



State of the Art Report on Mathematical Methods for Groundwater Pollution Source Identification

Juliana Atmadja* and Amvrossios C. Bagtzoglou†

Departments of Civil Engineering & Engineering Mechanics and Earth & Environmental Engineering, Columbia University, 500 W. 120th Street, New York, NY 10027, U.S.A.

(Received 8 March 2001, Revised manuscript accepted 10 June 2001)

The reliable assessment of hazards or risks arising from groundwater contamination problems and the design of efficient and effective techniques to mitigate these problems require the capability to predict the behavior of chemical contaminants in flowing water. Most attempts at quantifying contaminant transport have relied on a solution of some form of a well-known governing equation referred to as advection-dispersion-reaction equation. To choose an appropriate remediation strategy, knowledge of the contaminant release source and time release history becomes pertinent. As additional contaminated sites are being detected, it is almost impossible to perform exhaustive drilling, testing, and chemical fingerprint analysis every time. Moreover, chemical fingerprinting and site records are not sufficient to allow a unique solution for the timing of source releases. The purpose of this paper is to present and review mathematical methods that have been developed during the past 15 years to identify the contaminant source location and recover the time release history.

© 2001 AEHS

Keywords: inverse problem; contaminant transport; heat conduction; groundwater pollution.

Introduction

As the world's population continues to grow, the demand for fresh water will continue to increase. Seventy-five percent of the Earth's surface is covered with water. Of this almost 97% of the world's water is salty and not readily drinkable. The other 2% is locked as solids, in the form of ice caps and glaciers, leaving us with about 1% of freshwater that is available for all of humanity's needs. This small amount of freshwater can be found in the form of surface water and groundwater. Of the 1% of freshwater available, 96% is in the form of groundwater, which accounts for about half of the U.S. population's source of drinking water. From 1940 to 1990, withdrawals of fresh water from rivers, lakes, reservoirs, and other sources have increased fourfold (Engelking, 1999). However, quantity is not the only problem; the quality of the drinking water is also a concern since this vital resource is vulnerable to contamination. With the advances of technology, more and more human activities are polluting the water system.

Water pollutants may originate from a point source or a non-point source. A point source is characterized by the presence of identifiable, small-scale sources while the non-point problem refers to larger-scale, more diffuse/spread contamination originating from many smaller sources. The United States Environmental Protection Agency (USEPA) reported that as of 1996, about 40% of the nation's surveyed rivers, lakes, and estuaries are too polluted for basic uses, such as fishing and swimming (USEPA, 1998a). In most cases, pollution in the surface water is caused by runoff and/or groundwater seepage. In the USEPA 305(b) Report

(USEPA, 1998a), 37 states reported that they found potential sources of groundwater contamination. Major sources of point-source pollution are leaking underground storage tanks (LUSTs), landfills, septic systems, and hazardous waste sites. Among the major non-point sources, pesticides, leaks or spills of industrial chemicals at manufacturing facilities, runoff of salt and other chemicals from roads and highways, and fertilizers on agricultural land are high on the list. As of March 1996, more than 300,000 releases from LUSTs had been confirmed (USEPA, 1998b). Fertilizers, pesticides, and agricultural chemical facilities also contribute to the contamination in the water as non-point sources. Most of water contamination cases occur in highly developed areas, agricultural areas, and industrial zones. In addition to affecting human health, pollution is also detrimental to natural resources and ecosystems. Once a contaminant is detected, an effort should be, and is being made, to clean up the groundwater system. In 1980, the U.S. Congress passed the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) along with an accompanying legislation called the Superfund Tax Act. Due to limited resources available for the cost recovery program, the EPA must set priorities and select plan of actions in a manner and at a time that will provide for the maximum return to the Fund (USEPA, 1998c). After the priority selection, the clean up process begins according to the list. This procedure is followed by the recovery of the cost from the identified responsible private parties. In case a private party is not identified, the money comes from the Superfund "bank", which essentially is taxpayers' money. The legislation has not been an unqualified success in assigning the liability and the cost of cleanup

*Fax: (212)854-6267. E-mail: juliana@civil.columbia.edu
†Fax: (212)854-6267. E-mail: abagtzog@civil.columbia.edu

is staggering. This can be seen by the fact that in 1994, the National Academy of Sciences estimated that over a trillion dollars, or approximately \$4000 per person in the U.S., would be spent in the next 30 years on clean up of contaminated soil and groundwater (USEPA, 1999). In many cases, it is hard to find out which companies or parties are responsible for the contamination due to lack of tools to identify the pollution source. Chemical fingerprinting, state and federal agency records, and private parties' history records of handling hazardous substances are seldom sufficient to allow a unique solution for the location of sources and the timing of source releases. Therefore, there exists a need for tools that can reconstruct the plume's spatial and temporal history (NRC, 1990).

Reliable assessment of contamination problems and the design of efficient and effective techniques to mitigate them require the capability to predict the behavior of chemical contaminants in groundwater. Quantitative predictions of contaminant movement can be made only if we understand the processes controlling transport, advection, hydrodynamic dispersion, chemical, physical, and biological reactions that affect solute concentrations in the water. To predict the behavior of contaminants, the effects of each of these influences must be adequately represented in a model or group of models. One area of contaminant behavior that has been researched for quite some time is the chemical behavior aspect. Chemical fingerprinting has been a major element in environmental investigation to identify the contamination source in groundwater (Rachdawong and Christensen, 1997; Dempster, Sherwood-Lollar and Feenstra, 1997; Mansuy, Philp and Allen, 1997). However, finger printing alone is not always sufficient to provide answers to questions of source and responsibility (Stout *et al.*, 1998). In most cases, additional information such as the release and spill history of the contaminant, chemical parent product types from which contaminants have originated, and site historical records are needed. Furthermore, by accurately identifying pollution sources, the complex and lengthy process of remediation that is often initiated by finding the contamination source can be cut down.

Groundwater Contaminant Transport

Contaminants that are dissolved in the subsurface environment are transported by the following three processes: advection, mechanical dispersion, and molecular diffusion. Advection refers to the contaminant being carried along with the flow of subsurface water. This process is a result of the large-scale gradients in fluid energy (head) and it is the most significant mass transport process. The velocity of the water is described by the average speed of the water movement through the pores of the soil. Mechanical dispersion and molecular diffusion collectively are referred to as hydrodynamic dispersion. Mechanical dispersion is the main process that causes the contaminant to spread out and become diluted. The contaminant spreads due to the spatial variation of flow paths and the variation of velocity in the groundwater movement. The amount of mechanical dispersion increases with the heterogeneity of the aquifer and the scales involved. Molecular

diffusion is the process in which the contaminants move from high concentration areas to low concentration areas due to concentration gradients according to Fick's Law. This process can take place in the absence of groundwater flow. All three mechanisms mentioned above cause the contaminant to spread in the direction of flow both longitudinally and transversally (Figure 1).

In the process of being transported, many contaminants react with other compounds in the solution or the soil grains. The contaminant can be adsorbed onto or desorbed off the solid matrix. These mechanisms are known as reactions. Various types of reaction, ranging from simple sorption/desorption to radioactive decay, are evident in contaminant transport. Reactive contaminant transport is an area of active and intense research. In this paper, non-reactive contaminant transport will be introduced and used to show how it can be used to identify pollution sources.

Advection-dispersion equation

Non-reactive contaminant transport may be represented by the advection-dispersion equation (ADE) in three-dimensional (3-D) form:

$$\frac{\partial C}{\partial t} = \nabla[D \cdot \nabla C] - \nabla[v C] \quad (1)$$

where ∇ is the gradient operator, C is the solute concentration, D is the dispersion coefficient tensor, v is the transport velocity vector, and t is time.

To understand the behavior of the movement of a contaminant in groundwater, one solves equation (1) forward in time where it is assumed that the plume originates from some location and will travel through the porous media due to advection and dispersion. The conventional procedure used to solve equation (1) is to use finite difference or finite element methods. For simple cases, closed-form solutions exist. Quantitative descriptions of the processes forward in time are well understood. Multidimensional models of these processes have been used successfully in practice (NRC, 1990). Numerical solute transport models were first developed about 25 years ago. When properly applied, these models can provide useful information about transport processes and can assist in the design of remedial programs.

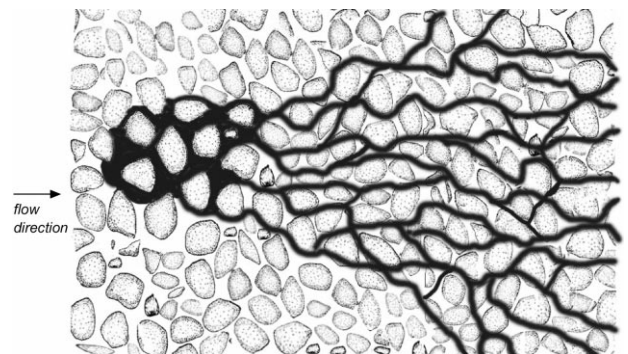


Figure 1. Contaminant transported in porous media.

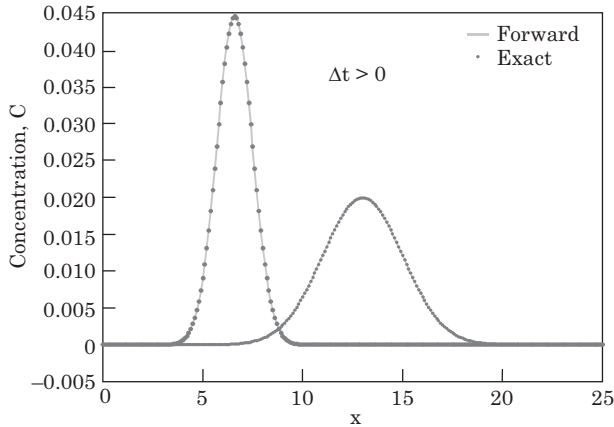


Figure 2. Finite difference solution of the ADE for forward simulation with $\Delta t > 0$.

Pollution source identification

One-dimensional (1-D) heterogeneous non-reactive contaminant transport in a semi-infinite domain, described by the ADE is:

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left[D(x) \frac{\partial C}{\partial x} \right] - \frac{\partial}{\partial x} [u(x)C] \quad (2)$$

$$C(x_1, t) = C_{in}(t) \quad 0 \leq t \leq T_{obs} \quad (3)$$

$$C(\infty, t) = 0 \quad 0 \leq t \leq T_{obs} \quad (4)$$

$$C(x, T_{obs}) = C_T(x) \quad 0 \leq x \leq \infty \quad (5)$$

where $u(x)$ is the transport velocity in the x -direction and x is distance. In the pollution source location identification, the source (x_1) information is not known, but measurements of the spatial distribution of the plume are given at time T_{obs} . For the release history reconstruction, usually the source location is assumed to be known, but the contaminant source function, $C_{in}(t)$, is unknown.

Finding the source location and the time history of the solute in groundwater can be categorized as a problem of time inversion. This means that we have to solve the governing equations backward in time. Modeling contaminant transport using reverse time is an ill-posed problem since the process, being dispersive is irreversible. Because of this, the solutions have discontinuous dependence on data and are sensitive to errors in the data.

A problem is categorized as a well-posed problem if (1) the solution exists; (2) the solution is unique; and (3) the solution is stable (Tikhonov and Arsenin, 1977). Problems that do not satisfy these criteria are called ill-posed. For the groundwater contamination problem, the plume has to have originated from someplace, therefore, physically, the plume exists. However, in rigorous mathematical terms, the fact that there exists a present day plume concentration, does not necessarily mean that we satisfy the existence criterion. The solution exists only when we have perfect and consistent model and data that satisfy extremely restrictive

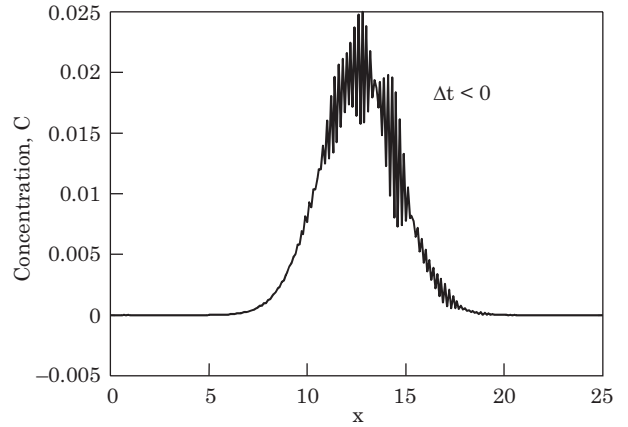


Figure 3. Finite difference solution of the ADE for backward simulation with $\Delta t < 0$.

conditions. Satisfying the stability criterion is a difficult task to accomplish since numerical schemes, which are usually implemented as a marching procedure are unstable for negative time steps, and make it impossible to solve contaminant transport problems backward in time. This can be seen from a numerical example of the propagation of a rectangular pulse of solute. Using a Crank-Nicolson finite difference scheme, the plume is propagated forward in time for 1600 time-steps, from time $t = 0.4$ to $t = 2$ with $\Delta x = 0.1$ and $\Delta t = 0.001$ and the numerical results are in good agreement with the exact solutions (Figure 2). On the other hand, starting at $t = 2$ and using the same scheme with $\Delta t = -0.001$ leads to instability after just 89 time-steps (Figure 3).

As for the non-uniqueness of the solution, there is no method that can bypass this inherent problem. In inverse problems, one of the common practices to overcome the stability and non-uniqueness criteria is to make assumptions about the nature of the unknown function so that the finite amount of data in observations is sufficient to determine that function. This can be achieved by converting the ill-posed problem to a properly posed one by stabilization or regularization methods. In the case of groundwater pollution source identification, additional information such as potential release sites and chemical fingerprints of the plume are usually available to make this possible.

Methods for Source Identification

Heat transport inversion methods

The physical and mathematical models of heat and mass transfer are similar. The governing equation for the heat conduction problem is similar to equation (1), without the advective or convective term on the right hand side of the equation. Research in the inverse heat transfer area has been going on for about three decades and numerous methods have been proposed. Therefore, it would be impossible to mention all of them in this paper. An extensive review has been presented by Beck, Blackwell and St Clair (1985) and Hensel (1991). Early methods included approaches using linear programming by Cannon (1966), who presented numerical inversion of the heat equation in his paper, Lattes and Lions (1969), who proposed the quasi-reversibility

(QR) method to solve heat conduction problems, and Ewing (1975), who used the Sobolev equations to approximate the backward solution of parabolic equations. The QR method has also been used for the solution of ill-posed heat conduction problems in a quasi-boundary-value-problem (QBVP) variant of the original formulation (Clark and Oppenheimer, 1994). The conjugate gradient method with minimization procedures was also used for the linear and non-linear inverse heat transfer (Beckman, 1960; Fletcher and Reeves, 1964; Lasdon, Mitter and Waren, 1967; Kammerer and Nashed, 1972). While Huang and Ozisik (1992) employed the combination of the regular and modified conjugate gradient methods to determine the unknown wall heat flux for laminar flow inside a parallel plate duct, Silva Neto and Ozisik (1993) used the conjugate gradient method and the adjoint equation simultaneously to solve for the timewise-varying strength of a two-plane heat source.

Different methods to eliminate the ill-posedness of the problem were developed, such as Tikhonov regularization (TR) and stabilization (Alifanov and Artyukhin, 1976; Tikhonov and Arsenin, 1977; Li and Zheng, 1997; Ames and Epperson, 1997). Ames *et al.* (1998) presented a comparison of regularization approaches for the backward heat equation problem. A finite element approach without stabilization procedures was introduced by Bayo *et al.* (1992). The computational procedure was made efficient by using the fast Fourier transform. Although Bayo and co-workers presented the application of the inverse heat conduction problem for the finite element method in their paper, the method can be used for finite differences or any other technique for spatial discretization. Marquardt and Auracher (1990) introduced the observer-based technique for solving the inverse heat conduction problem. Carasso (1992) constructed consistent space marching difference schemes to solve the nonlinear inverse heat conduction problem that blow-up more slowly than the counterpart analytical problem. Carasso examined various combinations of space and time differencing that lead to a total of 18 different algorithms.

Contaminant transport inversion methods

Optimization approaches. For the past 15 years, several attempts have been made to solve the ADE to identify the pollution source. One of the early methods to backtrack the pollution source location was to run forward simulations and check the solutions with the measured/current spatial data observed. Due to the non-uniqueness of the solution and the infinite number of plausible combinations, one needs to follow an optimization method to obtain the best-fitted solution. Among the first people to tackle pollution source identification problems using optimization approaches are Gorelick, Evans and Remson (1983). They formulated the groundwater pollution source identification problem as forward-time simulations in conjunction with an optimization model using linear programming and multiple regressions. In their approach, they incorporated the solute transport model as a series of constraints in the form of a concentration response matrix. Two examples were

used: (1) a steady state case, where the method is used to locate the unknown pollutant sources constructed from concentration data collected at few locations by means of a subsurface pipe that lies in the unsaturated zone with a selected leak location, and (2) a transient case, where the concentration history was used to identify the source schedule for a complex two-dimensional (2-D) system in which several wells are responsible for the observed pollution. In the steady state case example, because the data were sparse and there were more unknowns than constraining equations, an additional restriction to the linear programming method was needed. This was done by utilizing a mixed integer programming model such as in a stepwise multiple regression solution. While using the multiple regressions, the results displayed spurious negative values that detracted from the true solution. For the transient case, both methods identified the pollution source and the disposal episodes, but contained some errors in determining the disposal flux magnitude. In both models, Gorelick and co-workers assumed no uncertainty in the physical parameters of the aquifer. The method is restricted to cases where data are available in the form of breakthrough curves.

Another optimization approach was followed by Wagner (1992). Wagner developed a methodology for performing simultaneous model parameter estimation and source characterization, in which he used an inverse model as a non-linear maximum likelihood estimation problem. The hydrogeologic and source parameters were estimated based on hydraulic head and contaminant concentration measurements. Wagner assumed a steady confined groundwater flow and transient non-reactive single-species solute transport for the example problem. His method is capable of providing accurate and reliable estimates of the unknown model parameters, which are within 10% of the true parameter values. While adding the magnitude of disposal flux as unknown, the method became less accurate. In this case, the estimated parameters are within 30% of the true values and the estimated contamination flux is 20% greater than the true value. Wagner did go a step further, that is assuming an unknown contaminant release history. Assuming that the aquifer parameters, exact time span, and magnitude of disposal flux are unknown, the method gives fairly accurate results, however, the estimated standard deviations for the aquifer parameters increased by as much as 37%. The contaminant flux estimate remains the same as before. The method is also capable of quantifying the history of the disposal accurately. The last model used by Wagner was identifying the contaminant disposal source together with the parameters being estimated. For this example, it was assumed that there existed two potential disposal areas. The method estimated the parameters fairly accurately, but the standard deviations increased up to 55% compared to the case of unknown disposal history. As for the source characterization, the maximum likelihood estimated the true source location and the true disposal rate fairly accurately. The contaminant flux estimate at the true location was approximately 6% greater than the true disposal flux.

Along the same line of work, Mahar and Datta (1997; 2001) developed a methodology that combines

the concepts of optimal identification of a pollutant source with the optimal design of a groundwater quality monitoring network for an efficient identification process. They applied their method to a hypothetical 2-D homogeneous, isotropic, and saturated aquifer with a conservative pollutant plume. The first step in their methodology is to utilize an optimization model for preliminary identification of an unknown pollution source based on observation data. The next step is to simulate different realizations of pollution plume using perturbed sources. The solution of the simulation model for a given set of source fluxes represents a probable realization of the pollutant plume. Once the realizations are obtained, integer-programming is used to determine the optimal locations of the water quality monitoring wells. Finally, the concentration measurements made at the designed monitoring well locations are used in the non-linear optimization model to obtain a more accurate estimation of the sources. [Mahar and Datta \(2000\)](#) also used a nonlinear optimization technique to estimate the magnitude, location and duration of groundwater pollution sources under transient conditions.

Probabilistic and geostatistical simulation approaches. [Bagtzoglou \(1990\)](#), [Bagtzoglou, Tompson and Dougherty \(1991\)](#) and [Bagtzoglou, Dougherty and Tompson \(1992\)](#) are among the first to attempt solving the ADE backwards in time without relying on optimization approaches. In their work, they modeled the reversed time transport equation using the random walk particle method. In this approach, the advective part of the transport model is reversed while the dispersive part is left unchanged. They presented a probabilistic framework to identify solute sources in heterogeneous media. The example problem used for this study was a heterogeneous aquifer system with three potential sources of contamination. In all simulations the spill incidents are assumed to be instantaneous and occurring simultaneously. A particle representation of the concentration field (at the present time) was evaluated. Assuming that the hydraulic conductivity field is known with certainty, the velocity field was reversed. Repeated reversed time solute transport realizations were conducted and the first two moments of the concentration probability density function (pdf) were obtained. Using geostatistical techniques, [Bagtzoglou](#) and co-workers successfully assessed the relative importance of each potential source. Their approach was also extended to cases where the hydraulic conductivity field is not known with certainty with the help of conditional geostatistical simulations.

In the line of probabilistic approaches, [Wilson and Liu \(1994\)](#) solved the transport equation using stochastic differential equations backwards in time. Like in the method developed by [Bagtzoglou](#) and co-workers, [Wilson and Liu](#) also kept the dispersion part positive and reversed the advection part. They provided two types of maps, namely travel time probability and location probability. [Wilson and Liu](#) showed that both location and travel time probabilities could be calculated directly, using a backward-in-time version of traditional continuum advection-dispersion modeling. In addition, they claimed that by choosing the boundary conditions properly, the method can be

readily generalized to include linear adsorption with kinetic effects and 1st order decay. An extension of their study for a 2-D heterogeneous aquifer was reported in [Liu and Wilson \(1995\)](#). The results for travel time probability are in very close agreement with the simulation results from traditional forward-in-time methods. Results from this model were verified by [Neupauer and Wilson \(1999\)](#) using the adjoint method. In this method, the forward governing equation, with concentration as the dependent variable, is replaced by the adjoint equation, with the adjoint state as the dependent variable. They showed that backward-in-time location and travel time probabilities are adjoint states of the forward-in-time resident concentration. In this and the follow-up paper, [Neupauer and Wilson \(1999; 2001\)](#) presented the adjoint method as a formal framework for obtaining the backward-in-time probabilities for multidimensional problems and more complex domain geometries.

[Snodgrass and Kitanidis \(1997\)](#) also used a probabilistic approach combining Bayesian theory and geostatistical techniques. In their method, the source function to be estimated is discretized into components that are assigned a known stochastic structure with unknown stochastic parameters. The method incorporates uncertainty in contaminant concentration. [Snodgrass and Kitanidis'](#) method is an improvement from some other methods in that the solutions are more general and make no blind assumptions about the nature and structure of the unknown source function. Limitation to this approach is that the location of the potential source must be known *a priori*.

Analytical solution and regression approaches. [Sidauruk, Cheng and Ouazar \(1998\)](#) presented an inverse method based on analytical solutions of contaminant transport problems. Using their method, one only needs a well-planned set of concentrations, and no other prior information. The method provides a complete estimate of the dispersion coefficient, flow velocity, amount of pollutant, its initial location, and time origin. However, since this method is based on an analytical solution, it works only for a very limited number of cases. The method is applicable only to homogeneous aquifers with simple geometries and flow conditions. In their paper, [Sidauruk](#) and co-workers looked at two simple cases of instantaneous and continuous point releases in a 2-D uniform groundwater flow field. They took the exact solutions and obtained the linear relations between the logarithm of concentration and certain combination of parameters. The parameters were then obtained by minimizing the correlation coefficient of the linear regression.

Another inverse analytical technique was developed by [Ala and Domenico \(1992\)](#) to determine parameters such as the source strength and size, and the advective position of the contaminant front for the instantaneous contaminant plumes at Otis Air Force Base (AFB), Massachusetts. The contaminant plumes contained chloride, biodegradable and nonbiodegradable detergents, trichloroethene, and tetrachloroethene. To obtain the uniquely determined parameters, [Ala and Domenico](#) structured the analytical technique to solve several equations simultaneously. The results of their study indicated that the method works reasonably well

with a limited amount of variation in parameters, especially the dispersivities.

Butcher and Gauthier (1994) used inverse modeling to estimate the residual DNAPL mass. The flux was estimated from observed down-gradient concentrations. For the inverse model, Butcher and Gauthier used a tractable analytical approximation to the problem and developed additional simplifications to yield a form that is easily solved for the parameters of interest. From the simplified set of analytical equations, the parameters desired were solved using a least-squares estimator. The authors concluded that the methods presented in the paper can be used to provide an indication of the presence of the residual DNAPL, but not an accurate estimate of the residual volume. Improvement of the residual volume estimated can be effected using a longer period of observation.

In a different field of study, namely physics, Macdonald (1995) successfully applied the Nonlinear Least Squares (NLS) method to invert a 1-D pure diffusion problem. He applied the method to recover a number of Dirac-delta sources with large measurement errors in the data. Alapati and Kabala (2000) employed the NLS method without regularization to recover the release history of a groundwater contaminant plume from its current measured spatial distribution. In a number of synthetic numerical examples the solution was found to be very sensitive to noise and to the extent to which the plume is dissipated, especially for the gradual release scenario but not in the catastrophic release scenarios even in the presence of moderate measurement errors. Similar to the TR method, the noise level used was 5% and 20 and 50% for moderate and high levels, respectively.

Direct approaches. A different approach was proposed by Skaggs and Kabala (1994). They attempted to reconstruct the history of the plume using TR. Skaggs and Kabala studied a 1-D solute transport through a saturated homogeneous medium problem with a complex contaminant release history and assumed no prior knowledge of the release function. The authors pointed out that the accuracy of the regularized solution depends on finding a good value for the regularization parameter. The results showed that the solutions are insensitive to round off errors but their accuracy is affected by plume measurement errors. The authors observed that as long as the plume is not significantly dissipated, it is possible to adequately recover its history even when the data contains moderate random measurement errors. In their simulations, they used 5 and 20% noise for moderate and high levels, respectively. However, when the plume's original details are masked by dispersion, the presence of even moderate noise in the data significantly reduces the accuracy of the recovered history. This method was very effective when adequate data were available. In a similar effort, Samarskaia (1995) applied the TR with fast Fourier transforms to a groundwater contamination source reconstruction problem. Later, Liu and Ball (1999) used Skaggs and Kabala's modified TR technique to study a contaminant release at Dover AFB, Delaware. They used field measured concentration profiles in low-permeability porous media that underlie a contaminated aquifer at Dover, AFB. For

the two principal chemical contaminants, PCE and TCE, the Tikhonov method gave similar results as the measured data for PCE. In the case of TCE, however, the technique gave similar results only for the near-surface measurements. For deeper TCE concentrations the method showed anomalous behavior. Liu and Ball (1999) suggested that the interpretation of the estimated results requires careful consideration in the context of other available information.

Skaggs and Kabala (1995) employed the QR method for the same problem solved in their TR method. In the QR method, Skaggs and Kabala solved an equation that is close to the original equation and which is stable with a negative time step. The diffusion operator $\partial/\partial t - \nabla^2$ was replaced by $\partial/\partial t - \nabla^2 - \varepsilon \nabla^4$. A moving coordinate system was used to account for the velocity term of the ADE. A problem similar to the one presented in Skaggs and Kabala (1994) is used for the QR study. The results are less accurate than that of the TR approach, but it is computationally less expensive. The authors claimed that it is much easier to incorporate heterogeneous parameters in the QR method. However, until today, heterogeneous parameters have not been incorporated either in the QR method or in the Tikhonov method.

Another approach, namely the Minimum Relative Entropy (MRE) inversion, was used by Woodbury and Ulrych (1996). The MRE inversion is a method of statistical inference. Given prior information in terms of a lower and upper bound, a prior bias, and constraints in terms of measured data, MRE provides exact expressions for the posterior pdf and expected value of the inverse problem. The plume source is also characterized by a pdf. The problem solved in their study is the same as Skaggs and Kabala's problem. For the noise-free data, MRE was able to reconstruct the plume evolution history indistinguishable from the true history. As for data with noise, the MRE method managed to recover the salient features of the source history. Another advantage using the MRE approach is that once the plume source history is reconstructed, future behavior of the plume can be easily predicted due to the probabilistic framework of MRE. Woodbury *et al.* (1998) extended the MRE approach to reconstruct a 3-D plume source within a 1-D constant velocity field and constant dispersivity system.

Recently, Neupauer, Borchers and Wilson (2000) performed a study to compare the TR and the MRE methods. They found that both methods perform well in reconstructing a smooth source function. For an error free step function source history, the MRE performs better than the TR. On the other hand, the TR method is more robust in handling data that contain measurement errors. Neupauer and co-workers used measurement errors of 1%, 5% and 25%.

Skaggs and Kabala (1998) extended their study of TR using Monte Carlo simulation to answer the question of how far back one can use the Tikhonov procedure in recovering the release history of the plume, since the procedure always produces the recovered release curve that accurately reproduces the data. A table containing the percentage of test function recovery accuracy was produced. From the table, one can extract the information on how likely is that the recovered release history correctly resolves the details of

Table 1. Summary of methods

Specific method	Reference(s)	Classification	Limitation(s)
Linear programming	Cannon, 1966	Optimization	No advection, poor performance more than 50% back in time
Quasi-reversibility (QR & QBVP)	Lattes and Lions, 1969; Skaggs and Kabala, 1995; Clark and Oppenheimer, 1994;	1. Heat transport inversion 2. Direct	Homogeneous parameters, stabilization parameter not easy to predict <i>a priori</i>
Conjugate gradient method	Beckman, 1960; Fletcher and Reeves, 1964; Lasdon, Mitter and Waren, 1967; Kammerer and Nashed, 1972; Huang and Ozisik, 1992; Silva Neto and Ozisik, 1993	Optimization	No advection
Regularization (TR) and stabilization	Neupauer, Borchers and Wilson 2000; Alifanov and Artyukhin 1976; Tikhonov and Arsenin, 1977; Skaggs and Kabala, 1994; Liu and Ball, 1999; Samarskaia, 1995; Ames and Epperson, 1997; Ames <i>et al.</i> , 1998; Li and Zheng, 1997	1. Heat transport inversion 2. Direct	Homogeneous parameters
Space marching finite difference	Carasso, 1992	1. Heat transport inversion 2. Direct	No advection, homogeneous parameters
Linear programming and multiple regression	Gorelick, Evans and Remson, 1983	Optimization	Displayed spurious negative values
Non-linear maximum likelihood estimation	Wagner, 1992	Optimization	Accuracy decreases with increasing number of unknown parameters
Integer programming	Mahar and Datta, 1997; 2000; 2001;	Optimization	Homogeneous parameters
Reverse time random walk particle method	Bagtzoglou, 1990; Bagtzoglou <i>et al.</i> , 1991; 1992	Probabilistic and geostatistical simulation	Spill incidents are assumed to be instantaneous and occurring simultaneously, dispersion part kept positive
Stochastic differential equations backward in time	Wilson and Liu, 1994; Liu and Wilson, 1995; Neupauer and Wilson, 1999; 2001	Probabilistic and geostatistical simulation	Dispersion part kept positive
Bayesian theory and geostatistical techniques	Snodgrass and Kitaniidis, 1997	Probabilistic and geostatistical simulation	Potential source must be known <i>a priori</i>
Inverse method based on analytical solution	Sidauruk, Cheng and Ouazar, 1998; Ala and Domenico, 1992; Butcher and Gauthier, 1994	Analytical solution/regression	Simple geometries and flow conditions
Sobolev equations	Ewing, 1975	Direct	Perturbation parameter is not known <i>a priori</i>
Observer-based method	Marquardt and Auracher, 1990	Direct	Introduces phase-lag, sensitive to noise
Nonlinear least squares method	Macdonald, 1995; Alapati and Kabala; 2000	Analytical solution/regression	Homogeneous parameters, very sensitive to noise
Minimum relative entropy	Woodbury and Ulych, 1996; Woodbury <i>et al.</i> , 1998; Neupauer, Borchers and Wilson, 2000	Direct	Homogeneous parameters
Fourier series-based inverse technique	Birchwood, 1999	Direct	Results are highly sensitive to the accuracy of the inferred location
Backward beam equation	Carasso, 1972; 1975; Buzbee and Carasso, 1973; Carasso, Sanderson and Hyman, 1978; Elden, 1982	Direct	Excessive computational effort
Marching-jury backward beam equation	Atmadja and Bagtzoglou, 2000; 2001	Direct	Partial recovery, sensitive to noise

Table 2. Cases applied

Method classification	Cases applied
Heat transport inversion	1. Heat conduction problems, unknown wall heat flux for laminar flow inside a parallel plate duct, 2. Timewise-varying strength of a two-plane heat source
Optimization	1. Steady state and transient pollution source identification, 2. Hydrogeologic and source parameters identification, 3. Pollutant source identification and groundwater quality monitoring
Probabilistic and geostatistical simulation	1. Assessed the relative importance of each potential source, 2. Produced maps of time and location probability
Analytical solution and regression	Aquifer with very simple geometries and flow conditions
Direct	Spatial and temporal plume release history in homogeneous and heterogeneous media

the true release. Readers are referred to the original paper by Skaggs and Kabala (1998) for detailed results. This approach can also be applied to solution methodologies other than TR. In addition, other factors such as transport parameter and source location uncertainty can be included in the analysis to increase the accuracy of the recovered release history.

Birchwood (1999) used a Fourier-based inverse technique to recover the source location and release history of the groundwater contaminant plume from breakthrough curve data obtained from a single monitoring well. The Birchwood model is capable of finding the potential release sites by imposing a causality requirement upon the release history and carefully interpreting the multiple solutions that arise thereafter. In this method, the breakthrough curve is represented as a Fourier series. The coefficients of the Fourier series are expressed as functions of the release location and the Fourier coefficients of the release history. The release location is determined by imposing a condition of causality upon the release history. Recovery of the source location and profile of the unit pulse release history with homogeneous parameters was used to demonstrate the method. Birchwood concluded that preliminary results suggest that the method is capable of quantifying the location coordinate of a pollution source to a reasonable degree of accuracy, but the results are highly sensitive to the accuracy of the inferred source location.

Recently, Atmadja (2001) and Atmadja and Bagtzoglou (2000; 2001) proposed a method called the marching-jury backward beam equation (MJBBE) method. The backward beam equation (BBE) was first developed by Carasso (1972) to solve mixed parabolic initial boundary problems over long time intervals. The BBE method implementation for heat equation problems was originally proposed by Buzbee and Carasso (1973). Elden (1982) compared the BBE method with the Tikhonov-Phillips regularization method for parabolic equation problems. The BBE method has also been used in the past by Buzbee and Carasso (1973) to solve linear self-adjoint parabolic problems backwards in time. In addition, the method was also used by Carasso (1975) in obtaining the solutions to the final value problem of Burger's equation and Carasso, Sanderson and Hyman (1978) for image restoration. For large time intervals, that is as $t \rightarrow \infty$, one can assume that the solution reaches a known steady state value and one might have an approximate value of the solution at the forward

terminal time, $t = T_{ft}$. Unfortunately, for this type of problem, not all unconditionally stable numerical schemes, such as the implicit and Crank–Nicolson schemes, give correct solutions. Carasso suggested that with an extra data point at $t = T_{ft}$, it is feasible to consider the use of elliptic boundary value techniques as a solution. His interest in the BBE method stemmed from the fact that a solution of the heat equation, $f_t = f_{xx}$ also satisfies $f_{tt} = f_{xxxx}$. The method of backward beam equation was modified and enhanced to solve the 1-D ADE for homogeneous and heterogeneous problems within a contamination source identification context by Atmadja and Bagtzoglou (2000; 2001). They were able to cut down the computational time requirements by about three times by modifying the method to be a hybrid between the marching and jury methods for the examples they presented in their work. Most importantly, the MJBBE method is robust enough to handle heterogeneity. In addition, the method was able to preserve the shape and salient features of the initial data.

A recent paper by Morrison (2000a) presented an extensive literature review of commonly used environmental forensic methods for age dating and source identification. In this paper, Morrison discussed commonly used environmental forensic techniques and their possible applications so that readers can decide which technique is most appropriate for their cases. A more in-depth review of these techniques can be found in Morrison (2000b; c).

Conclusions

Inverse modeling for source identification purposes has been around for more than three decades. Its origin was based on the needs of solving heat conduction problems backwards in time. In identifying the contaminant source location and time release history, one needs to solve the governing equation backwards in time. Due to the similarity between the heat and mass transfer processes, numerous methods used in heat conduction problems are applicable to groundwater pollution source identification. Early methods to backtrack the pollution source location were to run forward simulations and check the solutions with the measured/current spatial data observed. Solving the ADE backwards in time without relying on optimization approaches was only started in the 1990s. In these methods, there exist three classes of approaches in identifying groundwater pollution sources. One class is

employing probabilistic techniques such as geostatistics to deduce the probability of the location of the sources. The second class is based on either analytical solutions or nonlinear regression approaches. The third class is using deterministic direct methods to solve the governing equations backward in time to reconstruct the release history of the contaminant plumes. This paper presented a review of the literature on mathematical methods used in pollution source identification. Table 1 summarizes and classifies the methods discussed herein, gives the pertinent references and finally presents the methods' limitations. Table 2 summarizes the classes of the methods discussed and presents the cases where these methods have been applied before. It is hoped that Tables 1 and 2 can provide the reader with a quick summary and guide to the applicability and usefulness of the methods presented in this paper.

References

- Ala, N.K. and Domenico, P.A. 1992. Inverse analytical techniques applied to coincident contaminant distributions at Otis Air Force Base, Massachusetts. *Ground Water* **32**, 212–218.
- Alapati, S. and Kabala, Z.J. 2000. Recovering the release history of a groundwater contaminant via the non-linear least-squares estimation. *Hydrol. Process.* **14**, 1003–1016.
- Alifanov, O. and Artyukhin, E. 1976. Regularized numerical solution of nonlinear inverse heat-conduction problem. *J. Eng. Phys.* **29**, 934–948.
- Ames, K.A. and Epperson, J.F. 1997. A kernel-based method for the approximate solution of backward parabolic problems. *SIAM J. Numer. Anal.* **34**, 1357–1390.
- Ames, K.A., Clark, G.W., Epperson, J.F. and Oppenheimer, S.F. 1998. A comparison of regularizations for an ill-posed problem. *Math. Comput.* **67**, 1451–1471.
- Atmadja, J. 2001. *The marching-jury backward beam equation method and its application to backtracking non-reactive plumes in groundwater*. Ph.D. Dissertation, Columbia University, 121.
- Atmadja, J. and Bagtzoglou, A.C. 2000. Groundwater pollution source identification using the backward beam equation method. In: *Computational Methods for Subsurface Flow and Transport*, 397–404. Rotterdam, The Netherlands, A.A. Balkema.
- Atmadja, J. and Bagtzoglou, A.C. 2001. Pollution Source Identification in Heterogeneous Porous Media. *Water Resour. Res.* **37**, 2113–2125.
- Bagtzoglou, A.C. 1990. *Particle-grid methods with application to reacting flows and reliable solute source identification*. Ph.D. Dissertation, University of California Irvine, 246.
- Bagtzoglou, A.C., Tompson, A.F.B. and Dougherty, D.E. 1991. Probabilistic simulation for reliable solute source identification in heterogeneous porous media. *Water Resources Engineering Risk Assessment, NATO ASI Series*, Vol. **G29**, 189–201.
- Bagtzoglou, A.C., Dougherty, D.E. and Tompson, A.F.B. 1992. Application of particle methods to reliable identification of groundwater pollution sources. *Water Res. Man.* **6**, 15–23.
- Bayo, E., Moulin, H., Crisalle, O. and Gimenez, G. 1992. Well-conditioned numerical approach for the solution of the inverse heat conduction problem. *Numer. Heat Tr.* **B, 21**, 79–98.
- Beck, J.V., Blackwell, B. and St Clair, C.R. Jr 1985. *Inverse Heat Conduction: Ill-Posed Problems*. New York, Wiley.
- Beckman, F.S. 1960. The solution of linear equations by conjugate gradient method. In: *Mathematical Methods for Digital Computers*, 62–72, Volume 1. (Ralston, A. and Wilf, H.S., Eds). New York, Wiley.
- Birchwood, R.A. 1999. Identifying the location and release characteristics of a groundwater pollution source using spectral analysis. *Proceedings of the 19th Annual American Geophysical Union Hydrology Days Conference*. Fort Collins, Colorado, Colorado State University, 37–50.
- Butcher, J.B. and Gauthier, T.D. 1994. Estimation of residual dense NAPL mass by inverse modeling. *Ground Water* **32**, 71–78.
- Buzbee, B.L. and Carasso, A. 1973. On the numerical computation of parabolic problems for preceding times. *Math. Comput.* **27**, 237–266. No. 122.
- Cannon, J.R. 1966. Some numerical results for the solution of the heat equation backwards in time. *Proceedings Adv. Symposium Numerical Solutions of Nonlinear Differential Equations*. New York, Madison, Wisconsin, Wiley, 21–54.
- Carraso, A. 1972. The backward beam equation: two A-stable schemes for parabolic problems. *SIAM J. Numer. Anal.* **9**, 406–434. No. 3.
- Carasso, A. 1975. The backward beam equation and the numerical computation of dissipative equations backwards in time. *MRC Technical Summary Report 1534*. University of Wisconsin-Madison.
- Carasso, A.S. 1992. Space marching difference schemes in the nonlinear inverse heat conduction problem. *Inverse Probl.* **8**, 25–43.
- Carasso, A.S., Sanderson, J.G. and Hyman, J.M. 1978. Digital removal of random media image degradations by solving the diffusion equation backwards in time. *SIAM J. Numer. Anal.* **15**, 344–367. No. 2.
- Clark, G.W. and Oppenheimer, S.F. 1994. Quasireversibility methods for non-well-posed problems. *Elec. J. Diff. Equations*, 1–9. No. 08.
- Dempster, H.S., Sherwood-Lollar, B. and Feenstra, S. 1997. Tracing organic contaminants in groundwater: a new methodology using compound-specific isotopic analysis. *Environ. Sci. Technol.* **31**, 3193–3197. No. 11.
- Elden, L. 1982. Time discretization in the backward solution of parabolic equations. II. *Math. Comput.* **39**, 69–84. No. 159.
- Engelking, P. 1999. Pollution. Microsoft® Encarta® Encyclopedia 99. ©. Microsoft Corporation, 1993–1998.
- Ewing, R.E. 1975. The approximation of certain parabolic equations backward in time by Sobolev equations. *SIAM J. Math. Anal.* **6**, 283–294. No. 2.
- Fletcher, R. and Reeves, C.M. 1964. Function minimization by conjugate gradients. *Brit. Comput. J.*, 149–154.
- Gorelick, S.M., Evans, B.E. and Remson, I. 1983. Identifying sources of groundwater pollution: an optimization approach. *Water Resour. Res.* **19**, 779–790. No. 3.
- Hensel, E. 1991. *Inverse Theory and Application for Engineers*. Englewood Cliffs, New Jersey, Prentice-Hall.
- Huang, C.H. and Ozisik, M.N. 1992. Inverse problem of determining unknown wall heat flux in laminar flow through a parallel plate duct. *Numerical Heat Transfer* **21**, 55–70. Part A.
- Kammerer, W.J. and Nashed, M.Z. 1972. On the convergence of the conjugate gradient method for singular linear operator equations. *SIAM J. Numer. Anal.* **9**, 165–181. No. 1.
- Lasdon, L.S., Mitter, S.K. and Waren, A.D. 1967. The conjugate gradient method for optimal control problem. *IEEE Trans. Automat. Control* **12**, 132–138.
- Lattes, R. and Lions, J.L. 1969. *The Method of Quasi-Reversibility, Applications to Partial Differential Equations*. New York, Elsevier.
- Li, Z. and Zheng, K. 1997. An inverse problem in a parabolic equation. *Elec. J. Diff. Equations*, Conference 01, 193–199.
- Liu, C. and Ball, W.P. 1999. Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware. *Water Resour. Res.* **35**, 1975–1985. No. 7.
- Liu, J. and Wilson, J.L. 1995. Modeling travel time and source location probabilities in two-dimensional heterogeneous aquifer. *Proc. 5th WERC Technology Development Conference*, Las Cruces, New Mexico, 59–76.
- Macdonald, J.R. 1995. Solution of an “impossible” diffusion-inversion problem. *Computers in Physics* **9**, 546–553.
- Mahar, P.S. and Datta, B. 1997. Optimal monitoring network and ground-water pollution source identification. *J. Water Res. Pl.* **123**, 199–207.
- Mahar, P.S. and Datta, B. 2000. Identification of pollution sources in transient groundwater systems. *Water Res. Man.* **14**, 209–227. No. 3.
- Mahar, P.S. and Datta, B. 2001. Optimal identification of groundwater pollution sources and parameter identification. *J. Water Res. Pl.* **127**, 20–29.
- Mansuy, L., Philp, R.P. and Allen, J. 1997. Source identification of oil spills based on the isotopic composition of individual components in weathered oil samples. *Environ. Sci. Technol.* **31**, 3417–3425. No. 12.
- Marquardt, W. and Auracher, H. 1990. An observer-based solution of inverse heat conduction problems. *Int. J. Heat Mass Transfer.* **33**, 1545–1562. No. 7.
- Morrison, R.D. 2000a. Application of forensic techniques for age dating and source identification in environmental litigation. *Environ. Forens.* **1**, 131–153.

- Morrison, R.D. 2000b. Critical review of environmental forensics techniques: part I. *Environ. Forens.* **1**, 157–173.
- Morrison, R.D. 2000c. Critical review of environmental forensics techniques: part II. *Environ. Forens.* **1**, 173–195.
- NRC (National Research Council). 1990. *Groundwater Models, Scientific and Regulatory Applications*. Washington, D.C., National Academy Press.
- Neupauer, R.M. and Wilson, J.L. 1999. Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. *Water Resour. Res.* **35**, 3389–3398. No. 11.
- Neupauer, R.M. and Wilson, J.L. 2001. Adjoint-derived location and travel time probabilities for a multi-dimensional groundwater system. *Water Resour. Res.* **37**, 1657–1668. No. 6.
- Neupauer, R.M., Borchers, B. and Wilson, J.L. 2000. Comparison of inverse methods for reconstructing the release history of a groundwater contamination source. *Water Resour. Res.* **36**, 2469–2475. No. 9.
- Rachdawong, P. and Christensen, E. 1997. Determination of PCB sources by a principal component method with nonnegative constraints. *Environ. Sci. Technol.* **31**, 2686–2691. No. 9.
- Samarskaia, E. 1995. Groundwater contamination modeling and inverse problems of source reconstruction. *SAMS* **18–19**, 143–147.
- Sidauruk, P., Cheng, A. and Ouazar, D. 1998. Ground water contaminant source and transport parameter identification by correlation coefficient optimization. *Ground Water* **36**, 208–214. No. 2.
- Silva Neto, A.J. and Ozisik, M.N. 1993. Inverse problem of simultaneously estimating the timewise-varying strength of two plane heat sources. *J. Appl. Phys.* **73**, 2132–2137. No. 5.
- Skaggs, T.H. and Kabala, Z.J. 1994. Recovering the release history of a groundwater contaminant. *Water Resour. Res.* **30**, 71–79. No. 1.
- Skaggs, T.H. and Kabala, Z.J. 1995. Recovering the history of a groundwater contaminant plume: Method of quasi-reversibility. *Water Resour. Res.* **31**, 2669–2673. No. 11.
- Skaggs, T.H. and Kabala, Z.J. 1998. Limitations in recovering the history of a groundwater contaminant plume. *J. Contam. Hydrol.* **33**, 347–359.
- Snodgrass, M.F. and Kitanidis, P.K. 1997. A geostatistical approach to contaminant source identification. *Water Resour. Res.* **33**, 537–546. No. 4.
- Stout, S.A., Uhler, A.D., Naymik, T.G. and McCarthy, K.J. 1998. Environmental forensics unraveling site liability. *Environ. Sci. Technol.* **32**, 260A–264A. No. 11.
- Tikhonov, A.N. and Arsenin, V.Y. 1977. *Solutions of Ill-Posed Problems*. Washington, DC, V.H. Winston & Sons.
- USEPA (U.S. Environmental Protection Agency). 1998a. *Report Brochure: National Water Quality Inventory: 1996 Report to Congress, Background Section*. Washington, DC. Office of Water, (28 March 2000) [online] <http://www.epa.gov/OW/resources/brochure/broch2.html>.
- USEPA (U.S. Environmental Protection Agency). 1998b. *National Water Quality Inventory: 1996 Report to Congress, Ground Water Chapters*. Washington, DC. Office of Water, Available: (28 March 2000) [online] <http://www.epa.gov/OW/resources/9698/chap6.html>.
- USEPA (U.S. Environmental Protection Agency). 1998c. This is *Superfund* (20 March 2000) <http://www.epa.gov/superfund/whatis/sfguide.htm>.
- USEPA (U.S. Environmental Protection Agency). 1999. *Safe Drinking Water Act, Section 1429. Groundwater Report to Congress*. Washington, DC. Office of Water, EPA-816-R-99-016. (28 March 2000) [online] <http://www.epa.gov/safewater/gwr/finalgw.pdf>.
- Wagner, B.J. 1992. Simultaneously parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modeling. *J. Hydrol.* **135**, 275–303.
- Wilson, J.L. and Liu, J. 1994. Backward tracking to find the source of pollution. *Waste Man. Risk Remed.* **1**, 181–199.
- Woodbury, A.D. and Ulrych, T.J. 1996. Minimum relative entropy inversion: theory and application to recovering the release history of groundwater contaminant. *Water Resour. Res.* **32**, 2671–2681. No. 9.
- Woodbury, A., Sudicky, E., Ulrych, T.J. and Ludwig, R. 1998. Three-dimensional plume source reconstruction using minimum relative entropy inversion. *J. Contam. Hydrol.* **32**, 131–158.