Application of simulated annealing and hybrid methods in the solution of inverse heat and mass transfer problems

Antônio José da Silva Neto[1],
Jader Lugon Junior[2,5], Francisco José da Cunha Pires Soeiro[1],
Luiz Biondi Neto[1], Cesar Costapinto Santana[3],
Fran Sérgio Lobato[4] and Valder Steffen Junior[4]

Universidade do Estado do Rio de Janeiro[1],
Instituto Federal de Educação, Ciência e Tecnologia Fluminense[2],
Universidade Estadual de Campinas[3],
Universidade Federal de Uberlândia[4],
Centro de Tecnologia SENAI-RJ Ambiental[5]
Brazil

1. Introduction

The problem of parameter identification characterizes a typical inverse problem in engineering. It arises from the difficulty in building theoretical models that are able to represent satisfactorily physical phenomena under real operating conditions. Considering the possibility of using more complex models along with the information provided by experimental data, the parameters obtained through an inverse problem approach may then be used to simulate the behavior of the system for different operation conditions. Traditionally, this kind of problem has been treated by using either classical or deterministic optimization techniques (Baltes et al., 1994; Cazzador and Lubenova, 1995; Su and Silva Neto, 2001; Silva Neto and Özişik 1993ab, 1994; Yan et al., 2008; Yang et al., 2009). In the recent years however, the use of non-deterministic techniques or the coupling of these techniques with classical approaches thus forming a hybrid methodology became very popular due to the simplicity and robustness of evolutionary techniques (Wang et al., 2001; Silva Neto and Soeiro, 2002, 2003; Silva Neto and Silva Neto, 2003; Lobato and Steffen Jr., 2007; Lobato et al., 2008, 2009, 2010).

The solution of inverse problems has several relevant applications in engineering and medicine. A lot of attention has been devoted to the estimation of boundary and initial conditions in heat conduction problems (Alifanov, 1974, Beck *et al.*, 1985, Denisov and Solov'yera, 1993, Muniz et al., 1999) as well as thermal properties (Artyukhin, 1982, Carvalho and Silva Neto, 1999, Soeiro et al., 2000; Su and Silva Neto, 2001; Lobato et al., 2009) and heat source intensities (Borukhov and Kolesnikov, 1988, Silva Neto and Özisik, 1993ab, 1994, Orlande and Özisik, 1993, Moura Neto and Silva Neto, 2000, Wang *et al.*, 2000)

in such diffusive processes. On the other hand, despite its relevance in chemical engineering, there is not a sufficient number of published results on inverse mass transfer or heat convection problems. Denisov (2000) has considered the estimation of an isotherm of absorption and Lugon et al. (2009) have investigated the determination of adsorption isotherms with applications in the food and pharmaceutical industry, and Su et al., (2000) have considered the estimation of the spatial dependence of an externally imposed heat flux from temperature measurements taken in a thermally developing turbulent flow inside a circular pipe. Recently, Lobato et al. (2008) have considered the estimation of the parameters of Page's equation and heat loss coefficient by using experimental data from a realistic rotary dryer.

Another class of inverse problems in which the concurrence of specialists from different areas has yielded a large number of new methods and techniques for non-destructive testing in industry, and diagnosis and therapy in medicine, is the one involving radiative transfer in participating media. Most of the work in this area is related to radiative properties or source estimation (Ho and Özisik, 1989, McCormick, 1986, 1992, Silva Neto and Özisik, 1995, Kauati et al., 1999). Two strong motivations for the solution of such inverse problems in recent years have been the biomedical and oceanographic applications (McCormick, 1993, Sundman et al., 1998, Kauati et al., 1999, Carita Montero et al., 1999, 2000).

The increasing interest on inverse problems (IP) is due to the large number of practical applications in scientific and technological areas such as tomography (Kim and Charette, 2007), environmental sciences (Hanan, 2001) and parameter estimation (Souza et al., 2007; Lobato et al., 2008, 2009, 2010), to mention only a few.

In the radiative problems context, the inverse problem consists in the determination of radiative parameters through the use of experimental data for minimizing the residual between experimental and calculated values. The solution of inverse radiative transfer problems has been obtained by using different methodologies, namely deterministic, stochastic and hybrid methods. As examples of techniques developed for dealing with inverse radiative transfer problems, the following methods can be cited: Levenberg-Marquardt method (Silva Neto and Moura Neto, 2005); Simulated Annealing (Silva Neto and Soeiro, 2002; Souza et al., 2007); Genetic Algorithms (Silva Neto and Soeiro, 2002; Souza et al., 2007); Artificial Neural Networks (Soeiro et al., 2004); Simulated Annealing and Levenberg-Marquard (Silva Neto and Soeiro, 2006); Ant Colony Optimization (Souto et al., 2005); Particle Swarm Optimization (Becceneri et al, 2006); Generalized Extremal Optimization (Souza et al., 2007); Interior Points Method (Silva Neto and Silva Neto, 2003); Particle Collision Algorithm (Knupp et al., 2007); Artificial Neural Networks and Monte Carlo Method (Chalhoub et al., 2007b); Epidemic Genetic Algorithm and the Generalized Extremal Optimization Algorithm (Cuco et al., 2009); Generalized Extremal Optimization and Simulated Annealing Algorithm (Galski et al., 2009); Hybrid Approach with Artificial Neural Networks, Levenberg-Marquardt and Simulated Annealing Methods (Lugon, Silva Neto and Santana, 2009; Lugon and Silva Neto, 2010), Differential Evolution (Lobato et al., 2008; Lobato et al., 2009), Differential Evolution and Simulated Annealing Methods (Lobato et al., 2010).

In this chapter we first describe three problems of heat and mass transfer, followed by the formulation of the inverse problems, the description of the solution of the inverse problems with Simulated Annealing and its hybridization with other methods, and some test case results.

2. Formulation of the Direct Heat and Mass Transfer Problems

2.1 Radiative Transfer

Consider the problem of radiative transfer in an absorbing, emitting, isotropically scattering, plane-parallel, and gray medium of optical thickness τ_0, between two diffusely reflecting boundary surfaces as illustrated in Fig.1. The mathematical formulation of the direct radiation problem is given by (Özişik, 1973)

$$\mu\frac{\partial I(\tau,\mu)}{\partial\tau} + I(\tau,\mu) = \frac{\omega}{2}\int_{-1}^{1} I(\tau,\mu')d\mu', \qquad 0 < \tau < \tau_0, \quad -1 \le \mu \le 1 \tag{1}$$

$$I(0,\mu) = A_1 + 2\rho_1\int_{0}^{1} I(0,-\mu')\mu' d\mu', \quad \mu > 0 \tag{2}$$

$$I(\tau_0,\mu) = A_2 + 2\rho_2\int_{0}^{1} I(\tau_0,\mu')\mu' d\mu', \quad \mu < 0 \tag{3}$$

where $I(\tau,\mu)$ is the dimensionless radiation intensity, τ is the optical variable, μ is the direction cosine of the radiation beam with the positive τ axis, ω is the single scattering albedo, and ρ_1 and ρ_2 are the diffuse reflectivities. The illumination from the outside is supplied by external isotropic sources with intensities A_1 and A_2.

No internal source was considered in Eq. (1). In radiative heat transfer applications it means that the emission of radiation by the medium due to its temperature is negligible in comparison to the strength of the external isotropic radiation sources incident at the boundaries $\tau = 0$ and/or $\tau = \tau_0$.

In the direct problem defined by Eqs. (1-3) the radiative properties and the boundary conditions are known. Therefore, the values of the radiation intensity can be calculated for every point in the spatial and angular domains. In the inverse problem considered here the radiative properties of the medium are unknown, but we still need to solve problem (1-3) using estimates for the unknowns.
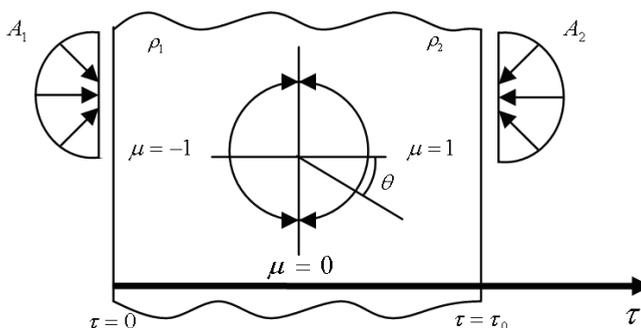
Fig. 1. The geometry and coordinates.

2.2 Drying (Simultaneous Heat and Mass Transfer)

In Fig. 2, adapted from Mwithiga and Olwal (2005), it is represented the drying experiment setup considered in this section. In it was introduced the possibility of using a scale to weight the samples, sensors to measure temperature in the sample, and also inside the drying chamber.
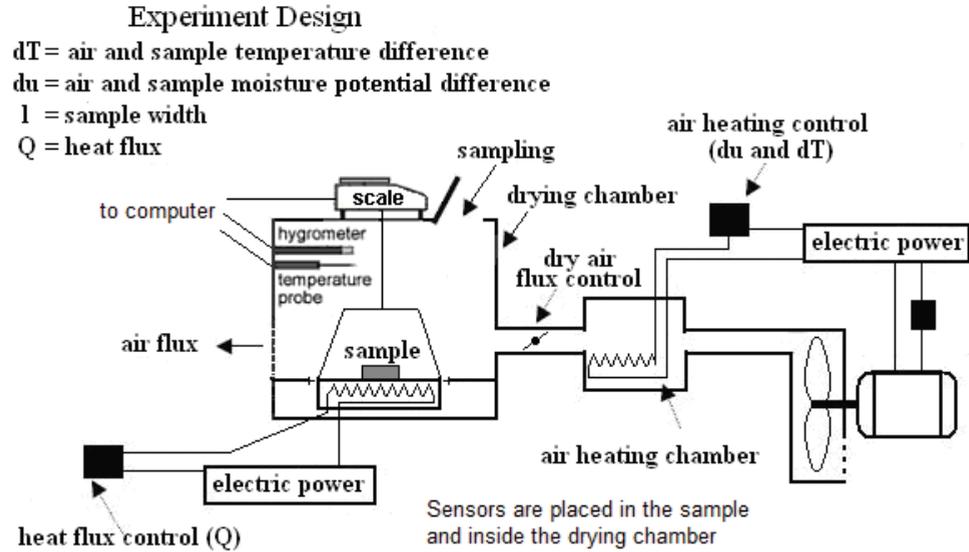
Fig. 2. Drying experiment setup (Adapted from Mwithiga and Olwal, 2005).

In accordance with the schematic representation shown in Fig. 3, consider the problem of simultaneous heat and mass transfer in a one-dimensional porous media in which heat is supplied to the left surface of the porous media, at the same time that dry air flows over the right boundary surface.
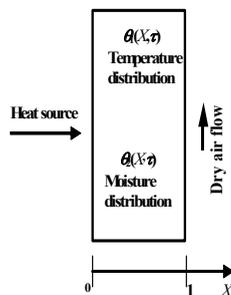
Fig. 3. Drying process schematic representation.

The mathematical formulation used in this work for the direct heat and mass transfer problem considered a constant properties model, and in dimensionless form it is given by (Luikov and Mikhailov, 1965; Mikhailov and Özisik, 1994),

$$\frac{\partial \theta_1(X,\tau)}{\partial \tau} = \alpha \frac{\partial^2 \theta_1}{\partial X^2} - \beta \frac{\partial^2 \theta_2}{\partial X^2} , \quad 0 < X < 1, \quad \tau > 0 \tag{4}$$

$$\frac{\partial \theta_2(X,\tau)}{\partial \tau} = Lu \frac{\partial^2 \theta_2}{\partial X^2} - Lu\,Pn \frac{\partial^2 \theta_1}{\partial X^2} , \quad 0 < X < 1, \quad \tau > 0 \tag{5}$$

subject to the following initial conditions, for $0 \le X \le 1$

$$\theta_1(X,0) = 0 \tag{6}$$

$$\theta_2(X,0) = 0 \tag{7}$$

and to the boundary conditions, for $\tau > 0$

$$\frac{\partial \theta_1(0,\tau)}{\partial X} = -Q \tag{8}$$

$$\frac{\partial \theta_2(0,\tau)}{\partial X} = -Pn\,Q \tag{9}$$

$$\frac{\partial \theta_1(1,\tau)}{\partial X} + Bi_q \theta_1(1,\tau) = Bi_q - (1-\varepsilon) KoLuBi_m \left[1 - \theta_2(1,\tau)\right] = 0 \tag{10}$$

$$\frac{\partial \theta_2(1,\tau)}{\partial X} + Bi_m^* \theta_2(1,\tau) = = Bi_m^* - Pn\,Bi_q \left[\theta_1(1,\tau) - 1\right] \tag{11}$$

where

$$\alpha = 1 + \varepsilon\,Ko\,Lu\,Pn \tag{12}$$

$$\beta = \varepsilon\,Ko\,Lu \tag{13}$$

$$Bi_m^* = Bi_m \left[1 - (1-\varepsilon) Pn\,Ko\,Lu\right] \tag{14}$$

and the dimensionless variables are defined as

$$\theta_1(X,\tau) = \frac{T(x,t) - T_0}{T_s - T_0} , \text{ temperature} \tag{15}$$

$$\theta_2(X,\tau) = \frac{u_0 - u(x,t)}{u_0 - u^*} , \text{ moisture potential} \tag{16}$$

$$X = \frac{x}{l} , \text{ spatial coordinate} \tag{17}$$

$$\tau = \frac{at}{l^2} , \text{ time} \tag{18}$$

$$Lu = \frac{a_m}{a} , \text{ Luikov number} \tag{19}$$

$$Pn = \delta \frac{T_s - T_0}{u_0 - u^*} , \text{ Possnov number} \tag{20}$$

$$Ko = \frac{r}{c}\frac{u_0 - u^*}{T_s - T_0} \text{ , Kossovitch number} \tag{21}$$

$$Bi_q = \frac{hl}{k} \text{ , heat Biot} \tag{22}$$

$$Bi_m = \frac{h_m l}{k_m}, \text{ mass Biot} \tag{23}$$

$$Q = \frac{ql}{k(T_s - T_0)} \text{ , heat flux} \tag{24}$$

When the geometry, the initial and boundary conditions, and the medium properties are known, the system of equations (4-11) can be solved, yielding the temperature and moisture distribution in the media. The finite difference method was used to solve the system (4-11).

Many previous works have studied the drying inverse problem using measurements of temperature and moisture-transfer potential at specific locations of the medium. But to measure the moisture-transfer potential in a certain position is not an easy task, so in this work it is used the average quantity

$$\bar{u}(t) = \frac{1}{l} \int_{x=0}^{x=l} u(x,t)dx \tag{25}$$

or

$$\bar{\theta}_2(\tau) = \int_{X=0}^{X=1} \theta_2(X,\tau)dX \tag{26}$$

Therefore, in order to obtain the average moisture measurements, $\bar{u}(t)$, one have just to weight the sample at each time (Lugon and Silva Neto, 2010).

2.3 Gas-liquid Adsorption

The mechanism of proteins adsorption at gas-liquid interfaces has been the subject of intensive theoretical and experimental research, because of the potential use of bubble and foam fractionation columns as an economically viable means for surface active compounds recovery from diluted solutions, (Özturk et al., 1987; Deckwer and Schumpe, 1993; Graham and Phillips, 1979; Santana and Carbonell, 1993ab; Santana, 1994; Krishna and van Baten, 2003; Haut and Cartage, 2005; Mouza et al., 2005; Lugon, 2005).

The direct problem related to the gas-liquid interface adsorption of bio-molecules in bubble columns consists essentially in the calculation of the depletion, that is, the reduction of solute concentration with time, when the physico-chemical properties and process parameters are known.

The solute depletion is modeled by

$$\frac{dC_b}{dt} = -\frac{6v_g}{(1-\varepsilon_g)Hd_b}\Gamma \tag{27}$$

where C_b is the liquid solute concentration (bulk), d_b is the bubble diameter, H is the bubble column height, v_g is the superficial velocity (gas volumetric flow rate divided by the area of the transversal section of the column A), and Γ is the surface excess concentration of the adsorbed solute.

The symbol ε_g represents the gas volumetric fraction, which can be calculated from the dimensionless correlation of Kumar (Özturk et al., 1987),

$$\varepsilon_g = 0.728U - 0.485U^2 + 0.095U^3 \tag{28}$$

where

$$U = v_g \left[\frac{\rho_l^2}{\gamma(\rho_l - \rho_g)g} \right]^{\frac{1}{4}} \tag{29}$$

ρ_l is the liquid density, γ is the surface tension, g is the gravity acceleration, and ρ_g is the gas density.

The quantities Γ and C are related through adsorption isotherms such as:

(i) Linear isotherm

$$\Gamma = B + KC \tag{30}$$

(ii) Langmuir isotherm

$$\Gamma_1 = \frac{1}{\hat{a}} \left[\frac{K_1(T)C}{1 + K_1(T)C} \right] \tag{31}$$

(iii) Two-layers isotherm

$$\Gamma_t = \Gamma_1 + \Gamma_2 = \frac{K_1(T)\exp(-\lambda\Gamma_1)C\left[1 + K_2(T)\hat{a}C\right]}{\hat{a}\left[1 + K_1\exp(-\lambda\Gamma_1)C\right]} \tag{32}$$

where Γ_1 and Γ_2 are the excess superficial concentration in the first and second adsorption layers respectively (see Fig. 4).
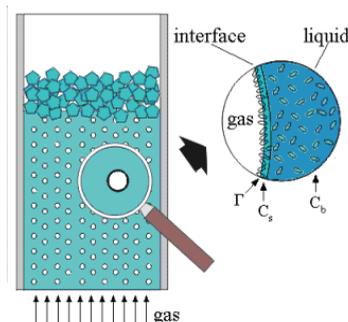
Fig. 4. Schematic representation of the gas-liquid adsorption process in a bubble and foam column.

Considering that the superficial velocity, bubble diameter and column cross section are constant along the column,

$$\frac{\partial \Gamma(z,t)}{\partial z} = \frac{(k_l a) d_b \left[C_b(t) - C_s(z,t) \right]}{6 v_g}$$ (33)

where z represents the spatial coordinate along the column, C_s is the solute concentration next to the bubbles and $(k_l a)$ is the volumetric mass transfer coefficient.

There are several correlations available for the determination of $(k_l a)$ but following the recommendation of Deckwer and Schumpe (1993) we have adopted the correlation of Öztürk et al. (1987) in the solution of the direct problem:

$$Sh = 0,62 \, Sc^{0,5} \, Bo^{0,33} \, Ga^{0,29} \left(\frac{v_g}{\sqrt{g d_b}} \right)^{0,68} \left(\frac{\rho_g}{\rho_l} \right)^{0,04}$$ (34)

where

$$Sc = \left(\frac{v_l}{D_i} \right), \text{ Schmidt number}$$ (35)

$$Sh = \frac{(k_l a) d_b^2}{D_i}, \text{ Sherwood number}$$ (36)

$$Bo = \frac{v_l}{D_i}, \text{ Bond number}$$ (37)

$$Ga = \frac{g d_b^3}{v_l^2}, \text{ Galilei Number}$$ (38)

D_i is the tensoactive diffusion coefficient and v_l is the liquid dynamic viscosity.

Combining Eqs. (27) and (33) and using an initial condition, such as $C_b = C_{b0}$ when $t = 0$, and a boundary condition, like $\Gamma = 0$ at $z = 0$, the solute concentration can be calculated as a function of time, $C_b(t)$. Santana and Carbonell (1993ab) developed an analytical solution for the direct problem in the case of a linear adsorption isotherm and the results presented a good agreement with experimental data for BSA (Bovine Serum Albumin).

In order to solve Eq. (27) a second order Runge Kutta method was used, known as the mid point method. Given the physico-chemical and process parameters, as well as the boundary and initial conditions, the solute concentration can be calculated for any time t (Lugon et al., 2009).

3. Formulation of Inverse Heat and Mass Transfer Problems

The inverse problem is implicitly formulated as a finite dimensional optimization problem (Silva Neto and Soeiro, 2003; Silva Neto and Moura Neto, 2005), where one seeks to minimize the cost functional of squared residues between the calculated and experimental values for the observable variable,

$$S(\mathbf{P}) = \left[\mathbf{G}_{calc}(\mathbf{P}) - \mathbf{G}_{meas}(\mathbf{P}) \right]^T \mathbf{W} \left[\mathbf{G}_{calc}(\mathbf{P}) - \mathbf{G}_{meas}(\mathbf{P}) \right] = \mathbf{F}^T \mathbf{F}$$ (39a)

where $\mathbf{G}_{meas}$ is the vector of measurements, $\mathbf{G}_{calc}$ is the vector of calculated values, $\mathbf{P}$ is the vector of unknowns, $\mathbf{W}$ is the diagonal matrix whose elements are the inverse of the measurement variances, and the vector of residues $\mathbf{F}$ is given by

$$\mathbf{F} = \mathbf{G}_{calc}(\mathbf{P}) - \mathbf{G}_{calc}(\mathbf{P}) \tag{39b}$$

The inverse problem solution is the vector $\vec{P}^*$ which minimizes the norm given by Eq. (39a), that is

$$S(\mathbf{P}^*) = \min_{\mathbf{P}} S(\mathbf{P}) \tag{40}$$

Depending on the direct problem, different measurements are to be taken, that is:

a) Radiative problem
Using calculated values given by Eq. (1) and experimental radiation intensities at the boundaries $\tau = 0$ and $\tau = \tau_0$, as well as at points that belong to the set Ω (points inside the domain τ - internal detectors) we try to estimated the vector of unknowns $\mathbf{P}$ considered. Two different vectors of unknowns $\vec{P}$ are possibly considered for the minimization of the difference between the experimental and calculated values: (i) τ_0, ω, ρ_1 and ρ_2; (ii) τ_0, ω, A_1 and A_2.

b) Drying problem
Using temperature measurements, T, taken by sensors located inside the medium, and the average of the moisture-transfer potential, $\bar{u}$, during the experiment, we try to estimate the vector of unknowns $\mathbf{P}$, for which a combination of variables was used: Lu (Luikov number), δ (thermogradient coefficient), r/c (relation between latent heat of evaporation and specific heat of the medium), h/k (relation between heat transfer coefficient and thermal conductivity), and h_m/k_m (relation between mass transfer coefficient and mass conductivity).

c) Gas-liquid adsorption problem
Different vectors of unknowns $\mathbf{P}$ are possibly considered, which are associated with different adsorption isotherms: (i) K and B (Linear isotherm); (ii) $K_1(T)$ and $\hat{a}$ (Langmuir isotherm); (iii) $K_1(T)$, $K_2(T)$, λ and $\hat{a}$ (two-layers isotherm). Here the BSA (Bovine Serum Albumin) adsorption was modeled using a two-layer isotherm.

4. Solution of the Inverse Problems with Simulated Annealing and Hybrid Methods

4.1 Design of Experiments
The sensitivity analysis plays a major role in several aspects related to the formulation and solution of an inverse problem (Dowding et al., 1999; Beck, 1988). Such analysis may be performed with the study of the sensitivity coefficients. Here we use the modified, or scaled, sensitivity coefficients

$$SC_{P_jV(t)} = P_j \frac{\partial V(t)}{\partial P_j}, \quad j = 1,2,...,N_p \tag{41}$$

where V is the observable state variable (which can be measured), P_j is a particular unknown of the problem, and N_p is the total number of unknowns considered.

As a general guideline, the sensitivity of the state variable to the parameter we want to determine must be high enough to allow an estimate within reasonable confidence bounds. Moreover, when two or more parameters are simultaneously estimated, their effects on the state variable must be independent (uncorrelated). Therefore, when represented graphically, the sensitivity coefficients should not have the same shape. If they do it means that two or more different parameters affect the observable variable in the same way, being difficult to distinguish their influences separately, which yields to poor estimations.

Another important tool used in the design of experiments is the study of the matrix

$$\mathbf{SC} = \begin{bmatrix} SC_{P_1V_1} & SC_{P_2V_2} & ... & SC_{P_{N_P}V_1} \\ SC_{P_1V_2} & SC_{P_2V_2} & ... & SC_{P_{N_P}V_2} \\ ... & ... & ... & ... \\ SC_{P_1V_m} & SC_{P_2V_m} & ... & SC_{P_{N_P}V_m} \end{bmatrix} \tag{42}$$

where V_i is a particular measurement of temperature or moisture potential and m is the total number of measurements.

Maximizing the determinant of the matrix $\mathbf{SC}^T\mathbf{SC}$ results in higher sensitivity and uncorrelation (Beck, 1988).

4.2 Simulated Annealing Method (SA)

Based on statistical mechanics reasoning, applied to a solidification problem, Metropolis et al. (1953) introduced a simple algorithm that can be used to accomplish an efficient simulation of a system of atoms in equilibrium at a given temperature. In each step of the algorithm a small random displacement of an atom is performed and the variation of the energy ΔE is calculated. If $\Delta E < 0$ the displacement is accepted, and the configuration with the displaced atom is used as the starting point for the next step. In the case of $\Delta E > 0$, the new configuration can be accepted according to Boltzmann probability

$$P(\Delta E) = \exp(-\Delta E / k_B T) \tag{43}$$

A uniformly distributed random number p in the interval [0,1] is calculated and compared with $P(\Delta E)$. Metropolis criterion establishes that the new configuration is accepted if $p < P(\Delta E)$, otherwise it is rejected and the previous configuration is used again as a starting point.

Using the objective function $S(\mathbf{P})$, given by Eq. (39a), in place of energy and defining configurations by a set of variables $\{P_i\}, i = 1,2,.N_p$, where N_p represents the number of unknowns we want to estimate, the Metropolis procedure generates a collection of

configurations of a given optimization problem at some temperature T (Kirkpatric et al., 1983). This temperature is simply a control parameter. The simulated annealing process consists of first "melting" the system being optimized at a high "temperature", then lowering the "temperature" until the system "freezes" and no further change occurs.

The main control parameters of the algorithm implemented ("cooling procedure") are the initial "temperature", T_0, the cooling rate, r_t, number of steps performed through all elements of vector $\mathbf{P}$, N_s, number of times the procedure is repeated before the "temperature" is reduced, N_t, and the number of points of minimum (one for each temperature) that are compared and used as stopping criterion if they all agree within a tolerance ε, N_ε.

4.3 Levenberg-Marquardt Method (LM)

The Levenberg-Marquardt is a deterministic local optimizer method based on the gradient (Marquardt, 1963). In order to minimize the functional $S(\mathbf{P})$ we first write

$$\frac{dS}{d\mathbf{P}} = \frac{d}{d\mathbf{P}}\left(\mathbf{F}^T\mathbf{F}\right) = 0 \rightarrow \mathbf{J}^T\mathbf{J} = 0 \tag{44}$$

where J is the Jacobian matrix, with the elements $J_{ps} = \partial C_{bp}/\partial P_s$ being $p = 1,\ 2,\ ...,\ M$, and $s = 1,\ 2,\ ...,\ N_p$, where M is the total number of measurements and N_p is the number of unknowns. It is observed that the elements of the Jacobian matrix are related to the scaled sensitivity coefficients presented before.

Using a Taylor's expansion and keeping only the terms up to the first order,

$$\mathbf{F}(\mathbf{P} + \Delta\mathbf{P}) \cong \mathbf{F}(\mathbf{P}) + \mathbf{J}\Delta\mathbf{P} \tag{45}$$

Introducing the above expansion in Eq. (44) results

$$J^T J \Delta\vec{P} = -J^T \vec{F}\left(\vec{P}\right) \tag{46}$$

In the Levenberg-Marquardt method a damping factor γ^n is added to the diagonal of matrix $\mathbf{J}^T\mathbf{J}$ in order to help to achieve convergence.

Equation (46) is written in a more convenient form to be used in the iterative procedure,

$$\Delta\mathbf{P}^n = -\left[(\mathbf{J}^n)^T \mathbf{J}^n + \gamma^n\mathbf{I}\right]^{-1}(\mathbf{J}^n)^T \mathbf{F}(\mathbf{P}^n) \tag{47}$$

where $\mathbf{I}$ is the identity matrix and n is the iteration index.

The iterative procedure starts with an estimate for the unknown parameters, $\mathbf{P}^0$, being new estimates obtained with $\mathbf{P}^{n+1} = \mathbf{P}^n + \Delta\mathbf{P}^n$, while the corrections $\Delta\mathbf{P}^n$ are calculated with Eq. (46). This iterative procedure is continued until a convergence criterion such as

$$\left|\Delta P_k^n / P_k^n\right| < \varepsilon, \qquad n = 1,\ 2,\ \cdots,\ N_p \tag{48}$$

is satisfied, where ε is a small number, e.g. 10^{-5}.

The elements of the Jacobian matrix, as well as the right side term of Eq. (47), are calculated at each iteration, using the solution of the problem with the estimates for the unknowns obtained in the previous iteration.

4.4 Artificial Neural Network (ANN)

The multi-layer perceptron (MLP) is a collection of connected processing elements called nodes or neurons, arranged in layers (Haykin, 1999). Signals pass into the input layer nodes, progress forward through the network hidden layers and finally emerge from the output layer (see Fig. 5). Each node i is connected to each node j in its preceding layer through a connection of weight, w_{ij}, and similarly to nodes in the following layer.
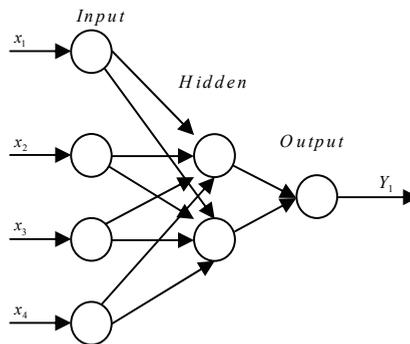
Fig. 5. Multi-layer perceptron network.

A weighted sum is performed at i of all the signals x_j from the preceding layer, yielding the excitation of the node; this is then passed through a nonlinear activation function, f, to emerge as the output of the node x_i to the next layer, as shown by the equation

$$y_i = f\left(\sum_j w_{ij} x_j \right) \qquad (49)$$

Various choices for the function f are possible. Here the hyperbolic tangent function $f(x) = \tanh(x)$ is used.

The first stage of using an ANN to model an input-output system is to establish the appropriate values for the connection weights w_{ij}. This is the "training" or learning phase.

Training is accomplished using a set of network inputs for which the desired outputs are known. These are the so called patterns, which are used in the training stage of the ANN. At each training step, a set of inputs are passed forward through the network yielding trial outputs which are then compared to the desired outputs. If the comparison error is considered small enough, the weights are not adjusted. Otherwise the error is passed backwards through the net and a training algorithm uses the error to adjust the connection weights. This is the back-propagation algorithm.

Once the comparison error is reduced to an acceptable level over the whole training set, the training phase ends and the network is established. The parameters of a model (output) may be determined using the real experimental data, which are inputs of the established neural network. This is the generalization stage in the use of the ANN. More details can be found in (Soeiro et al., 2004).

4.5 Differential Evolution

The Differential Evolution (DE) is a structural algorithm proposed by Storn and Price (1995) for optimization problems. This approach is an improved version of the Goldberg's Genetic Algorithm (GA) (Goldberg, 1989) for faster optimization and presented the following advantages: simple structure, easiness of use, speed and robustness (Storn and Price, 1995).

Basically, DE generates trial parameter vectors by adding the weighted difference between two population vectors to a third vector. The key parameters of control in DE are the following: N, the population size, CR, the crossover constant and, D, the weight applied to random differential (scaling factor). Storn and Price (1995) have given some simple rules for choosing key parameters of DE for any given application. Normally, N should be about 5 to 10 times the dimension (number of parameters in a vector) of the problem. As for D, it lies in the range 0.4 to 1.0. Initially, $D = 0.5$ can be tried, and then D and/or N is increased if the population converges prematurely.

DE has been successfully applied to various fields such as digital filter design (Storn, 1995), batch fermentation process (Chiou and Wang, 1999), estimation of heat transfer parameters in a bed reactor (Babu and Sastry, 1999), synthesis and optimization of heat integrated distillation system (Babu and Singh, 2000), optimization of an alkylation reaction (Babu and Gaurav, 2000), parameter estimation in fed-batch fermentation process (Wang et al., 2001), optimization of thermal cracker operation (Babu and Angira, 2001), engineering system design (Lobato and Steffen, 2007), economic dispatch optimization (Coelho and Mariani, 2007), identification of experimental data (Maciejewski et al., 2007), apparent thermal diffusivity estimation during the drying of fruits (Mariani et al., 2008), estimation of the parameters of Page's equation and heat loss coefficient by using experimental data from a realistic rotary dryer (Lobato et al., 2008), solution of inverse radiative transfer problems (Lobato et al., 2009, 2010), and other applications (Storn et al., 2005).

4.6 Combination of ANN, LM and SA Optimizers

Due to the complexity of the design space, if convergence is achieved with a gradient based method it may in fact lead to a local minimum. Therefore, global optimization methods are required in order to reach better approximations for the global minimum. The main disadvantage of these methods is that the number of function evaluations is high, becoming sometimes prohibitive from the computational point of view (Soeiro et al., 2004).

In this chapter, different combinations of methods are used for the solution of inverse heat and mass transfer problems, involving in all cases Simulated Annealing as the global optimizer:

   a) when solving radiative inverse problems, it was used a combination of the LM and SA;
   b) when solving adsorption and drying inverse problems, it was used a combination of ANN, LM and SA.

Therefore, in all cases it was run the LM, reaching within a few iterations a point of minimum. After that we run the SA. If the same solution is reached, it is likely that a global

minimum was reached, and the iterative procedure is interrupted. If a different solution is obtained it means that the previous one was a local minimum, otherwise we could run again the LM and SA until the global minimum is reached.

When using the ANN method, after the training stage one is able to quickly obtain an inverse problem solution. This solution may be used as an initial guess for the LM. Trying to keep the best features of each method, we have combined the ANN, LM and SA methods.

5. Test Case Results

5.1 Radiative Transfer
5.1.1 Estimation of { τ_0 , ω , ρ_1 , ρ_2 } using LM-SA combination

The combined LM-SA approach was applied to several test problems. Since there were no real experimental data available, they were simulated by solving the direct problem and considering the output as experimental data. These results may be corrupted by random multipliers representing a white noise in the measuring equipment. In this effort, since we are developing the approach and trying to compare the performance of the optimization techniques involved, the output was considered as experimental result without any change.

The direct problem is solved with a known vector { τ_0 , ω , ρ_1 , ρ_2 }, which will be considered as the exact solution for the inverse problem. The correspondent output is recorded as experimental data. Now we begin the inverse problem with an initial estimate for the above quantities, obviously away from the exact solution. The described approach is, then, used to find the exact solution.

In a first example the exact solution vector was assumed as {1.0,0.5,0.95,0.5} and the initial estimate as {0.1,0.1,0.1,0.1}. Using both methods the exact solution was obtained. The difference was the computational effort required as shown in Table 1.

Method	Iterations/Cycles	Number of function evaluations	Final value of the objective function
LM	8 iterations	40	2.265E-13
SA	90 cycles	36000	2.828E-13

Table 1. Comparison LM – SA for the first example.

In a second example the exact solution was assumed as {1.0,0.5,0.1,0.95} and the starting point was {5.0,0.95,0.95,0.1}. In this case the LM did not converge to the right answer. The results are presented in Table 2.

Iteration	τ_o	ω	ρ_1	ρ_2	Obj. Function
0	5.0	0.95	0.95	0.1	10.0369
1	5.7856	9.63E-1	6.60E-2	1.00E-4	1.7664
:	:	:	:	:	:
20	9.2521	1.0064	1.00E-4	1.00E-4	2.4646
Exact Solution	1.0	0.5	0.1	0.95	0.0

Table 2. Results for Z_{exact} = {1.0,0.5,0.1,0.95} and Z_o = {5.0,0.95,0.95,0.1} using LM.

The difficulty encountered by LM in converging to the right solution was due to a large plateau that exists in the design space for values of τ_o between 6.0 and 10.0. In this interval the objective function has a very small variation. The SA solved the problem with the same performance as in the first example. The combination of both methods was then applied.

SA was let running for only one cycle (400 function evaluations). At this point, the current optimum was {0.94,0.43,0.61,0.87}, far from the plateau mentioned above. With this initial estimate, LM converged to the right solution very quickly in four iterations, as shown in Table 3.

Iteration	τ_0	ω	ρ_1	ρ_2	Obj. Function [Eq. (39a)]
0	0.94	0.43	0.61	0.87	1.365E-2
1	1.002	0.483	0.284	0.945	5.535E-5
:	:	:	:	:	:
4	0.999	0.500	0.100	0.9500	9.23E-13
Exact Sol.	1.0	0.5	0.1	0.95	0.0

Table 3. Results for Z_{exact} = {1.0,0.5,0.1,0.95} and Z_o = {5.0,0.95,0.95,0.1} using LM after one cycle of SA.

5.1.2 Estimation of { ω , τ_0 , A_1 , A_2 } using SA and DE

In order to evaluate the performance of the methods of Simulated Annealing and Differential Evolution for the simultaneous estimation of both the single scattering albedo, ω, and the optical thickness, τ_0, of the medium, and also the intensities A_1 and A_2 of the external sources at $\tau = 0$ and $\tau = \tau_0$, respectively, of a given one-dimensional plane-parallel participating medium, the four test cases listed in Table 4 have been performed (Lobato et al., 2010).

Parameter	Meaning	Problem			
		1	2	3	4
ω	Single scattering albedo	0.1	0.1	0.9	0.9
τ_0	Optical thickness of the layer	0.5	5.0	0.5	5.0
A_1	Intensity of external source at $\tau = 0$	1.0	1.0	1.0	1.0
A_2	Intensity of external source at $\tau = \tau_0$	0.0	0.0	0.0	0.0
n	Number of experimental data points	20	20	20	20

Table 4. Parameters used to define the illustrative examples.

It should be emphasized that 20 points were used for the approximation of the variable μ, and 10 collocation points were taken into account to solve the direct problem. All test cases were solved by using a microcomputer PENTIUM IV with 3.2 GHz and 2 GB of RAM. Both the algorithms were executed 10 times for obtaining the values presented in the Tables (6-9).

The parameters used in the two algorithms are presented in Table 5.

Parameter		SA	DE	
Iteration number	N_{gen}	100	100	
Population size	N	-	10	
Crossover probability	CR	-	0.8	
Perturbation rate	D	-	0.8	
Strategy	-	-	DE/rand/1/bin	
Temperature number for each temperature	N_{temp}	50	-	
Temp. initial/final	T_i/T_f	0.5/0.01	-	
Initial Estimate	Case 1	[0.25 0.25 0.5 0.5]	Randomly generated	$0\le w,\tau_0\le 1$; $1\le A_1\le 1.5$; $0\le A_2\le 1$
	Case 2	[0.25 0.45 0.5 0.5]		$0\le w,A_2\le 1$; $3\le \tau_0\le 5$; $1\le A_1\le 1.5$
	Case 3	[0.75 0.25 0.5 0.5]		$0\le w\le 1.0$; $0\le \tau_0,A_2\le 1$; $1\le A_1\le 1.5$
	Case 4	[0.75 0.45 0.5 0.5]		$0\le w\le 1.0$; $3\le \tau_0\le 5$; $1\le A_1\le 1.5$; $0\le A_2\le 1$

Table 5. Parameters used to define the illustrative examples.

			ω	τ_0	A_1	A_2	Objective Function [Eq. (39a)]
Exact	Error in experimental data		0.1	0.5	1.0	0.0	-
DE*	0.0	Worst	0.1003	0.5002	1.0000	0.0001	1.5578x10⁻⁶
		Average	0.0998	0.4999	0.9999	0.0000	5.7702x10⁻⁷
		Best	**0.1000**	**0.4999**	**0.9999**	**0.0000**	**4.4564x10⁻⁹**
	0.5%	Worst	0.1015	0.4991	0.9980	0.0012	8.4403x10⁻⁴
		Average	0.1007	0.4985	0.9976	0.0010	8.4244x10⁻⁴
		Best	**0.1006**	**0.4983**	**0.9974**	**0.0011**	**8.4144x10⁻⁴**
	5.0%	Worst	0.0876	0.5018	0.9992	0.0058	0.0842
		Average	0.0876	0.5018	0.9992	0.0058	0.0842
		Best	**0.0870**	**0.5017**	**0.9990**	**0.0057**	**0.0842**
SA**	0.0	Worst	0.0994	0.4999	1.0001	0.0000	5.3920x10⁻⁷
		Average	0.0996	0.4998	0.9999	0.0000	3.4741x10⁻⁷
		Best	**0.0999**	**0.4999**	**0.9999**	**0.0000**	**2.1496x10⁻⁷**
	0.5%	Worst	0.0944	0.4917	0.9922	0.0001	9.6060x10⁻⁴
		Average	0.0962	0.4959	0.9970	0.0000	8.5299x10⁻⁴
		Best	**0.0984**	**0.4976**	**0.9974**	**0.0000**	**8.4058x10⁻⁴**
	5.0%	Worst	0.0885	0.5012	0.9991	0.0059	0.0849
		Average	0.0880	0.5010	0.9990	0.0059	0.0844
		Best	**0.0879**	**0.5010**	**0.9989**	**0.0056**	**0.0842**

* *NF*=1010, *cputime*=4.1815 min and ** *NF*=7015, *cputime*=30.2145 min.

Table 6. Results obtained for case 1.

			ω	τ_o	A_1	A_2	Objective Function [Eq. (39a)]
Exact	*Error in experimental data*		0.1	5.0	1.0	0.0	-
DE*	0.0	*Worst*	0.1024	4.9982	0.9988	0.0013	6.3559x10^{-6}
		Average	0.1004	4.9976	0.9992	0.0000	2.6107x10^{-6}
		Best	**0.0998**	**5.0036**	**1.0008**	**0.0000**	**1.1856x10^{-7}**
	0.5%	*Worst*	0.0978	4.9438	0.9844	0.0007	8.0356x10^{-4}
		Average	0.0984	4.9470	0.9847	0.0008	8.0333x10^{-4}
		Best	**0.0983**	**4.9494**	**0.9850**	**0.0010**	**8.0310x10^{-4}**
	5.0%	*Worst*	0.0453	4.9678	0.9683	0.0000	0.0878
		Average	0.0454	4.9675	0.9682	0.0000	0.0878
		Best	**0.0455**	**4.9674**	**0.9680**	**0.0000**	**0.0878**
SA**	0.0	*Worst*	0.0997	5.0097	1.0026	0.0004	8.6468x10^{-7}
		Average	0.0998	4.9981	0.9995	0.0003	7.7231x10^{-7}
		Best	**0.0994**	**4.9956**	**0.9988**	**0.0005**	**7.1664x10^{-7}**
	0.5%	*Worst*	0.0929	4.9487	0.9789	0.0009	9.4786x10^{-3}
		Average	0.0971	4.9256	0.9848	0.0005	8.0999x10^{-3}
		Best	**0.0987**	**4.9390**	**0.9841**	**0.0004**	**8.0645x10^{-4}**
	5.0%	*Worst*	0.0483	4.9578	0.9689	0.0001	0.0892
		Average	0.0484	4.9575	0.9685	0.0001	0.0890
		Best	**0.0485**	**4.9554**	**0.9680**	**0.0001**	**0.0888**

* *NF=1010, cputime=21.4578 min and ** NF=8478, cputime=62.1478 min.*
Table 7. Results obtained for case 2.

			ω	τ_o	A_1	A_2	Objective Function [Eq. (39a)]
Exact	*Error in experimental data*		0.9	0.5	1.0	0.0	-
DE*	0.0	*Worst*	0.8998	0.5001	1.0000	0.0000	4.0332x10^{-9}
		Average	0.8999	0.5000	1.0000	0.0000	2.1772x10^{-9}
		Best	**0.9000**	**0.5000**	**1.0000**	**0.0000**	**2.0152x10^{-9}**
	0.5%	*Worst*	0.9028	0.4978	0.9979	0.0001	8.9999x10^{-3}
		Average	0.9020	0.4980	0.9984	0.0000	8.8788x10^{-4}
		Best	**0.9018**	**0.4988**	**0.9994**	**0.0000**	**8.6296x10^{-4}**
	5.0%	*Worst*	0.9020	0.4700	0.9864	0.0000	0.0776
		Average	0.9022	0.4790	0.9870	0.0000	0.0746
		Best	**0.9032**	**0.4807**	**0.9871**	**0.0000**	**0.0736**
SA**	0.0	*Worst*	0.8998	0.5000	1.0000	0.0001	8.3002x10^{-8}
		Average	0.8998	0.5000	1.0000	0.0000	4.7782x10^{-8}
		Best	**0.8999**	**0.5000**	**1.0000**	**0.0000**	**2.0152x10^{-8}**
	0.5%	*Worst*	0.9039	0.4981	0.9981	0.0000	8.7988x10^{-4}
		Average	0.9025	0.4980	0.9986	0.0000	8.7744x10^{-4}
		Best	**0.9021**	**0.4990**	**0.9994**	**0.0000**	**8.7014x10^{-4}**
	5.0%	*Worst*	0.9049	0.4790	0.9859	0.0000	0.0760
		Average	0.9024	0.4792	0.9860	0.0000	0.0756
		Best	**0.9030**	**0.4800**	**0.9864**	**0.0000**	**0.0738**

* *NF=1010, cputime=3.8788 min and ** NF=8758, cputime=27.9884 min.*
Table 8. Results obtained for case 3.

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

➢ HTML (Free /Available to everyone)

➢ PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)

➢ Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below