial x}\left(K_{zz}\frac{\partial h}{\partial z}\right) + W = S_s\frac{\partial h}{\partial t} \tag{2}$$

Where

$K_{xx}, K_{yy},$ and K_{zz} are the values of hydraulic conductivity (LT^{-1}) along the x, y and z co-ordinate axes respectively;

H is the potentiometric head (L);

W is the volumetric flux per unit volume representing sources and/or sinks of water (T^{-1});

S_s is the specific storage of the porous media (L^{-1}); and

t is time (T).

The flow equation describes transient groundwater flow in three dimensions in a homogeneous anisotropic medium, provided the principal axes of hydraulic conductivity are aligned with the co-ordinate directions. A computer code called MODFLOW is used to solve this groundwater flow equation. MODFLOW was developed by United States Geological Survey (USGS) and is one of the most popular computer programs being used to simulate groundwater flow today. MODFLOW is based on modular finite-difference method which discretizes the study area into a grid of cells. The potentiometric head is calculated at the center of each cell.

To solve the three dimensional ground water transport equation, another computer code called MT3DMS is used. This is also a very popular computer program developed by the USGS and uses modular finite-difference just like MODFLOW. The transport simulation model (MT3DMS) utilizes flow field generated by the flow model (MODFLOW) to compute the velocity field used by the transport simulation model. [34]

2.2. Formulation of the optimization problem

It is assumed in this study that information on a set of potential source locations are available. The objective of simulation-optimization method then reduces to regenerating the source release histories at these potential source locations. Spatial and temporal contaminant concentration(C) is known at specific monitoring locations at various point of time. Candidate source fluxes are generated by the optimization algorithm. These values are used for forward transport simulations in MT3DMS. The difference between simulated and observed contaminant concentrations are then used to calculate the objective function. The objective function for this optimization problem is defined as:

$$Minimize F1 = \sum_{k=1}^{nk} \sum_{iob=1}^{nob} \left(cest_{iob}^k - cobs_{iob}^k\right)^2 . w_{iob}^k \tag{3}$$

Where,

$cest_{iob}^k$ = Concentration estimated by the identification model at observation well location iob and at the end of time period k.

nk = Total number of concentration observation time periods;

nob = Total number of observation wells;

$cobs_{iob}^k$ = Observed concentration at well iob and at the end of time period k;

w_{iob}^k = Weight corresponding to observation location iob, and the time period k.

The weight w_{iob}^k can be defined as follows:

$$w_{iob}^k = \frac{1}{(cobs_{iob}^k + n)^2} \tag{4}$$

Where n is a constant, sufficiently large, so that errors at low concentrations do not dominate the solution [18]. It is possible to include other forms of this weight.

2.3. Optimization algorithms

Of the various simulated annealing implementations, it is evident in literature that the adaptive simulated annealing algorithm converges faster [16] while maintaining the reliability of results and hence it was preferred over traditional Boltzmann annealing implementation [19]. Its application to the unknown pollutant source identification has been limited but it is potentially a good alternative because its convergence curve is steep, thereby producing better results when execution time is limited.

Currently, the most widely used optimization algorithm for solving groundwater source identification problem using linked simulation-optimization model is Genetic Algorithm and its variants. The effectiveness of ASA in solving this problem is compared against the effectiveness of GA. Genetic algorithms (GAs) are population based search strategies which are popular for many difficult to solve optimization problems including inverse problems. GAs emulate the natural evolutionary process in a population where the fittest survive and reproduce [12]. GA-based search performs well because of its ability to combine aspects of solutions from different parts of the search space. Real coded genetic algorithm was used with a population size of 100, crossover probability of 0.85 and a mutation probability of 0.05. The values were chosen based on a series of numerical experiments.

3. Performance evaluation

In order to evaluate the performance of two different optimization algorithms involving comparison of solutions obtained, it is vital to first ensure that only one solution exists. In other words, a unique solution has to be guaranteed. This is possible only under the following idealized assumptions [33]:

1. The numerical models used for simulation of groundwater flow and transport are able to provide exact solution of the governing equations in forward runs.

2. All the model parameters and concentration measurements are known without any associated errors.

3. The unknown parameter is piecewise constant.

The first assumption is valid for cases where grid size and time step used in the numerical solution tends to zero. However since the groundwater simulation models used in this study have been proven to be stable and convergent, this assumption approximately holds. The second assumption however, cannot hold in real life scenarios. Hence it becomes necessary to use synthetically generated observation values initially which can be considered free of measurement errors. The third condition is implemented by assuming that the unknown fluxes are constant in every stress period. In such conditions it approximately resembles a well posed problem. Therefore these evaluations are initially carried out for synthetic data (simulated data) with known parameter values. There is another related issue of unique solutions. Whenever numerical simulation and optimization models are used, the convergence of the solutions may be another issue related to unique solutions. These issues are discussed in [4]. In this study the use of synthetic observation data, with known hydro-geologic parameter values reduces the ill-posed nature of the problem. The uniqueness of the solution cannot be guaranteed. However, sufficient iterations were allowed to ensure convergence to the optimal solution. Performance of the source identification methodology is evaluated using synthetic data from a three dimensional aquifer study area. The synthetic contaminant concentration data is obtained by solving the numerical flow and transport simulation models.

3.1. Simulating errors in concentration measurement data

Once the global optimal solution has been obtained for the idealistic assumption, the performance evaluation of developed methodology can take into account the effects of contaminant concentration measurement errors as well as uncertainty associated with the determination of hydro-geological parameters. To test the performance for realistic scenarios, concentration measurement errors are incorporated by introducing varied amounts of synthetically generated statistical noise in the simulated concentration values. The perturbed simulated concentrations represents erroneous measurements and is defined as follows:

$$C_{pert} = C_{ns} + S_{ud} \times a \times C_{ns} \tag{5}$$

Where,

C_{pert}= Perturbed Concentration values

C_{ns}= Simulated Concentration

S_{ud}= a uniform random number between -1 and +1

a = a fraction between 0 and 1.0.

3.2. Performance evaluation criteria

The execution time of the algorithms is compared based on convergence curves which represent the value of objective function achieved versus time. To compare the ability of

competing linked simulation-optimization approaches to produce accurate source histories, the errors in estimating source fluxes accurately is also used as a performance criterion. Normalized absolute error of estimation (NAEE) is used as the measure of errors in estimation of the sources. It can be represented as:

$$NAEE(\%) = \frac{\sum_{i=1}^{S} \sum_{j=1}^{N} \left| \left(q_i^j \right)_{est} - \left(q_i^j \right)_{act} \right|}{\sum_{i=1}^{S} \sum_{j=1}^{N} \left(q_i^j \right)_{act}} \times 100 \qquad (6)$$

Where,

NAEE = Normalized Absolute Error of Estimation

S = Number of Sources = 2 in this case.

N = number of transport stress periods = 5 in this case.

$\left(q_i^j \right)_{act}$ = Actual source flux for source number i in stress period j

$\left(q_i^j \right)_{est}$ = Estimated source flux for source number i in stress period j

4. Discussion of solution results

The developed methodology was applied to a hypothetical illustrative study area with synthetically generated concentration measurements over space and time. Advantage of using a hypothetical study area lies in the fact that unknown data errors do not distort the performance evaluation of the methodology. This helps in understanding the drawbacks of developed methodology and improving it further.

4.1. Study area

The hypothetical study area is a heterogeneous aquifer measuring 2100m x 1500m x 30m and consisting of three unconfined layers as shown in Figure 2.

The East and west boundaries are constant head boundaries, whereas north and south boundaries are no flow boundaries. There are two sources (S1 and S2) of contamination. S1 is located in the top layer and S2 in middle layer. Five monitoring location (M1 through M5) are located in the first layer as shown in Figure 3. A grid size of 30m x 30m x 10m is used for finite difference based numerical calculation of groundwater flow and transport equations. Transport time step used for MT3DMS is 36.5 days. Other model parameters are listed in Table 1. Only a conservative contaminant is considered. There are two point sources of contaminants. One in the top layer and another one in the middle layer. A time horizon of 16 years is considered. Entire time horizon is divided into 5 different stress periods. The first four stress periods are each 1.5 years long and the final stress period is of 10 years duration. Sources are assumed to be active only in the first four stress periods or in the initial 6 years. Original source fluxes are presented in Table 2. It is assumed that groundwater contamination is detected at five different locations in the study area at the
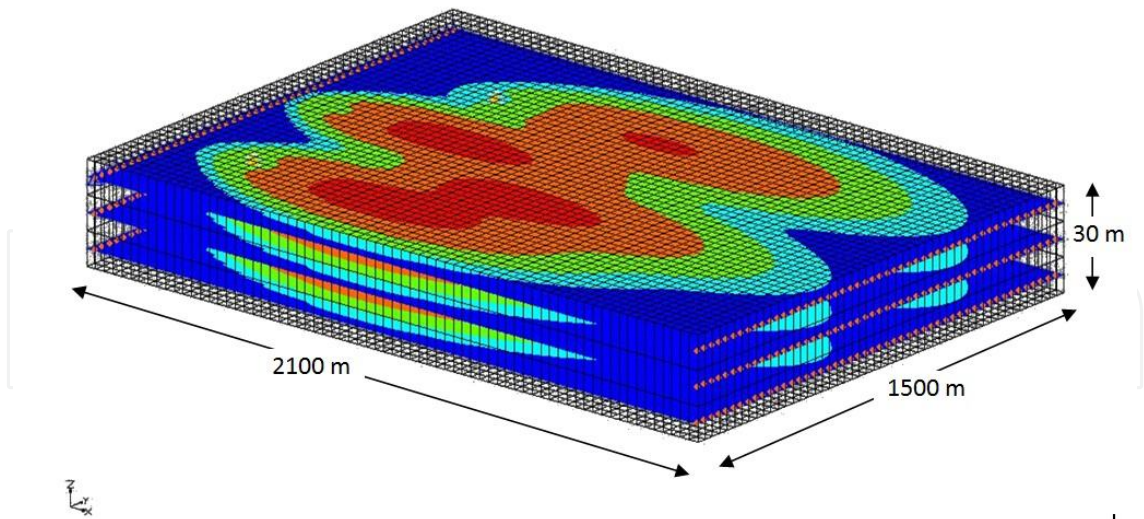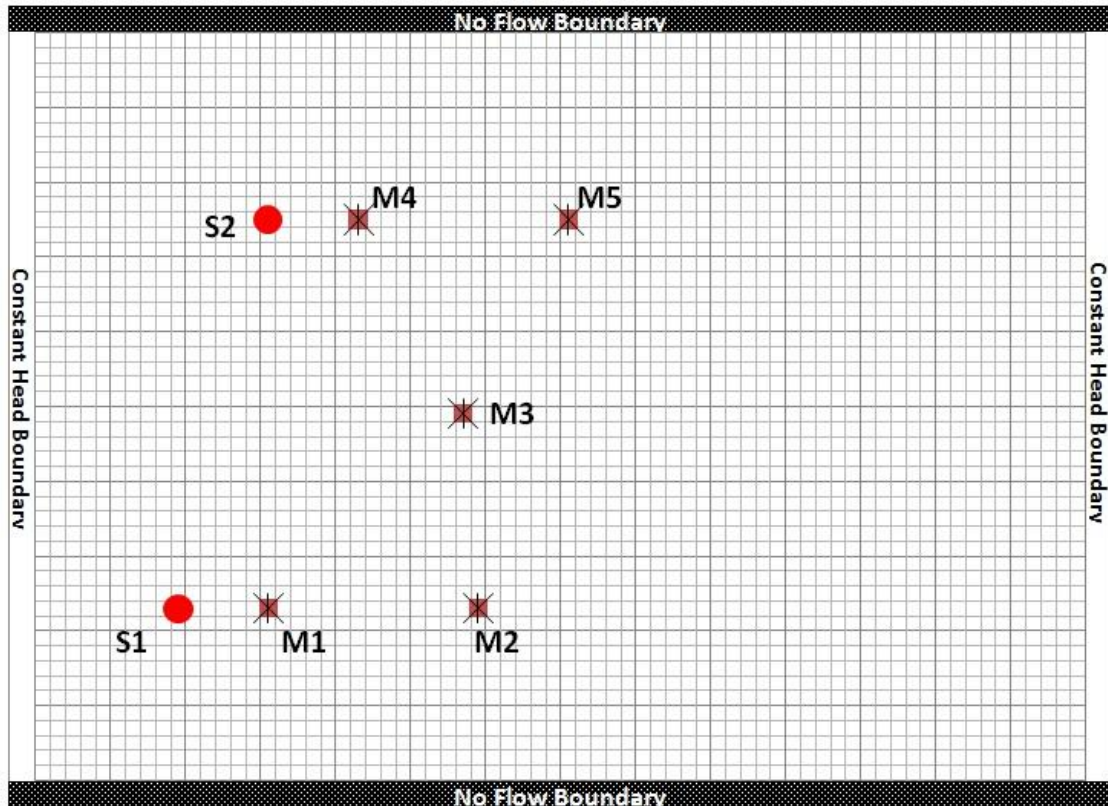
Figure 2. Illustrative Study Area

Figure 3. Top View of Study Area Showing Sources and Monitoring Locations

end of 8 th year, that is two years after the sources had ceased to exist. The observation wells are monitored for a period of 8 years starting from year 9 at an interval of 36.5 days. Observed contaminant concentration measurements at the designated monitoring locations are generated using MT3DMS as transport simulation model followed by perturbation as per Equation 5.

Parameter	Values
Length of Study Area (m)	2100
Width of Study Area (m)	1500
Saturated thickness, b(m)	30
Grid spacing in x-direction, Δx (m)	30
Grid spacing in y-direction, Δy (m)	30
Grid Spacing in z-direction, Δz (m) 10	
Hydraulic conductivity in x-direction, K_{xx} (m/day)	12
Hydraulic conductivity in y-direction, K_{yy} (m/day)	8
Vertical Anisotropy	5
Hydraulic Gradient (m/m)	0.002
Effective porosity, θ	0.3
Longitudinal dispersivity, α_L (m)	18
Transverse dispersivity, α_T (m)	4
Initial contaminant concentration (mg/l)	0.00

Table 1. Model Parameters

Sources	Layer	Row	Column	Contaminant Flux (g/sec)	
				Stress Period 1 1.5 years	6.000
				Stress Period 2 1.5 years	4.000
Source 1	1	12	15	Stress Period 3 1.5 years	8.000
				Stress Period 4 1.5 years	5.000
				Stress Period 5 10 years	0.000
				Stress Period 1 1.5 years	7.000
				Stress Period 2 1.5 years	9.000
Source 2	2	38	9	Stress Period 3 1.5 years	6.000
				Stress Period 4 1.5 years	7.300
				Stress Period 5 10 years	0.000

Table 2. Original Source Fluxes

4.2. Release history estimation with error free data

A set of error free observation data is generated. These observations are then used to evaluate the developed linked simulation-optimization methodology based on both GA and ASA. Input parameters used for GA and ASA are presented in Table 3. Every iteration of ASA based method uses one run of the groundwater transport simulation model (MT3DMS) whereas every generation of GA based method uses 100 (population size) runs of the same simulation model. Irrespective of the method, one run of the groundwater transport simulation model takes 4.281 sec to run on a Dell Optiplex®running an Intel®Core™2 Duo Processor at 2.93GHz. The execution time for one transport simulation run is however dependent on the computing platform. In order to keep the comparison independent of computing platform, both the methods were compared based on number of transport simulation runs which is directly proportional to the execution time. Both the methods were used to estimate source release histories using the error free data. In order to verify the convergence of each optimization method, time of run was made practically unconstrained. It was found that eventually both the optimization algorithms were able to achieve an objective function value very close to zero and identified the release history accurately. The objective function

convergence profile as well as estimated fluxes were plotted at the end of 40,000 simulation runs of the groundwater transport model. Minimum value of objective function achieved is plotted against number of runs of the transport simulation model. The estimated flux values for both the sources in each stress period is also plotted against actual source fluxes. Convergence profile and source flux estimates are shown in Figure 4. Convergence profile
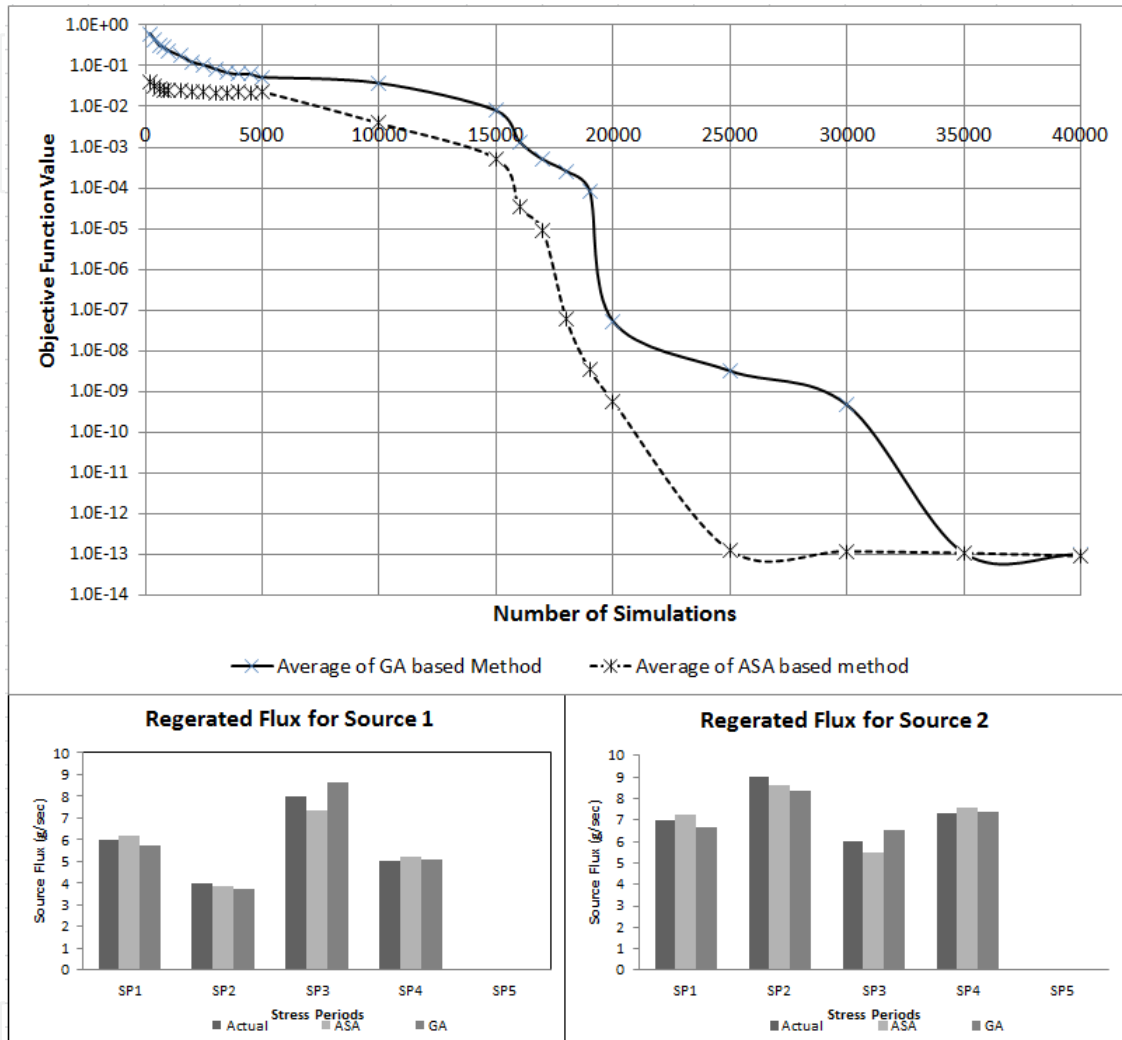

Figure 4. Convergence Profile and Estimated Release History with Error Free Data

shows that the objective function value for source identification model converges to a value very close to zero with about 5,000 simulation runs. However, further convergence is accelerated when using ASA algorithm. From these results, it can be concluded that the developed methodology is able to achieve optimal solution for an ideal error free scenario which resembles a well-posed problem.

4.3. Release history estimation with erroneous data

Five sets of erroneous observation data are generated with the formulation described in Equation 5. The value of fraction 'a' is specified as 0.1. These erroneous observations are used to reconstruct the release histories of contaminant sources. Linked simulation

optimization method using ASA is compared with the one using GA as the optimization algorithm. Parameters used for both the optimization algorithms is presented in Table 3. Unlike the case with error free measurement data, in this case both the methods were used

Parameters of ASA		Parameters of GA	
Accepted to generated ratio	1.00E-06	Mutation Strategy: Polynomial Mutation Variable Boundaries : Rigid	
Cost Precision	1.00E-15	Population size	80
Maximum Cost Repeat	5	Total no. of generations	600
Temperature Ratio Scale	1.00E-05	Cross over probability	0.8
Temp. Anneal Scale	100	Mutation probability	0.05

Table 3. Parameters used in Optimization Algorithms

to reconstruct source release histories using the erroneous data with a limit on execution time. In order to make the comparison consistent by ensuring same number of simulation runs in the ASA and GA based methodologies, the number of simulation runs are restricted to 40,000. This restriction was based on the fact that increasing the number of simulation runs even to 80,000 resulted in very little improvement in the objective function value. Minimum value of objective function achieved is averaged over five solutions and is plotted against number of runs of the transport simulation model. The plot is presented in Figure 5. This
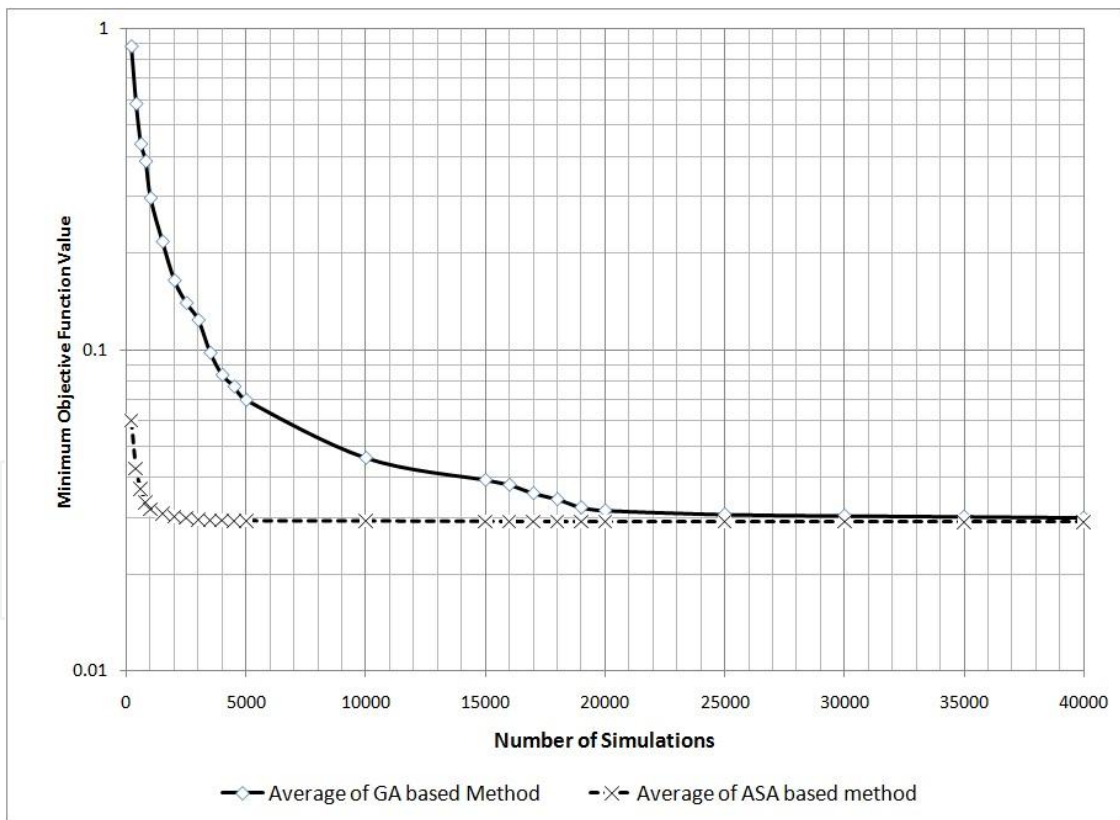

Figure 5. Convergence Plot

plot clearly shows that ASA based method converges much faster in the beginning and the GA based method is able to achieve comparable objective function values only after a much

larger simulation runs. Because of the erroneous measurement data this problem may be ill-posed and the solution may not be unique. Therefore, lower objective function values do not always mean accurate reconstruction of the release histories.

In order to test the effectiveness of the competing methods based on accuracy of solutions produced, reconstructed release histories were compared to the actual release history after every set of 10,000 transport simulation runs. The results are shown in Figure 6. It can be seen that ASA based method is more efficient compared to GA based method after 10,000 and 20,000 simulation runs. However, as the execution time increases further with increase in number of simulation runs, the release histories produced by both methods become similar. This is also confirmed from the calculated values of NAEE presented in Table 4. As the execution time increases, the NAEE of ASA based method appears to increase only slightly. This could be due to statistical variation in the five different solutions and may be attributed to the input data error. Averaging over larger number of solutions may modify this inference. NAEE of GA based method consistently improves. However, the NAEE values obtained using ASA is still better in comparison.
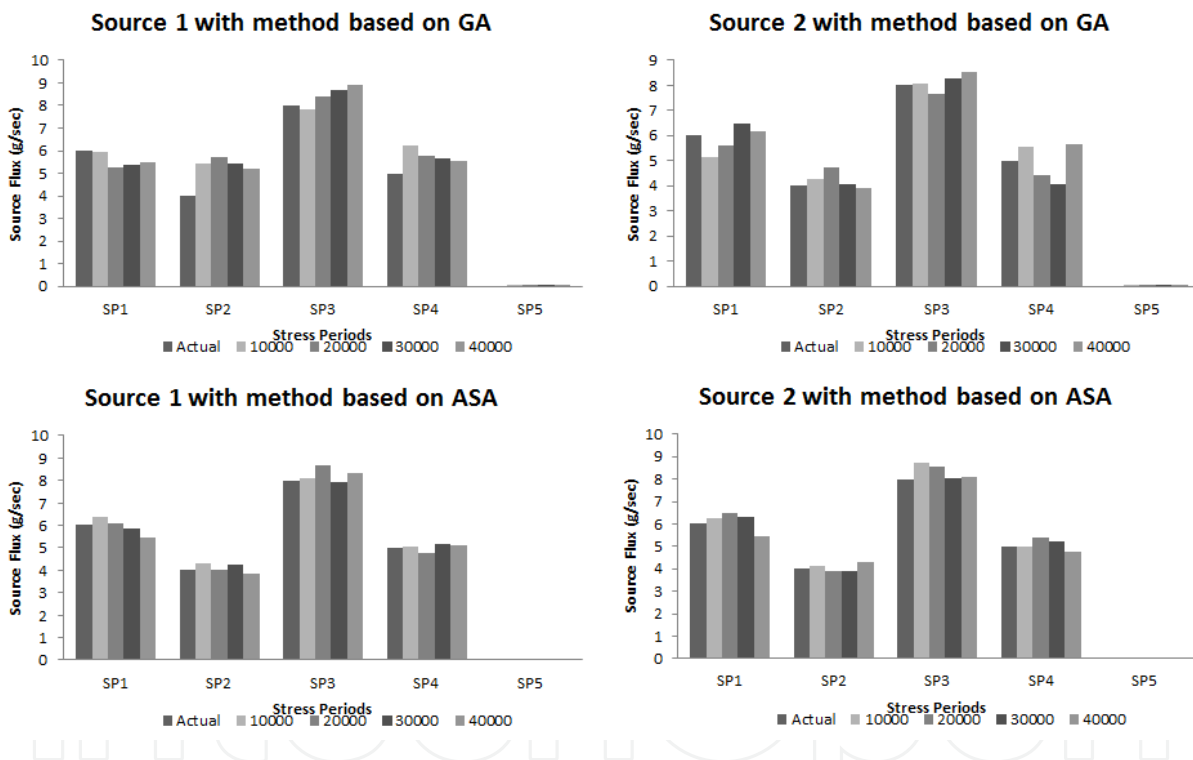

Figure 6. Reconstructed Release Histories using the competing methods

No. Of Simulation Runs	NAEE (%)	
	GA	ASA
10000	6.86	4.25
20000	6.53	4.18
30000	5.82	3.83
40000	4.26	3.62

Table 4. Normalized Absolute Error of Estimation

5. Application of simulated annealing for monitoring network design

Monitoring network design in the context of groundwater quality management essentially means specifying the spatial location of monitoring wells and frequency of sampling. Since this is one of the most cost intensive part of most contaminated groundwater remediation problems, an efficient and cost effective design of monitoring network is essential. Monitoring of groundwater quality may be necessitated by a variety of objectives such as:

1. Unknown groundwater source characterization

2. Compliance monitoring for limiting the effects of groundwater contamination

3. Better aquifer characterization

4. Hydro-geological parameter estimation

Irrespective of the various objectives, the problem of monitoring network design can be formulated as an optimization problem [8, 20]. While designing a monitoring network for estimating unknown groundwater source characteristics, the objective of optimization can be to maximize the reliability of estimated source characteristics or to minimize the total number of monitoring locations in the network or both. Compliance monitoring is aimed at minimizing the area of contamination when the contamination is first detected at monitoring network or maximizing the probability of detection of contaminant in groundwater. Often, only the average values of hydro-geological parameters of the aquifer are known. This results in uncertainty in the modeling results. In order to better characterize an aquifer, spatial distribution of hydro-geological properties should be specified. This objective can be achieved by sampling hydro-geologic parameter at sufficient locations such that the interpolated values can represent actual hydrological parameters accurately. The objective of optimization in this case is to find the minimum number of samples required to accurately represent a population of random hydro-geological parameter values. In all such cases, Adaptive Simulated Annealing can be efficiently used as the tool for optimization. Our attempts to develop classical simulated annealing algorithm for optimal design of a dedicated monitoring network for enhancing the efficiency of source identification was successful to a large extent. However, the mixed integer nature of the decision variables in a monitoring network design problem makes the application of classical simulated annealing algorithm a bit constraining. Adaptive Simulated Annealing is more suitable to solve this monitoring network design problems.

6. Suitability and sensitivity of adaptive simulated annealing

In the application discussed here, Simulated Annealing is utilized for finding the global minimum of a cost function that characterizes large and complex systems such as transport of pollutants in groundwater.Simulated Annealing, as an algorithm, is very efficient in solving non-convex optimization problems by ensuring that it does not always move downhill on a complex non-convex search space and hence avoids getting trapped in local minimum. Simulated annealing also differs significantly from conventional iterative optimization algorithms in that gross features of the final state of the system are seen at higher temperatures whereas the finer details of the state appear at lower temperatures [10]. The fact that simualted annealing ensures a global optimal solution enhances its suitability for solving

ill-posed inverse problems in general and the problem of unknown groundwater pollutant source characterization in particular.

Its ease of use and remarkable efficiency in handling complex objective functions and constraints has made simulated annealing an attractive choice for solving a wide range of complex optimization problems. However, the slow convergence and hence long time of execution of standard Boltzmann-type simulated annealing has been a constraint. Adaptive Simulated Annealing removes that constraint by making the annealing schedules decrease exponentially in annealing-time, thereby making the convergence much faster. A major difference between ASA and traditional Boltzamnn Annealing algorithms is that the ergodic sampling takes place in terms of n parameters and the cost function. In ASA the exponential annealing schedules permit resources to be spent adaptively on re-annealing and on pacing the convergence in all dimensions, ensuring ample global searching in the first phases of search and ample quick convergence in the final phases[15].

Another major advantages of using Adaptive Simulated Annealing is also the fact that the parameters of algorithm are adjusted adaptively and hence the solutions do not vary widely if parameter values are changed within reasonable limits. This is in contrast with Genetic Algorithm where even minor changes to parameters such as mutation probability, cross over probability or population size causes a significant difference in the solutions.

7. Conclusion

A linked simulation-optimization method for source identification was developed based on adaptive simulated annealing. It was applied to an illustrative study area. The results obtained were compared with those obtained using genetic algorithm, a more commonly used optimization approach. It is evident from the limited numerical experiments that adaptive simulated annealing algorithm based solutions converge to the actual source fluxes faster than genetic algorithm based solutions. This results in substantial saving in computational time. The source fluxes identified by using adaptive simulated annealing are closer to actual fluxes when compared to the results obtained using genetic algorithm, even when the observation data are erroneous and the hydro-geological parameters are uncertain. It can be concluded that adaptive simulated annealing is computationally more efficient for use in simulation-optimization based methods for identification of unknown groundwater pollutant sources, specially in a time constrained environment. Use of ASA has the potential to reduce CPU time required for solution by an order of magnitude. However, with very large number of iterations in the linked simulation-optimization approach, it is possible that the solutions obtained using GA could converge to a marginally better solution compared to that ASA based algorithm. However, it appears that ASA based solutions converge very close to the optimal solution using only a small fraction of iterations required while using GA. The relevance of contaminant monitoring locations is demonstrated. Further studies are required to develop dedicated monitoring networks which can increase the efficiency of source identification.

Author details

Manish Jha and Bithin Datta
James Cook University, Townsville and CRC CARE, Adelaide, Australia

8. References

[1] Aral, M., Guan, J. & Maslia, M. [2001]. Identification of contaminant source location and release history in aquifers, *Journal of hydrologic engineering* 6(3): 225–234.

[2] Atmadja, J. & Bagtzoglou, A. [2001]. State of the art report on mathematical methods for groundwater pollution source identification, *Environmental Forensics* 2(3): 205–214.

[3] Bagtzoglou, A. & Atmadja, J. [2005]. Mathematical methods for hydrologic inversion: The case of pollution source identification, *in* T. Kassim (ed.), *Water Pollution*, Vol. 3 of *The Handbook of Environmental Chemistry*, Springer Berlin / Heidelberg, pp. 65–96. 10.1007/b11442.
URL: *http://dx.doi.org/10.1007/b11442*

[4] Datta, B. [2002]. Discussion of "identification of contaminant source location and release history in aquifers" by Mustafa M. Aral, Jiabao Guan, and Morris L. Maslia, *Journal of Hydrologic Engineering* 7(5): 399–400.

[5] Datta, B., Chakrabarty, D. & Dhar, A. [2009a]. Optimal dynamic monitoring network design and identification of unknown groundwater pollution sources, *Water Resources Management* 23(10): 2031–2049.

[6] Datta, B., Chakrabarty, D. & Dhar, A. [2009b]. Simultaneous identification of unknown groundwater pollution sources and estimation of aquifer parameters, *Journal of Hydrology* 376(1-2): 48–57.

[7] Datta, B., Chakrabarty, D. & Dhar, A. [2011]. Identification of unknown groundwater pollution sources using classical optimization with linked simulation, *Journal of Hydro-Environmental Research* 5(1): 25–36.

[8] Dhar, A. & Datta, B. [2010]. Logic-based design of groundwater monitoring network for redundancy reduction, *Journal of water resources planning and management* 13(1): 88–94.

[9] Gorelick, S., Evans, B. & Remson, I. [1983]. Identifying sources of groundwater pollution: an optimization approach, *Water Resources Research* 19(3): 779–790.

[10] Haykin, S. [1999]. *Neural networks: a comprehensive foundation*, Prentice Hall.

[11] Hilton, A. & Culver, T. [2005]. Groundwater remediation design under uncertainty using genetic algorithms, *Journal of water resources planning and management* 131: 25.

[12] Holland, J. [1975]. Adaptation in natural and artificial systems, *Ann Arbor, MI: University of Michigan Press* 31.

[13] Ingber, L. [1989]. Very fast simulated re-annealing, *Mathematical and Computer Modelling* 12(8): 967–973.

[14] Ingber, L. [1993]. Adaptive simulated annealing (asa), *Global optimization C-code, Caltech Alumni Association, Pasadena, CA* .

[15] Ingber, L. [1996]. Adaptive simulated annealing (asa): Lessons learned, *Control and Cybernetics* 25: 33–54.

[16] Ingber, L. & Rosen, B. [1992]. Genetic algorithms and very fast simulated reannealing: A comparison, *Mathematical and Computer Modelling* 16(11): 87–100.

[17] Javandel, I., Doughty, C. & Tsang, C. [1984]. *Groundwater transport: Handbook of mathematical models*, DOE/SF/00098-T8, American Geophysical Union, Washington, DC; Lawrence Berkeley Lab., CA (USA).

[18] Keidser, A. & Rosbjerg, D. [1991]. A comparison of 4 inverse approaches to groundwater-flow and transport parameter-identification, *Water Resources Research* 27(9): 2219–2232.

[19] Kirkpatrick, S. [1984]. Optimization by simulated annealing - quantitative studies, *Journal of Statistical Physics* 34(5-6): 975–986.

[20] Loaiciga, H. [1989]. An optimization approach for groundwater quality monitoring network design, *Water Resources Research* 25(8): 1771–1782.

[21] Mahar, P. & Datta, B. [1997]. Optimal monitoring network and ground-water-pollution source identification, *Journal of water resources planning and management* 123(4): 199–207.

[22] Mahar, P. S. & Datta, B. [2000]. Identification of pollution sources in transient groundwater systems, *Water Resources Management* 14(3): 209–227.

[23] Mahar, P. S. & Datta, B. [2001]. Optimal identification of ground-water pollution sources and parameter estimation, *Journal of Water Resources Planning and Management-Asce* 127(1): 20–29.

[24] Mahinthakumar, G. & Sayeed, M. [2005]. Hybrid genetic algorithmï£¡local search methods for solving groundwater source identification inverse problems, *Journal of water resources planning and management* 131: 45.

[25] Michalak, A. M. & Kitanidis, P. K. [2004]. Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling, *Water Resour. Res.* 40(8): W08302.

[26] Pinder, G. [2009]. Optimal search strategy for the definition of a dnapl source, *Technical report*, Defense Technical Information Center, United States of America.

[27] Singh, R. & Datta, B. [2006]. Identification of groundwater pollution sources using ga-based linked simulation optimization model, *Journal of hydrologic engineering* 11: 101.

[28] Singh, R. & Datta, B. [2007]. Artificial neural network modeling for identification of unknown pollution sources in groundwater with partially missing concentration observation data, *Water Resources Management* 21: 557–572.

[29] Singh, R., Datta, B. & Jain, A. [2004]. Identification of unknown groundwater pollution sources using artificial neural networks, *Journal of water resources planning and management* 130: 506.

[30] Singh, R. M. & Datta, B. [2004]. Groundwater pollution source identification and simultaneous parameter estimation using pattern matching by artificial neural network, *Environmental Forensics* 5(3): 143–153.

[31] Sun, A. Y., Painter, S. L. & Wittmeyer, G. W. [2006a]. A constrained robust least squares approach for contaminant release history identification, *Water Resour. Res.* 42(4): W04414.

[32] Sun, A. Y., Painter, S. L. & Wittmeyer, G. W. [2006b]. A robust approach for iterative contaminant source location and release history recovery, *Journal of Contaminant Hydrology* 88(3-4): 181–196.

[33] Sun, N.-Z. [1994]. *Inverse Problems in Groundwater Modeling*, Kluwer Academic Publishers, pp. 12–37.

[34] Zheng, C., Wang, P. P., Zheng, C. & Wang, P. P. [1999]. Mt3dms: A modular three-dimensional multi-species transport model for simulation of advection, dispersion, and chemical reactions of contaminants in ground-water systems. documentation and user's guide, *Contract Report SERDP-99-1, U.S. Army Engineer Research and Development*.