

# Inverse Green element evaluation of source strength and concentration in groundwater contaminant transport

Ednah Onyari and Akpofure Taigbenu

## ABSTRACT

Evaluating pollution sources from limited concentration data of their contaminant plumes in groundwater systems is an inverse problem that presents computational challenges because of its ill-posedness. In this work, the Green element method (GEM) is used to solve inverse contaminant transport problems in 2-D orthotropic homogeneous aquifers. The GEM discretization of the differential equation produces an over-determined, ill-conditioned global matrix that is decomposed by the singular value decomposition method and solved by the least square method with Tikhonov regularization. Five test cases of pollution sources of unknown strengths are used to evaluate the performance of GEM. The results indicate that the current methodology is capable of correctly predicting the strength of pollution sources and the historical concentration plumes that they produce. However, the computation accuracy is influenced by the location of the observation points in relation to the source, data errors of observed concentrations, the transport mode, and the value of the Courant number used in the simulations. Data errors at observation points influence more significantly the GEM prediction of the pollution source strength than the concentration plume, while numerical artefacts of dispersion and oscillations observed in direct solutions of advection-dominant transport cases are also evident in inverse modelling.

**Key words** | concentration plume recovery, Green element method, inverse contaminant transport, pollution source strength recovery, singular value decomposition, Tikhonov regularization

**Ednah Onyari**

**Akpofure Taigbenu** (corresponding author)  
School of Civil and Environmental Engineering,  
University of the Witwatersrand,  
P. Bag 3,  
Johannesburg WITS 2050,  
South Africa  
E-mail: akpofure.taigbenu@wits.ac.za

**Ednah Onyari**

Department of Civil & Chemical Engineering,  
University of South Africa,  
P. Bag 6,  
Florida 1710,  
South Africa

## INTRODUCTION

The quality of groundwater, which accounts for the water source and productive activities of a significant population of the world, can degrade due to the release of contaminants into aquifers from pollution sources. These sources include landfill sites, mine tailings, septic tanks, underground repositories of toxic and radioactive materials, unintentional spills of oils, hydrocarbons, and toxic substances on land surfaces, pesticides and fertilizers applied on agricultural lands, and seepages from polluted streams. Contamination may result from a single source or a combination of sources. In most instances, when contamination is detected, the locations and characteristics of the pollution sources are unknown.

Therefore, a paramount step in addressing groundwater contamination is precise identification of pollution sources, their strength, spatial and temporal distributions, and this in essence constitutes an inverse groundwater contaminant transport problem.

There are various classes of inverse groundwater contaminant transport problems that have been presented in the literature. These include identifying the contaminant source location (Gorelick *et al.* 1985; Wagner 1992; Mahar & Datta 2000), estimating the number of possible pollution sources (Ayvaz 2010), estimating the release history or source strength (Skaggs & Kabala 1994), and recovering

the historical spatial distribution of contaminant plumes (Michalak & Kitanidis 2004). In this work, we seek to simultaneously recover the signal of boundary pollution sources and the resulting spatial and temporal distributions of the plume.

Unlike in direct groundwater contaminant transport where the problem is well-posed with known boundary and initial conditions, inverse problems may seek to determine these input variables and their solutions may exhibit non-uniqueness, non-existence and instability (Sun 1996). Generally, for an inverse groundwater contaminant transport problem to be solvable additional information is provided by concentration measurements at some observation points.

Various numerical methods and their variants have previously been applied to inverse groundwater contaminant transport problems. Examples include the finite difference method (Mahar & Datta 2000; Neupauer & Wilson 2004; Ayvaz 2010; Prakash & Datta 2014), the finite element method (FEM) (Gorelick *et al.* 1983; Wagner 1992), the boundary element method (BEM) (Lesnic *et al.* 1998; Katsifarakis *et al.* 1999) and the adjoint method (Neupauer & Wilson 1999; Michalak & Kitanidis 2004). The methodology proposed in this study is based on the Green element method (GEM) which is founded on the singular integral theory of BEM, but implements the theory in an element-by-element manner as in FEM. By this approach, the resultant coefficient matrix is banded, the singular integral theory of BEM is amenable to solving nonlinear problems in heterogeneous domains, and point and distributed singularities are easier to handle because of their being aligned with the free-space Green's function which is in itself singular. GEM has successfully been previously applied to inverse modelling of heat conduction problems (Taigbenu 2015), and this is the first time it is being applied to inverse groundwater contaminant transport problems.

The GEM formulation followed in this work was recently developed (Taigbenu 2012), in which a second-order polynomial expression is used to approximate the internal normal fluxes so that only the solution for the concentration or the flux is calculated at external nodes and the concentration at internal nodes. Although this formulation is slightly less accurate than a flux-based one, its ability to generate fewer degrees of freedom per node than the latter

makes it more attractive for inverse modelling. The resultant coefficient matrix from applying this formulation to the inverse contaminant transport is ill-conditioned, and is decomposed by the singular value decomposition (SVD) method and solved by the least square method with the Tikhonov regularization (Tikhonov & Arsenin 1977).

The objectives of this study therefore are: (1) to demonstrate the use of the GEM and a regularization approach to recover a contaminant's release history and historical distribution from limited observation data; (2) to investigate the influence of the Peclet and Courant numbers on source identification; and (3) to evaluate the influence of the location of observation points and measurement errors on source identification.

## INVERSE DISPERSION-ADVECTION-REACTION EQUATION

Contaminant transport in groundwater systems involves processes of hydrodynamic dispersion which causes the spread of the contaminant, advection which enables the contaminant to be carried along by the ambient flow as well as a chemical reaction between the contaminant and the soil matrix. The mathematical statement, referred to as the dispersion-advection-reaction equation, which describes these processes under transient condition in a 2-D orthotropic homogeneous medium is (Bear 1972):

$$D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} = R \frac{\partial C}{\partial t} + \mathbf{V} \cdot \nabla C + \mu C \quad (1)$$

where  $C$  is the contaminant concentration [ $\text{ML}^{-3}$ ],  $\mathbf{V} = i\mathbf{u} + j\mathbf{v}$  is the velocity vector of the transporting medium [ $\text{LT}^{-1}$ ],  $D_x$  and  $D_y$  are the hydrodynamic dispersion coefficients in the spatial dimensions [ $\text{L}^2\text{T}^{-1}$ ],  $R$  is the retardation factor,  $\mu$  is the first-order decay rate coefficient that captures that reaction process [ $\text{T}^{-1}$ ],  $t$  is the time dimension [ $\text{T}$ ] and  $x$  and  $y$  are the Cartesian coordinates [ $\text{L}$ ]. In forward modelling where the concentration distribution in the domain is sought, Equation (1) applies to a domain  $\Omega$  over which the initial and boundary conditions are given as:

$$C(x, y, t = 0) = C_0(x, y) \text{ on } \Omega \quad (2)$$

$$C(x, y, t) = f_1(t) \text{ on } \Gamma_1 \quad (3)$$

$$D_x \frac{\partial C}{\partial x} n_x + D_y \frac{\partial C}{\partial y} n_y = Q(t) \text{ on } \Gamma_2 \quad (4)$$

where  $n_x$  and  $n_y$  are the direction cosines of the outward pointing normal  $\mathbf{n}$  to the surface  $\Gamma_2$  in the  $x$  and  $y$  directions. When the conditions in Equations (2) and (3) are specified with known medium and flow variables, the solution is that of the direct problem which is tractable and can be solved by conventional numerical methods. However, the inverse problem addressed in this paper has an additional boundary segment  $\Gamma_3$  where neither the concentration nor the flux is specified, and it is expected that the release history of the pollutant from this boundary be evaluated, making use of the concentration plume that it produces in the aquifer. In essence, the domain  $\Omega$  has a boundary  $\Gamma$  that comprises three parts,  $\Gamma_1$ ,  $\Gamma_2$  and  $\Gamma_3$ . To determine the release history of the unknown boundary information, additional data of the concentration are required at observation points  $\gamma_b = (x_b, y_b)$  within the aquifer, and these are

represented as:

$$C(\gamma_b) = C(x_b, y_b, t); \quad b = 1, 2, \dots, P \quad (5)$$

Essentially the inverse solution of Equation (1) is required in the domain  $\Omega$  with the boundary  $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$  (see Figure 1).

One of the necessary conditions for the solvability of the inverse problem is that the number of discrete equations generated by the numerical model should be at least equal to the number of unknowns, and that means that there should be an adequate number of observation points  $P$ . In many real life circumstances, these observed concentration measurements are tainted with measurement errors, and it is expected that the influence of these errors on the numerical solutions is assessed. The actual measured concentrations can be expressed as:

$$\tilde{C}(\gamma_b) = C(\gamma_b)[1 + \sigma \times \text{RAN}(\gamma_b)] \quad b = 1, 2, \dots, P \quad (6)$$

where  $\sigma$  is the error or noise level on the observed concentration  $C(\gamma_b)$  and  $\text{RAN}$  are random numbers that are generated from a standard normal distribution.

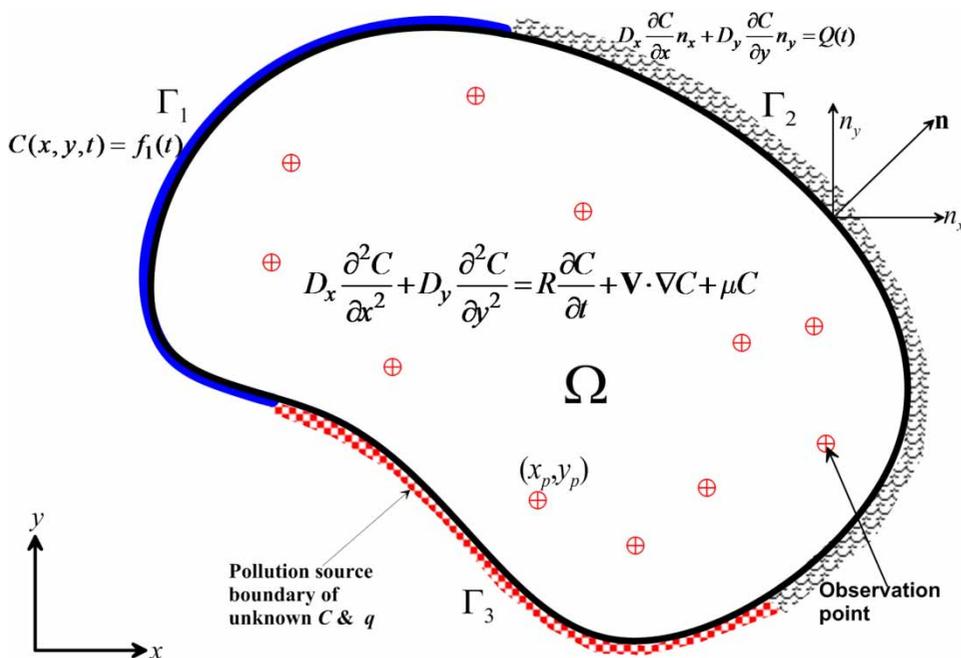


Figure 1 | Schematics of problem statement.

The governing partial differential Equation (1) is transformed into an integral one by applying Green’s second identity. The integral equation is given as (Brebbia *et al.* 1984):

$$\begin{aligned}
 & -\lambda C_{(r_i)} + \int_{\Gamma} C G^* ds - \int_{\Gamma} G q ds \\
 & + \iint_{\Omega} G \left[ \left( R \frac{\partial C}{\partial t} + \mu C + \mathbf{V} \cdot \nabla C \right) \right] dA \\
 & = 0
 \end{aligned} \tag{7}$$

where  $r_i = (x_i, y_i)$  is the source or collocation point,  $\lambda =$  nodal angle at  $r_i$ ,  $q$  is the normal contaminant flux which has an expression similar to Equation (4). Equation (7) consists of integrals along the boundary  $\Gamma$  with the integration variable  $s$  and integrals over the domain  $\Omega$  with integration variable  $A$ . The first three terms in Equation (7) represent the dispersion transport or the terms on the left-hand side of Equation (1), while the terms in the domain integral represent the temporal variation, advection and reaction transport or the right-hand side of Equation (1). The variable  $G$  in Equation (7) is the fundamental solution of:

$$D_x \frac{\partial^2 G}{\partial x^2} + D_y \frac{\partial^2 G}{\partial y^2} = \delta(r - r_i) \tag{8}$$

in the infinite space that has the expression:

$$G(r, r_i) = \frac{1}{2D} \ln \left[ \frac{(x - x_i)}{D_x} + \frac{(y - y_i)}{D_y} \right]; \quad D = \sqrt{D_x D_y} \tag{9}$$

and  $G^*$  is the normal derivative of  $G$  that is expressed as:

$$G^* = D_x \frac{\partial G}{\partial x} n_x + D_y \frac{\partial G}{\partial y} n_y \tag{10}$$

The discretized form of Equation (7) is achieved by discretizing the computational domain into polygonal elements (rectangular elements are used in this work) and interpolating the quantities  $C$ ,  $q$ , and  $\mathbf{V}$  by basis functions of the Lagrange family, that is  $C \approx N_j(r)C_j(t)$  (where  $N_j(r)$  are linear interpolating functions in space and  $j$  denotes the nodes of the element). The discrete element equation for

each element  $\Omega^e$  with boundary  $\Gamma^e$  is:

$$\begin{aligned}
 & S_{ij}C_j + L_{ij}q_j + W_{ij} \left( R \frac{dC_j}{dt} + \mu C_j \right) + U_{ikj}u_k C_j + Y_{ikj}v_k C_j \\
 & = 0
 \end{aligned} \tag{11}$$

in which the indices  $i, j$  and  $k$  in Equation (11) denote nodes of the element  $\Omega^e$ , while the elemental matrices are given as:

$$\begin{aligned}
 & S_{ij} = \int_{\Gamma^e} N_j G^*(r, r_i) ds - \delta_{ij}\lambda; \\
 & L_{ij} = \int_{\Gamma^e} N_j G(r, r_i) ds, \quad W_{ij} = \iint_{\Omega^e} N_j G(r, r_i) dA, \\
 & U_{ikj} = \iint_{\Omega^e} G(r, r_i) N_k \frac{\partial N_j}{\partial x} dA, \quad Y_{ikj} = \iint_{\Omega^e} G(r, r_i) N_k \frac{\partial N_j}{\partial y} dA
 \end{aligned} \tag{12}$$

The matrices in Equation (11) have been derived from the integral Equation (7) with  $S_{ij}$  representing the first two terms,  $L_{ij}$  the third term,  $W_{ij}$  the first two terms in the domain integrals, and  $U_{ijk}$  and  $Y_{ijk}$  account, respectively, for the advection in the  $x$  and  $y$  directions. The elemental integrations over rectangular elements in Equation (12) are done analytically and their expressions are presented for the first time in the Appendix (available with the online version of this paper). The discrete element equations are then aggregated over all elements used to discretize the computational domain, and this gives a matrix equation of the form:

$$E_{ij}C_j + L_{ij}q_j + W_{ij}R \frac{dC_j}{dt} = 0 \tag{13}$$

in which  $E_{ij} = S_{ij} + \mu W_{ij} + U_{ikj}u_k + Y_{ikj}v_k$ . At internal boundaries between the elements, the normal flux  $q$  is approximated by a second order polynomial function that is expressed in terms of the concentration as described in Taigbenu (2012). This results in two variables at the boundary nodes ( $C$  and  $q$ ) and only the concentration  $C$  at the interior nodes. The temporal derivative term is approximated by a finite difference expression:  $dC/dt \approx [C^{(2)} - C^{(1)}]/\Delta t$  that is evaluated at  $t = t_1 + \theta \Delta t$ , where  $0 \leq \theta \leq 1$ , is the difference weighting factor, and  $\Delta t$  is the time step between the current time  $t_2$  and the previous one  $t_1$ . Introducing the approximation for the temporal derivative into Equation (13) and

weighting the other terms in the equation by  $\theta$ , gives:

$$\begin{aligned} & \left( \theta E_{ij} + R \frac{W_{ij}}{\Delta t} \right) C_j^{(2)} + \theta L_{ij} q_j^{(2)} \\ & = \left( R \frac{W_{ij}}{\Delta t} - \beta E_{ij} \right) C_j^{(1)} - \beta L_{ij} q_j^{(1)} \end{aligned} \quad (14)$$

where  $\beta = 1 - \theta$  and the bracketed superscripts 1 and 2 represent the previous and current times, respectively. Incorporating the known initial, boundary data and concentration data at observation points into Equation (14) yields the matrix equation:

$$\mathbf{A}\mathbf{w} = \mathbf{b} \quad (15)$$

where

$$\mathbf{A} = \begin{bmatrix} \left( \theta E_{ij} + R \frac{W_{ij}}{\Delta t} \right) \\ \theta L_{ij} \end{bmatrix} \text{ and } \mathbf{w} = \begin{Bmatrix} C_j^{(2)} \\ q_j^{(2)} \end{Bmatrix} \quad (16)$$

in which  $\mathbf{A}$  is an  $M \times N$  matrix, with  $M$  being the number of nodes in the computational domain which is the same as the number of discrete equations generated and  $N$  is the number of unknowns which are  $C$  and/or  $q$  at external nodes, and  $C$  at internal nodes. The vector  $\mathbf{w}$  is  $N \times 1$  in dimension and represents the unknown quantities that should be calculated, and  $\mathbf{b}$  is an  $M \times 1$  vector of known quantities which consists of the terms on the right-hand side of Equation (14), as well as the contributions from the available concentration measurements that are given by Equation (5). For inverse problems, the matrix Equation (15) is over-determined and ill-conditioned. As an over-determined equation, a least square solution is sought, while  $\mathbf{A}$  is decomposed by the SVD technique (Heath 1997):

$$\mathbf{A} = \mathbf{HDP}^T = \sum_{i=1}^N \varphi_i h_i p_i^T \quad (17)$$

where the superscript  $T$  denotes the transpose,  $\mathbf{H}$  is an  $M \times M$  orthogonal matrix,  $\mathbf{P}$  is an  $N \times N$  orthogonal matrix, and  $\mathbf{D}$  is an  $M \times N$  diagonal matrix with  $N$  non-negative diagonal elements  $\varphi_1, \varphi_2, \dots, \varphi_N$  which satisfy the condition:  $\varphi_1 > \varphi_2 > \dots > \varphi_N > 0$ . The least square solution of Equation (15)

requires the minimization of the Euclidean norm  $\|\mathbf{A}\mathbf{w} - \mathbf{b}\|^2$  which gives:

$$\mathbf{w} = (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{b}) = \sum_{i=1}^N \frac{h_i^T \mathbf{b} p_i}{\varphi_i} \quad (18)$$

The small singular values  $\varphi_i$  cause instability in the solution of  $\mathbf{w}$ , and this can be ameliorated by using the zero order Tikhonov regularization technique which requires the minimization of the Euclidean norm  $\|\mathbf{A}\mathbf{w} - \mathbf{b}\|^2 + \alpha^2 \|\mathbf{w}\|^2$  that yields the solution for  $\mathbf{w}$ :

$$\mathbf{w}(\alpha) = \sum_{i=1}^N \frac{\varphi_i}{\alpha^2 + \varphi_i^2} h_i^T \mathbf{b} p_i \quad (19)$$

The factor  $\varphi_i / (\alpha^2 + \varphi_i^2)$  serves to dampen the contribution of the small singular values. The choice of the regularization parameter,  $\alpha$ , is crucial to obtaining meaningful solutions. Various methods have been proposed for selecting the regularization parameter and they include the L-curve (Hansen 1992), generalized cross validation (Golub *et al.* 1979) and the discrepancy principle (Goncharskii *et al.* 1972). In this work, we utilize the L-curve method. It is a continuous curve consisting of all the points  $(\|\mathbf{A}\mathbf{w}_\alpha - \mathbf{b}\|, \|\mathbf{w}_\alpha\|)$  for  $\alpha \in [0, \infty)$ . It shows the compromise between the perturbation error and the regularization error (over-smoothing). A suitable regularization is the one corresponding to a regularized solution near the corner of the L-curve, at which both errors are minimized. The values of  $\alpha$  used in the GEM simulations are obtained by a computer code which automatically identifies the corner of the L-curve.

## RESULTS AND DISCUSSION

In this section five numerical test cases are presented to demonstrate the accuracy of the GEM formulation that was described in the previous section. The first three test cases have analytical solutions that are used to benchmark the numerical results, while the last two cases mimic real life situations that utilize solutions obtained from direct GEM simulations. In the first test case, the

influences on the inverse solutions arising from the mode of contaminant transport, the Courant number and location of observation points relative to the pollution source are evaluated. The mean error is used to evaluate the accuracy of the numerical solutions for all the numerical test cases, and it is given by:

$$\epsilon = \frac{1}{M} \sqrt{\frac{\sum_{i=1}^M (C_i^{(e)} - C_i^{(n)})^2}{\sum_{i=1}^M C_i^{(e)2}}} \times 100 \quad (20)$$

where  $M$  is the number of unknowns of the variable  $C$  whose solution is sought and the superscripts  $e$  and  $n$  denote the exact and numerical solutions, respectively.

### Test case 1

This test case is the classical transient 1-D contaminant transport problem that arises as a result of advection and dispersion of a continuous pollution source  $C_s$  introduced along  $x=0$  for infinite time duration in a uniform velocity field ( $\mathbf{V} = iu$ ). The contaminant transport problem is in one spatial direction  $x$ , and the exact solution for the conservative transport case with  $\mu=0$  is well known (Ogata & Banks 1961):

$$C(x, t) = \frac{C_s}{2} \left\{ \operatorname{erfc} \left( \frac{x-ut}{2\sqrt{Dt}} \right) + \exp \left( \frac{ux}{D} \right) \operatorname{erfc} \left( \frac{x+ut}{2\sqrt{Dt}} \right) \right\} \quad (21)$$

Posed as an inverse problem, the strength of the pollution source  $C_s$  is unknown. Although this is a 1-D problem, it is implemented with the GEM formulation in a 2-D rectangular domain with  $D = D_x = D_y$ . The top and bottom boundaries of the domain are considered as no-flux boundaries, and the lateral extent in the  $x$ -direction for the right boundary,  $L$ , is set as a Dirichlet boundary with the concentration values obtained from Equation (21). To minimize the influence of the imposed no-flux condition on the top and bottom boundaries on the 1-D solution, the boundaries are kept as far apart from each other in the  $y$  direction by an arbitrary dimension of  $10^3$  km. Along the left boundary ( $x=0$ ) the source concentration and flux are not known. The concentration everywhere in the domain at the initial time  $t=0$  is zero.

The influence of the location of the observation point in relation to the pollution source at  $x=0$  is assessed with this test example using the parameters:  $u = 1.0$  km/yr,  $D = 0.025$  km<sup>2</sup>/yr,  $L = 1.0$  km,  $\Delta x = 0.025$  km and  $\Delta t = 0.0125$  yr. This corresponds to a marginal advection transport mode with local Peclet number  $Pe = u\Delta x/D = 1$  and a Courant number  $Cr = u\Delta t/\Delta x = 0.5$ . Forty uniform rectangular elements are used in the GEM simulation, and each simulation is run until  $t = 0.5$  yr with two observation points ( $\gamma_1 = \gamma_2$ ) that are chosen at nodes with the same  $x$ -coordinate and located at the top and bottom boundaries of the computational domain.

The mean error for the simulated release history of the pollution source in relation to the location of the observation point  $\gamma_1$  is presented in Figure 2, which, as expected, indicates better prediction of the pollution source strength with smaller distance between the observation point and the source. The value of  $\epsilon$  is less than 2% when  $\gamma_1$  does not exceed 0.15 km, but for larger distances of the observation point, GEM is unable to give any meaningful prediction of the source strength. Figure 3 shows the plots of the GEM predicted pollution source strength for  $\gamma_1 = 0.025, 0.1$  and  $0.15$  km against the exact  $C_s$  value of 1.0 g/L (gram/litre). To provide indicative values of the regularization parameter  $\alpha$  at each time step when  $\gamma_1 = 0.025$  km, these range between  $1.6 \times 10^{-4}$  to  $4.0 \times 10^{-5}$

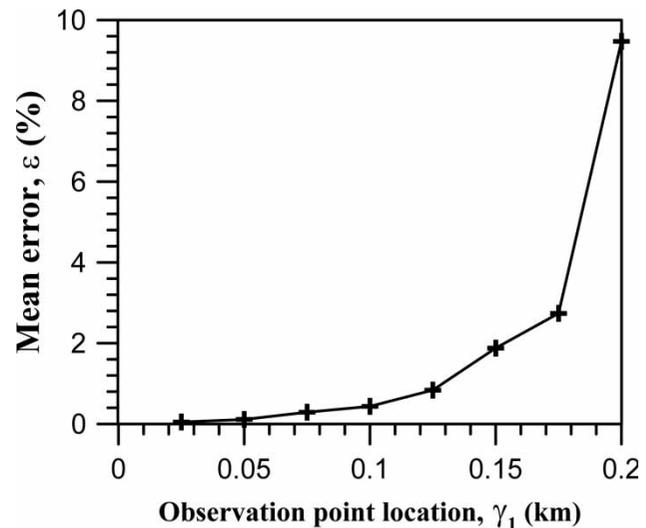
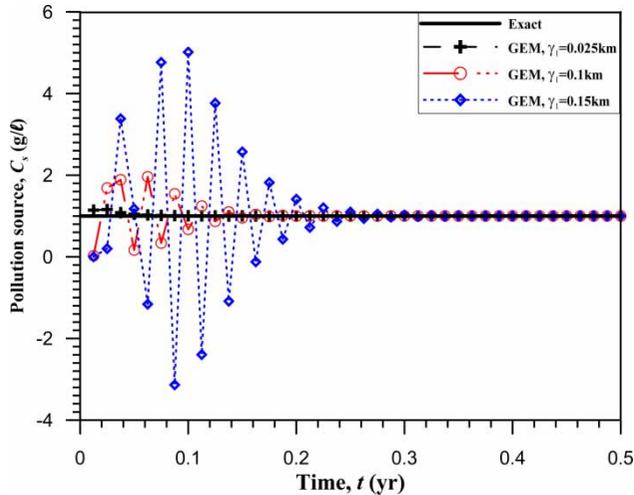


Figure 2 | Mean error of estimated source strength in relation to observation point location for Test case 1.



**Figure 3** | GEM prediction of the pollution source strength for various locations of observation points.

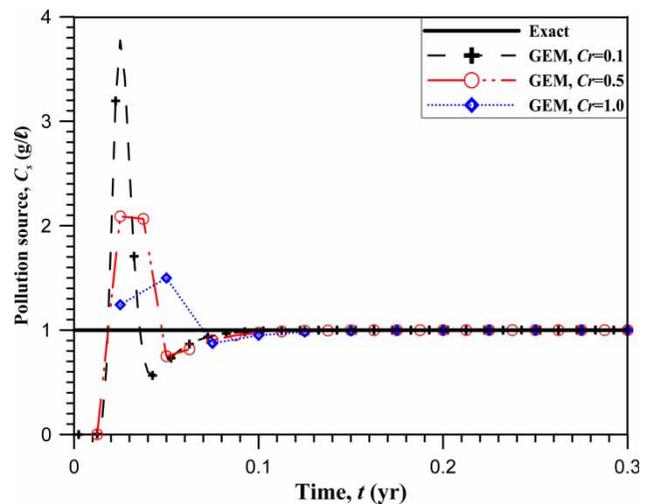
with an average value of  $6.5 \times 10^{-5}$ . This suggests a narrow band for the variation of the regularization parameter. The results in Figure 3 confirm our earlier finding of enhanced solution accuracy of the pollution source strength with closeness of the observation points to it. This is expected and has been confirmed by previous numerical investigations. The integrity of inverse pollution source solutions relies on the quality of the available information at observation points which depend on their number, spatial distribution, and positioning in relation to the flow field (Mahar & Datta 1997; Sun *et al.* 2006; Li & Mao 2011; Jha & Datta 2013). The CPU time for this simulation on a single CPU desktop 64-bit Intel® Core™ i7-2600, 3.40 GHz, 8 GB memory is 0.26 seconds which represents a CPU time of  $1.63 \times 10^{-4}$  seconds/element/time step.

The influence of two transport modes on the source identification is assessed with this test case. The first transport mode (i) represents an advection dominant transport with  $Pe = 50$ , and the second example (ii) is one of marginal advection transport with  $Pe = 1$ . The simulation parameters for the two transport cases are given in Table 1, while the fully implicit scheme with  $\theta = 1$  is used in all the simulations. The influence of the time step on the numerical solutions is assessed for three values of the Courant number,  $Cr = u\Delta t/\Delta x$ , namely  $Cr = 0.1$ ,  $0.5$  and  $1.0$ . Additional concentration data are provided at two observation points  $\gamma_1$  and  $\gamma_2$  that are located along  $x = 0.025$  km for both transport modes.

**Table 1** | Simulation parameters for Test case 1

Transport case	$Pe$	$D$ ( $\text{km}^2/\text{yr}$ )	$u$ ( $\text{km}/\text{yr}$ )	$\Delta x$ ( $\text{km}$ )	$L$ ( $\text{km}$ )
Advection dominant (i)	50.0	0.0005	1.0	0.025	1.0
Marginal advection (ii)	1.0	0.025	1.0	0.025	1.0

With the exact value of the pollution source strength  $C_s = 1.0$  g/L for all times, its computed value for the three values of the Courant number for the advection dominant case (i) is presented in Figure 4. From Figure 4, it is observed that there is correct prediction of the pollution source strength by GEM for  $t \geq 0.08$  yr. At early times the numerical solutions oscillate about the exact value, with the simulation with  $Cr = 1.0$  exhibiting the least amount of oscillation compared to those with  $Cr = 0.1$  and  $0.5$ . It should be pointed out that unlike in direct numerical modelling with GEM of contaminant transport where the value of the Courant number  $Cr$  influences the numerical stability (Taigbenu 1999), its value in inverse modelling plays a dual role of influencing the numerical stability and determining the frequency at which data are available to the numerical scheme from the observation points. With respect to the latter role, because the transport process with  $Pe = 50$  is predominantly advection driven, the GEM simulation with  $Cr = 0.1$  takes about ten time steps before information of



**Figure 4** | Exact and inverse GEM solutions for the pollution source strength  $C_s$  for  $Pe = 50$ .

the concentration from the pollution source along  $x=0$  becomes available at the two observation points along  $x=0.025$  km, hence the wild oscillations of the numerical results at the early simulation times, but at later times information of the concentration is frequently available for the numerical calculations as the plume passes these points. Conversely, the GEM simulation with  $Cr=1.0$  takes only one time step before concentration information from the pollution source is available at the two observation points along  $x=0.025$  km, hence the more subdued oscillations at early times, but information of the concentration is less frequently available for the numerical calculations at later times. Thus, the GEM simulation with  $Cr=1.0$  gives better prediction of  $C_s$  at early times than that with  $Cr=0.1$ , and conversely at later times. However, it should be pointed out that it is difficult to ascertain whether these two influences of  $Cr$  on the overall GEM solution are conflicting or complementing, and which influence is more significant for any contaminant transport mode. The GEM solutions for all  $Cr$  values correctly predict the concentration strength  $C_s$  at later times, and this indicates that GEM converges to the exact solution as the numerical calculations proceed with time. Another reason for the poor numerical solutions at early times is the discontinuity in the concentration along  $x=0$  which is zero at  $t=0$  and  $C_s$  for  $t>0$ .

The exact and numerical solutions for the concentration fronts at various times are presented in Figure 5, and it

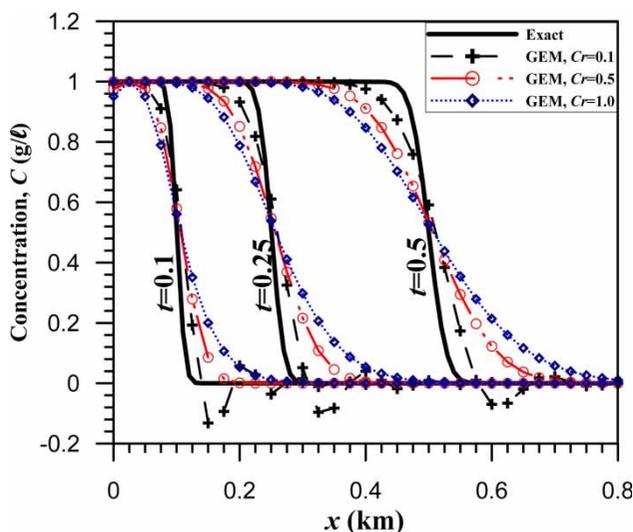


Figure 5 | Concentration fronts for  $Pe=50$  of Test case 1.

shows that GEM solutions for the contaminant front are more dispersed with  $Cr=1.0$  and  $Cr=0.5$  than with  $Cr=0.1$ , while the latter has trailing oscillations which are not observed for  $Cr=0.5$  and  $Cr=1.0$ . The mean error  $\epsilon$  for the computed concentration for  $Cr=0.1$ ,  $Cr=0.5$  and  $Cr=1.0$  is 0.48%, 0.52% and 0.53%, respectively.

For the transport case (ii) of marginal advection ( $Pe=1$ ), the GEM solutions for the pollution source strength  $C_s$  and its exact value are presented in Figure 6. The GEM simulation with  $Cr=0.1$  correctly predicts  $C_s$  for times  $t \geq 0.05$  yr, while that with  $Cr=0.5$  does so correctly for times  $t \geq 0.075$  yr, and that with  $Cr=1.0$  correctly predicts  $C_s$  for times  $t \geq 0.1$  yr. Because of the greater dispersion of the contaminant that occurs in this case than the previous, a wider spread of the plume takes place, indicating that concentration information at the observation points along  $x=0.025$  km is more readily available at early times than with the earlier case. Furthermore, we do not observe the same level of oscillations of the numerical solutions for  $C_s$  at early times as in the advection-dominant case (i), and this is most probably due to the better numerical stability characteristics observed with the direct modelling of this transport case (Taigbenu 1999). Figure 7 shows the concentration fronts at times  $t=0.1$ , 0.25 and 0.5 yr for Courant numbers values of  $Cr=0.1$ , 0.5 and 1.0. The GEM solution of the concentration fronts with  $Cr=0.1$  is excellent and better than those with

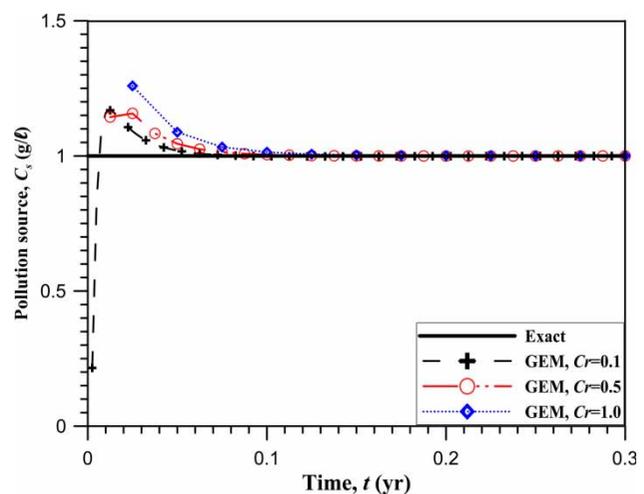


Figure 6 | Exact and inverse GEM solutions for the pollution source strength  $C_s$  for  $Pe=1$ .

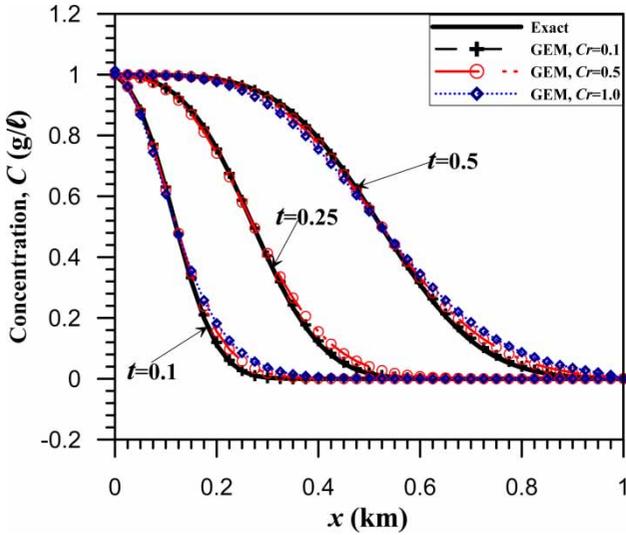


Figure 7 | Concentration fronts for  $Pe = 1$  of Test case 1.

$Cr = 0.5$  and  $Cr = 1.0$ , and this is confirmed by the values of the mean error:  $\varepsilon = 0.04\%$  for  $Cr = 0.1$ ,  $\varepsilon = 0.10\%$  for  $Cr = 0.5$  and  $\varepsilon = 0.17\%$  for  $Cr = 1.0$ .

The influence of observation errors on the pollution source recovery is investigated for the marginal advection transport case ( $Pe = 1$ ) with noise levels of  $\sigma = 2\%$  and  $5\%$ . With  $Cr = 1$  used in the GEM simulation, the results for the pollution source strength are presented in Figure 8. It is observed that GEM predicts the pollution source strength  $C_s$  reasonably well up to noise level of  $2\%$  but the solution oscillates more about the exact for noise level of  $5\%$ .

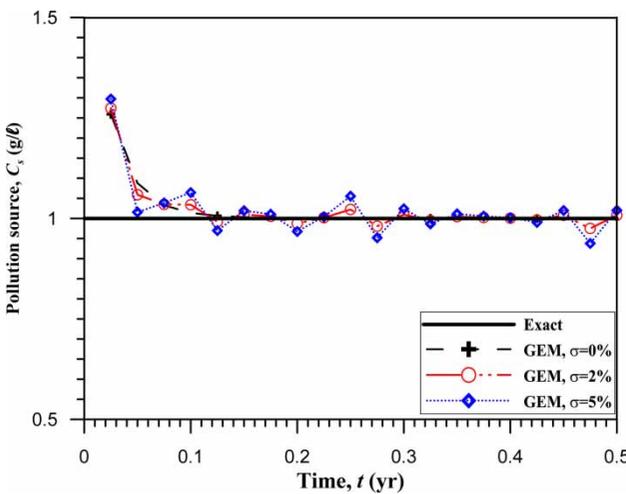


Figure 8 | Source strength for  $Pe = 1$  and  $Cr = 1.0$  of Test case 1 for various noise levels in the observation data.

### Test case 2

This test case is a 1-D contaminant transport problem with an exponentially decaying source in a homogeneous domain. It is governed by Equation (1) in the  $x$ -dimension with  $R = 1$ ,  $\mu = 0$  and  $\mathbf{V} = iu$ . The velocity flow field is uniform, and the initial and boundary conditions for the direct problem, which had been solved by Mariño (1974), are given as:

$$C(x, 0) = 0; \quad x \geq 0 \tag{22}$$

$$C(0, t) = C_s e^{\gamma t}; \quad t > 0 \tag{23}$$

$$C(L, t) = 0; \quad t \geq 0 \tag{24}$$

The analytical solution is given as:

$$C(x, t) = \frac{C_s e^{\gamma t}}{2} \left\{ \exp\left[\frac{x(u - \xi)}{2D}\right] \operatorname{erfc}\left(\frac{x - \xi t}{2\sqrt{Dt}}\right) + \exp\left[\frac{x(u + \xi)}{2D}\right] \operatorname{erfc}\left(\frac{x + \xi t}{2\sqrt{Dt}}\right) \right\} \tag{25}$$

in which  $\gamma$  is a constant, and  $\xi = (u^2 + 4D\gamma)^{1/2}$ .

The inverse modelling of this example is carried out in a 2-D rectangular domain, and requires the calculation of the source concentration along  $x = 0$ , while the right boundary is specified as a Dirichlet boundary and the top and bottom boundaries are no-flux boundaries. A value of  $\gamma = -1.0 \text{ yr}^{-1}$  is used in the exact solution. The parameters used for this example are:  $D = 6.0 \text{ km}^2/\text{yr}$ ,  $u = 6.0 \text{ km}/\text{yr}$  and  $\Delta x = 1.0 \text{ km}$  which correspond to the case of marginal advection transport with  $Pe = 1$ . There are two observation points where concentration measurements are available and these are on the top and bottom boundaries along  $x = 1.0 \text{ km}$ , and a uniform time step  $\Delta t = 3.33 \times 10^{-2} \text{ yr}$ , corresponding to a value of  $0.2$  for the Courant number. The range of values of the regularization parameter  $\alpha$  at each time step when there is no noise in observation data is between  $5.5 \times 10^{-6}$  and  $5.7 \times 10^{-6}$  with an average value of  $1.9 \times 10^{-5}$ , indicating small variations in regularization parameter over the simulation period. Figure 9 shows the exact and computed source strength, and the results

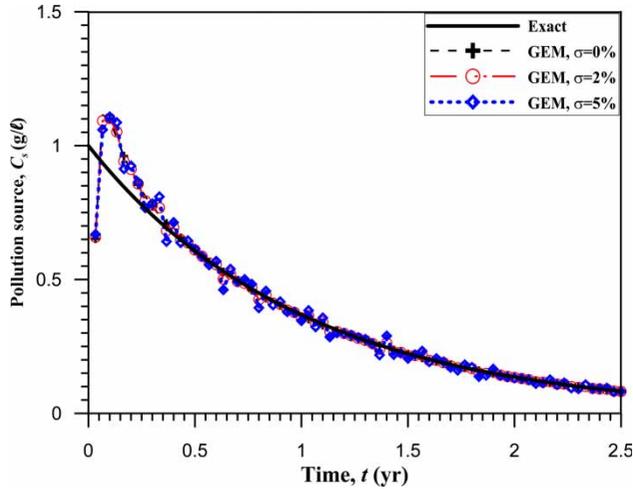


Figure 9 | Exact and computed source strength for Test case 2.

indicate that GEM overestimates the source strength for the first ten time steps and gives correct estimates after  $t \geq 0.4$  yr. The mean error  $\varepsilon$  for the computed source strength with and without errors randomly introduced into the observed data is 0.1566% for  $\sigma = 0\%$ , 0.1568% for  $\sigma = 2\%$  and 0.1660% for  $\sigma = 5\%$ . For the same noise levels, their influence on the concentration distribution is more muted, with  $\varepsilon = 0.0508\%$  for  $\sigma = 0\%$ ,  $\varepsilon = 0.0518\%$  for  $\sigma = 2\%$  and  $\varepsilon = 0.0563\%$  for  $\sigma = 5\%$ . Figure 10 shows a comparison of the GEM predicted concentration profiles at times of 0.5 yr, 1.0 yr, 1.5 yrs and 2.0 yrs and their exact

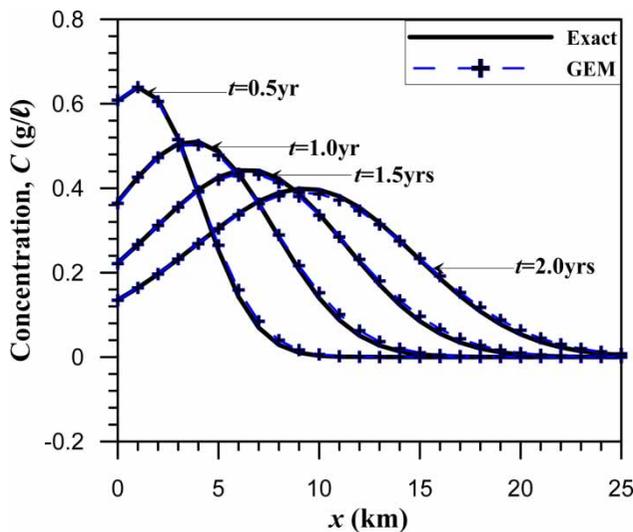


Figure 10 | Concentration profiles at various times for Test case 2.

profiles. The results indicate that GEM correctly predicts the exact profiles.

### Test case 3

This test case represents a steady contaminant transport in a 2-D unit square domain that occurs in a uniform velocity field in the  $x$ -direction that is governed by Equation (1) with  $R = 0$ ,  $\mu = 0$  and  $\mathbf{V} = i\mathbf{u}$ . For the direct problem, solved by Qiu *et al.* (1998), all the boundaries have Dirichlet boundary conditions, with the left boundary along  $x = 0$  having a sinusoidal concentration distribution  $C_s = \sin(\pi y)$  while all other boundaries have a zero concentration value. The analytical solution to this problem is given as:

$$C(x, y) = C_s \left[ \frac{e^{m_1(1-x)} - e^{m_2(1-x)}}{e^{m_1} - e^{m_2}} \right] \quad (26)$$

where

$$\begin{aligned} m_1 &= \frac{1}{2}(\sqrt{u^2 + 4\pi^2} - u) \text{ and } m_2 \\ &= -\frac{1}{2}(\sqrt{u^2 + 4\pi^2} + u) \end{aligned} \quad (27)$$

Posed as an inverse problem, this problem is simulated considering the boundary along  $x = 0$  as a  $\Gamma_3$  boundary where neither the concentration nor the flux is known, and boundary along  $x = 1.0$  km as a Neumann boundary. On the top and bottom boundaries,  $C = 0$ . The [1 km  $\times$  1 km] domain is discretized uniformly in the  $x$  and  $y$  directions with  $\Delta x = 0.1$  km and  $\Delta y = 0.1$  km, the dispersion coefficient for the isotropic medium,  $D = 1.0 \text{ km}^2/\text{yr}$ , and a uniform velocity  $u = 10.0 \text{ km}/\text{yr}$ , thus corresponding to the transport case with local Peclet number  $Pe = 1$ . Using this 2-D example, the influence of the location of observation points is examined as with the first 1-D test case. The examined locations of observation points are nine internal points in the  $y$ -direction along  $x_p = 0.1, 0.2, 0.3, 0.4$  and  $0.5$  km ( $p = 1, 2, \dots, 9$ ). The plot of the mean error  $\varepsilon$  in predicting the source strength  $C_s$  in relation to the location of the observation points is presented in Figure 11. The results exhibit a

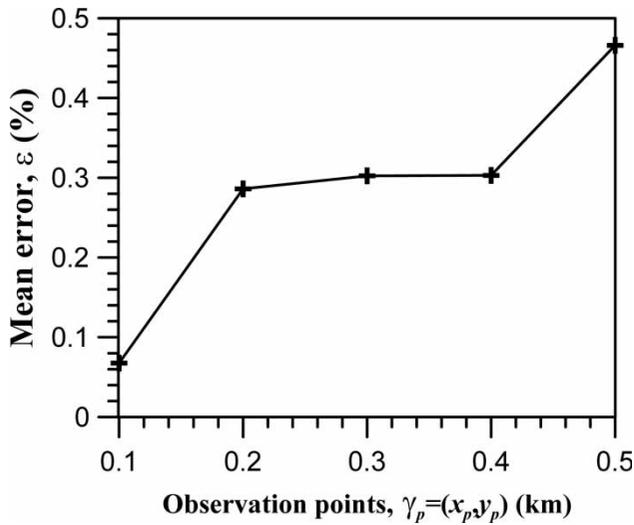


Figure 11 | Mean error of estimated source strength in relation to location of observation points for Test case 3.

trend similar to that observed for the first 1-D test case which confirms that closeness of the observation points to the source of the pollution signal enhances the prediction capability of its strength by GEM.

Using the nine observation measurements at the internal points along  $x = 0.1$  km, the GEM inverse solution for the pollution source signal, obtained with a regularization parameter value of  $4.6 \times 10^{-4}$ , is presented in Figure 12 along with the exact solution for noise levels of 0, 2 and 5%. We observe that the recovered pollution source is correctly

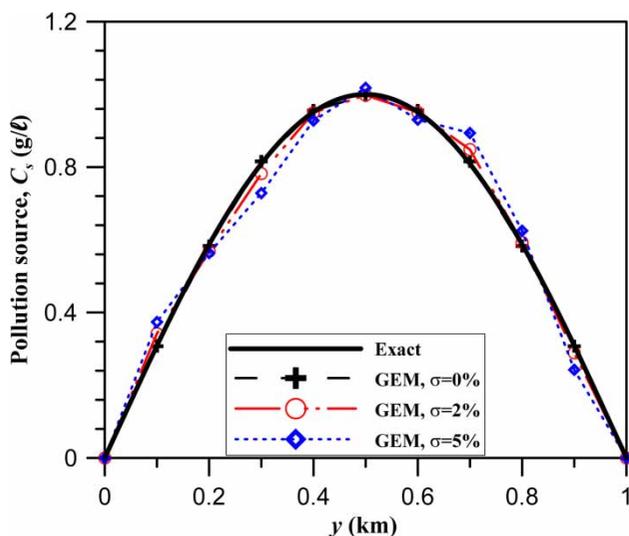


Figure 12 | Exact and GEM calculated source strength along  $x = 0$  for Test case 3.

predicted by GEM for noise-free data and with noise level of 2%, but for noise level of 5% the numerical solution oscillates about the exact value. Figure 13 shows the concentration plume obtained by the exact solution and that by GEM with  $\sigma = 0\%$ , and both solutions are in good agreement. The mean error  $\epsilon$  for the concentration distribution obtained by GEM is  $1.51 \times 10^{-2}\%$  with  $\sigma = 0\%$ ,  $2.40 \times 10^{-2}\%$  with noise  $\sigma = 2\%$  and  $4.91 \times 10^{-2}\%$  with noise  $\sigma = 5\%$ , while for the source strength prediction, the mean error values are  $6.76 \times 10^{-2}\%$  for  $\sigma = 0\%$ ,  $3.17 \times 10^{-1}\%$  for  $\sigma = 2\%$  and  $7.94 \times 10^{-1}\%$  for  $\sigma = 5\%$ , indicating that the influence of noise on the concentration distribution is less significant compared to predicting the pollution source strength.

The steady nature of this test case makes it amenable to assess the influence of the regularization parameter  $\alpha$  on the numerical solutions for the spatial concentration distribution  $C(x,y)$  and the source strength  $C_s$  when there is no noise in the observation data. The variations of the mean error  $\epsilon$  for  $C(x,y)$  and  $C_s$  in relation to the regularization parameter are presented in Figure 14, and the results indicate that minimum value of  $\epsilon$  for the concentration is attained when  $\alpha = 4.1 \times 10^{-4}$  and for the source strength when

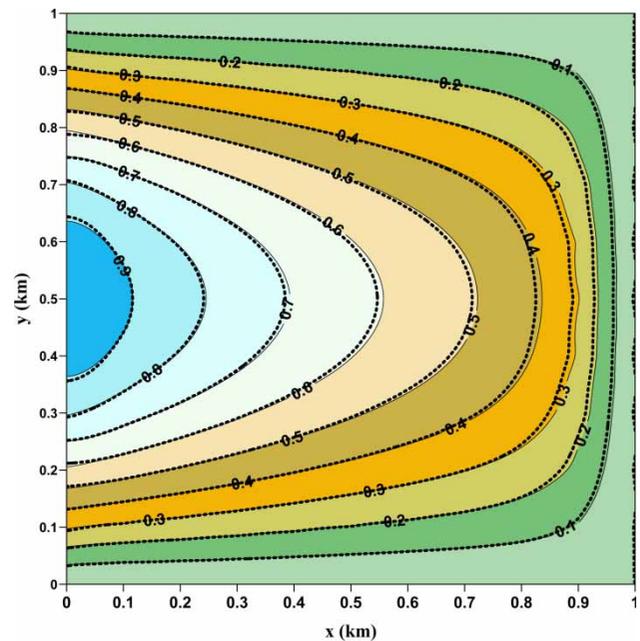


Figure 13 | Exact and GEM concentration (g/L) contours of Test case 3 – solid line is exact and dotted is GEM with no noise in the data.

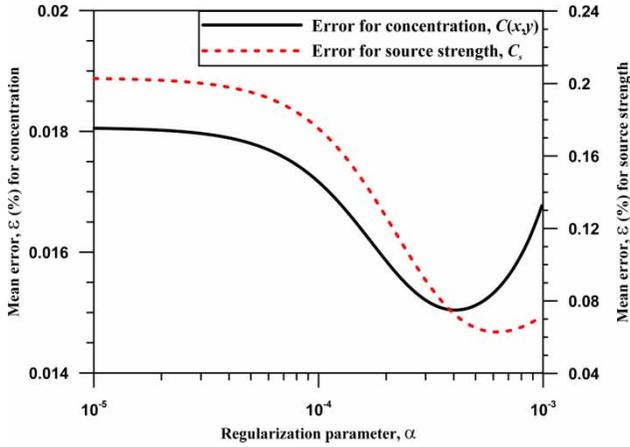


Figure 14 | Error plots of solutions for Test case 3 against regularization parameter.

$\alpha = 6.2 \times 10^{-4}$ . The L-curve gave an optimum value of  $4.6 \times 10^{-4}$  which lies between those two values.

**Test case 4**

This is a 2-D contaminant transport case, which is similar to that addressed by Huang *et al.* (2008). The flow is uniform in the  $x$ -direction with a velocity  $u = 6$  m/day and contaminant transport takes place in an orthotropic, homogeneous medium with thickness  $H = 25$  m and length  $L = 1,000$  m. The contaminant source is located at the top boundary

within the section  $L_2 - L_1$ , as shown in Figure 15. When the space and time dependent contaminant flux release function  $q(x,t)$  is known, the solution procedure follows the direct approach. For the inverse model  $q(x,t)$  is not known and additional information on the concentration is provided at the observation points shown in Figure 15. The problem is governed by Equation (1) with the initial condition:

$$C(x, y, t = 0) = 0, \quad 0 < x < L, \quad 0 < y < H \tag{28}$$

$$C(0, y, t) = c_0, \quad 0 < y < H, \quad t > 0 \tag{29}$$

$$\frac{\partial C(1, y, t)}{\partial x} = 0, \quad 0 < y < H, \quad t > 0 \tag{30}$$

$$D_y \frac{\partial C(x, 0, t)}{\partial y} = 0, \quad 0 < x < L, \quad t > 0 \tag{31}$$

$$D_y \frac{\partial C(x, 1, t)}{\partial y} = \begin{cases} q(x, t) & L_1 \leq x \leq L_2 \\ 0 & \text{elsewhere} \end{cases} \tag{32}$$

The parameters that are used in the simulation are:  $L_1 = 200$  m,  $L_2 = 600$  m,  $c_0 = 0.2$  g/L,  $D_x = 2 \times 10^5$  m<sup>2</sup>/day,  $D_y = 10^4$  m<sup>2</sup>/day,  $\mu = 2.0$  day<sup>-1</sup>,  $R = 1.0$ ,  $\Delta x = 25$  m,

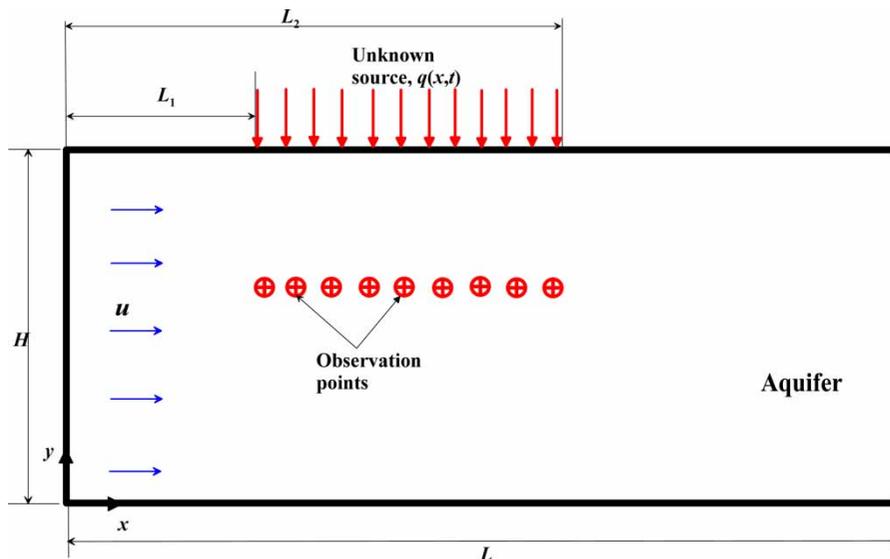


Figure 15 | Schematic representation of Test case 4.

$\Delta y = 1.25$  m and  $\Delta t = 0.025$  day. The computational domain is discretized into 800 rectangular uniform elements. The expression for the release history of the contaminant source is given by:

$$q(x, t) = 8 \times 10^4 [7 + 0.02x \times \sin(4\pi t)] g/L^{-1} \text{day}^{-1} \quad (33)$$

It should be noted that the release history is unrestricted in time in contrast to that used by Huang *et al.* (2008) and Li & Mao (2011) who both used global space-time optimization techniques in their inverse modelling of this example which places a limitation on the time dimension to enable the size of the matrix that is solved to be manageable. GEM inverse solution to this example does not suffer from this limitation in time, as it allows for the numerical scheme to march forward to any time level. The direct solution procedure uses the expression for  $q(x, t)$  that is given by Equation (33) to compute the concentrations at the observation points ( $P = 17$ ) that are located at a depth of 2.5 m from the top of the aquifer over a simulation period of 1 day. The inverse model then considers these concentration values at the observation points as the actual ones to calculate the release history of the pollution source.

The inverse solutions by GEM for the release history of pollution flux  $q(x, t)$  and the concentration distribution  $C(x, y, t)$  are presented as error plots that are based on the mean error  $\varepsilon$  that is given by Equation (20). The error plots for the pollution flux when  $\sigma = 0\%$ ,  $\sigma = 2\%$  and  $\sigma = 5\%$  are presented in Figure 16, while the error plots for the concentration distribution are presented in Figure 17. Comparing Figures 16 and 17, the errors for the spatial concentration distribution are about three orders of magnitude lower than those from predicting the pollution source flux. Furthermore, introduction of randomized noise into observation data tends to produce a more muted influence on the inverse solution for the concentration distribution than for the pollution flux (Taigbenu 2015). The CPU time for running each of the simulations with a given noise level is 13.61 seconds, representing a CPU time of  $4.3 \times 10^{-4}$  seconds/element/time step. The values of the regularization parameter that were used in the simulation of the noise free data range from  $1.03 \times 10^{-4}$  to  $1.23 \times 10^{-4}$  with an average value of  $1.06 \times 10^{-4}$ .

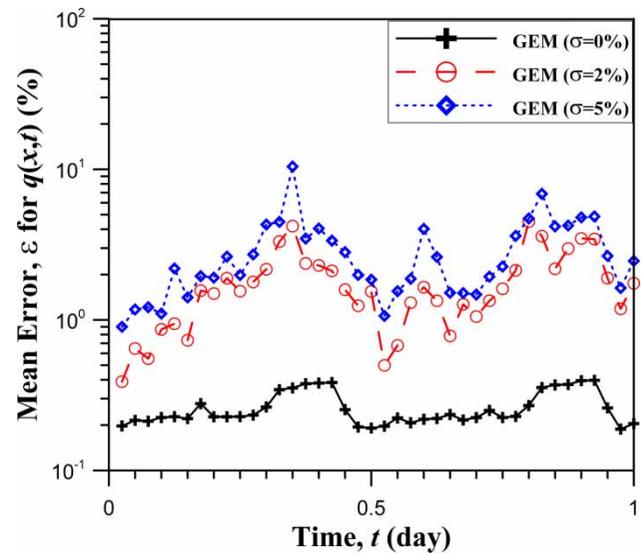


Figure 16 | Error plots of the inverse GEM solutions for the pollution source of Test case 4 for various noise levels of observation data.

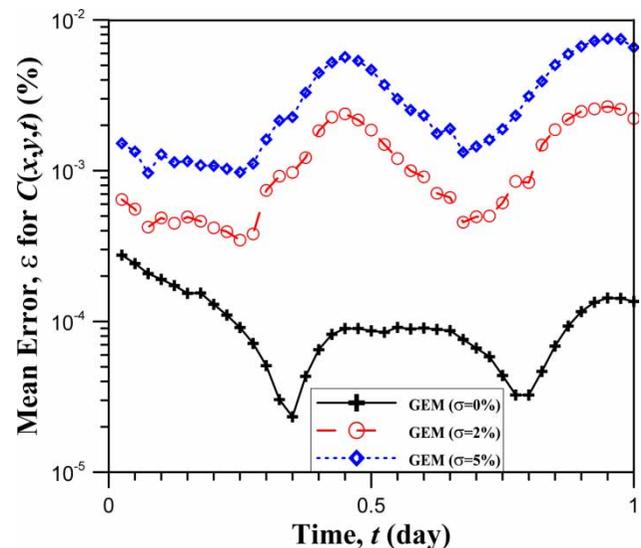


Figure 17 | Error plots of the inverse GEM solutions for the spatial concentration distribution of Test case 4 for various noise levels of observation data.

### Test case 5

Another test case that is simulated is similar to the previous one but represents a contaminant transport situation that is commonly encountered in real life where pollutant source is active for a while and then becomes inactive for a variety of reasons that may include, stoppage of the pollution

source after being identified or legal prosecution of the polluter which halts his activities or decommissioning of the factory or industry that is the source of the pollution. Over a time frame which extends beyond the active period of the pollutant source, the contaminant plume will continue to evolve within the aquifer. The solution to the inverse problem addresses how correctly GEM can predict the strength and active period of the pollution source. The domain is similar to the previous one with the following parameters used for the simulations:  $L_1 = 200$  m,  $L_2 = 600$  m,  $c_0 = 0.35$  g/L,  $D_x = 2 \times 10^5$  m<sup>2</sup>/day,  $D_y = 4 \times 10^4$  m<sup>2</sup>/day,  $\mu = 2.0$  day<sup>-1</sup>,  $R = 1.0$ ,  $\Delta x = 25$  m,  $\Delta y = 2.5$  m and  $\Delta t = 0.025$  day. The true flux of the pollution source occurs over an active period of 1 day, and it is given by the expression:

$$q(x, t) = \begin{cases} 1600x \times e^{-2t} & 0 < t \leq 1 \text{ day} \\ 0 & t > 1 \text{ day} \end{cases} \text{ g/L}^{-1} \text{ day}^{-1} \quad (34)$$

This expression for the flux is used in the direct formulation to generate the concentrations over a period of 2 days at the 17 observation points located at a depth of 5 m from the top where the pollution source is located. These concentration values at the observation points are considered as the actual values. With these observed concentration values, the inverse GEM simulation is carried on the domain with the top boundary between  $x = 200$  m and  $x = 600$  m considered as a  $\Gamma_3$  boundary where the concentration and the flux are not specified. The computational domain was discretized into 800 uniform rectangular elements for both the direct and inverse GEM simulations. The inverse GEM solution for the release history of the pollution flux  $q(x, t)$  is presented in Figure 18. The results indicate that GEM correctly predicts the pollution source flux evolution, as well as the active period of 1 day over which the pollution occurs. The error plots of the inverse GEM solutions for the release history of pollution flux  $q(x, t)$  and the concentration distribution  $C(x, y, t)$  are presented in Figures 19 and 20, respectively, for noise levels of  $\sigma = 0\%$ ,  $\sigma = 2\%$  and  $\sigma = 5\%$ . It should be noted that the error plots for the pollution flux are presented only over the active period of the pollutant  $t \leq 1$  day. As with Test case 4, the errors for the spatial concentration distribution are about three orders of magnitude lower than those from predicting the pollution source flux, and introduction of randomized noise in the observation data produces a

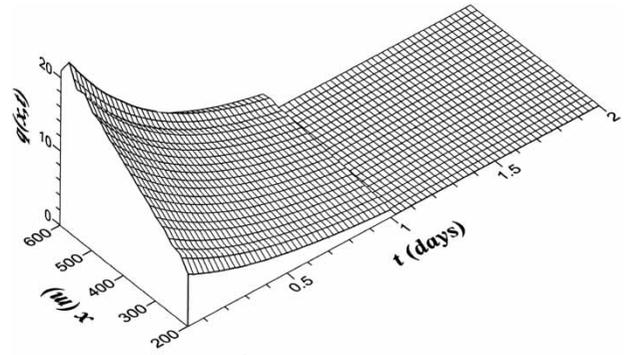


Figure 18 | Release history of pollution source  $q(x, t)$  g/L<sup>-1</sup> day<sup>-1</sup> m<sup>-1</sup> of Test case 5 by GEM with  $\sigma = 0\%$ .

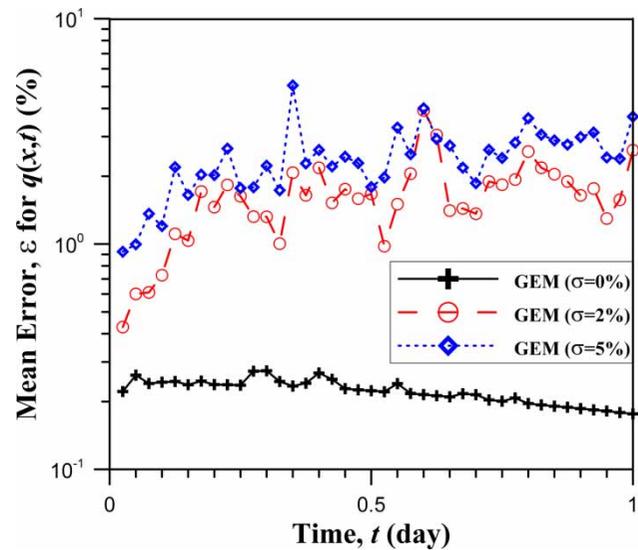
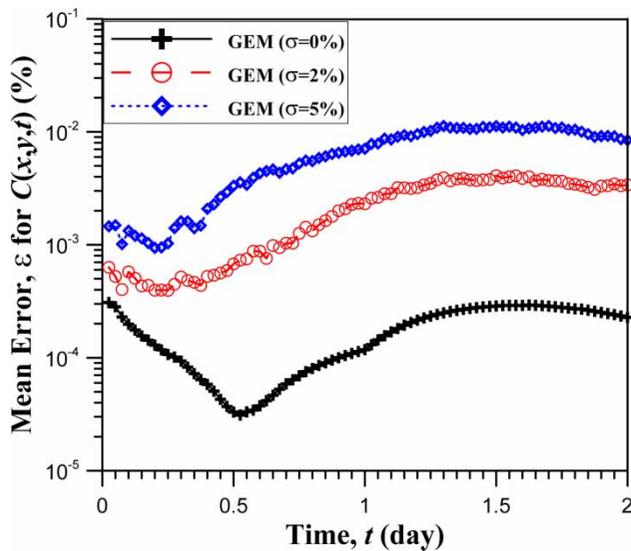


Figure 19 | Error plots of the inverse GEM solutions for the pollution flux  $q(x, t)$  of Test case 5 for various noise levels of observation data.

more muted influence on the inverse solution for the concentration distribution than for the pollution flux. The CPU time for the simulation run with no noise in the data is 24.36 seconds, and this represents a CPU time of  $3.8 \times 10^{-4}$  seconds/element/time step. The values of the regularization parameter for the noise free data range from  $1.03 \times 10^{-4}$  to  $1.16 \times 10^{-4}$  with an average value of  $1.05 \times 10^{-4}$ .

## CONCLUSION

This paper has presented the Green element solutions of the inverse dispersion–advection contaminant transport



**Figure 20** | Error plots of the inverse GEM solutions for the spatial concentration distribution of Test case 5 for various noise levels of observation data.

equation in 2-D orthotropic homogeneous aquifers. The GEM discretization of the differential equation gives rise to an over-determined, ill-conditioned global matrix that is solved by the least square method with Tikhonov regularization, and aided by the SVD of the matrix. Using five examples, the influences on the GEM simulation results are evaluated for various transport modes, values of the Courant number, location of observation points in relation to the pollution source, and errors in observation data. We found the following:

- (1) The prediction capability of the pollution source strength by GEM is enhanced by having observation points located in close proximity to the source. This confirms the results of previous numerical investigations and what should be expected intuitively, that the integrity of pollution source recovery depends on that of measured concentrations at observation points which in part is dependent on how close those points are to the source.
- (2) The GEM prediction of the source strength is generally better with more dispersion dominant transport than with the advection dominant one; for transient transport, oscillations are observed in the numerical solution at early simulation times but they die out at later times. The challenge in the GEM modelling of the advection-dominant transport problem lies with the inability of the

steep concentration profiles associated with such problems to be correctly approximated by the linear interpolation that is used in the current formulation. The use of higher interpolation functions like quadratic and Hermitian functions provides better approximation of the steep concentration fronts and, hence, improved solutions (Taigbenu 1998).

- (3) The two-fold influence of the Courant number  $Cr$  on the GEM solutions, namely numerical stability and frequency of available sampling data, makes it difficult to determine which influence is dominant in any particular transport case. However, small values of  $Cr$  in the GEM simulations generally produce oscillatory solutions at early simulation times when concentration information is not available at observation points, but give accurate solutions at later times when they are available frequently at those points. The oscillations are more pronounced in the GEM solutions for advection-dominant transport than for dispersion-dominant transport.
- (4) For the simulated noise levels in the data at observation points, their influence on the recovery of the source strength is significant and, as expected, the influence on the inverse solutions increases with increased noise level in the data.
- (5) Errors in observation measurements have greater influence on the recovery of the pollution source strength than on the historical concentration distribution.
- (6) GEM is capable of identifying the active period of pollution sources as well as their strengths when they are active.

The simulations of all the test cases were carried out on a single CPU desktop 64-bit Intel® Core™ i7-2600, 3.40 GHz, 8 GB memory computer, and the simulation times with respect to spatial discretization element and time step ranged from  $1.63 \times 10^{-4}$  seconds for the simple 1-D example to  $4.3 \times 10^{-4}$  seconds for the more challenging 2-D example. These small simulation times are indicative that the proposed methodology is computationally efficient.

## REFERENCES

- Ayvaz, M. T. 2010 [A linked simulation–optimization model for solving the unknown groundwater pollution source](#)

- identification problems. *Journal of Contaminant Hydrology* **117**, 46–59.
- Bear, J. 1972 *Dynamics of Fluids in Porous Media*. Elsevier Science, New York, USA.
- Brebbia, C. A., Telles, J. C. F. & Wrobel, L. C. 1984 *Boundary Element Techniques*. Springer-Verlag, Berlin, Heidelberg, New York, USA.
- Golub, G. H., Heath, M. & Wahba, G. 1979 Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* **21** (2), 215–223.
- Goncharskii, A. V., Leonov, A. S. & Yagola, A. G. 1972 A generalized discrepancy principle. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* **2** (13), 294–302.
- Gorelick, S. M., Evans, B. & Remson, I. 1983 Identifying sources of groundwater pollution: an optimisation approach. *Water Resources Research* **19** (3), 779–790.
- Hansen, P. C. 1992 Analysis of discrete ill-posed problems by means of the L curve. *SIAM Review* **34** (4), 561–580.
- Heath, M. T. 1997 *Scientific Computing: An Introductory Survey*. McGraw-Hill, New York, USA.
- Huang, C.-H., Li, J.-X. & Kim, S. 2008 An inverse problem in estimating the strength of contaminant source for groundwater systems. *Applied Mathematical Modelling* **32**, 417–431.
- Jha, M. & Datta, B. 2013 Three-dimensional groundwater contamination source identification using adaptive simulated annealing. *Journal of Hydrologic Engineering* **18**, 307–317.
- Katsifarakis, K. L., Karpouzou, D. K. & Theodosiou, N. 1999 Combined use of BEM and genetic algorithms in groundwater flow and mass transport problems. *Engineering Analysis with Boundary Element* **23**, 555–565.
- Lesnic, D., Elliot, L. & Ingham, D. B. 1998 A boundary element method for the determination of the transmissivity of a heterogeneous aquifer in groundwater flow systems. *Engineering Analysis with Boundary Elements* **21**, 223–234.
- Li, Z. & Mao, X. 2011 Global multiquadric collocation method for groundwater contaminant source identification. *Environmental Modelling & Software* **26**, 1611–1621.
- Mahar, P. S. & Datta, B. 1997 Optimal monitoring network and ground-water pollution source identification. *Water Resources Planning and Management* **123**, 199–207.
- Mahar, P. S. & Datta, B. 2000 Identification of pollution sources in transient groundwater system. *Water Resources Management* **14**, 209–227.
- Mariño, M. A. 1974 Longitudinal dispersion in saturated porous media. *Journal of the Hydraulics Division, ASCE* **100**, 151–157.
- Michalak, A. M. & Kitanidis, P. K. 2004 Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. *Water Resources Research* **40**, W08302.
- Neupauer, R. M. & Wilson, J. L. 1999 Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contamination. *Water Resources Research* **35** (11), 3389–3398.
- Neupauer, R. M. & Wilson, J. L. 2004 Numerical implementation of a backward probabilistic model of groundwater contamination. *Ground Water* **42** (2), 175–189.
- Ogata, A. & Banks, R. B. 1961 A solution of the differential equation of longitudinal dispersion in porous media. *U.S Geological Survey Professional Paper* **411-A**, A1–A9.
- Prakash, O. & Datta, B. 2014 Characterization of groundwater pollution sources with unknown release time history. *Journal of Water Resource and Protection* **6**, 337–350.
- Qiu, Z. H., Wrobel, L. C. & Power, H. 1998 Numerical solution of the convection-diffusion problems at high Peclet number using Boundary elements. *International Journal for Numerical Methods in Engineering* **41**, 899–914.
- Skaggs, T. H. & Kabala, Z. J. 1994 Recovering the release history of a groundwater contaminant. *Water Resources Research* **30**, 71–79.
- Sun, N.-Z. 1996 *Mathematical Modelling of Groundwater Pollution*. Springer-Verlag, New York, USA.
- Sun, A. Y., Painter, S. L. & Wittmeyer, G. W. 2006 A constrained robust least squares approach for contaminant release history identification. *Water Resources Research* **42** (4), W04414.
- Taigbenu, A. E. 1998 Hermitian Green element calculations of contaminant transport. *Water South Africa* **24**, 303–307.
- Taigbenu, A. E. 1999 Three Green Element Models for the diffusion-advection equation and their stability characteristics. *Engineering Analysis Boundary Elements* **23**, 577–589.
- Taigbenu, A. E. 2012 Enhancement of the accuracy of the Green element method: application to potential problems. *Engineering Analysis Boundary Elements* **36**, 125–136.
- Taigbenu, A. E. 2015 Inverse solutions of temperature, heat flux and heat source by the Green element method. *Applied Mathematical Modelling* **39**, 667–681.
- Tikhonov, A. N. & Arsenin, V. Y. 1977 *Solutions of Ill-Posed Problems*. Winston and Sons, New York, USA.
- Wagner, B. J. 1992 Simultaneous parameter estimation and contaminant source characterisation for coupled groundwater flow and contaminant flow and contaminant transport modelling. *Journal of Hydrology* **135**, 275–303.

First received 2 March 2016; accepted in revised form 24 August 2016. Available online 25 October 2016