Available online at www.sciencedirect.com
ScienceDirect
Physics and Chemistry of the Earth 32 (2007) 491–506
PHYSICS and CHEMISTRY of the EARTH
www.elsevier.com/locate/pce

Application of inverse modeling methods to thermal and diffusion experiments at Mont Terri Rock laboratory

A. Cartalade [a,*], P. Montarnal [a], M. Filippi [a], C. Mugler [b], M. Lamoureux [a], J.-M. Martinez [c], F. Clément [d], Y. Wileveau [e], D. Coelho [f], E. Tevissen [g]

[a] CEA/Saclay, DEN-DM2S-SFME-MTMS, 91191 Gif-sur-Yvette cedex, France
[b] CEA/Saclay, DSM-LSCE, 91191 Gif-sur-Yvette cedex, France
[c] CEA/Saclay, DEN-DM2S-SFME-LETR, 91191 Gif-sur-Yvette cedex, France
[d] INRIA/Rocquencourt, BP105, 78153 Le Chesnay cedex, France
[e] ANDRA, Meuse/Haute-Marne Laboratory, BP 3, 55290 Bure, France
[f] ANDRA, DS/TR, Parc de la croix blanche, 92298 Châtenay-Malabry, France
[g] CEA/Saclay, DEN-DPC-SECR-L3MR, 91191 Gif-sur-Yvette cedex, France

Received 19 April 2005; received in revised form 19 January 2006; accepted 19 August 2006
Available online 27 October 2006

Abstract

The Opalinus clay parameter identification at Mont Terri Underground Rock Laboratory (URL) is performed with inverse modeling. This article focuses on a comparison between the global and local inverse approaches, in a deterministic framework, applied to the thermal (HE-C) and diffusion (DI) experiments respectively. Each experiment presents a similar diffusion process described by the same equation. This simple forward model makes easier the comparison of each inverse approach and allows to easily apply different methods such as the neural networks and the adjoint state method. The synthesis yields to valuable information about their merits and flaws, and highlights the importance of the parametrization. The singular value decomposition of the Jacobian matrix is presented to identify the best parametrization. In each experiment, the numerically found parameter values are in good agreement with the experimental methods. In DI experiment a spatial variability description of the medium is a hypothesis to explain the rapid decrease of the tracer in the injection chamber.

Keywords: Mont Terri Underground Rock Laboratory; HE-C and DI experiments; Inverse modeling; Sensitivity; Optimization; Global and local approaches; Neuronal network; Adjoint state; Singular value decomposition; CAST3M; NeMo; Kalif

1. Introduction

Various countries are considering consolidated clay formations as suitable host rocks for the deep disposal of radioactive waste. The main purpose of experiments performed at the Mont Terri Underground Research Laboratory (Mont Terri Project – URL), in Switzerland, is to develop experimental tools and modeling methods to characterize the properties of the clay formations. The Mont Terri Laboratory is located in a tunnel and the porous medium is composed of Opalinus Clay.

Among all the experiments carried out, HE-C is dedicated to the characterization of the thermal behaviour of the rock, and DI is a diffusion tracer test. The HE-C experiment has consisted in measuring the time evolution of the rock temperature submitted to a heating source during 250 days in order to determine the thermal conductivity parameters of the Opalinus clay. The main components of the experiment (Fig. 2) were a heating source with a heater power regulation unit and several temperature sensors with the required data acquisition and control components. The heating source was first put in a long steel tube and then set

* Corresponding author. Fax: +33 169 085 242.
E-mail address: alain.cartalade@cea.fr (A. Cartalade).

1474-7065/$ - see front matter © 2006 Published by Elsevier Ltd.
doi:10.1016/j.pce.2006.08.043

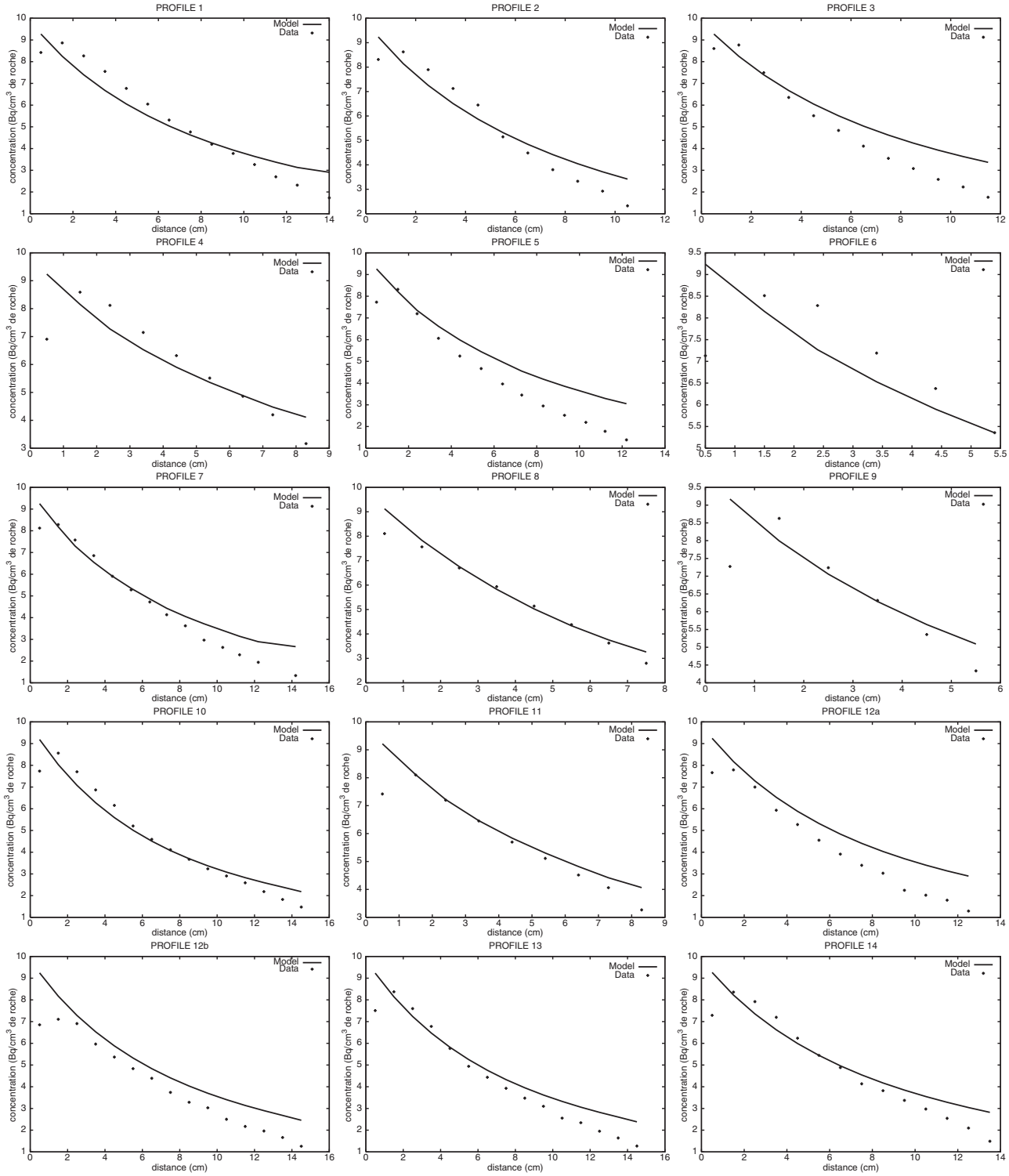

Fig. 1. Set of profiles for the 3D model DI (Montarnal et al., 2002).

up in a deep vertical borehole located close to the gallery wall. Temperature sensors were set up in the surrounding rock mass in two vertical boreholes. These two boreholes were drilled parallel to the main borehole containing the heating source. The design of the experiment and the first results are described in Wileveau (2002).

In DI experiment, tritiated water (HTO) and stable iodine ($^{127}I^{-}$) were injected in a borehole between two packers (3). The tracer concentration was weakly measured and several readjustments of concentrations in the test interval were necessary. Concentration has been monitored in the injection system for 324 days. An over-coring of this

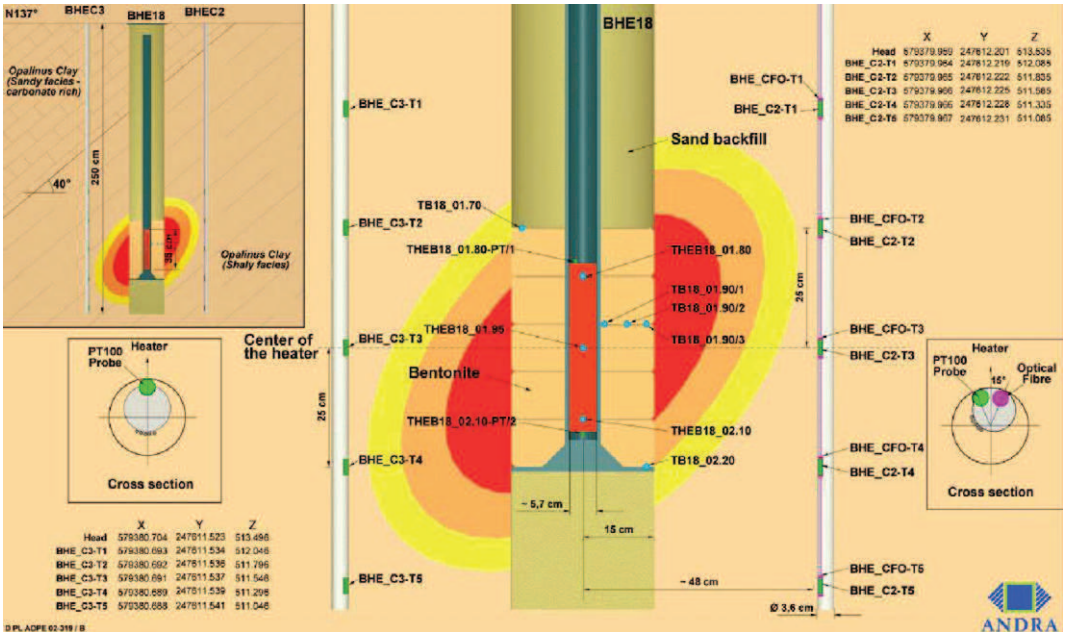

Fig. 2. Design of the HE-C experiment (Wileveau, 2002).

diffusion experiment has been performed at the end of the experiment to recover the whole rock volume around the experimental area. The aim was to sample it, to measure several concentration profiles of HTO and I in different directions. These profiles were used as the basis for modeling an in-situ diffusion tensor for HTO and I. The field work and the monitored data are presented in Fierz (1999). A description of the sampling method and procedure carried out in the over-core and the concentration profile extractions can be found in Möri et al. (2000). A synthesis on the modeling of DI experiment is presented in Tevissen and Soler (2003) (see Fig. 3).

Phenomena such as temperature and concentration propagation in the porous medium are described by a partial differential equation which links parameters to state variables. Initial and boundary conditions being given, we consider the forward problem as solved when the parameters are known and we wish to compute the measurable quantity such as the temperature or the concentration. However, in several problems, and particularly in

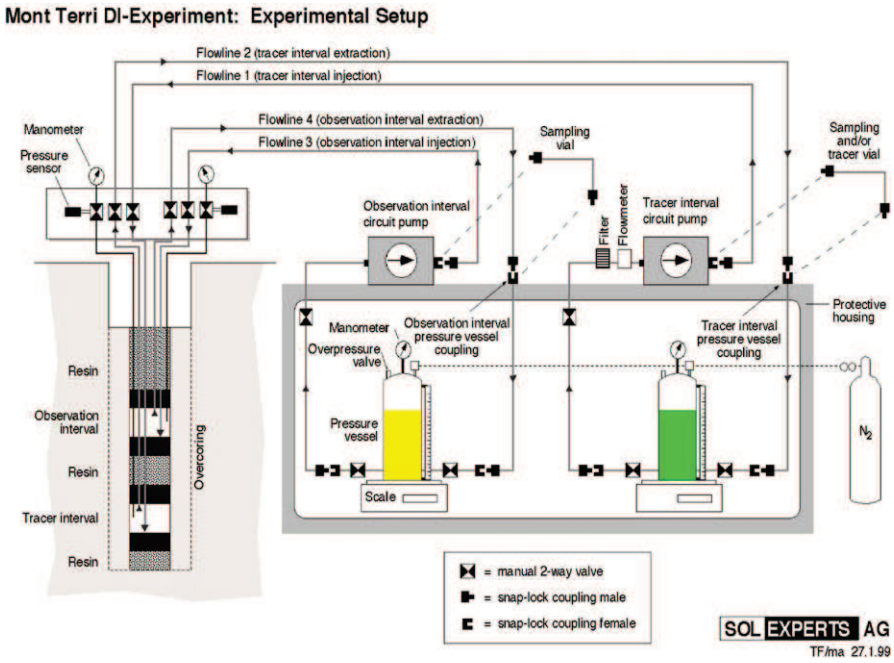

Fig. 3. Design of the DI experiment (Tevissen and Soler, 2003).

geosciences, it is easier to measure the state variables than to measure parameters. So in HE-C and DI experiments, the temperature and concentration values are known and we search to estimate the best parameter values to carry out reliable simulations. This is what we call an inverse problem.

Inverse methods are applied in several scientific area and a lot of works have been done and many books published on this subject. Let us cite Sun (1994) and de Marsily et al. (1999) for the groundwater flow and contaminant transport problems, and Özisik et al. (2000) for the inverse heat transfer problems. A lot of approaches are well documented for solving the inverse problem. They can be classified in two main frameworks: probabilistic or deterministic. In the probabilistic one, all the variables are seen as random functions. The solution of the inverse problem is to determine an *a posteriori* probability density function. All the variables of the problem (mean value, correlation,...) can be deduced from this function. A description of probabilistic inverse approach can be found in Tarantola (1987) and OFTA (1999). A comparison on seven geostatiscally based inverse methods is presented in Zimmerman et al. (1998). Another framework we use in this paper is the maximum likelihood estimator (Carrera and Neuman, 1986). The approach is inspired from the Bayesian approach but the interpretation is quite different because there is no underlying stochastic model. The solution of the inverse problem is a deterministic function, the most probable in relation to the data.

In spite of these two different frameworks to solve the inverse problem, the solution is usually found using an identical approach. A criterion is defined to measure the difference between computed and measured values and the aim is to find the set of parameters associated to the lowest criterion value. The methods used to find these «best parameters» can be local or global. In the local approach a starting point is given and a sequence of new parameters is generated from this point until the criterion is minimized. The stationary point found by the local method is a local minimum depending on initial values. To deal with the problem of local minima, we can use a global approach based on a great number of simulations using several values from the parameter space. This work focuses on a comparison between global and local approaches to find the lowest criterion value in the deterministic framework. Each of them has been applied on HE-C and DI experiments respectively. After a presentation of the forward models (Section 2), the main steps of each method are described (Section 3). A critical discussion on advantages and inconvenients is exposed in Section 4. The discussion will focus next on the parametrization and the sensitivity analysis using the Singular Value Decomposition tool.

HE-C and DI experiments are interesting to perform such a comparison because the physical situations are described by the same partial differential equation, well known and easy to solve. This simplicity for solving the forward problem will allow us to focus on each step of the global and local approaches. In spite of this simplicity, the equations exhibit a non-linearity between the state variables and the parameters and complicates the computation of the gradients. Modelings are performed on data measured with two different experimental protocols. As a consequence we will have to deal with different forward modelings and different inverse approaches. The description of each step will make the designation «inverse modeling» more comprehensible.

This work is a synthesis of several CEA and INRIA reports about the inverse modeling of these experiments. More details about the models and the numerical tools will be found in Filippi (2003) for the neural network building and the presentation of the global approach. The three following reports are dedicated to DI inverse modeling. A 3D local approach is described in Montarnal et al. (2002). A complete description of the adjoint state method applied in a one-dimensional case can be found in Cartalade et al. (2003) and the Singular Value Decomposition method is presented in Clément et al. (2004).

2. Forward model

This section presents the common characteristics of the two physical phenomena and the equation to be solved for each experiment. The same notations will be used in both modelings. Physical meanings of each term and hypothesis applied in HE-C and DI models will be described in next subsections. For more details about the numerical tools (number of meshes, convergence, physical values for each models and so on) the reader can refer to Mugler (2003) for the HE-C forward modeling using the Finite Element method, Montarnal and Lamoureux (2000) for a DI Mixed Hybrid Finite Element modeling and Cartalade et al. (2003) for the DI Finite Difference modeling.

2.1. General equation

DI and HE-C experiments were designed so that the main process is the diffusion of temperature or concentration in the porous medium. This process is described by a partial differential equation similar for the two phenomena. Two main properties of the medium, the anisotropy and the presence of a bedding plane, require a three-dimensional modeling. An energy (resp. mass) balance is carried out in a thermal problem (resp. diffusion). The equation to be solved is a parabolic-type equation, linear with respect to the state variable $F(\mathbf{x},t)$. It has the following form:

$$\Psi(\mathbf{x})\partial_t F(\mathbf{x},t) = -\vec{\nabla}\cdot\vec{q}(\mathbf{x},t) + S(\mathbf{x},t). \quad (1)$$

Integrated on a volume, Eq. (1) expresses the decrease in time of the total conservative quantity, $\Psi(\mathbf{x})F(\mathbf{x},t)$, which is equal to the outgoing flux through the surface surrounding the volume. In (1), $F(\mathbf{x},t)$ is the state variable depending upon position $\mathbf{x} = (x,y,z)^{\mathrm{T}}$ and time t [s]. The symbol ∂_t means the partial derivative with respect to the time,

$S(\mathbf{x},t)$ is a source term indicating a disappearance of the conservative quantity if it is negative, and an addition if it is positive. In HE-C and DI, this source term has two different forms and will be described more precisely below, $\Psi(\mathbf{x})$ is the first parameter of the equation. It is a function of space. The flux vector, $\vec{q}$, is given by the Fourier's law in the thermal problem and by the Fick's law in the diffusion problem. The mathematical form of these laws is identical. The flux is given by the product of a medium property $\bar{\bar{\chi}}(\mathbf{x})$ and the gradient of the state variable (2). The negative sign means that the flux direction is the opposite of the gradient direction

$$\vec{q}(\mathbf{x},t) = -\bar{\bar{\chi}}(\mathbf{x})\vec{\nabla}F(\mathbf{x},t). \tag{2}$$

In expression (2), $\bar{\bar{\chi}}(\mathbf{x})$ is a second order symmetric tensor to take into account the anisotropy of the medium. It is possible to establish this tensor form by using the transformation rules of the tensor components in a rotation of the cartesian coordinate system. Merging the Oy axis of the two coordinate systems this tensor takes the following form:

$$\bar{\bar{\chi}}(\mathbf{x}) = \begin{pmatrix} \chi_l \cos^2\theta + \chi_t \sin^2\theta & 0 & (\chi_l - \chi_t)\sin\theta\cos\theta \\ 0 & \chi_l & 0 \\ (\chi_l - \chi_t)\sin\theta\cos\theta & 0 & \chi_l \sin^2\theta + \chi_t \cos^2\theta \end{pmatrix}, \tag{3}$$

where θ is the angle between the axis Ox and Ox' (Fig. 4), $\chi_l(\mathbf{x})$ and $\chi_t(\mathbf{x})$ the longitudinal and transverse properties, respectively. $\chi_l(\mathbf{x})$ and $\chi_t(\mathbf{x})$ are the other parameters and also depend on the $\mathbf{x}$-position.

To ensure the uniqueness of the solution, Eq. (1) must be supplied with initial and boundary conditions. In the two problems considered in this paper, the initial condition is supposed to be constant on the whole domain Ω

$$F(\mathbf{x},0) = F_0. \tag{4}$$

Boundary conditions are the following:

$$\begin{cases} \text{boundary condition depending on the experiment on } \Gamma_1, \\ \Phi = 0 \quad \text{on } \Gamma_2, \end{cases} \tag{5}$$

the second boundary condition in relation (5), applied to the external surface of the system Γ_2, is null for the two problems. The surface Γ_1 and the boundary condition applied on it are different in the two models. They will be both described in the next subsections.

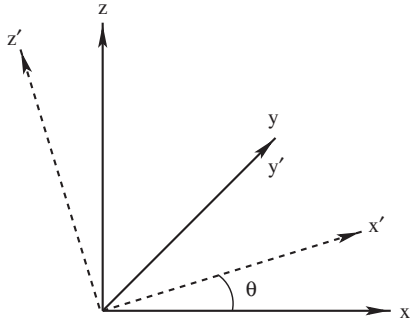

Fig. 4. Rotation of the cartesian coordinate system.

Let us finish this section by indicating an important property of diffusion problem. In isotropic and homogeneous media, it is possible to introduce a new quantity $\sqrt{\kappa t}$[m] where κ is the diffusivity [$m^2\,s^{-1}$] and t the time [s]. It is possible to show that the solution of the diffusion equation depends only on this quantity. It means that an initial perturbation propagates on a distance of the order $l = \sqrt{\kappa t}$ after a time t. The linear ratio between the mean diffusion distance and the time square root is an essential characteristic of a diffusion problem. Such a property can be represented by an adimensional number called the Fourier number and defined by the ratio between $\kappa\tau$ and L^2 where τ is a characteristic time and L a characteristic length. The Fourier number characterizes the propagation of the diffusion (heat or concentration) in the medium.

The general equation, similar for the two experiments, being introduced, let us emphasize now the main differences between HE-C and DI models.

2.2. Thermal model applied to the HE-C experiment

In the HE-C modeling, $F(\mathbf{x},t)$ is the temperature [K], $\Psi(\mathbf{x})$ is the product of the mass per volume unit ρ [$kg\,m^{-3}$] and specific heat C_p [$J\,kg^{-1}\,K^{-1}$] of the porous medium comprising water plus solid, $\bar{\bar{\chi}}(\mathbf{x})$ is the thermal conductivity tensor [$W\,K^{-1}\,s^{-1}$], and $S(\mathbf{x},t)$ is the power per volume unit given by the source.

The HE-C modeling supposes an homogeneous medium, so the parameters Ψ and $\bar{\bar{\chi}}$ are not space-dependent. According to this modeling, $S(\mathbf{x},t)$ is given by the product of a dimensionless coefficient describing a loss of power α and a time-dependent power per volume unit $Q(\mathbf{x},t)$. The equation has the following form:

$$\Psi\partial_t F(\mathbf{x},t) = \vec{\nabla}\cdot(\bar{\bar{\chi}}\vec{\nabla}F(\mathbf{x},t)) + \alpha Q(\mathbf{x},t). \tag{6}$$

The initial temperature is supposed to be the same in the whole domain and is equal to the mean value of the temperature measured by the various probes at time $t = 0$: $F_0 = 285.9$ K (12.756 °C). The first boundary condition is applied on the surface Γ_1 separating the porous medium and the tunnel gallery. Some measurements have shown the temperature gallery influence on the porous medium temperature, this is why the following boundary condition is used:

$$\Phi(\mathbf{x},t) = h(F(\mathbf{x},t) - F_g(t)) \quad \text{on } \Gamma_1, \tag{7}$$

with h the transfer coefficient [$W\,m^{-2}\,K^{-1}$] and $F_g(t)$ the time-dependent temperature in the gallery.

Eqs. (6) and (7) are solved numerically with Cast3M, a CEA code. A finite element method is used to discretize the spatial term and a Crank-Nicholson approximation for the temporal term.

2.3. Diffusion models applied to the DI experiment

In the diffusion problem, $F(\mathbf{x},t)$ is the concentration [mol m^{-3}], $\Psi(\mathbf{x})$ is the accessible porosity (dimensionless), $\bar{\bar{\chi}}(\mathbf{x})$ is the effective diffusion tensor [m^2 s^{-1}] and $S(\mathbf{x},t)$ expresses the tracer disappearance in the medium due to the radioactive decay of the tracer symbolized by the coefficient λ [s^{-1}]: $S(\mathbf{x},t) = -\lambda\Psi(\mathbf{x})F(\mathbf{x},t)$.

The boundary condition on Γ_1 is given by a mass balance performed in the injection chamber taking into account the radioactive decay:

$$F(\mathbf{x},t) = F(\mathbf{x},t_0) - \frac{1}{V}\int_{t_0}^{t}\int_{\Gamma_1}\vec{q}(\mathbf{x},s)\cdot\vec{n}\,\mathrm{d}\Gamma\,\mathrm{d}s - \lambda\int_{t_0}^{t}F(\mathbf{x},s)\,\mathrm{d}s, \tag{8}$$

where V is the total volume of the mixture water plus tracer in the whole system comprising the injection chamber, the pipes and the tanks. The concentration $F(\mathbf{x},t)$ at the time t is obtained from the initial concentration $F(\mathbf{x},t_0)$ minus the quantity disappeared in the porous medium (second term in the right hand side) and the quantity desappeared due to the radioactive decay (last term). Two different models have been applied for DI inverse modeling.

2.3.1. 3D modeling

The first one was performed in a three-dimensional homogeneous case. The equation to be solved is then similar to Eq. (6) with another expression for the source term:

$$\Psi\partial_t F(\mathbf{x},t) = \vec{\nabla}\cdot(\bar{\bar{\chi}}\vec{\nabla}F(\mathbf{x},t)) - \lambda\Psi F(\mathbf{x},t). \tag{9}$$

Computations were performed with Cast3M using the Mixed Hybrid Finite Element method (Dabbene, 1995). At the initial time the concentration is supposed to be zero in the whole domain ($F(\mathbf{x},0) = 0$). The boundary condition (8) is applied. A simulation result using a 40° bedding plane is presented in Fig. 5. We notice the bedding plane effect on the concentration field at the top and the bottom of the injection interval.

2.3.2. 1D modeling

Although the Mont Terri geometry requires a three-dimensional forward modeling because of the bedding plane presence, the second model applied was only one dimensional. This is because the inverse approach used for this problem requires to develop a new system which is much more easily and rapidly deduced and implemented on a one-dimensional model. The 1D approximation is however sufficient to introduce the parametrization concept and to show the improvement on the results brought by the zonation. The model is now one dimensional but takes into account the spatial variability of parameters $\Psi(r)$ and $\chi(r)$.

In a one-dimensional cylindrical problem the diffusion equation can be stated $\forall r \in [r_1, r_2]$:

$$\Psi(r)\partial_t F(r,t) = \frac{1}{r}\partial_r(r\chi(r)\partial_r F(r,t)) - \lambda\Psi(r)F(r,t), \tag{10}$$

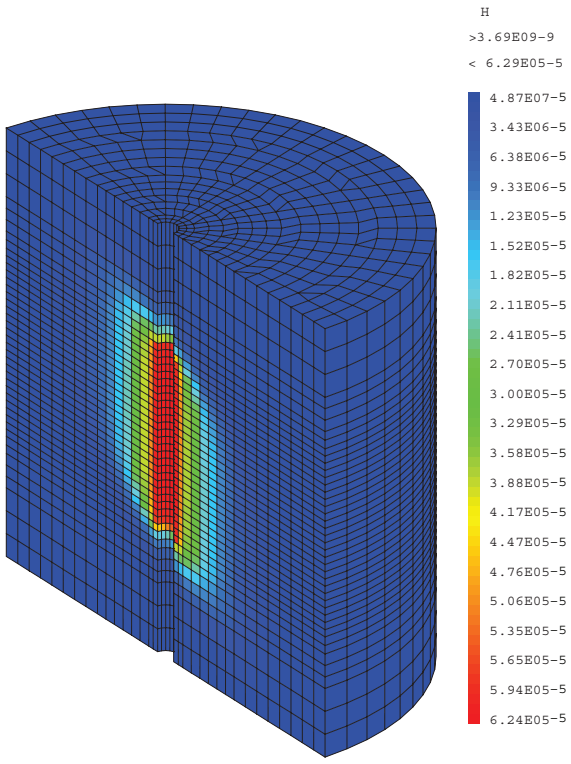

Fig. 5. Concentration field at the final time of the DI experiment.

with the null initial condition ($F(r,0) = 0$). The boundary condition on Γ_1 expressed in one dimension is given by the following relation:

$$F(r_1,t) = F(r_1,t_0) + \sigma\chi(r_1)\int_{t_0}^{t}\partial_r F(r_1,s)\,\mathrm{d}s - \lambda\int_{t_0}^{t}F(r_1,s)\,\mathrm{d}s, \tag{11}$$

where Γ_1 is the surface between the injection chamber in the borehole and the porous medium. In the Eq. (11), σ is the ratio between the injection chamber surface and the volume V. Eqs. (10) and (11) are solved with a finite difference scheme for the spatial term and an implicit scheme for the temporal term.

Let us mention that a semi-analytical solution in a one-dimensional homogeneous case, taking into account this boundary condition, can be found in Novakowski and ven der Kamp (1996). We did not use this analytical solution in this work in order to take into account the variability of parameters with the position.

3. Inverse models

Before beginning the description of the inverse approaches, let us summarize the parameters to be identified in each model. For HE-C, three scalar values have to be estimated: the longitudinal and transverse thermal conductivities χ_l and χ_t and the α coefficient. In the three-dimensional model for DI, three scalar parameters have also to be estimated: the effective porosity Ψ and the longi-

tudinal and transverse effective diffusion χ_l and χ_t. For the DI one dimensional model, two functions of the position have to be identified: $\Psi(r)$ and $\chi(r)$.

3.1. Parametrization

One of the practical and difficult aspects in inverse problems is related to the description of the spatial variability of the parameters $\Psi(r)$ and $\chi(r)$. The number of parameters to estimate depends upon the dimension of the problem and the parametrization chosen to approximate the functions $\Psi(r)$ and $\chi(r)$. The amount of data is often not sufficient to identify one parameter per mesh and the method used to decrease the number of parameters is called the parametrization of the problem. Several methods have been reviewed in Sun (1994) and an enlightening discussion on advantages and inconveniences on different parametrizations can be found in McLaughlin and Townley (1996). The parametrization presents two main difficulties: the parameter structure and its corresponding values inside it. In most of inverse modelings the parameter structure is supposed to be known and only the parameter values are identified. However the parameter structure is important because a wrong structure leads to a wrong parameter estimation (Sun and Yeh, 1985).

In this work, the parametrization used is the classical zonation method. The parameter structure is the number and the position of each zone. For example inside a system comprising N mesh points, the parameters $\overline{\Psi}_N = (\psi_1, \psi_2, \ldots, \psi_N)^{\mathrm{T}}$, $\bar{\chi}_N^l = (\chi_1^l, \chi_2^l, \ldots, \chi_N^l)^{\mathrm{T}}$ and $\bar{\chi}_N^t = (\chi_1^t, \chi_2^t, \ldots, \chi_N^t)^{\mathrm{T}}$ will be replaced by new vectors $\overline{\Psi}_M = (\psi_1, \psi_2, \ldots, \psi_M)^{\mathrm{T}}$, $\bar{\chi}_M^l = (\chi_1^l, \chi_2^l, \ldots, \chi_M^l)^{\mathrm{T}}$ and $\bar{\chi}_M^t = (\chi_1^t, \chi_2^t, \ldots, \chi_M^t)^{\mathrm{T}}$ with M very small compared to N. M is the number of zones, and each zone contains several mesh points.

In practice, the choice of M, called the dimension of parametrization, and the positions of each zone are difficult and depend upon the physical problem and the available data. In this work the parameter structure is supposed to be known, and only the parameter values inside them are identified. A solution to get the best parametrization using the measured data is given by the Singular Value Decomposition (SVD) method and will be presented in Section 4.

The importance of the zonation, number and position of each zone, will not be highlighted on the three-dimensional HE-C and DI models because they are considered homogeneous. So the vectors $\overline{\Psi}$, $\bar{\chi}^l$ and $\bar{\chi}^t$ are reduced to a scalar value ψ and χ_l and χ_t applied on the whole domain. The difficulties of parametrization will be more easily pointed out with the one dimensional model.

In all the rest of this paper the parameter to identify will be noticed by a vector $\mathbf{p}$. For HE-C modeling the components are the following: $\mathbf{p} = (\chi_l, \chi_t, \alpha)^{\mathrm{T}}$. For the three dimensional model of DI $\mathbf{p}$ expresses the vector $(\psi, \chi_l, \chi_t)^{\mathrm{T}}$ and for the one dimensional model $\mathbf{p} = (\overline{\Psi}_M, \bar{\chi}_M)^{\mathrm{T}}$. Although the DI forward model is one dimensional compared to the three dimensional models, the dimension of $\mathbf{p}$ could be greater if $M \geqslant 2$.

3.2. Global approach applied on the HE-C experiment

3.2.1. Objective function

Inverse modeling requires to define a criterion called an objective function (or performance function) measuring the difference between experimental and computed data. In HE-C and DI experiments, the objective functions are both based on the least squares but are slightly different due to the difference between the available data and the hypothesis made in each modeling. In the thermal problem, the objective function has the following form:

$$j = \frac{1}{\sum_{j=1}^{ns} \mu_j} \sum_{i=1}^{ns} \mu_i (F(\mathbf{x}_i) - \widehat{F}(\mathbf{x}_i))^2, \tag{12}$$

where ns is the number of sensors, $\widehat{F}(\mathbf{x}_i)$ is the ith measured temperature, $\widehat{F}(\mathbf{x}_i)$ is the ith computed temperature, $\mathbf{x}_i$, is the position. μ_i is a positive weight, different for each experimental value, describing the degree of confidence on this measure. We emphasize the fact that the objective function is not time-dependent. A steady state inverse problem has been solved in this study. Once the objective function is defined, the problem to solve is the following: find the three components of $\mathbf{p} = (\chi_l, \chi_t, \alpha)^{\mathrm{T}}$ to obtain the lowest value of J.

The global strategy to find $\mathbf{p}$ is simple and almost intuitive. Lower and upper bounds of each component are fixed and a great number of $\mathbf{p}_k$ $(k = 1, \ldots, N)$ are chosen inside this parameter space. The choice of a parameter $\mathbf{p}_{k+1}$ is independent from the previous one $\mathbf{p}_k$. A statistical law can be used to generate all the $\mathbf{p}_k$ in order to have an uniform repartition inside the parameter space. The objective function is then computed for each parameter $\mathbf{p}_k$ and when all simulations are achieved, the set of parameters giving the lowest cost function value is the solution of the problem.

3.2.2. Neural network

The large computation time required by the global approach leads us to build a Cast3M approximation using a neural network, especially in this application because the forward model is solved by a numerical method (3D finite element). For each parameter $\mathbf{p}_k$, the neural network computes a temperature according to the following equation:

$$F^s = \zeta_0^s + \sum_{i=1}^{N_h} \zeta_i^s f\left(\sum_{j=1}^{N_{in}} w_{ij}^s p_j + w_{i0}^s \right), \tag{13}$$

with F^s the computed temperature at the considered sensor s. The superscript s indicates that one neural network is built for each temperature sensor. Ten sensors were available in the experiment. ζ_i^s, w_{ij}^s and w_{i0}^s are some weights, N_h is the number of hidden neurones, p_j indicates the input parameter ($p_j = \chi_l, \chi_t, \alpha$), and N_{in} their number. $f(x) = \frac{1}{1+e^{-x}}$ is the logistic activation function.

The neural networks were built with NeMo (CEA software, Dreyfus et al. (2002)) from a sample training of 1100 Cast3M simulations. These simulations give some indications to choose the number of hidden neurones and for each of them to fit the weights ζ_i, w_{ij} and w_{i0}. A simple neural network includes one hidden layer made of three hidden neurones (6). The most simple neural network with no hidden layer has been tested but gave no satisfactory results. (Fig. 6)

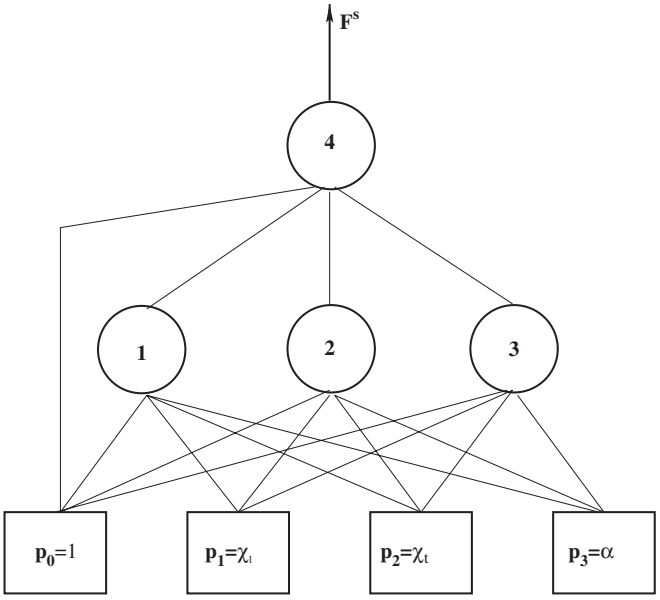

Fig. 6. Neural network with one hidden layer.

The very fast neural networks allowed to approximate the Cast3M code and to compute the objective function for 500,000 different values for the parameter $\mathbf{p} = (\chi_l, \chi_t, \alpha)^T$. 100 values of χ_l in the range $[1,4]$, 100 values of χ_t between $[0.5,3]$ and 50 values of α in the range $[0.7,1]$. This number of 500,000 instantaneous computations performed with the neural network has to be compared with the 1100 Cast3M preliminary direct simulations necessary to the training of the neural network.

3.2.3. Results for HE-C experiment

The values of the thermal conductivities were estimated separately using the probes located on each side of the heating source. The triplet of optimized values $\mathbf{p}$ is quite different on each side of the heating source. On one side, we obtain $\chi_l = 1.84 \pm 6\%$, $\chi_t = 055 \pm 9\%$ and $\alpha = 077 \pm 3\%$, although we obtain on the other side $\chi_l = 1.90 \pm 7\%$, $\chi_t = 1.07 \pm 11\%$ and $\alpha = 0.77 \pm 3\%$.

In order to check the validity of this method, some forward numerical Cast3M simulations were performed with various values of parameters $\mathbf{p}$. Fig. 7 displays the temperature evolution versus time measured with the five probes on each side of the heating source, and calculated with three value triplets on each side, equal to $\mathbf{p}_1^{C_2} = (1.727, 0.525, 0.718)^T$, $\mathbf{p}_2^{C_2} = (1.848, 0.550, 0.773)^T$ and $\mathbf{p}_3^{C_2} = (1.970, 0.601, 0.816)^T$ on one side (C_2 line) and equal to $\mathbf{p}_1^{C_3} = (1.788, 1.005, 0.718)^T$, $\mathbf{p}_2^{C_3} = (1.939, 1.081, 0.773)^T$ and $\mathbf{p}_3^{C_3} = (2.030, 1.131, 0.816)^T$ on the other side (C_3 line). Experimental and numerical results are in good agreement. The additional direct simulations prove that the neural net-

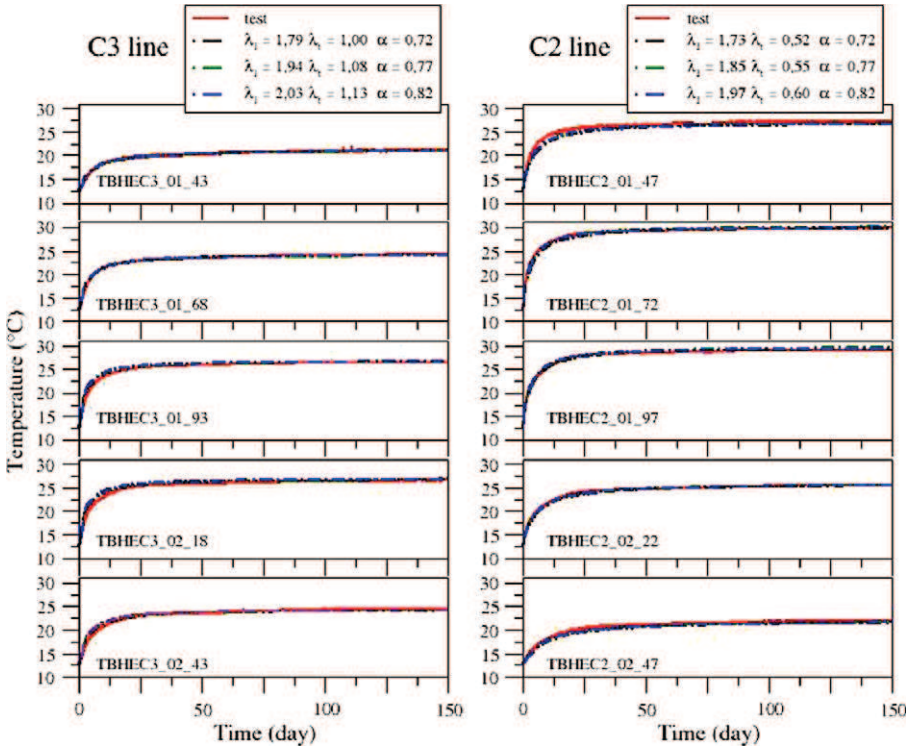

Fig. 7. Set of temperatures for the 3D model HE-C (Filippi, 2003).

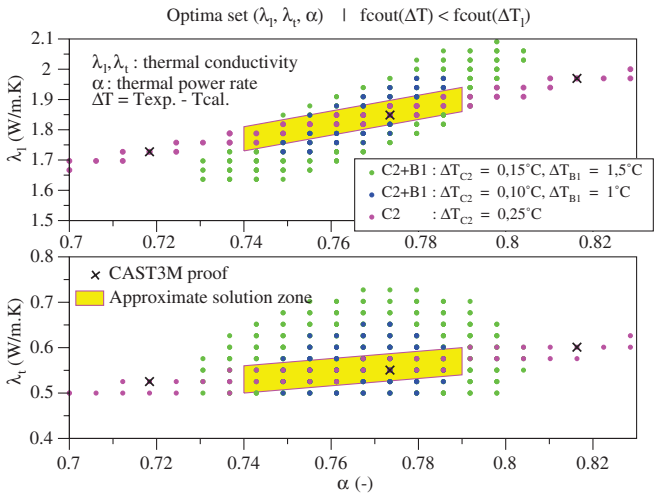

Fig. 8. Functional values in the parameter space (Filippi, 2003).

work correctly simulates Cast3M computations and that parameters **p** obtained from the global approach yield to the correct measured temperatures.

The functional values in the parameter space are presented in Fig. 8. The intervals of parameter values giving the lowest objective function is represented by the yellow zone.

3.3. Local approach applied on the DI experiment

3.3.1. Objective function

The inverse approach used in DI experiment is based on the same deterministic theory but two main differences exist between DI and HE-C and modify the inverse modeling. The first one concerns the experimental protocol and the data sampling. Two sets of concentration data are available in DI experiment. The concentrations $\widehat{F}(\mathbf{x}_i, t_f)$ measured at the final time t_f of the experiment and varying with the position $\mathbf{x}_i$ belong to the first set of data. The concentrations $\widehat{F}(\mathbf{x}_1, t_j)$ measured in the injection chamber $\mathbf{x}_1$ and varying with the time t_j belong to the second set of data. The relation (14) gives the objective function for the DI problem:

$$J = \sum_{i=1}^{I} \mu_i^2 (F(\mathbf{x}_i, t_f) - \widehat{F}(\mathbf{x}_i, t_f))^2 + \sum_{j=1}^{J} \beta_j^2 (F(\mathbf{x}_1, t_j) - \widehat{F}(\mathbf{x}_1, t_j))^2. \tag{14}$$

Two terms are embedded in the objective function, each one describing one set of experimental data. In the first term of the right hand side, the experimental concentrations are described by a function depending on the position, and in the second term the concentrations are described by a function depending on the time. I is the number of space data, J the number of temporal data, μ_i^2 and β_j^2 are positive weights.

3.3.2. Optimization

The minimization of criterion (14) is performed with an optimization algorithm called an optimizer. A lot of books were published on this subject: Fletcher (1987), Bonnans et al. (1997), and Nocedal and Wright (1999). Many optimization methods exist in the literature and it is out of range of this paper to describe all of them. We only mention the gradient algorithms, more robust and less time-consuming than the others. Let us mention that the optimization methods are found in many problems where a stationary point of an objective function is searched for. These methods are sometimes extended to the resolution of non-linear systems. For example, another application is given by the coupling between the transport equation and the chemistry (Bouillard et al., 2005).

The principle of these algorithms is simple. A search sequence $(\mathbf{p}_1, \ldots, \mathbf{p}_k, \ldots, \mathbf{p}_N)$ is generated from a starting point $\mathbf{p}_0$, such as for the iterate $k+1$, $J(\mathbf{p}_{k+1}) < J(\mathbf{p}_k)$, $\mathbf{p}_k$ is the parameter vector at iterate k. The search sequence between the iterate k and $k+1$ is not independent contrary to the global approach. The relation is given by $\mathbf{p}_{k+1} = \mathbf{p}_k + \xi_k \mathbf{d}_k$ where ξ_k is a positive scalar indicating the length of the step and $\mathbf{d}_k$ is the kth descent direction.

A first class of gradient methods are the conjugate gradient algorithms. The descent direction is given by $\mathbf{d}_k = -\mathbf{r}_k + \beta_k \mathbf{d}_{k-1}$ where $\mathbf{r}_k$ is the residual and the scalar β_k is to be determined by requirements that $\mathbf{d}_k$ and $\mathbf{d}_{k+1}$ must be conjugate with respect to the discretization matrix of the problem. Several forms of β_k depending upon $\nabla J_k = \nabla J(\mathbf{p}_k)$ exist and distinguish the methods. The Fletcher–Reeves and the Polak–Ribière (PR) methods are the most popular conjugate gradient algorithms. In practice, the PR method with automatic restarts out performs the other methods (Bonnans et al., 1997; Nocedal and Wright, 1999).

In a second class of gradient methods, $\mathbf{d}_k$ is given by the relation $\mathbf{d}_k = -\mathbf{B}_k^{-1} \nabla J_k$ with $\mathbf{B}_k$ a symmetrical matrix and ∇J_k the gradient of the objective function. In the steepest descent method, $\mathbf{B}_k$ is simply the identity matrix, whereas in the Newton method, $\mathbf{B}_k$ is exactly the Hessian matrix $\nabla^2 J_k$. In the quasi-Newton methods, $\mathbf{B}_k$ is an approximation of the Hessian matrix. The BFGS formula is one of the most popular Hessian approximation. The Byrd et al. algorithm (1995) was applied in DI inverse modeling. The algorithm uses a BFGS method and a limited memory to approximate directly the inverse Hessian matrix. Lower and upper bounds can also be specified to impose a positive constraints on parameters. Optimization with bound constraints is more difficult.

The Levenberg–Marquardt method is a very efficient algorithm but it requires the computation of the sensitivity matrix (or Jacobian matrix). It is possible to compute it for the one-dimensional model but it is difficult for two or three space dimensions. In Section 4 we will compute the Jacobian matrix but in few points, whereas the Levenberg–Marquardt optimization method needs to compute it for each iteration. A good review of all these algorithms (conjugate gradients, BFGS, Newton methods, Levenberg–Marquardt ...) with or without bound constraints, and a critical discussion on the advantages and inconvenients about the main

classes of many methods (trust-region and line search methods) can be found in Nocedal and Wright (1999).

3.3.3. Results for the 3D inverse model

A first application of the local approach with a BFGS algorithm has been performed on the 3D modeling using Cast3M. The gradients of the cost function were computed by the finite difference method and the optimization software used was Kalif (Martinez et al., 2002). We refer Tevissen et al. (2004) for a complete comparison between the *in situ* and laboratory diffusion studies from the Mont Terri. The laboratory measurements on the centimetric core samples are described. We present in this paper only the results for the HTO tracer in order to focus the discussion on the parametrization in the following of this paper.

The simulations performed have shown that the entire set of HTO data is consistent with due account to only a diffusive process in the rock (Palut et al., 2003). The 3D model explains the disappearance of the tracer from the injection system over the time (Fig. 9), and its distribution in the rock versus the distance from the borehole (Fig. 10).

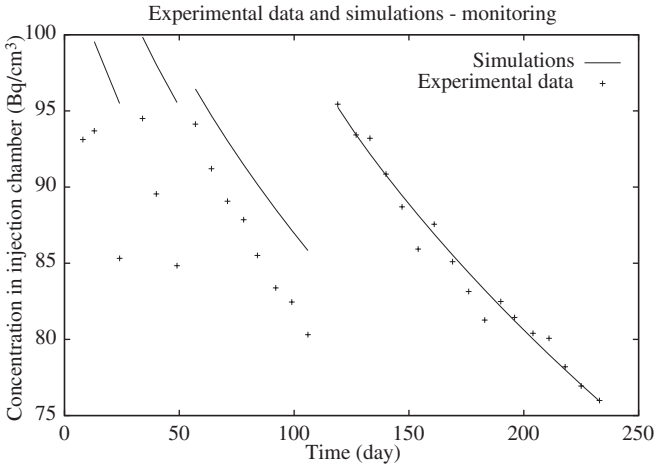

Fig. 9. Experimental and computed monitoring (Montarnal et al., 2002).

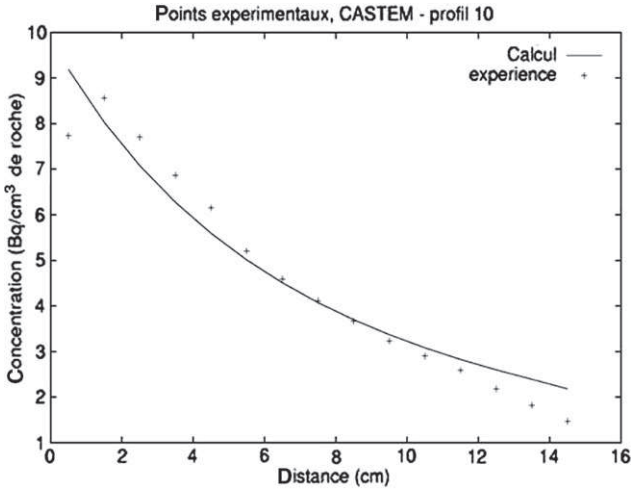

Fig. 10. Experimental and computed profile.

The best estimate values for HTO are $\chi_l = 18 \times 10^{-11}\ \mathrm{m^2\ s^{-1}}$, $\chi_t = 1.2 \times 10^{-11}\ \mathrm{m^2\ s^{-1}}$ and $\Psi = 12\%$. The appendix presents the comparisons between all the experimental profiles and the computed concentrations.

On the monitoring curve, only the first data are not well fitted. Those differences between the experimental data and the model, could be explained by the effect of a disturbed zone around the injection borehole where the porosity and diffusion coefficients are higher. As the tracers diffuse further into the rock, the diffusion slows down. The tracer disappearance in the injection borehole is more affected by the disturbed zone than profiles in the rock, because the tracer is mostly located near the borehole as illustrated by the shape of the profiles. This hypothesis will be tested in Section 3.3.5.

3.3.4. Adjoint state method

Another way to compute the gradients of the cost function, required by the BFGS algorithms, is to apply the adjoint state method. It is based on the optimal control theory (Lions, 1968) and was used in parameter identification problems by Chavent (1971, 1975). The method computes accurate gradients with respect to the parameters to identify. A complete description of the adjoint state method is presented in Carrera and Neuman (1986) and Sun (1994). The application is performed on a classical diffusion equation and a convection–dispersion equations.

In DI problem, the adjoint state is slightly different due to the boundary condition, which depends on the diffusion to identify and to the time. Another difference for the adjoint deduction comes from the objective function formulation. Let us remember that the cost function in this problem presents two terms, the first one for the monitoring data and the second one for the profile data. The adjoint system, deduced from the forward problem, has to be solved according to the following equations:

$$-\Psi(r)\partial_t \Upsilon(r,t) = \frac{1}{r}\partial_r(r\chi(r)\partial_r \Upsilon(r,t)) - \lambda\Psi(r)\Upsilon(r,t), \quad (15)$$

$$\Upsilon(r,t_f) = \sum_{i=1}^{I} \frac{2\mu_i^2}{\Psi r}(F(r_i,t_f) - \widehat{F}(r_i,t_f)), \quad (16)$$

$$\begin{cases} \frac{1}{\alpha}\partial_t \Upsilon + \chi\partial_r \Upsilon|_{r_1} = -\sum_{j=1}^{J} \frac{2\beta_j^2}{r_1}(F(r_1,t_j) - \widehat{F}(r_1,t_j)), \\ \chi\partial_r \Upsilon|_{r_2} = 0. \end{cases} \quad (17)$$

This system is adjoint-qualified because the operators in its definition are adjoint of state equation operators: $-\partial_t$ as adjoint of ∂_t and $\frac{1}{r}\partial_r(r\chi(r)\partial_r)$ is the same because it is self adjoint. $\gamma(r,t)$ is called the adjoint state of the concentration $F(r,t)$. A final time (and not initial time) is given in Eq. (16), consequently Eq. (15) must be solved backwards in time. Because data are measured at the end of the experiment (profile data) and in the injection chamber (monitoring data), differences between computed and observed concentrations are present in final condition and in the first boundary condition respectively.

When $F(r,t)$ and (r,t) are computed, the gradients of the objective function can be deduced from the following relations

$$\frac{\delta J}{\delta \chi} = -\int_0^{t_f}\int_{r_1}^{r_2} \partial_r F(r,t)\partial_r \Upsilon(r,t) r\,\mathrm{d}r\,\mathrm{d}t, \tag{18}$$

$$\frac{\delta J}{\delta \Psi} = -\int_0^{t_f}\int_{r_1}^{r_2} \Upsilon(r,t)(\partial_t F(r,t) + \lambda F(r,t)) r\,\mathrm{d}r\,\mathrm{d}t. \tag{19}$$

Two possibilities are available to deduce the discretized adjoint system. The first one is to discretize the relations (15)–(17) and (18), (19). Another way is to find the adjoint system with the variational method from the primary problem already discretized. This latter method is better as indicated in Chavent (1979) and has been applied in this work.

The success of the gradient method depends strongly on the possibility to accurately compute the gradients. The adjoint state method is based upon an analytical derivation of the gradients and allows to compute them efficiently. This is the main advantage of adjoint state method compared to the finite difference method. Indeed, in a comparison between the gradients computed by an adjoint state and a finite difference method several significant values of the steps required by the finite difference, must be adjusted to obtain the same values of the gradients. In this work, this comparison has been performed to verify the development of the relations (15)–(19).

3.3.5. Results for the 1D inverse model

The one-dimensional inverse model using the adjoint state method has been applied on the DI experimental data. Several inverse simulations have been performed using one, two and three zones respectively. The lowest values of the cost function have been obtained with three zones with two parameters per zone: one diffusion and one porosity coefficient. Six parameters were optimized. The inverse simulations performed with three zones gave several admissible solutions. Figs. 11 and 12 present one of them chosen from some physical considerations. As already mentioned in Section 3.3.3, the diffusion and porosity parameters could be higher near the injection borehole.

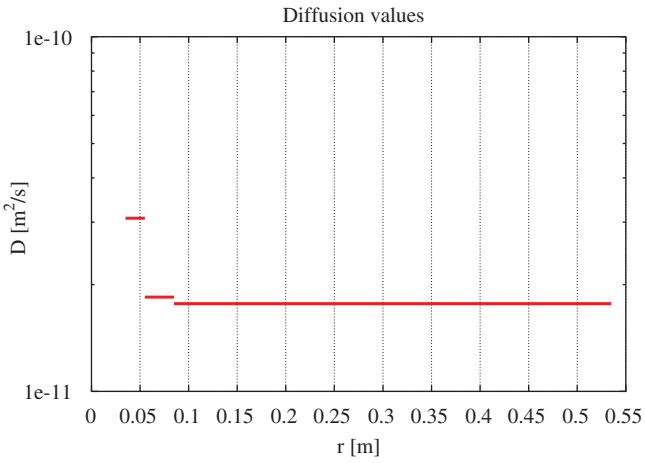

Fig. 11. Diffusion parameters estimated (Cartalade et al., 2003).

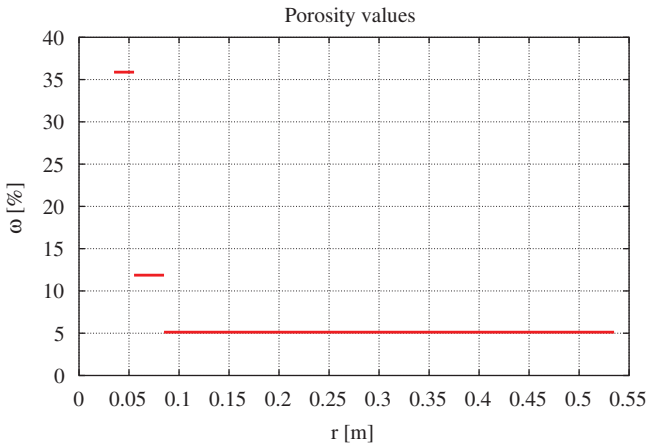

Fig. 12. Porosity parameters estimated.

The forward simulation results performed with these parameter values are presented in Figs. 13 and 14. The improvements brought by the spatial variability of

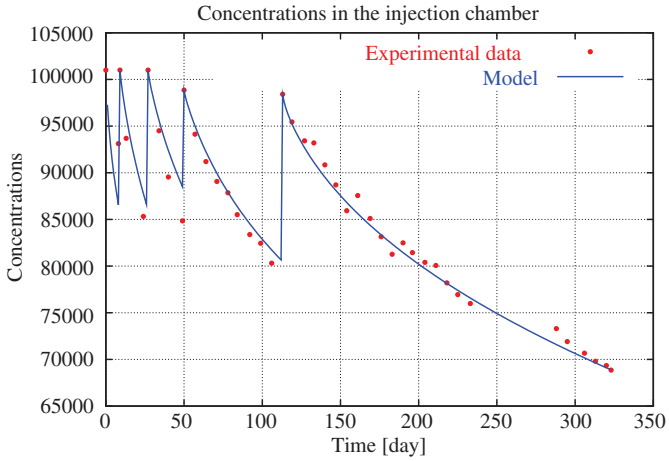

Fig. 13. Experimental and computed monitoring.

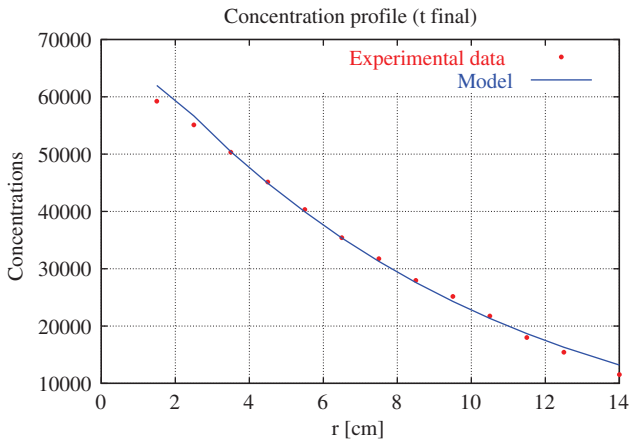

Fig. 14. Experimental and computed profile.

parameters $\chi(r)$ and $\Psi(r)$ on the monitoring are presented in Fig. 13.

In spite of the one-dimensional modeling, the parameter values found seem to be coherent with the experimental data and the 3D inverse modeling. The range of the diffusion is between 10^{-11} and 10^{10} m^2 s^{-1} for the diffusion. The porosity presents a mean value of 5% high value over the first two centimeters. The first two centimeters present a high value of the porosity (greater than 35%) which could be explained by the disturbed zone.

4. Discussion

The way to explore the parameter space $\wp$ in order to find the solution of the minimization problem is the main difference between the global and local approaches.

In the global strategy $\wp$ is scanned with a large number of parameters. The larger is the set of $\{\mathbf{p}_k/k = 1,\ldots,N\}$, the better is the chance to find the solution. However when the simulations are performed with a numerical code, the approach increases the computation time because one simulation must be carried out for each $\mathbf{p}_k$ to obtain the corresponding value of the cost function.

On the contrary, the local strategy searches to reduce the computation time by scanning only a region of $\wp$, around an initial point $\mathbf{p}_0$ given by the modeler. The choice of $\mathbf{p}_{k+1}$ at iterate $k+1$, is not independent from the previous one $\mathbf{p}_k$. The local terminology is now clear. All the sequence $\mathbf{p}_k$ $(k = 1,\ldots,N)$ is dependent from the starting point $\mathbf{p}_0$ and its choice is important for the final solution. Indeed, if the cost function presents more than one stationary point, an initial point $\mathbf{p}_0$ could lead to a local minimum. The flaw of this approach is the local character of the solution. To ensure the finding of the global solution, several optimizations should be performed with different values of $\mathbf{p}_0$.

Let us emphasize that the problem is ill-posed: it may not have an unique solution, or the solution may depend non continuously on the data. Hence it leads to numerical oscillations when the number of parameters is too high, and Tichonov regularization by adding a penalizing term cures the problem.

Another issue about the efficiency of the optimization algorithms is the computation of the derivatives of the cost function, or of the forward model. There are several ways to compute derivatives: approximated or analytical, direct or reverse mode, automatic or manual implementation. The approximate computation of derivatives by finite differences is very simple as it does not need any further code development, but it provides the direct mode. Hence, it is advisable to restrict its use to the validation of analytical derivatives computations. The complexity of the direct (or forward) mode is proportional to the number of inputs, e.g. to the number of parameters, and the complexity of the reverse (or backward, or adjoint) mode is proportional to the number of outputs, eg to the number of measures. So, when dealing with gradients, the reverse mode is obviously the most efficient way since there is only one output: the value of the cost function. But when computing a Jacobian (or sensitivity matrix), the choice of the cheapest mode depends on which is higher between the number of lines and the number of columns. Automatic differentiation opens either through code transformation as Tapenade (Hascoët and Pascual, 2004), or Adifor (Bischof et al., 1992) or through operator overloading as AdolC (Walther and Griewank, 2004). These high-level tools are very promising and their use is highly recommended for the direct mode, but there are still some limitations: choice of the programming language, access to the reverse mode, source code requirements. Manual differentiation of the discrete equations in the reverse mode is not difficult in principle when using variational numerical schemes, but its implementation, and its validation, can be a long and tedious task for large codes. Furthermore, it is even possible to discretize the continuous adjoint equations, but then one has to be careful, eg by refining the mesh when getting closer to the solution. We have chosen to compute the derivatives in the reverse mode, and to implement them manually. Even when computing the full Jacobian matrix for the singular value decomposition in the next section.

No supplementary development is required in the global approach. The direct simulation code is sufficient to obtain a good estimation of the parameters. However the interpretation is difficult especially when the dimension of the parameter space is greater than 3. It is then not possible to interpret graphically the results and it can be extremely hard to distinguish several small values of the cost function and to choose between a set of parameters and another. Moreover the exploration of the parameter space with a large number of $\mathbf{p}$ does not guarantee that the found solution is the global solution of the problem. But it highlights rapidly the family of probable solutions and can be tested afterwards with an optimization algorithm. A genetic algorithm with multicriteria optimization techniques, could be used to avoid these difficulties and interpret the results (Gaudier and Dumas, 2005).

Because the neural network is efficient, it allows to scan a large region of the parameter space and gives rapidly some preliminary results. However its implementation is not fast because four steps have to be performed: (1) training simulations, (2) neural network building, (3) verify the suitability of the neural network with the simulation code, (4) finding the minimum of the objective function, and (5) validation of solutions with a simulation code. The first step is the longest because it requires a lot of direct simulations. Each neural network is next quite easy to build. Let us emphasize that the relation (13) is a combination of nonlinear function depending on the parameters. It is a supplementary approximation of a code which is already a numerical approximation of physical equations. The initial choice of the parametrization and the stationarity/transient problem is important because the built neural network will be dependent from these hypothesis. A modification of the zonation requires to perform again the training simulations

and to build new neural networks. Nevertheless, let us remember that the neural network is an interesting tool to obtain some results quickly. The solution of the minimization problem should be next validated with the simulation code.

4.1. Singular value decomposition

The parametrization, introduced in Section 3.1, raises several difficulties. As already mentioned, two notions are hidden in the parametrization concept: the structure of parameters and the values inside it. In this paragraph, the position and the number of zones (in DI inverse modeling), were supposed to be known. The presentation of methods and the discussion were focused on the identification of parameter values. However, in practice, the parameter structure is not always known. So, to overcome this difficulty, two possibilities exist. The first one tries to estimate at the same time the structure of the zonation and the parameter values inside it. Such a method has been applied by Ben Ameur et al. (2002) in a two-dimensional case using the refinement and coarsening indicators for an adaptive parametrization.

A second possibility to find a good zonation, or even to compare several parametrizations, is based on a sensitivity analysis of the concentration measures with respect to the parameters. The study of singular values and vectors, given by the decomposition (SVD) of the Jacobian matrix, is a powerful tool to define the identifiable degrees of freedom and is very efficient to quantify the sensitivity of the system. It gives indications to choose the best parametrization, to test a new acquisition system, and to evaluate the non linearity of the model. However, conclusions are valid only locally around the considered point.

In DI experiment, the components of the Jacobian matrix are the first partial derivatives of the concentration with respect to the porosity and diffusion parameters. The Jacobian matrix is computed row by row with the adjoint state method. The difference between the adjoint equation giving the gradient of the cost function is the right hand side of the equation. The singular value decomposition is a generalization of the diagonalization concept applied to rectangular matrices. The singular values of the Jacobian matrix $\mathbf{A}$ with I rows and J columns are the positive square roots of the eigenvalues of the symmetric matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. More precisely, the SVD of $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}}, \tag{20}$$

with $\mathbf{V}$, the singular vector matrix in the parameter space, $\mathbf{U}$, the singular vector matrix in the measure space, and $\mathbf{S}$, the diagonal singular value matrix in decreasing order. Singular values are nonnegative and $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices.

Many simulations have been performed to test the sensitivity of different parameter values and zonations. Parametrization with 20 and 3 zones have been studied. In the last case, the sensitivity of zone positions have been studied too. We focus the discussion in this paper on the interpretation of singular value decomposition. We refer Clément et al. (2004) for a presentation of other results.

We consider a parametrization with three zones. Fig. 15 displays the porosity and diffusion values associated with each zone. 175 values of monitoring values and 22 profile values (Fig. 16) are next computed with these 6 parameters using the 1D simulation code. The derivatives of the 197 ($=175+22$) measures with respect to the 6 parameters are then computed by the adjoint state method. A matrix $\mathbf{A}$ with 197 rows and 6 columns is obtained. The singular value decomposition (20) gives the unitary square matrices $\mathbf{U}$ and $\mathbf{V}$. The order of $\mathbf{U}$ and $\mathbf{V}$ is 197 and 6, respectively. The dimension of the matrices $\mathbf{A}$ and $\mathbf{S}$ are identical. But only the main diagonal of the matrix $\mathbf{S}$ is potentially non null. It is composed of positive terms classified in decreasing order.

The interpretation of the SVD is the following: for $i = 1, 2, \ldots, 6$, a modification of the parameter in the v_i direction will affect the measures in the u_i direction, proportionally to the singular value s_i. Fig. 17 displayed on the logarithmic scale presents the six relative singular values. The ratio between the values are 30 meaning that the component of parameter on v_1 will be found with the same noise level than the measures. The component on v_6 will

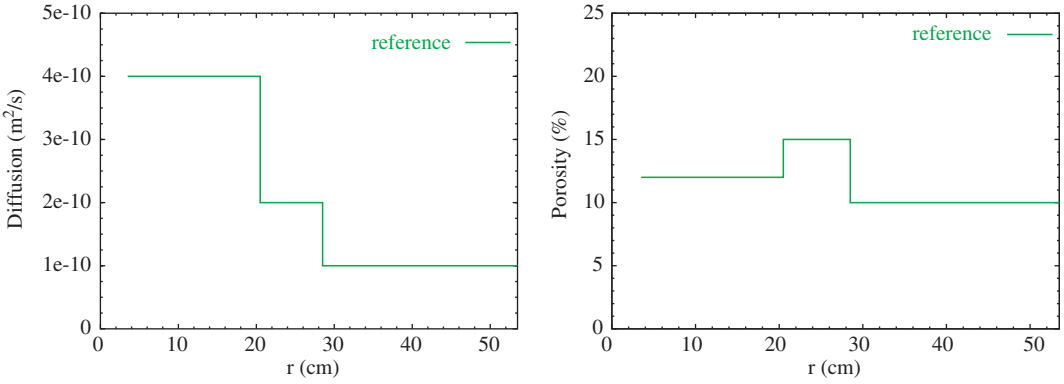

Fig. 15. Diffusion (left) and porosity (right) values (Clément et al., 2004).

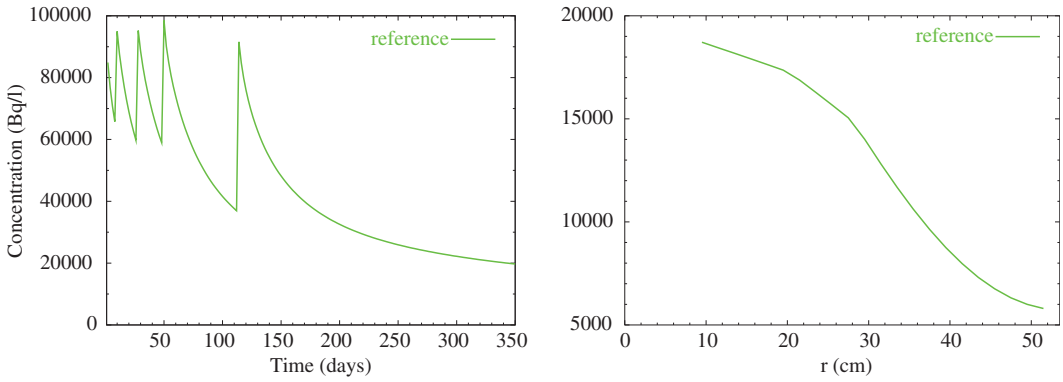

Fig. 16. Concentration monitoring (left) and profile (right) simulated.

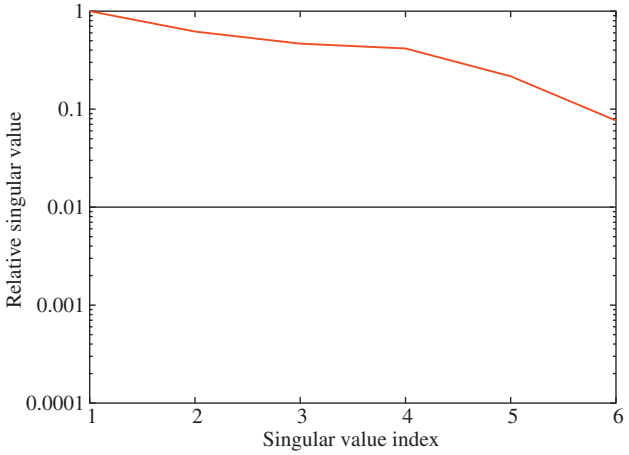

Fig. 17. Singular values.

be found with a noise level 30 times higher. So the more the singular value curve is near the constant curve 1, the more the inversion will be better. In Fig. 17 the first four are in a ratio 6.

Fig. 18 displays the singular vectors in the parameter space $\mathbf{V}^{\mathrm{T}}$. The influence of the parameters is the following: porosity in the zone 1, porosity in the zone 3, diffusion in the zone 1, the porosity in the zone 2, difference of the diffusion in the zones 2 and 3; mean of the diffusion in the zones 2 and 3. Fig. 19 displays the first six singular vectors in the measure space. The classification in the decrease order of the sensitivity of measures is the following: difference between the mean in the zone 1 of the profile at initial time and the injection peaks in the monitoring; mean value in the zone 3 of the profile at final time; mean between the mean in the zone 1 of the profile at the final time and the injection peaks in the monitoring; mean in the zone 2 of the profile at the final time; difference between a mean at the beginning of the zone 2 of the profile at the final time and one value at the end of this zone; mean between the beginning and the end of the zone 2 of the profile at the final time.

Let us mention that the singular value decomposition results depend greatly upon the weights μ and ν in the cost function (14). They also depend upon normalization values of porosity ω_0 and D_0. So, before beginning the sensitivity analysis, several simulations have been performed to choose the best weight values μ and ν and the best normalization values of diffusion and porosity. The simulation

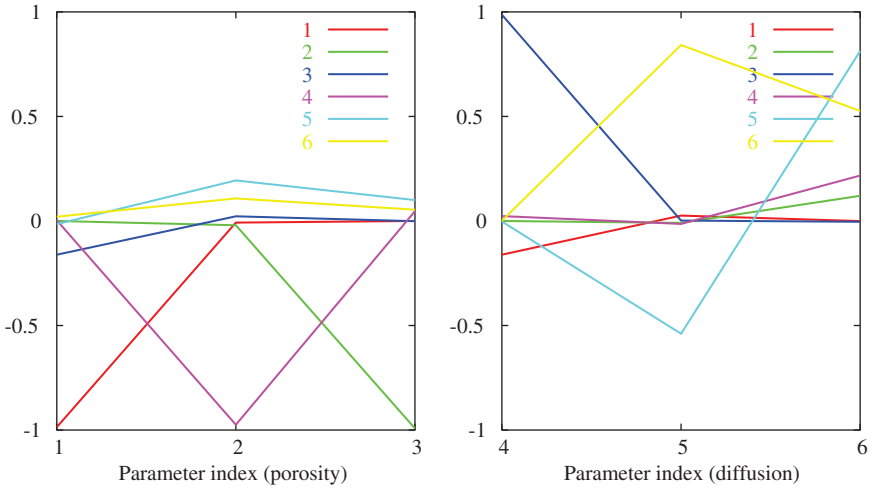

Fig. 18. Singular vectors in the parameter space.

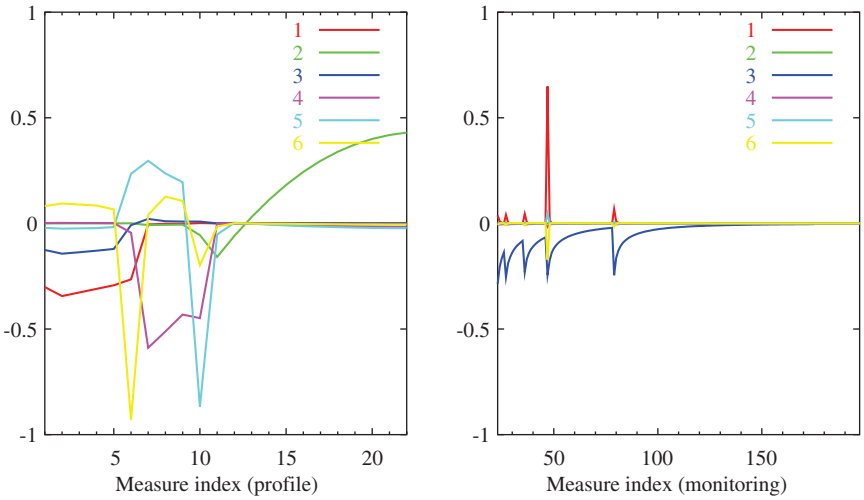

Fig. 19. Singular vectors in the measure space.

results shown in this paper are obtained with $\mu = 1$, $v = 10^{-1}$, $\omega_0 = 10^{-1}$ and $D_0 = 10^{-10}\,m^2/s$. The choice of these weight and normalization values is important and is computation time consuming.

5. Conclusion

The existence of various methods to identify numerically the parameters in a partial derivative equation complicates the manner to proceed to solve the inverse problem. The approach to use, global or local, and the optimization methods to apply (algorithms with or without gradients and methods of computation associated) depends on the physical problem, the framework and the numerical tools for modeling. The objective function is written differently according to the experimental design and the available data. Of course the overview is far to be exhaustive and many methods have not been investigated. Let us mention the genetic algorithms and all the methods based on the Bayesian theory.

Global and local approaches have been applied to HE-C and DI experiments, respectively. Rock properties were determined using complementary laboratory measurements on rock samples. Parameter values found with the inverse modeling are in good agreement with the experimental methods. This contribution allows us to synthesize different approaches for inverse modeling and yields valuable information about their advantages and weaknesses. In the global approach, interpretation in the parameter space becomes difficult when the number of parameters is greater than three. The global approach requires a great number of computations (500,000). In order to drastically reduce the CPU time of computations, a neural network, using NeMo (software CEA), is built to approximate the forward model. The neural network training requires 1100 forward simulations but a new parametrization of forward model requires to build a new neural network. In the local approach, the set of parameters is a local solution and is not necessarily the global minimum. To accelerate convergence it is better to have gradients of the objective function. It is time consuming when computed by finite differences, particularly when the number of parameters is important. Adjoint state method gives more accurate gradients, but the adjoint system building requires a long programming effort.

Appendix. Comparison between the data and models for the DI experiment

See (Fig. 1).

References

Ben Ameur, H., Chavent, G., Jaffré, J., 2002. Refinement and coarsening indicators for adaptive parametrization: application to the estimation of hydraulic transmissivities. Inverse Problems 18, 775–794.

Bischof, C., Carle, A., Corliss, G., Griewank, A., Hovland, P., 1992. ADIFOR -Generating Derivative Codes from Fortran Programs. Scientific Programming (1), 1–29.

Bonnans, J.F., Gilbert, J.-C., Lemaréchal, C., Ghidaglia, J.M. (Eds.), Optimisation numérique aspects théoriques et pratiques. Mathématiques et applications, vol. 27. Springer, Berlin.

Bouillard, N., Montarnal, P., Herbin, R., 2005. Development of numerical methods for the reactive transport of chemical species in a porous media: a nonlinear conjugate gradient method. In: Proceedings accepted in International Conference on Computational Method for Coupled Problems in Science and Engineering, COUPLED PROBLEMS 2005.

Byrd, R.H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. Siam Journal of Scientific Computing 16 (5), 1190–1208.

Cartalade, A., Montarnal, P., Cavanna, B., Blum, J., 2003, Paramétrisation automatique des coefficients de transport d'un milieu poreux, approche par état adjoint, CEA Technical Report DM2S/SFME/MTMS/03-002/A.

Carrera, J., Neuman, S.P., 1986. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. Water Resources Research 22 (2), 211–227.

CAST3M. Available from: <http://www.cast3m.cea.fr>.

Chavent, G., 1971. Analyse fonctionnelle et identification de coefficients répartis des équations aux dérivées partielles, thèse de l'Université Paris VI.

Chavent, G., 1975. History matching by use of optimal control theory. Society of Petroleum Engineers Journal 15 (1), 74–86.

Chavent, G., 1979. Identification of distributed parameter systems: about the output least square method, its implementation and identifiability. In: Proceedings of the 5th IFAC Symposium on Identification and System Parameter Estimation. Pergamon Press, pp. 85–97.

Clément, F., Khvoenkova, N., Cartalade, A., Montarnal, P., 2004, Analyse de sensibilité et estimation de paramètres de transport pour une équation de diffusion, approche par état adjoint. Research Report INRIA no. 5132. Available from: <http://www.inria.fr/rrrt/rr-5132.html>.

Dabbene, F., 1995. Schémas de diffusion-convection en éléments finis mixtes hybrides, CEA Technical Report, DMT/95/613.

de Marsily, G., Delhomme, J.-P., Delay, F., Buoro, A., 1999. Regards sur 40 ans de problèmes inverses en hydrogéologie. Comptes Rendus de l Academie Des Sciences, Paris, Sciences de la terre et des planètes 329, 73–87.

Dreyfus, G., Martinez, J.-M., Samuelides, M., Gordon, M.B., Bardan, F., Hérault, L., 2002. Réseaux de neurones, méthodologie et applications. Eyrolles.

Fierz, T., 1999. DI-Experiment Field activities, Tracer Injection, Sampling, Maintenance of the Test Equipment. Mont Terri Project. Technical Note 99-39.

Filippi, M., 2003. Détermination de la conductivité thermique de l'argile opalinus de l'experience HE-C: simulation numérique inverse, CEA Technical Report DM2S/SFME/MTMS/RT/03-03 1/A.

Fletcher, R., 1987. Practical Methods of Optimization. John Wiley & Sons, New York.

Gaudier, F, Dumas, M., 2005. Uncertainties Analysis by Genetic Algorithms. Application to Wall Friction Model. The 11th International Topical Meeting on Nuclear Thermal-Hydraulics Paper 355, October 2-6 Avignon, France.

Hascoët, L., Pascual, V., 2004. Tapenade 2.1 User's Guide, INRIA Technical Report, France.

Lions, J.-L., 1968. Contrôle optimal de systemes gouvernés par des équations aux dérivées partielles. Dunod/Gauthier-Villars, Paris.

Martinez, J.-M., Arnaud, G., Gaudier, F., 2002. Kalif: outil d'aide à la qualification et à l'optimisation de codes. CEA Technical Report DEN/DM2S/SFME/LETR/TR/02-03 1/A.

McLaughlin, D., Townley, L.R., 1996. A reassessment of groundwater inverse problem. Water Resources Research 32 (5), 1131–1161.

Mont Terri Project. Available from: <http://www.mont-terri.ch/>.

Montarnal, P., Lamoureux, M., 2000, DI experiment: Numerical modelling of the experiment. CEA Technical Report DMT/SEMT/MTMS/PU/00-031.

Montarnal, P., Traynard, E., Martinez, J.-M., Arnaud, G., Dumas, M., 2002. Identification des paramètres de transport dans un milieu poreux et analyse multicritère á l'aide de réseaux de neurones: application á l'expérience DI du Mont Terri. CEA Technical Report DE/DM2S/SFME/MTMS/RT/02-021/A.

Möri, A., Steiger, H., Bossart, P., 2000. DI experiment: sampling documentation of boreholes BDI-1 and BDI-3. Geotechnical Institute Ltd., Mont Terri Project, Technical note 2000-16.

Mugler, C, Transfert thermique dans un milieu argileux: simulation directe 3D de l'expérience HE-C du Mont Terri. CEA Technical report DEN/DM2S/SFME/MTMS/TR/03-006/A.

Nocedal, J., Wright, S., 1999. Numerical optimization. Springer Series in Operations Research.

Novakowski, K.S., ven der Kamp, G., 1996. The radial diffusion method 2. A semianalytical model for the determination of effective diffusion coefficients, porosity, and adsorption. Water Resources Research 32 (6), 1823–1830.

OFTA (Observatoire Français des Techniques Avancées), 1999. Problèmes inverses de l'expérimentation à la modélisation. ARAGO 22. Editions TEC & DOC.

Özisik, M., Necati, O., Helcio, R.B., 2000. Inverse Heat Transfer Fundamentals and Applications. Taylor & Francis, New York (lieu d'édition).

Palut, J.M., Montarnal, P., Gautschi, A., Tevissen, E., Mouche, E., 2003. Characterisation of HTO diffusion by an in-situ tracer experiment in Opalinus Clay at Mont Terri. Jounal of Contaminant Hydrology 61, 405.

Sun, N.-Z., Yeh, W.W.-G., 1985. Identification of parameter structure in groundwater inverse problem. Water Resource Research 21 (6), 869–883.

Sun, N.-Z., 1994. Inverse Problems in Groundwater Modeling. Kluwer Publishers, The Netherlands.

Tarantola, A., 1987. Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation. Elsevier, New York.

Tevissen, E., Soler, J.M., 2003. In situ diffusion experiment (DI): Synthesis Report. Mont Terri Project, Technical Report 2001-05.

Tevissen, E., Soler, J.-M., Montarnal, P., Gautschi, A., Van Loon, L.R., 2004. Comparison between in situ and laboratory diffusion studies of HTO and halides in Opalinus Clay from the Mont Terri. Radiochimica Acta 92, 781–786.

Walther, A., Griewank, A., 2004. ADOL-C: Computing higher-order derivatives and sparsity patterns for functions written in C/C++. In: Neittaanmäaki, P. et al. (Eds.), Proceeding of ECCOMAS Conference. Paper 577 (14 p).

Wileveau, Y., 2002. Expérimentation HE-C, Conception, installation et premiers résultats. ANDRA Technical Report nADPE 02 388.

Zimmerman, D.A., de Marsily, G., Gotway, C.A., Marietta, M.G., Axness, C.L., Beauheim, R.L., Bras, R.L., Carrera, J., Dagan, G., Davies, P.B., Gallegos, D.P., Galli, A., Gomez-Hernandez, J., Grindrod, P., Gutjahr, A.L., Kitanidis, P.K., Lavenue, A.M., McLaughlin, D., Neuman, S.P., RamaRao, B.S., Ravenne, C., Rubin, Y., 1998. A comparison of seven geostatiscally based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. Water Resources Research 34 (6), 1373–1413.