Proceedings of IMECE'03
2003 ASME International Mechanical Engineering Congress & Exposition
Washington, D.C., USA, November 16-21, 2003

IMECE2003-42058

A COMPARISON OF TWO SOLUTION TECHNIQUES FOR THE INVERSE PROBLEM OF SIMULTANEOUSLY ESTIMATING THE SPATIAL VARIATIONS OF DIFFUSION COEFFICIENTS AND SOURCE TERMS

Marcelo J. Colaço[1]
Department of Mechanical Engineering, EE/COPPE
Federal University of Rio de Janeiro, UFRJ
Cid Universitaria, Cx. Postal 68503
Rio de Janeiro, RJ, 21945-970
BRAZIL
colaco@ufrj.br

Helcio R. B. Orlande[2]
Department of Mechanical Engineering, EE/COPPE
Federal University of Rio de Janeiro, UFRJ
Cid Universitaria, Cx. Postal 68503
Rio de Janeiro, RJ, 21945-970
BRAZIL
helcio@serv.com.ufrj.br

George S. Dulikravich[3]
Department of Mechanical and Aerospace Engineering
Multidisciplinary Analysis, Inverse Design & Optimization (MAIDO) Institute, UTA Box 19018
The University of Texas at Arlington
Arlington, TX 76019, U.S.A.
gsd@uta.edu

Fabio A. Rodrigues[4]
Department of Mechanical Engineering, EE/COPPE
Federal University of Rio de Janeiro, UFRJ
Cid Universitaria, Cx. Postal 68503
Rio de Janeiro, RJ, 21945-970
BRAZIL
fabio@lmt.coppe.ufrj.br

ABSTRACT

This work deals with the simultaneous estimation of the spatially varying diffusion coefficient and of the source term distribution in a one-dimensional nonlinear diffusion problem. This work can be physically associated with the detection of material non-homogeneities such as inclusions, obstacles or cracks, heat conduction, groundwater flow detection, and tomography. Two solution techniques are applied in this paper to the inverse problem under consideration, namely: the conjugate gradient method with adjoint problem and a hybrid optimization algorithm. The hybrid optimization technique incorporates several of the most popular optimization modules; the Davidon-Fletcher-Powell (DFP) gradient method, a genetic algorithm (GA), the Nelder-Mead (NM) simplex method, quasi-Newton algorithm of Pshenichny-Danilin (LM), differential evolution (DE), and sequential quadratic programming (SQP). The accuracy of the two solution approaches was examined by using simulated transient measurements containing random errors in the inverse analysis.

[1] Postdoctoral Fellow. Lecturer. Member of ASME.
[2] Associate Professor.
[3] Professor and Director of MAIDO. Fellow of ASME. Currently Professor and Chair, Mech. & Materials Eng. Dept, Florida International University, Miami, Florida.
[4] Graduate Research Assistant.

NOMENCLATURE

$D(x)$	spatially dependent diffusion coefficient
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,..., M

Greek symbols

α_1	first order regularization parameter
β_D^k , β_μ^k	search step sizes for $D(x)$ and $\mu(x)$, respectively
$\Delta D(x)$, $\Delta\mu(x)$	variations in $D(x)$ and $\mu(x)$, respectively
$\Delta U_D(x,t)$, $\Delta U_\mu(x,t)$	sensitivity functions for $D(x)$ and $\mu(x)$, respectively
γ_D^k, γ_μ^k, ψ_D^k , ψ_μ^k	conjugation coefficients
$\lambda(x,t)$	Lagrange multiplier
$\mu(x)$	spatial distribution of the source term
σ	standard deviation of the measurements
∇S	gradient components for the functional

Subscripts

m	refers to the number of sensors, m=1,.., M

Superscripts

k	iteration number
*	dimensional quantities

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 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 [1-36]).

In this work we present a comparison of two solution techniques, as applied to the inverse problem of simultaneous estimation of spatially dependent diffusion coefficients and source terms in a non-linear diffusion problem. The techniques examined here are the conjugate gradient method together with an adjoint problem formulation [2,3,6,9,13,14] and a hybrid optimization algorithm [7,8,31-35]. The conjugate gradient method with adjoint problem is a function estimation approach [1-4], where no information is *a priori* assumed 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]. The hybrid optimization technique is a parameter estimation scheme, which consists of the use of the Davidon-Fletcher-Powell (DFP) gradient method, a genetic algorithm (GA), the Nelder-Mead (NM) simplex method, quasi-Newton algorithm of Pshenichny-Danilin (LM), differential evolution (DE), and sequential quadratic programming (SQP) [31]. It is robust, stable and can avoid local minima [7,8,31-35]. Highlights of the basic steps of these two solution techniques are presented below, as applied to the inverse problem of interest.

DIRECT PROBLEM

In this work we consider problems governed by the following non-linear diffusion equation

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

where $\mathbf{r}^*$ denotes the vector of coordinates and the superscript * denotes dimensional quantities.

Equation (1) can be used for the modeling of several physical phenomena, such as heat conduction [1-8], groundwater flow [9-13] and tomography [14-29]. We focus our attention here to a one-dimensional version of equation (1) written in dimensionless form as

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

and subject to the following boundary and initial conditions.

$$\frac{\partial U}{\partial x} = 0 \qquad \text{at } x=0 \text{ for } t>0 \qquad (2.b)$$

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

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

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 $U(x,t)$.

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. These measurement errors 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.

The basic steps of the two solution techniques examined in this work are discussed below.

CONJUGATE GRADIENT METHOD WITH ADJOINT PROBLEM

For the simultaneous estimation of the functions $D(x)$ and $\mu(x)$ with the conjugate gradient method with adjoint problem we consider the minimization of the following objective functional

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

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

The use of the conjugate gradient method with an adjoint problem for the minimization of the objective functional (3)

requires the solution of auxiliary problems, known as *sensitivity and adjoint 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 $\varepsilon\Delta D(x)$, where ε is a real number. The sensitivity problem for $D(x)$ 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 equation (4) results in the following sensitivity problem.

$$\frac{\partial \Delta U_D}{\partial t}=\frac{\partial}{\partial x}\left(D(x)\frac{\partial \Delta U_D}{\partial x}+\Delta D(x)\frac{\partial U}{\partial x}\right)+\mu(x)\Delta U_D$$

$$\text{in } 0<x<1 \quad \textit{for } t>0 \tag{5.a}$$

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

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

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

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

$$\frac{\partial \Delta U_\mu}{\partial t}=\frac{\partial}{\partial x}\left(D(x)\frac{\partial \Delta U_\mu}{\partial x}\right)+\mu(x)\Delta U_\mu+\Delta\mu(x)U$$

$$\text{in } 0<x<1 \quad \text{for } t>0 \tag{6.a}$$

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

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

A Lagrange multiplier $\lambda(x,t)$ is utilized in the minimization of the functional (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 equations (5.a-d) and (6.a-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-d) and (6.a-d), is obtained.

In order to derive the adjoint problem, the governing equation of the direct problem, equation (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 (3). The following extended functional is obtained

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

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 (7) written for perturbed $D(x)$ and $\mu(x)$, respectively.

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

$$-\frac{\partial \lambda}{\partial t}-\frac{\partial}{\partial x}\left(D(x)\frac{\partial \lambda}{\partial x}\right)-\mu(x)\lambda+\sum_{m=1}^{M}[U-Y]\delta(x-x_m)=0$$

$$\text{in } 0<x<1 \ , \quad \text{for } t>0 \tag{9.a}$$

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

$$\lambda=0 \qquad \text{in } 0\le x\le 1 \quad \text{for } t=t_f \tag{9.d}$$

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.

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

$$\Delta S_\mu[D,\mu] = -\int_{x=0}^{1}\int_{t=0}^{t_f} \Delta\mu(x)\lambda(x,t)U(x,t)dt\,dx \tag{10.b}$$

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]:

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

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

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

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

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

An analysis of equations (12.a) and (9.b.c) 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.

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; and 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

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

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

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 equations (14.a,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 [30].

Different versions of the conjugate gradient method can be found in the literature, depending on how the conjugation coefficients are computed. In this work we use the so-called Powell-Beale's version of the conjugate gradient method because of its superior robustness and convergence rate in non-linear problems [6,30]. The conjugation coefficients for this version of the conjugate gradient method are given by [6,30]

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

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

$$\psi_D^k = \frac{\int_{x=0}^{1}\left\{\nabla S[D^{qD+1}(x)] - \nabla S[D^{qD}(x)]\right\}\nabla S[D^k(x)]\,dx}{\int_{x=0}^{1}\left\{\nabla S[D^{qD+1}(x)] - \nabla S[D^{qD}(x)]\right\}d_D^{qD}(x)\,dx} \tag{16.a}$$

$$\psi_\mu^k = \frac{\int_{x=0}^{1}\left\{\nabla S[\mu^{q\mu+1}(x)] - \nabla S[\mu^{q\mu}(x)]\right\}\nabla S[\mu^k(x)]\,dx}{\int_{x=0}^{1}\left\{\nabla S[\mu^{q\mu+1}(x)] - \nabla S[\mu^{q\mu}(x)]\right\}d_\mu^{q\mu}(x)\,dx} \tag{16.b}$$

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 [6, 30].

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 $\mu(x)$, equations (13.a,b), respectively, are obtained by minimizing the objective functional at each iteration along the

specified directions of descent. If the objective functional given by equation (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]

$$\beta_d^k = \frac{F_1 A_{22} - F_2 A_{12}}{A_{11} A_{22} - A_{12}^2} \quad ; \quad \beta_\mu^k = \frac{F_2 A_{11} - F_1 A_{12}}{A_{11} A_{22} - A_{12}^2} \tag{17.a,b}$$

where

$$A_{11} = \int_{t=0}^{t_f} \sum_{m=1}^{M} [\Delta U_D^k(x_m, t)]^2 dt \tag{18.a}$$

$$A_{22} = \int_{t=0}^{t_f} \sum_{m=1}^{M} [\Delta U_\mu^k(x_m, t)]^2 dt \tag{18.b}$$

$$A_{12} = \int_{t=0}^{t_f} \sum_{m=1}^{M} \Delta U_D^k(x_m, t) \Delta U_\mu^k(x_m, t) dt \tag{18.c}$$

$$F_1 = \int_{t=0}^{t_f} \sum_{m=1}^{M} [Y_m^k - U^k(x_m, t)][\Delta U_D^k(x_m, t)] dt \tag{18.d}$$

$$F_2 = \int_{t=0}^{t_f} \sum_{m=1}^{M} [Y_m^k - U^k(x_m, t)][\Delta U_\mu^k(x_m, t)] dt \tag{18.e}$$

In equations (18.a-e), $\Delta U_D^k(x,t)$ and $\Delta U_\mu^k(x,t)$ are the solutions of the sensitivity problems given by equations (5.a-d) and (6.a-d), respectively, obtained by setting $\Delta D^k(x) = d_D^k(x)$ and $\Delta \mu^k(x) = d_\mu^k(x)$.

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 adapted from those found in reference [3]. The conjugate gradient method of function estimation belongs to the class of *iterative regularization methods* [2]. For this class of methods, the stopping criterion for the computational procedure is chosen 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 use in this work the *discrepancy principle* [2].

HYBRID OPTIMIZATION APPROACH

The hybrid optimization algorithm [31-35] utilized in this work incorporates some of the most popular optimization algorithms: genetic algorithm, a quasi-Newton method, modified Nelder-Mead simplex method, sequential quadratic programming, Davidon-Fletcher-Powell gradient search algorithm and differential evolution. Each technique provides a unique approach to optimization with varying degrees of convergence, reliability and robustness at different cycles during the iterative optimization procedure. A set of analytically formulated rules and switching criteria were coded into the program to automatically switch back and forth among the different optimization algorithms as the iterative minimization process proceeded.

The evolutionary hybrid algorithm handles the existence of equality and inequality constraint functions in three ways: Rosen's projection method, feasible searching, and random design generation. Rosen's projection method provided search directions that guided descent-directions tangent to active constraint boundaries. In the feasible search, designs that violated constraints were automatically restored to feasibility via the minimization of the active global constraint functions. If at any time this constraint minimization failed, random designs were generated about the current design until a new feasible design was reached.

Gradients of the objective and constraint functions with respect to the design variables, also called design sensitivities, were calculated using finite differencing formulas. The population matrix was updated every iteration with new designs and ranked according to the value of the objective function. During the optimization process, local minima can occur and halt the process before achieving an optimal solution. In this case, the optimizer switches to another method. The user can also stop the iterative process, switch manually to another method and restart the optimizer from the previous iteration.

The population matrix was updated every iteration with new designs and ranked according to the value of the objective function. The optimization problem was completed when the maximum number of iterations or objective function evaluations were exceeded, or when the optimization program tried all individual optimization algorithms but failed to produce a non-negligible decrease in the objective function. The latter criterion was the primary qualification of convergence and it usually indicated that a global minimum had been found.

The hybrid optimization algorithm was applied to estimate the unknown parameters, correspondent to the values of the functions $D(x)$ and $\mu(x)$ at each of the grid points. Tikhonov's first-order regularization technique [5] was used in this case, in order to avoid solution instabilities resulting from the ill-posed character of the inverse problem. Note that we have also examined the use of zero-th and second order regularization schemes, but the best results were obtained with first order regularization. Then, the objective functional given by equation (3) was rewritten to accommodate the first order regularization term as

$$S[D(x), \mu(x)] = \frac{1}{2} \int_{t=0}^{t_f} \sum_{m=1}^{M} \{U[x_m, t; D(x), \mu(x)] - Y_m(t)\}^2 dt + \alpha_1 \sum_{m=1}^{M-1} \left\{ [D_{i+1}(x) - D_i(x)]^2 + [\mu_{i+1}(x) - \mu_i(x)]^2 \right\} \tag{19}$$

where α_1 is the first order regularization parameter, which was assumed to be the same for the unknown functions $D(x)$ and $\mu(x)$. In equation (19), the first term on the RHS is the original

functional given by equation (3), while the second term on the RHS is the first order regularization penalty function, which minimizes the oscillations on the estimated functions.

RESULTS AND DISCUSSIONS

The accuracies of the solution methods described above were examined by using simulated transient measurements containing random errors in the inverse analysis. Functions containing discontinuities, which are the most difficult to be recovered by the inverse analysis, were used to generate the simulated measurements.

For the test cases examined, the direct, sensitivity and adjoint problems were numerically solved with 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 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 base values of $D(x) = 54\ W\ m^{-1}\ K^{-1}$ and $\mu(x) = 10^5\ W\ m^{-3}\ K^{-1}$, which result in dimensionless base values of $D_c = 1$ and $\mu_c = 5$, 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 seconds, resulting in a dimensionless value of $t_f = 0.36$, and 50 measurements are supposed available per temperature sensor.

Figure 1 shows the results obtained with the conjugate gradient method and the measurements of two non-intrusive sensors, for a step variation of $D(x)$ and for constant $\mu(x)$. The simulated measurements in this case contained random errors with standard deviation $\sigma = 0.01\ Y_{max}$, 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 and for the hybrid optimizer were $D(x) = 0.9$ and $\mu(x) = 4.5$. We note in Figure 1 that quite good results were obtained for such a strict test case involving a discontinuous variation for $D(x)$, even with only two non-intrusive sensors, by using the conjugate gradient method of function estimation.

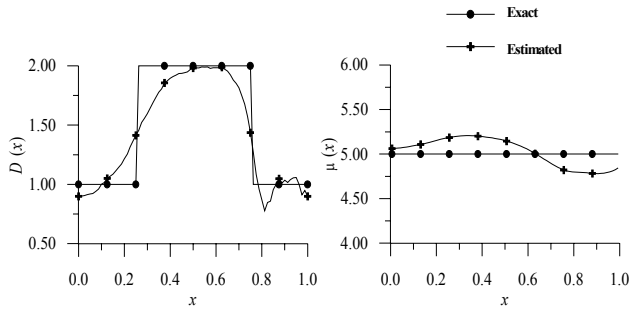

Figure 1 Estimation of $\mu(x)$ and $D(x)$ obtained by the CGM with two non-intrusive sensors with standard deviation ($\sigma = 0.01\ Y_{max}$).

Figure 2 shows the results obtained for the same test case shown in Figure 1, but with 10 sensors. We note in Figure 2 that more accurate results are obtained for both $D(x)$ and $\mu(x)$ when more sensors are used in the inverse analysis.

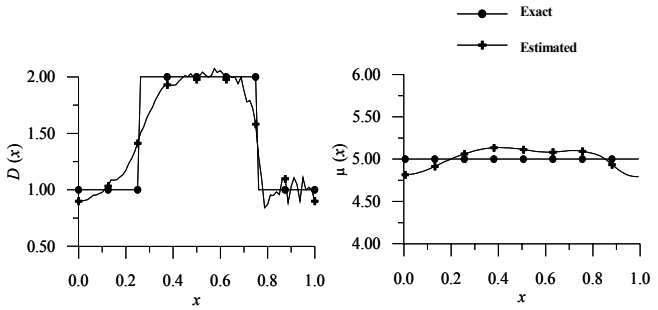

Figure 2. Estimation of $\mu(x)$ and $D(x)$ obtained by the CGM with ten sensors with standard deviation ($\sigma = 0.01\ Y_{max}$).

Figure 3 shows results similar to those presented in Figure 1, obtained with the hybrid optimization algorithm using 2 non-intrusive sensors, with the regularization parameter α_1 appearing in equation (19) set as zero. The *x*-axis in this figure represents the i^{th} value of the control volume, that is, $i = 80$ represents $x = 1.0$. Note that the estimated parameters suffer from large oscillations, resulting from the ill-posed character of the inverse problem.

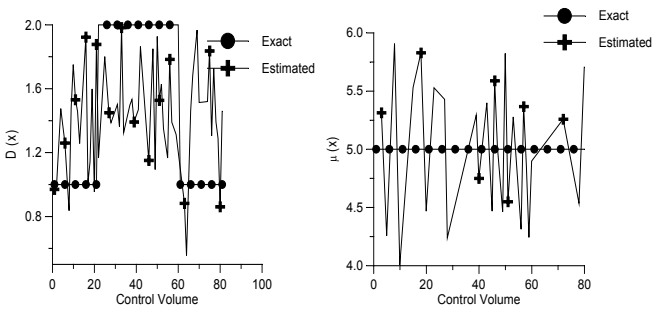

Figure 3. Estimation of $\mu(x)$ and $D(x)$ by the hybrid optimizer without regularization and with 2 sensors ($\sigma = 0.01\ Y_{max}$).

In order to find the best value of the regularization parameter α_1, we used the L-shape curve [36] as shown in Figure 4. In this figure, the *x*-axis represents the second term appearing on the RHS of equation (19) and the *y*-axis represents the first term appearing on the RHS of equation (19). The best choice for α_1 is the one that minimizes both terms represented by the *x* and *y*-axis [36]. Note that the first term appearing in equation (19) is the original functional given by equation (3), while the second term is the regularization penalty function. The optimum value for α_1 in this case was 0.0001.

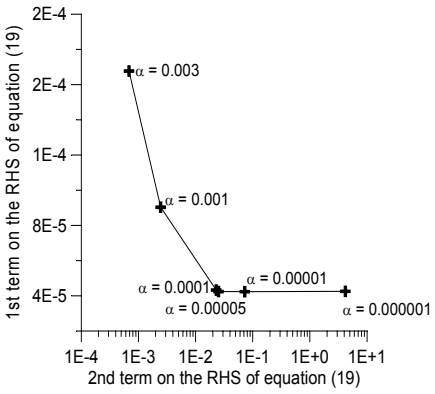

Figure 4. L-shape curve for choice of the regularization parameter using 2 sensors.

Figure 5 shows the estimated functions obtained with such value of the regularization parameter and the hybrid optimization approach, with the measurements of two non-intrusive sensors. Note that the oscillations are eliminated because of the stabilization introduced by the first order regularization term. However, the estimated function for $D(x)$ is in very bad agreement with the exact one. This is probably due to the fact that this case, involving 50 measurements per sensor and only two sensors, is underdetermined, that is, the number of unknown parameters is less than the number of measurements. Note that for the hybrid optimization approach, 2 parameters are estimated for each of the control-volumes used for the discretization of the domain, corresponding to the values of $D(x)$ and $\mu(x)$ at the control volume. Therefore, a total of 160 parameters were estimated in this case. A comparison of the functions estimated for $D(x)$ with the conjugate gradient method and with the hybrid optimization approach by using only two sensors (see figures 1 and 5, respectively) shows that the conjugate gradient method is not as sensitive to the fact that the problem is undetermined. This is probably because of the fact that the measured data is used in the source-function term of the adjoint problem in the function estimation approach with the conjugate gradient method, so that the information from the sensors at the boundaries is extended to the rest of the domain through the adjoint function $\lambda(x,t)$. On the other hand, the accuracies of the functions estimated for $\mu(x)$ with the conjugate gradient method and with the hybrid optimization approach, with the measurements of two non-intrusive sensors (see figures 1 and 5, respectively), are similar. This is due to the fact that the exact $\mu(x)$ is constant and the initial guess used for both approaches was relatively near the exact function.

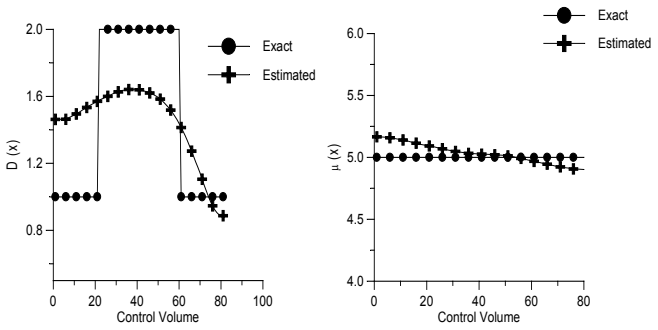

Figure 5. Estimation of $\mu(x)$ and $D(x)$ by the hybrid optimizer with regularization and with 2 sensors ($\sigma = 0.01\ Y_{max}$).

Figure 6 shows the estimated functions obtained with the hybrid optimization approach and the measurements of 10 sensors evenly spread in the medium. For such a case, no regularization was used, that is, $\alpha_1 = 0$. One can see the large oscillations in the estimated functions, but it is interesting to note that, even without any regularization, the hybrid optimization approach is capable of locating the discontinuities in $D(x)$.

The L-shape curve [36] was also used in this case to choose the regularization parameter, as shown in Figure 7, where we again choose a value of 0.0001 for the regularization parameter. Figure 8 shows the results obtained with $\alpha_1 = 0.0001$. Note that the oscillations are reduced and the function estimated for $D(x)$ is in much better agreement with the exact one, than that estimated with only two sensors (Figure 5). It is interesting to note that the function estimated for $\mu(x)$ with 10 sensors is in much better agreement with the exact one than that obtained with 2 sensors (see figure 5).

A comparison of figures 2 and 8 shows that similar results are obtained for the simultaneous estimation of $D(x)$ and $\mu(x)$, by using either the conjugate gradient method or the hybrid optimization approach, when ten sensors are used in the inverse analysis.

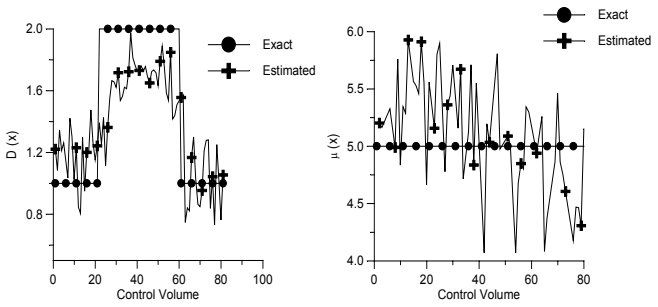

Figure 6. Estimation of $\mu(x)$ and $D(x)$ by the hybrid optimizer without regularization and with 10 sensors ($\sigma = 0.01\ Y_{max}$).

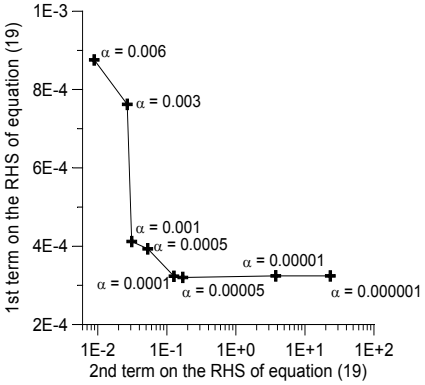

Figure 7. L-shape curve for choice of the regularization parameter using 10 sensors.

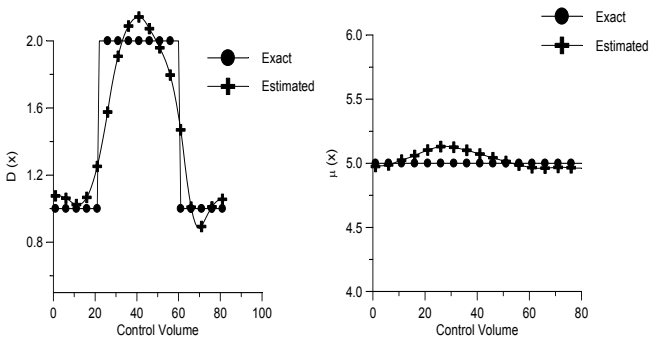

Figure 8. Estimation of $\mu(x)$ and $D(x)$ by the hybrid optimizer with regularization and with 10 sensors ($\sigma = 0.01\ Y_{max}$).

We note that the use of the regularization parameter, based on the Tikhonov's technique, produced the regularization necessary to obtain stable results for the estimation of both functions with the hybrid optimization approach. In fact, completely unstable results were obtained if the regularization technique was not used, as a result of the ill-posed character of the inverse problem under picture. Also, the results presented above show that the two solution approaches examined in this paper are not sensitive to measurement errors. In fact, qualitatively similar results were obtained by using errorless simulated measurements.

Finally, Figure 9 shows the convergence history of the hybrid optimizer for the estimation of the functions presented in Figure 8, where one can see that no further reduction in the cost function is obtained after approximately 500 iterations.

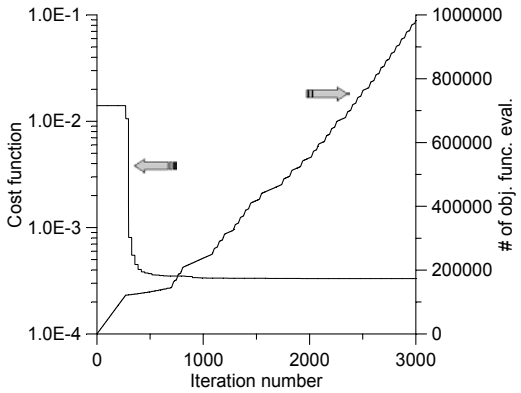

Figure 9. Convergence history for the simultaneous estimation of $\mu(x)$ and $D(x)$ by the hybrid optimizer with regularization and with 10 sensors.

CONCLUSIONS

In this article we compared the conjugate gradient method and a hybrid optimization technique, as applied to the simultaneous estimation of the spatially dependent diffusion coefficient and of the source term distribution, in a one-dimensional non-linear diffusion problem.

Quite accurate results could be obtained with the conjugate gradient method by using the measurements of ten sensors, even for a step variation of the diffusion coefficient, when the source term distribution was constant. The hybrid optimization technique produced equivalent results to those obtained with the conjugate gradient method, with the same number of sensors. On the other hand, with the use of only two sensors, when the estimation problem examined in this work was underdetermined, the results obtained with the conjugate gradient method of function estimation were in much better agreement with the exact functions than those obtained with the hybrid optimization approach.

The use of the regularization technique based on Tikhonov's regularization scheme produced the regularization necessary to obtain stable results for the estimation of both functions with the hybrid optimization approach. Similarly, the use of the discrepancy principle to select the stopping criterion for the iterative procedure of the conjugate gradient method resulted on stable solutions for both functions.

AKNOWLEDGEMENTS

This work was partially funded with resources from FAPERJ, CAPES and CNPq, Brazilian agencies for scientific and technological development.

The third author is grateful for the partial support provided for this research on the grant NSF DMS-0073698 administered through the Computational Mathematics program.

REFERENCES

1. Beck, J. V., 1999, "Sequential Methods in Parameter Estimation", 3rd International Conference on Inverse Problems in Engineering, Tutorial Session, Port Ludlow, WA.

2. Alifanov, O. M., 1994, *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin.
3. Ozisik, M. N. and Orlande, H. R. B., 2000, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor & Francis, New York.
4. Beck, J.V., Blackwell, B. and St. Clair, C. R., 1985, *Inverse Heat Conduction: Ill-Posed Problems*, Wiley Interscience, New York.
5. Tikhonov, A. N. and Arsenin, V. Y., 1977, *Solution of Ill-Posed Problems*, Winston & Sons, Washington, DC.
6. Colaço, M. J. and Orlande, H. R. B., 1999, "Comparison of Different Versions of the Conjugate Gradient Method of Function Estimation", *Numerical Heat Transfer*, Vol. 36, Part A, pp. 229-249.
7. Dulikravich, G. S. and Martin, T. J., 1996, "Inverse Shape and Boundary Condition Problems and Optimization in Heat Conduction", *Chapter 10 in Advances in Numerical Heat Transfer,* Vol. 1, 381-426, Minkowycz, W. J. and Sparrow, E. M. (eds.), Taylor and Francis.
8. Martin, T. J. and Dulikravich, G. S., 1996, "Inverse Determination of Boundary Conditions in Steady Heat Conduction with Heat Generation", *ASME J. Heat Transfer*, Vol. 118, 546-554.
9. Sun, N-Z., *Inverse Problems in Groundwater Modeling*, Kluwer Academic Publishers, Dordrecht, 1994.
10. Sun, N-Z., Tsai, F. and Yeh, W., 2002, "Parameter Structure Identifiability and Experimental Design In Groundwater Modeling", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
11. Lesnic, D., Eliott, L., Ingham, D.B., Clennel, B. and Knipe, R. J., 1997, "A Mathematical Model and Numerical Investigation for Determining the Hydraulic Conductivity of Rocks", *Int. J. Rock. Mech. Min. Sci*., vol. 34, pp. 741-759.
12. Le Dimet, F-X. and Ngnepieba, P., 2002, "Second-order Methods for Inverse Problems: An Application in Hydrology", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
13. Onishi, K., Yasuhara, K., Murakami, S., Ohura, Y. and Iijima, K., 2002, "Identification of Aquifer Transmissivity from Interior Point Observation", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
14. Hielscher, A. H., Klose, A. D. and Hanson, K. M., 1999, "Gradient-Based Iterative Image Reconstruction Scheme for Time-Resolved Optical Tomography", *IEEE Trans. Med. Imaging*, Vol. 18, pp. 262-271.
15. Boas, D., Gaudette, T., Strangman, G., Cheng, X., Marota, J. and Mandeville, J., 2001, "The Accuracy of Near Infrared Spectroscopy and Imaging During Focal Changes in Cerebral Hemodynamic"s, *Neuroimage*, vol. 13, pp. 76-90.
16. Roy, R. and Muraca, E., 2001, "Three-dimensional Unconstrained and Constrained Image-Reconstruction Techniques Applied to Fluorescence, Frequency-Domain Photon Migration", *Applied Optics*, vol. 40, pp.2206-2215.
17. Holboke, M., Tomberg, B., Li, X., Shah, N., Fishkin, J., Kidney, D., Butler, J., Chance, B. and Yodh, A., 2000, "Three-dimensional Diffusive Optical Mammography with Ultrasound Localization in a Human Object", *J. Biomedical Optics*, vol. 5., pp. 237-247.
18. Zhu, Q., Conant, E. and Chance, B., 2000, "Optical Imaging as an Adjunct to Sonograph in Differentiating Benign from Malignant Breast Lesions", *J. Biomedical Optics*, vol. 5., pp. 229-236.
19. Mcbride, T., Pogue, B., Jiang, S., Osterberg, U. and Paulsen, K., 2001, "Initial Studies of *in vivo* Absorbing and Scattering Heterogeneity in Near-infrared Tomographic Breast Imaging", *Optical Letters*, vol. 26, pp. 822-824.
20. Culver, J., Ntziachristos, V., Holboke, M. and Yodh, A., 2001, "Optimization of Optode Arrangements for Diffusive Optical Tomography: A Singular-Value Analysis", *Optical Letters*, vol. 26, pp. 701-703.
21. Ntziachristos, V. and Weissleder, R., 2001, "Experimental Three-dimensional Fluorescence Reconstruction of Diffusive Media by Use of a Normalized Born Approximation", *Optical Letters*, vol. 26, pp. 893-895.
22. Kress, R., 2001, Electromagnetic Waves Scattering: Scattering by Obstacles, in *Scattering*, P. Sabatier (ed.), pp. 191-210, Academic Press, London.
23. Kirsch, A. and Kress, R., 1993, "Uniqueness in Inverse Obstacle Problems", *Inverse Problems*, Vol. 9, pp. 285-299, 1993
24. Potthast, R., 2001, *Point-Sources and Multipoles in Inverse Scattering Theory*, Chapman & Hall, London.
25. Kim, K.Y., Kang, S., Kim, S., Kim, M., Kang, S. and Lee, J., 2002, "Dynamic Impedance Imaging of Binary-Mixture Fields with External and Internal Electrodes", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
26. Seppanen, A., Vauhkonem, M., Kaipio, J. and Somersalo, E., 2002, "Inference of Velocity Fields Based on Tomographic Measurements in Process Industry", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
27. Fraguela, A., Gamio, C. and Hinestroza, D., 2002, "An Inverse Algorithm for Capacitance Tomography of Two-Phase Flow Regimes", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
28. Rolnik, V. and Seleghim, P., 2002, "On-Site Calibration of a Phase Fraction Meter by an Inverse Technique", *4th Int. Conf. Inv. Problems in Engineering: Theory and Practice*, (ed: Orlande, H. R. B.) Rio de Janeiro, Brazil, May 26-31, 2002.
29. Cheney, M., Isaacson, D. and Newell, J., 1999, "Electrical Impedance Tomography", *SIAM Review*, Vol. 41, pp. 85-101.
30. Powell, M. J. D., 1977, "Restart Procedures for the Conjugate Gradient Method", *Mathematical Programming*, Vol. 12, pp. 241-254.

31. Dulikravich, G. S., Martin, T. J., Dennis, B. H. and Foster, N. F., 1999, "Multidisciplinary Hybrid Constrained GA Optimization", Chapter 12 in *EUROGEN'99 - Evolutionary Algorithms in Engineering and Computer Science: Recent Advances and Industrial Applications,* (eds: K. Miettinen, M. M. Makela, P. Neittaanmaki and J. Periaux), John Wiley & Sons, Ltd., Jyvaskyla, Finland, May 30 - June 3, 1999, pp. 231-260.
32. Dulikravich, G. S., Colaço, M. J., Martin, T. J. and Lee, S., 2003, "Optimization of Intensities, and Orientations of Magnets Controlling Melt Flow During Solidification", *J. of Materials and Manufacturing Processes*, to appear.
33. Dulikravich, G. S., Colaço, M. J., Martin, T. J. and Lee, S., 2003, "Magnetized Fiber Orientation and Concentration Control in Solidifying Composites, *Journal of Composite Materials"*, to appear.
34. Colaço, M. J., Dulikravich, G. S. and Martin, T. J., 2003, "Reducing Convection Effects in Solidification by Appling Magnetic Fields Having Optimized Intensity Distribution", ASME paper HT2003-47308, ASME Summer Heat Transfer Conference, Las Vegas, NV, July 21-23, 2003.
35. Kanevce, G., Kanevce, Lj. and Dulikravich, G. S., 2003, "An Inverse Method for Drying Processes at High Mass Transfer Biot Number", ASME paper HT2003-47146, ASME Summer Heat Transfer Conference, Las Vegas, NV, July 21-23, 2003.
36. Trujillo, D. and Busby, H., 1997, *Practical Inverse Analysis in Engineering*, CRC Press, Boca Raton, Florida.