Available online at www.sciencedirect.com
SCIENCE DIRECT®

Mathematics and Computers in Simulation 66 (2004) 409–424

MATHEMATICS AND COMPUTERS IN SIMULATION
www.elsevier.com/locate/matcom

Simultaneous estimation of spatially dependent diffusion coefficient and source term in a nonlinear 1D diffusion problem

F.A. Rodrigues [a], H.R.B. Orlande [a,*], G.S. Dulikravich [b]

[a] *Department of Mechanical Engineering, POLI/COPPE, Federal University of Rio de Janeiro, UFRJ, Cidade Universitária, Cx. Postal 68503, 21945-970 Rio de Janeiro, RJ, Brazil*
[b] *Department of Mechanical and Materials Engineering, Florida International University, College of Engineering, Room EC 3474, 10555 West Flagler Street, Miami, FL 33174, USA*

Available online 14 April 2004

Abstract

This work deals with the use of the conjugate gradient method in conjunction with an adjoint problem formulation for the simultaneous estimation of the spatially varying diffusion coefficient and of the source term distribution in a one-dimensional nonlinear diffusion problem. In the present approach, no a priori assumption is required regarding the functional form of the unknowns. This work can be physically associated with the detection of material non-homogeneities, such as inclusions, obstacles or cracks, in heat conduction, groundwater flow and tomography problems. Three versions of the conjugate gradient method are compared for the solution of the present inverse problem, by using simulated measurements containing random errors in the inverse analysis. Different functional forms, including those containing sharp corners and discontinuities, are used to generate the simulated measurements and to address the accuracy of the present solution approach.

Keywords: Conjugate gradient method; Function estimation; Physical properties; Detection of obstacles

1. Introduction

Despite being considered in the past as not of physical interest because of their ill-posed character, inverse problems play nowadays an important role in the solution of a number of practical problems. The use of inverse methods represent a new research direction, where the results obtained from numerical simulations and from experiments are not simply compared a posteriori, but a close synergism exists between experimental and theoretical researchers during the course of the study, in order to obtain the

* Corresponding author.
E-mail addresses: fabio@lttc.coppe.ufrj.br (F.A. Rodrigues), helcio@serv.com.ufrj.br (H.R.B. Orlande), dulikrav@fiu.edu (G.S. Dulikravich).

0378-4754/$30.00 © 2004 IMACS. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.matcom.2004.02.005

Nomenclature

$D(x)$	spatially dependent diffusion coefficient
$d_D^k(x)$ and $d_\mu^k(x)$	directions of descent for $D(x)$ and $\mu(x)$, respectively
M	number of sensors
$S[D(x), \mu(x)]$	objective functional
t	time
t_f	final time
$U(x, t)$	estimated field variable
x	spatial variable
$Y_m(t)$	transient measurement obtained with sensor m ($m = 1, \ldots, M$)
Greek symbols	
β_D^k and β_μ^k	search step sizes for $D(x)$ and $\mu(x)$, respectively
γ_D^k, γ_μ^k, ψ_D^k and ψ_μ^k	conjugation coefficients
$\Delta D(x)$ and $\Delta\mu(x)$	variations in $D(x)$ and $\mu(x)$, respectively
$\Delta U_D(x, t)$ and $\Delta U_\mu(x, t)$	sensitivity functions for $D(x)$ and $\mu(x)$, respectively
$\lambda(x, t)$	Lagrange multiplier
$\mu(x)$	spatial distribution of the source term
σ	standard deviation of the measurements
∇S	gradient components for the functional
Subscript	
m	refers to the number of sensors ($m = 1, \ldots, M$)
Superscripts	
k	iteration number
qD and $q\mu$	iteration number where a restarting strategy is applied in the iterative procedure for $D(x)$ and $\mu(x)$, respectively
$*$	dimensional quantities

maximum of information regarding the physical problem under consideration [1]. Most of the methods for the solutions of inverse problems, which are currently in common use, were formalized in the last four decades in terms of their capabilities to treat ill-posed unstable problems. The basis of such formal methods resides on the idea of reformulating an inverse problem in terms of an approximate well-posed problem, by utilizing some kind of regularization (stabilization) technique (see, e.g., references [1–31]).

In this work we examine the use of different versions of the conjugate gradient method together with an adjoint problem formulation [2,3,7,8,11,15,16,32] for the purpose of simultaneous estimation of spatially dependent diffusion coefficients and source terms in a nonlinear diffusion problem. The inverse problem considered in this work is solved by using a function estimation approach [1–4], where no information is a priori available regarding the forms of the unknown functions, except for the functional space that they belong to. It is assumed that the unknowns belong to the Hilbert space of square integrable functions in the spatial domain of interest [2,3,6].

The solution of inverse problems by using the conjugate gradient method with adjoint problem for function estimation consists of the following basic steps [2,3]: (i) direct problem formulation, (ii) inverse problem formulation, (iii) sensitivity problems formulation, (iv) adjoint problem formulation, (v) gradient equations, (vi) iterative solution procedure, (vii) iterative procedure stopping criterion, and (viii) computational algorithm. Highlights of such steps are presented below, as applied to the inverse problem of interest. For further details on such steps, the reader should consult references [2,3].

2. Direct problem

In this work we consider problems governed by the following nonlinear diffusion equation:

$$C^*(r^*)\frac{\partial U^*(r^*,t^*)}{\partial t^*} = \nabla\cdot[D^*(r^*)\nabla U^*] + \mu^*(r^*)U^* \tag{1}$$

where r^* denotes de vector of coordinates and the superscript $*$ denotes dimensional quantities.

Eq. (1) can be used for the modeling of several physical phenomena, such as heat conduction [1–10], groundwater flow [5,11–15] and optical tomography [16–23]. The physical significance of the different quantities appearing in Eq. (1) is summarized in Table 1 for each of these physical problems.

We focus our attention on a one-dimensional version of Eq. (1) written in dimensionless form as

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\frac{\partial U}{\partial x}\right) + \mu(x)U \quad \text{in } 0 < x < 1 \text{ for } t > 0 \tag{2.a}$$

that is subject to the following boundary and initial conditions:

$$\frac{\partial U}{\partial x} = 0 \quad \text{at } x = 0 \text{ for } t > 0 \tag{2.b}$$

$$D(x)\frac{\partial U}{\partial x} = 1 \quad \text{at } x = 1 \text{ for } t > 0 \tag{2.c}$$

$$U = 0 \quad \text{for } t = 0 \text{ in } 0 < x < 1 \tag{2.d}$$

Notice that in the *direct problem* the diffusion coefficient function $D(x)$ and the source term distribution function $\mu(x)$ are regarded as known quantities, so that a direct (analysis) problem is concerned with the computation of $U(x, t)$.

Table 1
Physical significance of the quantities appearing in Eq. (1)

Physical problem	$U^*(r^*, t^*)$	$C^*(r^*)$	$D^*(r^*)$	$\mu^*(r^*)$
Heat conduction	Temperature	Volumetric heat capacity	Thermal conductivity	Spatial distribution of source term
Groundwater flow	Piezometric head	Storage coefficient	Transmissivity	Spatial distribution of source term
Optical tomography[a]	Fluence rate	–	Diffusion coefficient	Absorption distribution

[a] For optical tomography problems $C^*(r^*) = 1$, the diffusion coefficient is given by $D^*(r^*) = c^*/(3[\mu_a^*(r^*) + \mu_s'^*(r^*)])$ and $\mu^*(r^*) = -c^*\mu_a^*(r^*)$, where c^* is the speed of light in the medium, $\mu_a^*(r^*)$ is the absorption coefficient and $\mu_s'^*(r^*)$ is the reduced scattering coefficient [16].

3. Inverse problem

For the *inverse problem* of interest here, the functions $D(x)$ and $\mu(x)$ are regarded as unknown. Such functions will be simultaneously estimated by using measurements of $U(x, t)$ taken at appropriate locations in the medium or on its boundaries. Such measurements may contain random errors that are assumed to be uncorrelated, additive, normally distributed, with zero mean and with a known constant standard deviation.

Practical applications of this inverse problem include the identification of non-homogeneities in the medium, such as inclusions, obstacles or cracks, determination of thermal diffusion coefficients and distribution of heat sources, groundwater flow and tomography physical problems, in which both $D(x)$ and $\mu(x)$ vary [1–31].

For the simultaneous estimation of the functions $D(x)$ and $\mu(x)$ we make use of a minimization procedure involving the following objective functional:

$$S[D(x), \mu(x)] = \frac{1}{2}\int_{t=0}^{t_f}\sum_{m=1}^{M}\{U[x_m, t; D(x), \mu(x)] - Y_m(t)\}^2\,dt \tag{3}$$

where $Y_m(t)$ are the transient measurements of $U(x, t)$ taken at the positions x_m $(m = 1, \ldots, M)$. The estimated dependent variable $U[x_m, t; D(x), \mu(x)]$ is obtained from the solution of the direct problem (Eqs. (2.a)–(2.d)) at the measurement positions x_m $(m = 1, \ldots, M)$ with estimates for $D(x)$ and $\mu(x)$.

The conjugate gradient method with an adjoint problem is used for the minimization of the objective functional (Eq. (3)). Such minimization procedure requires the solution of auxiliary problems, known as sensitivity and adjoint problems.

4. Sensitivity problems

The *sensitivity function*, solution of the sensitivity problem, is defined as the directional derivative of $U(x, t)$ in the direction of the perturbation of the unknown function [2,3]. Since the present problem involves two unknown functions, two sensitivity problems are required for the estimation procedure, resulting from perturbations in $D(x)$ and $\mu(x)$.

The sensitivity problem for $U_D(x, t)$ is obtained by assuming that the dependent variable $U(x, t)$ is perturbed by $\varepsilon\Delta U_D(x, t)$ when the diffusion coefficient $D(x)$ is perturbed by $\varepsilon\Delta D(x)$. Here, ε is a real number. The sensitivity problem for $U_D(x, t)$ is then obtained by applying the following limiting process:

$$\lim_{\varepsilon\to 0}\frac{L_\varepsilon(D_\varepsilon) - L(D)}{\varepsilon} = 0 \tag{4}$$

where $L_\varepsilon(D_\varepsilon)$ and $L(D)$ are the direct problem formulations written in operator form for perturbed and unperturbed quantities, respectively. The application of the limiting process given by Eq. (4) results in the following sensitivity problem:

$$\frac{\partial \Delta U_D}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\frac{\partial \Delta U_D}{\partial x} + \Delta D(x)\frac{\partial U}{\partial x}\right) + \mu(x)\Delta U_D \quad \text{in } 0 < x < 1 \text{ for } t > 0 \tag{5.a}$$

$$\frac{\partial \Delta U_D}{\partial x} = 0 \quad \text{at } x = 0 \text{ for } t > 0 \tag{5.b}$$

$$\Delta D(x)\frac{\partial U}{\partial x} + D(x)\frac{\partial \Delta U_D}{\partial x} = 0 \quad \text{at } x = 1 \text{ for } t > 0 \tag{5.c}$$

$$\Delta U_D = 0 \quad \text{in } 0 \le x \le 1 \text{ for } t = 0 \tag{5.d}$$

A limiting process analogous to that given by Eq. (4), obtained from the perturbation $\varepsilon\Delta\mu(x)$, results in the following sensitivity problem for $\Delta U_\mu(x, t)$:

$$\frac{\partial \Delta U_\mu}{\partial t} = \frac{\partial}{\partial x}\left(D(x)\frac{\partial \Delta U_\mu}{\partial x}\right) + \mu(x)\Delta U_\mu + \Delta\mu(x)U \quad \text{in } 0 < x < 1 \text{ for } t > 0 \tag{6.a}$$

$$\frac{\partial \Delta U_\mu}{\partial x} = 0 \quad \text{at } x = 0 \text{ and } x = 1 \text{ for } t > 0 \tag{6.b,c}$$

$$\Delta U_\mu = 0 \quad \text{in } 0 \le x \le 1 \text{ for } t = 0 \tag{6.d}$$

5. Adjoint problem

A Lagrange multiplier $\lambda(x, t)$ is utilized in the minimization of the functional (Eq. (3)) because the estimated dependent variable $U[x_m, t; D(x), \mu(x)]$ appearing in such functional needs to satisfy a constraint, which is the solution of the direct problem. Such Lagrange multiplier, needed for the computation of the gradient equations (as will be apparent below), is obtained through the solution of problems adjoint to the sensitivity problems, given by Eqs. (5.a)–(6.d) [2,3]. Despite the fact that the present inverse problem involves the estimation of two unknown functions, thus resulting in two sensitivity problems as discussed above, one single problem, adjoint to problems (5.a)–(6.d), is obtained.

In order to derive the adjoint problem, the governing equation of the direct problem (Eq. (2.a)) is multiplied by the Lagrange multiplier $\lambda(x, t)$, integrated in the space and time domains of interest and added to the original functional (Eq. (3)). The following extended functional is obtained:

$$S[D(x), \mu(x)] = \frac{1}{2}\int_{x=0}^{1}\int_{t=0}^{t_f}\sum_{m=1}^{M}[U - Y]^2\delta(x - x_m)\,dt\,dx + \int_{x=0}^{1}\int_{t=0}^{t_f}\left[\frac{\partial U}{\partial t} - \frac{\partial}{\partial x}\left(D(x)\frac{\partial U}{\partial x}\right) - \mu(x)U\right]\lambda(x, t)\,dt\,dx \tag{7}$$

where δ is the Dirac delta function.

Directional derivatives of $S[D(x), \mu(x)]$ in the directions of perturbations in $D(x)$ and $\mu(x)$ are respectively defined by

$$\Delta S_D[D, \mu] = \lim_{\varepsilon\to 0}\frac{S[D_\varepsilon, \mu] - S[D, \mu]}{\varepsilon} \tag{8.a}$$

$$\Delta S_\mu[D, \mu] = \lim_{\varepsilon\to 0}\frac{S[D, \mu_\varepsilon] - S[D, \mu]}{\varepsilon} \tag{8.b}$$

where $S[D_\varepsilon, \mu]$ and $S[D, \mu_\varepsilon]$ denote the extended functional (Eq. (7)) written for perturbed $D(x)$ and $\mu(x)$, respectively.

After performing some lengthy but straightforward manipulations [2,3] and letting the directional derivatives of $S[D(x), \mu(x)]$ go to zero, which is a necessary condition for the minimization of the extended functional (Eq. (7)), the following adjoint problem for the Lagrange multiplier $\lambda(x, t)$ is obtained:

$$-\frac{\partial\lambda}{\partial t} - \frac{\partial}{\partial x}\left(D(x)\frac{\partial\lambda}{\partial x}\right) - \mu(x)\lambda + \sum_{m=1}^{M}[U - Y]\delta(x - x_m) = 0 \quad \text{in } 0 < x < 1 \text{ for } t > 0 \tag{9.a}$$

$$\frac{\partial\lambda}{\partial x} = 0 \quad \text{at } x = 0 \text{ and } x = 1 \text{ for } t > 0 \tag{9.b,c}$$

$$\lambda = 0 \quad \text{in } 0 \leq x \leq 1 \text{ for } t = t_\mathrm{f} \tag{9.d}$$

6. Gradient equations

During the limiting processes used to obtain the adjoint problem, applied to the directional derivatives of $S[D(x), \mu(x)]$ in the directions of perturbations in $D(x)$ and $\mu(x)$, the following integral terms are respectively obtained:

$$\Delta S_D[D, \mu] = \int_{x=0}^{1}\int_{t=0}^{t_\mathrm{f}} \Delta D(x)\frac{\partial U}{\partial x}\frac{\partial\lambda}{\partial x}\,\mathrm{d}t\,\mathrm{d}x \tag{10.a}$$

$$\Delta S_\mu[D, \mu] = -\int_{x=0}^{1}\int_{t=0}^{t_\mathrm{f}} \Delta\mu(x)\lambda(x, t)U(x, t)\,\mathrm{d}t\,\mathrm{d}x \tag{10.b}$$

By invoking the hypotheses that $D(x)$ and $\mu(x)$ belong to the Hilbert space of square integrable functions in the domain $0 < x < 1$, it is possible to write [2,3]

$$\Delta S_D[D, \mu] = \int_{x=0}^{1} \nabla S[D(x)]\Delta D(x)\,\mathrm{d}x \tag{11.a}$$

$$\Delta S_\mu[D, \mu] = \int_{x=0}^{1} \nabla S[\mu(x)]\Delta\mu(x)\,\mathrm{d}x \tag{11.b}$$

Hence, by comparing Eqs. (10.a), (10.b) and (11.a), (11.b) we obtain the gradient components of $S[D, \mu]$ with respect to $D(x)$ and $\mu(x)$, respectively, as

$$\nabla S[D(x)] = \int_{t=0}^{t_\mathrm{f}} \frac{\partial U}{\partial x}\frac{\partial\lambda}{\partial x}\,\mathrm{d}t \tag{12.a}$$

$$\nabla S[\mu(x)] = -\int_{t=0}^{t_\mathrm{f}} \lambda(x, t)U(x, t)\,\mathrm{d}t \tag{12.b}$$

An analysis of Eqs. (9.b,c) and (12.a) reveals that the gradient component with respect to $D(x)$ is null at $x = 0$ and $x = 1$. As a result, the initial guess used for $D(x)$ is never changed by the iterative procedure of the conjugate gradient method at such points, which can create instabilities in the inverse problem solution in their neighborhoods.

7. Iterative procedure

For the simultaneous estimation of $D(x)$ and $\mu(x)$, the iterative procedure of the conjugate gradient method is written respectively as [2,3]

$$D^{k+1}(x) = D^k(x) + \beta_D^k d_D^k(x) \tag{13.a}$$

$$\mu^{k+1}(x) = \mu^k(x) + \beta_\mu^k d_\mu^k(x) \tag{13.b}$$

where $d_D^k(x)$ and $d_\mu^k(x)$ are the directions of descent for $D(x)$ and $\mu(x)$, respectively; β_D^k and β_μ^k are the search step sizes for $D(x)$ and $\mu(x)$, respectively; k is the number of iterations.

For the iterative procedure for each unknown function, the direction of descent is obtained as a linear combination of the gradient direction with directions of descent of previous iterations. The directions of descent for the conjugate gradient method for $D(x)$ and $\mu(x)$ can be written respectively as

$$d_D^k(x) = -\nabla S[D^k(x)] + \gamma_D^k d_D^{k-1}(x) + \psi_D^k d_D^{qD}(x) \tag{14.a}$$

$$d_\mu^k(x) = -\nabla S[\mu^k(x)] + \gamma_\mu^k d_\mu^{k-1}(x) + \psi_\mu^k d_\mu^{q\mu}(x) \tag{14.b}$$

where γ_D^k, γ_μ^k, ψ_D^k and ψ_μ^k are the conjugation coefficients. The superscripts qD and $q\mu$ in Eqs. (14.a) and (14.b) represent the iteration numbers where a restarting strategy is applied to the iterative procedure for the estimation of $D(x)$ and $\mu(x)$, respectively [32].

Different versions of the conjugate gradient method can be found in the literature, depending on how the conjugation coefficients are computed. The conjugation coefficients for Powell-Beale's version of the conjugate gradient method are given by [8,32]

$$\gamma_D^k = \frac{\int_{x=0}^{1}\{\nabla S[D^k(x)] - \nabla S[D^{k-1}(x)]\}\nabla S[D^k(x)]\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[D^k(x)] - \nabla S[D^{k-1}(x)]\}d_D^{k-1}(x)\,\mathrm{d}x} \tag{15.a}$$

$$\gamma_\mu^k = \frac{\int_{x=0}^{1}\{\nabla S[\mu^k(x)] - \nabla S[\mu^{k-1}(x)]\}\nabla S[\mu^k(x)]\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[\mu^k(x)] - \nabla S[\mu^{k-1}(x)]\}d_\mu^{k-1}(x)\,\mathrm{d}x} \tag{15.b}$$

$$\psi_D^k = \frac{\int_{x=0}^{1}\{\nabla S[D^{qD+1}(x)] - \nabla S[D^{qD}(x)]\}\nabla S[D^k(x)]\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[D^{qD+1}(x)] - \nabla S[D^{qD}(x)]\}d_D^{qD}(x)\,\mathrm{d}x} \tag{15.c}$$

$$\psi_\mu^k = \frac{\int_{x=0}^{1}\{\nabla S[\mu^{q\mu+1}(x)] - \nabla S[\mu^{q\mu}(x)]\}\nabla S[\mu^k(x)]\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[\mu^{q\mu+1}(x)] - \nabla S[\mu^{q\mu}(x)]\}d_\mu^{q\mu}(x)\,\mathrm{d}x} \tag{15.d}$$

where $\gamma_D^k = \gamma_\mu^k = \psi_D^k = \psi_\mu^k = 0$ for $k = 0$.

Powell-Beale's version of the conjugate gradient method is restarted by making the conjugation coefficient $\psi_D^k = 0$ (or $\psi_\mu^k = 0$) if gradients at successive iterations are too far from being orthogonal (which is a measure of the nonlinearity of the problem) or if the direction of descent is not sufficiently downhill. For further details, the reader is referred to Refs. [8,32].

For Fletcher-Reeves' version of the conjugate gradient method, the conjugation coefficients are computed as [2,3,8,32]

$$\gamma_D^k = \frac{\int_{x=0}^{1}\{\nabla S[D^k(x)]\}^2\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[D^{k-1}(x)]\}^2\,\mathrm{d}x}, \qquad \gamma_\mu^k = \frac{\int_{x=0}^{1}\{\nabla S[\mu^k(x)]\}^2\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[\mu^{k-1}(x)]\}^2\,\mathrm{d}x} \tag{16.a,b}$$

For Polak-Ribiere's version of the conjugate gradient method, the conjugation coefficients are computed as [2,3,8,32]

$$\gamma_D^k = \frac{\int_{x=0}^{1}\nabla S[D^k(x)]\{\nabla S[D^k(x)] - \nabla S[D^{k-1}(x)]\}\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[D^{k-1}(x)]\}^2\,\mathrm{d}x} \tag{17.a}$$

$$\gamma_\mu^k = \frac{\int_{x=0}^{1}\nabla S[\mu^k(x)]\{\nabla S[\mu^k(x)] - \nabla S[\mu^{k-1}(x)]\}\,\mathrm{d}x}{\int_{x=0}^{1}\{\nabla S[\mu^{k-1}(x)]\}^2\,\mathrm{d}x} \tag{17.b}$$

In Fletcher-Reeves' and Polak-Ribiere's versions of the conjugate gradient method $\gamma_D^k = \gamma_\mu^k = 0$ for $k = 0$. Furthermore, in these two versions of the conjugate gradient method $\psi_D^k = \psi_\mu^k = 0$ for any k, so that a restarting strategy is not taken into account as in Powell-Beale's version.

The search step sizes β_D^k and β_μ^k appearing in the expressions of the iterative procedures for the estimation of $D(x)$ and $\mu(x)$ (Eqs. (13.a) and (13.b), respectively) are obtained by minimizing the objective functional at each iteration along the specified directions of descent. If the objective functional given by Eq. (3) is linearized with respect to β_D^k and β_μ^k, closed form expressions can be obtained for such quantities as follows [2,3]:

$$\beta_D^k = \frac{F_1A_{22} - F_2A_{12}}{A_{11}A_{22} - A_{12}^2}, \qquad \beta_\mu^k = \frac{F_2A_{11} - F_1A_{12}}{A_{11}A_{22} - A_{12}^2} \tag{18.a,b}$$

where

$$A_{11} = \int_{t=0}^{t_\mathrm{f}}\sum_{m=1}^{M}[\Delta U_D^k(x_m, t)]^2\,\mathrm{d}t \tag{19.a}$$

$$A_{22} = \int_{t=0}^{t_\mathrm{f}}\sum_{m=1}^{M}[\Delta U_\mu^k(x_m, t)]^2\,\mathrm{d}t \tag{19.b}$$

$$A_{12} = \int_{t=0}^{t_\mathrm{f}}\sum_{m=1}^{M}\Delta U_D^k(x_m, t)\Delta U_\mu^k(x_m, t)\,\mathrm{d}t \tag{19.c}$$

$$F_1 = \int_{t=0}^{t_\mathrm{f}}\sum_{m=1}^{M}[Y_m^k - U^k(x_m, t)][\Delta U_D^k(x_m, t)]\,\mathrm{d}t \tag{19.d}$$

$$F_2 = \int_{t=0}^{t_\mathrm{f}}\sum_{m=1}^{M}[Y_m^k - U^k(x_m, t)][\Delta U_\mu^k(x_m, t)]\,\mathrm{d}t \tag{19.e}$$

In Eqs. (19.a)–(19.e), $\Delta U_D^k(x,t)$ and $\Delta U_\mu^k(x,t)$ are the solutions of the sensitivity problems given by Eqs. (5.a)–(5.d) and (6.a)–(6.d), respectively, obtained by setting $\Delta D^k(x) = d_D^k(x)$ and $\Delta\mu^k(x) = d_\mu^k(x)$.

8. Stopping criterion

The use of the conjugate gradient method for the simultaneous estimation of $D(x)$ and $\mu(x)$ can be suitably arranged in a systematic and straightforward computational procedure, which is omitted here for the sake of brevity, but can be readily found in Ref. [3]. The conjugate gradient method of function estimation belongs to the class of *iterative regularization methods* [2]. In such class of methods, the stopping criterion for the computational procedure is used as a regularization parameter, so that sufficiently accurate and smooth solutions are obtained for the unknown functions. Although different approaches can be used for the specification of the tolerance for the stopping criterion, we illustrate in this work the use of the *discrepancy principle* [2].

With the use of the discrepancy principle, the iterative procedure of the conjugate gradient method is stopped when the difference between measured and estimated variables is of the order of the standard deviation, σ, of the measurements, that is, when

$$|U(x_m, t; D, \mu) - Y_m(t)| \approx \sigma \quad (20)$$

Therefore, the iterative procedure is stopped when

$$S[D(x), \mu(x)] < \varphi \quad (21)$$

where the tolerance, φ, is obtained by substituting Eq. (20) into the definition of the functional given by Eq. (3), that is

$$\varphi = \tfrac{1}{2} M\sigma^2 t_\mathrm{f} \quad (22)$$

9. Results and discussions

The accuracy of the present solution approach was examined by using simulated transient measurements containing random errors in the inverse analysis. Different functional forms, including those containing sharp corners and discontinuities that are the most difficult to be recovered by the inverse analysis, were used to generate the simulated measurements. The functions were supposed to vary from a base value (D_c or μ_c) to a maximum value (D_0 or μ_0) within a specified range (ε_D or ε_μ), as illustrated in Fig. 1.

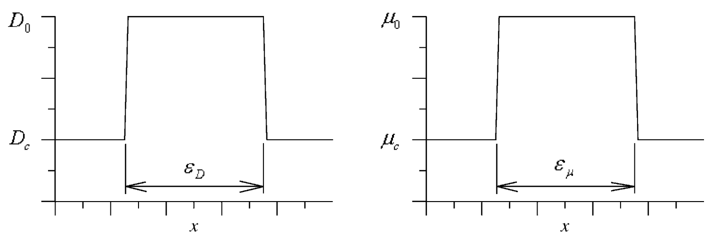

Fig. 1. Functions used to generate the simulated measurements.

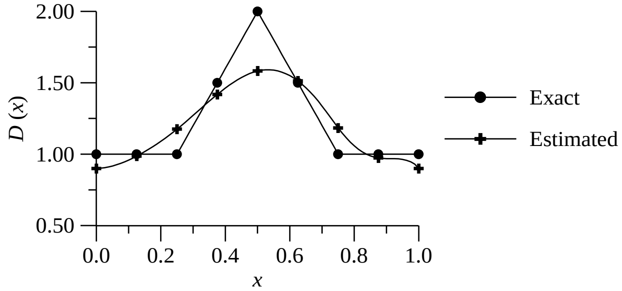

Fig. 2. Estimation of $D(x)$ for known $\mu(x)$ obtained with errorless measurements of two non-intrusive sensors.

For the test cases examined, the direct, sensitivity and adjoint problems were numerically solved using finite volumes. The numerical solution of the direct problem was validated with a test case with a known analytical solution. The discrepancy between numerical and analytical solutions of the direct problem was less than 1%, by using 80 equal volumes and a time step of 0.0072 for the discretization. This number of volumes and time steps were used for all test cases considered in this work.

The test cases examined below in dimensionless form are physically associated with a heat conduction problem in a homogeneous steel bar of length 0.050 m. The diffusion coefficient and the spatial distribution of the source term are supposed to vary from $D(x) = 54\,\mathrm{W/m\,K}$ and $\mu(x) = 10^5\,\mathrm{W/m^3\,K}$, which result in dimensionless base values of $D_c = 1$ and $\mu_c = 5$ (see Fig. 1), respectively. The base values for the diffusion coefficient and source term distribution are associated with solid–solid phase transformations in steels. The final time is assumed to be 60 s, resulting in a dimensionless value of $t_\mathrm{f} = 0.36$, and 50 measurements are supposed available per sensor.

Before addressing the simultaneous estimation of $D(x)$ and $\mu(x)$, let us consider special cases involving the estimation of either of such functions, by considering the other function as exactly known. Fig. 2 presents the exact and estimated diffusion coefficients obtained with the use of errorless transient measurements of two temperature sensors, for a triangular variation with $\varepsilon_D = 0.5$ and $D_0 = 2D_c$, by assuming $\mu(x) = 5$. It should be pointed out that, qualitatively, the results are not affected by the functional form of $\mu(x)$ if $\mu(x)$ is assumed as exactly known. The sensors were located at each of the boundaries and the initial guess for the iterative procedure of the conjugate gradient method was taken as $D(x) = 0.9$. The results presented in Fig. 2 were obtained with Powell-Beale's version of the conjugate gradient method. We note in this figure that reasonably accurate results were obtained with the measurements of non-intrusive sensors located at the boundaries. Although the peak value of the function could not be exactly recovered, the locations of the discontinuities in the first derivative of $D(x)$ were resolved, despite the fact that the initial guess for $D(x)$ at the boundaries is not changed by the iterative procedure of the conjugate gradient method. The estimated function $D(x)$ is in much better agreement with the exact one when the number of sensors is increased as illustrated in Fig. 3, which was obtained with the errorless measurements of 10 temperature sensors evenly located inside the medium.

On the other hand, difficulties were encountered for the estimation of $\mu(x)$, even when the function for $D(x)$ was regarded as known, as illustrated in Fig. 4 for a triangular variation with $\varepsilon_\mu = 0.5$ and $\mu_0 = 2\mu_c$. The results shown in this figure were obtained with errorless measurements of two non-intrusive sensors, by using Powell-Beale's version of the conjugate gradient method. The initial guess for the conjugate gradient method was taken as $\mu(x) = 4.5$, and the function $D(x) = 1$ was specified. However, as for

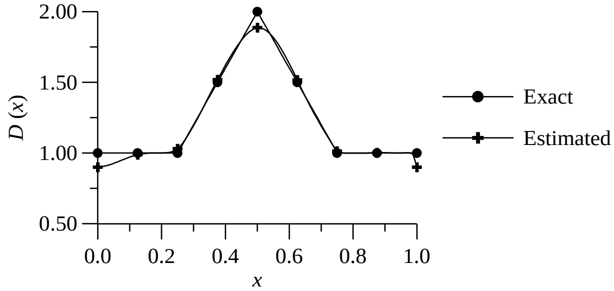

Fig. 3. Estimation of $D(x)$ for known $\mu(x)$ obtained with errorless measurements of 10 sensors equally spaced in the medium.

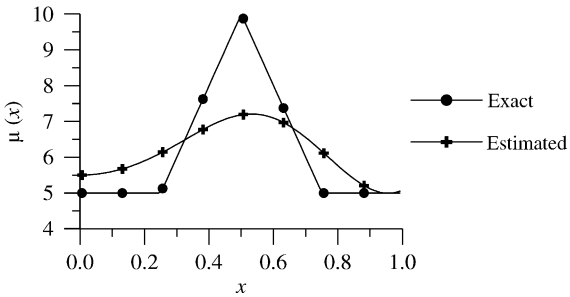

Fig. 4. Estimation of $\mu(x)$ for known $D(x)$ obtained with errorless measurements of two non-intrusive sensors.

the estimation of $D(x)$ when $\mu(x)$ was treated as known, the function specified for $D(x)$ did not affect the results. The estimated function is in quite poor agreement with the exact one (Fig. 4). Qualitatively such results were not affected by increasing the number of sensors, as shown in Fig. 5, which was obtained with the errorless measurements of 10 sensors evenly located in the medium. Such is the case because the measurements are much less affected by changes in $\mu(x)$ than in $D(x)$, because of the low magnitude of the sensitivity function $\Delta U_{\mu}(x, t)$.

We now examine the case of simultaneous estimation of $D(x)$ and $\mu(x)$. Fig. 6 shows the results obtained with the measurements of two non-intrusive sensors, for a step variation of $D(x)$ ($\varepsilon_D = 0.5$ and $D_0 = 2D_c$) and for constant $\mu(x)$ ($\mu_0 = \mu_c$). The results presented in Fig. 6 were obtained with Powell-Beale's version

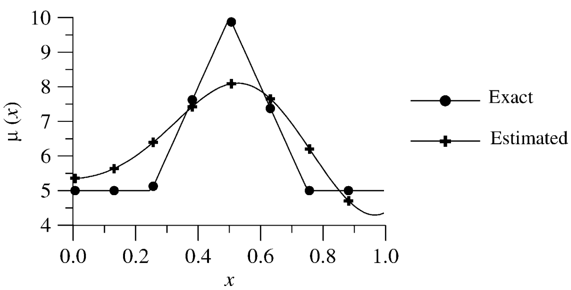

Fig. 5. Estimation of $\mu(x)$ for known $D(x)$ obtained with errorless measurements of 10 sensors equally spaced in the medium.

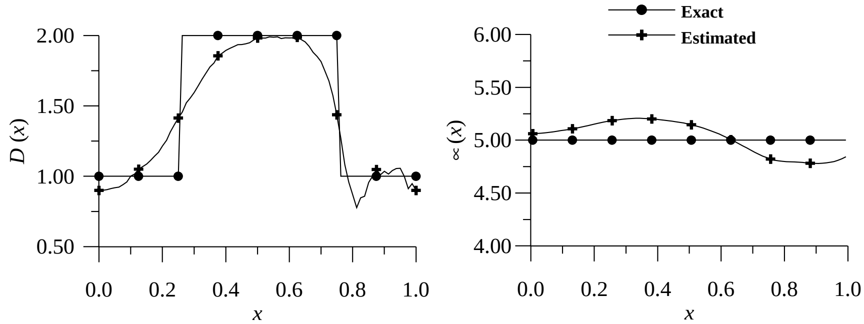

Fig. 6. Simultaneous estimation of $\mu(x)$ and $D(x)$ obtained with measurements of two non-intrusive sensors with standard deviation $\sigma = 0.01Y_{\max}$: results obtained for a step variation of $D(x)$ ($\varepsilon_D = 0.5$ and $D_0 = 2D_c$) and for constant $\mu(x)$ ($\mu_0 = \mu_c$).

of the conjugate gradient method. The simulated measurements in this case contained random errors with standard deviation $\sigma = 0.01Y_{\max}$, where $Y_{\max}$ is the maximum absolute value of the measured variable. The initial guesses used for the iterative procedure of the conjugate gradient method were $D(x) = 0.9$ and $\mu(x) = 4.5$. We note in Fig. 6 that quite accurate results were obtained for such strict test case, involving a discontinuous variation for $D(x)$ and only non-intrusive measurements. Although some blurring is observed near the discontinuity of $D(x)$ at $x = 0.25$, the locations of the discontinuities and the maximum value of the function are quite accurately estimated. Furthermore, the estimated function for $\mu(x)$ oscillates about its constant exact value with an amplitude smaller than the original distance of the initial guess to the exact function. The accuracy of the estimated functions improve when measurements of more sensors are used in the inverse analysis, as illustrated in Fig. 7, which was obtained with measurements containing random errors ($\sigma = 0.01Y_{\max}$) of 10 sensors evenly located inside the medium.

Fig. 8 presents the results obtained for a second-degree polynomial variation for $D(x)$ ($\varepsilon_D = 1$ and $D_0 = 2D_c$) and for constant $\mu(x)$ ($\mu_0 = \mu_c$). The results presented in Fig. 8 were obtained with measurements containing random errors ($\sigma = 0.01Y_{\max}$) of 10 sensors evenly located inside the medium, by using Powell-Beale's version of the conjugate gradient method. A comparison of Figs. 7 and 8 shows

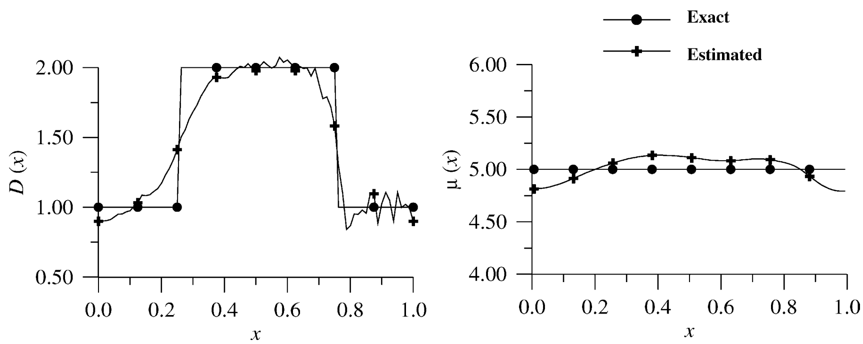

Fig. 7. Simultaneous estimation of $\mu(x)$ and $D(x)$ obtained with measurements of 10 sensors equally spaced in the medium with standard deviation $\sigma = 0.01Y_{\max}$: results obtained for a step variation of $D(x)$ ($\varepsilon_D = 0.5$ and $D_0 = 2D_c$) and for constant $\mu(x)$ ($\mu_0 = \mu_c$).

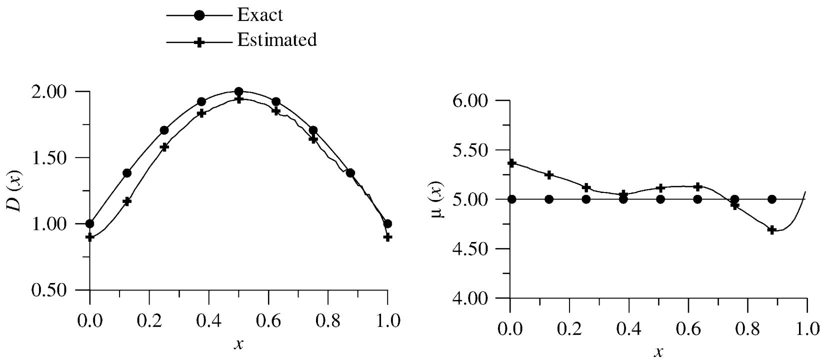

Fig. 8. Simultaneous estimation of $\mu(x)$ and $D(x)$ obtained with measurements of 10 sensors equally spaced in the medium with standard deviation $\sigma = 0.01Y_{\max}$: results obtained for a second-degree polynomial variation for $D(x)$ ($\varepsilon_D = 1$ and $D_0 = 2D_c$) and for constant $\mu(x)$ ($\mu_0 = \mu_c$).

that the accuracy of estimated functions generally improves for continuous and smooth variations of the unknown diffusion coefficient.

Fig. 9 illustrates the results obtained for the simultaneous estimation of $D(x)$ and $\mu(x)$, for a constant exact functional form for $D(x)$ ($D_0 = D_c$) and a triangular variation for $\mu(x)$ ($\varepsilon_\mu = 0.5$ and $\mu_0 = 2\mu_c$). The results presented in Fig. 9 were obtained with measurements containing random errors ($\sigma = 0.01Y_{\max}$) of 10 sensors evenly located inside the medium, by using Powell-Beale's version of the conjugate gradient method. Differently from the results shown above in Figs. 7 and 8, we note in Fig. 9 that the present solution approach fails to estimate the peak value of the exact triangular function for $\mu(x)$. The locations of the discontinuities in the first derivative of the exact function, which characterize the change of $\mu(x)$ from its base value, could not be accurately estimated. Furthermore, the function estimated for $D(x)$ is characterized by large oscillations. We note that, generally, results analogous to those presented in Fig. 9 were obtained whenever the exact function for $\mu(x)$ was not constant. Similar behavior was observed by Hielscher et al. [16] with the use of Polak-Ribiere's version of the conjugate gradient method in an

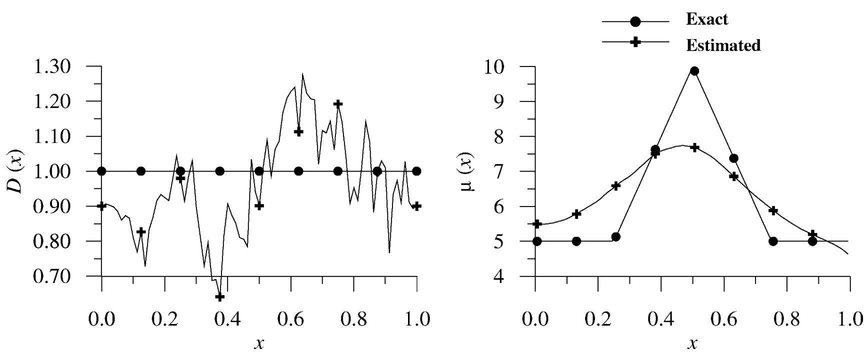

Fig. 9. Simultaneous estimation of $\mu(x)$ and $D(x)$ obtained with measurements of 10 sensors equally spaced in the medium with standard deviation $\sigma = 0.01Y_{\max}$: results obtained for constant $D(x)$ ($D_0 = D_c$) and a triangular variation for $\mu(x)$ ($\varepsilon_\mu = 0.5$ and $\mu_0 = 2\mu_c$).

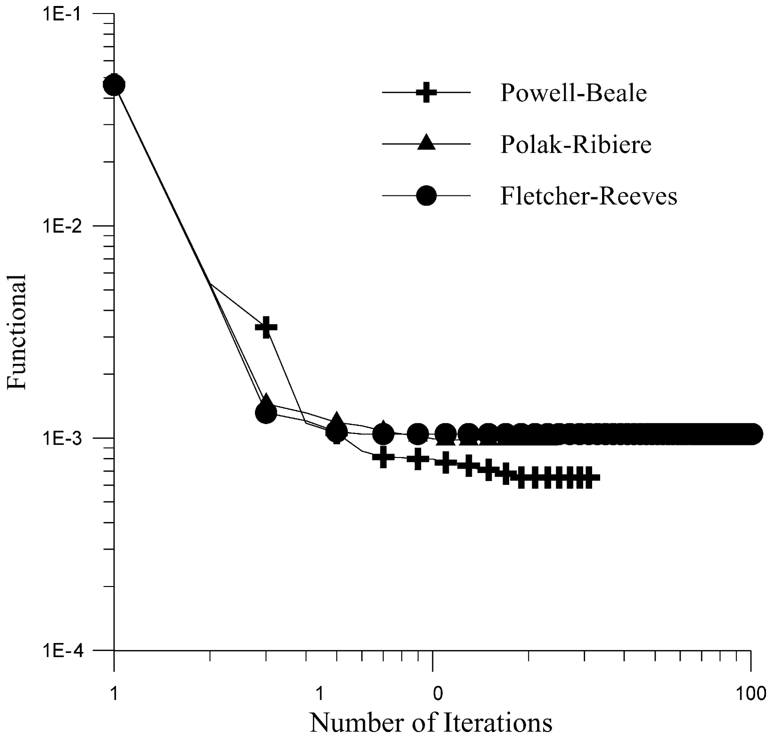

Fig. 10. Comparison of different versions of the conjugate gradient method for the simultaneous estimation of $\mu(x)$ and $D(x)$: results obtained for constant $D(x)$ ($D_0 = D_c$) and a triangular variation for $\mu(x)$ ($\varepsilon_\mu = 0.5$ and $\mu_0 = 2\mu_c$).

optical tomography problem. This is due to the lower sensitivity of the measured variable with respect to $\mu(x)$ as compared to the sensitivity with respect to $D(x)$.

Fig. 10 presents the reduction of the objective functional with respect to the number of iterations, obtained with Powell-Beale's, Polak-Ribiere's and Fletcher-Reeves' versions of the conjugate gradient. The results presented in Fig. 10 correspond to the test case shown in Fig. 9, involving a constant functional form for $D(x)$ ($D_0 = D_c$) and a triangular variation for $\mu(x)$ ($\varepsilon_\mu = 0.5$ and $\mu_0 = 2\mu_c$). Fig. 10 shows that the prescribed tolerance for the iterative procedure of the conjugate gradient method was reached only with Powell-Beale's version; the other two versions did not effectively reduce the objective functional and the iterative procedure was stopped when the specified maximum number of iterations (100) was reached. The results presented in Fig. 10 are representative of those obtained with the other test cases examined in this paper, that is, Powell-Beale's version of the conjugate gradient method resulted in the largest rate of reduction of the objective functional, so that the tolerance prescribed for the stopping criterion was reached in the smallest number of iterations. However, for some test cases the use of Polak-Ribiere's version of the conjugate gradient method resulted in reduction rates for the functional comparable to those obtained with Powell-Beale's version, but unexpected oscillations were observed on the values of the functional. Similar results were reported in Ref. [8], where these three versions of the conjugate gradient method were applied to the estimation of timewise- and spacewise-dependent heat transfer coefficients at the surface of a three-dimensional body.

10. Conclusions

In this article, a function estimation approach based on the conjugate gradient method was applied for the simultaneous estimation of the spatially dependent diffusion coefficient and of the source term distribution, in a one-dimensional nonlinear diffusion problem.

The accuracy of the proposed solution approach was addressed by using simulated transient measurements containing random errors. For the test cases examined, involving heat conduction in a homogeneous steel bar, acceptable results could be obtained with the measurements of two non-intrusive sensors, even for a step variation of the diffusion coefficient, when the source term distribution was constant. On the other hand, the present solution approach was not able to accurately recover the unknown functions for other functional forms of the source term distribution.

A comparison of three versions of the conjugate gradient method was performed, as applied to the solution of the inverse problem under consideration. For the test cases examined in this paper, the use of Powell-Beale's version of the conjugate gradient method resulted in the largest rates of reduction of the objective functional. Generally, Polak-Ribiere's and Fletcher-Reeves' versions of the conjugate gradient method did not effectively reduce the objective functional.

References

[1] J.V. Beck, Sequential methods in parameter estimation, in: Proceedings of the Third International Conference on Inverse Problems in Engineering, Tutorial Session, Port Ludlow, WA, 1999.
[2] O.M. Alifanov, Inverse Heat Transfer Problems, Springer, Berlin, 1994.
[3] M.N. Ozisik, H.R.B. Orlande, Inverse Heat Transfer: Fundamentals and Applications, Taylor & Francis, New York, 2000.
[4] J.V. Beck, B. Blackwell, C.R. St. Clair, Inverse Heat Conduction: Ill-Posed Problems, Wiley Interscience, New York, 1985.
[5] J. Gottlieb, P. Duchateau, Parameter identification and inverse problems in hydrology, geology and ecology, in: Water Science and Technology Library, vol. 23, Kluwer Academic Publishers, Dordrecht, 1996.
[6] A. Kirsch, An Introduction to the Mathematical Theory of Inverse Problems, Springer, New York, 1996.
[7] H.R.B. Orlande, M.N. Ozisik, Determination of the reaction function in a reaction-diffusion parabolic problem, ASME J. Heat Transfer 116 (1994) 1041–1044.
[8] M.J. Colaço, H.R.B. Orlande, Comparison of different versions of the conjugate gradient method of function estimation, Numer. Heat Transfer A 36 (1999) 229–249.
[9] G.S. Dulikravich, T.J. Martin, Inverse shape and boundary condition problems and optimization, in: W.J. Minkowycz, E.M. Sparrow (Eds.), Heat Conduction: Advances in Numerical Heat Transfer, vol. 1, Taylor & Francis, 1996, Chapter 10, pp. 381–426.
[10] T.J. Martin, G.S. Dulikravich, Inverse determination of boundary conditions in steady heat conduction with heat generation, ASME J. Heat Transfer 118 (1996) 546–554.
[11] N.-Z. Sun, Inverse Problems in Groundwater Modeling, Kluwer Academic Publishers, Dordrecht, 1994.
[12] N.-Z. Sun, F. Tsai, W. Yeh, Parameter structure identifiability and experimental design in groundwater modeling, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 1, e-Papers, Rio de Janeiro, 2002, pp. 37–48.
[13] D. Lesnic, L. Eliott, D.B. Ingham, B. Clennel, R.J. Knipe, A mathematical model and numerical investigation for determining the hydraulic conductivity of rocks, Int. J. Rock Mech. Miner. Sci. 34 (1997) 741–759.
[14] F.-X. Le Dimet, P. Ngnepieba, Second-order methods for inverse problems: an application in hydrology, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 1, e-Papers, Rio de Janeiro, 2002, pp. 129–135.
[15] K. Onishi, K. Yasuhara, S. Murakami, Y. Ohura, K. Iijima, Identification of aquifer transmissivity from interior point observation, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 1, e-Papers, Rio de Janeiro, 2002, pp. 349–356.
[16] A.H. Hielscher, A.D. Klose, K.M. Hanson, Gradient-based iterative image reconstruction scheme for time-resolved optical tomography, IEEE Trans. Med. Imaging 18 (1999) 262–271.

[17] D. Boas, T. Gaudette, G. Strangman, X. Cheng, J. Marota, J. Mandeville, The accuracy of near infrared spectroscopy and imaging during focal changes in cerebral hemodynamics, Neuroimage 13 (2001) 76–90.
[18] R. Roy, E. Muraca, Three-dimensional unconstrained and constrained image-reconstruction techniques applied to fluorescence, frequency-domain photon migration, Appl. Opt. 40 (2001) 2206–2215.
[19] M. Holboke, B. Tomberg, X. Li, N. Shah, J. Fishkin, D. Kidney, J. Butler, B. Chance, A. Yodh, Three-dimensional diffusive optical mammography with ultrasound localization in a human object, J. Biomed. Opt. 5 (2000) 237–247.
[20] Q. Zhu, E. Conant, B. Chance, Optical imaging as an adjunct to sonograph in differentiating benign from malignant breast lesions, J. Biomed. Opt. 5 (2000) 229–236.
[21] T. Mcbride, B. Pogue, S. Jiang, U. Osterberg, K. Paulsen, Initial studies of in vivo absorbing and scattering heterogeneity in near-infrared tomographic breast imaging, Opt. Lett. 26 (2001) 822–824.
[22] J. Culver, V. Ntziachristos, M. Holboke, A. Yodh, Optimization of optode arrangements for diffusive optical tomography: a singular-value analysis, Opt. Lett. 26 (2001) 701–703.
[23] V. Ntziachristos, R. Weissleder, Experimental three-dimensional fluorescence reconstruction of diffusive media by use of a normalized born approximation, Opt. Lett. 26 (2001) 893–895.
[24] R. Kress, Electromagnetic waves scattering: scattering by obstacles, in: P. Sabatier (Ed.), Scattering, Academic Press, London, 2001, pp. 191–210.
[25] A. Kirsch, R. Kress, Uniqueness in inverse obstacle problems, Inverse Probl. 9 (1993) 285–299.
[26] R. Pottahast, Point-Sources and Multipoles in Inverse Scattering Theory, Chapman & Hall, London, 2001.
[27] K.Y. Kim, S. Kang, S. Kim, M. Kim, S. Kang, J. Lee, Dynamic impedance imaging of binary-mixture fields with external and internal electrodes, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 2, e-Papers, Rio de Janeiro, 2002, pp. 413–420.
[28] A. Seppanen, M. Vauhkonem, J. Kaipio, E. Somersalo, Inference of velocity fields based on tomographic measurements in process industry, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 2, e-Papers, Rio de Janeiro, 2002, pp. 453–460.
[29] A. Fraguela, C. Gamio, D. Hinestroza, An inverse algorithm for capacitance tomography of two-phase flow regimes, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 2, e-Papers, Rio de Janeiro, 2002, pp. 461–468.
[30] V. Rolnik, P. Seleghim, On-site calibration of a phase fraction meter by an inverse technique, in: H.R.B. Orlande (Ed.), Inverse Problems in Engineering: Theory and Practice, vol. 2, e-Papers, Rio de Janeiro, 2002, pp. 363–370.
[31] M. Cheney, D. Isaacson, J. Newell, Electrical impedance tomography, SIAM Rev. 41 (1999) 85–101.
[32] M.J.D. Powell, Restart procedures for the conjugate gradient method, Math. Program. 12 (1977) 241–254.