# NONLINEAR PARAMETER ESTIMATION IN LAMINAR FORCED CONVECTION INSIDE A CIRCULAR SECTOR TUBE 

J. B. APARECIDO<br>Departamento de Engenharia Mecânica<br>Universidade Estadual Paulista, SP, BR

M. N. OZISIK<br>Mechanical and Aerospace Engineering<br>Department<br>North Carolina State University, NC, US


#### Abstract

In this paper we examine an inverse heat convection problem of estimating unknown parameters of a variable boundary heat flux. The function estimation is reduced to a parameter estimation problem through a parameterization in terms of some trial functions. The physical problem treated here is a hydrodynamically developed, thermally developing, three-dimensional steady state laminar flow of a fluid inside a circular sector duct, insulated in the flat walls and subject to unknown wall heat flux at the curved wall. Results are presented for polynomial and sinusoidal trial functions and the unknown surface heat fluxes are determined. Depending on the nature of the flow, on the trial functions used, on the position of experimental points, the inverse problem sometimes could not be solved. Therefore an identification condition is defined in order to specify a condition under which the inverse problem can be solved. Once the parameters have been computed is possible to obtain the statistical significance of the inverse problem solution. Therefore, approximate confidence bounds, based on standard statistical linear procedure, for the estimated parameters are analyzed and presented.


## INTRODUCTION

Inverse problems have been reported in a variety of studies to determine unknown parameters or unknown functions by using measurements of some quantities and a mathematical model [1-3]. Several numerical experiments in inverse thermal problems have been done to estimate many aspects such as inlet (or initial) condition, boundary conditions and physical properties. The reported works deal with inverse conduction [4-7], radiation [8-9] and convection [10-13] problems. The instability inherent to the inverse problems is overcome by data over-specification. In this paper we examine the estimation of unknown parameters for a variable boundary
heat flux. The parameterization of the unknown functions is done by using some trial functions such as polynomial and sinusoidal curves. Physically the problem treated here is a constant properties, three-dimensional thermally developing, hydrodynamically developed, steady state laminar flow inside a duct with cross-section shaped as a circular sector, as described in Figure 1.

## NOMENCLATURE

$\mathrm{A}_{\mathrm{c}} \quad$ circular sector cross sectional area $=\phi_{\mathrm{o}} \mathrm{b}^{2} / 2$
b circular sector radius
c fluid specific heat capacity
$\mathrm{D}_{\mathrm{h}} \quad$ hydraulic diameter $=4 \mathrm{~A}_{\mathrm{c}} / \mathrm{P}=2 \phi_{\mathrm{o}} \mathrm{b} /\left(2+\phi_{\mathrm{o}}\right)$
$\mathrm{f}(\mathrm{r}, \phi) \quad$ initial condition
$F(R, \phi) \quad$ dimensionless initial condition $=\left[f(r, \phi)-T_{0}\right] / T_{0}$
$\mathbf{g}(\mathbf{p}) \quad$ gradient of the function $\mathrm{S}(\mathbf{p})$, eq.(6c)
G Hessian matrix, eq.(8b)
$\mathrm{k} \quad$ thermal conductibility
p unknown parameter vector, eq.(5b)
$\hat{\mathbf{p}} \quad$ the least square estimator of $\mathbf{p}$
$\mathrm{P} \quad$ perimeter $=\left(2+\phi_{\mathrm{o}}\right) \mathrm{b}$
$\mathrm{q}_{\mathrm{o}}(\phi) \quad$ curved wall heat flux
$\mathrm{Q}(\phi) \quad$ dimensionless heat flux $=\mathrm{q}_{\mathrm{o}}(\phi) \mathrm{D}_{\mathrm{h}} / \mathrm{kT}_{\mathrm{o}}$
$\mathrm{u}(\mathrm{r}, \phi) \quad$ fully developed velocity distribution in the polar coordinate ( $\mathrm{r}, \phi, \mathrm{x}$ ), given by [14]
$\overline{\mathrm{u}} \quad$ average flow velocity $=\begin{gathered}\mathrm{b} \\ 0\end{gathered} \phi_{\mathrm{o}} \mathrm{u}(\mathrm{r}, \phi) \operatorname{rdrd} \phi / \mathrm{A}_{\mathrm{c}}$
$\mathrm{U}(\mathrm{R}, \phi) \quad$ dimensionless flow velocity $=\mathrm{u}(\mathrm{r}, \phi) / \overline{\mathrm{u}}$
$r$ radial coordinate
$R \quad$ dimensionless radial coordinate $=r / D_{h}$

Re $\quad$ Reynolds number $=\rho \bar{u} D_{h} / \mu$
$\mathrm{R}_{\mathrm{w}} \quad$ dimensionless radius $=\mathrm{b} / \mathrm{D}_{\mathrm{h}}=\left(2+\phi_{\mathrm{o}}\right) / 2 \phi_{\mathrm{o}}$
$\mathrm{S}(\mathbf{p}) \quad$ function of the unknown parameters $\mathbf{p}$, eq. (5a)
$\mathrm{T}(\mathrm{r}, \phi, \mathrm{x})$ temperature at ( $\mathrm{r}, \phi, \mathrm{x}$ )
$\mathrm{T}_{\mathrm{o}} \quad$ average initial temperature $=\begin{gathered}\mathrm{b} \\ 0\end{gathered} \phi_{\mathrm{o}} \mathrm{f}(\mathrm{r}, \phi) \mathrm{rdrd} \phi / \mathrm{A}_{\mathrm{c}}$ x axial coordinate
$\mathrm{X} \quad$ dimensionless axial coordinate $=x / D_{h} \operatorname{Re} \operatorname{Pr}$ $\mathbf{X} \quad$ sensitivity matrix, eq.(7a)

## Greek symbols

| $\rho$ | fluid density |
| :--- | :--- |
| $\phi$ | dimensionless angular coordinate |
| $\phi_{0}$ | circular sector opening angle $=2 \pi / 3$ |
| $\Theta(\mathrm{R}, \phi, \mathrm{X})$ | dimensionless temperature $=\left[\mathrm{T}(\mathrm{R}, \phi, \mathrm{X})-\mathrm{T}_{\mathrm{o}}\right] / \mathrm{T}_{\mathrm{o}}$ |



Figure 1 - Geometry and coordinate system

## DIRECT PROBLEM

The direct problem related to the inverse problem considered here is thermally developing, hydrodynamically developed laminar flow, inside circular sector tube of radius $b$ by azimuthal angle $\phi_{\mathrm{O}}=2 \pi / 3$, subject to prescribed heat flux at boundaries. At the plane surfaces the heat flux is zero and at the curved surface heat flux varies with the azimuthal angle $\phi$. Figure 1 shows the geometry and coordinates.

The energy equation for constant properties, thermally developing, hydrodynamically developed laminar forced convection in a circular cross-section tube is given in a dimensionless form as

$$
\begin{equation*}
\mathrm{U}(\mathrm{R}, \phi) \frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{X}}=\frac{\partial}{\mathrm{R} \partial \mathrm{R}}\left|\mathrm{R} \frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{R}}\right|+\frac{\partial^{2} \Theta(\mathrm{R}, \phi, \mathrm{X})}{\mathrm{R}^{2} \partial \phi^{2}} \tag{1a}
\end{equation*}
$$

in

$$
\begin{equation*}
0<\mathrm{R}<\mathrm{R}_{\mathrm{w}}, \quad 0<\phi<\phi_{\mathrm{o}}, \quad \mathrm{X}>0 \tag{1b}
\end{equation*}
$$

The boundary conditions are given by

$$
\begin{align*}
& \left.\frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \phi}\right|_{\phi=0}=0, \quad 0<\mathrm{R}<\mathrm{R}_{\mathrm{w}}, \quad \mathrm{X}>0 ; \\
& \left.\frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \phi}\right|_{\phi=\phi_{\mathrm{o}}}=0, \quad 0<\mathrm{R}<\mathrm{R}_{\mathrm{w}}, \quad \mathrm{X}>0 ;  \tag{2a-d}\\
& \left.\frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{R}}\right|_{\mathrm{R}=0}=\text { finite }, \quad 0 \leq \phi \leq \phi_{\mathrm{o}}, \quad \mathrm{X}>0 ; \\
& \left.\frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{R}}\right|_{\mathrm{R}=\mathrm{R}_{\mathrm{w}}}=\mathrm{Q}(\phi), \quad 0 \leq \phi \leq \phi_{\mathrm{o}}, \quad \mathrm{X}>0 .
\end{align*}
$$

The initial condition is

$$
\begin{equation*}
\Theta(\mathrm{R}, \phi, 0)=\mathrm{F}(\mathrm{R}, \phi), \quad 0 \leq \mathrm{R} \leq \mathrm{R}_{\mathrm{w}}, \quad 0 \leq \phi \leq \phi_{\mathrm{o}} \tag{3}
\end{equation*}
$$

when $\mathrm{f}(\mathrm{r}, \phi)$ is constant, like in this work, $\mathrm{F}(\mathrm{R}, \phi)$ is zero.
This direct problem was solved in each step of the inverse problem solution by using the generalized integral transform technique[36]. The three-dimensional temperature distribution $\Theta(R, \phi, X)$ for such kind of problem is given by Aparecido and Ozisik[37] as the following explicit equation

$$
\Theta(R, \phi, X)={ }_{i=1}^{\infty} \mathrm{B}_{\mathrm{i}} \cos \left(\mu_{\mathrm{i}} \phi\right){ }_{\mathrm{m}=1}^{\infty} \mathrm{B}_{\mathrm{im}}^{*} \mathrm{~J}_{\mu_{\mathrm{i}}}\left(\beta_{\mathrm{im}} \mathrm{R}\right) \overline{\widetilde{\Theta}}_{\mathrm{im}}(\mathrm{X})+\frac{\mathrm{Q}(\phi) \mathrm{R}^{2}}{2 \mathrm{R}_{\mathrm{w}}}
$$

## INVERSE PROBLEM

The inverse analysis considered here is defined as follows:
Suppose the applied wall heat flux $\mathrm{Q}(\phi)$ is not known, instead some temperature readings taken in the fluid at different locations are available. Our objective is by utilizing these measured data, to estimate the unknown heat flux $\mathrm{Q}(\phi)$.

In nonlinear parameter estimation a priori insight is necessary to consider the unknown heat flux $\mathrm{Q}(\phi)$ as a nonlinear multivariate function of $\phi$ and some unknown parameters $\mathrm{p}_{\mathrm{i}}(\mathrm{i}=1,2, \ldots, \mathrm{I})$, as follow

$$
\begin{equation*}
\mathrm{Q}(\phi) \equiv \mathrm{F}\left(\mathrm{p}_{1}, \mathrm{p}_{2}, \ldots, \mathrm{p}_{\mathrm{I}}, \phi\right) . \tag{4}
\end{equation*}
$$

In the problem considered here, the parameters $p_{i}(i=1,2, \ldots, I)$ are unknown, to be determined by inverse analysis from the knowledge of temperature measurements taken at locations in the down stream region of the flow. The analysis is now recast as a problem of optimization for finding the unknown parameters $\mathrm{p}_{\mathrm{i}}(\mathrm{i}=1,2, \ldots, \mathrm{I})$ by minimizing the function

$$
\begin{equation*}
S(\mathbf{p})={ }_{\mathrm{m}=1}^{\mathrm{M}}\left[\Theta_{\mathrm{m}}(\mathbf{p})-\mathrm{Y}_{\mathrm{m}}\right]^{2} \tag{5a}
\end{equation*}
$$

where the vector of unknown parameters $\mathbf{p}$ is define as

$$
\begin{equation*}
\mathbf{p} \equiv\left[\mathrm{p}_{1}, \mathrm{p}_{2}, \ldots, \mathrm{p}_{\mathrm{I}}\right]^{\mathrm{T}} \tag{5b}
\end{equation*}
$$

and $Y_{m}$ are the measured temperatures taken by thermocouples at the positions $(R, \phi, X)_{m}=\left(R_{m}, \phi_{m}, X_{m}\right),(m=1,2, \ldots, M) ; \Theta_{m}(\mathbf{p})$ are the temperatures at corresponding positions computed by solving the direct problem and using the estimated values of the unknown parameters.

The function $S(\mathbf{p})$ is differentiated with respect to each unknown parameter, to yield

$$
\mathrm{g}_{\mathrm{i}}(\mathbf{p}) \equiv \frac{\partial \mathrm{S}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{i}}}=2_{\mathrm{m}=1}^{\mathrm{M}} \frac{\partial \Theta_{\mathrm{m}}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{i}}}\left[\Theta_{\mathrm{m}}(\mathbf{p})-\mathrm{Y}_{\mathrm{m}}\right],(\mathrm{i}=1,2, \ldots, \mathrm{I}) .(6 \mathrm{a})
$$

Defining the gradient of the function $S(\mathbf{p})$ as

$$
\begin{equation*}
\mathbf{g}(\mathbf{p}) \equiv \mathrm{S}(\mathbf{p}) \equiv\left|\frac{\partial \mathrm{S}(\mathbf{p})}{\partial \mathrm{p}_{1}}, \frac{\partial \mathrm{~S}(\mathbf{p})}{\partial \mathrm{p}_{2}}, \ldots, \frac{\partial \mathrm{~S}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{I}}}\right|^{\mathrm{T}} \equiv \frac{\partial \mathrm{~S}(\mathbf{p})}{\partial \mathbf{p}} \tag{6b}
\end{equation*}
$$

the system of equations (6a) can be written in matrix form as

$$
\begin{equation*}
\mathbf{g}(\mathbf{p})=\frac{\partial \mathrm{S}(\mathbf{p})}{\partial \mathbf{p}}=2 \mathbf{X}^{\mathrm{T}}(\mathbf{T}-\mathbf{Y}) \tag{6c}
\end{equation*}
$$

where

$$
\begin{gather*}
\mathbf{T}=\mathbf{T}(\mathbf{p}) \equiv\left[\Theta_{1}(\mathbf{p}), \Theta_{2}(\mathbf{p}), \ldots, \Theta_{\mathrm{M}}(\mathbf{p})\right]^{\mathrm{T}},  \tag{6d}\\
\mathbf{Y}=\left[\mathrm{Y}_{1}, \mathrm{Y}_{2}, \ldots, \mathrm{Y}_{\mathrm{M}}\right]^{\mathrm{T}},  \tag{6e}\\
\mathbf{X}=\mathbf{X}(\mathbf{p}) \equiv \frac{\partial \mathbf{T}}{\partial \mathbf{p}^{\mathrm{T}}}=\left|\frac{\partial \Theta_{\mathrm{m}}}{\partial \mathrm{p}_{\mathrm{i}}} ; \quad \mathrm{m}=1,2, \ldots, \mathrm{M} ; \mathrm{i}=1,2, \ldots, \mathrm{I}\right| \tag{7a}
\end{gather*}
$$

The matrix $\mathbf{X}$, defined by equation (7a), is the so called sensitivity matrix and their elements $\mathbf{X}_{\mathrm{mi}} \quad(\mathrm{m}=1,2, \ldots, \mathrm{M}$; $\mathrm{i}=1,2, \ldots, \mathrm{I})$ are the sensitivity coefficients. This sensitivity coefficient $\mathbf{X}_{\mathrm{mi}}$ is the first derivative of the dependent variable $\Theta_{\mathrm{m}}(\mathbf{p})$ with respect the unknown parameter $p_{i}$, in other words the sensitivity matrix $\mathbf{X}$ defined by equation (7a) is the Jacobian matrix $\mathbf{J}$ of this transformation

$$
\begin{equation*}
\mathbf{J}=\mathbf{J}(\mathbf{p}) \equiv \frac{\partial \mathbf{T}}{\partial \mathbf{p}^{\mathrm{T}}}=\mathbf{X} \tag{7b}
\end{equation*}
$$

Differentiating the equation (6a) in relation to the unknown parameters gives the $\mathrm{G}_{\mathrm{ij}}(\mathbf{p})$ elements of the Hessian of $\mathrm{S}(\mathbf{p})$

$$
\begin{align*}
\mathrm{G}_{\mathrm{ij}}(\mathbf{p}) \equiv \frac{\partial^{2} \mathrm{~S}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{i}} \partial \mathrm{p}_{\mathrm{j}}} & =2_{\mathrm{m}=1}^{\mathrm{M}} \frac{\partial \Theta_{\mathrm{m}}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{i}}} \frac{\partial \Theta_{\mathrm{m}}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{j}}}+ \\
& +\left[\Theta_{\mathrm{m}}(\mathbf{p})-\mathrm{Y}_{\mathrm{m}}\right] \frac{\partial^{2} \Theta_{\mathrm{m}}(\mathbf{p})}{\partial \mathrm{p}_{\mathrm{i}} \partial \mathrm{p}_{\mathrm{j}}} \tag{8a}
\end{align*}
$$

The complete Hessian matrix $\mathbf{G}(\mathbf{p})$ of $\mathrm{S}(\mathbf{p})$ can be written as

$$
\begin{equation*}
\mathbf{G}=\mathbf{G}(\mathbf{p})=2\left(\mathbf{J}^{\mathrm{T}} \mathbf{J}+\mathbf{W}\right), \mathbf{W}=\mathbf{W}(\mathbf{p})={ }_{\mathrm{m}=1}^{\mathrm{M}}\left[\Theta_{\mathrm{m}}(\mathbf{p})-\mathrm{Y}_{\mathrm{m}}\right] \mathbf{G}_{\mathrm{m}}^{*} \tag{8b,c}
\end{equation*}
$$

where $\mathbf{G}_{\mathrm{m}}^{*}$ is the Hessian matrix of $\Theta_{\mathrm{m}}(\mathbf{p})$ given by

$$
\begin{equation*}
\mathbf{G}_{\mathrm{m}}^{*} \equiv \mathbf{G}_{\mathrm{m}}^{*}(\mathbf{p}) \equiv\left|\frac{\partial^{2} \Theta_{\mathrm{m}}}{\partial \mathrm{p}_{\mathrm{i}} \partial \mathrm{p}_{\mathrm{j}}} ; \quad \mathrm{i}, \mathrm{j}=1,2, \ldots, \mathrm{I}\right| \tag{8d}
\end{equation*}
$$

Following the same idea when defining the so called sensitivity coefficients is possible also to call the elements of $\mathbf{G}_{\mathrm{m}}^{*}$ as second order sensitivity coefficients.

## The sensitivity problem

To compute the gradient of $S(\mathbf{p})$ using equation (6b) it is necessary to calculate the sensitivity coefficient that can be obtained as follow.

Differentiating the equations (1-3) with respect to each of the parameters $p_{i}(i=1,2, \ldots, I)$ we obtain the sensitivity problem as

$$
\begin{equation*}
\mathrm{U}(\mathrm{R}, \phi) \frac{\partial \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{X}}=\frac{1}{\mathrm{R}} \frac{\partial}{\partial \mathrm{R}}\left|\mathrm{R} \frac{\partial \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{R}}\right|+\frac{1}{\mathrm{R}^{2}} \frac{\partial^{2} \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \phi^{2}} \tag{9a}
\end{equation*}
$$

where

$$
\begin{equation*}
\Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X}) \equiv \frac{\partial \Theta(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{p}_{\mathrm{i}}}, \quad(\mathrm{i}=1,2, \ldots, \mathrm{I}) \tag{9b}
\end{equation*}
$$

the boundary conditions

$$
\begin{align*}
& \left.\frac{\partial \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \phi}\right|_{\phi=0}=0, \quad 0<\mathrm{R}<\mathrm{R}_{\mathrm{w}}, \quad \mathrm{X}>0  \tag{10a,b}\\
& \left.\frac{\partial \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \phi}\right|_{\phi=\phi_{0}}=0, \quad 0<\mathrm{R}<\mathrm{R}_{\mathrm{w}}, \quad \mathrm{X}>0
\end{align*}
$$

$$
\begin{aligned}
& \left.\frac{\partial \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{R}}\right|_{\mathrm{R}=0}=0, \quad 0 \leq \phi \leq \phi_{\mathrm{o}}, \quad \mathrm{X}>0 ; \\
& \left.\frac{\partial \Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X})}{\partial \mathrm{R}}\right|_{\mathrm{R}=\mathrm{R}_{\mathrm{w}}}=\frac{\partial \mathrm{F}(\mathbf{p}, \phi)}{\partial \mathrm{p}_{\mathrm{i}}}, \quad 0 \leq \phi \leq \phi_{\mathrm{o}}, \mathrm{X}>0 ;
\end{aligned}
$$

and the initial condition

$$
\begin{equation*}
\Gamma_{\mathrm{i}}(\mathrm{R}, \phi, 0)=0, \quad 0 \leq \mathrm{R} \leq \mathrm{R}_{\mathrm{w}}, \quad 0 \leq \phi \leq \phi_{\mathrm{o}} . \tag{11}
\end{equation*}
$$

The above equations (9-11) form a decoupled second order system of linear partial differential equations defining the sensitivity problem. The unknown variables in this system are the sensitivity functions $\Gamma_{\mathrm{i}}(\mathrm{R}, \phi, \mathrm{X}),(\mathrm{i}=1,2, \ldots, \mathrm{I})$ if we consider that the $\mathrm{p}_{\mathrm{i}}$ parameters are a priori known.

After solution of the sensitivity problem we obtain the sensitivity functions for the whole domain. Actually the sensitivity coefficients are equal to the sensitivity functions evaluated at the m-positions. If we evaluate the sensitivity functions in those positions $(R, \phi, X)_{m} \equiv\left(R_{m}, \phi_{m}, X_{m}\right)$, i.e. where the experiments were done, we obtain the sensitivity coefficients

$$
\begin{equation*}
\mathbf{X}_{\mathrm{mi}}=\Gamma_{\mathrm{i}}\left(\mathrm{R}_{\mathrm{m}}, \phi_{\mathrm{m}}, \mathrm{X}_{\mathrm{m}}\right), \quad(\mathrm{m}=1,2, \ldots, \mathrm{M} ; \mathrm{i}=1,2, \ldots, \mathrm{I}) \tag{12}
\end{equation*}
$$

## MINIMIZATION METHOD

A point $\mathbf{p}^{*}$ is said to be a strong minimum of a function $\mathrm{S}(\mathbf{p})$ if a scalar $\delta>0$ exists such that $S\left(\mathbf{p}^{*}\right)<\mathrm{S}\left(\mathbf{p}^{*}+\Delta \mathbf{p}\right)$ for all $\Delta \mathbf{p}$ such that $0<\|\Delta \mathbf{p}\| \leq \delta$. It is important to have conditions to know whether a given point $\mathbf{p}$ is a minimum of $\mathrm{S}(\mathbf{p})$ or not.

If $S(\mathbf{p})$ has continuous first and second derivatives, it is feasible to approximate the function at an arbitrary point $\mathbf{p}+\Delta \mathbf{p}$ by using information about $S(\mathbf{p}), \mathbf{g}(\mathbf{p})$ and $\mathbf{G}(\mathbf{p})$. When retaining just the first two terms in the Taylor series expansion for $\mathrm{S}(\mathbf{p})$, it becomes

$$
\begin{equation*}
\mathrm{S}(\mathbf{p}+\Delta \mathbf{p})=\mathrm{S}(\mathbf{p})+\Delta \mathbf{p}^{\mathrm{T}} \mathbf{g}(\mathbf{p}) \tag{13a}
\end{equation*}
$$

Considering equation (13a), the only way to not contradict the definition for a strong minimum is setting [35]

$$
\begin{equation*}
\mathbf{g}\left(\mathbf{p}^{*}\right)=\left.\frac{\partial \mathrm{S}(\mathbf{p})}{\partial \mathbf{p}}\right|_{\mathrm{p}=\mathrm{p}^{*}}=2 \mathbf{X}^{\mathrm{T}}(\mathbf{T}-\mathbf{Y})=\mathbf{0} \tag{13b}
\end{equation*}
$$

which is the first condition. These points which obey equation (13b) are called stationary points. Other kind of points $\mathbf{p}^{*}$ also satisfy the above condition, maximum points for example. Then it is necessary to establish an extra condition for determining whether a point is a minimum instead just another stationary point.

Retaining the first three terms in the Taylor series expansion for $S(\mathbf{p})$, it gives

$$
\begin{equation*}
\mathrm{S}(\mathbf{p}+\Delta \mathbf{p})=\mathrm{S}(\mathbf{p})+\Delta \mathbf{p}^{\mathrm{T}} \mathbf{g}(\mathbf{p})+\frac{1}{2} \Delta \mathbf{p}^{\mathrm{T}} \mathbf{G}(\mathbf{p}) \Delta \mathbf{p} \tag{13c}
\end{equation*}
$$

Considering that in a stationary point $S\left(\mathbf{p}^{*}\right)<S\left(\mathbf{p}^{*}+\Delta \mathbf{p}\right)$ and $\mathbf{g}\left(\mathbf{p}^{*}\right)=\mathbf{0}$, then equation (13c) becomes

$$
\begin{equation*}
\Delta \mathbf{p}^{\mathrm{T}} \mathbf{G}\left(\mathbf{p}^{*}\right) \Delta \mathbf{p}>\mathbf{0} \tag{13d}
\end{equation*}
$$

which is the second condition, commonly called positive definiteness. First and second conditions are sufficient to ensure that $\mathbf{p}^{*}$ is a strong minimum [35]. If $\mathbf{G}\left(\mathbf{p}^{*}\right)$ is singular it is necessary to hold more terms in the Taylor series to get the extra condition. It is assumed here that $\mathbf{G}\left(\mathbf{p}^{*}\right)$ is not singular.

The unknown parameter vector $\mathbf{p}^{*}$ can now be determined from the solutions of equations (13b) subject to condition (13d), by means of several iterative methods such as: modified-Newton [15-17], quasi-Newton [15], Levenberg-Marquardt [20,21], conjugate gradient [18], and direct search [19]. About characteristics of such methods see [15,35]. In this work we used the IMSL routine DUMING [34] that is based on the quasi-Newton algorithm.

## Quasi-Newton method

Modified-Newton method requires computation of complete Hessian $\mathbf{G}^{\mathrm{k}}$ and subsequently, when $\mathbf{G}^{\mathrm{k}}$ is not positive definite, computation of a pseudo Hessian $\mathbf{B}^{k}$ of $\mathbf{G}^{k}$. In present work to calculate $\mathbf{G}^{\mathrm{k}}$ is not an easy task, it requires, at each iteration, computation of $\left(\mathrm{I}^{2}+\mathrm{I}\right) / 2$ second order partial differential equations. Following the formulae in Newton and Modified-Newton methods it is possible to write for quasi-Newton method [15]

$$
\begin{gather*}
\Delta \mathbf{p}^{\mathrm{k}}=-\left(\mathbf{B}^{\mathrm{k}}\right)^{-1} \mathbf{g}^{\mathrm{k}}  \tag{14a}\\
\mathbf{p}^{\mathrm{k}+1} \equiv \mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}, \quad \alpha^{\mathrm{k}}>0 \tag{14b}
\end{gather*}
$$

where $\mathbf{B}^{k}$ is a positive definite matrix that approximate in some way the Hessian $\mathbf{G}^{\mathbf{k}}$, and $\alpha^{\mathbf{k}}$ can be determined using linear search minimization in the direction of $\Delta \mathbf{p}^{\mathrm{k}}$, see Appendix A. The matrix $\mathbf{B}^{k}$ is updated by the application of the BFGS (Broyden-Fletcher-Goldfarb-Shanno) formula [15]

$$
\begin{equation*}
\mathbf{B}^{\mathrm{k}+1}=\mathbf{B}^{\mathrm{k}}-\frac{\mathbf{B}^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}} \mathbf{B}^{\mathrm{k}}}{\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}} \mathbf{B}^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}}+\frac{\Delta \mathbf{g}^{\mathrm{k}}\left(\Delta \mathbf{g}^{\mathrm{k}}\right)^{\mathrm{T}}}{\left(\Delta \mathbf{g}^{\mathrm{k}}\right)^{\mathrm{T}}\left(\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\right)} \tag{14c}
\end{equation*}
$$

where $\Delta \mathbf{g}^{\mathrm{k}} \equiv \mathbf{g}^{\mathbf{k + 1}}-\mathbf{g}^{\mathbf{k}}$. The initial approximation $\mathbf{B}^{0}$ is any positive definite matrix and could be the unit matrix.

## COMPUTATIONAL PROCEDURE

It is now possible to establish the steps of the computational procedure in the solution of this convective inverse problem using procedures for the quasi-Newton method as follow.

## Computational procedure using quasi-Newton method

1. Choose the initial guess for $\mathbf{p}^{0}$, set the initial Hessian approximation $\mathbf{B}^{0}$ for $\mathrm{k}=0$, for example ; $\mathbf{B}^{0}=\mathbf{I}$;
2. Calculate the temperature distribution $\Theta\left(\mathrm{R}, \phi, \mathrm{X}, \mathbf{p}^{\mathrm{k}}\right)$ through solution of the direct problem (1-3);
3. Calculate the sensitivity functions $\Gamma_{i}\left(R, \phi, X, \mathbf{p}^{k}\right)$, $(\mathrm{i}=1,2, \ldots, \mathrm{I})$ through solution of the sensitivity problem (9-11) and consequently the sensitivity coefficient $\mathbf{X}$ matrix, using the equations (12);
4. $\mathbf{g}^{\mathrm{k}}=2\left(\mathbf{J}^{\mathrm{T}}\right)^{\mathrm{k}}\left(\mathbf{T}^{\mathrm{k}}-\mathbf{Y}\right)$;
5. $\Delta \mathbf{p}^{\mathrm{k}}=-\left(\mathbf{B}^{\mathrm{k}}\right)^{-1} \mathbf{g}^{\mathrm{k}}$;
6. Compute $\alpha^{k}$ using univariate optimization in the direction of $\Delta \mathbf{p}^{\mathrm{k}}$ (see Appendix A);
7. $\mathbf{p}^{\mathrm{k}+1} \equiv \mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}$;
8. Calculate the temperature distribution $\Theta\left(R, \phi, X, \mathbf{p}^{k+1}\right)$ through solution of the direct problem (1-3);
9. Calculate the sensitivity functions $\Gamma_{i}\left(R, \phi, X, \mathbf{p}^{k+1}\right)$, ( $\mathrm{i}=1,2, \ldots, \mathrm{I}$ ) through solution of the sensitivity problem (9-11) and consequently the sensitivity coefficients $\mathbf{X}$ matrix using the equations (12);
10. $\mathbf{S}\left(\mathbf{p}^{\mathrm{k}+1}\right)=\left(\mathbf{T}^{\mathrm{k}+1}-\mathbf{Y}\right)^{\mathrm{T}}\left(\mathbf{T}^{\mathrm{k}+1}-\mathbf{Y}\right)$;
11. $\mathbf{g}^{\mathrm{k}+1}=2\left(\mathbf{J}^{\mathrm{T}}\right)^{\mathrm{k}+1}\left(\mathbf{T}^{\mathrm{k}+1}-\mathbf{Y}\right)$;
12. $\Delta \mathbf{g}^{\mathrm{k}}=\mathbf{g}^{\mathrm{k}+1}-\mathbf{g}^{\mathrm{k}}$;
13. $\mathbf{B}^{\mathrm{k}+1}=\mathbf{B}^{\mathrm{k}}-\frac{\mathbf{B}^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}} \mathbf{B}^{\mathrm{k}}}{\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}} \mathbf{B}^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}}+\frac{\Delta \mathbf{g}^{\mathrm{k}}\left(\Delta \mathbf{g}^{\mathrm{k}}\right)^{\mathrm{T}}}{\left(\Delta \mathbf{g}^{\mathrm{k}}\right)^{\mathrm{T}}\left(\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\right)}$;
14. Increment iteration, $k:=k+1$;
15. Check if the stopping criterion is achieved, if yes stop; if not return to 5 .

## STOPPING CRITERION

The function $S(\mathbf{p})$ is to be minimized, consequently the gradient of the function $S(\mathbf{p})$ must be close to zero and the iterative algorithm generates the following sequences $\mathbf{p}^{0}, \mathbf{p}^{1}, \mathbf{p}^{2}, \ldots ; \mathrm{S}\left(\mathbf{p}^{0}\right), \mathrm{S}\left(\mathbf{p}^{1}\right), \mathrm{S}\left(\mathbf{p}^{2}\right), \ldots ; \mathbf{g}^{0}, \mathbf{g}^{1}, \mathbf{g}^{2}, \ldots$ A necessary condition to achieve a solution for the inverse problem is that these sequences must be convergent. In this sense the distance between one step ( $k+1$ ) and the last ( $k$ ) must be less than a arbitrary small positive number $\varepsilon$. The concept of distance is measured using the Euclidean norm. If the experimental data contains no measurement errors and taking in consideration the aspects stated above the following stopping criterion can be defined, among others:

$$
\begin{equation*}
\mathrm{S}\left(\mathbf{p}^{\mathrm{k}}\right)<\varepsilon_{1}, \quad\left\|\mathbf{g}^{\mathrm{k}}\right\|<\varepsilon_{2} \tag{15a,b}
\end{equation*}
$$

The more traditional check condition is that specified by the equation (15a) [22]. However, the IMSL routine DUMING [34] used in this work, uses the gradient criterion, equation (15b).

However, the real temperature results contain measurement errors, then the inverse solution will converge to the experimental data and the inverse problem solution will lose a regular character more and more as the number of iterations is increased [23]. Some researches [24,12] show that is adequate to use the discrepancy principle for terminating the iterations and avoid that unsuitable behavior. The idea of application of discrepancy principle (or error principle) to the inverse problem analysis appears to have originated firstly with [25]. A similar work was published in the same year by [26]. More detailed mathematical proofs about this subject can be found in [27] and [28]. Let the standard deviation $\sigma$ of the measurement errors be nearly the same for all sensors and measurements, that is

$$
\begin{equation*}
\left|\mathrm{Y}_{\mathrm{m}}-\Theta_{\mathrm{m}}(\mathbf{p})\right| \approx \sigma, \quad(\mathrm{m}=1,2, \ldots, \mathrm{M}) \tag{15c}
\end{equation*}
$$

Introducing the equation (15c) into (5a) we have a residual amount for the function

$$
\begin{equation*}
\mathrm{S}(\mathbf{p}) \approx{ }_{\mathrm{m}=1}^{\mathrm{M}} \sigma^{2}=\mathrm{M} \sigma^{2} \tag{15d}
\end{equation*}
$$

Then the discrepancy principle applied as stopping criteria is taken as

$$
\begin{equation*}
S\left(\mathbf{p}^{k}\right)<\varepsilon_{3} \equiv M \sigma^{2} \tag{15e}
\end{equation*}
$$

## STATISTICAL ANALYSIS OF CONFIDENCE BOUNDS

Once the parameters are computed is important to obtain the statistical significance of the inverse problem solution. Approximate confidence bounds, based on linearization procedure, for the estimated parameters are here analyzed.

## Assumption on errors

The random measurement errors are generally the major source of error in estimates made by the inverse analysis. A statistical description of such random errors is useful in their analysis. Some standard assumptions regarding the temperature measurements are listed below [1,3]:

1. The errors are additive, that is

$$
\begin{equation*}
\mathrm{Y}_{\mathrm{m}}=\Theta_{\mathrm{m}}(\mathbf{p})+\varepsilon_{\mathrm{m}}, \quad \mathrm{~m}=1,2, \ldots, \mathrm{M} \tag{16a}
\end{equation*}
$$

where $Y_{m}$ is the measured temperature, $\Theta_{m}(\mathbf{p})$ is the estimated temperature and $\varepsilon_{\mathrm{m}}$ is the random error.
2. The temperature errors $\varepsilon_{i}$, have a zero mean, that is

$$
\begin{equation*}
\mathrm{E}\left(\varepsilon_{\mathrm{m}}\right)=0 \tag{16b}
\end{equation*}
$$

where E() is the expected value operator.
3. The errors have constant variance $\sigma^{2}$, that is

$$
\begin{equation*}
\operatorname{var}\left(\varepsilon_{\mathrm{m}}\right)=\mathrm{E}\left\{\left[\varepsilon_{\mathrm{m}}-\mathrm{E}\left(\varepsilon_{\mathrm{m}}\right)\right]^{2}\right\}=\mathrm{E}\left(\varepsilon_{\mathrm{m}}^{2}\right)=\sigma^{2} \tag{16c}
\end{equation*}
$$

4. The measurement errors $\varepsilon_{\mathrm{m}}$ and $\varepsilon_{\mathrm{n}}, \mathrm{m} \neq \mathrm{n}$, are uncorrelated, i.e. the covariance of $\varepsilon_{\mathrm{m}}$ and $\varepsilon_{\mathrm{n}}$ is zero

$$
\begin{align*}
& \operatorname{cov}\left(\varepsilon_{\mathrm{m}}, \varepsilon_{\mathrm{n}}\right)=\mathrm{E}\left\{\left[\varepsilon_{\mathrm{m}}-\mathrm{E}\left(\varepsilon_{\mathrm{m}}\right)\right]\left[\varepsilon_{\mathrm{n}}-\mathrm{E}\left(\varepsilon_{\mathrm{n}}\right)\right]\right\}=  \tag{16~d}\\
&=\mathrm{E}\left(\varepsilon_{\mathrm{m}} \varepsilon_{\mathrm{n}}\right)=0, \text { for } \mathrm{m} \neq \mathrm{n}
\end{align*} .
$$

5. The statistical parameters such as $\sigma^{2}$ describing the variance of $\varepsilon_{i}$ are known.
6. The measurement positions and the thermal properties are all accurately known.

## The standard statistical linear model

We assumed above that the elements of $\varepsilon=\left(\varepsilon_{1}, \varepsilon_{2}, \ldots, \varepsilon_{M}\right)^{T}$ were independently normally distributed with mean zero

$$
\begin{equation*}
\mathrm{E}(\varepsilon)=\mathrm{E}\left[\varepsilon_{1}, \varepsilon_{2}, \ldots, \varepsilon_{\mathrm{M}}\right]^{\mathrm{T}}=\left[\mathrm{E}\left(\varepsilon_{1}\right), \mathrm{E}\left(\varepsilon_{2}\right), \ldots, \mathrm{E}\left(\varepsilon_{\mathrm{M}}\right)\right]^{\mathrm{T}}=\mathbf{0} \tag{17a}
\end{equation*}
$$

and the variance

$$
\mathrm{E}\left(\varepsilon \varepsilon^{\mathrm{T}}\right)=\mathrm{E}\left|\begin{array}{cccc}
\varepsilon_{1}^{2} & \varepsilon_{1} \varepsilon_{2} & \cdots & \varepsilon_{1} \varepsilon_{\mathrm{M}}  \tag{17b}\\
\varepsilon_{2} \varepsilon_{1} & \varepsilon_{2}^{2} & \cdots & \varepsilon_{2} \varepsilon_{\mathrm{M}} \\
\vdots & \vdots & \ddots & \vdots \\
\varepsilon_{\mathrm{M}} \varepsilon_{1} & \varepsilon_{\mathrm{M}} \varepsilon_{2} & \cdots & \varepsilon_{\mathrm{M}}^{2}
\end{array}\right|=\sigma^{2} \mathbf{I} .
$$

As stated in assumption (1) the errors are additive and are expressed in vector form as

$$
\begin{equation*}
\mathbf{Y}=\mathbf{T}+\varepsilon . \tag{18a}
\end{equation*}
$$

The estimated temperature vector $\mathbf{T}(\mathbf{p})$ can be linearized by using the Taylor expansion

$$
\begin{equation*}
\mathbf{T}(\mathbf{p})=\mathbf{T}\left(\mathbf{p}^{\mathrm{o}}\right)+\frac{\partial \mathbf{T}\left(\mathbf{p}^{\mathrm{o}}\right)}{\partial \mathbf{p}^{\mathrm{T}}}\left(\mathbf{p}-\mathbf{p}^{\mathrm{o}}\right) \tag{18b}
\end{equation*}
$$

If we choose $\quad \mathbf{p}^{\mathbf{o}}=\mathbf{0}, \quad \mathbf{T}\left(\mathbf{p}^{\mathbf{o}}\right)=\mathbf{0}, \quad$ and $\partial \mathbf{T}\left(\mathbf{p}^{\mathrm{o}}\right) / \partial \mathbf{p}^{\mathrm{T}}=\partial \mathbf{T}(\mathbf{p}) / \partial \mathbf{p}^{\mathrm{T}}=\mathbf{X}=$ constant matrix, the equation (18b) reduces to

$$
\begin{equation*}
\mathbf{T}(\mathbf{p})=\frac{\partial \mathbf{T}(\mathbf{p})}{\partial \mathbf{p}^{T}} \mathbf{p}=\mathbf{X} \mathbf{p} \tag{18c}
\end{equation*}
$$

Combining the previous statistical assumptions with the linearized equation (18c), we have the standard linear statistical model [29]

$$
\begin{equation*}
\mathbf{Y}=\mathbf{X} \mathbf{p}+\varepsilon, \quad \mathrm{E}(\varepsilon)=\mathbf{0}, \quad \mathrm{E}\left(\varepsilon \varepsilon^{\mathrm{T}}\right)=\sigma^{2} \mathbf{I} \tag{18d-f}
\end{equation*}
$$

where the elements of $\varepsilon$ may be independently normally distributed but are not necessarily so distributed.

Let $\hat{p}_{i}$, for $i=1,2, \ldots, I$, be an estimator of $p_{i}$ with Expectation

$$
\begin{equation*}
\mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{i}}\right), \tag{19a}
\end{equation*}
$$

deviance

$$
\begin{equation*}
\hat{\delta}_{i}=\hat{p}_{i}-\mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{i}}\right) \tag{19b}
\end{equation*}
$$

variance

$$
\begin{equation*}
\operatorname{var}\left(\hat{\mathrm{p}}_{\mathrm{i}}\right)=\mathrm{E}\left(\hat{\delta}_{\mathrm{i}}^{2}\right)=\mathrm{E}\left\{\left[\hat{\mathrm{p}}_{\mathrm{i}}-\mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{i}}\right)\right]^{2}\right\} \tag{19c}
\end{equation*}
$$

and covariance

$$
\begin{equation*}
\operatorname{cov}\left(\hat{\mathrm{p}}_{\mathrm{i}}, \hat{\mathrm{p}}_{\mathrm{j}}\right)=\mathrm{E}\left(\hat{\delta}_{\mathrm{i}} \hat{\delta}_{\mathrm{j}}\right)=\mathrm{E}\left\{\left[\hat{\mathrm{p}}_{\mathrm{i}}-\mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{i}}\right)\right]\left[\hat{\mathrm{p}}_{\mathrm{j}}-\mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{j}}\right)\right]\right\} \tag{19d}
\end{equation*}
$$

The concepts of expectation, deviance, variance and covariance can be expressed in a matrix for as:
Expectation

$$
\begin{equation*}
\mathrm{E}(\hat{\mathbf{p}})=\mathrm{E}\left[\hat{\mathrm{p}}_{1}, \hat{\mathrm{p}}_{2}, \ldots, \hat{\mathrm{p}}_{\mathrm{I}}\right]^{\mathrm{T}}=\left[\mathrm{E}\left(\hat{\mathrm{p}}_{1}\right), \mathrm{E}\left(\hat{\mathrm{p}}_{2}\right), \ldots, \mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{I}}\right)\right]^{\mathrm{T}} \tag{19e}
\end{equation*}
$$

deviance

$$
\begin{equation*}
\hat{\delta}=\hat{\mathbf{p}}-\mathrm{E}(\hat{\mathbf{p}}) \tag{19f}
\end{equation*}
$$

and variance-covariance

$$
\begin{align*}
& \operatorname{var}-\operatorname{cov}(\hat{\mathbf{p}})=\mathrm{E}\left(\hat{\delta} \hat{\delta}^{\mathrm{T}}\right)= \\
& \quad=\left|\begin{array}{cccc}
\operatorname{var}\left(\hat{\mathrm{p}}_{1}\right) & \operatorname{cov}\left(\hat{\mathrm{p}}_{1}, \hat{\mathrm{p}}_{2}\right) & \cdots & \operatorname{cov}\left(\hat{\mathrm{p}}_{1}, \hat{\mathrm{p}}_{\mathrm{I}}\right) \\
\operatorname{cov}\left(\hat{\mathrm{p}}_{2}, \hat{\mathrm{p}}_{1}\right) & \operatorname{var}\left(\hat{\mathrm{p}}_{2}\right) & \cdots & \operatorname{cov}\left(\hat{\mathrm{p}}_{2}, \hat{\mathrm{p}}_{\mathrm{I}}\right) \\
\vdots & \vdots & \ddots & \vdots \\
\operatorname{cov}\left(\hat{\mathrm{p}}_{\mathrm{I}}, \hat{\mathrm{p}}_{1}\right) & \operatorname{cov}\left(\hat{\mathrm{p}}_{\mathrm{I}}, \hat{\mathrm{p}}_{2}\right) & \cdots & \operatorname{var}\left(\hat{\mathrm{p}}_{\mathrm{I}}\right)
\end{array}\right| \tag{19~g}
\end{align*}
$$

The least square estimator $\hat{\mathbf{p}}$ of $\mathbf{p}$ in the standard statistical linear model is given by

$$
\begin{equation*}
\hat{\mathbf{p}}=\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \mathbf{Y} \tag{20a}
\end{equation*}
$$

Substituting $\mathbf{Y}$, defined by equation (18d), in equation (20a) we have

$$
\begin{equation*}
\hat{\mathbf{p}}=\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}}(\mathbf{X p}+\varepsilon)=\mathbf{p}+\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \varepsilon \tag{20b}
\end{equation*}
$$

As in this model $\mathbf{X}$ and $\mathbf{p}$ are not functions of $\varepsilon$, then the expectation of $\hat{\mathbf{p}}$ is expressed as

$$
\mathrm{E}(\hat{\mathbf{p}})=\mathrm{E}(\mathbf{p})+\mathrm{E}\left[\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \varepsilon\right]=\mathbf{p}+\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \mathrm{E}(\varepsilon)=\mathbf{p} .(20 \mathrm{c})
$$

Further, the bias of $\hat{\mathbf{p}}$ is obtained by

$$
\begin{equation*}
\operatorname{bias}(\hat{\mathbf{p}})=\mathrm{E}(\hat{\mathbf{p}})-\mathbf{p}=\mathbf{0} \tag{20d}
\end{equation*}
$$

then $\hat{\mathbf{p}}$ is an unbiased estimator of $\mathbf{p}$. Similarly for the deviance we obtain

$$
\begin{equation*}
\hat{\delta}=\hat{\mathbf{p}}-\mathrm{E}(\hat{\mathbf{p}})=\hat{\mathbf{p}}-\mathbf{p}=\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \varepsilon \text { and } \mathrm{E}(\hat{\delta})=\mathbf{0} \tag{20e,f}
\end{equation*}
$$

and $E\left(\varepsilon \varepsilon^{T}\right)=\sigma^{2} \mathbf{I}$, hence the variance-covariance of $\hat{\mathbf{p}}$ is given by

$$
\begin{array}{r}
\operatorname{var}-\operatorname{cov}(\hat{\mathbf{p}})=\mathrm{E}\left(\hat{\delta} \hat{\delta}^{\mathrm{T}}\right)=\mathrm{E}\left\{\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \varepsilon\left[\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \varepsilon\right]^{\mathrm{T}}\right\}= \\
=\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} \mathbf{X}^{\mathrm{T}} \sigma^{2} \mathbf{I} \mathbf{X}\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1}=\sigma^{2}\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1} . \tag{20~g}
\end{array}
$$

## Confidence bounds

Comparing equation $(19 \mathrm{~g})$ and equation $(20 \mathrm{~g})$ for the $\operatorname{var}-\operatorname{cov}(\hat{\mathbf{p}})$ we conclude that the ii-th element represents the variance of $\hat{\mathrm{p}}_{\mathrm{i}}$ and the ij -th element represents the covariance of $\hat{p}_{i}$ and $\hat{p}_{j}$, for $i \neq j$. Therefore the variance $\sigma_{\hat{\mathbf{p}}}^{2}$ of the estimated parameter vector $\hat{\mathbf{p}}$ is given by

$$
\begin{equation*}
\sigma_{\hat{\mathbf{p}}}^{2}=\sigma^{2} \operatorname{diag}\left\{\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)^{-1}\right\} \tag{20h}
\end{equation*}
$$

Assuming now a normal distribution for the measurement errors, since the exact distribution is unknown, and the $99 \%$ confidence bounds for the computed parameters $\hat{\mathrm{p}}_{\mathrm{i}},(\mathrm{i}=1,2, \ldots, \mathrm{I})$ are expressed as [30]

$$
\begin{equation*}
\text { Probability of }\left\{\hat{\mathrm{p}}_{\mathrm{i}}-2.576 \sigma_{\hat{\mathrm{p}}_{\mathrm{i}}}<\mathrm{E}\left(\hat{\mathrm{p}}_{\mathrm{i}}\right)<\hat{\mathrm{p}}_{\mathrm{i}}+2.576 \sigma_{\hat{\mathrm{p}}_{\mathrm{i}}}\right\}=0.99 \tag{21a}
\end{equation*}
$$

where the number 2.576 arises from the fact that 0.99 (or $99 \%$ ) of a normal probability distribution lie in between $\pm 2.576$ of that distribution. Considering negligible the deterministic error, i.e. $E(\hat{\mathbf{p}})=\mathrm{p}$, between the mean estimated parameter $\hat{\mathrm{p}}_{\mathrm{i}}$ and the true value of $\mathrm{p}_{\mathrm{i}}$, the equation (21a) becomes

Probability of $\left\{\hat{\mathrm{p}}_{\mathrm{i}}-2.576 \sigma_{\hat{\mathrm{p}}_{\mathrm{i}}}<\mathrm{p}_{\mathrm{i}}<\hat{\mathrm{p}}_{\mathrm{i}}+2.576 \sigma_{\hat{\mathrm{p}}_{\mathrm{i}}}\right\}=0.99$.

The approximate statistical confidence bounds for the estimated parameters $p_{i}$ are given by equation (21b). Similar analysis were used by Flach and Ozisik[31], Huang and Ozisik[32] to determine the confidence bounds for thermal
conductivity and heat capacity, and by Ho and Ozisik[33] to compute the confidence bounds for single scattering albedo and optical thickness.

## IDENTIFICATION CONDITION

Depending on the nature of the flow, on the trial functions used, on the position and the number of experimental points, the inverse problem sometimes could not be solved. Therefore, it is necessary to specify some conditions under which the inverse problem can be solved. In equation (13b,d) was settled the conditions for the existence of a minimum of a multivariate function. When using the Newton method of minimization (not shown here) a key step is defined by inversion of the Hessian matrix $\mathbf{G}^{\mathrm{k}}$. It is clear that a condition to have a solution for the inverse problem is to set

$$
\begin{equation*}
\operatorname{det}\left(\mathbf{G}^{\mathrm{k}}\right) \neq 0 \tag{22a}
\end{equation*}
$$

The equation (22a) establishes the so called identification condition. When dealing with minimization methods other than the Newton method the identification condition takes specific forms. In some methods the Hessian matrix $\mathbf{G}$ can be approximated in a form expressed by setting $\mathbf{W}=\mathbf{0}$ in equation (8b), in this case the identification condition becomes

$$
\begin{equation*}
\operatorname{det}\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right) \neq 0 \tag{22b}
\end{equation*}
$$

Further, when we inspect the equation (20h) for the confidence bound of the estimated parameters, evidently the determinant of $\left(\mathbf{X}^{\mathrm{T}} \mathbf{X}\right)$ must be different from zero but also be as big as possible to provide smaller confidence bounds.

## RESULTS AND DISCUSSION

An inverse analysis is done for estimating the unknown azimuthally varying wall heat flux $\mathrm{Q}(\phi)$. The inverse problems, in a broad classification, are of two kinds: parameter estimation and function estimation. Here we treat the case of parameter estimation, for both linear and non-linear problems. We show results for three cases, one linear and two non-linear. Also we present results for the inverse problem identification factor and for the parameter confidence bound as well.

We present the inverse problem results for the three cases:
Case A- $\mathrm{Q}(\phi)=\mathrm{a}_{\mathrm{o}}+\mathrm{a}_{1}\left(\frac{\phi}{\phi_{\mathrm{o}}}\right)+\mathrm{a}_{2}\left(\frac{\phi}{\phi_{\mathrm{o}}}\right)^{2}$,
Case B $-\mathrm{Q}(\phi)=\mathrm{a}_{\mathrm{o}}+\mathrm{a}_{1} \sin \left(\mathrm{a}_{2} \frac{\phi}{\phi_{o}}+\mathrm{a}_{3}\right)$,
(23a-c)
Case C $-\mathrm{Q}(\phi)=\mathrm{a}_{\mathrm{o}} \sin \left(\mathrm{a}_{1} \frac{\phi}{\phi_{\mathrm{o}}}\right)+\mathrm{a}_{2} \sin \left(\mathrm{a}_{3} \frac{\phi}{\phi_{\mathrm{o}}}\right)$,

The Case A represents a linear parameter estimation problem, while the Cases B and C are non-linear problems.

We used the identification factor to locate the measurement sensors. As discussed previously a good place to put the sensor is where the identification factor is different of zero and is as big as possible. Figure 2 shows the identification factor calculated along the axis X , with varying number of experimental points, for the linear case. The sensors are placed along the heated surface of a cross-section, with an uniform distance among the experimental points. The number of sensors in each measurement set is $M=5,10,15$ and 20. In linear cases the sensitivity matrix does not depend of the unknown parameters and so the identification factor.


Figure 2 - Identification factor for the linear inverse problem varying along the axis X with $\mathrm{M}=5,10,15$ and 20.

Inspecting Figure 2 we can see that the identification factor decreases towards zero as the measurement positions approaches the inlet $(\mathrm{X} \rightarrow 0)$, or the number of experimental points decreases. Therefore, according to Figure 2 and the identification condition a good place to put the measurement sensors are as far as possible from flow entrance and with as much as possible experimental points. In Figure 3 we keep the number of experimental points constant, equal to 20, and show the behavior of the identification factor related to the axis X and the radial positions $\mathrm{R} / \mathrm{R}_{\mathrm{w}}=1,0.9,0.8$, and 0.7 .

Analyzing Figure 3 we see that the identification factor decreases as the experimental positions go towards the r-coordinate origin, so a good place to put the sensors is in the heated surface or as near as possible of it.

Combining the information in Figures 2 and 3 we concluded, in according with the identification factor values, that a good place to settle the experimental sensors is at the cross-section heated surface (or so near of it) and as distant as possible from the flow entrance. Physically this conclusion can be understood in the following manner: It comes to the flow two types of information, that one from the inlet and other one from the heated surface.


Figure 3 - Identification factor for the linear inverse problem varying along the axis X with radial positions $\mathrm{R} / \mathrm{R}_{\mathrm{w}}=1.0,0.9,0.8$ and 0.7 , and $\mathrm{M}=20$.

In a cross-section near the inlet the information coming from the entrance is proportionally greater than that one coming from the heated surface. Otherwise in a cross-section distant from the inlet the information coming from the heated surface is proportionally greater than that coming from the entrance. As we are interested in recovering as much as possible of information about the unknown surface heat flux we should place the experimental sensors as near as possible to the surface and as far as possible from the entrance. Also the number of experimental points should be as big as possible. For the nonlinear cases it is more difficult to determine good places to put the sensors since in these cases the sensitivity matrix, and consequently the identification factor, depend of the unknown parameters. Is necessary to inspect the identification factor values during all the iterative process of solving the non-linear inverse problems in order to identify if the identification condition is being obeyed, otherwise the solution could not be achieved.

The simulated experimental temperature data, $\mathrm{Y}_{\mathrm{m}}$ $(\mathrm{m}=1,2, \ldots \mathrm{M})$, are generated by adding a random error term, $\omega \sigma$, to the computed exact temperature, $\Theta_{\text {exact }}$, obtained from the solution of the direct problem (1-3) as

$$
\begin{equation*}
\mathrm{Y}_{\mathrm{m}}=\Theta_{\mathrm{exact}}+\omega \sigma \tag{24a}
\end{equation*}
$$

where $\sigma$ is the standard deviation of the measurement errors and $\omega$ is the normally distributed random variable, calculated by the IMSL routine RNOPT[34]. The values of $\omega$ are generated by a pseudo-random algorithm and then selected to lie in the range $-2.576<\omega<2.576$, which represent a confidence bound of 0.99 (or $99 \%$ ) for the simulated measurement errors.

Now we analyze some results obtained for the inverse problem. All experimental sensors are placed in the wall of a cross-section located at $X=0.5$. The experimental sensors are uniformly distributed over the entire opening angle $\phi_{0}$.

Figures 4-6 show results for the unknown surface heat flux $Q(\phi)$, in which the corresponding expression for the exact surface heat flux $\mathrm{Q}(\phi)_{\text {exact }}$ is given by

$$
\begin{equation*}
\mathrm{Q}(\phi)_{\text {exact }}=0+6\left(\frac{\phi}{\phi_{\mathrm{o}}}\right)-6\left(\frac{\phi}{\phi_{\mathrm{o}}}\right)^{2} \tag{25a}
\end{equation*}
$$

the used standard deviation values are $\sigma=0.01,0.05$ and 0.1 , and the number of experimental sensors considered is $\mathrm{M}=5,10$, and 15 . Equation (25a) corresponds to a specific linear Case A. Figures 4-6 keep, in each one, a constant value for the number of experimental sensors $M$, and vary the standard deviation $\sigma$. Figure 4 shows the influence of the standard deviation on the estimated surface heat flux $\mathrm{Q}(\phi)$ for the number of experimental sensors, $M=5$. As the standard deviation decreases the results for the estimated heat flux $\mathrm{Q}(\phi)$ becomes more close to the exact distribution. In Figures 5 and 6 are shown similar results for the number of sensors $M=10$ and $\mathrm{M}=15$. Figure 6 shows that for $\mathrm{M}=15$, even with $\sigma=0.1$, is possible to achieve a good estimation of the heat flux $\mathrm{Q}(\phi)$. Comparing Figures $4-6$ we conclude that as the standard deviation decreases and the number of experimental sensors increases the estimated heat flux becomes more and more close to the exact distribution. This means that to increase the accuracy of the estimated heat flux is necessary to increase the number of experimental points and have more accurate temperature experimental data, i.e. decreasing the standard deviation $\sigma$ of the measurement errors.


Figure 4 - Estimated surface heat flux, Case A, for different values of standard deviation $\sigma=0.01,0.05$ and 0.1 and the number of experimental sensors, $M=5$.

For $\sigma=0.01$ the estimation is so close to the exact distribution. Increasing the number of experimental sensors to $\mathrm{M}=10$ and 15 , the estimation for $\sigma=0.05$ and 0.1 , becomes less dispersed when compared with the exact distribution, as shown in Figures 5 and 6.

Figures 7 shows results for the non-linear Case B for which the corresponding exact surface heat flux distribution $Q(\phi)_{\text {exact }}$ is expressed by

$$
\begin{equation*}
\mathrm{Q}(\phi)_{\text {exact }}=\frac{\pi}{2+\pi}+\frac{\pi \sqrt{2}}{2+\pi} \sin \left(\pi \frac{\phi}{\phi_{\mathrm{o}}}-\frac{\pi}{4}\right), \tag{25b}
\end{equation*}
$$

using a fixed number of experimental sensors, $\mathrm{M}=15$.


Figure 5 - Estimated surface heat flux, Case A, for different values of standard deviation $\sigma=0.01,0.05$ and 0.1 and the number of experimental sensors, $M=10$.


Figure 6 - Estimated surface heat flux, Case A, for different values of standard deviation $\sigma=0.01,0.05$ and 0.1 and the number of experimental sensors, $\mathrm{M}=15$.

Figures 7 presents results for the estimated surface heat flux using the number of experimental data $M=15$, for each value of the standard deviation $\sigma$. Again the estimated results are more accurate as the value of the deviation $\sigma$ decreases. In Figures 7, the estimated results for $\sigma=0.01$ are very close to the exact distribution, but the results for $\sigma=0.05$ and 0.1 are less accurate.

Figure 8 shows estimation results for the non-linear case $C$ which the related exact surface heat flux $\mathrm{Q}(\phi)_{\text {exact }}$ is

$$
\begin{equation*}
\mathrm{Q}(\phi)_{\text {exact }}=\frac{\pi}{2} \sin \left(\pi \frac{\phi}{\phi_{\mathrm{o}}}\right)-\frac{\pi}{2} \sin \left(2 \pi \frac{\phi}{\phi_{\mathrm{o}}}\right), \tag{25c}
\end{equation*}
$$

and the number of experimental points is $\mathrm{M}=15$.


Figure 7 - Estimated surface heat flux, Case B, for different values of standard deviation $\sigma=0.01,0.05$ and 0.1 and the number of experimental sensors, $\mathrm{M}=15$.

Here also the better results for the estimated surface heat flux $\mathrm{Q}(\phi)$ are given by using the more accurate temperature experimental data, i.e. for that ones showing a standard deviation $\sigma=0.01$. As the deviation increases the estimates become worse. This behavior could be studied through the concept of the confidence bounds of the estimated parameters defined by equation (21b).


Figure 8 - Estimated surface heat flux, Case C, for different values of standard deviation $\sigma=0.01,0.05$ and 0.1 and the number of experimental sensors, $\mathrm{M}=15$.

For the previous case Figures 9 and 10 present the exact values of the parameters $a_{0}, a_{1}, a_{2}$ and $a_{3}$ as horizontal solid lines, the corresponding estimated parameters as dash lines, and the respective confidence bounds as vertical bars. The
confidence bounds computations are performed using ten runs of temperature experimental data sets. Figure 9 corresponds to standard deviation $\sigma=0.01$ and Figure 10 to $\sigma=0.05$. In both Figures 9 and 10 the estimated parameters oscillate around the exact value parameter value. For small values of the standard deviation, say $\sigma=0.01$, Figure 9 , the confidence bounds are closer than for $\sigma=0.05$, Figure 10 where confidence bounds are wider.


Figure 9 - Estimated confidence bounds, Case C, for the standard deviation $\sigma=0.01$, the number of experimental sensors, $M=15$, and using ten runs.


Figure 10 - Estimated confidence bounds, Case C, for the standard deviation $\sigma=0.05$, the number of experimental sensors, $M=15$, and using ten runs.

Inspecting Figures 9 and 10 it seems that the linearized statistical model (18d-f) can provide some useful insights about the behavior of the estimated confidence bounds, through the use of the equation (21b).

## CONCLUSION

The heat flux estimation through the solution of an inverse problem for laminar forced convection inside a circular sector
tube was successfully done for three cases. Based on the values for the identification factor we concluded that the best points to locate the thermocouples is as far as possible from the inlet and as near as possible from the heated wall. As the number of experimental points increases more reliable are the results. When the standard deviation of the measured temperatures is small the results are better than when it is big. The linear statistical model gives useful confidence bounds for the estimated parameters.

## ACKNOWLEDGMENTS

The first author wishes to acknowledge the financial support provided by CAPES and UNESP/Ilha Solteira Campus, both from BRASIL, during his stay at NCSU/Raleigh/USA. This work was supported in part by the National Science Foundation Grant Number INT. 9416874 and by FAPESP Grant Number 99/02383-5.

## REFERENCES

[1] Beck, V. J., Blackwell, B. and St. Clair, C. A. (1985), Inverse Heat Conduction, Wiley, New York.
[2] Hensel, E. (1991), Inverse Theory and Application for Engineers, Prentice Hall, Englewood Cliffs, N.J.
[3] Beck, J.V. and Arnold, K. J. (1977), Parameter Estimation in Engineering and Science, Wiley, New York.
[4] Truffart, B., Jarny, Y. and Delaney, D. (1993), A general Optimization Algorithm to Solve 2-D Boundary Inverse Heat Conduction Problems Using Finite Elements, Inverse Problem in Engineering: Theory and Practice, ASME, New York, pp.53-60.
[5] Orlande, H.R.B. and Ozisik, M.N. (1993), Inverse Problem of Estimating Interface Conductance Between Periodically Contacting Surfaces, J. of Thermophysics and Heat Transfer, vol.7, n.2, pp.319-325.
[6] Silva Neto, A.J. and Ozisik, M.N. (1992), Two-dimensional Inverse Heat Conduction Problem of Estimating Time-varying Strength of a Line Heat Source, J. Applied Physics, vol.70, pp.5357-5362.
[7] Silva Neto, A.J. and Ozisik, M.N. (1994), The Estimation of Space and Time Dependent Strength of a Volumetric Heat Source in a One Dimensional Plate, Int. J. Heat Mass Transfer, vol.37, n.6, pp.909-915.
[8] Li, H.Y. and Ozisik, M.N. (1992), Estimation of Radiation Source Term with a Conjugate Gradient Method of Inverse Analysis, J. Quart. Spectrosc. Radiat. Transfer, vol.48, n.3, pp.237-244.
[9] Silva Neto, A.J. and Ozisik, M.N. (1992), An Inverse Analysis of Simultaneously Estimating Phase Function, Albedo and Optical Thickness, HTD-vol.203, Development in Radiative Transfer, ASME.
[10 Huang, C.H. and Ozisik, M.N. (1992), Inverse Problem of Determining Unknown Wall Heat Flux in Laminar Flow
through a Parallel Plate Duct, Num. Heat Transfer, Pt A, 21, 101-106.
[11 Raghumath, R. (1993), Determining Entrance Conditions from Downstream Measurements, Int. Comm. Heat Transfer, 20, 173-183.
[12] Moutsoglou, A. (1989), An Inverse Convection Problem, J. Heat Transfer, 111, 37-43.
[13] Bokar, J.C. and Ozisik, M.N. (1995), An Inverse Analysis for Estimating the Time-Varying Inlet Temperature in Laminar Flow Inside a Parallel Plate Duct, Int. J. Heat and Mass Transfer, vol.38, n.1, 39-45.
[14] Eckert, E. R. G. and Irvine Jr., T. F., (1956), Flow in Corners of passages with noncircular cross sections, Trans. ASME 78, 709-718.
[15] Dennis, J. E., Jr., and Schnabel, R. B. (1983), Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice Hall, Englewood Cliffs, New Jersey.
[16] Gay, D. M. (1981), Computing optimal locally constrained steps, SIAM Journal on Scientific and Statistical Computing, 2, 186-197.
[17] Gay, D. M. (1983), Algorithm 611: Subroutine for unconstrained minimization using a model/trust-region approach, ACM Transaction on Mathematical Software, 9, 503-524.
[18] Powell, M. J. D. (1977), Restart procedures for the conjugate gradient method, Mathematical Programming, 12, 241-254.
[19] Nelder, J. A. and Mead, R. (1965), A simplex method for function minimization, Computer Journal, 7, 308-313.
[20] Levenberg, K. (1944), A method for the solution of certain problems in least squares, Quarterly of Applied Mathematics, 2, 164-168.
[21] Marquardt, D. (1963), An algorithm for least squares estimation of nonlinear parameters, SIAM Journal on Applied Mathematics, 11, 431-441.
[22] Hestenes, M. R. and Stiefel, E. (1952), Methods of conjugate gradients for solving linear systems, J. Res. N.B.S, 49, pp.409-436.
[23] Fletcher, R. and Reeves, C. M. (1964), Function minimization by conjugate gradients, Computer Journal, 7, pp,149-154.
[24] Polak, E. and Ribière, G. (1969), Note sur la convergence de methodes des directions conjugées, Revue Fr. Inf. Rech. Oper., 16-R1, pp.35-43.
[25] Morozov, A. (1966), On the Solution of Function Equations by the Method of Regularization, Soviet Mathematics Doklady, Vol.7, 414-417, 1966.
[26] Arcangeli, R. (1966), Pseudo-Solution de l'Equation Ax=y, Comptes Rendus des Seances de l'Academie des Sciences, Paris, Serie A, Vol.263, 282-285.
[27] Morozov, A., The Error Principle in the Solution of Operational Equations by the Regularization Method, Zh. Vychislit. Matem. i Matem. Fiz. Doklady, 8, 295-309, 1968.
[28] Engl, H. W. (1987), Discrepancy Principle for Tikhonov Regularization of Ill-Posed Problems Leading to Optimal Convergence Rates, J. Optimization Theory and Appl., 52, No.2, 209-215.
[29] Farebrother, R. W. (1988), Linear Least Squares Computations, Marcel Decker, New York,156-160.
[30] Gallant, A. R. (1987), Nonlinear Statistical Models, pp.16,17, 104, 105,259-266, Wiley, New York.
[31] Flach, G. P. and Ozisik, M. N. (1989), Inverse Heat Conduction Problems of Simultaneously Estimating Spatially Varying Thermal Conductivity and Heat Capacity per Unit Volume, Numerical Heat Transfer, Vol.16, No,2, 249-266.
[32] Huang, C. H. and Ozisik, M. N. (1990), A Direct Integration Approach for Simultaneously Estimating Spatially Varying Thermal Conductivity and Heat Capacity, Int. J. Heat and Fluid Flow, v.11, n.3, pp.262-268.
[33] Ho, C. H. and Ozisik, M. N. (1989), An Inverse Radiation Problem, Int. J. Heat Mass Transfer, v.32, n.2, pp.335-341.
[34] Visual Numerics (1994), IMSL Math/Library, Edition 10, Version 2.0, Suite 400, 9990 Richmond Avenue, Houston, TX-77042.
[35] L.E. Scales (March 1985), Introduction to Non-Linear Optimization, Springer Verlag, New York, NY.
[36] Cotta, R. M. (1993), Integral Transforms in Heat and Fluid Flow, CRC Press, Boca Raton.
[37] Aparecido, J. B. and Ozisik, M. N. (1998), Laminar Forced Convection Inside Circular Sector Tube with Variable Boundary Heat Flux, ENCIT98-7 ${ }^{\text {th }}$ Brazilian Congress of Engineering and Thermal Sciences, vol.1, pp.37-42, Rio de Janeiro.

## Appendix A: Linear Search Minimization.

In some minimization methods it is necessary to compute the scalar value of $\alpha^{k}$, that minimizes a generic function $S\left(\mathbf{p}^{k+1}\right)$, where $\mathbf{p}^{\mathbf{k}+1}$ is

$$
\begin{equation*}
\mathbf{p}^{\mathrm{k}+1} \equiv \mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}, \quad \alpha^{\mathrm{k}}>0 \tag{A1}
\end{equation*}
$$

once $\mathbf{p}^{\mathrm{k}}$ and $\Delta \mathbf{p}^{\mathrm{k}}$ are known. If $\hat{\alpha}^{k}$ is the value for $\alpha^{\mathrm{k}}$ that does this, then using the chain rule

$$
\left.\frac{\partial \mathrm{S}\left(\mathbf{p}^{\mathrm{k}+1}\right)}{\partial \alpha^{\mathrm{k}}}\right|_{\alpha^{\mathrm{k}}=\hat{\alpha}^{\mathrm{k}}}=\left.\frac{\partial \mathrm{S}\left(\mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\right)}{\partial\left(\mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}}} \frac{\partial\left(\mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\right)}{\partial \alpha^{\mathrm{k}}}\right|_{\alpha^{\mathrm{k}}=\hat{\alpha}^{\mathrm{k}}}=0(\mathrm{~A} 2)
$$

Considering that the two terms in the right side of equation (A2) are developed as $\left(\mathbf{g}^{\mathrm{k}+1}\right)^{\mathrm{T}}$ and $\Delta \mathbf{p}^{\mathrm{k}}$, thus it becomes

$$
\begin{equation*}
\left.\left(\mathbf{g}^{\mathrm{k}+1}\right)^{\mathrm{T}} \Delta \mathbf{p}^{\mathrm{k}}\right|_{\alpha^{\mathrm{k}}=\hat{\alpha}^{\mathrm{k}}}=0 \tag{A3}
\end{equation*}
$$

meaning that when the linear search minimization is completed, the gradient vector at $\mathbf{p}^{k+1}$ must be orthogonal to $\Delta \mathbf{p}^{k}$.

Formula (A3) is useful for theoretical purposes and when the function $\mathrm{S}(\mathbf{p})$ is generic. In specific cases, as in least squares function, it is possible to devise another way to do that. The formula for the least squares function (5a) at $\mathbf{p}^{k+1}$ is

$$
\begin{equation*}
S\left(\mathbf{p}^{k+1}\right)={ }_{m=1}^{M}\left[\Theta_{m}\left(\mathbf{p}^{k+1}\right)-Y_{m}\right]^{2} \tag{A4}
\end{equation*}
$$

where $\Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}+1}\right)$ can be expanded in a Taylor series which the first two terms are

$$
\Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}+1}\right)=\Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}+\alpha^{\mathrm{k}} \Delta \mathbf{p}^{\mathrm{k}}\right)=\Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right)+\alpha^{\mathrm{k}}\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}} \frac{\partial \Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right)}{\partial \mathbf{p}^{\mathrm{k}}}
$$

Substituting equation (A5) into (A4) it gives

$$
\begin{equation*}
\mathrm{S}\left(\mathbf{p}^{\mathrm{k}+1}\right)={ }_{\mathrm{m}=1}^{\mathrm{M}}\left|\Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right)-\mathrm{Y}_{\mathrm{m}}+\alpha^{\mathrm{k}}\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}} \frac{\partial \Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right)}{\partial \mathbf{p}^{\mathrm{k}}}\right|^{2} \tag{A6}
\end{equation*}
$$

If $\hat{\alpha}^{k}$ is the value of $\alpha^{k}$ that minimizes equation (A6), thus doing the derivative of $S\left(\mathbf{p}^{k+1}\right)$ relative to $\alpha^{k}$ and solving the resulting equation, it becomes

$$
\begin{equation*}
\hat{\alpha}^{\mathrm{k}}=-\frac{\operatorname{m}_{\mathrm{m}=1}^{\mathrm{M}}\left[\Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right)-\mathrm{Y}_{\mathrm{m}}\right]\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}}\left[\partial \Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right) / \partial \mathbf{p}^{\mathrm{k}}\right]}{\mathrm{M}_{\mathrm{m}=1}^{\mathrm{M}}\left[\left(\Delta \mathbf{p}^{\mathrm{k}}\right)^{\mathrm{T}}\left[\partial \Theta_{\mathrm{m}}\left(\mathbf{p}^{\mathrm{k}}\right) / \partial \mathbf{p}^{\mathrm{k}}\right]\right]^{2}} \tag{A7}
\end{equation*}
$$

where $\partial \Theta_{m}\left(\mathbf{p}^{\mathrm{k}}\right) / \partial \mathbf{p}^{\mathrm{k}}$ is the m -th column of the Jacobian matrix $\mathbf{J}\left(\mathbf{p}^{\mathrm{k}}\right)$.

