26

Application of Artificial Neural Networks and Hybrid Methods in the Solution of Inverse Problems

Jader Lugon Junior[1,5], Antônio J. da Silva Neto[2], Luiz Biondi Neto[2], Francisco José da Cunha Pires Soeiro[2], Cesar Costapinto Santana[3] and Haroldo F. Campos Velho[4]

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

1. Introduction

In this chapter Artificial Neural Networks are presented and used to solve different parameter estimation inverse problems, that is, Gas-liquid Adsorption Mass Transfer, Radiative Transfer Problems, and Simultaneous Heat and Mass Transfer. Besides, results obtained using hybrid methods are also presented, combining the Artificial Neural Network (ANN) method to other inverse problem solutions techniques, such as Simulated Annealing (SA) and Levenberg-Marquardt (LM).

The first problem studied is the radiative transfer phenomenon, modeled with an integro-differential equation known as Boltzmann equation. This equation describes mathematically the interaction of the radiation with the participating medium, i.e., a medium that may absorb, scatter and emit radiation. The inverse radiative transfer problem considered the simultaneous estimation of the absorption and scattering coefficients of a two-layer medium, using measured exit radiation intensities. In this sense, a study is presented regarding the estimation of radiative properties using ANN and hybrid methods combining ANN and LM.

Then, the inverse problem of simultaneous heat and mass transfer modeled by Luikov equations is studied using a hybrid combination of the ANN, LM and SA. Direct and inverse problems are presented, formulated and solved. An ANN was used to generate the initial guess for the LM, another ANN to approximate the gradient needed by LM, and finally the global minimum was searched using the SA. The experimental data used was generated using the solution for the direct problem with the addition of artificial noise.

The gas-liquid interface adsorption isotherm identification is also investigated using the same hybrid approach, that is, the combination of an ANN, LM and SA methods. The bubble and foam fractionation columns system works basically through the injection of a gas at the base of a column containing the solution. The gas bubbles formed in the

distributor rise and along this path adsorb the solute, which is extracted in the foam region, formed above the bubble column. The inverse problem approach described allows the determination of the adsorption isotherms needed to solve the mathematical and numerical models developed.

2. Formulation of the direct heat and mass transfer problems

2.1 Radiative transfer

Consider the problem of radiative transfer in a composite medium with two plane-parallel, isotropically scattering, gray layers, with diffusely reflecting boundary surfaces and interface, as shown in Fig. 1. The medium is subjected to external irradiation at both sides with intensity $f_1(\mu)$ at $x=0$ and $f_2(\mu)$ at $x=L_1+L_2$, where μ is the cosine of the polar angle, and L_1 and L_2 represent the thickness of layers 1 and 2, respectively.

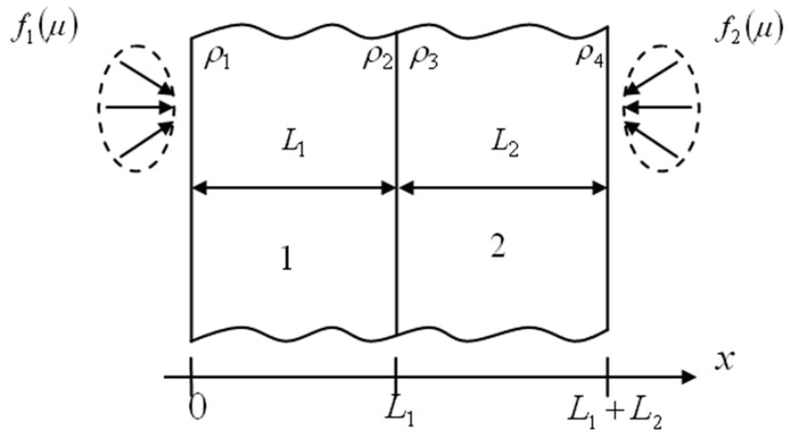


Fig. 1. Two-layer semitransparent medium.

The mathematical formulation of the direct steady-state radiative transfer problem with azymuthal symmetry is given by (Özisik, 1973)

Region 1

$$\mu\frac{\partial I_1(x,\mu)}{\partial x}+\beta_1 I_1(x,\mu)=\frac{\sigma_{s_1}}{2}\int_{-1}^{1} I_1(x,\mu')d\mu' \text{ in } 0<x<L_1,\ -1\le\mu\le 1 \tag{1a}$$

$$I_1(0,\mu)=f_1(\mu)+2\rho_1\int_0^1 I_1(0,-\mu')\mu' d\mu',\ \mu>0 \tag{1b}$$

$$I_1(L_1,\mu)=(1-\rho_3)I_2(L_1,\mu)+2\rho_2\int_0^1 I_1(L_1,\mu')\mu' d\mu',\ \mu<0 \tag{1c}$$

Region 2

$$\mu\frac{\partial I_2(x,\mu)}{\partial x}+\beta_2 I_2(x,\mu)=\frac{\sigma_{s_2}}{2}\int_{-1}^{1} I_2(x,\mu')d\mu' \text{ in } L_1<x<L_1+L_2,\ -1\le\mu\le 1 \tag{2a}$$

$$I_2(L_1,\mu) = (1-\rho_2) I_1(L_1,\mu) + 2\rho_3 \int_0^1 I_2(L_1,-\mu')\mu' d\mu', \ \mu > 0 \quad (2b)$$

$$I_2(L_1+L_2,\mu) = f_2(\mu) + 2\rho_4 \int_0^1 I_2(L_1+L_2,\mu')\mu' d\mu', \ \mu < 0 \quad (2c)$$

where $I_i(x,\mu)$ represents the radiation intensity in layer i, with $i = 1 or 2$, β_i, is the total extinction coefficient

$$\beta_i = k_{a_i} + \sigma_{s_i} \quad (3)$$

k_{a_i} is the absorption coefficient, σ_{s_i} is the scattering coefficient, and ρ_j are the diffuse reflectivities, with $j = 1,\ldots,4$.

When the geometry, the radiative properties, and the boundary conditions are known, problem (1-2) may be solved yielding the values of the radiation intensities $I_1(x,\mu)$, for $0 \le x \le L_1$ and $-1 \le \mu \le 1$, and $I_2(x,\mu)$, for $L_1 \le x \le L_1 + L_2$ and $-1 \le \mu \le 1$. This is the direct problem. For the solution of the direct problem we use in the present work a combination of Chandrasekhar's discrete ordinates method (Chandrasekhar, 1960) with the finite difference method (Soeiro and Silva Neto, 2006).

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 the approach considered it was introduced the possibility of using a scale to weight the samples, and sensors to measure temperature in the sample, as well as inside the drying chamber.

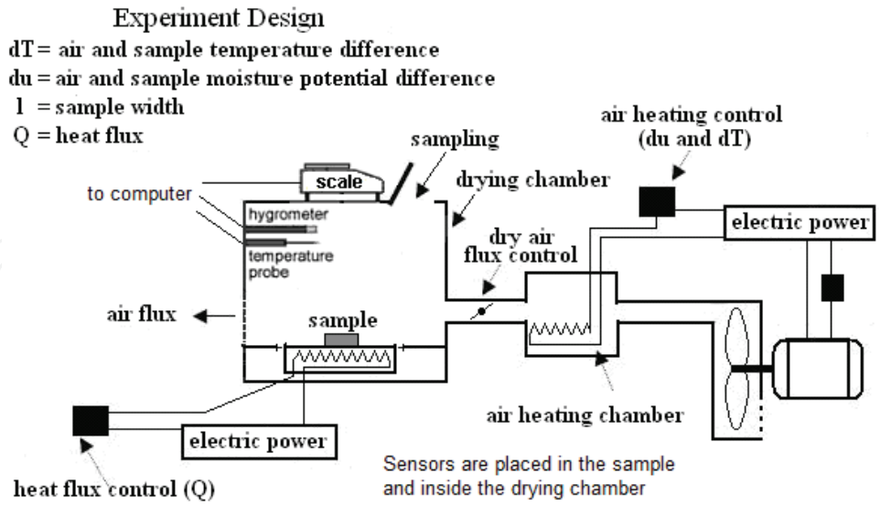


Fig. 2. Drying experiment setup (Adapted from Mwithiga and Olwal, 2005).

In accordance to 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.

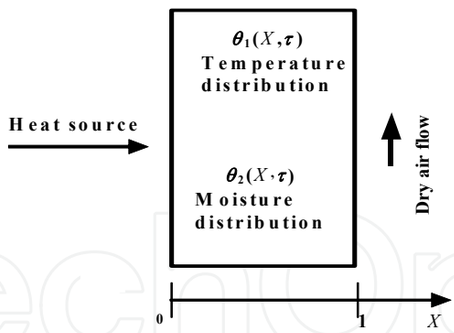


Fig. 3. Drying process schematic representation.

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),

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

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

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 $\tau>0$

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

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

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

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

where

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

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

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

and the dimensionless variables are defined as

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

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

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

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

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

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

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

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

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

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

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

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

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, Silva Neto et al., 2010).

2.3 Gas-liquid adsorption

The mechanism of proteins adsorption at gas-liquid interfaces represented in Fig. 4 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, 1993a,b; Santana, 1994; Krishna and van Baten, 2003; Haut and Cartage, 2005; Mouza et al., 2005; Lugon, 2005).

The system works basically through the gas injection at the base of a column containing the solution. The gas bubbles formed in the distributor rise and along this path adsorb the solute. In the foam region, formed above the bubble column, the extraction of the material of interest is made (see Fig. 4).

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.

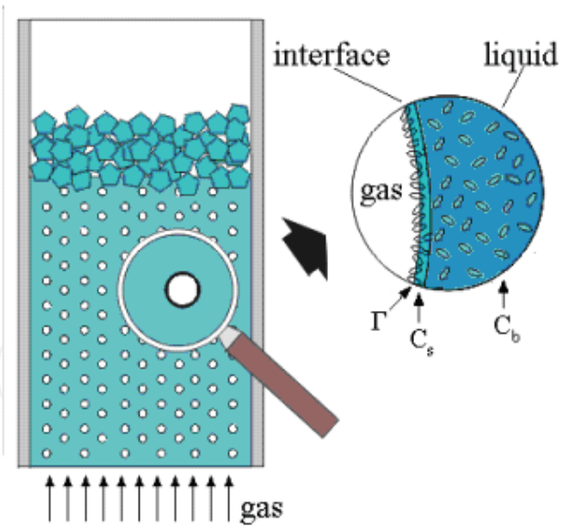


Fig. 4. Gas-liquid adsorption process in a bubble and foam column.

The solute depletion is modeled by

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

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

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

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-layer isotherm

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

where Γ_1 and Γ_2 are the excess superficial concentration in the first and second adsorption layers respectively (see Fig. 4).

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

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

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:

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

where

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

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

$$Bo = \frac{\nu_l}{D_i}, \text{ Bond number} \tag{37}$$

$$Ga = \frac{g d_b^3}{\nu_l^2}, \text{ Galilei Number} \tag{38}$$

where D_i is the tensoactive diffusion coefficient and ν_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 (1993a,b) 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; Silva Neto and Becceneri, 2009), where one seeks to minimize the cost functional of squared residues between the calculated and experimental values for the observable variable,

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

where $\mathbf{V}_{meas}$ is the vector of measurements, $\mathbf{V}_{calc}$ 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

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

The inverse problem solution is the vector $\mathbf{P}^*$ 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 considered, as described in sections 2.1 - 2.3, different measurements are to be taken, that is:

a. Radiative problem

We are interested in obtaining estimates for the vector of unknowns $\mathbf{P}$, given by: σ_{s_1}, k_{a_1}, σ_{s_2} and k_{a_2}. Measured data were used on the emerging radiation intensity acquired at the boundary surfaces $x=0$ and $x=L_1+L_2$, and the interface $x=L_1$, Y_i, with $i=1,2,\ldots,N$, being N the total number of experimental data.

b. Drying problem

Using temperature measurements, T, taken by sensors located inside the medium, and the average of the moisture-transfer potential, $\overline{u}$, 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 phenomenon was modeled using a two-layer isotherm.

4. Solution of the inverse problems with Artificial Neural Networks, simulated annealing and hybrid methods

Instead of going directly to the description of the inverse problem solution methods, we opted for presenting first the approach considered in the analysis of the sensitivity of the observable variables with respect to the unknown parameters to be determined with the inverse problem solution.

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

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

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

$$\mathbf{Y} = \begin{bmatrix} Y_{P_1V_1} & Y_{P_2V_2} & \dots & Y_{P_{N_P}V_1} \\ Y_{P_1V_2} & Y_{P_2V_2} & \dots & Y_{P_{N_P}V_2} \\ \dots & \dots & \dots & \dots \\ Y_{P_1V_m} & Y_{P_2V_m} & \dots & Y_{P_{N_P}V_m} \end{bmatrix} \tag{42}$$

where V_i is a particular measurement of the observable variable, i.e. radiation intensity, concentration, temperature or moisture potential, and m is the total number of measurements. Maximizing the determinant of the matrix $\mathbf{Y}^T\mathbf{Y}$ results in higher sensitivity and uncorrelation (Beck, 1988).

4.2 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.

In order to solve the inverse problem we use here a multi-layer perceptron (MLP) neural network (Soeiro et al., 2004). In Fig. 5 is given a representation of the MLP with the input and output layers, and one hidden layer for the solution of the inverse problem of determining the vector of unknowns $\mathbf{P}$. By providing $\mathbf{V}_{meas}$ at the input layer we expect that the ANN will provide at the output layer an estimate for $\mathbf{P}$.

Each neuron j, with $j = 1,2,\dots,N_H$, in the hidden layer performs a linear combination of the input values provided at the input layer

$$p_j = \sum_{i=1}^{N} w_{ji}^{(1)} x_i + w_{j0}^{(1)} = \sum_{i=1}^{N} w_{ji}^{(1)} Y_i + w_{j0}^{(1)}, \quad j = 1,2,\dots,N_H \tag{43}$$

where $w_{ji}^{(1)}$, $j = 1,2,...,N_H$, $i = 1,2,...,N$ are the weights of the connections between the nodes of the input layer and the neurons of the hidden layer, N is the number of nodes in the input layer, and N_H is the number of neurons in the hidden layer.

The weighted sum p_j given by Eq. (43) is viewed as an excitation to neuron j of the hidden layer, which provides in response

$$q_j = f\left(p_j\right), \quad j = 1,2,...,N_H \tag{44}$$

where $f(.)$ is an activation function. Various choices for the function $f(.)$ are possible (Haykin, 1999).

Each neuron k, $k = 1,2,...,N_{un}$ of the output layer performs a linear combination of the response q_j, $j = 1,2,...,N_H$, of the neurons of the hidden layer

$$s_k = \sum_{j=1}^{N_H} w_{kj}^{(2)} q_j + w_{k0}^{(2)}, \qquad k = 1,2,...,N_{un} \tag{45}$$

where $w_{kj}^{(2)}$, $k = 1,2,...,N_u$, $j = 1,2,...,N_H$, are the weights of the connections between the neurons of the hidden layer and the neurons of the output layer, and N_{un} is the number of

neurons in the output layer, which coincides with the number of unknowns of the inverse problem..

The weighted sum s_k given by Eq. (45) is viewed as an excitation to neuron k of the output layer, which provides in response

$$t_k = g(s_k), \qquad k = 1,2,...,N_{un} \tag{46}$$

where $g(.)$ is an activation function. Various choices for the function $g(.)$ are possible (Haykin, 1999).

Combining Eqs. (43-46) we get

$$t_k = g\left(\sum_{j=1}^{N_H} w_{kj}^{(2)} f\left(\sum_{i=1}^{N} w_{ji}^{(1)} Y_i + w_{j0}^{(1)} \right) + w_{k0}^{(2)} \right) \qquad k = 1,2,...,N_{un} \tag{47}$$

Considering available the experimental data Y_i, $i = 1,2,...,N$, we observe in Eq. (47) that t_k, $k = 1,2,...,N_{un}$, are estimates for the unknowns Z_k, $k = 1,2,...,N_{un}$, obtained by the ANN. But before we can use Eq. (47) we must determine the weight parameters $w^{(1)}$ and $w^{(2)}$.

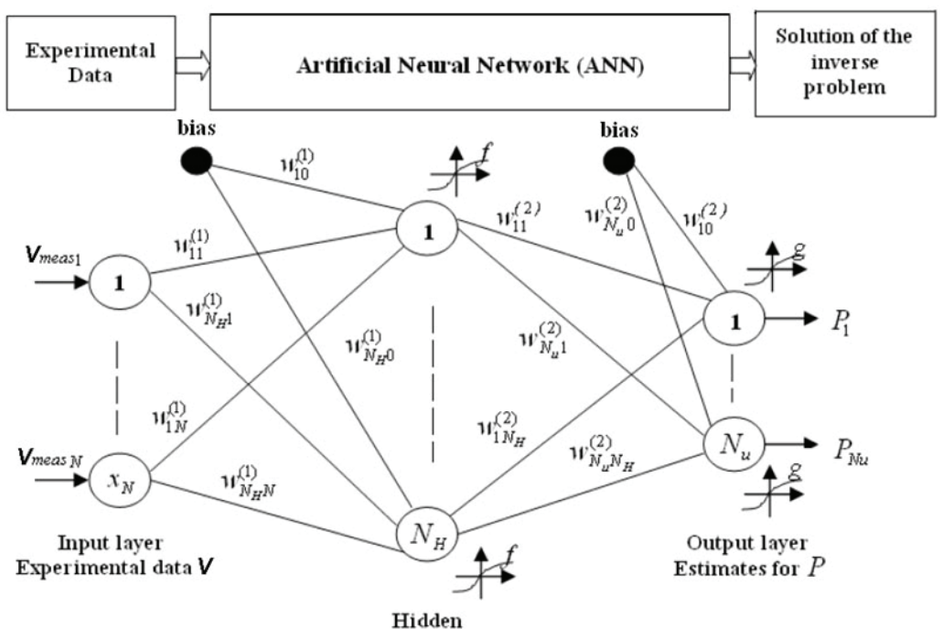


Fig. 5. Multi-layer perceptron network with one hidden layer for the inverse radiative transfer problem.

The determination of the weights $w^{(1)}$ and $w^{(2)}$ is accomplished by presenting a set of patterns (known input $\mathbf{P}_{exact}$ and outputs $\mathbf{V}_{exact}$) and calculating the weights that provides the best match between the calculated values $\mathbf{t}$ and the target values $\mathbf{P}_{exact}$. The patterns used in this supervised training stage of the ANN were generated by calculating the values $\mathbf{V}_{exact}$ from known sets $\mathbf{P}_{exact}$ with the discrete ordinates and finite difference solution (Soeiro and Silva Neto, 2006; Silva Neto and Becceneri, 2009).

For the determination of $w^{(1)}$ and $w^{(2)}$ we used the back propagation algorithm. We start with an initial guess for the weights, $w^{(1)n}$, $w^{(2)n}$, with $n = 0$, and the set of inputs $\mathbf{V}$ is

passed forward through the network yielding trial outputs $\mathbf{t}^{n=0}$ which are compared with the desired outputs $\mathbf{P}_{exact}$ leading to the errors,

$$e_k^n = P_{k\,exact} - t_k^n, \quad k = 1,2,...,N_{un} \tag{48}$$

The weights are then adjusted using the information provided by the output error (Haykin, 1999)

$$w_{kj}^{(2)n+1} = w_{kj}^{(2)n} + \eta^{(2)} \delta_k^{(2)n} q_j^n \tag{49a}$$

$$w_{ji}^{(1)n+1} = w_{ji}^{(1)n} + \eta^{(1)} \delta_j^{(1)n} Y_i \tag{49b}$$

where

$$\delta_k^{(2)n} = e_k^n g'\left(s_k^n\right) \tag{50a}$$

$$\delta_j^{(1)n} = f'\left(p_j^n\right) \sum_{k=1}^{N_u} \delta_k^{(2)n} w_{kj}^{(2)n} \tag{50b}$$

$\eta^{(1)}$ and $\eta^{(2)}$ are the learning rates, which can assume different values for the weights between input-hidden layers (1) and hidden-output layers (2).

The forward and backward sweeps procedure is continued until a convergence criterion related to errors e_k, $k = 1,2,...,N_{un}$, is satisfied.

The presentation of a full set of patterns is denominated epoch. After one epoch is completed the set of patterns is presented again, in a different (random) order. After a number of epochs, once the comparison error is reduced to an acceptable level over the whole training set, the training phase ends and the ANN is established.

4.3 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,

$$P\left(\Delta E\right) = \exp\left(-\Delta E / k_B T\right) \tag{51}$$

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 $S\left(\mathbf{P}\right)$, given by Eq. (39a), in place of energy and defining configurations by a set of variables $\left\{P_i\right\}, i = 1,2,...,N_p$ 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.

4.4 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

$$\frac{dS}{d\mathbf{P}} = \frac{d}{d\mathbf{P}}(\mathbf{F}^T\mathbf{F}) = 0 \rightarrow \mathbf{J}^T\mathbf{F} = 0 \tag{52}$$

where J is the Jacobian matrix, with the elements $J_{ps} = \partial V_p / \partial P_s$ 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,

$$\mathbf{F}(\mathbf{P} + \Delta\mathbf{P}) \cong \mathbf{F}(\mathbf{P}) + \mathbf{J}\Delta\mathbf{P} \tag{53}$$

Introducing the above expansion in Eq. (52) results

$$\mathbf{J}^T\mathbf{J}\Delta\mathbf{P} = -\mathbf{J}^T\mathbf{F}(\mathbf{P}) \tag{54}$$

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

$$\Delta\mathbf{P}^n = -\left[(\mathbf{J}^n)^T\mathbf{J}^n + \gamma^n\mathbf{I}\right](\mathbf{J}^n)^T\mathbf{F}(\mathbf{P}^n) \tag{55}$$

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. (55). Actually in most cases the vector $\Delta\mathbf{P}^n$ is obtained directly from the solution of the linear system of equations (54). This iterative procedure is continued until a convergence criterion such as

$$\left|\Delta P_k^n / P_k^n\right| < \varepsilon, \qquad n = 1, 2, \cdots, N_p \tag{56}$$

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. (54), are calculated at each iteration, using the solution of the direct problem with the estimates for the unknowns obtained in the previous iteration.

In order to calculate the gradient, a central difference approximation was used (Lugon Jr. and Silva Neto, 2010; Silva Neto et al., 2010)

$$\frac{\partial V}{\partial P} = \frac{V\left(P+\Delta P\right) - V\left(P-\Delta P\right)}{2\times\Delta P} \tag{57}$$

In the beginning of the process, an ANN trained to solve the direct problem was used to approximate $V\left(P+\Delta P\right)$ and $V\left(P-\Delta P\right)$. This faster scheme proved to be accurate enough to begin the process. Afterwards, the direct problem solution itself was used in Eq. (57) and although being slower, it offers better results at the final stages of the LM method.

4.5 Hybrid 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 Artificial Neural Networks:

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

Therefore, in all cases studied ANN was used after the training stage in order to quickly provide an inverse problem solution. This solution was used as an initial guess for the LM.

In order to improve the solution, we have also studied the combination of ANN, LM and SA methods. After using LM, reaching within a few iterations a point of minimum, 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.

5. Test case results

As real data for the three problems considered were not available we simulated the experimental data using

$$V_{p_{meas}} = V_{p_{calc}} + r_p\sigma \ , \ p = 1,2,...,N_p \tag{58}$$

where r_p is a random number in the range $\left[-1,1\right]$ and σ simulates the standard deviation of the measurements error.

5.1 Radiative transfer problem

In Table 1 we present the results obtained with the LM method starting with the initial guess: $\sigma_{s_1} = 0.10\ cm^{-1}$, $k_{a_1} = 0.8\ cm^{-1}$, $\sigma_{s_2} = 0.10\ cm^{-1}$ and $k_{a_2} = 0.8\ cm^{-1}$ for the particular case with the exact values for the unknowns $\sigma_{s_1} = 0.45\ cm^{-1}$, $k_{a_1} = 0.05\ cm^{-1}$, $\sigma_{s_2} = 0.45\ cm^{-1}$ and $k_{a_2} = 0.05\ cm^{-1}$ (maximum noise in the experimental data = 8%,

i.e. σ = 0.002 in Eq. (58). Note that LM does not converge with such initial guesses for the unknown parameters.
For the test case presented it is also considered $L_1 = L_2 = 2cm$, $\rho_1 = 0.1$, $\rho_2 = \rho_3 = 0$, $\rho_4 = 0.9$, $f_1 = 1.0$ and $f_2 = 0$, which represents a difficult test case.

Iteration	$\sigma_{s_1}\ (cm^{-1})$	$k_{a_1}\ (cm^{-1})$	$\sigma_{s_2}\ (cm^{-1})$	$k_{a_2}\ (cm^{-1})$	*Obj. Func. [Eq.(39a)]*
0	0.10	0.8	0.10	0.8	7.439
5	0.52	0.049	2.1E09	0.01	6.86E-01
10	0.45	0.05	5.5E07	3.7E07	1.216

Table 1. Estimates obtained with LM (10 iterations). Noisy data (8%)

In Table 2 are shown the results for the same test case using ANNs, and in Tables 3 and 4 are presented the results obtained when the ANN is used to generate the initial guess for the LM method. Here we used noisy data (maximum 8%), i.e., σ = 0.002 in Eq. (58). The experimental data used for the solution of the inverse problem consisted of a set of 40 radiation intensities measured at different polar angles, 20 intensities measured by external detectors and 20 intensities measured by internal detectors located at the interface between the two-layers, i.e. $x = L_1$. Therefore, there are N = 40 entries in the input layer of the ANN. For the hidden layer we considered $N_H = N = 40$. We used 500 patterns (N_P) and a decreasing number of epochs (N_E) in order to save computational time.
In this work the Neural Network Toolbox of the software MATLAB (Mathworks, Inc.) was used with the following neuron model in the backpropagation network: 40 elements in the input vector, log-sigmoid (logsig) transfer (activation) function between the input layer and the hidden layer (with 40 elements) and a linear transfer function (purelin) in the output layer (with 4 elements in the output vector).

N_E	CPU *time(min)*	Estimates (ANN)			
		σ_{s_1} (cm^{-1})	k_{a_1} (cm^{-1})	σ_{s_2} (cm^{-1})	k_{a_2} (cm^{-1})
500	120	0.40	0.01	0.43	0.01
200	48	0.34	0.01	0.33	0.01
100	22	0.35	0.09	0.32	0.09
30	7,5	0.50	0.10	0.60	0.05

Exact values: $\sigma_{s_1} = \sigma_{s_2} = 0.45cm^{-1}$, $k_{a_1} = k_{a_2} = 0.05cm^{-1}$

Table 2. Neural Network solutions for the inverse problem and CPU time considering different number of epochs (N_H = *40*, N_P = *500)* and noisy data (8%)

It can be observed from Table 2 that the ANN did not provide good estimates for the unknowns. An improvement can be obtained, but at the expense of a higher CPU time requirement. A different strategy is then adopted with a hybridization ANN-LM. In Table 3 are presented the results obtained with such hybridization in which the former method provides an initial guess for the latter. The solution of the ANN was obtained considering 100 epochs in the training stage of the ANN. An improvement in the results of the inverse problem is then observed.

Noise	ANN estimates				Results (LM)			
	σ_{s_1} (cm^{-1})	k_{a_1} (cm^{-1})	σ_{s_2} (cm^{-1})	k_{a_2} (cm^{-1})	σ_{s_1} (cm^{-1})	k_{a_1} (cm^{-1})	σ_{s_2} (cm^{-1})	k_{a_2} (cm^{-1})
8%	0.35	0.09	0.32	0.09	0.447	0.050	0.455	0.050
4%	0.40	0.03	0.45	0.04	0.450	0.050	0.445	0.049
2%	0.39	0.04	0.42	0.05	0.450	0.049	0.449	0.050
0%	0.37	0.04	0.40	0.05	0.45	0.05	0.45	0.05

Table 3. Combined method results with ANN to obtain estimates for the LM, with noisy data and number of epochs N_E = 100. Exact values $\sigma_{s_1} = 0.450$, $k_{a_1} = 0.05$, $\sigma_{s_2} = 0.450$ and $k_{a_2} = 0.050$

In Table 4 are shown the results obtained using also the hybridization ANN-LM, but now with only 30 epochs in the training stage of the ANN. It can also be observed that very good results are obtained for the inverse problem.

Noise	ANN estimates				Results (ANN-LM)			
	σ_{s_1} (cm^{-1})	k_{a_1} (cm^{-1})	σ_{s_2} (cm^{-1})	k_{a_2} (cm^{-1})	σ_{s_1} (cm^{-1})	k_{a_1} (cm^{-1})	σ_{s_2} (cm^{-1})	k_{a_2} (cm^{-1})
8%	0.50	0.10	0.60	0.05	0.447	0.050	0.455	0.050
4%	0.49	0.01	0.47	0.01	0.450	0.050	0.445	0.049
2%	0.54	0.03	0.53	0.04	0.449	0.050	0.449	0.050
0%	0.51	0.02	0.49	0.04	0.45	0.05	0.450	0.050

Table 4. Combined method results with ANN providing estimates for the LM, with noisy data and number of number of epochs N_E = 100. Exact values $\sigma_{s_1} = 0.450$, $k_{a_1} = 0.05$, $\sigma_{s_2} = 0.450$ and $k_{a_2} = 0.050$

It must be stressed that the solution of the inverse problem with either the LM or ANN methods using only external detectors led to non-unique solutions of the inverse radiative transfer problem. That is the reason why internal detectors located at the interface of the two layers were also considered for the solution of the inverse problem.

5.2 Drying (simultaneous heat and mass transfer)

Much research effort has already been made in order to estimate the Possnov, Kossovitch, heat Biot and mass Biot numbers (Dantas et al., 2003; Huang and Yeh, 2002; Lugon Jr. and Silva Neto, 2004), but it was considered the possibility of optimizing the number and location of temperature sensors, experiment duration, etc. In this work instead, δ, r/c, h/k and h_m/k_m are estimated using an "optimum" experiment (Dowding et al., 1999 and Beck, 1988) for wood drying, and doing so, it was also considered the following process control parameters: heat flux, Q, the medium width, l, the difference between the medium and the air temperatures, $dT = T_s - T_0$, and the difference between the medium and the air moisture potential, $du = u_0 - u^*$.

There is no difference between the sensitivity coefficients for the two sets of variables, that is, the scaled sensitivity coefficients are exactly the same for both vectors $\{Lu, Pn, Ko, Bi_q, Bi_m, \varepsilon\}^T$ and $\{Lu, \delta, r/c, h/k, h_m/k_m, \varepsilon\}^T$,

$$SC_{\delta}(X,\tau)=\delta\frac{\partial V(X,\tau)}{\partial\delta}=Pn\frac{\partial V(X,\tau)}{\partial Pn}=SC_{Pn}(X,\tau) \tag{59a}$$

$$SC_{r/c}(X,\tau)=r/c\frac{\partial V(X,\tau)}{\partial r/c}=Ko\frac{\partial V(X,\tau)}{\partial Ko}=SC_{Ko}(X,\tau) \tag{59b}$$

$$SC_{h/k}(X,\tau)=h/k\frac{\partial V(X,\tau)}{\partial h/k}=Bi_q\frac{\partial V(X,\tau)}{\partial Bi_q}=SC_{Bi_q}(X,\tau) \tag{59c}$$

$$SC_{h_m/k_m}(X,\tau)=h_m/k_m\frac{\partial V(X,\tau)}{\partial h_m/k_m}=Bi_m\frac{\partial V(X,\tau)}{\partial Bi_m}=SC_{Bi_m}(X,\tau) \tag{59d}$$

The reasons for changing the estimated variables are the use of the design of experiment tools and interpretation. Consider the heat and mass Biot numbers for example. If one changes the media width, l, both heat and mass Biot numbers change. The mathematical problem would be different, even though the material is still the same, because one is estimating two different heat and mass Biot numbers. In order to solve this problem, it was decided to estimate the relation between heat transfer coefficient and thermal conductivity, h/k, and the relation between mass transfer coefficient and mass conductivity, h_m/k_m, so that we could change the media width and continue with the same value for both variables to be estimated.

The same idea was used, choosing to estimate the thermogradient coefficient (δ) and the relation between latent heat of evaporation and specific heat of the medium (r/c), instead of the Possnov (Pn) and Kossovitch (Ko) numbers. Doing so, one is able to optimize the experiment considering the difference between the medium and the air temperatures, $dT=T_s-T_0$, and the difference between the moisture-transfer potential between the medium and the air, $du=u_0-u^*$, without affecting the estimated parameters values.

In Fig. 7 is represented the variation of the value of the matrix $\mathbf{Y}^T\mathbf{Y}$ determinant as a function of the temperature differences and moisture potential differences between the medium and the air flowing over it. It is not difficult to understand that one could not build such a graph using a vector of unknown parameters containing Possnov (Pn) and Kossovitch (Ko) numbers. In order to achieve greater sensitivities, while the temperature difference has to be the lowest, the moisture potential difference has to be the highest possible. The solid square represents the chosen designed experiment, considering the existence of practical difficulties that may limit our freedom of choice.

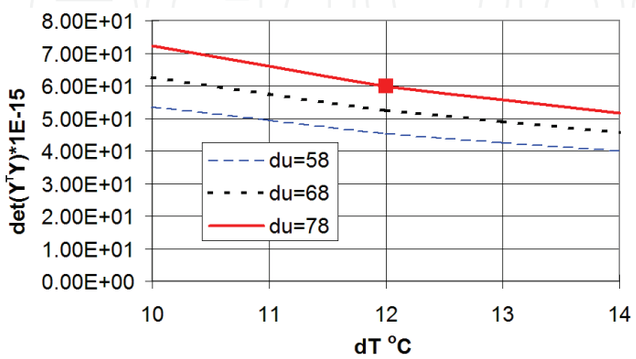


Fig. 7. Determinant of matrix $\mathbf{Y}^T\mathbf{Y}$ as a function of temperature (dT) and moisture potential ($d\overline{u}$) differences.

In Fig. 8 are represented the values of the determinant of matrix $\mathbf{Y}^T\mathbf{Y}$ for different values of the heat flux Q and media thickness l. It is also easy to understand that one could not build such a graph using a vector of unknown parameters containing heat and mass Biot numbers. For practical reasons it was chosen to limit the sample temperature to 130° C. In Fig. 8 the same curve has a continuous-line part and a dashed-line one, when the sample temperature exceeds the limit of 130° C. The solid square shows the chosen designed experiment.

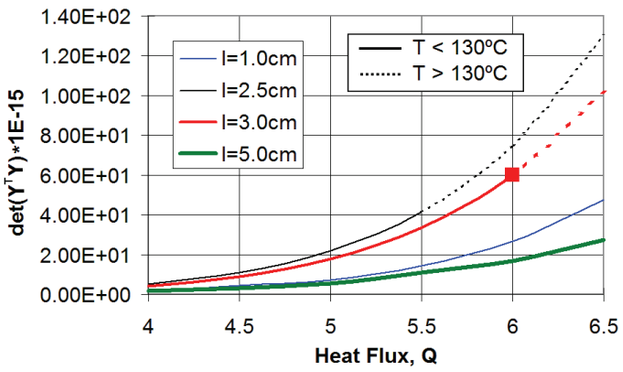


Fig. 8. Determinant of $\mathbf{Y}^T\mathbf{Y}$ matrix for different values of the heat flux Q and medium thickness l.

Considering the previous analysis of the sensitivity graphs and matrix the $\mathbf{Y}^T\mathbf{Y}$ determinant, it was designed the experiment whose geometric and process parameters are shown in Table 5. Since the average moisture potential, $\overline{u}$, is more difficult to measure than the temperature, θ_1, the measurement interval for the average moisture potential, $\Delta\tau_{\overline{u}}$, was considered larger than the interval for the temperature $\Delta\tau_{\theta_1}$.

Geometric or process parameter	Values	Geometric or process parameter	Values
$dT = T_s - T_0$	12 °C	Q	6.0
T_0	24 °C	l	0.03 m
T_s	36 °C	τ_0	0
$du = u_0 - u^*$	78 °M	τ_f	20
u_0	86 °M	$\Delta\tau_{\theta_1}$	0.2
u^*	8 °M	$\Delta\tau_{\overline{u}}$	1
ε	0.2	-	-

τ_0 and τ_f represent the initial and sampling times, respectively.

Table 5. Reference values for the designed experiment.

An experiment was designed to perform the simultaneous estimation of Lu, δ, r/c, h/k and h_m/k_m. In order to study the proposed method, since real experiment data were not available, we generated synthetic data using

$$\theta_{1meas_i} = \theta_{1calc_i}(\vec{P}_{exact}) + \sigma_{\theta_1} r_i,\ i = 1,\ 2,\ \ldots,\ M_{\theta_1} \tag{60a}$$

$$\overline{u}_{meas_i} = \overline{u}_{calc_i}(\vec{P}_{exact}) + \sigma_{\overline{u}} r_i \ , \ i = 1,\ 2,\ ...,\ M_{\overline{u}} \tag{60b}$$

where r_i are random numbers in the range [-1,1], M_{θ_1} and $M_{\overline{u}}$ represent the total number of temperature and moisture-transfer potential experimental data, and σ_{θ_1} and $\sigma_{\overline{u}}$ emulates the standard deviation of measurement errors. It was established a standard deviation of $\sigma_{\theta_1} = 0.03$ considering 100 temperature measurements ($\Delta\tau = 0.2$), resulting in a maximum error of 2%, and $\sigma_{\overline{u}} = 0.001$ considering 20 moisture measurements ($\Delta\tau = 1.0$), resulting in a maximum error of 4%.
In Fig. 9 the temperature (θ_1) and moisture potential ($\overline{\theta}_2$) measurements are presented. The continuous line represents the direct problem solution and the squares represent noisy data. In order to show a better representation, only 20 temperature (θ_1) measurements were represented.

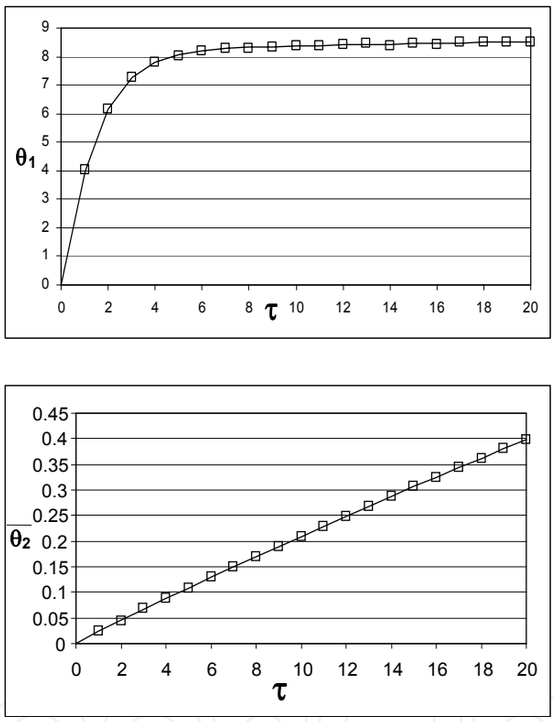


Fig. 9. Temperature (θ_1) and moisture potential ($\overline{\theta}_2$) artificially simulated data.

The results obtained using the methods LM 1 (gradient approximated by FDM - Finite Difference method), LM 2 (gradient approximated by Artificial Neural Networks), ANN, SA and hybrid combinations, for different levels of noise represented by different values of the standard deviation of measurements errors in temperature and average moisture potential, σ_T and $\sigma_{\overline{u}}$, respectively in Eqs. (60a,b) are shown in Table 6.
One observes that when there is no noise, that is, the standard deviation of measurements errors are zero, the LM method was able to estimate all variables very quickly (see test cases 1 and 2). When noise is introduced, the LM is retained by local minima (test cases 3 and 4); the ANN did not reach a good solution, but quickly got close to it (test case 5). The ANN solution was then used as a first guess for the LM method with good performance in test

cases 6 and 7. The SA reached a good solution but required the largest CPU time, and finally the combination of all methods was able to reach a good solution, without being retained by local minima, and also without taking too much time, i.e. one sixth of the SA time. The time shown in the eleventh column of Table 6 corresponds to the CPU time on a Pentium IV 2.8 GHz processor (Lugon Jr., 2005; Silva Neto et al., 2010).

Case	Method	σ_{θ_1}	σ_{ε}	Information	Lu	δ	r/c	h/k	h_m/k_m	Time (s)	S Eq. (39a)
-	-	-	-	Exact values	0.0080	2.0	10.83	34.0	114.0	-	-
1	LM 1 (grad. FDM)	0	0	Initial guess	0.0040	1.50	8.00	20.0	80.0	15	0
				Result $\vec{Z}_{LM\,FDM}$	0.0080	2.00	10.83	34.0	114.0		
2	LM 2 (grad. ANN)	0	0	Initial guess	0.0040	1.50	8.00	25.0	80.0	10	0
				Result $\vec{Z}_{LM\,ANN}$	0.0080	2.00	10.83	34.0	114.0		
3	LM 1 (grad. FDM)	0.03	0.001	Initial guess	0.0040	1.50	8.00	20.0	80.0	15	977
				Result $\vec{Z}_{LM\,FDM}$	0.0076	2.09	10.76	34.1	121.2		
4	LM 2 (grad. ANN)	0.03	0.001	Initial guess	0.0040	1.50	8.00	20.0	80.0	11	897
				Result $\vec{Z}_{LM\,ANN}$	0.0093	1.71	10.73	34.1	95.7		
5	ANN (without initial guess)	0.03	0.001	Result $\vec{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1	1	3190
6	LM 1 (grad. FDM)	0.03	0.001	Initial guess $\vec{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1	16	974
				Result $\vec{Z}_{LM\,FDM}$	0.0083	1.92	10.75	34.1	110.0		
7	LM 2 (grad. ANN)	0.03	0.001	Initial guess $\vec{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1	11	903
				Result $\vec{Z}_{LM\,ANN}$	0.0082	1.79	9.89	35.1	114.5		
8	SA (SA 20,000 evaluations)	0.03	0.001	Initial guess	0.0040	1.50	8.00	25.0	80.0	300	856
				Result $\vec{Z}_{SA}$	0.0094	1.58	9.96	35.0	98.2		
9	ANN-LM 2-SA (SA 2,000 evaluations)	0.03	0.001	Initial guess $\vec{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1	47	760
				Result $\vec{Z}_{LM\,ANN}$	0.0082	1.79	9.89	35.1	114.5		
				Result $\vec{Z}_{SA}$	0.0079	2.01	11.00	33.9	113.8		
				Result $\vec{Z}_{LM\,ANN}$	0.0080	2.05	10.93	33.8	113.9		

Table 6. Results obtained using LM 1 (partial derivatives obtained with finite differences), LM 2 (partial derivatives obtained with Artificial Neural Network), ANN, and, and hybrid combinations.

5.3 Gas-liquid adsorption

Recently, the inverse problem of interface adsorption has attracted the attention of an increasing number of researchers (Lugon Jr., 2005; Forssén et al., 2006; Garnier et al., 2007; Voelkel and Strzemiecka, 2007; Ahmad and Guiochon, 2007).

Based on sensitivity analysis we concluded that in order to solve the inverse problem of gas-liquid adsorption, considering the two-layer isotherm given by Eq. (32), it was necessary to design two different experiments. One to estimate $K_2(T)$ and $\hat{a}$, called experiment 1, and another one to estimate λ, called experiment 2. In all cases studied the sensitivity to $K_1(T)$ is low and therefore this parameter was not estimated with the inverse problem solution.

In Fig. 10 are shown the sensitivity coefficients related to the parameters $K_1(T)$, $K_2(T)$, λ and $\hat{a}$ in experiment 1. It is observed that the sensitivity to $K_2(T)$ and $\hat{a}$ for BSA (Bovine

Serum Albumin) are higher than the sensitivity to the other parameters and their shapes are different, indicating that these variable are uncorrelated.

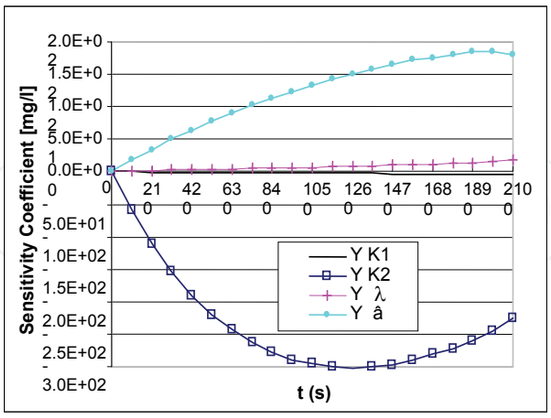


Fig. 10. Scaled sensitivity coefficients for BSA – Experiment 1.

In Fig. 11 are shown the sensitivity coefficients related to the parameters $K_1(T)$, $K_2(T)$, λ and $\hat{a}$ for BSA in experiment 2. It is observed that the sensitivity to λ is higher than the sensitivity to the other parameters.

Another important tool used in the design of experiments is the study of the matrix $\mathbf{Y}^T\mathbf{Y}$, that is, maximizing the determinant of the matrix $\mathbf{Y}^T\mathbf{Y}$ results in higher sensitivity and uncorrelation (Dowding et al., 1999).

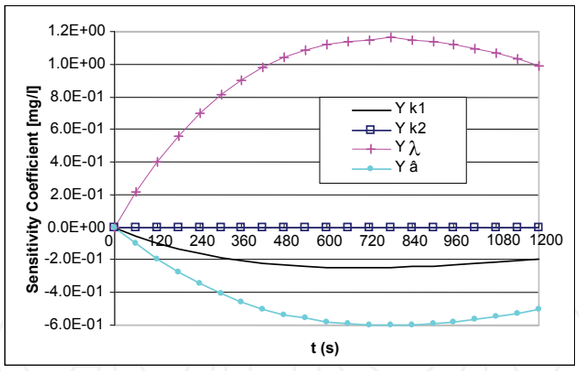


Fig. 11. Scaled sensitivity coefficients for BSA – Experiment 2.

The difference between the two experiments is related to the BSA concentration, being larger in the first experiment (see Table 7).

In Fig. 12 are shown the values of the determinant of the matrix $\mathbf{Y}^T\mathbf{Y}$ for BSA in experiment 1. The designed experiment is marked with a full square. Its choice is justified by the small gain in sensitivity considering the operational difficulties in using a longer column or a higher superficial velocity.

Considering the analysis of the sensitivity graphs and the determinant of the matrix $\mathbf{Y}^T\mathbf{Y}$, two experiments were designed, one to estimate $K_2(T)$ and $\hat{a}$, and another to estimate λ, as shown in Table 7.

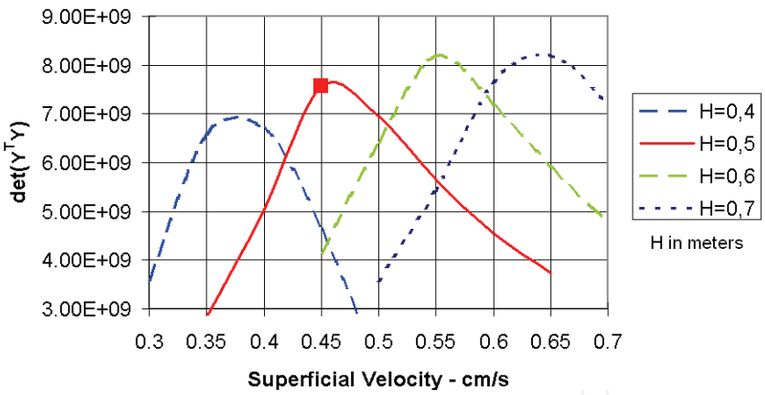


Fig. 12. Matrix $Y^T Y$ determinant for BSA – Experiment 1.

		Experiment	
	Units	1	2
Estimated parameters	-	K_2, $\hat{a}$	λ
Initial solute concentration, C_{b0}	g/m³	1,000	10
Bubble column height, H	m	0.50	0.80
Superficial velocity, v_g	m/s	4.50E-3	1.00E-3
First measurement	s	210	120
time measurement steps	s	210	120
Last measurement	s	2100	1200

Table 7. Reference values for the designed experiment (Lugon, 2005, Lugon et al., 2009).

The results achieved using the ANN, LM 1 (gradient approximated by FDM), LM 2 (gradient approximated by ANN), SA and hybrid combinations, for different standard deviations for the measurements errors, σ, are shown in Tables 8 and 9.

In Table 8 are presented the results obtained for the estimation of $K_2(T)$ and $\hat{a}$, using the designed experiment number 1. Test cases 3-9 used simulated artificial data generated with

Case	Method	σ	Information	k_2	$\hat{a}$	Time (s)	S [mg²/l²] Eq. (39a)
1	LM 1 (grad. FDM)	0	Result $\vec{Z}_{LM^{FDM}}$	0.01040	0.322	169	0
2	LM 2 (grad. ANN)	0	Result $\vec{Z}_{LM^{ANN}}$	0.01040	0.322	80	0
3	LM 1 (grad. FDM)	10	Result $\vec{Z}_{LM^{FDM}}$	0.00790	0.158	170	8.39
4	LM 2 (grad. ANN)	10	Result $\vec{Z}_{LM^{ANN}}$	0.00805	0.157	78	8.64
5	RNA	10	Result $\vec{Z}_{ANN}$	0.01101	0.377	1	6.81
6	LM 1 (grad. FDM)	10	Result $\vec{Z}_{LM^{FDM}}$	0.01080	0.335	172	6.27
7	LM 2 (grad. ANN)	10	Result $\vec{Z}_{LM^{ANN}}$	0.01058	0.314	79	5.68
8	SA (2.000 evaluations)	10	Result $\vec{Z}_{SA}$	0.01050	0.312	6034	4.22
9	ANN-LM-SA SA (200 evaluations)	10	Result $\vec{Z}_{LM^{ANN}}$	0.01101	0.377	682	4.16
			Result $\vec{Z}_{LM^{ANN}}$	0.01058	0.335		
			Result $\vec{Z}_{SA}$	0.01054	0.314		

The exact values used are: $k_2 = 0.0104 mg/(m^2 wt\%)$ and $\hat{a} = 0.322\, m^2/mg$.

Table 8. Results obtained using ANN, LM 1, LM 2, SA and hybrid combinations for experiment 1.

the direct problem solution corrupted with white gaussian noise with standard deviation $\sigma = 10 mg/l$, which corresponds to measurement errors of the order of 4%. While in test cases 1, 2, 3, 4 and 8 the initial guesses are $K_2 = 0.0080 mg/(m^2 wt\%)$ and $\hat{a} = 0.100 m^2/mg$, in test cases 6, 7 and 9 the initial guesses are the estimates obtained with the ANN.
In Table 9 are presented the results obtained for the estimation of λ, using the designed experiment number 2. Test cases 3-9 used simulated artificial data generated with the direct problem solution corrupted with white gaussian noise with standard deviation $\sigma = 0.10 mg/l$, which corresponds to measurement errors of the order of 3%. While in test cases numbers 1, 2, 3, 4 and 8 the initial guess is $\lambda = 0.700 m^2/mg$, in test cases 6, 7 and 9 the initial guesses are the estimates obtained with the ANN.

Case	Method	σ	Information	λ	Time (s)	S [mg²/l²] Eq. (39a)
1	LM 1 (grad. FDM)	0	Result $\vec{Z}_{LM\,FDM}$	1.117	40	0
2	LM 2 (grad. ANN)	0	Result $\vec{Z}_{LM\,ANN}$	1.117	29	0
3	LM 1 (grad. FDM)	0.1	Result $\vec{Z}_{LM\,FDM}$	1.159	45	7.96
4	LM 2 (grad. ANN)	0.1	Result $\vec{Z}_{LM\,ANN}$	1.159	30	7.96
5	RNA	0.1	Result $\vec{Z}_{ANN}$	1.432	1	202.9
6	LM 1 (grad. FDM)	0.1	Result $\vec{Z}_{LM\,FDM}$	1.159	6	7.96
7	LM 2 (grad. ANN)	0.1	Result $\vec{Z}_{LM\,ANN}$	1.159	4	7.96
8	SA (2.000 evaluations)	0.1	Result $\vec{Z}_{SA}$	1.099	5937	10.12
9	ANN-LM-SA SA (200 evaluations)	0.1	Result $\vec{Z}_{ANN}$	1.432	601	7.92
			Result $\vec{Z}_{LM\,ANN}$	1.159		
			Result $\vec{Z}_{SA}$	1.156		

The exact value used is: $\lambda = 1.117 m^2/mg$.

Table 9. Results obtained using ANN, LM 1, LM 2, SA and hybrid combinations for experiment 2.

6. Conclusions

6.1 Radiative transfer

In this case, Artificial Neural Networks (ANN), Levenberg-Marquardt (LM) and hybrid combinations of methods were used to solve the inverse radiative transfer problem. The solution with ANN and LM methods using only external detectors led to non-unique solutions of the inverse radiative transfer problem. It was demonstrated that the hybid combination of ANN-LM obtained better results than using either methods alone.

6.2 Drying (simultaneous heat and mass transfer)

The direct problem of simultaneous heat and mass transfer in porous media modeled with Luikov equations can be solved using the finite difference method, yielding the temperature and moisture distribution in the media, when the geometry, the initial and boundary conditions, and the medium properties are known.
Inverse problem techniques can be useful to estimate the medium properties when they are not known. After the use of an experiment design technique, the hybrid combination ANN-LM-SA resulted in good estimates for the drying inverse problem using artificially generated data.
The design of experiments technique is of great importance for the success of the estimation efforts, while previous works studied the estimation of Lu, Pn, Ko, Bi_q and Bi_m, here is

considered Lu, δ, r/c, h/k and h_m/k_m. The main advantage of such approach is to be able to design an "optimum" experiment using different medium width, l, porous medium and air temperature difference, $T_s - T_0$, and porous medium and air moisture potential difference, $u_0 - u^*$.

The combination of deterministic (LM) and stochastic (ANN and SA) methods achieved good results, reducing the time needed and not being retained by local minima. The use of ANN to obtain the derivatives in the first steps of the LM method reduced the time required for the solution of the inverse problem.

6.3 Gas-liquid adsorption

After the use of an experiment design technique, the hybrid combination ANN-LM-SA resulted in good solutions for the gas-liquid adsorption isotherm inverse problem.

The use of the ANN to obtain the derivatives in the first step of the LM method reduced the time necessary to solve the inverse problem.

7. Acknowledgements

The authors acknowledge the financial support provided by CNPq, *Conselho Nacional de Desenvolvimento Científico e Tecnológico*, FAPERJ, *Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro*, and FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo.

8. References

Ahmad, T. and Guiochon, G, 2007, Numerical Determination of the Adsorption Isotherms of Tryptophan at Different Temperatures and Mobile Phase Composition, J. of Chromatography A, v. 1142, pp. 148-163.

Beck, J. V., 1988, Combined Parameter and Function Estimation in Heat Transfer with Application to Contact Conductance, J. Heat Transfer, v. 110, pp. 1046-1058.

Chandrasekhar, S, 1960, Radiative Transfer, Dover Publications Inc., New York.

Dantas, L. B., Orlande, H.R.B. and Cotta, R. M., 2003, An Inverse Problem of Parameter Estimation for Heat and Mass Transfer in Capillary Porous Media, Int. J. Heat Mass Transfer, v. 46, pp. 1587-1598.

Deckwer, W. R. and Schumpe, A., 1993, Improved Tools for Bubble Columns Reactor Design and Scale-up, Chem. Eng. Sc., v.48, No., pp. 889-911.

Dowding, K. J., Blackwell, B. F. and Cochran, R. J., 1999, Applications of Sensitivity Coefficients for Heat Conduction Problems, Numerical Heat Transfer, Part B, v. 36, pp. 33-55.

Forssén, P., Arnell, R. and Fornstedt, T., 2006, An Improved Algorithm for Solving Inverse Problems in Liquid Chromatography, Computers and Chemical Engineering, v.30, pp.1381-1391.

Garnier, C., Görner, T., Villiéras, F., De Donato, Ph., Polakovic, M., Bersillon, J. L. and Michot, L. J., 2007, Activated Carbon Surface Heterogeneity Seen by Parallel Probing by Inverse Liquid Chromatography at the Solid/Liquid Interface and by Gas Adsorption Analysis at the Solid/Gas Interface, Carbon, v. 45, pp. 240-247.

Graham, D. E. and Phillips, M. C., 1979, Proteins at Liquid Interfaces, II. Adsorption Isotherms, J. Colloid and Interface Science, v. 70, No. 3, pp. 415-426.

Haut, B. and Cartage, T., 2005, Mathematical Modeling of Gas–liquid Mass Transfer Rate in Bubble Columns Operated in the Heterogeneous Regime, Chem. Eng. Science, v. 60, n. 22, pp. 5937-5944.

Haykin, S., 1999, Neural Networks - A Comprehensive Foundation, Prentice Hall.

Huang, C. H. and Yeh, C. Y., 2002, An Inverse Problem in Simultaneous Estimating the Biot Number of Heat and Moisture Transfer for a Porous Material, Int. J. Heat Mass Transfer, v. 45, pp. 4643-4653.

Kirkpatrick, S., Gellat, Jr., C. D. and Vecchi, M. P., 1983, Optimization by Simulated Annealing, Science, v. 220, pp. 671-680.

Krishna, R. and van Baten, J. M., 2003, Mass Transfer in Bubble Columns, Catalysis Today, v. 79-80, pp. 67-75.

Lugon Jr., J. and Silva Neto, A. J., 2004, Deterministic, Stochastic and Hybrid Solutions for Inverse Problems in Simultaneous Heat and Mass Transfer in Porous Media, Proc. 13th Inverse Problems in Engineering Seminar, pp. 99-106, Cincinatti, USA.

Lugon Jr., J., 2005, Gas-liquid Interface Adsorption and One-dimensional Porous Media Drying Inverse Problems Solution, D.Sc Thesis, Universidade do Estado do Rio de Janeiro (in Portuguese).

Lugon Jr., J, Silva Neto, A. J. and Santana, C. C., 2009, A Hybrid Approach with Artificial Neural Networks, Levenberg-Marquardt and Simulated Annealing Methods for the Solution of Gas-Liquid Adsorption Inverse Problems, Inverse Problems in Science and Engineering, Vol. 17, No. 1, pp.85-96.

Lugon Jr., J and Silva Neto, A. J., 2010, Solution of Porous Media Inverse Drying Problems Using a Combination of Stochastic and Deterministic Methods, Journal of the Brazilian Society of Mechanical Sciences and Engineering, Accepted for publication.

Luikov, A. V. and Mikhailov, Y. A., 1965, Theory of Energy and Mass Transfer, Pergamon Press, Oxford, England.

Marquardt, D. W., 1963, An Algorithm for Least-Squares Estimation of Nonlinear Parameters, J. Soc. Industr. Appl. Math., v. 11, pp. 431-441.

Metropolis, N., Rosenbluth, A. W., Teller, A. H. and Teller, E., 1953, Equation of State Calculations by Fast Computing Machines, J. Chem. Physics, v.21, pp.1087-1092.

Mikhailov, M. D. and Özisik, M. N., 1994, Unified Analysis and Solutions of Heat and Mass Diffusion, Dover Publications, Inc.

Mouza, A. A., Dalakoglou, G. K. and Paras, S. V., 2005, Effect of Liquid Properties on the Performance of Bubble Column Reactors with Fine Pore Spargers, Chem. Eng. Science, v. 60, n. 5, pp. 1465-1475.

Mwithiga, G., and Olwal, J. O., 2005, The drying kinetics of kale (Brassica oleracea) in a convective hot air dryer, J. of Food Engineering, v. 71, No. 4, pp. 373-378.

Özişik, M. N., 1973, Radiation Transfer and Iterations with Conduction and Convection, John Wiley.

Özturk, S. S., Schumpe, A. and Deckwer, W.D., 1987, Organic Liquids in a Bubble Column: Holdups and Mass Transfer Coefficients, AIChE Journal, v. 33, No. 9, pp. 1473-1480.

Santana, C. C. and Carbonell, R. G., 1993a, Waste Minimization by Flotation: Recovery of Proteins and Other Surface-Active Compounds, 3rd Int. Conf. Waste Management, Bahia, Brazil.

Santana, C. C. and Carbonell, R. G., 1993b, Adsorptive Bubble Separation as a Means of Reducing Surface-Active Contaminants in Industrial Wastewaters, Proc. Int. Symp. on Heat and Mass Transfer, Cancun, Mexico, pp. 1-11.

Santana, C. C., 1994, Adsorptive Bubble Separation Process as a Means of Reducing Surface-Active Contaminants in Industrial Wastewaters, Brazilian J. Engineering-Chemistry, v. 5, pp. 7-74.

Silva Neto, A. J. and Becceneri, J. C. (Eds.), 2009, Nature Inspired Computational Intelligence Techniques – Application in Inverse Radiative Transfer Problems, SBMAC, São Carlos, Brazil. (in portuguese).

Silva Neto, A. J., Lugon Jr., J., Soeiro, F. J. C. P., Biondi Neto, L, Santana, C. C., Lobato, F. S. and Steffen Jr., V., 2010, Application of Simulated Annealing and Hybrid Methods in the Solution of Inverse Heat and Mass Transfer Problems, In Simulated Annealing, Theory with Applications, Ed. Chibante, R., Chapter 2, pp. 17-50, ISBN 978-953-307-134-3, Sciyo, Croatia

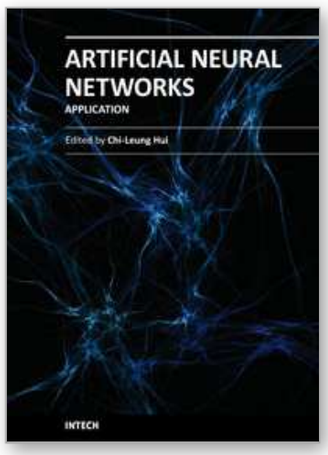
Silva Neto, A. J. and Moura Neto, F. D., 2005, Inverse Problems: Fundamental Concepts and Applications, EdUERJ, Rio de Janeiro (in Portuguese).

Silva Neto, A. J. and Soeiro, F. J. C. P., 2003, Solution of Implicitly Formulated Inverse Heat Transfer Problems with Hybrid Methods, Mini-Symposium Inverse Problems from Thermal/Fluids and Solid Mechanics Applications - 2nd MIT Conference on Computational Fluid and Solid Mechanics, Cambridge, USA.

Soeiro, F.J.C.P., Soares, P.O. and Silva Neto, A.J., 2004, Solution of Inverse Radiative Transfer Problems with Artificial Neural Networks and Hybrid Methods, Proc. 13th Inverse Problems in Engineering Seminar, pp. 163-169, Cincinnati, USA.

Soeiro, F. J. C. P. and Silva Neto, A. J., 2006, Inverse Radiative Transfer Problems in Two-Layer Participating Media, Proc. III European Conference on Computational Mechanics, Lisbon, Portugal.

Voelkel, A. and Strzemiecka, B., 2007, Characterization of Fillers Used in Abrasive Articles by Means of Inverse Gas Chromatography and Principal Components Analysis, Int. J. of Adhesion and Adhesives, v. 27, pp. 188-194.



Artificial Neural Networks - Application

Edited by Dr. Chi Leung Patrick Hui

ISBN 978-953-307-188-6
Hard cover, 586 pages
Publisher InTech
Published online 11, April, 2011
Published in print edition April, 2011

This book covers 27 articles in the applications of artificial neural networks (ANN) in various disciplines which includes business, chemical technology, computing, engineering, environmental science, science and nanotechnology. They modeled the ANN with verification in different areas. They demonstrated that the ANN is very useful model and the ANN could be applied in problem solving and machine learning. This book is suitable for all professionals and scientists in understanding how ANN is applied in various areas.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jader Lugon Junior, Antônio J. da Silva Neto, Luiz Biondi Neto, Francisco José da Cunha Pires Soeiro, Cesar Costapinto Santana and Haroldo F. Campos Velho (2011). Application of Artificial Neural Networks and Hybrid Methods in the Solution of Inverse Problems, Artificial Neural Networks - Application, Dr. Chi Leung Patrick Hui (Ed.), ISBN: 978-953-307-188-6, InTech, Available from: http://www.intechopen.com/books/artificial-neural-networks-application/application-of-artificial-neural-networks-and-hybrid-methods-in-the-solution-of-inverse-problems

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen