Geophys. J. Int. (2006) **166,** 907–922 doi: 10.1111/j.1365-246X.2006.02926.x

Efficient sensitivity analysis for flow and transport in the Earth's crust and mantle

C. A. Hier-Majumder,[1] B. J. Travis,[2] E. Bélanger,[3] G. Richard,[4] A. P. Vincent[5] and D. A. Yuen[5,6]

[1]*Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, DC* 20015, *USA. E-mail: catherine@dtm.ciw.edu*
[2]*Atmospheric, Climate, and Environmental Dynamics Group, Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM* 87545, *USA*
[3]*Département de physique, Université de Montréal, Montréal, Québec* H3C 3J7, *Canada*
[4]*Department of Geology and Geophysics, Yale University, New Haven, CT* 06520, *USA*
[5]*Department of Geology and Geophysics, University of Minnesota, Minneapolis, MN* 55455, *USA*
[6]*Minnesota Supercomputing Institute, University of Minnesota, Minneapolis, MN* 55455, *USA*

Accepted 2006 January 17. Received 2005 December 13; in original form 2005 May 4

SUMMARY

This paper focuses on the role of sensitivity analysis in studies of water transport in the environmental and geophysical sciences. Sensitivity analysis is useful in model development to determine how changes in parameter input will affect a model's response. It is also used in data assimilation to incorporate observations into a model. The adjoint method allows an efficient calculation of a model's sensitivity when the number of parameters exceeds the number of observations. In this paper we present studies demonstrating the use of the adjoint method for modelling water transport in both the Earth's crust and mantle.

Key words: fluid dynamics, fluids in rocks, hydrology, inverse problems, mantle convection, sensitivity.

1 INTRODUCTION

It is generally not possible to obtain an exactly true answer when calculating a numerical solution of a physical system model because there are always several sources of uncertainty. All physical input parameters are associated with some degree of error, and it is impossible to include all the parameters that actually influence the system in the model. Mathematical simplifications generate additional uncertainties in a model. A major limit on controlling the uncertainty of a model is the ability to isolate the system from its surroundings (Ronen 1988a). Sensitivity analysis looks at how uncertainties in its input parameters propagate through a model. As geophysical models become more complicated it becomes vital that efficient methods be developed for understanding the sensitivity of these models to hundreds or even thousands of model parameters (Cohen 2005).

This paper will focus on physical systems described by sets of partial differential equations. We are interested in how the solutions of these differential equations vary in response to parameter variation. The sensitivity coefficient is defined as the change in an output variable, ψ_i (**p**), of the system in response to a change in an independent parameter, p_j (Ronen 1988a; Cacuci 2003):

$$S_{ij} = \frac{\Delta \psi_i}{\Delta p_j}, \tag{1}$$

where S_{ij} is the sensitivity coefficient matrix. Examples of parameters include material properties, initial conditions, boundary conditions, values on interfaces, interface location and numerical discretization factors. Examples of common output variables are velocities and temperatures. A sensitivity coefficient is basically a derivative, that is, the rate of change of some cost function with respect to another variable, mathematically, a Fréchet or even a Gâteau derivative; see, for example, (Tarantola 1987; Marchuk 1995; Cacuci 2003). The differentiation method involves solving the forward differential sensitivity equation for the sensitivity coefficients (Tomović 1963; Cacuci 1988; Cacuci & Hall 1984).

Often a cost function is chosen to give a representative idea of the state of the system so that changes in it measure the cost paid by the system for a change in a given parameter. The cost function is also sometimes referred to as the response or measure and is defined by some function of which the parameters are the independent variables. The cost function depends on the dependent variables of a model, which are related to the independent parameters. It is given by $\mathcal{J}(\psi(\mathbf{p}))$, where $\psi(\mathbf{p})$ is a dependent variable, the results of the model. Cost functions can be peak or average values over time intervals, or they can be a value at a fixed point in space or over an interval within a domain or along a boundary; they can be virtually any statement that involves a dependent variable. Some possible cost functions are travel times, pressures, concentrations or fluxes. In this case $\psi(\mathbf{p})$ usually represents the solution to a set of differential equations. The derivative of $\mathcal{J}$ can be separated into two parts: a derivative of the cost function with respect to the model dependent variables and then a derivative of model dependent variables with

© 2006 The Authors
Journal compilation © 2006 RAS

respect to model independent variables:

$$S_{ij} = \frac{\partial \mathcal{J}}{\partial p_j} = \frac{\partial \mathcal{J}}{\partial \psi_i}\frac{\partial \psi_i}{\partial p_j}. \tag{2}$$

The first part of eq. (2) depends on the sensitivity cost function chosen, which can be almost anything, but which is usually easy to deal with. The second part is just the derivative of the model (usually existing as a computer code) with respect to its parameters; this part is much more difficult to compute but is always the same for any cost function.

Often we want to find relative importance of a single parameter over a given time interval, $[0, \tau]$. This is known as the relative sensitivity. A cost function is defined as:

$$\mathcal{J}(\mathbf{x}, \mathbf{p}) = \int_0^\tau f(t; \mathbf{x}, \mathbf{p})\, dt. \tag{3}$$

In this case the Gâteaux differential is defined as:

$$\delta\mathcal{J}\left(\mathbf{x}^0, \mathbf{p}^0; \boldsymbol{\delta}_\mathbf{x}, \boldsymbol{\delta}_\mathbf{p}\right) = \int_0^\tau \frac{\partial f}{\partial \mathbf{x}}|_{\mathbf{x}^0,\mathbf{p}^0} \cdot \boldsymbol{\delta}\mathbf{x}\, dt + \int_0^\tau \frac{\partial f}{\partial \mathbf{p}}|_{\mathbf{x}^0,\mathbf{p}^0} \cdot \boldsymbol{\delta}\mathbf{p}\, dt, \tag{4}$$

where $\boldsymbol{\delta}\mathbf{x}$ and $\boldsymbol{\delta}\mathbf{p}$ are perturbations around $\mathbf{x}$ and $\mathbf{p}$. The integrand is the tangent linear model (TLM) (Kalnay 2003). To find the relative sensitivity of a given parameter, p_j, we define the perturbation as:

$$\boldsymbol{\delta}\mathbf{p} = (0, \ldots, \delta p_j, \ldots, 0). \tag{5}$$

We then use the Gâteaux differential with respect to this perturbation, $\delta\mathcal{J}_j$, to compute the relative sensitivity (Zou *et al.* 1993):

$$s_j = \frac{\delta\mathcal{J}_j}{\mathcal{J}}\frac{\boldsymbol{\delta}\mathbf{p}}{p_j^0}. \tag{6}$$

The magnitude of s_j can then be used as a guide to the importance of a given parameter.

Knowledge of sensitivities can be a valuable adjunct to models per se. Sensitivities can be and should be used in experimental and observational program design to determine where to monitor output and how much data to collect. In process design, sensitivities are used to locate critical points. For inverse problems, sensitivities are part of any algorithm to find initial conditions, boundary conditions, and/or material properties that best match a given set of observations. Sensitivities are also needed to determine a model's predictability, by identifying areas in space and/or time where rapid growths in error may occur. The state of the art in sensitivity analysis can be found in (Cacuci *et al.* 2005; Cacuci 2003).

The calculation of model sensitivities is important to many scientific and engineering fields, such as in atmospheric models (Hall & Cacuci 1983; Zhu & Navon 1998; Li & Navon 1998) and in models for prediction of water flow and floods in systems involving rivers, canals and estuaries (Sanders & Katopodes 1999; Sanders & Katapodes 2000; Bélanger *et al.* 2002; Bélanger & Vincent 2004). It has been crucial to nuclear engineering since its inception (Wigner 1945; Stacey 1974; Greenspan 1975; Cacuci 1990; Cacuci *et al.* 2005). Sensitivity analysis is used in oil reservoir engineering for economic analysis and determining well placement (Wu *et al.* 2003; Li *et al.* 2003). It is important in process control both in chemical engineering (Whitecombe *et al.* 2003) and other industries (Homescu & Navon 2003; Gunzburger 2003). An accurate knowledge of the sensitivity of nuclide migration time is essential in models of nuclear waste disposal sites (Lee *et al.* 1994). Sensitivities are needed for porous flow and hydrology models used in water resource control (Christiaens & Feyen 2002; Kunstmann *et al.* 2002; Stauffer *et al.* 2002; Francos *et al.* 2003). Sensitivities are also important in assimilating data into and solving inverse problems in atmospheric transport and ocean circulation (Bennett 1992; Kalnay *et al.* 2000; Yang & Hamrick 2002; Le Dimet *et al.* 2002; Bennett 2002; Kalnay 2003; Erofeeva *et al.* 2003), weather forecasting (Cacuci *et al.* 2005), atmospheric chemistry (Daescu *et al.* 2000; Vautard *et al.* 2003; Hess & Vukicević 2003), porous media characterization (Hughson & Jim Yeh 2000; Jim Yeh *et al.* 2002; Altmann-Dieses *et al.* 2002; Delay & Porel 2003), astrophysics (Reynolds *et al.* 1999), mantle convection (Bunge *et al.* 2003; Ismail-Zadeh *et al.* 2004), chemical transport in biosystems (Kaminski *et al.* 2003), and medicine (Li & Xing 2000). Efficient sensitivity analysis is essential for global climate modelling due to the unusually large number of significant parameters (Cacuci & Hall 1984; Cacuci *et al.* 2005).

In this paper we will focus on differential sensitivity analysis (DSA) and the use of the adjoint method to determine sensitivities in environmental and geophysical flow calculations. First, we review various ways to compute sensitivities.

2 CALCULATING SENSITIVITIES

The simplest procedure used to determine the sensitivity of a model is the 'brute force' method. This involves varying a selected input parameter, recomputing the model, and comparing how the output changes. This method requires the model to be completely recomputed to test the sensitivity to each individual parameter over a defined range. The total number of computer runs will be equal to the number of parameters, plus one. It is only practical to use this method on a very limited number of parameters over a very specific range of values if rerunning the model requires considerable computation time. Although this method gives useful results about a specific parameter of interest, it is generally not feasible for large systems with distributed parameters. A major limitation of the method is that the parameters to be studied must be identified *a priori*. This means that important sensitivities that were not expected may be missed.

The goal of interval analysis is to find the variation in a response if the parameters are allowed to vary around a central value within certain bounds (Hansen 1969; Deif 1986). This method involves using interval arithmetic to calculate the bounds of responses that are returned when the parameters are within a given range. The method has two main drawbacks, both related to the necessity of using interval arithmetic (Deif 1986). The first is that computer hardware is not designed to perform interval arithmetic efficiently. The second is that, because of the subdistributive property of interval arithmetic, the method tends to give unnecessarily large bounds, resulting in an exaggerated estimate of a model's sensitivity. However, improvements have been made to reduce this limitation (Neumaier 1990; Gravilliers *et al.* 2004).

Statistical and stochastic techniques have been developed to calculate sensitivity for systems whose parameters are distributed according to a probability distribution (Gaussian, Levy and the like). For laboratory measurements, one usually receives different answers when repeating the same measurement. This variation is due to stochastic error and statistical methods are used to gain maximum information from the results. There is a rich literature on stochastic processes (McKay 1988), but we are not concerned with stochastic variables here.

Perturbation analysis is an additional method that is used to calculate sensitivities. The classical perturbation theory (CPT) provides sensitivity for the linearized analogue of the model under question (Gandini 1988). The method estimates the change of the

system's eigenvalues with respect to a parameter perturbation. The main drawback to the perturbation method is that it is based on linear functional analysis (Gandini 1988; Cacuci 1988) so a linearized form must be used. Since the perturbation theory involves a Taylor Series expansion, it may be difficult to determine whether the first term in the series is adequate to completely describe the sensitivity without knowledge of how the system converges (Cacuci 1988). This approach can also be time consuming when many perturbations are to be studied because a calculation must be computed for each perturbation (Delay & Porel 2003).

The generalized perturbation theory (GPT) is also known as variational sensitivity analysis (Henryson *et al.* 1974). This method was developed in nuclear reactor theory to take into account effects arising both directly from the perturbation of system parameters and indirect effects in changes to the state function due to system alteration without needing to recalculate the state function (Cacuci 1988). It requires the existence, uniform continuity, and derivability of derivatives of operators and state functions with respect to system parameters and space–time variables (Cacuci 1988). The GPT has many of the same drawbacks as the CPT, including relying on the linear equations, the necessity of Taylor Series expansion and the need to solve many additional equations.

Functional analysis is also based on perturbation theory but avoids solving the perturbation equations. It is used to estimate an upper bound for the uncertainty by analysing the mathematical properties of the uncertainty operator (Ronen 1988b). This is done by finding an upper bound, which is the norm of the uncertainty operator, in a given space (Ronen 1988b). The main drawback of the method is that it does not give exact values of the sensitivity coefficients, only estimates of the uncertainty in a model based on a given perturbation.

Let a system be described by differential equations of the form:

$$\mathcal{F}(\ddot{\psi}(\mathbf{p}, t), \dot{\psi}(\mathbf{p}, t), \psi(\mathbf{p}, t)) = 0. \quad (7)$$

If δp_j is a small perturbation of parameter p_j, the sensitivity coefficient for the output variable ψ_i is defined as:

$$\lim_{\delta p_j \to 0} \left(\frac{\psi_i(t, p_j + \delta p_j) - \psi(t, p_j)}{\delta p_j} \right) = \frac{d\psi_i(t, p_j)}{dp_j} = s(t, p_j). \quad (8)$$

The forward sensitivity equation can then be formulated using the chain rule:

$$\frac{\partial \mathcal{F}}{\partial \ddot{\psi}_i}\ddot{s} + \frac{\partial \mathcal{F}}{\partial \dot{\psi}_i}\dot{s} + \frac{\partial \mathcal{F}}{\partial \psi_i}s = \frac{\partial \mathcal{F}}{\partial p_j}. \quad (9)$$

This equation can be solved for the sensitivity coefficient, s. Note that this gives the sensitivity of a specific output variable, ψ_i with respect to a specific input variable, p_j. This can then be used in eq. (2) to the find the sensitivity of the cost function with respect to p_j. This approach can also be used for non-linear systems (Cacuci 1988, 2003). Its main drawback is that it requires solving the forward sensitivity equation for each parameter in turn for each output variable.

The Green's function method for sensitivity analysis involves replacing differential equations for the sensitivity with a set of integrals through definition of a Green's function. The approach is useful in applications such as chemical kinetics where a set of stiff differential equations can be replaced by a set of smooth integrands (Hwang *et al.* 1978). One drawback to this approach is that a time grid must be chosen prior to the calculation. This could lead to problems if rapid variations in the integrand occur at unexpected times.

Statistical methods are often combined with 'response surface methods' to glean as much information as possible from recomputations done for sensitivity analysis (Cacuci 1988). The modeller selects a small number of important parameters and design points in the parameter space where the model will be recalculated. These recalculated results are used to construct 'response surfaces' that approximate the behaviour of the response measure as a function of the chosen parameters. These response surfaces are then used in statistical studies to estimate sensitivities and uncertainty distributions. This method, however, has several drawbacks (Cacuci 1988). Only a small subset of parameters can be tested due to the cost of recalculation. Therefore, assumptions must be made about parameter sensitivity and importance prior to conducting the sensitivity analysis. This makes it relatively easy to miss important behaviours. Another statistical technique is the Fourier amplitude sensitivity test (FAST) (Cacuci 1988) that calculates statistical mean values and standard deviations for model responses. Each parameter is varied as a periodic function of a search variable. The system solutions are then periodic, and the sine coefficients of the resulting Fourier series are approximately related to the sensitivity coefficients averaged over the uncertainties of all the parameters in the system (Hwang *et al.* 1978). FAST is perhaps a misnomer, however, as this method requires a substantial number of recalculations. A method for calculating operator responses with the adjoint method has been developed that avoids the need for large numbers of recalculations (Zou *et al.* 1993).

The adjoint method, which will be discussed in more detail in Section 2.1, avoids the need to solve the sensitivity equations individually for each parameter (Cacuci 1981a,b, 2003). It involves the use of a functional derivative to derive an auxiliary set of equations, the adjoint equations that can be solved for the adjoint variables (Bunge *et al*. 2003; Cacuci 2003). The sensitivity coefficients then are determined from a defined convolution of forward and adjoint variables. The adjoint approach is very efficient, especially if the number of parameters involved becomes large. When the adjoint equations are manipulated appropriately, the adjoint of a model can be computed by using the original model coding, with minor modifications, even when the model is non-linear. The modifications result in a new source term for the adjoint equations. Many adjoint methods use only the first derivative and therefore have some of the same drawbacks as CPT and GPT, but second order adjoint methods have been developed (Ngodock 1996; Le Dimet *et al.* 1997).

Another use of sensitivity coefficients is in variational data assimilation, which involves incorporating observational data into numerical simulations. The word variational is used to imply that the method can incorporate data of 'varying' trustworthiness (Le Dimet & Navon 1988). Sensitivity studies are essential to this process to adjust the simulation to best match the given observations. Variational data assimilation has been extensively researched in the meteorology and oceanography community in order to understand how to best introduce large quantities of data with widely varying qualities and spatial resolutions into numerical models attempting to forecast the complicated coupled behaviour of the ocean and atmosphere (Bennett 1992, 2002). One common and powerful algorithm is known as 4-D variational assimilation (4D-VAR). A forward forecast model, such as a weather prediction based on noisy initial data, is compared with observable data throughout the runtime. Measurements of the error between the predicted and observed values are then taken, and the adjoint sensitivity coefficients are used to adjust the initial data and obtain an improved forecast. This involves finding the initial conditions that minimize the error and is therefore, actually an optimal control problem (Gunzburger 2003). Minimizing the error involves the need to know not only the sensitivity coefficients but also their concavity. This has lead to the development of

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

second-order adjoint sensitivity analysis (Ngodock 1996; Le Dimet *et al.* 1997).

Advances in algorithmic differentiation (AD) have greatly enhanced the ability to calculate sensitivity coefficients (Griewank 2000). Traditionally, derivatives have been calculated using the difference quotient. This method, however, is severely hampered by truncation errors and loss of significant digits. AD, however, allows derivatives to be calculated with no truncation error and without resorting to symbolic differentiation. It works by systematically applying the chain rule to the actual numerical values throughout the evaluation trace of the function, where the evaluation trace is a record of floating point values and the operations that computed them (Griewank 2000). Basically, the sensitivity of the cost function is calculated for each intermediate step in the evaluation trace with respect to a given observable. This requires augmenting the evaluation trace with the necessary extra code lines and variables. This can be done through the use of a pre-processor. The AD is especially powerful when used with the adjoint method since the augmented evaluation trace does not need to be modified for each parameter. Several pre-processors have been developed that use AD, including GRESS (Horwedel 1991, 1992), ADIFOR (Bischof & Griewank 1992), and TAMC (Giering & Kaminski 1988). These codes read in application codes and create new codes that calculate sensitivity coefficients for specified sets of variables by adding the additional lines and variables needed to the evaluation trace. The primary drawback of this method is that it tends to be not as efficient as a chain rule or adjoint version of the application code written by someone with intimate knowledge of the application code's structure.

2.1 Adjoint methods

Our goal is to provide an efficient way of computing the sensitivity of some pre-defined cost function with respect to a distributed parameter. In the context of fluid flow and mass and energy transport, this means determining the sensitivity of a functional of an observable quantity, such as concentrations, fluxes, pressures, or temperatures, to changes or uncertainties in system parameters such as diffusivity, topography or numerical discretization parameters. For example, one might want to know the sensitivity of arrival time of a particular concentration level at some distance from a source to the spatial variability of sorption properties. The methodology is applicable to a wide range of disciplines.

An advantage of using the adjoint sensitivity approach is that it is well understood, having been used by a variety of disciplines since its first introduction during the Manhattan Project (Wigner 1945). The adjoint operators, originally defined by Lagrange, have been thoroughly explored theoretically (Marchuk 1995). The adjoint method uses a reverse or backward mode of differentiation to find the sensitivity coefficients. This makes it much easier to implement AD methods to find sensitivity coefficients. Instead of calculating the sensitivity of the measure at each intermediate step with respect to an observable, one calculates the sensitivity of the measure with respect to each intermediate value (Griewank 2000). This means we have only one dependent variable rather than a dependent variable for each step in the evaluation trace so we can use the same augmented evaluation trace in the AD method to find the sensitivity with respect to all the observables rather than constructing a new augmented evaluation trace for each observable.

The method has been broadly applied in solving many problems in the physical sciences, including nuclear reactor safety issues (Wigner 1945; Cacuci *et al.* 1980; Oblow & Pin 1986; Cacuci 1990; Cacuci *et al.* 2005), oil reservoir engineering for well placement studies and economic analyses (Chavent Duprey & Lemornier 1975; Li *et al.* 2003), oceanography and meteorology (Errico & Vukicević 1992; Errico 1997; Li & Navon 2001; Bennett 2002), fluid flow control and optimization (Gunzburger 2003; Homescu & Navon 2003), chemical engineering for process control (Rabitz 1981; Scott 1988; Van Noorden *et al.* 2003), economics (Seierstad 1982; de Jongh 1997) and water resources management (Sykes *et al.* 1985; Kool *et al.* 1987; Knopman & Voss 1987, 1988; Katopodes & Piasecki 1996). It has also been used for estimating and controlling error in numerical solutions (Cao & Petzold 2004). Navon (1998) provides a survey of adjoint methods of parameter estimation in meteorology and oceanography. The adjoint equations combined with variational data assimilation to create the 4D-VAR method has become essential to meteorologic and oceanographic forecasting (Bennett 2002). The method was first introduced in the late 1980s and early 1990s (Le Dimet & Talagrand 1986; Talagrand & Courtier 1987; Le Dimet & Navon 1988; Tziperman & Thacker 1989; Courtier & Talagrand 1990). One of the first applications to an operational meteorological model was by Navon *et al.* (1992).

Here we will focus on the development of adjoint models for use in computational fluid dynamics (CFD). In this case, we generally have a non-linear model, $\mathcal{M}$ that maps a set of input parameters, $\mathbf{p} = (p_1, \ldots, p_k)$, in k-dimensional space into a set of output variables, $\psi = (\psi_1, \ldots, \psi_n)$, in n-dimensional space. We can then define a scalar cost function, $\mathcal{J}(\psi) = \mathcal{J}(\mathcal{M}(\mathbf{p}))$ that is designed to give a representation of the state of the model. In CFD, heat or concentration fluxes are commonly used as cost functions. If we add a small perturbation to the input, the resulting perturbation to the cost function can then be approximated as:

$$\delta\mathcal{J} = \langle \nabla_\psi \mathcal{J}, \delta\psi \rangle = \langle \nabla_\psi \mathcal{J}, \mathbf{M}\delta\mathbf{p} \rangle = \langle \nabla_\mathbf{p} \mathcal{J}, \delta\mathbf{p} \rangle, \tag{10}$$

where $\mathbf{M}$ is the Jacobian:

$$M_{ij} = \partial \mathcal{M}_j / \partial p_i, \tag{11}$$

and the inner product is defined such that:

$$\langle \nabla_\psi \mathcal{J}, \delta\psi \rangle = \left(\frac{\partial \mathcal{J}}{\partial \psi_1}, \ldots, \frac{\partial \mathcal{J}}{\partial \psi_n} \right) \cdot (\delta\psi_1, \ldots, \delta\psi_n). \tag{12}$$

This gives the TLM (Kalnay 2003). We can use eq. (10) to find the sensitivity. This is known as the forward method because it finds the sensitivities by starting with the input parameters and moving forwards.

The backward or adjoint method starts with the output variables and works backwards. It is much more computationally efficient because it is more amenable to AD (Section 2), when there are generally many less output variables than input parameters. The adjoint of an operator is defined as:

$$\langle \mathbf{A}^* \mathbf{x}, \psi \rangle = \langle \mathbf{x}, \mathbf{A}\psi \rangle. \tag{13}$$

If we are limited to a finite-dimensional real vector space as in most CFD problems, the adjoint of an operation is simply the transpose:

$$\mathbf{A}^* = \mathbf{A}^{\mathrm{T}}. \tag{14}$$

Therefore:

$$\langle \nabla_\psi \mathcal{J}, \mathbf{M}\delta\mathbf{p} \rangle = \langle \mathbf{M}^* \nabla_\psi \mathcal{J}, \delta\mathbf{p} \rangle. \tag{15}$$

From eq. (10) we get:

$$\nabla_\mathbf{p} \mathcal{J} = \mathbf{M}^* \nabla_\psi \mathcal{J} = \mathbf{M}^{\mathrm{T}} \nabla_\psi \mathcal{J}. \tag{16}$$

We then define the adjoint variables such that:

$$\nabla_\psi \mathcal{J} = \delta\psi^*. \tag{17}$$

The adjoint variable, $\delta\psi^*$, can then be used to find the sensitivity coefficients:

$$\nabla_{\mathbf{p}}\mathcal{J} = \mathbf{M}^{\mathrm{T}}(\nabla_{\psi}\mathcal{J}) = \mathbf{M}^{\mathrm{T}}\delta\psi^*, \tag{18}$$

and the gradient of the cost function with respect to the parameters gives the sensitivity coefficients. If we have an integration of Λ time steps, we can find the gradient of the cost function with respect to the input variables by moving backwards from the gradient of the cost function with respect to the output:

$$\begin{aligned}\nabla_{\mathbf{p}}\mathcal{J} &= \mathbf{M}^{\mathrm{T}}\nabla_{\psi}\mathcal{J} = \mathbf{M}_0^{\mathrm{T}}\ldots\mathbf{M}_{\Lambda}^{\mathrm{T}}\nabla_{\psi}\mathcal{J}\\ &= \mathbf{M}^{\mathrm{T}}\nabla_{\psi}\mathcal{J} = \mathbf{M}_0^{\mathrm{T}}\ldots\mathbf{M}_{\Lambda}^{\mathrm{T}}\delta\psi^{(\Lambda)*}.\end{aligned} \tag{19}$$

Note that although we actually began with a non-linear model both the TLM and adjoint models used to find the sensitivity coefficients are linear.

2.2 4D-VAR

When simulating geophysical fluid flows, it is never possible to define exact deterministic equations controlling the evolution of the system. One reason for this is that there are always uncertainties in our knowledge of the system. Another reason is that there are many processes such as conduction of heat and viscous dissipation that occur at subgrid scales. These processes must be estimated by diffusivity terms. The equations, therefore, are one piece of data available about the system. Other data available to the modeller are the observations collected both at the start of the model time (initial conditions) and throughout the model runtime. These observations themselves have two main types of uncertainties. The first is the measurement error. The observations are likely to be of different types with widely varying uncertainties. For example, for an ocean simulation one may have temperatures inverted from satellite data and buoy measurements of wave speeds. For a mantle convection simulation one may have seismic tomographic data showing temperature anomalies in the mantle, surface heat flux measurements, and past plate positions. Also, the available observations will likely have widely varying spatial and time coverage. This motivates a search for a method that can account for the heterogeneous nature of the available data. One way to formulate the problem is as an optimal control problem. We want to find the optimal prediction for the natural fluid flow that both fits the available observations and is constrained by the flow equations (Le Dimet *et al.* 1997, 2002). The adjoint method for solving optimal control problems has been well developed in CFD engineering in order to solve problems such as the optimal shape of a wing for minimizing drag (Gunzburger 2003). The 4D-VAR method uses optimal control adjoint methods to find the trajectory in space–time of the output variables, ψ that minimize the cost function while obeying the physical equations that act as constraints. This gives a problem of minimization with constraints (Talagrand & Courtier 1987).

In an optimization problem, the goal is to find the parameters, $\mathbf{p}$ that minimize a given cost function, $\mathcal{J}(\mathbf{p}, \mathbf{obs})$, subject to the constraints of the physical equations, $\mathcal{E}$ (Gunzburger 2003). The physical equations of fluid flow in the system are given by:

$$\mathcal{E}(\mathbf{x}, t, \mathbf{p}) = 0. \tag{20}$$

We can then define a cost function that measures the error between the prediction calculated by the physical equations using the chosen parameters and the observations:

$$\mathcal{J} = \int_0^{\tau}\int_{\Omega} \mathrm{f}(\mathbf{obs}, \mathbf{p})\, d\Omega\, dt, \tag{21}$$

integrated over the model volume, Ω, and total time, τ.

We now have an optimization problem that requires minimizing eq. (21) subject to the constraints of eq. (20). Usually, when resolving a problem with constraints, one forms the Lagrangian:

$$\mathcal{L}(\mathbf{p}, \boldsymbol{\lambda}) = \mathcal{J}(\mathbf{p}) + \int_0^{\tau}\int_{\Omega} \boldsymbol{\lambda}(\mathbf{x}, t)\cdot\mathcal{E}(\mathbf{p}, \mathbf{x}, t)\, d\mathbf{x}\, dt, \tag{22}$$

where $\boldsymbol{\lambda}(\mathbf{x}, t)$ are the undetermined Lagrangian multipliers (Sanders & Katopodes 1999). One of the shortcomings of the 4D-VAR method is that the use of a strong constraint assumes that the model is perfect. There are methods to account for this problem by including the model error in the cost function.

We want to find the extrema of the Lagrangian (Le Dimet & Talagrand 1986). In order to accomplish this, we use the variational operator, δ, of the Lagrangian. Taking the variation of the Lagrangian, we obtain:

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\mathbf{p}}\boldsymbol{\delta}\mathbf{p} + \frac{\partial\mathcal{L}}{\partial\boldsymbol{\lambda}}\boldsymbol{\delta}\boldsymbol{\lambda}. \tag{23}$$

We have also formed the TLM of our problem in the above step (Ehrendorfer 1992). For an arbitrary displacement, $(\boldsymbol{\delta}\mathbf{p}, \boldsymbol{\delta}\boldsymbol{\lambda})$, we have reached an extremum when $\delta\mathcal{L} = 0$ (Daley 1991). The derivative of the Langrangian with respect to each displacement direction must be zero at the extremum:

$$\frac{\partial\mathcal{L}}{\partial\boldsymbol{\lambda}} = \mathcal{E}(\mathbf{p}, \mathbf{x}, t) = 0, \tag{24}$$

and

$$\frac{\partial\mathcal{L}}{\partial\mathbf{p}} = \frac{\partial\mathcal{J}}{\partial\mathbf{p}} + \frac{\partial\mathcal{G}}{\partial\mathbf{p}} = 0, \tag{25}$$

where $\mathcal{G}$ is:

$$\int_0^{\tau}\int_{\Omega} \boldsymbol{\lambda}(\mathbf{x}, t)\cdot\mathcal{E}(\mathbf{p}, \mathbf{x}, t)\, d\mathbf{x}\, dt. \tag{26}$$

We can see, however, from eq. (18) that if we integrate the time integral from τ to 0, $\mathcal{G}$ is simply the adjoint equations, and the Lagrange undetermined multipliers are the adjoint variables (Schröter *et al.* 1993). Of course it is also important to know whether one is approaching a minimum or maximum in the Lagrangian. This requires the use of the second-order adjoint method (Ngodock 1996; Le Dimet *et al.* 1997). Eqs (24) and (25) are the Euler–Lagrange equations and can be solved using the method of representers (Bennett 1992, 2002).

3 CONTAMINANT TRANSPORT EXAMPLES

Two simple examples illustrate the utility of the adjoint approach, in the context of steady porous flow and transient transport. The transport equation for a sorbing tracer with saturated flow is:

$$\epsilon(1+R)\frac{\partial C}{\partial t} + \nabla\cdot(\mathbf{v}C) - \nabla\cdot(\epsilon D\nabla C) = \dot{\mathrm{S}}(x, y, t), \tag{27}$$

where C is concentration; $\mathbf{v}$ is water velocity; D is diffusivity/dispersivity; $\dot{S}$ is a source/sink function; t is time; R is retardation; and ϵ is the porosity. R, $\mathbf{v}$, and ϵ may be spatially varying. The equation governing flow is given by:

$$\nabla\cdot(\rho\mathbf{v}) = -\nabla\cdot\left(\rho\frac{k}{\mu}\nabla p\right) = \dot{M}, \tag{28}$$

where p is pressure; k is permeability; $\dot{M}$ is a mass source/sink; ρ is fluid density; and μ is fluid viscosity. In eq. (28), Darcy's law relates velocity, $\mathbf{v}$, and the pressure gradient in soils at low Reynolds

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

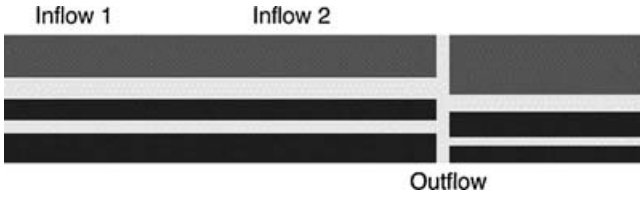

Figure 1. Faulted rock sequence for example demonstrating sensitivity of contaminant transport model to permeability distribution. Low-permeability rock is dark grey, intermediate permeability rock is medium grey and high-permeability rock is light grey. Outflow refers to the narrow strip where the vertical fault intersects the bottom boundary.

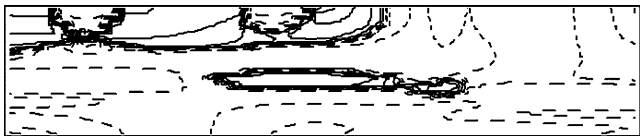

Figure 2. Values of $d\mathcal{J}/d\ln k$ at each point in the domain as computed with the adjoint algorithm. The absolute values are not important since this is an artificial problem.

number. Boundary and initial conditions are given as appropriate for both equations. The sensitivity of C and any functional involving C to any of the other variables appearing in eqs (27) and (28) along with the boundary and initial conditions can be determined efficiently through use of the adjoint concept.

The sensitivity of a functional of C with respect to variations in permeability, dispersion, and sorption, is important in studies of solute transport. An example for a cost function measuring cumulative vertical contaminant flux across a plane at a particular depth in a domain is:

$$\mathcal{J} = \int dA_h \int_0^\tau w_l C \delta(y - y_0)\, dt, \tag{29}$$

where A_h is the area; w_l is the liquid mass flux; and y_0 is the monitoring depth. We emphasize, though, that $\mathcal{J}$ can have a great variety of forms, involving, either explicitly or implicitly, the dependent variables. For the forward problem defined here, we will naturally be interested in $d\mathcal{J}/dR$ or $d\mathcal{J}/dk$, since R and k (the permeability field) are the major distributed independent variables. The particular forms for these will be given in the next section.

Following the optimization procedure in Section 2.2, the adjoint equations corresponding to the forward model (eqs 27 and 28) were obtained from integration (eq. 25):

$$-\epsilon(1+R)\frac{\partial C^*}{\partial t} - \nabla\cdot(\mathbf{v}C^*) - \nabla\cdot(\epsilon D\nabla C^*) = -C^*\nabla\cdot\mathbf{v} + \frac{\partial\mathcal{J}}{\partial C}, \tag{30}$$

$$-\nabla\cdot\left(\rho\frac{k}{\mu}\nabla P^*\right) = \nabla\cdot\left(C\frac{k}{\mu}\nabla C^*\right), \tag{31}$$

and then discretized. The last term in eq. (30) is easily derived from eq. (29) and is given by:

$$\frac{\partial\mathcal{J}}{\partial C} = w_l\delta(y - y_0). \tag{32}$$

This will of course depend on the particular definition of $\mathcal{J}$.

For each dependent variable, there is a forward equation and a corresponding adjoint variable (indicated by *) and adjoint equation. Boundary conditions for the adjoint variables have the same form as for the forward model, except that they are homogeneous. Further, instead of an initial condition, the time-dependent adjoint C^* satisfies an end time condition, $C^*(\mathbf{x}, \tau) = 0$, allowing eq. (30) to be solved stably, even though running backwards in time. eq. (30) has been manipulated to have the same form on the left side of the equal sign as the forward model, eq. (27), except for sign. This allows the use of the same numerical discretization treatment used in the forward model. The right-hand side of eq. (30) contains new source terms compared to eq. (27). The first source term results from the manipulation of the original adjoint equation, and the second derives from the cost function $\mathcal{J}$.

The numerical procedure consists of several steps. First, eq. (28) is solved for pressure, and then the steady state velocity field $\mathbf{v}$ is computed from Darcy's law. Eq. (27) provides the forward solution for $C(\mathbf{x}, t)$. Next the adjoint solution, C^*, of eq. (30) is computed, running backwards from the end time, τ, to the initial time. Then eq. (31) provides P^*. Finally the forward and adjoint solutions are used to compute the sensitivity of $\mathcal{J}$ from integration of eq. (25).

3.1 Sensitivity of solute concentration to permeability

The first example applies the adjoint method to computation of the sensitivity of solute outflow to the permeability distribution for contaminant transport in a faulted rock sequence. The 2-D domain consists of three materials (Fig. 1), a low-permeability rock (dark grey), an intermediate permeability soil (medium grey), and a high-permeability fractured layer (light grey) with a vertical fault. Side boundaries are no-flow. Inflow occurs along two narrow segments of the upper boundary as indicated in Fig. 1, with the Inflow 1 rate

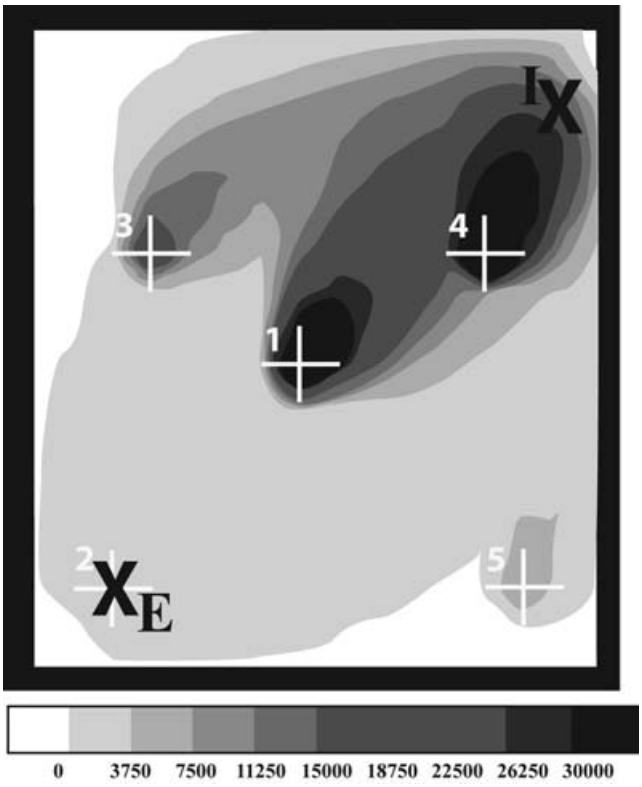

Figure 3. Distribution of $d\mathcal{J}/dR(x, y)$ for the case with a five-spot monitoring arrangement. Injection site is shown by black X labelled I and extraction site is shown by black X labelled E. Monitoring locations are shown by numbered white crosses.

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

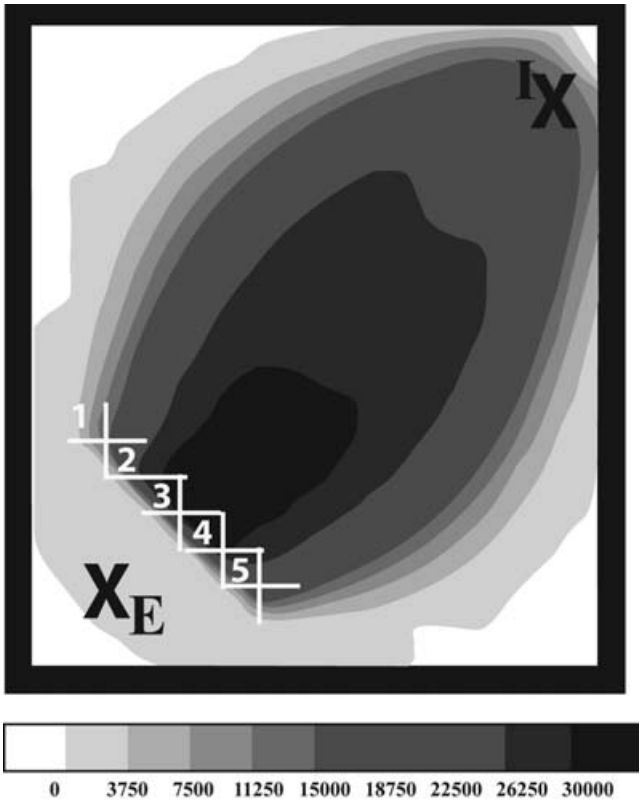

Figure 4. Distribution of $d\mathcal{J}/dR(x, y)$ for the case with collinear monitoring locations. Injection site is shown by black X labelled I and extraction site is shown by black X labelled E. Monitoring locations are shown by numbered white crosses.

larger than that of Inflow 2. Outflow is tracked through the indicated segment at lower right. After a steady flow has been established, a contaminant is injected below the centre of the Inflow 1 region, at half the depth of the top soil layer. Here, the contaminant is assumed to be non-sorbing, that is, $R = 0$ in eq. (27). We define cumulative contaminant transport, $\mathcal{J}$, through the outflow as:

$$\mathcal{J} = \int dA_h \int_0^T w_l C_{li} \delta(y - y_0)[\mathrm{H}(x - x_1) - \mathrm{H}(x - x_2)]\, dt, \quad (33)$$

where H is the Heaviside function, defined from (x_1, y_0) to (x_2, y_0) that gives the horizontal location of the outflow region, and y_0 defines the depth of the outflow measurements. Integration of eq. (25) shows that the sensitivity of $\mathcal{J}$, with respect to the permeability field, is:

$$\frac{d\mathcal{J}}{d\ln k} = \rho \mathbf{v} \cdot \nabla P^* - \int_0^\tau C\mathbf{v} \cdot \nabla C^*\, dt. \quad (34)$$

It is convenient to take this derivative with respect to the log of permeability because permeability can vary over many orders of magnitude. Note that the derivative depends on the solution of both forward equations and on both adjoint variables.

Eq. (34) is derived in the following way. Eq. (33) is perturbed with respect to C, permeability k and pressure P. The forward model (eqs 27 and 28) are also perturbed with respect to C, P, and k. Since the forward perturbed equations are satisfied by the perturbed fields, subtracting the unperturbed equations from the perturbed equations is still 0. These equation differences are weighted by as yet undefined functions and added to the perturbation of eq. (33). We can do this, since the added terms, being 0, do not change eq. 33 or its perturbed form. That sum is then integrated over time and space. Integration by parts allows recombination of terms, and almost magically, the perturbation of $\mathcal{J}$ simplifies to an integrand involving a term times dk, *if* a term multiplied by dC is taken to be 0, and another term multiplied by dP is taken to be 0. These so-constrained terms involve the previously unknown weighting functions. Since they were unassigned, we are free to choose them so that the coefficients of dP and dC become 0. The forms of these coefficients are the adjoint equations (eqs 30 and 31), and the unassigned variables are the adjoint variables C^* and P^*. In the process of performing the integration by parts mentioned earlier, so-called natural boundary conditions arise. If we further choose the adjoint variables to satisfy them, and we are free to do so, then the only term that remains in the expression for $d\mathcal{J}$ is a coefficient times dk, and that coefficient is simply what we seek, $d\mathcal{J}/dk$. The cost of simplifying the expression for $d\mathcal{J}$ is that now we must solve the adjoint equations for C^* and P^*. But these are easy to solve; numerically the same coding used to solve the forward equations for C and P can be used, with minor adjustments.

Fig. 2 shows contours of $d\mathcal{J}/d\ln k$, at each point in the domain as computed with the adjoint algorithm. Actual values are not important here, since the goal is to illustrate the kind of information that the adjoint solution can provide. In this example, the outflow is most sensitive to permeabilities in the region of the inflows, to the low-permeability layer between the two high-permeability layers, and to the intersection of that layer with the fault.

The adjoint algorithm allows us to quantify these sensitivities in an efficient manner. Traditional methods of determining permeability sensitivity require that the permeability be perturbed at each node of interest followed by recomputation of the forward model for each perturbation. Since the number of nodes usually ranges from hundreds to millions, it is generally prohibitively expensive to use such a method for studying the sensitivity at more than a few nodes. In contrast, the adjoint formulation requires only two model simulations, the forward and the adjoint. An additional simple auxiliary calculation completes the process.

3.2 Sensitivity of contaminant concentration to retardation

The second study addresses the sensitivity of contaminant concentration, C, to the retardation factor, $R(\mathbf{x})$. The governing equations are the same as in the previous example, but the measure and sensitivity parameters are different. In this case the cost function is taken to be the least-squares sum of differences between observed concentrations and those calculated from best estimates of the distribution of retardation, $R(x, y)$:

$$\mathcal{J} = \int_0^\tau \sum_i^n (C - C_{\mathbf{obs}})_i^2\, dt, \quad (35)$$

where i represents an individual monitoring point, and n is the total number of monitoring locations. The measure in eq. (35) is a typical measure that would be used in an inverse problem to determine $R(x, y)$ in a region (Hagelberg & Travis 1997). In this example, we have generated a set of 'data', that is, histories of C at the n monitoring points, assuming that we already know the permeability and sorptivity fields. The goal is to determine how well R can be recovered using these 'data' histories and assuming that we now initially do *not* know how the R field is distributed spatially. With this definition of $\mathcal{J}$, we find that

$$\frac{\partial \mathcal{J}}{\partial C} = 2 \int_0^\tau \sum_{i=1}^n (C(x, t) - C_{\mathbf{obs}}(t))\, dt. \quad (36)$$

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

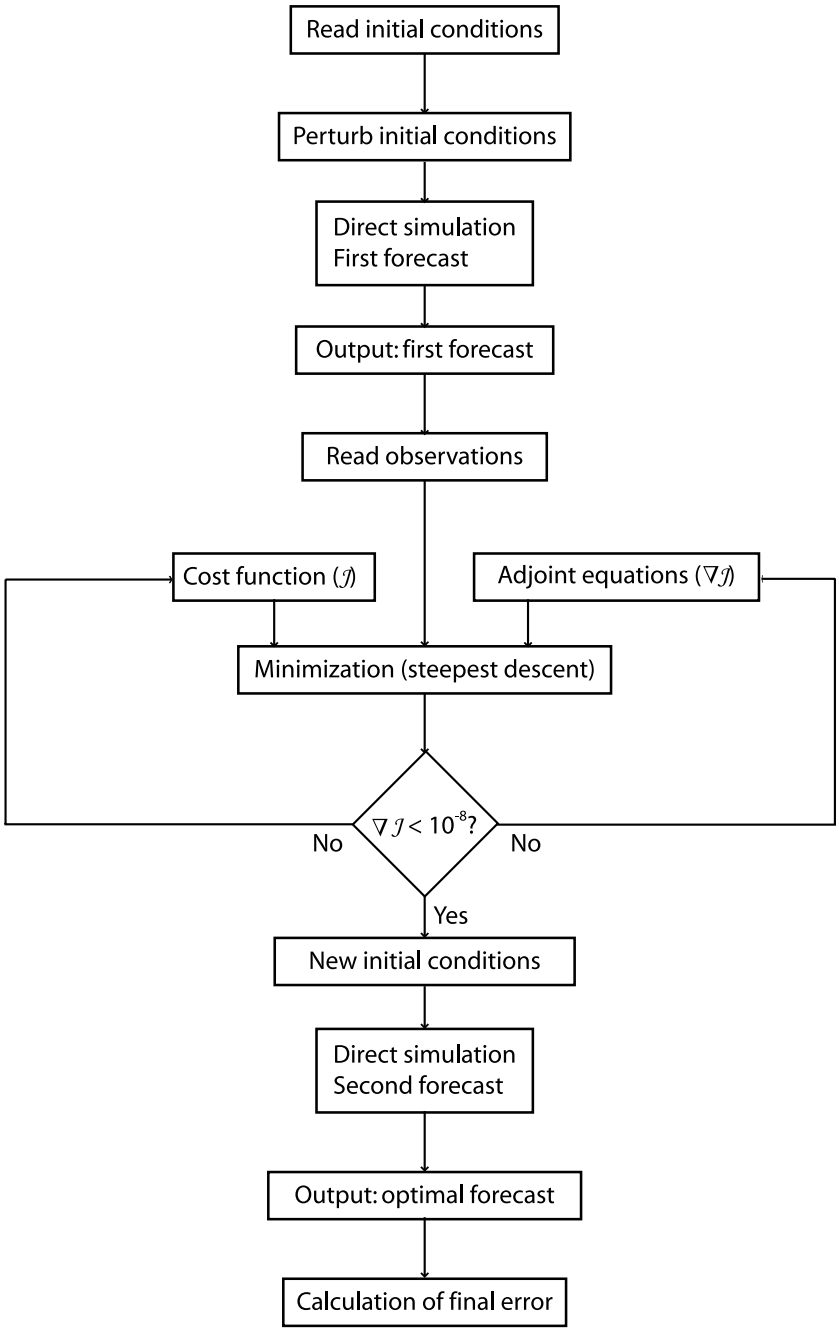

Figure 5. Algorithm for the 4D-VAR method.

This expression is required for the last term in eq. (30). Our intention here is not to solve this inverse problem but to show how the adjoint approach can provide the critical information, $d\mathcal{J}/dR$, needed to solve it. In particular, our interest here is in how $d\mathcal{J}/dR$ varies with the location of monitoring points. Since we assume that the permeability field is known, and since R does not affect flow, only the transport equation (eq. 27) is perturbed, and there is only one adjoint, C^*, governed by eq. (30). The optimization procedure in Section 2.2 then gives:

$$\frac{d\mathcal{J}}{dR}(\mathbf{x}) = \int_0^\tau C\frac{\partial C^*}{\partial t}\,dt. \tag{37}$$

Since $\mathbf{v}$ is known, we need only solve the forward equation (eq. 27) and its adjoint (eq. 30) then compute $d\mathcal{J}/dR$ from eq. (37).

The study region is shown in Fig. 3. Water is injected continuously near the upper right corner of a closed permeable domain and extracted at the lower left. This creates a steady flow through a region of variable but known permeability. Then a pulse of solute is introduced at I, and the subsequent concentration histories at five monitoring locations are recorded and taken as 'data'. In the inverse problem, we would make a guess at the initial R distribution, usually an average constant value based on the data, and proceed to improve on that initial guess through an iterative process, using the values of $d\mathcal{J}/dR$ to update the R field at each iteration. But here we only

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

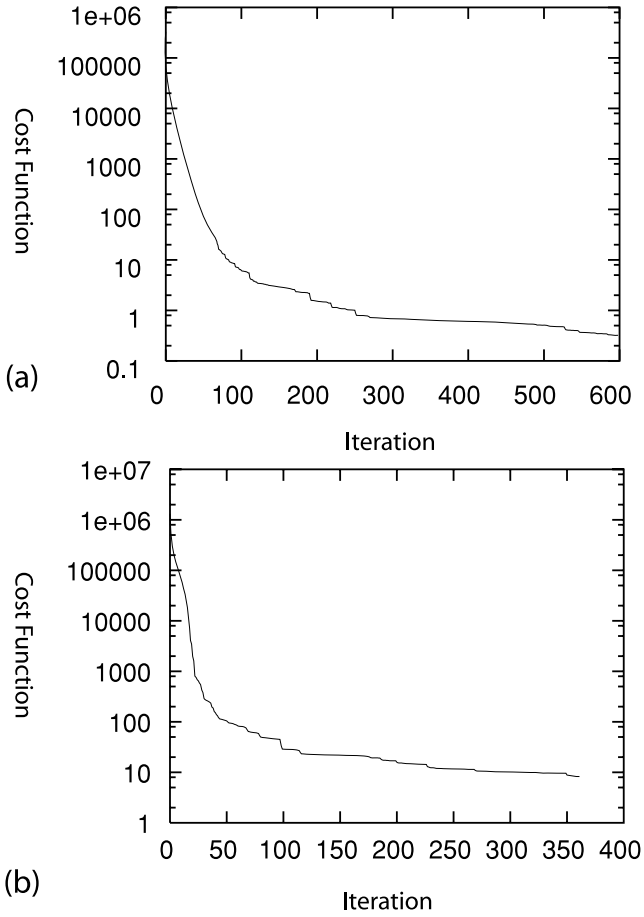

Figure 6. Minimization of the cost function with the steepest descent method. The cost function decreases exponentially. (a) small initial perturbation. (b) large initial perturbation. A larger initial perturbation results in a larger cost function.

want to show how the sensitivity of $\mathcal{J}$ to R can vary dramatically depending on the monitoring strategy.

Fig. 3 shows the distribution of $d\mathcal{J}/dR$ for a five-spot monitoring location, while Fig. 4 shows $d\mathcal{J}/dR$ when the monitoring locations are collinear near the extraction point. In the first case, the monitoring observations are sensitive to R mostly in the vicinity of their locations, and the domain is not very well covered. In the second case, however, high R sensitivity covers much more of the domain. The patterns will of course depend on **v**, and **v** depends on permeability, k, assumed known here. Clearly the location of monitors can strongly affect the value of data collected. This behaviour is seen in other disciplines. A singular vector approach has been used in meteorology in conjunction with the 4D-VAR method to find the optimal locations for collection observations by finding the directions where errors in the state vector will propagate most at the time of interest (Daescu & Navon 2004).

4 4D-VAR FOR WATER TRANSPORT IN A CONVECTIVE MEDIUM

Our final example examines the use of 4D-VAR as applied to water transport in a 2-D porous convective medium. The 4D-VAR method is an algorithm that uses adjoint sensitivity analysis to assimilate observations into a numerical model during the model run. The incorporation of observations at times throughout the model run leads to less reliance on the initial conditions and improvements in model forecasts. We have applied this technique to a case of water transport by convective flow in the Earth's mantle.

The first step in the 4D-VAR method is to define a cost function to measure the error between the forecast and the observations (Talagrand & Courtier 1987). Then the adjoint equations will be used to evaluate the gradient of the cost function obtained by applying a variational procedure to the Lagrangian of the problem Section 2.2 (Courtier & Talagrand 1990). The cost function and its gradient are given to a minimization algorithm (Burden & Faires 1993) in order to find the initial conditions that will give the optimal forecast. The 4D-VAR method has been applied to problems in both meteorology (Courtier & Talagrand 1990), hydrology (Bélanger & Vincent 2004), and convection (Hier Majumder *et al.* 2005). We use the following cost function:

$$\mathcal{J} = \frac{1}{2}\int_{t_1}^{t_2}\int_{\Omega}([H]-[H_{\rm obs}])^2\,d\mathbf{x}\,dt, \tag{38}$$

where $[H]$ is the predicted water concentration and $[H_{\rm obs}]$ is the observed water concentration. In general the cost function requires an interpolation operator to match the time and space locations of the observations with those of the forecast. In our case since we are dealing with twin experiments, we have neglected this operator in eq. (38). It is also usually necessary to use a weighting factor in the cost function to account for the varying reliability of different observations (Navon *et al.* 2005). In our case this was not necessary since we are dealing with only one observation. In general, a cost function is often also comprised of several other terms not considered in eq. (38), including a background error term, an observation error covariance term, and a term accounting for model error (Bunge *et al.* 2003).

In our case, we do not need to solve the full Euler–Lagrange system (eqs 24 and 25) since we can reformulate our problem as one without constraints (Talagrand & Courtier 1987). Since physical equations of the model are deterministic, it is evident that the state of the system at the time of observation depends only on the initial conditions, $\mathbf{p}_0$, of the system. Since the cost function is an implicit function of the initial conditions, varying the initial conditions will allow us to find a solution of the physical equations that minimizes the cost function (Ehrendorfer 1992). According to the theory of optimal control (Lions 1968), the control variables of the problem are the initial conditions. Therefore, in our problem we can remove the constraints since no restrictions are applied to the initial conditions. The gradient of the cost function (eq. 38) with respect to the initial concentration of water is given by the adjoint variables evaluated at time $\tau = t_2$ (Courtier & Talagrand 1990):

$$\nabla\mathcal{J}_{\mathbf{p}_0} = \lambda(x, z, \tau = t_2). \tag{39}$$

Our forward model for convection in the mantle is described by the dimensional incompressible Navier–Stokes equations in the infinite Prandtl approximation (Richard *et al.* 2002):

$$\nabla\cdot\mathbf{u} = 0, \tag{40}$$

$$\frac{\partial^4\Phi}{\partial x^4} + 2\frac{\partial^4\Phi}{\partial x^2 z^2} + \frac{\partial^4\Phi}{\partial z^4} = -\frac{\rho\alpha g}{\eta}\frac{\partial T}{\partial x}, \tag{41}$$

$$\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T = \nabla\kappa\cdot(\nabla T) + \Theta, \tag{42}$$

where **u** is the velocity; Φ is the stream function; ρ is the density; α is the coefficient of thermal expansivity; g is the gravitational acceleration; T is the temperature; η is the viscosity; κ is the thermal

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

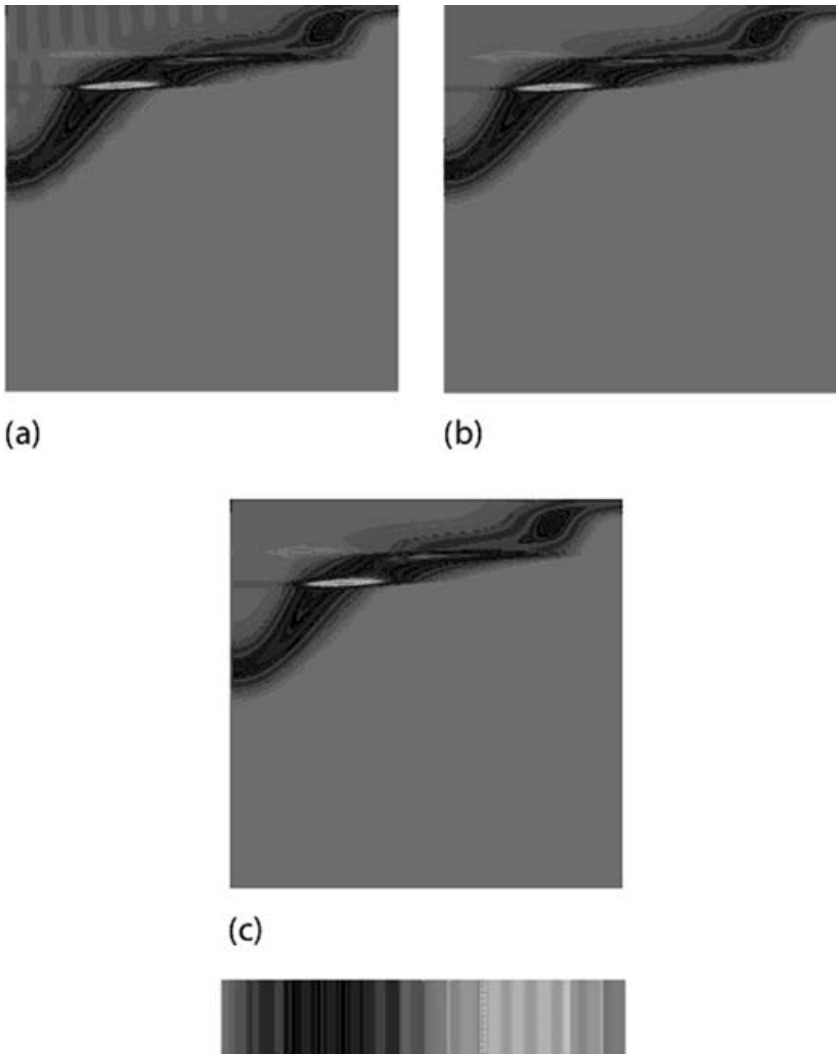

Figure 7. Comparison of forecasts of water concentration with a direct simulation and the 4D-VAR method for a small initial perturbation. (a) Prediction from direct simulation. (b) Observation. (c) 4D-VAR prediction. In the 4D-VAR forecast one can see that the undulations in the right-hand corner of the prediction have disappeared, and that the 4D-VAR method has reproduced the observations. Greyscale varies from minimum to maximum water concentration. The absolute values are not important since this is an artificial problem.

diffusivity; and Θ is the internal heat generation. The conservation of water content is given by Richard *et al.* (2002):

$$\frac{\partial [H]}{\partial t} + \mathbf{u} \cdot \nabla [H] = \kappa_H \nabla^2 [H] + \kappa_H \nabla \cdot \left([H] \nabla \frac{\mu}{RT} \right) + S, \quad (43)$$

where $[H]$ is the water concentration; κ_H is the water diffusivity; μ is the pure material chemical potential; R is the gas constant; and S is the source/sink term.

The adjoint equations corresponding to eqs (40)–(43) are found by differentiating (Gunzburger 2003):

$$u_x^* = [H] \frac{\partial [H]^*}{\partial x} + T \frac{\partial T^*}{\partial x} - \frac{\partial \mathrm{J}}{\partial u_x}, \quad (44)$$

$$u_z^* = [H] \frac{\partial [H]^*}{\partial z} + T \frac{\partial T^*}{\partial z} - \frac{\partial \mathrm{J}}{\partial u_z}, \quad (45)$$

$$\frac{\partial^4 \Phi^*}{\partial x^4} + 2 \frac{\partial^4 \Phi^*}{\partial x^2 z^2} + \frac{\partial^4 \Phi^*}{\partial z^4} = - \frac{\partial u_x^*}{\partial z} + \frac{\partial u_z^*}{\partial x} - \frac{\partial \mathrm{J}}{\partial \Phi}, \quad (46)$$

$$\frac{\partial T^*}{\partial \tau} = u_x \frac{\partial T^*}{\partial x} + u_z \frac{\partial T^*}{\partial z} + \kappa \nabla^2 T^* - \frac{\partial \kappa}{\partial x} \frac{\partial T^*}{\partial x} - \frac{\partial \kappa}{\partial z} \frac{\partial T^*}{\partial z} + \frac{\rho \alpha g}{\eta} \frac{\partial \Phi^*}{\partial x} - \frac{\mathrm{J}}{\partial T}, \quad (47)$$

$$\frac{\partial [H]^*}{\partial \tau} = u_x \frac{\partial [H]^*}{\partial x} + u_z \frac{\partial [H]^*}{\partial z} + \kappa_H \nabla^2 [H]^* - \kappa_H \left(\frac{\partial}{\partial z} \frac{\mu}{RT} \right) \frac{\partial [H]^*}{\partial z} - \frac{\partial \mathrm{J}}{\partial [H]}, \quad (48)$$

where * indicates the adjoint variable, and $\tau = t_2 - t$ is the inverse time. eqs (44)–(48) are then discretized. The initial conditions are:

$$\delta H \left(x, z, t|_{t_1} \right) = 0, \quad (49)$$

$$H^* \left(x, z, t|_{t_2} \right) = 0, \quad (50)$$

$$\delta T \left(x, z, t|_{t_1} \right) = 0, \quad (51)$$

$$T^* \left(x, z, t|_{t_2} \right) = 0. \quad (52)$$

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

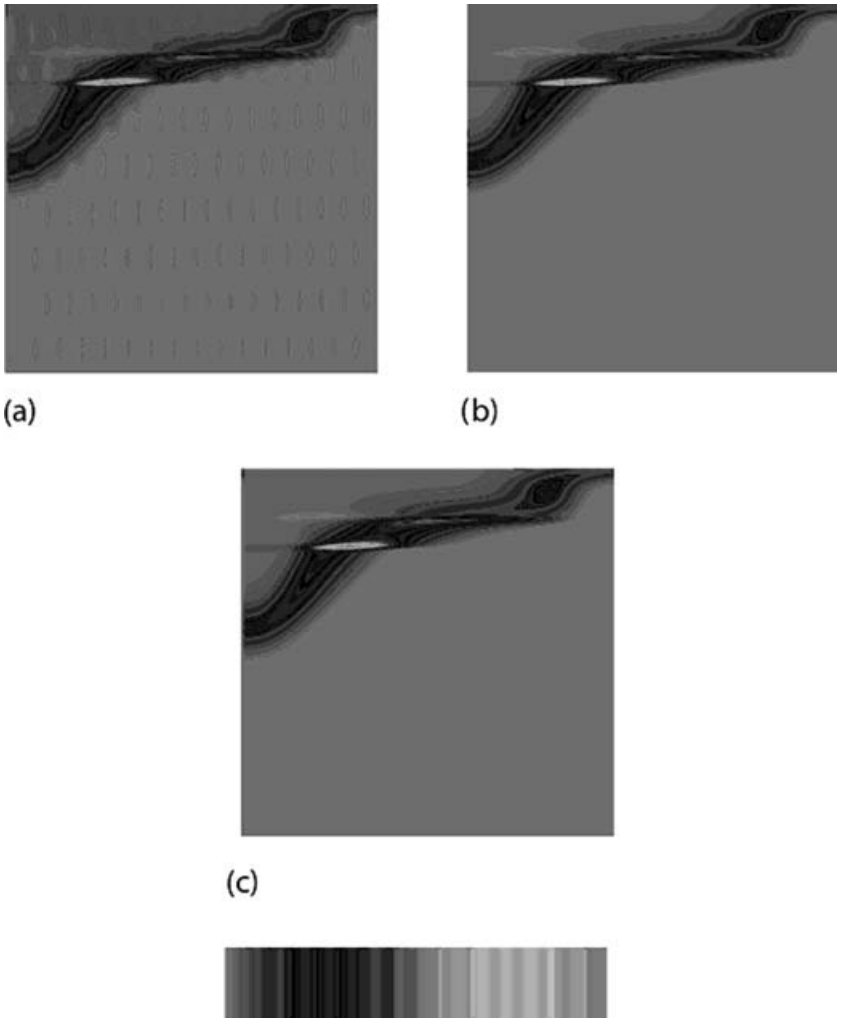

Figure 8. Comparison of the water concentration forecast between a direct simulation and the 4D-VAR method for a large initial perturbation. (a) Prediction from direct simulation. (b) Observation. (c) Prediction from 4D-VAR forecast. In the 4D-VAR prediction the noise has disappeared and the structure has been recovered. Greyscale varies from minimum to maximum water concentration.

The boundary conditions are set so that the adjoint variables along with their first and second derivatives with respect to x and z are zero at all boundaries. The gradient of the cost function (eq. 38) with respect to the water concentration is given by the adjoint variables evaluated at $\tau = t_2$ (Courtier & Talagrand 1990):

$$\nabla \mathcal{J}_{\Phi_0} = \Phi^*(x, z, \tau = t_2), \tag{53}$$

$$\nabla \mathcal{J}_{\mathbf{u_0}} = \mathbf{u}^*(x, z, \tau = t_2), \tag{54}$$

$$\nabla \mathcal{J}_{H_0} = H^*(x, z, \tau = t_2), \tag{55}$$

$$\nabla \mathcal{J}_{T_0} = T^*(x, z, \tau = t_2). \tag{56}$$

The algorithm used in conducting the 4D-VAR study is shown in Fig. 5. Since we do not have any experimental data to use as observations to assimilate into the forecast, we used a twin experiment set-up. We conducted an initial direct simulation that we considered as our observations. We then perturbed the initial conditions of this direct simulation with a sinusoidal function. A direct simulation with the perturbed initial conditions allowed us to obtain a first forecast. Using the observations obtained previously, we calculated the error between the initial forecast and the observations. Then we minimized the cost function (eq. 38) using its gradient (eqs 53–56) and the steepest descent method. We chose the steepest descent method for simplicity but use of the conjugate gradient would have been potentially more efficient (Zou *et al.* 1993; Wang *et al.* 1995). After the minimum of the cost function is found, which is defined as occurring when the norm of its gradient reaches a value of less than 10^{-8}, we have obtained the optimal initial conditions that provide an optimal forecast when used in a second direct simulation. The final error between the forecast and the observations is then minimal.

We present the results obtained for two cases, a small and large perturbation of the initial conditions. The small perturbation is on the order of 1 per cent for the temperature and 19 per cent for the water concentration. The large perturbation is on the order of 9 per cent for the temperature and 39 per cent for the water concentration. The forecast time is 5 million yr.

Fig. 6 shows the value of the cost function with respect to each iteration during the course of minimization with the steepest descent method. A large number of iterations are required to obtain convergence. Nevertheless, the cost function diminishes rapidly during the first few iterations. We have used a logarithmic scale since the convergence curves resemble exponential curves. Fig. 6 shows that the

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

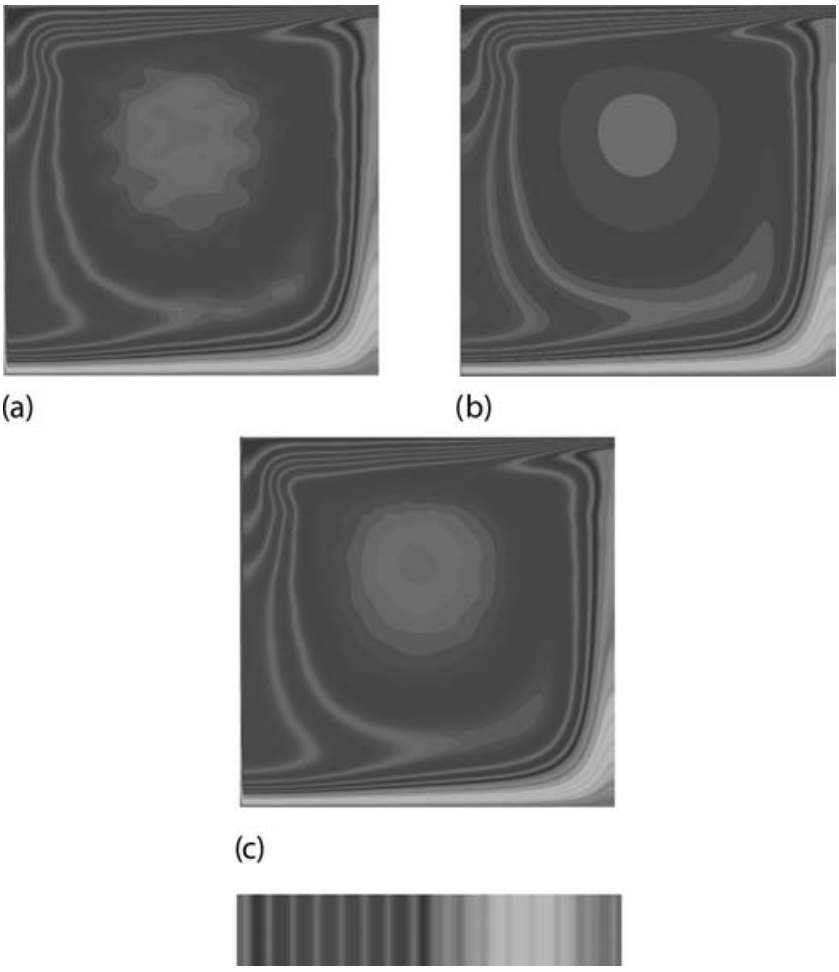

Figure 9. Comparison of temperature forecast between a direct simulation and the 4D-VAR method with a small perturbation in the initial conditions. (a) Forecast from direct simulation. (b) Observations. (c) 4D-VAR forecast. In the 4D-VAR forecast the undulations have disappeared in the centre of the structure, but there is still some blurring. The principal vortex structure, however, has been retrieved. Greyscale varies from minimum to maximum temperature. The absolute values are not important since this is an artificial problem.

larger the initial perturbation, the more difficult it is to minimize the cost function. With a small perturbation (Fig. 6a), the cost function reaches a value of less than 1; whereas with the larger perturbation (Fig. 6b) the convergence stabilizes around 10. In both cases, the cost function diminishes several orders of magnitude.

We now bring our attention to the moment when the water has already begun to pass from the upper to lower mantle via the transition zone. There is a large concentration of water at the boundary between the transition zone and the lower mantle (Figs 7 and 8). When the initial conditions are perturbed slightly, the 4D-VAR prediction succeeds in correcting the noise in the prediction (Fig. 7). Even with a larger initial perturbation, the 4D-VAR method reproduces the observations (Fig. 8). In studying the temperature distribution for the two preceding cases, we see that the temperature forecast is improved with the 4D-VAR method (Figs 9 and 10). This improvement occurs despite the fact that only the water concentration is considered in the error function (eq. 38). This is explained because the coupling of the physical equations (eqs 40–43) forces a correction in the initial temperature in order to obtain better results in the water concentration forecast. Fig. 9 demonstrates that the 4D-VAR method can retrieve the correct vortex for the convection cell even with undulations caused by a small perturbation in the initial conditions. However, the forecast is still blurred. The effect of the initial perturbation is most evident in Fig. 10 where there is significant noise in the prediction. Despite this, the 4D-VAR method succeeded in reconstructing the principal characteristics of the convection cell.

The results obtained for the two cases considered indicate that the 4D-VAR method is capable of correcting erroneous initial conditions and producing improved forecasts. Since the minimization required a large number of iterations, it is necessary to use the least number of grid points possible in order to produce results without an unreasonable delay for forecasts of longer time intervals. The 4D-VAR runs took about four times as long to compute and required twice as much memory as the DNS runs. It is possible to stop the convergence process before the completion of minimization. This will improve the calculation time, but the forecast will not be optimal. In future work the replacement of the steepest descent method with a conjugate gradient or quasi-Newton method will also reduce the convergence time. Use of pre-conditioning, limited memory BFGS (L-BFGS) or truncated Newton methods will also speed up convergence rate (Zou *et al.* 1993; Li *et al.* 2003). Scaling of the cost function, finding the condition number of the Hessian with respect to the control variables, and regularizing the problem using a Tikhonov regularization method can also improve the convergence. It is also possible to improve predictions by changing the cost

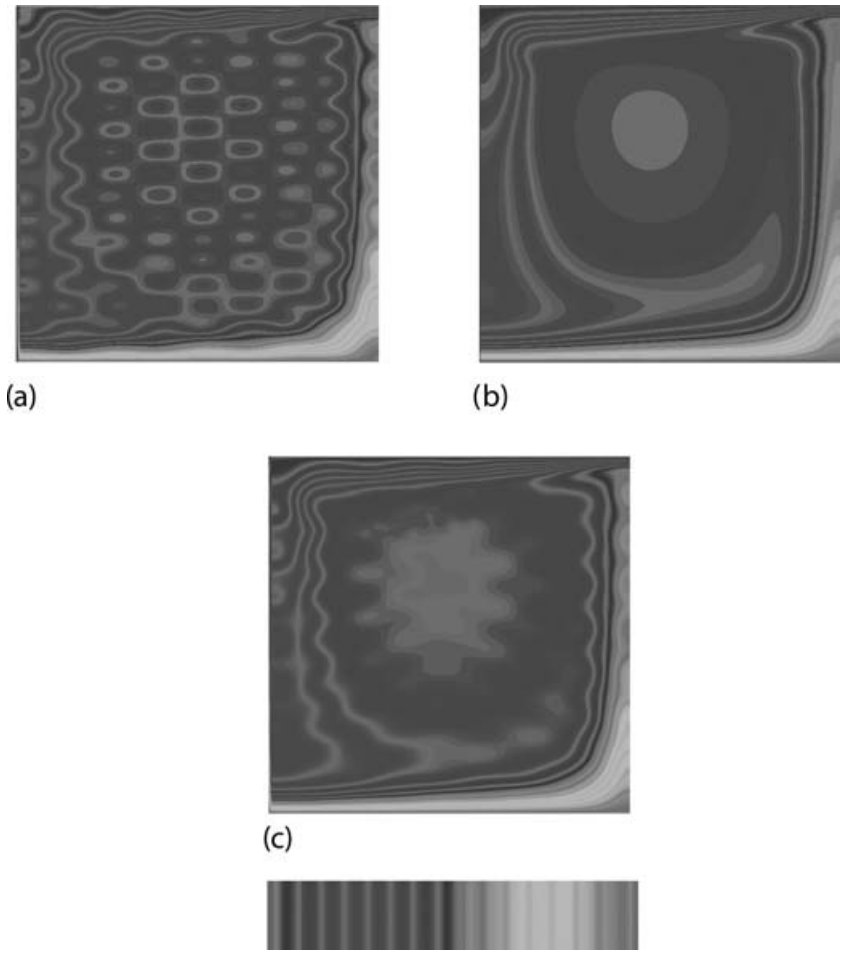

Figure 10. Comparison of temperature forecast between a direct simulation and the 4D-VAR method with a large perturbation in the initial conditions. (a) Forecast from direct simulation. (b) Observations. (c) 4D-VAR forecast. The direct simulation forecast is significantly affected by the perturbation and is very noisy. In the 4D-VAR forecast the noise has been reduced and the gross structure of the principal forecast has been recovered. Greyscale varies from minimum to maximum temperature. The absolute values are not important since this is an artificial problem.

function. For example the error in temperature could also be considered in the cost function.

5 CONCLUSIONS

In this paper we demonstrated the use of the adjoint sensitivity method for complementing numerical models in the environmental and geophysical sciences. The adjoint method provided a simple and efficient means of calculating the sensitivity coefficients of a model. Several cases were presented showing the use of the method for modelling water and contaminant transport in the Earth's crust and mantle.

We demonstrated the use of the adjoint sensitivity analysis in models of contaminant transport through porous media. The first example consisted of a faulted rock sequence into which a contaminant was injected. The contaminant concentration was then measured at an outflow region. The adjoint sensitivity method was used to determine the regions in space where the model is most sensitive to the permeability function. The adjoint method allowed us to quantify these sensitivities in a very efficient manner, requiring the computational burden equal to approximately twice the computer time needed for a single forward simulation.

The second example of the use of the adjoint method in contaminant transport models addresses the best placement of monitoring locations in a simulation that seeks to determine spatial distribution of retardation of contaminant from measurements of concentration at a few selected monitoring points. We found that the location of monitoring points can strongly affect the quality of the data collected. When the monitoring points were placed according to a conventional five-spot spacing in the domain, they were only sensitive to retardation in their immediate vicinity. However, when the monitoring points were placed collinearly near the extraction point, they provided more extensive sensitivity coverage to retardation throughout the domain. This example shows how using the adjoint method can greatly improve the quality of data collected by indicating the optimal placement of monitors.

The final example looked at the use of the adjoint method to assimilate data into a model for water transport in the convecting mantle. The method was used to incorporate observations into the model throughout the model run. This allowed us to obtain accurate forecasts of the water concentration, temperature, and velocity distributions in the mantle even with an imperfect knowledge of the initial conditions. The development of models that do not rely too heavily on the initial conditions is important in many fields in the geosciences as it is impossible to ever have a completely accurate

© 2006 The Authors, *GJI*, **166,** 907–922
Journal compilation © 2006 RAS

knowledge of parameters such as initial velocities and temperatures in a natural system. The adjoint approach can also be used to help better estimate the possible initial conditions when the final state is known, such as has been done in mantle convection (Bunge *et al.* 2003).

ACKNOWLEDGMENTS

We thank Monica D. Christiansen for her help with conducting simulations. CAH-M was supported by NASA and the Carnegie Institution of Washington. She was also supported during parts of this study by the LLNL post-doc program. BJT and DAY acknowledge the support of DOE under the grant 'Fluid Flow in the Earth's Crust'. DAY was also supported by NSF. APV thanks NSERC of Canada for support, and Guillaume Richard thanks CNRS of France. We express our sincere thanks to I. M. Navon for an extremely thorough and helpful review. We also thank an additional anonymous reviewer.

REFERENCES

Altmann-Dieses, A.E., Schlöder, J.P. & Bork, H.G., 2002, Optimal experimental design for parameter estimation in column outflow experiments, *Water Resour. Res.,* **38,** doi:10.1029/2001WR000358.

Bélanger, E. & Vincent, A.P., 2004, Data assimilation (4D-VAR) to forecast flood in shallow waters with sediment erosion, *J. Hydrol.,* **300,** 114–125.

Bélanger, E., Vincent, A.P. & Fortin, A., 2002, Data assimilation (4D-VAR) for shallow-water flow: the case of the Chicoutimi River, *Visual Geosciences,* **8,** doi:10.1007/s10069-003-0009-7.

Bennett, A.F., 1992. *Inverse Methods in Physical Oceanography,* Cambridge Univ. Press, Cambridge.

Bennett, A.F., 2002. *Inverse Modeling of the Ocean and Atmosphere,* Cambridge Univ. Press, Cambridge.

Bischof, C. & Griewank, A., 1992, ADIFOR: A FORTRAN System for portable automatic differentiation, in *Proc 4th Symp on Multidisciplinary Analysis and Optimization,* AIAA Paper 92-4744, pp. 433–441.

Bunge, H.-P., Hagelberg, C.R. & Travis, B.J., 2003, Mantle circulation models with variational data assimilation: inferring past mantle flow and structure from plate motion histories and seismic tomography, *Geophys. J. Int.,* **152,** 280–301.

Burden R. L. & Faires J.D., 1993. *Numerical Analysis,* PWS Publishing Company, Boston.

Cacuci, D.G., 1981a. Sensitivity theory for nonlinear systems. I. Nonlinear function analysis approach, *J. math. Phys.,* **12,** 2794–2802.

Cacuci, D.G., 1981b. Sensitivity theory for nonlinear systems. II. Extensions to additional classes of responses, *J. math. Phys.,* **22,** 2803–2812.

Cacuci, D.G., 1988. The forward and adjoint methods of sensitivity analysis, in *Uncertainty Analysis,* pp. 71–144, ed. Ronen Y., CRC Press, Boca Raton, Florida.

Cacuci, D.G., 1990. Global optimization and sensitivity analysis, *Nuc. Sci. Eng.,* **104,** 78–88.

Cacuci, D.G., 2003. *Sensitivity and Uncertainty Analysis. Volume I: Theory,* Chapman & Hall, Boca Raton, Florida.

Cacuci, D.G. & Hall, M.C. G., 1984. Efficient estimation of feedback effects with application to climate models, *J. atm. Sci.,* **41,** 2063–2068.

Cacuci, D.G., Ionescu-Bujor, M. & Navon, I.M., 2005. *Sensitivity and Uncertainty Analysis. Volume II: Applications to Large-Scale Systems,* Chapman & Hall, Boca Raton, Florida.

Cacuci, D.G., Weber, C.F., Oblow, E.F. & Marable, J.H., 1980. Sensitivity theory for general systems of nonlinear equations, *Nucl. Sci. Eng.,* **75,** 88–110.

Cao, Y. & Petzold, L., 2004. A posteriori error estimation and global error control for ordinary differential equations by the adjoint method, *SIAM J. sci. Comput.,* **26,** 359–374.

Chavent, G., Duprey, M. & Lemornier, P., 1975. History matching by use of optimal control theory, *Soc. Petrol. Eng. J.,* **15,** 74–86.

Christiaens, K. & Feyen, J., 2002. Use of sensitivity and uncertainty measures in distributed hydrological modeling with an application to the MIKE SHE model, *Water Resour. Res.,* **38,** doi:10.1029/2001WR000478.

Cohen, R.E., ed., 2005. *High-Performance Computing Requirements for the Computational Solid Earth Sciences,* http://www.geoprose.com/computational_SES.html.

Courtier, P. & Talagrand, O., 1990. Variational assimilation of meteorological observation with the direct and adjoint shallow water equations, *Tellus,* **42A,** 531–549.

Daescu, D. & Navon, I.M., 2004. Adaptive observations in the context of 4D-Var, *Meterol. atmos. Phys.,* **85,** 205–226.

Daescu, D., Carmichael, G.R. & Sandu, A., 2000. Adjoint implementation of Rosenbrook methods applied to variational data assimilation problems, *J. comp. Phys.,* **165,** 496–510.

Daley, R., 1991. *Atmospheric Data Analysis,* Cambridge Univ. Press, Cambridge.

de Jongh, D.C.J., 1997. Structural parameter sensitivity of the 'limits to growth' world model, *Appl. math. Model.,* **2,** 77–80.

Deif, A.S., 1986. *Sensitivity Analysis in Linear Systems,* Springer-Verlag, Berlin.

Delay, F. & Porel, G., 2003. Inverse modeling in the time domain for solving diffusion in a heterogeneous rock matrix, *Geophys. Res. Lett.,* **30,** doi:10.1029/2002GL016428.

Ehrendorfer, M., 1992. Four-dimensional data assimilation: Comparison of variational and sequential algorithms, *Q.J. R. meteor. Soc.,* **118,** 673–713.

Erofeeva, S.Y., Egbert, G.D. & Kosro, P.M., 2003. Tidal currents on the central Oregon shelf: Models, data, and assimilation, *J. geophys. Res.,* **108,** doi:10.1029/2002JC001615.

Errico, R., 1997. What is an adjoint model?, *Bull. Amer. meteor. Soc.,* **78,** 2577–2591.

Errico, R.M. & Vukicević, T., 1992. Sensitivity analysis using an adjoint of the PSU-NCAR mesoscale model, *Mon. Weather Rev.,* **120,** 1644–1660.

Francos, A., Elorza, F.J., Bouraoui, F., Bidoglio, G. & Galbiati, L., 2003. Sensitivity analysis of distributed environmental simulation models: Understanding the model behavior in hydrological studies at the catchment scale, *Reliability Eng. &Sys. Safety,* **79,** 205–218.

Gandini, A., 1988. Uncertainty analysis and experimental data transposition methods based on perturbation theory, in *Uncertainty Analysis,* pp. 217–258, ed. Ronen Y., CRC Press, Boca Raton, Florida.

Giering, R. & Kaminski, T., 1988. Recipes for adjoint code construction, *ACM Trans. Math. Soft.,* **24,** 437–474.

Gravilliers, L., Cruz, J. & Barahona, P., 2004. Parameter estimation using interval computations, *SIAM J. sci. Comput.,* **26,** 591–612.

Greenspan, E., 1975. Generalized perturbation theory and variational principle for multiple rations of linear and bilinear functionals, *Nuc. Sci. Eng.,* **56,** 107–112.

Griewank, A., 2000. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation,* SIAM, Philadelphia.

Gunzburger, M.D., 2003. *Prespectives in Flow Control and Optimization,* SIAM, Philadelphia.

Hagelberg, C.R. & Travis, B.J., 1997. First and second order adjoint analysis of tracer transport in a fracture with sorption, *Los Alamos Natl. Lab. Rep.,* LA-UR-97-1295.

Hall, M.C. & Cacuci, D.G., 1983. Physical interpretation of the adjoint function for sensitivity analysis of atmospheric models, *J. atm. Sci.,* **50,** 2537–2546.

Hansen, E., ed., 1969. *Topics in Interval Analysis,* Clarendon Press, Oxford.

Henryson, H., II, Hummel, H.H., Hwang, R.N., Stacey, W.M., Jr. & Toppel, B.J., 1974. Variational sensitivity analysis—theory and application, *Argonne Natl. Lab. Rep.,* CONF-740903-11.

Hess, P.G. & Vukicević, T., 2003. Intercontinental transport, chemical transformations, and baroclinic systems, *J. geophys. Res.,* **108,** doi:10.1029/2002JD002798.

Hier Majumder, C.A., Bélanger, E., DeRosier, S., Yuen, D.A. & Vincent, A.P., 2005. Data assimilation for plume models, *Nonlinear Processes in Geophysics,* **12,** 257–267.

Homescu, C. & Navon, I.M., 2003. Numerical and theoretical considerations for sensitivity calculations of discontinuous flow, *Sys. &Control Lett.,* **48,** 97–104.

Horwedel, J.E., 1991. GRESS version 2.0 user's manual, *Oak Ridge Natl. Lab. Rep.,* ORNL/TM-11591.

Horwedel, J.E., 1992. Reverse automatic differentiation of modular FROTRAN programs, *Oak Ridge Natl. Lab. Rep.,* ORNL/TM-12050.

Hughson, D.L. & Jim Yeh, T.C., 2000. An inverse model for three dimensional flow in variably saturated porous media, *Water Resour. Res.,* **36,** 829–839.

Hwang, J.-T., Dougherty, E.P., Rabitz, S. & Rabitz, H., 1978. The Green's function method of sensitivity analysis in chemical kinetics, *J. chem. Phys.,* **69,** 5180–5191.

Ismail-Zadeh, A., Schubert, G., Tsepelev, I. & Korotkii, A., 2004. Inverse problem of thermal convection: numerical approach and application to mantle plume restoration, *Phys. Earth planet. Inter.,* **145,** 99–114.

Jim Yeh, T.-C., Liu, S., Glass, R.J., Baker, K., Brainard, J.R., Alumbaugh, D. & LaBrecque, D., 2002. A geostatistically based inverse model for electrical resistivity surveys and its application to vadose zone hydrology, *Water Resour. Res.,* **38,** 10.1029/2001WR0010204.

Kalnay, E., 2003. *Atmospheric Modeling, Data Assimilation and Predictability,* Cambridge Univ. Press, Cambridge.

Kalnay, E., Park, S.K., Pu, Z.-X. & Gao, J., 2000. Application of the quasi-inverse method to data assimilation, *Mon. Weather Rev.,* **128,** 864–875.

Kaminski, T., Knorr, W., Rayner, P.J. & Heimann, M., 2003. Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, *Global Biogeochemical Cycles,* **16,** doi:10.1029/2001GB001463.

Katopodes, N.D. & Piasecki, M., 1996. Site and size optimization of contaminant sources in surface water systems, *J. environ. Eng.,* **122,** 917–923.

Knopman, D.S. & Voss, C.I., 1987. Behavior of sensitivities in the one-dimensional advection-dispersion equation: Implications for parameter estimation and sampling design, *Water Resour. Res.,* **23,** 253–272.

Knopman, D.S. & Voss, C.I., 1988. Further comments on sensitivities, parameter estimation, and sampling design in one-dimension analysis of solute transport in porous media, *Water Resour. Res.,* **24,** 225–238.

Kool, J.B., Parker, J.C. & Van Genuchten, M., 1987. Parameter estimation for unsaturated flow and transport models—a review, *J. Hydrol.,* **91,** 255–293.

Kunstmann, H., Kinzelbach, W. & Siegfried, T., 2002. Conditional first order second moment method and its application to the quantification of uncertainty in groundwater modeling, *Water Resour. Res.,* **38,** doi:10.1029/2000WR000022.

Le Dimet, F.-X. & Navon, I.M., 1988. Variational and optimization methods in meteorology: a review, *Supercomputer Computations Research Institute,* Florida State University, Technical Report, **144**.

Le Dimet, F.-X. & Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus,* **38A,** 97–110.

Le Dimet, F.-X., Ngodock, H.-E. & Luong, B., 1997. Sensitivity analysis in variational data assimilation, *J. meteor. Soc. Jap.,* **75,** 245–255.

Le Dimet, F.-X., Navon, I.M. & Daescu, D.N., 2002. Second-order information in data assimilation, *Mon. Weather Rev.,* **130,** 629–648.

Lee, D.W., Yambert, M.W. & Kocher, D.C., 1994. Uncertainty analysis for low level radioactive waste disposal performance assessment at Oak Ridge National Laboratory, *Oak Ridge Natl. Lab. Rep.,* CONF-940815-114.

Li, J.G. & Xing, L., 2000. Inverse planning incorporating organ motion, *Med. Phys.,* **27,** 573–578.

Li, R., Reynolds, A.C. & Oliver, D.S., 2003. Sensitivity coefficients for three-phase flow history matching *J. Can. pet. Technol.,* **42,** 70–76.

Li, Z. & Navon, I.M., 1998. Sensitivity analysis of outgoing radiation at the top of the atmosphere in the NCEP/MRF model, *J. geophys. Res.,* **13,** 3801–3814.

Li, Z. & Navon, I.M., 2001. Optimality of 4D-Var and its relationship with the Kalman filter and Kalman smoother, *Q.J. R. meteor. Soc.,* **127,** 661–684.

Li, Z., Navon, I.M., Hussaini, M.Y. & Le Dimet, F.-X., 2003. Optimal control of cylinder wakes via suction and blowing, *Comput. Fluid.,* **32,** 149–171.

Lions, J.L., 1968. *Contrôle optimale de systèmes gourvernés par des equations aux dérivées partielles,* Dumond, Paris.

Marchuk, G.I., 1995. *Adjoint Equations and Analysis of Complex Systems,* Kluwer Academic Press, Dordrecht.

McKay, M.D., 1988. Sensitivity and uncertainty analysis using a statistical sample of input values, in *Uncertainty Analysis,* pp. 145–186, ed. Ronen Y., CRC Press, Boca Raton, Florida.

Navon, I.M., 1998. Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, *Dyn. Atmos. Oceans,* **27,** 55–79.

Navon, I.M., Zou, X., Derber, J. & Sela, J., 1992. Variational data assimilation with an adiabatic version of the NMC spectral model, *Mon. Weather Rev.,* **120,** 1433–1446.

Navon, I.M., Daescu, D.N. & Liu, Z., 2005. The impact of background error on incomplete observations for 4D-Var data assimilation with the FSU GSM, in *ICCS 2005, LNCS 3515,* pp. 837–844, eds Sunderam V.S., van Albada, G.D., Sloot, P.M.A. & Dongarra, J.J., Springer-Verlag, Berlin.

Neumaier, A., 1990. *Interval Methods for Systems of Equations,* Cambridge Univ. Press, Cambridge.

Ngodock, H.E., 1996. Assimilation de données et Analyse de sensibilité: Une application à la circulation océanique, *Doctoral Thesis*, Université Joseph Fourier—Grenoble I, France.

Oblow, E.M. & Pin, F.G., 1986. Sensitivity analysis using computer calculus: A nuclear waste isolation application, *Nucl. Sci. Eng.,* **94,** 46.

Rabitz, H., 1981. Chemical sensitivity analysis theory with applications to molecular dynamics and kinetics, *Comp. &Chem.,* **5,** 167–180.

Reynolds, R., Eichler, P., Bade, A., Krisko, P. & Johnson, N., 1999. Sensitivity analysis of the orbital debris environment using the evolve 4.0 model, *Adv. Space Res.,* **23,** 175–185.

Richard, G., Monnereau, M. & Ingrin, J., 2002. Is the transition zone an empty water reservoir? Inferences from numerical model of mantle dynamics, *Earth planet. Sci. Lett.,* **205,** 37–51.

Ronen, Y., 1988a. The role of uncertainties, in *Uncertainty Analysis,* pp. 1–40, ed. Ronen, Y., CRC Press, Boca Raton, Florida.

Ronen, Y., 1988b. Uncertainty analysis based on functional analysis, in *Uncertainty Analysis,* pp. 41–69, ed. Ronen, Y., CRC Press, Boca Raton, Florida.

Sanders, B.F. & Katopodes, N.D., 1999. *J. irr. drain. Eng.* **125,** 287–297.

Sanders, B.F. & Katopodes, N.D., 2000. Adjoint sensitivity analysis for shallow-water wave control, *J. eng. Mech.,* **126,** 909–919.

Schröter, J., Seiler, U. & Wenzel, M., 1993. Variational assimilation of geostat data into a eddy-resolving model of the gulf stream extension area, *J. phys. Oceanogr.,* **23,** 925–953.

Scott, H.A., 1988. Sensitivity analysis of time-dependent non-LTE kinetics, *Lawrence Livermore Natl*. Lab. Rep., UCRL-99666.

Seierstad, A., 1982. Differentiability properties of the optimal value function in control theory, *J. Economic Dynamics &Control,* **4,** 303–310.

Stacey, W.M., Jr., 1974. *Variational Methods in Nuclear Reactor Physics,* Academic Press, New York.

Stauffer, F., Attinger, S., Zimmermann, S. & Kinzelbach, W., 2002. Uncertainty estimation of well catchments in heterogeneous aquifers, *Water Resour. Res.,* **38,** 10.1029/2001WR000819.

Sykes, J.F., Wilson, J.L. & Andrews, R.W., 1985. Sensitivity analysis for steady state groundwater flow using adjoint operator, *Water Resour. Res.,* **21,** 359–371.

Talagrand, O. & Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: theory, *Q.J.R. meteor. Soc.,* **113,** 1311–1328.

Tarantola, A., 1987. *Inverse Problem Theory,* Elsevier, Amsterdam.

Tomović, R., trans.Tornquist, D., 1963. *Sensitivity Analysis of Dynamical Systems,* McGraw-Hill Book Co., Inc., New York.

Tziperman, E. & Thacker, W.C., 1989. An optimal-control/adjoint-equations approach to studying the oceanic general circulation, *J. phys. Oceanogr.,* **19,** 1471–1485.

Van Noorden, T.L., Verduyn, S.M. & Bliek, A., 2003. Optimization of cyclically operated reactors and separators, *Chem. Eng. Sci.,* **58,** 4115–4127.

Vautard, R. *et al.*, 2003. A synthesis of the air pollution over the Paris region (ESQUIF) field campaign, *J. geophys. Res.,* **108,** doi:10.1029/2003JD003380.

Wang, Z., Navon, I.M., Zou, X., Le Dimet, F.X., 1995. A truncated Newton optimization Algorithm in meteorology applications with analytic Hessian/vector products, *Comput. optim. Appl.,* **4,** 241–262.

Whitecombe, J.M., Cropps, R.A., Braddock, R.D. & Agranovski, I.E., 2003. Application of sensitivity analysis to oil refinery emission, *Reliability Eng. &Sys. Safety,* **79,** 219–224.

Wigner, E.P., 1945. Effects of small perturbations on pile period, *Chicago Rep.,* CP-G 3048.

Wu, Z., Alpak, F.O. & Torres-Verdin, C., 2003. A quantitative study to assess the value of pressure data acquired with in-situ permanent sensors in complex 3D reservoir models subject to two-phase fluid flow, in *Proceedings–SPE Annual Technology Conference and Exhibition,* 2718–2798.

Yang, Z. & Hamrick, J.M., 2002. Variational inverse parameter estimation in a long term tidal transport model, *Water Resour. Res.,* **38,** doi:10.1029/2001WR001121.

Zou, X., Barcilon, A., Navon, I.M., Whitaker, J. & Cacuci, D.G., 1993. An adjoint sensitivity study of blocking in a two-layer isentropic model, *Mon. Weather Rev.,* **121,** 2833–2857.

Zhu, Y. & Navon, I.M., 1998. FSU-GSM forecast error sensitivity to initial conditions: applications to Indian summer monsoon, *Meteorol. atmos. Phys.,* **68,** 35–41.