Jader Lugon Junior
jljunior@iff.edu.br
Department of Research and Innovation
Fluminense Federal Institute of Education Science
and Technology
Macaé, Brazil

Antônio J. Silva Neto
ajsneto@iprj.uerj.br
Depart. of Mechanical Engineering and Energy
Polytechnic Institute
State University of Rio de Janeiro
Rio de Janeiro, Brazil

Solution of Porous Media Inverse Drying Problems Using a Combination of Stochastic and Deterministic Methods

In the present work the inverse problem of simultaneous heat and mass transfer modeled by Luikov equations is studied using a hybrid combination of the Levenberg-Marquardt (LM), Simulated Annealing (SA) and Artificial Neural Network (ANN) methods. The direct and inverse problems are described, formulated and solved. After the use of an experiment design technique, the hybrid combination ANN-LM-SA yielded good estimates for the heat and mass transfer problem of interest. The proper choice of the set of parameters to be estimated allowed the design of an experiment with higher sensitivity coefficients. One ANN was used to generate the initial guess for the LM, another one to approximate the gradient needed by LM, and, finally, the global minimum was searched using the SA. The experimental data considered in the inverse problem was generated using the solution for the direct problem with the addition of noise.
Keywords: inverse problem, design of experiment, drying, Luikov equations, heat and mass transfer

Introduction

The modeling and analysis of the simultaneous heat and mass transfer phenomena in porous media has been tackled by many researchers, and different formulations have been developed (Whitaker, 1977; Pavón-Melendez et al., 2002; Mwithiga and Olwal, 2005; Kanevce et al., 2002a). In this work the mathematical model used is based on Luikov's equations (Mikhailov and Özisik, 1994; Luikov and Mikhailov, 1965).

More recently, the inverse problem of simultaneous heat and mass transfer has attracted the attention of an increasing number of researchers (Kanevce et al., 2002a,b; Dantas et al., 2003; Huang and Yeh, 2002; Lugon and Silva Neto, 2004; Pedreño-Molina et al., 2005a,b,c; Bialobrzewski, 2007). The Levenberg-Marquardt (LM) method was used in Kanevce et al. (2002a), and Dantas et al. (2003). Huang and Yeh (2002) used Alifanov's Iterative Regularization method. A hybrid combination of LM and Simulated Annealing (SA) was used in Lugon and Silva Neto (2004). Artificial Neural Networks (ANN) were used in Pedreño-Molina et al. (2005a,b,c) and Bialobrzewski (2007).

In the present work we extend the results of Lugon and Silva Neto (2004) using one ANN (Haykin, 1999; Soeiro et al., 2004a,b; Silva Neto and Soeiro, 2003) to generate the initial guess for the LM, and another one to approximate the gradient needed by LM. Another improvement introduced in the present work is the use of a different choice of the parameters to be estimated, with respect to the usual set of unknowns considered in previous works, allowing the design of an experiment with higher sensitivity coefficients (Dowding et al., 1999; Beck, 1988).

Nomenclature

a = thermal diffusivity of the porous medium, m^2/s
a_m = moisture diffusivity of the porous medium, m^2/s
c = specific heat of the porous medium, J/kg K
h_m = mass transfer coefficient between the porous medium and the air, $kg/m^2\, s\, ^oM$
h = heat transfer coefficient between the porous medium and the air, $W/m^2\, K$
k_m = moisture conductivity, $kg/m\, s\, ^oM$
k = thermal conductivity, W/m K
l = thickness of the medium, m

q = thermal flux supplied at the left side of the porous medium, W/m^2
R = latent heat of evaporation, J/kg
T_0 = initial uniform temperature of the porous medium, K
T_s = temperature of the surrounding air, K
u_0 = initial moisture potential, ^oM
u^* = moisture potential equilibrium with the surrounding air, ^oM
X = coordinate axis, m

Greek Symbols

δ = thermogradient coefficient, $^oM/K$
ε = phase indicator (i.e., $\varepsilon = 1 \rightarrow$ vapor, $\varepsilon = 0 \rightarrow$ liquid), dimensionless

Direct Problem

In Fig. 1, adapted from Mwithiga and Olwal (2005), it is represented the drying experiment setup considered in the present work. It was introduced the possibility of using a scale to weigh the samples and sensors to measure temperature in the sample, as well as inside the drying chamber. In order to obtain the variation of moisture content with time, from time to time samples are extracted from the drying chamber and weighed.
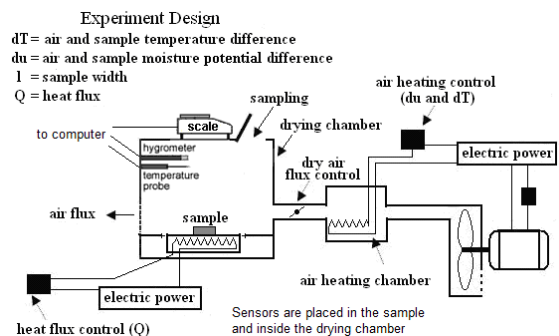

Figure 1. Drying experimental setup adapted from Mwithiga and Olwal (2005).

In accordance with the schematic representation shown in Fig. 2, consider the problem of simultaneous heat and mass transfer in a one-dimensional porous medium in which heat is supplied to the lower surface of the porous media, at the same time that dry air flows over the upper boundary surface.
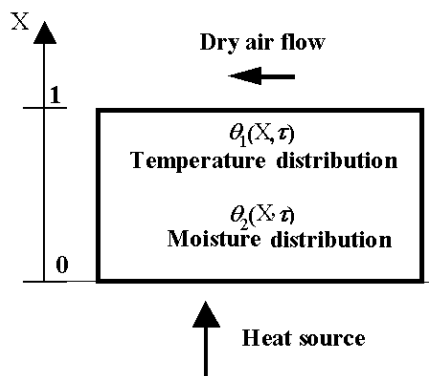
Figure 2. Drying process schematic representation.

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

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

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

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

$$\theta_1(X,0) = 0 \tag{3}$$

$$\theta_2(X,0) = 0 \tag{4}$$

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

$$\frac{\partial \theta_1(0,\tau)}{\partial X} = -Q \tag{5}$$

$$\frac{\partial \theta_2(0,\tau)}{\partial X} = -Pn\, Q \tag{6}$$

$$\frac{\partial \theta_1(1,\tau)}{\partial X} + Bi_q \theta_1(1,\tau) = Bi_q - (1-\varepsilon)KoLuBi_m[1-\theta_2(1,\tau)] = 0 \tag{7}$$

$$\frac{\partial \theta_2(1,\tau)}{\partial X} + Bi_m^* \theta_2(1,\tau) = Bi_m^* - Pn\, Bi_q[\theta_1(1,\tau)-1] \tag{8}$$

where

$$\alpha = 1 + \varepsilon\, Ko\, Lu\, Pn \tag{9}$$

$$\beta = \varepsilon\, Ko\, Lu \tag{10}$$

$$Bi_m^* = Bi_m[1 - (1-\varepsilon)Pn\, Ko\, Lu\,] \tag{11}$$

and the dimensionless variables are defined as

$$\theta_1(X,\tau) = \frac{T(x,t)-T_0}{T_s - T_0}, \ \text{temperature} \tag{12}$$

$$\theta_2(X,\tau) = \frac{u_0 - u(x,t)}{u_0 - u^*}, \ \text{moisture potential} \tag{13}$$

$$X = \frac{x}{l}, \ \text{spatial coordinate} \tag{14}$$

$$\tau = \frac{at}{l^2}, \ \text{time} \tag{15}$$

$$Lu = \frac{a_m}{a}, \ \text{Luikov number} \tag{16}$$

$$Pn = \delta \frac{T_s - T_0}{u_0 - u^*}, \ \text{Possnov number} \tag{17}$$

$$Ko = \frac{r}{c} \frac{u_0 - u^*}{T_s - T_0}, \ \text{Kossovitch number} \tag{18}$$

$$Bi_q = \frac{hl}{k}, \ \text{heat Biot} \tag{19}$$

$$Bi_m = \frac{h_m l}{k_m}, \ \text{mass Biot} \tag{20}$$

$$Q = \frac{ql}{k(T_s - T_0)}, \ \text{heat flux} \tag{21}$$

When the geometry, the initial and boundary conditions, and the medium properties are known, the system of equations (1)-(8) can be solved, yielding the temperature and moisture distribution in the media. In this work the finite difference method is used to solve such system.

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

$$\bar{u}(t) = \frac{1}{l} \int_{x=0}^{x=l} u(x,t)dx \tag{22}$$

or

$$\bar{\theta}_2(\tau) = \int_{X=0}^{X=1} \theta_2(X,\tau)dX \tag{23}$$

Therefore, in order to obtain the average moisture measurements, $\bar{u}(t)$, one has just to weigh the sample at each time of observation.

Inverse Problem Formulation

The inverse problem considered is implicitly formulated as a finite dimensional optimization problem where one seeks to minimize the objective function of squared residues (Silva Neto and Soeiro, 2003).

$$S(\mathbf{P}) = [\mathbf{G}_{calc}(\mathbf{P}) - \mathbf{G}_{meas}]^T \mathbf{W}[\mathbf{G}_{calc}(\mathbf{P}) - \mathbf{G}_{meas}] = \mathbf{F}^T \mathbf{W} \mathbf{F} \tag{24a}$$

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

$$\mathbf{F} = \mathbf{G}_{calc}(\mathbf{P}) - \mathbf{G}_{meas} \qquad (24b)$$

The inverse problem solution is the vector $\mathbf{P}^*$ which minimizes the norm given by Eq. (24a), i.e.

$$S(\mathbf{P}^*) = \min_{\mathbf{P}} S(\mathbf{P}) \qquad (25)$$

Using temperature measurements, θ_1, acquired by sensors located inside the medium, and the average of the moisture potential, $\bar{u}$, during the experiment, we try to estimate the vector of unknowns $\mathbf{P}$. In the present work a combination of unknown 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 moisture transfer coefficient and mass conductivity).

Experiment Design

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 and Silva Neto, 2004), and 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; Beck, 1988) designed for wood drying, and doing so it was considered the following process control parameters: heat flux, Q, the medium width, l, difference between the medium and the air temperatures, $dT = T_s - T_0$, and difference between the moisture potential in the medium and the air, $du = u_0 - u^*$.

The sensitivity analysis plays a major role in several aspects related to the formulation and solution of an inverse problem. Such analysis may be performed with the study of the sensitivity coefficients. Here we use the modified, or scaled sensitivity coefficients

$$SC_{P_j}(X,\tau) = P_j \frac{\partial V(X,\tau)}{\partial P_j}, \qquad j = 1, 2, \cdots, N_p \qquad (26)$$

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.

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 sets of unknowns $\{Lu, Pn, Ko, Bi_q, Bi_m, \varepsilon\}^T$ and $\{Lu, \delta, r/c, h/k, h_m/k_m, \varepsilon\}^T$,

$$SC_{\delta}(X,\tau) = \delta \frac{\partial V(X,\tau)}{\partial \delta} = Pn \frac{\partial V(X,\tau)}{\partial Pn} = SC_{Pn}(X,\tau) \qquad (27)$$

$$SC_{r/c}(X,\tau) = r/c \frac{\partial V(X,\tau)}{\partial r/c} = Ko \frac{\partial V(X,\tau)}{\partial Ko} = SC_{Ko}(X,\tau) \qquad (28)$$

$$SC_{h/k}(X,\tau) = h/k \frac{\partial V(X,\tau)}{\partial h/k} = Bi_q \frac{\partial V(X,\tau)}{\partial Bi_q} = SC_{Bi_q}(X,\tau) \qquad (29)$$

$$SC_{h_m/k_m}(X,\tau) = h_m/k_m \frac{\partial V(X,\tau)}{\partial h_m/k_m} = Bi_m \frac{\partial V(X,\tau)}{\partial Bi_m} = SC_{Bi_m}(X,\tau) \qquad (30)$$

The reason for choosing the second set of unknowns to be estimated is based on the use of experiment design tools and interpretation. Consider the heat and mass Biot numbers for example. If one changes the medium 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 avoid this situation, 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 medium width and continue with the same value for both variables to be estimated.

The same reasoning was used in the decision of estimating 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 potential between the medium and the air, $du = u_0 - u^*$, without affecting the values of the unknown parameters.

As a general guideline, the sensitivity of the state (observable) variable to the unknown parameter we want to estimate must be high enough to allow estimation 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 state variable in the same way, being difficult to distinguish their influences separately, which yields to poor estimations.

Since it was observed that the temperature sensor location did not influence significantly the sensitivity, we fixed its location at $X = 0.5$.

In Figs. 3 and 4, it is represented the sensitivity coefficients for temperature and moisture potential for a specific case in which are considered the following geometric and process parameters: $l = 0.03$ m, $Q = 6.0$, $dT = 12$ K and $du = 78$ $°M$. Since the sensitivity to the parameter ε is too low for both the temperature and moisture measurements, the estimation of this parameter is not considered in this work.
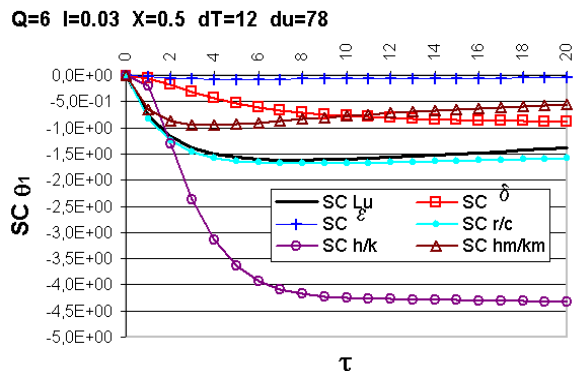

Figure 3. Sensitivity coefficients for Temperature.
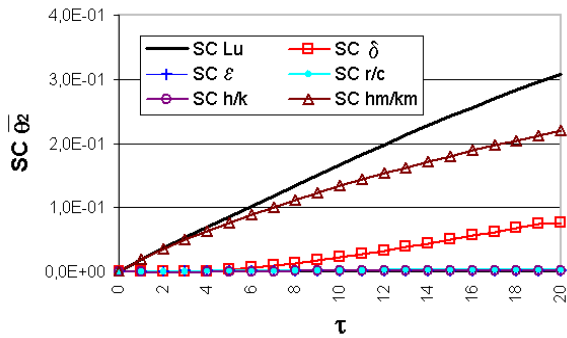
Q=6 l=0.03 dT=12 du=78

Figure 4. Sensitivity coefficients for moisture potential.

Another important tool used in this work to design the experiment consists on the study of the matrix

$$\mathbf{Y} = \begin{bmatrix} SC_{P_1 V_1} & SC_{P_2 V_1} & ... & SC_{P_{N_P} V_m} \\ SC_{P_1 V_2} & SC_{P_2 V_2} & ... & SC_{P_{N_P} V_m} \\ ... & ... & ... & ... \\ SC_{P_1 V_m} & SC_{P_2 V_m} & ... & SC_{P_{N_P} V_m} \end{bmatrix} \quad (31)$$

where V_i is a particular measurement of temperature or moisture potential and m is the total number of measurements. With m_{θ_1} data points for the temperature and $m_{\bar{\theta}_2}$ for the moisture potential, one has $m = m_{\theta_1} + m_{\bar{\theta}_2}$.

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

In Fig. 5 it 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 dry air flowing over it. 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.
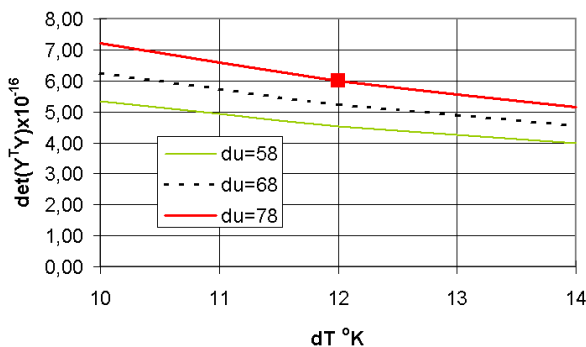

Figure 5. Matrix $\mathbf{Y}^T\mathbf{Y}$ determinant as a function of temperature (dT) and moisture potential ($\bar{du}$) differences.

In Fig. 6 it is represented the values of the determinant of matrix $\mathbf{Y}^T \mathbf{Y}$ for different values of the heat flux Q and medium thickness l. For practical reasons it was chosen to limit the sample temperature to 130°C. In Fig. 6 when the sample temperature exceeds the limit of 130°C it is used a dashed-line representation. The solid square represents the chosen designed experiment.
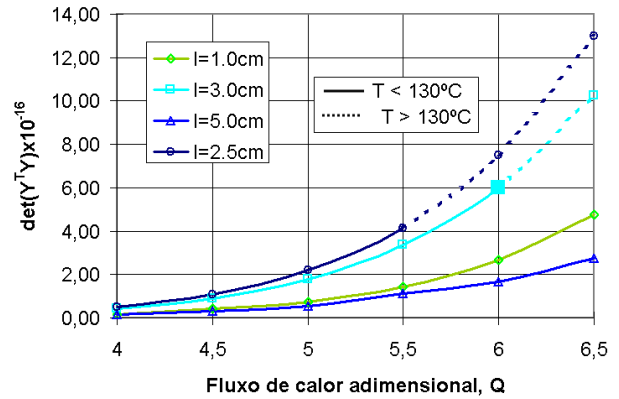

Figure 6. $\mathbf{Y}^T\mathbf{Y}$ matrix determinant for different values of the heat flux Q and medium thickness l.

Considering the previous analysis of the sensitivity graphs and of the matrix $\mathbf{Y}^T \mathbf{Y}$ determinant, it was designed the experiment whose geometric and process parameters are shown in Table 1. Since the average moisture potential, $\bar{\theta}_2$ or $\bar{u}$, is more difficult to measure than temperature, θ_1, the measurement interval for the average moisture potential, $\Delta \tau_{\bar{u}}$, was considered larger than the interval for the temperature $\Delta \tau_{\theta_1}$.

Table 1. Reference values for the designed experiment. τ_0 and τ_f represent the initial and final sampling times, respectively.

Geometric or process parameter	Values
$dT = T_s - T_0$	12 K
T_0	297 K
T_s	309 K
$du = u_0 - u^*$	78 ºM
u_0	86 ºM
u^*	8 ºM
ε	0.2
Q	6.0
l	0.03 m
τ_0	0
τ_f	20
$\Delta \tau_{\theta_1}$	0.2
m_{θ_1}	100
$\Delta \tau_{\bar{\theta}_2}$	1
$m_{\bar{\theta}_2}$	20

Inverse Problem Solution

In this work we used one Artificial Neural Network (ANN) to generate the initial guess for the Levenberg-Marquardt (LM) method, another ANN to approximate the gradient needed by LM, and, finally, the global minimum was searched using the SA for the minimization of the objective function given by Eq. (24a).

The Levenberg-Marquardt method (LM)

The Levenberg-Marquardt method is a deterministic local optimization method based on the gradient of the objective function. In order to minimize the functional S we first write

$$\frac{dS}{d\mathbf{P}} = \frac{d}{d\mathbf{P}}(\mathbf{F}^T\mathbf{WF}) = 0 \quad \longrightarrow \quad \mathbf{J}^T\mathbf{WF} = 0 \qquad (32)$$

where $\mathbf{J}$ is the Jacobian matrix, whose elements are $J_{ps} = \partial\theta_{1p}/\partial P_s$, for $p = 1,2,...m_{\theta_1}$, $J_{ps} = \partial\bar{\theta}_{2p}/\partial P_s$, for $p = m_{\theta_1}+1,...,m_{\theta_1}+m_{\bar{\theta}_2}$ and $s = 1,2,...N_p$. Observe 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:

$$\mathbf{F}(\mathbf{P}+\Delta\mathbf{P}) \cong \mathbf{F}(\mathbf{P}) + \mathbf{J}\Delta\mathbf{P} \qquad (33)$$

and introducing the above expansion in Eq. (32) results

$$\mathbf{J}^T\mathbf{WJ}\Delta\mathbf{P} = -\mathbf{J}^T\mathbf{WF}(\mathbf{P}) \qquad (34)$$

In the Levenberg-Marquardt method it is added to the diagonal of matrix $\mathbf{J}^T\mathbf{WJ}$ a damping factor λ to help to achieve convergence. The value of λ is varied along the iterative process with $\lambda \to 0$ when convergence is achieved.

Equation (34) is then written in a more convenient form to be used in the iterative procedure:

$$\Delta\mathbf{P}^n = -\left[(\mathbf{J}^n)^T\mathbf{WJ}^n + \lambda^n\mathbf{I}\right]^{-1}(\mathbf{J}^n)^T\mathbf{WF}(\mathbf{P}^n) \qquad (35)$$

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

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. (35). This iterative procedure is continued until a convergence criterion such as

$$\left|\Delta P_k^n / P_k^n\right| < \varepsilon, \ \ k = 1,2,...,N_P \qquad (36)$$

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

The elements of the Jacobian matrix as well as the right hand side term of Eq. (34) are updated at each iteration, using the solution of the direct problem with the estimates for the unknowns obtained in the previous iteration.

The Artificial Neural Network (ANN)

The multi-layer perceptron, MLP, (Haykin, 1999) is a collection of connected processing elements called nodes or neurons, arranged in layers (see Fig. 7). Signals pass into the input layer nodes, progress forward through the network hidden layers and finally emerge from the output layer. 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.

A weighted sum is performed at node 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 i to the next layer

$$y_i = f\left(\sum_j w_{ij}x_j\right) \qquad (37)$$

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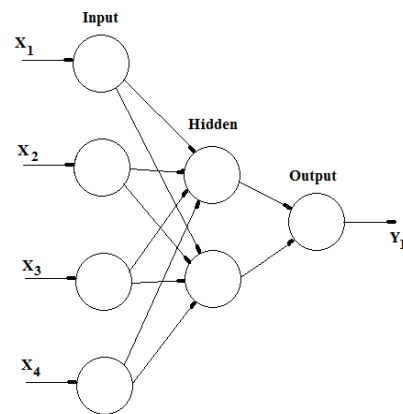

Figure 7. Multi-layer perceptron network.

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 used in the present work.

First of all, the training patterns are generated using the direct problem solution. For that purpose ten intervals were taken for each parameter, ranging from 50% up to 150% of the exact parameter values. The ANN used to solve the direct problem was trained to calculate the temperature and average moisture potential values, being informed in the input layer the following quantities: Lu, δ, r/c, h/k, h_m/k_m, τ. This ANN was also used to approximate the derivatives necessary to calculate the elements of the Jacobian matrix for LM, with respect to each parameter, using a central difference approximation.

The other ANN, dedicated to solve the inverse problem, used a set of five values of temperature measurements, m_{θ_1}, and five values of average moisture potential, $m_{\bar{\theta}_2}$, to estimate simultaneously the parameters Lu, δ, r/c, h/k, h_m/k_m. The patterns used in the training stage of the ANN for the inverse problem were also calculated using the direct problem solution. After being trained, given the five temperature measurements and five average moisture potential measurements, the ANN provided an estimate for the desired

parameters. The number of patterns generated was 3,125 for the inverse problem, and 31,250 for the direct problem.

The training of each ANN consisted of 2,000 epochs varying randomly the order of patterns presentation.

Once the comparison error is reduced to an acceptable level over the whole training set, the training phase ends and the network is established. After that, the parameters of a model (output) may then be determined using the real experimental data, which are the inputs of the established neural network developed for the problem solution. This is the generalization stage in the use of the ANN.

The 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 may be accepted according to the Boltzmann probability:

$$P(\Delta E) = \exp(-\Delta E / k_B T) \tag{38}$$

A uniformly distributed random number p is generated in the interval [0,1] 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.

Kirkpatrick et al. (1983) developed an optimization algorithm inspired in the cooling problem described above using the Metropolis criterion, the so called Simulated Annealing (SA) method. Using the objective function $S(\mathbf{P})$, given by Eq. (24a), 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. 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 the stopping criterion if they all agree within a tolerance ε, N_ε.

Combination of ANN, LM and SA optimizers

After the training stage, an ANN is able to quickly obtain an inverse problem solution. This solution is then used as an initial guess for the LM.

Due to the complexity of the design space, if convergence is achieved with a gradient based method it may in fact correspond to a local minimum. Therefore, global optimization methods are required in order to reach the global minimum, or at least its vicinity. 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., 2004a,b).

Trying to keep the best features of each method, we have combined the ANN developed for the solution of the inverse problem, the LM and the SA methods. First we used the ANN, obtaining quickly an initial guess for the LM. We then ran the LM, reaching within a few iterations a point of minimum. After that we ran the SA. If the same solution was reached, it was likely that a global minimum was found, and the iterative procedure was interrupted. If a different solution was obtained, it meant that the previous one was a local minimum. In that case we could run again the LM and SA until the global minimum was reached.

The canonical LM depends on the calculation of the gradient, which in many cases must be approximated by finite differences. In practice it means that the direct problem has to be solved many times. As previously described, in this work an ANN was trained to solve the direct problem and then it was used to approximate the gradient in the first steps of the LM. After this initial stage a refinement in the solution was implemented, with the use of finite differences for the calculation of the gradient.

Results

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

$$\theta_{1\,meas_i} = \theta_{1\,calc_i}(\vec{P}_{exact}) + \sigma_{\theta_1} r_i, \; i=1,2,...,m_{\theta_1} \tag{39a}$$

$$\bar{\theta}_{2\,meas_i} = \bar{\theta}_{2\,calc_i}(\vec{P}_{exact}) + \sigma_{\bar{\theta}_2} r_i, \; i=1,2,...,m_{\bar{\theta}_2} \tag{39b}$$

where r_i are random numbers in the range [-1,1], m_{θ_1} and $m_{\bar{\theta}_2}$ represent the total number of temperature and moisture potential experimental data, and σ_{θ_1} and $\sigma_{\bar{\theta}_2}$ emulates the standard deviation of measurement errors. It was used $\sigma_{\theta_1}=0.03$ considering 100 temperature measurements ($\Delta\tau=0.2$), resulting in a maximum error of 2%, and $\sigma_{\bar{\theta}_2}=0.001$ considering 20 moisture measurements ($\Delta\tau=1.0$), resulting in a maximum error of 4%.

In Figs. 8 and 9, it is shown the temperature (θ_1) and the moisture potential ($\bar{\theta}_2$) measurements, respectively, represented by squares, and the solid lines correspond to the direct problem solution using the parameters estimated with the inverse problem solution. In order to give a clearer representation, only 20 temperature (θ_1) measurements were represented.
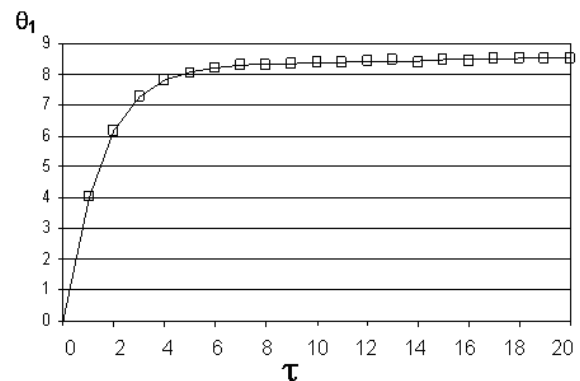

Figure 8. Temperature (θ_1) artificially simulated data.
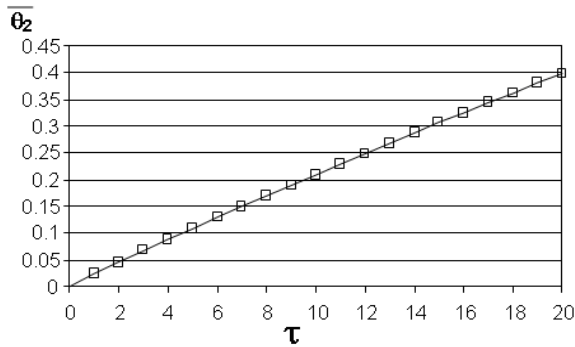
$\bar{\theta}_2$

Figure 9. Moisture potential ($\bar{\theta}_2$) artificially simulated data.

The results obtained using the LM 1 (gradient approximated by finite differences), LM 2 (gradient approximated by ANN), ANN, SA and hybrid combinations, for different levels of noise represented by different values for the standard deviation of measurements errors in temperature and average moisture potential, σ_{θ_1} and $\sigma_{\bar{\theta}_2}$, are shown in Table 2. In parenthesis are shown the percentage errors between the estimates and the exact values of the unknowns.

One observes that when there is no noise, that is, the standard deviation of measurements errors is null, the LM method was able to estimate all variables very quickly (test cases 1 and 2). When noise is introduced, the LM is retained by local minima (test cases 3 and 4). The ANN alone did not reach a good solution, but quickly got close to it (test case 5). The ANN solution was used as a first guess for the LM method with good performance in test cases 6 and 7. The SA alone reached a reasonable solution (test case 8), but it required a high computational time. Finally, the combination of all methods was able to reach a good solution (test case 9), without being retained by local minima, and without taking too much time, i.e. six times less than SA.

The time shown in the sixth column of Table 2 corresponds to the CPU time on a Pentium IV 2.8 GHz processor.

Conclusions

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 as well as 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 experiment technique is of great importance for the success of the estimation. While previous works studied the estimation of Lu, Pn, Ko, Bi_q and Bi_m, in this work it was considered Lu, δ, r/c, h/k and h_m/k_m. The main advantage 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.

The next step of this research is to study the impact of larger measurement deviation and to use real experiment data to estimate medium properties for industrial cases of interest.

References

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

Bialobrzewski, I, 2007, "Determination of the Mass Transfer Coefficient During Hot-Air-Drying Celery Root", *J. of Food Engineering*, Vol. 78, pp. 1388-1396.

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*, Vol. 46, pp. 1587-1598.

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

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*, Vol. 45, pp. 4643-4653.

Kanevce, L.P. Kanevce, G.H. and Dulikravich, G.S., 2002a, "Estimation of Drying Food Thermophysical Properties by Using Temperature Measurements", 4th International Conference on Inverse Problems in Engineering, Vol. II, pp. 299-306, Angra dos Reis, Brazil.

Kanevce, G.H. Kanavce, L.P. Dulikravich, G.S. and Orlande, H.R.B., 2002b, "Estimation of Thermophysical Properties of Moist Materials Under Different Drying Conditions", 4th International Conference on Inverse Problems in Engineering, Vol. II, pp. 331-338, Angra dos Reis, Brazil.

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

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.

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

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

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

Mwithiga, G., and Olwal, J.O., 2005, "The Drying Kinetics of Kale (Brassica Oleracea) in a Convective Hot Air Dryer", *J. of Food Engineering*, Vol. 71, No. 4, pp. 373-378.

Pavón-Melendez, G., Hernández, J.A., Salgado, M.A. and Garcia, M.A., 2002, "Dimensionless Analysis of the Simultaneous Heat and Mass Transfer in Food Drying", *J. of Food Engineering*, Vol. 51, pp. 347-353.

Pedreño-Molina, J.L., Monzó-Cabrera, J. and Sánchez-Hernández, D., 2005a, "A New Predictive Neural Architecture for Solving Temperature Inverse Problems in Microwave-assisted Drying Process", *Neurocomputing*, Vol. 64, pp. 521-528.

Pedreño-Molina, J.L., Monzó-Cabrera, J., Toledo-Moreo, A. and Sánchez-Hernández, D., 2005b, "A Novel Predictive Architeture for Microwave-assisted Drying Process Based on Neural Networks", *Int. Comm. in Heat and Mass Transfer*, Vol. 32, pp. 1026-1033.

Pedreño-Molina, J.L., Monzó-Cabrera, J., Toledo-Moreo, A. and Sánchez-Hernández, D., 2005c, "Estimation of Microwave-Assisted Drying Parameters Using Adaptative Optimization Techniques", *Int. Comm. in Heat and Mass Transfer*, Vol. 32, pp. 323-331.

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., 2004a, "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., Soares, P.O., Campos Velho, H.F. and Silva Neto, A.J., 2004b, "Using Neural Networks to Obtain Initial Estimates for the Solution of Inverse Heat Transfer Problems", Proc. Inverse Problems, Design and Optimization Symposium, Vol. I, pp. 358-363, Rio de Janeiro, Brazil.

Whitaker S., 1977, "Simultaneous Heat, Mass and Momentum Transfer in Porous Media: A Theory of Drying", *Advances in Heat and Mass Transfer*, Vol. 13, pp. 110-203.

Table 2. Results obtained using LM 1, LM 2, ANN, and hybrid combinations. σ_{θ_1} = 0.03 corresponds to a 2% noise level in temperature data. $\sigma_{\bar{\theta}_2}$ = 0.001 corresponds to a 4% noise level in moisture potential data.

Case	Method	σ_{θ_1}	$\sigma_{\bar{\theta}_2}$	S Eq. (24a)	Time (s)	Information	Lu	δ	r/c	h/k	h_m/k_m
-	-	-	-	-	-	Exact values	0.0080	2.0	10.83	34.0	114.0
1	LM 1 (grad. FDM)	0	0	0	15	Initial guess	0.0040	1.50	8.00	20.0	80.0
						Result $\mathbf{Z}_{LM^{FDM}}$	0.0080 (0%)	2.00 (0%)	10.83 (0%)	34.0 (0%)	114.0 (0 %)
2	LM 2 (grad. ANN)	0	0	0	10	Initial guess	0.0040	1.50	8.00	25.0	80.0
						Result $\mathbf{Z}_{LM^{ANN}}$	0.0080 (0%)	2.00 (0%)	10.83 (0%)	34.0 (0%)	114.0 (0%)
3	LM 1 (grad. FDM)	0.03	0.001	977	15	Initial guess	0.0040	1.50	8.00	20.0	80.0
						Result $\mathbf{Z}_{LM^{FDM}}$	0.0076 (5%)	2.09 (4.5%)	10.76 (0.6%)	34.1 (0.3%)	121.2 (6.3%)
4	LM 2 (grad. ANN)	0.03	0.001	897	11	Initial guess	0.0040	1.50	8.00	20.0	80.0
						Result $\mathbf{Z}_{LM^{ANN}}$	0.0093 (16%)	1.71 (14%)	10.73 (0.9%)	34.1 (0.3%)	95.7 (16%)
5	ANN (without initial guess)	0.03	0.001	3190	1	Result $\mathbf{Z}_{ANN}$	0.0083 (3.8%)	2.10 (5%)	10.04 (7.3%)	35.0 (2.9%)	117.1 (2.7%)
6	LM 1 (grad. FDM)	0.03	0.001	974	16	Initial guess $\mathbf{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1
						Result $\mathbf{Z}_{LM^{FDM}}$	0.0083 (3.8%)	1.92 (4%)	10.75 (0.7%)	34.1 (0.3%)	110.0 (3.5%)
7	LM 2 (grad. ANN)	0.03	0.001	903	11	Initial guess $\mathbf{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1
						Result $\mathbf{Z}_{LM^{ANN}}$	0.0082 (2.5%)	1.79 (10.5%)	9.89 (8.7%)	35.1 (3.2%)	114.5 (0.4%)
8	SA (SA 20,000 function evaluations)[1]	0.03	0.001	856	300	Initial guess	0.0040	1.50	8.00	25.0	80.0
						Result $\mathbf{Z}_{SA}$	0.0094 (17.5%)	1.58 (21%)	9.96 (8%)	35.0 (2.9%)	98.2 (14%)
9	ANN-LM 2-SA (SA 2,000 function evaluations)[1]	0.03	0.001	760	47	Initial guess $\mathbf{Z}_{ANN}$	0.0083	2.10	10.04	35.0	117.1
						Result $\mathbf{Z}_{LM^{ANN}}$	0.0082 (2.5%)	1.79 (19.5%)	9.89 (8.7%)	35.1 (3.2%)	114.5 (0.4%)
						Result $\mathbf{Z}_{SA}$	0.0079 (1.3%)	2.01 (0.5%)	11.00 (1.5%)	33.9 (0.3%)	113.8 (0.2%)
						Result $\mathbf{Z}_{LM^{ANN}}$	0.0080 (0%)	2.05 (2.5%)	10.93 (0.9%)	33.8 (0.6%)	113.9 (0.1%)

Note 1: One function evaluation corresponds to one solution of the direct problem.