DESIGN OF EXPERIMENTS AND PROPERTY IDENTIFICATION

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

RESEARCH OF THERMAL CONDUCTIVITY FOR DIFFERENT MATERIALS BY SOLVING OF HEAT CONDUCTION INVERSE PROBLEMS

Konstantyn B. Isayev
Institute for Problems of Materials Science
National Academy of Science, NASc
Kyiv, K, Ukraine
isayev_k@gala.net

ABSTRACT

The technology for research of thermal conductivity of various materials is offered. It consists of experimental determination of one-dimensional temperature fields in heat-receivers (samples with thermal sensors) from a researched materials and use of these fields for the solving of inverse problems of heat conductivity. Investigation of materials with low (0.6 $Wm^{-1}K^{-1}$) and high (160 $Wm^{-1}K^{-1}$) thermal conductivity by the application of offered technology to research an influence of different factors on temperature dependence of materials thermal conductivity is considered.

INTRODUCTION

Thermal conductivity (**k**) is the fundamental characteristic of substances and materials. Without its knowledge it is impossible to create new materials, to design different products, to simulate their thermal conditions etc. Now there are a plenty traditional methods for determination of materials **k**. However, only some of them are realized in practice, for example, stationary methods, methods of pulse heating and some other. There is a huge quantity of the publications devoted the description of these methods, for example [1-3]. These methods for last some tens years have allowed to determine thermal conductivity of significant amount of different substances and materials. However, it is necessary to note, that the majority of these methods are labor-consuming, equipment, as a rule, is made in individual variant, they have no universality. The sample for determination of materials thermal conductivity usually has the insignificant sizes (except for stationary methods) that complicate research high porous and powder materials.

Any of traditional method does not allow to determine materials **k** in wide ranges of heating rate and values of thermal conductivity, to spend researches materials **k**, which at heating undergo to various physical-chemical transformations accompanying with allocation gas products (for example, metal hydrides, heat-shield materials with organic resins). The installations on the basis of stationary and pulse laser methods for determination of **k** have the high price (especially laser method).

From all of these lacks the methods for determination of thermophysical properties based on solving of inverse problems of heat conduction (IPHC) are free. A plenty of these methods are developed now. However, only some of them were used for research of thermophysical properties of various materials. As a rule, authors of these methods are limited with demonstration of opportunities of the methods on model temperature fields.

THE ENGINEERING FOR RESEARCH k OF MATERIALS

The engineering for research of materials thermal conductivity consists of several stages, which are shown in Fig.1. The first stage consists of manufacturing of heat receiver (HR), which is represented a sample of a researched material with thermal sensors (thermocouples). Further there is a stage of its tests in conditions of one-sided heating. Result of these tests is the experimental temperature fields in samples of a researched material. This information is necessary for solving of IPHC with use of the additional information about a researched material and heat receiver.

As a result of the solving IPHC is received thermal conductivity temperature dependence in frameworks of the originally chosen mathematical model of heat and mass transfer in the material. The experimental information (temperature fields

in samples of researched materials), additional information (data about specific heat and density of a material, distance of thermocouples from heating surface of heat receiver etc.) and also the results of the solving IPHC come in a database. All this information is used for the analysis of different factors influence on thermophysical properties of a material, and also for correction of mathematical model of heat and mass transfer in a researched material in case of need.

Investigation of materials with low (0.6 $Wm^{-1}K^{-1}$) and high (160 $Wm^{-1}K^{-1}$) thermal conductivity by the application of offered engineering for research an influence of different factors on temperature dependence of materials thermal conductivity is considered.

As a first asbestos cloth phenolic composite (ACPC) was chosen. The material contains in the structure phenol-formaldehyde resin (48%) and asbestos cloth. The analysis of influence both of kind and heating rate on effective thermal conductivity of this material was carried out. It was supposed, that specific volumetric heat (product of density on specific heat) of material is constant in all of investigated range of temperatures. In mathematical model of heat and mass transfer the presence of various physical-chemical processes (which take place at heating of ACPC) was not taken into account.

As materials with high thermal conductivity the composites on a basis of aluminium-silicon alloy AL25 were investigated. The alloy AL25

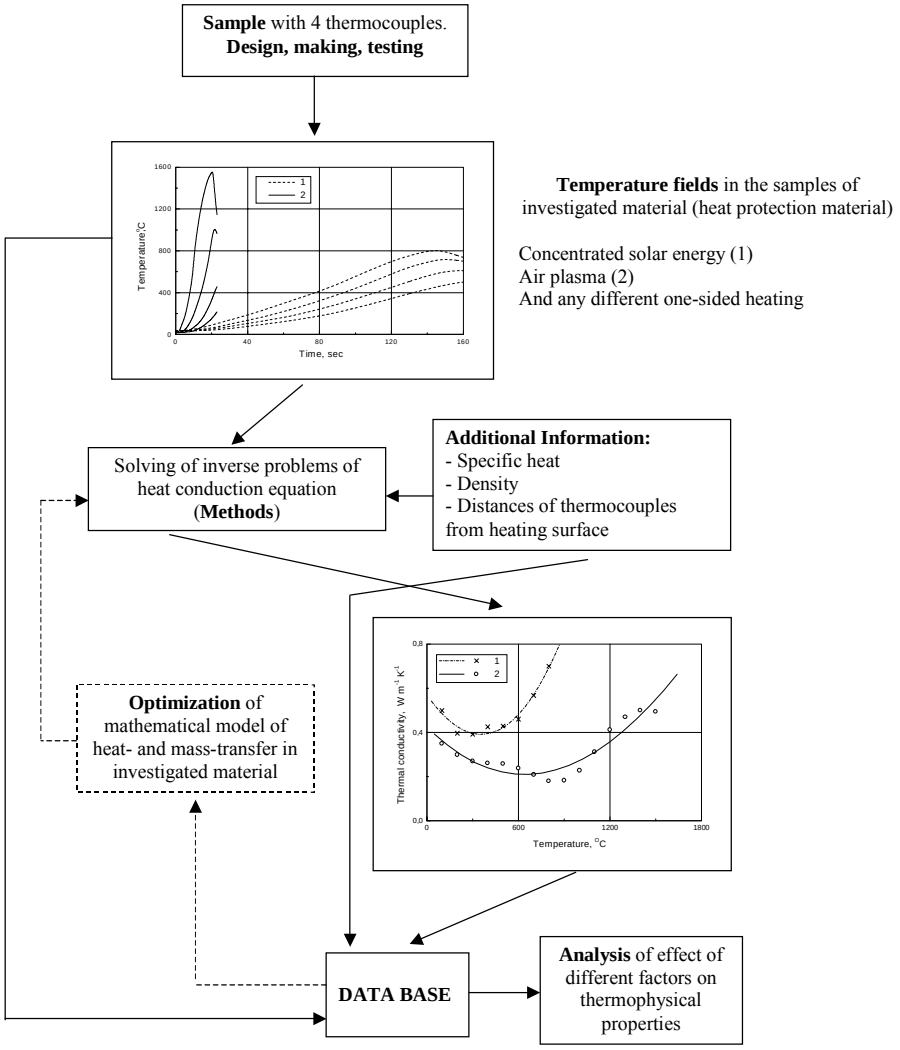


Fig.1. Engineering for research of thermal conductivity of materials.

consists basically of aluminium and silicon (11-13 %) and also small amount of the additives of metals (~ 4.5%), such as Cu, Mg, Ni. The research of influence of three powder additives: aluminium dioxide (22%), silicon carbide (18%) and mullitesilica (19%), on thermophysical properties of composites on the basis of alloy AL25 was carried out. The markings of compositions with these powders have the following kind: AL25+22, AL25+18, AL25+19. Samples of the composites were made by liquid punching with the further turning processing after crystallization and cooling. The temperature dependence of specific heat of these composites [4] was taken into account on the solving of IPHC. This characteristic was determined on standard installation, in which the method of monotonous heating is realized.

DESIGNS OF HEAT–RECEIVER

For experimental determination of one-dimensional temperature fields in samples of solid, powder and super hard materials the designs of heat-receiver [5] were developed. In Fig.2 the designs of heat-receiver (HR) for materials with high (a) and low (b) thermal conductivity are submitted. In both designs, the sample of a researched material represents the step cylinder (1), in which four cuts parallel to heated face (Fig.3) were made by a diamond disk (thickness 0.1mm) in depth half of diameter of the smaller cylinder (1). In these cuts the thermocouples (4) (diameter of electrodes is 0.1mm) with insulating covering of yttrium dioxide are installed. Thickness of a covering is some microns. The cuts are filled with a high thermal conductivity powder (6) for thermal resistance decreasing (Fig.3). The material of a powder should be chemically inert to a researched material. It can be, for example, the powders of crystal quartz or copper. The powder, which filling cuts in a sample of a researched material, does not resist for gas products penetration, if those are available. These products are the result of a material thermal destruction. The alternative to such way can be served a thermocouples closing in apertures, drilled on radius the smaller diameter cylinder instead of cuts. These apertures after installing of thermocouples also are filled with high thermal conductivity powder.

These two ways of thermocouples closing were tested on samples of the alloy AL25. For filling of cuts and apertures the copper powder was used. As a result of solving IPHC it is not influence of a way for thermocouples closing on thermal conductivity of this alloy.

Lateral surface of a smaller diameter cylinder is covered by silica filament. Atop of the filament the thermocouple wires are installed. They covered by the second layer of a silica filament. The sample is located in the bushing (5) of a researched material (or material with higher thermal conductivity). It prevents heat outflow from a lateral surface of sample (1) and hinders

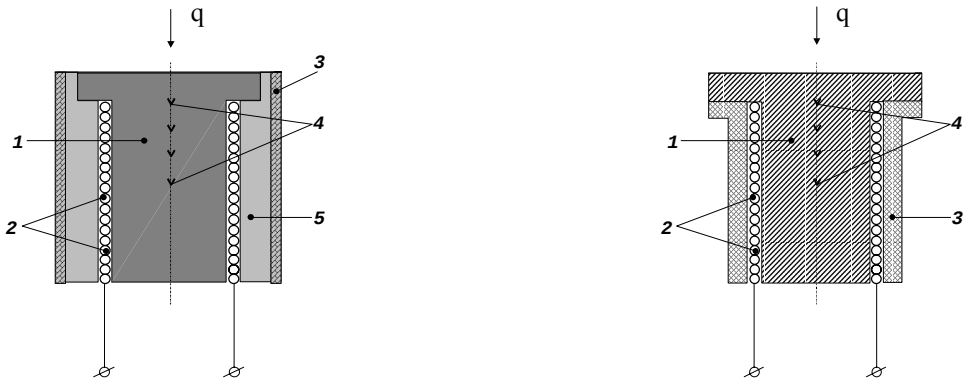


Fig. 2. Heat-receiver designs for solid materials with:
a - high thermal conductivity; b - low thermal conductivity.
1 - researched material; 2 - two layers of silica filament; 3 - bushing of heat-insulated material;
4 - thermocouples; 5 - bushing of researched material.

with violation of one-dimensionality of a temperature field in this sample. Further all design is located in the bushing (3) which is made of heat-insulated material (Fig.2a). For materials with low thermal conductivity the bushing (5) is not used (Fig.2b).

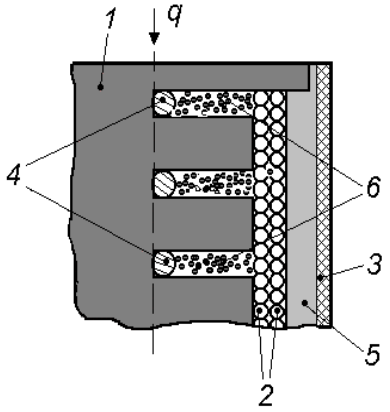


Fig.3. Engineering of thermocouples closing in samples of solid materials. The designations are similar Fig.2. 6 - powder of high thermal conductivity material.

The one-dimensionality of a temperature field in a sample of a researched material is provided with conditions of uniform heating of its surface and design features of heat receiver.

For investigated materials smaller diameter of a sample (1) made 18 mm, external diameter of bushing (3) made 25 mm, height of a sample (1) made 30-40 mm. The developed designs of heat receivers can be used both for determination of thermal conductivity of various materials and in structure of different products for determination of boundary conditions on its heating surface with the help of inverse problems solving of heat conduction.

EXPERIMENTAL TEMPERATURE FIELDS

The experimental temperature fields in samples of researched materials are the base information for determination of their **k** by solving IPHC. One-dimensional temperature fields in samples of asbestos material and alloy AL25, received in conditions of one-sided heating, are given below. The error of temperature measurement made 7-10 %.

Asbestos composite

Samples ACPC were tested at convective and radiation one-sided heating. The conditions of their tests are submitted in table 1. The design of HR of this material for experimental determination of temperature fields is submitted in Fig.2b. After closing of thermocouples in cuts they were filled by quartz powder.

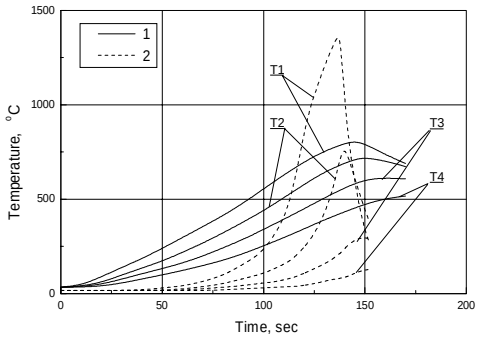


Fig.4. Temperature fields in samples of ACPC. 1 - N19, 2 - N28. T1-T4 are four thermocouples indications.

Table 1. Test specifications of samples ACPC and depths of thermocouples closing.

N	Test conditions of samples	Heat flux, MW/m^2	Number of sample	Depths of thermocouples from heating surface			
				h_1, mm	h_2, mm	h_3, mm	h_4, mm
1	Concentrated solar radiation	0,50	19	5,40	6,35	7,30	8,30
2	“	6,3	28	15,3	17.3	19,6	21,7
3	Air plasma	2,32	22	2,40	3,40	4,35	5,30
4	Combustion products of kerosene in oxygen	4,0	5	2,00	3,05	4,10	5,10

ACPC sample prepared with thermocouples was located in water-cooling holder. The holder

with heat-reseiver was entered into a jet of high-temperature gas (or in a focal volume of a solar energy concentrator) on beforehand selected distance from a nozzle exit section before combining of axes of a jet and sample.

The registration of thermocouple indications was realized with the help of the loop oscillograph. Cold junctions of four working thermocouples (tungsten-rhenium 5/20) were heat-insulated and their temperature was measured with the help of copper-constantan thermocouple. Its indications also were registered by loop oscillograph. The experimental results are submitted in Fig.4 and Fig.5. The sample N28 ACPC was tested in conditions of quasistationary process of heating. This process of heating is characterized by constancy both of heating surface temperature and linear rate of its movement. Surface temperature of all investigated samples ACPC was controlled with the help of photoelectric pyrometer. About the presence of quasistationary process of heating in sample N28 indicate the equidistance of curves T1 and T2 in Fig.4 since temperature 200^0C to disconnect moment of heating. All other samples of this material were tested in conditions of non-stationary heating.

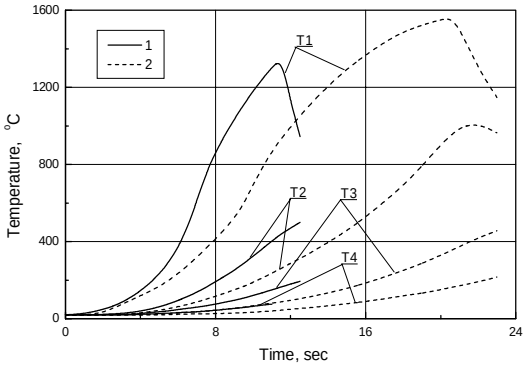


Fig.5. Temperature fields in samples of ACPC. 1 – N5, 2 – N22.

The indications of 4-th thermocouple of sample N5 coincide with the indications of 3-d thermocouple of sample N22 (Fig.5).

Composites on AL25 basis

The experimental determination of temperature fields in samples of the composites was realized at intensive one-sided heating. As a source of heat was used melt of lead overheat to temperature approximately 400^0C.

Sample of a researched material sharply lowered in melt on depth approximately 5mm. The registration of thermocouple indications was realized also with the help of loop oscillograph. Distance from heating surface of tested samples to first thermocouple was 2mm, and distance between thermocouples was 2.4mm for all composites.

It was tested three samples for each composite. As an example, in Fig.6 the experimental temperature field in sample N1 of alloy AL25 is given. This field is typical for all materials on the basis of this alloy, tested in the given conditions of one-sided heating. The character of the thermocouple indications is growth of temperature took place. Then slow

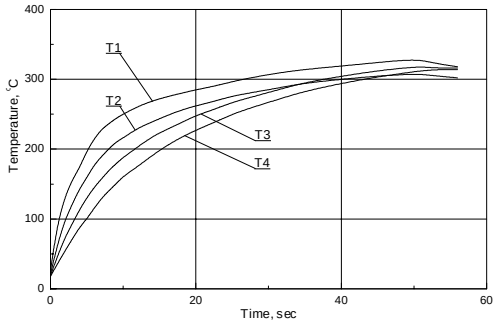


Fig.6. Experimental temperature field in sample N1 of alloy AL25.

increase of temperature to the moment of HR extraction from the melt.

METHODS FOR DETERMINATION k MATERIALS

Two methods for solving of inverse problems of heat conduction were used for determination of temperature dependence of materials thermal conductivity in conditions of one-sided heating. The methods of solving IPHC are realized in the programs KV [6] and FRIEND [7].

In KV is used a method based on laws of heat transfer at quasistationary process of sample heating. This sample simulates a semi-bounded body. The registration of a time changing of temperature in one section of a sample is needed for realization of this method. In addition, it is necessary to measure linear rate of material ablation simultaneously (for example, with the help of filming). If independent measurement of linear rate is absent it is necessary the time changing of temperature register in two sections of a sample. In the latter case, linear rate of

material is determined as quotient from division of distance between thermocouples on constant time passage of this distance by any temperature. All range of indications of the first thermocouple is broken on N intervals, in each of them thermal conductivity of a material is supposed constant. The received ratio [6] allows calculate temperature dependence of thermal conductivity of a researched material. In this ratio the temperature dependence of specific volumetric heat, absorption of heat both at destruction of resin and by gases, formed as a result of this process, took into account.

FRIEND is used an iterative method for solving of inverse problems of heat conduction, based on the theory of sensitivity functions. The mathematical model of non-linear non-stationary equation of heat conduction for unlimited plate with boundary conditions of the first kind is used. The unknown temperature dependence of thermal conductivity is presented as a functional series, with basis functions of cubic B-splines. For searching of unknown parameters vector the iterative identification is used. Thus the increment of unknown parameters vector is determined from the solution of linear algebraic equations system with the help of method of least squares with method of regularization of A.N. Tikhonov [8]. The regularization parameter equal 100 was used in activity. For realization of this method the registration of a time changing of temperature in 3-4 sections of a sample of a researched material is necessary.

The described above methods supplement each other. At high heating rates and at presence of ablation of a material, when the temperature field is compressed and practically it is impossible to measure change of temperature in several sections of a sample, the first method is used, and at low heating rate is used second.

k OF INVESTIGATED MATERIALS

Asbestos cloth fenolic composite and composites on the basis of an alloy AL25 were chosen for demonstration of the technology application for research of their thermal conductivity. For these materials the values of thermal conductivity differ approximately in 270 times at a room temperature. The first material is used as thermal protection of various products of rocket engineering and second is used in designs of internal combustion engines. The error of **k** determination depends from an error of temperature measurements, specific heat, density etc and on our evaluation makes 15-25 %.

Asbestos Material

The influence both of kind and heating rate on effective thermal conductivity ACPC is shown in Fig.7. The temperature field in a sample N28 was processed by two described above methods of solving IPHC. The results are given on Fig.7 by points 4 and points 5. These points practically coincide with other except for high temperatures, where values **k** (received with the help of FRIEND) slump. It is probably connect with the large difference of the thermocouple indications at high temperatures (Fig.5). The thermal conductivity ACPC in a range of resin destruction temperatures has a minimum ($\mathbf{k}_m$), which is displaced in area of high temperatures and values $\mathbf{k}_m$ decrease at the same time (Fig.7, Table 2). This minimum of thermal conductivity ACPC is caused by the maximal rate of resin thermal decomposition. It is accompanied by the maximal rate of gas products allocation at this decomposition. By this reason the maximal porosity took place in coke ACPC at temperature of minimum thermal conductivity ($\mathbf{T}_{\mathbf{k}\,m}$).

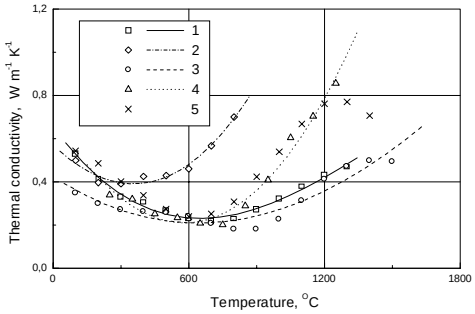


Fig7. Influence both kind and rate of heating on thermal conductivity of ACPC.
1 - N5, 2 - N19, 3 - N22, 4 - N28 (KV),
5 - N28 (FRIEND).

The methodical calculations have shown that the account of temperature dependence of specific volumetric heat and the absorption of heat at decomposition and by gases do not influence on temperature $\mathbf{T}_{\mathbf{k}\,m}$.

With increasing of temperature the pores in coke “are healed” by carbon, which is falling out from gas products of resin destruction. This process and other (for example, growth of radiation component **k**) promote for growth of thermal conductivity of coke. The decreasing of $\mathbf{k}_m$ value with increasing of heating rate is

connected with change of orientation of pores and channels in coke of material. With increasing of heating rate pores and channels in coke have a tendency to orient parallel heating surface. The analysis of pores structure in coke ACPC was carried out with the help of metallurgical microscope.

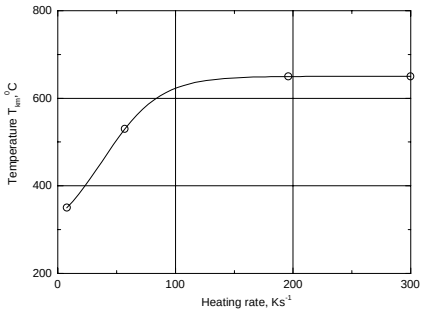


Fig.8. Influence of heating rate on temperature, at which **k** of ACPC has a minimum.

The introduction of parameter $\mathbf{T_{km}}$ allows analyze influence of heating rate on process of resin destruction of heat-sheild materials with different fillers simultaneously.

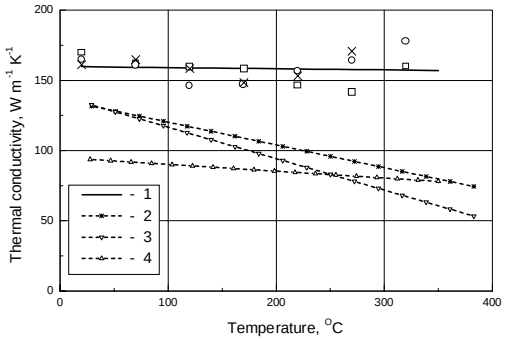


Fig.9. Thermal conductivity temperature dependence for composites on the basis of alloy AL25.
1 - AL25; 2 - AL25+18; 3 - AL25+22; 4 - AL25+19.

Table 2. Influence of heating rate on the minimal value of thermal conductivity ACPC.

N	N of sample	Max. heating rate, K/s	$\mathbf{k}_m$, $Wm^{-1}K^{-1}$	$\mathbf{T_{km}}$, ^0C
1	19	7.6	0.39	350
2	28	57	0.27	530
3	22	196	0.21	650
4	5	300	0.23	650

The influence of heating rate on temperature $\mathbf{T_{km}}$ of asbestos material is shown in Fig.8. The curve in this figure is constructed with the help of the least squares method with approximation of $\mathbf{T_{km}}$ values by Boltzman function. The behaviour of temperature $\mathbf{T_{km}}$ shows, that displacement of thermal conductivity minimum of this material take place up to 100K/s. This circumstance, in turn, speaks that the displacement of thermal decomposition process of resin in area of high temperatures has a place up to this heating rate.

In table 2 there are the maximal heating rate of samples, the minimal values of thermal conductivity of asbestos material and temperature, at which **k** of this material has minimum.

The maximal heating rate of samples ACPC was estimated under the indications of the first thermocouple (from the side of a heating surface). At the solving IPHC with the help of FRIEND the part of a sample between the first and fourth thermocouples is considered only.

Composites on AL25 basis

The temperature dependence of thermal conductivity of the investigated composites are submitted on Fig.9. They are received by linear approximation of values of this characteristic (results of solving IPHC) for three samples of each composite with the help of the least squares method. As an example, in this figure the dispersion of the data **k** of an alloy AL25 is shown. The symbols designate values of thermal conductivity AL25 for 3 samples. The injection in alloy AL25 the additives of silicon carbide and aluminium dioxide is resulted to decreasing of its thermal conductivity. This decreasing becomes greater with increasing of temperature. It is generated first of all with that circumstance, that thermal conductivity of the additives in some times below this characteristic AL25 [1,2].

It is necessary also to note that for silicon carbide and aluminium dioxide sharp decreasing of values of their thermal conductivity has a place

with growth of temperature. So, according to data [1,2] temperature increasing from 100^0C up to 400^0C leads to decreasing of thermal conductivity values of these materials in 3-4 times. For a composite with mullitesilica such sharp decreasing of thermal conductivity is absent.

CONCLUSIONS

The engineering of thermal conductivity research of different materials is developed. It has allowed analyze the influence of the different factors on **k** of materials with low and high thermal conductivity.

The increase of heating rate for asbestos material (it is the typical representative of heat-shield materials with organic resins) results to displacement in area of high temperature of its effective thermal conductivity minimum. The analysis of **k** temperature dependence of this material (**k** received by solving IPHC with use of experimental data at various kinds of one-sided heating) has shown, that the **k** minimum is displaced in area of high temperature approximately up to 100K/s. This circumstance speaks that the displacement of thermal decomposition process of resin in area of high temperatures has a place up to this heating rate.

The injection in alloy AL25 of 3 various powder additives results in decreasing its thermal conductivity. The injection in this alloy of silicon carbide and aluminium dioxide results in sharp fall of temperature dependence of its thermal conductivity. For a composite with mullitesilica such sharp decreasing of thermal conductivity is absent.

REFERENCES

1. Kingery W.D. Thermal conductivity: VI, Determination of conductivity of Al_2O_3 by spherical envelope and cylinder methods. *J. Amer. Ceram. Soc.,* **37**, 2 (part II), 88 (1954).
2. Vasilos T. and Kingery W.D. Thermal conductivity: XI. Conductivity of some refractory carbides and nitrides. *J. Amer. Ceram. Soc.*, **37**, 9, 409 (1954).
3. Compendium of thermophysical property measurement methods. Vol.1. Survey of measurement techniques. *Ed. Maglic K.D. e.a. New York; London: Plenum Press*, 1984. XVI, 789pp.
4. Isayev K.B. Thermophysical properties of composites on basis of alloy AL25. *Industrial Heat Engineering*." **3**, 1-2, 77 (2001).
5. Isayev K.B. Designs of heatreceiver for determination of one-dimensional temperature fields in compact and powder materials. *IV Minsk Int. Forum: Heat/Masstransfer-2000,* **3**, 394 (2000) (in Russ.).
6. Isayev K.B. Heat transfer in the destroying under one-sided heating composite materials. *Engineering Physical J*., **65**, 6, 645 (1993) (in Russ.).
7. Krukovsky P.G. A universal approach to solution of inverse heat transfer problems (method and software). *30th National Heat Transfer Conf., N. Y. (1995), ASME (UEC)*, **10**, 107 (1995).
8. Tikhonov A.N. About solving of ill-posed problems and regularization method. *Reports AN SSSR*. **151**, 3, 501 (1963) (in Russ.).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INTEGRATING THE ERROR IN THE INDEPENDENT VARIABLE FOR OPTIMAL PARAMETER ESTIMATION

PART I: DIFFERENT ESTIMATION STRATEGIES ON ACADEMIC CASES

Denis Maillet, Thomas Metzger, Sophie Didierjean
LEMTA - CNRS UMR 7563 - INPL - UHP-Nancy I
Vandoeuvre-lès-Nancy, 54, France
dmaillet@ensem.inpl-nancy.fr

ABSTRACT

The classical approach to the linear – or non-linear – parameter estimation problem deals with three types of variables: the dependent variable, several independent variables and unknown parameters. These are estimated starting from experimental values of the signal. In many inverse heat transfer problems the location of the detectors is usually not perfectly known and the estimated parameters can be severely biased if this is not taken into account. In order to get a better insight into this case, the academic linear least squares problem is revisited for high signal over noise ratio. The independent variable (nominal abscissa) is considered as both a signal and a parameter, which requires the minimization of a modified two terms least squares functional. The stochastic properties of the resulting estimator are reviewed and its efficiency is tested. This technique has been successfully implemented for thermophysical property measurement in a porous medium and is presented in the second part of this paper.

NOMENCLATURE

b	bias
cov	variance – covariance matrix
E	expected value
n	number of points
N	number of simulations
Q	variance ratio ($= \sigma'^2/\sigma^2$)
r	linear correlation coefficient between y and t or τ
s	plus or minus sign ($= r/\lvert r \rvert$)
s_u	statistical standard deviation of u
s_{uv}	statistical covariance between u and v
S	two terms modified least square sum
S_t	least squares sum for 'signal' t
S_y	least squares sum for y
t	dependent noised variable (location)
X	space coordinate or time
$\boldsymbol{X}$	matrix of the sensitivity coefficients or sensitivity vector
y	dependent noised variable
Z	coefficient

Greek symbols

α	extended parameter vector
β_1, β_2	straight line parameters
β	parameter vector
δ	noise on τ
δ_{ij}	Kronecker symbol
ε_{ij}	noise on a time – space signal
ε	noise on η
η	dependent variable
ρ_{12}	correlation coefficient between parameters
σ, σ'	standard deviations of ε and η
$\nabla_{\boldsymbol{u}}$	gradient with respect to vector $\boldsymbol{u}$

Subscripts

cor	corrected value
lin	associated to the linear estimation

Superscripts

k	number of a Monte Carlo inversion
(m)	order of a Taylor series expansion (m = 0, 1 or 2)
-	average value
$*$	reduced quantity
$\wedge$	estimator or estimated quantity
t	transpose of a matrix or vector

INTRODUCTION

The usual approach to the linear, or non-linear, parameter estimation problem [1] consists in working with models involving three types of variables. The first one is the dependent variable η, whose experimental value y, corrupted by additional noise, is the signal, which is a stochastic variable. This dependent variable is linked, through a model, to both one or more independent variables X_1, X_2,..,X_v and to a certain number of unknown parameters that have to be estimated once experimental values of the signal have been recorded.

In many inverse heat transfer problems, where the signal y is produced by a set of detectors (j = 1, 2,…, d) at different measurement times t_i (i = 1, 2,..., n), one can write:

$$y_{ij} = \eta\,(X_{1j},\, X_{2j},\, X_{3j},\, t_i;\, \boldsymbol{\beta}) + \varepsilon_{ij} \qquad (1)$$

where X_{1j}, X_{2j}, X_{3j} designate the coordinates of the j^{th} detector and $X_{4i} = t_i$ the i^{th} time of measurement, ε_{ij} being the noise associated with this measurement. $\boldsymbol{\beta}$ is the parameter vector here.

Usually y and ε are considered as stochastic variables while time X_4 and location (X_1, X_2, X_3) are assumed to be known without any error (deterministic variables). The problem of estimating $\boldsymbol{\beta}$ is solved by classical regression in that case. If modern data acquisition systems yield data associated with usually errorless times, any heat transfer experimentalist knows that the location of the detectors cannot be known with an infinite precision. That means that X_1, X_2 and X_3 should also be considered as stochastic variables in order to prevent any estimation of $\boldsymbol{\beta}$ with a too large bias. This type of model, which differs from the classical regression model, is called measurement error model in the statistical literature, and has been extensively studied for more than a century [1 p. 199, 2, 3 p. 492-496, 4].

It is this problem that will be presented next, in a very academic linear case, the straight line model, where only one independent variable (v = 1) will be considered.

THE STRAIGHT LINE CASE

The following linear model is considered:

$$\eta = \eta\,(\tau;\, \boldsymbol{\beta}) = \beta_1 + \beta_2\,\tau \qquad (2)$$

where η and τ are the dependent and independent variable respectively and β_1 and β_2 the two parameters. One first assumes that experimental noised data y_i are available at 'locations' τ_i (i = 1 to n):

$$y_i = \eta\,(\tau_i;\, \boldsymbol{\beta}) + \varepsilon_i \qquad (3)$$

where ε_i is an uncorrelated unbiased noise of constant standard deviation σ and $\boldsymbol{\beta} = [\beta_1 \;\; \beta_2]^t$.

$$E\,(\varepsilon_i) = 0 \qquad E\,(\varepsilon_i\,\varepsilon_j) = \sigma^2\,\delta_{ij} \qquad (4)$$

EXACT INDEPENDENT VARIABLE

If the τ_i's are known exactly, the best estimation technique consists in minimizing the ordinary least squares sum:

$$S_y\,(\beta_1, \beta_2) = (\boldsymbol{y} - \boldsymbol{X}\,\boldsymbol{\beta})^t\,(\boldsymbol{y} - \boldsymbol{X}\,\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[y_i - \eta\,(\tau_i;\, \boldsymbol{\beta})\right]^2 \qquad (5)$$

where $\boldsymbol{X}$ is the sensitivity matrix and $\boldsymbol{y}$ the measurement vector:

$$\boldsymbol{X} = \begin{bmatrix} 1 & 1 & .... & 1 \\ \tau_1 & \tau_2 & .... & \tau_n \end{bmatrix}^t \qquad \boldsymbol{y} = \begin{bmatrix} y_1 & y_2 & ... & y_n \end{bmatrix}^t \qquad (6)$$

In this case the solution is explicit:

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^t\,\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t\,\boldsymbol{y} \qquad (7)$$

$$\Rightarrow \quad \hat{\beta}_2 = s_{\tau y} / s_\tau^2 \quad ; \quad \hat{\beta}_1 = \bar{y} - \hat{\beta}_2\,\bar{\tau}$$

with $\bar{\tau} = \dfrac{1}{n}\sum_{i=1}^{n} \tau_i$; $\bar{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$

and

$$s_{\tau y} = \frac{1}{n}\sum_{i=1}^{n} \tau_i\, y_i - \bar{\tau}\,\bar{y} \quad ; \quad s_\tau^2 = \frac{1}{n}\sum_{i=1}^{n} \tau_i^2 - \bar{\tau}^2$$

In these expressions s_τ^2 is the statistical variance of the locations (a measure of the dispersion of these locations) and $s_{\tau y}$ the statistical covariance between signal and location of measurement.

The covariance matrix of the estimation error can also be calculated:

$$\mathrm{cov}\,(\hat{\boldsymbol{\beta}}) = \sigma^2 \left(\boldsymbol{X}^{\mathrm{t}}\,\boldsymbol{X}\right)^{-1} = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

with

$$\sigma_1 = \frac{\sigma}{\sqrt{n}}\,(1 + \bar{t}^{\,2}/s_\tau^2)^{1/2} \quad ; \quad \sigma_2 = \frac{\sigma}{s_\tau\,\sqrt{n}}$$
$$\rho_{12} = -\frac{1}{(1 + \bar{t}^{\,2}/s_\tau^2)^{1/2}} \tag{8}$$

where σ_1 and σ_2 are the standard deviations of the estimators of β_1 and β_2, and ρ_{12} their correlation coefficient.

UNKNOWN INDEPENDENT VARIABLE

If no information is available for the τ_i's, the first idea is to integrate these unknowns into a new extended parameter vector $\boldsymbol{\alpha} = \left[\boldsymbol{\beta}^{\mathrm{t}}\ \boldsymbol{\tau}^{\mathrm{t}}\right]^{\mathrm{t}} = \left[\beta_1\ \beta_2\ \tau_1\ \tau_2\ \ldots.\ \tau_n\right]^{\mathrm{t}}$.

The matrix of the new sensitivity coefficients can be calculated:

$$\boldsymbol{X}_\alpha = \begin{bmatrix} 1 & \tau_1 & \beta_2 & 0 & \ldots & 0 \\ 1 & \tau_2 & 0 & \beta_2 & 0 & 0 \\ \vdots & \vdots & & 0 & \ddots & \vdots \\ 1 & \tau_n & 0 & \ldots & 0 & \beta_2 \end{bmatrix}^{\mathrm{t}} \tag{9}$$

It can easily be shown that this matrix is singular since the sensitivity vectors $\boldsymbol{X}_{\alpha j}$ (the column vectors of $\boldsymbol{X}_\alpha$) are linearly dependent:

$$\beta_2\ \boldsymbol{X}_{\alpha 1} + 0\ \boldsymbol{X}_{\alpha 2} - \sum_{j=1}^{n} \boldsymbol{X}_{\alpha\,(j+2)} = 0 \tag{10}$$

This means that it is impossible to estimate β_1 and β_2 without any prior information about the τ_i's.

UNCERTAIN INDEPENDENT VARIABLE

Usually only approximate (nominal) values t_i of the dependent variables τ_i are available:

$$t_i = \tau_i + \delta_i \tag{11}$$

where δ_i is an uncorrelated unbiased noise, of constant standard deviation σ', which is not correlated with ε_j.

$$E\,(\delta_i) = 0 \qquad E\,(\delta_i\ \delta_j) = \sigma'^2\ \delta_{ij}$$
$$E\,(\delta_i\ \varepsilon_j) = 0 \tag{12}$$

In the Measurement Error Model terminology, model (3, 4, 11, 12), where the unknown dependent variable τ_i is deterministic, is called a functional model, while cases where τ_i are stochastic variables with either fixed or variable (*i* dependent) expected values are called a structural or a ultrastructural model [2, 4]. We will deal here with the functional model only.

With this new information two kinds of measurements are available, the y_i's and the t_i's, and three kinds of parameters are sought after, parameters β_1 and β_2 and the unknown 'locations' τ_i. These locations are therefore considered both as a 'signal' (t_i's) and as unknown parameters (τ_i's). In order to take into account the two different kinds of signals, that do not have the same unit, under the form of a least squares sum, the two kinds of residuals have to be normalized by the two different standard deviations σ and σ'. The sum that has to be minimized with respect to parameter vector $\boldsymbol{\alpha}$ becomes:

$$S\,(\boldsymbol{\beta},\boldsymbol{\tau}) = \sum_{i=1}^{n}\left(\frac{y_i - \eta\,(\tau_i;\,\boldsymbol{\beta})}{\sigma}\right)^2 + \sum_{i=1}^{n}\left(\frac{t_i - \tau_i}{\sigma'}\right)^2$$
$$= \frac{1}{\sigma^2} S_y\,(\boldsymbol{\beta},\boldsymbol{\tau}) + \frac{1}{\sigma'^2} S_t\,(\boldsymbol{\tau}) \tag{13}$$

or, in a matrix – column vector form:

$$S\,(\boldsymbol{\beta},\boldsymbol{\tau}) = \frac{1}{\sigma^2}\left(\boldsymbol{y} - \boldsymbol{X}\,(\boldsymbol{\tau})\boldsymbol{\beta}\right)^{\mathrm{t}}\left(\boldsymbol{y} - \boldsymbol{X}\,(\boldsymbol{\tau})\boldsymbol{\beta}\right)$$
$$+ \frac{1}{\sigma'^2}\left(\boldsymbol{t} - \boldsymbol{\tau}\right)^{\mathrm{t}}\left(\boldsymbol{t} - \boldsymbol{\tau}\right) \tag{14}$$

This estimation problem is not linear any more. The estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ are obtained through solution of the following system:

$$\nabla_{\boldsymbol{\beta}}\ S_y(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\tau}}) = 0 \tag{15}$$

$$\nabla_{\boldsymbol{\tau}}\ S_y(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\tau}}) + \frac{1}{Q}\nabla_{\boldsymbol{\tau}}\ S_t\,(\hat{\boldsymbol{\tau}}) = 0 \tag{16}$$

with $Q = (\sigma'/\sigma)^2$

Equation (16) allows the calculation of the estimator of τ in terms of the estimator of β:

$$\hat{\tau}_i = \frac{t_i + Q\,(y_i - \hat{\beta}_1)\,\hat{\beta}_2}{1 + Q\,\hat{\beta}_2^2} \qquad (17)$$

One can notice at that point that errorless *t* measurements (Q = 0) lead to $\hat{\tau} = \boldsymbol{t}$ while no knowledge on t ($Q \to \infty$), or, equivalently errorless *y*'s, leads to $y_i = \hat{\beta}_1 + \hat{\beta}_2\,\hat{\tau}_i$. This means that in these two extreme cases either S_y or S_t in the least square sum S – equation (13) - is equal to zero. Even if it is not a function that has to be estimated, coefficient Q behaves like a regularization coefficient. One can also say that the second S_t term brings some information, which corresponds to a Bayesian maximum likelihood approach [1, 4, 5].

Equation (15) provides the classical linear estimator of β in terms of $\hat{\tau}$:

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^t(\hat{\tau})\,\boldsymbol{X}(\hat{\tau})\right)^{-1}\boldsymbol{X}^t(\hat{\tau})\,\boldsymbol{y} \qquad (18)$$

Substitution of the $\hat{\tau}_i$'s given by equation (17) into equation (18) provides the following equations (for $Q \neq 0$):

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2\,\bar{t} \qquad (19)$$

$$\hat{\beta}_2^2 - 2\,Z\,\hat{\beta}_2 - 1/Q = 0 \qquad (20a)$$

with $$Z = \frac{Q\,s_y^2 - s_t^2}{2\,Q\,s_{ty}} \qquad (20b)$$

where the upper bar designates the statistical average, s_y^2 and s_t^2 the statistical variances of *y* and *t* respectively, and s_{ty} the statistical covariance of *t* and *y* - see equation (7) for the definition of these coefficients. Equation (20a) has two solutions of opposite signs that can be discriminated using the linear correlation coefficient *r* between the *t*'s and the *y*'s:

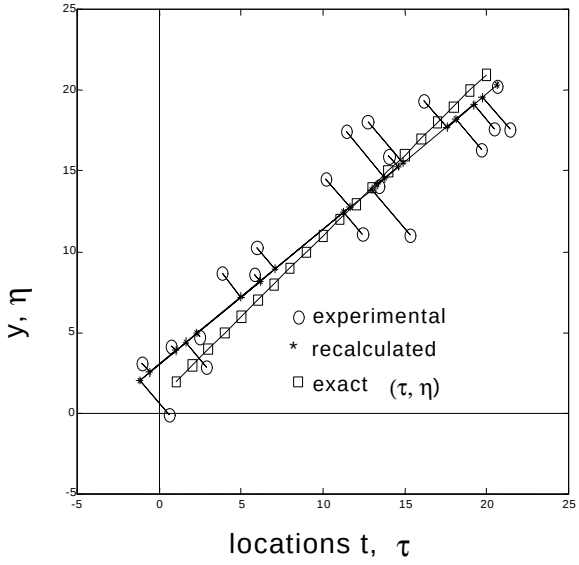


Figure 1. Example of estimation (*S* minimization)

$$\hat{\beta}_2 = Z + s\left(Z^2 + 1/Q\right)^{1/2}$$
$$\text{with } r = \frac{s_{ty}}{s_t\,s_y} \text{ and } s = \frac{r}{|r|} \qquad (21)$$

If σ' tends to zero, one can show, using a first order series expansion of $\hat{\beta}_2$ with respect to Q, that $\hat{\beta}_2$ and, consequently, $\hat{\beta}_1$ tends to the classical linear estimator given by equation (7).

If it is σ that tends to zero, it can be shown, using the same kind of series expansion with respect to $1/Q$, that $1/\hat{\beta}_2$ tends to the value of $\hat{\beta}_2$ given by the linear estimator – equation (7). This means that in this case we meet the classical linear model where τ has to be expressed as a function of η (permutation of the dependent and independent variables).

This shows that the non linear estimator given by equations (21), (19) and (17) is only a generalization of the classical linear estimator (7).

This estimator – equations (17, 21) - has been previously derived (for the $\beta_1 = 0$ case), starting from a maximum likelihood estimation (MLE) associated with a Lagrange multipliers approach [1, p.198-199]. It is also derived using MLE in [3, p. 23]. In our approach, we assume that *Q*, the

variance ratio is known, and we do not try to estimate it using MLE.

An example of estimation is given in Figs. 1 and 2, for n = 20 locations distributed uniformly between 1 and 20, for the case $\beta_1 = \beta_2 = 1$ and $\sigma = \sigma' = 3$.

It can be noted that if a normalization of both dependent and independent variables and on their associated noises is done, according to:

$$\eta^* = \eta/\sigma \;;\; \tau^* = \tau/\sigma' \qquad \varepsilon^* = \varepsilon/\sigma \;;\; \delta^* = \delta/\sigma' \tag{22a}$$

$$\Rightarrow \quad y^* = y/\sigma \quad ; \quad t^* = t/\sigma' \tag{22b}$$

model (2) becomes:

$$\eta^* = \beta_1^* + \beta_2^* \tau^* \tag{23}$$

with $$\beta_1^* = \beta_1/\sigma \quad ; \quad \beta_2^* = \beta_2 (\sigma'/\sigma) \tag{24}$$

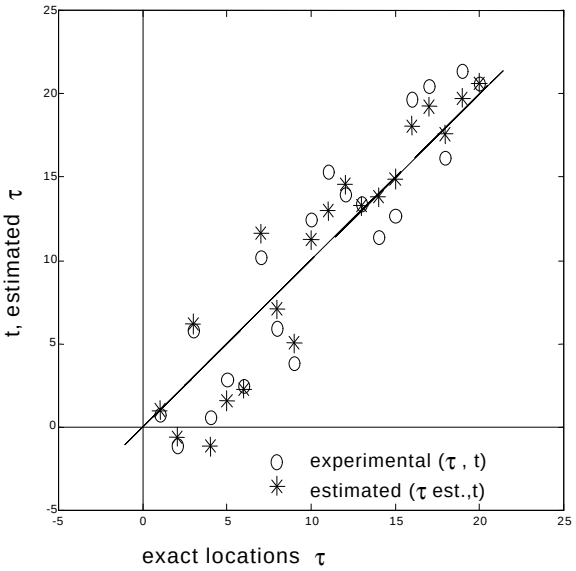


Figure 2. Example of estimation of the dependent variable (S minimization)

and the normalized sum:

$$S^*(\boldsymbol{\beta}^*, \boldsymbol{\tau}^*) = \sum_{i=1}^{n} \left(y_i^* - \eta^*(\tau_i^*; \boldsymbol{\beta}^*)\right)^2 + \sum_{i=1}^{n} \left(t_i^* - \tau_i^*\right)^2 = S_y^*(\boldsymbol{\beta}^*, \boldsymbol{\tau}^*) + S_t^*(\boldsymbol{\tau}^*) \tag{25}$$

In the normalized (t^*, y^*) plane, the least square sum S^* represents the minimum of the square of the distance between the experimental points (t_i^*, y_i^*) and the recalculated points $(\hat{\tau}_i^*, \hat{\eta}_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* \hat{\tau}_i^*)$. This means that equation (17) represents an orthogonal projection (for Q = 1) of the experimental points onto the least square straight line. This is the reason why minimization of sum (13) is called orthogonal regression [4, p. 9]. The classical linear estimator - equation (6) - is associated with a corresponding vertical projection (parallel to the y axis).

BIAS AND VARIANCE OF THE NON LINEAR ESTIMATOR

Since the previous estimator $\hat{\boldsymbol{\alpha}} = \left[\hat{\boldsymbol{\beta}}^t \; \hat{\boldsymbol{\tau}}^t\right]^t$ is not a linear combination of the data (the y's and t's), it will be biased, which means that its expected value will a priori differ from its exact value:

$$E(\hat{\boldsymbol{\alpha}}) \neq \boldsymbol{\alpha} \tag{26}$$

A good way to find an estimate of this bias, as well as an estimate of the standard deviation of the estimator, is to use a second order Taylor expansion of the estimate with respect to the noise. Since two types of noise are now considered (ε for y and δ for t), the previous normalization (based on the two standard deviations σ and σ') will be used. In order to alleviate the notation, the star superscript will be dropped, which corresponds to a use of the dimensional notation with $Q = \sigma = \sigma' = 1$.

Replacing $\boldsymbol{y}$ by $(\boldsymbol{\eta} + \boldsymbol{\varepsilon})$ and $\boldsymbol{t}$ by $(\boldsymbol{\tau} + \boldsymbol{\delta})$ in equation (20b) defining Z and written in a column vector form, and later on in equation (21) defining $\hat{\boldsymbol{\beta}}_2$, will produce such a second order development. To indicate how this procedure can start, the denominator of Z is expanded:

$$s_{ty} = s_{ty}^{(0)} + s_{ty}^{(1)} + s_{ty}^{(2)} + o(\theta^t\theta) \quad \text{with } \theta = \begin{bmatrix} \delta \\ \varepsilon \end{bmatrix} \tag{27}$$

and

$$s_{ty}^{(0)} = s_{\tau\eta} = \beta_2 \, s_\tau^2 \quad ; \quad s_{ty}^{(1)} = s_{\tau\varepsilon} + \beta_2 \, s_{\tau\delta} \qquad s_{ty}^{(2)} = s_{\delta\varepsilon} \tag{28}$$

where the upper subscript (m) designates the m^{th} order term (in ε and δ) in the development of s_{ty} and where notation o (x) designates a term of higher order than x.

Such an approach yields for both parameters $\hat{\beta}_j$ (j = 1, 2):

$$\hat{\beta}_j = \hat{\beta}_j^{(0)} + \hat{\beta}_j^{(1)} + \hat{\beta}_j^{(2)} + o\,(\theta^t\theta) \tag{29}$$

This allows the calculation of the bias b_j and the standard deviation σ_j of $\hat{\beta}_j$ (within the above second order approximation):

$$b_j = E\,(\hat{\beta}_j) - \beta_j = \hat{\beta}_j^{(0)} + E\,(\hat{\beta}_j^{(2)}) - \beta_j \tag{30}$$

$$\text{since } E\,(\hat{\beta}_j^{(1)}) = 0 \text{ and } \sigma_j^2 = \text{var}\,(\hat{\beta}_j^{(1)}) \tag{31}$$

The results for the dimensional parameters can be calculated, taking into account the fact that the two kinds of noise δ and ε are not correlated:

$$b_2 = \frac{\beta_2 / n}{(s_\tau / \sigma')^2} \quad ; \quad b_1 = -\bar{\tau}\,\frac{\beta_2 / n}{(s_\tau / \sigma')^2} \tag{32}$$

$$\sigma_2 = \frac{1}{\sqrt{n}}\,\frac{\left(\beta_2^2 + 1/Q\right)^{1/2}}{s_\tau / \sigma'}$$
$$\sigma_1 = \frac{\sigma}{\sqrt{n}}\left(1 + Q\,\beta_2^2\right)^{1/2}\left(1 + \frac{\bar{\tau}^2}{s_\tau^2}\right)^{1/2} \tag{33}$$

The two standard deviations have limit values, as Q tends to zero, that correspond to the linear case – see equations (8). It can also be noticed that the estimators of the slope and of the intercept are consistent since both their bias and standard deviations tend towards zero for large n values.

These asymptotic values for the standard deviations and bias for the estimators of the slope and of the intercept have been obtained here for any value of n and for small normalized noise over signal ratios (σ/s_η and $\sigma'/s_\tau < 1$) and for noises δ an ε not correlated with the distribution of the exact dependent variable τ. Our standard deviations (33) differ, in principle, from the asymptotic standard deviations given in [4, p. 52-53] which correspond to large samples (large n), for any value of the signal over noise ratios.

The common limiting factor for the four parameters (32, 33) of the probabilistic distributions of the estimators, is the 'signal over noise ratio' (s_τ / σ') of the (deterministic) distribution of the independent variable. If this ratio becomes small, the bias grows faster than the standard deviation and the sought after parameters will not be found. In the example presented in Fig. 1 and 2, this ratio is close to 2 which can be seen in Fig. 2 where the experimental locations are not always an increasing function of the exact location (ordered according to its increasing value): this ordering is nearly blurred by the high standard deviation of the nominal locations (and the too low dispersion of the corresponding exact locations).

Detailed characteristics of this inversion are given in Table 1 where the non linear estimates $\hat{\beta}_j$ are given (see eqs. (19) and (21)), as well as their bias b_j, calculated according to eqs. (32), taking either the exact values of β_2 and s_τ (theoretical) or their counterparts $\hat{\beta}_2$ and s_t that are affected by noise (noised). The bias b_j, either theoretical – see equation (32) – or estimated (noised) are also shown, as well as the linear estimates $\hat{\beta}_{j\,\text{lin}}$, calculated according to equation (7). A corrected estimate, that takes into account the bias – eqs (30) – is also given, with:

$$\hat{\beta}_{j\,\text{cor}} = \hat{\beta}_j - b_j \tag{34}$$

The two components of the modified least square sum, S_t and S_y – see eq. (14) – are also given, as well as the average residuals:

$$res_t = (S_t / n)^{1/2} \quad ; \quad res_y = (S_y / n)^{1/2} \tag{35}$$

In this extreme situation the bias correction has no positive effect and the uncorrected non linear estimates yields better results for β_1 and β_2. This bias correction is not very effective for one single inversion since the bias over standard deviation ratio is rather low. Even if the linear estimation of β_2 is not too bad, it is clear that its S_y residuals do not compare well with the non linear case.

Table 1. Characteristics of the estimation presented in Figs. 1 and 2

j	1	2	
β_j	1	1	$n = 20$
$\hat{\beta}_j$	3.04	0.837	$\sigma = \sigma' = 3$
b_j (theo.)	-0.142	0.0135	$s_\tau / \sigma' = 1.92$
$\hat{b}_j$ (noised)	-0.077	0.239	$S_y = 52.0$ $S_t = 36.4$
σ_j	6.54	0.165	$res_y = 1.61$
$\hat{\beta}_{j\,\mathrm{cor}}$ (noised)	3.130	0.598	$res_t = 1.35$ $S_{y\,\mathrm{lin}} = 145$
$\hat{\beta}_{j\,\mathrm{lin}}$	3.79	0.765	
b_j / σ_j	0.0217	0.0823	

MONTE CARLO VALIDATION OF THE THEORETICAL BIAS AND STANDARD DEVIATION

In order to validate the stochastic estimates of both bias and standard deviation of the estimates, it is necessary to implement a repetitition of the realizations of the simulated measurements by generation of new δ and ε noises, at each realization, before a corresponding inversion. This "crude" Monte Carlo approach [1, p. 125] has been adopted for a number $N = 10000$ realizations in a case where the bias effect is most effective (high value of β_2). The corresponding results are shown in Table 2. If $\hat{\beta}_j^k$ designates the estimates for the k^{th} inversion, the averages and statistical standard deviation for the N inversions are noted:

$$\bar{\hat{\beta}}_j = \frac{1}{N}\sum_{k=1}^{N} \hat{\beta}_j^k \;;\; s_{\hat{\beta}_j} = \left(\frac{1}{N}\sum_{k=1}^{N} \left(\hat{\beta}_j^k\right)^2 - \left(\bar{\hat{\beta}}_j\right)^2 \right)^{1/2} \tag{36}$$

A corrected value for $\bar{\hat{\beta}}_j$ is also given, which takes into account the estimated bias

$$\bar{\hat{\beta}}_{j\,\mathrm{cor}} = \bar{\hat{\beta}}_j - \hat{b}_j \tag{37}$$

This table shows that the bias correction is very effectient and that the statistical standard deviations s match very well their theoretical counterparts σ. The linear estimator is unable to reach the exact parameter values.

Table 2. Statistics on 10000 Monte Carlo simulations of inversion

j	1	2	
β_j	1	100	
$\bar{\hat{\beta}}_j$	-14.0	101.3	
$\bar{\hat{\beta}}_j - \beta_j$	-15.0	1.33	
b_j (theo.)	-14.2	1.35	$n = 20$ points
$\bar{\hat{\beta}}_{j\,\mathrm{cor}}$	1.11	100.0	$\sigma = \sigma' = 3$
$s_{\hat{\beta}_j}$	144.9	12.2	$N = 10000$ realizations
σ_j	139.4	11.6	$s_\tau / \sigma' = 1.92$
$\bar{\hat{\beta}}_{j\,\mathrm{lin}}$	200.4	80.9	
$s_{\hat{\beta}_{j\,\mathrm{lin}}}$	100.1	8.01	
$\sigma_{j\,\mathrm{lin}}$	1.39	0.116	
b_j / σ_j	-0.102	0.116	

In fact these analytical and simulated results show that the bias of the estimator decrease in $1/n$ while the standard deviations of the non linear estimate decrease in $1/n^{1/2}$, which means that in most practical situations the bias over standard deviation ratio (b_j/σ_j) remains low and consequently the bias can be neglected.

CONCLUSION

It has been shown within the simple case of the linear straight line model that both the errors on the ordinates (dependent variables) and on the abscissa (independent variables) could be taken into account by the minimization of a two term functional. The weighing factor, that has to account for the non identical units of these

variables, can be chosen in a unique way, as the ratio $1/Q$ of the variances of the variables.

On a theoretical point of view, this non linear estimator, that relates to the family of the "Measurement Error Models" [4] constitutes a generalization of the classical linear estimator that does not consider errors in the abscissa. Its approximate bias and standard deviation can be analytically expressed thanks to a second order Taylor expansion for high signal over noise ratios. It has been shown that the bias remains negligible in most practical situations. The effect on the estimations of an only approximate knowledge of Q remains to be studied.

On a practical point of view, the use of such a minimization with a penalization term that takes into account a prior but uncertain information on the independent variable, with a non arbitrary weighing term, can be very interesting in the particular case where this variable correspond to the uncertain locations of some temperature sensors. This case will be studied in the second part of this article, in the particular case of the estimation of dispersion coefficients in a porous medium.

Such an optimization technique could also be used in applications where a material is heated by a Joule effect power whose measurements are noised, with simultaneous noised temperature measurements for thermal property estimation (two sources of noise).

REFERENCES

1. J.V. Beck and K.J. Arnold, *Parameter Estimation in Engineering and Science*, John Wiley & Sons, New York, 1977.

2. W.A. Fuller, *Measurement Error Models*, John Wiley & Sons, New York, 1987.

3. G.A.F. Seber and C.J. Wild, *Nonlinear Regression*, John Wiley & Sons, New York, 1988.

4. Chi-Lun Cheng and J.W. Van Ness, *Statistical Regression with Measurement Errors*, Arnold, London, 1999.

5. A. F. Emery, Parameter estimation in the presence of uncertain parameters and with correlated errors, *EUROTHERM Seminar No. 68.*, *Inverse Problems and Experimental Design*, Poitiers (France), march 5-7, 2001.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INTEGRATING THE ERROR IN THE INDEPENDENT VARIABLE FOR OPTIMAL PARAMETER ESTIMATION

PART II: IMPLEMENTATION TO EXPERIMENTAL ESTIMATION OF THE THERMAL DISPERSION COEFFICIENTS IN POROUS MEDIA WITH NOT PRECISELY KNOWN THERMOCOUPLE LOCATIONS

Thomas Metzger, Sophie Didierjean, Denis Maillet
LEMTA – CNRS UMR 7563 – INPL – UHP Nancy 1
Vandoeuvre-lès-Nancy, France
thomas.metzger@ensem.inpl-nancy.fr

ABSTRACT

This experimental work aims to estimate the thermal dispersion coefficients for a packed bed of glass spheres through which water is flowing. Thermocouples in the downstream neighborhood of a linear heat source measure the temperature response to a step heat input. Thermal dispersion coefficients can then be obtained by the classical estimation method of least squares. However, thermocouple locations and water velocity in the set-up are not precisely known, which leads to incorrect estimates. Only one of these two parameters can be estimated additionally by the classical method. A two terms least squares functional, as introduced in part I of this paper, which takes the thermocouple positions as normal stochastic variables can simultaneously give estimates for both parameters. Experimental estimation results for the longitudinal thermal dispersion coefficient are presented for a range of Péclet numbers up to 100.

NOMENCLATURE

a	thermal diffusivity, m^2/s
$\mathbf{a}$, $\mathbf{b}$	estimators for parameters
c_p	heat capacity, J/kg/K
d	particle diameter, m
$\mathbf{I}$	identity matrix
N	number of thermocouples for estimation
Pe	particle Péclet number, ud/a_f
Q	variance ratio, $(\sigma_x/\sigma_T)^2$
$\mathbf{R}$	regularization matrix, see Eq. (11)
S	sum of squares
s	integration variable
T	temperature, K
t	time, s
$\mathbf{u}_D$, u	Darcy velocity, m/s
W	linear heating power, W/m
$\mathbf{x}$	vector defined in Eq. (11)
$\mathbf{X}$	sensitivity matrix
x,y,z	space coordinates, m

Greek symbols

α, β	parameter vectors
ε	porosity
λ, λ_x,…	thermal dispersion tensor, W/m/K
μ	mean value
Ω	weighting matrix
ρ	density, kg/m^3
σ	standard deviation

Subscripts and superscripts

f	fluid phase
s	solid phase
t	total, or transposed vector/matrix
0	equilibrium

INTRODUCTION

The understanding of heat transport in a porous medium through which a fluid is flowing is essential for many chemical engineering applications, for all devices where the exchange or storage of thermal energy is improved by a porous matrix and for the design of nuclear waste storage sites.

The situation is characterized by diffusion in the solid matrix and both convection and diffusion in the fluid phase. Pore geometry and the microscopic velocity field make the problem extremely complicated. Up-scaling techniques such as volume averaging or homogenization allow to describe the phenomenon on a macro-

scopic scale. The simplest model describes the evolution of one average temperature by a convection-diffusion equation with a (velocity-dependent) thermal dispersion tensor characterizing the system [1]. More complete models involve two energy equations for the average temperatures of fluid and solid phases which are coupled by many parameters [2]. In thermal dispersion, fluid velocity is measured by the dimensionless Péclet number which is the ratio of the characteristical times of diffusion and convection on the pore scale. Classically, the one-temperature model was supposed to work only for small Péclet numbers, where diffusion can guarantee local thermal equilibrium ($\langle T_s \rangle = \langle T_f \rangle$). But Moyne et al. [1] have shown that a local average temperature $\langle T \rangle$ can always be defined so that the one-equation model keeps being valid.

The present experimental work is set out to measure the thermal dispersion coefficients of the one-temperature model, and to determine their velocity dependence for Péclet numbers up to 100. Furthermore, the applicability of this simple model is to be tested. In literature, only few experimental results are reported [3-5].

THEORETICAL MODEL

The one-temperature model for the local average temperature $\langle T \rangle$ in a porous medium through which a fluid is flowing is given by

$$(\rho c_p)_t \frac{\partial \langle T \rangle}{\partial t} = \nabla \cdot (\boldsymbol{\lambda} \nabla \langle T \rangle) - (\rho c_p)_f \mathbf{u}_D \cdot \nabla \langle T \rangle \quad (1)$$

where the total heat capacity of the porous medium is calculated by the mixing law

$$(\rho c_p)_t = \varepsilon (\rho c_p)_f + (1-\varepsilon)(\rho c_p)_s \quad (2)$$

and where $\boldsymbol{\lambda}$ is the thermal dispersion tensor and $\mathbf{u}_D$ the Darcy velocity of the fluid (volumetric flux per total cross section). (For simplicity, we will further denote the average temperature by T.)

For the case of an isotropic homogeneous porous medium with fluid flow in x-direction, Eq. (1) simplifies to

$$(\rho c_p)_t \frac{\partial T}{\partial t} = \left(\lambda_x \frac{\partial^2 T}{\partial x^2} + \lambda_y \frac{\partial^2 T}{\partial y^2} + \lambda_z \frac{\partial^2 T}{\partial z^2} \right) - (\rho c_p)_f u \frac{\partial T}{\partial x} \quad (3)$$

where λ_x is the longitudinal, and $\lambda_y = \lambda_z$ the lateral thermal dispersion coefficient; u is the x-component of $\mathbf{u}_D$.

In the experiments we use a linear heat source along the z-axis which is switched on at $t = 0$ to dissipate the power per unit length W. Before excitation, the medium is at thermal equilibrium, i.e. $T = T_0$ everywhere. For an infinite porous medium (with $T \to T_0$ for $x,y \to \infty$) the evolution of the rise in temperature $\Delta T = T - T_0$ can be calculated by successive use of an exponential transform [6] and Green's functions [7] to obtain

$$\Delta T(x,y,t) = \frac{W}{4\pi\sqrt{\lambda_x \lambda_y}} \cdot e^{\frac{(\rho c_p)_f u x}{2\lambda_x}} \cdot \int_0^{\frac{(\rho c_p)_f^2 u^2 t}{4(\rho c_p)_t \lambda_x}} e^{-\left(\frac{x^2}{\lambda_x} + \frac{y^2}{\lambda_y}\right)\frac{(\rho c_p)_f^2 u^2}{16\lambda_x}\frac{1}{s} - s} \frac{ds}{s} \quad (4)$$

This integral is then solved numerically.

EXPERIMENTS

In this section we introduce the set-up used for the experiments and show a typical temperature response diagram.

Experimental set-up

For the experiments, a consolidated porous medium of glass beads with a diameter $d = 2$ mm is used; its porosity ε is about 36%. Water is flowing vertically downwards through this packed bed. Table 1 gives the thermal properties of the experimental system.

Table 1. Thermal properties of the system

	water	glass
ρc_p (kJ/m^3)	4170	2080
λ (W/m/K)	0.607	1

A high-resistance electrical heating wire (diameter 260 µm) is set perpendicular to fluid flow and excited with a step power function. Altogether 13 thermocouples of type E (diameter 127 µm) are used to measure the temperature response. Figure 1 shows the dimensions of the porous medium as well as the locations of the heating wire and the thermocouples.

In a previous work [8] we have shown that the sensitivity of the temperature signal to the thermal dispersion coefficients is highest on the positive

x-axis, and that axial temperature response measurements allow for an estimation of both, longitudinal and lateral, thermal dispersion coefficients; therefore most of the thermocouples (1 to 7) are located there. Thermocouples 8 to 11 are set off-axis to determine the lateral extent of the temperature signal and to verify that wall effects may be neglected (assumption of infinite geometry). The thermocouples upstream of the heat source (12 and 13) are used to monitor T_0 and to assess the effect of *natural convection* for very small Péclet numbers. (Only for Pe ≤ 1, thermocouple 12 measures a change in temperature. But velocity estimation, as described later, seems to show that natural convection lowers water velocity for Péclet numbers up to 10.)

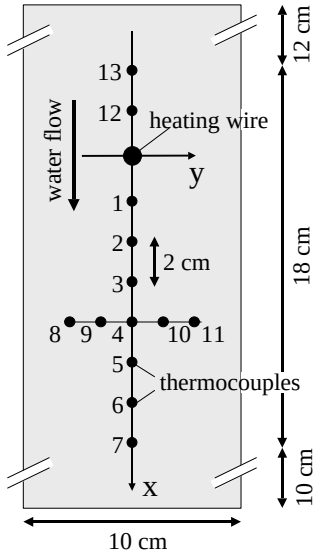


Figure 1. Experimental set-up: dimensions of porous medium (20 cm in *z*-direction), heating wire and thermocouple positions and numbers

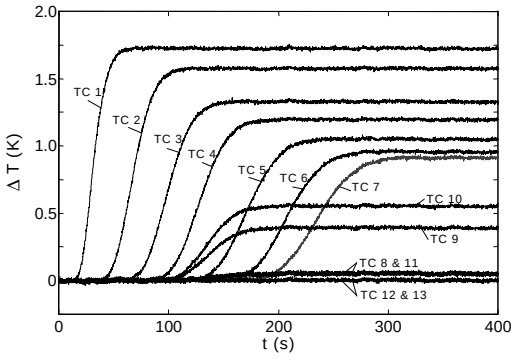


Figure 2. Temperature response signal (Pe ≈ 6)

The heating power level needs to be chosen in a way that temperature differences stay small enough, so that all physical properties can be taken as constant and boiling is excluded in the immediate surroundings of the heating wire, but that the signal to noise ratio is high enough for the temperature measurements. Generally, it is set such that ΔT takes values of the order of 1 K at the thermocouple positions. The thermoelectrical voltage of the thermocouples (between a thermally insulated reference junction and the measurement junction) is amplified 2000 times and measured with a frequency of 2 KHz. To reduce electric noise, we use a first-order low pass filter of 44 Hz and average data groups of 80 points (corresponding to 2 periods of 50 Hz noise!). This allows us to record temperature at an effective frequency of 8 Hz with a noise of about 0.01 K. Additionally, any offset of the data acqui-sition system or the thermocouples is corrected by the signals measured at $t < 0$. Heating power is measured electrically.

Average water velocity is monitored by a flowmeter outside the porous medium. This introduces a systematical error because of the channeling effect near the box's walls: typically, velocity in the center of the packed bed will be a few percent less than the average value [9].

It also has to be mentioned, that the exact locations of heating wire and thermocouples are not known, since the glass beads are filled in the box after the wires have been set: they probably displace them, typically by the diameter of a sphere, and visual estimation of their position is not possible anymore.

Temperature response diagrams

Experiments were carried out in the range of Péclet numbers up to 100 which corresponds to Darcy velocities of about 7 mm/s. For each water velocity, thermal equilibrium of the porous medium was established, then the heating was switched on and all temperature response signals where recorded. Figure 2 gives an example for one such experiment at low water velocity. From these response curves, the thermal dispersion coefficients λ_x and λ_y are to be estimated, which constitutes an inverse problem. But imprecise knowledge of thermocouple locations and water velocity needs to be accounted for.

PARAMETER ESTIMATION

Ordinary least squares estimator

The classical method of parameter estimation involves the minimization of the sum of squares

$$S_T(\boldsymbol{\beta}) = \sum_i \sum_k \left(\mathbf{T}_{exp,ik} - \mathbf{T}(x_i, y_i, t_k, \boldsymbol{\beta})\right)^2 \qquad (5)$$

where $\mathbf{T}_{exp}$ are the experimental temperature values, for positions (x_i, y_i) and times t_k, arranged in a vector; $\mathbf{T}$ is the corresponding theoretical temperature vector, depending on parameters $\boldsymbol{\beta}$. The Gauss algorithm [10] given in Eq. (6) leads to the ordinary least squares estimator $\mathbf{b}_{OLS}$ for the searched parameters $\boldsymbol{\beta}$.

$$\mathbf{b}_{n+1} = \mathbf{b}_n + \left(\mathbf{X}(\mathbf{b}_n)^{\mathrm{t}}\,\mathbf{X}(\mathbf{b}_n)\right)^{-1}\mathbf{X}(\mathbf{b}_n)^{\mathrm{t}}\left(\mathbf{T}_{exp} - \mathbf{T}(\mathbf{b}_n)\right) \qquad (6)$$

Here $\mathbf{X}_{mj} = (\partial \mathbf{T}_m / \partial \boldsymbol{\beta}_j)$ is the sensitivity matrix, the columns of which are partial derivatives of the theoretical temperature [10].

The parameter vector $\boldsymbol{\beta}$, which is to be estimated, necessarily contains the thermal dispersion coefficients λ_x and λ_y. In [8] we showed that it may be extended to include *either* the Darcy velocity u *or* the thermocouple locations x_i. However, both of them cannot be estimated simultaneously: then the parameters become correlated. For central measurement (y = 0), neither locations in y-direction nor the heating power W can be estimated. This leaves us with the following three estimation modes:

$$\begin{aligned} \boldsymbol{\beta}_1 &= [\lambda_x \;\; \lambda_y]^{\mathrm{t}} \\ \boldsymbol{\beta}_2 &= [\lambda_x \;\; \lambda_y \;\; u]^{\mathrm{t}} \\ \boldsymbol{\beta}_3 &= [\lambda_x \;\; \lambda_y \;\; x_i]^{\mathrm{t}} \end{aligned} \qquad (7)$$

(Note, that in $\boldsymbol{\beta}_3$-mode thermocouple position x changes its role from known independent variable to estimated parameter.) With these three modes the limits of the classical approach are reached.

Regularization

Both Darcy velocity and thermocouple positions can be estimated, if the formalism of part I of this article is used. Let us therefore define

$$\boldsymbol{\alpha}_4 = [\lambda_x \;\; \lambda_y \;\; u \;\; x_i]^{\mathrm{t}} \qquad (8)$$

as the parameters to be estimated. Starting from estimation mode $\boldsymbol{\beta}_2$ and applying the formalism to the independent variable x, one obtains the modified sum of squares S given in Eq. (9).

The sum of squares for the temperature signal is hence completed by a second sum which contains the thermocouple positions. Here the nominal positions x_i^{nom} are the positions of the thermocouples before the beads were filled in the box. This filling randomly displaces the thermocouples. The nominal positions are treated as a measurement signal for they are the most accurate information available. In order to combine the two sums, a proper ponderation is needed. The sum for the temperatures is weighed by the standard deviation of temperature measurement noise σ_T, which is obtained by data acquisition at thermal equilibrium. The statistical properties of the positions, however, are not so easily available and a typical displacement of the order of the particle diameter will be taken for σ_x. (Later we will see that the estimation results are only affected if σ_x is changed several orders of magnitude.) By introducing this additional information on the positions, x_i^{nom} and σ_x, the estimation of both velocity and positions becomes possible. When estimating the parameters, the second sum will penalise a large deviation of the estimated positions from the nominal values.

$$\begin{aligned} S(\boldsymbol{\alpha}_4) &= \frac{1}{\sigma_T^{\,2}} S_T + \frac{1}{\sigma_x^{\,2}} S_x \\ &= \frac{1}{\sigma_T^{\,2}} \sum_i \sum_k \left(\mathbf{T}_{exp,ik} - \mathbf{T}(x_i, y_i, t_k, \boldsymbol{\beta}_2)\right)^2 \\ &\quad + \frac{1}{\sigma_x^{\,2}} \sum_i \left(x_i^{nom} - x_i\right)^2 \end{aligned} \qquad (9)$$

The appropriate algorithm to minimize the modified sum of squares S and obtain an estimator for $\boldsymbol{\alpha}_4$ is given in two alternative forms in Eqs. (10). The signal vector has been extended to include the nominal positions x_i^{nom}, the sensitivity matrix has been extended accordingly to $\mathbf{X}_{ext}$, and a weighting matrix $\boldsymbol{\Omega}$ with the two weighting factors on its diagonal is introduced [10]:

$$\mathbf{a}_{n+1} = \mathbf{a}_n + \left(\mathbf{X}_{ext}(\mathbf{a}_n)^{\mathrm{t}}\,\boldsymbol{\Omega}\,\mathbf{X}_{ext}(\mathbf{a}_n)\right)^{-1} \left(\mathbf{X}_{ext}(\mathbf{a}_n)^{\mathrm{t}}\,\boldsymbol{\Omega}\right)\left(\begin{bmatrix}\mathbf{T}_{exp}\\ \mathbf{x}^{nom}\end{bmatrix} - \begin{bmatrix}\mathbf{T}(\mathbf{a}_n)\\ \mathbf{x}_n\end{bmatrix}\right) \qquad (10a)$$

$$= \mathbf{a}_n + \left(\mathbf{X}(\mathbf{a}_n)^{\mathrm{t}}\mathbf{X}(\mathbf{a}_n) + \frac{1}{Q}\mathbf{R}^{\mathrm{t}}\mathbf{R}\right)^{-1} \left(\mathbf{X}(\mathbf{a}_n)^{\mathrm{t}}\left(\mathbf{T}_{exp} - \mathbf{T}(\mathbf{a}_n)\right) + \frac{1}{Q}\mathbf{R}^{\mathrm{t}}\left(\mathbf{x}^{nom} - \mathbf{x}_n\right)\right) \qquad (10b)$$

$$\text{where } \mathbf{x} = [x_1 \ldots x_N]^{\mathrm{t}},\ \mathbf{X}_{ext} = \begin{bmatrix}\mathbf{X}\\ \mathbf{R}\end{bmatrix},$$

$$\mathbf{R} = [\mathbf{0} \quad \mathbf{0} \quad \mathbf{0} \quad \mathbf{I}_N] \text{ and } Q = \frac{\sigma_x^2}{\sigma_T^2} \qquad (11)$$

The matrix **R** is the sensitivity matrix for the position signal only, but in Eq. (10b) its role as a regularization matrix – in combination with the variance ratio Q – is stressed. N is the number of thermocouples and $\mathbf{I}_N$ the identity matrix of dimension $N \times N$. The structure of the vectors $\mathbf{x}^{nom}$ and $\mathbf{x}_n$ is defined generally.

This algorithm is a generalization of Eq. (6). As can easily be seen, the additional terms vanish for $\sigma_x \to \infty$, as does the additional sum S_x (in this limit, the algorithm breaks down; it corresponds to a classical approach for estimating α_4).

Monte Carlo comparison of estimation modes

In order to compare this new estimation procedure (α_4) with the classical modes (β_i) and to assess its quality, experimental temperature data were simulated by the Monte Carlo method [10] for a model porous medium ($(\rho c_p)_s$ = 2 kJ/m^3 and $\varepsilon = 0.4$) with water flowing at a Darcy velocity of

$$u^{nom} = 1 \text{ mm/s} \qquad (12a)$$

(corresponding to Pe = 14). Linear heating power was chosen to be Q = 75 W/m, and for the thermal dispersion coefficients approximate values were taken as

$$\begin{aligned} \lambda_x &= 5 \text{ W/m/K} \\ \lambda_y &= 2 \text{ W/m/K} \end{aligned} \qquad (12b)$$

The nominal thermocouple positions were

$$\begin{aligned} (x_1,y_1)^{nom} &= (2,0) \text{ cm} \\ (x_2,y_2)^{nom} &= (4,0) \text{ cm} \\ (x_3,y_3)^{nom} &= (6,0) \text{ cm} \end{aligned} \qquad (12c)$$

and the times of temperature measurement t_k = 0.5, 1.0, 1.5, …, 120 s. The data points T_{ik} were generated by adding uncorrelated normal noise with a standard deviation σ_T = 0.1 K. All these values are taken rather pessimistically compared to what can be achieved in experiments in order to show that the algorithm is robust.

Four cases, as defined in Table 2, were treated to investigate the influences of the channeling effect and thermocouple displacement by filling in the glass beads. First, temperature data were simulated for the nominal velocity and positions; second, water velocity was taken 5% less than the nominal value for the simulations; third, thermocouple positions were changed by normal-distributed uncorrelated δx_i ($\mu_x = 0$, σ_x = 1 mm); and last, velocity *and* positions were both modified. (The absence of correlation in position error cannot be checked experimentally).

Table 2. Definition of the four cases of Monte Carlo simulation, and results for the different estimation modes as mean values and standard deviations of estimated parameters

estimation mode ↓	parameters used for signal generation →	case 1 u^{nom} x_i^{nom}	case 2 0.95 u^{nom} x_i^{nom}	case 3 u^{nom} $x_i^{nom} + \delta x_i$	case 4 0.95 u^{nom} $x_i^{nom} + \delta x_i$
β_1	λ_x (W/m/K)	5.004 ± 0.375	5.729 ± 0.373	5.027 ± 0.369	5.795 ± 0.397
	λ_y (W/m/K)	1.999 ± 0.015	1.922 ± 0.013	1.985 ± 0.067	1.907 ± 0.064
β_2	λ_x (W/m/K)	5.007 ± 0.383	4.971 ± 0.344	5.025 ± 0.393	5.078 ± 0.345
	λ_y (W/m/K)	1.998 ± 0.018	2.000 ± 0.016	1.983 ± 0.085	1.981 ± 0.085
	u (mm/s)	1.001 ± 0.007	0.950 ± 0.006	1.002 ± 0.015	0.953 ± 0.014
β_3	λ_x (W/m/K)	4.998 ± 0.378	5.499 ± 0.369	4.966 ± 0.355	5.548 ± 0.351
	λ_y (W/m/K)	2.001 ± 0.025	1.806 ± 0.020	1.998 ± 0.025	1.806 ± 0.023
	Δx_1 (mm)	–0.017 ± 0.198	1.043 ± 0.205	0.016 ± 0.206	1.036 ± 0.218
	Δx_2 (mm)	–0.025 ± 0.376	2.079 ± 0.320	0.024 ± 0.382	2.069 ± 0.401
	Δx_3 (mm)	–0.020 ± 0.497	3.142 ± 0.507	0.040 ± 0.460	3.118 ± 0.477
α_4	λ_x (W/m/K)	5.002 ± 0.383	4.968 ± 0.344	4.970 ± 0.380	5.025 ± 0.345
	λ_y (W/m/K)	1.999 ± 0.021	2.000 ± 0.021	1.996 ± 0.055	1.995 ± 0.056
	u (mm/s)	1.000 ± 0.007	0.950 ± 0.006	1.000 ± 0.014	0.952 ± 0.013
	Δx_1 (mm)	–0.008 ± 0.150	–0.001 ± 0.171	0.025 ± 0.306	0.027 ± 0.314
	Δx_2 (mm)	–0.007 ± 0.265	–0.009 ± 0.238	0.039 ± 0.564	0.035 ± 0.582
	Δx_3 (mm)	0.005 ± 0.191	0.005 ± 0.178	0.049 ± 0.811	0.049 ± 0.786

Table 3. Results of parameter estimation of 1000 MC simulations for fixed positions as in Eq. (13)

estimation mode		case 3a u^{nom} $x_i^{nom} + \delta x_i$	case 4a 0.95 u^{nom} $x_i^{nom} + \delta x_i$
β_1	λ_x (W/m/K)	5.130 ± 0.406	5.947 ± 0.407
	λ_y (W/m/K)	2.022 ± 0.016	1.975 ± 0.015
β_2	λ_x (W/m/K)	5.154 ± 0.414	5.146 ± 0.370
	λ_y (W/m/K)	2.052 ± 0.019	2.053 ± 0.018
	u (mm/s)	1.002 ± 0.007	0.952 ± 0.006
β_3	λ_x (W/m/K)	4.985 ± 0.392	5.538 ± 0.389
	λ_y (W/m/K)	2.000 ± 0.024	1.806 ± 0.022
	Δx_1 (mm)	–0.007 ± 0.208	1.100 ± 0.226
	Δx_2 (mm)	0.009 ± 0.361	2.099 ± 0.347
	Δx_3 (mm)	0.006 ± 0.495	3.110 ± 0.494
α_4	λ_x (W/m/K)	5.066 ± 0.403	5.079 ± 0.362
	λ_y (W/m/K)	1.971 ± 0.022	1.971 ± 0.022
	u (mm/s)	1.008 ± 0.007	0.958 ± 0.006
	Δx_1 (mm)	0.119 ± 0.176	0.118 ± 0.177
	Δx_2 (mm)	0.303 ± 0.262	0.316 ± 0.238
	Δx_3 (mm)	0.528 ± 0.195	0.521 ± 0.179

For each case, 1000 measurements were simulated, choosing new thermocouple positions (for case 3 and 4) at each simulation. Then the four estimation modes were tested on these simulated measurements, using the nominal values for velocity and positions. In the modes, where these are not estimated, a systematical error is committed. The results for the estimated parameters are given in Table 2 as mean values and standard deviations for one estimation. Position estimation is described by mean value and standard deviation of the differences $\Delta x_i \equiv x_i^{estimated} - x_i$. Hence for a good estimation, thermal dispersion coefficients should be found as in Eq. (12b), velocity as u^{nom} (case 1 and 3) or 0.95 u^{nom} (case 2 and 4), and Δx_i close to zero.

As a first result, case 1 shows that there is no loss in accuracy if more than just the thermal dispersion coefficients are estimated. If velocity is not well known (case 2), β_2 and α_4 cope equally well, leaving only a small bias on the estimated λ_x. In case 3, randomly displaced thermocouple positions are well estimated by both β_3 and α_4. Case 4, closest to reality, clearly shows the suitability of α_4, which gives the best results, with nearly no bias on all parameters. It combines the advantages of β_2 and β_3, loosing a little in position accuracy. Here, emphasis has to be put on the *reduction* of uncertainty of position: the standard deviation of Δx_i as compared to σ_x = 1mm.

It has to be mentioned that in a real experimental set-up thermocouple positions are determined only once when filling the box, while measurement noise will be different for each experiment carried out. Hence, strictly speaking, the results of Table 2 decribe the statistics of many box fillings.

For this reason, we pick out one situation with the positions as given in Eq. (13) and simulate 1000 measurements without changing them.

$$\begin{aligned} x_1 &= 2.11 \text{ cm} \\ x_2 &= 3.96 \text{ cm} \\ x_3 &= 5.91 \text{ cm} \end{aligned} \qquad (13)$$

Hence we get case 3a for the nominal velocity and case 4a for the biased one.

The results in Table 3 show that in α_4-mode for both cases all position estimates are biased towards positive x-values (and velocity is overestimated for compensation); the mean position values are

$$\begin{aligned} x_1^{estimated} &= 2.12 \text{ cm} \\ x_2^{estimated} &= 3.99 \text{ cm} \\ x_3^{estimated} &= 5.96 \text{ cm} \end{aligned} \qquad (14)$$

which describes reality much better than the nominal values of Eq. (12c). In case 4a, α_4 gives better estimation results for the thermal dispersion coefficients than the β_i-modes. Residuals also are best for α_4 (nearly uncorrelated), whereas those of β_1 and β_2 contain artifacts due to wrong positions (not shown here).

So α_4 proved to be the best estimation mode, be it for one particular or a whole ensemble of set-ups (see Table 2).

EXPERIMENTAL ESTIMATION

This new mode shall now be used for thermal dispersion estimation from the temperature data of thermocouples 2 to 7 as shown in Fig. 2.

Influence of σ_x

The difference of real experimental data to the simulations of the previous section is that the standard deviation σ_x of the thermocouple positions is not known, but can only be guessed. However, a definite value needs to be chosen for the estimation procedure in Eq. (10). (σ_T can be measured in thermal equilibrium.)

If the estimation results are plotted as a function of σ_x (see Figure 3), one can clearly distinguish two domains: for σ_x < 0.1 mm the esti-

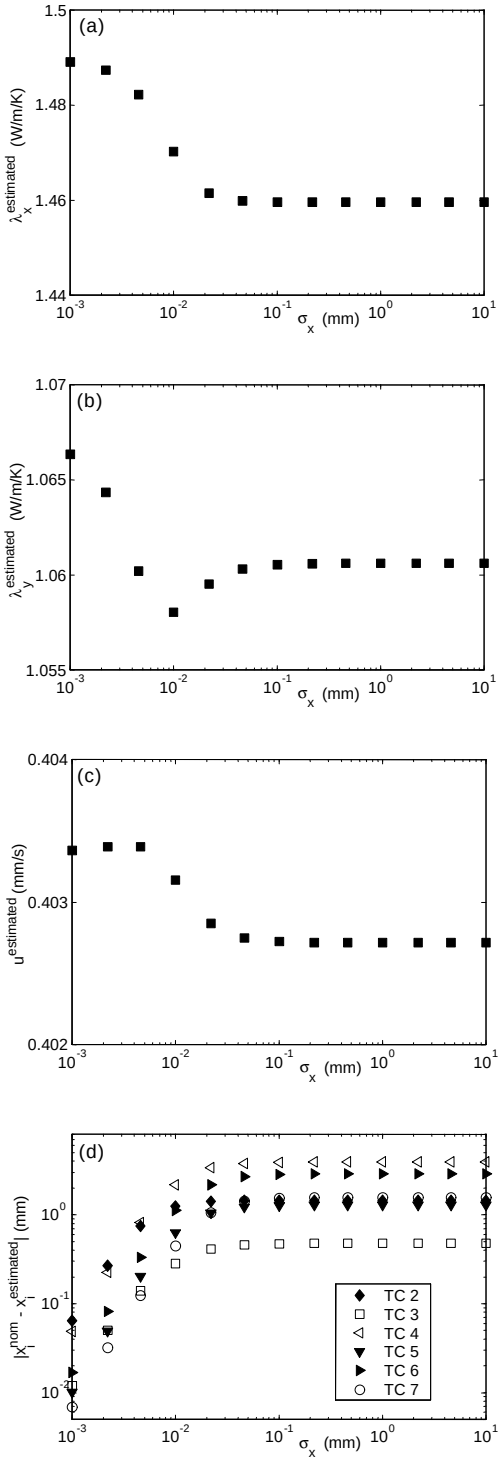


Figure 3. Estimation results for data of Fig. 2 as a function of σ_x. The nominal value in (c) is u^{nom} = 0.436 mm/s. In (d), $x_i^{nom} - x_i^{estimated}$ is positive for i = 2,5,6 and negative for the rest.

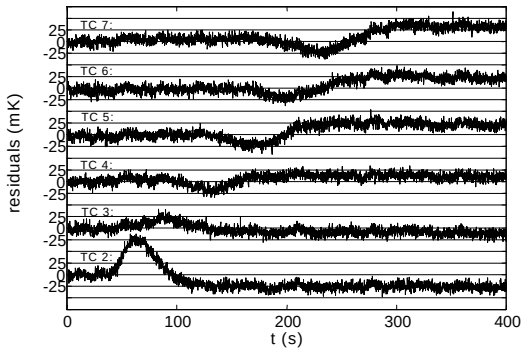


Figure 5. Residuals of parameter estimation corresponding to data of Fig. 2

mated parameter values depend on σ_x (in the limit $\sigma_x \to 0$ they approach the results for β_2). For $\sigma_x >$ 0.1 mm, parameter estimation is independent on the choice of σ_x, (for $\sigma_x \to \infty$ the algorithm breaks down). As pointed out in the previous section, σ_x should be of the order of a particle diameter of the packed bed, hence in the millimeter range. This means that any physically reasonable value in the second domain is safe to be chosen for the estimation algorithm.

Experimental results

For the exemplary measurement of Fig. 2 the estimator for α_4 was calculated with the algorithm in Eq. (10) and σ_x = 1 mm. The result is shown in Fig. 4 as largely magnified residuals that are the remaining differences between experimental and best theoretical temperatures. The measurement noise of less than 0.01 K can be seen as well as a remaining structure which suggests that the physical modeling of the problem can still be improved.

In a similar way, thermal dispersion parameters have been estimated from the experimental temperature signals in the whole range of Péclet numbers. The results for the longitudinal dispersion coefficient λ_x are given in Fig. 6 (the arrow shows the example of Fig. 2). As common in literature, we give the velocity-dependence of λ_x as a power law (the two coefficients were obtained by classical nonlinear least squares):

$$\lambda_x = \lambda_{eq} + 0.0414\ \mathrm{Pe}^{1.62}\ \mathrm{W/m/K} \qquad (15)$$

The equivalent thermal conductivity λ_{eq} = 0.837 W/m/K was not estimated but calculated by the model of Zehner [11]. The reason for not showing the lateral thermal dispersion coefficient is given in the next section.

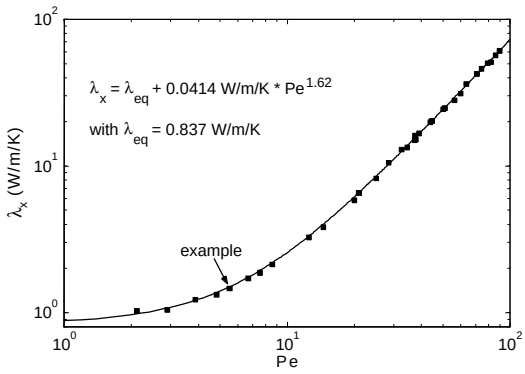


Figure 6. Longitudinal thermal dispersion coefficient as a function of Péclet number

REMAINING CHALLENGES AND CONCLUSION

The experimental results for the longitudinal thermal dispersion coefficient are already very promising. The estimated values have little noise and yield a power law, which corresponds well to what is measured or expected by other research groups.

However, we have so far neglected any uncertainty in thermocouple coordinates y_i. Of course, this is no realistic formulation of the problem, and introduces bias effects. In fact, for increasing Péclet number, the experimental temperature data become more difficult to fit with the theoretical model used, and the structure in the residuals becomes more important. We could already show that an error in y_i can explain these effects. Also, we could see that this only affects the estimation result for λ_y, which is overestimated, and that λ_x is practically left unchanged. This is why we cannot give any definitive results for λ_y yet.

The experimental set-up and the estimation procedure are currently modified in order to take into account this error in *y*-direction. We do this by slightly setting the thermocouples off-axis ($y_i \neq 0$), so that the sensitivity of the temperature signal to the *y*-location is not zero any more. First modifications of the estimation procedure indicate that with this change in set-up design it should be possible to estimate the positions in *y*-direction as well.

In addition, the possibilities of a plane heat source, realized by several heating wires, are explored. This might give an alternative to estimate the Darcy velocity in the center of the packed bed and verify the values obtained in the linear heating mode.

REFERENCES

1. C. Moyne, S. Didierjean, H.P. Amaral Souto, O.T. da Silveira, Thermal dispersion in porous media: One-equation model, *Int. J. Heat Mass Transfer*, **43**, 3853 (2000).
2. M. Quintard, M. Kaviany, S. Whitaker, Two-medium treatment of heat transfer in porous media: numerical results for effective properties, *Advances in Water Resources*, **20**, 77 (1997).
3. D. Vortmeyer, W. Adam, Steady-state measurements and analytical correlations of axial effective thermal conductivities in packed beds at low gas flow rates, *Int. J. Heat Mass Transfer*, **27**, 1465 (1984).
4. D.W. Green, R.H. Perry, R.E. Babcock, Longitudinal dispersion of thermal energy through porous media with a flowing fluid, *AIChE Journal*, **10**, 645 (1964).
5. J. Levec, R.G. Carbonell, Longitudinal and lateral thermal dispersion in packed beds, *AIChE Journal*, **31**, 581 (1985).
6. M.N. Özisik, *Heat Conduction*, John Wiley & Sons, New York, 1980, p. 74-75.
7. J.V. Beck, K.D. Cole, A. Haji-Sheikh, B. Likouhi, *Heat Conduction Using Green's Functions*, Hemisphere Publishing Corporation, London, 1992.
8. T. Metzger, S. Didierjean, D. Maillet, Experimental design for thermal dispersion estimation in porous media, *EUROTHERM Seminar No. 68*, Poitiers, France (2001).
9. M. Winterberg, *Modellierung des Wärme- und Stofftransports in durchströmten Festbetten mit homogenen Einphasenmodellen*, Fortschritt-Berichte VDI, Reihe 3, Nr. 654, 2000.
10. J.V. Beck, K.J. Arnold, *Parameter estimation in engineering and science*, John Wiley & Sons, New York, 1977.
11. P. Zehner, E.-U. Schlünder, Wärmeleitfähigkeit von Schüttungen bei mäßigen Temperaturen, *Chemie-Ing.-Techn.*, **14**, 933 (1970).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

DESIGN OF EXPERIMENTS TO ESTIMATE THERMAL PROPERTIES AND BOUNDARY CONDITIONS SIMULTANEOUSLY

Mikhail Romanovski
CAD/CAE Department, POINT Ltd.
79-1-334, Schelkovskoe shosse, 107497, Moscow, Russia
mromanovski@netscape.net
http://mywebpage.netscape.com/mromanovski/IP.htm

ABSTRACT

The goal is the condition analysis to estimate a maximum of unknown quantities with limited volume of observation data. Simultaneous identification of thermal properties and boundary conditions are based on the uniqueness solution of inverse problems and optimal design of measurements. A number of new theoretical problems are solved. The main result is the substantiation of the existence of the maximal informativeness of experimental data processing even with very small volume of observation data.

NOMENCLATURE

a	Vector of unknown properties
c	Specific heat coefficient
f	Volumetric heat sources
ℓ	Specimen length
m	Number of samples
n	Sample volume
p	Number of sought properties
t	Time
t_h	Heat flux duration
u	Temperature field
u^{δ}	Sampling
u_0	Initial temperature
$v_{1,2}$	Boundary temperatures
x	Spatial variable
A	Domain of admissible solutions
$\mathcal{R}$	Indeterminacy power of identification
δ	Absolute error of measurements
ε	Noise of measurements
ρ	Specimen density
λ	Thermal conductivity coefficient
ϑ	Dimensionless temperature
μ	Relative identification error
$\mu^{(C)}$	*C*-optimal relative estimation error
$\mu^{(R)}$	*R*-optimal relative estimation error
ν_{abs}	Absolute estimation error
ν_{rms}	Root-mean-square estimation error
Φ	Form-factor of estimation error
Ξ	Observation design

INTRODUCTION

In the theory of inverse problems many authors studied the conditions to identify more than one model parameter. Similar investigations were highly stimulated by the study on the simultaneous identification of specific heat and thermal conductivity coefficients [1-5]. The results obtained indicate that several estimation methods can be suggested, which choice depends on the volume of the initial information. Together with these investigations the uniqueness solution of inverse problems and identifiability of mathematical models were examined [6-10]. Such studies are aimed to substantiate the correct mapping of the sought functions on an observation domain.

In contrast to these studies, we are dealing with the extension of the number of desired quantities as much as possible. From a theoretical viewpoint we want to determine the upper bound of the useful information on the input signal components containing in a received signal and unambiguously reconstructed during observation data processing. From a practical viewpoint the problem under study is regarded with the simultaneous estimation as test object properties and its boundary conditions.

Namely, the minimal volume of input data received from the sole experiment and single sensor observation are the subject of our investigation. Substantially, we want to answer the next question. *Is it possible to analyze a test object state on a whole using observation data received from a single interior point of a specimen*?

The positive answer expands the formulation of the experimental data processing and gives ample opportunities to identify many properties without numerous measurements. In addition, the approach offered elicits the most informative measurements and best experimental conditions.

EXPERIMENTAL DESIGN METHOD

The results obtained in [11] attest that a sole experiment can evince the information on all the test object properties. Now, let us expand the investigation and determine the most informative measurements at a given experiment. Namely, we will seek the optimal one sensor location and minimal number of measurements on time that guarantee the inverse problem solution with maximal precision grounded on the sole experiment.

To solve the design experimental problem we use the method [12]. The one is based methodologically on the theory of ill-posed problems. The approach developed formulates the experimental design from a standpoint of the error analysis of the inverse problem solution.

According to the method proposed the estimation error $\nu = \|a - \overline{a}\|$ with respect to the given $\overline{a}$ and noise level δ of measurements is sought as the solution of the observation matching equation

$$\left\| u_i^{\delta} - u_i\big|_{a^{(\nu)}} \right\| = \delta_i, \quad i = \overline{1, m}, \qquad (1)$$

where the functions $u_i\big|_{a^{(\nu)}}$ are determined as a direct problem solution at a given observation design $\{\Xi_i\}_{i=\overline{1,m}}$ subjected that the solution $a^{(\nu)}$ minimizes the stabilizing functional $\Omega[a]$ on a set whose elements differ from the true value $\overline{a}$ at most by ν:

$$a^{(\nu)} = Arg \inf_{a \in A_\nu} \Omega[a], \quad A_\nu = \{a : \|a - \overline{a}\| \le \nu\} \quad (2)$$

An essential feature of formulation (1), (2) is the type and character of the estimation error. They are determined as *a guaranteed error* [12]. The one expresses an operation of the worse measurement error. Consequently, the solution of the experimental design problem conveys the evaluation with the worse measurement conditions. The further details of the approach developed are considered in [13, 14].

Special significance has the fact that the estimation methods without sufficient regularization and matching with observation give rise to the large errors despite a stability of a solution obtained [13]. Therefore, the guaranteed estimation error is not a substantiation of any method exactitude, but generally evinces satisfactory identification conditions.

To specify the numerical criterion of the guaranteed estimation error we introduce the *rms* estimator $\nu_{rms} = \sqrt{\sum_{k=1}^{p} \nu_k^2 / p}$, where ν_k denotes the absolute estimation error of k –th property. Then the experimental design consists in the finding of the minimization problem solution $\nu^{(R)} = \min_{\Xi} \nu_{rms}$. The observation design $\Xi^{(R)}$ ensuring the minimal estimation error $\nu^{(R)}$ is said to be *R-optimal* [12].

In addition to the *rms* estimator other norms of the estimation error can be considered. Here the most practical form could be the absolute-error estimator. By analogy with the preceding case, the corresponding design $\Xi^{(C)}$ is said to be *C-optimal* [12]. The one is based upon the estimator $\nu_{abs} = \max_{1 \le k \le p} |\nu_k|$. Its minimization is the purpose of the experimental design. In this case the *min-max* estimate $\nu^{(C)} = \min_{\Xi} \nu_{abs}$ is sought.

The main difference between the experimental design method developed and traditional approach grounded on the sensitivity analysis is expressed as the transition to the direct study of the estimation errors. The approach based on the Fisher matrix can determine the optimal sensor location and conditions of the linear dependence between the sensitivity coefficients. Formulation (1), (2) ensures the determination of the error distribution on each point of the specimen. Also, the existence and uniqueness of the inverse problem solution are analyzed not only with the fixed initial data, but on a wide range of their values. As a result the explicit picture of the estimation error behavior is elicited. Due to this fact one can determine new singularities of the identification. Theoretically, the sensitivity analysis can reveal many features of the experimental design. However, such singularities as δ-unidentifiability, threshold and singular conditions, self-compensation loading have been revealed in [14], grounding on the approach (1), (2).

An important feature of formulation (1), (2) is the structural decomposition of experimental conditions from a viewpoint of their effect on the estimation error behavior. As it was indicated [14], there are two types of factors that comprise all the experimental conditions and completely convey their effect. The indeterminacy power $\mathcal{R}$ of the identification and form-factor Φ of the estimation errors are these factors. The first one determines the condition due to the estimation errors are decreased to zero. The second one shapes the estimation error distribution.

So, our solution of the experimental design problem refers two main goals. Namely, the optimal measurement scheme and its dependence upon the experimental conditions will be brought to a fixed proportion. As a result, the regularities of the optimal identification will be revealed.

MAXIMUM ESTIMATES WITH MINIMAL INPUT INFORMATION

Let us analyze typical thermal experiments from a viewpoint of the maximal informativeness being grounded on the minimal volume of input data.

First of all, we will investigate the cases of the constant parameter models. This will allow us to carry out the analytical treatment of the experiment informativeness and therefore the general regularities of the problem studied will be revealed. The nonlinear properties and transient boundary conditions can be also analyzed by the offered approach. These cases will be investigated after the general regularities determination.

Simultaneous Estimation of Thermal Properties and Boundary Temperatures

Let us specify the following model

$$\begin{aligned} &c\rho\frac{\partial u}{\partial t}=\lambda\frac{\partial^2 u}{\partial x^2}+f,\quad 0<x<\ell, t>0;\\ &u\big|_{t=0}=u_0,\ 0<x<\ell;\\ &u\big|_{x=0}=v_1,\ u\big|_{x=\ell}=v_2,\quad t>0 \end{aligned} \tag{3}$$

It is required to establish the existence and features of the optimal observation design $\Xi^{(R)}=\{x_i^{(opt)},t_j^{(opt)}\}_{i=1}^{j=\overline{1,n}}$ with one observation point and n measurement times, for which the known sampling

$$u_{i,j}^{\delta}=\bar{u}(x_i,t_j)+\varepsilon_j,\quad i=1,\ j=\overline{1,n}$$

provides the identification of the constant unknowns $a=\{c,\lambda,v_1,v_2\}$ with the minimal guaranteed *rms*-error

$$\mu^{(R)}=\min_{x,t}\mu_{rms},\quad \mu_{rms}=\sqrt{\frac{\mu_1^2+\mu_2^2+\mu_3^2+\mu_4^2}{4}}.$$

The relative estimation errors

$$\mu_1=(\bar{c}-c)/\bar{c},\ \mu_2=(\bar{\lambda}-\lambda)/\bar{\lambda},$$
$$\mu_3=(\bar{v}_1-v_1)/\bar{v}_1,\ \mu_4=(\bar{v}_2-v_2)/\bar{v}_2$$

are the functions of the variables x and t, $\bar{a}=\{\bar{c},\bar{\lambda},\bar{v}_1,\bar{v}_2\}$ denotes the actual values of sought quantities.

The solution of the experimental design problem is sought with the assumption that the upper bound of measurement errors is known, $\max_{1\le j\le n}|\varepsilon_j|\le\delta$. A similar estimate of the noise level can be obtained as a rule for every experiment.

We especially underline that any other information about the noise properties is not required and not used below. In particular, the law of the noise distribution may be arbitrary. Below it will be proved that such volume of initial information is sufficient to solve the inverse problem.

Formulation (1), (2) gives rise to the determination of $\{\mu_k\}_{k=\overline{1,4}}$ and requires the analysis of their behavior over the variation of the thermal loading and measurement conditions. The guaranteed estimation errors $\{\mu_k\}_{k=\overline{1,4}}$ are defined as the solution of the observation matching equation

$$\max_{x_i,t_j}\left|u_{i,j}^{\delta}-u(x_i,t_j)\big|_{a^{(\nu)}}\right|=\delta.$$

The functions $u\big|_{a^{(\nu)}}$ are determined from model (3) at the given observation design Ξ subject to the condition that the vector of unknowns $a=\{c,\lambda,v_1,v_2\}$ minimizes the stabilizing functional $\Omega[a]=\max|a_k|$ on a set, whose elements differ from the true value at most by ν_k

$$\min\max_k|a_k|$$
$$|a_k-\bar{a}_k|\le\nu_k,\ k=\overline{1,4}.$$

For model (3) the observation matching equation as to the relative errors $\mu_k=\nu_k/\bar{a}_k$ is of the following form

$$\begin{aligned} &\max_{1\le j\le n}\Big|\frac{1}{1-\mu_2}F_1(\xi_1,\tau_k,\mu_1,\mu_2)+\\ &+(2\vartheta_0-\vartheta_1-\vartheta_2+\vartheta_1\mu_3+\vartheta_2\mu_4)F_2(\xi_1,\tau_k,\mu_1,\mu_2)+\\ &+(\vartheta_2-\vartheta_1+\vartheta_1\mu_3-\vartheta_2\mu_4)F_3(\xi_1,\tau_k,\mu_1,\mu_2)-\\ &-F_1(\xi_1,\tau_k)\big|_{\mu_{1,2}=0}-(\vartheta_2-\vartheta_1)F_3(\xi_1,\tau_k)\big|_{\mu_{1,2}=0}-\\ &-(2\vartheta_0-\vartheta_1-\vartheta_2)F_2(\xi_1,\tau_k)\big|_{\mu_{1,2}=0}-\\ &+\mu_3\vartheta_1-(\vartheta_1\mu_3-\vartheta_2\mu_4)\xi_1+\frac{\varepsilon_j}{\delta}\mathcal{R}\Big|=\mathcal{R},\quad k=\overline{1,n}, \end{aligned} \tag{4}$$

where

$$F_1=\frac{4}{\pi^3}\sum_{k=1}^{\infty}\frac{\sin(2k-1)\pi\xi}{(2k-1)^3}\left\{1-\exp\left[-\frac{1-\mu_2}{1-\mu_1}(2k-1)^2\pi^2\tau\right]\right\},$$

$$F_2=\frac{2}{\pi}\sum_{k=1}^{\infty}\frac{\sin(2k-1)\pi\xi}{2k-1}\exp\left[-\frac{1-\mu_2}{1-\mu_1}(2k-1)^2\pi^2\tau\right],$$

$$F_3=\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\sin 2k\pi\xi}{k}\exp\left(-4\frac{1-\mu_2}{1-\mu_1}k^2\pi^2\tau\right)$$

$$\xi=\frac{x}{\ell},\ \tau=\frac{\bar{\lambda}}{\bar{c}\rho\ell^2}t,$$

$$\mathcal{R}=\frac{\delta\bar{\lambda}}{f\ell^2},\ \vartheta_0=\frac{u_0\bar{\lambda}}{f\ell^2},\ \vartheta_1=\frac{\bar{v}_1\bar{\lambda}}{f\ell^2},\ \vartheta_2=\frac{\bar{v}_2\bar{\lambda}}{f\ell^2}.$$

To identify the constant unknowns $\{c,\lambda,v_1,v_2\}$ four measurement times (n = 4) for any sensor

location ξ_1 should be known. The sampling $u^{\delta} = \{u^{\delta}_{i,j}\}^{j=1,\dots 4}_{i=1}$ conveys the minimal volume of measurements for the simultaneous identification studied.

The observation scheme with the single sensor is not unique from a viewpoint of the system (4) solvability. The two $u^{\delta} = \{u^{\delta}_{i,j}\}^{j=1,2}_{i=1,2}$ and four $u^{\delta} = \{u^{\delta}_{i,j}\}^{j=1}_{i=1,\dots,4}$ point samples assure the existence of the inverse problem solution also. All these schemes are the subject of the further investigation. Below the attention is paid to the main question regarded to the analysis of the object properties, if one has a single sensor sampling.

The system (4) solution does not exist, if the observation is set in one of the points $x^* = \{0, \ell/2, \ell\}$. This means that the corresponding measurements are unidentifiable for any external loadings and every specimen described by model (3). In these cases the general estimation error $\mu_{rms} \to \infty$. If the measurements are carried out in the middle of a specimen, $x^* = \ell/2$, then the inverse problem solution depends on the combination $\vartheta_1 + \vartheta_2$ of the boundary temperatures and the sought values μ_3 and μ_4 become the arbitrary magnitudes. Note, that this affirmation is valid for the constant model parameters $\{c, \lambda, v_1, v_2\}$. Other inverse problem formulation can reconstruct the solution uniqueness at the point $x = \ell/2$.

Among the main design features it should be determined, first of all, the optimal measurement scheme structure and the factors that able to decrease the estimation errors.

As it follows from (4) the values $\{\mu_k\}_{k=\overline{1,4}}$ depend on the criterion $\mathcal{R}$ and factors

$$\Phi_1 = \vartheta_0 - \frac{\vartheta_1 + \vartheta_2}{2},$$

$$\Phi_2 = \vartheta_2 - \vartheta_1, \Phi_3 = \vartheta_0 - \vartheta_1, \Phi_4 = \vartheta_0 - \vartheta_2 .$$

The ones determine the significance of the system (4) terms. Due to these factors, the decreasing or increasing of the magnitudes $\{\mu_k\}_{k=\overline{1,4}}$ are occurring and the corresponding optimal sensor location is established.

The first factor $\mathcal{R}$ directly determines the values of the estimation errors. The factors $\{\Phi_k\}_{k=\overline{1,4}}$ render the mode of the estimation error distribution. Variations of Φ_1 and Φ_2 denote four distinctive cases of the estimation error behavior. Their bounds are defined by zero values of factors Φ_1, Φ_2 (Fig.1*b, d*) and $\Phi_3 = 0$ (Fig.1*c*) or $\Phi_4 = 0$.

Thus, the factors $\mathcal{R}$ and $\{\Phi_k\}_{k=\overline{1,4}}$ entirely convey all the cases of the experiment effect on the solution accuracy of the inverse problem studied.

Being grounded on the system (4) numerical analysis, we will elicit the further experimental design peculiarities. The *rms*-error was minimized on time for each point of the specimen. The results obtained are shown in Fig. 1. The solution $\tilde{\mu}_{opt} = \min_t \mu_{rms}(x,t)$ depicts the typical character of the optimal design and elicit its main regularities. In particular, the conditions of the simultaneous estimation, mode of the estimation error and factors decreased the measurement noise effect are revealed.

As it seems, there are a number of the optimal sensor locations, namely, in the middle area of the specimen and in the neighborhood of a specimen boundary. The optimal position depends on the variance between the initial and boundary temperatures. In general, there are up to five local minimums (Fig. 1*b*, curves 8 and 12). Their existence is explained by the non-uniqueness of the minimization problem solution. The concurrences of curves signify the transition from one set of optimal times to other ones.

The character of the function $\tilde{\mu}_{opt}$ becomes simpler, if the estimation errors μ_1 and μ_2 are taken into account only. In this case the point $x = \ell/2$ is not singular [14].

From a theoretical viewpoint it is interesting to note the optimal solution existence near the specimen boundary (Fig.1*a*, curves 1-3). These cases take place for the small boundary temperatures, $|v_k| \le 1, k = 1,2$. Similar experiment conditions are the worst thermal loadings. For $|v_k| > 1, k = 1,2$ the optimal sensor location very quickly moves to the middle area of the specimen (Fig. 1*a*, curves 4-7). This area is the position of the global minimum for many cases of the thermal loadings. The optimal sensor location can be approximately termed as $x^{(opt)} = 0.45\ell$ and $x^{(opt)} = 0.55\ell$. For the case $|v_k| \le 1, k = 1,2$ the optimal sensor location is the neighborhood of the point $x^{(opt)} = 0.1\ell$ or $x^{(opt)} = 0.9\ell$.

An essential feature of the optimal time determination is the dense localization of the values $\{\tau_i\}_{j=\overline{1,4}}$ near the certain dimensionless magnitudes. Their magnitudes depend, first of all, on the factor Φ_1. The largest variations of these values are held in the case $\Phi_1 \ge 0$ (Fig. 1*a*). In all other cases the optimal time variations are not significant. So, the inverse problem studied has the

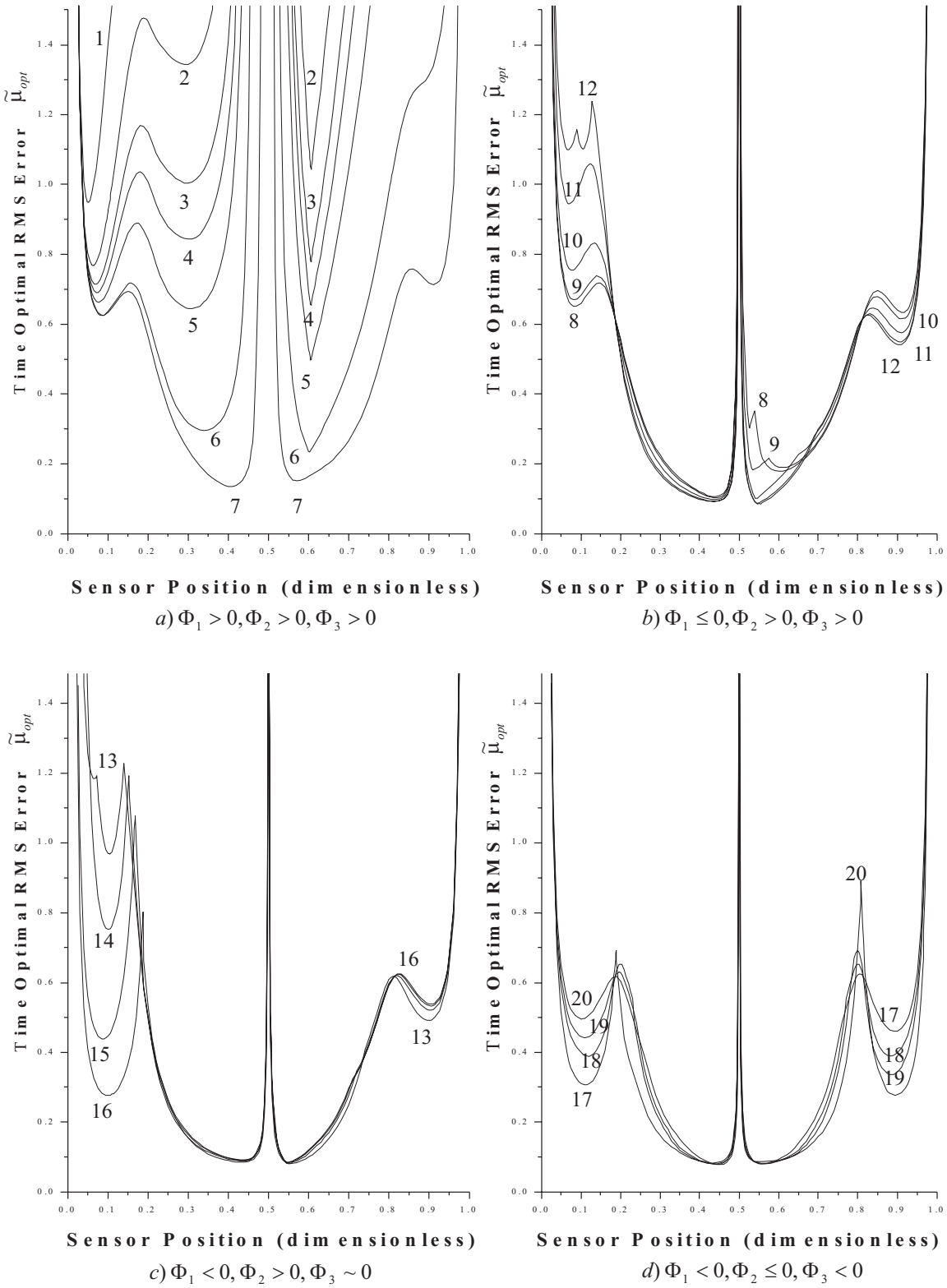


Figure 1 Estimation Error Dependence on Sensor Location and Thermal Loading Variation (in all cases $v_2 > u_0$)

1) $v_1^{(1)} = 1$; 2,...,7) $1 < v_1^{(2)} < v_1^{(3)} < \ldots < v_1^{(7)} < 2u_0 - v_2$; 8) $(v_1^{(8)} + v_2)/2 = u_0$; 9,...,14) $v_1^{(8)} < v_1^{(9)} < \ldots < v_1^{(14)} < u_0$; 15) $v_1^{(15)} = u_0$; 16,17) $u_0 < v_1^{(16)} < v_1^{(17)} < v_2$; 18) $v_1^{(18)} = v_2$; 19) $v_1^{(19)} > v_2$; 20) $v_1^{(20)} >> v_2$

strictly defined scheme of the optimal sensor location and measurement time.

Notice that the above consequence was revealed for the strongly nonlinear dependence between the sought parameters and input/output signals. Similar outcomes are impossible for the traditional design methods. Thereto, a number of design peculiarities are elicited below.

There are brightly denominated areas of the unsatisfactory accuracy of the inverse problem solution. The ones are located near the points $x = \ell / 6$, $x = \ell / 2$ and $x = 5 / 6\ell$. The measurements in these points and their neighborhoods result in the large estimation errors. While the first and third locations result in the large errors as for the thermal properties and boundary temperatures, in the middle of the specimen only the boundary temperatures are reconstructed with the large errors. Note, that the sensor bias from the singular point $x^* = \ell / 2$ only on the value 0.05ℓ provides the optimal identification as the thermal properties and boundary temperatures.

For all optimal observation designs the estimation errors of the boundary temperatures appear less the estimation errors of the thermal properties. The change of the measurement scheme [14] can significantly improve the precision of the thermal properties identification.

It is necessary to make a point on the $\{\Phi_k\}_{k=\overline{1,4}}$ unlimited growth. The curve 20 (Fig.1*d*) shows that the temperature gap between u_0 and $\{v_k\}_{k=1,2}$ does not reduce the estimation error to zero. However the variations of factors $\{\Phi_k\}_{k=\overline{1,4}}$ considerably affect on the shape of the estimation error distribution (Fig.1). Consequently, these factors are the form-factors of the inverse problem studied. The largest variations of the estimation errors take place near the specimen boundary. Here the periodical decreasing and increasing of the local optimum values are occurring, when the boundary temperature v_1 or v_2 sequentially increases from the small value to very large magnitude (Fig.1).

The condition of the absolute decreasing of the estimation errors is the requirement $\mathcal{R} \to 0$. The one can be satisfied due to the noise level diminution and volumetric source magnification.

As it follows from (4) the factor $\mathcal{R}$ growing has the upper admissible bound $\mathcal{R}_{\max}$. The one conveys the condition, when the magnitude of the minimal estimation error is unallowable, $\mu^{(R)} > 1$. The growing of the factor $\mathcal{R}$ means that the noise level δ is increasing. Therefore, the criterion $\mathcal{R}_{\max}$ determines some threshold level of the measurement noise, greater of which the inverse problem solution does not exist. In this case the variation of the noise level δ does not change the character of the estimation error distribution. Consequently, the factor $\mathcal{R}$ determines the indeterminacy power of the mathematical model identification. Its existence shows that the observation uncertainties affect the mathematical model identifiability.

Thus, only one optimal observation of a specimen temperature ensures a test object analysis on a whole and allows to elicit its properties, temperature distribution at every point and, evidently, can determine the boundary heat flux.

Simultaneous Thermal Properties and Boundary Flux Identification

Let us consider the commonly known inverse problem with a heat flux loading scheme. We specify the following mathematical model

$$\begin{aligned} &c\rho \frac{\partial u}{\partial t} = \lambda \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < \ell, t > 0; \\ &u\big|_{t=0} = u_0, \quad 0 < x < \ell; \\ &\left.\frac{\partial u}{\partial x}\right|_{x=0} = 0, \ \lambda \left.\frac{\partial u}{\partial x}\right|_{x=\ell} = \begin{cases} q, t \le t_h, \\ 0, \ t > t_h. \end{cases} \end{aligned} \tag{5}$$

In the frame of the informativeness investigation the following question is raised. *Is it possible to reconstruct as the thermal conductivity coefficient λ and heat flux q simultaneously*?

The question response will be sought as the experimental design problem solution. Namely, the existence and features of the optimal observation design $\Xi^{(C)} = \{x_i^{(opt)}, t_j^{(opt)}\}_{i=1}^{j=1,2}$ with one sensor and two measurement times will be analyzed. Accordingly, the minimum of the guaranteed error

$$\mu^{(C)} = \min_{x,t} \max_i |\mu_i|$$

should be sought, where $\mu_1 = (\overline{\lambda} - \lambda)/\overline{\lambda}$ and $\mu_2 = (\overline{q} - q)/q$ denote the relative errors of the inverse problem solution, $\overline{\lambda}, \overline{q}$ are the actual values.

For the design problem studied the observation matching equation is expressed as follows

$$\max_j \left| \frac{1-\mu_2}{1-\mu_1} \left[\frac{1}{2}\left(\frac{x_1}{\ell}\right)^2 - \frac{1}{6} + (1-\mu_1)\tau_1 + F_1(x_1, \tau_1, \mu_1) \right] - \left[\frac{1}{2}\left(\frac{x}{\ell}\right)^2 - \frac{1}{6} + \tau_1 + F_1\big|_{\mu_1=0}(x, \tau_1) \right] + \frac{\varepsilon_j}{\delta}\mathcal{R} \right| = \mathcal{R}, \ t \le t_h \tag{6}$$

$$\max_j \left| -\mu_2 \Phi + F_2(x_1, \tau_2, \mu_1) - F_2\big|_{\mu_1 = 0}) + \frac{\varepsilon_j}{\delta} \mathcal{R} \right| = \mathcal{R}, \; t > t_h \; ,$$

where

$$F_1 = \frac{2}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} \exp\left[-k^2 \pi^2 (1-\mu_1)\tau \right] \cos k\pi\xi,$$

$$F_2 = \frac{2}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{k}}{k^2} \exp\left[-k^2 \pi^2 (1-\mu_1)(\tau - \Phi) \right] \times$$

$$\times \left\{ 1 - \exp\left[-k^2 \pi^2 (1-\mu_1)\Phi \right] \right\} \cos k\pi\xi \, ,$$

$$\mathcal{R} = \frac{\delta \bar{\lambda}}{\bar{q} \ell}, \Phi = \frac{t_h \bar{\lambda}}{c\rho\ell^2}, \tau = \frac{\bar{\lambda}}{c\rho\ell^2} t, \xi = \frac{x}{\ell} \, .$$

Analysis of the observation matching equations (6) results in the following outcomes.

First of all, two factors, $\mathcal{R}$ and Φ, determine the experimental design character and its behavior. The factor $\mathcal{R}$ is the indeterminacy power of the model (5) identification. There always exists a value $\mathcal{R}_{\max}$, beginning from which $\mu_1 \to \infty$. Again, if $\mathcal{R} \to 0$, then $\mu_k \to 0, k = 1,2$. The factor Φ determines the shape of the error distribution.

Secondly, the estimation error $\mu^{(C)}$ as a function of the variable x has the global and one local minimums. The position of the local minimum depends on the factors $\{\mathcal{R}, \Phi\}$ and covers the closed interval $0 \le x \le 0.9\ell$, i.e., every point of this segment including the opposite specimen boundary, $x = 0$, can be optimal also. For the small observation time t_1 and small $\mathcal{R}$ the local minimum is situated at the neighborhood of $x = 0.9\ell$. This position moves to the opposite boundary $x = 0$, if the value of observation time t_1 is increased.

Being grounded on the use of the majorizing series, we obtain for the small values $\{\mu_k\}_{k=1,2}$ the following estimate

$$\mu_1 \sim \frac{\pm 2\mathcal{R}}{\pm 2\mathcal{R}\left(1 + \frac{t_1}{t_h} \right) + C\left(\frac{\xi^2}{2} - \frac{1}{6} \right)}, \qquad (7)$$

where $C < 1$ is a certain parameter. Also, the estimate $\min_{x,t} |\mu_2| = 2\mathcal{R}/\Phi$ is revealed. From it follows that the magnification of the flux duration t_h provides for the absolute decreasing of the flux estimation error. However, for the fixed $\delta \ne 0$ the estimation error of the specific heat cannot be reduced to zero and $\min_{x,t} |\mu_1| \ne 0$, if $t_h \to \infty$.

As is seen, the kind of function (7) is defined by the factor $\mathcal{R} > 0$. Other parameter t_1 / t_h has the restricted variation because of its limiting in accordance with (5), where $t_1 \le t_h$. Analyzing the asymptotic behavior $\mu_1 \to 1$ at the point $x = \ell$, we obtain $\mathcal{R}_{\max} = 1/12$. Also, formula (7) indicates that the estimation error is the function with the asymptote at the point

$$x^* = \ell \sqrt{\frac{1}{3} \pm \frac{2\mathcal{R}}{C}\left(1 + \frac{t_1}{t_h} \right)}$$

Here only one of these two magnitudes determines the asymptote location. Their choice is defined from the condition of the worst estimation error obtaining. For the optimal measurement time on the interval $|x| < x^*$ the function μ_1 has the minimum at the point $x = 0$. For $|x| > x^*$ the estimation error is a steadily decreasing function over the variable x and in this case $|\mu_1|_{x=\ell}| < |\mu_1|_{x=0}|$.

So, for every constant thermal properties and non-zero external impact the global minimum of the function $\mu^{(C)}$ is located at the point $x^{(opt)} = \ell$. For this location the sought optimal time is the stationary state observation. Other optimal time, generally speaking, is defined as $t^{(opt)} \le t_h$. However, if the heat flux duration t_h is not large, then the estimation error μ_1 at the point $x = \ell$ is the steadily decreasing function, and $t^{(opt)} = t_h$. Thus, for the fixed $\mathcal{R} < 1/12$ and $\Phi > 1/6$ the observation design $\Xi^{(C)} = \{\ell, t_1^{(opt)} \le t_h, t_2^{(opt)} \to \infty\}$ guarantees the optimal reconstruction of the quantities $\{\lambda, q\}$.

It is very remarkable fact that for the optimal observation time the opposite boundary is the location of the second minimum. This means that the isolated specimen side despite the absence of the external impact possesses enough informativeness and the inverse problem can be solved with the sufficient precision.

From a theoretical viewpoint the next peculiarity is important. There exists a such sensor location $0 < x^* < \ell$, depending on the factors $\{\mathcal{R}, \Phi\}$ and observation time t_1, for which the function μ_1 has the unlimited growth. For the optimal observation time the sensor location near the point $x_a = \ell / \sqrt{3}$ gives the worst observation. In this area the error of the inverse problem solution go beyond the permissible bound $\mu_1 > 1$. Hence, between the global and local minimums there always exist unidentifiable and poor informativeness areas.

Thus, the main regularities of the inverse problem solution are revealed analytically with the help of the experimental design method developed.

As it seems, one can identify not only the thermal conductivity, but simultaneously the boundary heat flux. There are two optimal sensor locations. The best measurement is the temperature observation at the boundary with heat flux. The isolated specimen boundary can be alternative local optimal area of measurements. The heat flux duration determines the optimal moment of the temperature measurement, if its value is not large. Otherwise, the optimal observation time appears less the duration that the heating is applied.

CONCLUSIONS

The principal questions of the simultaneous estimation of the model coefficients and boundary conditions were considered. The main theoretical outcome is that the measurements into the single observation point convey the vast information on the object properties and the ones can be sought uniquely and with sufficient precision.

This means that the interpretation of thermal experiments allows one to identify simultaneously as the specimen properties and the heat-exchange parameters, the thermal properties and boundary conditions. As a result, the properties of complex heat objects can be studied without additional numerous temperature measurements.

The comprehensive analyses of two types of inverse problems were accomplished. The effect of the experiment conditions on the experimental design problem solution was traced. It was shown that two types of factors define the behavior of the estimation errors. The optimal observation schemes were detected for any constant thermal properties and external impact. The complete picture of the estimation error distribution was revealed.

The results obtained expand the frame of the inverse problem formulation. Grounded on the approach developed the actual thermal objects can be analyzed.

REFERENCES

1. J. V. Beck, and K. Arnold, *Parameter Estimation in Engineering and Science*, John Wiley and Sons, Inc., New York, 1977.

2. M. R. Romanovskii, Identification of the Thermal Conductivity and Heat Capacity by Regularization Method Based on Partial Agreements, *J. of Engineering Physics and Thermophysics,* **42**, 110-118 (1982).

3. C. H. Huang, and M. N. Ozisik, Direct Integration Approach for Simultaneously Estimating Temperature Dependent Thermal Conductivity and Heat Capacity, *Numerical Heat Transfer*, *Part A,* **20**, 95-110 (1991).

4. K. J. Dowding, J. V. Beck, A. Ulbrich, B. F. Blackwell, and J. Hayes, Estimation of Thermal Properties and Surface Heat Flux in a Carbon-carbon Composite Material, *J. of Thermo-physics and Heat Transfer,* **9**, 345-351 (1995).

5. D. Lesnic, L. Elliott, and D. B. Ingham, Identification of the Thermal Conductivity and Heat Capacity in Unsteady Nonlinear Heat Conduction Problems Using the Boundary Element Method, *J. of Computational Physics*, **126**, 410-420 (1996).

6. J. R. Cannon, and P. DuChateau, Determination of Unknown Physical Properties in Heat Conduction Problems, *Int. J. Engineering Sci.,* **11**, 783-794 (1973).

7. S. Kitamura, and S. Nakagiri, Identifiability of Spatially-Varying and Constant Parameters in Distributed Systems of Parabolic Type, *SIAM J. Control and Optimization,* **15**, 785-800 (1977).

8. L. Carotenuto, and G. Raiconi, Identifiability and Identification of a Galerkin Approximation for a Class of Distributed Parameter Systems, *Int. J. Syst. Sci.,* **11**, 1035-1049 (1980).

9. M. Courdesses, M. Polis, and M. Amouroux, On Identifiability of Parameter in a Class of Parabolic Distributed Systems, *IEEE Trans., Automatic Control,* **26**, 474-477 (1981).

10. V. Isakov, *Inverse Problems for Partial Differential Equations*, Applied Mathematical Science, **127**, Springer Verlag, Berlin, 1998.

11. M. R. Romanovskii, Mathematical Modeling of Experiments With the Help of Inverse Problems, *J. of Engineering Physics and Thermophysics,* **57**, 1112-1117 (1989).

12. M. R. Romanovskii, Experimental Design Method with General Assumptions About Form of the Model of Experimental Object, *Industrial Laboratory*, **59**, 89-96 (1993).

13. M. R. Romanovskii, Regularization of Inverse Problems, *High Temperature*, **18**, 135-140 (1980).

14.M. R. Romanovski, Optimal Experimental Design with Inverse Problems, *Inverse Problems In Engineering*, Ed. K. A. Woodbury, 501-508 (2000).

THE DETERMINATION OF THE THERMAL PROPERTIES OF A HEATED FLAT PLATE

Daniel Lesnic
Department of Applied Mathematics, University of Leeds,
Leeds, West Yorkshire, LS2 9JT, UK.
Email: amt5ld@amsta.leeds.ac.uk

ABSTRACT

This paper considers the inverse determination of the thermal properties of a flat plate heat conductor. The governing equation for the unknown temperature in given by the nonlinear transient heat conduction equation. In addition to prescribed initial and/or boundary values, specified continuously differentiable temperature data with non-zero derivative at a single sensor location is given. When the thermal properties obey a certain relationship which enables one to linearise exactly the nonlinear heat equation, then their dependence upon the temperature is obtained explicitly, whilst the unknown temperature is obtained implicitly. Since the resulting linear heat conduction equation cannot be, in general, solved analytically, a boundary element method is developed in order to obtain the solution numerically.

NOMENCLATURE

A	Constant
A_n	Thermal amplitudes
C	Heat capacity
G	Fundamental solution
H	Heaviside function
K	Thermal conductivity
L	Half length of the plate
T	Temperature
a	Moving boundary
g	Function relating $K(T)$ and $C(T)$
q_L	Applied heat flux
α	Thermal diffusivity
η	Coefficient function

1 INTRODUCTION

We consider the heat conduction in a flat plate $(-L, L)$ of length $2L$ such that at the boundary a non-zero heat flux $q_L(t) \neq 0$ is applied and, due to symmetry, at the centre $x = 0$ of the flat plate the heat flux is zero. Also, at $t = 0$ the initial temperature is uniform and is taken to be zero. Then the mathematical model requires solving the unsteady nonlinear heat conduction equation, namely,

$$\frac{\partial}{\partial x}\left(K(T)\frac{\partial T}{\partial x}\right) = C(T)\frac{\partial T}{\partial t}, \ t > 0, \ 0 < x < L \tag{1}$$

where T is the temperature, $K(T) > 0$ is the thermal conductivity and $C(T) > 0$ is the heat capacity, subject to the initial condition

$$T(x, 0) = 0, \quad 0 \leq x \leq L \tag{2}$$

and the boundary conditions

$$-K(T(0, t))\frac{\partial T}{\partial x}(0, t) = 0, \quad t > 0 \tag{3}$$

$$K(T(L, t))\frac{\partial T}{\partial x}(L, t) = q_L(t) \neq 0, \quad t > 0 \tag{4}$$

2 LINEARIZATION

It has been shown, see Storm [1], that a sufficient condition for the nonlinear eqn.(1) to be transformed into a linear form is given by the constancy of the term

$$(KC)^{-1/2}(d/dT)ln(C/K)^{1/2} = A = \text{constant}. \tag{5}$$

Under the assumption (5), the transformations

$$\overline{Q}(T) = \int_0^T [K(\lambda)C(\lambda)]^{1/2} d\lambda \tag{6}$$

$$Q(x,t) = \overline{Q}(T(x,t)) = Q^*(X(x,t),t) \qquad (7)$$

where

$$X(x,t) = \int_0^x [C(T(x',t))/K(T(x',t))]^{1/2} dx' \qquad (8)$$

and finally, if $A \neq 0$, the transformation

$$Q^*(X,t) = -(1/A) ln(\zeta(X,t)) \qquad (9)$$

recast the nonlinear heat conduction eqn.(1) into the linear form, see Storm [1],

$$\frac{\partial^2 \zeta}{\partial X^2} = \frac{\partial \zeta}{\partial t}. \qquad (10)$$

At this stage it is worth noting that the most general forms for $K(T)$ and $C(T)$ which satisfy eqn.(5) are given by

$$K(T) = K_0 g(T) exp\left[-A(K_0 C_0)^{1/2} \int_0^T g(\lambda) d\lambda\right] \qquad (11)$$

$$C(T) = C_0 g(T) exp\left[A(K_0 C_0)^{1/2} \int_0^T g(\lambda) d\lambda\right] \qquad (12)$$

where $K_0 = K(0)$, $C_0 = C(0)$ are positive constants, and the otherwise arbitrary positive function $g(T)$ satisfies the normalization relation $g(0) = 1$.

Observations:

(i) It was remarked, see for example Hills and Hensel [2] and Wrobel and Brebbia [3], that many materials, in addition to having the property that the thermal conductivity $K(T)$ and the heat capacity $C(T)$ are temperature dependent, also have direct proportionality between these two physical quantities, i.e. their thermal diffusivity $\alpha(T) = K(T)/C(T)$ is constant. In such a situation, the constant A in eqn.(5) is zero, and therefore only the transformation (6) needs to be employed to recast eqn.(1) into a linear form. Inverse determinations of $K(T)$ and $C(T)$ in this quasi-linear situation can be found in Cannon and Duchateau [4] and Lesnic *et al.* [5].

(ii) Conditions of the type (5) were originally obtained by Storm [1] in an investigation of heat conduction in simple monoatomic metals. There, the validity of the approximation was examined for aluminium, silver, sodium, cadium, zinc, copper and lead. It was shown that $K(T)C(T)$ and $(d/dT) ln(C/K)^{1/2}$ exhibit only small variation over wide temperature ranges within which, the approximation (5) is justified. Similar conclusions were reached for iron and 80% carbon steel. A good account of the experimental evidence on the variation of thermal conductivity and specific heat with temperature is given in Tslaf [6].

(iii) For a variety of simple metals, Storm [1] found that the product $K(T)C(T)$ is essentially constant for a wide range of temperatures and therefore $g(T) \equiv 1$, which on using eqns (11) and (12) give

$$K(T) = K_0 exp\left[-A(K_0 C_0)^{1/2} T\right] \approx K_0(1 - bT) \qquad (13)$$

$$C(T) = C_0 exp\left[A(K_0 C_0)^{1/2} T\right] \approx C_0(1 + bT) \qquad (14)$$

where for simple metals the quantity $b = A(K_0 C_0)^{1/2}$ is typically $\mathbf{O}(10^{-4})$. The linear assumption given by eqns (13) and (14) was previously adopted by Huang and Ozisik [7] for simultaneously estimating numerically the thermal conductivity $K(T)$ and the heat capacity $C(T)$ from interior temperature measurements.

(iv) By taking $g(T) = \frac{\gamma(T+1)^3}{\gamma - 1 + (1+T)^4}$, where $\gamma = 4A^{-1}(K_0 C_0)^{-1/2}$, we obtain $K(T) = \frac{K_0 \gamma^2 (1+T)^3}{(\gamma - 1 + (1+T)^4)^2}$ and $C(T) = C_0(1+T)^3$ which are the thermal properties of solid crystalline hydrogen under certain conditions, see Rosen [8].

(v) By taking $C(T) \equiv 1$, Knight and Philip [9] remarked that $g(T) = (1 + AK_0^{1/2} T)^{-1}$, and then the thermal conductivity (diffusivity) satisfying eqn.(5) is given by $K(T) = K_0(1 + AK_0^{1/2} T)^{-2}$. Further, this is the only form of thermal conductivity for which eqn.(1) is invariant under Lie-Backlund transformations, see Bluman and Kumei [10]. General inverse determinations of $K(T)$ when $C(T) \equiv 1$ can be found in Cannon and Duchateau [11,12] and Duchateau [13].

3 INVERSE PROBLEM

In an inverse coefficient identification problem formulation we may require finding the triplet $(K(T) > 0, C(T) > 0, T)$ satisfying (1)-(4) and the additional boundary temperature measurement at $x = 0$, namely,

$$T(0,t) = f(t), \quad t > 0. \qquad (15)$$

Perhaps at this point we should emphasize the distinction between what we refer to as intial and boundary data such as (2)-(4) and data that we refer to as overspecified (additional) such as (15). In the quenching of a heat conductor for which eqn.(1) is a mathematical model, we can force the conditions (2)-(4) to be satisfied by appropriately arranging the experiment as described. At the same time we can measure the temperature $T(0,t)$ and record the fact that it is equal to $f(t)$. We cannot, however, arbitrarily require $T(0,t)$ to attain specific values independent of (2)-(4). In other words, the specified data is that over which we exercise control and thus force to be satisfied, whilst the overspecified (additional) data is simply observed and recorded. Clearly, any hypothesis on the overspecified data $f(t)$ cannot be completely independent of whatever hypotheses are placed on the specified data $q_L(t)$.

There are two main approaches to solving the type of inverse problem aforementioned. These can be classified into either parameterisation schemes, e.g. linear variations of the thermal properties, see Huang and Ozisik [7], or quadratic, see Milano and Scarpa [14] and Lesnic *et al.* [15], or coefficient identification schemes, see Cannon [16], Cannon and Duchateau [4,11,12], Duchateau [13] and Lesnic *et al.* [5]. The parameterisation schemes reduce the inverse problem to finding a finite-dimensional space of unknowns and then seeking to adjust these parameters such that the gap between the computed output generated by an associate mathematical model and the experimentally measured input data is minimized in a least-squares manner. These methods have the advantage that the inverse problem can be recast as a (constrained) minimization problem for which efficient computational routines are available. However, the solution obtained using parameterisation based on a minimization procedure is not necessarily the mathematical solution of the inverse problem. Therefore, the purpose of this paper is to develop a reconstructional approach in which the unknown thermal properties can be expressed directly in terms of the overspecified data. For this purpose we investigate a special space of functions to which the coefficients $K(T)$ and $C(T)$ should belong, see eqn.(5), in order to linearise eqn.(1). With this respect this study generalizes the analyses of Cannon and Duchateau [4] and Lesnic *et al.* [5] who considered only the case when the constant A in eqn.(5) is zero.

In this paper we investigate the inverse determination of the triplet $(K(T) > 0, C(T) > 0, T)$ satisfying the problem (1)-(4) and (15), when $K(T)$ and $C(T)$ have the special form given by eqns (11) and (12) with K_0, C_0 and A specified; a situation in which we have to identify only the pair $(g(T), T)$. This provides insight into what data is required for the more general and yet unsolved inverse problem when $K(T)$ and $C(T)$ are not connected by any relationship and need to be determined simultaneously.

Since the thermal properties are continuous and positive, we also require that $g(T)$ is a strictly positive continuous function with $g(0) = 1$. Further, we assume that the measurement $f(t)$ is a continuously differentiable function with non-zero derivative. This assumption should be taken as a definition for the *admissible data set* in which the inverse problem is formulated and similar considerations have been made in Cannon [16] and Muzylev [17]. The dependence upon T of the coefficients $K(T)$ and $C(T)$ is obtained explicitly, whilst the unknown temperature $T(x,t)$ is obtained implicitly and can be then calculated numerically.

4 INVERSE ANALYSIS

Using eqns (11) and (12), the transformations (6)-(9) become

$$-\frac{1}{A} ln(\zeta(X(x,t),t)) = Q^*(X(x,t),t)$$
$$= Q(x,t) = \int_0^{T(x,t)} (K_0 C_0)^{1/2} g(\lambda) d\lambda \qquad (16)$$

where

$$X(x,t) = \int_0^x (C_0/K_0)^{1/2}$$
$$exp\left(A(K_0 C_0)^{1/2} \int_0^{T(x',t)} g(\lambda) d\lambda \right) dx'$$
$$= \int_0^x \frac{(C_0/K_0)^{1/2}}{\zeta(X(x',t),t)} dx'. \qquad (17)$$

Inverting eqn.(17) we obtain

$$x = (K_0/C_0)^{1/2} \int_0^X \zeta(X',t) dX'. \qquad (18)$$

It should be noted that, in particular, $x = 0$ if and only if $X(0,t) \equiv 0$. Applying the condition (15) in eqn.(16), results in

$$Q(x_0,t) = \int_0^{f(t)} (K_0C_0)^{1/2} g(\lambda) d\lambda \qquad (19)$$

Differentiating expression (19) with respect to t, yields

$$(K_0C_0)^{1/2} g(f(t)) f'(t) = \frac{\partial Q}{\partial t}(0,t), \quad t > 0 \qquad (20)$$

Now using eqns (7)-(9) it can be seen that

$$\frac{\partial Q}{\partial t}(0,t) = \frac{\partial^2 Q^*}{\partial X^2}(0,t) = \left(\frac{\zeta_X^2 - \zeta\zeta_{XX}}{A\zeta^2} \right)(0,t) \qquad (21)$$

where $\zeta_X = \partial\zeta/\partial X$ and $\zeta_{XX} = \partial^2\zeta/\partial X^2$. If $f'(t) \neq 0$ for all $t > 0$, then f is invertible and therefore using eqns (20) and (21) we obtain the expression for $g(t)$ in the form

$$g(t) = \frac{(f^{-1})'(t)}{(K_0C_0)^{1/2}A} \left(\frac{\zeta_X^2 - \zeta\zeta_{XX}}{\zeta^2} \right)(0, f^{-1}(t)) \qquad (22)$$

Once the function g, and thus K and C, as given by eqns (11) and (12), are obtained, using eqns (6), (7), (16) and (22) we obtain

$$\overline{Q}(T) = \frac{1}{A} \int_0^T (f^{-1})'(\lambda) \left(\frac{\zeta_X^2 - \zeta\zeta_{XX}}{\zeta^2} \right)(0, f^{-1}(\lambda)) d\lambda. \qquad (23)$$

Inverting eqn.(7), we finally obtain

$$T(x,t) = \overline{Q}^{-1}(Q(x,t)) \qquad (24)$$

where $\overline{Q}$ is given by eqn.(23).

5 DISCUSSION

In what follows, we consider only the case $A \neq 0$, as the case $A = 0$ was previously treated by Cannon and Duchateau [4] and Lesnic *et al.* [5].

Based on the transformations (16) and (17), in terms of the function ζ, the problem given by the eqn.(10) subject to the boundary conditions (2)-(4) becomes

$$\frac{\partial\zeta}{\partial t}(X,t) = \frac{\partial^2\zeta}{\partial X^2}(X,t), \quad t > 0, \quad 0 < X < a(t) \qquad (25)$$

$$\zeta(X,0) = 1, \quad 0 \leq X \leq a(0) \qquad (26)$$

$$\frac{\partial\zeta}{\partial X}(0,t) = 0, \quad t > 0 \qquad (27)$$

$$\frac{\partial\zeta}{\partial X}(a(t),t) = -Aq_L(t)\zeta(a(t),t), \quad t > 0 \qquad (28)$$

where $a(t) = X(L,t)$. From eqn.(17) we have $X(0,t) = 0$, $a(0) = X(L,0) = L(C_0/K_0)^{1/2}$ and from eqn.(18) we have

$$L = (K_0/C_0)^{1/2} \int_0^{a(t)} \zeta(X',t) dX'. \qquad (29)$$

Differentiating eqn.(29) with respect to t and using eqns (25), (27) and (28) we obtain

$$\begin{aligned} 0 &= \zeta(a(t),t)a'(t) + \int_0^{a(t)} \zeta_t(X',t) dX' \\ &= \zeta(a(t),t)a'(t) + \zeta_X(a(t),t) \\ &= \zeta(a(t),t)[a'(t) - Aq_L(t)]. \end{aligned} \qquad (30)$$

Since $\zeta \neq 0$, we deduce that the moving boundary $a(t)$ is given by

$$a(t) = a(0) + A \int_0^t q_L(t') dt'. \qquad (31)$$

It can be remarked that the problem (25)-(28) is an inverse Stefan type moving boundary problem which appears not to have an analytical solution even in the simplest case when $q_L \equiv$ *constant* $\neq 0$. However, it can be transformed into an integral form using Green's identities, see Wrobel [18], namely,

$$\begin{aligned} \eta(X)\zeta(X,t) &= \int_0^t G(X,t;a(\tau),\tau) \frac{\partial\zeta}{\partial X}(a(\tau),\tau) d\tau \\ &\quad + \int_0^t G(X,t;0,\tau) \frac{\partial\zeta}{\partial X}(0,\tau) d\tau \\ &\quad - \int_0^t \frac{\partial G}{\partial X}(X,t;a(\tau),\tau)\zeta(a(\tau),\tau) d\tau \\ &\quad - \int_0^t \frac{\partial G}{\partial X}(X,t;0,\tau)\zeta(0,\tau) d\tau \\ &\quad + \frac{1}{2}\left[erf\left(\frac{X}{2t^{1/2}} \right) - erf\left(\frac{X - a(0)}{2t^{1/2}} \right) \right] \\ &\quad + \int_{a(0)}^{a(t)} G(X,t;y,\tau(y))\zeta(y,\tau(y)) dy, \\ &\quad t > 0, \quad 0 \leq X \leq a(t) \end{aligned} \qquad (32)$$

where $erf(\xi) = 2\pi^{-1/2}\int_0^{\xi} e^{-\sigma^2}d\sigma$ is the error function, $\eta(X) = 1$ if $X \in (0, a(t))$, $\eta(0) = \eta(a(t)) = 1/2$, $\tau(y) = a^{-1}(y)$ which exists if $A \neq 0$ and $q_L(t) \neq 0$, see eqn.(31), and G is the fundamental solution of the heat equation, namely,

$$G(X,t;\xi,\tau) = \frac{H(t-\tau)}{(4\pi(t-\tau))^{1/2}} exp\left(-\frac{(X-\xi)^2}{4(t-\tau)}\right) \tag{33}$$

where H is the Heaviside function.

Further, using the boundary conditions (27) and (28) and the transformation of variables $y = a(\tau)$, we can recast eqn.(32) in the form

$$\begin{aligned}\eta(X)\zeta(X,t) &= -\int_0^t \frac{\partial G}{\partial X}(X,t;a(\tau),\tau)\zeta(a(\tau),\tau)d\tau \\ &\quad -\int_0^t \frac{\partial G}{\partial X}(X,t;0,\tau)\zeta(0,\tau)d\tau \\ &\quad +\frac{1}{2}\left[erf\left(\frac{X}{2t^{1/2}}\right) - erf\left(\frac{X-a(0)}{2t^{1/2}}\right)\right], \\ &\quad t>0, \quad 0 \le X \le a(t)\end{aligned} \tag{34}$$

since the first two terms and the last term of the right-hand side of (32) are equal and of opposite signs.

A possible way to check the accuracy of the results is to solve the problem (25)-(28) when instead of (31) the moving boundary $a(t)$ is given by

$$a_s(t) = a(0) + sA\int_0^t q_L(t')dt' \tag{35}$$

where $s \in [0,1]$. If ζ_s denotes the solution of this modified problem, eqn.(32) yields

$$\begin{aligned}\eta(X)\zeta_s(X,t) &= -\int_0^t \frac{\partial G}{\partial X}(X,t;a_s(\tau),\tau)\zeta_s(a_s(\tau),\tau)d\tau \\ &\quad -\int_0^t \frac{\partial G}{\partial X}(X,t;0,\tau)\zeta_s(0,\tau)d\tau \\ &\quad +\frac{1}{2}\left[erf\left(\frac{X}{2t^{1/2}}\right) - erf\left(\frac{X-a(0)}{2t^{1/2}}\right)\right] \\ &\quad +A(s-1)\int_0^t G(X,t;a_s(\tau),\tau) \\ &\quad q_L(\tau)\zeta_s(a_s(\tau),\tau)d\tau,\ t>0,\ 0 \le X \le a(t)\end{aligned} \tag{36}$$

The reason for this is that when $s = 1$, eqn.(36) reduces to (34) and when $s = 0$ we can compare the numerical results with the analytical solution of the problem (25)-(28) which is available when $q_L = constant \neq 0$ and $a(t)$ is replaced by $a(0) = a_0$. This analytical solution, see Luikov [19], is given by

$$\zeta_0(X,t) = \sum_{n=1}^{\infty} A_n cos(\mu_n X/a_0) exp(-\mu^2 t/a_0^2) \tag{37}$$

where the thermal amplitudes A_n are given by

$$A_n = \frac{2sin(\mu_n)}{\mu_n + a_0 sin(\mu_n)cos(\mu_n)} \tag{38}$$

and μ_n are the roots of the transcendental equation

$$\mu_n tan(\mu_n) = Bi \tag{39}$$

where $Bi = Aq_L a_0$ is the Biot number. For example for $Bi \in [0,2]$, within a maximum error of 0.08%, the approximations for μ_n are given by, see Beck *et al.* [20],

$$\begin{aligned}\mu_1 &\approx \left\{\frac{3Bi}{3+Bi}\left[1 - \frac{Bi^2}{5(3+Bi)^2}\right]\right\}^{1/2} \\ \mu_n &\approx \frac{(n-1)\pi}{2(3+Bi)} \\ &\quad \times \left[2Bi + 3 + 3\left(1 + \frac{4Bi(Bi+3)}{3(n-1)^2\pi^2}\right)^{1/2}\right], \\ &\quad n \ge 2.\end{aligned} \tag{40}$$

To obtain even more accurate values one can use the Newton-Raphson formula

$$\mu_n = \mu_n' - \frac{\mu_n' tan(\mu_n') - Bi}{tan(\mu_n') + \mu_n' sec^2(\mu_n')}, \quad n \ge 1 \tag{41}$$

where μ_n' is the value from eqn.(40).

The boundary integral eqn.(36) can be solved numerically using the boundary element method (BEM), as described in Wrobel [18] and Chuang and Szekely [21].

Since $\zeta_{XX}(0,t) = \zeta_t(0,t)$ and using eqn.(26), eqns (20) and (21) yield

$$g(f(t))f'(t) = -(K_0C_0)^{-1/2}A^{-1}\frac{\zeta_t}{\zeta}(0,t) \tag{42}$$

where the value of $\zeta_t(0,t)$ can be evaluated numerically using finite differences.

As additional temperature measurements (15) we can consider the simplest continuously differentiable function with $f'(t) \neq 0$, namely $f(t) =$

$t-\beta$, which is a strictly increasing function, and then eqn.(42) gives

$$g(T) = -(K_0 C_0)^{-1/2} A^{-1} \frac{\zeta_t}{\zeta}(0, T+\beta) \quad (43)$$

where $\beta > 0$ is chosen such that $g(0) = 1$.

The thermal properties $K(T)$ and $C(T)$ can also be found from eqns (11) and (12), namely,

$$K(T) = -(K_0/C_0)^{1/2} A^{-1} \zeta_t(0, T+\beta) \quad (44)$$

$$C(T) = -(C_0/K_0)^{1/2} A^{-1} \frac{\zeta_t}{\zeta^2}(0, T+\beta) \quad (45)$$

and the temperature T from eqns (23) and (24).

6 CONCLUSIONS

In this study the inverse determination of the thermal properties of heat conductors which are temperature dependent has been investigated. Since it is too much to hope for an analytical solution of the nonlinear heat equation, when the thermal coefficients are represented as general functions of the temperature, the thermal properties were sought in a special class of functions. Then, not only can the nonlinear heat equation be exactly linearised but also the thermal properties and the temperature can be determined uniquely from a single time-dependent boundary temperature measurement function, provided that its continuous derivative never vanishes. In practice this can be achieved by fitting a smooth function with non-zero derivative through the noisy measurement data. The analytical method developed in this study yields exact solutions, in implicit form, which can then be solved numerically. The numerical solutions have been obtained elsewhere, see Lesnic [22], and will be presented at the conference. The main thrust of the paper was on establishing the uniqueness of solution, hoping that afterwards the stability issue can be resolved by using, for example, Tikhonov's regularization method with the choice of the regularization parameter based on Morozov's discrepancy principle. Future work will be concerned with extending the inverse analysis to situations when the thermal properties may depend on both x and T, which are encountered especially in porous media flows.

Acknowledgement

The author would like to acknowledge some financial support received from the UK Royal Society and the University of Leeds for attending the conference.

REFERENCES

1. M.L. Storm, Heat conduction in simple metals, *J. Appl. Phys.*, **22**, 940 (1951).

2. R.G. Hills and E. Hensel, One-dimensional nonlinear inverse heat conduction technique, *Numer. Heat Transfer*, **10**, 369 (1986).

3. L.C. Wrobel and C.A. Brebbia, The dual reciprocity boundary element method for nonlinear diffusion problems, *Comput. Meth. Appl. Mech. Eng.*, **65**, 147 (1987).

4. J.R. Cannon and P. Duchateau, Determining unknown coefficients in a nonlinear heat conduction problem, *SIAM J. Appl. Math.*, **24**, 298 (1973).

5. D. Lesnic, L. Elliott and D.B. Ingham, A note on the determination of the thermal properties of a material in a transient nonlinear heat conduction problem, *Int. Commun. Heat Mass Transfer*, **22**, 475 (1995).

6. A. Tslaf, *Combined Properties of Conductors*, Elsevier, Amsterdam, 1981.

7. C.H. Huang and M.N. Ozisik, Direct integration approach for simultaneously estimating temperature dependent thermal conductivity and heat capacity, *Numer. Heat Transfer*, **20**, 95 (1991).

8. G. Rosen, Nonlinear heat conduction in solid H_2, *Phys. Rev. B*, **19**, 2398 (1979).

9. J.H. Knight and J.R. Philip, Exact solutions in nonlinear diffusion, *J. Eng. Math.*, **8**, 219 (1974).

10. G. Bluman and S. Kumei, On the remarkable nonlinear diffusion equation $(\partial/\partial x)[a(u+b)^{-2}(\partial u/\partial x)] - (\partial u/\partial t) = 0$, *J. Math. Phys.*, **21**, 1019 (1980).

11. J.R. Cannon and P. Duchateau, An inverse problem for a nonlinear diffusion equation, *SIAM J. Appl. Math.*, **39**, 272 (1980).

12. J.R. Cannon and P. Duchateau, Indirect determination of hydraulic properties, in *Flow in Porous Media*, (Douglas Jr, J. and Hornung, U., editors), Birkhauser Verlag, Berlin, 1993, 37-50.

13. P. Duchateau, Monotonicity and uniqueness results in identifying an unknown coefficient in a nonlinear diffusion equation,

SIAM J. Appl. Math., **41**, 310 (1981).

14. G. Milano and F. Scarpa, Numerical experiments of thermophysical properties identification from transient temperature data, in *Proc. 3rd Ann. Inverse Probl. Eng. Seminar*, (Beck, J.V., editor), 1990, 1-13.

15. D. Lesnic, L. Elliott and D.B. Ingham, Identification of the thermal conductivity and heat capacity in unsteady nonlinear heat conduction problems using the boundary element method, *J. Comput. Phys.*, **126**, 410 (1995).

16. J.R. Cannon, Determination of the unknown coefficient $k(u)$ in the equation $\nabla \bullet k(u)\nabla u = 0$ from overspecified boundary data, *J. Math. Anal. Appl.*, **18**, 112 (1967).

17. N.V. Muzylev, Uniqueness of the determination of the temperature dependence of thermophysical properties in piecewise monotonic processes, *Inzhenerno Fizicheskii Zhurnal*, **49**, 1009 (1985) (in Russian).

18. L.C. Wrobel, A boundary element solution to Stefan's problem, in *Boundary Elements V*, (Brebbia, C.A., Futagami, T. and Tanaka, M., editors), Springer-Verlag, Berlin, 1983, 173-182.

19. A.V. Luikov, *Analytical Heat Diffusion Theory*, Academic Press, London, 1968.

20. J.V. Beck, K.D. Cole, A. Haji-Sheikh and B. Litkouhi, *Heat Conduction Using Green's Functions*, Hemisphere Publ., London, 1992.

21. Y.K. Chuang and J. Szekely, On the use of Green's functions for solving melting or solidification problems, *Int. J. Heat Mass Transfer*, **14**, 1285 (1971).

22. D. Lesnic, The determination of the thermal properties of a heat conductor in a nonlinear heat conduction problem, *ZAMP*, **52**, 2002 (in press).

EXTENSION OF METHODOLOGY FOR THE DETERMINATION OF TWO-PHASE POROUS MEDIA PROPERTY FUNCTIONS

Randi Valestrand
University of Bergen
Department of Physics
Bergen, Norway
Randi.Valestrand@fi.uib.no

Alv-Arne Grimstad
RF-Rogaland Research
Bergen, Norway
Alv-Arne.Grimstad@rf.no

Kristofer Kolltveit
University of Bergen
Department of Physics
Bergen, Norway
stoffen@fi.uib.no

Jan-Erik Nordtvedt
University of Bergen
Department of Physics
Bergen, Norway
jen@petec.com

Jack Phan
Colorado State University
Dept. of Chemical Engineering
Fort Collins, Colorado, USA
jackphan@engr.colostate.edu

A. Ted Watson
Colorado State University
Dept. of Chemical Engineering
Fort Collins, Colorado, USA
atw@engr.colostate.edu

ABSTRACT

This paper addresses estimation of porous media property functions, absolute and relative permeabilities, from experiments on core samples. Typically, one has assumed homogeneous and isotropic property functions. However, recent studies have shown that such an assumption frequently is erroneous, and may indeed lead to large errors in the estimated properties. The property functions are inaccessible to direct measurements, and have to be estimated through an inverse problem. This paper describes an algorithm capable of successfully determining the property functions taking the heterogeneous nature of the porous medium into account. In this paper, this methodology is extended in two ways - the estimation algorithm is made more robust and saturation data has been included, and, for the first time, experimental data is used for verification purposes.

NOMENCLATURE

$\boldsymbol{a}$ = Parameter vector
B = Spline basis function
$\boldsymbol{b}$ = Parameter vector
$\boldsymbol{c}$ = Parameter vector
$\mathbf{F}$ = Model output
J = Objective function
$\tilde{J}$ = Predicted attainable obj. function
k = Absolute permeability [Darcy=10^{-12}m^2]
kr = Relative permeability
p = Pressure
P_c = Capillary pressure
S = Saturation
$\boldsymbol{x}$ = Spatial coordinate vector
$\mathbf{Y}$ = Measured data
$\boldsymbol{y}$ = Spline knot vector

Greek symbols
μ = Viscosity
ϕ = Porosity
σ = Standard deviation
$\tilde{\sigma}$ = Approximate standard deviation

INTRODUCTION

The mathematical simulation of the flow of fluids through permeable media is of vital importance to the management of underground resources, such as aquifiers and petroleum reservoirs. In order to perform simulations of field-scale processes, the absolute and relative permeabilities of the porous media must be specified. Such data are typically obtained through experiments conducted on small samples extracted from the reservoir considered. These analysis are referred to as Special Core Analysis (SCAL) in the industry.

One of the main problems within SCAL is that the permeability and the relative permeability functions are not directly measurable

quantities, but quantities that need to be inferred from the mathematical model of flow through porous media, and some data associated with an experiment conducted on the sample - i.e. through solution of an inverse problem.

Recent studies show that the modeling error resulting from the failure to adequately describe the spatial variations in the permeability distribution can lead to significant errors in the estimation of the relative permeability functions [1]. To mitigate these errors, it is desirable to determine the actual absolute permeability distribution for inclusion into the mathematical model of the laboratory experiment [1]. Mejia et al., [2] included variations in absolute permeability when estimating the relative permeability and capillary pressure functions. Independent determination of the permeability distribution may be made (see Seto et al., [3] where nuclear magnetic resonance is utilized), however; the necessary equipment is typically not available within SCAL. Valestrand et al., [1] presented, for the first time, a method for simultaneous determination of the permeability distribution and the relative permeabilities from SCAL data measured during a two-phase displacement experiment.

In this paper the methodology in Valestrand et al., [1] is extended in two ways: (1) We refine the simultaneous estimation procedure by using an algorithm for absolute permeability estimation which provides more flexibility in selection of basis functions, and is more computationally efficient. We also utilize saturation data in the estimation procedure. Such data has shown to be very advantageous for relative permeability estimation [2,4]. (2) We attempt to verify the estimation procedure using experimental data from SCAL type experiments.

THEORY AND PROCEDURE

This section describes the mathematical model (forward model), the inverse model and its solution.

Forward Model

Flow in porous media is determined by the distribution of pores, pore sizes and pore throats, the fluid-rock interaction, the fluid-fluid interaction, and properties of the fluids. On the macroscopic scale [5], the rock properties are described by the permeability $k(\boldsymbol{x})$, and the porosity $\phi(\boldsymbol{x})$, which vary spatially in the core. The fluid properties are described by the viscosities μ, and densities ρ, both known to vary with temperature. The rock-fluid interactions are described by the capillary pressure P_c, and the empirical functions k_{ri}, which take care of the reduced flow when immiscible fluids flow together. P_c and k_{ri} are usually given as functions of the wetting phase saturation, the fraction of pore space occupied by the wetting phase. The mathematical model of flow through porous media, assuming isotropic media fully saturated with two immersible fluids, e.g., oil and water (i=o,w), and constant viscosities and densities of the fluids, is

$$\phi(\boldsymbol{x})\frac{\partial S_i}{\partial t} = \nabla \cdot (\frac{k(\boldsymbol{x})k_{ri}(S_w)}{\mu_i}\nabla p_i) \quad (1)$$

$$p_o - p_w = P_c(S_w) \quad (2)$$

$$S_o + S_w = 1 \quad (3)$$

where S_i is the saturation of phase i and p_i is the phase pressure. Eqs.(1)-(3) are solved numerically with specified initial and boundary conditions [5]. A fully implicit black-oil simulator is used in this work.

To simulate dynamic data such as fluid production, in-situ pressure, pressure drop, and in-situ saturation, the reservoir parameters $k(\boldsymbol{x})$, $\phi(\boldsymbol{x})$, $P_c(S_w)$, and $\mathrm{k}_{ri}(S_w)$ must be specified. These parameters are generally not known, and are inaccessible to direct measurements. Moreover, they are known to differ from one porous medium to another. In this paper we want to demonstrate an inverse methodology for estimating $k(\boldsymbol{x})$ and $k_{ri}(S_w)$ assuming $\phi(\boldsymbol{x})$ and $P_c(S_w)$ to be known.

The Inverse Problem

Note that while discrete data are measured, the absolute and relative permeabilities are actually functions, so that their determination is an ill-posed problem. Regularization is effectively accomplished by finite-dimensional representations of the unknown functions [6]. The relative permeabilities are represented with third-order B-spline functions, which are known to be capable of representing any smooth function ar-

bitrarily accurately [7]:

$$k_{ri}(S_w) = \sum_{j=1}^{M_i} a_j^i B_j^m(S_w, \boldsymbol{y}^i), \quad i = o, w \qquad (4)$$

where o and w denote oil and water respectively. The functions are specified by the order m, the spline coefficients a_j, and the extended partition $\boldsymbol{y}$. The vector of the unknown parameters, $\boldsymbol{a}$, composed of the coefficients within the spline representation of the relative permeability functions, is given by

$$\boldsymbol{a} = [a_1^o, \ldots, a_{M_o}^o, a_1^w, \ldots, a_{M_w}^w] \qquad (5)$$

.

The absolute permeability is a rock property which is a function of position. In this paper the absolute permeability is represented by a hierarchical multi-scale basis function [8]

$$k_N(\boldsymbol{x}) = \boldsymbol{c}_N^T \boldsymbol{\psi}_N(\boldsymbol{x}) \qquad (6)$$

where $\boldsymbol{c}_N \in \Re^N$ is the parameter vector and $\boldsymbol{\psi}_N$ is an N-vector of real-valued piecewise constant basis functions spanning the space within which k is to be estimated.

An objective function is formulated as a weighted sum of squared differences between the measured data, $\mathbf{Y}$, and the corresponding simulated values, $\mathbf{F}$. Examples of such data in SCAL are: Fluid production, in-situ pressure, pressure drop, and in-situ saturation. In the estimation process, one seeks to minimize the objective function,

$$J(\boldsymbol{b}) = [\mathbf{Y} - \mathbf{F}(\boldsymbol{b})]^T W [\mathbf{Y} - \mathbf{F}(\boldsymbol{b})] \qquad (7)$$

the idea being that the simulated data should reconcile those actually measured. Here W is the inverse of the covariance matrix of the measurement errors. In the simultaneous estimation algorithm, $\boldsymbol{b}$ is represented by either relative permeability coefficients ($\boldsymbol{a}$) or permeability coefficients ($\boldsymbol{c}_N$).

Both the relative permeability and the absolute permeability are estimated through a regression based approach, in which the number of parameters is successively increased until the measured data are adequately represented by their corresponding calculated values [6]. The value of $\boldsymbol{b}$ that minimizes Eq.(7) is obtained through an implementation of the Levenberg-Marquardt optimization algorithm, which incorporates linear inequality constraints to enforce monotonous relative permeability functions and nonnegativity for the permeability [9].

Simultaneous Estimation

The estimation of absolute and relative permeability is combined by a successive estimation of $\mathbf{a}$ and $\boldsymbol{c}_N$. First the relative permeability is estimated by representing the entire sample by a single absolute permeability value. (A single value can easily be measured in a one-phase flow experiment, thus, the value is generally known in core analysis.) The absolute permeability is then estimated keeping the relative permeabilities fixed at their last estimate. The successive estimation of absolute and relative permeability is repeated until convergence. Thus, successively more accurate estimates are determined. Note that both the absolute and the relative permeabilities are estimated by the same estimation algorithm (solution of Eq.(7)).

The estimation of relative permeability has been reported several times before, see e.g. [4,6,10], and will not be discussed in more detail here. Next, the multiscale estimation of the absolute permeability will be described.

Multi-scale Estimation

In the multiscale estimation the number of parameters to be estimated is increased successively. In general, the total number of parameters available equals the number of grid cells used in solving the model equations, but estimating all these parameters is prohibitive as the data usually does not contain sufficient information to do so. Also, estimating more parameters requires more computation time, and increasing the number of parameters generally increases the parameter uncertainty.

An alternative is to evaluate all the possible combinations of parameters when increasing the number of available parameters successively. Performing all of these estimations will take inordinately long time as the number of combinations increases rapidly with the number of parameters. Also, each estimation requires a lot of computational work as the forward problem is typically calculated several

times during each estimation. Naturally, it would be preferable to predict the number of parameters and which parameters to use without performing all the estimations. Such a method has been developed in [11].

The methodology described in [11], utilizes the predicted objective function value, $\tilde{J}(P_Q)$, and its associated uncertainty $\tilde{\sigma}(\tilde{J}(P_Q))$. These functions are given by

$$\tilde{J}(P_Q) = \Delta \boldsymbol{e}^T (W - G) \Delta \boldsymbol{e}, \qquad (8)$$

$$G = \left[W A_Q \left(A_Q^T W A_Q \right)^{-1} A_Q^T W \right], \qquad (9)$$

$$\tilde{\sigma}(\tilde{J}(P_Q)) = \sqrt{4\tilde{J}(P_Q) - 2(M - Q)} \qquad (10)$$

where P_Q is the parameterization containing Q parameters. The sensitivity matrix for the parameterization P_Q is denoted A_Q, $\Delta \boldsymbol{e}$ is the current residual vector ($\Delta \boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{F}(\boldsymbol{b})$), and M is the number of observations. Eqs.(8)-(10) are functions based on linearization of the model $\boldsymbol{F}(\boldsymbol{b})$, and are therefore available by analytical calculations.

Which and how many basis functions to include in the estimation is performed in the following way. First a single value of the domain of definition is estimated, $Q = L = 1$. Then, from this point in the parameter space, reduction potentials are calculated. For each fixed number of parameters, one particular combination of parameters will produce the lowest $\tilde{J}$; this parameterization is called the local winner. $\tilde{J}$ and $\tilde{\sigma}$ of the local winners are used to decide how many parameters to include. We evaluate more parameters, $L = L + 1$, until $\tilde{J}_{L+1} + 2\tilde{\sigma}_{L+1} \geq \tilde{J}_L$. Then an estimation with $Q = L$ is performed. From this new point in the parameter space, more parameters are evaluated, see Fig.1 for an example of choosing parameters. For more details, see [11]. Note that $2\tilde{\sigma}$ is used as uncertainty measure in this paper. Next, we discuss the criteria to be fulfilled to accept an estimate.

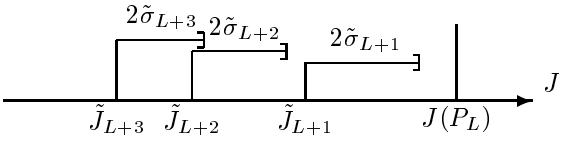


Figure 1: Example of ranking of predictions from refinement analysis. The original number of parameters is L. In this case Q=L+2, since $\tilde{J}_{L+3} + 2\tilde{\sigma}_{L+3} \geq \tilde{J}_{L+2}$ and $\tilde{J}_{L+2} + 2\tilde{\sigma}_{L+2} < \tilde{J}_{L+1}$.

Solution and Termination Criteria

In general, a solution of the inverse problem should not be accepted as an estimate unless it produces a match of the experimental data. To discard solutions where the experimental data are not reconciled, solution criteria to validate estimates of absolute and relative permeabilities from laboratory data have been established, see e.g. [4,6,12].

The objective function, J, given in Eq.(7), should be close to $M - Q$, where M is the number of measurements, and Q is the number of estimated parameters [6,13]. In addition the number of runs, (number of times the residuals change signs as one move through the data set) R, should be close to $M/2$ [12]. In addition to these acceptance criteria, we calculate the impact of measurement noise on the estimated functions by utilizing a linearized covariance analysis [4,12].

From the acceptance criteria given above, the criteria for termination of the simultaneous estimation follows. The sequence of estimations is terminated when the value of the objective function and the calculated number of runs can be explained by random measurement errors, i.e., the sequence is terminated when

$$J(P_Q) < (M - Q) + 2\sqrt{2(M - Q)} \qquad (11)$$

$$R(P_Q) < (M/2) + \sqrt{M} \qquad (12)$$

(see, e.g., [6]). The square root terms in Eq.(11) and Eq.(12) are just 2 standard deviations of $M - Q$ and $M/2$, respectively. For the absolute permeability, termination is also achieved whenever the criterion for introduction of new parameters advises against introducing more parameters.

For both the absolute and the relative permeabilities, the successive estimation sequence is terminated when the calculated objective function for an estimation is within two standard deviations of the objective function for

the previous estimation. The successive estimation is also terminated if the estimated permeability is within two decimal places of the previous estimate, and the same parameterization was used.

Estimation Algorithm

The simultaneous estimation algorithm can be summarized as follows:

Initialization:
(**1**) As the first estimate of the absolute permeability, use one single value for the entire sample. Estimate the relative permeability while keeping the absolute permeability fixed.

Estimate the absolute permeability:
(**2**) Estimate a constant value of the permeability. Set L=1.
(**3**) If the sequence is terminated, see Eqs.(11)-(12), go to step 6.
(**4**) Refinement analysis: Let L=L+1 until $\tilde{J}_{L+1} + 2\tilde{\sigma}_{L+1} \geq \tilde{J}_L$ then estimate the permeability with L parameters.
(**5**) If the sequence is terminated, see Eqs.(11)-(12), or if the absolute permeability value estimated is close to the previous estimate of the successive approach, or if the objective function is close to the one from the previous calculated in the successive approach, perform step 6 and stop. If the criterion advises against introducing more parameters, go to step 6. If not go to step 4.

Estimating the relative permeability:
(**6**) Estimate the relative permeability. If the solution criteria are satisfied, or if the new objective function value is within 2 standard deviations of the last relative permeability estimation objective function, stop, else return to step 2.

TEST OF THE METHOD

The estimation procedure described above is performed on one synthetic example to evaluate the performance of the estimation procedure. The true absolute permeability used is shown in Fig.2. The core is represented by 32 equally sized grid blocks in the horizontal direction. The absolute permeability is constant within the following groups of grid blocks: 1-10, 11-12, 13-14, 15-16, 17-18, 19-20, 21-22, 23-32. (1-10 corresponds to the value 50mD, 11-12 corresponds to the value 60mD and so on, in 10mD steps.)

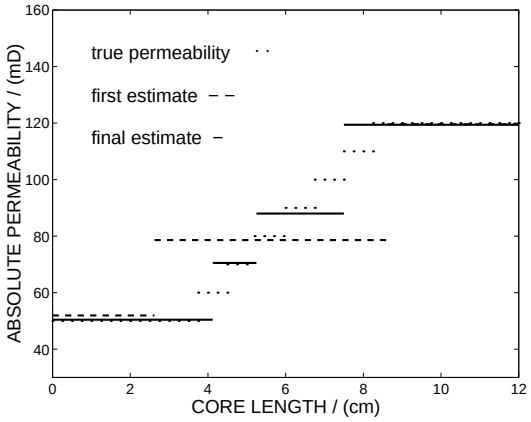


Figure 2: Absolute permeability.The true functions and the final estimate.

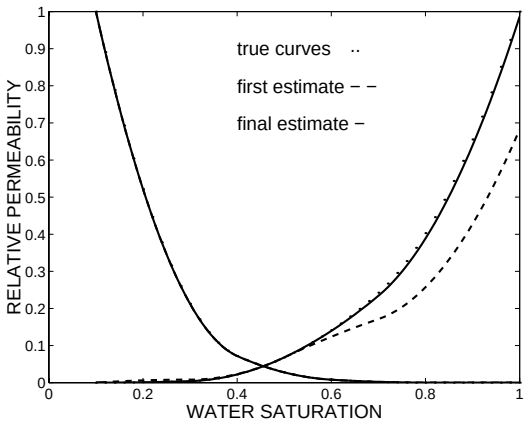


Figure 3: Relative permeability. The true curves, the first estimate and the final estimate.

The true relative permeability curves are shown in Fig.3. These relative permeability functions are used in all the grid blocks. The true capillary pressure is also equal for all the grid blocks. In practice the capillary pressure will vary in a heterogeneous core. We have here idealized the situation, to concentrate on the effect of permeability heterogeneity alone.

The core and fluid properties listed in Table 1 are together with the true absolute and

Table 1: Fixed medium and fluid properties.

Reservoir length	12 cm
Cross section area	10 cm^2
Water viscosity	$0.85 \cdot 10^{-3}$ Pa·s
Oil viscosity	$0.66 \cdot 10^{-3}$ Pa·s
Water injection rates	0.9, 0.5 ml/min
Oil injection rates	0.1, 0.5, 1 ml/min

relative permeability functions implemented in the core flow simulator. Synthetic experimental data are produced using no-flow boundary conditions, fixed pressure at the outlet, and constant injection rates. A primary drainage experiment is simulated where the water saturation in the core is reduced in steps due to the oil injection rate. Gaussian noise with zero mean is then added to the data to mimic measurement error. In this work the synthetic data consist of: Water saturation, water production, differential pressure, and well pressure in grid block 1, 8, and 24. These data can be seen in Fig.4.

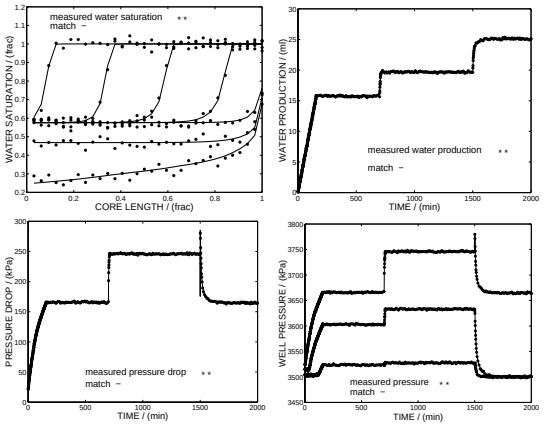


Figure 4: The synthetic experimental data, and the match produced by the final estimates.

The simultaneous estimation is performed according to the estimation algorithm given above. For this example the relative permeability was estimated seven times, and the absolute permeability was estimated six times. The estimation terminated while the calculated objective function of the sixth absolute permeability estimate was within two standard deviations of the calculated objective function of the fifth absolute permeability estimate.

To show how the absolute permeability estimation has evolved, both the first and the final estimates are shown in Fig.2. The final estimate does not coincide with the true permeability partition-wise, but it does provide a better match of the true distribution than the first estimate. For the final estimate, J and R are within three standard deviations of their limits. By keeping this estimate fixed, the final relative permeability estimation is made.

To show how the relative permeability estimate has improved throughout the successive estimation, both the first and the final estimates are shown in Fig.3. J and R are within 3 standard deviations of their limits for the final estimate. The match of the synthetic experimental data produced by the final estimates can be seen in Fig.4.

USE OF EXPERIMENTAL DATA

An experiment is designed and performed to test the simultaneous estimation method on experimental data.

Experimental Setup

When conducting a flooding experiment, it is important to arrive at experimental conditions that leads to accurate property identification. An experimental design procedure has been developed to do this. Through the experimental design [4], where the accuracy of the flow functions can be determined for different experiments, we determined the injection rates (see Table 2) and the corresponding injection times (see Fig.5). (The details of the experimental design will not be given here.)

Table 2: Fixed medium and fluid properties.

Reservoir length	4.7 cm
Cross section area	2.5 cm^2
Water viscosity	$1.0 \cdot 10^{-3}$ Pa·s
Oil viscosity	$1.125 \cdot 10^{-3}$ Pa·s
Oil injection rates	0.04, 0.3, 1.0, 2.0 ml/min

The core sample was initially fully saturated with water. The pressure drop across the sample was monitored by a differential pressure

gauge, and the water production from the core was found by collecting the water in a cylinder and measuring the increasing hydraulic pressure of the column.

The two-phase experiment was performed on a composite sample. The composite was constructed by joining two Bentheimer sandstone samples. Before conducting the samples, a one phase experiment was performed on both to determine the absolute permeability of the two parts in the composite. A single phase experiment was also conducted on the composite sample. The permeability of the composite sample is 992mD±157mD, while the permeabilities of the two cores in the composite are 1199mD±119mD and 1348mD±101mD. The permeability of the composite is less than the values of the two cores due to the sand layer placed between the two cores in the composite to maintain capillary contact. We also determined the average porosity (gravimetrically), which was measured to be 24%. Additional core and fluid properties are listed in Table 2.

Simultaneous Estimation From Experimental Data

The pressure drop and water production data collected from the two-phase experiment are shown in Fig.5. We attempt to use these data to verify our estimation method.

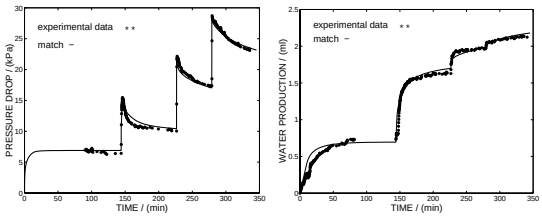


Figure 5: The experimental data and the match produced by the estimated functions.

The simultaneous estimation starts by estimating the relative permeability functions keeping the absolute permeability fixed at 992mD, the value determined by the single-phase experiment on the composite. Since we do not know the capillary pressure, that also needs to be estimated initially. The initial estimates of the relative permeability and the capillary pressure are shown in Fig.6 and Fig.7,

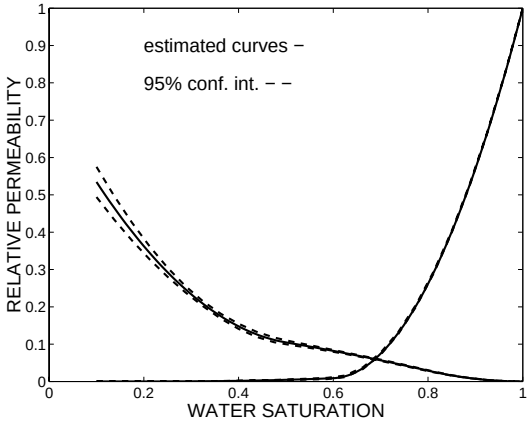


Figure 6: Estimated relative permeability.

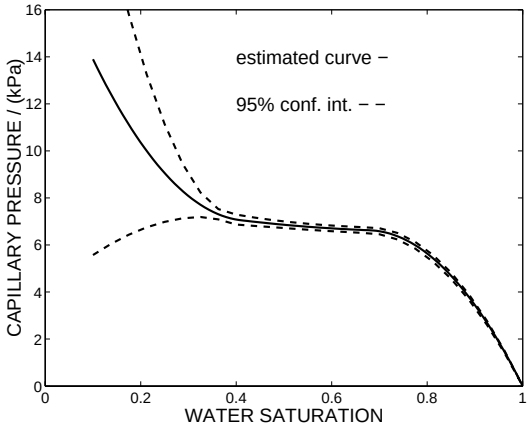


Figure 7: Estimated Capillary pressure.

respectively. To arrive at these estimates, the regression based method has been used [6].

The analysis of how many and which parameters to include when estimating the absolute permeability is performed. We arrive at estimating a single parameter. We then return to the relative permeability. The new estimate of the relative permeabilities has a objective function within two standard deviation of the objective function of the previous relative permeability estimate. Thus, the simultaneous estimation is terminated.

Fig.6 shows the final relative permeability estimate, and the constant value of 985mD is the final absolute permeability estimate. The match of the experimental data produced by these estimates is shown in Fig.5.

The methodology works well for the experimental case in the sense that the methodology successfully estimated a permeability close to the measured single value, 992mD±157mD, and flow properties with narrow confidence bands. It would be preferable to include saturation data, and the advantages of the method could be demonstrated better with a greater difference in the permeability values.

CONCLUSIONS

(**1**) The method for successive estimation of absolute and relative permeability has been extended in two ways: We have included a more flexible and faster algorithm for absolute permeability estimation, and we have included saturation data in the estimation problem.
(**2**) An algorithm for selection of parameterization of the absolute permeability has been used. The method has earlier proved successful in estimating absolute permeability (2D) from pressure data [11]. The method significantly reduces the risk of over-parameterization in multiscale representation, and saves the computational work associated with use of an excessive number of parameters in the estimation.
(**3**) We have performed an experimental test of the successive estimation algorithm. We were able to successfully estimate absolute and relative permeability from the acquired data.

ACKNOWLEDGMENTS

The authors thank the Norwegian Research Council's petroleum research program, Petro-Forsk, and the US Department of Energy for financial support of this work. The authors also thank Petec AS for use of the black-oil simulator Sendra.

REFERENCES

[**1**] Valestrand, R., Grimstad, A.-A., Kolltveit, K., Nævdal, G., Nordtvedt, J.-E., "Simultaneous Determination of Absolute and Relative permeabilities." *Proceedings of the 68 Eurotherm International conference,* 2000.
[**2**] Meija, G.M., Mohanty, K.K., Watson, A.T., "Use of In-Situ Saturation Data in Estimation of Two-Phase Flow Functions in Porous Media," *J. of Petroleum Science and Engineering,* **12**, pp. 233–245, 1995.
[**3**] Seto, K., Hollenshead, J.T., Watson, A.T., Chang, C.T.P., Slattery, J.C., "Permeability Determination Using NMR Imaging," *Transport in Porous Media,* **42**, pp. 351–388, 2001.
[**4**] Urkedal, H., *"Design of Multiphase Flow Experiments,"* Dr. Scient thesis, Department of Physics, University of Bergen, 1998.
[**5**] Aziz, K., and Settari, A.: *"Petroleum Reservoir Simulation"*, Applied Science, London, 1979.
[**6**] Grimstad, A.-A., Kolltveit, K.K., Nordtvedt, J.-E., Watson, A.T., Mannseth, T., Sylte, A., "The Uniqueness and Accuracy of Porous Media Multiphase Properties Estimated from Displacement Experiments," *Proc. SCA International Symposium,* SCA-9709, 1997.
[**7**] Scumaker, L.L., *"Spline Functions: Basic Theory"*, J. Wiley & Sons Inc., New York City, 1981.
[**8**] Chavent, G., Liu, J., "Multiscale parameterization for the estimation of a diffusion coefficient in elliptic and parabolic problems," *Proc. in fifth IFAC Symposium on Consol of Distributed Parameter Systems,* 1998.
[**9**] Gill, P.E., Murray, W., Wrigth, M.H., *"Practical Optimization."* Academic Press, London, 1981.
[**10**] Kulkarni, K., Watson, A.T., Nordtvedt, J.-E., Sylte, A., "Two Phase Flow in Porous Media: Property Identification and Model Validation," *AIChE Journal,* Vol.44, No.11, pp.2337-2350, 1998.
[**11**] Grimstad, A.-A., Mannseth, T., Nævdal, G., Urkedal, H., "Adaptive multiscale permeability estimation", *Computational Geosciences,* to appear.
[**12**] Bard, Y., *"Nonlinear Parameter Estimation,"* Academic Press. Washington DC, p. 201, 1974.
[**13**] Weisberg, S., *"Applied Linear Regression,"* John Wiley & Sons, New York, 1995.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

USE OF A SINGLE HEATED SURFACE FOR THE ESTIMATION OF THERMAL CONDUCTIVITY COMPONENTS OF ORTHOTROPIC 3D SOLIDS

F. A. Rodrigues, H. R. B. Orlande and M. M. Mejias
Department of Mechanical Engineering, DEM/PEM - POLI/COPPE
Federal University of Rio de Janeiro, UFRJ
Rio de Janeiro, RJ, Brazil
fabio@lttc.coppe.ufrj.br, helcio@serv.com.ufrj.br and mejias@lmt.coppe.ufrj.br

ABSTRACT

This paper deals with the experimental design and with the solution of an inverse parameter estimation problem, for the identification of the three thermal conductivity components of an orthotropic solid. The experimental setup examined here involves the heating of a specimen in the form of a parallelepiped, through part of one of its surfaces. Different experimental variables are chosen by using the D-optimum criterion, including the number and location of sensors and the size of the heating surface. The unknown parameters are estimated with the Levenberg-Marquardt method of minimization of the least-squares norm.

NOMENCLATURE

a, b, c	body dimensions along the x, y and z directions, respectively
$\mathbf{F}_I$	information matrix
F_x, F_y, F_z	normalized sensitivity coefficients for the thermal conductivity components along the x, y and z directions, respectively
k_x, k_y, k_z	thermal conductivity components along the x, y and z directions, respectively
$\mathbf{P}$	vector of unknown parameters
$S(\mathbf{P})$	ordinary least squares norm
$\mathbf{T}$	vector of estimated temperatures
$\mathbf{Y}$	vector of measured temperatures

INTRODUCTION

Due to the importance of non-isotropic materials in nowadays engineering, such as composites, a lot of attention has been devoted in the recent past for the identification of their thermal properties through inverse analysis techniques of parameter estimation [1-11].

The choice of various optimal experimental variables, including the number and locations of sensors, as well as the heating and final times, for the model used by Sawaf and Özisik [1] for the estimation of thermal conductivity components of orthotropic solids, was addressed by Mejias et al [8]. In these two works[1,8], the solid was considered in the form of a cube, with three of its surfaces heated by fluxes of equal magnitude, while the other three surfaces were assumed as insulated. Later, Mejias et al [10,11] have examined in the experimental design the choice of the boundary conditions for the non-heated surfaces of the cube, as well as the choice of the magnitudes of the applied heat fluxes and of the body dimensions. It was found [10] that the use of constant temperature boundary conditions, instead of insulated boundary conditions, could result on more accurate estimates for the unknown parameters. It was also found [11] that the body dimensions and the magnitudes of the applied heat fluxes should be chosen in accordance with a transformation that maps the orthotropic solid into an isotropic solid [12]. In this case, the three thermal conductivity components could be estimated with identical relative accuracies. Despite such an important feature for the experimental design, the implementation of the experimental setup devised was not simple, because it involved the uniform heating of three surfaces of the solid and the magnitudes of the applied heat fluxes were assumed as known for the inverse analysis.

In order to overcome such difficulties, we examine in this paper an alternative experimental setup for the estimation of the three thermal conductivity components of orthotropic solids. In such setup, only one of the surfaces of the solid would be partially heated. It would consist of a heater placed between two identical specimens of

the material with unknown thermal conductivity components, thus allowing for better control of the magnitude of the applied heat flux, as well as of its distribution. We note that a 2D version of this setup was examined by Taktak [3]. The D-optimum criterion [13-16] is used here for the experimental design, involving the choice of different variables like the number and location of sensors and the size of the heating surface. The *Levenberg-Marquardt Method* of minimization of the least-squares norm [13,17-19] is applied as the estimation procedure by using simulated measurements with random errors, as described below.

DIRECT PROBLEM

The physical problem considered in this work consists of a parallelepiped with dimensions a^*, b^* and c^*, in the x^*, y^* and z^* directions, respectively. The solid is initially at the uniform temperature T_0^*. For times $t^* > 0$, the boundary at $z^*=0$ is partially heated in the surface defined by $0 < x^* < a_1^*$ and $0 < y^* < b_1^*$, while the remaining surface of this boundary is supposed to be insulated. The boundary at $z^*=c^*$ is kept at a constant and uniform temperature T_0^* and the other four boundaries are supposed insulated. The solid is assumed to be orthotropic, with thermal conductivity components k_x^*, k_y^* and k_z^* in the x^*, y^* and z^* directions, respectively, and the physical properties are assumed constant. The mathematical formulation of such physical problem is given in dimensionless form as

$$k_x \frac{\partial^2 T}{\partial x^2} + k_y \frac{\partial^2 T}{\partial y^2} + k_z \frac{\partial^2 T}{\partial z^2} = \frac{\partial T}{\partial t}$$

$$\text{in } 0 < x < a; \quad 0 < y < b; \quad 0 < z < c; \quad t > 0 \qquad (1.a)$$

$$\frac{\partial T}{\partial x} = 0 \text{ at } x = 0 \text{ and } x = a, \text{ for } t > 0 \qquad (1.b,c)$$

$$\frac{\partial T}{\partial y} = 0 \text{ at } y = 0 \text{ and } y = b, \text{ for } t > 0 \qquad (1.d,e)$$

$$-k_z \frac{\partial T}{\partial z} = q(x,y,t) \text{ at } z = 0, \text{ for } t > 0 \qquad (1.f)$$

$$T = 0 \quad \text{at} \quad z = c, \text{ for } t > 0 \qquad (1.g)$$

$$T = 0 \quad \text{for} \quad t = 0, \text{ in the region} \qquad (1.h)$$

where

$$q(x,y,t) = \begin{cases} \overline{q_0}(t), \text{ in } 0 < x < a_1 \text{ and } 0 < y < b_1 \\ 0 \quad , \quad \text{outside this region} \end{cases} \qquad (2)$$

The time-dependent heat flux in equation (2) is given by:

$$\overline{q_0}(t) = \begin{cases} q_0 \quad , \quad \text{for } 0 < t < t_h \\ 0 \quad , \quad \text{for } t > t_h \end{cases} \qquad (3)$$

where t_h is the dimensionless heating time.

The following dimensionless variables were defined in order to write the physical problem in dimensionless form:

$$T = \frac{(T^* - T_0^*)k_{ref}^*}{q_{ref}^* l^*}; \; t = \frac{k_{ref}^* t^*}{r\, c_p\, l^{*2}}; \; t_h = \frac{k_{ref}^* t_h^*}{r\, c_p l^{*2}}$$

$$q_0 = \frac{q_0^*}{q_{ref}^*}; x = \frac{x^*}{l^*}; \; y = \frac{y^*}{l^*}; \; z = \frac{z^*}{l^*}$$

$$a = \frac{a^*}{l^*}; b = \frac{b^*}{l^*}; \; c = \frac{c^*}{l^*}; \; a_1 = \frac{a_1^*}{l^*}; \; b_1 = \frac{b_1^*}{l^*}$$

$$k_x = \frac{k_x^*}{k_{ref}^*}; \; k_y = \frac{k_y^*}{k_{ref}^*}; \; k_z = \frac{k_z^*}{k_{ref}^*} \qquad (4.a\text{-}o)$$

In equations (4.a-o), k_{ref}^*, q_{ref}^* and l^* are characteristic values for thermal conductivity, heat flux and length, respectively. The superscript "*" appearing above denotes dimensional quantities, while r and c_p are the density and specific heat of the body, respectively.

In the *direct problem* associated with the physical problem described above, the three thermal conductivity components k_x, k_y and k_z, as well as the solid geometry, initial and boundary conditions, are known. The objective of the direct problem is to determine the transient temperature field $T(x,y,z,t)$ in the body.

INVERSE PROBLEM

For the *inverse problem* considered here, the thermal conductivity components k_x, k_y and k_z are regarded as unknown, while the other quantities appearing in the formulation of the direct problem described above are assumed to be known with high degree of accuracy.

For the estimation of the *vector of unknown parameters* $\mathbf{P}^T = [k_x, k_y, k_z]$, we assume available the transient readings of M temperature sensors. The temperature measurements may contain random errors. Such errors are assumed to be

additive, uncorrelated and normally distributed, with zero mean and a known constant standard-deviation s. Also, for the estimation procedure we consider that no previous information regarding the parameter values is available. Therefore, the solution of the present parameter estimation problem can be obtained through the minimization of the ordinary least-squares norm, which is a minimum variance estimator in this case [13]. The ordinary least-squares norm is given by

$$S(\mathbf{P})=[\mathbf{Y}-\mathbf{T}(\mathbf{P})]^T[\mathbf{Y}-\mathbf{T}(\mathbf{P})] \qquad (5)$$

where

$$[\mathbf{Y}-\mathbf{T}(\mathbf{P})]^T=[\vec{Y}_1-\vec{T}_1(\mathbf{P}),\vec{Y}_2-\vec{T}_2(\mathbf{P}),\cdots,\vec{Y}_I-\vec{T}_I(\mathbf{P})] \qquad (6.a)$$

and each element $[\vec{Y}_i-\vec{T}_i(\mathbf{P})]$ is a row vector of length equal to the number of sensors M, that is,

$$[\vec{Y}_i-\vec{T}_i(\mathbf{P})]=[Y_{i1}-T_{i1}(\mathbf{P}),Y_{i2}-T_{i2}(\mathbf{P}),\cdots,Y_{iM}-T_{iM}(\mathbf{P})] \quad \text{for } i=1,..,I \qquad (6.b)$$

We note that Y_{im} and $T_{im}(\mathbf{P})$ are the measured and estimated temperatures, respectively, for time t_i, i=1,…,I, and for the sensor m, m=1,…,M. The estimated temperatures are obtained from the solution of the direct problem by using the current available estimate for the vector of unknown parameters $\mathbf{P}^{\mathrm{T}} = [k_x , k_y , k_z]$.

For the minimization of the least squares norm (5), we apply the *Levenberg-Marquardt Method* [13,17-19]. The iterative procedure of such a method is given by:

$$\mathbf{P}^{k+1}=\mathbf{P}^k+[(\mathbf{J}^k)^T\mathbf{J}^k+m^k\Omega^k]^{-1}(\mathbf{J}^k)^T[\mathbf{Y}-\mathbf{T}(\mathbf{P}^k)] \qquad (7)$$

where μ^k is a positive scalar named *damping parameter*, $\mathbf{\Omega}^k$ is a *diagonal matrix* and $\mathbf{J}^k$ is the *sensitivity matrix*.

The purpose of the matrix term $\mu^k\mathbf{\Omega}^k$ in equation (7) is to damp oscillations and instabilities due to the ill-conditioned character of the problem, by making its components large as compared to those of $\mathbf{J}^T\mathbf{J}$, if necessary [13]. The damping parameter is made large in the beginning of the iterations. With such an approach, the matrix $\mathbf{J}^T\mathbf{J}$ is not required to be non-singular in the beginning of iterations and the Levenberg-Marquardt Method tends to the *Steepest Descent Method*, that is, a very small step is taken in the negative gradient direction. The parameter μ^k is then gradually reduced as the iteration procedure advances to the solution of the parameter estimation problem and then the Levenberg-Marquardt Method tends to the *Gauss Method* [13]. However, if the errors inherent to the measured data are amplified, generating instabilities on the solution as a result of the ill-conditioned character of the problem, the damping parameter is automatically increased. Such an automatic control of the damping parameter makes the Levenberg-Marquardt method a quite robust and stable estimation procedure, so that it does not require the use of the *Discrepancy Principle* in the stopping criterion to become stable, like the conjugate gradient method [19].

STATISTICAL ANALYSIS

By performing a statistical analysis it is possible to assess the accuracy of $\hat{P}_j$, j = 1, 2 and 3, which are the values estimated for the unknown parameters P_j, j = 1, 2 and 3. By taking into account the statistical hypotheses described above, the *covariance matrix* for the ordinary least-squares estimator is given by [13]:

$$\mathbf{V}\equiv\begin{bmatrix}\mathrm{cov}(\hat{P}_1,\hat{P}_1) & \mathrm{cov}(\hat{P}_1,\hat{P}_2) & \mathrm{cov}(\hat{P}_1,\hat{P}_3)\\ \mathrm{cov}(\hat{P}_2,\hat{P}_1) & \mathrm{cov}(\hat{P}_2,\hat{P}_2) & \mathrm{cov}(\hat{P}_2,\hat{P}_3)\\ \mathrm{cov}(\hat{P}_3,\hat{P}_1) & \mathrm{cov}(\hat{P}_3,\hat{P}_2) & \mathrm{cov}(\hat{P}_3,\hat{P}_3)\end{bmatrix}=(\mathbf{J}^T\mathbf{J})^{-1}s^2 \qquad (8)$$

where $\mathbf{J}$ is the sensitivity matrix and σ is the standard deviation of the measurement errors, which is assumed to be constant. We note that equation (8) is exact for linear estimation problems and is approximately used for nonlinear parameter estimation problems.

The *standard deviations* for the estimated parameters can thus be obtained from the diagonal elements of $\mathbf{V}$ as

$$s_j\equiv\sqrt{\mathrm{cov}(\hat{P}_j,\hat{P}_j)}=\sqrt{V_{jj}} \quad \text{for } j = 1, 2 \text{ and } 3 \qquad (9)$$

where V_{jj} is the j^{th} element in the diagonal of $\mathbf{V}$.

Confidence intervals at the 99% confidence level for the estimated parameters can be obtained as

$$\hat{P}_j-2.576\,s_j\le P_j\le\hat{P}_j+2.576\,s_j \quad \text{for } j = 1, 2 \text{ and } 3 \qquad (10)$$

The *joint confidence region* for the estimated parameters is given by [13]:

$$(\hat{\mathbf{P}}-\mathbf{P})^T \mathbf{V}^{-1}(\hat{\mathbf{P}}-\mathbf{P}) \le c_3^2 \qquad (11)$$

where c_3^2 is the value of the chi-square distribution with 3 degrees of freedom for a given probability.

DESIGN OF OPTIMUM EXPERIMENTS

Optimum experiments can be designed by minimizing the hypervolume of the confidence region of the estimated parameters, in order to ensure minimum variance for the estimates. The minimization of the confidence region given by equation (11) can be obtained by maximizing the determinant of $\mathbf{V}^{-1}$, in the so-called *D-optimum design* [13-16]. Since the covariance matrix $\mathbf{V}$ is given by equation (8), we can then design optimum experiments by maximizing the determinant of the so-called Fisher's Information Matrix, $\mathbf{J}^T\mathbf{J}$ [13-16]. Therefore, optimal experimental variables are chosen based on the criterion

$$\max \left| \mathbf{J}^T \mathbf{J} \right| \qquad (12)$$

For cases involving a single sensor, each element $\mathbf{F}_{r,s}$, r,s = 1, 2 and 3, of the matrix $\mathbf{F} \equiv \mathbf{J}^T\mathbf{J}$ is given by:

$$\mathbf{F}_{r,s} \equiv [\mathbf{J}^T\mathbf{J}]_{r,s} = \sum_{i=1}^{I} \left(\frac{\partial T_i}{\partial P_r} \right) \left(\frac{\partial T_i}{\partial P_s} \right) \quad \text{for } r,s = 1, 2 \text{ and } 3 \qquad (13)$$

where I is the number of measurements.

If we take into account constraints, such as a large but fixed number of transient measurements of M sensors and also the maximum temperature in the region, T_{max}, we can choose to maximize the determinant of a normalized form of $\mathbf{F}$, here denoted as $\mathbf{F}_1$ [13], the elements of which are given by:

$$[\mathbf{F}_1]_{r,s} = \frac{1}{M t_f} \sum_{m=1}^{M} \int_{t=0}^{t_f} \left(P_r \frac{\partial T_m}{\partial P_r} \right) \left(P_s \frac{\partial T_m}{\partial P_s} \right) \left(\frac{1}{T_{\max}} \right)^2 dt \quad \text{for } r,s = 1, 2 \text{ and } 3 \qquad (14)$$

where t_f is the duration of the experiment.

RESULTS AND DISCUSSIONS

For the results presented below we assume the solid with unknown thermal conductivity components in the form of a cube, so that the characteristic length used to obtain the dimensionless problem is taken as $l^*=a^*=b^*=c^*$. Also, the characteristic heat flux is taken as $q_0^* = q_{ref}^*$, so that q_0=1 in equation (3).

We note that for nonlinear estimation problems, such as the one under picture in this work, the analyses of the sensitivity coefficients and of the determinant of $\mathbf{F}_1$ are not global, because these quantities are functions of the unknown parameters. Therefore, *a priori* estimated values for the parameters are required for the design of optimum experiments. For the test-cases examined in this paper, we used the following dimensionless values for the thermal conductivity components in order to design the experiment: k_x=1, k_y=15 and k_z=15. Such values were chosen based on the work of Dowding et al [5] for the estimation of the thermal conductivity components of two-dimensional carbon-carbon composites. The dimensionless thermal conductivity k_x=1 implies that $k_{ref}^* = k_x^*$. For the design of optimum experiments, involving the maximization of the determinant of ($\mathbf{J}^T\mathbf{J}$), we took into account the constraints of a large and fixed number of measurements for each sensor, as well as the maximum temperature in the region, so that the matrix $\mathbf{F}_1$ with elements given by equation (14) was used in the analysis.

Before proceeding to the analysis of the transient variation of the determinant of the information matrix, let us examine the transient variation of the normalized sensitivity coefficients for different sensor positions. The sensor positions examined in this work are presented in Table 1. They include sensors located at the center of the lateral thermally insulated surfaces at y=1 (sensor 1) and x=1 (sensor 2), as well as sensors located on the surface in contact with the heater at z=0 (sensors 3, 4 and 5).

Table 1. Sensor Locations

Sensor	Sensor Location
1	(0.5, 1, 0.5)
2	(1, 0.5, 0.5)
3	(0.25, 0.25, 0)
4	(0.5, 0.5, 0)
5	(0.75, 0.75, 0)

The normalized sensitivity coefficients were obtained by multiplying the original sensitivity coefficients by the parameters that they are

referred to. The normalized sensitivity coefficients with respect to the parameters k_x , k_y and k_z are defined respectively as:

$$F_x = k_x \frac{\partial T}{\partial k_x} \ , \quad F_y = k_y \frac{\partial T}{\partial k_y} \ , \ F_z = k_z \frac{\partial T}{\partial k_z} \quad (15)$$

Figures 1.a-e present the normalized sensitivity coefficients for sensors 1 to 5, respectively, by assuming the size of the heating surface as a_1=b_1=0.75 (see equation (2)). The heating time was supposed to be t_h=1(see equation (3)). Figures 1.a-e show a strong tendency towards linear dependence of the sensitivity coefficients for all sensor positions. Therefore, it is not possible to estimate the three unknown parameters with the readings of a single sensor. On the other hand, it may be possible to estimate the parameters with the readings of more than one sensor, because the constants of proportionality between the sensitivity coefficients are not the same, at the different sensor positions. Note in figures 1.a-e that the sensitivity coefficients tend to zero after the heating is stopped at t=1. However, before becoming zero, the sensitivity coefficient with respect to k_z undergoes a small increase for sensor 2, as depicted in figure 1.b. It can be noticed in figures 1.a-e that, generally, the sensitivity coefficient with respect to k_z has the largest magnitude among the sensitivity coefficients, for any sensor position. The magnitude of F_z is larger for the sensors 3, 4 and 5, located at z=0. However, note in figures 1.c-e that the magnitude of F_z decreases when the sensor is located towards the edge of the heating surface. The sensitivity coefficient F_z is negative for sensors 3, 4 and 5, because an increase in the thermal conductivity component k_z tends to decrease the temperature over the heating surface. Figures 1.c-e also show that the sensitivity coefficients with respect to k_x and k_y are negative for sensors 3 and 4, but positive for sensor 5. Such is the case because an increase in the thermal conductivity components in the x and y directions tends to decrease the temperature of sensors 3 and 4 located over the heating surface, but tends to increase the temperature of sensor 5 located at the edge of the heating surface. The largest magnitude of F_x takes place for sensor 2, located at the surface x=1. Similarly, the largest magnitude of F_y takes place for sensor 1, located at the surface y=1, although the magnitude of F_y for sensor 3 is practically the same as for sensor 1.

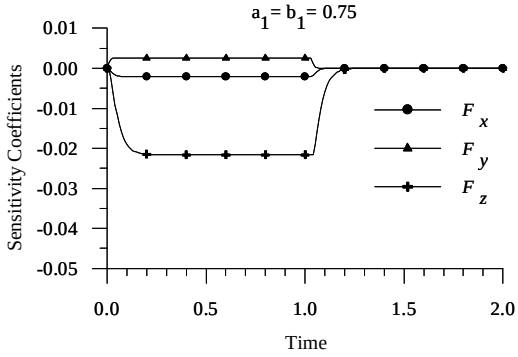


Figure 1.a – Normalized sensitivity coefficients for sensor 1

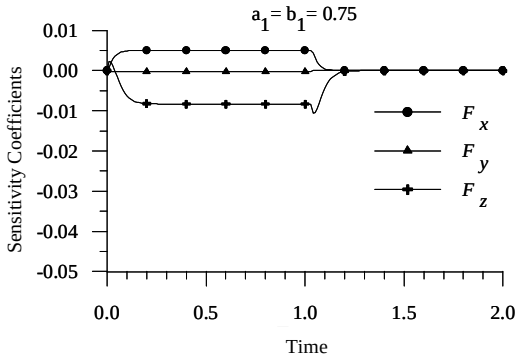


Figure 1.b – Normalized sensitivity coefficients for sensor 2

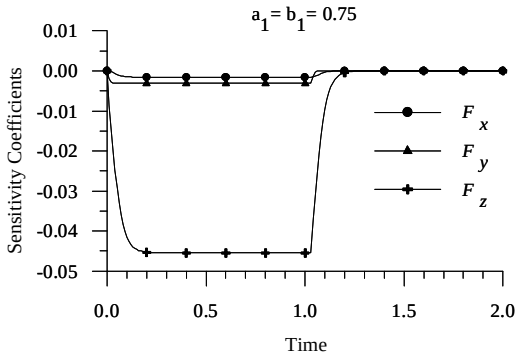


Figure 1.c – Normalized sensitivity coefficients for sensor 3

Figures 2.a-d present the transient variation of the determinant of the information matrix for different heating times, by assuming available the measurements of four different sensor configurations, which are presented in table 2. Configuration 1 consists of the sensors located at the insulated surfaces, sensor 1 at y=1 and sensor 2 at x=1, which have the largest magnitudes for F_y and F_x, respectively. Configurations 2-4 were

composed by sensors 1, 2 and of one of the sensors at the boundary z=0, i.e., sensor 3 for configuration 2, sensor 4 for configuration 3 and sensor 5 for configuration 4. Sensors 3, 4 and 5 have the largest magnitudes for F_z. The results shown in figures 2.a-d were obtained for a_1=b_1=0.75.

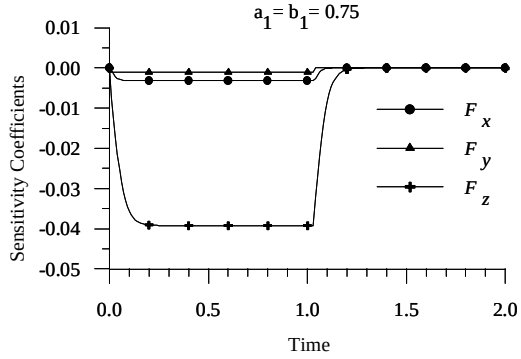


Figure 1.d – Normalized sensitivity coefficients for sensor 4

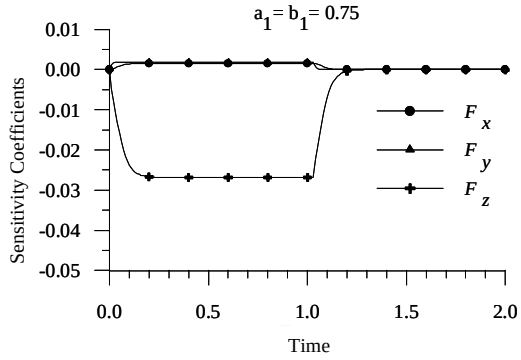


Figure 1.e – Normalized sensitivity coefficients for sensor 5

Table 2 – Configurations for the sensors

Configuration	Sensors
1	1 and 2
2	1, 2 and 3
3	1, 2 and 4
4	1, 2 and 5

Figure 2.a shows that the determinant of the information matrix undergoes a sudden increase at the moment that the heating is stopped, for configuration 1. This is a result of the change of the shape of the sensitivity coefficient with respect k_z for sensor 2 when the heating is stopped, which undergoes a small increase before tending to zero (see figure 1.b). For configuration 1, the optimum heating and final times are 0.2 and 0.24, respectively, when the determinant of the information matrix is maximum.

A comparison of figures 2.a-d shows that the magnitude of the determinant of the information matrix increases when a sensor at z=0 is used in the analysis, in addition to sensors 1 and 2, with the exception for configuration 4 (see figure 2.d). The configuration with the maximum determinant of the information matrix is configuration 2, because sensor 3 has the largest magnitude for the sensitivity coefficient with respect to k_z (see figures 1.c-e). As the sensitivity coefficient with respect to k_z decreases, when the sensor at z=0 is moved towards the edge of heating surface, the maximum determinant of the information matrix decreases and, for configuration 4, the effects of the use of the measurements of a sensor at z=0 are negligible. We note in figures 2.b,c that, for configurations 2 and 3, the determinant of the information matrix decreases at the moment that the heating is stopped. Such is the case because the sensitivity coefficients tend to zero very fast when the heating is stopped, with the exception of F_z for sensor 2, as depicted in figures 1.a-e. Therefore, for these configurations the effect of the sudden decrease of the sensitivity coefficients when the heating is stopped is more significant than the change in the shape of F_z for sensor 2. As a result, the heating time shall be chosen identical to the final time for configurations 2 and 3. Note also in figures 2.b,c that the determinant of the information matrix does not increase significantly for $t > 1$.

Configuration 2 is the best among those examined, because of the maximum value obtained for the determinant of the information matrix. Hence, more accurate estimates are expected with configuration 2 than with the other configurations. For the results presented below we have used the measurements of sensors arranged in accordance with configuration 2 and assumed $t = t_h = 1$.

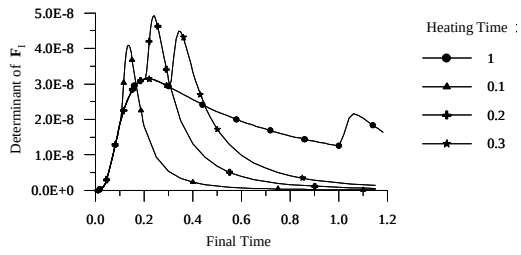


Figure 2.a – Determinant of the information matrix for the sensor configuration 1.

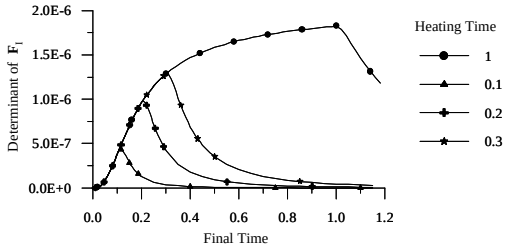


Figure 2.b – Determinant of the information matrix for the sensor configuration 2.

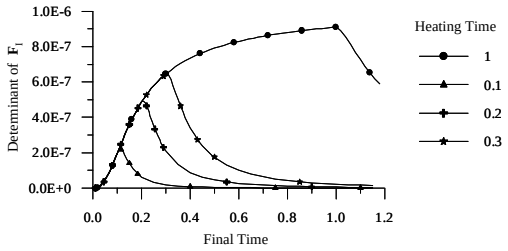


Figure 2.c – Determinant of the information matrix for the sensor configuration 3.

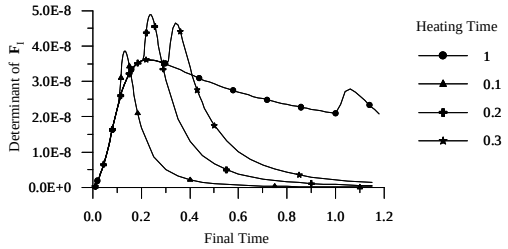


Figure 2.d – Determinant of the information matrix for the sensor configuration 4.

Let us now examine the effects of the size of the heating surface on the maximum determinant of the information matrix. Figure 3 shows the maximum values of such determinant, obtained with configuration 2 for $t = t_h = 1$ and for different sizes of the heating surface. Figure 3 shows that the optimum size of the heating surface is $a_1=b_1=0.75$. For smaller sizes of such surface, the sensitivity coefficients decrease in magnitude and become more linearly-dependent. For larger sizes of the heating surface, the physical problem tends towards a one-dimensional problem, irrespective of the thermal conductivity components in the x and y directions. We note that the best sensor configuration was not changed for the different sizes of the heating surface.

An alternative arrangement of the thermal conductivity components along the three principal axes would be $k_x=15$, $k_y=15$ and $k_z=1$, instead of $k_x=1$, $k_y=15$ and $k_z=15$ as examined above. However, the maximum determinant of the information matrix was smaller for this alternative arrangement because of smaller magnitudes of the sensitivity coefficients. The effects of the body dimensions and of the magnitude of the applied heat flux were also examined, but they are not discussed here for the sake of brevity.

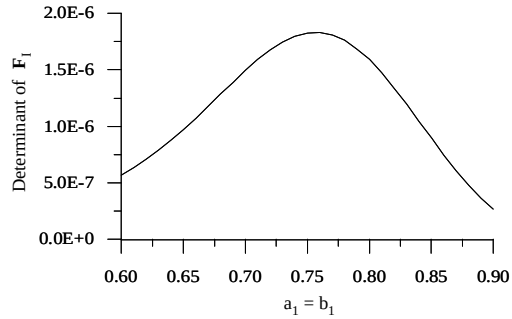


Figure 3 – Effects of the size of the heating surface on the determinant of the information matrix for $t_f = t_h = 1$ and sensor configuration 2.

We now examine the estimation of the three thermal conductivity components of orthotropic solids, by using 100 transient simulated measurements per sensor, of three sensors arranged in accordance with configuration 2. The measurements were taken during $0 < t < 1$. The initial guesses used for the Levenberg-Marquardt method were $k_x=k_y=k_z=0.5$. The parameters $k_x=1$, $k_y=15$ and $k_z=15$ were exactly recovered when errorless measurements were used in the inverse analysis. The parameters estimated with measurements containing random errors with standard-deviation $\sigma=0.01Y_{max}$ (where Y_{max} is the maximum measured temperature) and their corresponding 99% confidence interval, were:

$$\hat{k}_x = 0.99 \pm 0.03$$

$$\hat{k}_y = 14.89 \pm 0.54$$

$$\hat{k}_z = 15.02 \pm 0.05$$

Therefore, quite accurate estimates were obtained for the unknown thermal conductivity components with the present approach. However, note that the parameters were estimated with different relative accuracies. Relatively, the most accurate estimate was obtained for k_z, which has the largest value of the sensitivity coefficients, as illustrated in figures 1.a-e.

CONCLUSIONS

In this work we discuss the estimation of the three thermal conductivity components of an orthotropic solid, by using one single heated surface. The D-optimum criterion was used for the design of the experiment. The optimum size of the heating surface, as well as the heating and final times, were obtained for the best sensor configuration. The inverse problem of parameter estimation was solved by using the Levenberg-Marquardt method of minimization of the least-squares norm.

Accurate results were obtained by using in the inverse analysis simulated transient measurements containing random errors. However, the parameters could not be estimated with the same relative accuracy, as for the case involving the experimental setup with three heated surfaces.

REFERENCES

1. B.W. Sawaf and M.N. Özisik, Determining the Constant Thermal Conductivities of Orthotropic Materials, *Int. Comm. Heat Mass Transfer*, **22**, pp. 201-211 (1995)

2. B. Sawaf, M.N. Özisik and Y. Jarny, An Inverse Analysis to Estimate Linearly Temperature Dependent Thermal Conductivity Components and Heat Capacity of an Orthotropic Medium, *Int. J. Heat Mass Transfer*, **38**, No. 16, pp. 3005-3010 (1995)

3. R. Taktak, *Design and Validation of Optimal Experiments for Estimating Thermal Properties of Composite Materials*, Ph.D. dissertation, Michigan State University (1992)

4. R. Taktak, J.V. Beck and E.P. Scott, Optimal Experimental Design for Estimating Thermal Properties of Composite Materials, *Int. J. Heat Mass Transfer*, **36**, No. 12, pp. 2977-2986 (1993)

5. K. J. Dowding, J.V. Beck and B. Blackwell, Estimation of Directional-Dependent Thermal Properties in a Carbon-carbon Composite, *Int. J. Heat Mass Transfer*, **39**, No. 15, pp. 3157-3164 (1996)

6. K.J. Dowding, and B. Blackwell, Design of Experiments to Estimate Temperature Dependent Thermal Properties, *3rd Int. Conf. on Inverse Problems in Engineering: Theory and Practice*, pp. 509-518, Port Ludlow, WA (1999)

7. K. Dowding, J.V. Beck, A. Ulbrich, B. Blackwell and J. Hayes, Estimation of Thermal Properties and Surfaces Heat Flux in Carbon-Carbon Composite, *J. of Termophysics and Heat Transfer*, **9**, No. 2, pp. 345-351 (1995)

8. M.M. Mejias, H.R.B. Orlande and M.N. Özisik, Design of Optimum Experiments for the Estimation of the Thermal Conductivity Components of Orthotropic Solids, *Hybrid Methods in Engineering*, **1**, No 1, pp. 37-53 (1999)

9. M.M. Mejias, H.R.B. Orlande and M.N. Özisik, A Comparison of Different Parameter Estimation Techniques for the Identification of the Thermal Conductivity Components of Orthotropic Solids, *3rd Int. Conf. on Inverse Problems in Engineering: Theory and Practice*, pp. 325-332, Port Ludlow, WA, June 13-18 (1999)

10. M.M. Mejias, H.R.B. Orlande and M.N. Özisik, On the Choice of Boundary Conditions for the Estimation of the Thermal Conductivity Components of Orthotropic Solids, *ASME National Heat Transfer Conference, Paper NHTC 2000 – 12062*, Pittsburgh, August 20-22 (2000)

11. M.M. Mejias, H.R.B. Orlande and M.N. Özisik, Effects of the Heating Process and Body Dimensions on the Estimation of the Thermal conductivity Components of Orthotropic Solids, *Inv. Problems Engineering, in press*, (2002)

12. M.N. Özisik, *Heat Conduction*, 2nd ed., Wiley, New York, 1993.

13. J.V. Beck and K.J. Arnold, *Parameter Estimation in Engineering and Science*, Wiley, New York, 1977.

14. E. Artyukhin, Experimental Design of Measurement of the Solution of Coefficient-type Inverse Heat Conduction Problem, *J. Eng. Phys.*, **48**, pp. 372-376 (1985)

15. T.D. Fadale, A.V. Nenarokomov and A.F. Emery, Two Approaches to Optimal Sensor Locations, *ASME J. Heat Transfer*, **117**, pp.373-379 (1995)

16. A.F. Emery and T.D. Fadale, The Effects of Imprecisions in Thermal Sensor Location and Boundary Conditions on Optimal Sensor Location and Experimental Accuracy, *ASME J. Heat Transfer*, **119**, pp. 661-665 (1997)

17. K. Levenberg, A Method for the Solution of Certain Non-linear Problems in Least Squares, *Quart. Appl. Math.*, **2**, pp.164-168 (1944)

18. D.W. Marquardt, An Algorithm for Least Squares Estimation of Nonlinear Parameters, *J. Soc. Ind. Appl. Math*, **11**, pp.431-441 (1963)

19. M.N. Özisik and H.R.B. Orlande, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor & Francis, New York, 2000.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

OPTIMAL EXPERIMENT DESIGN AND SIMULTANEOUS IDENTIFICATION OF THERMO-PHYSICAL PROPERTIES OF ORTHOTROPIC SOLIDS

F. Mzali, L. Sassi, A. Jemni and S. Ben Nasrallah
Laboratoire d'Etude des Systèmes Thermiques et Energétiques, Ecole Nationale d'Ingénieurs de Monastir
Avenue Ibn El Jazzar, Monastir 5019, Tunisia
E-Mail : Foued.mzali@isbas.rnu.tn

D. Petit
Laboratoire d'Etudes Thermiques
UMR-CNRS 6608
Ecole Nationale Supérieure de Mécanique et d'Aérotechnique
Poitiers- Futuroscope 86960
France

ABSTRACT
This numerical study deals with the identification of several thermal properties of orthotropic solids by using an optimal experiment design. The system under investigation is a cylindrical vertical sample, which is submitted to an irradiation flux on the upper face on a central circular area. The temperature response, during and after the irradiation, is measured at the opposite face. In this paper, the system is particularly developed to the identification of the Fourier number including the axial diffusivity, the ratio of axial and radial conductivities, a Biot number and a heat flux term from one temperature evolution versus time. The temperature sensor position is adjusted simultaneously with the identification of the above four parameters.

NOMENCLATURE
Bi	Biot number
$c.$	Heat capacity [J/(kg·K)]
$e.$	Sample thickness [m]
$h.$	Heat transfer coefficient [W/(K·m^2)]
J	Ordinary least square norm
P	Vector of unknown parameters
$q.$	Heat flux density [W/m^2]
Q	Heat flux [W]
f	Normalized form of the flux shape (r,t)
R^+	Radial coordinates [m]
$r_m.$	Temperature sensor position [m]
$r_p.$	Heat flux radius [m]
$r_s.$	Sample radius [m]
T'	Temperature [K]
T_0	Heat pulse duration [s]
T_0'	Initial temperature [K]
X	Sensitivity matrix
$\overline{X}$	Reduced sensitivity matrix
$\overline{X}_{i,k}$	Reduced sensitivity coefficients
n	Number of measurement points
Y	Vector of estimated temperatures
Y_m	Vector of measured temperatures
z^+	Axial coordinates [m]

Greek letters
α	Thermal diffusivity [m^2/s]
λ	Thermal conductivity [W/(m·K)]
ρ	Density [kg/m^3]

INTRODUCTION
The knowledge of thermo-physical properties of materials is important for thermal design and numerical simulation. In the past decades many methods have been presented to determine these properties [1].
The flash method is one of the most used methods. It consists to apply a brief heat pulse on the front face of a small cylindrical sample. The resulting temperature rise on the opposite face is recorded versus time. Thus, the thermal diffusivity is computed using the measured temperature evolution and an identification method based on an analytical model [2].
In the origin, this method was established to characterize isotropic solids, then, it has been extended to orthotropic ones. Namely, Lachi and Degiovanni [3] proposed an identification method based on an analytical model and the recorded temperature at two positions on the opposite face, to determine the axial and radial thermal diffusivities of a semi -infinite solid. The major difficulty of this method lies in the inaccurate knowledge of the sensor positions

[4]. Therefore, it has been shown that the use of many temperature sensors or an infrared system on the opposite face can reduce the errors due to the measurement positions [4, 5, 6, 7].
The advantage of using an instantaneous pulse in the flash method lies in the high sensitivity of the temperature to the different parameters. However the high temperature rise at the exposed face of the sample limits their use in the characterization of poorly conductive materials and large samples. Only a limited amount of solutions have been used to overcome these problems [8, 9]. Therefore, we propose to reduce the intensity of the energy source simultaneously with increasing the exposure time (pulse extension) just enough to decrease the high temperature gradient in the sample, keeping the temperature enough sensitive to the different parameters.
The first aim of this numerical study is to define the optimal values of heat flux radius and temperature sensor position using condition numbers and sensitivity analysis.
The found optimal values are used for identifying simultaneously the Fourier number including the axial diffusivity, the ratio of axial and radial conductivities, a Biot number and a heat flux term from one temperature evolution versus time. At the same time, the temperature sensor position is adjusted in order to reduce the uncertainty in the sensor position.

SYSTEM DESCRIPTION

The system under investigation is a cylindrical sample, supposed to be homogeneous, opaque and having constant thermal properties and density (figure 1). This sample is submitted to an irradiation flux on the upper face on a central circular area. The radius r_p of the considered circular area is smaller than the sample radius r_s. The sample exchanges heat throw lateral and bases areas (h_r, h_e, h_0).

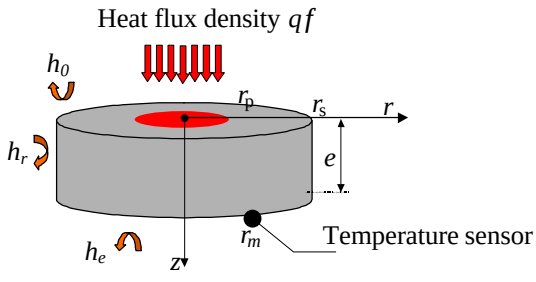


Figure 1 radial heat flow method

Taking into account the symmetry about the principal axis z, the temperature evolution at position (z, r) of the sample can be obtained by solving the two-dimensional heat conduction equation in cylindrical coordinates [10].
Considering the following variables:

$$r = r^{+}/e \; ; \; z = z^{+}/e \; ; \; t = t/t_0 \; ; \; T=(T' - T_0')/T_m \quad (1)$$

Where T_0 is the initial temperature and Tm is the heat flux term (the maximal temperature in the adiabatic case) :

$$T_m = q^{x} t_0 / r^{x} c^{x} e; \quad (2)$$

The heat equation is then written, in its dimensionless form, as:

$$\frac{1}{F_{oz}}\frac{\partial T}{\partial t} = \frac{K_{rz}}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + \frac{\partial^2 T}{\partial z^2} \quad (3.a)$$

Subject to the boundary conditions:

$$\frac{\partial T}{\partial z} = -\frac{f(r,t)}{F_{oz}} + Bi_o T \quad at\ z=0 \quad (3.b)$$

$$\frac{\partial T}{\partial r} = 0 \quad at \quad r=0 \quad (3.c)$$

$$\frac{\partial T}{\partial z} = -Bi_e T \qquad at\ z=1 \quad (3.d)$$

$$\frac{\partial T}{\partial r} = -Bi_r T \quad at\ r=r_s \quad (3.e)$$

and the initial condition is taken as :

$$T = 0 \quad for\ t = 0 \quad (3.f)$$

Where F_{0z} is the axial Fourier number defined as :

$$F_{oz} = \frac{a_z \cdot t_0}{e^2} \; ; \quad (4)$$

K_{rz} is the ratio of axial and radial conductivities:

$$K_{rz} = \frac{l_r}{l_z} \; ; \quad (5)$$

the Biot numbers, corresponding to the lateral area and the bases areas at $z=0$ and at $z=e$, are respectively defined as :

$$Bi_r = \frac{h_r \cdot r_s}{l_r} \; ; \; Bi_o = \frac{h_o \cdot e}{l_z} \; ; \; Bi_e = \frac{h_e \cdot e}{l_z} \; ; \quad (6)$$

all the heat transfer coefficients are taken equal:

$$h_r = h_0 = h_e \quad (7)$$

Then, we introduce the axial Biot Number Bi_z, wich corresponds to the bases areas at $z=0$ and at $z=e$:

$$Bi_z = Bi_e = Bi_0 \quad (8)$$

The radial Biot number can be written as:

$$Bi_r = \frac{r_s / e}{K_{rz}} Bi_z \qquad (9)$$

The heat flux on the base area at $z=0$ is assumed to be a pulse of finite duration τ and limited in a surface of a radius r_p, the shape of which is described in normalized form :

$$f(r,t) = \begin{cases} 1 \ \text{if } t<1 \ \text{and} \ 0<r<r_p \\ 0 \ \text{if } t>1 \ \text{or} \ r \geq r_p \end{cases} \qquad (10)$$

The system of the presented equations is solved numerically by the finite volume method [11].

OPTIMIZATION METHOD

Ordinary least squares norm

The problem consists of minimizing the ordinary least square norm defined as :

$$J = [Y(P) - Y_m]^T [Y(P) - Y_m] \qquad (11)$$

J is optimal if :

$$\tilde{N}J(P) = 0 \qquad (12)$$

That means :

$$2X^T [Y(P) - Y_m] = 0 \qquad (13)$$

Where $X = \left[\frac{\partial Y(P)}{\partial P}\right]$ is the sensitivity matrix which is written as :

$$X = \begin{bmatrix} \frac{\partial Y^1}{\partial r_m} & \frac{\partial Y^1}{\partial T_m} & \frac{\partial Y^1}{\partial Bi_z} & \frac{\partial Y^1}{\partial F_{0z}} & \frac{\partial Y^1}{\partial K_{rz}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial Y^n}{\partial r_m} & \frac{\partial Y^n}{\partial T_m} & \frac{\partial Y^n}{\partial Bi_z} & \frac{\partial Y^n}{\partial F_{0z}} & \frac{\partial Y^n}{\partial K_{rz}} \end{bmatrix} \qquad (14)$$

In non linear problems, the sensitivity coefficients $\frac{\partial Y^i}{\partial p_j}$ depends on the unknown parameters p_j .

Conjugate gradient method

Many techniques have been presented, in the literature[12], to minimize the least squares norm. Among these, the Levenberg- Marquard and the conjugate gradient methods are very efficient non linear estimation techniques [13,14].

In this work, we use the conjugate gradient method[15], in which a suitable step size is taken along a direction of descent in order to minimize the objective function.

The unknown parameter vector is given at each iteration as follows :

$$p^{(k+1)} = p^{(k)} + g_k d^{(k)} ; \qquad (15)$$

with k is the iteration subscript.

The search step size is defined as :

$$g_k = -\frac{d^{(k)^T} g^{(k)}}{d^{(k)^T} H d^{(k)}} \qquad (16)$$

The direction of descent is given by :

$$d^{(k+1)} = g^{(k+1)} - b^{(k)} d^{(k)} \qquad (17)$$

where $g^{(k)}$ is the gradient of J at the k^{th} iteration $(g^{(k)} = \tilde{N} J(p^{(k)}))$.

$b^{(k)}$ is a conjugation coefficient and is obtained from the following expression :

$$b^{(k)} = \frac{g^{(k+1)^T} H d^{(k)}}{d^{(k)^T} H d^{(k)}} \qquad (18)$$

Where H is the symmetric Hessien defined as :

$$H(P) = 2\chi X^T X + 2\chi [\partial X^T / \partial P] \chi [Y - Y_m] \qquad (16)$$

which coefficients (index j_p ,j_q) are written as :

$$H_{j_p, j_q} = \frac{\partial^2 J}{\partial p_{j_p} \cdot \partial p_{j_q}} = 2 \cdot \sum_{i=1}^{n} \frac{\partial Y^i}{\partial p_{j_p}} \cdot \frac{\partial Y^i}{\partial p_{j_q}} + 2 \cdot \sum_{i=1}^{n} \frac{\partial^2 Y^i}{\partial p_{j_p} \cdot \partial p_{j_q}} \cdot \left(Y^i - Y_m^i\right) \qquad (19)$$

According to the version of the algorithm of conjugate gradient, $b^{(k)}$ can be written in different manners. While considering all gradients $g^{(0)},\ldots, g^{(k)}$ are orthogonal, one gets the Fletcher – Reeves version, where :

$$b^{(k)} = \frac{\left\| g^{(k+1)} \right\|^2}{\left\| g^{(k)} \right\|^2} \qquad (20)$$

While considering only the gradients $g^{(0)},..,g^{(k-1)}$ are orthogonal, one has the version proposed by Polak and Ribiere, where :

$$b^{(k)} = \frac{g^{(k+1)^T} \left(g^{(k+1)} - g^{(k)} \right)}{\left\| g^{(k)} \right\|^2} \qquad (21)$$

The software that we have used [15] seems to give better results with non quadratic functions. The stopping criterion used for this algorithm is the following :

$$|J(p^{(k)}) - J(p^{(k-1)})| \leq \tfrac{1}{2} ftol\,(|J(p^{(k)})| + |J(p^{(k-1)})| + e) \qquad (22)$$

with $e = 10^{-10}$ and $ftol = 10^{-8}$ are chosen for all the identifications.

OPTIMAL EXPERIMENT DESIGN

Sensitivity analysis

The sensitivity analysis is classically used for the optimal experimental design. It consists of analyzing the evolution of the different reduced sensitivity coefficients versus an explicative variable (time).

$$\overline{X}_{i,k} = p_k \frac{\partial Y^i}{\partial p_k} \qquad (23)$$

These reduced sensitivity coefficients are calculated using a central finite difference approximation:

$$\overline{X}_{i,k} = p_k \cdot \frac{Y^i(p_k + \delta p_k) - Y^i(p_k - \delta p_k)}{2 \cdot \delta p_k} \qquad (24)$$

Where δp_k is taken $10^{-3} \cdot p_k$ for the parameters (T_m, Bi_z, F_{0z}, K_{rz}) and it is equal to the ratio of mesh step to mesh size for the parameter r_m/r_s.

$$\delta(r_m/r_s) = 1/(maxr-1) \qquad (25)$$

Where maxr=101 is the radial mesh size.

For these studies, the reduced sensitivity matrix $\overline{X}$ is computed for a known parameter vector.

Optimality criteria

The most common criteria for the optimal experiment design are based on the information matrix X^TX which summarizes the information content of an experiment or the precision of the parameter estimates. It is the approximation of the Hessien matrix H defined by equation (19) near the optimal solution.

First it is necessary to search for a great determinant of X^TX as a guaranty for great sensitivity coefficients and no correlated parameters. This criterion is named D-optimality. Its physical motivation is that it minimizes the volume of the ellipsoid which represents the maximum confidence region.[18] This allows us to identify the parameters. Second, it is necessary to search for a little condition number expressed here as the ratio of the largest eigenvalue to the smallest one, in order to obtain a matrix as well conditioned as possible. This is the modified E-optimality criterion. It optimizes the shape of the objective function. This means that the sensitivity coefficients of each parameter should have the same order of magnitude while varying in different manners (no linear dependence) [16,17]. In consequence, the sensitivity of the identification to the measurement errors is low.

In order to validate the results, we use the E-optimality criterion which maximizes the minimum eigenvalue. The physical motivation is to maximize the minimum diameter of the ellipsoid. It is used if one parameter has a very large variance compared to the others.[18]

RESULTS AND DISCUSSIONS

Choice of an optimal position

We analyses the three optimality criteria in order to chose the optimal heat flux radius and the sensor position. Then we identify the dependence of this criteria on the thermo-physical properties (F_{0z} and K_{rz}) and the sample configuration (r_s/e).

The influence of the flux radius and the sensor position on the determinant, the minimum eigenvalue and the condition number are respectively plotted on the figures 2a,b,c.

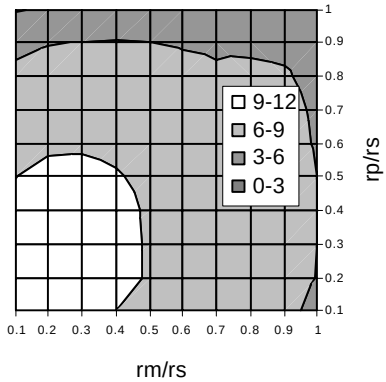


Figure 2a Influence of r_p/r_s and r_m/r_s on the determinant (exponent value) with $r_s/e = 1/2$ and $F_{oz}=0.1$ $K_{rz}=0.1$ $T_m=0.25$ $Bi_z=0.1$ $t_0=1s$

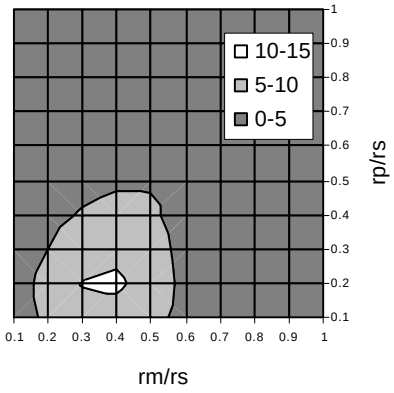


Figure 2b Influence of r_p/r_s and r_m/r_s on the minimum eigenvalue with $r_s/e=1/2$ and $F_{oz}=0.1$ $K_{rz}=0.1$ $T_m=0.25$ $Bi_z=0.1$ $t_0=1s$

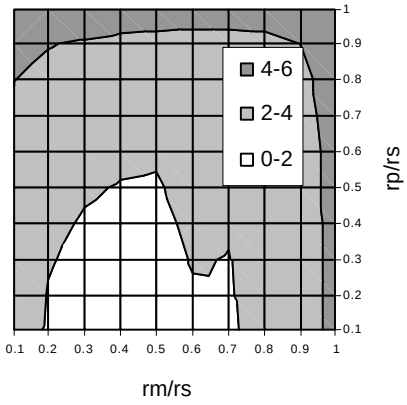


Figure 2c Influence of r_p/r_s and r_m/r_s on the condition number (exponent value) with $r_s/e=1/2$ and $F_{oz}=0.1$ $K_{rz}=0.1$ $T_m=0.25$ $Bi_z=0.1$ $t_0=1s$

We remark that there is a significant region where the determinant (10^9-10^{12}) and the minimum eigenvalue (5-15) are great and the condition number (1-10^2) is low (intersection of white areas). In this region the identification is possible and is less sensitive to measurement errors. We chose $r_p/r_s=0.2$ and $r_m/r_s=0.3$ as optimal values, then, we perform a sensitivity analysis of the temperature to all parameters. We see (figure 3) that the parameters are not correlated because the correspondent sensitivity coefficients vary in different manner and they do not reach their maximums at the same time. We see also that the parameters of interest have approximately the same order of magnitude (~0.2). The identification is then optimal.

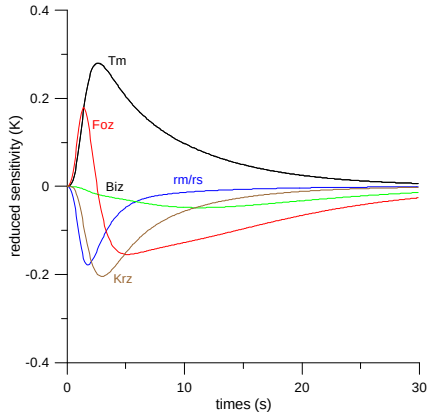


Figure 3 Reduced sensitivity coefficients of r_m/r_s , Tm , B_{iz}, F_{0z} and K_{rz} with $r_p/r_s=0.2$; $r_m/r_s=0.3$ $r_s/e=1/2$ and $F_{oz}=0.1$ $K_{rz}=0.1$ $T_m=0.25$ $Bi_z=0.1$ $t_0=1s$

From the dark region of the figures 2a,b,c, we take an other position $r_p/r_s=0.5$ and $r_m/r_s=0.7$ as an example of bad optimality. we perform a sensitivity analysis of the temperature to all parameters.

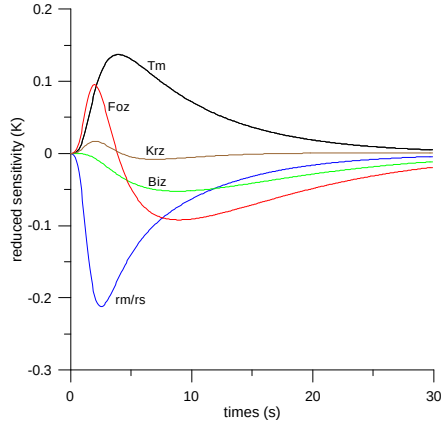


Figure 4 Reduced sensitivity coefficients of r_m/r_s , Tm , B_{iz}, F_{0z} and K_{rz} with $r_p/r_s=0.5$; $r_m/r_s=0.7$ $r_s/e=1/2$ and $F_{oz}=0.1$ $K_{rz}=0.1$ $T_m=0.25$ $Bi_z=0.1$ $t_0=1s$

We see (figure 4) that the parameters are near to be correlated because the maximums of the correspondent sensitivity coefficients are closer to each other and do not have the same order of magnitude (0.02 to 0.2). The identification will be consequently very sensitive to measurement noise.

In this way, the results of quantitative analysis of the optimality criteria has been, then, confirmed by a qualitative and visual sensitivity analysis.

Dependence on K_{rz} and F_{0z}

We analyze the dependence of the optimality criteria on the Fourier number and the ratio of radial to axial thermal conductivities. The determinant, the minimum eigenvalue and the condition number are respectively plotted on the figures 5a,b,c.

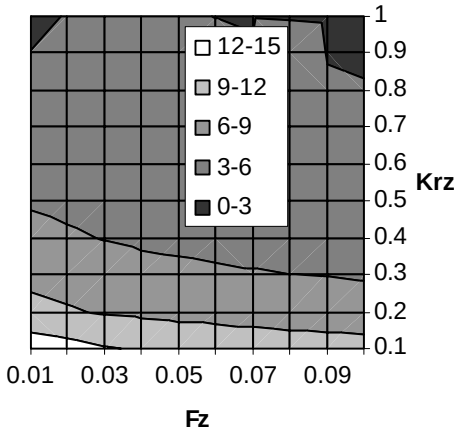


Figure 5a Influence of K_{rz} and F_{0z} on the determinant (exponent value) with r_s/e =1/2 and r_p/r_s=0.2 r_m/r_s=0.3 T_m=0.25 Biz=0.1 t_0=1s

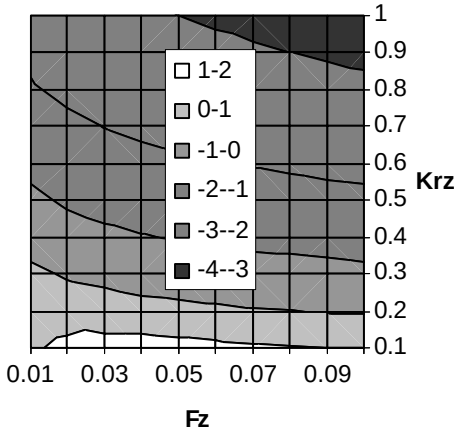


Figure 5b Influence of K_{rz} and F_{0z} on the minimum eigenvalue (exponent value) with r/e=1/2 et r_p/r_s=0.2 r_m/r_s=0.3 T_m=0.25 Biz=0.1 t_0=1s

We remark that the determinant and the minimum eigenvalue are maximum when F_{0z} or K_{rz} decreases. However the condition number has a different aspect. It is minimum for a great F_{0z} and a low K_{rz}.

Consequently, K_{rz} should not be greater than 0.1 and F_z should be within 0.05-0.1.

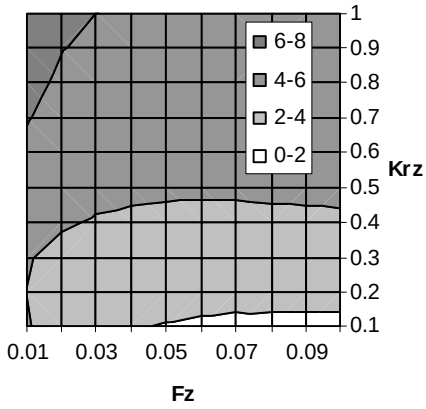


Figure 5c Influence of K_{rz} and F_{0z} on the condition number (exponent value) with r/e=1/2 et r_p/r_s=0.2 r_m/r_s=0.3 T_m=0.25 Biz=0.1 t_0=1s

Dependence on r_s/e

We analyze the dependence of the optimality criteria on the geometrical sample configuration represented by r_s/e. The determinant, the minimum eigenvalue and the condition number are respectively plotted on figure 7.

We see that the determinant and the minimum eigenvalue are maximized and the condition number is minimized, when r_s/e increases.

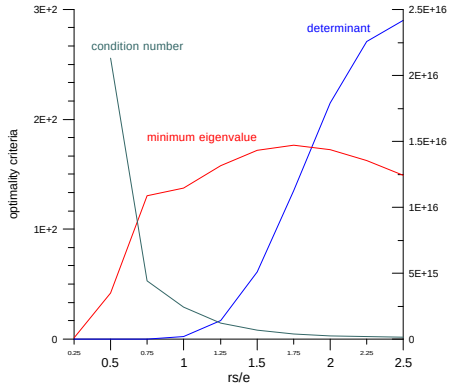


Figure 6 Influence of r_s/e on the different optimality criteria with F_{oz}=0.1 K_{rz}=0.1 r_p/r_s=0.2 r_m/r_s=0.3 T_m=0.25 Biz=0.1 t_0=1s

Identification

We perform an identification of all parameters. The simulated measured temperature $Y_{measured}$ is

obtained by adding a noise term ws to the computed exact temperature T_{exact} as :

$$Y_{measured} = T_{exact} + ws \qquad (25)$$

The noise is gaussian distributed and s is the standard deviation of measurement errors. Assuming 99% confidence for the measured data, w lies in the range $-2.576 \pounds w \pounds 2.576$ and it is calculated by a random generator [19,20] (figure 7).

The results of the identification, using different standard deviations of noise, are shown on the tables 1 and 2. We perform an identification with the chosen optimal radius r_p/r_s=0.2, r_m/r_s=0.3. Initial values are taken at 50% from the expected values. The parameters are identified with an error less than 8% even with measurement noise.

We perform an other identification with the non optimal radius r_p/r_s=0.5, r_m/r_s=0.7. When measurement are altered, another try with an increased time step using only 30 measurement points seems to give better results (figure 8). When initial values are taken close (+10%) to the expected values, the parameters are identified with a good accuracy (error less then 7%). However, when the initial values are taken far from the expected values (+50%), the identification error become more significant and can reach 80%. We note that the initial value of the sensor position is not taken far from the expected value because it is known and need only be adjusted.

So, we have demonstrated that the values of r_p/r_s=0.2, r_m/r_s=0.3 chosen by the optimal experiment design are effectively optimal for the identification. The other values r_p/r_s=0.5, r_m/r_s=0.7, rejected by the optimal experiment design, are not optimal for the identification, which is very dependant to initial values and very sensitive to measurement noise.

Table 1 Results of the identification
(r_p/r_s=0.2 r_m/r_s=0.3 r_s/e =1/2 and F_{oz}=0.1
K_{rz}=0.1 T_m=0.25 Bi_z=0.1 t_0=1s)

	Initial values	σ=0 K 300 pts	σ=0.001K 30 pts	σ=0.01K 300 pts	Expected Values
r_m/r_s	0.297	0.302	0.301	0.314	0.3
Tm	0.375	0.250	0.249	0.270	0.25
Biz	0.15	0.100	0.099	0.106	0.1
Foz	0.15	0.099	0.100	0.100	0.1
Krz	0.15	0.100	0.100	0.108	0.1
Iterations Nb		80	31	42	

Table 2 Results of the identification
r_p/r_s=0.5 ; r_m/r_s=0.7 r_s/e =1/2 and F_{oz}=0.1
K_{rz}=0.1 T_m=0.25 Bi_z=0.1 t_0=1s

	Initial values	σ=0 K 300 pts	σ=0.001K 300 pts	σ=0.01K 300 pts	Expected Values
r_m/r_s	0.693	0.700	0.704	0.704	0.7
Tm	0.275	0.249	0.255	0.251	0.25
Biz	0.090	0.099	0.107	0.103	0.1
Foz	0.110	0.100	0.097	0.098	0.1
Krz	0.110	0.099	0.104	0.107	0.1
Iterations Nb		59	9	5	
r_m/r_s	0.693	0.722	0.725	0.726	0.7
Tm	0.375	0.273	0.274	0.282	0.25
Biz	0.15	0.141	0.138	0.147	0.1
Foz	0.15	0.089	0.091	0.088	0.1
Krz	0.15	0.178	0.179	0.171	0.1
Iterations Nb		39	22	11	

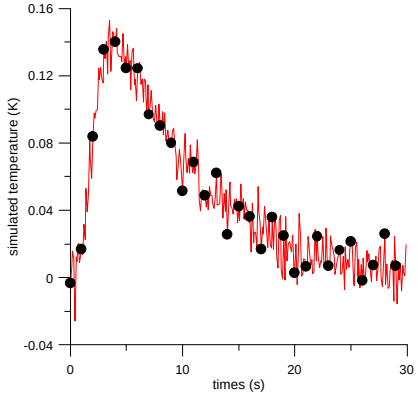


figure 7: measurement points (30 pts) shown on the simulated temperature curve

CONCLUSION

This numerical study has been performed to estimate simultaneously the Fourier number including the axial diffusivity, the ratio of axial and radial conductivities, a Biot number and a heat flux term from one temperature evolution versus time. At the same time, the temperature sensor position is adjusted in order to reduce the uncertainty in the sensor position.

In order to consider the optimal system configuration, we have used different optimality criteria related to the sensitivity matrix. They are the determinant, the minimum eigenvalue and the condition number of an approximation of the Hessien matrix build on the sensitivity matrix. The corresponding analysis has been confirmed by a qualitative and visual sensitivity analysis.

So, an identification using the optimal design of the experiment has been performed. The results

are accurate and show that the axial and radial thermal diffusivities can be simultaneously determined using a temperature evolution measured at one optimal position.

REFERENCES

1. J. Hladik, Métrologie des propriétés thermophysiques des matériaux, Masson collection mesures physiques, Paris, 1990.
2. A. Degiovanni, Diffusivité et méthode flash, Rev. Gén. Therm., 185, pp. 417 (1977).
3. M. Lachi, A. Degiovanni, Détermination des diffusivités thermiques des matériaux anisotropes par méthode flash bidirectionnelle, J. Phys. III, France 1, pp. 2027-2046 (1991).
4. A. Degiovanni, J.C. Batsale, D. Maillet, Mesure de la diffusivité longitudinale de matériaux anisotropes, Rev. Gén. Therm., 35, pp. 141-147 (1996).
5. I. Philippi, J.C. Batsale, D. Maillet, A. Degiovanni, Traitement d'images infrarouges par transformation intégrale, Rev. Gén. Therm., 392-393, pp. 486-496 (1994).
6. L. Vozar, T. Sramkova, Simultaneous measurement of axial and radial thermal diffusivities of anisotropic media by the radial-heat-flow step-heating method, High Temp.-High Press., 29, pp. 183-190 (1997).
7. D. Demange, P. Beauchêne, M. Bejet, R. Casuleras, Mesure simultanée de la diffusivité termique selon les deux directions principales d'un matériau, Rev. Gén. Therm., 36, pp.755-770 (1997).
8. L. Vozar, G. Groboth, Thermal diffusivity measurement of poorly conductive materials : comparison of the step-heating and flash methods , High Temp.-High Press., 29, pp. 191-199 (1997).
9. L. Vozar, W. Hohenauer, Measurement of thermal diffusivity by the laser-flash method with repeated pulses, High Temp.-High Press. Vol. 33, pp. 9-16 (2001).
10. M. N. Ozisik, Heat transfert, a Wiley - interscience publication, USA, 1985.
11. S.V. Patankar, Numerical Heat transfer and fluid flow, Hemisphere Mc. Graaw-Hill, New York, 1980.
12. J.V. Beck, K. Arnold, Parameter estimation in engineering and science, Wiley, NY, 1977.
13. M.M. Mejias, H.R.B. Orlande, M.N. Ozisik, A comparision of different parameter estimation techniques for the identification of thermal conductivity components of orthotropic solids, Proceeding of the 3rd Int. Conference on inverse Problems in Engineering, Port Ludlow, WA, USA (1999).
14. M.J. Colaço, H.R.B. Orlande, Comparison of different versions of the conjugate gradient method of function estimation, Num. Heat Transfer, Part A, 36, pp.229-249 (1999).
15. W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical recipes in PASCAL, Cambrigde University Press, 1990.
16. M. Raynaud, Strategy for experimental design and the estimation of parameters, High Temp.-High Press., 31, pp. 1-15 (1999).
17. R. Taktak, J.V. Beck, E.P. Scott, Optimal experimental design for estimating thermal properties of composite materials", Int. J. Heat Mass Transfer., 36, 12, pp. 2977-2986 (1993).
18. A. F. Emery and A. V. Nenarokomov, Optimal experiment design, Meas. Sci. Technol., 9, pp. 864-876 (1998).
19. B. Sawaf, M.N. Ozisik, Y. Jarny, An inverse analysis to estimate linearly temperature dependent thermal conductivity components and heat capacity of an orthotropic medium, Int. J. Heat Mass Transfer, 38, 16, pp. 3005-3010 (1995).
20. R.W. Hamming, Numerical methods for scientists and engineers, Dover Publications, New York, 1973, Chap. 8.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSE ANALYSIS FOR IDENTIFICATION OF ELECTROMAGNETIC PARAMETERS

Yann Favennec
Cemef – Ecole des Mines de Paris
Rue Claude Daunesse – B.P. 207
Sophia Antipolis, 06904, France
favennec@cemef.cma..fr

Valérie Labbé,
François Bay
Cemef – Ecole des Mines de Paris

ABSTRACT

Some magnetic properties, namely the magnetic Frölich Kenelly coefficients and its temperature sensitivity coefficients are identified through temperature measurements. We use an inverse analysis approach, based on induction heating processes. A gradient type method is used to minimize an assumed convex objective function – this functional being based on experimental temperature data. Two distinct approaches are presented depending on the nature of the magnetic properties to be identified: a finite difference method and an adjoint method.

INTRODUCTION

The measurement of the main magnetic properties of a material are often difficult to perform. Moreover, modeling of coupled electromagnetic – heat transfer problems requires a precise knowledge of the evolution of these properties with respect to temperature. A new procedure for measuring these values is presented here. It is based on an inverse analysis coupled with a direct finite element modeling of the experimental set-up. One of the main innovations of this procedure is that the measured quantities are only based on temperatures in order to identify electromagnetic parameters. A cost function integrates differences between some experimentally measured temperature (due to induced eddy currents) and the ones computed at the same time and location by a general coupled electromagnetic – heat transfer – mechanical finite element code developed in our laboratory. This direct model solves the non-linear transient direct equations. Since characteristic times between heat transfer and electromagnetic models can differ of a factor reaching 10**6 or more, a weak coupling strategy has been adapted. The cost function is then minimized using an iterative gradient type method.

Two approaches are presented for the sensitivity analysis calculation: a global finite difference approach, and an adjoint approach based on optimal control theory. For this latter, reverse electromagnetic and thermal adjoint equations are then integrated. An originality of the adjoint approach comes from the adaptation of the reverse adjoint equations to include the features of the algorithm for solving the direct weakly coupled problem. The reverse adjoint equations have been modified with no approximation. The whole procedure has been successfully tested and results on the magnetic permeability properties identification will be presented.

THE MAGNETIC PROPERTIES

All electromagnetic systems are modeled using the Maxwell equation. Both finite element and the boundary element approaches are commonly used. Two physical parameters are involved in the Maxwell equations, they are the electrical conductivity s and the magnetic permeability m. This latter parameter is, most of the time, function of the internal magnetic field strength. Moreover, in most electromagnetic devices, the induced currents heat materials, and it is well known that the magnetic permeability highly depends on temperature. One therefore gets $m = m(H,T)$. For most steels – which are ferromagnetic materials – the magnetic permeability is directly derived – at a given temperature – from the well-known magnetization curve $B(H)$ through relation (1):

$$m(H,T)=\frac{\partial B}{\partial H}(H,T). \qquad (1)$$

A typical magnetization curve may be divided in three main parts. In the first part, at low fields, the curve starts with a weak finite slope, and rises so that it is concave upwards. The second part has the greatest slope, which then decreases from saturation onwards.

As far as we are concerned, we shall focus on the last two parts, when involved fields are very high – for instance when dealing with high power electromagnetic devices. Then, at high fields, the saturation curves, for several temperature, are as shown in *Figure 1*.

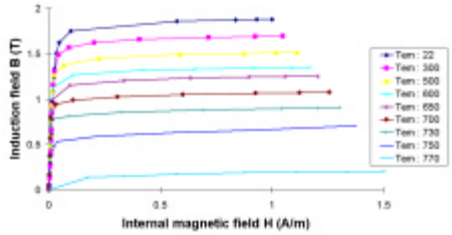


Figure 1: Saturation curves at several temperature

It is well assumed that the well-known Frölich-Kenelly formulation is a good approximation for modeling the magnetization curves up to the critical Curie temperature T_c. The relative magnetic permeability may then be expressed as:

$$m_r\left(\left|\vec{H}\right|,T\right)=1+\left(\frac{a}{m_0\left(b+\left|\vec{H}\right|\right)}-\frac{a\left|\vec{H}\right|}{m_0\left(b+\left|\vec{H}\right|\right)^2}\right)f(T), \qquad (2)$$

with

$$f(T)=1-\left(\frac{T}{T_c}\right)^g. \qquad (3)$$

The magnetic parameters to be identified are then both Frölich coefficients a and b and the temperature sensitivity coefficient g.

When dealing with powerful electromagnetic processes modeling, it is very common to linearize the magnetic field dependency. Among several different approaches, the linearization method based on energetic considerations seems to be the most powerful one (see [1]). Then, the relative magnetic permeability involves only two coefficients to be identified a_{lin} and g:

$$m(T)=m_0+a_{lin}\left(1-\left(\frac{T}{T_c}\right)^g\right), \qquad (4)$$

THE OPTIMIZATION PROBLEM

As has been said previously, the main objective is to identify magnetic parameters (either a, b and g when dealing with the H nonlinear formulation (2) or a_{lin} and g when dealing with the H linear formulation (4)) through some temperature measurements. An inverse analysis approach has been used. The experimental set-up consists in heating, by electromagnetic way, the material to be characterized. In order to evaluate magnetic properties within a wide temperature range, an induction heating set-up has been used, input power within external coils being easily changeable. Temperature are experimentally measured on the heated workpiece surface thanks to pre-welded thermocouples.

We have developed in our laboratory a complete 2D electro-thermo-mechanical model which is able to simulate the experimental induction heating set-up used for identification. This model has been validated thanks to comparisons with experimental data (see [2]). The model writes in its very condensed matricial form as $R=0$.

The inverse analysis approach consists in minimizing errors between calculated and experimental data. In order to minimize these errors, these are integrated within an objective function to be minimized. This adimensionnalized objective function, with no regularization term, writes:

$$J(T)=\sum_{i=1}^{nbdata}\frac{e_i^2}{T_i^{\exp 2}}=\sum_{i=1}^{nbdata}\left(\frac{T_i^{\exp}-T_i^{cal}}{T_i^{\exp}}\right)^2, \qquad (5)$$

where $T_i^{\exp}$ and T_i^{cal} are respectively experimental and calculated (simulated)

temperature at certain locations and times. For the sake of clarity, the objective function given in (5) will be rewritten as:

$$J(T) = J_1\left(T\left(t_{fin}\right)\right) + J_2\left(T(t)\right), \qquad (6)$$

Let us denote by P the physical parameters ($P = \{a, b, g\}$ in case of use of (2) and $P = \{a_{lin}, g\}$ in case of use of (4)), and write classically $j(P) = J(T)$. The optimization problem is formulated as:

$$Find\ \overline{P} \in P_{ad} : j(\overline{P}) = \min_{R=0} j(P). \qquad (7)$$

THE DIRECT ELECTRO-THERMAL COMPUTATION

The full mathematical model developed in our laboratory couples the three following physical phenomena: electromagnetism, heat transfer and a mechanical computation (see [2]). Here, only the two first physical phenomena will be of concern.

The Electromagnetic Model

The global system of equations modeling wave propagation is based on the four Maxwell equation.

$$\vec{\nabla}.\vec{B} = 0, \qquad (8)$$

$$\vec{\nabla}.\vec{D} = 0, \qquad (9)$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \qquad (10)$$

$$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}, \qquad (11)$$

where $\vec{H}$ is the magnetic field, $\vec{B}$ the magnetic induction, $\vec{E}$ the electric field, $\vec{D}$ the electric flux density, $\vec{J}$ the electric current density associated with free charges and $\times$ denotes the vector product. Most electro-thermal processes work at frequencies lower than 1 GHz. It is then much convenient to use the Magneto-quasi-static approximation. This approximation consists in neglecting propagation phenomena – neglecting the term $\partial D/\partial t$ in (11). This system of equations is coupled with relations associated to material properties, material being assumed to be isotropic:

$$\vec{B} = m(T, |\vec{H}|)\vec{H}, \qquad (12)$$

$$\vec{J} = s(T)\vec{E}, \qquad (13)$$

where σ the electrical conductivity, the magnetic permeability m depending on temperature and on the magnetic field strength through (2) or (4).

As it was our intention to simulate electromagnetic processes with any shape of exciting input current (sinusoidal or squared signal, pulse,…), we have decided to integrate numerically in time the electromagnetic equation. Choosing, as the state variable, the electrical field E on axi-symmetrical configurations, one gets the following weak electromagnetic formulation:

$$\begin{aligned} & a^E(\dot{E}, y) + b^E(E, y) = l^E(y) \quad \forall y \in V, t > 0 \\ & E(0) = 0 \end{aligned} \qquad (14)$$

where the functional space V is defined as $V = \{v \in H_0^1(\Omega), v = 0 \text{ on } \Gamma_0\}$ and where:

$$a^E(\dot{E}, y) = \int_\Omega s\, \dot{E}(t) y \; ds, \qquad (15)$$

$$\begin{aligned} b^E(E, y) = & \int_\Omega \frac{1}{m} \vec{\nabla} E(t).\vec{\nabla} y \; ds + \\ & \int_\Omega \frac{1}{m\, r^2} E(t).y \; ds + \\ & \int_\Omega \frac{1}{m\, r} \frac{\partial}{\partial r} (E(t)\, y) \; ds, \end{aligned} \qquad (16)$$

$$\begin{aligned} l^E(y) = & -\int_\Omega \frac{\partial J_s}{\partial t}(t) y \; ds \\ & + \int_{\Gamma - \Gamma_0} \frac{1}{m} \left(\frac{\partial E}{\partial n}(t) + \frac{E(t)}{r} \vec{e}_r.\vec{n} \right) y \; d\Gamma. \end{aligned} \qquad (17)$$

The Heat Transfer Model

The eddy currents created within the workpiece heat the material due to Joule effects. Temperature evolutions $u(t)$ are governed by the classical heat transfer problem:

$$\frac{\partial u}{\partial t} - div(a\ gradu) = f \qquad (18)$$

where a is the temperature dependent material diffusivity and the loading term is $f = sE^2 / rC$. Different boundary conditions can be involved in electro-thermal processes. We present here the case of a prescribed temperature on Γ. The functional space V is here defined as $V = \{v \in H_0^1(\Omega), v = 0 \text{ on } \Gamma\}$. The associated weak formulation is:

$$\begin{aligned} &a^T(\dot{u}, y) + b^T(u, y) = l^T(y), \forall y \in V, t > 0 \\ &u(0) = u_0 \end{aligned} \qquad (19)$$

where

$$a^T(\dot{u}, y) = \int_\Omega \dot{u}(t) y\,, \qquad (20)$$

$$b^T(u, y) = \int_\Omega a(u(t)) \nabla u(t) . \nabla y \qquad (21)$$

$$l(y) = \int_\Omega f(t) y = \int_\Omega sE^2(t) y \qquad (22)$$

Time integration – coupling strategy

Both electromagnetic and thermal problems ((14) and (19)) being transient, both problems must be numerically integrated in time. For both problems, we have used a second order implicit integration scheme as detailed in [2].

A weak coupling strategy has been chosen. This means that the electromagnetic problem is solved at a given temperature field. In the same way, the heat transfer problem is solved at a given electrical field – calculated by the electromagnetic problem –.

A good choice for the electromagnetic time step integration is not straight forward. However, a time step value of around $T/32$ – where T stand for the electromagnetic period – seems to be a good compromise between computation time costs and results precision.

Moreover, it can be proved that, at a given temperature field, all electromagnetic fields taken on a full period of time stabilize. As a consequence, it is not necessary, at a given temperature field, to go on computing the electromagnetic problem as soon as the electromagnetic fields have stabilized.

Next, a good choice for the heat transfer time integration is neither straight forward. Here again, the time integration step is determined throughout a compromise between computation time costs and results precision. The optimal choice for the time step is given by time discretization error estimates as presented – for automatic time step control – by [4].

The diffusivity $a = k / rC_p : \Re \to \Re^+$ being bounded such that:

$$c_0 \le a(r) \le C, |a'(r)| \le C \ \forall r \in \Re \qquad (23)$$

and choosing a tolerance d between the exact solution u and the approached solution T so that the error is bounded by:

$$\max_{t \le t_n} \|u(t) - T(t)\| \le d \qquad (24)$$

with $\|.\| = \|.\|_2$,.then the algorithm an optimal time integration step is as follows:

1. Choose Δt_n
2. Δt_n being given, calculate the corresponding solution T_n
3. if $\frac{d}{gC} \le \|U_n - U_{n-1}\| \le \frac{d}{C}$ then Δt_n is accepted else Δt_n is increased or decreased, go back to 2.

Algo. 1: choice for a good thermal time step integration

This algorithm, based on error estimates, leads to consider a so-called ultra-weak coupling strategy rather than a weak one. The ultra-weak coupling is performed as follows:

1. Electromagnetic fields are computed at a given temperature map using time steps $\Delta t = T/32$ until convergence over two successive periods. The source term involved in the heat transfer equation is averaged.
2. The heat transfer problem is solved with an integration time step determined by the previously presented algorithm (*Algo. 1*). In practical, the heat transfer time step may reach a second. Go to first step 1.

Algo. 2: choice for a good thermal time step integration

OPTIMIZATION PROBLEM RESOLUTION

We now focus on the optimization problem (7). The direct electro-thermal model being highly nonlinear, we have used gradient type methods to solve (7). At each iteration l the solver searches the descent direction dP^l and a direction step a^l such that the objective function is lowered:

$$P^{l+1} = P^l + a^l . dP^l : j\left(P^{l+1}\right) < j\left(P^l\right) \quad (25)$$

Two different approaches are considered whether one wants to identify the three coefficients a, b and g or only coefficients a_{lin} and g. Indeed, it is possible to develop an analytical sensitivity calculation when dealing with both coefficients a_{lin} and g, but this seems very difficult when dealing with all coefficients. Therefore, a finite difference approach has been developed in this latter case, and an adjoint method in the first case.

m(H,T) = m(a,b,g) : A Gauss-Newton / Finite difference approach

A general Gauss-Newton method has been used to calculate, at each iteration l the descent direction dP^l. This method is based on the minimization of the Taylor developed objective function gradient, neglecting all second (and higher) order derivatives:

$$\frac{dj}{dP}\left(P^l + dP^l\right) \cong \frac{dj}{dP}\left(P^l\right) + \frac{d^2 j}{dP^2}\left(P^l\right) dP^l \quad (26)$$

Combining (5) and (26), the descent direction is calculated, for all parameters $P_k : P_1 = a, P_2 = b, P_3 = g$, by:

$$\left[\sum_{i=1}^{nbdata}\left(\frac{1}{\left(T_i^{ex}\right)^2}\frac{dT_i^c}{dP_j}\frac{dT_i^c}{dP_k}\right)\right]\{dP_k\} + \left\{\sum_{i=1}^{nbdata}\left(\frac{\left(T_i^c - T_i^{ex}\right)}{T_i^{ex}}\frac{dT_i^c}{dP_k}\right)\right\} = 0 \quad (27)$$

where sensitivity components are calculated using the finite difference method which consists in calculating, using the direct model, the values of each $T_i^c(P)$ and $T_i^c(P+dP)$ where dP represents a small perturbation of parameter P_k.

m(T)= m(a lin,g) : A conjugate gradient / Adjoint approach

For this, we could also consider a Gauss-Newton approach coupled with a finite difference sensitivity analysis but, for powerfulness reasons, we shall rather consider, here, a sensitivity analysis based on the optimal control approach. We first consider discretized systems $R^E(P,E) = 0$ and $R^T(E,T) = 0$ respectively related to electromagnetic and thermal problems posed in their weak form in (14) and (19). To solve the optimization problem (7), we have decided to use a conjugate gradient method coupled with a parabolic interpolation for the linear research algorithm. A sensitivity analysis must then be performed. We thus build the lagrangian defined as:

$$L\left(P,E,T,l^E,l^T\right) = J(T) + \left\langle R^E, l^E\right\rangle_{\Omega\times[t_0,t_f]} + \left\langle R^T, l^T\right\rangle_{\Omega\times[t_0,t_f]} \quad (28)$$

where scalar products involved in (28) are defined as:

$$\left\langle u,v\right\rangle_{\Omega\times[t_0,t_f]} = \int_{t_0}^{t_f}\int_{\Omega} uv\,d\Omega dt \quad (29)$$

We are now interested in finding the lagrangian saddle point. This point, where

derivations of the lagrangian with respect to physical parameters $P_k : P_1 = a_{lin}, P_2 = g$ equal zero, also corresponds to the minimum of the objective function.

Adjoint equations

The adjoint equations associated to (28) are given in (30) and (31)

$$\left\langle \frac{\partial R^E}{\partial E}\frac{dE}{dP_k}, l^E \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \frac{\partial R^T}{\partial E}\frac{dE}{dP_k}, l^T \right\rangle_{\Omega\times[t_0,t_f]} = 0 \tag{30}$$

$$\left\langle \frac{\partial R^T}{\partial T}\frac{dT}{dP_k}, l^T \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \frac{\partial J_2}{\partial T}, \frac{dT}{dP_k} \right\rangle_{\Omega\times[t_0,t_f]} + \left\langle \frac{\partial J_1}{\partial T}(t_f), \frac{dT}{dP_k}(t_f) \right\rangle_{\Omega} = 0 \tag{31}$$

Both adjoint equations (30) and (31) are well posed in the backward time integration. Thus, they are to be solved from the terminal time t_f down to the initial time t_0.

The objective function being involved only in the heat transfer adjoint problem, one has first to solve the thermal adjoint equation expressed in (32), and use the thermal adjoint variable to solve the electromagnetic adjoint problem expressed in (34).

$$\begin{aligned} &-a^T(l^T, y) + b^T(l^T, y) = l^{lT}(y), \forall y \in V, t > 0 \\ &l^T(t_f) = \frac{\partial J_1}{\partial T}(t_f) \end{aligned} \tag{32}$$

where the second hand side of the thermal adjoint problem (32) is expressed as:

$$l^{lT}(y) = \int_\Omega \frac{\partial J_2}{\partial T}(t) y \, ds \tag{33}$$

The electromagnetic adjoint problem is written as:

$$\begin{aligned} &-a^E(l^E, y) + b^E(l^E, y) = l^{lE}(y) \ \forall y \in V, t > 0 \\ &l^E(t_f) = 0 \end{aligned} \tag{34}$$

where the second hand side (34) is expressed as:

$$l^{lE}(y) = \int_\Omega 2sE(t) l^T(t) y \, ds \tag{35}$$

Objective function gradient

Both adjoint equations being verified within the whole crossed space – time domains $\Omega \times [t_0, t_f]$, the objective function gradient is determined by:

$$\frac{dJ}{dP_k} = \left\langle \frac{\partial R^E}{\partial P_k}, l^E \right\rangle_{\Omega\times[t_0,t_f]} \tag{36}$$

Determination of vectors $\partial R^E / \partial P_k$, for each P_k is done in an analytical way, these parameters being continuous physical parameters as shown in (16).

Inverse coupling strategy

The inverse coupling strategy has been designed to include the main features of the algorithm used in the direct electro-thermal mathematical model.

Indeed, both adjoint problems, which are to be solved in the backward time, involve the same capacity and rigidity matrices as in the direct model. Only loadings and initial conditions differ.

Moreover, the thermal adjoint problem does not depend on any electromagnetic quantity. This has lead us to integrate in time the thermal adjoint problem using exactly the opposite of the time step used within the direct thermal problem.

Next, as for the direct model, electromagnetic adjoint fields – taken on a period of time – stabilize for given thermal adjoint and temperature fields. This means that it is not necessary to compute the electromagnetic adjoint problem on the whole time range $[t_f, t_0]$ as is shown in *Figure 2*.

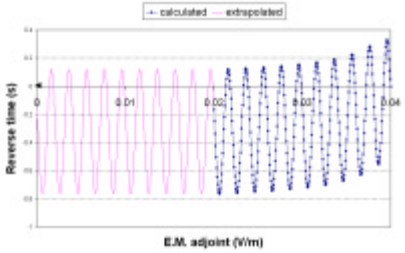


Figure 2: Calculated electromagnetic adjoint variable on the workpiece surface. It shows that 8 periods of time are necessary for converge over two successive periods. Afterwards, the electromagnetic adjoint field is extrapolated.

IDENTIFICATION RESULTS

The experimental set-up

The experimental set-up consists in heating, by induction way, a round magnetic billet of dimensions 0.03 meter diameter and 0.08 meter long. A seven turns coil surrounds the billet. Five thermocouples are embedded on the billet surface at equal distance from one to another. The corresponding mesh is shown in *Figure 3*.

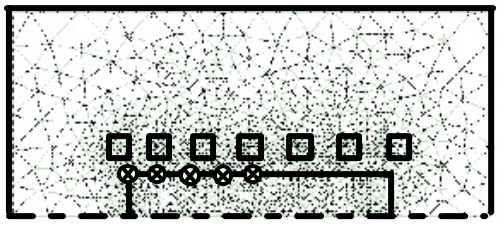


Figure 3: The set-up corresponding finite element mesh

The exciting current within the seven turns coil is a sine wave of frequency 50 hertz and of amplitude 10^9 amperes per meter squared.

Temperature evolutions are measured by the five thermocouples embedded on the surface at the frequency of one measure every half second, the whole experimental test lasting twenty seconds in total, and the maximal temperature increase reaching 800°C.

The global approach

As a first approach, all measured temperature points are involved in the objective function (5) to be minimized, and also in the sensitivity analysis. Obtained results were disappointing, since no test converged to a very low objective function. The problem seemed to be ill-posed. Taking a larger number of observables or using regularization methods did not improve the convergence.

The uncoupling approach

Therefore, a new uncoupled approached has been tested. This approach consists in uncoupling the identification of magnetic parameters with the identification of the temperature sensitivity component. Indeed, at low temperature, sensitivities of temperature with respect to the temperature exponent coefficient g is very low. Thus, temperature evolutions at beginning of heating are mostly due to magnetic parameters.

The procedure we set is the following. We first evaluate the magnetic parameters a_{lin} or $(a+b)$ taking into account only the first temperature data within the objective function (5). As a second stage the objective function is minimized with respect to the temperature coefficient g – assuming fixed magnetic parameters. This whole procedure is repeated several times until the objective function gets very low.

Hereafter are presented the main results on both cases, that is when dealing with the full magnetic permeability dependency -–when three coefficient are to be identified – and on the single magnetic permeability dependency – when two coefficient are to be identified.

m(T,H)=m(a,b,g) identification

The uncoupling identification approached with a Gauss-Newton minimization algorithm and a finite difference approach for the sensitivity components calculation gave the following results as shown in *Figure 4* and *Figure 5*.

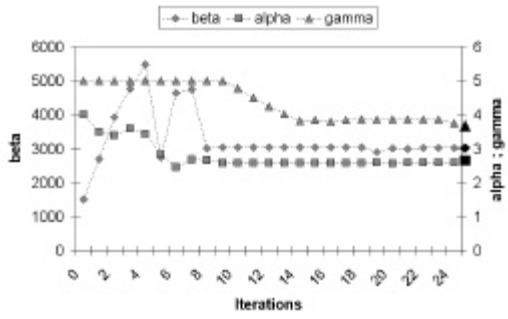


Figure 4: evolution of the calculated physical parameters (Frölich coefficients a and b) and

temperature sensitivity g with respect to identification iterations

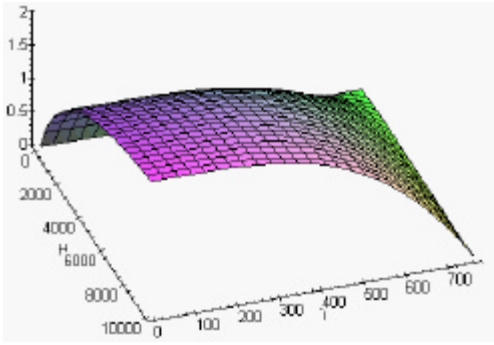


Figure 5: The full identified magnetization curve B(H,T). B is expressed in Teslas, H in Amperes per meter and temperature in °Celcius.

$\mu(T)=\mu(a_{lin},g)$ identification

Figure 6 below represents evolutions of the physical parameters with respect to identification iterations. The identification test shown hereafter has started with initial guessed values of $a_{ini} = 1000$ and $g_{ini} = 10$.

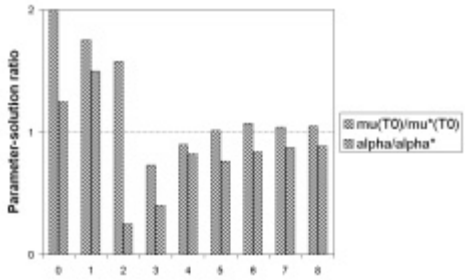


Figure 6: Physical parameters evolution with respect to identification iterations

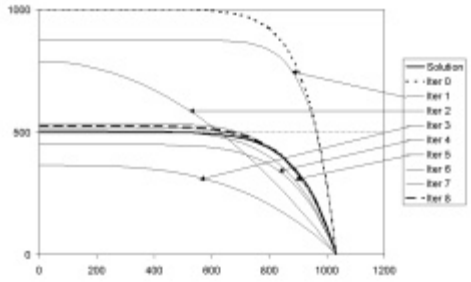


Figure 7: Evolution of the magnetic permeability as a function of temperature (°C) for several identification iterations

CONCLUSION

A new strategy for identification of some magnetic properties through temperature measurements has been presented. This method enables initial measurements in a quite simple and inexpensive way. The main mathematical innovation of this approach resides in the fact that the sensitivity analysis is performed on coupled distinct problems where typical time integration steps are much different for both. Identification results, which have been presented for both magnetic formulations – linear and nonlinear magnetic field strength dependencies – are in very good accordance with some classical magnetic properties measurements (made at DTU – Danish Technical University). This proves the good applicability of magnetic properties identification by an inverse method coupled with an electromagnetic – heat transfer finite element modeling.

REFERENCES

1 D. Labridis, P. Dokopoulos, “Calculation of eddy currents losses in non-linear ferromagnetic materials”, *IEEE Trans. Magn.*, 33, 1, pp. 746-752, 1997

2 F. Bay, V. Labbé, Y. Favennec & J.L. Chenot, “A numerical model for induction heating processes coupling electromagnetism and thermomechanics”, Submitted in *Int. J. Num. Meth. Engrg.*, October 2001

3 H. Bleuvin, “Analyse par la méthode des éléments finis des phénomènes magnéto-thermiques. Application aux systèmes de chauffage par induction”, *PhD Thesis, INPG*, France, 1984, in French

4 K. Eriksson, C. Johnson, “Error estimates and automatic time step control for nonlinear parabolic problems”, *SIAM I. Num. Anal.*, 24, pp. 12-23, 1987

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ESTIMATION OF INSULATING MATERIAL'S THERMOPHYSICAL AND RADIATION PROPERTIES

Aleksey V. Nenarokomov
Department of Space Systems Engineering
Moscow Aviation Institute
4 Volokolamskoe Sh., 125993, Moscow, Russia
nenar@cosmos.rcnet.ru

Oleg M. Alifanov
Department of Space Systems Engineering
Moscow Aviation Institute
4 Volokolamskoe Sh., 125993, Moscow, Russia
alf@cosmos.rcnet.ru

ABSTRACT

In many practical situations it is impossible to measure directly such properties of analyzed materials (for example, composite) as thermal and radiation characteristics. The only way that can often be used to overcome these difficulties is indirect measurements. This type of measurements is usually formulated as the solution of inverse heat transfer problems. Such problems are ill-posed in mathematical sense and their main feature shows itself in the solution instabilities. That is why special regularizing methods are needed to solve them. The experimental methods of identification of the mathematical models of heat transfer based on solving of the inverse problems are one of the modern effective solving manners. The objective of this paper is to estimate thermal and radiation properties of advanced materials using the approach based on inverse methods.

INTRODUCTION

In order to determine the thermal characteristics of modern structural, thermal-protective and thermal-insulating materials as a function of temperature, the inverse heat transfer problem should be solved. These problems have been recognized as highly effective in the analysis of heat transfer processes in manufacturing of high temperature composite materials, in many applications including in car industry, aerospace industry, metallurgy, thermal control systems and others. New methodology under development combines accurate measurements of thermal quantities, which can be experimentally observable in real conditions and accurate data processing, based on the solutions of inverse heat transfer problems.

In the present paper, the development of methods for estimating thermal and radiation properties is carried out for thermally stable high temperature composite materials. For such materials the goal is to estimate the characteristics as temperature functions by using results of measuring boundary conditions and temperature histories inside the body under consideration. The conditions of uniqueness usually define the minimum volume of measurements needed in one experiment. For example: to simultaneously determine the dependencies of thermal conductivity and volumetric heat capacity on temperature it is necessary to measure the non-zero heat flux entering into a specimen at least at one boundary and make transient temperature measurements at least than in two internal points. The boundary conditions of the first kind or any condition of heat insulation on both boundaries can be preset, but in particular case a specimen should be multi-layered and contain one layer of the material with known thermophysical characteristics and the number of temperature measurement points in the material layer under study should also be not less than two.

The most promising direction in further development of research methods for non-destructive composite materials using the procedure of inverse problems is the simultaneous determination of a combination of material's thermal and radiation characteristics (thermal conductivity $k(T)$, heat capacity $C(T)$ and total hemispherical emissivity $\varepsilon(T)$, where $T(x,\tau)$ is the temperature) [1]. Such problems are of great practical importance in the study of properties of composite materials used as non-destructive surface coating in objects of space technology, power engineering etc.

The mathematical model of heat transfer in specimen for the case of isothropic material is

$$C(T)\frac{\partial T}{\partial \tau}=\frac{\partial}{\partial x}\left(k(T)\frac{\partial T}{\partial x}\right),\ x\in(X_0,X_1),\tau\in(\tau_{\min},\tau_{\max}] \quad (1)$$

$$T(x,\tau_{\min})=T_0(x),\quad x\in[X_0,X_1] \quad (2)$$

$$-\beta_1 k(T)\frac{\partial T(X_0,\tau)}{\partial x}+\alpha_1 T(X_0,\tau)=q_1(\tau),\quad \tau\in(\tau_{min},\tau_{\max}] \quad (3)$$

$$-k(T)\frac{\partial T(X_1,\tau)}{\partial x}=-q_2(\tau)+\varepsilon(T)\sigma T^4(X_1,\tau),\quad \tau\in(\tau_{min},\tau_{\max}] \quad (4)$$

where $\beta_1=\{0,1\}$ and $\alpha_1=\{0,1\}$ are parameters, which provide using boundary condition of the first ($\beta_1=0,\ \alpha_1=1$) and the second ($\beta_1=1,\ \alpha_1=0$) kinds. Heat fluxes q_1 and q_2 have their positive directions along OX axis.

In equations (1)-(4) the quantities $C(T)$, $k(T)$ and $\varepsilon(T)$ are unknown. The experimental equipment and the method described below could be applied for estimating of material's three characteristics; the availability of a few specimens of the material allows us to provide uniqueness of the solution. This paper is concerned with modification of the approach, presented at [1] (further it will be called approach A). The main difference from the approach A is another way in interpretation of input data of the experiment, namely, the heat flux from the metallic heater was considered at [1] as Stefan-Boltsman irradiation with known (measured by thermocouple) temperature of the heater $T_h(\tau)$ and a-priory known (theoretically) emissivity of heater $\varepsilon_h(T)$. If emissivity of the heater material is known a-priori, and its temperature is measured, the heat flux from the heater can be derived as

$$q_2(\tau)=\varepsilon_h(T_h)\sigma T_h^4(\tau) \quad (4a)$$

At this paper the emissivity of the heater is considered as additional (fourth) estimating functions. Therefore the accuracy of the inverse problems, considered bellow, will not depend to the a-priory information about the radiation properties of the heater's material.

The results of temperature measurements inside the specimen are assigned as necessary additional information to solve an inverse problem

$$T^{\exp}(x_m,\tau)=f_m(\tau),\quad m=\overline{1,M} \quad (5)$$

In the inverse problem (1)-(5) it is necessary first of all to indicate as a temperature range $[T_{\min},T_{\max}]$ of the unknown functions, which is general for all experiments, and for which the inverse problem analysis has a unique solution. For T_{min} the minimum value of initial temperature is used. Of much greater importance is a correct sampling of value T_{max}. Proceeding from the necessity to provide uniqueness of solution, it seems possible to sample, for $T_{max,}$ a minimum among maximum temperature values gained on the thermocouple positioned on the heated surface at every testing specimen. The same should be done for the heater temperature range $[T_{h,\min},T_{h,\max}]$.

The execution of the single experiment is not enough to provide the conditions of uniqueness of the inverse problem solving by simultaneous estimating of thermal conduction, heat capacity and emissivity of the testing material. To solve this problem the data of several N (in partial three) similar experiments with equal material specimen and different heating regimes were processed simultaneously.

INVERSE PROBLEM ALGORITHM

At this section the algorithm of inverse problem is presented for the case, when $q_2(\tau)$ is a-priori known to provide independence of algorithm from experimental merits. Suppose then that the unknown characteristics are given in their parametric form. With this purpose introduce in the interval $[T_{\min},T_{\max}]$, three uniform difference grids with the number of nodes N_i, i=1,2,3.

$$\omega_i=\left\{T_k=T_{\min}+(k-1)\Delta T,\ \ k=\overline{1,N_i-1}\right\},\quad i=\overline{1,3} \quad (6)$$

The function

$$B^{(j-1)}=B^{(j-1)}(T_k,T_{k+1},\ldots,T_{k+j},\tau)=\sum_{s=k}^{k+j}\frac{(T_s-T)_+^{j-1}}{\omega_k'(T_s)}$$

where

$$\omega_k'=(T-T_k)(T-T_{k+1})\ldots(T-T_{k+j})$$

and

$$(T_s-T)_+^{j-1}=\max\left\{0,(T_s-T)^{j-1}\right\}$$

is called *B*-spline of the *(j-1)* degree relatively with nodes $T_k, T_{k+1}, ..., T_{k+j}$. When solving practical problems, *B*-splines are used with so-called "natural" boundary conditions

$$u''(T_{\min}) = u''(T_{\max}) = 0$$

where *u* is desired function.

Then, in case of cubic B-splines (*(j-1=3),* the unknown function is presented as

$$u_i(\tau) = \sum_{k=1}^{N_i} u_k \varphi_{i,k}(T)$$

$$\varphi_1(T) = 2B_0(\overline{T} + \Delta T) + B_0(\overline{T})$$

$$\varphi_2(T) = -B_0(\overline{T} + \Delta T) + B_0(\overline{T} - \Delta T)$$

$$\varphi_k(T) = B_{k-1}(\overline{T}),\ k = 3, ..., N_i - 2$$

$$\varphi_{N_i-1}(T) = B_0(\overline{T} - (N_i - 2)\Delta T) - B_0(\overline{T} - N_i \Delta T)$$

$$\varphi_{N_i}(T) = B_0(\overline{T} - (N_i - 1)\Delta T) + 2B_0(\overline{T} - N_i \Delta T)$$

where

$$B_k(T) = B_0(\overline{T} - k\Delta T)$$

$$\overline{T} = T - T_{\min}$$

$$B_0(T) = ((T + 2\Delta T)_+^3 - 4(T + \Delta T)_+^3 + 6(T)_+^3$$

$$-4(T - \Delta T)_+^3 + (T - 2\Delta T)_+^3)/(6\Delta T^3)$$

The function $B_0(T)$ has the property

$$B_0(T) = \begin{cases} >0, & if\ -2\Delta T < T < 2\Delta T \\ =0, & if\ |T| \geq 2\Delta T \end{cases}$$

This property makes the computational algorithm simpler.

In solving any particular inverse problem the choice of the number of parameters of approximation for unknown characteristics should be justified. Approximating the unknown functions on grids (6) using the cubic B-splines

$$C(T) = \sum_{k=1}^{N_1} C_k \varphi_{1,k}(T)$$

$$k(T) = \sum_{k=1}^{N_2} k_k \varphi_{2,k}(T) \qquad (7)$$

$$\varepsilon(T) = \sum_{k=1}^{N_3} \varepsilon_k \varphi_{3,k}(T)$$

where C_k, *k=1,N1,* k_k, *k=1,N2,*, ε_k, *k=1,N3* - parameters.

As a result of approximation, the inverse problem is reduced to the search of a vector of unknown parameters

$$\overline{p} = \{C_1, ..., C_{N_1}, k_1, ..., \varepsilon_{N_3}\} = \{p_k\}, k=1,N_p$$

with dimensions $N_p = N_1+N_2+N_3$. Writing down a least-square error of the design and experimental temperature values in points of thermal sensors positioning

$$J(C(T), k(T), \varepsilon(T)) == J(\overline{p}) =$$

$$= \sum_{n=1}^{N} \sum_{m=1}^{M_n} \int_{\tau_{\min}^m}^{\tau_{\max}^m} \left(T^n(x_m^n, \tau) - f_m^n(\tau)\right)^2 d\tau \qquad (8)$$

where $T^n(x, \tau)$ is determined from a solution of the boundary-value problem (1)-(4) for *n*-th experiument using the approximations of (7). It is assumed here that the conditions of uniqueness of the inverse problem solving are satisfied. Bellow to simplify the niotation equations index *n* will be excluded.

In processing the results of real experiments and tests there are always errors, depending on a number of reasons. First of all, the errors in the experimentally measured data are both random and systematic by nature. Random errors in input data are stipulated by a spread of thermal and electrical characteristics in thermal sensors and other measuring devices, by inaccuracy of their calibration, etc.; as a rule, these errors show a large enough value. Systematic errors in the input data f are usually connected with inaccuracy in determining the co-ordinates of thermal sensors, with temperature pattern distortion in the specimen by thermal sensors, with displacement of the very specimen during filming or X-ray survey etc. That is why the term f_m is always known with some error.

The second group of errors - the errors of finite-difference approximation of differential operator in the initial problem and round-off errors in the computer. Besides, there are errors because of uncertainties in the a-priori assigned characteristics of a mathematical model Eqs.(1)-(4), which are determined either through calculations or from a solution of the corresponding inverse problems.

Due to mathematical ill-posedness of the problem in defining the heat transfer

characteristics, the availability of errors in the input information leads to a solution, which can have nothing in common with an unknown one. To overcome this circumstance it was suggested to make use of the regularized properties of gradient methods of optimization, which permit to begin quickly an iterative process from a distant estimate of characteristics to be determined and it slow down abruptly while approaching the functional minimum. As the number of iterations increases, the solution of the inverse problem can start oscillating, gradually loosing its smooth nature. A necessity occurs to stop the iterative process, falling short of solution oscillations. An important task here is to sample a stop criterion for the iterative process. By principle of generalized discrepancy such criterion can be a constraint on the value of minimized functional $J(\overline{p})$.

A successive regularized method realised here, according to general definition of Tikhonov's regularizing operators, gives rise to a regularizing set of operators, in which a parameter of regularization is the number of the last iteration [3]. For linear ill-posed problems the iterative regularization method has received a rigorous mathematical proofs and practical verification through data of mathematical modelling. Applicable to non-linear problems, there are at present no complete theoretical results on the substation of stability of iterative algorithms. However, the results of computational experiments already made in solving the inverse heat transfer problems of different types prove high efficiency of the iterative regularization method and possibility to analyze a wide range of nonlinear problems.

So, proceeding from the principle of iterative regularization [1,2], the unknown vector can be determined through minimization of functional (8) by gradient methods of the first order prior to a fulfillment of the condition

$$J(\overline{p}) \le \delta_f \tag{9}$$

where $\delta_f = \sum_{m=1}^{M} \int_{\tau_{\min}}^{\tau_{\max}} \sigma_m(\tau)\, d\tau$ is integral error of temperature measurements $f_m(\tau)$, m=1,M, and σ_m - measurement variance.

To construct an iterative algorithm of the inverse problem solving a conjugate gradient method was used. The successive approximation process is constructed as follows:

1) a-priori, an initial approximation of the unknown parameter vector $\overline{p}^0$ are set

2) a value of the unknown vector at the next iteration are calculated

$$\begin{aligned} &\overline{p}^{s+1} = \overline{p}^{s} + \gamma^s \overline{g}^{s} \\ &\overline{g}^{s} = -\overline{J'}^{s} + \beta^s \overline{g}^{s-1} \\ &\beta^0 = 0, \\ &\beta^s = \left\langle \left(\overline{J}_p'^{(s)} - \overline{J}_p'^{(s-1)}\right), \overline{J}_p'^{(s)} \right\rangle_{R^{N_P}} \Big/ \left\| \overline{J}_p'^{(s)} \right\|_{R^{N_P}} \end{aligned} \tag{10}$$

where $\overline{J}_p'^{(s)}$ - value of the functional gradient at current iteration.

The greatest difficulties in realizing the gradient methods are connected with calculation of the minimized functional gradient. In the approach being developed the methods of calculus of variations are used. Here an analytic expression for the minimized functional gradient can be obtained

$$J'_{C_k} = -\int_{\tau\max}^{\tau\max} \int_{x_0}^{x_1} \psi(x,\tau)\cdot \varphi_{1,k}(T) \frac{\partial T}{\partial \tau}\, dx d\tau \qquad k = \overline{1, N_1} \tag{11}$$

$$\begin{aligned} J'_{k_k} = &-\int_{\tau\max}^{\tau\max} \int_{x_0}^{x_1} \psi(x,\tau)\cdot \Big(\frac{\partial^2 T}{\partial x^2}\cdot \varphi_{2,k}(T) \\ &+ \left(\frac{\partial T}{\partial x}\right)^2 \cdot \frac{\partial \varphi_{2,k}}{\partial T} \Big)\, dx d\tau \\ &- \beta_1 \int_{\tau\max}^{\tau\max} \psi(X_0,\tau) \frac{\partial T}{\partial x}(X_0,\tau) \varphi_{2,k}(T(X_0,\tau))\, d\tau \\ &+ \int_{\tau\max}^{\tau\max} \psi(X_1,\tau) \frac{\partial T}{\partial x}(X_1,\tau) \varphi_{2,k}(T(X_1,\tau))\, d\tau \end{aligned} \qquad k = \overline{1, N_2} \tag{12}$$

$$J'_{\varepsilon_k} = -\int_{\tau\max}^{\tau\max} \psi(X_1,\tau)\cdot \varphi_{3,k}(T)\, \sigma T^4(X_1,\tau)\, d\tau \qquad k = \overline{1, N_3} \tag{13}$$

where $\psi(x,\tau)$ is solution of a boundary-value problem adjoint to a linearized form of the initial problem (1)-(4).

$$-c(T)\frac{\partial\psi_m}{\partial\tau}= k\frac{\partial^2\psi_m}{\partial x^2}, x\in(x_{m-1},x_m), x_0 = X_0,$$
$$x_{M+1} = X_1,\ m=\overline{1,M+1},\ \tau\in(\tau_{min},\tau_{max}] \quad (14)$$

$$\psi_m(x,\tau_{\max})=0,\ x\in[x_{m-1},x_m],\ m=\overline{1,M+1} \quad (15)$$

$$-\beta_1 k\frac{\partial\psi_1(X_0,\tau)}{\partial x}+\alpha_1\psi_1(X_0,\tau)=0, \quad (16)$$

$$-k\frac{\partial\psi_{m+1}(X_1,\tau)}{\partial x} \quad (17)$$
$$-\varepsilon(T)\sigma T^3(X_1,\tau)\ \psi_{m+1}(X_1,\tau)=0,$$

$$\psi_m(x_m,\tau)=\psi_{m+1}(x_m,\tau),\ \ m=\overline{1,M} \quad (18)$$

$$\lambda(T)\cdot\left(\frac{\partial\psi_m(x_m,\tau)}{\partial x}-\frac{\partial\psi_{m+1}(x_m,\tau)}{\partial x}\right) \quad (19)$$
$$=2\left(T(Y_m,\tau)-f_m(\tau)\right),\quad m=\overline{1,M}$$

EXPERIMENTAL VERIFICATION

As an example, the results of data processing on experimental study of material specimens are used, which involves the thermo-vacuum facility developed in MAI at the Department of Space System Engineering. The test facility consists of a horizontally cylindrical set vacuum chamber with a water-cooling system.

A heating element consists of a molybdenum foil with dimensions 80x70x0.05 mm (Fig. 1). The tested specimen is located on a thermo-insulating base made of thermal insulation material so that its heating surface is parallel to heater's plane and at certain distance from it (3-5 mm). A ceramic plate is put over the heater, which reduces the heat withdrawal from it in the direction opposite to the test specimen improving its heating uniformity. Control of the specimen heating regime is performed by temperature on the exposed surface in accordance with a preset regime. Control over the heating regime was maintained in the experiment using the results of three trial starts with specimens from test materials, the structure of which being similar to the structure used in tests. In making the trial starts a criterion of choosing a suitable heating regime is the coincidence of specimen external surface temperature measured with the preset temperature.

With the proposed scheme of test performance, the data gained in one experiment are not sufficient for simultaneous recovery of two thermal characteristics (thermal conduction and heat capacity), because data by values of the heat flux applied to a specimen are also needed. If emissivity of the heater material (ε_h) is known a-priori, and its temperature is measured via control temperature, the heat flux from the heater can be calculated according (4a). In presented paper a case is considered, when $\varepsilon_h(T)$ is not known a-priori and should be estimated, this approach will be called as approach B (to compare with presented in INTRODUCTION and in [1] approach A). Therefore some parts of inverse problems algorithm should be modify.

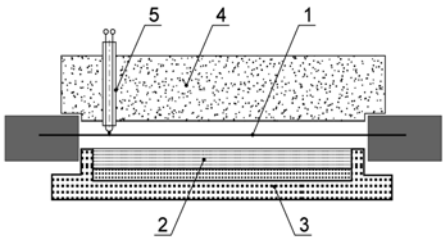


Figure 1. Experimental module: 1 – heater, 2 - test specimen, 3 - insulating basement, 4 - insulating cover, 5 - control thermocouple.

Lets introduce in the interval $[T_{h\min},T_{h\max}]$ uniform difference grids with the number of nodes N_4

$$\omega_4=\left\{T_k=T_{h\min}+(k-1)\Delta T_h, k=\overline{1,N_4-1}\right\} \quad (6a)$$

and approximate the unknown function $\varepsilon_h(T)$ on grids (6a) using the cubic B-splines

$$\varepsilon_h(T)=\sum_{k=1}^{N_4}\varepsilon_{hk}\varphi_{4,k}(T) \quad (7a)$$

where ε_{hk}, $k=1,N_4$ - parameters. Than the residual function (8) will depends to four functions $J(\overline{p})=J(C(T),k(T),\varepsilon(T),\varepsilon_h(T))$ and its gradient can be presented as (11-13) and

$$J'_{\varepsilon_{hk}}=-\int_{\tau\max}^{\tau\max}\psi(X_1,\tau)\cdot\varphi_{4,k}(T)\sigma T_h^{\,4}(\tau)\,d\tau$$
$$k=\overline{1,N_4} \quad (13a)$$

It should be noted here, that adjoint problem (14)-(19) remains without any transformations.

An example to how apply the approach suggested is presented bellow. Given are the results of data processing of specimen experimental investigations with modern

composite materials. The investigations were carried out on a set of pairs of specially manufactured specimens (the first in the pair for simultaneous estimating material's heat capacity per and thermal conductivity and the second for determining boundary conditions).

The models of test material are the square plates of 50x50x15 mm (Fig. 2) with four thermocouples installed in the specimen. A test material 1 is deposited on support 2 of 1mm thickness made of metallic alloy. Thermoelectrode wires of thermocouples were brought out via the fabric-based laminate and placed in glass braids. Length of the thermocouple isothermal zone was 7-8 mm. An installation of thermosensors in specimens was chosen from a solution of the problem of optimal experiment design. The co-ordinates of thermocouple positioning in the first set of specimens, for estimating the material's thermal characteristics, had the following values: X_o = 0 mm (for a boundary condition of the first kind sensor readings on the internal surface were used), x_1 = 7.5 mm, x_2 = 11.8 mm, $x_3 = X_1$ = 15 mm (positioned on the exposed surface). The second set of specimens, for defining the emissivity, has the thermocouples at points X_o = 0 mm, x_1 = 12.65 mm, x_2 = 13.65 mm, $x_3 = x_1$ = 15 mm.

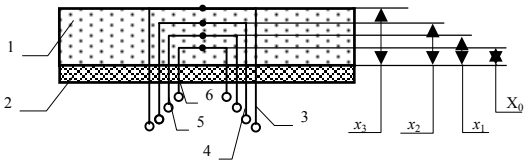


Figure 2. Test specimen: 1 - test materials, 2 - metallic basement, 3-6 - thermocouples.

The number of approximation parameters N_1, N_2, N_3 and N_4 for every characteristic was assumed to be 5. During specimen heating a theoretically preset time dependence of surface temperature (Fig. 3) was provided. The measurement errors were estimated as 5 %. A comparison of experimentally measured and calculated (with the help of thermal characteristics obtained from a solution of the inverse problem) temperature values in points of thermocouple positioning is shown on Fig. 3 (only for one specimen). The results are in agreement, which shows the robustness of the inverse problem algorithm.

The results proper of the inverse problem solving - the composite material thermal characteristics and emissivity are given on Fig. 4,5 (the results for two sets by three experiments in vacuum and air conditions). The accuracy of these results of the inverse heat conduction problem was ere verified using different (quite distinct from each other) initial approximations for an iterative process. The results show reasonable agreement.

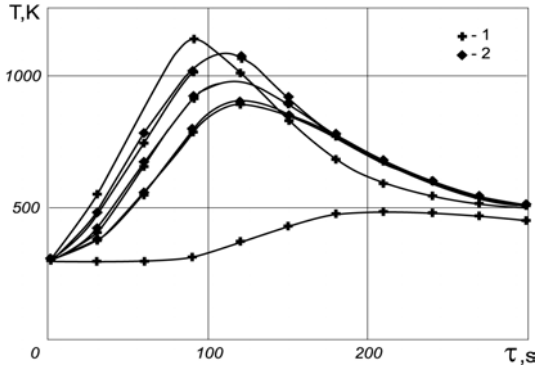


Figure 3. Temperature values in points of thermocouple positioning: 1 - calculated, 2 – experimental

CONCLUSIONS

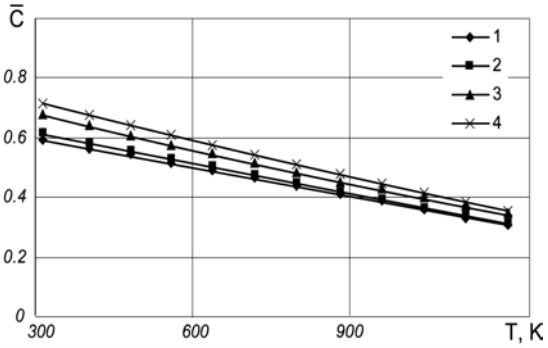
The paper seeks to describe the algorithm developed to process the data of unsteady-state thermal experiments. The algorithm is suggested for determining these unknown on the surface of a slab as a solution of the nonlinear inverse heat conduction problem in an extreme formulation.

The following main factors have an influence on the accuracy of the inverse heat conduction problem (in sequence of significance): the errors in coordinates of thermosensor positions; the errors in values of different characteristics; the errors in estimating the residual level. It was shown that in the cases considered the accuracy of the inverse problems solution is compatible with the errors of the simulated "experimental measurements".

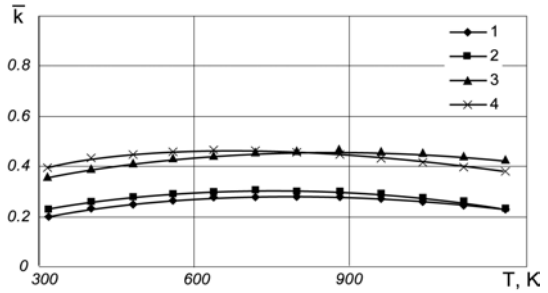
Next step in the development of the proposed approach is to consider an estimating interface conductance between periodically contacting surface of specimen and heater foil using the approach similar [13].

ACKNOWLEDGMENTS

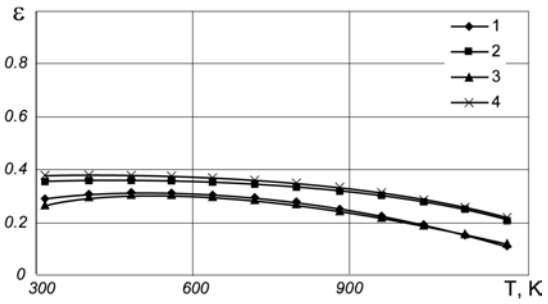
This work was done with the financial support support of RFBR (grant # 02-02-17304).



a



b



c

Figure 4. Results of testing: a - heat capacity, b - conductivity, c – emissivity.
1- approach A (vacuum), 2 – approach B (vacuum), 3- approach A (air), 4 – approach B (air).

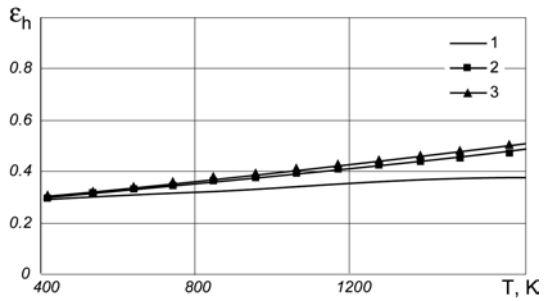


Figure 5. A-priori known and estimated emissivity of the heater foil. 1- a-priori known value (used in approach A), 2 – approach B (vacuum), 3 – approach B (air).

REFERENCES

1. A.V.Nenarokomov and O.M.Alifanov, Inverse radiative-conductive problems for estimating material properties, M.P.Menguc, N.Selcuk ed. *Proc. of the Third International Symposium on Radiative Transfer,* Begell House, New York\Wallingford (UK), 2001, p. 8.

2. O.M.Alifanov, *Inverse Heat Transfer Problems,* Springer-Verlag, Berlin \ Heidelberg, 1994, p. 348.

3. O.M.Alifanov, E.A.Artyukhin and S.V.Rumyantsev, *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Problems,* Begell House, New York \ Wallinford (UK), 1995, p. 306.

4. E.A.Artyukhin, G.A.Ivanov and A.V.Nenarokomov, Determination of a complex of materials thermophysical properties through data of nonstationary temperature measurements, *High Temperature*, **31**, 235 (1993).

5. E.A.Artyukhin and A.V. Nenarokomov, Coefficient Inverse heat conduction problem, *Journal of Engineering Physics*, **53**, 1085 (1987).

6. E.A. Artyukhin and A.V.Nenarokomov, Inverse Problem of Determining the Integral Emissivity of a Solid, *High Temperature*, **24**, 222, (1986).

7. D.Baillis, M.Raynaud and J.F.Sacadura, Spectral radiative properties of open-cell foam insulation, *Journal of Thermophysics and Heat Transfer*, **13**, 292, (1999).

8. C.-H.Ho and M.N.Ozisik, An inverse radiation problem, *Int. Journal Heat and Mass Transfer*, **32**, 335 (1989).

9. H.Y.Li, Estimation of thermal properties in combined conduction and radiation, *Int. Journal Heat and Mass Transfer*, **42,** 565, (1999).

10. H.Y.Li, Ozisik, M.N. Inverse radiation problem for simultaneous estimation of temperature profile and surface reflectivity, *Journal of Thermophysics and Heat Transfer*, **7**, No. 1, pp 88-93, 1993.

11. Li, H.Y. and M.N.Ozisik, Identification of the temperature profile in a absorbing, emitting and isotropically scattering medium by inverse

analysis, *Trans. ASME. Journal of Heat Transfer*, 1060 (1992).

12. H.Y.Li and C.Y.Yang, A genetic algorithm for inverse radiation problems, *Int. Journal Heat and Mass Transfer*, **40**, 1545, (1997).

13. H.R.B.Orlande and M.N.Ozisik, Inverse problem of estimating interface conductance between periodically contacting surfaces, *Journal of Thermophysics and Heat Transfer* , **7**, 319 (1993).

14. X.Sun and D.L.Jaggard, The inverse blackbody radiation problem: a regularization solution, *Journal of Applied Physics*, **62**, 4382 (1987).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

IDENTIFICATION OF THERMAL PROPERTIES OF MATERIALS WITH APPLICATIONS FOR SPACECRAFT STRUCTURES

Oleg M. Alifanov
Sergey A. Budnik
Valeriy V. Michailov
Aleksey V. Nenarokomov
Department of Space Systems Engineering
Moscow Aviation Institute
4 Volokolamskoe Sh., 125993, Moscow, Russia
alf@cosmos.rcnet.ru

Valeriy M. Yudine
Central Aero-HydroDynamic Institute
TsAGI
1 Zukovskiy Str., 140160,
Zukovskiy, Moscow Reg., Russia

ABSTRACT

This paper presents the results of computational and experimental studies of particular thermal processes in composite materials. The considered approach is based on the mathematical theory of ill-posed problems of mathematical physics. In the majority of cases this methodology is used for optimisation, but in a number of practical situations it is the sole technique available, as, for example, in measuring the transient heat fluxes and heat transfer coefficients. Owing to the structural version and homogeneous surface heating in specimens a one-dimensional heat transfer process is realized. A complex of thermal properties of the composite material (thermal conductivity $\lambda(T)$ and heat capacity coefficient $C(T)$) is estimated. The initial data for such problems are formed grounded on the results of measurements and include the boundary conditions and temperature-time values in several internal points of specimens. The type of boundary conditions and the number of points of temperature measurement should meet the conditions of uniqueness of the inverse problem solution under analysis.

INTRODUCTION

In space rocketry many problems deal with the structures operating in the conditions of intensive, often extreme thermal effects. The general tendency in the development of technology is connected with the increasing of the number of thermo-loaded engineering systems with the simultaneous increasing of their reliability and safe life. For space vehicles and reusable transportation systems the support of thermal conditions is one of the most important aspects of design and in determining the main design solutions. Of great importance is the thermal condition support for different engines (gas-turbine, liquid/solid propellant rocket engines, etc.), power plants, heat-exchange apparatus etc. The distinctive features of modern thermo-loaded structures in space engineering are the non-stationarity, non-linearity, multi-dimensionality and conjugate nature of heat-and-mass transfer processes. These distinctions confine a possibility of using many traditional design-and-theory and experiment methods. So, in developing the flight vehicles of different mission and type, traditionally there were both the development of new approaches and the improvement of already available research techniques. Similar problems exist in other branches of industry.

The optimal design of heat-loaded objects is actually an important problem in the presence of weight and cost restrictions on the structures developed for any mission, for example, the aircraft and rocket-space systems, internal combustion engines, nuclear power plants, metallurgical and chemical equipment etc. The problem of reducing the specific consumption of materials in machines and vehicles, their power consumption, harmful action on the environment is extremely urgent in civil ingineering. Here, it is impossible to create the thermal protection systems, meeting modern requirements, without carrying out the extensive theoretical and experimental studies of properties of the materials used also to design new materials with given properties. The priority and general orientation in the development of theoretical and experimental foundations of research, support and optimization

of thermal conditions of the materials and structures of all prototypes of modern technology should be given to the development and broad application of methods of mathematical and physical simulation of the thermal state of the objects under study and, especially, to the experimental-and-theoretical methods of diagnostics and identification based on the solution of inverse heat-and-mass transfer problems. The investigations of recent years showed that the utilization of such an approach is the most perspective and fruitful. It allows to take into consideration the actually existing effects of non-stationarity and non-linearity of heat-and-mass transfer processes, it displays high information efficiency and gives a possibility to conduct investigations close, to the maximum, to a full-scale or directly during the use of objects.

One of the most important directions of studies where the application of methods of inverse problems is not only expected but also extremely obligatory is the identification of heat transfer process in partially decomposed composite materials. An essential effect in exerted on the heat transfer in such material by the course of thermal destruction of the binder (matrix) and related processes. Chemical reactions of the destruction process are accompanied by thermal effects with the resulting formation in the material of the distributed heat sources and heat outflows, the power of which is determined by the very process of decomposition. The released gaseous products begin moving by a system of gas pockets formed in the direction of the least pressure loss (heated surface) thus cooling the more heated layers of the material. As this takes place, the chemical processes could run both in gaseous phase and at interaction with a solid residual of the partially decomposed binder or with a filler also accompanied by thermal effects. A change in chemical composition and formation of pores in the binder, shrinkage or swelling of the material results in the change of its thermal properties.

Depending on the type and mission of the material, the action of one or the other of the pointed factors on heat transfer in the material could be different. By this is meant that the availability in literature of a large number of mathematical models of heat transfer in different materials explains the situation. The mathematical models now available can be divided into four groups:

1) Models of a heat conduction in a solid with the effective thermal properties depending on temperature, temperature and time, temperature and its variable, as well as with the so called functional characteristics;

2) Models with a front of the destruction process and division of the material into virgin and completely charred zones with own thermal properties;

3) Models with regard to the kinetics of destruction and related processes;

4) Models with regard to thermomechanical effects.

Without going into details of advantages and disadvantages of these models, we will note that most of them are not provided in the metrological aspect, that is, with the data on their parameters and methods of determination of these parameters. The more sophisticated is the model: the more is the number of parameters, the greater are the difficulties of their determination.

The availability of such great number of models complicates the supporting of industries with reference information and design methods. That is why it is necessary to develop a base group of models and methods for determination of their parameters. Such group should consist of a number of complicated models, each successive one including a new range of essential, for heat transfer, processes with the continuous, by some parameters or system, experimental data for their determination. This provides the use, for each material, of the simplest model reflecting the most essential processes thus reducing the volume of experimental data.

As the initial model, it is expedient to adopt a model of one dimensional heat conductiomn in the slab (the coordinate of the left boundary is X_0 and the coordinate of the right boundary is X_1) with the temperature-dependent thermal properties, therefore the mathematical model of heat transfer is like

$$C(T)\frac{\partial T}{\partial \tau} = \frac{\partial}{\partial x}\left(\lambda(T)\frac{\partial T}{\partial x}\right), \quad x \in (X_0, X_1), \quad \tau \in (\tau_{\min}, \tau_{\max}] \tag{1}$$

$$T(x, \tau_{\min}) = T_0(x), \quad x \in [X_0, X_1] \tag{2}$$

$$-\beta_1 \lambda(T)\frac{\partial T(X_0, \tau)}{\partial x} + \alpha_1 T(X_0, \tau) = q_1(\tau), \quad \tau \in (\tau_{min}, \tau_{\max}] \tag{3}$$

$$-\beta_2\lambda(T)\frac{\partial T(X_1,\tau)}{\partial x}+\alpha_2 T(X_1,\tau)=q_2(\tau),\quad \tau\in(\tau_{min},\tau_{\max}] \qquad (4)$$

where $\beta_1=\{0,1\}$, $\beta_2=\{0,1\}$, $\alpha_1=\{0,1\}$ and $\alpha_2=\{0,1\}$ are parameters, which provide using boundary condition of the first ($\beta_1=0,\ \alpha_1=1$ or $\beta_2=0,\ \alpha_2=1$) and the second ($\beta_1=1,\ \alpha_1=0$ or $\beta_2=1,,\ \alpha_2=0$) kinds. Heat fluxes q_1 and q_2 have their positive directions along OX axis.

This model is sufficiently exact for a wide class of non-decomposed materials and could be applied as an approximate model for materials, the decomposition depth and destruction effect of which on the heat transfer are relatively not large [5]. Permissibility of the use of this model for each particular material and conditions of its application should be proven experimentally.

The modern approaches to the development of composite materials assume broad application of mathematical and physical simulation methods. But the mathematical simulation is impossible if there is no true information available on the characteristics (properties) of objects thus analyzed. In defining the thermal properties of modern structural, thermal-protective and thermal-insulating materials - as temperature-dependent - the most effective methods are based on the solution of the coefficient inverse heat conduction problems (CIHCP) [1-7]. Methods of Inverse Heat Transfer Problems (IHTP) were developed to increase the amount of information from thermal experiments and tests, to improve the accuracy and reliability of experimental data processing and interpretation, and also for investigating and control of Heat Transfer in the manufacturing processes. In the majority of cases in practice the direct measurement of materials' thermal properties, especially of complex composition, is impossible. There is only one way, which permits to overcome these complexities - the indirect measurement. Mathematically, such an approach is usually formulated as a solution of the inverse problem: through direct measurements of system's state (temperature, component concentration, etc.) the properties of the system analyzed are defined, for example, the materials' thermal characteristics. Violation of the cause-and-effect relations in the statement of these problems results in their correctness in mathematical sense (i.e., the absence of existence and/or uniqueness and/or stability of the solution). Hence to solve such problems special methods, usually called the regularized methods [4-6], were developed.

In developing new thermal-protective materials quite a great number of comparative heat tests are carried out, the purpose of which is clear from the analysis of thermal-protective properties of materials in different heating conditions corresponding to service conditions. The experimental specimens for such tests are manufactured in the form of a flat plate of the material analyzed. To control the assigned heating condition the temperature of the external heated surface is measured and to estimate the thermal properties of the material in study the temperature in two internal points of specimens and on the internal surface is measured (the temperature of the external surface is also used). In addition, the heat flux density is assumed to be known for the warm-up side of a specimen. Realization of this condition is possible through experimental means. The internal surface temperature is used as a boundary condition. In practice it is difficult to realize the uniform initial temperature distribution in specimens, hence the initial temperature distribution is approximated through the reading data at zero time. Thermal conductivity $\lambda(T)$ and heat capacity coefficient $C(T)$ are the two characteristics to be recovered unambiguously and simultaneously for the described procedure in conducting the experiment and by measuring the data obtained in one experiment [8].

IDENTIFICATION ALGORITHM

In the model (1)-(4) the functions $C(T)$ and $\lambda(T)$ are unknown. For additional information necessary to solve an inverse problem the results of temperature measurements inside a specimen are assigned

$$T^{\exp}(x_m,\tau)=f_m(\tau),\quad m=\overline{1,M} \qquad (5)$$

Suppose then that the unknown characteristics are given in their parametric form. With this purpose introduce in the interval $[T_{\min},T_{\max}]$ two uniform difference grids with the number of nodes *Ni, i=1,2* and the unknown functions are approximated on this grids using cubic B-splines

$$C(T)=\sum_{k=1}^{N_1}C_k\varphi_{1,k}(T),\ \lambda(T)=\sum_{k=1}^{N_2}\lambda_k\varphi_{2,k}(T) \qquad (6)$$

where C_k, $k=1,N_1$, λ_k, $k=1,N_2$ - parameters.

As a result of approximation, the inverse problem is reduced to a search of vector of unknown parameters $\overline{p} = \{p_k\}$, $k=1,N_p$, with the dimensions $N_p = N_1+N_2$. Lets introduce a mean-square error of the design and experimental temperature values in points of thermal sensors positioning

$$J(C(T),\lambda(T)) = \sum_{m=1}^{M} \int_{\tau_{\min}}^{\tau_{\max}} (T(x_m,\tau) - f_m(\tau))^2 d\tau \quad (7)$$

where $T(x_m,\tau)$ is determined from the solution of the boundary-value problem (1)-(4) using the approximations of (6). Proceeding from the principle of iterative regularization, the unknown vector $\overline{p}$ can be determined through minimization of functional (7) by gradient methods of the first order prior to a fulfilment of the condition

$$J(\overline{p}) \le \delta_f \quad (8)$$

where $\delta_f = \sum_{m=1}^{M} \int_{\tau_{\min}}^{\tau_{\max}} \sigma_m(\tau) d\tau$ is an integral error of temperature measurements $f_m(\tau)$, $m=1,M$, and σ_m - measurement variance. The Conjigate Gradient Algorithm is presented in details in [6].

EXPERIMENTAL STUDY

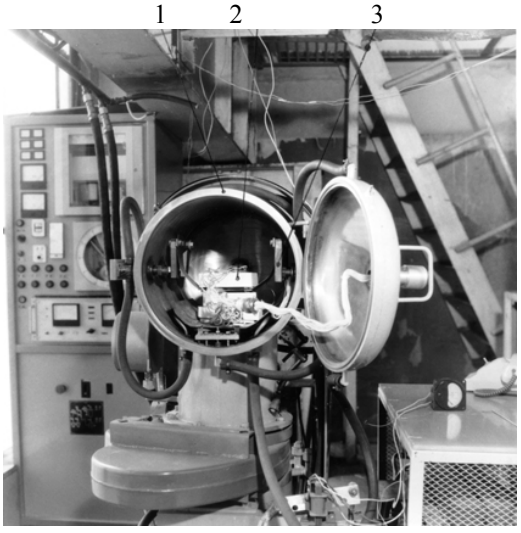


Figure 1. Test Stand: 1 - vacuum chamber, 2 - experimental module, 3 - heater.

By the way of experimental study of the material specimens the results of experiments are used on the automatic thermo-vacuum stand (TVS-1) manufactured in MAI at the Department of Space System Engineering. A TVS-1 is meant to carry out the experimental investigations on specimens of TPS materials in the condition of non-stationary radiation heating.

The test stand (Fig. 1) consists of units and systems, described below. A horizontally set vacuum chamber 1 of 0.1 m^3 volume is a cylinder with double walls between which the cooling water is circulating. A cylinder is sealed at both ends by spherical covers, the rear cover is set fixed, the front cover is hinged and provided with the quick-opening locks. Both covers are water-cooled too. A vacuuming system consists of a mechanic vacuum pump, a diffusion pump, a vacuum seal and valves. A power supply system includes a control desk, a thyristor voltage regulator, a control unit and a power transformer. A control unit is for voltage control fed to a heater. The control can be in manual mode or in automatic mode from a computer.

Two specimens made of the same test material (Fig. 2) were located symmetrically about the heating element on a thermoinsulating base made of thermal insulation material so that its heating surface is parallel to the heater and at a certain distance from it (δ = 4-5 mm). The other surfaces of specimens were heat-insulated by a layer of heat-insulating material. A heating element from a tantalum foil with dimensions 80x70x0.1 mm allowed: 1) to increase the specimen heating rate till the desired values; 2) to avoid destruction of heating elements till the completion of specimen tests, this having occurred sometimes in attempting to provide a corresponding heating rate. Control of the specimen heating condition is performed by temperature on the heated surface measured by a thermocouple, made from a thermocouple wire BP_5 - BP_{20} of 0.1 mm diameter in accordance with a prescribed regime. A control system of the heating regime includes a thermocouple, set up on the specimen surface, a control-point setting device, operating jointly with a self-balancing electronic potentiometer, an analogy regulator device in the set with the control units. Control over the heating regime was maintained in the experiment by the results of three trial starts with specimens from test materials, the structure of which is similar to the structure used in tests. In making the trial starts a criterion in choosing the suitable heating regime is the coincidence of the specimen external surface temperature measured with a prescribed

temperature. Measurement and recording of non-steady temperatures in the test specimen are made by means of an automatic system for experimental information gathering and processing based on PC.

This module is implemented to provide a sufficiently exact definition of the heat flow fed inside the specimens of materials, this in turn, permits to provide the conditions of uniqueness of the inverse problem solving by simultaneous definition of heat conduction and material's heat capacity per unit volume.

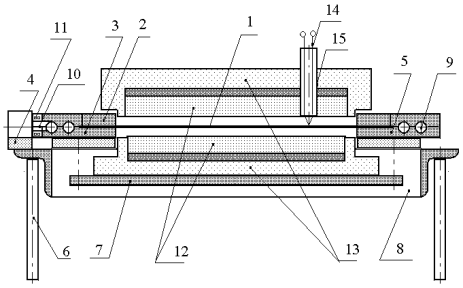


Figure 2. Experimental module: 1 - heater, 2 - pressed plates of current tap, 3 - mobile current taps, 4 - insulator, 5 - immobile current taps, 6 - struts, 7 - plate, 8 - frame, 9 - cooling tubes, 10 - guides, 11- springs, 12 - test specimen, 13 - insulating basement, 14 - control thermocouple, 15 - ceramic tube.

In the process of non-steady heating of specimens by means of an automatic system, recording of temperatures inside the specimen in places of thermocouple positioning, heater's temperature and also electric power released on it were performed:

$$Q_{electr} = U * I \qquad (9)$$

where U - r.m.s. voltage on the heater, I - r.m.s. strength of current, transmitted through the heating element. The heat flux supplied to a specimen due to symmetry is determined as

$$q_2(\tau) = Q_{electr}/(2A) = U*I/(2A) \qquad (10)$$

where A - heater's surface area.

Below there are the results of data processing for experimental study of a high-porous composite thermo-insulating material. The models of test material are the square slabs of dimensions 50x50x5 mm. Thermoelectrode wires of thermocouples were brought out via the fabric-based laminate and placed in glass tubes. The length of the thermocouple isothermal zone was 7-8 mm. The installation of thermosensors in specimens was determined based on the solution of the optimal experiment design problem [2].

OPTIMAL EXPERIMENT DESIGN

In the given case, the search problem of optimal plan or optimal layout of temperature measurements can be formulated as the following extreme problem:

$$\xi^* = Arg \max_{\xi \in \Xi} \Psi(\xi), \quad \xi = \left\{ M, \bar{x} \right\}, \quad \bar{x} = \{x_m\}_1^M,$$

$$X_0 \le x_m \le X_1, \quad m = \overline{1,M} \qquad (11)$$

where ε – possible layouts of measurements out of the set Ξ in the following manner:

$$\Xi = \left\{ (M, \bar{x}) : \quad M \ge M_{\min}; \right.$$

$$\left. x_m \in [X_0, X_1], \ m = \overline{1,M} \ \right\} \qquad (12)$$

where M_{min} = 2 is the minimum number of measurement points necessary for uniqueness of the solution of the analyzed inverse problem [8]; $\Psi(\xi) = \det[F(\xi)]$ is the experiment performance quality characterizing the accuracy of the inverse problem solution. Here $F(\varepsilon)$ is Fischer's information matrix, which is defined by

$$F(\xi) = \frac{1}{M} \left\{ \Phi_{k,j}; k, j = \overline{1,N_p} \right\}; \qquad (13)$$

$$\Phi_{k,j} = \sum_{m=1}^{M} \int_{\tau_{\min}}^{\tau_{\max}} \nu_m(\tau) \Theta_k(x_m, \tau) \Theta_j(x_m, \tau) d\tau \ ;$$

where $\Theta_k(x\ ,\tau) \equiv \partial T(x,\tau)/\partial p_k, \quad k = \overline{1,N_p}$ is the sensitivity function; $\nu_m(\tau), m = \overline{1,M}$ is the functions of measurement efficiency enabling us to consider the errors of experimental data. For equally accurate measurements $\nu_m(\tau) = \sigma^{-2}$, where σ^{-2} is the measurement deviation. Matrix $F(\xi)$ characterizes the total sensitivity of the system in the measurement points $x_m, \ m = \overline{1,M}$, to small variations of the whole set of parameters $p_k, k = \overline{1,N_p}$. It is required to find such an experiment design ξ^* for which the criterion $\Psi(\xi)$ attains the maximum value. Here the domain of admissible values of thermosensors positions is determined by the Eq. (12). The solution of the given problem is constructed on the realization of iterative procedure in which at every iteration we solve the search problem for an

optimal vector of coordinates of the fixed number of thermosensors *M:*

$$\bar{x}^* = Arg \max \det F(M, \bar{x}), \quad X_0 \le x_m \le X_1, \quad m = \overline{1,M} \tag{14}$$

and sequentially we increase by a unit the number of thermosensors. The iterative process ends on condition that the following relation is fulfilled:

$$\left|\left(\Psi\left(l+1, \bar{x}^*\right) - \Psi\left(l, \bar{x}^*\right)\right)\right| / \left|\Psi\left(l+1, \bar{x}^*\right)\right| \le \rho \tag{15}$$

where l – the iteration number; $\rho>0$ – the given value.

The elements of matrix $F(\varepsilon)$ and the unknown layout of temperature measurements ε^* depend on the unknown parameters $p_k, k = 1,...,N_P$. This is due to the fact that the temperature field $T(x, \tau)$, $X_0 \le x \le X_1$, $\tau_{min} \le \tau \le \tau_{max}$, which is determined from the solution of the boundary-value problem (1) – (4) is nonlinearly dependent of the vector of unknown parameters. In these cases we can construct only the approximate locally optimal plans satisfying the solution of problem (11) using the a-priori information about the values of vector $\bar{p}$. In this connection when solving the problem of planning it is of reason to estimate the effect of uncertainties in the values of thermophysical characteristics used in problem solving on the results of planning.

The practice of carrying the experiment studies shows that it is not always the case when in the experiment we manage to accurately provide the given mode of heating of the specimen. The actual coordinates of the points of thermosensors installation can deviate from the optimal points because of errors admitted at installation. Besides, on the coordinates of installation we can impose the constraints conditioned by structural and technological requirements to the sensors and specimens, in which we make measurements. That is why along with the selection of optimal measuring circuits it is necessary to make an analysis of the effect of uncertainties in the specimens heating modes and deviations from optimal coordinates of thermosensors installation on the sensitivity of the heat system under study. Such analysis allows to define some domains where we can install the thermosensors and estimate the losses in the accuracy of solution of the inverse problem. For this purpose it is convenient to consider $\Psi\{F(M, \bar{x})\}$ as a function of the thermosensors positions. The analysis of the experiment design sensitivity to possible deviations of the measurement values should be considered as a part of the considered problem.

The problem of experiment design was solved by the following values of the initial data: the specimen thickness b=5 mm; the process duration $\tau_{\max}$ = 25 seconds; the initial temperature in the specimen $T_0(x)$ = 293 K; the boundary condition on the specimen internal surface is $q_1(\tau)=0$; the boundary condition on the specimen heated surface $q_2(\tau)$ was assumed corresponding to that on Fig. 5 and later on was considered as the rated conditions of heating; the a-priori information about the thermal properties of the materials are shown on Fig. 6; the number of parameters in the spline-approximation of unknown functions N_1 = 4 for $\lambda(T)$ and N_2 = 4 for $C(T)$; the dispersion of equally accurate measurements is $\sigma^{-2} = 1$;

With the purpose to estimate the effect of uncertainty in the specimens heating mode on the measuring circuit the problem of experiment design was also solved for another mode of heating $q_2^-(\tau) = 0.8(q_2(\tau))$. The effect on the experiment design of uncertainty in thermal conductivity $\lambda(T)$ was estimated by the results of solving the design problem with the assignment of values $\lambda^-(T) = 0.8(\lambda(T))$ and $\lambda^+(T) = 1.2(\lambda(T))$.

The coordinates of thermosensors optimal location and the corresponding maximum values $(\Psi(\bar{x}))_{\max}$ obtained as a result of calculations are shown in Table 1. Fig. 3 and 4 illustrate dependencies of the criterion value $\overline{\Psi}(\bar{x}) = \Psi(\bar{x}) / (\Psi(\bar{x}))_{\max}$ on the location of thermosensors for M = 2 и M = 3, correspondingly. It should be noted here that the coordinates and numbers of measurement points are counted off the insulated surface of the slab.

The presented results show that at the analyzed conditions of heating for experiment design with M = 2 и M = 3 one thermosensor should be installed on the heated surface and the rest – in the depth of the specimen. Actually, the optimal coordinates of thermosensors location x_1^* and x_2^* coincide. This allows to assume that for simultaneous determination of $\lambda(T)$ and $C(T)$ we can use the experiment design with M = 2. However, the comparison of values $(\Psi(\bar{x}))_{\max}$ for

M = 3 и M = 2 shows that in the second case $(\Psi(\bar{x}))_{\max}$ is less almost by four orders of magnitude, that demonstrating a more lower sensitivity of the heat system at M = 2 as compared with M = 3. The decrease of values of the heat flux density by 20% results in some lesser value of $(\Psi(\bar{x}))_{\max}$ and in the displacement of sensor location x_2^* onto the specimen heated surface. The coordinate of point x_1^* is not changed.

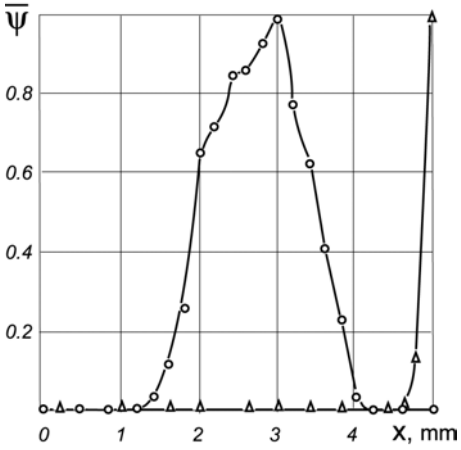


Figure 3. Dependences of the criterion value $\overline{\Psi}(\bar{x})$ on the location of thermosensors for M = 2:

○ - $\overline{\Psi}(x_1), 0 \le x_1 \le 5\ mm, x_2 = x_2^*;$

Δ - $\overline{\Psi}(x_2),\ x_1 = x_1^*, 0 \le x_2 \le 5\ mm;$

The increase of values $\lambda(T)$ by 20% results in the displacement of coordinate of the measurement point x_2^* onto the heated surface and in some displacement of point x_1^* inside the specimen. At the increase of values $\lambda(T)$ by 20% the optimal coordinates of measurement points x_1^* and x_2^* coincide, while the of $(\Psi(\bar{x}))_{\max}$ are by two orders of magnitude greater. In both cases with variations of $\lambda(T)$ the optimal coordinate of measurement point x_3^* is not changed. As a result of the analysis carried out for simultaneous estimating of $\lambda(T)$ and $C(T)$ the design of temperature measurements with the coordinates X_o = 0 mm (for a boundary condition of the first kind the readings of the sensor on the internal surface were used), x_1^* = 2.4 mm, x_2^* = 3.0 mm, x_3^* = 5.0 mm (positioned on the exposed surface) were recommended.

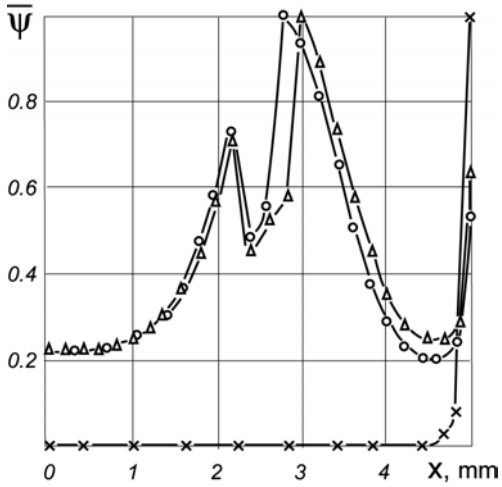


Figure 4. Dependences of the criterion value $\overline{\Psi}(\bar{x})$ on the location of thermosensors for M = 3:

○ - $\overline{\Psi}(x_1), 0 \le x_1 \le 5\ mm, x_2 = x_2^*; x_3 = x_3^*;$

Δ - $\overline{\Psi}(x_1),\ x_1 = x_1^*, 0 \le x_2 \le 5\ mm, x_2 = x_2^*;$

x - $\overline{\Psi}(x_3), x_1 = x_1^*, x_2 = x_2^*, 0 \le x_3 \le 5\ mm,;$

Table 1. The criterion of experiment design $(\Psi(\bar{x}))_{\max}$ and optimal thermosensors location

Mode of Heating	Values of $\lambda(T)$	x_1^*, mm	x_2^*, mm	x_3^*, mm	$(\Psi(\bar{x}))_{\max}$
$q_2(\tau)$	$\lambda(T)$	2,8	3,0	5,0	$0{,}43*10^{-5}$
$0.8(q_2(\tau))$	$\lambda(T)$	2,8	5,0	5,0	$0{,}23*10^{-6}$
$q_2(\tau)$	$1.2(\lambda(T))$	2,4	5,0	5,0	$0{,}63*10^{-6}$
$q_2(\tau)$	$0.8(\lambda(T))$	3,0	3,0	5,0	$0{,}16*10^{-3}$
$q_2(\tau)$	$\lambda(T)$	3,0	5,0	-	$0{,}99*10^{-9}$

RESULT OF EXPERIMENTAL DATA PROCESSING

During the specimen heating on the test facility a theoretically prescribed time dependence of heat flux (Fig. 5) was provided. The number of approximation parameters N_1 and N_2 for desired functions were equal 4. The comparison of experimentally measured and calculated (using thermal characteristics obtained from a solution of the inverse problem) temperature values in points

of thermocouple positioning are shown on Fig. 5 (only for one specimen). A quite well coincidence of them take place here.

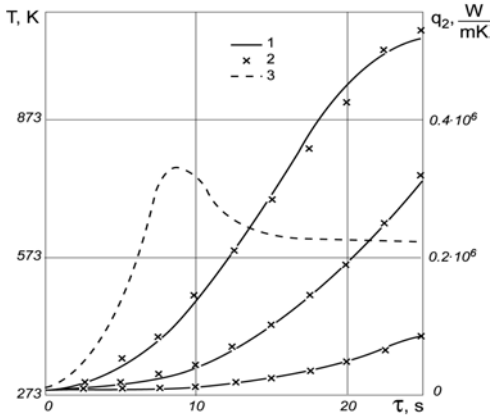


Figure 5. Temperature values in points of thermocouple locations:
1- calculated, 2 – experimental, and heat flux from the heater - 3.

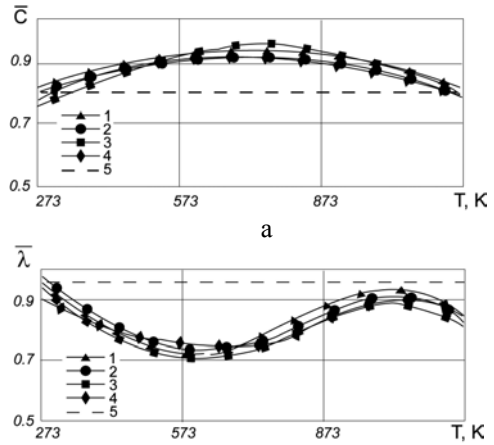


Figure 6. Results of testing:
a - heat capacity, b - conductivity.

The results of the inverse problem solving - the composite material thermophysical characteristics are given on Fig. 6 (results of four experiments- curves 1,…,4). These results were verified by using different (quite distinct from each other) initial values for an iterative process. The results obtained depend rather weakly on the initial approximations.

CONCLUSIONS

For partially decomposed materials the model of heat conduction with temperature-dependent thermal characteristics is approximate, and characteristics are effective, since the heat transfer in such material is provided not only by heat conduction but also by different transformation processes depended on conditions of heating. A deviation of of calculated and experimental temperature values in the experiments did not exceed 8 K, that confirms the possibility of using, for the given material, a model of heat conduction with the effective thermal characteristics. But the presented method can be used used only for determining the effective thermal characteristics of composite materials for partial heating conditions.

ACKNOWLEDGMENTS

This work was done with the support of ISTC and ESTEC/ESA in the frame of ISTC grant 804f.

REFERENCES

1. O.M.Alifanov, Mathematical and experimental simulation in aerospace system verification, *Acta Astronautica,* , **41**, 43 (1997).

2. O.M.Alifanov, E.A.Artyukhin and S.V.Rumyantsev, *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Problems,* Begell House, New York\Wallinford(UK), 1995, p. 306.

3. E.A.Artyukhin, S.A.Budnik and A.S.Okhapkin, Temperature measurement optimization and numerical inverse conduction-treatment solution, *Journal of Engineering Physics*, **55**, 924 (1988).

4. E.A.Artyukhin, G.A.Ivanov and A.V.Nenarokomov, Determination of a complex of materials thermophysical properties through data of nonstationary temperature measurements, *High Temperature*, **31**, 235 (1993).

5. E.A.Artyukhin and A.V. Nenarokomov, Coefficient inverse heat conduction problem, *Journal of Engineering Physics*, **53**, 1085 (1987).

6. K.Dowding, J.V.Beck, B.Blackwell et.al., Estimation of thermal properties and surface heat flux in carbon-carbon composite, *Journal of Thermophysics and Heat Transfer*, **9**, 345 (1995).

7. C.-H.Huang, J-Y.Yan, An inverse problem in simultaniously measuring temperature-dependent thermal conductivity and heat copacity, *Int. Journal Heat and Mass Transfer*, **38**, 3433 (1995).

8. N.V.Muzylev Uniqueness of simultaneous determining of coefficients of thermal conductivity and volumetric heat capacity, *USSR Comput. Math. and Math. Phys.,* **23**, 102 (1983).

HEAT AND MASS TRANSFER

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ESTIMATION OF POWER DISSIPATION IN ELECTRONIC CIRCUITS

Marcin Janicki **Andrzej Napieralski**
Department of Microelectronics & Computer Science
Technical University of Lodz
Al. Politechniki 11, 93-590 Lodz, Poland
e-mail: janicki@dmcs.p.lodz.pl

ABSTRACT

This paper presents an approach to the estimation of hybrid circuit temperature solving an inverse problem. First, based on the infrared temperature measurements, an adequate thermal model of the circuit is created. Then, fully three-dimensional analytical time dependent solution of the model is determined using the Green's functions. Next, in the numerical experiments, temperature values of the active transistors are estimated from remote diode temperature measurements. The quality of the obtained estimates is improved owing to the application of the sequential function specification algorithm. The required for the algorithm sensitivity coefficients, which relate the amount of dissipated power to the temperature rise at sensor locations, are computed from the circuit thermal model. Finally, the estimation results are compared with infrared measurements.

NOMENCLATURE

A –matrix of thermal influence coefficients
J - number of sensors
T - temperature
T_∞ - surrounding fluid temperature
a, b, d – first layer dimensions
c_v – specific heat at constant volume
g_v – generated heat per unit volume
h - heat exchange coefficient
p – number of heat sources
q - heat flux
r - number of future values
t – time
x – position vector
α - thermal diffusivity
β, γ - eigenvalues
ϕ – eigenfunctions
λ – thermal conductivity

INTRODUCTION

Continuous miniaturisation of electronic devices and increase of their operating frequency augmented significantly dissipated power density. For this reason, cooling problems occur now even in apparently low power circuits. Therefore, nowadays, many modern applications require on-line monitoring of their temperature. Obviously, the best solution would be to place temperature sensors directly in the locations where heat is generated. Unfortunately, this is usually impossible to realise in electronic circuits due to some technological limitations. Then, circuit temperature can be computed solving an inverse heat conduction problem (IHCP). Since, in most cases, heat is generated close to structure surface, the problem of determining dissipated power density might be reduced to the estimation of unknown surface heat flux.

The thermal response in a structure is both delayed and damped, hence inverse problems are extremely sensitive to measurement errors and special techniques must be applied to obtain robust estimates of unknown quantities. There are many possible ways to deal with such ill-conditioned IHCPs, but none of them is univeral and each method must be suited to a particular application. The method used in this publication is the sequential form of the function specification method.

The theoretical part of this paper consists of two sections covering briefly the methodology of solving the heat equation in non-homogenous solids and the principles of the sequential function specification algorithm. Then, the presented mathematical apparatus is applied in the experimental part for the solution of the circuit thermal model and the temperature estimation of the heat sources (power transistors) from the temperature measurements of the diodes placed at certain distance from the heat sources.

The transient analytical solution of the heat equation resulting from the circuit thermal model is found using the Green's functions (GFs). Next, the obtained solution is used to compute the relevant sensitivity coefficients required for the function specification algorithm. Then, the power dissipated in the heating transistors is estimated based on the temperature measurements from remote diode locations. Once the power is known, the temperature of the transistors is computed solving a direct problem. The simulation results are compared with infrared transient thermal measurements.

THERMAL ANALYSIS METHOD

This section presents briefly the methodology of obtaining analytically solutions of time-dependent heat conduction problems in non-homogenous materials where the thermophysical properties (e.g. thermal conductivity, diffusivity or specific heat) depend on space co-ordinates. The presented approach cannot be applied if the properties vary with temperature.

For heat diffusion processes obeying the Fourier law of conduction, which states that the heat flux is proportional to the temperature gradient, the energy balance for a unitary volume leads to the following three-dimensional heat equation [1]-[4]:

$$\nabla\left[\lambda(x)\nabla T\right] + g_v(x,t) = c_v(x)\frac{\partial T}{\partial t} \qquad (1)$$

Equation 1 can be solved given material thermophysical properties as well as initial and boundary conditions. The initial conditions determine temperature distribution in a structure at the starting time of the analysis. The boundary conditions describe the heat exchange with the surrounding environment at the outer limits of the structure. Heat generated in electronic circuits is conducted towards structure edges where it is removed by means of convection or radiation. Therefore, the heat removal at circuit surface is represented the best by the so-called third kind of boundary condition. Based on the Fourier's and Newton's laws, this boundary condition can be expressed by the following equation:

$$q = -\lambda\frac{\partial T}{\partial n} = h(T - T_\infty) \qquad (2)$$

The solution of Equation 1 is assumed to have the following form:

$$T(x,t) = \sum_n A_n(t)\,\varphi_n(x)\exp(-\gamma_n t) \qquad (3)$$

Assuming temporarily that the coefficients A_n are time independent, the solution is substituted into the complementary equation which is basically the same as Equation 1 but without internal heat generation and the equation has homogenous boundary conditions of the same type. This leads to an eigenvalue problem, in which ϕ_n and γ_n are called eigenfunctions and eigenvalues respectively. In general case, each eigenfunction ϕ_n is a combination of certain linearly independent basis functions satisfying the homogenous boundary conditions. When determining exact solutions, which is true in the case considered here, the basis functions are orthogonal and identical to the eigenfunctions. Then, the eigenvalue problem can be solved using a Galerkin-like procedure multiplying both sides of the equation resulting from the substitution by appropriate eigenfunction and integrating over the whole structure volume. The entire procedure leads to the following formula to compute the eigenvalues [5]:

$$\gamma_n = \frac{-\int_v \varphi_n\,\nabla\left(\lambda(x)\nabla\varphi_n\right)dV}{\int_v c_v(x)\varphi_n^2\,dV} \qquad (4)$$

When the eigenfunctions and the eigenvalues are computed, the same procedure with the substitution of the solution and the integration over volume is repeated for Equation 1. This time however it is assumed that the coefficients A_n in the solution are time dependent. When the initial temperature of the whole structure is equal to 0 and the generated heat density g_v is equal to $c_v(r)\delta(r-r')\delta(t-\tau)$, where δ is the Dirac delta function, a Green's function can be found using the following formula:

$$G(r,t\,|\,r',\tau) = c_v(r)\sum_n \frac{\varphi_n(r)\,\varphi_n(r')}{\int_v c_v(r)\varphi_n^2\,dV}\exp\left(-\gamma_n\,(t-\tau)\right) \qquad (5)$$

There exist two possible interpretations of Green's functions. First, they can be regarded as a transient temperature response at a point r in time t caused by instantaneous heat generation at a point r' in time τ. Thus, in order to obtain the temperature response in time, it is required to integrate a GF over all volume and time at which heat is generated. Alternatively, a GF is the transient temperature response at a point r in time t due to the initial temperature rise at a point r'. Then, the temperature distribution in time is computed as an integral of a GF evaluated at time 0 over the whole analysis domain. More information on the GFs and methods for deriving them can be found in [5]-[7].

The particular hybrid circuit considered in this publication has rectangular shapes and can be approximated, as many other electronic circuits, by a multilayered parallelepiped. Since power is dissipated only in a small region close to the structure top surface, this phenomenon is represented in the model by the heat flux penetrating into the structure. Furthermore, taking into account that circuits are relatively thin in comparison to their area, it is often assumed that the four lateral surfaces of the structure are adiabatic and the heat is removed only at the two remaining surfaces, which is modelled by the heat exchange coefficient. Such a circuit thermal model is visualised in Figure 1.

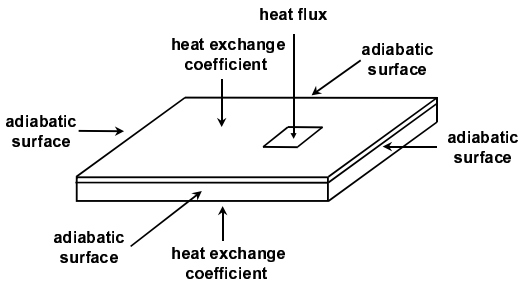


Figure 1: Circuit thermal model

For the above circuit thermal model consisting of two layers, its time dependent solution can be found in accordance with the earlier presented outline of the method. First, the three-dimensional eigenfunctions ϕ_n, shown in Equation 6, were found applying the specific boundary conditions consistent with the thermal model, i.e. prescribed heat flux in x and y horizontal directions and convective condition in the vertical z direction.

$$\varphi_{m,n,k}(x,y,z) = C_{ik}\cos\left(\frac{m\pi}{a}x\right) * \cos\left(\frac{n\pi}{b}y\right) * \left(\sin\left(\frac{\beta_k}{d}z\right) + \Gamma_{ik}\cos\left(\frac{\beta_k}{d}z\right)\right) \tag{6}$$

The values of C, Γ and β are chosen so that all the boundary conditions and the perfect material contact conditions are satisfied. In particular, the eigenvalues in the vertical direction β are the positive solutions of Equation 7. The next step in the solution was to compute the eigenvalues γ using the formula from Equation 4. The details of the solution method can be found in [8].

$$\lambda_2\left(tg\,\beta_k + \Gamma_{1k}\right)\left(1 - \Gamma_{2k}\,tg\,\beta_k\right) = \lambda_1\left(1 - \Gamma_{1k}\,tg\,\beta_k\right)\left(tg\,\beta_k + \Gamma_{2k}\right) \tag{7}$$

When the temperature rise does not result from the initial temperature distribution and there is no internal energy generation, which is the case considered here, the only one non-homogenous boundary condition is the heat flux diffusing into the structure through its top surface. Then, the formula to compute temperature rise in the circuit can be expressed in terms of Green's function as:

$$c_v(r)T(r,t) = \int_{\tau=0}^{t}\int_S (q\,G)_{S'}\,dS'\,d\tau \tag{8}$$

FUNCTION SPECIFICATION ALGORITHM

For linear problems, if the heat fluxes for each surface are constant, the direct heat conduction problem discussed in the previous section can be re-written in the matrix form as shown in Equation 9. Then, the structure temperature vector T is determined multiplying the vector of surface heat fluxes q and the matrix of thermal influence coefficients A. The elements of the matrix are computed evaluating the appropriate integrals according to Equation 9.

$$T = A * q \tag{9}$$

The particular inverse problem in this case would consist in determining the unknown heat fluxes from temperature measurements in selected locations. From mathematical point of view, the solution of the inverse problem requires performing the inversion of the thermal influence

coefficient matrix A. However, due to the fact that the transient thermal response in the structure is damped and delayed, the matrix is usually poorly conditioned thus, for numerical reasons, it is virtually impossible to invert it. One of the methods of improving problem conditioning is the function specification method.

The function specification method is based on the assumption that the variation with time of the unknown quantity to be estimated can be described in the functional form. The most obvious and the most commonly used assumption is that the function is constant, which implies here that a few subsequent heat flux samples are equal. The method can be further expanded allowing the introduction of additional temperature sensors, i.e. more than the number of heat sources. The method can be also adapted for simultaneous multiple heat flux estimation. In the sequential estimation the values of heat fluxes in time are found successively with each new arriving data. The surface heat fluxes in all iterations are computed knowing current sensor temperature values and *r-1* "future" temperature values, where *r* is the number of heat fluxes assumed to be equal. Then, for *p* heat fluxes and *J* temperature sensors, the unknown heat flux vector in the *k*-th iteration can be found using the following equation [9]-[10]:

$$q_k = \left((XB)^T XB\right)^{-1} (XB)^T T \qquad (10)$$

where:

$$X = \begin{bmatrix} a_1 & 0 & 0 & \vdots & 0 \\ a_2 & a_1 & 0 & \vdots & 0 \\ a_3 & a_2 & a_1 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_r & a_{r-1} & a_{r-2} & \vdots & a_1 \end{bmatrix} \quad a = \begin{bmatrix} a_{11} & \vdots & a_{1p} \\ \cdots & \cdots & \cdots \\ a_{J1} & \vdots & a_{Jp} \end{bmatrix}$$

$$B = \begin{bmatrix} I_1 \\ \cdots \\ \cdots \\ I_r \end{bmatrix} \quad q_k = \begin{bmatrix} q_{1k} \\ \cdots \\ \cdots \\ q_{pk} \end{bmatrix} \quad T = \begin{bmatrix} T_1 \\ \cdots \\ \cdots \\ T_r \end{bmatrix}$$

I - [p x p] identity matrices

TEMPERATURE ESTIMATION

The function specification algorithm presented in the previous section was applied for the estimation of power dissipated in a hybrid power module. The module is a three-phase AC motor driver manufactured in the Insulated Metal Substrate hybrid technology. The circuit, shown in Figure 2, consists of IGBT power transistors T1-T7, rectifying diodes D1-D6 and protection diodes DT1-DT7. All the silicon devices together with molybdenum heat spreaders are mounted on a large 50 mm x 75 mm x 1.5 mm aluminium plate. The devices considered in the estimation procedure are encircled in Figure 2.

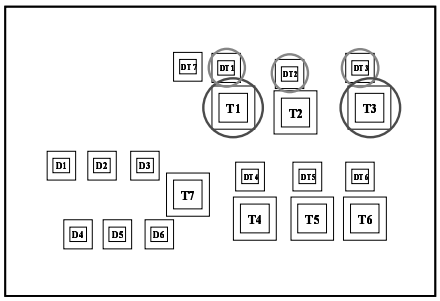


Figure 2: IGBT module layout

For the measurements the circuit was placed horizontally on a large block of aluminium. The circuit transient temperature measurements were taken using the AGEMA Thermovision 900 infrared camera. Before the measurements, in order to assure uniform surface emissivity, the circuit was sprayed with black matt paint. The ambient temperature was equal to 24 °C.

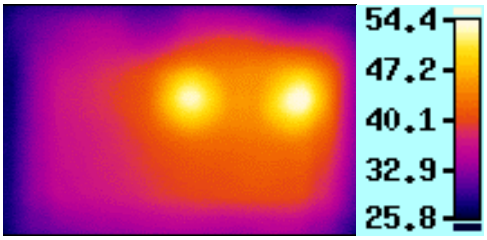


Figure 3: Infrared picture

Initially, the power of 8.4 W was dissipated only in the transistor T3. Then, after 63 s the transistor T1 was also switched on. Finally, after 187 s the transistor T3 was switched off and the power was dissipated only in the transistor T1. An exemplary infrared picture is shown in Figure 3. The particular temperature curves obtained for the hot spots (transistors T1 and T3) and the temperature sensor locations (diodes DT1, DT2 and DT3) are printed in Figure 4.

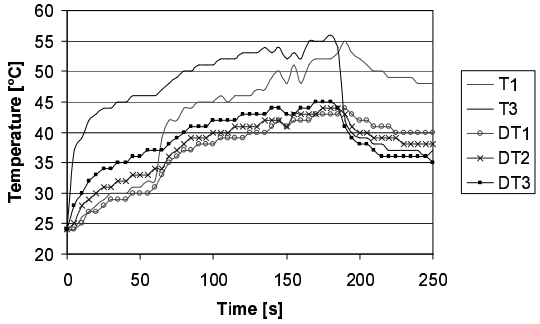


Figure 4 Measured temperature curves

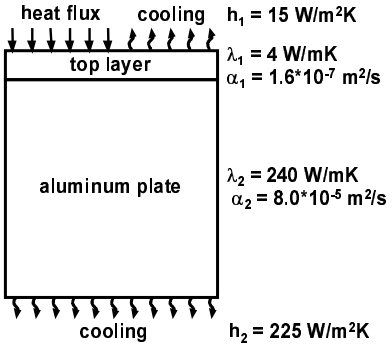


Figure 5 Circuit thermal model

Originally, a full five-layer three-dimensional thermal model of the circuit was considered. The model was solved using the numerical Finite Difference Method. The preliminary simulations showed that the temperature drop in the top layers (silicon dies, molybdenum heat spreaders and copper frame) is not significant, thus the model was eventually reduced to only two layers: the top layer representing the other layers and the aluminium base plate.

This simplified thermal model of the circuit was employed for the numerical simulations. The model, shown in Figure 5, permits the application of both the GF heat equation solution method and the function specification algorithm. During the simulations, the bottom layer was assumed to be made of pure aluminium. The thermal properties of the top layer and the value of the heat exchange coefficient were set experimentally minimising the relative error with respect to the infrared measurements. The thermal conductivity and the heat exchange coefficient were determined from the steady state asymptotic temperature values. The thermal diffusivity value, being the ratio of the thermal conductivity to the thermal capacity, was found from the transient temperature curve so as to get the correct temperature rise rate.

The finally accepted values of the thermal parameters for the top layer might seem to be underestimated, however it should be kept in mind that this layer represents in the model many other layers, thus the values have no direct physical interpretation. Moreover, the base plate might not be necessarily made of pure aluminium and consequently the real values of the parameters for this layer might be significantly different. More detailed information about the circuit itself and its thermal model can be found in [11]-[12].

Next, given the circuit thermal model, the required thermal influence coefficients were computed and placed in the appropriate locations in the matrices presented in Equation 10. Then, the diode temperature values were supplied to the function specification algorithm. The sequential estimation of transistor temperature values was conducted as a single heat flux estimation when only one transistor was dissipating power and as a multiple heat flux estimation when both transistors were active. Thus, for the algorithm, the information on the transistor switching times was necessary. The estimation results obtained for 1 and 3 future values are presented in Figures 6-7.

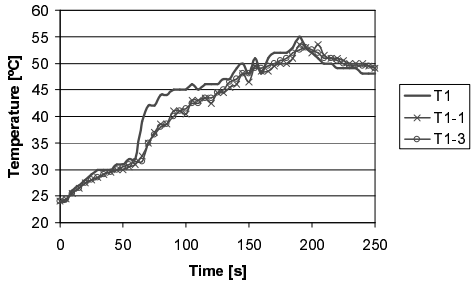


Figure 6: Transistor T1 temperature estimation

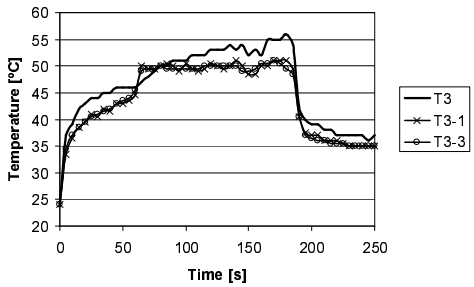


Figure 7: Transistor T3 temperature estimation

As can be seen from the charts, the estimates follow quite closely the real transistor temperature values. The temperature differences between the estimates and the measurements do not exceed 5 K. The obtained results are visibly worse in the case when both heat fluxes are estimated. Then, the heat flux associated with the transistor T3 is overestimated and the other heat flux is underestimated.

The increase of the number of future values taken into account removes fluctuations in the estimates, but on the other hand, in the case of oscillations present in the real heat flux, it causes the estimates to be erroneous for longer time (see the peaks after 150 s). The incorporation of future values improves also the problem conditioning since the condition number value drops from 2.9 to 2.6.

CONCLUSIONS

The results of the research presented in this paper are very promising. Although the proposed circuit thermal model is extremely simple, the obtained estimation accuracy of 10 % is fairly acceptable for thermal simulations in electronic applications.

The diode temperature values introduced into the algorithm were quite inaccurate (the accuracy of infrared measurements is 3K), thus it should be expected that in the case of real diode measurements the estimation accuracy would improve. Obviously, if more of the information contained in the infrared image is used, the results will be much better, but the main objective here was to verify the possibility of temperature estimation knowing temperature values only in the limited number of remote locations.

Another issue, which should be commented on, is the question of the uniformity of the heat transfer coefficient value. Obviously, its local values are different but the solution method does not allow the coefficient value to vary across a boundary surface, so some average value of the coefficient should be assumed in the simulations.

However, the change of the coefficient value in time could be taken into account by updating its value with each time step. This should significantly improve the estimation accuracy, especially when the coefficient value changes rapidly. For example, in the beginning of the heating process, the value of the coefficient is less than the half of the one in the steady state, thus in reality the circuit heats up faster than it is shown in the simulations.

The simulations have proved that the function specification algorithm could be an efficient instrument for real-time electronic circuit temperature estimation. Moreover, the fact that the algorithm could be synthesised as a digital filter and fully integrated with the monitored circuit increases its attractiveness for electronic applications.

ACKNOWLEDGEMENTS

This work is partly supported by the grant of Polish State Committee for Scientific Research No. 8T11B 021 19 and partly by the Internal University Grant K-25/1/2002/Dz.St.

REFERENCES

1. Y. Bayazitoglu, M.N. Ozisik, Elements of heat transfer, McGraw-Hill, 1988
2. M.N. Ozisik, Heat Conduction, John Wiley & Sons Inc., 1993
3. J.P. Holman, Heat transfer, McGraw-Hill, 1985
4. H.S. Carslaw, J.S. Jaeger, Conduction of Heat in Solids, Clarendon Press, Oxford, 1947
5. J.V. Beck, K.D. Cole, A. Haji-Sheikh, B. Litkouhi, Heat Conduction Using Green's Functions, Hemisphere Publishing, 1992
6. M.D. Greenberg, Application of Green's Functions in Science and Engineering, Prentice-Hall, 1971
7. I. Stakgold, Green's Functions and Boundary Value Problems, Willey-Interscience, 1979
8. M. Janicki, G. De Mey, A. Napieralski, Transient Thermal Analysis of Multilayered Structures Using Green's Functions, accepted for publication in Microelectronics Reliability
9. J. V. Beck, B. Blackwell, and C. R. St. Clair, Inverse Heat Conduction - Ill-posed Problems, John Wiley & Sons, 1985
10. M. Janicki, M. Zubert, and A. Napieralski, Application of inverse heat conduction problem solution with error correction to estimation of IC temperature, Proceeding of The 4th THERMINIC International Workshop, Cannes, France, pp. 191-196, 1998
11. M. Janicki, A. Napieralski, IGBT Module Thermal Simulation, Proceedings of 9th European Conference on Power Electronics and Applications EPE 2001, Graz, Austria, 27-29 August 2001, DS2.2-2
12. M. Janicki, P. Kawka, G. De Mey, A. Napieralski, IGBT Hybrid Module Thermal Measurements and Simulations, Proceedings of 8th International Conference Mixdes 2001, Zakopane, Poland, 21-23 June 2001, pp. 249-252

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

SIMULATION OF HEAT TRANSFER PROCESSES AND OPTIMIZATION OF COMPONENTS IN MOBILE CLIMATE CONTROL SYSTEM BY MEANS OF THERMAL PARAMETER IDENTIFICATION

Alexander V. Moultanovsky
Development & Validation Department
Bergstrom Climate Systems, Inc.
Rockford, Illinois, USA
amoult@bergstrominc.com

ABSTRACT

Intention of this presentation is to combine and sum up our research on inverse heat transfer approach to mobile climate control systems simulation of thermal processes and predicting heat transfer, to compact heat exchangers inverse material design, preliminary design optimization (shape & dimensions) and capacity calculation. A common approach to the solution of different types of problems such as optimization problem and identification of thermal and geometrical parameters requires common method that allows to utilize the Inverse Heat Transfer Problem (IHTP) tactic to solve all above-mentioned objectives. The latter is possible due to a direct connection between oprimization and parameter identification problems. An identification approach to the solution of both problems makes possible by means of the well-known method of Adaptive Iterative Filter (AIF). This method was adapted here for the solving numerous inverse problems that appeared in the area of mobile Heating, Ventilating, and Air-Conditioning (HVAC) systems.

NOMENCLATURE

T	temperature
τ, t	time
k	thermal conductivity of the heat exchanger material
c_p	specific heat at constant pressure
q	heat flux between outer evaporator surface and air
G	total heat transfer
$\vec{X}$	stochastic state vector
$\vec{U}_k$	stochastic vector of controls
$\hat{\vec{Z}}$	estimated vector of unknown parameters
$\hat{\vec{Z}}^{(j)}_{k+1/k+1}$	unbiased, with minimum dispersion, estimate obtained for the vector of the parameters which are defined at the *j*-th iteration of the *k+1*-st time step on the basis of the vector of measurements $\tilde{\vec{Y}}_{k+1}$
$\tilde{\vec{Y}}$	vector of measurements
$\hat{\vec{Y}}^{j}_{k}$	estimated vector of "measurements" obtained at the *j*-th iteration of *k*- th time step
P	covariance matrix of estimate errors
R	covariance matrix of measurement errors
K	weight matrix
W, V	uncorrelated white noise
H	measurement, sensitivity matrix
σ	standard deviation of measurement errors or mean square error of measurements
h[*]	operator of relations between estimated and measured parameters

Subscripts and Superscripts

i, j	iteration numbers
k	number of time steps
$\rightarrow$	vector sign
$\wedge$	estimation sign

Symbols

IHTP	Inverse Heat Transfer Problem
AIF	Adaptive Iterative Filter
HVAC	Heating, Ventivating, Air-Conditioning
TC	Thermocouple
SH, SC	Superheat, Subcooling

INTRODUCTION

The estimation of heat transfer as a whole, identification of specific heat transfer parameters, and optimization of the geometry of device under

investigation, all by means of inverse heat transfer methods, makes it possible to get precise results of heat transfer boundary conditions because one identifies them on the basis of temperature measurements, statistically taking into account the measurement errors. The identification is based on the solution of external, internal, or combined IHTP, where some uniqueness conditions, initially unknown, are identified or the mathematical model of the phenomenon under study is refined by the limited and rather approximate data on the temperature field.

MOBILE HVAC SYSTEM

Mobile HVAC system has three fluid streams: air, coolant, and refrigerant. Each of these streams defines heat transfer between its corresponding heat exchanger and fluid. The automotive climate control system includes three main heat exchangers. They are the evaporator, condenser, and heater. An Air Conditioning thermal system consists of the evaporator and condenser with two heat transfer fluid streams, refrigerant and air. A heating thermal system consists of the heater with two heat transfer streams, as well, coolant and air. Each of these two thermal systems affects the human body comfort level in the summer or wintertime, correspondingly. Heat transfer process in the thermal system is defined as exchange of heat between the heat receiver ("heat sink"), which is refrigerant in AC thermal system and air in Heat thermal system, and heat supplier ("heat source"), which is air in AC system and coolant in Heat system.

An evaporator acts as a device to lower the temperature of ambient or recirculated air that passes through it. Physics of the process is explained in the following manner: the low temperature and low pressure liquid refrigerant enters the liquid line of an evaporator, where it is boiled (evaporated) by the air coming through. The evaporator must provide stable refrigerant flow under every possible operating condition, and should have sufficient capacity for rapidly lowering the vehicle cabin temperature.

The necessary conditions for the evaporation process that follow immediately from the refrigerant cycle (or reversed Carnot cycle) are constant pressure and temperature of the refrigerant on the evaporation line of the pressure-enthalpy diagram. The next very important consideration that greatly affects the evaporators design is that the refrigerant circuiting must be designed to make the outer surface temperature as uniform as possible.

The heating process is less complicated than evaporation. The high temperature liquid (usually coolant from engine cooling system) enter the heater core, where it is cooled by the air coming through. The heater must provide stable coolant flow under every possible operating condition. The heater core must be large enough to ensure that the widnshield is defrosted and cab area is defogged within the requirements and conditions defined by the Society of Automotive Engineering (SAE). In addition to acceptable defrost, demist, and defog performance, the heater must make driver and passengers comfortable in relatively short period of time after a cold start-up at temperatures as low as -30°C and even much lower in some cold areas of the US or world. The above-mentioned consideration about uniformity of the surface temperature that greatly affects the evaporator design has influences on the selection of heater core that also must be designed to make the outer surface temperature as even as possible.

One more point that should be mentioned is the packaging constraints. Although, it seems these constraints are not so important; for the automotive application the packaging constraints end up to be the most decisive factors in the design of the heat exchangers. Since the evaporator and heater are a part of HVAC module, they should be packaged as a part of the whole unit, where space is very limited, and, thereby, the selection of the heat exchanger type, in great part, depends upon this space availability. The same note has to be made in regard to the frontal area of the device. The frontal area should be separated from the whole apparatus because, on the one hand, the surface area is a very great factor in influencing the evaporator and heater capacity. That is, the greater this area, the better the heat exchanger performance. On the other hand, the frontal area of the heat exchanger is limited, again, by the space available for the device.

Currently all heat transfer calculations and rough evaluation of heat fluxes (or heat transfer coefficients) for the purpose of HVAC system design or performance evaluation are based on simplified heat (energy) balance equations. It is obviously that such rough estimation of the most important thermal parameters for heat exchangers performance evaluation, thermal system simulation, and subsequent climate control

module design is completely unacceptable. As a result of such inaccuracies, the HVAC heat exchanger's geometrical parameters, material selection or required capacity calculation, which are performed on the basis of those heat transfer calculations, would be obtained roughly as well.

With all of the preceding, it is no doubt that the solution of the IHTP for the identification of boundary conditions and device material's thermophysical characteristics is the most appropriate approach to get precise estimates of desired thermal parameters.

RECCURENT METHOD FOR PARAMETERS IDENTIFICATION

It results from Bellman principle of optimality and theory of dynamic programming, that, in the problems with discrete time, the process of determining the control strategy (controlling sequence) may be reduced to the recurrent computation of separate members of this sequence $\{\vec{U}_k\} = \{\vec{U}_1, \vec{U}_2, ..., \vec{U}_k\}$. Adopting Bellman principle of optimality, the quality functional that takes into consideration both the model fit and control vector behavior and allows at each time step to search for conditionally-optimum control, can be written in the following form:

$$\hat{I}_k^{(3)}\left[\vec{Z}_k, \tilde{\tilde{Y}}_k, j\right] = E\left\{ \sum_{i=1}^{k} \left[\left\| h\ [\hat{\vec{Z}}_{i/i}^{(j)}] - \tilde{\tilde{Y}}_i \right\|^2_{R_i^{-1}} + \frac{1}{j(i)} \left\| \vec{Z}_i - \hat{\vec{Z}}_{i/i}^{(j)} \right\|^2_{[P_{i-1/i-1}^{(j)}]^{-1}} \right] \right\}. \quad (1)$$

The required optimum control strategy $z_k = \{\ \vec{z}_i\}$ searches the best estimation in the sense of the estimate mean square error minimum.

The minimization of the functional (1) for the identification of the vector $\vec{Z}_k$ supposes that the search of the estimates of the desired parameters is carried out with the aid of the AIF. This method has been chosen as a mathematical and numerical instrument for identification of thermophysical and technological parameters, solving IHTP and optimum control problems, and thermal system simulation. The fundamental numerical algorithm of AIF that allows direct calculation of the vector of parameters being defined on the basis of vector measurements can be written as follow:

$$\hat{\tilde{Z}}_{k+1/k+1}^{(j)} = \hat{\tilde{Z}}_{k+1/k+1}^{(j-1)} + K_{k+1}^{(j)} \left[\tilde{\tilde{Y}}_{k+1} - \hat{H}_{k+1}^{(j)} \hat{\tilde{Z}}_{k+1/k+1}^{(j-1)} \right]; \quad (2)$$

$$K_{k+1}^{(j)} = P_{k/k}[\ \tilde{H}_{k+1}^{(j)}]^T \{\ \tilde{H}_{k+1}^{(j)} P_{k/k} [\ \tilde{H}_{k+1}^{(j)}]^T + R_{k+1}\}^{-1}; \quad (3)$$

$$P_{k/k} = [\ I - K_k^{(i)} \tilde{H}_k^{(i)}]\ \ P_{k-1/k-1}[\ I - K_k^{(i)} \tilde{H}_k^{(i)}]^T + K_k^{(i)} R_{k-1} [\ K_k^{(i)}]^T; \quad (4)$$

$$\hat{H}_{k+1}^{(j)} = \left\{ \frac{\partial \vec{Y}}{\partial \vec{Z}} \right\}_{k+1}^{(j)}. \quad (5)$$

Here, $\hat{H}_{k+1}^{(j)}$ is the non-stationary artificial matrix of measurements (sensitivity matrix), which is calculated, and respectively changed, "corrected" (is refined) at each current iteration *j*. This matrix incorporates all internal coupling of the thermal system and takes into account the estimates obtained. The $\hat{H}_{k+1}^{(j)}$ matrix terms represent the partial derivatives of the measured parameters with respect to the estimated (identified) parameters. The calculation of the artificial measurement matrix by the numerical method requires the solution of the equations of the process under study several times at each iteration, that is solving many direct heat transfer problems at each iteration.

The mathematical model of the thermal system formalized in the form of a matrix-vector equation, relating the system state vector $\vec{X}_{k+1}^{(j)}$ (in fact, estimate of this vector at *j*-th iteration) to the estimates of vectors of state and control, and vector of perturbation all at the previous *j-1*-st iteration, describes the Markov process that could be represented by a non-linear stochastic equation, namely,

$$\hat{\vec{X}}_k^{(j)} = \vec{f}\ \left(\hat{\vec{X}}_k^{(j-1)}\right) + \varphi\ \left(\hat{\vec{X}}_k^{(j-1)}\right) \hat{\vec{U}}_{k+1}^{(j-1)} + \psi\ \left(\hat{\vec{X}}_{k+1}^{(j-1)}\right) \vec{W}_{k+1}. \quad (6)$$

For the solution of the equation (6), the estimates $\hat{\tilde{Z}}_{k+1/k+1}^{(j-1)}$ substitute control vector $\hat{\vec{U}}_{k+1}^{(j-1)}$ in the right-hand side of this equation. The results of solving equation (6) represent the estimate of the state vector, starting from $\hat{\vec{X}}_1$.

The iteration number *"i"* (equation (4)) is used as the regularizing (tolerant) factor or the stop-criterion of the iteration process at each time step. This number *"i"* is selected from the

condition in terms of the agreement between the mean square errors (σ) of the measurements and the value of the general discrepancy, both over k moments of time, namely,

$$\left\| \hat{B}_k^{(j)} \hat{Z}_k^{(j)} - \tilde{\vec{Y}}_k \right\|^2 + \left\| \hat{\vec{Z}}_{k/k}^{(j)} - \hat{\vec{Z}}_{k/k}^{(j-1)} \right\|^2 \le f_1(\sigma_k). \quad (7)$$

This condition seems to be optimal, and it leads to regularization of the solution, and, therefore, it does not allow estimates to get into the tube of an eventual stability loss.

Because of ill-posedness, the solution obtained by means of algorithm (2) - (5) & (7) represents an optimization with restriction. The optimization is herein taken as the extremalization of a quality functional, such as the functional $\hat{I}_k^{(3)}$ from expression (1), by estimate (2).

Since the behavior of the thermal system can be most totally described in terms of probabilistic characteristic of its parameters, and because the information obtained from experiment always bears a random character, the initial model of heat transfer process under study should be written as non-linear stochastic equation

$$\tilde{\vec{Y}}_k = \vec{h}\ [\vec{Z}_k] + \vec{V}_k, \quad (8)$$

where $\vec{V}_k$ is the stochastic (white Gaussian sequence) errors of measurements that include: 1. direct errors of measurements, 2. any other errors or accuracy of determination of constraints $\vec{\alpha}_k$ imposed on the state or control parameters, 3. accuracy of discreteness of the mathematical model of process under study, etc.

After set up a problem, the equation (8) can be obtained from equation (6). On this basis, in the real thermal system, the case in point should be the stochastic approach to the solving of the identification, simulation, optimization, or control problems.

APPLICATIONS

General Statements

Each heat exchanger under study (Fig. 1) was installed on an adapter plate which was mounted on the test bench (calorimeter). Each device was equipped with a number of thermocouples (TC's) uniformly mounted on its outer surface to measure surface metal temperature (Fig. 1).

The test conditions were taken from actual vehicle HVAC system operation. For each flow rate, a set point has been allowed a long waiting period (from 0.5 to 1 hour) to make sure of stabilization of all refrigerant or coolant parameters. After the stabilization time, 10 measurements of device surface temperatures for each test set point have been taken. Simultaneously, the ambient temperature and inner surface temperature have been measured and recorded.

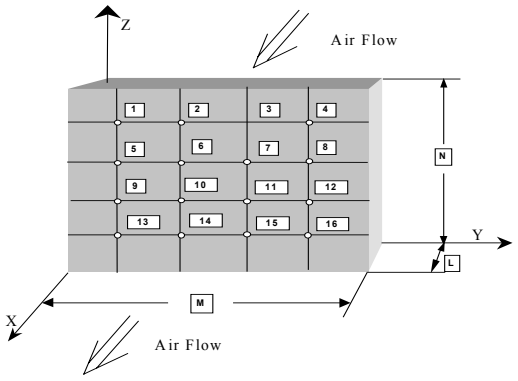


Fig.1 Schematic of Heat Exchanger Core with Air Flow, Axes of Coordinates, and Mounted TC's

Our experiments confirmed the uniformity of the heat exchanger discharge temperature (difference between highest and lowest temperatures measured by uniformly mounted TC's have not exceeded 0.75°C). In regard to the test conditions in the test bench, the air coming into the inner surface of the evaporator or heater attached to the calorimeter also had an even velocity and temperature distribution profile. Therefore, due to very stable conditions during the measurement, the local heat transfer coefficient or heat flux to or from the outlet side of the device under study can be represented as a constant value at constant air flow. Thus, we can solve the one-dimensional inverse heat transfer problem, that can be portrayed by the rod of the length equal to the coil's thickness L, with measurement data using an average of outer surface measured temperatures (all 16 TC's in Fig. 1.), where boundary conditions of 2nd or 3rd kind were defined, and with the 1st kind of boundary conditions on the inner surface of the rod (after all process' conditions were stabilized, inner surface temperature was equal to air inlet temperature that was measured).

One more very important consideration must be noted at this point. The proposed inverse heat exchanger material design or inverse

optimization approach on the preliminary stage of the HVAC system development should be as simple as possible and should require minimum time at the minimal cost. That is why the opportunity to solve the one-dimension problem is a good chance to simplify the suggested approach.

One more very important condition that must be satisfied during the testing process when measurements are taken is that the airside/tubeside heat transfer balance should be within 3%. During real HVAC system operations, the airside heat transfer is balanced on the tubeside of the coil with an equal amount of heat being either absorbed or given up by the fluid flowing through the tubes. So, the measuring procedure should begin only after stable air/refrigerant sides heat transfer balance has been achieved.

Evaporator Material Design

As it was mentioned under the "Mobile HVAC System" section, the frontal area of the heat exchanger is limited by the space available for the device. Thus the frontal area can not be varied much to satisfy the requirements for heat transfer. That is why the only way for HVAC system to significantly change the heat transfer between the evaporator and the air coming through it, is to produce the item from another material, thereby changing the heat transfer between the device and the surroundings. The material used in evaporator construction is of utmost importance because it greatly affects the thermal conductivity to the evaporation heat transfer process, and, thus the evaporator capacity. Estimation of heat transfer for evaporator produced from alternative prospective material requires the evaporator's thermal system mathematical simulation. The latter is impossible without very accurate information about boundary conditions between the existing evaporator and air.

Fin plate, four pass evaporator applied to minivan rear unit was equipped with twelve surface-mounted TC's (Fig. 1). The evaporator's dimensions were: Width: M = 0.188 m, Depth: L = 0.075 m, Height: N = 0.2 m. The test measured and calculated conditions are presented in Table 1 where the measured temperatures (last column) represent the average of twelve thermocouples and ten data points taken at each air flow rate.

The one-dimensional rod in hand (0.075 m in length) was divided into 10 sections with a pitch of $7.5*10^{-3}$ m, and the first kind of boundary conditions at the 11th node. The outer surface temperature $\tilde{T}_{os}$ is measured at the first node.

Table 1. Measured and Calculated Parameters

Air Flow, Kg/hr	Refrigerant Calculated Parameters,°C		Evaporator Surfaces Measured Temp.,°C	
	SH	SC	Inlet	Outlet
599	10.5	9.6	30.4	15.67
498	9.4	9.9	30.4	14.48
398	8.3	10.4	30.5	7.94
297	8.3	10.9	30.7	5.29
246	7.8	9.2	30.3	4.5

The initial approximation of the desired heat flux and covariance matrix required for the solution of IHTP were taken as $\hat{q}^{(0)} = 7 \cdot 10^4$, W/m^2 and $P=15000 \div 30000$, respectively. The ambient temperature was 28°C±.1°C. Initially, the problem was solved with thermal conductivity $k = 221$, $W/(m \cdot K)$.

In relation to the identification of only parameter q, constant for each air flow tested, the main AIF equation (2) and expression for the calculation of "measurement" matrix (5) take the forms:

$$\hat{q}^{(j)} = \hat{q}^{(j-1)} + K(j)\left[\tilde{T}_{os} - \hat{H}^{(j)}\hat{q}^{(j-1)}\right] \qquad (9)$$

$$\hat{H}^{(j)} = \left\{\frac{\partial T_{os}}{\partial q}\right\}^{(j)}. \qquad (10)$$

In such a manner the scalars $\tilde{T}_{os}$, H, P, R, & I are used for the identification of the heat flux q, that is scalar as well. Measured temperature $\tilde{T}_{os}$, that is changing with air flow rate, is defined as a normal Gaussian sequence with math expectation $E\left[\tilde{T}_{os}\right] = \bar{T}_{os}$, where $\bar{T}_{os}$ is the exact value of this temperature.

The identified heat flux as a function of coil outer surface temperature and air mass flow rate is shown in Fig 2 & 3 (curves with thermal conductivity k=221). The obtained results of the heat flux agree with the real heat transfer process in the automotive HVAC module.

Numerical simulation of the evaporation process with the evaporators produced from different materials leads to the introduction of the nomograms (Fig. 2 & 3), which serve as an interconnection between the evaporator material

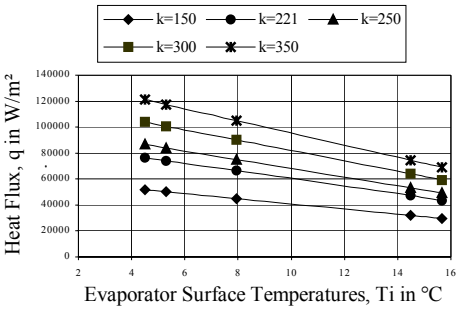


Fig. 2 Heat Flux between Evaporator & Ambient Air (Function of Surface Temperature & Material Thermal Conductivity, k, W/(m*K))

(thermal conductivity), air-device heat transfer, and the evaporator outer surface temperature (or air flow rate). The given in Fig. 2 & 3 curves have been obtained for all possible ranges of outer surface temperatures that correspond with the ranges of air flow (Table 1).

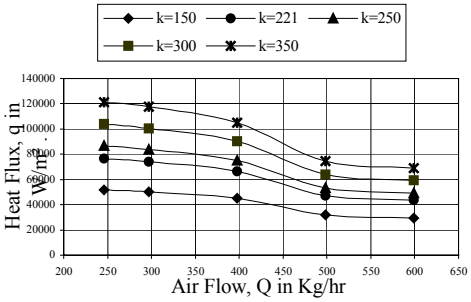


Fig. 3 Heat Flux between Evaporator & Ambient Air (Function of Air Flow & Material Thermal Conductivity, k, W/(m*K))

These nomograms (Fig. 2 & 3) enable us to make the preliminary inverse selection and inverse design of evaporator material. Indeed, if it is given the minimal possible heat flux for the the specified size of the evaporator, it is easy to find the evaporator's material, which provides the required heat transfer. For example, if the minimal required heat flux is 60,000, W/m^2 (Fig. 2), then material for the evaporator should be chosen with thermal conductivity no less than 300, W/(m*K) for all possible range of surface temperature. Another example, for the given ranges of air flow from 300, Kg/hr to 600, Kg/hr and required heat fluxes from 70,000, W/m^2 to 130,000, W/m^2 (Fig. 3) the evaporator's alloy should have thermal conductivity no less than 350, W/(m*K).

Additional area where could be applied the obtained nomograms consists in prediction of the true temperature of the evaporator's surface and air from the unit for the given (or calculated) heat flux. This application is very important for development of new material evaporator and preliminary evaluation of performance of current coil design both by heat transfer process mathematical simulation.

Evaporator Design Optimization

Inverse Dimensions Selection. The approach to the solution of the problem of design optimization of the HVAC system's evaporator has been proposed. The four pass, drawn cup-type vehicle front unit evaporator was equipped with sixteen thermocouples (Fig. 1). The evaporator's dimensions were: M = 0.21 m, L = 0.08 m, N = 0.2 m. The one-dimensional rod under study (0.08 m in length) was divided into 8 sections with a pitch of 10^{-2} m, and the first kind of boundary conditions at the ninth node. The outer surface temperature $(\widetilde{T}_{os})$ is measured at the first node.

The ambient temperature was 26.1°C. All initial measured and calculated system parameters that required for IHTP solution were the same as in material design example.

The results of measured and calculated parameters are presented in Table 2 where the measured temperatures (last column) represent the average of sixteen thermocouples and ten data points taken at each air flow rate.

Table 2. Measured and Calculated Parameters

Air Vol. Flow, V, M^3/s	Air Mass Flow, m, Kg/hr	Refrigerant Calculated Parameters, °C		Evaporator Surfaces Measured Temperatures, °C	
		SH	SC	Inlet	Outlet
0.138	600	11.1	8.3	30.2	11.4
0.115	501	6.7	8.3	30.2	10.0
0.092	398	5.6	9.4	30.1	8.8
0.069	296	6.7	10.6	30.1	7.9
0.046	198	6.7	10.0	30.3	6.5

The heat flux as a function of coil outer surface temperature was identified by means of a AIF (algorithm (9), (3), (4), (10), and (7)).

The total heat transfer of the heat exchanger depends upon the air mass flow coming through the device and its frontal area. So, as it was mentioned above, at the same airflow rate the heat transfer is defined by the surface area of the evaporator. Taking into consideration that the

evaporator frontal area is equal to $S=0.21*0.2=0.042(m^2)$, one can obtain the total heat transfer G of the evaporator under study as a function of either air flow rate or coil outer surface temperature (Fig. 4, curve "Total Heat Transfer").

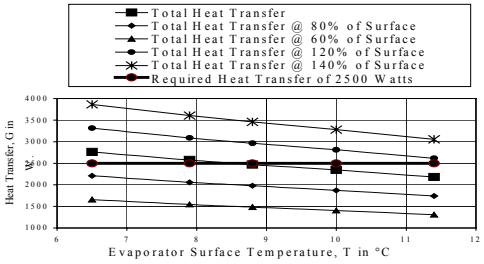


Fig. 4. Total Heat Transfer between Evaporator & Ambient Air (Function of Surface Temperature & Area)

Utilizing the same approach and statistically treating the obtained data, we can introduce the nomogram, which serves as an interconnection between the dimensions or, to be more specific, size of the frontal area of the evaporator, air-device total heat transfer, and the evaporator outer surface temperature (Fig. 4). Actually, our experiments have confirmed that the relationship between the surface temperature (or air flow rate), heat transfer, and discharge surface area bear the linear character because at the observed range of temperature change of density of air mass flow is negligible. That is why different curves in Fig. 4 can be obtained by multiplication with the surface ratio. The given curves have been obtained for all possible ranges of outer surface temperatures that correspond with the ranges of air flow (Table 2).

This nomogram enables us to choose the preliminary inverse selection (design optimization) of evaporator dimensions, which satisfy the total heat transfer value. For example, if total heat transfer requirement is 2,500 Watts, one draws a horizontal line (Fig. 4) to meet these requirements for all ranges of surface temperatures. Evidently, at this case, the frontal area of the device should not be less than 120% of the surface (0.042 m^2) of tested evaporator (Fig. 4) to satisfy required heat transfer at all temperature ranges.

Capacity Calculation. With the help of enthalpy analysis and on the basis of total heat transfer results (Fig. 4), the evaporator capacity can be calculated. The amount of heat transfer in Watts that takes place while the air stream passes through a cooling coil is the product of the air mass flow rate and the change of the air heat content (enthalpy). During a cooling and dehumidifying process, the air stream undergoes changes in both sensible and latent heat contents. Sensible heat is the heat associated with the change in the air's dry bulb temperature. Latent heat is the amount released by the water vapor as it condenses. Sensible and latent heats are usually combined and expressed as total heat, or enthalpy. A change in enthalpy includes changes in both sensible and latent heat, but it does not account for the small amount of heat in the condensed water that has left the air stream. However, since the total heat transfer between the evaporator outer surface and ambient air is identified by means of the IHTP solution on the basis of real surface temperature measurements, the value G already includes all three heats: sensible, latent, and condensed water. This statement is true, simply because the measured evaporator surface temperature is the final product of influence on the device and, specifically, on its frontal area, all of the above-mentioned heats. As for the coil's capacity calculations, since one identifies it on the basis of the results of total heat transfer, this capacity will include all kinds of heat that affects the performance of the cooling device. Based on evaporator total heat transfer G calculated in the previous section (Fig. 4) that can be rebuilt as a function of the air volume $V(\tau)$ through the evaporator, the coil capacity can be estimated by the equation: $Q = \int_0^t (G \times \frac{V(\tau)}{S}) d\tau, \quad t = 1s$. The calculated evaporator capacity as a function of airflow rate presented in Fig. 5 (middle curve).

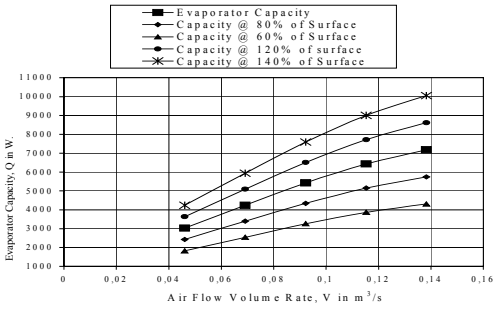


Fig. 5. Evaporator Capacity as a Function of Air Flow Rate

Using the same approach and arguments that have been used to construct graphs in Fig. 4, the nomogram for different evaporator dimensions can be created (Fig. 5). This nomogram serves as an interconnection between the size of the frontal

area of the evaporator, coil capacity, and airflow rate through the device. Using this nomogram, one is able to choose the preliminary inverse selection (design) of evaporator dimensions, which will satisfy the capacity requirements. For example, for a given maximum required capacity value, we can draw a horizontal line so as to meet these requirements either for all ranges of airflow, or to satisfy the maximum required capacity at the highest airflow. Evidently, for instance, if the required maximum evaporator capacity is 5.8 kW at airflow rate of $0.138 m^3/s$=600kg/hr=296 ft^3/min, the frontal area of the device should be selected not to exceed 80% of the surface (0.042 m^2) of tested evaporator to satisfy the required evaporator performance.

Heater Boundary Conditions Identification

Automotive Heating thermal system mostly operates in cold weather HVAC modes. Heat fluxes or heat transfer coefficients between an automotive heater core and fluids are supposed to be identified on the basis of IHTP solution. The airside of the heater core is much more "controlling" than the coolant side because of the much lower air side heat transfer coefficient (at least 1.5 times). This airside heat transfer coefficient greatly affects the transfer of heat from the coolant to the air and therefore its accurate measurement or identification is of extreme importance.

The typical two pass aluminum fins and copper/aluminum tubes heater core (Fig. 1) was equipped with seven TC's (instead of 16 as in Fig. 1 for the evaporators testing) that were uniformly fixed on its exterior surface (Fig. 1). Two more thermocouples were located at depth of 2 mm and 8 mm from the middle of the desired heater surface. These TC's were used thereafter for the identified results verification.

The heat transfer and capacity of the apparatus are governed by the rates of coolant flow and air flow. Another parameter that greatly affects the heater performance is the inlet air temperature. Due to these facts, the special test procedure was designed to obtain measuring (supporting) parameters for the inverse problem solution. The coolant/air flows were held constant, whereas the inlet air temperature is varied while measurements were taken. Then the air/glycol flows change and experiment was repeated, and so on. The fluid flow rates were chosen as maximum and minimum values of the actual vehicle HVAC system operation conditions. In such a manner the selected rates of heat transfer streams span all possible range of heat transfer in the heat exchanger under investigation. The following pairs of glycol/air flows were chosen: Q_{liq} = 1000±5 L/hr & Q_{air} = 400±5 Kg/hr; Q_{liq} = 1000±5 L/hr & Q_{air} = 200±5 Kg/hr; Q_{liq} = 300±5 L/hr & Q_{air} = 400±5 Kg/hr; and Q_{liq} = 300±5 L/hr & Q_{air} = 200±5 Kg/hr. The fresh air inlet temperature range was selected from -15^0C to 27^0C. Glycol inlet temperature was 85 ± 1.5^0C, which is consistent with real operation of a mobile HVAC system. The heater core dimensions were: M = 0.233 m, N = 0.155 m, L = 0.043 m. To solve the one-dimensional problem, the rod of 0.043 m in length, was divided into 42 sections with a pitch of 10^{-2} m, and the first kind boundary conditions at the 42nd node. The identification results of heat flux as a function of air inlet temperature are presented in Fig. 6.

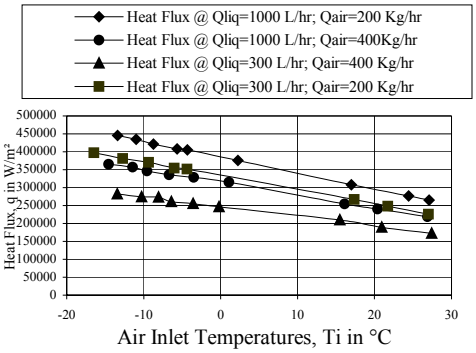


Fig. 6. Heat Flux between Automotive Heater Core and Ambient Air

Firstly, the obtained boundary conditions can be used for the simulation of heater thermal system. Another possible application is the use these parameters for the calculation of heater capacity and evaluation of the coolant side heater core performance. The heater capacity and performance data are applied to the heater core manufacturing process, as well as for the HVAC system performance evaluation. Other applications are is the same as discussed in previous sections such as heater material design and coil design optimization.

CONCLUDING REMARK

The proposed comprehensive approach and methodology of AIF is usable and perspective for the solution of various thermal optimization and identification problems occurred during mobile HVAC system design and development.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ADJOINT APPROACH TO INVERSE ANALYSIS OF CONVECTION HEAT TRANSFER

Kazunari Momose and Hideo Kimoto
Graduate School of Engineering Science
Osaka University
Osaka, Japan
momose@me.es.osaka-u.ac.jp

ABSTRACT

An inverse analysis based on adjoint formulation of forced convection heat transfer is proposed to obtain the optimal thermal boundary conditions for heat transfer characteristics; such as a total heat transfer rate or a temperature at a specific location. In the inverse analysis via adjoint formulation, the heat flow is reversed in both time and space. Thus, using the numerical solution of the adjoint problem, we can inversely predict the boundary condition effects on the heat transfer characteristics. As a result, we can obtain the optimal thermal boundary conditions in both time and space to control the heat transfer at any given time. A numerical solution of the adjoint problem in a two-dimensional enclosure is presented to illustrate the capability of the present method.

NOMENCLATURE

$\mathcal{L}$	differential operator defined in Eq.(9)
ℓ	dimensionless length
$\boldsymbol{n}$	unit normal vector to the boundary
Pr	Prandtl number
Q	dimensionless heating rate
q	dimensionless heat flux
T	dimensionless temperature
t	dimensionless time
$\boldsymbol{u}$	dimensionless velocity vector
$\boldsymbol{x}$	dimensionless space vector
x	dimensionless horizontal coordinate
y	dimensionless vertical coordinate
δ	Dirac's delta function
Γ	boundary
θ	dimensionless temperature defined in Eq.(2)
τ	dimensionless specific time
Ω	forced convection domain
$\boldsymbol{\xi}$	dimensionless space vector at a specific location

Subscripts

b	bottom wall
h	heating area on wall
in	inlet
l	left wall
opt	optimal heater location
out	outlet
r	right wall
s	steady state
u	upper wall
q	boundary specified via heat flux
θ	boundary specified via temperature

Superscripts

$*$	adjoint operator or adjoint variable
-	given or controllable value

INTRODUCTION

Since forced convection heat transfer is one of the fundamental processes in heat transfer, a large number of analytical and experimental studies have been made for various configurations [1, 2, 3]. However, the thermal boundary conditions employed in most studies have been limited to an isothermal or a uniform heat flux condition. Although the uniform thermal boundary condition is useful to simplify heat transfer phenomena and sometimes gives a good approximation for practical conditions, much of heat transfer problems that we encounter in practical applications have nonuniform and also unsteady thermal boundary conditions.

With recent progress in computer hardware and numerical simulation techniques [4, 5, 6], numerical prediction of heat transfer characteristics has

been possible. The numerical methods can be applied to heat transfer problems under arbitrary boundary conditions, which cannot be treated by analytical approaches. Thus, a number of packages of software for the convection heat transfer have been developed and are already used as a design tool. However, each result obtained by a numerical simulation gives a particular solution under a specific boundary condition; this means that we need to solve the problems again if the boundary conditions are modified. From the viewpoint of design or control, the boundary conditions to achieve the desired heat transfer characteristics are more significant than the heat transfer characteristics under specific boundary conditions. Thus, the problem becomes inverse one [7, 8].

From the viewpoint mentioned above, we have proposed the use of a numerical solution of an adjoint problem for a steady-state forced convection problem, in which the adjoint problem was derived from the linearity of energy equation. As a result, it was shown that the mean heat transfer characteristic can be expressed as a function of the surface temperature or the surface heat flux distribution; thus, the mean heat transfer characteristic can generally be evaluated under arbitrary steady thermal boundary conditions [9, 10].

In this paper, the adjoint approach mentioned above is extended to the unsteady forced convection problems under unsteady thermal boundary conditions. A numerical solution of the adjoint problem enables us to predict the heat transfer characteristic, such as the total heat transfer rate or the temperature at a specific location, for arbitrary time changes of thermal boundary conditions.

GOVERNING EQUATION AND BOUNDARY CONDITIONS

Consider a forced convection field Ω with boundary Γ. When the fluid is incompressible and the properties are not temperature dependent, the energy equation in forced convection problem can be written in a dimensionless form as

$$\frac{\partial T(\boldsymbol{x},t)}{\partial t} + \boldsymbol{u}(\boldsymbol{x})\cdot\nabla T(\boldsymbol{x},t) = \frac{1}{Pr}\nabla^2 T(\boldsymbol{x},t) \tag{1}$$

where $\boldsymbol{x}$ is the space vector, t is the time, T is the temperature, $\boldsymbol{u}$ is the velocity vector, and Pr is the Prandtl number. In this study, we assume that the flow field is steady and is given by a numerical solution; thus the energy equation (1) is a linear differential equation with a space-dependent coefficient $\boldsymbol{u}(\boldsymbol{x})$.

Here, we introduce a temperature θ defined as the difference from a known initial temperature distribution, such as

$$\theta(\boldsymbol{x},t) \equiv T(\boldsymbol{x},t) - T(\boldsymbol{x},0) \quad t \geq 0 \tag{2}$$

Further, we suppose that the boundary Γ consists of the temperature-specified boundary Γ_θ and the heat-flux-specified boundary Γ_q, namely

$$\Gamma = \Gamma_\theta \cup \Gamma_q \tag{3}$$

Under these assumptions, the governing equation and the boundary conditions adopted in this study can be summarized as follows:

$$\frac{\partial \theta(\boldsymbol{x},t)}{\partial t} + \boldsymbol{u}(\boldsymbol{x})\cdot\nabla\theta(\boldsymbol{x},t) = \frac{1}{Pr}\nabla^2\theta(\boldsymbol{x},t) \tag{4}$$

$$\theta(\boldsymbol{x},0) = 0 \quad \boldsymbol{x}\in\Omega \tag{5}$$

$$\theta(\boldsymbol{x},t) = \bar{\theta}(\boldsymbol{x},t) \quad \boldsymbol{x}\in\Gamma_\theta \tag{6}$$

$$q(\boldsymbol{x},t) \equiv \frac{1}{Pr}\frac{\partial\theta}{\partial n} - \theta\boldsymbol{u}\cdot\boldsymbol{n} = \bar{q}(\boldsymbol{x},t) \quad \boldsymbol{x}\in\Gamma_q \tag{7}$$

where $\bar{\theta}(\boldsymbol{x},t)$ and $\bar{q}(\boldsymbol{x},t)$ represent the boundary temperature and the boundary heat flux, both of which are given or controllable, and $\boldsymbol{n}$ denotes the outward unit vector normal to the boundary.

INTEGRAL EQUATION

For convenience in representing the integral equations derived later, we write Eq.(4) as

$$\mathcal{L}[\theta(\boldsymbol{x},t)] = 0 \tag{8}$$

where $\mathcal{L}$ is the linear operator defined as

$$\mathcal{L} \equiv \frac{\partial}{\partial t} + \boldsymbol{u}(\boldsymbol{x})\cdot\nabla - \frac{1}{Pr}\nabla^2 \tag{9}$$

Let us now consider the weak form of Eq.(8), which can be expressed as

$$\int_0^\tau\int_\Omega \theta^*\mathcal{L}[\theta]d\Omega dt = 0 \tag{10}$$

where θ^* is a test function or an adjoint temperature and τ is the specific time. Applying the divergence theorem to Eq.(10), we obtain the following

integral equation:

$$\int_0^\tau \int_\Omega \theta \mathcal{L}^*\left[\theta^*\right] d\Omega dt = \int_0^\tau \int_\Gamma \left(\theta^* q - q^* \theta\right) d\Gamma dt - \int_\Omega \left[\theta^* \theta\right]_0^\tau d\Omega \tag{11}$$

where q^* is the adjoint heat flux defined as

$$q^* = \frac{1}{Pr}\frac{\partial \theta^*}{\partial n} \tag{12}$$

and $\mathcal{L}^*$ is the adjoint operator for $\mathcal{L}$, and $\mathcal{L}^*$ possesses the form as

$$\mathcal{L}^* \equiv -\frac{\partial}{\partial t} - \boldsymbol{u}(\boldsymbol{x})\cdot\nabla - \frac{1}{Pr}\nabla^2 \tag{13}$$

In the adjoint operator derived above, the signs of unsteady and convection terms become negative, while that of diffusion term is unchanged, when compared with those in Eq.(9). This means that any adjoint problems derived later should be solved backward in time, and the solutions will be reversed in both time and space.

In order to eliminate the last term in the right-hand side of Eq.(11), we set the adjoint temperature at $t = \tau$ as

$$\theta^*(\boldsymbol{x},\tau) = 0 \quad \boldsymbol{x} \in \Omega \tag{14}$$

which can be regarded as an initial condition for the adjoint problem. In addition, for convenience, defining an integral operator as

$$\left\langle f * g \right\rangle_\Gamma^\tau \equiv \int_0^\tau \int_\Gamma f(\boldsymbol{x},\tau - t) g(\boldsymbol{x},t) d\Gamma dt \tag{15}$$

we arrive at the following integral equation:

$$\int_0^\tau \int_\Omega \theta \mathcal{L}^*\left[\theta^*\right] d\Omega dt = \left\langle \theta^* * q \right\rangle_\Gamma^\tau - \left\langle q^* * \theta \right\rangle_\Gamma^\tau \tag{16}$$

BOUNDARY INTEGRAL RELATIONSHIPS

Equation (16) suggests that if we can eliminate the left-hand side integral, which is a time-domain integral, we can obtain several time-boundary integral relationships.

First, let us adopt the simplest adjoint problem, such as

$$\mathcal{L}^*\left[\theta^*(\boldsymbol{x},t)\right] = 0 \tag{17}$$

and let us set the boundary conditions for the adjoint problem, such that

$$\theta^*(\boldsymbol{x},t) = \delta(\tau - t) \quad \boldsymbol{x} \in \Gamma_\theta,$$
$$q^*(\boldsymbol{x},t) = 0 \quad \boldsymbol{x} \in \Gamma_q \tag{18}$$

where $\delta(\cdot)$ is Dirac's delta function. Then we obtain the following boundary integral relationship:

$$Q(\tau) \equiv \int_{\Gamma_\theta} q(\boldsymbol{x},\tau) d\Gamma = \left\langle q^* * \overline{\theta} \right\rangle_{\Gamma_\theta}^\tau - \left\langle \theta^* * \overline{q} \right\rangle_{\Gamma_q}^\tau \tag{19}$$

Equation (19) indicates that if we numerically solve the adjoint equation (17) under the initial and boundary conditions in Eqs.(14) and (18), we can predict the total heat transfer rate on Γ_θ at $t = \tau$ under arbitrary thermal boundary conditions.

In a similar fashion, if we choose

$$\mathcal{L}^*\left[\theta^*(\boldsymbol{x},t)\right] = \delta(\boldsymbol{\xi})\delta(\tau - t) \tag{20}$$

under

$$\theta^*(\boldsymbol{x},t) = 0 \quad \boldsymbol{x} \in \Gamma_\theta,$$
$$q^*(\boldsymbol{x},t) = 0 \quad \boldsymbol{x} \in \Gamma_q \tag{21}$$

we get the following relationship:

$$\theta(\boldsymbol{\xi},\tau) = \left\langle \theta^* * \overline{q} \right\rangle_{\Gamma_q}^\tau - \left\langle q^* * \overline{\theta} \right\rangle_{\Gamma_\theta}^\tau \tag{22}$$

where $\boldsymbol{\xi}$ is a specific location at which we want to predict the temperature. Equation (22) implies that the solution of Eq.(20) under Eq.(21) enables us to evaluate arbitrary thermal boundary condition effects on the temperature at $\boldsymbol{\xi}$. Moreover, if we replace the point impulse in the right-hand side of Eq.(20) with that in a finite area, we can predict the mean temperature within the area.

For steady-state problems, the time-dependent notations can be simplified as follows:

$$\theta(\boldsymbol{x},t) \to \theta_s(\boldsymbol{x}), \quad q(\boldsymbol{x},t) \to q_s(\boldsymbol{x}),$$
$$\mathcal{L} \to \mathcal{L}_s \equiv \boldsymbol{u}\cdot\nabla - \frac{1}{Pr}\nabla^2 \tag{23}$$
$$\theta^*(\boldsymbol{x},t) \to \theta_s^*(\boldsymbol{x}), \quad q^*(\boldsymbol{x},t) \to q_s^*(\boldsymbol{x}),$$
$$\mathcal{L}^* \to \mathcal{L}_s^* \equiv -\boldsymbol{u}\cdot\nabla - \frac{1}{Pr}\nabla^2 \tag{24}$$
$$\delta(\tau - t) \to 1,$$
$$\left\langle f * g \right\rangle_\Gamma^\tau \to (f_s, g_s)_\Gamma \equiv \int_\Gamma f_s(\boldsymbol{x}) g_s(\boldsymbol{x}) d\Gamma \tag{25}$$

Then the relationships obtained in Eqs.(19) and (22) become

$$Q_s \equiv \int_{\Gamma_\theta} q_s(\boldsymbol{x})\,d\Gamma = \left(q_s^*, \overline{\theta}_s\right)_{\Gamma_\theta} - \left(\theta_s^*, \overline{q}_s\right)_{\Gamma_q} \tag{26}$$

$$\theta_s(\boldsymbol{\xi}) = \left(\theta_s^*, \overline{q}_s\right)_{\Gamma_q} - \left(q_s^*, \overline{\theta}_s\right)_{\Gamma_\theta} \tag{27}$$

for steady-state problems.

COMPUTATIONAL IMPLEMENTATION

Let us now consider Eq.(20) as an example for computational implementation of the adjoint problem. The implementation is surprisingly easy. Reversing the velocity vector and shifting time, such that

$$\boldsymbol{u}' = -\boldsymbol{u}, \qquad t' = \tau - t \tag{28}$$

we obtain

$$\frac{\partial \theta^*}{\partial t'} + \boldsymbol{u}' \cdot \nabla \theta^* = \frac{1}{Pr} \nabla^2 \theta^* + \delta(\boldsymbol{\xi})\delta(t') \tag{29}$$

This is the same form as an energy equation having a point source term. Thus, the adjoint problem obtained in Eq.(29) can easily be solved by a standard heat transfer code based on a finite difference method or a finite element method.

It should be noted that the solution of the adjoint problem described in Eq.(29) converges rapidly in time, because it is an impulse response from a unit impulse at $\boldsymbol{\xi}$. In other words, only the early stage of the solution is significant and the late stage is negligible. Thus, we can predict $\theta(\boldsymbol{\xi},\tau)$ in Eq.(22) at arbitrary time τ by recycling a truncated solution in the convolution integral with respect to the time, which is defined in Eq.(15).

In addition, using a relationship

$$\theta_s^*(\boldsymbol{x}) = \int_0^\infty \theta^*(\boldsymbol{x}, t')\,dt' \tag{30}$$

we can also obtain the steady-state solution of the adjoint problem from the truncated unsteady solution.

ILLUSTRATIVE EXAMPLE

As an illustrative example of the present method, we computed an adjoint problem in a two-dimensional enclosure with one inlet and one outlet as shown in Fig.1. The purpose of this example is to predict the influence of the inlet temperature and the left- and bottom-wall heating on the temperature at $\boldsymbol{\xi}$ ($x = 0.25$, $y = 0.25$), when the temperatures on the other walls and at the outlet are fixed at zero. This can be done by solving Eq.(20) under the boundary conditions (21), which become

$$\theta^*(\boldsymbol{x},t) = 0 \quad \boldsymbol{x} \in \Gamma_{in} \cup \Gamma_u \cup \Gamma_r \cup \Gamma_{out} \tag{31}$$

$$q^*(\boldsymbol{x},t) = 0 \quad \boldsymbol{x} \in \Gamma_l \cup \Gamma_b \tag{32}$$

in this example. Then, using the adjoint solution, the temperature at $\boldsymbol{\xi}$ and at τ can be predicted by

$$\theta(\boldsymbol{\xi},\tau) = \left\langle \theta^* * q \right\rangle^{\tau}_{\Gamma_l \cup \Gamma_b} - \left\langle q^* * \theta \right\rangle^{\tau}_{\Gamma_{in}} \tag{33}$$

In the computation, we assumed that the Prandtl number is 0.71, and employed a standard finite difference method for both the flow and adjoint temperature computations, in which the computational domain was nonuniformly divided into 100 grid points in both horizontal and vertical directions. Although the computational time step was 2×10^{-7} to make computation stable, the computed results were saved in files at intervals of 2×10^{-3} until $t = 0.6$, at which all the adjoint variables were converged to almost zero, and were enough to calculate the convolution integral in Eq.(33).

We should note that all the results presented below, except some direct simulation results only for the confirmation, can be obtained from a single numerical solution of the adjoint problem mentioned above. The computation time for the adjoint problem is almost equal to that for the conventional forced convection problem under a specific boundary condition.

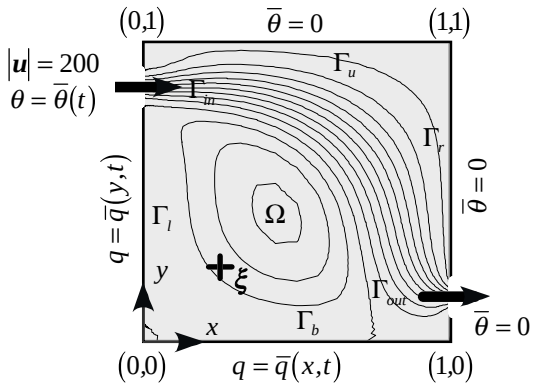


Fig.1 Configuration of model forced convection field in a two-dimensional enclosure with one inlet and one outlet

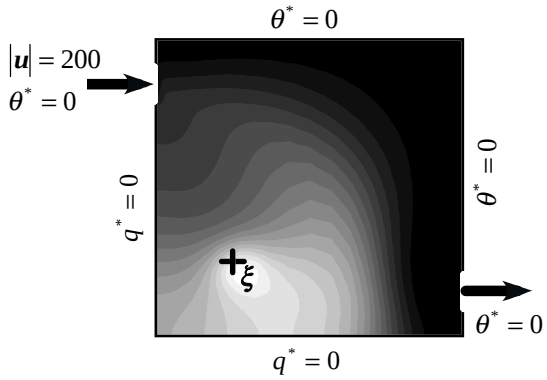
Optimization of Wall Heating

Let us consider the influence of steady inlet temperature and steady wall heating on the temperature at $\boldsymbol{\xi}$. From Eq. (27), the temperature at $\boldsymbol{\xi}$ under steady-state condition can be obtained by

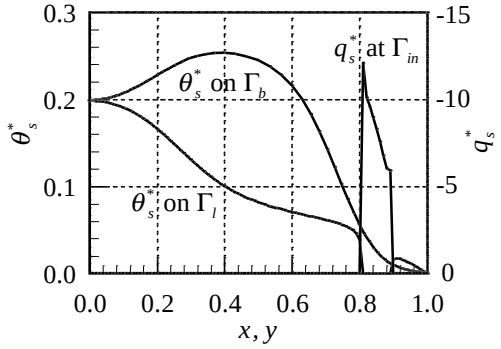
$$\theta_s(\boldsymbol{\xi}) = \left(\theta_s^*, \bar{q}_s\right)_{\Gamma_l \cup \Gamma_b} - \left(q_s^*, \bar{\theta}_s\right)_{\Gamma_{in}} \tag{34}$$

where θ_s^* and q_s^* are the adjoint temperature and adjoint heat flux in steady state, which can be calculated from the truncated unsteady solution of adjoint problem using the relationship in Eq.(30). The steady-state adjoint temperature field obtained by the time integration is shown in Fig.2(a), and the adjoint temperature distributions on left and bottom walls and the adjoint heat flux distribution at inlet are shown in Fig.2(b).

From Fig.2(b), we can predict the most effective heating location to increase the temperature at $\boldsymbol{\xi}$; in this example, this location is around x = 0.4 on the bottom wall. Thus, if we install a heater on a wall in a steady-state situation, the location around x = 0.4 on the bottom wall is the best to raise the temperature at $\boldsymbol{\xi}$.



(a) adjoint temperature field



(b) influence of inlet temperature and wall heating on temperature at $\boldsymbol{\xi}$

Fig.2 Steady-state results obtained from a unsteady adjoint computation

To demonstrate the results in Fig.2, let us consider the minimization of wall heating rate to achieve $\theta_s(\boldsymbol{\xi}) = 1$ when the inlet temperature is unity within the inlet ($\bar{\theta}_{in} = 1$). From Eq.(34), if the heating area on the wall is Γ_h and the heat flux $\bar{q}_h$ is uniform on the heating area, the required heating rate $\bar{Q}_h$ can be obtained as

$$\bar{Q}_h \equiv \bar{q}_h \ell_h = \frac{1 + \int_{\Gamma_{in}} q_s^*(\boldsymbol{x}) d\Gamma}{\int_{\Gamma_h} \theta_s^*(\boldsymbol{x}) d\Gamma} \ell_h \tag{35}$$

where ℓ_h is the heater length. In this optimization (minimization) problem, we assume ℓ_h to be 0.2 and set the optimal heating area on Γ_{opt} ($0.3 \le x \le 0.5$).

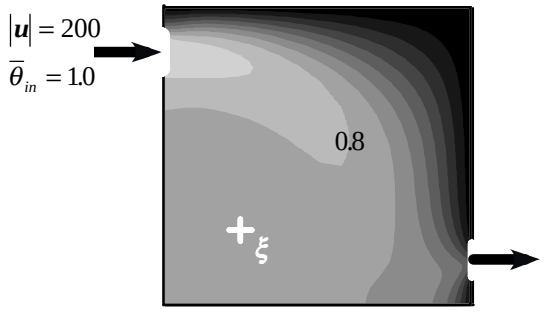
The optimal heating rate, which is determined by applying Γ_{opt} to Eq.(35), is indicated in Tab.1, in which the heating rate of uniform heating on both left and right walls are also indicated. Table 1 suggests that if we install a heater on Γ_{opt}, the heating rate to achieve $\theta_s(\boldsymbol{\xi}) = 1$ is less than half of that required in the uniform heating. The results in Tab.1 were confirmed by direct numerical simulations. The resulting temperature fields are shown in Fig.3; the agreement of predicted temperatures at $\boldsymbol{\xi}$ with the temperatures by the direct simulations can be seen to be quite good.

Optimal Control of Wall Heating

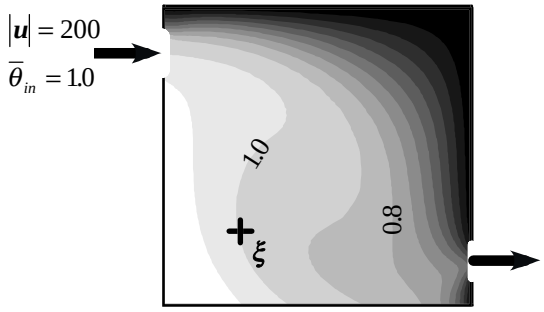
As an unsteady optimization problem using the adjoint solution obtained, let us consider an optimal control of a heater installed on a wall. The objective of this example is to maintain $\theta(\boldsymbol{\xi}, t) = 1$ by controlling the heater input $\bar{Q}_{opt}(t)$, when the inlet temperature $\bar{\theta}_{in}(t)$ is changed in time. For simplicity, the heater location is fixed at Γ_{opt} ($0.3 \le x \le 0.5$) determined in the above steady-state problem, and the inlet temperature and the heat flux on the heater are uniform in space as well as those assumed in the steady-state problem.

Table 1 Predicted optimal heating rate

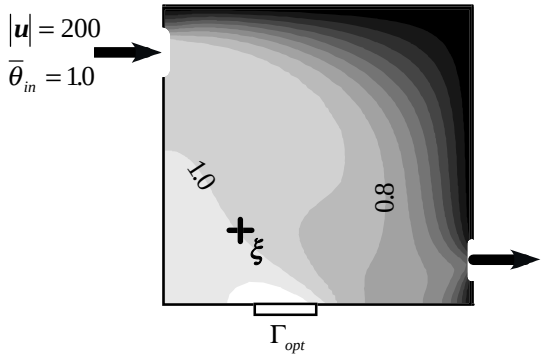
Heating Area Γ_h	Γ_{opt}	$\Gamma_l \cup \Gamma_b$	None
Heating Rate $\bar{Q}_h$	0.845	1.791	—
Predicted $\theta_s(\boldsymbol{\xi})$	1.0	1.0	0.751



(a) without wall heating
($\overline{Q}_h = 0.0$, $\theta_s(\boldsymbol{\xi}) = 0.742$)



(b) uniform heating on left and bottom walls
($\overline{Q}_h = 1.791$, $\theta_s(\boldsymbol{\xi}) = 0.989$)



(c) optimal heating on Γ_{opt}
($\overline{Q}_h = 0.845$, $\theta_s(\boldsymbol{\xi}) = 0.993$)

Fig.3 Direct numerical simulation results associated with Table 1

Under these assumptions and using the relationship derived in Eq.(33), the temperature at $\boldsymbol{\xi}$ can be expressed as

$$\theta(\xi,\tau) = \int_0^\tau \theta^*_{opt}(\tau - t)\overline{Q}_{opt}(t)dt - \int_0^\tau Q^*_{in}(\tau - t)\overline{\theta}_{in}(t)dt \tag{36}$$

where

$$\theta^*_{opt}(t) = \frac{1}{\ell_h}\int_{\Gamma_{opt}} \theta^*(\boldsymbol{x},t)d\Gamma\,,$$

$$Q^*_{in}(t) = \int_{\Gamma_{in}} q^*(\boldsymbol{x},t)d\Gamma \tag{37}$$

which are indicated in Fig.4. On the basis of the property of convolution integral in Eq.(36), both $\theta^*_{opt}(t)$ and $Q^*_{in}(t)$ at small *t* indicate the immediate effect, while those at large *t* indicate the effect after some delay. Thus, Fig.4 shows that the heat flow from the heater at Γ_{opt} reaches $\boldsymbol{\xi}$ more quickly than that from the inlet.

In accordance with the relationship in Eq.(36) with the known impulse responses $\theta^*_{opt}(t)$ and $Q^*_{in}(t)$ shown in Fig.4, the optimal heater input $\overline{Q}_{opt}(t)$ must satisfy the following relationship to maintain $\theta(\boldsymbol{\xi},t) = 1$ under arbitrary inlet temperature variation $\overline{\theta}_{in}(t)$.

$$\int_0^\tau \theta^*_{opt}(\tau - t)\overline{Q}_{opt}(t)dt = f(\tau) \tag{38}$$

where $f(\tau)$ is the function of τ, such as

$$f(\tau) = 1 + \int_0^\tau Q^*_{in}(\tau - t)\overline{\theta}_{in}(t)dt \tag{39}$$

which can be calculated if the inlet temperature variation is specified. Thus, if we can evaluate Eqs.(38) and (39) by a certain numerical manner, the optimal heating rate $\overline{Q}_{opt}(t)$ can be explicitly obtained at any τ. In this example, we employed a time discretization method to evaluate the convolution integrals in Eqs.(38) and (39); in time discretization, stepwise profiles at time intervals of 2×10^{-3} were used to evaluate all functions appearing in the above convolution integrals.

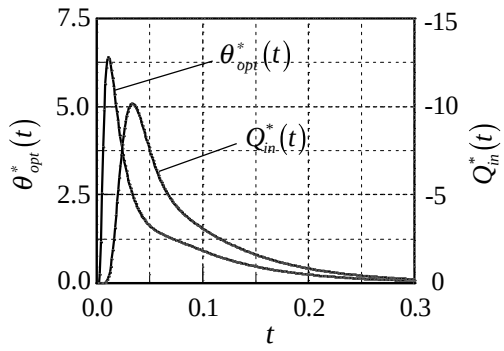
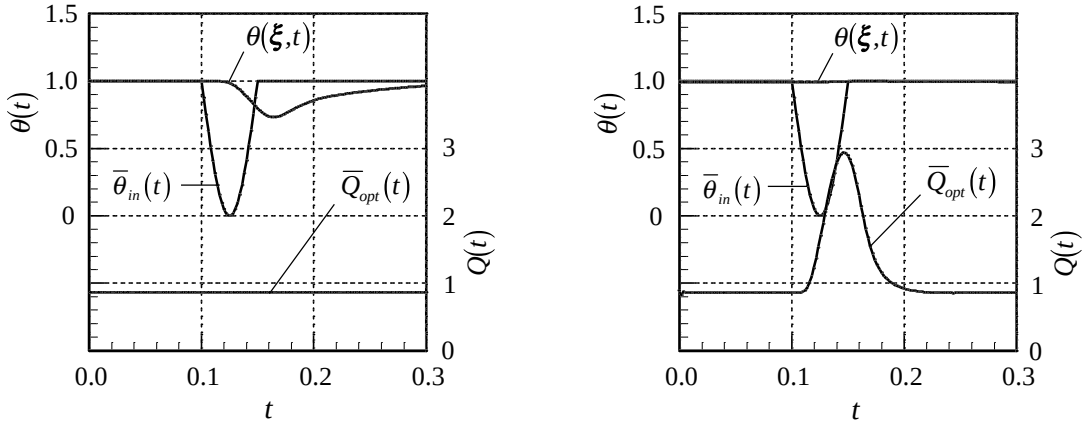


Fig. 4 Impulse responses to $\boldsymbol{\xi}$ from inlet (Γ_{in}) and from optimal heating area (Γ_{opt})

As an example, let us assume that the inlet temperature $\overline{\theta}_{in}(t)$ suddenly drops for a short period as shown in Fig.5(a) and 5(b), in both of which the initial temperature fields are the same as that shown in Fig.3(c); thus, in both cases, the temperatures at ξ are initially maintained at unity ($\theta(\xi,t)=1, t \le 0$).

Figure 5(a) shows that if the heater input is not controlled, the temperature at ξ decreases with some delay after the decrease of the inlet temperature. On the other hand, with optimally controlled heater input shown in Fig.5(b), the temperature at ξ can be maintained at unity against the decrease of the inlet temperature.

In order to confirm the present optimal control result, the corresponding numerical simulation was carried out using the optimal heater input suggested in Fig.5(b). The temperature fields obtained are displayed in Fig.6 at typical times. As shown in Fig.6, the temperature at ξ can satisfactorily be maintained with optimally controlled heater input.



(a) without control (b) with optimal control

Fig.5 Optimal control of heating rate on Γ_{opt} to maintan $\theta(\xi,t)=1$ against temperature variation at inlet

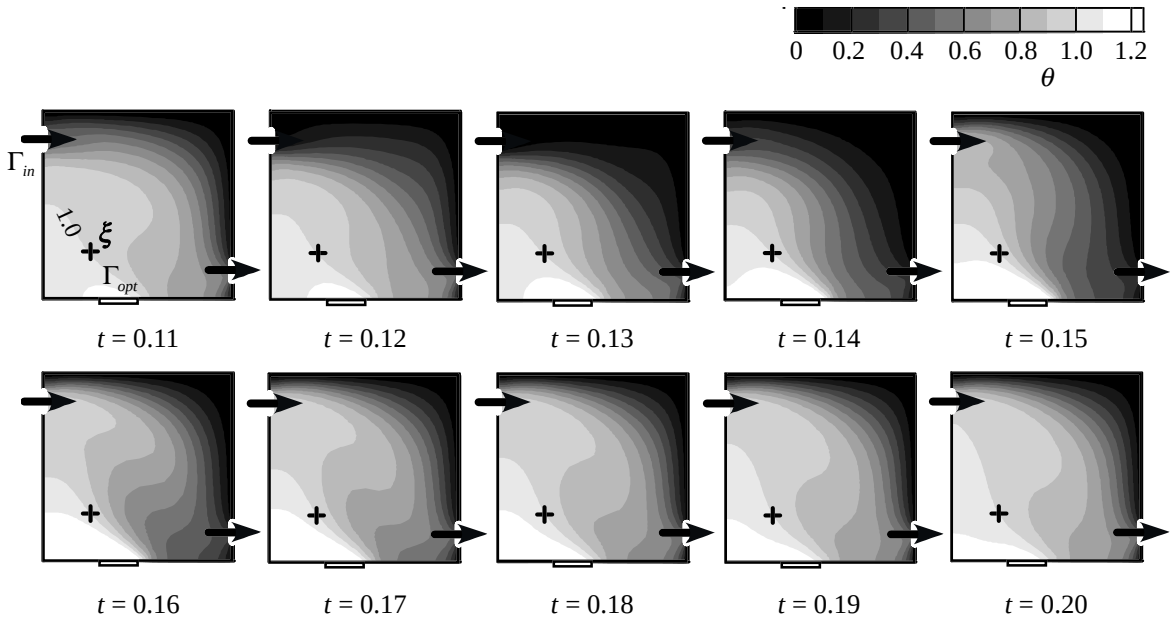


Fig.6 Time change of temperature field with optimal contorol indicated in Fig.5(b)

CONCLUDING REMARKS

For general evaluation of thermal boundary condition effects on forced convection heat transfer, we propose a noniterative inverse method based on an adjoint formulation of the energy equation. The main features of the present approach can be summarized as follows:

(1) A single adjoint computation enables us to predict the heat transfer characteristic, such as the total heat transfer rate or the temperature at a specific location, for arbitrary time changes of thermal boundary conditions.

(2) Using the adjoint solution, we can obtain the optimal thermal boundary conditions both in time and space to control the heat transfer at any given time.

(3) The computation time for the adjoint problem is almost equal to that required in an ordinary numerical simulation of forced convection problem under a specific boundary condition.

Acknowledgments

This work was supported by the Ministry of Education, Culture, Sports, Science and Technology through the Grant-in-Aid for Scientific Research (Basic Research (c)(2), No.12650204). The authors would like to thank Mr. Kouhei Abe, a graduate student of Osaka University, for the assistance on computational work.

REFERENCES

1. W. H. MacAdams *Heat Transfer Transmission* (*3rd Ed*.), MacGraw-Hill, 1954.

2. A. Zukauskas, *Advances in Heat Transfer*, Vol. 8, 1972.

3. A. Bejan, *Convection Heat Transfer* (*2nd Ed.*), John Wiley & Sons, 1995.

4. S. V. Patankar, *Numerical Heat Transfer and Fluid Flow*, MacGraw-Hill, 1980.

5. C. A. J. Fletcher, *Computational Techniques for Fluid Dynamics*, Springer, 1991.

6. J. H. Ferziger and M. Peric, *Computational Methods for Fluid Dynamics*, Springer, 1997.

7. JSME Edition, *Computer Analyses of Inverse Problems*, Corona-sha, 1991. (in Japanese)

8. K. Momose and H. Kimoto, Green's Function Approach to Multidimensional Inverse Heat Conduction Problems. *Advances in Computational Engineering Science* (Edited by Atluri SN, Yagawa G). Tech Science Press; 1997.

9. K Momose, K. Sasoh and H. Kimoto, Thermal Boundary Condition Effects on Forced Convection Heat Transfer (Application of Numerical Solution of Adjoint Problem), *JSME Int. J., B*, **42**-2, 293 (1999)

10. K. Momose and H. Kimoto, Numerical Prediction of Convection Heat Transfer under Arbitrary Thermal Boundary Conditions, *Proc. 5th ASME/JSME Joint Thermal Eng. Conf.*, CD-ROM, AJTE99-6381 (1999)

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brasil, 2002

RESEARCH A PHYSICAL PICTURE OF HEAT AND MASS TRANSFER PROCESSES IN VIBRATING HEAT PIPES

Melikayev Yu.,
Dniepropetrovsk National University;
Nauchnaya,13; 49050, Dniepropetrovsk; Ukraine
ymelikaev@mail.ru

Prisniakov K., Prisniakov V.
Institute Geomechanics of Nat. Ac. Sc. Simpher-opolskaya 2A; 49005; Dniepropetrovsk, Ukraine;
kprisn@a-teleport.com

Nikolaenko Yu., Kravez Yu.
Kiev; Ukraine

ABSTRACT

Results of research of heat and mass transfer processes in the thermal pipes subjected to vibrating influence are submitted. The vibrations influence on thermal resistance of thermal pipes and the maximal transferred ability was experimentally investigated. Characters of this influence determine the incomplete adequacy of the existing now representations to a real picture of processes of heat and mass transfer in porous systems. In particular, it is experimentally established, that thermal resistance of a thermal pipe can both decrease and rise at action of vibration. On the basis of experimentally established dependences the new concept of influence of vibrations on processes in thermal pipes is offered. The statement of a task of the wide theoretical and experimental researches of influence of vibrations on heat and mass transfer intensity in the porous systems is submitted.

INTRODUCTION

Energy stations and units of modern technology are characterized by high values of thermal capacity and mechanical pressure. These characteristics have a main influence on a choice of a material and a design data of this devices [2, 6]. Thus, joint influence of thermal capacity and mechanical pressures is shown, as a rule, only in the account of change of properties of the constructional materials at the increased temperatures [5].

However, other aspect of such interference exists: variable pressure causes the vibration of power unit elements, that can result in change of intensity of heat and mass transfer processes, proceeding in it under certain conditions. Effect of vibrations on the intensity of heat exchange, including in heat pipes, was investigated and earlier [1,3, 6]. Results of these researches testify to ambiguity of such influence - vibration of a heat-transmitting surface or the liquid can both to increase, and to reduce effective factor of heat exchange. In represented work preliminary results of experimental researches of influence of vibrations on the thermal resistance and the maximal transmitted capacity of heat pipes are given.

EXPERIMENTAL RESEARCH

For realization of experimental researches thermal pipes with the gas channel diameter from 1 up to 20 mm made of a thin-walled metal pipe and supplied with matches made from metal grid were selected. Working bodies of thermal pipes are water, acetone and spirit at the various degree of filling.

The heat pipe was settled on the vibrating stand. For a heat supply and heat removal from a heat pipe external heat-carriers - the hot and cold water, directed in the heating and cooling chambers were used accordingly. In case of need creations of the big gradients of temperatures as a source of heat the electric heater (spiral from nichrome on a warmed site of a heat pipe) was used. A transport site of a thermal pipe with the purpose of reduction of losses of heat was covered with a thick layer of thermoresistance material.

For an estimation of a thermal condition of a pipe on an external surface of transport site were welded the thermocouple conductors, allowing to measure absolute values and differences of temperatures of a working body in different areas of a thermal pipe. Besides to measurement the difference between the temperature of cold water on an input and on an output from cooling chamber was subject. For this purpose the differential thermo-

couple of graduation cuprum - copelium was used.

On a platform of the vibrating stand the frame of fastening of a thermal pipe and the acceleration gauge, connected to the specialized measuring device was settled. Model calculations of ther-mal resistance of a pipe wall (results here are not represented) have shown, that at intensity of heat transfer at a level up to 100 W/sm^2 temperature drop on thickness of a metal wall and a match of a thermal pipe does not exceed 0.1K. At the same time the error of measurement of a differe-nce of temperatures of surface heat pipe was estimated by us at a level 0.2 K. Therefore, we neglected by thermal resistance of pipe wall and believed, that the measured temperature of pipe surface is approximately equal to working body vapor temperature of the appropriate zone.

During experiments the following parameters were measured and registered:

- temperature increase of cool water on a cooled site of thermal pipe;
- temperatures of thermal pipe surface in places, where the thermocouples was setting, at the established mode of heat exchange;
- the mass speed of current of cold water was determined by measurement of time which was necessary for the filling of measuring volume;
- vibration acceleration of fastening frame of thermal pipe ;
- the frequency of vibrations, which was generated by vibration stand.

Processing results of measurements was consisted in calculation of thermal resistance of a heat pipe and the maximal transmitted capacity at the various parameters of vibrating influence. Further, among the parameters of system describing its thermal and mechanical condition, were searched such, for which influence on the thermal resistance and transmitted capacity appeared the strongest and unequivocal.

Definition of parameters of vibrating influence (amplitude of fluctuation and vibration speed) was carried out on known dependences

$$a = uw, \quad u = w\,A, \quad w = 2p\,f,$$

where a is vibration acceleration,- u is vibration speed; A is amplitude of vibrations; f is frequency*;* w is circular frequency.

Thermal resistance of a thermal pipe was understood as size, return to effective factor of heat conductivity

$$R = \frac{(T_1 - T_2)\cdot}{W \cdot l},$$

where R is thermal resistance, K / Wm; T_1, T_2- are temperatures of a surface of warmed and cooled sites, accordingly, K; W is transmitted thermal capacity, W; l is length of a transport site of a thermal pipe, m.

Some results of experimental researches are submitted on fig. 1-3.

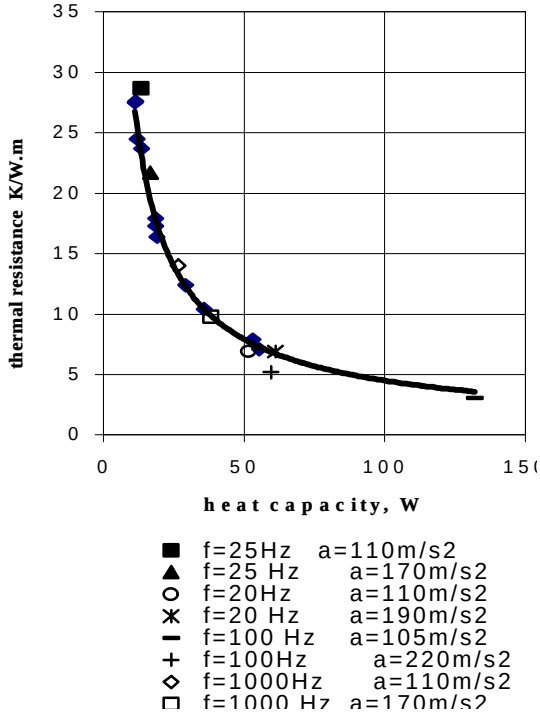


Fig 1. Dependence of thermal pipe resistance on transmitted capacity

On fig. 1 continuous curve shows dependence of thermal resistance of a heat pipe with diameter of 6 mm from the transmitted thermal capacity, taken at absence of vibrations. Points of different color show some characteristic modes of a vibrating heat pipe.

The kind of dependence of thermal resistance from transmitted capacity and character of vibrations effect for a pipe with diameter of 6 mm is a differs from the results received for heat pipes of the greater size very not enough. However for more tiny pipes this dependence differs from usual (fig. 2). The positive inclination of diagram *R=f (W)* in a range of small values of transmitted capacity first of all is evident. Character of effects of vibration of different frequencies on this dependence is determined, basically frequency.

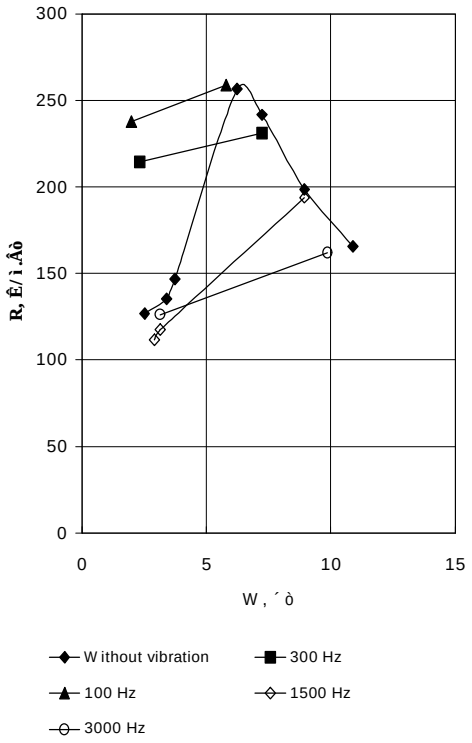


Fig 2. Dependence of heat pipe resistance on transmitted capacity at various frequencies of vibration

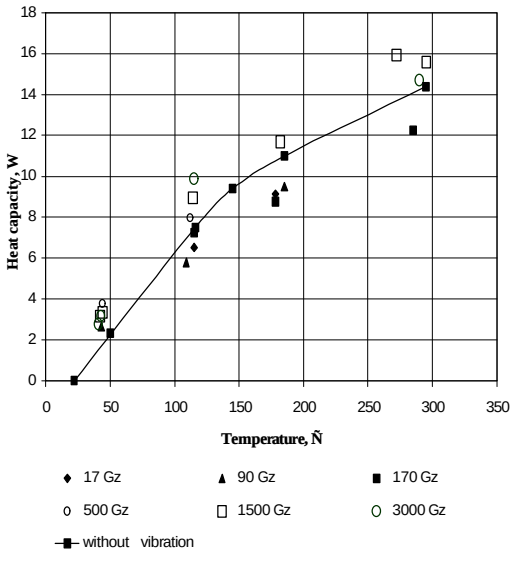


Fig. 3. Dependence of transmitted capacity of a heat pipe on temperature of a heated up site under various frequency of vibration

Rather low frequencies of vibration (100-300 Hz) sharply worsen efficiency of a heat pipe in the field of small values of transmitted thermal capacity. Vibrations of higher frequencies (1500-3000 Hz), on the contrary, reduce thermal resistance in the field of average and high values of transmitted capacity. Obviously, qualitative distinctions of characteristics of pipes in diameter of 6 and 4 mm shows strong influence of edge effects on a stream parameters in the thermal pipes of small diameter.

The characteristic of a heat pipe is submitted as dependence of transmitted capacity from temperature of the hot end on fig. 3.. These data qualitatively correlate with the data fig. 2: high-frequency vibrations increase transmitted capacity at the fixed temperature gradient, and low-frequency vibrations, on the contrary, reduce it.

Thus, vibrations of various frequencies have differently an effect for efficiency of heat pipes of the different characteristic size. The greatest effect (positive) is rendered vibrations with frequency about 100 Hz for a thermal pipe with diameter of 6 mm, this characteristic frequency grew up to several thousand Hz for a pipe in diameter of 4 mm.

The given here experiments results have not by universal character unfortunately. The strong experiment dependence of the characteristic frequencies causing abnormal changes of characteristics of a thermal pipe was detected not only for its characteristic size, but also for orientation of a pipe concerning a direction of vibrations, for type of wick and for other parameters. The correlation analysis on the limited volume of experimental data has not allowed determining types of this dependence. For example, in fig.4 dependence of thermal resistance of a heat pipe on parameters of vibrating influence are given.

The given in figure 4 data are received for a heat pipe with diameter 20 mm. In this series of experiments the fluctuations with frequency 0 - 30 Hz and acceleration of 0-300 m/s^2 were used. On fig. 4 (a) influence of vibration of various frequencies on the thermal resistance is shown at the fixed vibrating acceleration equal of 100 m/s^2.

Apparently from figure, appreciable influence on a mode of operation of this heat pipes the vibrations of a narrow frequency range - about 17 - 25 Hz is showing. On fig. 4 (b-d) dependence of thermal resistance from the amplitude of fluctuations, vibrating speed and vibrating acceleration for this frequency range are represented.

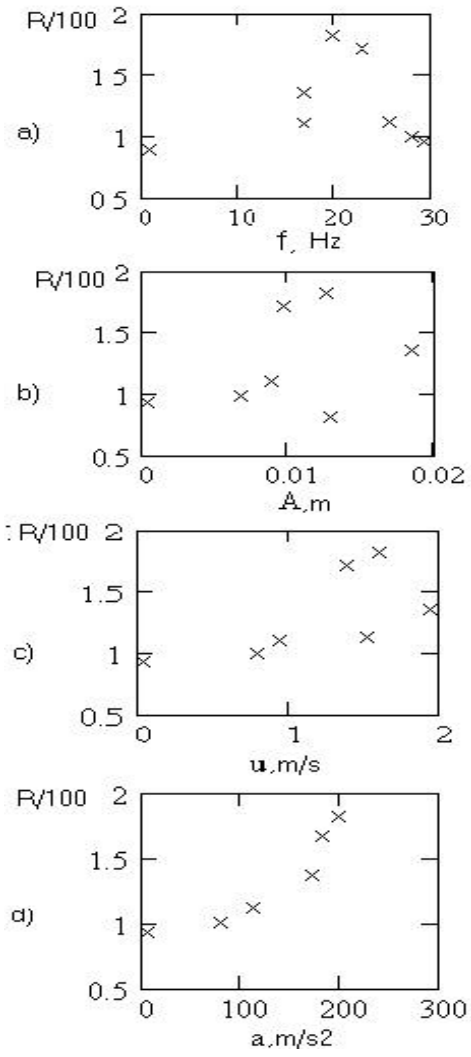


Fig. 4. Thermal resistance of a heat pipe as function of parameters of vibration

Analyzing these data, it is possible to conclude, that the independent parameters determining a mode of operation of a vibrating heat pipe are frequency and vibrating acceleration. Thus, influence of vibrating acceleration is shown only in some narrow enough range of frequencies, which is various for pipes of a different design. So for a similar heat pipe with diameter 6 mm the characteristic frequency has made 110 Hz, and for a pipe with diameter 3 mm has made near 800-1000 Hz.

Such type of dependence of a mode of operation of a heat pipe causes the assumption of the resonant phenomena. The nature of this resonance can not be connected to mechanical fluctuations of a design of a thermal pipe or acoustic waves in a steam phase of the heat-carrier. A liquid phase of the heat-carrier in capillaries is a unique probable element for which such own frequencies can be characteristic.

Dependence of thermal resistance from vibration acceleration appeared rather complex and ambiguous. At small-transmitted capacities the increase of vibrating acceleration resulted in growth of thermal resistance of a thermal pipe. On the contrary, at the big heat demands, amplification of vibration reduced thermal resistance. Value of thermal resistance generally grew with growth of transmitted capacity. But depending on a level of thermal capacity and frequency of vibration it could be as higher, as below thermal resistance, which was registered for a not vibrated heat pipe (fig. 2).

DISCUSSION

The possible mechanism of influence of vibrations on heat and mass transfer processes with boiling in thermosyphons was presented in [1, 3, 4]. In the present report we shall not repeat its interpretation, referring the reader to our published works. This mechanism is consisted of account of change of a angle of wetting owing to moving of a solid wall relatively of vapour bubbles. In a heat pipe unlike thermosyphon as the evaporation occurs not in a bubble, but on a surface of a meniscus. Therefore as a first approximation this mechanism can be used for an estimate of influence of vibrations on heat and mass transfer processes in heat pipes. But it is necessary to expect, that this mechanism is added by change of a surface of a meniscus owing to movement of a solid wall along an axis "meniscus - fluid". The change of quantity of an evaporating fluid from a surface of a meniscus of capillary - porous systems under effect of vibrations is determined by the relevant solution of a hydrodynamic problem. The law of oscillations determines the motion of a solid surface concerning the interphase boundary of a meniscus

$$= A \cdot \sin(wt)$$

that is

$$w_w = A \cdot w \cdot \cos(wt)$$

The solid rough wall, driving in one side, catches a fluid at the expense of watering. The direction of vibration acceleration (in comparison with a force direction of gravity) determines value of "parachute extraction", which increases at the growing of evaporating surface.

Unwatering a part of a surface takes place at motion in the opposite side. Thus, the surface of a meniscus is distorted depending on values of vibration acceleration, roughness of a solid wall, viscosity of a fluid, angle of wetting, surface tension, force direction of gravity. Such quality mechanism of influence of vibrations follows from the analysis of the experimental data on influence of value of amplitude of vibrations on change of efficiency of heat-transfer properties of heat pipe. Let's consider a problem in common statement. Under the assumption of small thickness by a carried away wall of a fluid at tip sections of a meniscus, the flow of a fluid is flat and covered by equation of the Navier-Stokes:

$$n\,\partial^2 w/\partial y^2 = \partial w/\partial t \qquad (1)$$

where w is speed of displacing layer of liquid, n is kinematical viscosity, t is time, y is coordinate.

The boundary conditions of a considered problem are recorded in such view:

at $t=0$: $w=0$ for $0 \le y \le \infty$;

at $t>0$: $w = w_R = Aw\cos(wt)$ for $y=0$;

at $t>0$: $w=0$ for $y=\infty$.

The solution of an equation (1) can be found through a Laplace transformation by analogy with the solution E.Dwyer:

$$w(y,t) = \frac{Ayw}{2\sqrt{pn}} \int_0^t \frac{\cos(w\cdot t)\cdot \exp\left[-y^2/4n(t-t)\right]}{\sqrt{(t-t)^3}} dt \qquad (2)$$

where t is a integration variable.

Following Cooper M.G. -Lloyd A.J.P. [8], we shall enter an allowance about a flow continuity of a mass of a fluid which is flowing past through a boundary layer, that is we use a requirement of equality of rates of flow on the right and at the left for a plane problem:

$$d_0 w_w = \int_0^\infty w(y,t)\,dy \qquad (3)$$

where d_0 is thickness of carried away layer.

The left-hand part of this equation represents a rate of flow through microlayer, carried away by a moving wall, with velocity of a wall w_w (without slipping). The right part represents a rate of liquid which is running in this microlayer. The value d according to the theory of a boundary layer is equale displacement thickness of a boundary layer J_w and it determines thickness of microlayer at the moment of its formation. Parameter J_w is determined by a velocity profile $w(y)$ at flow of a semi-infinite mass of a fluid near a wall, runing on with rate w_w. Let's rewrite (3) as follows

$$d = (1/w_w)\int_0^\infty w\,dy \qquad (4)$$

Substituting here by solution (2) we obtain a ratio

$$d = \frac{Aw}{2w_w\sqrt{pn}} \int_0^\infty y \int_0^t \frac{\cos(w\cdot t)\exp\left[-y^2/4n(t-t)\right]}{\sqrt{(t-t)^3}} dt\,dy \qquad (5)$$

Calculation of this integral does not represent the special difficulties. Final relation (5) is resulted in a following view

$$d = \frac{Aw}{w_w}\sqrt{n/p}\int_0^t \frac{\cos(w\cdot t)}{\sqrt{t-t}}dt \qquad (6)$$

We enter a new variable $z = t-t$. Then (6) is rewritten as follows.

$$d = \frac{Aw}{w_w}\sqrt{n/p}\int_0^t (\cos(w\cdot t)\cos(w\cdot z)/\sqrt{z} + \\ + \sin(w\cdot t)\sin(w\cdot z)/\sqrt{z})dz \qquad (7)$$

This integral can be divides into two parts

$$d = \frac{Aw}{w_w}\sqrt{n/p}\cdot\cos(wt)\int_t^0 \frac{\cos(w\cdot z)}{\sqrt{w\cdot z}}d(w\cdot z) + \\ + \frac{Aw}{w_w}\sqrt{n/p}\cdot\sin(wt)\int_t^0 \frac{\sin(w\cdot z)}{\sqrt{w\cdot z}}d(w\cdot z) \qquad (8)$$

This integral is reduced to a sine and cosine to an Frenel' integral (or can be computed approximately after expansion trigonometric function into a series)

$$d \approx \frac{2A\sqrt{wt}}{w_w}\sqrt{2n\left[1-(wt)^2\right]}$$

The surface of a meniscus grows, if the liquid film does not come back in an initial position. lt is possible to find the change of a watering angle and surface of a meniscus from geometrical constructions, if value of d is known.

Approximately efficiency of vibration action is determined by the formula

$$h = 1 + \frac{A\cos^2 q}{r_{\tilde{n}}(2 - q/45)}$$

where r_c is a capillary radius; q is a watering angle, in degrees.

CONCLUSION

Thus, in the submitted work on the basis of the analysis of experimental data the consistent physical model of influence of vibration for work of a thermal pipe is offered. At the construction of this model the classical inverse problem was solved: the most rational explanation of behaviour of complex heat and mass transfer systems under influence of vibrations was chosen from set of possible.

Unfortunately, the physical model given here, does not describe all features of registered experimental results. In particular, this model does not give the answer to a question about the mechanism of influence of transmitted capacity on a sign of effect of influence of vibrations. Therefore, the calculated model for definition of thermal resistance and the maximal transmitted capacity of thermal pipes, which are subject to vibration, can be constructed only after accumulation of a plenty of experimental data.

As it is visible, the problem of definition of influence of vibrations on heat and mass transfer of heat pipe and thermosyphon very composite and requires the further study both theoretically and experimentally.

REFERENCES

1. Prisniakov K., Prisniakov V. “The Impact of the Vibration on Heat and Mass Transfer in the oscillating Heat Pipes and Two- Phase Thermosyphons”, Proc. 11th Intern. Heat Pipes Conference. - Tokyo.- 1999.

2. Prisniakov K. “Vibrations in Space Propulsion Systems: + and –“.-.Rapport IAF-00-S.1.06 // Proc. of 51st International Astronautical Congress. - Rio de Janeiro (Brasil)- 2000.- Abstracts Book. P. 69.

3. Prisniakov K., Prisniakov V. “Mathematical Modelling of Thermal-Hydraulic Characteristics of Boiling Process in Thermosyphons” // Proc. of the Sixth International Heat Pipe Symposium “6IHPS– 2000”.- Chiang Mai (Thailand). 2000.

4. Prisniakov K., Prisniakov V. ‘The Operating Conditions of Heat Pipes under Vibration Actions” // Proc. of the Sixth International Heat Pipe Symposium “6IHPS– 2000”.- Chiang Mai (Thailand). - 2000.

5. Prisniakov K., Prisniakov V. Impact of vibration on heat and mass transfer under boiling. –Journal of Engineering Physics and thrmophysics. V. 74, No. 4. 2001, p.128-133.

6. Prisniakov V, Bondarenko S, Luzenko V, Prisniakov K. And other. “Heat and Mass Transfer. Odessa. Neptun – Technology. 2001.208 p.

7. Dwyer E. “Boiling Liquid-Metal Heat Transfer”. American Nuclear Society. 1976.

8. Prisniakov V. Boiling. –Kiev: Naukova dumka. 1988.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

Inverse finite element technique for identification of thermal resistance of gas-gap between the ingot and mould in continuous casting of metals

Aleksander Nawrat
Institute of Mathematics
Silesian Technical University
Kaszubska 23, 44-100 Gliwice,Poland
olek@zeus.polsl.gliwice.pl

Janusz Skorek
Institute of Thermal Technology
Silesian Technical University
Konarskiego 22, 44-100 Gliwice,Poland
skorek@itc.ise.polsl.gliwice.pl

ABSTRACT

Identification of thermal resistance of the gas-gap between the ingot and mould during continuous casting of metals is the subject of presented paper. Analysed parameter estimation problem of the steady-state heat conduction belongs to the group of inverse problems. Presented work shows the possibility of applying of the least square adjustment method with *a priori data* for identification of thermal resistance as well as the interphase location and the temperature field within the ingot and mould. The most important feature of the approach is that the unknown thermal resistance of the is obtained from the temperature measurements at the number of sensor located in the wall of the mould by solving appropriate inverse problem. The validity of the solution of the inverse problem is checked by comparison with the results of direct problem. In the present work, a front-tracking method with an automatic mesh generation finite element technique (so called deforming or moving finite-elements) for steady-state equation conduction-advection problems is developed. Such an approach makes possible to find exact solid-liquid interface location, because it is integral part of the solution of the considered problem. Due to the strong convective nature of the boundary problem (casting velocity, thermal parameters) the special up-wind technique is applied.

INTRODUCTION

One of the most important technological parameters, which influence the process of solidification of the metal in the ingot, is thermal resistance of the gap between the mould and ingot. The inverse technique is proposed by the authors for identification of thermal resistance on the base of temperature measurements within the wall of the mould. The process of continuous casting is analysed.

Numbers of theoretical models and experiments have been evaluated to describe thermal resistance in processes of continuous casting [7], [5]. Those models should however deal with a very complex phenomenon's taking place within the ingot. We can mention here phase change, solidification shrinkage, thermal contraction etc. It makes the problem very difficult to be described and solved. Experiments, for example measurements of temperature in the cast and in the mould cannot be directly used to describe thermal resistance but can be used for validation of the mathematical models.

Thermal resistance between the cast and mould can be identified using inverse technique. Presented work shows the possibility of applying of the least squares adjustment method within *a priori data* for identification of thermal resistance. To apply the method the mathematical model of the process has to be evaluated. Mathematical models of the physical processes usually involve several input data, which directly or indirectly come from the measurements. In the situation when mathematical models of the process is evaluated and also certain quantities, which appear in the model, are known the surplus of information arises. That surplus of information can be used for evaluating the most likelihood values of unknown quantities and measurements. It is due to the fact that either mathematical model or measurements are not absolutely exact. Inaccuracies of the mathematical models usually result from the incomplete knowledge about the process. Very often mathematical difficulties force to simplify the model. Inaccuracies of measurements generally result from the accuracy of measurement. Surplus of information gives the opportunity to improve accuracy of unknown

quantities and parameters comparing to the situation when minimum number of information concerning analysed process is available. The problem of calculation the most likelihood estimates of the sough for quantities (in the sense of given criterion) is often called co-ordination technique. The most effective and widely used co-ordination technique is *the least square adjustment method LSAM* that belongs to the group statistically optimal estimation methods. The least square adjustment technique refers to two groups of quantities: unknown (which are not measured) and measured. These quantities are interrelated by the equations of mathematical model (so called constraint equations). In contrary to the classical algebraic problems all the quantities are here treated as stochastic. Essential aim of calculations is to evaluate the most likelihood estimates of unknown and measured quantities.

In this paper problem of identification of thermal resistance between the mould and ingot is solved using least squares approach with *a priori data*.

FORMULATION OF THE PROBLEM

First stage of formulation of boundary problem is to evaluate mathematical model of transient temperature field within the ingot and mould.

The Finite Element Model was adopted for solving analysed Stefan problem under following assumptions:

1. the cylindrical geometry of the calculation domain (ingot and mould) is considered,
2. the problem is steady state in the co-ordinate system attached to the mould,
3. the temperature distribution in the ingot and mould are axially symmetrical,
4. the phase change occurs at a constant temperature,
5. the convection in liquid metal is neglected,
6. the velocity of the ingot is constant,
7. the material parameters, like thermal conductivity or heat capacity are temperature independent but different in liquid and solid phases,
8. the heat is transferred through the gap between the ingot and mould.

The geometry of the problem is shown in Fig.1.

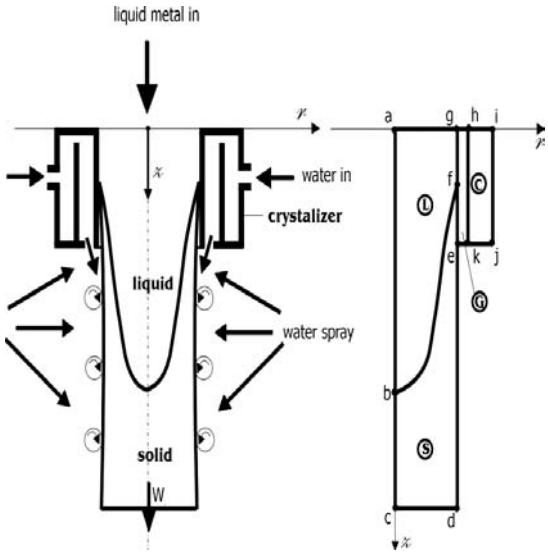


Fig.1 Geometry of the problem

Metal is moving at the constant velocity w_z through a region Ω. The region Ω is defined as a control volume, where $\Omega=\Omega_L \cup \Omega_S$ and $\Omega_L \cap \Omega_S=0$. Where Ω_L - liquid domain, Ω_S - solid domain, W_C - mould domain, Ω_G - gas gap between the mould and ingot.

Solidification of metal takes place within the region Ω. The governing equation in the Eulerian frame (the co-ordinate system is attached to the mould) without heat sources has a form [1],[4]:

$$\underset{i=L,S,C,G}{\forall} \quad \underset{(r,z)\in\Omega_i}{\forall} \; -r\,\rho_i c_{pi} w_{zi} \frac{\partial T^i}{\partial z} + \frac{\partial}{\partial r}\left(r\lambda_i \frac{\partial T^i}{\partial r}\right) + \frac{\partial}{\partial z}\left(r\lambda_i \frac{\partial T^i}{\partial z}\right) = 0 \tag{1}$$

where: $w_{zL}=w_{zS}=w_z \wedge w_{zC}=w_{zG}=0$, r - radial co-ordinate, z - axial co-ordinate, $T(r, z)$ - is a temperature at spatial point (r, z), ρ_i - density, c_{pi}- specific heat, λ_i- heat conductivity, $i = L, S$ or C indicate the solid or liquid phases, respectively, w_z- is the casting velocity.

The temperature of the solid/liquid interface, is known:

$$\underset{(r,z)\in\partial\Omega_{bf}}{\forall} \quad T^S(r,z) = T^L(r,z) = T_m, \tag{2}$$

where T_m is the solidification temperature and $\sum\Omega_{bf}$ indicates the solid/liquid interface. In the Lagrangian co-ordinate system (the co-ordinate system is attached to the ingot), the energy

balance at the solid/liquid interface can be expressed as [1], [4](Stefan condition):

$$\underset{(r,z)\in\partial\Omega_{bf}}{\forall} \lambda_S \frac{\partial T^S}{\partial n} - \lambda_L \frac{\partial T^L}{\partial n} = \rho_S \chi (\vec{w} \circ \vec{n}), \tag{3}$$

where χ - is the latent heat of fusion, $\vec{w}$ - is the velocity vector of the solid/liquid interface in the Lagrangian frame, $\vec{n}$ - is a unitary vector.

The set of the most important external boundary have a form:

$$\underset{(r,z)\in\partial\Omega_{ab}}{\forall} \lambda_L \frac{\partial T^L}{\partial r}\bigg|_{r=0} = 0, \qquad \underset{(r,z)\in\partial\Omega_{bc}}{\forall} \lambda_S \frac{\partial T^S}{\partial r}\bigg|_{r=0} = 0,$$

$$\underset{(r,z)\in\partial\Omega_{cd}}{\forall} T^S(r,z)\big|_{z=Hz} = T_k,$$

$$\underset{(r,z)\in\partial\Omega_{de}}{\forall} \lambda_L \cdot \frac{\partial T^L}{\partial r}\bigg|_{r=Hr} = \alpha_1 \left[T^S(r,z) - T^\infty\right],$$

$$\underset{(r,z)\in\partial\Omega_{ij}}{\forall} \lambda_C \frac{\partial T^C}{\partial r}\bigg|_{r=HCr} = \alpha_2 \left[T^C(r,z) - T^\infty\right],$$

where T_0 - specified temperature at the upper level of the mould (casting temperature of the liquid metal), T_k - specified final temperature of the solidified (and cooled) metal, T^∞ - ambient temperature, a_I - heat transfer coefficient *(i=1,2)*, $Hz = \left|\vec{ac}\right|$, $HCr = \left|\vec{ai}\right|$, $HCz = \left|\vec{ij}\right|$. Indexes *S, L* and *C* refer to the solid phase, liquid phase and mould domain respectively.

FINITE ELEMENT FORMULATION

To solve the boundary Stefan problem the finite element approach has been used. Using the Galerkin method [3], [6] we obtain the algebraic system of linear equations of the following form

$$\mathbf{A} \cdot \mathbf{T} - \mathbf{Q} = \mathbf{0}, \tag{4}$$

where **A** is stiffness matrix, **T** is a one-column matrix of unknown temperatures , and **Q** is a one-column matrix of boundary conditions.

To verify the solid/liquid interface position, we have to check the energy balance at the solid/liquid interface (3). In a steady state condition and in the fixed frame (Eulerian) the solid/liquid interface is fixed but the body is moving at the constant velocity w_z. Notice that in the Eulerian frame material movement in the tangential direction doesn't affect the position of the interface. So we can rewrite the energy balance (3) to the following form:

$$\underset{(r,z)\in\partial\Omega_{bf}}{\forall} \lambda_S \frac{\partial T^S}{\partial n} - \lambda_L \frac{\partial T^L}{\partial n} = \rho_S \chi w_z. \tag{5}$$

After integrating over an element boundary and without loosing the generality we will take under consideration only the boundary over a single element because $\rho_S \chi w_z$ is a constant value:

$$\underset{(r,z)\in\partial\Omega_{bf}}{\forall} \int_{\Gamma^e} \left(\lambda_S \frac{\partial T^{eS}}{\partial n} - \lambda_L \frac{\partial T^{eL}}{\partial n} \right) d\Gamma = \rho_S \chi w_z \int_{\Gamma^e} d\Gamma, \tag{6}$$

or

$$\underset{(r,z)\in\partial\Omega^e_{bf}}{\forall} q_S^e - q_L^e - dS^e = 0, \tag{7}$$

where

$$dS^e := \rho_S \chi w_z \int_{\Gamma^e} d\Gamma = \rho_S \chi w_z h_e. \tag{8}$$

Using Galerkin method with quadratic shape function we can calculate the heat fluxes on the interphase elements. After transformations of the integrals (6) we obtain:

$$q_S^e := \int_{\Gamma^e} \lambda_S \frac{\partial T^{eS}}{\partial n} d\Gamma = \frac{\lambda_S h_e \left(2A_2B_2 + (A_2B_1 - A_1B_2) \log\left(\frac{A_1 + A_2}{A_1 - A_2} \right) \right)}{2A_2^2}, \tag{9}$$

$$q_L^e := \int_{\Gamma^e} \lambda_L \frac{\partial T^L}{\partial n} d\Gamma = \frac{\lambda_L h_e \left(2C_2D_2 + (C_2D_1 - C_1D_2) \log\left(\frac{C_1 + C_2}{C_1 - C_2} \right) \right)}{2C_2^2}, \tag{10}$$

where A_1, A_2, B_1, B_2, C_1, C_2, D_1, D_2 are the coefficients.

The left side of equation (7) we can define as:

$$\Delta q^e := q_S^e - q_L^e - dS^e, \tag{11}$$

and it is the discrepancy of energy balance on the liquid/solid interphase segment. During calculation the equation (7) has to be satisfied with respect to given accuracy ε:

$$\left| q_S^e - q_L^e - dS^e \right| < \varepsilon. \tag{12}$$

Iteration procedure starts from the arbitrary position of the solid/liquid interface defined by the user. Then we obtain the temperature field by solving the linear system of algebraic equations

obtained form (1) using Galerkin method with the isothermal boundary condition on the solid/liquid interface. To check if the position of the interphase is in the right place we search the discrepancy (12) for each finite element located on the interphase. When the condition (12) is not fulfilled then we have to update the interface position. When $\Delta q^e > 0$ then we move the interface location in the positive direction, otherwise in negative direction (Fig. 2).

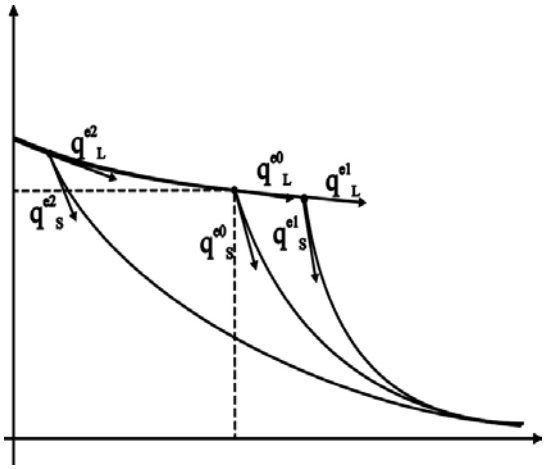


Fig. 2 Analysis of the interface location update

To update the position of the interphase we assume, that the temperature distribution close to the interphase location can be interpolate by a priori given function $T^e(z)$ e.g. polynomial of the shape:

$$T^e(z) = A^e z^2 + B^e z + C^e, \tag{13}$$

which cross the points (T_i^e, z_i^e), $i=a,b,c$. The heat flux in the solid phase is equal to

$$\Delta\widetilde{q}_S^e := \lambda_S \frac{\partial T^e}{\partial z} = \lambda_S\left(2A^e z + B^e\right). \tag{14}$$

Hence, the inequality of the heat flux on the boundary interphase we be expressed as follows:

$$\Delta\widetilde{q}_S^e - \Delta q^e = \lambda_S\left(2A^e\left(z_a^e + dE^e\right) + B^e\right). \tag{15}$$

where dE^e is the interface update on each element from the boundary interface. The value of dE^e is calculated from relationship:

$$dE^e = \frac{\Delta q^e}{2A^e\lambda_S}, \tag{16}$$

where

$$A^e = \frac{T_a^e\left(z_b^e - z_c^e\right) + T_b^e\left(z_c^e - z_a^e\right) + T_c^e\left(z_a^e - z_b^e\right)}{z_a^{e2}\left(z_b^e - z_c^e\right) + z_b^{e2}\left(z_c^e - z_a^e\right) + z_c^{e2}\left(z_a^e - z_b^e\right)}. \tag{17}$$

The next difficulty, which should be overcome, is the convective nature of the analysed problem. When applying the standard Galerkin method, the matrix associated with the convective term is strongly non-symmetric, and we can obtain non-realistic results of calculations. In practice, solutions are often corrupted by spurious node-to-node oscillations. Oscillations are more likely to appear in convection-dominated cases (high Peclet numbers) when a downstream boundary conditions forces rapid change in the solution. It is known that for a Peclet number greater then 1:

$$\underset{i=L,S}{\forall} \quad P_e = \frac{\rho_i c_{pi} w_z Hz}{2\lambda_i} > 1, \tag{18}$$

solution has strong oscillations. To remedy for this problem is to refine the mesh in such a way that the element Peclet number is less then 1. In a finite element framework, up winded convective terms can be developed in several different ways. The initial upwind finite element formulation employs modified weighting functions to achieve the upwind effect. In essence, the element upstream of a node is weighted more heavily than the convective downstream of a node. We use the method proposed in [2], based on modification of the weighting function. In the case when only movement of the ingot is considered in the axial direction z with the constant velocity w_z, we select the upwind parameters as follows:

$$\widetilde{\alpha}_1 = \widetilde{\alpha}_3 = 0; \qquad \widetilde{\alpha}_2 = \widetilde{\alpha}_4 = 1. \tag{19}$$

ALGORITHM OF THE LEAST SQUARES ADJUSTMENT

Numerous physical problems (including numerical representation of boundary problems of heat conduction) can be described in the form of linear matrix constraint equation:

$$\mathbf{Ax} + \mathbf{c} = \mathbf{0}, \tag{20}$$

where **A** are rectangular matrices of coefficients (corresponding to the equation (4)), **x=** [**T**,**λ**] denotes vector of nodal temperatures **T** (few of them are known from measurements) and unknown thermal conductivity of the gap **λ**, **c** is the vector of boundary conditions.

Solution of equation (20) can be derived applying least square approach. Equation (20) is satisfied by the exact values of quantities **x**. When we put into equation (20) results of $\mathbf{x_0}$ measurements and initially estimated values of unknowns we get

$$\mathbf{A}\mathbf{x}^{\mathbf{0}} + \mathbf{c} = \mathbf{w}, \tag{21}$$

where **w** denotes the vector of discrepancies. The aim of the least square adjustment technique is to find such corrections $\boldsymbol{\delta}\tilde{\mathbf{x}}$ which satisfy constraint equation

$$\mathbf{A}\left(\mathbf{x}^{\mathbf{0}} + \boldsymbol{\delta}\tilde{\mathbf{x}}\right) + \mathbf{c} = \mathbf{0}, \tag{22}$$

and minimise quadratic form

$$\boldsymbol{\delta}\mathbf{x}^{\mathrm{T}}\mathbf{V}^{-1}\boldsymbol{\delta}\mathbf{x} = \sum_{j=1}^{s} \delta x_{ii}^{2}\, V_{ii}^{-1} \Rightarrow \min, \tag{23}$$

where **V** is the covariance matrix of the results of measurements.

As the result of solution of an optimisation problem (23) and (22) one obtains following relationship describing estimates of unknown quantities $\tilde{\boldsymbol{\lambda}}$ and covariance matrix of $\mathbf{G}_{\tilde{\lambda}}$?:

$$\tilde{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^{\mathbf{0}} + \boldsymbol{\delta}\tilde{\boldsymbol{\lambda}} = -\left(\mathbf{B}^{\mathrm{T}}\mathbf{F}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{F}^{-1}\mathbf{A}\left(\mathbf{x}^{\mathbf{0}} + \mathbf{C}\right), \tag{24}$$

$$\mathbf{G}_{\tilde{\lambda}} = \left(\mathbf{B}^{\mathrm{T}}\mathbf{F}^{-1}\mathbf{B}\right)^{-1}, \tag{25}$$

where $\mathbf{F} = \mathbf{A}\mathbf{V}\mathbf{A}^{\mathrm{T}}$.

Relationships (24) and (25) are widely used for co-ordination of the wide range of problems including the thermal processes and inverse heat conduction problems [9]. On the other hand, standard least squares adjustment technique often leads to the wrong results when solving inverse boundary heat conduction problems [10]. Accuracy of estimation can be improved by the modification of optimisation procedure by adding stabilizing term. This concept is realized by least squares adjustment method with *a priori data*.

When analysing physical or technical problems it is usually possible to estimate values of sough for unknown quantities $\boldsymbol{\lambda_0}$ and covariance matrix $\mathbf{G}_{\boldsymbol{\lambda}\mathbf{0}}$ of initial estimate of unknown quantities. This *a priori* information lies at the base of formulation of enhanced least squares method. This concept has been successfully used to stabilize solution of inverse boundary heat conduction problems [8].

Basic matrix equation describing a problem (constraint equation) has a form (20) and include unknown quantities **λ** and results of measurement **T**. Similarly to the standard least squares approach we have to find such a corrections $\boldsymbol{\delta}\tilde{\mathbf{T}}$ and $\boldsymbol{\delta}\tilde{\boldsymbol{\lambda}}$ which satisfy constraint equation (22) and minimise certain quadratic form. In the case of proposed approach that quadratic form involves either unknown or measured quantities:

$$\left[\boldsymbol{\delta}\mathbf{T}, \boldsymbol{\delta}\boldsymbol{\lambda}\right]\begin{bmatrix}\mathbf{V} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}_{\boldsymbol{\lambda}\mathbf{0}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\delta}\mathbf{T}\\ \boldsymbol{\delta}\boldsymbol{\lambda}\end{bmatrix} = \boldsymbol{\delta}\mathbf{T}^{\mathrm{T}}\mathbf{V}\boldsymbol{\delta}\mathbf{T} + \boldsymbol{\delta}\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{G}_{\boldsymbol{\lambda}\mathbf{0}}\boldsymbol{\delta}\boldsymbol{\lambda} \Rightarrow \min, \tag{26}$$

where **V** is the covariance matrix of measurements and $\mathbf{G}_{\boldsymbol{\lambda}\mathbf{0}}$ is the covariance matrix of initially estimated unknowns $\boldsymbol{\lambda_0}$.

For LSAM with a priori data matrix $\mathbf{G}_{\boldsymbol{\lambda}\mathbf{0}}^{-1}$ decrease sensitivity of the algorithm on inaccuracies of input data and improve accuracy of the method.

IDENTIFICATION OF THERMAL RESISTANCE BETWEEN THE CAST AND MOULD

Presented method was used for identification of thermal resistance between copper ingot and steel mould. Thermal resistance in the gap between the ingot and mould is strongly dependent from the heat conductivity in the gas-gap. This relation can be expressed in the form:

$$R_i = \frac{H_G}{\lambda_i}, \tag{27}$$

where λ_i - heat conductivity, H_G- thickness of the gas-gap, R_i- thermal resistance of the gas-gap. Since the thickness of the gas-gap H_G is assumed at the mathematical model to be constant, then in the identification analysis will be considered the virtual heat conduction coefficient λ_i on the place of thermal resistance R_i calculated from (27).

An ingot has 1 m length and 0.1 m of radius. The mould (made from steel) has length 0,22 m and its wall is of 0,025 m of thickness. Thermo-physical properties of metal are: thermal conductivity λ_L = 226 W/mK, λ_S = 394 W/mK, specific heat c_L = 475 J/kgK, c_S = 380 J/kgK, density ρ_L = 8300 kg/m³, ρ_S = 8930 kg/m³, the temperature of phase change T_m = 1083 °C, latent heat of solidification χ = 209340 J/kg and the velocity of the ingot w_z = 0,002 m/s. Due to the axial symmetry, only half of the domain is

discretized. Along the surface *ag* and *cd* first kind boundary condition were assumed, i.e. T_0 = 1100 °C and T_k = 20 °C. On the surface *ed, ek, kj, ij* boundary condition of the third kind is assumed, where α_1 = 9000 W/m²K, α_2 = 1500 W/m²K and T^{∞} = 20 °C. Relative accuracy for convergence of calculations calculation is selected as 0.1 %. The values of temperature measurements for identification of thermal resistance of the gas-gap where selected from the solution of direct problem. The size of the gas-gap in the direct problem was calculated on the analysis of the simplified analysis of thermal shrinkage of the solidified part of ingot.

Sample temperature distribution in the ingot and mould is shown in the Fig. 3 and 4.

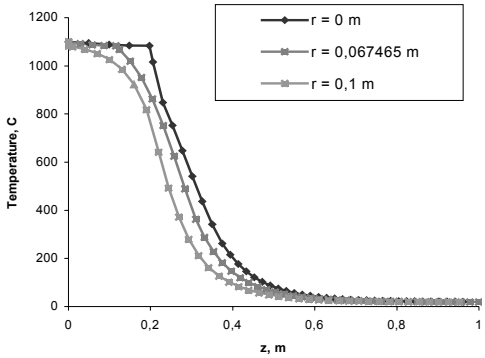


Fig. 3 Temperature distribution in ingot

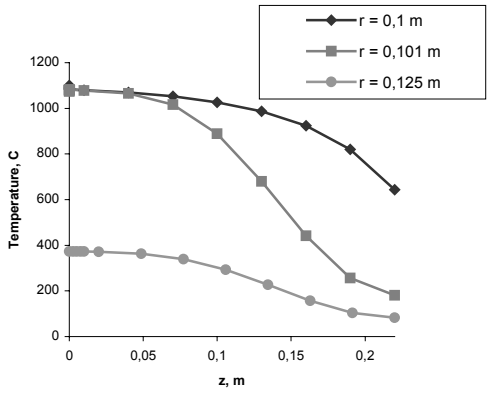


Fig. 4 Temperature distribution in mould

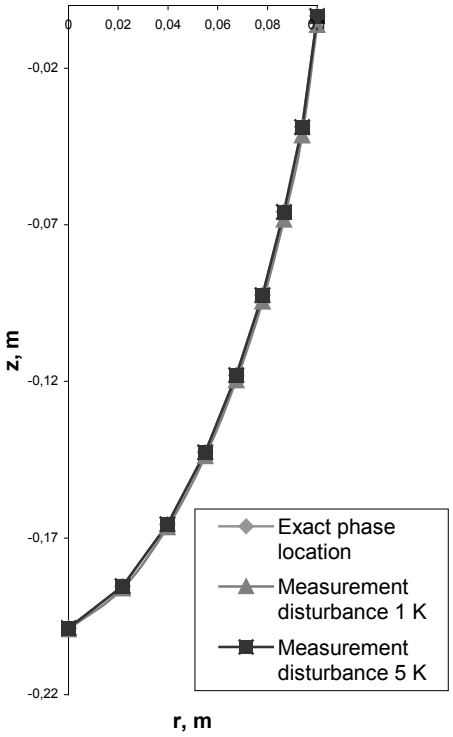


Fig. 5 Phase change location

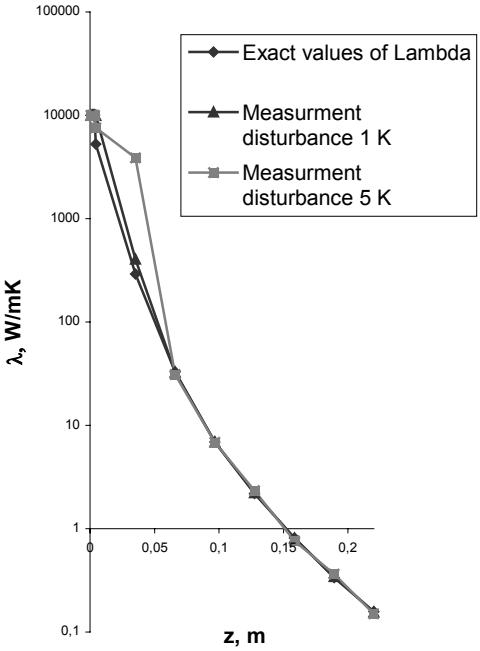


Fig. 6 Thermal resistance identification

Tab. 1 Results of identification

Thermal conductivity λ_i (exact) W/mK	Measurements disturbance			
	δ=1 K		δ=5 K	
	Identified λ_i	Identification relative error, %	Identified λ_i	Identification relative error, %
5264,39	9954,29	50,01	4802,42	65,90
291,92	405,21	50,29	11626,96	25,86
32,83	32,57	5,44	30,35	26,49
6,92	6,92	1,95	6,83	9,94
2,19	2,22	1,48	2,32	7,20
0,81	0,80	1,76	0,76	9,29
0,33	0,34	2,23	0,36	10,62
0,15	0,15	1,92	0,15	10,02

On the basis of the sensitive coefficient analysis were observed that the major influence on the identified values have the temperature measurements in the mould. Measurements points where located at the nodes of the finite element mesh in mould.

Since the first four elements in the mould have direct contact with the liquid metal it have been assumed that the thermal resistance of this elements is equal to zero (therefore corresponding thermal conductivity tends to infinity on this elements). As size of these elements is close to 10^{-10} m so this assumption require physical aspect of the phenomena.

In this paper the measurement results were simulated on the basis of the solution of direct problem ("exact" solution) corresponding to the inverse problem. Inaccuracy of measurements was simulated by adding to the exact solution the random disturbance:

$$T_i^0 = T_i + 2 \cdot (0{,}5 - \xi) \cdot \delta \,, \qquad (28)$$

where

T_i^0 - simulated measurement result,

T_i - exact value (in the sense of the solution of direct problem),

δ - maximal measurement disturbance,

ξ - random number form the range [0,1].

To check the influences of measurement error on the accuracy of identification, following maximal measurement discrepancy were chosen: δ = 1 K and δ = 5 K.

The results of the identification are shown in the Fig. 6 and absolute error of identification is given in the Tab. 1. It can be seen that differences between exact and identified values are relatively low and appears just at the beginning of the gas-gap. The interface location is presented in the Fig. 5. It is easy to see that for the analysed measurement discrepancy δ the difference between exact solution and results of identification is not observed.

CONCLUSIONS

Inverse problem was formulated and algorithm of least square adjustment method with a priori data was used for identification of thermal resistance between ingot and mould during permanent casting of metals. Calculation and numerical tests have proved that proposed method can be effectively used for solving inverse boundary heat conduction Stefan problem. Important future of least squares adjustment method is the possibility of direct statistical estimation of results. It is important advantage because in practice there is usually no possibility to verify accuracy of calculations. Result of the present inverse analysis can be used for optimisation of the process of casting metals.

ACKNOWLEDGMENTS

The financial assistance of the National Committee for Fundamental Research, Poland, within the grant no. 4T11A 012 22 and no. 8T10B 069 21 is gratefully acknowledged herewith.

REFERENCES

[1] Carslaw, H. J. and Jäger, J. C. ,*Conduction of heat in solids'*, Oxford University Press, Oxford, (1959).

[2] Heinrich, J.C. , Zienkiewicz, O. C., *Quadratic finite element schemes for two-dimensional convective - transport problems*, Computer Methods in Applied Mechanics & Engineering, vol. 11, pp. 1831-1844, North Holland, 1977.

[3] Kikuchi, N., *Finite element methods in mechanics*, Cambridge University Press, Cambridge, (1987).

[4] Özisik, M.N. , *Heat Conduction*, Wiley, New York, (1980).

[5] Prinz B., *Measuring and Modelling of the Influence of Coatings on the Heat Transfer*

between Melt/Casting and Mould during the Permanent Mould Casting of Aluminium Alloys, COST 504/III Report, Aachen, Germany, 1994.

[6] Reddy, J. N., *An Introduction to The Finite Element Method*, McGraw-Hill, USA, (1993).

[7] Schmidt P., *Heat transfer in permanent mould casting*, Proceedings of the COST 504 Conference Advanced Casting and Solidification Technology, 1994.

[8] Skorek J., *Applying stochastic method for solving inverse heat conduction problems* (in polish), Proceedings of the Symposium on Eat and Mass transfer, Warsaw, 1983.

[9] Szargut J., *et al*, *The least square adjustment technique in thermal engineering*, (in Polish), PAN, Wrocław, 1984.

[10] Szargut J., Skorek J., *Influence of the Preliminary Estimation of Unknowns on the Results of Co-ordination of Substances and Energy Balances*, Bulletin, of Polish Academy of Sciences, Thermodynamics, vol. 39, No. 2, p. 335, 1991.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

AN INVERSE APPROACH FOR PIPING NETWORKS MONITORING

Antonio C. Caputo **Pacifico M. Pelagagge**
Faculty of Engineering, University of L'Aquila, Monteluco, L'Aquila, Italy
caputo@ing.univaq.it; pelmar@ing.univaq.it

ABSTRACT

In fluid distribution the problem of monitoring the network status in order to identify abnormal conditions and locate leakages arises. In the paper an inverse approach resorting to a multi-layer perceptron back-propagation Artificial Neural Network (ANN) is proposed in order to locate leakages based on pressure and flow rate information. Strategies for generating input data and for correlating by ANN such data to the fluid distribution system status are presented. A two-level architecture is selected, composed by a main ANN at the first level and several branch-specific second-level ANNs in cascade to the main one. The branch in which the leakage occurs is identified resorting to the ANN operating at the first level while the specific second-level ANN is activated to accurately estimate the magnitude and location of the leakage in the selected branch.

NOMENCLATURE

D	Diameter (m)	*Subscripts*	
f	Friction factor	i	i-th node
k	Loss coefficient (kg/m^7)	j	j-th node
L	Length	F	Fanning
p	Pressure	*Superscripts*	
q	Volumetric flow rate	in	Inlet
Re	Reynolds number	ex	Outlet
v	Velocity	*Symbols*	
		ε	Roughness
		Δp	Pressure loss (Pa)
		ρ	Density (kg/m^3)

INTRODUCTION

Piping networks have found wide diffusion in industrial and civil applications where fluids have to be distributed to several consumers or processes. In order to carry out such task in an effective and safe manner the occurrence of unauthorized withdrawal, spills and leakages should be identified and promptly eliminated. Therefore, the problem of monitoring the network status in order to identify abnormal conditions and locate leakages arises. In the literature, despite the significant efforts made, an effective solution to the networks monitoring problem has not been accomplished and it still presents a significant challenge.

Current technologies for fault detection and identification (FDI) are based on methods which are both dependent or not-dependent on process parameters. Not-dependent technologies include:

- injection of tracing substances in the fluid stream;
- analysis of acoustic emissions from leaks;
- pressure waves transmission;
- visual inspection;
- detection of leak-generated temperature variations resorting to infrared thermography;
- identification of radar or radio frequencies emitted from transmitters located inside the pipes and permeating outside through pipe cracks.

At the moment, however, such classes of techniques are expensive, often intrusive and perturbating the whole fluid network, or difficult to monitor by telemetry.

Methods dependent on process parameters, instead, are based on establishing mass and energy balances at different network nodes using measured values of flow rates (q) and pressure levels (p). However, it is well known that computation of the network boundary conditions starting from measured q and p values at the moment presents mathematical difficulties that are far from being solved. On the contrary evaluation of p and q at each node of the network is straightforward when boundary conditions are known. Therefore the FDI in piping networks based on process parameters results in a classical inverse engineering problem, that is determination of causes (leaks) starting from knowledge of their effects (p-q maps) [1].

Control of fluid distribution networks requires monitoring of the current state of the system. Although telemetry systems represent a promising solution for evaluating state variables in the network (pressure and flowrates), their

widespread diffusion is impaired by the high costs. State estimation techniques are therefore utilized to ascertain system status from a limited number of information. However, both metering and telemetry systems are subject to errors having a negative impact on the accuracy of the state estimation calculation. Results may be therefore very inaccurate when compared to the actual system state. Moreover, the analytical techniques adopted are rather complex and computationally intensive. Therefore, a robust, efficient and low time consuming means to evaluate the entire fluid network status starting from a limited amount of state knowledge is required.

In this paper at first a recursive mathematical model of a generic network has been presented, enabling the direct computation of the system status in any condition. The inverse problem is successively stated discussing difficulties involved in its analytical solution, also presenting the available techniques to correlate state variables during perturbated operations of the network. A solution approach based on the utilization of an ANN is then presented examining possible ANN architectures and selecting a multi-layer perceptron back-propagation network. The strategy for coding of input and output data with reference to the real system and the utilization of this method are discussed. A two-level architecture is selected composed by a main ANN at the first level and several branch-specific second-level ANN in cascade to the main one. The branch in which the leakage occurs is identified resorting to the global ANN operating at the first level while the specific second-level ANN is activated to precisely estimate the magnitude and location of the leakage in the selected branch.

LITERATURE REVIEW

A considerable amount of research effort has focused on the identification of leaks in fluid distribution networks addressing either the compressible or incompressible fluid flow cases. As already mentioned available techniques rely either on process-dependent or process-independent variables. Silk and Carter [2] presented an in depth review of available techniques related to process-independent methods. As far as process-dependent methods are concerned, the most common is the line volume balance technique utilizing data supplied from flow rate meters located at the extremity of each branch. Usually such approaches prevent the exact pinpointing of the leak location. Ellul [3] proposed a method in which measured pressure and flow rate data are compared with those computed on the basis of a mathematical model of the system and the leak amount is made proportional to the obtained difference. Leaks lower than 5% may be discovered with processing times in the order of minutes. Stouffs and Giot [4] present methods based on mass balances able to treat transient states of compressible fluids highlighting how the practical detectable limit is 2-3% for compressible fluids. Billman and Isermann [5] propose an adaptive non-linear observation method of the pipeline dynamic behaviour, along with a special correlation technique, based on the measurement of pressure and flow rate data at the inlet and exit of a branch. Leaks around 2% for liquids and 10% for gases have been detected. Siebert [6] developed a method based on the statistical analysis (cross-correlation) of signals sampled over a time interval of 1.7 s, being able to locate leaks in the order of 0.2% for liquids, while for gases the value of 5% was reached. Also Zhang [7] adopted statistical signal analysis techniques to address the problem, detecting 1% leaks on a 37 km pipeline. Whang et al. [8] developed a method able to identify leaks of 0.5% in 120 m long water pipes in negligible time delay utilizing an autoregressive modeling approach requiring only 4 pressure measurements with a sampling time of 20 ms. Hamande et al. [9] tested a leak detection system on an actual ethylene pipeline operated since 1989. The system relied on 21 pressure measurement stations and 2 flowrate meters associated with a mathematical model of the pipeline, enabling detection of 7% leaks in a 60 minutes time span. Parry et al. [10] describe another system adopted for LPG pipelines interfaced with a measurement instrumentation enabling detection of 2% leaks over times ranging from 46 minutes to 9 hours depending on the distance of the leak from the nearest measuring station. Another leak detection system for pipelines was developed by using artificial neural networks for leak sizing and location [11]. This system can detect and locate leaks down to 1% of flow rate in pipelines in about 100 s with a probability of success that is greater than 50% for the smallest leak. In the paper specific signal processing techniques able to reject spurious alarms due to compressibility effects subsequent to operational transients are developed. However, this method too involves leaks identification in

single–branch pipelines only. Generally speaking, besides such approaches, usually aimed at long pipelines or single branches of a network, in case of inverse problem solution, calculation of all pressures and flow rates in a fluid distribution system may be accomplished by formulating and solving the mass and/or energy conservation equations. Such balances have to be consistent with the measurements carried out in the system provided that there are as many independent equations as there are variables that are to be calculated [12]. However, neglecting numerical difficulties, the analytical solution to such equations may become impossible if measurements are corrupted by noise or are unavailable as often happens in practice. State estimation techniques may overcome this problem [13] but introducing a much higher computational effort, even if utilization of sparsity exploiting techniques and numerically stable factorisation besides parallel and distributed computation [13,14,15,16] may leviate this difficulty. In recent times artificial intelligence techniques, in particular Neural Networks, have been proposed to analyze fluid distribution networks in order to exploit their intrinsically parallel structure, leading to high computational efficiency, but also their robustness as global state estimators. The main utilization of ANNs has been by formulating the state estimation problem in terms of analog neural networks and utilize such networks to solve the systems of linear equations involved [17,18,19,20]. Since a water distribution system is described by an overdetermined set of nonlinear equations, Gabrys and Barjela [18,19] utilized ANNs to solve the nonlinear set of differential equations resulting from the application of the Newton Raphson iterative method to the solution of the linearised water network equations. An optimization problem may be then defined which enables to minimize the discrepancies between the actual measurements and the values calculated from the mathematical model of the network itself. To account for uncertainties in measurements some authors [14,18,19,21,22,23] also developed a methodology called Confidence Limit Analysis which allows for constructing mathematical models of the system, taking into account effects of measurement uncertainty on accuracy of the derived state estimates. The problem of leak detection has been mainly faced resorting to analytical techniques based on the analysis of residuals (i.e. errors on the mass balances at the nodes) and the paths linking them [13,24,25]. As far as neural networks are concerned, recently Barjela and Gabrys [20] proposed an approach based on pattern recognition, applied to the calculated state estimates and residuals with their corresponding confidence limits, carried out by a specifically developed General Fuzzy Min-Max (GFMM) neural network for clustering and classification. The adopted neuro-fuzzy recognition approach to leak detection is based on the examination of patterns of state estimates and has been successfully applied to a small but realistic water distribution network and tested using data covering a 24-h period of operation [23]. In this paper, instead, a simplified approach to leak detection and location is suggested using an ANN as a global classifier, among the possible pressure/flow rate patterns, in order to directly correlate state variables values to network states.

THE PROPOSED METHODOLOGY

The methodology is based on two main phases (figure 1):

- evaluation of pressure/flow rate conditions (effects) by simulation, imposing the piping network's boundary conditions (causes);
- correlation of effects with causes by Neural Networks.

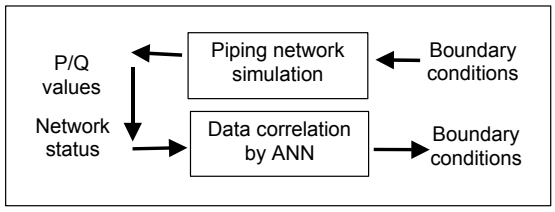


Figure 1. Inverse approach for piping fault detection and identification.

In order to carry out the training step of the ANN the direct problem is solved at first resorting to the mathematical model of the network. In particular, variables at the nodes are computed in case of either normal operation and when artificial leakages are introduced, enabling to determine a set of patterns characterizing the system status in several conditions of normal or abnormal operation. Normal operating data could be in fact obtained on the field but only for existing networks, moreover this is a time consuming task. Furthermore no data are readily available for atypical events (i.e. leakages) or operation in non standard situations while it is not

feasible to obtain such data from physical simulation on the real network. Therefore, a mathematical model of the network has been utilized to predict the status of the network consequent to normal and abnormal flow situations and generate learning data sets. Successively the ANN is trained on these data sets representing the system status (i.e. pressures and flow rates) in different operating conditions either with or without leakages. In this way a robust correlation between leak effects (p-q maps) and location-amount of leaks (causes) may be defined. However, the ANN should be able to reject the noise represented by the inevitable measurement errors affecting the sensors. In order to increase robustness of the ANN, the input data sets may be therefore altered respect the accurate values exactly representing the system status by superimposing random deviation to the pressure drop and the leak flow readings, comprised in the instruments uncertainty range. During the running phase the ANN is fed with actual measurements of pressure and flow rate from a limited number of nodes and performs clustering and classification tasks giving in output the location (branch) and magnitude of the leak.

In more detail the following main steps have to be performed in the proposed methodology:

1. development of a mathematical model of the network for p and q evaluation starting from the actual boundary conditions;
2. simulation of leakages occurring and computation of resulting p and q values in the network (data generation);
3. correlations of leaks patterns to p and q data resorting to ANN.
4. integration of the ANN system in the piping network in order to carry out real-time and on-line monitoring and leak identification tasks based on actual fluid network measured operational status.

The entire process is schematized in figure 2.

Implementation of the ANN-based pattern classification system in particular involves:

a) ANN architecture definition;
b) selection of data structure and ANN input and output layer;
c) ANN training;
d) ANN testing;
e) running phase.

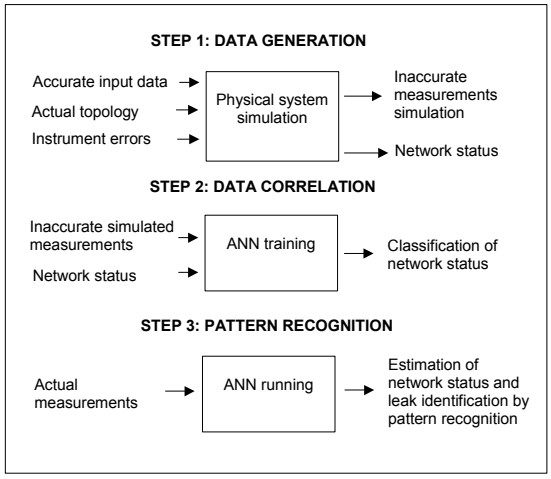


Figure 2. Main steps of the proposed methodology.

ESTIMATION OF THE NETWORK STATUS

Several well proven methods for the solution of a direct fluid distribution network problem are available in the literature [12,26]. In this work an efficient and easy to use algorithm has been developed able to simulate any fluid distribution network of arbitrary topology.

Pressure drop due to friction losses of the flow rate q_{ij} across a generic branch connecting nodes i and j may be expressed as

$$\Delta p_{ij} = k_{ij} (q_{ij})^2 \qquad (1)$$

where k_{ij} is the loss coefficient

$$k_{ij} = \frac{32 f_F \rho L_{ij}}{\pi^2 D_{ij}^5} \qquad (2)$$

The Fanning friction factor f_F, a function of flow regime and wall roughness, has been expressed by the following empirical correlation:

$$f_F = \frac{1}{16\left\{\log\left[\frac{\varepsilon}{3.7D} - \frac{5.2}{Re}\log\left(\frac{\varepsilon}{3.7D} + \frac{14.5}{Re}\right)\right]\right\}^2} \qquad (3)$$

It follows that the branch flow rate is

$$q_{ij} = \frac{p_i - p_j}{k_{ij}|q_{ij}|} = \frac{p_i - p_j}{\sqrt{k_{ij}|p_i - p_j|}} \qquad (4)$$

However, the mass balance asks that

$$\Sigma_i (q_{ij} + q^{in}_i + q^{ex}_i) = 0 \qquad \forall j \qquad (5)$$

meaning that both the mass balances at the nodes

$$\Sigma_i q_{ij} = 0 \qquad \forall j \qquad (6)$$

and the overall external flow rate balances

$$\Sigma_i (q^{in}_i + q^{ex}_i) = 0 \qquad (7)$$

have to be satisfied. Therefore eq. 5 may be rewritten as

$$\Sigma_i \left[\frac{p_i}{\sqrt{k_{ij}|p_i - p_j|}} + q^{in}_i - q^{ex}_i \right] = p_j \Sigma_i \frac{1}{\sqrt{k_{ij}|p_i - p_j|}} \qquad (8)$$

Indicating by ΣI the summatory at left hand member and with ΣII the summatory at right hand member one gets the expression for the pressure at the *j-th* node as

$$p_j = \frac{\Sigma I}{\Sigma II} \qquad (9)$$

The algorithm is as follows:

Initialization phase:

- inputs of tentative inlet flowrate values at each network node q^{in}_i =(0,1), the value is set to 1 to indicate that the node receives an input external flowrate, otherwise it is set to 0;
- input of actual exit flowrate values q^{ex}_i;
- input of tentative values of pressure p_i at nodes. Pressure at the network input nodes are considered to be known and to remain constant, while pressures at any other node is the simulation output;
- definition of network topology. This is carried out resorting to the pressure loss coefficient k_{ij}: it is initially assigned value 1 if nodes i and j are connected by a branch, or value 0 otherwise;
- definition of length L_{ij} and diameter D_{ij} for each network branch.

Network simulation phase

1. branch flow rates q_{ij} are computed by eq. 4;
2. the average fluid velocity in each branch $v_{ij} = q_{ij}/\pi D_{ij}^2/4$ is computed as well as the Reynolds number;
3. the Fanning friction factor is computed (eq.3);
4. values of loss factor k_{ij} are updated through eq. 2;
5. pressures at the nodes p_j are updated through eq. 9;
6. flowrates for the nodes receiving an external flow rate (those where the initial value was q^{in}_i = 1 are also updated through eq. 9 by maintaining constant the assigned p_j value and solving eq. 9 for q^{in}_i;
7. flow rates q_{ij} for the other generic branches are updated by eq. 4.

The entire procedure is iterated until convergence of computed flow rates q_{ij} is obtained.

The algorithm shows fast convergence and enables an easy configuration of the network topology during the data initialization phase. In fact it does not require to define a different set of equations to be solved according to the specific network analysed. Program outputs are pressure values at each nodes and flow rate values at each branch. This algorithm was compared with other traditional solution approaches (the Hardy-Cross iterative method and the solution of a set of equations describing conservation of flow rate at the nodes of the algebraic sum of pressure loss across any pipe loop [26]) obtaining very good agreement as shown in Table 1 for a simple 5-nodes network.

Table 1 Comparison of computation methods.

q^{in}_1 (l/s)			100
q^{ex}_4 (l/s)			50
q^{ex}_5 (l/s)			50
Branch 1-2 Length – Diameter (m)			300 - 0.3
Branch 1-3 Length – Diameter (m)			300 - 0.3
Branch 2-4 Length – Diameter (m)			200 - 0.3
Branch 2-3 Length – Diameter (m)			500 - 0.3
Branch 3-5 Length – Diameter (m)			400 - 0.3
Branch 4-5 Length – Diameter (m)			500 - 0.3
Computed flow rates (l/s)			
Branch	Hardy-Cross	Mass/energy balance	This work
1-2	47.30	47.14	47.17
1-3	52.70	52.85	52.83
2-3	10.10	10.08	10.01
2-4	57.50	57.22	57.18
3-5	42.70	42.77	42.81
4-5	7.30	7.22	7.19

The simulation of a leaking network may be simply carried out by inserting in the desired branch, fictious nodes from where the leakage flowrate spills. The user may specify the leak amount, its location and the interested branch. The program then rearranges automatically the network configuration consequent to the increase of the number of nodes and branches respect the original network and computes the new values for pressures and flow rates. Such values therefore may represent the readings supplied under

degraded operation by measuring instruments distributed on the networks, that shall be analysed to locate the leaks. According to the control method hypothesized, the input flow rate will increase in order to compensate for the leaking flowrate by maintaining the fixed pressure at the input nodes. Besides generating training data for the ANN, the program is a useful tool to explore the amount of pressure perturbations consequent to leakages and assess the required sensitivity of the measuring instruments so that the flow anomaly can be positively detected. Effect of signal noise may also be evaluated. On the contrary given a set of measuring instruments having a specified sensitivity, the minimum leak amount able to be detected may be computed. Optimal placement of sensors along the network may also be studied. Such issues are of great importance for the practical utilization of this method. In fact, the quality of available pressure sensors determines the accuracy that can be obtained in determining the location of the leak along the pipe.

UTILIZATION OF ANN FOR FDI

In order to solve the leak detection and identification problem in a generic fluid distribution network, the use of ANNs is proposed by exploiting their classification abilities. Main goal is to recognize patterns among the pressure and flowrate data supplied by a set of measuring instruments distributed across the piping network and as a consequence select the branch where leak occurs, location along the branch and amount (figure 3). ANN are basically a non linear transfer function applied to a set of input variables. Physically the ANN is composed by a set of elementary computing units, the neurons, connected each other according to the peculiar network structure adopted.

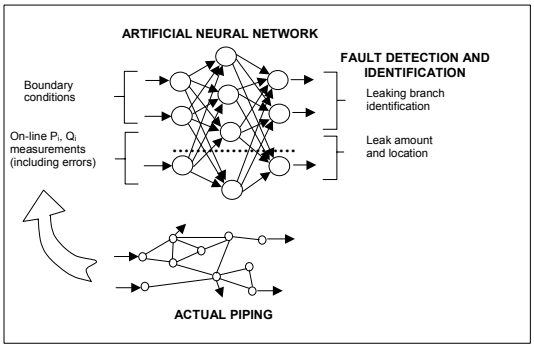


Figure 3. Leak identification through ANN.

Each neuron's output is the value of its non-linear activation function computed on the weighted sum of the input signals to the neuron itself. Assigned to the neurons connections are weights which are generally chosen during the network training phase in order to minimize the error between the output values computed by the ANN and the true values corresponding to the set of input data processed by the ANN. The error function may be considered as a surface in the space of weights leading to the utilization of gradient-based minimization techniques. Computational performances may be enhanced if multiple neuron layers are introduced between the input and output layer. ANN have proven succesful in approximating non linear multivariable functions and in classification tasks. Provided that a sufficiently large training data set is available, ANN may be useful when it is difficult or impossible to obtain a mathematical model of the system to be solved by analytical methods, when new data have to be processed at high speed in real time, and when the computation method should be robust and fairly insensitive to noise in the input data. Topologies of ANN which show best classification abilities are the Probabilistic Neural Network (PNN), the Radial Basis Function (RBF) network and the Multi Layer Perceptron (MLP).

Selection of neural network type

Characteristics of aforementioned ANNs will be briefly discussed in the following in order to make some consideration about the type of network which is more likely to represent the best candidate for this application. However, the final choice will depend on an extensive experimentation and the characteristics of the actual fluid network being studied. A PNN is basically a Bayesian classifier into a neural network architecture. This kind of ANN does not update the internal weights during the training phase. Rather, it stores the training sets and compares them with the actual input during the working phase in order to select the weight set which corresponds to the closest training set respect the actual data input (Nearest Neighbor method) and compute the output accordingly. PNN output in this case is simply the class of the leak, i.e. a code identifying one of the original training set that represent a known and specific status of the network (characterized by a given leak amount and location). PNN generally result the best classifier when inputs very similar to

those utilized for the training phase are utilized. Adding a noise to the input data respect the training data may result in some input cases not correctly classified. In case of incorrect classification the error is likely to be not on the leak amount or position on a branch but on the branch itself, making the computation output useless for the purposes of the study. It follows that the main drawback of the PNN lies in the fact that no estimation of the ANN output uncertainty can be made.Therefore even if the PNN shows perfect correspondence between known inputs and corresponding classes it is rather seriously affected by inevitable disturbances in the sensors reading, making it quite unreliable in this application. In fact a leak or an alteration of sensor signals generating a p/q pattern much different from those the ANN has been trained on, or very similar to the pattern generated by an entirely different leak condition may completely deceive the ANN which could indicate in output a leak class entirely different from the real one. The output layer may consist simply of a neuron indicating a number directly corresponding to one of the classes of training data. However, the required output coding (consisting in the association of the input to a class among those utilized for training) prevents from recognizing readily an erroneous result. RBF networks are made up by two external layers with the hidden layer composed by neurons having gaussian activation function while the output layer is composed by simple neurons carrying out a linear combination of the output of hidden neurons. This type of ANN showed interesting performances during preliminary trials, including significant robustness to noise, when applied to a single branch. In particular the uncertainty in leak location remained the same but the flowrate noise was nearly directly transferred to the new leak estimation value making RBF network even more robust than the MLP. However, it resulted highly unreliable when applied to a fluid distribution network. The ANN could not classify inputs altered by a noise while also in absence of noise it was not possible to minimize the ANN errors which involved indistinctly any component of the output vector. MLP are multilayer networks where the neuron activation function is sigmoidal. The MLP also shows good robustness respect alteration of input values. During the performed test MLP correctly identified the branch where the leak occurred in all the examined cases. Therefore the MLP may represent the best compromise among good previsional capability and good noise rejection. This latter type of ANN offered the best performance compromise and has been adopted in the following of the study.

Information coding

Different input-output coding schemes and number of neurons in the external layers may be utilized when MLP are utilized to correlate state variables in fluid networks. In particular, inputs may be the absolute or relative pressure loss respect the nominal value (better if normalized in the 0 to 1 range) at the nodes where such measurements are available as well as the boundary conditions (flow rate measurement at the inlet nodes and head at the fixed heads nodes). Outputs may be coded as binary 0-1 to identify the leaking branch (if an output neuron is associated to each branch) or as values comprised in the 0 to 1 range if the degree of leak severity (as a predefined percentage of nominal branch flow rate) and the leak location along a branch (as a predefined fraction of the branch length) need to be estimated. The kind of output coding may be selected in order to suit specific requirements, but the preferable input coding is the one based on relative variations of the input variables respect the value at nominal operating conditions.

Overall neural network architecture

MLP is the better suited architecture in this kind of application but a single MLP presents some output coding limitations that may prevent obtaining a good accuracy in real applications. On the other hand, given its proven function approximation capability, on a single branch application the MLP may guarantee a satisfying accuracy and a proper presentation of output data. This leads to the utilization of a hybrid network structure having two levels (figure 4). A first-level MLP is utilized to locate the branch where the leak occurs in order to subsequently activate specific second-level branch-dedicated MLPs having the task of precisely estimating the leak flow rate and location. In this way the complex task of identifying leaks in realistic fluid distribution networks may be subdivided in two sub-problems that may be fairly easily solved separately by dedicated MLP artificial neural networks. In case of unstationary distribution networks where periodic variations of the overall state are observed the entire structure is repeated and possibly selected by another ANN as suggested by Gabrys and Barjela [23].

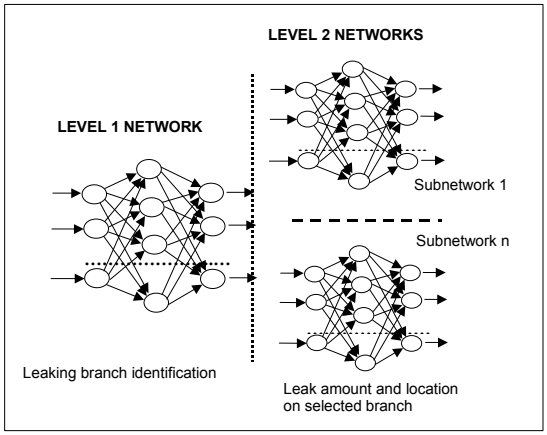


Figure 4. Two-level neural network architecture.

CONCLUSIONS

In this paper the feasibility of utilizing Artificial Neural Networks to locate leaks over a fluid distribution network has been discussed also analysing the relevant literature. A suitable overall two-level ANN architecture is presented where a first-level ANN determines the branch where the leak occurs and a specific second-level ANN estimates leak amount and location. The proposed architecture has been satisfactorily tested in a simplified case obtaining promising results. Details of the implementation phase on a real test case and a discussion of performances will be carried out in a forthcoming paper. As a future research work the complexity of problems which can be solved with this approach will be assessed and the robustness of the ANN to measurement errors of the sensors will be evaluated in order to define the conditions of applicability to actual fluid distribution networks of this method. More specifically some practical issues having economic relevance will be dealt with, such as the determination of the sensors' required number and accuracy level which minimizes the total monitoring system cost maintaining a required level of previsional capability.

REFERENCES

1. Ozisik, M.N., Orlande, H.R.B., 2000, Inverse Heat Transfer: Fundamentals and Applications, Taylor & Francis.

2. Silk, M., Carter, P., 1996, A Review of Means of Pipeline Leak Detection Independent of Flow Measurements, DEPIRE report TEC-T031-01, AEA Technology, United Kingdom.

3. Ellul, I., 1989, Pipeline Leak Detection, The Chemical Engineer, pp. 40-45, june.

4. Stouffs, P., Giot, M., 1993, Pipeline Leak Detection Based on Mass Balance: Importance of the Packing Term, J. Loss Prevention in Process Industry, vol. 6, no. 5, pp. 307-312.

5. Billmann, L., Isermann, R., 1987, Leak Detection Methods for Pipelines, Automatica, v. 23, no. 3, pp. 381-385.

6. Siebert, H., 1981, A Simple Method for Detecting and Locating Small Leaks in Gas Pipelines, Process Automation, pp. 90-95.

7. Zhang, X.J., 1992, Statistical Methods for Detection and Localization of Leaks in Pipelines, OMAE – Volume V-B, Pipeline Technology, ASME.

8. Whang, G., Dong, D., Fang, C., 1993, Leak Detection for Transport Pipelines Based on Autoregressive Modeling, IEEE Transactions on Instrumentation and Measurements, vol. 42, no. 1, pp. 68-70, February.

9. Hamande, A., Condacse, V., Modisette, J., 1995, New System Pinpoints Leaks in Ethylene Pipeline, Pipeline & Gas Journal, pp. 38-41, April.

10. Parry, B., MacTaggart, R., Toerper, C., 1992, Compensated Volume Balance Leak Detection on a Batched LPG Pipeline, OMAE – Volume V-B, pp. 501-507, Pipeline Technology, ASME.

11. Belsito, S., Lombardi, P., Andreussi, P., Barerjce, S., 1998, Leak Detection in Liquified Gas Pipelines by Artificial Neural Network, AIChE Journal, V. 44, n. 12, pp. 2675-2688.

12. Eggener, C.L., Polkowski, L.B., 1976, Network Models and the Impact of Modelling Assumptions, J. of the AWWA, v. 68, n. 4, April.

13. Barjela, A., 1984, On Line Monitoring of Water Distribution Networks, PhD Thesis, University of Durham.

14. Bargiela, A., Hainsworth, G., 1989, Pressure and Flow Uncertainty in Water Systems, ASCE J. of Water Resource Planning and Management, V. 115, n. 2, March, pp. 212-229.

15. Hartley, J.K., Barjela, A., 1995, Parallel Simulation of Large Scale Water Distribution System, Proc. Modelling and Simulation Conf. ESM'95, Prague, ISBN 1-56555-080-3, pp. 723-727.

16. Hosseinzaman, A., 1995, Parallel and Distributed Simulation of Network Systems, PhD Dissertation, The Nottingham Trent University, UK.

17. Chichocki, A., Barjela, A., 1997, Neural Networks for Solving Linear Inequality Systems, Parallel Computing, V. 22, n. 11, pp. 1455-1475.

18. Gabrys, B., Barjela, A., 1995, Neural Simulation of Water Systems for Efficient State Estimation, Proc. Modelling and Simulation Conf. ESM'95, Prague, ISBN 1-56555-080-3, pp. 775-779.

19. Gabrys, B., Barjela, A., 1996, Integrated Neural Based System for State Estimation and Confidence Limit Analysis in Water Networks, Proc. European Simulation Symposium ESS'96, Genoa, ISBN 1-565555-099-4, vol. 2, pp. 398-402.

20. Gabrys, B., 1997, Neural Network Based Decision Support: Modeling and Simulation of Water Distribution Networks, PhD Thesis, The Nottingham Trent University, UK.

21. Gabrys, B., Barjela, A., 1998, Simulation of Water Distribution Systems, Proc. European Simulation Symposium ESS'98, Nottingham, ISBN 1-56555-147-8, pp. 273-277.

22. Gabrys, B., Barjela, A., 1999a, Analysis of Uncertainties in Water Systems Using Neural Networks, Measurement + Control, V. 32, n. 5, pp. 145-147.

23. Gabrys, B., Barjela, A., 1999b, Neural Networks Based Decision Support in Presence of Uncertainties, ASCE J. of Water Resource Planning and Management, V. 125, n. 5, September/October, pp. 272-280.

24. Carpentier, P., Cohen, G., 1993, Applied Mathematics in Water Supply Network Management, Automatica, v. 29, n. 5,pp. 1215-1250.

25. Powell, R.S., 1992, On-line Monitoring for Operational Control of Water Distribution Networks, PhD Thesis, University of Durham, UK.

26. Roy, D.N., 1988, Applied Fluid Mechanics, Ellis Horwood Ltd.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

SIMULTANEOUS ESTIMATION OF SPATIALLY-DEPENDENT MASS AND HEAT TRANSFER COEFFICIENTS OF DRYING BODIES

Leonardo F. Saker and Helcio R. B. Orlande
Department of Mechanical Engineering, DEM/PEM - POLI/COPPE
Federal University of Rio de Janeiro, UFRJ
Rio de Janeiro, RJ, Brazil
saker@lmt.coppe.ufrj.br and helcio@serv.com.ufrj.br

ABSTRACT

In this paper, the conjugate gradient method with adjoint problem is applied for the identification of the heat and mass transfer coefficients at the surface of drying capillary-porous bodies. The unknown functions are supposed to vary along the surface open to the surrounding environment. The effects of temperature and moisture content measurements on the inverse analysis are examined. The inverse problem is solved by considering either the heat or the mass transfer coefficients as unknown, as well as by considering simultaneously both functions as unknown.

NOMENCLATURE

Bi_m	mass transfer coefficient
Bi_q	heat transfer coefficient
C	measured moisture content
Ko	Kossovitch number
Lu	Luikov number
M	measured temperature
Pn	Posnov number
r_a	aspect ratio
S	objective functional
X,Y	dimensionless Cartesian coordinates

Greeks

Df	sensitivity function for moisture content
Dq	sensitivity function for temperature
Y^k	conjugation coefficient def. by Eq. (6.b)
a	parameter defined by Eq. (2.m)
b	parameter defined by Eq. (2.n)
b^k	search step size at iteration k
f	dimensionless moisture content
g	parameter defined by Eq. (2.o)
g^k	conjugation coefficient def. by Eq. (6.a)
l	Lagrange multiplier
q	dimensionless temperature
t	dimensionless time

INTRODUCTION

The phenomena of coupled heat and mass transfer in capillary porous media has been drawing the attention of research groups for a long time, because of its importance in several practical applications, such as drying. For the mathematical modeling of such phenomena, Luikov [1] has proposed a model based on a system of coupled diffusion equations, which takes into account the effects of the temperature gradient on the moisture migration.

The computation of temperature and moisture content fields in capillary porous media, from the knowledge of initial and boundary conditions, as well as of the thermophysical properties appearing in the formulation, constitutes a *Direct Problem* of heat and mass transfer [1,2]. Appropriately formulated direct problems are mathematically classified as *well-posed*, that is, their solutions satisfy the requirements of existence, uniqueness and stability with respect to the input data [3-5]. On the other hand, the estimation of boundary conditions in Luikov's formulation, by using temperature and/or moisture content measurements taken in the medium, is an *Inverse Problem* of coupled heat and mass transfer [3-5]. Generally, inverse problems are mathematically classified as *ill-posed* [3-5]. Despite the ill-posed character, the solution of an inverse problem can be obtained through its reformulation in terms of a well-posed problem, such as a minimization problem associated with some kind of regularization (stabilization) technique. Different methods based on such an approach have been successfully used in the past for the estimation of parameters and functions, in linear and non-linear inverse problems [3-5]. Recently, several articles dealing

with the solution of inverse problems of coupled heat and mass transfer appeared in the literature [6-15].

In this paper, we examine the solution of inverse problems involving the identification of the heat and mass transfer coefficients at the surface of a capillary-porous body, as a function estimation approach. The unknown quantities are supposed to vary along the surface of the body, open to the surrounding environment. The conjugate gradient method with adjoint problem [3-5] is used for the identification of the unknown functions. The inverse problems of estimating one single function are examined, as well as the inverse problem of simultaneously estimating both functions. The effects of the use of temperature and moisture content measurements on the inverse analysis are examined, by using simulated measured data with random errors.

PHYSICAL PROBLEM AND MATHEMATICAL FORMULATION

The physical problem involves a two-dimensional capillary porous medium in Cartesian coordinates, initially at uniform temperature and uniform moisture content. The lateral surfaces of the body are impervious to moisture transfer and thermally insulated. The bottom boundary, which is impervious to moisture transfer, is in direct contact with a heater. The top boundary is in contact with the dry surrounding air, thus resulting in a convective boundary condition for both the temperature and the moisture content. The mass and heat transfer coefficients at this boundary may vary along the surface open to the surrounding environment. The linear system of equations proposed by Luikov [1], with associated initial and boundary conditions, for the modeling of such physical problem involving heat and mass transfer in capillary porous media, can be written in dimensionless form as:

$$\frac{\partial q}{\partial t} = a\left(\frac{\partial^2 q}{\partial X^2} + \frac{\partial^2 q}{\partial Y^2}\right) - b\left(\frac{\partial^2 q}{\partial X^2} + \frac{\partial^2 q}{\partial Y^2}\right)$$

in $0 < X < r_a$, $0 < Y < 1$ and $t > 0$ (1.a)

$$\frac{\partial f}{\partial t} = Lu\left(\frac{\partial^2 f}{\partial X^2} + \frac{\partial^2 f}{\partial Y^2}\right) - Lu\ Pn\left(\frac{\partial^2 q}{\partial X^2} + \frac{\partial^2 q}{\partial Y^2}\right)$$

in $0 < X < r_a$, $0 < Y < 1$ and $t > 0$ (1.b)

$$\frac{\partial q}{\partial Y} = -Q \qquad \text{at } Y = 0 \text{ , for } t > 0 \qquad (1.c)$$

$$\frac{\partial f}{\partial Y} = -Pn\,Q \qquad \text{at } Y = 0 \text{ , for } t > 0 \qquad (1.d)$$

$$\frac{\partial q}{\partial Y} = Bi_q(X)(1-q) - (1-e)Ko\,Lu\,Bi_m(X)(1-f)$$

at $Y = 1$, for $t > 0$ (1.e)

$$\frac{\partial f}{\partial Y} = Pn\ \frac{\partial q}{\partial Y} + Bi_m(X)(1-f)$$

at $Y = 1$, for $t > 0$ (1.f)

$$\frac{\partial q}{\partial X} = \frac{\partial f}{\partial X} = 0 \qquad \text{at } X = 0 \text{ and } X = r_a \text{, for } t > 0$$

(1.g,h)

$$q(X,Y,0) = f(X,Y,0) = 0$$

for $t = 0$, in $0 < X < r_a$, $0 < Y < 1$ (1.i,j)

where the following dimensionless variables were defined:

$$q = \frac{T - T_0}{T_s - T_0} \qquad f = \frac{u_0 - u}{u_0 - u^*} \qquad Q = \frac{qh}{k(T_s - T_0)}$$

$$t = \frac{at}{h^2} \qquad Lu = \frac{a_m}{a} \qquad Pn = d\frac{T_s - T_0}{u_0 - u^*}$$

$$Ko = \frac{r(u_0 - u^*)}{c(T_s - T_0)} \qquad X = \frac{x}{h} \qquad Y = \frac{y}{h}$$

$$Bi_q(X) = \frac{h_q(X)l}{k} \qquad Bi_m(X) = \frac{h_m(X)l}{k_m} \qquad r_a = \frac{L}{h}$$

$$a = 1 + e\,Ko\,Lu\,Pn \qquad b = e\,Ko\,Lu$$

$$g = 1 - (1-e)\,Ko\,Lu\,Pn \qquad (2.a\text{-}o)$$

The properties of the porous medium appearing above include the thermal diffusivity (a), the moisture diffusivity (a_m), the thermal conductivity (k), the moisture conductivity (k_m) and the specific heat (c). Other physical quantities appearing in the dimensionless groups of Eqs. (2) are the heat transfer coefficient (h_q), the mass transfer coefficient (h_m), the thickness of porous medium (h), the width of the porous medium (L), the prescribed heat flux (q), the latent heat of evaporation of water (r), the temperature of the surrounding air (T_s), the uniform initial temperature in the medium (T_0), the moisture content of the surrounding air (u^*), the uniform

initial moisture content in the medium (u_0), the thermogradient coefficient (d) and the phase conversion factor (e). *Lu*, *Pn* and *Ko* denote the Luikov, Posnov and Kossovitch numbers, respectively [1].

Problem (1) is referred to as a *Direct Problem* when initial and boundary conditions, as well as all parameters appearing in the formulation, are known. The objective of the direct problem is to determine the dimensionless temperature and moisture content fields, $q(X,Y,t)$ and $f(X,Y,t)$, respectively, in the capillary porous media.

INVERSE PROBLEM

For the *inverse problem* of interest here, the functions $Bi_q(X)$ and $Bi_m(X)$ are regarded as unknown quantities. For the estimation of such functions, we consider available the transient temperature measurements $M_i(t)$ taken at the locations (X_i,Y_i) i=1,…,I, as well as the moisture content measurements $C_n(t)$ taken at the locations (X_n^*,Y_n^*), n=1,…,N. We note that the measurements may contain random errors, but all the other quantities appearing in the formulation of the direct problem are supposed to be exactly known.

Inverse problems are ill-posed [3-5]. Several methods of solution of inverse problems, such as the one used here, involve their reformulation in terms of well-posed minimization problems. We consider our objective functional in the form:

$$S[Bi_m(X),Bi_q(X)] = \int_{t=0}^{t_f}\left\{\sum_{i=1}^{I}[q(X_i,Y_i,t;Bi_m,Bi_q)-M_i(t)]^2 w_q\right\}dt + \int_{t=0}^{t_f}\left\{\sum_{n=1}^{N}[f(X_n^*,Y_n^*,t;Bi_m,Bi_q)-C_n(t)]^2 w_f\right\}dt \quad (3)$$

where q and f are the estimated temperature and moisture content, respectively, which are obtained from the solution of the direct problem with estimates for the unknown functions. In equation (3) w_q and w_f are weights for the temperature and moisture content measurements, respectively.

For the minimization of such objective functional, we use here the conjugate gradient method with adjoint problem, as described next.

CONJUGATE GRADIENT METHOD

The iterative procedure of the conjugate gradient method is given by:

$$Bi^{k+1}(X) = Bi^k(X) + b^k d^k(X) \quad (4)$$

where the superscript k denotes the number of iterations, b^k is the search step size, $d^k(t)$ is the direction of descent and $Bi(X)$ may represent $Bi_q(X)$ and $Bi_m(X)$.

The direction of descent $d^k(t)$ is a conjugation of the gradient direction with previous directions of descent. It is given in the following general form:

$$d^k(X) = -\nabla S[Bi^k(X)] + g^k d^{k-1}(X) + y^k d^q(X) \quad (5)$$

where g^k and y^k are conjugation coefficients. The superscript q in equation (5) denotes the iteration number where a *restarting strategy* is applied to the iterative procedure of the conjugate gradient method.

Different versions of the conjugate gradient method can be found in the literature depending on the form used for the computation of the direction of descent given by equation (5) [3-5, 16,17]. In this paper, the so-called Powell-Beale's version is used due to its superior robustness for the solution of non-linear inverse problems [17].

Powell [16] suggested the following expressions for the conjugation coefficients, which gives the so-called *Powell-Beale's* version of the conjugate gradient method [16,17]:

$$g^k = \frac{\int_{X=0}^{r_a}\{\nabla S[Bi^k(X)]-\nabla S[Bi^{k-1}(X)]\}\nabla S[Bi^k(X)]dX}{\int_{X=0}^{r_a}\{\nabla S[Bi^k(X)]-\nabla S[Bi^{k-1}(X)]\}d^{k-1}(X)dX}$$

with $g^k = 0$ for $k = 0$ (6.a)

$$y^k = \frac{\int_{X=0}^{r_a}\{\nabla S[Bi^{q+1}(X)]-\nabla S[Bi^q(X)]\}\nabla S[Bi^k(X)]dX}{\int_{X=0}^{r_a}\{\nabla S[Bi^{q+1}(X)]-\nabla S[Bi^q(X)]\}d^q(X)dX}$$

with $y^k = 0$ for $k = 0$ (6.b)

In accordance with Powell [16], the application of the conjugate gradient method with the conjugation coefficients given by equations (6) requires restarting when gradients at successive iterations tend to be non-orthogonal

(which is a measure of the local non-linearity of the problem) and when the direction of descent is not sufficiently downhill. Restarting is performed by making $y^{k} = 0$ in equation (5).

The non-orthogonality of gradients at successive iterations is tested by using:

$$\text{ABS}\left(\int_{X=0}^{r_a} \nabla S[Bi^{k-1}(X)]\, \nabla S[Bi^{k}(X)]\, dX \right) \geq 0.2 \int_{X=0}^{r_a} \{\nabla S[Bi^{k}(X)]\}^2 dX \qquad (7.a)$$

where ABS (.) denotes the absolute value.

A non-sufficiently downhill direction of descent (i.e., the angle between the direction of descent and the negative gradient direction is too large) is identified if either of the following inequalities is satisfied:

$$\int_{X=0}^{r_a} d^{k}(X) \nabla S[Bi^{k}(X)]\, dX \leq -1.2 \int_{X=0}^{r_a} \{\nabla S[Bi^{k}(X)]\}^2 dX \qquad (7.b)$$

or

$$\int_{X=0}^{r_a} d^{k}(X) \nabla S[Bi^{k}(X)]\, dX \geq -0.8 \int_{X=0}^{r_a} \{\nabla S[Bi^{k}(X)]\}^2 dX \qquad (7.c)$$

For the numerical implementation of the iterative procedure of the conjugate gradient method, auxiliary problems are required, namely the sensitivity problems and the adjoint problem, as described next.

SENSITIVITY PROBLEMS AND SEARCH STEP SIZE

The sensitivity problem is used to determine the variation in temperature and in moisture content due to changes in the unknown quantity. Since the present work deals with the estimation of two unknown functions, two sensitivity problems are required in the analysis. They are derived by considering perturbations in the heat and mass transfer coefficients, each at a time, as described next.

Let us consider that the temperature $q(X,Y,t)$ and the moisture content $f(X,Y,t)$ undergo variations $\Delta q_1(X,Y,t)$ and $\Delta f_1(X,Y,t)$, respectively, when the mass transfer coefficient $Bi_m(X)$ is perturbed by $\Delta Bi_m(X)$. By substituting in the direct problem (1) $q(X,Y,t)$ by $[q(X,Y,t)+ \Delta q_1(X,Y,t)]$, $f(X,Y,t)$ by $[f(X,Y,t)+ \Delta f_1(X,Y,t)]$ and $Bi_m(X)$ by $[Bi_m(X)+ DBi_m(X)]$, and then subtracting from the resulting problem the original direct problem, we obtain the following sensitivity problem for the sensitivity functions $\Delta q_1(X,Y,t)$ and $\Delta f_1(X,Y,t)$:

$$\frac{\partial \Delta q_1}{\partial t} = a\left(\frac{\partial^2 \Delta q_1}{\partial X^2} + \frac{\partial^2 \Delta q_1}{\partial Y^2} \right) - b\left(\frac{\partial^2 \Delta q_1}{\partial X^2} + \frac{\partial^2 \Delta q_1}{\partial Y^2} \right)$$

in $0 < X < r_a$, $0 < Y < 1$ and $t > 0$ (8.a)

$$\frac{\partial \Delta f_1}{\partial t} = Lu\left(\frac{\partial^2 \Delta f_1}{\partial X^2} + \frac{\partial^2 \Delta f_1}{\partial Y^2} \right) - Lu\, Pn\left(\frac{\partial^2 \Delta q_1}{\partial X^2} + \frac{\partial^2 \Delta q_1}{\partial Y^2} \right)$$

in $0 < X < r_a$, $0 < Y < 1$ and $t > 0$ (8.b)

$$\frac{\partial \Delta q_1}{\partial Y} = \frac{\partial \Delta f_1}{\partial Y} = 0 \qquad \text{at } Y = 0 \text{ , for } t > 0 \quad (8.c.d)$$

$$\frac{\partial \Delta q_1}{\partial Y} = -Bi_q(X)\Delta q_1 + (1-e)Ko\, Lu\, \Delta Bi_m(X)[f-1] + (1-e)Ko\, Lu\, Bi_m(X)\, \Delta f_1$$

at $Y = 1$, for $t > 0$ (8.e)

$$\frac{\partial \Delta f_1}{\partial Y} = \{\Delta Bi_m(X) - Bi_m(X)\Delta f_1 - fBi_m(X)\}g - Bi_q(X)\, Pn\, \Delta q_1$$

at $Y = 1$, for $t > 0$ (8.f)

$$\frac{\partial \Delta q_1}{\partial X} = \frac{\partial \Delta f_1}{\partial X} = 0$$

at $X = 0$ and $X = r_a$, for $t > 0$ (8.g,h)

$$\Delta q_1(X,Y,0) = \Delta f_1(X,Y,0) = 0$$

for $t = 0$, in $0 < X < r_a$, $0 < Y < 1$ (8.i,j)

Similarly, the sensitivity problem for the sensitivity functions $\Delta q_2(X,Y,t)$ and $\Delta f_2(X,Y,t)$, resultant from a perturbation $DBi_q(X)$ in $Bi_q(X)$, can be obtained as:

$$\frac{\partial \Delta q_2}{\partial t} = a\left(\frac{\partial^2 \Delta q_2}{\partial X^2} + \frac{\partial^2 \Delta q_2}{\partial Y^2} \right) - b\left(\frac{\partial^2 \Delta q_2}{\partial X^2} + \frac{\partial^2 \Delta q_2}{\partial Y^2} \right)$$

in $0 < X < r_a$, $0 < Y < 1$ and $t > 0$ (9.a)

$$\frac{\partial \Delta f_2}{\partial t} = Lu\left(\frac{\partial^2 \Delta f_2}{\partial X^2} + \frac{\partial^2 \Delta f_2}{\partial Y^2} \right) - Lu\, Pn\left(\frac{\partial^2 \Delta q_2}{\partial X^2} + \frac{\partial^2 \Delta q_2}{\partial Y^2} \right)$$

in $0 < X < r_a$, $0 < Y < 1$ and $t > 0$ (9.b)

$$\frac{\partial \Delta q_2}{\partial Y} = \frac{\partial \Delta f_2}{\partial Y} = 0 \qquad \text{at } Y = 0 \text{ , for } t > 0 \qquad (9.c.d)$$

$$\frac{\partial \Delta q_2}{\partial Y} = \Delta Bi_q(X) - \left[Bi_q(X)\Delta q_2 + q\Delta Bi_q(X) \right] + (1-e)\,Ko\,Lu\,Bi_m(X)\Delta f_2$$

$$\text{at } Y = 1 \text{ , for } t > 0 \qquad (9.e)$$

$$\frac{\partial \Delta f_2}{\partial Y} = -Bi_m(X)\Delta f_2\, g + \Delta Bi_q(X)\, Pn - \left[Bi_q(X)\Delta q_2 + q\Delta Bi_q(X) \right] Pn$$

$$\text{at } Y = 1 \text{ , for } t > 0 \qquad (9.f)$$

$$\frac{\partial \Delta q_2}{\partial X} = \frac{\partial \Delta f_2}{\partial X} = 0$$

$$\text{at } X = 0 \text{ and } X = r_a \text{, for } t > 0 \qquad (9.g,h)$$

$$\Delta q_2(X,Y,0) = \Delta f_2(X,Y,0) = 0$$

$$\text{for } t = 0 \text{ , in } 0 < X < r_a \text{ , } 0 < Y < 1 \qquad (9.i,j)$$

Linearized expressions can be obtained for the search step-sizes for the iterative procedures for the estimation of $Bi_q(X)$ and $Bi_m(X)$, by minimizing the objective functional at each iteration with respect to these quantities. We omit details of such derivations, but they can be readily found in Ref. [5].

ADJOINT PROBLEM AND GRADIENT EQUATIONS

The adjoint problem is derived by multiplying the governing equations of the direct problem by Lagrange multipliers, integrating in the spatial and time domains that they are valid and then adding the resultant equation to the original functional (3). The directional derivative of the functional in the direction of the perturbation of each of the unknown functions is then obtained and the resultant expression, after some lengthy but straightforward manipulations, is allowed to go to zero. The same adjoint problem is obtained for perturbations in $Bi_q(X)$ and $Bi_m(X)$. The adjoint problem, for the computation of the Lagrange Multipliers $\lambda_1(X,Y,t)$ and $\lambda_2(X,Y,t)$, is given by:

$$\frac{\partial l_1}{\partial t} = a\left(\frac{\partial^2 l_1}{\partial X^2} + \frac{\partial^2 l_1}{\partial Y^2} \right) + Lu\ Pn \left(\frac{\partial^2 l_2}{\partial X^2} + \frac{\partial^2 l_2}{\partial Y^2} \right) - \sum_{i=1}^{I} 2\left(q_i - M_i\right) d\left(X - X_i\right) d\left(Y - Y_i\right) w_q$$

$$\text{in } 0 < X < r_a \text{ , } 0 < Y < 1 \text{ and } t > 0 \qquad (10.a)$$

$$\frac{\partial l_2}{\partial t} = Lu\left(\frac{\partial^2 l_1}{\partial X^2} + \frac{\partial^2 l_1}{\partial Y^2} \right) + b\left(\frac{\partial^2 l_2}{\partial X^2} + \frac{\partial^2 l_2}{\partial Y^2} \right) - \sum_{n=1}^{N} 2\left(f_n - C_n\right) d\left(X - X_n^*\right) d\left(Y - Y_n^*\right) w_f$$

$$\text{in } 0 < X < r_a \text{ , } 0 < Y < 1 \text{ and } t > 0 \qquad (10.b)$$

$$\frac{\partial l_1}{\partial Y} = \frac{\partial l_2}{\partial Y} = 0 \qquad \text{at } Y = 0 \text{ , for } t > 0 \qquad (10.c.d)$$

$$\frac{\partial l_1}{\partial Y} = -\,l_1\, Bi_q(X) - \frac{Lu\ l_2\, Bi_q(X)\, Pn}{a}$$

$$\text{at } Y = 1 \text{ , for } t > 0 \qquad (10.e)$$

$$\frac{\partial l_2}{\partial Y} = -a\ l_1 (1-e)\, Ko\, Bi_m(X) - l_2\, Bi_m(X) g$$

$$\text{at } Y = 1 \text{ , for } t > 0 \qquad (10.f)$$

$$\frac{\partial l_1}{\partial Y} = \frac{\partial l_2}{\partial Y} = 0$$

$$\text{at } X = 0 \text{ and } X = r_a \text{, for } t > 0 \qquad (10.g,h)$$

$$q(X,Y,t_f) = f(X,Y,t_f) = 0$$

$$\text{for } t = t_f \text{, in } 0 < X < r_a \text{ , } 0 < Y < 1 \qquad (10.i,j)$$

With the limiting process used to obtain the adjoint problem (10), we can also identify the following expressions for the gradient directions, where it was taken into account the hypotheses that $Bi_q(X)$ and $Bi_m(X)$ belong to the Hilbert space of square integrable functions in the domain $0 < X < r_a$:

$$\nabla S[Bi_m(X)] = \int_{t=0}^{t_f} \{a\, l_1(X,1,t)(1-e)KoLu[f(X,1,t)-1] + Lu\, l_2(X,1,t)[1-f(X,1,t)g]\}dt \qquad (11.a)$$

$$\nabla S[Bi_q(X)] = \int_{t=0}^{t_f} \{a\ l_1(X,1,t)[1-q(X,1,t)] + Lu\ l_2(X,1,t)\, Pn[1-q(X,1,t)]\}\ dt \qquad (11.b)$$

The use of the conjugate gradient method may result on stable solutions if the *Discrepancy Principle* [3-5] is used to specify the tolerance for the stopping criterion of the iterative procedure. In the Discrepancy Principle, the solution is

assumed to be sufficiently accurate when the difference between measured and estimated quantities is of the order of magnitude of the measurement errors.

RESULTS AND DISCUSSIONS

For the results presented below, we examined test-cases involving the drying of a capillary-porous body with dimensions h=0.05m and L=0.5m, made of ceramics, with properties [7]: $k=0.34W/mK$, $k_m=2.4\text{x}10^{-7}kg/ms^oM$, $c=607J/kgK$ $r=2.5\text{x}10^6J/kg$, $T_0=24^oC$, $u_0=80^oM$, $d = 0.56^oM/K$ and $e = 0.8$. The air conditions were taken as $T_s=30^oC$ and $u^*=40^oM$ and the applied heat flux as $q = 40\ W/m^2$. Therefore, the dimensionless numbers appearing in the formulation were Lu=0.2, Pn=0.084, Ko=49 and Q=0.9.

For the estimation of the unknown heat and mass transfer coefficients, we made use of simulated temperature and moisture content measurements. Two different levels of measurement errors were examined here: errorless measurements and measurements with standard-deviation of 1% of the maximum value of the measured quantity. Hypothetical functions containing discontinuities, which are the most difficult to be recovered through the solution of the inverse problem, were used to generate the simulated measurements.

Let us consider initially in the analysis the estimation of the heat transfer coefficient $Bi_q(X)$, by assuming that the mass transfer coefficient $Bi_m(X)$ is exactly known. For this case, $Bi_m(X)$ was taken as the same function as for $Bi_q(X)$. Figure 1 presents the functions estimated for $Bi_q(X)$, obtained with errorless temperature measurements (w_q=1 and w_f=0 in equation (3)) and with different sensor configurations, as well as the exact function used to generate the simulated measurements. The sensor configurations are given in table 1. The sensors were equally spaced along the X direction, at the Y position specified in this table. We also present in table 1, the RMS errors [5] obtained with the different sensor configurations. Table 1 shows that the accuracy of the estimated function deteriorates as the sensors are located farther from the top boundary, as well as when fewer sensors are used for the inverse analysis. However, note in figure 1 that quite accurate results could be obtained even with 13 sensors located at Y = 0.85 (which corresponds to 7.5 mm below the top surface).

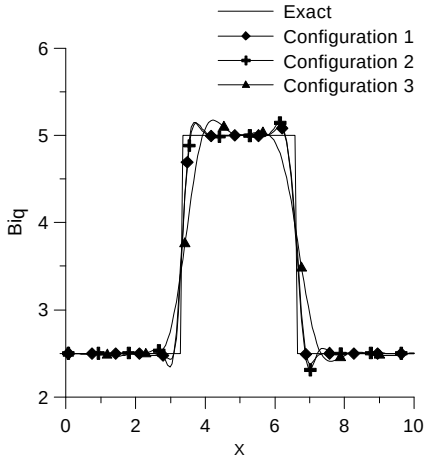


Figure 1. Results obtained for $Bi_q(X)$ with known $Bi_m(X)$ by using temperature measurements

Table 1. Sensor configurations used for the estimation of $Bi_q(X)$ with known $Bi_m(X)$.

Config.	Number of Sensors	Y Location	*RMS Error*
1	33	0.925	4.38×10^{-2}
2	33	0.85	4.58×10^{-2}
3	13	0.85	1.11×10^{-1}

We now consider the estimation of $Bi_m(X)$ by using temperature measurements (w_q=1 and w_f=0 in equation (3)), and assuming $Bi_q(X)$ as exactly known for the inverse analysis. For this case, $Bi_q(X)$ was taken as the same function as for $Bi_m(X)$. Figure 2 presents the estimated function obtained with measurements containing random errors, of 17 sensors equally spaced, at the position Y = 0.85. Figure 2 shows the interesting fact that temperature measurements provide useful information for the estimation of $Bi_m(X)$. This is an important result, because quite involved and inaccurate techniques for the measurement of moisture content can be avoided, in favor of inexpensive and accurate temperature measurements, for the estimation of $Bi_m(X)$, if $Bi_q(X)$ is known. On the other hand, the use of moisture content measurements can result on accurate estimations for $Bi_m(X)$, as illustrated in figure 3; but not for the estimation of $Bi_q(X)$. The estimated function shown in figure 3 was obtained with moisture content measured data containing random errors (w_q=0 and w_f=1 in equation (3)), of 17 sensors equally spaced, located at Y=0.85.

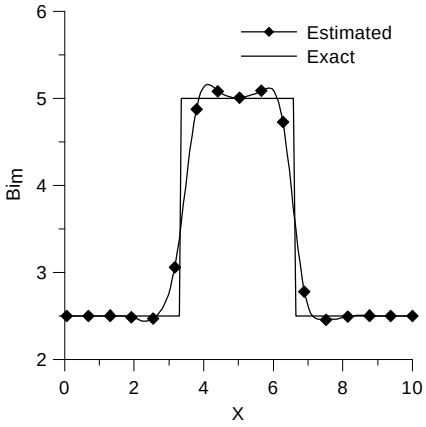


Figure 2. Results obtained for $Bi_m(X)$ with known $Bi_q(X)$ by using temperature measurements

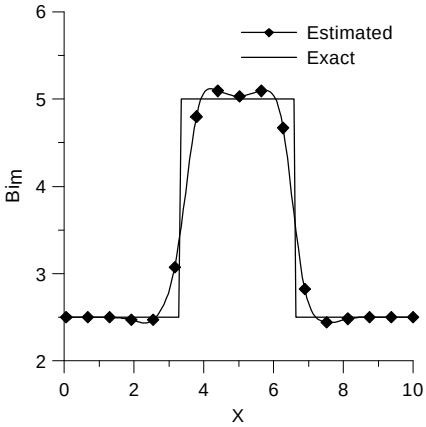


Figure 3. Results obtained for $Bi_m(X)$ with known $Bi_q(X)$ by using moisture content measurements

The simultaneous estimation of $Bi_q(X)$ and $Bi_m(X)$ is now examined. For this case, the use of only temperature measurements or only moisture content measurements resulted on quite inaccurate estimated functions. Therefore, temperature, as well as moisture content measurements, were required for the simultaneous estimation of $Bi_q(X)$ and $Bi_m(X)$. Figures 4.a,b present the results obtained for $Bi_q(X)$ and $Bi_m(X)$, respectively, by using in the inverse analysis simulated measurements containing random errors of 14 temperature sensors and 15 moisture content sensors. In this case, we used $w_q = 1/M_{\max}^2$ and $w_f = 1/C_{\max}^2$ in equation (3), where M_{max} and C_{max} are the maximum measured values of temperature and moisture content, respectively. The temperature sensors and the moisture content sensors were located at Y=0.9, evenly spaced along the X direction. Figures 4.a,b show that quite accurate results can be obtained for the simultaneous estimation of $Bi_q(X)$ and $Bi_m(X)$, if both temperature and moisture content measurements are used in the inverse analysis.

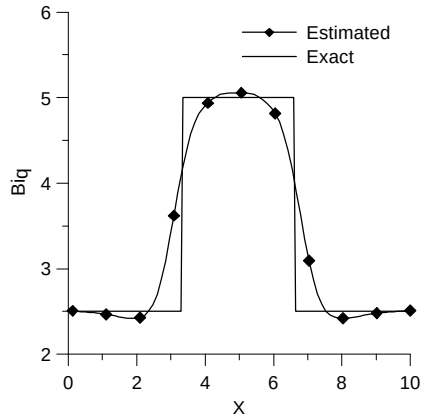


Figure 4.a. Results obtained for $Bi_q(X)$ by using temperature and moisture content measurements - simultaneous estimation of $Bi_m(X)$ and $Bi_q(X)$

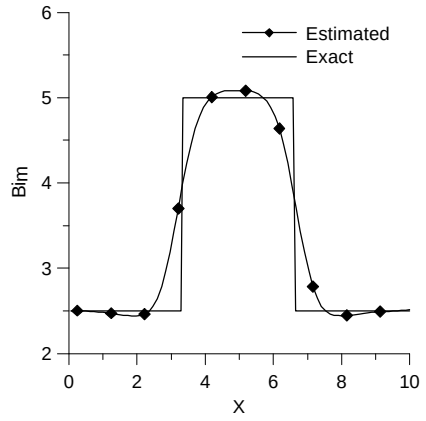


Figure 4.b. Results obtained for $Bi_m(X)$ by using temperature and moisture content measurements - simultaneous estimation of $Bi_m(X)$ and $Bi_q(X)$

CONCLUSIONS

In this paper we solve the inverse problem of simultaneously estimating the mass and heat transfer coefficient at the surface of a drying-capillary porous body. The unknown quantities are treated as functions of the position along the surface open to the surrounding environment. The present inverse problem is solved with the conjugate gradient method of function estimation with adjoint problem.

Results obtained with simulated measurements indicate that the present approach is capable of recovering the heat transfer coefficient, with the use of only temperature

measurements, if the mass transfer coefficient is known for the analysis. If the heat transfer coefficient is regarded as known for the inverse analysis, the mass transfer coefficient can be estimated by using either temperature measurements or moisture content measurements. Both temperature and moisture content measurements are required for the simultaneous estimation of the mass and heat transfer coefficients. The present solution approach is stable with respect to measurement errors. Quite accurate results were obtained even for functional forms containing discontinuities, which are the most difficult to be recovered by inverse analysis.

ACKNOWLEDGEMENTS

Useful discussions with Professors G. Kanevce and L. Kanevce from St. Kliment Ohridski University, Macedonia, are greatly appreciated. This work was partially funded by CNPq, CAPES and FAPERJ.

REFERENCES

1. Luikov, A. V., *Heat and Mass Transfer in Capillary-porous Bodies*, Pergamon Press, Oxford, 1966.

2. Mikhailov, M. D. and Özisik, M. N., *Unified Analysis and Solutions of Heat and Mass Diffusion*, John Wiley, New York, 1984.

3. Alifanov, O. M., *Inverse Heat Transfer Problems*, Springer-Verlag, New York, 1994.

4. Alifanov, O. M., Artyukhin, E. and Rumyantsev, A., *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems*, Begell House, New York, 1995.

5. Özisik, M. N. and Orlande, H. R. B., *Inverse Heat Transfer: Fundamentals and Applications*, Taylor and Francis, New York, 2000.

6. Dantas, L. B., Orlande H. R. B., Cotta R. M. and Lobo P. D. C., Parameter Estimation in Moist Capillary Porous Media by using Temperature Measurements, in *Inverse Problems in Engineering Mechanics II*, M. Tanaka and G. S. Dulikravich (eds), pp. 53-62, Elsevier, Oxford, 2000.

7. Dantas, L. B., H. R. B. Orlande, and R. M. Cotta., Effects of Lateral Heat Losses on the Parameter Estimation Problem in Moist Capillary Porous Media. *International Symposium on Inverse Problems in Engineering Mechanics – ISIP2001*, Nagano, Japan, February, (2001).

8. Dantas, L. B., Orlande, H. R. B. and Cotta, R. M., Estimation of Dimensionless Parameters of Luikov's System for Heat and Mass Transfer in Capillary Porous Media, *International Journal of Thermal Sciences*, (2001).

9. Kanevce, G., Kanevce, L. and Dulikravich, G. S., Moisture Diffusivity Estimation by Temperature Response of a Drying Body, in *Inverse Problems in Engineering Mechanics II*, M. Tanaka and G. S. Dulikravich (eds), pp. 63-72, Elsevier, Oxford, (2000).

10. Kanevce, G., Kanevce, L. and Dulikravich, G. S., Influence of Boundary Conditions on Moisture Diffusivity Estimation by Temperature Response of a Drying Body, *Proceedings of 34th ASME National Heat Transfer Conference*, paper NHTC2000-12296, Pittsburgh, Pennsylvania, August 20-22, (2000).

11. Kanevce, G., Kanevce, L. and Dulikravich, G. S., Simultaneous Estimation of Thermophysical Properties and Heat and Mass Transfer Coefficients of a Drying Body, *International Symposium on Inverse Problems in Engineering Mechanics – ISIP2001*, Nagano, Japan, February, (2001).

12. Kanevce, G., Kanevce, L. and Dulikravich, G. S. and Orlande, H. R. B., Estimation of Thermophysical Properties of Moist Materials Under Different Drying Conditions, *4th International Conference on Inverse Problems in Engineering*, (2002).

13. Chang W. J. and Weng C. I., Inverse Problem of Coupled Heat and Moisture Transport for Prediction of Moisture Distributions in an Annular Cylinder, *Int. J. Heat Mass and Transfer*, **42**, pp. 2661-2672, (1999).

14. Frias, J. M., Oliveira, J. C. and Schittkowski, K., Modeling and Parameter Identification of a Maltodextrin DE 12 Drying Process in a Convection Oven, *Applied Mathematical Modeling*, **25**, pp. 449-462, (2001).

15. Dantas, L. B., Orlande, H. R. B. and Cotta, R. M., An Inverse Problem Of Parameter Estimation For Heat And Mass Transfer In Capillary Porous Media, *Int. J. Heat Mass Transfer*, (in press).

16. M. J. D. Powell, Restart Procedures for the Conjugate Gradient Method, *Math. Program.*, **12**, pp. 241-254 (1977).

17. M. J. Colaço and H. R. B. Orlande, Comparison of Different Versions of the Conjugate Gradient Method of Function Estimation, *Numerical Heat Transfer*, **36**, pp. 229-249 (1999).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ESTIMATION OF THE HEAT FLUX AT THE SURFACE OF ABLATING MATERIALS BY USING TEMPERATURE AND SURFACE POSITION MEASUREMENTS

Alexandre P. de Oliveira and Helcio R. B. Orlande
Department of Mechanical Engineering, DEM/PEM - POLI/COPPE
Federal University of Rio de Janeiro, UFRJ
Rio de Janeiro, RJ, Brazil
alexandre@lttc.coppe.ufrj.br and helcio@serv.com.ufrj.br

ABSTRACT

In this paper, the conjugate gradient method with adjoint problem is applied for the identification of the heat flux at the surface of ablating materials, by assuming no *a priori* information regarding the functional form of the unknown. Simulated measurements of the position of the ablating surface are used in the inverse analysis, together with simulated temperature measurements. The accuracy of the conjugate gradient method with adjoint problem is examined for functions containing sharp corners, with different orders of magnitude for the peak-flux.

INTRODUCTION

An important requirement for the thermal design of heat shields of vehicles re-entering the atmosphere, which are subjected to extremely high heat loads, is to have prior accurate information regarding the thermal properties of the materials utilized [1]. With respect to the heat loads that the surface of the re-entering vehicle are subjected to, an usual approach for the design of the heat shield is to use simplified models, such as those described in reference [2]. On the other hand, the use of inverse analysis techniques can provide the important information on how accurate such simplified models are [3,4]. In reference [4], we examined the use of temperature measurements taken below the surface for the solution of such inverse problem. Accurate results were obtained for functions containing sharp corners and discontinuities, but for small magnitudes of the peak-flux, when the portion of material removed during ablation is small.

In this paper, we examine the use of surface position measurements, in addition to temperature measurements, for the estimation of the boundary heat flux at the surface of ablating materials. We show that the use of surface position measurements in the inverse analysis is specially important to recover heat fluxes with large magnitudes, when the portion of material removed during ablation is significant. Measurements of the position of the ablating surface can be performed with different experimental techniques, as described in [5]. For the solution of the inverse problem, we use the conjugate gradient method of function estimation with adjoint problem [6,7], also known as the Alifanov's variational method.

DIRECT PROBLEM

The physical problem under picture in this paper consists of a one-dimensional slab of thickness b_0 initially at the temperature $T_0(x)$, which is lower than the temperature of ablation T_{ab}. The surface of the slab at $x = b_0$ is heated with a heat flux $q(t)$, while the other surface at $x = 0$ is kept insulated. As the slab is heated, the temperature of the heated surface may reach the temperature of ablation; the material will then be removed and a moving boundary problem will be established. For vehicles re-entering the atmosphere, the heat flux increases during the initial stages of the re-entry procedure, passes through a peak value and then decreases, perhaps even to negative values, when the surface of the vehicle is cooled by convection in the lower portions of the atmosphere [2]. Thus, when solving the direct problem with this kind of heat flux, three distinct periods come into picture: (i) *the pre-ablation period*, which lasts from the beginning of the re-entry procedure until the heated surface reaches the ablating temperature; (ii) *the ablation period*, when a moving-boundary problem takes place because of ablation; and (iii)

the post-ablation period, which begins when the ablation process ends and the surface of the solid remains again fixed. More specifically, these three periods take place, respectively, during the following time intervals: $0 < t \le t_{ab}$, $t_{ab} < t \le t_{ab,f}$ and $t_{ab,f} < t \le t_f$, where t_{ab}, $t_{ab,f}$ and t_f denote the time when ablation begins, the time when ablation ends and the final time, respectively. For convenience in the analysis, the pre-ablation and post-ablation periods are generally referred to as *no-ablation* periods and take place during the following general time interval $t_0 < t \le t_0 + dt_0$, with duration dt_0.

For the physical problem under picture in this paper, the thermal decomposition of the material is neglected. The mathematical formulation of this problem can be written in dimensionless form as:

No-ablation periods ($t_0 < t \le t_0 + dt_0$):

$$\frac{\partial q(X,t)}{\partial t} = \frac{\partial^2 q(X,t)}{\partial X^2} \qquad (1.a)$$

$$\frac{\partial q}{\partial X} = 0 \qquad \text{at } X = 0 \qquad (1.b)$$

$$\frac{\partial q}{\partial X} = Q(t) \qquad \text{at } X = B(t) \qquad (1.c)$$

$$q = q_i(X) \qquad \text{for } t = t_0 \qquad (1.d)$$

Ablation period ($t_{ab} < t \le t_{ab,f}$):

$$\frac{\partial q(X,t)}{\partial t} = \frac{\partial^2 q(X,t)}{\partial X^2} \qquad (2.a)$$

$$\frac{\partial q}{\partial X} = 0 \qquad \text{at } X = 0 \qquad (2.b)$$

$$q = 1 \qquad \text{at } X = B(t) \qquad (2.c)$$

$$q = q_i(X) \qquad \text{for } t = t_{ab} \qquad (2.d)$$

with the energy balance at the heated surface given by:

$$n\,\frac{dB(t)}{dt} = -Q(t) + \frac{\partial q(B(t),t)}{\partial X} \qquad (2.e)$$

$$B(t_{ab}) = B(t_0) \qquad (2.f)$$

For the pre-ablation period, the subscript i in equation (1.d) refers to the initial temperature distribution. For the ablation and post-ablation periods, the subscript i in equations (2.d) and (1.d), respectively, refers to the known temperature distribution at the end of the preceding period. In order to write the problem above in dimensionless form, the following variables were used:

$$X = \frac{x}{b_0},\ t = \frac{a\,t}{b_0^2},\ q(X,t) = \frac{T(x,t) - T_0(x)}{T_{ab} - T_0(x)},$$

$$Q(t) = \frac{b_0\,q(t)}{k[T_{ab} - T_0(x)]},\ n = \frac{H}{c_P[T_{ab} - T_0(x)]} \qquad (3.a\text{-}e)$$

where a, k, c_p and H are the thermal diffusivity, thermal conductivity, specific heat and heat of ablation of the ablating material, respectively.

INVERSE PROBLEM

The inverse problem under picture in this paper is concerned with the estimation of the surface heat flux $Q(t)$ by using the transient measurements of M temperature sensors located inside the slab at positions X_m, $m = 1,\ldots,M$. Furthermore, we assume available for the inverse analysis the transient measurements of a sensor capable of measuring the position of the ablating surface, during the time interval $0 \le t \le t_f$. All other quantities appearing in the formulation of the physical problem are assumed to be exactly known; but, the measurements may contain random errors, which are assumed here to be additive, uncorrelated, normally distributed, with zero mean and known standard deviation.

The unknown surface heat flux is estimated separately for the pre-ablation, ablation and post-ablation periods. The solution of the inverse problem is obtained through the minimization of the following functionals, for the no-ablation periods and for the ablation period, respectively:

$$S_1[Q(t)] = \sum_{m=1}^{M} \int_{t=t_0}^{t_0+dt_0} \{q[X_m, t; Q(t)] - m_m(t)\}^2\, dt \qquad (4.a)$$

$$S_2[Q(t)] = \sum_{m=1}^{M} \int_{t=t_{ab}}^{t_{ab,f}} w_1(t)\{q[X_m, t; Q(t)] - m_m(t)\}^2 dt + \int_{t=t_{ab}}^{t_{ab,f}} w_2(t)\{B_{est}[t; Q(t)] - B_{\exp}(t)\}^2 dt \qquad (4.b)$$

where $m_m(t)$ are the measured temperatures and $q[X_m, t; Q(t)]$ are the estimated temperatures at the measurement locations X_m, $m = 1,\ldots,M$, while $B_{\exp}(t)$ and $B_{est}[t; Q(t)]$ are the measured and estimated surface positions, respectively. In equation (4.b), $w_1(t)$ and $w_2(t)$ are weighting functions for the temperature and surface position measurements, respectively.

ITERATIVE PROCEDURE

The iterative procedure of the conjugate gradient method, applied for the estimation of the unknown function $Q(t)$, is given by:

$$Q^{k+1}(t) = Q^k(t) + b^k d^k(t) \tag{5}$$

where the superscript k denotes the number of iterations, b^k is the search step size and $d^k(t)$ is the direction of descent.

The direction of descent $d^k(t)$ is a conjugation of the gradient direction with previous directions of descent. It is given in the following general form:

$$d^k(t) = -\nabla S[Q^k(t)] + g^k d^{k-1}(t) + y^k d^q(t) \tag{6}$$

where g^k and y^k are conjugation coefficients. The superscript q in equation (6) denotes the iteration number where a *restarting strategy* is applied to the iterative procedure of the conjugate gradient method [8]. In this paper, Powell-Beale's version of the conjugate gradient method is used due to its superior robustness for the solution of non-linear inverse problems [8,9].

For the numerical implementation of the iterative procedure of the conjugate gradient method, two auxiliary problems are required, namely the sensitivity problem and the adjoint problem, as described next.

SENSITIVITY PROBLEM AND SEARCH STEP SIZE

The *sensitivity problem* can be obtained by assuming that the temperature $q(X,t)$ is perturbed by an amount $\Delta q(X,t)$, when the unknown heat flux $Q(t)$ is perturbed by $\Delta Q(t)$. Note that $B(t)$ is also perturbed by $\Delta B(t)$ during the ablation period, as a result of the perturbation $\Delta Q(t)$. By replacing $q(X,t)$ by $[q(X,t) + \Delta q(X,t)]$, $B(t)$ by $[B(t) + \Delta B(t)]$ and $Q(t)$ by $[Q(t) + \Delta Q(t)]$ in the direct problem given by equations (1) and (2), and then subtracting the original direct problem from the resulting expressions, the following *sensitivity problems* are obtained for the no-ablation and ablation periods, respectively:

No-ablation periods ($t_0 < t \le t_0 + dt_0$):

$$\frac{\partial \Delta q(X,t)}{\partial t} = \frac{\partial^2 \Delta q(X,t)}{\partial X^2} \tag{7.a}$$

$$\frac{\partial \Delta q}{\partial X} = 0 \qquad \text{at } X = 0 \tag{7.b}$$

$$\frac{\partial \Delta q}{\partial X} = \Delta Q(t) \qquad \text{at } X = B(t) \tag{7.c}$$

$$\Delta q = \Delta q_i(X) \qquad \text{for } t = t_0 \tag{7.d}$$

Ablation period ($t_{ab} < t \le t_{ab,f}$):

$$\frac{\partial \Delta q(X,t)}{\partial t} = \frac{\partial^2 \Delta q(X,t)}{\partial X^2} \tag{8.a}$$

$$\frac{\partial \Delta q}{\partial X} = 0 \qquad \text{at } X = 0 \tag{8.b}$$

$$\Delta q = -\frac{\partial q}{\partial X}\Delta B(t) \qquad \text{at } X = B(t) \tag{8.c}$$

$$\Delta q = \Delta q_i(X) \qquad \text{for } t = t_{ab} \tag{8.d}$$

where the energy balance (2.e) for the perturbed functions is given by:

$$n\frac{d\Delta B(t)}{dt} - \frac{\partial^2 q}{\partial X^2}\Delta B(t) = -\Delta Q(t) + \frac{\partial \Delta q}{\partial X} \tag{8.e}$$

$$\Delta B(t_{ab}) = \Delta B(t_0) \tag{8.f}$$

Expressions for the search step size b^k can be obtained by minimizing the functional $S[Q^{k+1}(t)]$ with respect to b^k [6,7].

ADJOINT PROBLEM AND GRADIENT EQUATION

The adjoint problem is derived by multiplying the governing equations of the direct problem by Lagrange multipliers, integrating in the spatial and time domains that they are valid and then adding the resultant equation to the functional. Therefore, one single Lagrange multiplier, namely $l(X,t)$, is necessary for the no-ablation periods, so that the extended functional is given by:

$$S[Q(t)] = \sum_{m=1}^{M} \int_{t=t_0}^{t_0+dt_0} \{q[X_m,t;Q(t)] - m_m(t)\}^2 dt + \int_{X=0}^{B(t)} \int_{t=t_0}^{t_0+dt_0} l(X,t)\left[\frac{\partial q}{\partial t} - \frac{\partial^2 q}{\partial X^2}\right] dt\, dX \tag{9}$$

On the other hand, two Lagrange multipliers are required for the ablation period, because the surface position is calculated with the energy balance given by equation (2.e). Therefore, the extended functional for the ablation period is given by:

$$S[Q(t)] = \sum_{m=1}^{M} \int_{t=t_{ab}}^{t_{ab,f}} w_1(t)\{q[X_m,t;Q(t)] - m_m(t)\}^2 dt +$$
$$+ \int_{t=t_{ab}}^{t_{ab,f}} w_2(t)\{B_{est}[t;Q(t)] - B_{\exp}(t)\}^2 dt +$$
$$+ \int_{X=0}^{B(t)} \int_{t=t_{ab}}^{t_{ab,f}} l(X,t)\left[\frac{\partial q}{\partial t} - \frac{\partial^2 q}{\partial X^2}\right] dt\, dX +$$
$$+ \int_{t=t_{ab}}^{t_{ab,f}} f(t)\left[n\frac{dB(t)}{dt} + Q(t) - \frac{\partial q}{\partial X}\right]_{X=B(t)} dt \qquad (10)$$

Expressions for the variation $\Delta S[Q(t)]$ of the functional $S[Q(t)]$ can be developed by assuming that $q(X,t)$ is perturbed by $\Delta q(X,t)$, as well as $B(t)$ is perturbed by $\Delta B(t)$, when $Q(t)$ is perturbed by $\Delta Q(t)$. The variation $\Delta S[Q(t)]$ gives the directional derivative of $S[Q(t)]$ in the direction of the perturbation $\Delta Q(t)$ [6]. By replacing $q(X,t)$ by $[q(X,t) + \Delta q(X,t)]$, $Q(t)$ by $[Q(t) + \Delta Q(t)]$, $B(t)$ by $[B(t) + \Delta B(t)]$ and $S[Q(t)]$ by $\{S[Q(t)] + \Delta S[Q(t)]\}$ in equations (9,10), subtracting from the resulting expressions the original equations (9,10), performing some lengthy but straightforward manipulations and letting the terms containing $\Delta q(X,t)$ and $\Delta B(t)$ to go to zero, the following *adjoint problems* are obtained:

No-ablation periods ($t_0 < t \le t_0 + dt_0$):

$$\frac{\partial l(X,t)}{\partial t} + \frac{\partial^2 l(X,t)}{\partial X^2} = G(t) \qquad (11.a)$$

$$\frac{\partial l}{\partial X} = 0 \qquad \text{at } X = 0 \qquad (11.b)$$

$$\frac{\partial l}{\partial X} = 0 \qquad \text{at } X = B(t) \qquad (11.c)$$

$$l = 0 \qquad \text{for } t = t_0 + dt_0 \qquad (11.d)$$

where

$$G(t) = 2\sum_{m=1}^{M} \{q[X_m,t;Q(t)] - m_m(t)\}\, d(X - X_m) \qquad (11.e)$$

Ablation period ($t_{ab} < t \le t_{ab,f}$):

$$\frac{\partial l(X,t)}{\partial t} + \frac{\partial^2 l(X,t)}{\partial X^2} = w_1(t)G(t) \qquad (12.a)$$

$$\frac{\partial l}{\partial X} = 0 \qquad \text{at } X = 0 \qquad (12.b)$$

$$l = -f(t) \qquad \text{at } X = B(t) \qquad (12.c)$$

$$l = 0 \qquad \text{for } t = t_{ab,f} \qquad (12.d)$$

and

$$n\frac{df(t)}{dt} - \frac{\partial^2 q}{\partial X^2} f(t) = -\frac{\partial l}{\partial X}\frac{\partial q}{\partial X} + w_2(t)G^*(t) \qquad (12.e)$$

$$f(t_{ab,f}) = 0 \qquad (12.f)$$

where $G^*(t)$ is given by:

$$G^*(t) = 2\{B_{est}[t;Q(t)] - B_{\exp}(t)\} \qquad (12.g)$$

In the limiting process used to obtain the adjoint problem, the following expression is obtained for the gradient of the functional, by assuming that the unknown function $Q(t)$ belongs to the space of square integrable functions in the time domain of interest:

$$\nabla S[Q(t)] = -l[B(t),t] \qquad (13)$$

which is valid for both no-ablation and ablation periods.

The use of the conjugate gradient method may result on stable solutions if the *Discrepancy Principle* [6,7] is used to specify the tolerance for the stopping criterion of the iterative procedure. In the Discrepancy Principle, the solution is assumed to be sufficiently accurate when the difference between measured and estimated quantities is of the order of magnitude of the measurement errors.

RESULTS AND DISCUSSIONS

We now examine the present inverse analysis approach for the estimation of the unknown boundary heat flux at the surface of ablating materials. For all test-cases examined below, the material was assumed to be Teflon, with properties $r = 1.922\text{x}10^3$ *kg/m*^{3}, $k = 0.22$ *W/(m K)*, $c_P = 1.256\text{x}10^3$ *J/(kg K)*, $a = 9.11\text{x}10^{-8}$ *m*^{2}*/s*, $H = 2.326\text{x}10^6$ *J/kg* and $T_{ab} = 833$ *K* [10]. However, we note that, qualitatively, similar results were obtained with other ablating materials. The thickness of the slab was taken as 0.020 *m*. The duration of the experiment was supposed to be 400 *s* and the frequency of measurements, for both temperature and surface position, was 10 *Hz*. In order to simulate the heat flux at the surface of the vehicle during atmospheric re-entry, a triangular heat flux, with different peak-fluxes, was used to generate the simulated temperature and surface position measurements. Two different levels of

measurement errors examined in this work included: (i) errorless measurements and (ii) temperature measurements with standard deviation equal to 0.01 (1% of the ablation temperature) and surface position measurements with standard deviation equal to 0.008. For the standard deviation of the surface position measurements, it was taken into account the sensitivity of instruments available for this purpose [5].

Let us first consider the case involving a triangular heat flux with peak-flux of 50 kW/m^2. Figure 1.a presents a comparison of the exact heat flux used to generate the simulated measurements with the estimated functions obtained with errorless measurements, as well as with measurements containing random errors. The initial guess used for the iterative procedure of the conjugate gradient method is also presented in this figure. The results shown in figure 1.a were obtained by using only temperature measurements in the inverse analysis, of a sensor located at $X = 0.9$. Furthermore, such results were obtained by using a single objective functional for the whole-time domain, i.e., for $0 \le t \le t_f$ [4]. Figure 1.a shows that quite accurate results could be obtained in this case, by using only temperature measurements in the inverse analysis. Such is the case because, for a peak-flux of 50 kW/m^2, the portion of material removed by ablation is quite small. In fact, after ablation finishes, the surface position was at X=0.99 for this case.

Figure 1.b shows the transient variation of the temperature sensitivity function, $\Delta q(X,t)$, in the slab, for different positions. Such sensitivity function was obtained from the solution of the sensitivity problems (7,8) for a unitary perturbation of the heat flux, i.e., $\Delta Q(t) = 1$. Also, the solution of the direct problem with the exact heat flux was used to compute the surface position $B(t)$, as well as $\partial q / \partial X$ and $\partial^2 q / \partial X^2$ at $B(t)$, which are required for the solution of the sensitivity problem during the ablation period (see equations (8.a-f)). It is interesting to note in equations (7,8) that, during the pre-ablation period, the solution of the sensitivity problem is independent of the unknown function $Q(t)$. On the other hand, during the ablation period, and consequently during the post-ablation period, the solution of the sensitivity problem depends on several quantities obtained from the solution of the direct problem, because of the non-linear character of the function estimation problem under picture. Note in figure 1.b the beginning of ablation at $t = 0.043$, when the sensitivity at the position $X = 1$ vanishes. After ablation begins, a decrease of the magnitude of $\Delta q(X,t)$ takes place for the position $X = 0.8$ and the rate in the increase of the sensitivity at $X = 0.6$ is reduced. These quantities increase again at larger times, after ablation finishes.

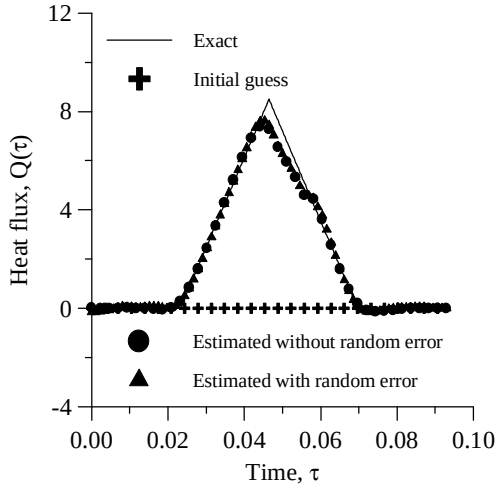


Figure 1.a. Solution of the inverse problem for a peak-flux of 50 kW/m^2.

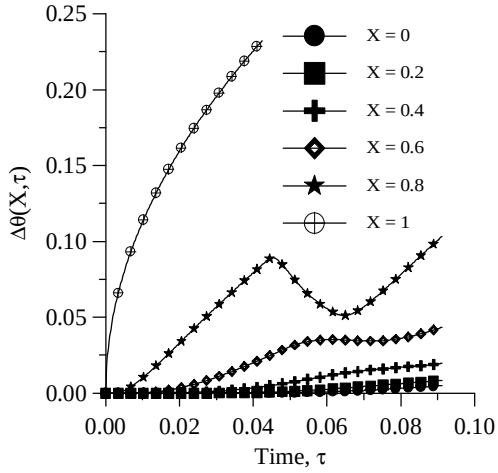


Figure 1.b. Temperature sensitivity function for a peak-flux of 50 kW/m^2.

Figure 2.a presents the transient variation of the temperature sensitivity function, $\Delta q(X,t)$, in the slab, for different positions, for a peak-flux of 100 kW/m^2. The behavior of the temperature sensitivity function for a peak-flux of 100 kW/m^2 is very similar to that for 50 kW/m^2 (see figure 1.b), because the portion of material removed

during ablation is also small. For 100 kW/m^2, the surface position after ablation finishes is at $X = 0.94$. Note also in figures 1.b and 2.a that the sensitivity curves for the two peak-fluxes are identical during the pre-ablation periods, because the estimation problem is linear until ablation begins, i.e., the sensitivity problem (7) does not depend on the unknown function $Q(t)$.

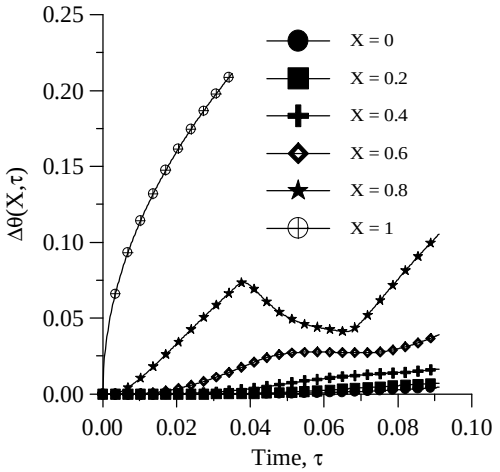


Figure 2.a. Temperature sensitivity function for a peak-flux of 100 kW/m^2.

It was also possible to obtain accurate estimates for the unknown function for a peak-flux of 100 kW/m^2, by using only temperature measurements in the inverse analysis, as illustrated in figure 2.b. The temperature sensor was located at $X = 0.9$. The results shown in figure 2.b were obtained by estimating the heat flux separately for the pre-ablation, ablation and post-ablation periods, with $w_1 = 1$ and $w_2 = 0$ in equation (4.b). Figure 2.b shows that stable results were obtained for measurements containing random errors, but the peak-flux was overestimated, even when errorless measurements were used in the inverse analysis.

Figure 2.c presents the sensitivity function, $\Delta B(t)$, of the position of the surface with respect to variations in the unknown heat flux, as well as the velocity of the moving surface, for a peak-flux of 100 kW/m^2. Note in this figure that the sensitivity of the position of the surface immediately increases in magnitude when ablation begins and suddenly becomes zero when ablation ends. As expected, the sensitivity of the position of the surface is zero during pre- and post-ablation periods.

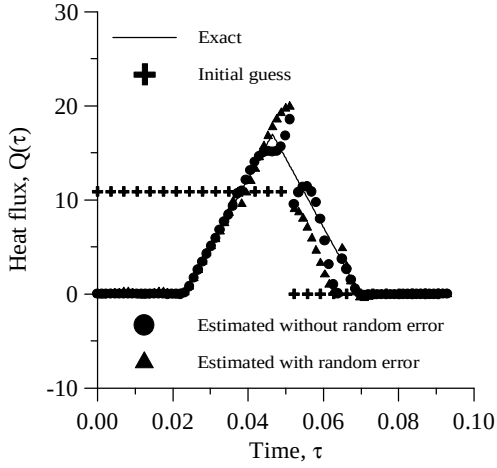


Figure 2.b. Solution of the inverse problem for a peak-flux of 100 kW/m^2 obtained with $w_1 = 1$ and $w_2 = 0$ in equation (4.b).

Figure 2.d presents the estimated functions for a peak-flux of 100 kW/m^2, obtained by using temperature measurements during the no-ablation periods, and surface position measurements during the ablation period, i.e., $w_1 = 0$ and $w_2 = 1$ in equation (4.b). The temperature sensor was located at $X = 0.9$. We note that the heat flux was exactly recovered when errorless measurements were used in the inverse analysis, which was not the case when only temperature measurements were assumed available (see figure 2.b). A comparison of figures 2.b and 2.d reveals an increase in the oscillations of the solution after ablation began, when surface position measurements containing random errors were used in the inverse analysis.

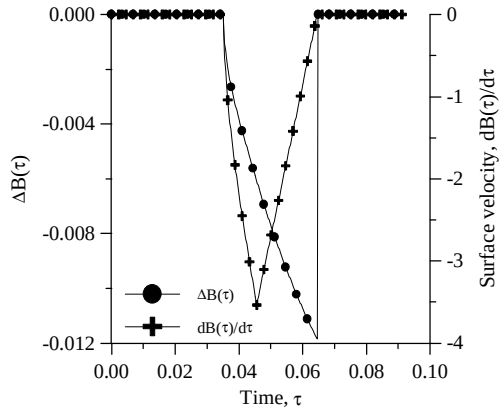


Figure 2.c. Surface position sensitivity function for a peak-flux of 100 kW/m^2.

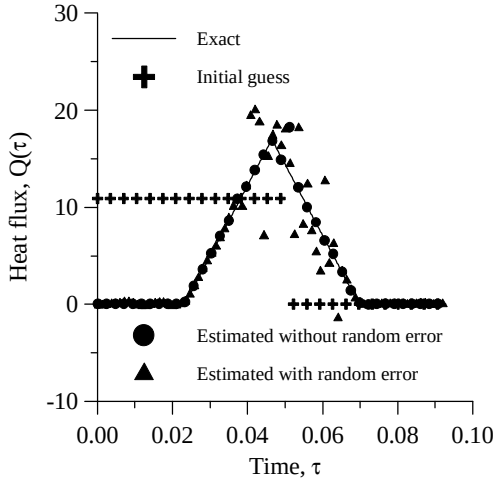


Figure 2.d. Solution of the inverse problem for a peak-flux of 100 kW/m^2 obtained with $w_1 = 0$ and $w_2 = 1$ in equation (4.b).

We now examine a more strict case involving a peak-flux of 500 kW/m^2. Figure 3.a presents the transient variation of the temperature sensitivity function, $\Delta q(X,t)$, in the slab, for such peak-flux, for different positions. For this case, the final surface position after ablation finished was at $X = 0.56$. A comparison of figures 2.a and 3.a shows an increase in the maximum magnitude of $\Delta q(X,t)$ when the peak-flux increases from 100 kW/m^2 to 500 kW/m^2. However, the sudden increase in $\Delta q(X,t)$ takes place when the ablating surface is quite close to the measurement position. Obviously, temperature measurements do not contribute with useful information for the estimation of the surface heat flux, after the ablating surface passes through the temperature sensor position and burns the sensor. Therefore, for large magnitudes of the heat flux, when the portion of the material removed during ablation is large, it appears that the use of a large number of temperature sensors, spread along the slab, would be required for the inverse analysis. However, even with such an approach the accuracy of the inverse problem solution could not be assured, because the temperature sensitivity function becomes large only for a very short period of time, exactly before the sensor gets burned. On the other hand, reasonably accurate results could be obtained when measurements of the surface position were used in the inverse analysis, as described below.

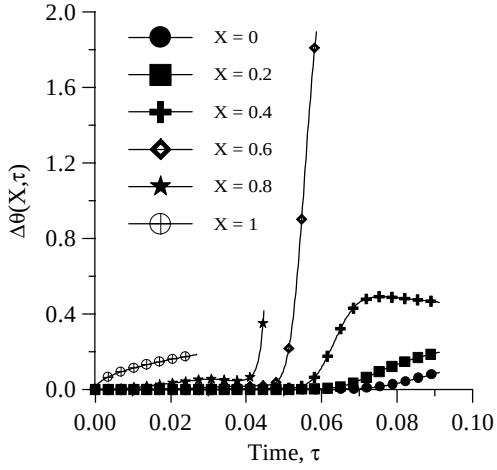


Figure 3.a. Temperature sensitivity function for a peak-flux of 500 kW/m^2.

Figure 3.b presents the sensitivity function, $\Delta B(t)$, of the position of the surface with respect to variations in the unknown heat flux, as well as the velocity of the moving surface, for a peak-flux of 500 kW/m^2. Note in figures 2.c and 3.b that, differently from the results for a peak-flux of 100 kW/m^2, the response of the sensitivity of the position of the surface is lagged with respect to the surface velocity for 500 kW/m^2. Because of such a lag, accurate results could be obtained for the estimated function in this case, only when the initial guess for the iterative procedure was reasonably close to the exact function, as illustrated in figure 3.c.

The initial guess used for the results shown in figure 3.c was obtained by neglecting the temperature gradient in equation (2.e) and by approximating the velocity of the surface in this equation by using the surface position measurements; the velocity $dB(t)/dt$ was calculated by using a backwards finite-differences approximation in time. The results presented in figure 3.c were obtained by using temperature measurements during the pre- and post-ablation periods, and surface position measurements during the ablation period, i.e., $w_1 = 0$ and $w_2 = 1$ in equation (4.b). Two temperature sensors were used in this case, located at the positions $X = 0.5$ and $X = 0.9$. The sensor at $X = 0.9$ was burned during the ablation period, and the sensor at $X = 0.5$ was required for the estimation of the heat flux during the post-ablation period. We note that the results obtained by using only temperature measurements for this case were completely inaccurate.

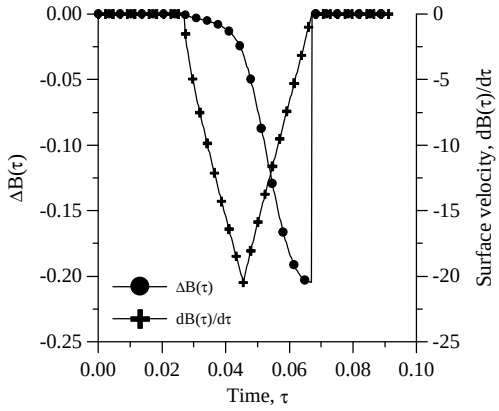


Figure 3.b. Surface position sensitivity function for a peak-flux of 500 kW/m^2.

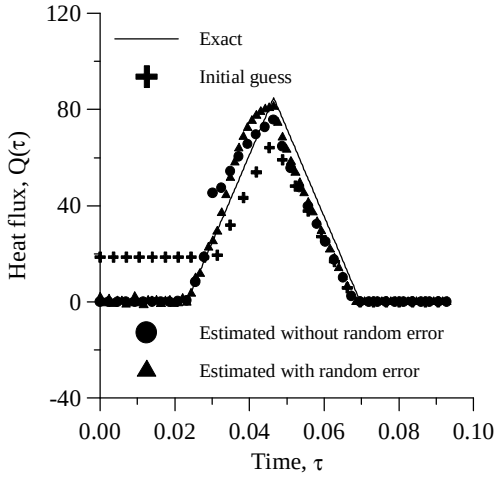


Figure 3.c. Solution of the inverse problem for a peak-flux of 500 kW/m^2 obtained with $w_1 = 0$ and $w_2 = 1$ in equation (4.b).

CONCLUSIONS

This paper addresses the use of surface position measurements, in addition to temperature measurements, for the solution of the inverse problem of estimating the heat flux at the surface of ablating materials. For heat fluxes of small magnitude, when the velocity of the ablating surface is small, the use of only temperature measurements may result in accurate estimations for the unknown function. On the other hand, for heat fluxes with large magnitudes, temperature measurements do not provide sufficient information for the estimation procedure. Measurements of the position of the ablating surface are then required in order to obtain accurate estimations for the surface heat flux.

ACKNOWLEDGEMENTS

This work was sponsored by AEB (Brazilian Space Agency). Partial support provided by CNPq, CAPEs and FAPERJ is also greatly appreciated. The authors would like to acknowledge useful discussions with Prof. Renato M. Cotta (PEM/COPPE/UFRJ) as well as with Dr. Luis Loures and Dr. José Bezerra (IAE/CTA).

REFERENCES

1. D. V. F. M. Rey Silva and H. R. B. Orlande, Estimation of the thermal properties of ablating materials, *Proceedings of the International Symposium on Inverse Problems in Engineering*, Nagano (2001)
2. F. R. DeJarnette, H. H. Hamilton, K. J. Weilmuenster and F. M. Cheatwood, A review of some approximate methods used in aerodynamic heating analyses, *AIAA J. Thermophysics*, **1**, No. 1, pp. 4-12 (1987)
3. D. Beneulen, G. Jahn and G. Reich, The surface protected ablator heat shield of Mirka: development, qualification and flight results, *3rd European Workshop on Thermal Protection Systems*, The Netherlands (1998)
4. A. P. de Oliveira and H. R. B. Orlande, Estimation of the heat flux at the surface of ablating materials, *Proceedings of the International Symposium on Inverse Problems in Engineering*, Nagano (2001)
5. G. K. Noffz and M. P. Bowman, Design and laboratory validation of a capacitive sensor for measuring the recession of a thin-layered ablator, *NASA Technical Memorandum 4777* (1996)
6. O. M. Alifanov, *Inverse heat transfer problems*, Springer-Verlag, New York, 1994.
7. M. N. Ozisik and H. R. B. Orlande, *Inverse heat transfer: fundamentals and applications*, Taylor & Francis, New York, 2000.
8. M. J. D. Powell, Restart procedures for the conjugate gradient method, *Math. Program.*, **12**. pp. 241-254 (1977)
9. M. J. Colaço and H. R. B. Orlande, Comparison of different versions of the conjugate gradient method of function estimation, *Numerical Heat Transfer*, **36**, pp. 229-249 (1999)
10. B. F. Blackwell, Numerical prediction of one-dimensional ablation using a finite control volume procedure with exponential differencing, *Numerical Heat Transfer*, **14**, pp. 17-34 (1988).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

EVALUATION OF PARAMETERS USED IN A MULTIMEDIA ENVIRONMENTAL MODEL – APPLICATION TO GUANABARA BAY

Mariella J. Berrocal Tito[1]
[1]*Departamento de Ciência dos Materiais e Metalurgia, Pontifícia Universidade Católica do Rio de Janeiro,Rua. Marquês do São Vicente, 225 – Gávea, C.P 22453-900, RJ, Brazil.*
mabet99@yahoo.com

Roberto José de Carvalho[1]
rjcar@dcmm.puc-rio.br
Nilson Costa Roberty[2]
[2]*Nuclear Engineering Program, COPPE/UFRJ, CP 68509, CEP 21945-970, Rio de Janeiro, RJ, Brazil, nilson@lmn.con.ufrj.br*

ABSTRACT

The Maximum Entropy and Levenberg-Marquardt methods were applied to assess the direct emissions of chemical pollutants and other input parameters used in multimedia environmental models. An environment consisting of air, water sediment and biota was considered. Results obtained by using experimental concentration data of some hydrocarbons measured in a type of mussel ("*mexilhão perna-perna*") in Guanabara Bay are presented.

NOMENCLATURE

A_S	sediment area, m^2.
A_W	water area, m^2.
C_A	chemical concentration in air, mol/m^3.
C_M	chemical concentration in mussel, mol/m^3
C_{SS}	chemical concentration in dry sediment, mol/m^3.
C_{ST}	total chemical concentration in sediment, mol/m^3.
C_{VFQ}	chemical concentration in aerosol particles, g/m^3.
C_{VPW}	chemical concentration in water particles, g/m^3.
C_{VPX}	chemical concentration in inflow water particles, g/m^3.
C_W	chemical concentration in water, mol/m^3.
C_{WI}	chemical concentration in inflow water, mol/m^3.
C_{WT}	total chemical concentration in water, mol/m^3.
D_{ENP}	density of water particles, g/m^3.
D_{ENQ}	density of aerosol particles, g/m^3.
D_{ENS}	density of sediment particles, g/m^3.
E_W	emission rate to water, mol/h.
G_{BN}	solids burial rate, g/m^{2}h.
G_{DN}	solids deposition rate, g/m^{2}h.
G_{RN}	solids resuspension rate, g/m^{2}h.
G_I	water inflow rate, m^3/h.
G_J	water outflow rate, m^3/h.
H	Henry's law constant, Pa.m^3/mol
K_{AW}	air-water partition coefficient.
K_{OC}	organic carbon-water partition coefficient.
K_{OW}	octanol-water partition coefficient.
K_{SW}	sediment-water partition coefficient.
K_T	sediment-water mass transfer coefficient, m/h.
K_V	volatilization mass transfer coefficient (air side), m/h.
K_{VW}	volatilization mass transfer coefficient (water side), m/h.
O_{RGI}	fraction of organic carbon in inflow water particles.
O_{RGP}	fraction of organic carbon in water particles.
O_{RGR}	fraction of organic carbon in resuspended sediment solids.
O_{RGS}	fraction of organic carbon in sediment solids.
Q	scavenging ratio.

R	gas constant, Pa.m^3/molK.
R_{AIN}	rain rate, m/h.
T_{DS}	degradation half life in sediment, h.
T_{DW}	degradation half life in water, h.
T_K	room temperature, K.
T_{MK}	chemical melting point, K.
V_{DEPA}	aerosol dry deposition velocity, m/h.
V_{FS}	volume fraction of particles in surface sediment.
V_S	active sediment volume, m^3.
V_W	water volume, m^3.

INTRODUCTION

Guanabara Bay is a highly urbanized tropical estuary located in the state of Rio de Janeiro, Brazil. The total population in the vicinity of the bay is close to 10 million inhabitants. Its drainage basin encompasses an area of about $4.0x10^9$ m^2 including 12 municipalities, the city of Rio de Janeiro and the second largest industrial comp lex in Brazil. Chemical, petrochemical, food processing, metallurgical plants and textile industries are the main economical activities. The bay presents anoxic conditions being heavily contaminated by organic and inorganic chemicals released daily in its waters. Its main sources of pollution are the intense discharges of domestic and industrial effluents [1]. Many industries produce effluents containing significant amounts of polycyclic aromatic hydrocarbons (PAHs). Several PAHs are mutagenics and/or carcinogenics and are particularly important due to the contamination of seafood and the consequent possibility of human exposure. Monitoring programs are necessary to assess concentration levels, persistence, fate, transport/transformation rates and partition of potentially harmful chemicals among the environmental media. In the absence of monitoring, multimedia environmental models often provide the only way to predict the approximate concentration of the chemicals in the media. The multimedia model proposed to determine the fate of PAHs in Guanabara Bay is the "Quantitative Water Air Sediment Interaction" (QWASI) developed by Mackay and coworkers [2,3]. The purpose of this work is to evaluate some input parameters of the model as mass transfer coefficients, emission rates to the media, etc. using the inverse problem technique. The direct problem consists of the solution of the advective-diffusive transport equations of the model, which given the contaminant concentrations in water and sediment. The solution of the inverse problem is accomplished by the Generalized Maximum Entropy and Levenberg-Marquardt methods using concentrations of different PAHs in water and sediments of Guanabara Bay. The values were approximated from the measured concentrations in a type of mussel ("*mexilhão perna-perna*").

MATHEMATICAL FORMULATION OF THE DIRECT PROBLEM

The direct problem corresponds to the advective-diffusive chemical transport described by the mass balance equations for the media or compartments that comprise the environment. These compartments are air (consisting of well-mixed air, aerosols and rain), water (consisting of well-mixed water and particles of organic and mineral matter origin), sediment (well mixed and homogeneous) and biota (only one aquatic specie), [2,3].

The overall steady states mass balance equations for the water and sediment compartments are given respectively by

$$V_W \frac{dC_{WT}}{dt} = [1]+[2]+[3]+[4]+[5]+[6]+[7]+[8]+ [9]-[10]-[11]-[12]-[13]-[14]-[15]=0 \quad (1)$$

$$V_S \frac{dC_{ST}}{dt} = [12]+[15]-[5]-[6]-[16]-[17] = 0 \quad (2)$$

The numbered terms on the right side of the equations, above, represent the processes, involving the chemical. See Fig. 1. They are:

[1] absorption $C_A A_W K_V / K_{AW}$

[2] wet particles deposition $6x10^6 C_A R_{AIN} A_W F_{RAT} D_{ENQ} Q/(V_P C_{VFQ})$

3] dry particles deposition $6x10^6 C_A V_{DEPA} A_W D_{ENQ} F_{RAT} /(V_P C_{VFQ})$

[4] rain dissolution $C_A R_{AIN} A_W / K_{AW}$

[5] sediment resuspension $C_{SS} O_{RGR} G_{RN} A_W /(O_{RGS} D_{ENS})$

[6] sediment-water diffusion $C_{SS} K_T A_S / K_{SW}$

[7] water inflow $C_{WI} G_I K_{WPI} (1+K_{OC}*O_{RG})/(1+K_{OC}*O_{RGP})$

[8] water particle inflow $C_{WI}G_I C_{VPX}/D_{ENP}$

[9] direct emissions E_W

[10] volatilization $C_W K_V A_W$

[11] water transformation $0.693 \cdot C_{WT} V_W / T_{DW}$

[12] sediment deposition $\ldots C_W G_{DN} A_W K_{OC} O_{RGP}$

[13] water outflow $C_W G_J$

[14] water particles outflow $C_W G_J C_{VPW} K_{OC} O_{RGP}$

[15] water-sediment diffusion $. C_W K_T A_S$

[16] sediment transformation $0.693 \cdot C_{ST} V_S / T_{DS}$

[17] sediment burial $C_{SS} G_{BN} A_W / D_{ENS}$

The solution of equations (1) and (2) can be solved, given the chemical concentrations in sediment and water according to equations (3) and (4).

$$C_{SS} = ([1]+[2]+[3]+[4]+[7]+[8]+E_W)(G_{DN}A_W K_{OC} O_{RGP} + K_T A_S)/\{(K_V A_W + G_J + G_{DN}A_W K_{OC} O_{RGP} + G_J C_{VPW} K_{OC} O_{RGP} + 0.693 \cdot \frac{F_{WT} V_W}{T_{DW}})(\frac{O_{RGR} G_{RN} A_W O_{RGS}}{D_{ENS}} + \frac{K_T A_S}{K_{SW}} + \frac{C_{ST} V_S}{T_{DS}} + \frac{C_{SS} G_{BN} A_W}{D_{ENS}})(K_V A_W + 0.693 * \frac{F_{WT} V_W}{T_{DW}} + G_{DN} A_W K_{OC} O_{RGP} + G_J + G_J C_{VPW} K_{OC} O_{RGP} + G_{DN} A_W K_{OC} O_{RGP} + V_S)\} \quad (3)$$

$$C_W = C_{SS}(\frac{O_{RGR} G_{RN} A_W}{O_{RGS} D_{ENS}} + \frac{K_T A_S}{K_{SW}} + 0.693 \cdot \frac{V_S F_{ST}}{T_{DS}} + \frac{G_{BN} A_W}{D_{ENS}})/(G_{DN} A_W K_{OC} O_{RGP} + K_T A_S) \quad (4)$$

The auxiliary correlations used were:

$K_{AW} = H / RT_K$

$K_{SW} = K_{OC} O_{RGS} D_{ENS}$

if $T_{MK} > T_K$, $F_{RAT} = \exp(6.79 \cdot (1 - T_{MK}/T_K))$

else $F_{RAT} = 1$

$F_{WT} = 1 - (C_{PW} + K_{OC} O_{RGI} D_{ENP} C_{PW}) / D_{ENP}$

$F_{ST} = V_{FS} + \frac{1 - V_{FS}}{K_{SW}}$

$K_{OC} = 4.1 \cdot 10^{-7} K_{OW}$

$K_V = K_{VW} K_{VA} K_{AW} / (K_{VA} K_{AW} + K_{VW})$

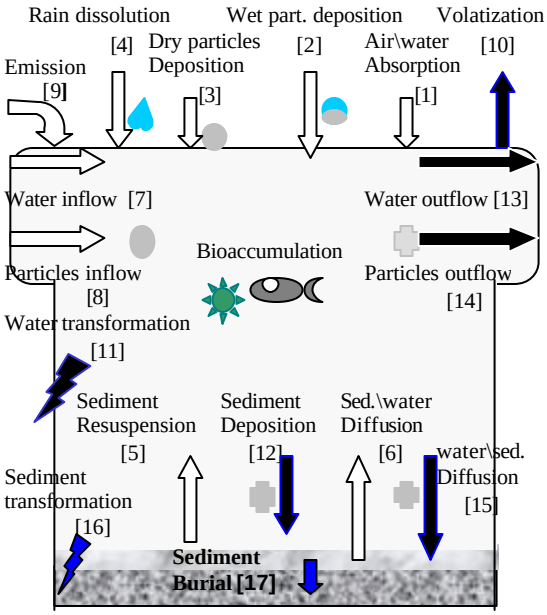


Figure 1. Processes considered in the direct problem.

INVERSE PROBLEM.

This work assumes the following considerations

1. The equations (2) and (3), we observed that the concentration chemical in the water and in the sediment depends on several parameters some related directly to the pollutant and other to the medium. Most of the parameters were obtained from chemical tables [4].
2. El value of the individual concentration chemical in water, was approximate by

$$C_W = C_M (\sum PAHs \text{ in water} / \sum PAHs \text{ in mussel}) \quad (5)$$

$\sum PAHs$ is the total concentrations of measured PAHs. The Table 1 presents the concentrations of measured PAHs in the mussel and the sediment, and calculated for the water in the Guanabara Bay.

3. There are five PAHs (anthracene, fluoranthene, chrysene, pyrene and benzo(a)pyrene.) for each PAHs we have the equations related to the concentration of the PAHs in the water and in the sediment, equations. (2) and (3).
4. For each PAHs we have two considered unknowns parameters: direct emission and the diffusion coefficient.

Table 1. Concentration of PAHs in the Guanabara Bay

	Mussel	Sediment	Water
A	35,08 ng/g	332482 ng/l	90,85 ng/l
F	70,15 ng/g	531376 ng/l	181,7 ng/l
C	23,20 ng/g	729530 ng/l	59,6 ng/l
P	21,46 ng/g	716512 ng/l	55,8 ng/l
B	11,7 ng/g	798232 ng/l	30,3 ng/l
$\sum PAHs$	271 ng/g		701 ng/l

A = Anthracene, F = Fluoranthene, C = Chrysene, P = Pyrene and B = Benzo(a)pyrene.

In general in a system: if n is it number of chemical strengths and p is the total number the unknowns therefore $2n$ is the number of the available equations and p is the sum of: $2n$ the number the unknowns related to the pollutants and m the number of unknown properties of the medium.

The problem of estimate p unknowns, having only $2n$ equations available may be solved with the Method of Generalized Maximum Entropy or the method of Tikhonov functional minimization. A variant of the Tikhonov functional minimization problem is the Levenberg Marquardt's method.

We can write z_j the unknowns like of the system, with $j = 1,2...\ p$. All considered unknowns are presented in the table 4, the last five are related to the properties physics and chemical of the Guanabara Bay

Method of Generalized Maximum Entropy

As the number of unknowns p is superior to the number of available equations or data (water and sediment concentration for each chemical) $2n$, there are a large number of possible solutions for the system represented by equations (1) and (2). Therefore, the inverse problem is solved as an optimization problem where a criterion is defined for the choice of one of the possible solutions.

Let Z be the vector formed by all unknowns of the system for n chemicals.

$$\mathbf{Z} = \{z_1, z_2, ...z_p\} \tag{6}$$

The known data, chemical concentrations in water and sediment, are denoted by C_J and can be calculated by a function $f_J(\mathbf{Z})$ as

$$C_J = f_J(\mathbf{Z}) \tag{7}$$

Defining

$$F_J(\mathbf{Z}) = C_J - f_J(\mathbf{Z}) \tag{8}$$

Considering the Lagrange multipliers method, the Lagrangian is expressed as

$$L_q(\mathbf{Z}, \mathbf{Z_0}, l_J) = D_{B,q}(\mathbf{Z},\mathbf{Z_0}) + \sum_{J=1}^{2n} l_J \cdot F_J(\mathbf{Z}) \tag{9}$$

l_J is the Lagrange multiplier for the J equation, $\mathbf{Z_O}$ is an initial or starting value of $\mathbf{Z}$, $D_{B,q}(\mathbf{Z},\mathbf{Z_0})$ is the Bregman distance between $\mathbf{Z}$ and $\mathbf{Z_0}$ [5], defined as

$$D_{B,q}(\mathbf{Z},\mathbf{Z_0}) = D_{B,q}(h_{B,q}(\mathbf{Z}), h_{B,q}(\mathbf{Z_0})) =$$
$$= h_{B,q}(\mathbf{Z}) - h_{B,q}(\mathbf{Z_0}) - \langle \nabla h_{B,q}(\mathbf{Z_0}), \mathbf{Z}-\mathbf{Z_0} \rangle \tag{10}$$

where $h_{B,q}$ is the functional momento of nth-order B of the q-discrepancy, defined as

$$h_{B,q}(\mathbf{Z}) = \sum_{i=1}^{p} z_i^B \frac{z_i^q - 1}{q} \tag{11}$$

an approximation when $q = 0$ gives

$$h_{B,q}(\mathbf{Z}) = \sum_{i=1}^{p} z_i^B \ln z_i \tag{12}$$

derived from the measurement of the Sharma and Mittall [6] directed divergence. This functional represents the deviation of an expected value of z, raised to the power q, of a priori information of the concentration measurement given by the inverse problem. The particular case for which $q \to 0$ and B=1 corresponds to the functional entropy

$$h_{1,0}(\mathbf{Z}) = \sum_{i=1}^{p} z_i \ln z_i \tag{13}$$

From equations (10), (11) and (12) results if $q = 0$

$$D_{B,q}(\mathbf{Z},\mathbf{Z_0}) = \sum_{i=1}^{p} z_i^B \ln(z_i) -$$
$$z_{oi}^{B-1}(z_{oi} - B(z_i - z_{oi}))\ln(z_{oi}) -$$

$$z_{oi}^{B+q-1}(z_i - z_{oi}) \qquad (14)$$

else

$$D_{B,q}(\mathbf{Z},\mathbf{Z_0}) = \sum_{i=1}^{p} z_i^B\left(\frac{z_i^q - 1}{q}\right) - z_{oi}^{B-1}(z_{oi} - B(z_i - z_{oi}))\left(\frac{z_{oi}^q - 1}{q}\right) - z_{oi}^{B+q-1}(z_i - z_{oi}) \qquad (15)$$

combining equations (11), (12) and (9), for the case of B = 1

if $q = 0$

$$L_q(\mathbf{Z}, \mathbf{Z_0}, \lambda) = \sum_{i=1}^{p} z_i \ln(z_i) - z_i \ln(z_{oi}) - (z_i - z_{oi}) + \sum_{J=1}^{2n} \lambda_J \cdot F_J(\mathbf{Z}) \qquad (16)$$

else

$$L_q(\mathbf{Z}, \mathbf{Z_0}, \lambda) = \sum_{i=1}^{p} z_i\left(\frac{z_i^q - z_{oi}^q}{q}\right) - z_{oi}^q(z_i - z_{oi}) + \sum_{J=1}^{2n} \lambda_J \cdot F_J(\mathbf{Z}) \qquad (17)$$

Solution of the Inverse Problem by the Generalized Maximum Entropy

As discussed in the previous section, the inverse problem is formulated as an optimization problem [7-10]. Using the Lagrangian given by equations (16) or (17) and varying the value of the parameter *q* a family of algorithms is built.

To obtain the minimum value of the Lagrangian at the critical point the following equation is written

$$\frac{\partial}{\partial z_i} L_q(\mathbf{Z},\mathbf{Z_0},\lambda) = \frac{\partial}{\partial z_i} D_{B,q}(\mathbf{Z},\mathbf{Z_0}) + \sum_{J=1}^{2n} \lambda_J \cdot \frac{\partial}{\partial z_i} F_J(\mathbf{Z}) = 0 \qquad (18)$$

from equations (16) or (17) the unknowns to be determined are the Lagrange multipliers

λ_J, with $J = 1,2,..2n$, $\lambda_J \in \lambda$

and the unknowns of the environmental model

z_i, with $i = 1,2,.. p$, $z_i \in \mathbf{Z}$

that are calculated solving the non-linear system $(2n+p)(2n+p)$ formed by

$$\frac{\partial}{\partial z_i} D_{B,q}(\mathbf{Z},\mathbf{Z_0}) + \sum_{J=1}^{2n} \lambda_J \frac{\partial}{\partial z_i} F_J(\mathbf{Z}) = 0$$

and $F_J(\mathbf{Z}) = 0$

with $i = 1,2...p$ and $J = 1,2\ldots 2n$ (19)

A code was written in the MATLAB environment in order to solve the system described by equation (19). This code provided the necessary tools for optimizing the results. The minimization of the system (19) was achieved with the medium-scale algorithm: SQP, Quasi-Newton, line-search [11].

It becomes evident that the problem of evaluating the unknowns depends on the experimental data (included in C_J), on an initial guess for the unknown, on the chemical properties and on the general physical properties of the environment.

Levenberg-Marquardt Method

Another alternative to solve the problem is to utilize the Tikhonov functional given by

$$\sum_{J=1}^{2n} [F_J(\mathbf{Z})]^2 + \alpha \cdot D_{B,q}(\mathbf{Z},\mathbf{Z_0}) \qquad (20)$$

where α represents the relation between the measured clean signal and the noise found in the measurements. Usually, the value of α is small.

A variant of the Tikhonov functional is the Levenberg Marquardt method. For the case in which B=1 and q =1, this method results in

$$\sum_{J=1}^{2n} [F_J(\mathbf{Z})]^2 + \alpha \cdot [\mathbf{Z} - \mathbf{Z_0}]^2 \rightarrow 0 \qquad (21)$$

The solution of the Levenberg-Marquardt method is based on the functional minimization given by equation (21).

RESULTS

Figures (2), (3) and (4) show the Bregman distance calculated varying *q* for different values of B, evaluated in $\mathbf{Z}$ when $F_J(\mathbf{Z}) = 0$, with $J = 1,2\ldots 2n$. In theses figures, we obtain the minimum *q* value for B=2, B=1 and B=0, considering which the Bregman distance should be greater and equal to zero. Table 2 gives the minimum *q* values obtained for different *B*.

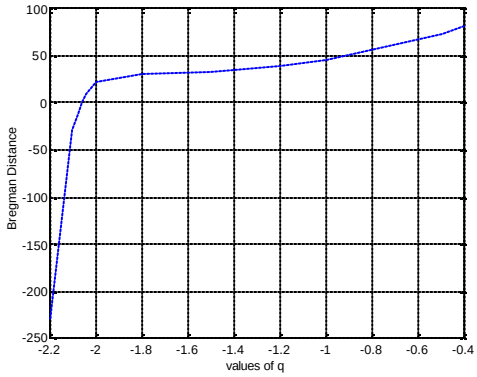


Figure 2. Bregman distance vs *q* for B=2.

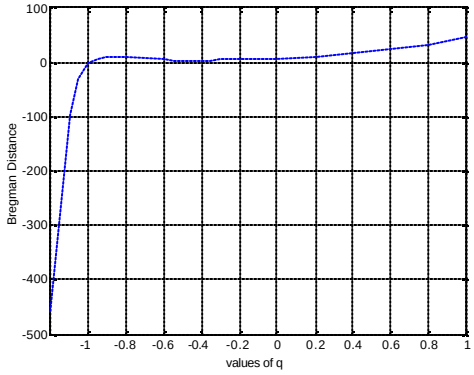


Figure 3. Bregman distance vs *q* for B=1.

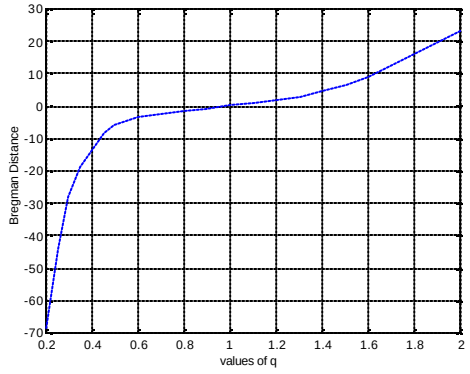


Figure 4. Bregman distance vs *q* for B=0.

Table 2. Minimum *q* values for different *B*

B momentum	q optimum
2	-2
1	-1
0	1

Results obtained by the maximum entropy method

Table 3 shows the physical-chemical properties of some PAHs studied in this work [4].

Table 4. Assumed physical-chemical properties of some PAHs.

Chem.	Molec.mass (g/mol)	Solubility (g/m^3)	Vapor Press. (Pa)
A	178.2	0.064	$3.5x10^{-4}$
F	202.3	0.245	$1.6x10^{-6}$
C	228	$1.8x10^{-3}$	$8.3x10^{-7}$
P	202.3	0.132	$6.1x10^{-4}$
B	252.3	$1.6x10^{-3}$	$7.3x10^{-7}$

Chem.	Half-life decay in sed. (h)	Half-life decay in water (h)	Melting point (^oC))	Log KOW
A	17000	550	216.2	5.33
F	55000	1700	111	5.33
C	55000	1700	255	5.61
P	55000	1700	156	5.32
B	55000	1700	175	6.3

Table 4. General characteristics of Guanabara Bay

Area (m^2)	$3.52x10^8$
Volume (m^3)	$2.70x10^9$
Water inflow rate (m^3/h)	$9.0x10^6$
Water outflow rate (m^3/h)	$7.12x10^6$
Solids deposition rate ($g/m^2.day$)	22.96
Solids resuspension rate ($g/m^2.day$)	2.65
Solids burial rate ($g/m^2.day$)	13.52
Conc. Of particles in water (mg/l)	40.39
Conc. of part. in inflow water (mg/l)	92.4255
Conc. Of aerosol particles ($\mu g/m^3$)	30
Volumetric fraction of part. in sed.	0.2895
Density of particles in water (kg/m^3)	2500
Density of particles in sed. (kg/m^3)	2500
Density of aerosol particles (kg/m^3)	1500
Rain rate (m/yr.)	1.15
Velocity of aerosol deposition (m/h)	7.2
Scavenging rate	$2.0x10^5$

Table 5. Maximum and minimum values for the unknowns.

	Unknown	Min. value	Max. value
Z_1	Direct emission rate anthracene (kg/yr)	10	$2.0\text{x}10^{4}$
Z_2	Direct emission rate fluoranthene (kg/yr)	10	$2.0\text{x}10^{4}$
Z_3	Direct emission rate chrysene (kg/yr)	10	$2.0\text{x}10^{4}$
Z_4	Direct emission rate pyrene (kg/yr)	10	$2.0\text{x}10^{4}$
Z_5	Direct emission rate benzo(a)pyrene (kg/yr)	10	$2.0\text{x}10^{4}$
Z_6	Sed–water mass transfer coeff. anthracene (m/h)	$1\text{x}10^{-5}$	0.1
Z_7	Sed–water mass transfer coeff. fluoranthene (m/h)	$1\text{x}10^{-5}$	0.1
Z_8	Sed.–water mass transfer coeff . chrysene (m/h)	$1\text{x}10^{-5}$	0.1
Z_9	Sed.–water mass transfer coeff . pyrene (m/h)	$1\text{x}10^{-5}$	0.1
Z_{10}	Sed.–water mass transfer coeff . BaP (m/h)	$1\text{x}10^{-5}$	0.1
Z_{11}	Active sed. layer (m)	$3.0\text{x}10^{-3}$	$1.0\text{x}10^{-2}$
Z_{12}	Fraction of organic carbon in sediment solids	0.03	0.063
Z_{13}	Fraction of organic carbon in inflow water particles	0.240	0.361
Z_{14}	Fraction of organic carbon in water particles	0.240	0.361
Z_{15}	Fraction of organic carbon in resuspended particles	0.03	0.240

Some general characteristics of Guanabara Bay are given in Table 4. These data were obtained from [12 – 15]. Table 5 presents the minimum and maximum values of the unknowns evaluated.

The q values for B=0 and B=1 shown in table 2 coincided to be the best in the solution of the system (19).

Tables 6 and 7 present the estimates of the chemical and media unknowns respectively, obtained with the better q for different B and the particular case for B=1 and q=0 (entropy).

Table 6. Obtained estimates of the unknown chemicals with the better q for different B.

Chem.		Emission rate (kg/yr)	Mass transf. coeff. (m/h)
A	B=1 q=–1	9019	0.0120
	B=1 q= 0	8954	0.0172
	B=0 q= 1	9000	0.0141
F	B=1 q=–1	12894	0.0191
	B=1 q= 0	12821	0.0291
	B=0 q= 1	12881	0.0229
C	B=1 q=–1	5831	0.00346
	B=1 q= 0	5756	0.00452
	B=0 q= 1	5826	0.00392
P	B=1 q=–1	5714	0.00060
	B=1 q= 0	5617	0.00081
	B=0 q= 1	5707	0.00070
B	B=1 q=–1	4116	0.000143
	B=1 q= 0	4051	0.000328
	B=0 q= 1	4114	0.000186

Table 7. Obtained estimates of the unknown media with the better q for different B.

Active sed. layer (m)	B=1 q=–1	0.0031
	B=1 q= 0	0.0030
	B=0 q= 1	0.0031
Fraction of organic carbon in sediment solids	B=1 q=–1	0.045
	B=1 q= 0	0.056
	B=0 q= 1	0.050
Fraction of organic carbon in inflow water particles	B=1 q=–1	0.355
	B=1 q= 0	0.356
	B=0 q= 1	0.355
Fraction of organic carbon in water particles	B=1 q=–1	0.354
	B=1 q= 0	0.360
	B=0 q= 1	0.360
Fraction of organic carbon in resuspended particles	B=1 q=–1	0.032
	B=1 q= 0	0.030
	B=0 q= 1	0.035

Table 8. Estimates of chemical unknowns obtained by the Levenberg-Marquardt method with $\alpha = 0.001$

Chem.	Emission rate (kg/yr)	Mass transf. coeff. (m/h)
A	8553	0.0134
F	12736	0.0219
C	5572	0.00317
P	5363	0.000348
B	3881	$1.0\text{x}10^{-6}$

Levenberg-Marquardt Method

The estimates of unknown chemicals will be presented solving the system of equations with 10 unknowns. It was assumed that the fractions of organic carbon: in sediment solids is 0.050, in resuspended particles is 0.033, in inflow water is 0.355 and in water particles is 0.356; the active sediment layer is equal to 0.0031 m The results are presented in Table 8.

CONCLUSIONS AND FUTURE WORK

The results obtained by the Generalized Maximum Entropy method presented convergence orders of 10^{-14} for B = 1 and q = -1 and 10^{-17} for B = 0 and q = 1. Better results were attained with the Levenberg-Marquardt method solving the system of equations with 10 unknowns and when it was tried to solve the system with 15 unknowns, convergence was not obtained in many cases, or the order of the function to be minimized was 10^4.

Finally, it should be emphasized that this work is only a first attempt to find the emission rates of pollutant chemicals in different regions of Guanabara Bay.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPERJ – Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro and CAPES – Comissão de Aperfeiçoamento de Pessoal de Nível Superior.

REFERENCES

1. FEEMA (Fundação Estadual de Engenharia do Meio Ambiente) Environmental quality of the ecosystem of the basin hydrographic of the of Guanabara Bay in 1990. General office in the state of Rio de Janeiro. Government in the state of Rio de Janeiro, (1991).

2. D. Mackay, Modeling the long-term behavior of an organic contaminant in a large lake: Application to PCBs in lake Ontario. *J. great Lakes Res*, **15,** pp. 283-297 (1989)

3. D. Mackay, (1989) Application of the QWASI. *Chesmosphere*, **18**. pp, 1343-1365 (1989).

4. S. R. Wild, and K. C. Jones, Polynuclear aromatic hydrocarbons in the United Kingdom environment: A preliminary source inventory and budget. *Environ. Polli.*, **88**, pp.91-108 (1995

5. L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, *USSR Computational and Mathematical Physics Journal*, 7, pp. 200-217, (1967)

6 B. D. Sharma and D. P. Mittall, New non additive Measures of entropy for discrete probability distribution *J. Math Sci*, **10**, pp. 18-40 (1975).

7. J. N. Kapur, and H. K. Kesavan, Entropy Optimization Principles with Applications, Academic Press, Inc., London, (1992).

8. M. L Reis, N. C. and Roberty, Maximum entropy algorithms for image reconstruction from projections, *Inverse Problems*, **8**, pp. 623-644. (1992)

9. R. F. Carita Motero, N. C. Roberty, and A. J. Silva Neto, Absorption coefficient estimation in heterogeneous media using a domain patition consist with divergent beams, *Inverse Problems in Engineering*, **9**, pp. 587-617. (2001)

10. A. Lent and Y. Censor. The primal-dual algorithm as a constraint-set-manipulation device, *Mathematical Programming*, **50**, pp. 343-357, (1991).

11. D. G. Luenberger, Introduction to linear and no linear programming, Addison – Wesley Iberoamericana, 1970.

12. L. A. Acevedo, Determinação de hidrocarbonetos em amostras de água e mexilhões da Baía de Guanabara, Master outside thesis, PUC_RIO, (1998).

13. E. Amador, Geologia da Baía de Guanabara. Doctor outside thesis, Geology UFRJ, (1996)

14. C. Hamacher, Determinação de hidrocarbonetos em amostras de água e sedimento da Baía de Guanabara. Master outside thesis, PUC_RIO, (1996).

15. A. Lima, Geocronologia de hidrocarbonetos poliaromáticos (PAHs). Master outside thesis, PUC_RIO, (1996).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

COMPARISON BETWEEN TWO METHODS FOR ESTIMATING THE CONVECTION HEAT TRANSFER COEFFICIENT DURING A METALLURGICAL JOMINY END-QUENCH TEST

Philippe Le Masson
LET2E
Université de Bretagne Sud
Lorient, France
Philippe.le-masson@univ-ubs.fr

Tahar Loulou
Centre Energétique Environnement
Ecole des Mines d'Albi
Carmaux, Albi, France
Loulou@enstimac.fr

Eugène Artioukhine
CREST - Université de Franche Comté
Belfort, France
eugene.artioukhine@univ-fcomte.fr

ABSTRACT

This paper compares two methods for estimating a heat transfer coefficient: the Iterative Regularization Method "IRM" [1] and the Function Specification Method "FSM" with a spatial regularization [2]. The study is a continuation of our previous investigations [3,4]. A cylindrical steel bar (L=100 mm and d= 25 mm) is heated up to 900°C and then quenched by a water jet sprayed on its lower end. The problem under analysis is multifaceted. First, during the exchange phase between water and solid surface, phenomena like vaporization, boiling or forced convection occur. Second, the metallurgical phase changes of the considered material have to be taken into account in the examination of the heat treatment problem. The study is conducted to examine and to solve a non linear thermo-metallurgical inverse problem of estimating time and space dependent convection heat transfer coefficient.

NOMENCLATURE

a	diffusivity ($m^2.s^{-1}$)
C_i	specific heat of the phase i
D^n	vector of descent direction
dT/dt	cooling speed ($°C.s^{-1}$)
$\Delta Fo=\frac{a\Delta t}{z^2}$	delta Fourier Number
H_γ, H_α	enthalpies of the phases γ and α (Jm^{-3})
H(r,t),H^{n+1}	heat transfer coefficient ($W.m^{-2}.K^{-1}$)
HRC	Rockwell C hardness
J(H)	residual functional
∇J^n	residual functional gradient
L_{ij}, $L_{\gamma\alpha}$	heat transformation of phase i to j (Jm^{-3})
n	iteration number
P_i, P, P_{eq}	proportion of metallurgic phase (volume fraction)
r	spatial coordinates (m)
S	sensitivity coefficient
T_{inf}	external Temperature (°C).
T_{water}	temperature of water (°C)
T,T_i	temperature in the sample (°C)
T_0, T_{max}	initial and maximal temperatures (°C)
Y,Y_i, Y_e	measured temperatures (°C)
z	spatial coordinates (m).
γ^n	descent parameter
δ^2	convergence criterion
$\varepsilon(T)$	emissivity of the sample
ε_1	stopping criterion
λ_i	conductivity of the phase i ($Wm^{-1}K^{-1}$)
ρ_i	density of the phase i ($Kg.m^{-3}$)
σ^2	root-mean-square error on the temperature
ω	random number
Δ_{max}	maximum amplitude of the noise
$\Psi(z,r,t)$	adjoint variable

INTRODUCTION

In the two past decades, the development of the 2D and 3D inverse methods is increased in different domains of technology. The final report of the "2nd Joint Russian-American Workshop on Inverse Problems in Engineering" [5] have defined several directions on the development of some points as: a) development of methods and algorithms for solving multidimensional inverse problems; b) Investigation and improvement of mathematical models of heat and mass transfer including physical and chemical transformations

in materials. For that, the inverse studies on the heat treatments are increased. The goal is the estimation of boundary conditions as heat flux, temperature or heat transfer coefficient [6,7,8]. In the last decade, several studies have treated the thermometallurgical problems but as linear and without coupling. Mainly used methods for multidimensional problems are: The function specification method "FSM" with spatial regularizations [2,9] and the iterative regularization method "IRM" with the conjugate gradient method[1,10,11]. These methods are well documented. Recently, J.V. Beck [12,13] have initiated comparisons between different method in 1D and linear cases.

We want to compare these above two methods for the 2D axisymmetric problem of the "Jominy" end-quench test. In fine, the difficulty of the study is in the resolution of the coupled thermometallurgical problem with a high nonlinearity (principally in the thermal conductivity during the phase transformations (austenite – martensite)).

In this paper, first, we present the experimental setup. Second, we describe the direct problem with the difficulties for the simulation of phase transformations and transfer between the water jet at 20°C and the quenching surface at 900°C. Third, we present the two estimation techniques. We insist on the fact that we use the same equation systems particularly for the direct and sensitivity (in variation) problems. For a first comparison, we apply the techniques on a material without phase transformations (we treat only the heat conduction equation). This work permit to define the parameters and the criterion for obtain the best results. At last, for the coupled thermometallurgical problem on the 16MND5 steel, we compare the two methods.

JOMINY END QUENCH TEST

Principle (Figure 1)

The Jominy end quench test is a standard test (NFA04-303) used to characterize the hardenability of steels [14,15]. A steel cylinder (diameter: 25mm ; length: 100mm) is heated within the austenitic domain during a preset time and cooled by a water jet on its lower end (T_{water} varying between 15 and 25°C) (Figure 1). After cooling, hardness measurements are carried out along the cylinder axis as a function of distance from the quenched extremity (the Jominy hardenability curve).

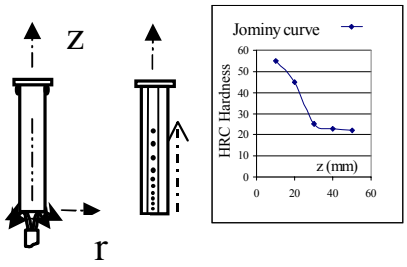


Figure 1: Principle of "Jominy" end-quench test.

Numerical simulation of the Jominy test

In a previous work [4], we have defined the direct problem. Here, we recall the basic equations. First, the numerical simulation is conducted from a fully austenitized 16MND5 steel sample heated at 880°C. The equations, implemented in both the SYSWELD code and the estimation code, are the heat conduction equation (1) and the metallurgical kinetic equations (2) of the type Leblond-Devaux [16] and Koistinen-Marburger [17]. The study domain (radial plane) is defined in cylindrical coordinates assuming the problem to be axisymmetric:

$$\sum_i P_i \rho_i(T) C_i(T) \frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(r \sum_i P_i \lambda_i(T) \frac{\partial T}{\partial r} \right) + \frac{\partial}{\partial z}\left(\sum_i P_i \lambda_i(T) \frac{\partial T}{\partial z} \right) - \sum_{i<j} P_{ij} L_{ij} \qquad (1)$$

$$\frac{dP}{dt} = \frac{P_{eq} - P}{\tau} f\left(\frac{dT}{dt}\right) \text{ and}$$

$$P = P_{max}\left(1 - \exp\left(-b(Ms - T)\right)\right) \qquad (2)$$

In the heat conduction equation, a source term allows to consider the phase change enthalpy $L\gamma\alpha$ according to the temperature during the sample cooling. The transformation heat is calculated according to the phase enthalpy: $L\gamma\alpha = H_\gamma - H_\alpha$ and by considering two metallurgical phases only: γ (austenite) and α (ferrite, pearlite, bainite or martensite). Moreover, the thermophysical characteristics $\rho(T)$, $C(T)$ and $\lambda(T)$ are calculated at all time steps by a law of mixtures according to the temperature. The boundary and initial conditions are the following:

Lower side:

$$-\sum_i P_i \lambda_i(T) \frac{\partial T(z=0,r,t)}{\partial z} = H(r,T)\left[T(0,r,t) - T_{water}\right] \qquad (3)$$

Upper side: $$\frac{\partial T(z_{max},r,t)}{\partial z} = 0 \qquad (4)$$

Lateral surface:

$$-\sum_i P_i \lambda_i(T)\frac{\partial T(z, r_{max}, t)}{\partial r} = \varepsilon(T)\sigma\left[T^4(z, r_{max}, t) - T_{inf}{}^4\right] \quad (5)$$

On the axis: $$\frac{\partial T(z, r = 0, t)}{\partial r} = 0 \quad (6)$$

Initial values: $T(z,r,0) = T_0$; $P\gamma(r, z, 0) = 1$ (7)

The direct problem is validated in Le Masson et al. [4,18].

ESTIMATIONS-MINIMIZATION PROCEDURES

The objective of this research is to determine the heat transfer coefficient on the lower side of the sample in function of both time and radius: H(r,t). The determination is conducted from the temperature values taken at a depth smaller or equal to z = 1mm. Two methods are used:

Iterative Regularization Method

The estimation of the heat transfer coefficient in view of the direct problem (equations (1) to (7)) can be formulated under the variational form which implies the residual functional minimization.

$$J(H) = \int_0^{t_f} \sum_{i=1}^{N} \left[T(z_i, r_i, t; H) - Y(z_i, r_i, t)\right]^2 dt \quad (8)$$

where $T(z_i, r_i, t; H)$ and $Y(z_i, r_i, t)$ represent the estimated and measured temperatures at N various points of the material, respectively. The inverse problem consists in minimizing this residual functional under constraints given by the equations of the direct system (equations (1) to (7)). The minimization is carried out by using the conjugate gradient method [1]. The function H(r,t) is considered here as an element of the Hilbert space, L^2 and the new functions are obtained after each iteration as follows:

$H^{n+1} = H^n + \gamma^n D^n$, n=1, ... (9)

where n is the iteration index, γ^n the descent parameter, H^{n+1} the unknown vector to be estimated and D^n the vector of descent direction given by:

$$D^n = -\nabla J^n + \beta^n D^{n-1} \quad (10)$$

where $\beta^n = \dfrac{\left(\nabla J^n - \nabla J^{n-1}, \nabla J^n\right)}{\left\|\nabla J^n\right\|^2}$, $\beta^0 = 0$. (11)

with (,) is the scalar product in Hilbert space L^2.

In the absence of noise, the iteration procedure is carried on until the following stopping criterion is verified:

$$\left|\frac{H^{n+1} - H^n}{H^n}\right| \le \varepsilon_1 \quad (12)$$

The gradient ∇J^n (Eq. (10) and (11)) of the residual functional is obtained for all values of r and t by the following analytical relationship [1]:

$$\nabla J^n = \Psi(0, r, t) * \left(T(0, r, t) - T_{inf}\right). \quad (13)$$

where $\Psi(z,r,t)$ is the solution of the adjoint problem. More details of the derivation of the adjoint problem, variation problem and computational algorithm are given in Le Masson et al.[3].

Regularization

The inverse problems are ill-posed and numerical solutions depend on the fluctuations occurring at the measurements. Small fluctuations at the measurements can generate big errors in the solution to be estimated. We use the iterative regularization method [1] in which the regularizing condition is the residual criterion:

$$J(H)^{n^*} \approx \delta^2 \quad (14)$$

where n* is the index of the last iteration at which the condition $J(H) < \delta^2$ is verified for the first time, δ^2 is the total (integrated) measurement error defined as:

$$\delta^2 = \int_0^{t_f} \sum_{i=1}^{N} \sigma^2(z_i, r_i, t)\, dt \quad (15)$$

$\sigma^2(z,r,t)$ is the root-mean-square error of the temperature measurements obtained by smoothing the measured temperature histories. The number n* is the regularization parameter of the method.

Function Specification Method

The method is based on the same principle than the function sequential specification method in time with spatial regularization [2]. For two dimensional heat transfer coefficient we must minimize the differences between the computed and measured temperatures at N various points of the specimen. The time regularization is obtained by adding a suitable number of future time steps 'ntf'. The Tikhonov regularization principle is employed to stabilize spatially the inverse problem. The inverse problem solution consists in minimizing the new functional:

$$J(H^{n+1}_{h\in[1,M]}) = \sum_{i=1}^{N}\sum_{j=1}^{ntf}\left[T^{n+j}\left(z_i, r_i, t; H^{n+j}\right) - Y^{n+j}(z_i, r_i, t)\right]^2 + R_1\sum_{h=2}^{M}\left(H^{n+1}_{h-1} - H^{n+1}_{h}\right)^2 + R_2\sum_{h=2}^{M-1}\left(H^{n+1}_{h-1} - 2H^{n+1}_{h} + H^{n+1}_{h+1}\right)^2 \quad (16)$$

where R_1 and R_2 are the coefficients of the first and second order spatial regularization, respectively. R_1 limits the variations of two consecutive spatial heat transfer coefficients. R2 is used for obtain smooth space variations between consecutive coefficients. Minimization of J(H) versus H^{n+1} using linearization ($H^{n+1} = H^n + \Delta H^{n+1}$) and assuming temporally that the heat tranfer coefficient is constant over "nft" time steps ($H^{n+j} = H^{n+1}$), leads to:

$$\left(\left[S^2_T\right] + \left[S^2_r\right]\right) * \left[\Delta H^{n+1}\right] = \left[D_T\right] + \left[D_r\right] \quad (17)$$

The matrice $\left[S^2_r\right]$ (MxM) and the vector $\left[D_r\right]$ (M) are defined for the regularization. The matrice $\left[S^2_T\right]$ (MxM) and the vector $\left[D_T\right]$ (M) come from the residual functional without regularization. Each terms (line l and column h) of the matrice $\left[S^2_T\right]$ and the vector $\left[D_T\right]$ are equal to:

$$S^2_{lh} = \sum_{i=1}^{N}\sum_{j=1}^{ntf}\left(S^{n+j}_{T_i, H_h^{n\to n+j}} * S^{n+j}_{T_i, H_l^{n\to n+j}}\right) \quad (18)$$

$$D_l = \sum_{i=1}^{N}\sum_{j=1}^{ntf}\left(Y_i^{n+j} - T_i^{n+j}\right) * S^{n+j}_{T_i, H_l^{n\to n+j}} \quad (19)$$

In these terms, $S^{n+j}_{T_i, H_l^{n\to n+j}}$ is the sensitivity coefficient of the temperature T_i^{n+j} with respect to the heat transfer coefficient at the point l. Here, we mention that the sensitivity equations system is precisely the variation system used in "IRM" [4] with $\delta H(r,t) = 1$ at the node l and =0 for all other points. In that case, we must solve M sensitivity problems which are very high computational time consuming. The direct and sensitivity systems are solved numerically using the control volume method and an implicit scheme. The estimation procedure consists in:

1)Defining the initial values of de $H^0(t_0)$.
2)Solving the direct system: equations (1) to (7) for $t_{n+1} \le t \le t_{n+j}$.
3)Solving all sensitivity problems [4]
4)Calculating $\left[S^2_T\right]$ and $\left[D_T\right]$
5)Adding the regularization: $\left[S^2_r\right]$ and $\left[D_r\right]$
6)Calculating the condition number C: eq. (20)
7)Solving the system (17)
8)Incrementing the vector $H^{n+1} = H^n + \Delta H^{n+1}$)
9)Calculating the quadratic criterion J(H): equation (16)
10) Verifying the convergence criterion (12).
11) And back to step 2) with $t = t + \delta t$ whether the convergence criterion is correct or t = t whether it is not correct.

Regularization

The inverse problems are ill-posed. So we must define the parameters "ntf" for the time regularization; R_1 and R_2 for the spatial regularization. We choose "ntf" versus the depth of the measured points and the noise on the measurements. With the deepest locations of sensors, the highest number of future time steps is needed for time regularization. J.V. Beck et al.[13] show that the accuracy of heat transfer coefficient estimation decreases with the increase of number of future times while that the sensitivity to measurements errors decreases with the same number.

We choose R_1 and R_2 versus the calculate of the condition number:

$$C = \left\|\left(\left[S^2_T\right] + \left[S^2_r\right]\right)\right\| * \left\|\left(\left[S^2_T\right] + \left[S^2_r\right]\right)^{-1}\right\| \quad (20)$$

G.Blanc et al.[9] show how to define these coefficients. We draw the condition number versus R_1 or R_2. We choose the best parameter when the curve is lower than a value (around 200 for exemple, Figure 2). We take the value around R` for measurements without noise and R`` for measurements with noise.

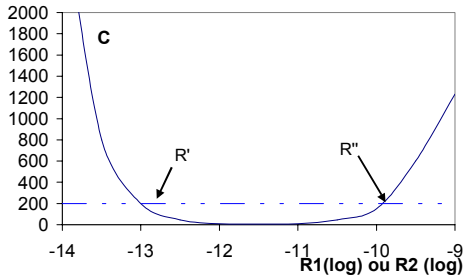


Figure 2 : Variation of the condition number.

APPLICATIONS

In order to verify the method, two materials have been used. One is nickel, without metallurgical transformations; Another one is 16MND5 steel, the thermophysical characteristics

ρ(T), λ(T) and C(T) of which in functional of the phases and the temperature, as well as the Continuous Cooling Temperature (C.C.T.) diagram considered as the basis of the transformations, are known. In the goal to perform the comparison between the two methods we introduce the delta Fourier number defined by $\Delta Fo = \frac{a \Delta t}{z_N^2}$ where a is the middle diffusivity for Nickel or the diffusivity at the beginning of the phase transformation of steel, Δt is the time step and z_N is the sensor locations. Numerical results presented here are obtained with computed temperatures taken at the sensor locations z_N =0.15mm and z_N =1mm for all thermocouples which result in four ΔFo numbers.

Table 1: Delta Fourier number

ΔFo	Nickel	steel
z=0.15mm	4.92	1.7
z=1mm	0.11	0.04

Case of nickel

As regards nickel only the heat conduction equation has to be solved. Two types of law have been tested in order to verify the code. The first law is H(r,t) = 15000 $W.m^{-2}.K^{-1}$ and the second is a coefficient, which varies along a dome ($H(r,t)= 1000 + 7500 * (\sin(\pi t/t_{max}) + \sin(\pi r/r_{max}))$). For "IRM", we have showed in a previous work [4] that in the absence of noise, the heat transfer coefficients are correctly found except near t_f where the initial condition of the adjoint problem sets the adjoint variable equal to zero and, therefore, implies that the gradient be always equal to zero. The more we are moving away from the surface (z = 0), the more the initial condition affects estimate and increases the error.

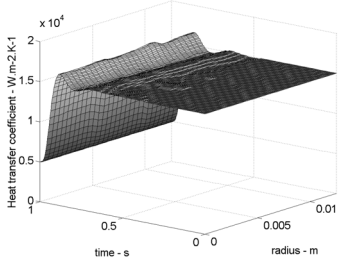


Figure 3: Estimation with "IRM" for z=1mm.

Figure 3 shows the result obtained for z = 1mm and N = 30.

The analysis of these estimates confirms the error when approaching $t = t_f$. So, the estimation domain is increased by 15 to 30% beyond the effective time. Figure 4 shows the result for an increased time of 15%.

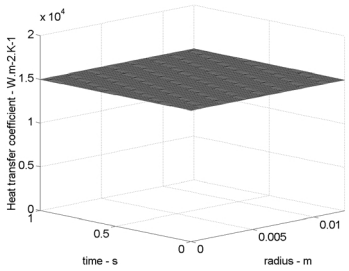


Figure 4: Estimation with "IRM" for z=1mm and an increased time.

We have the same problem for the dome. The Figure 5 shows the result for an increased time of 30%.

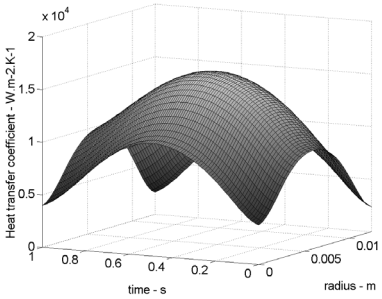


Figure 5: Estimation with "IRM" for z=1mm and an increased time of 30%.

Now for "FSM", we show here that in the absence of noise, the heat transfer coefficients equal to 15000 are perfectly found (Figure 6).

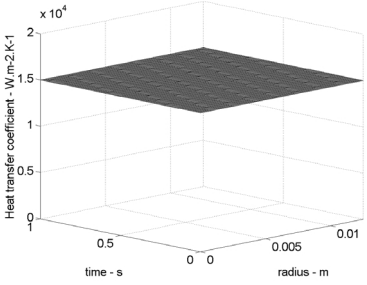


Figure 6: Estimation with "FSM" for z=1mm.

For the dome, with theoretical measurements in z = 0.15mm, the results are correctly found (Figure 7). With theoretical measurements in z=1mm, we have some problems at the beginning because the initial values of the heat transfer coefficients are equal to 1000 (Figure 8).

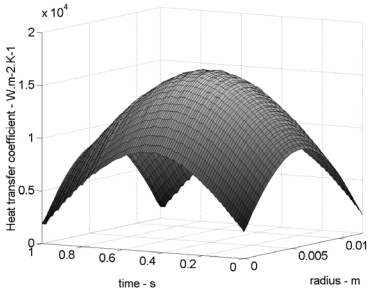


Figure 7: Estimations with "FSM" for z = 0.15mm

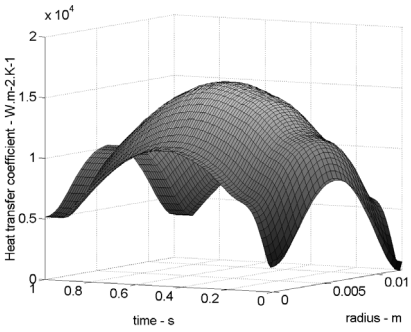


Figure 8: Estimations with "FSM" for z=1mm.

So, we must do several iterations to obtain the first values at initial time. At the end, near t_f, the values stay constant because the stop condition is verified at $t=t_f - ntf*\Delta t$. So, in the same manner as the "IRM", we must increase the estimation domain. We choose the parameters "ntf" and R_1 like as the values of the table 2.

Table 2: Values of "ntf" and R_1

	ntf	R_1
z = 0,15mm	2	$5\ 10^{-8}$
z = 1mm	8	10^{-7}

In a second phase, we verify that, even in the presence of noise, the convergence of the algorithm toward a good estimation is acceptable. The case of the dome only, for which a noise has been added to the theoretical thermal cycles, is presented here. The disturbance is defined by the relationship: $B = \omega\, T_{max}\, \Delta_{max}$ where ω is a random number in [-1, +1], T_{max} is the maximum value of the surface temperature and Δ_{max} is the maximum amplitude of the noise. For these numerical tests, we consider T_{max} = 880°C and Δ_{max} = 5%.

$$Y_i(r_i,z_i,t) = Y_e(r_i,z_i,t) + \omega\, T_{max}\, \Delta_{max} \quad i=1,\ldots,N \quad (21)$$

Figure 9 shows the results achieved in the case of the dome at z=0.15mm and for t_f = 1s. When the criterion J(H) is approximately equal to δ^2 the best result is reached. We have showed in the previous work [4] that when computing beyond the δ^2, the criterion decreases, while the solution starts to oscillate. Figure 10 shows the results in the case of the dome obtained with "FSM".

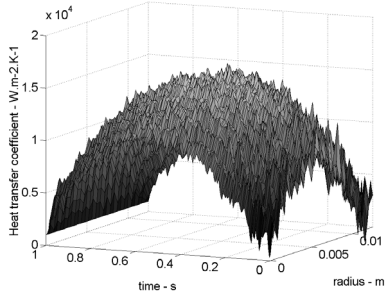


Figure 9: Estimations with "IRM" for z=0.15mm and noise.

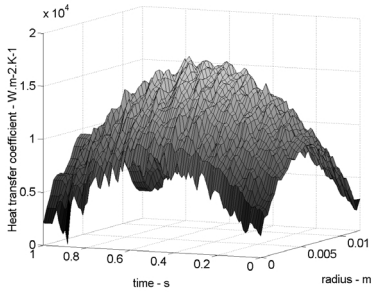


Figure 10: Estimations with "FSM" for z=0.15mm and noise.

With noised data, we must take more "ntf" (here ntf=4) and we take a bigger parameter of spatial regularization like as $R_2 = 10^{-6}$. We can see that the results are correct.

As a conclusion to this section, we can say that the two estimation methods, for the case of a material without phase change and for a Fourier number equal to 0.11, gives good enough results.

Case of 16MND5 steel

The following test cases are analyzed here: a) $H(r,t) = 15000\ W.m^{-2}.K^{-1}$ for $0 < r < r_{max}$ and $0 < t < 2s$. b) $H(r,t) = 5000 + 7500 * (\sin(\pi t/t_{max}) + \sin(\pi r/r_{max}))$ for $0 < r < r_{max}$ and $0 < t < 2s$.

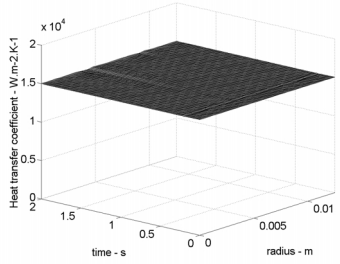


Figure 11: Estimations with "IRM" for z=1mm

Figure 11 displays with "IRM" the estimate of the heat transfer coefficient for the case « a » from the theoretical thermal curves considered in z = 1mm with an increased time of 30%. The results analysis show that the iterative regularization method is not disturbed by the phase changes. With "FSM" and for theoretical measurements in z = 0.15mm, we show that the results are completely found (Figure 12) (ntf=2 ; R_1=6 10^{-8} ; ΔFo=1.17).

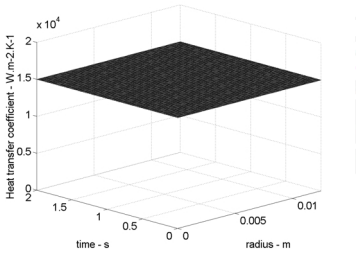


Figure 12: Estimations with "FSM" for z=0.15mm.

However, for theoretical measurements in z = 1mm (Figure 13)(ntf=8; R_1 = 5 10^{-8}; ΔFo=0.04), we show that we have an oscillation when the phase change begin.

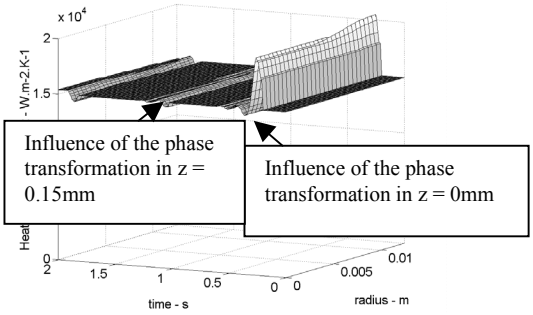


Figure 13: Estimations with "FSM" for z=1mm.

At each mesh, when the transformation begin, we have a bias on the estimate. The phase change play a role of perturbation and the measure can't dissociate the effects of heat transfer coefficient and the transformation of material. Archambault et al. [8] show the same problem for an estimation of a 1D heat flux and for a coupled thermometallurgical problem. During the transformation, we can see oscillations on the estimate. Figure 14 shows the evolution of the bias versus the depth of theoretical measurements. The more we are moving away from the surface (z = 0), the more the delta Fourier number decrease. We have more bias on the estimate. This results with this bias are however validate by the comparison between the estimated and theoretical temperatures. The errors are lower 0.1°C.

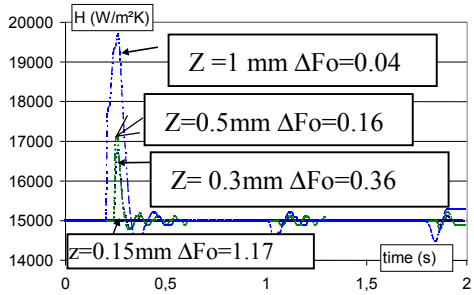


Figure 14 : Evolutions of the bias versus the depth of theoretical measurements.

We notice the same in case « b ». Results are displayed in Figures 15, 16 and 17.

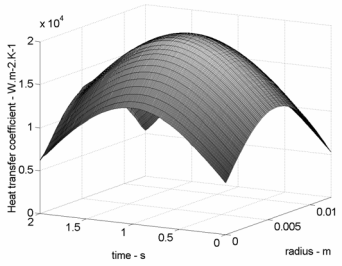


Figure 15: Estimations with "IRM" for z=1mm and an increased time of 30%.

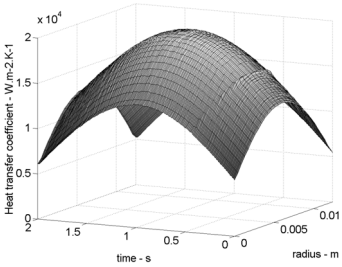


Figure 16: Estimations with "FSM" for z=0.15mm.

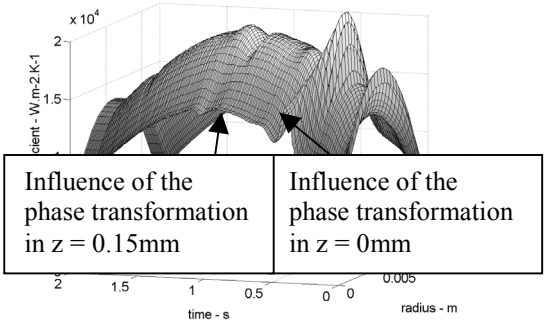


Figure 17: Estimations with "FSM" for z=1mm.

At last, we have compared the results of the "IRM" estimation and the 'FSM' estimation in the case of noise-added thermal cycles considered at z = 0.15mm and z=1mm. We have the same results for z = 0.15mm. But, we have oscillations on the "FSM" estimate for z=1mm.

CONCLUSION

Two methods are compared for the solution of the inverse coupled thermometallurgical problem of estimating the transient heat transfer coefficient from interior temperature histories.

At first, the two methods gave accurate results for nickel samples which have no phase transformations. The delta Fourier number is equal to 4.92 (z=0.15mm) or 0.11(z=1mm). So, there is no real problem to estimate the heat transfer coefficient by using both methods.

Second, we add for steel samples phase transformations. The corresponding delta Fourier number are equal to 1.7 and 0.04. For $\Delta Fo=1.7$, z=0.15mm, the two methods gave the same good results. But, when $\Delta Fo=0.04$ and the high nonlinearities come in the material, we have a biased estimation with the "FSM"(an oscillation). However, the "IRM" give an accurate result.

REFERENCES

1. O.M. Alifanov, E.A. Artyukhin, S.V. Rumyantsev, Extreme Methods for Solving Ill-posed Problems with Applications to inverse Heat Transfer Problems, Begell house, N-Y, 1995.

2. J.V. Beck, B. Blackwell and C.R. St Clair; Inverse Heat Conduction Ill Posed Problem, Wiley, New York, 1985.

3. P. Le Masson, D. Carron, P. Rogeon, J.J. Quemener, Identification du coefficient de transfert lors d'un essai Jominy instrumenté, SFT 99, 57-62, 1999.

4. P. Le Masson, T. Loulou, E. Artioukhine, P. Rogeon, D. Carron and J.J. Quemener, A new approach for the estimation of a convection heat transfer coefficient during a metallurgical « Jominy end-quench » test. *68th eurotherm seminar* - (Poitiers, France, 2001)

5. Final report of *2nd Joint Russian-American Workshop on inverse problems in Engineering;* St Petersburg, Russia, 21-27 August 1994.

6. A. Osman, J.V. Beck, Investigation of transient heat transfer coefficient in quenching experiments, J. Heat Transfer **112**(1990) 843-848.

7. S.G. Chen, C.I. Weng, J. Lin, Inverse estimation of transient temperature distribution in the end quenching test*, J. of Materials Processing Technology*, **86** (1999) 257-263

8. P. Archambault, S. Denis and A. Azim, Inverse resolution of the heat transfer equation with internal heat source: Application to the quenching of steels with phase transformations, J. of Mat. Eng. and Perf., **6,** 1997, 240-246

9. G. Blanc, M. Raynaud, The Hiep Chau ; A guide for the use of the function specification method for 2D Inverse heat conduction problems; *R. Générale de thermique*, **37**, 1998, 17-30

10. Y. Jarny, M.N. Ozisik and J.P. Bardon, A general optimization method using adjoint equation for solving multidimensional inverse heat conduction, , *Int.J.Heat Mass Transfer,* **34, n°11,** 1991, 2911-2919

11. M.J. Colaço, H.R.B. Orlande; Comparison of different versions of the conjugate gradient method of function estimation, *Numerical Heat transfer, Part A*, **36**, 1999, 229-249

12. J.V. Beck; Comparison of the iterative regularization and function specification algorithms for the inverse heat conduction problem; *Inverse Pb. in Engineering: Theory and Practice;* ASME 1993

13. J.V. Beck, B. Blackwell and A. Haji-Sheikh, Comparison of some inverse heat conduction methods using experimental data, *Int.J.Heat Mass Transfer,* **39, n°17,** 1996, 3649-3657

14. Essai de trempabilité Jominy ; NFA 04-303 ; juin 1979

15. Atlas de courbes Jominy (états bruts de trempe et revenus) d'aciers de la norme NFA 35-552 ; Collection ATS OTUA ; Paris 1985.

16. J. Leblond, J. Devaux, A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size ; Acta Metall., **32 n°1**, 137-146, 1984.

17. D.P. Koistinen, R.E. Marburger, A general equation prescribing the extent of austenite-martensite transformation in pure iron-carbon alloys and plain carbon steels; Acta. Metall., **7**, 59, 1959.

18. P. Le Masson, P. Rogeon, D. Carron, J.J. Quemener, Simulation numérique thermométallurgique : influence des paramètres d'entrée sur la modélisation d'un essai de trempabilité, Revue de Métallurgie – CIT/ Sciences et Génie des Matériaux, pp1055-1064, Sept. 2000.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

EXPERIMENTAL AND NUMERICAL ESTIMATIONS OF A TWO-DIMENSIONAL CONVECTION HEAT TRANSFER COEFFICIENT DURING A METALLURGICAL « JOMINY END-QUENCH » TEST

Philippe Le Masson
LET2E
Université de Bretagne Sud
Lorient, France
Philippe.le-masson@univ-ubs.fr

Tahar Loulou
Centre Energétique Environnement
Ecole des Mines d'Albi
Carmaux, Albi, France
Loulou@enstimac.fr

Eugène Artioukhine
CREST – Université de Franche Comté
Belfort, France
eugene.artioukhine@univ-fcomte.fr

ABSTRACT

This paper shows applications of a method for the estimation of a 2D heat transfer coefficient by the Iterative Regularization Method "IRM" [1]. The Jominy end-quench test considered here is used to measure the hardenability of steels. A cylindrical steel bar (L=100 mm and d= 25 mm) is heated up to 900°C and then quenched by a water jet sprayed on its lower end. The temperature histories are measured using thermocouples embedded in the sample. The reproducibility of test is provided by the automation of the experimental procedure. The study is conducted to examine and to solve a thermo-metallurgical inverse problem of estimating the time and space dependent convection heat transfer coefficient.

NOMENCLATURE

a	diffusivity ($m^2.s^{-1}$)
C_i	specific heat of the phase i
D^n	vector of descent direction
dT/dt	cooling speed (°C.s-1)
$\Delta Fo=\frac{a\Delta t}{z^2}$	delta Fourier number
H_γ, H_α	enthalpies of the phases γ and α ($J.m^{-3}$)
H(r,t), H^{n+1}	heat transfer coefficient ($W.m^{-2}.K^{-1}$)
HRC	Rockwell C hardness
J(H)	residual functional
∇J^n	residual functional gradient
L_{ij}, $L_{\gamma\alpha}$	heat transformation of phase i to j ($J.m^{-3}$)
n	iteration number
P_i, P, P_{eq}	proportion of metallurgic phase (volume fraction)
r	spatial coordinates (m)
T_{inf}	external Temperature (°C).
T_{water}	temperature of water (°C)
T, T_i	temperature in the sample (°C)
T_0, T_{max}	initial and maximal temperatures(°C)
Y,Y_i, Y_e	experimental temperatures (°C)
z	spatial coordinates (m).
γ^n	descent parameter
δ^2	total measurement error
$\varepsilon(T)$	emissivity of the sample
ε_1	stopping criterion
λ_i	conductivity of the phase i ($W.m^{-1}.K^{-1}$)
ρ_i	density of the phase i ($Kg.m^{-3}$)
σ^2	root-mean-square error of the temperature
ω	random number
Δ_{max}	maximum noise amplitude
$\Psi(z,r,t)$	adjoint variable

INTRODUCTION

Quenching is one of the most critical operations in the heat treatment of many metallic parts, affecting both mechanical and structural properties. The first stage of this study involved the prediction of the temperature history at the quenching surface. However, the main difficulty of these analyses is that the exact flux or heat transfer coefficient at the quenching surface are not easy to predict and they cannot be measured

directly. An inverse estimation of these transient parameters should be used with temperature measurements near the quenching surface. In this case, the estimation results are very sensitive to small errors in experimental data, which may lead to considerable errors in the estimated values.

Many methods have been proposed to solve the inverse heat conduction problems for the estimation of the unknown thermal boundary conductions in several material processes. Osman and Beck [4] use the sequential function specification method in a 1D quenching problem by solving only the heat conduction equation. Chen et al. [5] use the Iterative Regularization Method in a 2D quenching problem with a linear heat conduction equation. They estimate a 2D heat flux by fitting the experimental temperature data (30 values obtained with 4 thermocouples in the depth z = 2mm). Archambault and al. [6] use the function specification method in 1D quenching problem with a coupled thermo-metallurgical problem. In their results one can see oscillations on the estimation errors during the phase transformations.

In this paper, we use the Iterative Regularization Method "IRM" for the 2D axisymmetric problem for the "Jominy" end-quench test. The main difficulty of the problem is the simulation of the coupled thermometallurgical phenomena with a high nonlinearity of the thermal conductivity during the phase transformations (austenite - martensite).

First, we present the experiments. Second, we describe the direct problem with the simulation of phase transformations and heat transfer between a water jet at 20°C and a hot specimen surface at 880°C. Third, we present the estimation procedure. To verify the estimation code, we use a theoretical heat transfer coefficient which is similar to the experimental one available in the literature [12]. Finally, we present the experimental results for two materials. For the first, the Nickel, we solve only the nonlinear heat conduction equation. For the second, the 16MND5 steel, we treat the coupled thermometallurgical problem.

JOMINY END QUENCH TEST

Principle (Figure 1)

The JOMINY end quench test is a standard test (NF A 04-303) used to characterize the hardenability of steels [7,8]. A steel cylinder (diameter: 25mm ; length: 100mm) is heated within the austenite domain during a preset time and cooled by a water jet sprayed on its lower end (T_{water} varying between 15 and 25°C) (Figure 1).

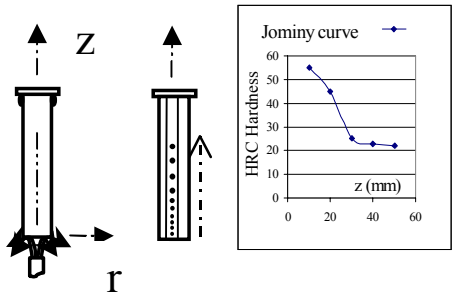


Figure 1: Principle of "Jominy" end-quench test

After cooling the Jominy bar two parallel polished plane bands are prepared on two sides of the test cylinder. Hardness measurements are carried out along the cylinder axis as a function of distance from the quenching surface (1.5mm, 3mm, 5mm, 7mm, 9mm, 11mm.....60mm, 70mm and 80mm). The curve representing the hardness variations in function of the distance from the quenched end of the cylinder is called the Jominy hardenability curve.

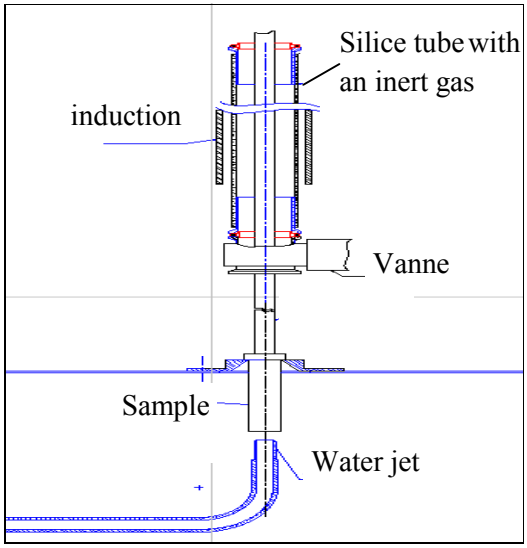


Figure 2: Sketch of experimental setup in the LET2E laboratory

The experimental setup built according to the standard gives following characteristics (Figure2):

- Induction heating in an enclosure filled by an inert gas (argon);
- Computer-aided control of various actions (heating, cylinder handling, quenching), thanks to which the test reproducibility is satisfactory.

Temperature histories are measured with the embedded thermocouples in a number of points near the quenching surface inside the specimen and at its lateral surface. That makes it possible, by using inverse methods, to identify parameters or functions at the quenching surface. The thermocouples (type K, $\phi = 70\mu m$) are embedded perpendicularly of the axis in holes ($\phi = 0.6mm$, depth between 0.5 mm and 12mm). They are isolated by alumine tubes (Photo 1). The thermal kinetic parameters are acquired by a "IOTECH-Daqbook/216" analyzer and a thermocouple card "DBK19".

Photo 1: sample with thermocouples

Numerical simulation of the JOMINY test

In our previous work [9], we have defined the direct problem. Here, we remind the system of equations. First, the numerical simulation is carried out from a fully austenitized 16MND5 steel sample (considered as an equivalent of ASTM A508 Cl. 3) heated at 880°C. The equations, implemented in both the SYSWELD code and the estimation code which we developed, are the energy equation (1) and the metallurgical kinetic equations (2) of the type Leblond-Devaux [10] and Koistinen-Marburger [11]. The integration domain (radial plane) is defined in cylindrical coordinates assuming that the problem is axisymmetric:

$$\sum_i P_i \rho_i(T) C_i(T) \frac{\partial T}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left(r \sum_i P_i \lambda_i(T) \frac{\partial T}{\partial r} \right) + \frac{\partial}{\partial z}\left(\sum_i P_i \lambda_i(T) \frac{\partial T}{\partial z} \right) - \sum_{i<j} P_{ij} L_{ij} \qquad (1)$$

$$\frac{dP}{dt} = \frac{P_{eq} - P}{\tau} f\left(\frac{dT}{dt}\right) \quad \text{and}$$

$$P = P_{max}\left(1 - \exp\left(-b(Ms - T)\right)\right) \qquad (2)$$

In the energy equation (1), the source term allows to consider the phase change enthalpy $L_{\gamma\alpha}$ according to the temperature evolution during the sample cooling. The phase change heat is calculated according to the phase enthalpy: $L_{\gamma\alpha} = H_\gamma - H_\alpha$ and by considering only two metallurgical phases: γ (austenite) and α (ferrite, pearlite, bainite or martensite). The enthalpies of the phase α and γ are defined by polynomial functions between 100°C and 1450°C. We compute the parameters of the metallurgical kinetic equations with the use of the Continuous Cooling Temperature (C.C.T.) diagram (Figure 3).

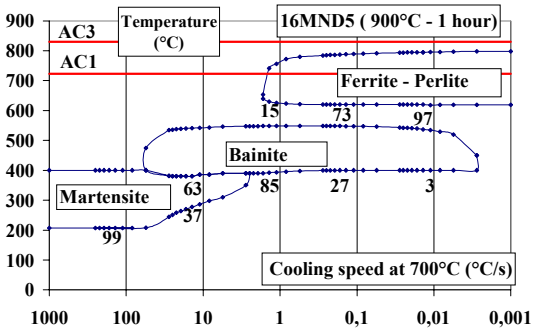


Figure 3: CCT diagram

We use the finite-difference method to simulate numerically the Jominy test. For each node of the finite-difference grid, the cooling speed and the temperature are computed. With these two parameters, we can calculate the percentage of the metallurgical phases. Then, the thermophysical characteristics $\rho(T)$, $C(T)$ and $\lambda(T)$ are calculated at all nodes by using the mixture law in function to the temperature (for example: $\lambda(T) = P_\gamma * \lambda_\gamma(T) + P_\alpha * \lambda_\alpha(T)$).

For the phase γ of 16MND5 and for $T < 1450°C$:

$\rho_\gamma * C_\gamma(T) = 3641440 + 4638.78*T - 11.7784*T^2 + 0.0155136*T^3 - 9.29165E\text{-}6*T^4 + 2.03093E\text{-}9*T^5$

$\lambda_\gamma(T) = 0.0148939+1.24115E\text{-}5*T\text{-}7.74533E\text{-}10*T^2+8.11438E\text{-}13*T^3$

For the phase α of 16MND5 and for T < 750°C:

$\rho_\alpha * C_\alpha(T) = 3620850+2955.1*T\text{-}7.52398*T^2+0.0249182*T^3\text{-}1.68686E\text{-}5*T^4$

$\lambda_\alpha(T) = 0.0525037\text{-}3.35115E\text{-}5*T+1.76665E\text{-}8*T^2\text{-}1.74307E\text{-}11*T^3$

The boundary and initial conditions are the following:

Lower side:

$$-\sum_i P_i \lambda_i(T)\frac{\partial T(z=0,r,t)}{\partial z} = H(r,T)\left[T(0,r,t)-T_{water}\right] \quad (3)$$

Upper side: $$\frac{\partial T(z_{max},r,t)}{\partial z} = 0 \quad (4)$$

Lateral surface:

$$-\sum_i P_i \lambda_i(T)\frac{\partial T(z,r_{max},t)}{\partial r} = \varepsilon(T)\sigma\left[T^4(z,r_{max},t)-T_{inf}{}^4\right] \quad (5)$$

On the axis: $$\frac{\partial T(z,r=0,t)}{\partial r} = 0 \quad (6)$$

Initial values: $T(z,r,0) = T_0$; $P\gamma(r,z,0) = 1$ (7)

We precise the hypothesis that we have used for the theoretical heat transfer coefficient:

The first hypotheses, given by Sorin [12] (Figure 4), involves the presence of vaporization, boiling and forced convection at the surface according to the temperature level.

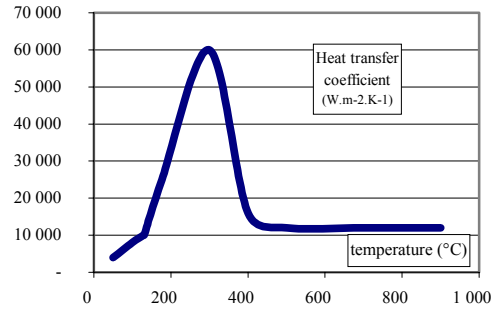


Figure 4: Hypothesis for the heat transfer Coefficient.

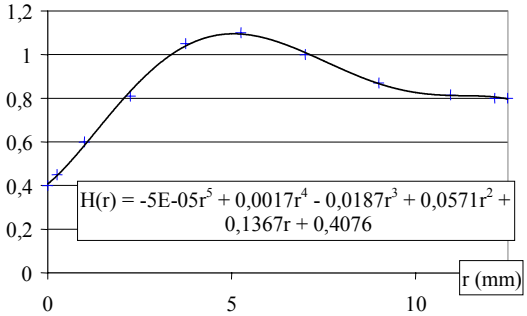


Figure 5: Weighting curve of the transfer coefficient versus radius.

The second hypotheses comes from the fact that when examining the water jet at the lower surface, we can consider the stagnation point or the stagnation region in the center (r = 0) and the heat transfer coefficient depending on the radius (Figure 5).

Figure 6 shows this heat transfer coefficient for the direct problem. It's one of the heat transfer coefficients that we have estimated for the verification of the estimation code [3].

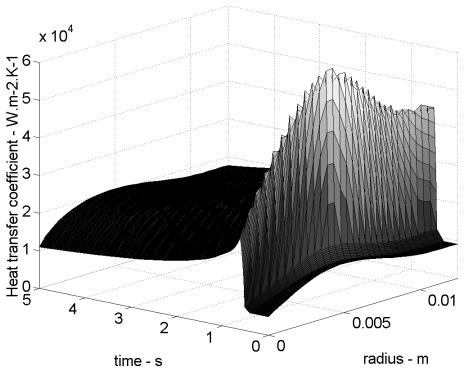


Figure 6: Theoretical heat transfer coefficient.

ESTIMATION-MINIMIZATION PROCEDURE

The objective of this research is to estimate the heat transfer coefficient at the quenching surface of the sample in function of both time and radius, H(r,t). The estimation is realized by using the temperature histories measured in the depth smaller or equal to z = 1mm.

The numerical algorithm built to estimate the heat transfer coefficient is based on the variational formulation of the inverse problem under analysis that leads to the minimization of the residual functional.

$$J(H) = \int_0^{t_f}\sum_{i=1}^{N}\left[T(z_i,r_i,t;H)-Y(z_i,r_i,t)\right]^2 dt \quad (8)$$

where $T(z_i,r_i,t;H)$ and $Y(z_i,r_i,t)$ represent the computed and measured temperature histories at N various points of the material, respectively. The inverse problem consists in minimizing this residual functional under constraints given by the equations of the direct problem (equations (1) to (7)). The minimization is carried out by using the conjugate gradient method [1]. The function H(r,t) is considered here as an element of the Hilbert space L^2 and the improved desired grid function is obtained at each iteration as follows:

$H^{n+1} = H^n + \gamma^n D^n$, n=1… (9)

where n is the iteration index, γ^n is the descent parameter, H^{n+1} is the unknown grid function to be estimated and D^n is the vector of descent direction given by:

$$D^n = -\nabla J^n + \beta^n D^{n-1} \qquad (10)$$

$$\text{where } \beta^n = \frac{\left(\nabla J^n - \nabla J^{n-1}, \nabla J^n\right)}{\left\|\nabla J^n\right\|^2},\ \beta^0=0. \qquad (11)$$

with (,) is the scalar product in Hilbert space L^2.

In the absence of noise, the iterative procedure is carried on until the following stopping criterion is verified:

$$\left|\frac{H^{n+1} - H^n}{H^n}\right| \le \varepsilon_1 \qquad (12)$$

The residual functional gradient ∇J^n in Eq. (10) and (11) is obtained for all values of r and t by the following analytical relationship [1]:

$$\nabla J^n = \Psi(0, r, t) * \left(T(0, r, t) - T_{inf}\right). \qquad (13)$$

where $\Psi(z,r,t)$ is the solution of the adjoint problem. More details of the derivation of the adjoint problem, variation problem and computational algorithm are given in Le Masson et al.[3].

Regularization

The inverse problems are ill-posed and numerical solutions depend on the fluctuations occurring at the measurements. Small fluctuations at the measurements can generate big errors in the solution to be estimated. We use the iterative regularization method [1] in which the regularizing condition is the residual criterion:

$$J(H)^{n^*} \approx \delta^2 \qquad (14)$$

where n* is the index of the last iteration at which the condition $J(H) < \delta^2$ is verified for the first time, δ^2 is the total (integrated) measurement error defined as:

$$\delta^2 = \int_0^{t_f} \sum_{i=1}^{N} \sigma^2\left(z_i, r_i, t\right) dt \qquad (15)$$

$\sigma^2(z,r,t)$ is the root-mean-square error of the temperature measurements obtained by smoothing the measured temperature histories. The number n* is the regularization parameter of the method.

APPLICATIONS

In order to verify the method, two materials have been analyzed. One is nickel, without metallurgical transformation; another one is 16MND5 steel, the thermophysical characteristics $\rho(T)$, $\lambda(T)$ and $C(T)$ of which in function of the phases and the temperature, as well as the C.C.T. diagram.

Theoretical case

In the paper, Le Masson et al. [3], we have used several theoretical heat transfer coefficients for the verification of the algorithm. Here, we show the results only for the theoretical heat transfer coefficient shown above (Figure 6). The theoretical temperatures are simulated for a 16MND5 steel and in the depth z=0.15mm and z=1mm. Figure 7 shows the heat transfer coefficient estimated from the thermal cycles simulated in z = 1mm. We note that the damping effect due to the depth generates a decrease of the boiling peak of H_{exp}.

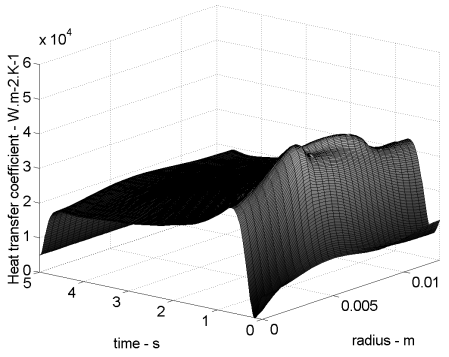


Figure 7: H(r,t) estimated from the thermal cycles simulated in z=1mm.

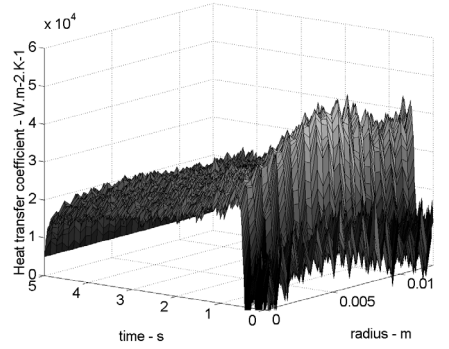


Figure 8: H(r,t) estimated from the noise-added thermal cycles simulated in z = 0,15mm.

When a noise is added to the thermal histories, the estimated values (Figure 8 for z = 0.15mm and Figure 9 for z = 1mm) remain satisfactory. A decrease of the boiling peak, however, is observed for the function H_{exp}. When the criterion

J(H) is approximately equal δ^2, the iterations are terminated.

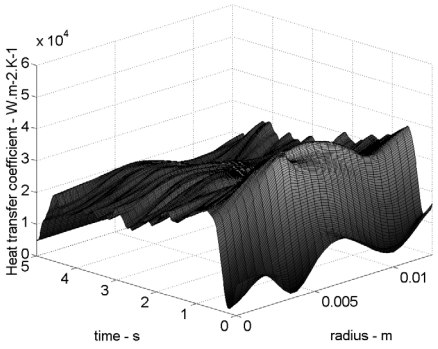


Figure 9: H(r,t) estimated from the noise-added thermal cycles simulated in z = 1mm.

The last result (Figure 10) we present here is a test without noise relating to the estimation for thermal cycles simulated in z = 0.6mm at a few measurement points in the radial direction N = 8 (eight thermocouples uniformly distributed with thirty space nodes)

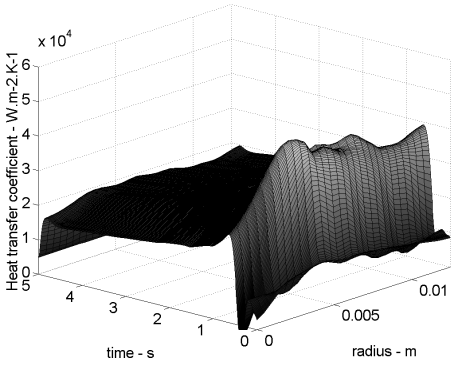


Figure 10: H(r,t) estimated from the thermal cycles simulated in z = 0.6mm.

Experimental cases

Table 1 : Thermocouple positions (radius r and depth z)

r (mm)	0.5	2.5	3.5	4.5	5.5	6.5
z (mm)	0.6	0.6	0.6	0.6	0.6	0.6
r (mm)	8.5	10.5	11.9	11.9		
z (mm)	0.6	0.6	1.5	3		

The two experimental cases have been treated first with a Nickel sample and second with a 16MND5 steel sample. A good reproducibility of experiments permits to compare estimation results. In the both experimental cases, the thermocouple positions were the same. The table 1 resume the positions of embedded thermocouples. We note the delta Fourier number (ΔFo) is around 0.11.

These K-type thermocouples (ϕ =70μm) are embedded perpendicularly of the axis in holes (ϕ = 0.6mm) and isolated by alumine tubes. After the heating during 30 minutes up to around 880°C, we obtain a complete austenite transformation (defined by the C.C.T. diagram – 16MND5 steel case) and the heat treatment can begin. The data acquisition begins before the sample cooling by the water jet and terminate after five minutes. Our principal goal is the estimation of the heat transfer coefficient at the beginning of the heat treatment. So, here, we present only these critical values (t < 10s for the 16MND5 steel and t < 20s for Nickel). The Figures 11 and 12 present the estimation results: Figure 11 for the Nickel sample and Figure 12 for the 16MND5 steel sample.

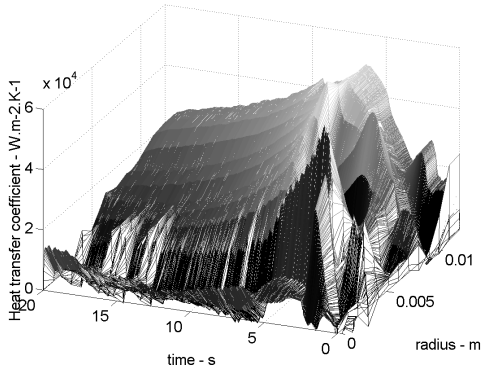


Figure 11: Estimation result for the Nickel sample ($0 \le t \le 20$s).

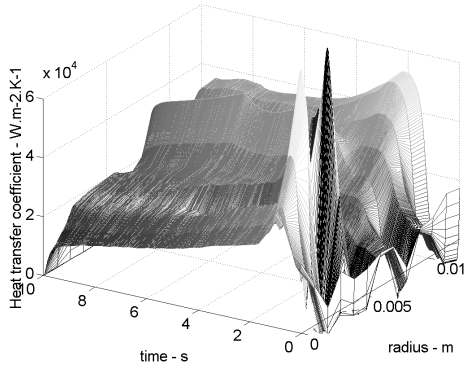


Figure 12: Estimation result for the 16MND5 steel sample ($0 \le t \le 10$s).

In the both cases, we can do the following remarks:

a) First, we can see that the forms of these heat transfer coefficients are the same.

b) During the cooling by the water jet, we have a highly transient and highly nonlinear phenomenon. We can see in the center of the sample an apparition of vapor and boiling phases (peaks for r lower of 5mm). In fact, the diameter of the water jet and that of the sample are 12.5mm and 25mm respectively. So, at the beginning the cooling at the center zone is more intensive.

c) After the stabilization of the water jet, we can remark that the heat transfer coefficient is really a function of radius and time. We can think, like as our second hypotheses, that the stagnation point or the stagnation region is established on the axis. The heat transfer coefficient grows up with the radius. So we have the highest heat transfer coefficient at r = 12.5mm.

d) We can see the evolution of the heat transfer coefficient in Figure 13. At the beginning, we have a big variation between all heat transfer coefficient values. We can remark that at t=1s for the small radius (r<5mm) we have an ebullition phenomenon. When the time increases, the variation on the heat transfer coefficient becomes smooth and regular.

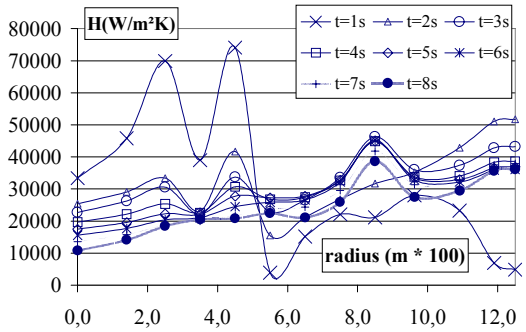


Figure 13: Heat transfer coefficient variations for the 16MND5 sample.

e) At last, we draw as in Figure 14 the heat transfer coefficient evolution versus the temperature of the quenching surface. We can remark that for all space nodes of the quenching surface the heat transfert coefficient has a maximum value between 100°C and 200°C. Then, for a temperature around 100°C, the heat transfer coefficient goes down. In fact, we suppose that we have a super-critical heat flux between 100°C and 200°C and forced convection under 100°C. (This result is similar of those given by Chen et al. [2]).

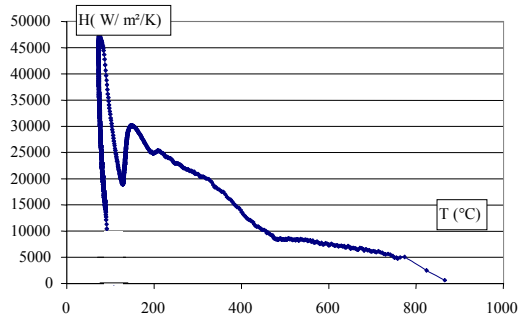


Figure 14: Heat transfer coefficient variations with temperature.

We note that the total measurement error δ^2 was computed by smoothing the measurd temperature histories. As the result we calculate the value of δ^2. The noise maximum amplitude was estimated to be around ΔT=10°C. Figure 15 shows the comparaison of three measured and estimated temperature histories for the nickel sample.

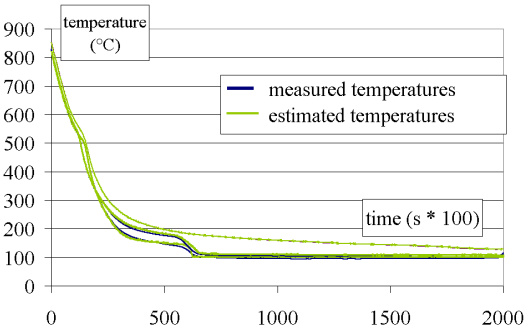


Figure 15: Comparison between measured and estimated temperatures.

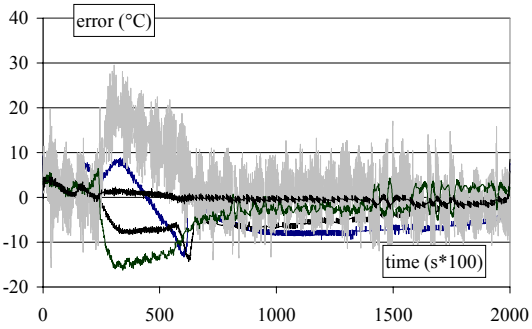


Figure 16: Error between five measured and estimated temperatures

In Figure 16, we show the residuals between the measured and estimated temperatures. We can see that the maximum residual values correspond to the value estimated above.

CONCLUSION

The objective of the presented experimental and numerical study was the estimation of a 2D heat tranfer coefficient during a heat treatment, the "Jominy" end-quench test. This 2D heat transfer coefficient is considered as a time and radius function. It is estimated at the beginning (20s) of the heat treatment of the sample heated up to 880°C by the water jet (20°C). During this period, we have a highly transient and highly nonlinear phenomenon. In the center of the sample, we have an apparition of vapor and boiling phases (peaks for r lower of 5mm). After the stabilization of the water jet, we estimate that the heat transfer coefficient is really a function of radius and time. In the center, a stagnation region is developed and at $r = r_{max}$ the heat transfer coefficient is maximum. The heat transfer coefficient values for the nickel sample and the steel sample are quite similar. We also analyzed a variation of the heat transfer coefficient versus temperature. A maximum value is obtained for a temperature of the quenching surface between 100°C and 200°C.

We have used the Iterative Regularization Method to carry out the estimation. The total measurement error δ^2 was obtained by smoothing the measured temperature histories. The noise maximum amplitude was estimated to be around to 10°C. The obtained results was verified by comparing the measured and estimated temperature histories.

In our future works we should pay more attention on the confirmations of the physical analysis and the CCT diagram constructions.

REFERENCE

1. O.M. Alifanov, E.A. Artyukhin, S.V. Rumyantsev, Extreme Methods for Solving Ill-posed Problems with Applications to inverse Heat Transfer Problems, Begell house, New-York, 1995.

2. P. Le Masson, D. Carron, P. Rogeon, J.J. Quemener, Identification du coefficient de transfert lors d'un essai Jominy instrumenté, SFT 99, 57-62, 1999.

3. P. Le Masson, T. Loulou, E. Artioukhine, P. Rogeon, D. Carron and J.J. Quemener, A new approach for the estimation of a convection heat transfer coefficient during a metallurgical « Jominy end-quench » test. *68th eurotherm seminar* - Inverse Problems and experimental Design in Thermal and Mechanical Engineering (Poitiers, France, 2001)

4. A. Osman, J.V. Beck, Investigation of transient heat transfer coefficient in quenching experiments, J. Heat Transfer **112**(1990) 843-848.

5. S.G. Chen, C.I. Weng, J. Lin, Inverse estimation of transient temperature distribution in the end quenching test*, J. of Materials Processing Technology*, **86** (1999) 257-263

6. P. Archambault, S. Denis and A. Azim, Inverse resolution of the heat transfer equation with internal heat source: Application to the quenching of steels with phase transformations, J. of Mat. Eng. and Perf., **6,** 1997, 240-246

7. Essai de trempabilité Jominy ; NFA 04-303 ; juin 1979

8. Atlas de courbes Jominy (états bruts de trempe et revenus) d'aciers de la norme NFA 35-552 ; Collection ATS OTUA ; Paris 1985.

9. P. Le Masson, P. Rogeon, D. Carron, J.J. Quemener, Simulation numérique thermométallurgique : influence des paramètres d'entrée sur la modélisation d'un essai de trempabilité, Revue de Métallurgie – CIT/ Sciences et Génie des Matériaux, pp1055-1064, Sept. 2000

10. J. Leblond J. Devaux, A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size ; Acta Metall., **32 n°1**, 137-146, 1984.

11. D. P. Koistinen, R. E. Marburger, A general equation prescribing the extent of austenite-martensite transformation in pure iron-carbon alloys and plain carbon steels; Acta. Metall., **7**, 59, 1959.

12. A. Sorin; Modélisation des conditions de refroidissement d'une éprouvette Jominy ; Rapport de stage - Centre d'Etude des Structures et des Matériaux Navals (CESMAN); Juin 1997.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSE TECHNIQUE APPLIED IN WELDING: A THEORETICAL AND EXPERIMENTAL APPROACH

Cristiene V. Gonçalves, Américo Scotti, Gilmar Guimarães
School of Mechanical Engineering (FEMEC)
Federal University of Uberlândia (UFU)
Uberlândia, MG, Brazil
cgoncalves@mecanica.ufu.br, ascotti@mecanica.ufu.br, gguima@mecanica.ufu.br

ABSTRACT

In this work, the application of the Simulated Annealing Inverse Technique to estimate the temperature history in GTA welded workpiece was assessed. The method uses transient temperature data from thermocouples attached at the opposite face of the welding bead and a quasi-steady model. In this case, a two-dimensional model with moving heat sources is used, considering the material thermal properties constant during all the welding process. The component of the heat flux input that goes into the workpiece and the efficiency of the GTAW process were obtained. The results presented a good agreement between the estimated and measured temperature cycle. It was shown a great potential of using inverse procedure in welding process.

NOMECLATURE

C	arbitrary constant, m^{-1}
C_p	specific heat, J/kg- ^oC
d	plate thickness, m
F	least square function, $[^oC]^2$
I	welding current, A
k	thermal conductivity, $W/m^2\ ^oC$
K_0	modified Bessel function of the first kind of order zero,
K_b	the Boltzmann´s constant.
q	heat flux input, W/m^2
T	temperature, ^oC
t	time, s
T_0	initial workpiece temperature
V	arc voltage, V
v	welding speed, m/s
x, y, z	moving Cartesian coordinates, m
x', y', z'	fixed coordinate system, m
Y	measured temperature, ^oC

GREEK SYMBOLS

α	thermal diffusivity, m^2/s
ρ	density, kg / m^3
η	efficiency of the arc

INTRODUCTION

Gas Tungsten Arc Welding (GTAW) is an arc welding process, in which a voltaic arc is formed between a non-consumable electrode and a metallic workpiece. The joint (weld bead) is achieved trough the fusion of the plate edges and material feeding is optional. For that, a very intense and moving heat source is applied over the workpiece. Part of this heat is used to melt the plate, part is sunk into the workpiece and part is lost in the environment and in the electrode cooling system. The prediction of the way that the heat reaches the plate (intensity, concentration, distribution, temperature gradient, etc.) becomes extremely useful for understanding welding phenomena, such as the bead formation (width and penetration depth) or microstructure changes in the heat-affected base metal. Kraus [1], for instance, cites the application of this predictions on the study of the metallurgical aspects of welds, e.g., grain growth, size, and orientation, and solid-state phase transformations, which in turn lead to the understanding of associated mechanical properties, e.g., residual stress, distortion, ductility, and strength.

Many investigators have studied welding heat flow problems, analytically, numerically, and experimentally. One can find studies about pulsed current GTA weld pool [2,3], penetration model of welding [4,5] or heat conduction analysis of bi-directional multipass welding [6], among others. The majority of the work, however, considers the heat flux input to be known for determining the temperature field. For example, Tsai [3] studied the pulsed current GTAW process considering a radial symmetric normal function for the heat transfer hate. Krauss [1] has formulated the steady state and transient heat transport for a thin-plate GTA welding using, in the same way, circular or elliptical Gaussian distribution for the welder arc heat flux.

In fact, the majority of the assumptions about the distribution are modeled based on the thermal efficiency of the welding process, that is, the net quantity of heat flux input to the workpiece. Usually, the values of this efficiency are either assumed to be known and taken from literature or determined by using calorimetric techniques.

Inverse Heat Conduction Method (IHCM) has recently been employed in welding process analyses. The choice of inverse procedure represents a good alternative to calorimeter methods (easier, cheaper and faster); actual heat flux can be obtained using experimental data from a thermocouple located in a region far from the welding zone. Most of applications treat the problem as a one-dimensional problem [7], stationary arc welding processes [8] or inverse phase-change in stationary conditions [9]. The Simulated Annealing Inverse Technique, one of the Inverse Technique methods, has

shown robustness and presented results with a very good degree of confidence.

There are, however, a large and promising field for studying new applications and enhancement of this technique when applied in welding, especially because there is the opportunity of using theoretical models applied over experimental data. Thus, in this work is aimed a contribution for the art state, through the application of the Simulated Annealing Inverse Technique on the estimation of the heat flux and, consequently, the determination of the efficiency and the temperature distribution during actual GTA welding.

THEORETIC FUNDAMENTALS

Gas Tungsten Arc Welding - GTAW

The GTAW process is commonly used in the fabrication of heat-sensitive materials, thin gage joints and dimensionally critical structures. The development of this process was accelerated early in 1940. Initially, the process was called Heliarc, since helium was used as shielding gas. Later, when argon became available, the process was renamed to Tungsten Inert Gas or "TIG". Nowadays, the designation Gas Tungsten Arc Welding (GTAW) turn out to be more generic and preferably accepted, since other gases rather than argon and helium, which are inert, can be mixed with them. Hydrogen, for example, is included for its special benefits [10].

The GTAW equipment basically consists of a power supply, a welding torch and connecting cables. The torch utilizes a non-consumable tungsten electrode, usually alloyed with oxides, such as thoria or ceria, to improve the electrode performance. Electrodes are attainable in various diameters, ranging from 0.25 mm up to 6.25 mm, and are accommodated in the torch by an adjustable collet. The torch also includes a shielding gas nozzle through which the shielding gas flows to cool the electrode and to shield the weld "puddle" or pool from oxygen (cooling of some torches are aided by water). Nozzles are selected from a range of sizes, materials and configurations. Ceramic (alumina) are the most popular nozzle material, yet metal and fused quartz can be used in special cases, the latter, which is translucent, to improve the weld pool visibility. Shape or configuration of the nozzles is based on aerodynamic principles and compatibly with special application requirements involving joint accessibility.

Power supplies are the constant or dropping output types and use either DC or AC current with transformers and rectifiers. Direct current electrode negative (DCEN) is the most often used. Some material, such as aluminum, demand alternate current, due to its oxide removal capability. Power supplies can also incorporate a pulsed generator to give special abilities for the output signal.

Gas tungsten arc welding is a thermal process that depends on the conducted heat to achieve penetration into the materials. In this respect, its conductance mode is similar to laser and electron beam (almost punctual heat source). They differ substantially from GTAW, however, in the thermal impingement area dimension and their rates of energy input. The thermal model considered here for the GTAW phenomena, however, is a moving punctual heat source with constant speed in a thin plate. The model is presented next.

Direct Problem

Solutions to the heat flux equation of a moving punctual heat-source were given by Rosenthal in 1935 [11]. He assumed that the energy of the heat source moves with a constant speed v along the x-axis of a fixed rectangular co-ordinate system, as shown in Figure 1. In this case, the heat flux (q) can be defined by,

$$q = \eta \cdot V \cdot I \tag{1}$$

where V and I represent the arc voltage and welding current, respectively. The variable η is the thermal efficiency of the arc and can be defined by

$$\eta = \cdot \frac{q}{V \cdot I} \tag{2}$$

Considering Figure 1, the general energy equation for an isotropic and homogenous material, with constant properties and negligible heat losses, is reduced to :

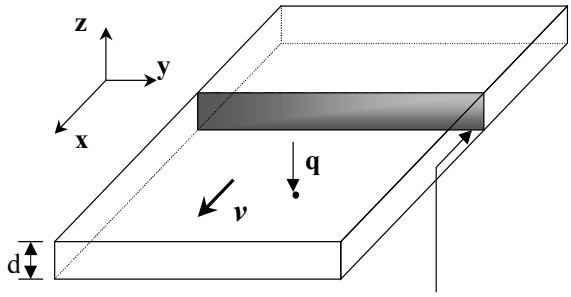


Figure 1. Schematic diagram of the welding process on a flat metal plate with a punctual heat source, q, and a constant velocity, v [11]

$$\frac{\partial^2 T}{\partial x'^2} + \frac{\partial^2 T}{\partial y'^2} + \frac{\partial^2 T}{\partial z'^2} = \frac{1}{\alpha}\frac{\partial T}{\partial t'} \tag{3}$$

where α is the thermal diffusivity and x'y'z' is the fixed coordinate system. Consider now the condition of a moving punctual heat source (q) traveling with a constant velocity (v) in the x direction (Fig. 1). If the source is assigned to the origin of a moving coordinate system xyzt, then the transformation of a point *P(x', y', z')* in the fixed system by $x = x' - vt$ becomes *P(x, y ,z, t)* in the moving system xyzt. The transformed form of Eq.(3) then becomes:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = \left(\frac{1}{\alpha}\cdot\frac{\partial T}{\partial t}\right) - \left(\frac{v}{\alpha}\frac{\partial T}{\partial x}\right) \tag{4}$$

Assuming that the workpiece is long enough, in such a way that the "quasi-steady-state" condition can be established, and that the workpiece thickness is much smaller than its width and length, so that the temperature gradients in the z direction can be neglected, Eq. (4) reduces to:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = \frac{v}{\alpha}\frac{\partial T}{\partial x} \tag{5}$$

Equation (5) can be solved using the separation of variable method, taking a solution of the form as follows:

$$T = T_0 + \exp(-C \cdot x) \cdot g(x, y) \tag{6}$$

in which T_0 is the initial workpiece temperature and C is an arbitrary constant. The solution that can also be found in the literature [12] is:

$$T = T_0 + \frac{q}{2 \cdot \pi \cdot k_o \cdot d} \exp\left(-\frac{v \cdot x}{2 \cdot \alpha}\right) \cdot K_0\left(\frac{v \cdot r}{2 \cdot \alpha}\right) \tag{7}$$

where, K_0 is the modified Bessel function of the first kind of order zero and $r = \left(x^2 + y^2 + z^2\right)^{1/2}$.

The assumptions involved with Eqs. (6) to (7) include: a two-dimensional coordinate system with no radial effects, negligible heat loss from the plate due to convection or radiation, the source being a punctual source, and a quasi-steady condition in the workpiece.

Inverse Problem: Simulated Annealing (SA)

Simulated Annealing is derived from an analogy with the annealing process of material physics [13]. In the process of annealing, the metal is heated up to a high temperature, causing the atoms to shake violently. Providing that the cooling is slow enough, the metal will eventually stabilize into an orderly structure. Otherwise, unstable atom structure is found. There is plenty of information on the current literature, so that in depth description of this technique is out of this work scope.

There are several inverse techniques besides of SA that can solve this welding problem. It can be cited, for example, the conjugate gradient with adjoint equation method [14], parameter estimation approaches [15], sequential time domain [16] or genetic algorithm [17]. All methods cited above uses experimentally determined histories in the sample studied (workpiece) to calculate the corresponding input heat flux for a given set of system parameters (welding). The Simulated Annealing was choice due, mainly, of his robust characteristic.

Simulated Annealing can be performed in optimization by randomly perturbing the decision variable and keeping track of the best objective function value for each randomized set of variables. After many trials, the set that produced the best objective function value is designed to be the center, over which perturbation will take place for the next temperature. The temperature, that in this technique is the standard deviation of the random number generator, is then reduced, and news trials performed.

The objective function can, then, be defined as the least-square residuals F between the computed temperature T_{ji} and the measured temperature Y_{ij} ,

$$F = \sum_{j=1}^{J}\sum_{i=1}^{M}\left(Y_{ji} - T_{ji}\right)^2 \tag{8}$$

were j represents the index for the thermocouple position and i the index for measurement time.

Let each configuration be defined by the set of atom positions where E represents the energy of the configuration and T is the temperature. In each step of the configuration, an atom is given a small random displacement and the resulting change, ΔE, in the energy of the system is computed. When the processes generating new states, this states is either accepted or rejected, according to the metropolis criterion [18]: if $\Delta E \leq 0$, the displacement is accepted and this configuration is used as the starting point of the next step. If $\Delta E \geq 0$, the probability that the configuration is accepted is given by the following equation:

$$P(\Delta E) = e^{(-E/K_b T)} \tag{9}$$

where K_b is the Boltzmann´s constant. The choice of the probability function given by Eq. (9) has the consequence that the system evolves, according to a Boltzmann distribution

EXPERIMENTAL PROCEDURE

Figure 2 represents the welding rig used in the experimental procedure. The TIG torch, representing the punctual heat sources, moves at a constant speed along a straight path, using a totally automated X-Y coordinate table. Attention was paid on the end effects, which results from either starting or of the heat source and cause a transient thermal state in the weldment. Some time after the heat source starting or some time before the heat source extinguishing, the temperature distribution is stationary with respect to the moving coordinates, which origin coincides with the center of the heat source. These regions were eliminated of the analysis. To prevent from test-plate dimension interference, the test-plate was hold suspended in air by 4 pointed cylindrical bars, so that, only very small contact area existed.

Figure 3 illustrates with more details how 10 thermocouples (type K) were located underneath the test-plate, that is, at the opposite face to the weld bead (heat source). They were equally distanced of each other (16,6 mm), being that the first one was 25 mm from of the test-plate edge. Capacitor discharge was used to stick the thermocouple ends to the plate surface.

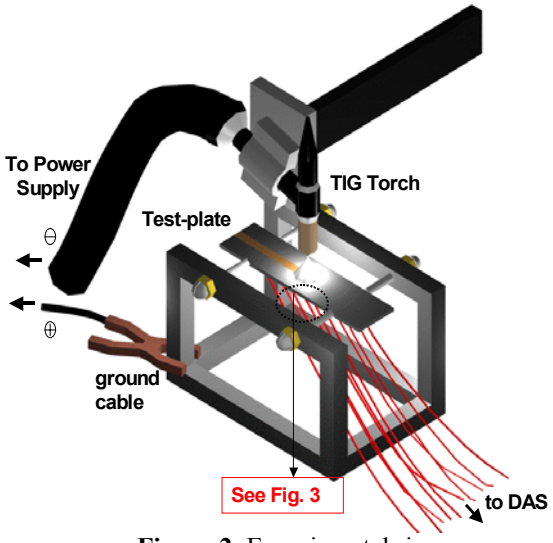


Figure 2. Experimental rig

The procedure used is, however, prone to error, since the accuracy of heat flux estimation procedure in most of the inverse heat conduction technique decreases with the increasing distance from the moving source [19]. Ohadi and Cheng,[12] showed how they managed with this behavior in a modeling of temperature distributions in a workpiece during abrasive waterjet machining. In that case, since the measurement target was the cutting face, to avoid this problem, they positioned the thermocouples slightly away from the cutting interface, but short enough to minimize the difference between the calculated and measured temperatures.

In welding applications the problem is even worse. Due to the high temperatures, the use of the same face on which the heat source is applied is very difficult. The thermocouple could be dammed. Supported by Ohadi and Cheng's approach, that is the reason why the opposite face was used for measuring but considering the heat source position.

The test-plate was made of stainless steel AISI304. Table 1 presents the thermal properties and dimensions of the test-plate. Those properties were the ones used in the calculations.

Table 1. Workpiece base configuration parameters

Material	AISI 304
Thermal Conductivity	15.0 (W/mK)
Thermal Diffusivity	3.96 10^{-6} (m²/s)
Melting Temperature	1427.0 (°C)
Welding Current	102 (A)
Welding Voltage	10.8 (V)
Plate Length	0.2 (m)
Plate Width	0.05 (m)
Plate Thickness	0.004 (m)

Collection and storage of the data from the thermocouples took place with the aid of a microcomputer-based data acquisition system (HP 75000 B E1326B), DAS for short. The DAS, under a software control, sampled (multiplexed) each thermocouple signals for 0.38 s (totaling 1028 points for each).

The first step in the analysis was the calculation the heat flux input to workpiece using the inverse technique (SA). Figure 4 presents the thermal cycles experimented by each of the 10 thermocouples. These temperature data are used in Eq. (8) in order to estimate the heat flux that goes into the workpiece.

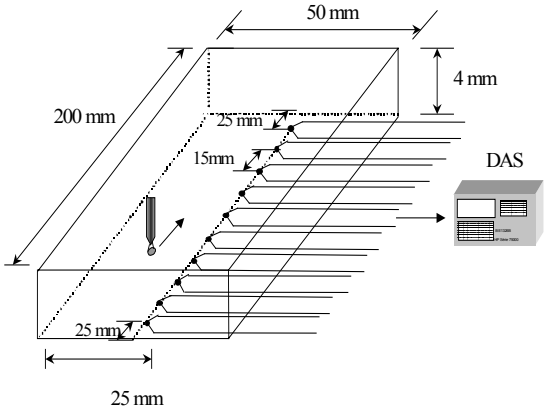


Figure 3. Thermocouple positioning on the test-plate

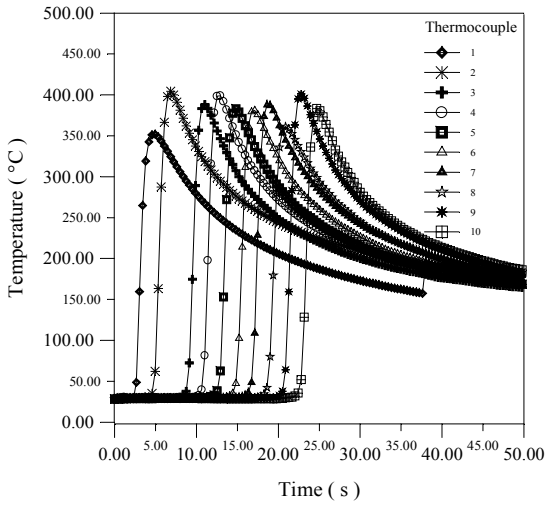


Figure 4. Thermal cycle registered during the experiments

RESULTS AND DISCUSSION

The procedure used to estimate each heat flux component could be applied for all measured temperatures over all measurement times. However, in this work, for minimization, only the temperatures from the thermocouples that the heat source had passed through were used to solve Eq. (8). It means that there is an optimal time for each thermocouple, which corresponds to the time that each thermocouple reached

its maximum temperature. It can be expected that, when the heat source is just above the thermocouple (Fig. 3), actually a little ahead (there is a diffusion time), the temperature for this thermocouple should be maximum. It can also be observed, in this case, that the thermocouple located downstream (the position of thermocouple is related to the time by the variable $x = x' - vt'$), has no or little information about the heat flux source. So, its use should be avoided. In other words, the number of measured temperatures used in Eq. (8) has an upper limit defined by the position of the heat source for each instant. These times are made discrete by the optimal times that each thermocouple took to reach the maximum temperature.

Table 2 presents the optimum times and the corresponding positions, calculated by the Simulated Annealing Technique. Table 2 also shows which thermocouples were used to minimize the objective function, F. The maximum temperatures reached by each thermocouple are shown in Table 3.

Table 2. Optimum times and the corresponding position

Time [s]	position x [mm]	thermocouple used
7.22	0.0250	1
9.12	0.0416	1 to 2
11.40	0.0582	1 to 3
13.30	0.0748	1 to 4
15.20	0.0914	1 to 5
17.48	0.1080	1 to 6
19.00	0.1246	1 to 7
21.28	0.1412	1 to 8
23.18	0.1578	1 to 9
25.08	0.1744	1 to 10

Table 3. Maximum temperatures for each thermocouple

Thermocouple	T_{max} [°C]
1	404.5
2	407.6
3	388.2
4	399.2
5	383.2
6	381.1
7	387.6
8	361.6
9	401.2
10	383.3

One consequence of using only the maximum times is to obtain a heat flux component for each maximum time used. Table 4 and Figure 6 show the heat flux to the workpiece estimation considering the optimal times and the thermocouples used.

Table 4. Estimated heat flux and arc efficiency for different maximum times (Thermocouple used according to Table 2)

Time [s]	Heat Flux [W]	Arc Efficiency
7.22	701.3641	0.6367
9.12	718.1497	0.6519
11.40	719.1958	0.6529
13.30	724.3361	0.6575
15.20	712.2252	0.6465
17.48	717.1894	0.6510
19.00	710.6900	0.6451
21.28	706.3974	0.6413
23.18	719.2376	0.6529
25.08	716.5645	0.6505

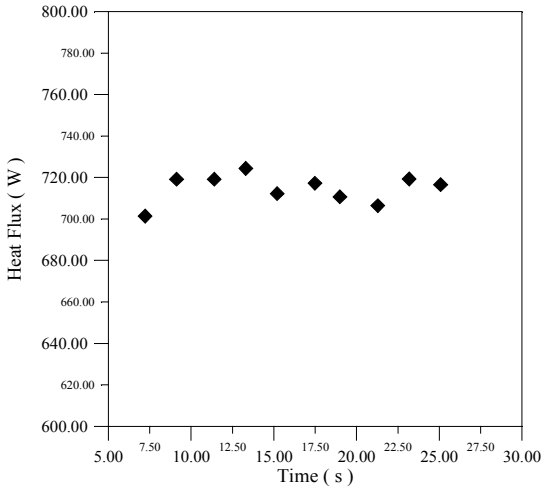


Figure 6. Estimated heat flux for different optimum times

If only the maximum temperatures are used in the inverse procedure, the estimated heat flux input to the workpiece is 677 W, with a corresponding efficiency of 0.62. In fact, this heat flux input represents the average of estimated heat flux components, in the case that only one maximum temperature was used at a time.

In this model, factors like metallurgical phase change, three-dimensional analysis and thickness effects are not considered. However, the results are good enough to indicate confidence values for the efficiency, heat flux and temperature estimations. Unfortunately, there is no way in this work to compare the estimated and measured values for heat flux and arc efficiency, so that a better assessment of the technique validation could be done. However the arc efficiency found here is situated in the range of typical values expected for GTWA (0.25 to 0.75) [20].

Once the heat flux to the workpiece is estimated, the temperature distribution can be obtained solving the direct problem given by Eq. (7). Figs. 7 and 8 presented some temperature evolutions at three different times in a two-dimensional and three-dimensional ways, respectively.

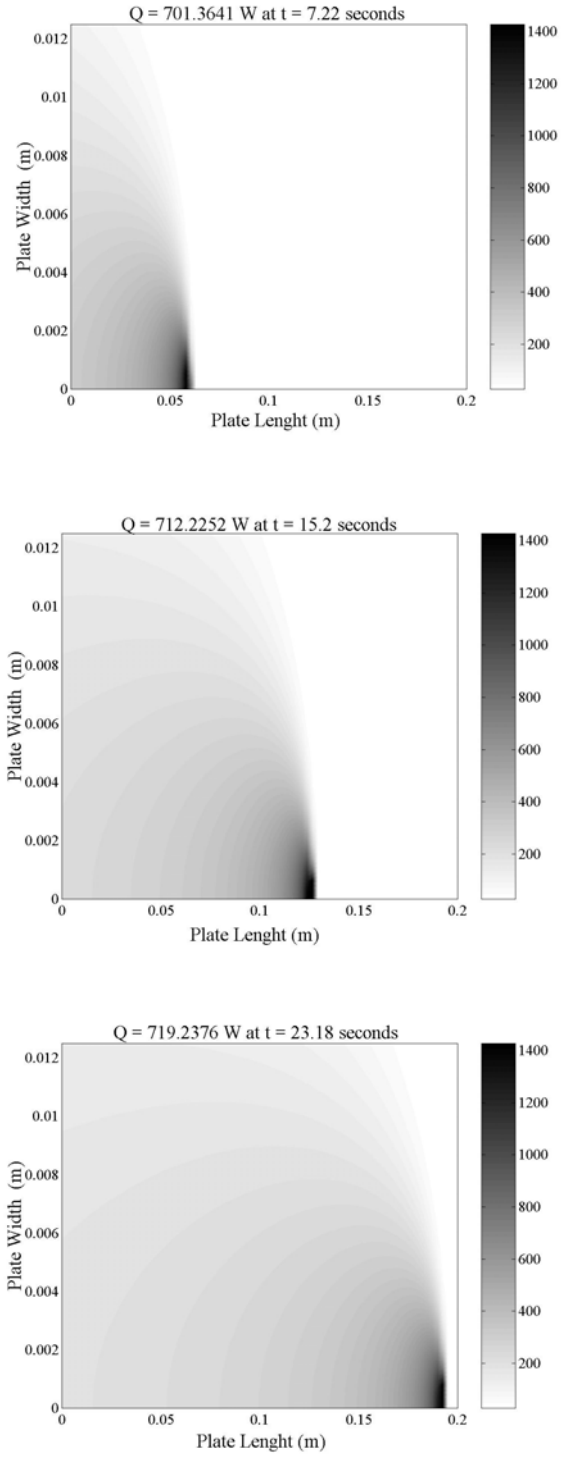


Figure 7. Two-dimensional presentation of temperature distribution in the test-plate at different times

The effect of the moving source can be observed in these figures. Since a punctual source is assumed in the model, the highest value of temperature is at the center of the torch.

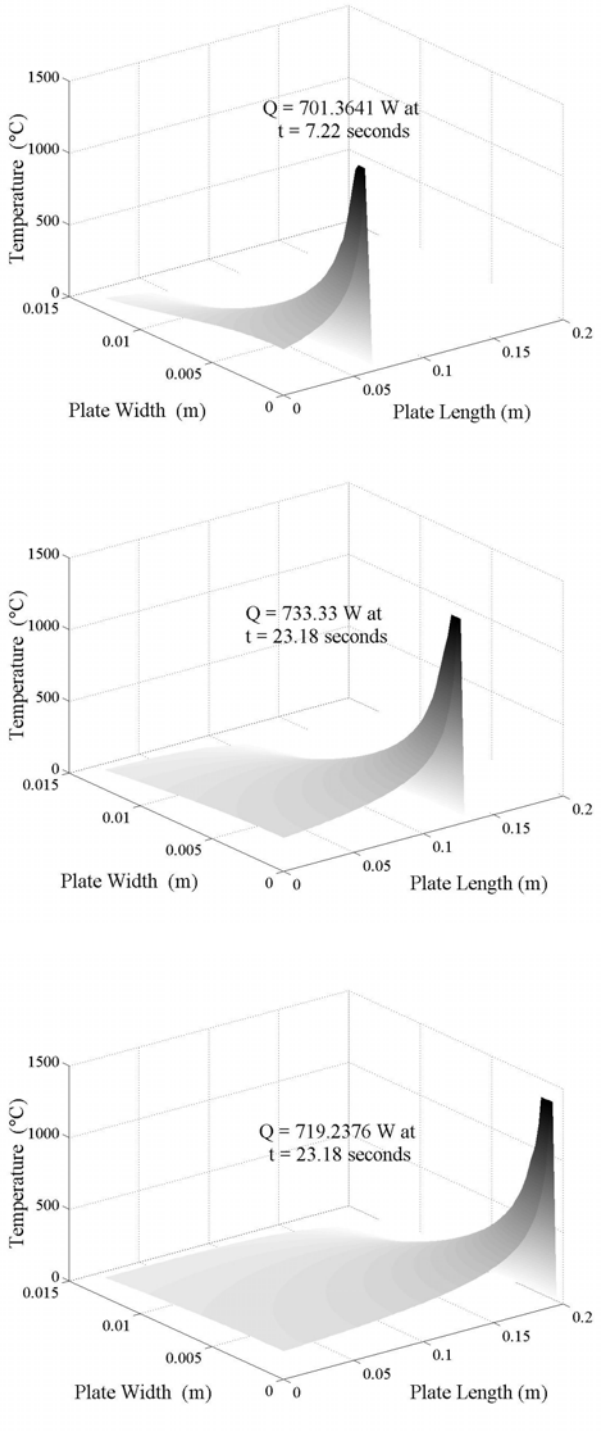


Figure 8 Three-dimensional presentation of temperature distribution in the test-plate at different times.

The comparison between the estimated and experimental temperatures is shown in Fig. 9. One can see a satisfactory agreement. The deviations can be attributed to various factors besides the common experimental difficulties, such as the assumption of a two-dimensional quasi-steady-state theoretical model. It can be noted, in this way, the systematic error caused by neglecting the influence of the plate thickness, radiation and convection at the surface variations of thermal

properties. A work that considers all these effects is in currently progress.

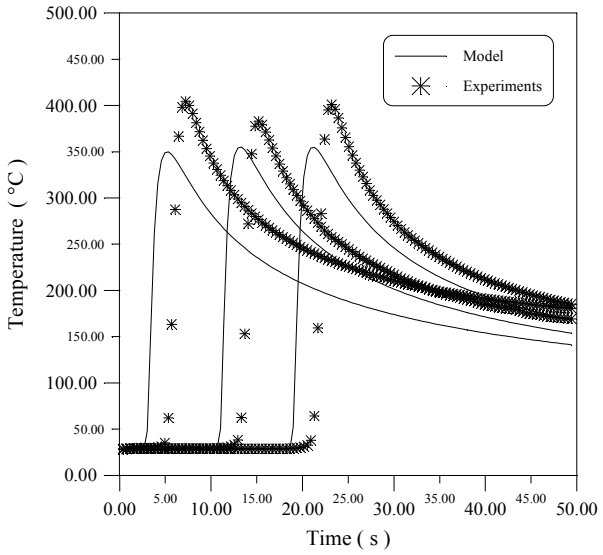


Figure 9. Comparison between estimated and experimental temperatures at the opposite face of the welding bead

6. CONCLUSION

The theoretical and experimental approach, based on the application of the Simulated Annealing Inverse Technique, showed to be a promising tool for calculating the thermal history experimented by a workpiece under welding. The agreement between the predicted temperature and the measured temperature are satisfactory, what leads for the reasoning that the estimated values for heat flux and arc efficient are reliable. However, new enhancements must be added to the models, so that more precise outcome can be expected.

ACKNOWLEDGEMENTS

The authors would like to thank Brazilians agents for research development, CAPES, CNPq and FAPEMIG, for financial support.

REFERENCES

1. Kraus, H. G., Thermal Finite Element Formulation and Solution Versus Experimental Results for Thin-Plate GTA Welding, *Journal of Heat Transfer*, **108**, pp. 591-596 (1986)

2. Kim, W.-H., Na, S.-J., Heat and Fluid Flow in Pulsed Current GTA Weld Pool, *Int. J. Heat and Mass Transfer*, **41**, 3213-3227 (1998)

3. Tsai, C. L. and Hou, C. A., Theoretical Analysis of Weld Pool Behavior in the Pulsed Current GTAW Process, *Journal of Heat Transfer*, **110**, pp. 160-165 (1988)

4. Liangyu L., and Ruidong, F., Penetration Analytical Model of Three Dimensional Welding Temperature Field and Experiment Rectification, *China Welding*, **9**, pp.161-166 (2000)

5. Lankalapalli, K., Model-Based Laser Weld Penetration Depth Estimation, *Welding in the World*, **39**, pp. 304-313 (1997)

6. Kasuya, T., Hashiba, H., Ohkita, S. and Fuji, M., Heat Conduction Analysis of Bidirectional Multipass Welding with Short Bead Lengths, *Science and Technology of Welding and Joining*, **5**, 215-220, 2000.

7. Katz, M. and Rubinsky, B., An Inverse Finite Technique to Determine the Change of Interface Location in One-Dimensional Melting Problem, *Numerical Heat Transfer*, **7**, pp. 269-283 (1984)

8. Hsu, Y. , Rubinsky, B., and Mahin, K., An Inverse Finite Element Method for the Analysis of Stationary Arc Welding Process, *ASME J. Heat Transfer*, **108**, pp. 734-741 (1986)

9. Al-Khalidy, Nehad, Application of optimization method for solving inverse phase-change problems, *Numerical Heat Transfer*, *part B*, pp.477 – 497 (1997).

10. Joining Technologies - Reference Center http://www.joiningtech.com/html/refgtaw.html 01.24.2002

11. Easterling, Kenneth, *Introduction to the Physical Metallurgy of Welding*, Butterworths&CoLtd. , London, UK (1983*)*

12. Ohadi, M. M. and Cheng, K. L., Modeling of Temperature Distributions in the Workpiece During Abrasive Waterjet Machining, *Journal of Heat Transfer*, **115**, pp. 446-45217 (1993),

13. Saramago, S. F. P., Assis, E. G., & Steffen, V., Simulated annealing: Some applications in mechanical systems optimization, *20th Iberian Latin – American Congress on Computational Methods in Engineering* (CD – ROM), São Paulo, Brazil (1999)

14. Ozisik, M. N. and Orlande, H. R. B., *Inverse Heat Transfer*, Taylor & Francis, NY, USA (2000)

15. Beck, J. V. e Arnold, K. J., Parameter estimation in engineering and science*, John Wiley and Sons Inc*., New York, 501 p. (1977)

16. Beck, J. V., B. Blackwell and C. R. St. Clair, *Inverse Heat Conduction, Ill-posed Problems*, Wiley Interscience Publication, New York (1985)

17. Gonçalves, C. V., Silva, L. A., Silveira Neto, A., Rade, D. A. and Guimarães, G., An inverse Technique Applied to Natural Convection Over a Heated Vertical Plate, *proceedings of National Heat Transfer Conference*, Pittsburg, USA (2000)

18. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. and Teller, A. H., Equations of State Calculations by Fast Computing Machines, *Journal of Chemical Physics*, 21, pp. 1087-1092 (1953)

19. Taktak, R., Beck, J. V. e Scott, E. P., Optimal Experimental Design for Estimating Thermal Properties, *Int. J. of Heat and Mass Transfer*, **36**, N° 12, pp. 2977-2986 (1993)

20. Grong, O., *Metallurgical Modelling of Welding*, The Institute of Materials, London (1994)

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

IDENTIFICATION OF VELOCITY DISTRIBUTION IN A TURBULENT FLOW INSIDE PARALLEL-PLATE DUCTS FROM WALL TEMPERATURE MEASUREMENTS

Manuel Girault, Daniel Petit, François Penot
Laboratoire d'Etudes Thermiques, U.M.R. C.N.R.S. 6608, Université de Poitiers, ENSMA, Téléport 2, 1 avenue Clément Ader, B.P. 41109, 86961 Futuroscope Cedex, France, manuel.girault@let.ensma.fr, daniel.petit@let.ensma.fr, francois.penot@let.ensma.fr

ABSTRACT

An Inverse Heat Convection Problem is investigated in this numerical study. Turbulent forced convection is considered, with a hydrodynamically fully developed, thermally developing, incompressible, constant property flow inside a parallel-plate duct. Velocity and effective diffusivity distributions of the turbulent model are characterised by a Reynolds number b^+ based on the shear stress velocity.

The identification of b^+ from simulated wall temperature measurements is presented. A heat flux density is applied along the channel length and temperature responses are taken at the wall external surface. Heat transfer has to act as a small but measurable perturbation of the flow.

The inverse problem is recast into an optimisation problem solved with a Quasi-Newton method. Reynolds numbers 10^5 and 10^6, corresponding to two different values of b^+, are considered. Effects of sensor number and location, as well as magnitude of measurement error on the estimation accuracy, are examined.

NOMENCLATURE

b	channel half-height, m
b^+	Reynolds number based on shear stress velocity $= \frac{bu_\tau}{\nu}$
De	equivalent diameter, m
e	wall thickness, m
f	squared residues functional to minimise
f', f''	first and second derivatives of f with respect to b^+
H	approximation to f''
L	channel length, m
L^*	dimensionless channel length
$Nmes$	number of measurements
Pe	Péclet number
P_m	motion pressure, Pa
Pr	Prandtl number
Pr_{turb}	turbulent Prandtl number
Re	Reynolds number based on mean velocity
S	absolute value of sensitivity
s	search direction of the minimisation algorithm
T	temperature, K
U	dimensionless velocity $= \frac{u}{u_m}$
u_τ	shear stress velocity, m.s^{-1}
u	velocity, m.s^{-1}
u_m	fluid bulk mean velocity $= \frac{1}{b}\int_{y=0}^{y=b} u(y)dy$, m.s^{-1}
u^+	dimensionless velocity $= \frac{u}{u_\tau}$
x	axial coordinate, m
y	transverse coordinate, m
y^+	dimensionless transverse coordinate $= (1-\eta)b^+$

Greek symbols

α	fluid thermal diffusivity, m^2.s^{-1}
α_{turb}	turbulent diffusivity for heat, m^2.s^{-1}
ΔT	measurement error, K
ΔT_{ref}	characteristic temperature difference, K
ε_{eff}	dimensionless effective diffusivity for heat
Φ	dimensionless applied heat flux =1
η	dimensionless transverse coordinate $= \frac{y}{b}$
φ	applied heat flux density, W.m^{-2}
λ	thermal conductivity, W.m^{-1}.K^{-1}
λ^*	solid/fluid thermal conductance ratio
ν	fluid kinematic viscosity, m^2.s^{-1}
ν_{turb}	turbulent viscosity, m^2.s^{-1}

θ dimensionless temperature
ρ fluid density, $kg.m^{-3}$
σ standard deviation of added noise, K
τ shear stress, Pa
ω random number
ξ dimensionless axial coordinate

Subscripts not previously introduced
f relative to fluid
in inlet
k index for current iteration of the minimisation algorithm
s relative to solid
surf at wall external surface
w at wall i.e. at solid/fluid interface

Superscripts not previously introduced
calc calculated
mes measured

INTRODUCTION

Forced convection in turbulent flow inside ducts is frequently encountered in industrial equipment. Many available studies are concerned with the direct problem, that is: having a mathematical model and knowing boundary conditions, one searches to compute the temperature field over the whole domain**.** Applications in which thermophysical properties, geometric characteristics or boundary conditions are unknown, but temperature measurements are locally available, belong to inverse problems research area, i.e. from the knowledge of experimental data, one searches to estimate the unknowns. The field of Inverse Heat Conduction Problems has already been intensively covered but inverse problems related to convective heat transfer have been much less studied. Turbulent regime has been investigated in a few studies involving the identification of boundary conditions. Estimation of timewise variation of wall heat flux applied to a parallel-plate duct has been performed by Liu and Özisik [1] using the conjugate gradient method with an adjoint problem. A non iterative sequential method using a low order model has been proposed by Girault et al. [2] to simultaneously estimate timewise variation of two wall heat fluxes applied to a parallel-plate duct. For the thermal steady state, estimation of spatially non uniform wall heat flux in circular pipe flow has been made by Su et al. [3] using the Levenberg-Marquardt algorithm. In the present work another class of inverse problems is investigated. A turbulent flow inside a parallel-plate duct is considered and the possibility to use heat transfer as a small perturbation of the flow to estimate a parameter characterising velocity and effective diffusivity distributions is explored in this numerical study. A known heat flux density is applied on the upper and lower plates of the duct. Temperature responses, that depend on the unknown, are taken at the wall external surface. A squared residues functional built with these data and the heat transfer model solution has then to be minimised to estimate the unknown. The basic idea of using temperature or concentration measurements to get information about velocity field has been under development for the last decades. Mass transfer is often used for a more detailed study of the near wall region. As an example, turbulent fluctuations in the velocity gradient at wall have been estimated by Mao and Hanratty [4] using the measure of a mass transfer coefficient. The fluid is kept isothermal and mass transfer does not induce perturbations in the flow. Heat transfer is used at larger scales. For instance, sea surface temperature fields and surface meteorological data have been used by Zavialov et al. [5] to obtain seasonal near-surface velocity fields in the southern Brazilian shelf. The present work is concerned with an industrial application. The system is first described, governing heat transfer equation and associated boundary conditions are written and the turbulent model is presented. The inverse problem is then formulated. Configurations under consideration are detailed and sensors sensitivities are studied. Finally, results for several inversion cases are presented.

SYSTEM DESCRIPTION

Physical Problem

Let us consider turbulent forced convection, with a hydrodynamically fully developed, thermally developing, incompressible, constant property, turbulent flow inside a parallel-plate smooth duct. Axial conduction in the fluid and viscous dissipation are neglected. Conduction along the flow in the wall material is disregarded. The assumption of uniform wall temperature across the whole thickness of the plates is made. Both plates are subjected to the same heat flux density. The channel axis being then a symmetry axis, only half of the duct is considered. A schematic view of the system is presented hereafter in Fig.1.

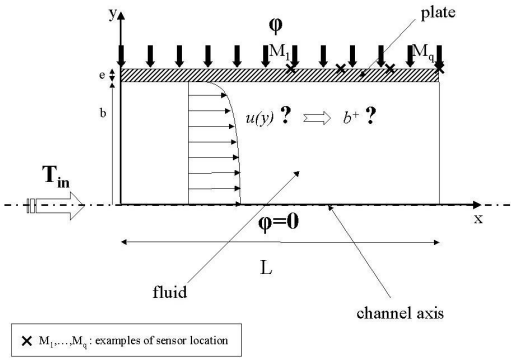


Figure 1: System description

Governing Equations

For such a problem, the characteristic heat flux density is the applied heat flux density φ. A characteristic temperature difference can then be defined as:

$$\Delta T_{ref} = \frac{\varphi b}{\lambda_f} \tag{1}$$

So, by choosing a dimensionless applied heat flux as:

$$\Phi = \frac{\varphi b}{\lambda_f \Delta T_{ref}} = 1$$

and introducing the following dimensionless quantities:

$$\theta = \frac{T - T_{in}}{\Delta T_{ref}}, \quad \xi = \frac{4x}{bPe}, \quad \eta = \frac{y}{b}, \quad U = \frac{u}{u_m}, \quad \varepsilon_{eff} = 1 + \frac{\alpha_{turb}}{\alpha},$$

$$Re = \frac{4bu_m}{\nu}, \quad Pr = \frac{\nu}{\alpha} \quad \text{and} \quad Pe = Re\,Pr,$$

the dimensionless governing steady state energy equation is written as [6]:

$$U(\eta)\frac{\partial\theta}{\partial\xi}(\xi,\eta) = \frac{\partial}{\partial\eta}\left(\varepsilon_{eff}(\eta)\frac{\partial\theta}{\partial\eta}(\xi,\eta)\right) \tag{2a}$$

with following boundary conditions, considering the channel axis symmetry:

$$\begin{cases} \dfrac{\partial\theta}{\partial\eta}(\xi,1) = 1 & \forall\xi \in [0,L^*] \\ \dfrac{\partial\theta}{\partial\eta}(\xi,0) = 0 & \forall\xi \in [0,L^*] \\ \theta(0,\eta) = 0 & \forall\eta \in [0,1] \end{cases} \tag{2b,c,d}$$

where $L^* = \dfrac{4L}{bPe}$ is the dimensionless length of the channel.

The direct problem consists in the resolution of Eqs.(2). A numerical solution approximation is obtained with the Finite Volume Method. Equation (2a) being linear and parabolic in ξ, the solution is obtained by using a spatially marching procedure in the ξ direction. An appropriate typical mesh uses 21 nodes in the ξ direction (regular mesh) and 30 in the η direction (non regular mesh, refined near wall).

Turbulent Model

The dimensionless fully developed turbulent velocity distribution $U(\eta)$ and the dimensionless effective diffusivity $\varepsilon_{eff}(\eta)$ in Eq.(2a) are determined by a pseudo-laminar turbulent model given by Kim and Özisik [7] and presented below, except for the turbulent Prandtl number: $Pr_{turb} = \dfrac{\nu_{turb}}{\alpha_{turb}}$, which has been taken by Jischa and Rieke [8] as:

$$Pr_{turb} = 0.85 + \frac{0.015}{Pr} \tag{3}$$

In a classical way, the turbulent model is expressed with new dimensionless variables.

$u_\tau = \sqrt{\dfrac{\tau_w}{\rho}}$ is called the shear velocity, where $\tau_w = -b\dfrac{dP_m}{dx}$ is the wall shear stress depending on the gradient of pressure $\dfrac{dP_m}{dx}$ which has generated the flow, and the fluid density ρ. A Reynolds number based on the shear stress velocity may be written as:

$$b^+ = \frac{bu_\tau}{\nu} \tag{4}$$

Defining a new dimensionless transverse coordinate $y^+ = (1-\eta)b^+$, b^+ may also be seen as the dimensionless transverse coordinate at the channel axis $(\eta = 0)$. A new dimensionless velocity $u^+ = \dfrac{u}{u_\tau}$ is introduced.

The *turbulent velocity* distribution is taken as the following three-layer model [7]:

- Viscous sublayer:

$$u^+ = y^+ \quad \text{for } 0 \le y^+ < 5 \tag{5a}$$

- Buffer layer:
  $u^+ = -3.05 + 5 \ln y^+ \quad for\ 5 \le y^+ < 30$ (5b)
- Turbulent core:
  $u^+ = 5.5 + 2.5 ln\left[y^+ \left(\frac{1.5(1+\eta)}{1+2\eta^2} \right) \right] \quad for\ y^+ \ge 30$ (5c)

The *turbulent viscosity* distribution is taken as the two-layer model [7]:

- $\frac{\nu_{turb}}{\nu} = \frac{k_1}{E}\left[e^{k_1 u^+} - 1 - k_1 u^+ - \frac{(k_1 u^+)^2}{2!} - \frac{(k_1 u^+)^3}{3!} \right]$ (5d)
  for $y^+ < 40$, where $k_1 = 0.407$ and $E = 10$
- $\frac{\nu_{turb}}{\nu} = \frac{k_2 b^+}{6}(1-\eta^2)(1+2\eta^2) \quad for\ y^+ \ge 40$ (5e)
  where $k_2 = 0.4$

The turbulent velocity and viscosity distributions are therefore characterised by the parameter b^+.

The Reynolds number is given by:

$$Re = 4b^+ \int_{\eta=0}^{\eta=+1} u^+(\eta) d\eta \tag{6}$$

$U(\eta)$ and $\varepsilon_{eff}(\eta)$ are given by:

$$U(\eta) = \frac{u(y)}{u_m} = \frac{u^+(\eta)}{\int_{\eta=0}^{\eta=+1} u^+(\eta) d\eta} \tag{7}$$

$$\text{and } \varepsilon_{eff}(\eta) = 1 + \frac{\alpha_{turb}(\eta)}{\alpha} = 1 + \frac{\nu_{turb}(\eta)}{\nu}\frac{Pr}{Pr_{turb}} \tag{8}$$

INVERSE PROBLEM

Formulation

Suppose that b^+ is unknown and inlet temperature is known. In order to identify b^+, a known heat flux density is applied on both upper and lower plates along the channel length and temperature responses are taken at the wall external surface. These measurements are dependent on the way that heat is transported in the flow. Thus they depend on b^+. The inverse procedure consists in the determination of b^+ from these simulated temperature data.

Note: the Reynolds number *Re* based on mean velocity could have been chosen as the dimensionless parameter to be identified in the inverse procedure, but the choice of b^+ has been made for the two following reasons:

- As $b^+ = \frac{b}{\nu}\sqrt{\frac{-b}{\rho}\frac{dP_m}{dx}}$ (c.f. Eq.(4)), the gradient of pressure which has generated the flow appears formally in the expression of b^+, allowing a direct access to this quantity (or the corresponding wall shear stress τ_w).
- If *Re* is the unknown parameter, then for each iteration of the inverse algorithm, *Re* has its current value and b^+ needs to be calculated to obtain u^+. This involves, for each iteration, the resolution of Eq.(6) which is transcendental in that case. A larger computing time for inversion is then needed.

The inverse problem consisting in the identification of b^+ may be formulated as an optimisation problem: find b^+ minimising the squared residues functional $f(b^+)$ defined as:

$$f(b^+) = \sum_{i=1}^{Nmes} \left(\theta_{surf\,i}^{calc}(b^+) - \theta_{surf\,i}^{mes} \right)^2 \tag{9}$$

where the $\theta_{surf\,i}^{mes}, i \in \langle 1, Nmes \rangle$ are the *Nmes* dimensionless temperature measurements taken at the wall external surface.

According to previous assumptions about the wall material, the $\theta_{surf\,i}^{calc}, i \in \langle 1, Nmes \rangle$ are obtained from the dimensionless temperatures $\theta_{w\,i}^{calc}, i \in \langle 1, Nmes \rangle$ at the fluid/solid interface coming from the resolution of Eqs.(2), by a simple heat flux balance:

$$\theta_{surf\,i}^{calc} = \theta_{w\,i}^{calc} + \frac{1}{\lambda^*} \tag{10}$$

where $\lambda^* = \frac{\lambda_s b}{\lambda_f e}$.

Magnitude Of Applied Heat Flux Density

The applied heat flux density φ has to be chosen as a satisfying compromise between two opposite considerations: a high φ to obtain high sensitivities of the sensors with respect to b^+, and a low φ to induce the slightest possible perturbation of the flow. In fact, when heating the channel, the flow may be modified by two ways: generation of natural convection and variation of the fluid properties. To prevent such events, the mean increase of temperature induced by φ should not be greater than a few Kelvin. This will be verified *a posteriori*.

OPTIMISATION ALGORITHM

A Quasi-Newton method [9] is used to minimise $f(b^+)$ defined by Eq.(9). This method is briefly presented in the following paragraphs.

The basic Newton method consists in an approximation of the objective function f around the current iterate b_k^+ by a quadratic function, that is a second order Taylor's expansion. The next iterate will be $b_{k+1}^+ = b_k^+ + s_k$ where the search direction s_k is the minimiser of the quadratic function and is obtained by writing the first order condition, leading to:

$$s_k = -\frac{f'(b_k^+)}{f''(b_k^+)} \tag{11}$$

The second order condition requires the positiveness of $f''(b_k^+)$, which is not always assured, depending on the initial guess. The convergence is guaranteed only if the starting point is sufficiently close to a local minimiser b^{+*} at which f'' is positive. A Quasi-Newton method is used to ensure the convergence when the initial guess is not close to a minimiser. It uses a line-search method: the next iterate is obtained by $b_{k+1}^+ = b_k^+ + \beta_k s_k$ where β_k is chosen so that $f(b_{k+1}^+) < f(b_k^+)$ and to satisfy the curvature condition, and the search direction s_k is not obtained by Eq.(11) but by:

$$s_k = -\frac{f'(b_k^+)}{H_k} \tag{12}$$

where H_k is an updated approximation to $f''(b_k^+)$.

The gradient $f'(b_k^+)$ is given by:

$$f'(b_k^+) = 2\sum_{i=1}^{Nmes}\left(\left[\frac{d\theta_{surf\,i}^{calc}}{db^+}(b_k^+)\right]\left(\theta_{surf\,i}^{calc}(b_k^+) - \theta_{surf\,i}^{mes}\right)\right) \tag{13}$$

Its calculation requires the computation of the sensitivities $\frac{d\theta_{surf\,i}^{calc}}{db^+}(b_k^+)$, $i \in \langle 1, Nmes \rangle$. These are approximated with Forward Finite Differences, yielding, for $i \in \langle 1, Nmes \rangle$, to:

$$\frac{d\theta_{surf\,i}^{calc}}{db^+}(b_k^+) = \frac{\theta_{surf\,i}^{calc}(b_k^+ + \gamma b_k^+) - \theta_{surf\,i}^{calc}(b_k^+)}{\gamma b_k^+} \tag{14}$$

where γ is a small perturbation factor.

The numerical evaluation of the gradient $f'(b_k^+)$ then requires a second resolution of the direct problem (Eqs.(2)) in addition to the one needed to compute $f(b_k^+)$. The computation of $f''(b_k^+)$ would then require to solve three direct problems. To avoid such a time-consuming operation, $f''(b_k^+)$ is approximated by H_k, using quantities $b_k^+ - b_{k-1}^+$ and $f'(b_k^+) - f'(b_{k-1}^+)$. Formulas for H_k and β_k can be found in [9].

The *general algorithm* may be summarised as:

1. Using the current iterate b_k^+ ,.compute u^+ with Eqs.(5a,b,c) and $\frac{\nu_{turb}}{\nu}$ with Eqs.(5d,e)
2. Compute *Re* with Eq.(6), *U* with Eq.(7) and ε_{eff} with Eq.(8) and Eq.(3).
3. Solve Eqs.(2a,b,c,d) to obtain θ_w^{calc}. Compute θ_{surf}^{calc} using Eq.(10) and $f(b_k^+)$ with Eq.(9).
4. Repeat steps 1 to 3 with $b_k^+ + \gamma b_k^+$ (γ small) and compute sensitivities using Eq.(14).
5. Compute $f'(b_k^+)$ using Eq.(13).
6. Compute H_k, then s_k with Eq.(12) and β_k.
7. Set the next iterate to: $b_{k+1}^+ = b_k^+ + \beta_k s_k$.
8. Go to step 1 and set $k=k+1$.

This iterative procedure is performed until the chosen stopping criterion (c.f. Eq.(15), Eq.(17) and following remark) is satisfied.

For a rectangular duct section, the laminar-turbulent transition corresponds to $Re \cong 2400$ (i.e. $b^+ \cong 50$). Therefore it seems logical to take 50 as lower bound for b^+. This value will also be taken as initial guess for the minimisation algorithm, according to sensitivity analysis.

THE STUDIED CONFIGURATIONS

The following case is considered: water at T_{in}=300K is entering through the region of the channel subjected to the applied heat flux density. Fluid properties are then taken as: $\nu = 8.58\ 10^{-7}$ $m^2.s^{-1}$, $\alpha = 1.47\ 10^{-7}$ $m^2.s^{-1}$ and $\lambda_f = 0.61$ $W.m^{-1}.K^{-1}$. Solid thermal conductivity is taken as $\lambda_s = 14.9$ $W.m^{-1}.K^{-1}$. The half channel height is $b = 5\ 10^{-2}$ m and the domain length used to heat the wall channel and to simulate the temperature experimental data is $L = De = 4b = 0.2$ m where $De = 4b$ is the equivalent diameter. The wall thickness is $e = 2\ 10^{-3}$ m. Two values of b^+ are considered: $b^+ = 1166$ and 9361, corresponding respectively to $Re \cong 10^5$ and 10^6. Corresponding test cases are called respectively case 1 and 2.

The applied heat flux density φ is chosen to be φ = 1750 W.m^{-2} for the case b^+=1166 and φ = 13000 W.m^{-2} for the case b^+=9361. According to Eq.(1), values of ΔT_{ref} are then 142.74 K and 1060.36 K for cases 1 and 2 respectively. Fig.2 shows dimensional temperatures at solid/fluid interface and wall external surface (21 positions in the axial direction) for case 1.

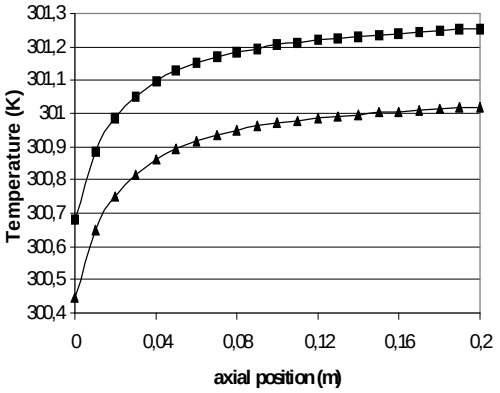


Figure 2: T_w^{calc} (▲) and T_{surf}^{calc} (■) with respect to axial position, for case 1.

It can be verified that dimensional temperature at solid/fluid interface and $x=L$, which is the hottest point in the fluid domain, is about 301K, corresponding to an increase of only 1K from T_{in}. The assumption of an unchanged flow is then expected to be valid. It is also verified for case 2.

SENSITIVITY ANALYSIS

The applied heat flux density φ = 1750 W.m^{-2} is considered. Temperatures are not linear functions of b^+, consequently sensitivities are dependent on b^+. Let us call S the absolute value of the sensitivity given by Eq.(14):

$$S = \left| \frac{d\theta_{surf}^{calc}}{db^+} \right|$$

In Fig.3 are shown the values of S at possible sensors locations for three values of b^+ (500, 1000 and 5000). Results were intuitively expected. It is verified that for a given b^+, S increases with the axial location x along the heated channel. If only one sensor is used, the best possible location is then at $x=L$. For a given location, the larger b^+, the lower S is. Hence, it seems useful to start with b^+=50 as the initial guess to ensure the highest possible sensitivities during the first iterations.

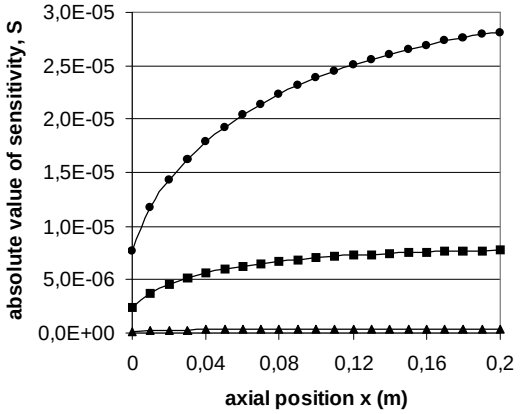


Figure 3: Values of S for the 21 nodes, for b^+=500 (●), 1000 (■) and 5000 (▲).

INVERSION RESULTS AND ANALYSIS

Validation With $\theta_{surf}^{mes} = \theta_{surf}^{calc}$

Inversion using "exact" temperatures gives excellent results whatever the number and position of sensors. Results obtained with a single sensor located in $x=L$ (node 21) are given for cases 1 and 2 (c.f. Table 1). One should note that the stopping criterion for the minimisation algorithm is:

$$f(b_{k+1}^+) < \text{a small number such as } 10^{-10} \quad (15)$$

Simulation Of Real Measurements

Simulated temperatures θ_{surf}^{mes} are generated by adding random errors to temperatures θ_{surf}^{calc} computed with Eqs.(2) and Eq.(10) for applied heat flux density φ. That is:

$$\theta_{surf}^{mes} = \theta_{surf}^{calc} + \omega\frac{\sigma}{\Delta T_{ref}} = \theta_{surf}^{calc} + \frac{\Delta T}{\Delta T_{ref}} \quad (16)$$

where σ is the dimensional standard deviation of measurement errors which is assumed to be the same for all measurements, ω is a normally distributed random number, meaning there is a 99% probability of the value of ω lying in the range $-2.576<\omega<+2.576$ and ΔT is the measurement error. Thus:
-10^{-1} K$<\Delta T<+10^{-1}$ K corresponds to $\sigma=3.88\ 10^{-2}$ K and $-2.576\ 10^{-1}$ K$<\Delta T<+2.576\ 10^{-1}$ K corresponds to $\sigma=10^{-1}$ K.

When a single sensor is used, $\Delta T=+\sigma$ is taken.

The iterative regularisation principle [10] is used to obtain the following stopping criterion:

$$f(b_{k+1}^{+}) \approx Nmes\left[\frac{\sigma}{\Delta T_{ref}}\right]^2 \qquad (17)$$

If this criterion is not satisfied in the iterative process, the algorithm is stopped after a chosen number of iterations without decreasing $f(b^+)$.

Considering the analysis of sensitivities to b^+ and the mesh sensitivity, nodes 1 to 5 have not been taken as sensors. In Fig.4 are shown both simulated exact and "measured" dimensional temperatures for case 1 with σ=3.88 10^{-2} K.

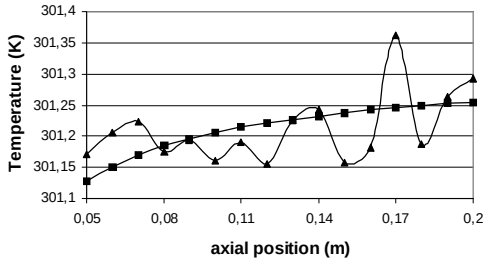


Figure 4: T_{surf}^{calc} (■) and T_{surf}^{mes} (▲) at nodes 6 to 21 (case 1 with σ=3.88 10^{-2} K).

Effect Of Sensor Number And Measurement Errors

Let us consider a single sensor measurement. The node 21 is chosen and a simulated error ΔT=+σ=3.88 10^{-2} K is added. One can note (Table 1) that the solution b^+ is underestimated (relative error ≈ 5 to 6%) for both cases 1 and 2. More sensors need to be used in the presence of measurement error. The inverse problem is then overdetermined since it has more data than unknowns. Table 1 shows results for case 1 and σ=3.88 10^{-2} K when 4, 6 and 16 sensors are used. With 4 sensors, the results quality is not clearly improved, but with 6 sensors, results become more satisfying, and with 16 sensors the estimation is accurate. For this last case, Figure 5 shows the convergence of $f(b^+)$ and b^+ with respect to iteration number.

When noise level is increased to σ=10^{-1} K, the estimation using nodes 6 to 21 as sensors is slightly deteriorated but remains very good (relative error with the exact value < 1%).

Case 2 is now considered. As b^+ is increased in comparison with case 1, sensitivities should be lower but this is compensated by a higher value for the applied heat flux φ. Moreover, wall external temperature data are higher and therefore less sensitive to additive noise. As a consequence, estimations are as good as those of case 1 (c.f. Table 1).

Table 1: Identification results for different test cases. Estimated values of b^+ and relative errors.

	CASE 1 *Exact value* b^+=1166 (*Re*≅10^5)	***CASE 2*** *Exact value* b^+=9361 (*Re*≅10^6)
Node 21 σ=0 K	**1166.107** *(0.009 %)*	**9361.001** *(0.000014%)*
Node 21 σ=3.88 10^{-2} K	**1097.682** *(5.859 %)*	**8889.011** *(5.042 %)*
4 Nodes (18 to 21) σ=3.88 10^{-2} K	**1098.981** *(5.748 %)*	**8989.315** *(3.971%)*
4 Nodes 6,11,16,21 σ=3.88 10^{-2} K	**1103.986** *(5.318 %)*	**9011.175** *(3.737%)*
6 Nodes (16 to 21) σ=3.88 10^{-2} K	**1145.635** *(1.747 %)*	**9172.593** *(2.013%)*
6 Nodes 6,9,12,15,18,21 σ=3.88 10^{-2} K	**1146.509** *(1.672 %)*	**9190.534** *(1.821%)*
16 Nodes (6 to 21) σ=3.88 10^{-2} K	**1169.558** *(0.305 %)*	**9378.425** *(0.186 %)*
16 Nodes (6 to 21) σ=10^{-1} K	**1174.952** *(0.768 %)*	**9406.627** *(0.487 %)*

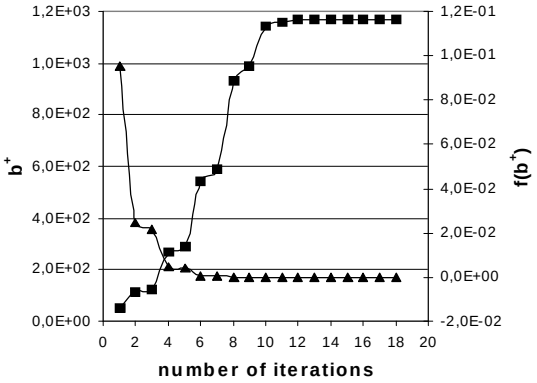


Figure 5: Convergence of $f(b^+)$ (▲) and b^+ (■) with respect to iteration number (case 1, sensors at nodes 6 to 21, σ=3.88 10^{-2} K)

Effect Of Sensor Location

As it can be seen in Table 1, for case 1 and a given number of sensors (4 or 6), the estimation with sensors close to $x=L$ is worse than the one obtained with sensors equally positioned between nodes 6 and 21. This seems strange because the closer the sensors are to L, the higher are their sensitivities. Nevertheless, this may be explained by the fact that data close to L present redundant information, especially in presence of noise.

Reconstruction Of Turbulent Quantities

Once b^+ has been identified, estimated profiles of velocity u and turbulent diffusivity α_{turb} are obtained using Eqs. (5a-e), (6), (7), (8) and (3). Fig.(6) shows reconstructed profiles for case 1 and two configurations of inversion: a single sensor (node 21) and 16 sensors (nodes 6 to 21), with $\sigma=3.88\ 10^{-2}$ K.

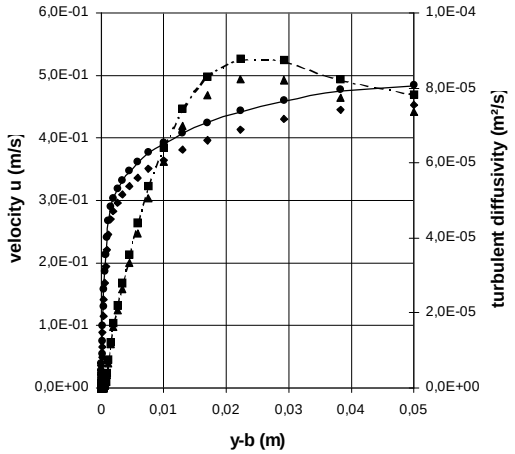


Figure 6: Reconstructed velocity and turbulent diffusivity profiles for case 1 with $\sigma=3.88\ 10^{-2}$ K. "Exact" values: u (——), α_{turb} (-·-·-). Estimations with node 21: u (♦), α_{turb} (▲). Estimations with nodes 6 to 21: u (●), α_{turb} (■).

CONCLUSIONS

An inverse problem of turbulent forced convection inside a parallel-plate duct has been investigated in this numerical study. In order to estimate a Reynolds number b^+ based on the shear stress velocity and characterising turbulent velocity and effective diffusivity distributions, a known heat flux density is applied on the upper and lower plates of the duct and temperature responses dependent on b^+, are taken at the wall external surface. A squared residues functional built on these data and a heat transfer model has been minimised to estimate the unknown. Results have shown that such a procedure allows to accurately estimate b^+ even if the measurements are altered with errors, provided that the number of sensors may be sufficiently large. Future works include cases in which no parametric laws describe the unknown distributions. Efforts will be done on hydrodynamically developing flows.

REFERENCES

1. F.B. Liu and M.N. Özisik, Inverse analysis of transient turbulent forced convection inside parallel-plate ducts, *International Journal of Heat and Mass Transfer*, **39**, 12 (1996).

2. M. Girault, D. Petit and E. Videcoq, Inverse problem of turbulent forced convection inside parallel-plate ducts using a reduced model, *12th International Heat Transfer Conference*, Grenoble, France, August 18-23, 2002 (full paper accepted).

3. J. Su, A.B. Lopes and A.J. Silva Neto, Estimation of unknown wall heat flux in turbulent circular pipe flow, *International Communications in Heat and Mass Transfer*, **27**, 7 (2000).

4. Z. Mao and T.J. Hanratty, Application of an inverse mass transfer method to the measurement of turbulent fluctuations in the velocity gradient at the wall, *Experiments in Fluids*, **11** (1991).

5. P.O. Zavialov, R.D. Ghisolfi and C.A.E. Garcia, An inverse model for seasonal circulation over the southern Brazilian shelf: near-surface velocity from the heat budget, *Journal of Physical Oceanography*, **28** (1998).

6. W.M. Kays and M.E. Crawford, *Convective Heat and Mass Transfer*, Third Edition, McGraw-Hill, Singapore, 1993.

7. W.S. Kim and M.N. Özisik, Turbulent forced convection inside a parallel-plate channel with periodic variation of inlet temperature, *Journal of Heat Transfer*, Trans ASME, **111** (1989).

8. M. Jischa and H.B. Rieke, About the prediction of turbulent Prandtl and Schmidt numbers from modeled transport equations, *International Journal of Heat and Mass Transfer*, **22** (1979).

9. P.E. Gill, W. Murray and M.H. Wright, *Practical Optimization*, Academic Press, New York, 1981.

10. O.M. Alifanov, *Inverse Heat Transfer Problems*, Springer-Verlag, New York, 1994.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

EXTENSION OF THE HOT WIRE METHOD TO THE CHARACTERISATION OF STRATIFIED SOILS WITH MULTIPLE TEMPERATURE ANALYSIS

Coment E.
Center for Energy and Environmental Processes,
Ecole des Mines d'Albi-Carmaux,
Albi, France
coment@enstimac.fr

Fudym O., Santander R.
Department of Mechanical Engineering,
Federal University of Santiago de Chile,
Santiago de Chile, Chile
ofudym@lauca.usach.cl

Batsale J.C.
LEPT,
Ecole Nationale des Arts et Métiers,
Talence, France
batsale@lept-ensam.u-bordeaux.fr

Ladevie B.
Center for Energy and Environmental Processes,
Ecole des Mines d'Albi-Carmaux,
Albi, France
ladevie@enstimac.fr

ABSTRACT :
The aim of this work is to develop a practical device able to estimate a thermal conductivity distribution in stratified media such as burned soils in Chile. The classical hot wire method consists in measuring the temperature response of a heat step imposed on a thin cylindrical probe by Joule effect. The main characteristic of the extension of the probe consists here in analysing the temperature response of multiple thermocouples equally spaced along the heating cylinder. A semi-analytical method (quadrupole method) is then developed in order to obtain a transfer matrix between the heat flux excitation and the temperature responses vector. Such method is suitable to obtain asymptotic expansions in order to investigate the sensitivity analysis and the estimation strategy. Some experiments and estimation results will be presented in a case where the characteristic diffusion times following the radial direction are small compared to the inter-strata diffusion time.

NOMENCLATURE

a	Thermal diffusivity
A, B, C, D	Generalized quadripole matrices
e	Stratum thickness
I_0, I_1	Modified Bessel functions
K_0, K_1	Modified Bessel functions
k	Thermal conductivity
l	Length of probe
$\mathbf{P}$, $\mathbf{P}^{-1}$	Eigenvectors matrix
r_c	Radius of central cylinder
r_r	Global radius of probe
R	Contact resistance
$\mathbf{R}$	Contact resistance matrix
s	Laplace variable
t	Time
T	Temperature
ρC	Specific heat
$\overline{T}$	Laplace temperature
$\overline{\mathbf{T}}$	Laplace temperature vector
φ	heat flux
$\overline{\varphi}$	Laplace heat flux
$\overline{\mathbf{\Phi}}$	Laplace heat flux vector
$\mathbf{\Omega}$	Diagonal eigenvalues matrix

INTRODUCTION

Numerous works have been devoted to the thermal characterization of global properties of soils [1-5]. In the case of stratified media a local characterization versus the depth can avoid a tedious sampling of the different variety of soils.

The method proposed here is an extension of the classical hot wire method (see Carslow [6]). The main characteristic of the extension of the probe consists here in analysing the temperature response of multiple thermocouples equally spaced along a heating cylinder.

The resolution of the direct heat transfer problem is often difficult in such case because the heterogeneities are made of radial and stratified layers. A semi-analytical extension of the quadrupole method [7] is here proposed in order to obtain a transfer matrix between the heat flux excitation and the temperature responses vector. Such method is convenient to obtain suitable asymptotic expansions in order to investigate the sensitivity analysis and the estimation strategy.

One experiment and estimation results will be presented in a case where the characteristic diffusion times following the radial direction are small compared to the inter-strata diffusion time.

POSITION OF THE PROBLEM

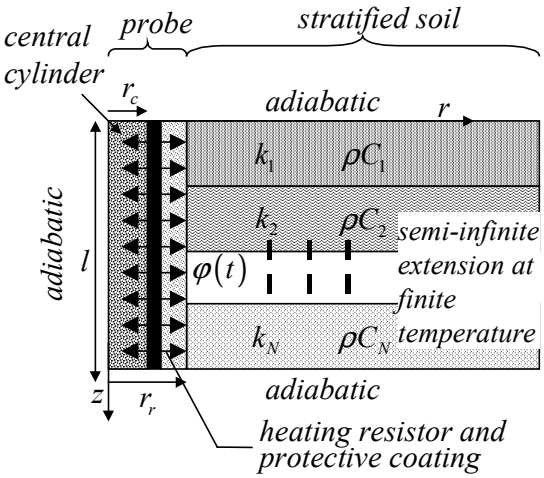


Figure 1 : Scheme of the heat transfer system and boundary conditions.

The main characteristics of the heat transfer problem are presented on figure 1. A central heating cylinder is installed vertically in the stratified medium. Only the response of several thermocouples equally spaced on the heating layer are recorded and analysed after a heating step.

The present study deals with the problem of finding a generalized intrinsic relationship between temperature and heat flux at the boundaries of an heterogeneous medium with one-dimensional varying properties versus layer direction. The extension of thermal quadrupole formalism is obtained from one-dimensional discretization versus *z*-direction, within a control volume formulation, coupled with the semi-analytical solution of the corresponding vectorial differential equation.

In a medium with one-dimensional varying thermal properties, such as the one depicted in figure 1, without volumetric source term, two-dimensional transient heat conduction is governed by the following equation :

$$\frac{k(z)}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right)+\frac{\partial}{\partial z}\left(k(z)\frac{\partial T}{\partial z}\right)=(\rho C)(z)\frac{\partial T}{\partial t} \quad (1)$$

Space discretization of equation. (1) is performed versus *z* direction. A number *N* of new variables is introduced as:

$$T_i(r,t)=\frac{1}{\Delta z}\int_{i^-}^{i^+} T(r,z,t).dz \quad (2a)$$

where i^- and i^+ indicate the i^{th} grid interfaces due to space discretization. Equation (1) is then integrated relative to *z:*

$$\frac{k_i\Delta z}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T_i}{\partial r}\right)+\varphi_{i^-}-\varphi_{i^+}=(\rho C)_i\,\Delta z\frac{\partial T_i}{\partial t} \quad (2b)$$

where φ is heat flux density in the *z* direction. The heat flux φ is linearized, and the interface thermal conductivity k_{i-} and k_{i+} are introduced in the conservative form:

$$k_{i-}=2\left(\frac{1}{k_{i-1}}+\frac{1}{k_i}\right)^{-1} \text{ and } k_{i+}=2\left(\frac{1}{k_i}+\frac{1}{k_{i+1}}\right)^{-1} \quad (2c)$$

Performing a Laplace transformation for *i = 2 to N-1,* such as:

$$\overline{T}_i(r,s)=\int_0^{\infty}\exp(-st)T_i(r,t).dt \quad (3a)$$

and substituting this expression into equation (2b) yields:

$$-k_{i^-}\overline{T}_{i-1}+(k_{i^-}+k_{i^+}+(\rho C)_i\Delta z^2 s)\overline{T}_i$$
$$-k_{i^+}\overline{T}_{i+1}-k_i\Delta z^2\frac{1}{r}\frac{d}{dr}\left(r\frac{d\overline{T}_i}{dr}\right)=0 \qquad (3b)$$

Both boundary conditions at $z = 0$ and $z = e$ have also to be expressed in the same way in order to obtain the corresponding equations at node $i = 1$ and $i = N$.

Introducing the vector $\overline{\mathbf{T}}$ of the N Laplace transformed temperatures at the position r, equation (3c) can be written in the matrix form:

$$\mathbf{M}_s\left(\mathbf{M}_{//}+\mathbf{G}s\right)\overline{\mathbf{T}}-\frac{1}{r}\frac{d}{dr}\left(r\frac{d\overline{\mathbf{T}}}{dr}\right)=\mathbf{0} \qquad (4)$$

with

$$\mathbf{M}_{//}=\begin{bmatrix} c_1+k_{1+} & -k_{1+} & 0 & & & & \\ -k_{2-} & k_{2-}+k_{2+} & -k_{2+} & & & & \\ 0 & \cdots & \cdots & \cdots & & & \\ & & -k_{i-} & k_{i-}+k_{i+} & -k_{i+} & 0 & \\ & & & \cdots & \cdots & \cdots & \\ & & & 0 & -k_{N-} & c_N+k_{N-} & \end{bmatrix}$$

$$\mathbf{G}=\Delta z^2\,\mathbf{diag}\left(\left[(\rho C)_1 \quad (\rho C)_2 \quad ..... \quad (\rho C)_N\right]\right)$$

$$\mathbf{M}_s=\frac{1}{\Delta z^2}\mathbf{diag}\left(\left[\frac{1}{k_1} \quad \frac{1}{k_2} \quad ..... \quad \frac{1}{k_N}\right]\right)$$

and $\overline{\mathbf{T}}=\left[\overline{T}_1 \quad \overline{T}_2 \quad ... \quad \overline{T}_N\right]^t$

The operator « **diag** » is used in order to build a diagonal matrix from the corresponding vector. The constant c_1 and c_N are related to the homogeneous boundary conditions at $z=0$ and $z=l$.

The square matrix $\mathbf{M}_s(\mathbf{M}_{//}+\mathbf{G}s)$ on the left side of equation (4) is independent of r. Equation (4) can be solved directly by the diagonalization of this matrix. The diagonalization of this system yields :

$$\mathbf{M}_s(\mathbf{M}_{//}+\mathbf{G}s)=\mathbf{P}\boldsymbol{\Omega}\mathbf{P}^{-1} \qquad (5a)$$

where $\boldsymbol{\Omega}$ is a diagonal matrix. Introducing the temperature vector in the new base :

$$\mathbf{V}=\mathbf{P}^{-1}\overline{\mathbf{T}} \qquad (5b)$$

Equation (4) is then written as

$$\boldsymbol{\Omega}\mathbf{V}-\frac{1}{r}\frac{d}{dr}\left(r\frac{d\mathbf{V}}{dr}\right)=\mathbf{0} \qquad (5c)$$

Each line of equation. (5c) can directly be solved in a scalar way, due to the fact that $\boldsymbol{\Omega}$ is a diagonal matrix, and the corresponding equation is homogeneous. The general results can be arranged in a matrix form, applying the quadrupole formalism results. Care must be taken in order to avoid the commutation of the matricial products:

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dv_k}{dr}\right)-d_k v_k=0 \qquad (6a)$$

The general solution of equation (6a) yields:

$$v_k=G_1 I_0(\sqrt{d_k}r)+G_2 K_0(\sqrt{d_k}r) \qquad (6b)$$

where I_0 and K_0 are the modified Bessel functions of order zero of the first and second kind, respectively. Coefficients d_k are function of the Laplace variable s, but is independent of r.

The scalar thermal quadrupole formalism is derived from the solution of equation (6a), by eliminating the integration constant G_1 and G_2 in equation (6b) in order to find a linear relationship between temperature and heat flux in the Laplace space.

Introducing the flux expressed in the eigenvalues basis as $j_k=-(2\pi r)\Delta z\frac{dv_k}{dr}$ or under vector notation :

$$\mathbf{J}_{\mathbf{V}}=-(2\pi r)\Delta z\frac{d\mathbf{V}}{dr} \qquad (6c)$$

The scalar thermal quadrupole at r_1 and r_2 location is written under the following form:

$$\begin{bmatrix} v_k \\ j_k \end{bmatrix}_{r_1}=\begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}\begin{bmatrix} v_k \\ j_k \end{bmatrix}_{r_2}$$ or under vector notation: $$\begin{bmatrix} \mathbf{V} \\ \mathbf{J}_{\mathbf{V}} \end{bmatrix}_{r_1}=\begin{bmatrix} \mathbf{A}_{\mathbf{V}} & \mathbf{B}_{\mathbf{V}} \\ \mathbf{C}_{\mathbf{V}} & \mathbf{D}_{\mathbf{V}} \end{bmatrix}\begin{bmatrix} \mathbf{V} \\ \mathbf{J}_{\mathbf{V}} \end{bmatrix}_{r_2} \qquad (6d)$$

where :

$$\mathbf{A}_V = r_2\boldsymbol{\Omega}^{1/2}\begin{bmatrix} I_0\left(\boldsymbol{\Omega}^{1/2}r_1\right)K_1\left(\boldsymbol{\Omega}^{1/2}r_2\right) \\ +I_1\left(\boldsymbol{\Omega}^{1/2}r_2\right)K_0\left(\boldsymbol{\Omega}^{1/2}r_1\right)\end{bmatrix},$$

$$\mathbf{B}_V = \frac{1}{2\pi\Delta z}\begin{bmatrix} I_0\left(\boldsymbol{\Omega}^{1/2}r_2\right)K_0\left(\boldsymbol{\Omega}^{1/2}r_1\right) \\ -I_0\left(\boldsymbol{\Omega}^{1/2}r_1\right)K_0\left(\boldsymbol{\Omega}^{1/2}r_2\right)\end{bmatrix},$$

$$\mathbf{C}_V = 2\pi r_1 r_2\boldsymbol{\Omega}\Delta z\begin{bmatrix} I_1\left(\boldsymbol{\Omega}^{1/2}r_2\right)K_1\left(\boldsymbol{\Omega}^{1/2}r_1\right) \\ -I_1\left(\boldsymbol{\Omega}^{1/2}r_1\right)K_1\left(\boldsymbol{\Omega}^{1/2}r_2\right)\end{bmatrix}$$

and

$$\mathbf{D}_V = r_1\boldsymbol{\Omega}^{1/2}\begin{bmatrix} I_0\left(\boldsymbol{\Omega}^{1/2}r_2\right)K_1\left(\boldsymbol{\Omega}^{1/2}r_1\right) \\ +I_1\left(\boldsymbol{\Omega}^{1/2}r_1\right)K_0\left(\boldsymbol{\Omega}^{1/2}r_2\right)\end{bmatrix}$$

$\mathbf{A_V}$, $\mathbf{B_V}$, $\mathbf{C_V}$ and $\mathbf{D_V}$ are diagonal matrices. Subscript V refers to the fact that the temperature and heat flux vectors are written in the eigenvalues basis of matrix $\mathbf{M}_s(\mathbf{M}_{//}+\mathbf{G}s)$.

The boundary conditions following the r direction are given as a relationship between temperature and heat flux, but are unknown in the V basis. It is thus convenient to express J_V as a function of heat flux vector $\bar{\boldsymbol{\Phi}}$:

$$\bar{\boldsymbol{\Phi}} = -(2\pi r)\Delta z\mathbf{K}\frac{d\bar{\mathbf{T}}}{dr} \tag{7}$$

where $\mathbf{K}$ is the diagonal matrix of thermal conductivity. Substituting equation (5b) into equation (7) yields

$$\bar{\boldsymbol{\Phi}} = \mathbf{KPJ}_V \tag{8}$$

Using equations (5b) and (8), it is now possible to express the V form-quadrupole given by equation (6d) as a generalized thermal quadrupole in terms of temperature and heat flux vectors in real space, such as :

$$\begin{bmatrix}\mathbf{P}^{-1}\bar{\mathbf{T}} \\ (\mathbf{KP})^{-1}\bar{\boldsymbol{\Phi}}\end{bmatrix}_{r_1} = \begin{bmatrix}\mathbf{A_V} & \mathbf{B_V} \\ \mathbf{C_V} & \mathbf{D_V}\end{bmatrix}\begin{bmatrix}\mathbf{P}^{-1}\bar{\mathbf{T}} \\ (\mathbf{KP})^{-1}\bar{\boldsymbol{\Phi}}\end{bmatrix}_{r_2} \text{ or}$$

$$\begin{bmatrix}\bar{\mathbf{T}} \\ \bar{\boldsymbol{\Phi}}\end{bmatrix}_{r_1} = \begin{bmatrix}\mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}\end{bmatrix}\begin{bmatrix}\bar{\mathbf{T}} \\ \bar{\boldsymbol{\Phi}}\end{bmatrix}_{r_2} \tag{9}$$

with: $\mathbf{A} = \mathbf{PA_V P}^{-1}$; $\mathbf{B} = \mathbf{PB_V(KP)}^{-1}$;
$\mathbf{C} = \mathbf{KPC_V P}^{-1}$; $\mathbf{D} = \mathbf{KPD_V(KP)}^{-1}$

When one of the boundary of the medium is located toward infinite in the r-direction, the corresponding coefficient G_1 in the scalar equation (6b) has to be zero, in order to get a finite solution. Then, it is possible to find a direct relationship between Laplace temperature vector $\bar{\mathbf{T}}_{r_1}$ and Laplace heat flux vector $\bar{\boldsymbol{\Phi}}_{r_1}$ at the position r_1, such as :

$$\begin{aligned}\bar{\mathbf{T}}_1 &= -\frac{1}{2\pi dzr}\mathbf{P}K_0(\boldsymbol{\Omega}^{1/2}r)\boldsymbol{\Omega}^{-1/2}K_1(\boldsymbol{\Omega}^{1/2}r)\mathbf{P}^{-1}\mathbf{K}^{-1}\bar{\boldsymbol{\Phi}}_1 \\ &= \mathbf{Z}_\infty\bar{\boldsymbol{\Phi}}_1\end{aligned} \tag{10}$$

Such expressions are very efficient compared to classical numerical methods because they avoid gridding in the semi-infinite r-direction.

Then the complete system, described on figure 1, can be modelled as successive radial layers related to successive transfer matrices such as:

-In the central cylinder (subscript : c)

$$\begin{bmatrix}\bar{\mathbf{T}} \\ \bar{\boldsymbol{\Phi}}_c\end{bmatrix}_{r_c} = \begin{bmatrix}\mathbf{D}_c & -\mathbf{B}_c \\ -\mathbf{C}_c & \mathbf{A}_c\end{bmatrix}\begin{bmatrix}\bar{\mathbf{T}} \\ \mathbf{0}\end{bmatrix}_0 \tag{11}$$

-In the external cylinder (subscript : e) (made of the resin coating (subscript : r), the contact resistance between the probe and the semi-infinite soil (subscript ∞)

$$\begin{aligned}\begin{bmatrix}\bar{\mathbf{T}} \\ \bar{\boldsymbol{\Phi}}_e\end{bmatrix}_{r_c} &= \begin{bmatrix}\mathbf{A}_r & \mathbf{B}_r \\ \mathbf{C}_r & \mathbf{D}_r\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{R} \\ 0 & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{Z}_\infty\bar{\boldsymbol{\Phi}} \\ \bar{\boldsymbol{\Phi}}\end{bmatrix} \\ &= \begin{bmatrix}\mathbf{A}_r & \mathbf{B}_r \\ \mathbf{C}_r & \mathbf{D}_r\end{bmatrix}\begin{bmatrix}(\mathbf{Z}_\infty+\mathbf{R})\bar{\boldsymbol{\Phi}} \\ \bar{\boldsymbol{\Phi}}\end{bmatrix}\end{aligned} \tag{12}$$

with :

$$\mathbf{Z}_\infty = \frac{1}{2\pi r_r\Delta z}\mathbf{P}K_0\left(\boldsymbol{\Omega}^{1/2}r_r\right)\boldsymbol{\Omega}^{-1/2}K_1\left(\boldsymbol{\Omega}^{1/2}r_r\right)^{-1}\mathbf{P}^{-1}\mathbf{K}^{-1}$$

It is then possible to express the interface temperature vector of the heating resistor versus the heating flux vector with the following expression :

$$\bar{\mathbf{T}} = \left(\mathbf{Z}_c^{-1}+\mathbf{Z}_e^{-1}\right)^{-1}\bar{\boldsymbol{\Phi}} \quad \text{with} \quad \bar{\boldsymbol{\Phi}} = \bar{\boldsymbol{\Phi}}_e - \bar{\boldsymbol{\Phi}}_c \tag{13}$$

where :

$\mathbf{Z}_c = \mathbf{D}_c \mathbf{C}_c^{-1}$ and

$\mathbf{Z}_e = \left(\mathbf{A}_r \left(\mathbf{Z}_\infty + \mathbf{R}\right) + \mathbf{B}_r\right)\left(\mathbf{C}_r \left(\mathbf{Z}_\infty + \mathbf{R}\right) + \mathbf{D}_r\right)^{-1}$

Such expression (13) is very similar to boundary element method, because it allows then to obtain a direct relationship between the boundary temperature and flux vector, instead of gridding a large domain. The numerical implementation of such expression is convenient with matrix solvers such as *Matlab* [8] and inverse Laplace transform such as *Stehfest* [9] (see also [7] for more details). Nevertheless, due to the necessary diagonalisation of the matrices at each time step, an iterative inversion method can be time consuming.

One other advantage of such semi analytical considerations is then to allow asymptotic expansions at short times (or in Laplace space when $(s \to \infty)$). Such « short times » can be considered as shorter than the inter-strata diffusion time (such as $t << e^2/a$, with e inter strata dimensions and a the thermal diffusivity). The matrix $\mathbf{M}_{//}$ is then neglected in expression (4) which significates that the heat transfer becomes 1D following radial direction. The matrices **A,B,C** and **D** from expressions (9), (11) and (12) becomes then diagonal and each term of the temperature vector of expression (13) is then related to each term of the flux vector by a similar scalar expression such as:

$$\bar{T}(s) = \frac{1}{\dfrac{C_c}{D_c} + \dfrac{C_r\left(Z_\infty + R\right) + D_r}{A_r\left(Z_\infty + R\right) + B_r}}\,\bar{\varphi}(s) \qquad (14)$$

with :

$$Z_\infty = \frac{1}{2\pi l k \sqrt{\dfrac{s}{a}} r_r} \frac{K_0\left(\sqrt{\dfrac{s}{a}} r_r\right)}{K_1\left(\sqrt{\dfrac{s}{a}} r_r\right)}$$

Thanks to an asymptotic development of expression (14) at long times (or in Laplace space when $(s \to 0)$), we notice that $C_c/D_c \to 0$, $C_r/D_r \to 0$ and $B_r/A_r \to \ln(r_r/r_c)/2\pi k l = R_r$. Therefore, we obtain a simplified expression of equation (14), only valid in a characteristic time window, as :

$$\bar{T}(p) = \left(Z_\infty + R'\right)\bar{\varphi}(p) \text{ with } R' = R + R_r \qquad (15)$$

To a heat step excitation φ, an analytical relation can be calculated (see Carslaw et Jaeger [6]), and give the classical one dimensional expression related to the hot wire method :

$$T(t) = \frac{\varphi}{4\pi k l} \ln(t) + C^{ste} \qquad (16)$$

Identification of thermal conductivity of each semi -infinite stratum is realized thanks to a simple linear regression on an appropriated time window.

The temperature evolutions calculated from expression (14) have been compared with a good agreement with the complete 2D-transient model given by equation (13) (with *N=50*).

Thanks to the 2D-transient model, the sensitivity of the temperatures $\mathbf{T}$ to the thermal conductivity k_i can be studied (see figure 2 where $i = 6$) and permit us to confirm (for a time $t \le 1000\,\text{s}$ and a $\Delta z \ge 0.02\,\text{m}$) that $\mathbf{T}$ is only sensitive to the $k_{i=6}$. Moreover, the thermal conductivity and the specific heat are correlated. The previous direct 2D model is then used to find the correct time window where the 1D-transient model is valid. Therefore, in this time window, estimation of thermal conductivity can be realized with equation (16).

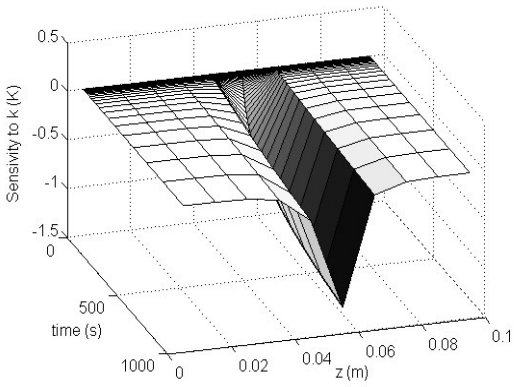


Figure 2 : Sensitivity of temperatures to *k* according to the time

The aim of this work is to identify *k* for each stratum corresponding to the number of temperature measurements. In order to verify and validate this method, we have tested a medium

constituted with only two layers but eight temperature measurements.

EXPERIMENT

Experimental device

The experimental device is made of a wood stick as central cylinder (of radius $r_c = 1.5\,10^{-3}$ m), the heating resistor is realised with a metallic wire which is wound round the stick, with multiple thermocouples of K-type regularly spaced ($\Delta z = 0.02$ m) in contact with the heating wire. Such device is coated with a protective polymer resin. The global length of the system is : $l = 0.24$ m and the global radius : $r_r = 2.1\,10^{-3}$ m.

The temperature responses to a heat step excitation are recorded and analysed with a classical electronic card linked to a personal computer.

The stick can be conveniently embedded in such stratified media as soils.

Test sample

In order to validate the methodology, a test sample is made with a two-layered medium. The first layer is a CMC-gel (CarboxylMéthylCellulosic). The second layer is a powder (semolina).

The thermophysical properties of the media are separately measured and listed in table 1.

Table 1. thermophysical properties of the media

Medium 1 (CMC- gel)	Medium 2 (semolina)
$k = 593$ mW.m^{-1}.K^{-1}	$k = 130$ mW.m^{-1}.K^{-1}
$\rho C = 4\,10^{6}$ J.m^{-3}.K^{-1}	$\rho C = 1.05\,10^{6}$ J.m^{-3}.K^{-1}

Results and discussion

The evolution of the experimental temperature responses is shown on figure 3. The power φ imposed on the heating wire is 2.08 W. The global experimentation time is about 600 s.

It can be observed that the temperatures responses are uniform and corresponding to 1D logarithmic evolution when the temperature sensors are far from the interface between the strata. The 2D transition (at z=0.07m o*r* z=0.09m) occurs only at long times as predicted by the asymptotic approximations.

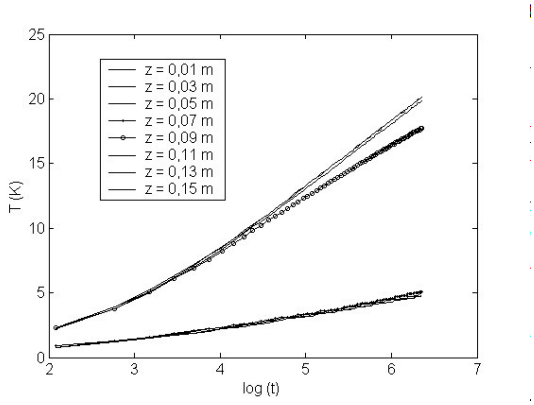


Figure 3 : Experimental temperature history versus natural logarithm of time.

From the previous observation, a simple least-square estimation based on 1D radial transfer assumption (expression (16)) is applied in order to estimate the thermal conductivity profiles. The result of such a very simple estimation process is reported on figure 4.

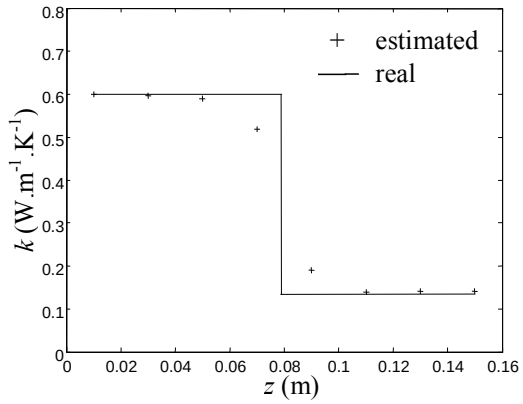


Figure 4 : Thermal conductivity profile (real and estimated).

The good agreement, shown on figure 4, is justifying the application of such simple estimation method in this particular situation. It can be observed that the discrepancy due to 2D effects is visible for estimations located near the interface (z=0.08 m).

CONCLUSIONS

The main interest of this study is to present a semi-analytical direct model. Such expressions are able to be declined in simple asymptotic expansions related to the classical hot wire

method for the characterisation of the thermal conductivity of homogeneous samples.

The complete model allow to validate the conditions of application of the simplified 1D models. Of course, such application is only valid if the dimensions and the time scales related to thermophysical properties are respected ($r_c^2/a_c << e^2/a$).

The perspective of this work are then to investigate the estimation method related to parameters linked to the 2D transfer, such as the interface depth. Such conditions as layers of small thickness compared to the stick radius can also be studied.

ACKNOWLEDGMENTS

This work was partially funded by the FONDECYT program (Award # 1020349) as well as the DICYT Dept. Of the Universidad de Santiago. The authors gratefully acknowledge this support.

REFERENCES

1. D.A. De Vries, A.J. Peck,On the cylindrical probe method of measuring thermal conductivity with special reference to soils, part I : extension of the theory and discussion of probe characteristies, *Aust. J. Phys.*, **11**, 255 (1958)

2. P. Laurent, D. Thomasset, M. Lallemand, Détermination des caractéristiques thermiques d'un sol au voisinage d'un échangeur enterré à partir de la mesure in situ des profils de température, *Entropie*, **120**, p 56 (1984)

3. J.P. Laurent, Evaluation des paramètres thermiques d'un milieu poreux : optimisation d'outils de mesure « in situ », *Int. J. Heat. Mass. Transfer.*, **32**, p 1247 (1989)

4. M. Rhachi, M. Boukalouch, B. Bourret, Etude d'une nouvelle technique de mesure des températures dans le sol, *Rev. Gén. Therm.*, **36**, p 851 (1997)

5. G. Achard, J.J. Roux, J.C. Sublet, Description d'une sonde de mesure des caractéristiques thermiques des couches superficielles du sol, Résultats d'une campagne de mesures, *Rev. Gen. Therm.*, **267**, p 177 (1984)

6. H.S. Carslaw, J.C. Jaeger, *Conduction of Heat in Solids*, 2nd Editions, Oxford University Press, 1959.

7. D. Maillet, .S. Andre, J.C. Batsale, A. Degiovanni, C. Moyne, *Thermal quadrupoles, solving the heat equation through integral transforms*, Editions Wiley, 2000.

8. Matlab user's guide :The Math Works, Inc.- 21 Eliot Street - South Natick

9. H. Stehfest, Remark on algorithm 368. Numerical inversion of Laplace transform, *A.C.M.*, **53**, p 624 (1970)

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

THE INVERSE ESTIMATION OF THE LOCAL THERMAL BOUNDARY CONDITIONS IN TWO DIMENSIONAL HEATED CYLINDER

Hasna Louahlia-Gualous, Eugéne Artioukhine and Prabodh K. Panday
Centre de recherche sur les écoulements, les surfaces et les transferts
CREST CNRS-UMR 6000, Site IGE, 2 avenue Jean Moulin
90000 Belfort, France
hasna.gualous@utbm.fr

ABSTRACT

A numerical algorithm for determining local thermal boundary conditions at the external surface of the cross-section of a cylinder is developed to analyze steady-state phenomena of condensation or evaporation by solving an inverse heat conduction problem (IHCP). The iterative regularization method is used to solve the inverse problem under study. The unknown functions are approximated by cubic B-splines. The gradient is computed by using the adjoint equations. First, the numerical simulations with the heat flux distribution given a priori are curried out to verify the developed algorithm. The numerical solution of the direct problem is used to simulate temperature measurements. The cases without and with measured errors are considered. The influence of the number of approximation parameters is analyzed. Second, the heat transfer coefficient for the evaporation of a liquid film flowing on a cylinder is estimated and analyzed by the use experimental data.

NOMENCLATURE

f	temperature function, K
h	heat transfer coefficient, W/m^2K
J	residual functional, K
N	angular nodes number
Q	flux density, W/m^2
r	radius, m
R_1	internal radius of cylinder, m
R_2	external radius of cylinder, m
r_n	co-ordinate of measurement point
T	calculated temperature, K
T_{meas}	measured temperature, K

Greek symboles

φ	co-ordinate of calculated point, rad
φ_n	co-ordinate of measured point, rad
λ	thermal conductivity, W/m^2K
θ	temperature variation
Ψ	lagrange multiplier
γ	descent parameter
δ	Dirac function

Subscripts

n	measured points
w	wall
it	iteration number
φ	local value

INTRODUCTION

The study of single phase or two phase flows requires the knowledge of thermal boundary conditions such as local surface heat fluxes, temperatures or heat transfer coefficients. These parameters are important for understanding the physics of phenomena in the field of evaporation condensation and other complex flows. Recently the IHCP have been used for analyzing problems where direct measurements of surface temperatures are difficult or impossible to curry out. For this purpose various numerical algorithms have been proposed [1-7].

A large number of studies relating to experimental determination of heat exchange in condensation, evaporation and boiling exist in the literature (see, for example [8]). In the majority of cases the mean heat transfer coefficient is measured. Some authors determine the local heat transfer coefficients based on temperature profiles of the cooling liquid or the wall [9,10]. These measurements often neglect the circumferential heat conduction in the wall.

In the present article the IHCP is applied to the problem of a hollow cylinder heated at the internal surface. The iterative regularization method is used to solve the inverse problem under analysis. We use a numerical solution of different

boundary-value problems. This method was chosen because it can easily be extended for nonlinear inverse problems which we are going to implement in future to analyze boiling phenomena. It is not the case when using an analytical solution of the direct problem and the Tikhonov regularization method (see, for example, [6,7]). The method is based on the conjugate gradient method to minimize the residual functional and the residual principal as the regularizing stopping criterion [11]. Two dimensional steady-state heat conduction is considered and polar coordinates are used.

In this paper, first, the inverse problem formulation is presented. Second, the numerical algorithm developed to solve the inverse problem under study is described. Third, the results of numerical simulations are discussed. Then, the experimental set up developed to study condensation and evaporation processes is described. Finally, the results of estimating the local heat transfer coefficients for an evaporating liquid film by using experimental data are discussed.

INVERSE PROBLEM FORMULATION

The physical model is presented in figure 1. An infinitely long cylinder of internal radius R_1 and external radius R_2 is considered. The temperatures at the internal radius are treated as known. The problem is considered symmetrical and the resolution of the inverse problem is carried out on half of the cylinder.

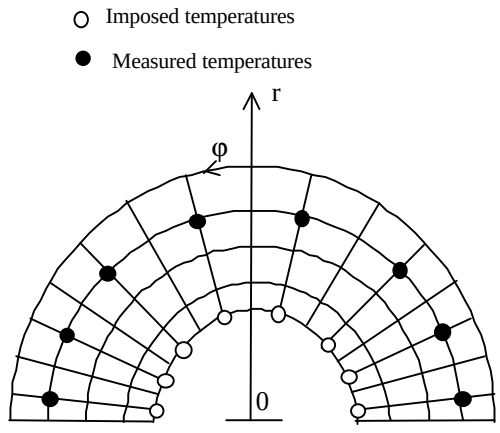


Figure 1: Physical model

The mathematical model of a heat conduction process in a hollow cylinder is given by the following system of equations:

$$\frac{\partial^2 T(r,\varphi)}{\partial r^2}+\frac{1}{r}\frac{\partial T(r,\varphi)}{\partial r}+\frac{1}{r^2}\frac{\partial^2 T(r,\varphi)}{\partial \varphi^2}=0\,,$$

$$R_1 < r < R_2\,,\ \Phi_1 < \varphi < \Phi_2 \qquad (1)$$

$$\frac{\partial T}{\partial \varphi}(r,\Phi_1)=0 \qquad (2)$$

$$\frac{\partial T}{\partial \varphi}(r,\Phi_2)=0 \qquad (3)$$

$$T(R_1,\varphi)=f(\varphi) \qquad (4)$$

$$\lambda\frac{\partial T}{\partial r}(R_2,\varphi)=Q_w(\varphi) \qquad (5)$$

In the model (1)-(5) the function $Q_w(\varphi)$, which is the heat flux density at the external surface of the cylinder, is unknown, but temperatures measured at a number N inside the cylinder are available to be used as additional information on the temperature distribution in the cylinder to estimate the function $Q_w(\varphi)$:

$$T_{meas}(r_n,\varphi_n)=f_n\,,\ n=1,2,...,N \qquad (6)$$

The inverse problem under study consists in estimating the function $Q_w(\varphi)$ from the conditions (1)-(6).

We use the variational formulation of the inverse problem under analysis. In such a case the solution of the inverse problem is based on the minimization of the residual functional defined by the following equation:

$$J(Q_w)=\frac{1}{2}\sum_{n=1}^{N}\left[T(r_n,\varphi_n;Q_w)-f_n\right]^2 \qquad (7)$$

where $T(r_n,\varphi_n;Q_w)$ are the temperatures computed at the sensor locations by solving the direct problem (1)-(5). Formally, the inverse problem the variational formulation consists in minimizing the residual functional (7) under constrains (1)-(5).

NUMERICAL ALGORITHM

The unknown function $Q_w(\varphi)$ is approximated in the form of a cubic B-spline as follows:

$$Q_w(\varphi) = \sum_{i=1}^{m} p_i \phi_i(\varphi) \tag{8}$$

where m is the number of approximation parameters, p_i are unknown approximation parameters and ϕ_i are given basis cubic B-spline. As a results, the IHCP is reduced to the estimation of a vector of parameters $p = [p_1, p_2, ..., p_m]$.

To minimize the residual functional (7) the conjugate gradient method is used. The function $Q_w(\varphi)$ is supposed to be an element of the $L_2[\Phi_1, \Phi_2]$ space.

At each iteration, the successive improvements of desired parameters are built as follows :

$$p_i^{it+1} = p_i^{it} - \gamma^{it} d_i^{it}, \ i = 1, 2, \ldots, m \tag{9}$$

where it is an iteration number. γ^{it} is the descent parameter and a positive scalar. d^{it} is the descent direction determined from the vector $d = [d_1, d_2, ..., d_m]$. It is given by the following expression:

$$d_i^{it} = -g_i^{it} + \beta^{it} d_i^{it-1}, \ i = 1, 2, \ldots, m \tag{10}$$

where the parameter β^{it} is computed as follows:

$$\beta^0 = 0 \tag{11}$$

$$\beta^{it} = \frac{\left\langle g^{it} - g^{it-1}, g^{it} \right\rangle}{\left\| g^{it} \right\|^2} \tag{12}$$

where $\langle , \rangle$ is the scalar product and $\| \cdot \|$ is the norm of the vector $g = [g_1, g_2, ..., g_m]$ in the space R^m of approximation parameters.

$$\beta^{it} = \frac{\sum_{i=1}^{m} \left(g_i^{it} - g_i^{it-1} \right) \times g_i^{it}}{\sum_{i=1}^{m} \left(g_i^{it-1} \right)^2} \tag{13}$$

To implement the iterative minimization procedure, the vector g is determined by the following equation:

$$G\, g = J'(p) \tag{14}$$

where G is the Gram's matrix for basis functions

$$G = \left\{ G_{i,j} = \left\langle \phi_i \, ; \phi_j \right\rangle, \ i, j = 1, 2, ..., m \right\} \tag{15}$$

$\left\langle \phi_i \, ; \phi_j \right\rangle$ is the scalar product in the L_2 space. The Gram's matrix is symmetric and positive. Let J'(p) be the vector gradient of the functional J(p). The most effective method for calculating the gradient J'(p) in the R^m space is based on the adjoint problem [11].

Adjoint problem

The adjoint problem is determined by the following boundary-value problem:

$$\frac{\partial^2 \psi}{\partial r^2} - \frac{1}{r}\frac{\partial \psi}{\partial r} + \frac{1}{r^2}\psi + \frac{1}{r^2}\frac{\partial^2 \psi}{\partial \varphi^2} = S(r, \varphi)$$

where :

$$S(r, \varphi) = -\sum_{n=1}^{N} \left\{ \delta(r, r_n; \varphi, \varphi_n) \times \left[T(r_n, \varphi_n; Q_w) - T_{meas}(r_n, \varphi_n) \right] \right\} \tag{16}$$

where : $R_1 < r < R_2$, $\Phi_1 < \varphi < \Phi_2$

$$\frac{\partial \psi}{\partial \varphi}(r, \Phi_1) = 0 \tag{17}$$

$$\frac{\partial \psi}{\partial \varphi}(r, \Phi_2) = 0 \tag{18}$$

$$\psi(R_1, \varphi) = 0 \tag{19}$$

$$\lambda \frac{\psi(R_2, \varphi)}{R_2} = \lambda \left. \frac{\partial \psi(r, \varphi)}{\partial r} \right|_{r=R_2} \tag{20}$$

The gradient of the residual functional is determined by solving the adjoint problem. The symbol $\delta(r, r_n, \varphi, \varphi_n)$ represents the Dirac function.

To compute the descent parameter we use a linear approximation of the descent parameter [11]:

$$\gamma^{it} = -\frac{\sum_{n=1}^{N}\left[T^{it}(r_n,\varphi_n;Q_w) - T_{meas}(r_n,\varphi_n)\right]\theta^{it}(r_n,\varphi_n;\delta Q_w)}{\sum_{n=1}^{N}\left[\theta^{it}(r_n,\varphi_n;\delta Q_w)\right]^2} \quad (21)$$

where $\theta^{it}(r_n,\varphi_n;\delta Q_w)$ is the solution of the problem in variation at the *it* iteration, resulting from the variation of the unknown function δQ_w. $\theta^{it}(r_n,\varphi_n;\delta Q_w)$ is defined at the sensor locations (r_n,φ_n). δQ_w is approximated from a cubic B-spline as follows:

$$\delta Q_w(\varphi) = \sum_{i=1}^{m} \delta p_i \phi_i(\varphi),\ i = 1, 2, \ldots, m.$$

Problem in variation

The problem in variations has the following form:

$$\frac{\partial^2\theta}{\partial r^2} + \frac{1}{r}\frac{\partial\theta}{\partial r} + \frac{1}{r^2}\frac{\partial^2\theta}{\partial\varphi^2} = 0, \quad (22)$$

where : $R_1 < r < R_2$, $\Phi_1 < \varphi < \Phi_2$

$$\frac{\partial\theta}{\partial\varphi}(r,\Phi_1) = 0 \quad (23)$$

$$\frac{\partial\theta}{\partial\varphi}(r,\Phi_2) = 0 \quad (24)$$

$$\theta(R_1,\varphi) = 0 \quad (25)$$

$$-\lambda\frac{\partial\theta}{\partial r}(R_2,\varphi) = \delta Q_p(R_2,\varphi) \quad (26)$$

The direct problem, the adjoint problem and the problem in variations are solved by using the Gauss Seidel procedure [12].

To obtain a stable solution of the inverse problem we use two approaches. The first one is the iterative regularization [11]. According to the method, the iterative improvements of the approximation parameters is terminated with the use of the residual criteria :

$$J(Q_w^{it}) \approx \delta^2 \quad (27)$$

δ^2 is the total measurement error defined as follows :

$$\delta^2 = \frac{1}{2}\sum_{n=1}^{N}\sigma^2(r_n,\varphi_n) \quad (28)$$

$\sigma^2(r_n,\varphi_n)$ is an estimate of the standard deviation for the temperature measured at the nth point.

If the number of measurement points N is relatively small, it is difficult to apply this method directly. The reason is that the sensitivity of the residual functional to variations of the approximation parameters decrease rapidly when N is decreased [13].

To increase the sensitivity we use the second approach based on smoothing the experimental data. The smoothing is realized by using B-splines with the use of the residual criterion [14]. This approach allows to interpolate the experimental temperature profiles and to compute temperatures, after smoothing, at all of the nodes on two measurement lines of used the finite difference grid.

RESULTS OF NUMERICAL SIMULATIONS

Numerical verification of the solution procedure

A copper horizontal tube (λ=389W/m^{2}K) having the radii R_1 and R_2 equal to 10 mm and 14 mm respectively is considered. A known heat flux (given a priori) is imposed at the external surface of the tube. The internal surface is maintained at the constant temperature 60°C. The direct problem is solved to determine the temperature at the measurement locations (figure 2).

In order to validate the estimation procedure we have assumed that the temperatures calculated from the direct problem solution at measurement points are the measured temperatures ($T_{meas}(r_n,\varphi_n) = f_n$). The heat flux is then estimated from the ICHP solution using these measured values. Figure 3 shows that the estimated heat flux practically coincides with the exact heat flux used to solve the direct problem. This validation is carried out for the number of approximation parameters equal to 10, 20 and 30.

The grid of 30 nodes along the circumference and 15 nodes in the radial direction has been used for these calculations.

For different values of m figure 4 shows the evolution of the residual functional $J(Q_w)$ as a function of the number of iterations. The

iterations are continued till the convergence criteria $J<10^{-10}$ is satisfied. It is observed that the number of iterations increases when the number of approximation parameters is increased.

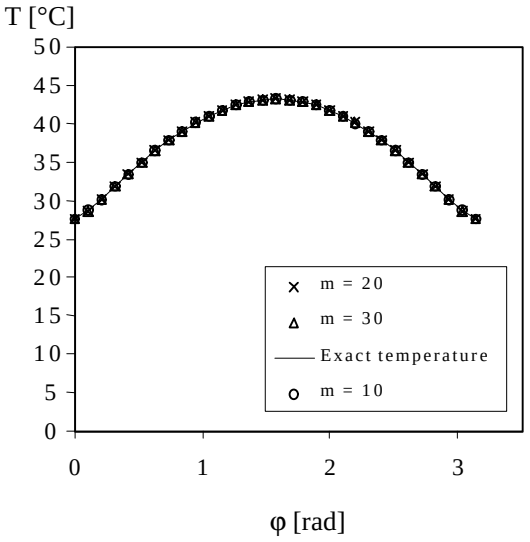


Figure 2: Comparison of calculated and measured temperatures

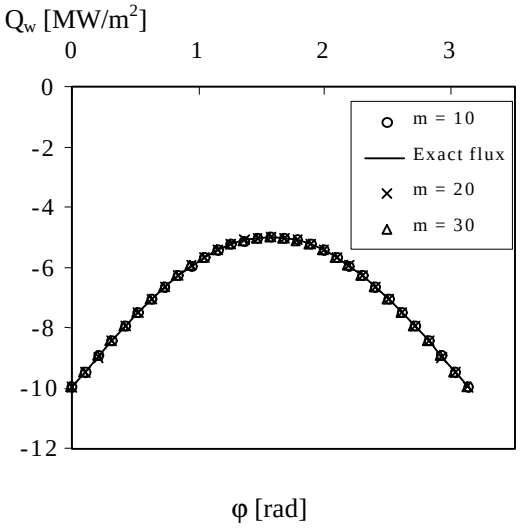


Figure 3: Comparison of the calculated and exact heat flux

Influence of measurement errors

The temperature measurements are always associated with errors whose magnitude is variable according to the measurement method and the means employed. We have studied the influence of the measurement errors on the variation of the heat flux by introducing temperature variations with respect to exact values obtained from the solution of the direct problem (figure 5).

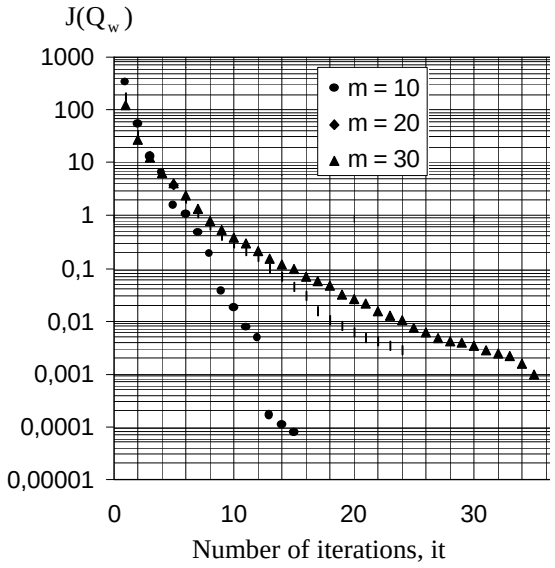


Figure 4: Evolution of the residual functional

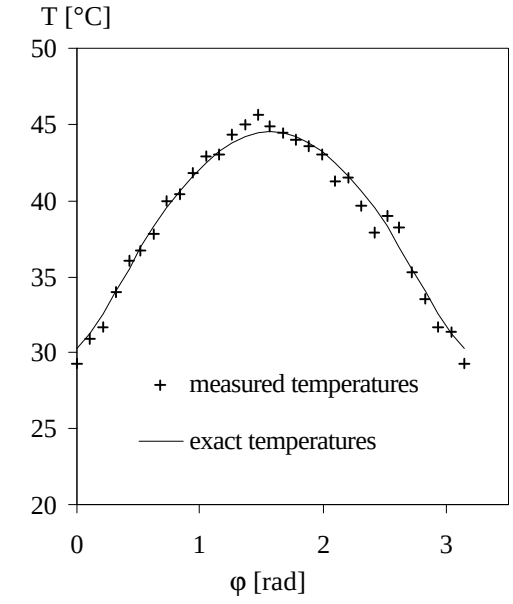


Figure 5: Evolution of calculated and measured temperatures

Figure 6 shows the variation of the local heat flux which results from these perturbations of the measured temperatures. It is observed that the error in the heat flux estimated from ICHP is ±4% With respect to the exact heat flux when the maximum error of the temperature measurement is ±6% (figure 7).

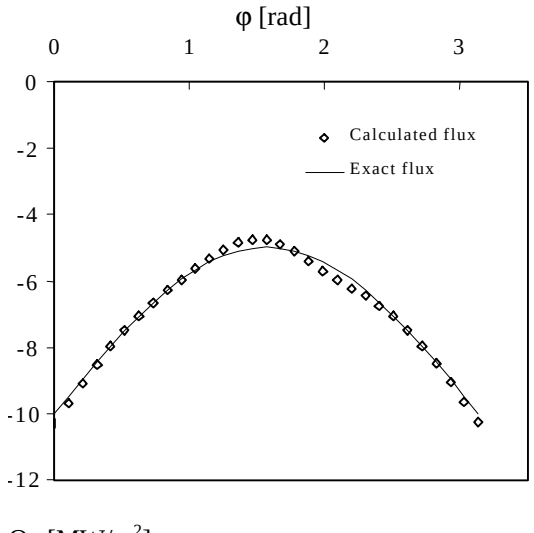


Figure 6: Evolution of calculated and measured heat flux ω

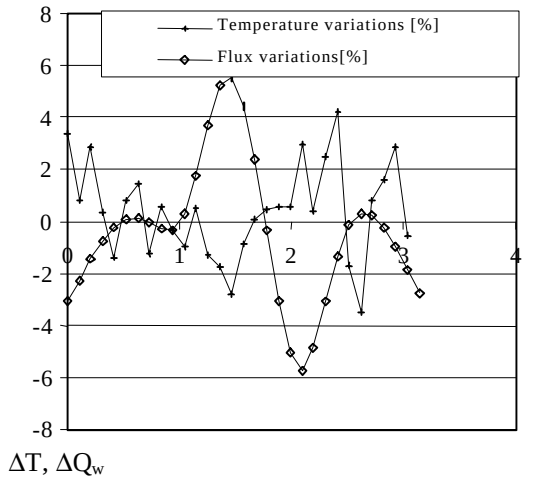


Figure 7 : Influence of temperature variations on heat flux

EXPERIMENTAL SET UP

Figure 8a shows a scheme of an experimental set up in our laboratory to study the interaction between a flowing liquid film and downward flow of air. The installation consists of a vertical wind tunnel having a rectangular section of 90x500 mm² and a centrifugal fan. In the experimental section a copper tube with a slit on the top is connected to a constant level overhead tank maintained at a controlled temperature. The experimental cylinder is placed below the liquid feed system to study different liquid flow patterns. A laser induced fluorescence technique is used to study the local liquid film thickness.

Figure 8b shows schematically the flow of a liquid film on a cylinder. Determination of the local heat flux in such a case when the surface is covered with liquid of variable thickness as is usually the case for condensation or evaporation process is a difficult problem. The thickness of the liquid film is generally less than 0.5 mm and it is difficult to measure directly the velocity and temperature distributions in the flowing film. Only the bulk temperatures at the inlet and outlet are easily accessible.

In order to treat such a problem we use the IHCP. The method of calculation developed above is proposed in order to determine the local heat flux and the surface temperature of the tube. In the present article the data of Bourouni [15] on natural convection evaporation of a liquid have been used. In his experiments the cylinder was placed directly below a rectangular liquid jet issuing from a liquid receiver.

ESTIMATION OF THE LOCAL HEAT TRANSFER COEFFICIENTS FOR THE EVAPORATION OF A LIQUID FILM

Figure 9 shows the advantage of using IHCP to determine the local heat transfer coefficients when the measurement of local wall temperatures is difficult without perturbing the flow. We present the variation of the local heat transfer coefficients for the evaporation of a liquid film flowing on a cylinder.

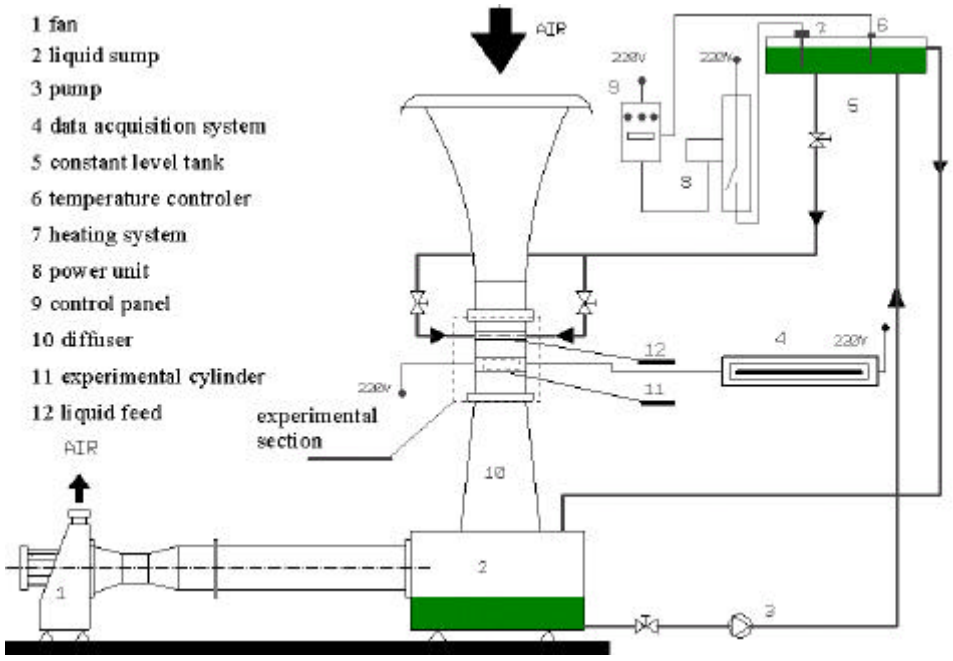


Figure 8a: Experimental set up

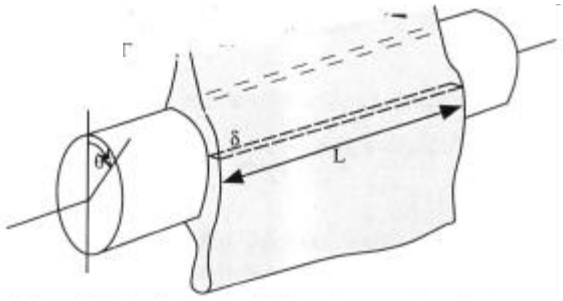


Figure 8b: Flow of a liquid film on a cylinder

For this problem the temperature distribution measured by Bourouni [15] in an electrically heated brass cylinder of internal and external radii equal to 5 mm and 11 mm respectively is used. Figure 9 shows the values of h_φ results as reported by Bourouni for film Reynolds number varying between 3000 and 5000. These values are based on the assumption that the heat flux is uniformly distributed on the inside surface of the cylinder and that the circumferential conduction is negligible.

We have evaluated the local heat transfer coefficients by solving the inverse problem using two temperatures distributions reported by Bourouni for the film Reynolds number equal to 4500. The computation is conducted by using the number of measurement points equal to the nodes number. The temperature profiles are measured at radius equal to 10.5 mm. Figure 9 shows the comparison of the values of h_φ obtained from the present model of inverse heat conduction. Figure 9 shows that the h_φ is maximum at the top of the tube (front stagnation point) and decreases continuously up to an angle of 143°. Then it increases at the base of the tube (near stagnation point). The local coefficient of heat transfer is obtained by dividing the local wall heat flux by the difference of local wall temperature and the constant saturation temperature of 80°C. The variation of the local heat flux is similar to the variation of h_φ. The general trend is similar to that obtained by Han and Fletcher [16] for the case of evaporation on a cylinder.

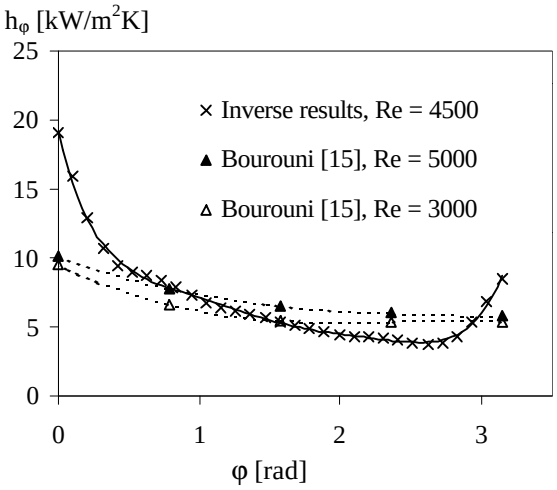


Figure 9: Local heat transfer coefficients for an evaporating film

CONCLUSIONS

This article shows the advantage of determining the surface conditions using the IHCP solution. The method of solution can be applied to other geometry and for steady-state or transient flows. The field of applications is vast. We have presented in this article the application of the solution of IHCP to a particular case of the evaporation of liquid film flowing on a cylinder where the measurements of precise local heat flux is difficult to obtain without perturbing the flow.

REFERENCES

1. R. Throne, L. Olson, The steady inverse heat conduction problem : a comparison of methods with parameter selection, *J. Heat Transfer, Transaction ASME*, **123**, 633 (2001).

2. H.T. Chen, S.Y. Linn, L.C. Fang, Estimation of surface temperature in two-dimensional inverse heat conduction problems, *Int. J. Heat Mass Transfer*, **44**, 1455 (2001).

3. S. Chantarasiriwan, Inverse determination of steady-state heat transfer coefficient, *Int. comm. Heat Mass Transfer*, **27**(8), 1155 (2000).

4. R. Abou Khachfe, Y. Jarny, Determination of heat sources and heat transfer coefficient for two-dimensional heat flow – numerical and experimental study, *Int. J. Heat Mass Transfer*, **44**, 1309 (2001).

5. C.H. Huang, J.Y. Wu, An inverse problem of determining two boundary heat fluxes in unsteady heat conduction of thick-walled circular cylinder, *Inverse Problems in Eng.*, **1**, 133 (1995).

6. D. Maillet, A. Degiovanni, R. Pasquetti, Inverse heat conduction applied to the measurement of heat transfer coefficient on a cylinder : comparison between an analytical and a boundary element technique, *J. of Heat Transfer, ASME*, **113**, 549 (1991).

7. T.J. Martin, G.S. Dulikavich, Inverse determination of steady heat convection coefficient distributions, *J. of Heat Transfer, ASME*, **120**, 328 (1998).

8. R.J. Goldstein et al., Heat transfer – a review of 1999 literature, *Int. J. Heat Mass Transfer*, **44**, 3579 (2001).

9. W.H. Parken, L.S. Fletcher, V. Sernas, J.C. Han, Heat transfer through falling film evaporation and boiling on horizontal tubes, *J. Heat Transfer, Transaction ASME*, **112**, 744 (1990).

10. S. Mochizuki, Y. Yagi, R. Tadano, W.J. Yang, Condensation heat transfer of nonazetropic binary mixtures (R113+R11) in a horizontal tube, *Transaction JSME, Part B*, **54**, 1796 (1990).

11. O.M. Alifanov, E.A. Artyukhin, S.V. Rumyantsev, *Extreme methods for solving Ill-posed problems with applications to inverse heat transfer problems*, Begell House, New York, 1995.

12. J.P. Holman, *Heat Transfer*, Mc Graw Hill Book Company, 1986.

13. G. Blanc, M. Raynaud and T.H. Bardon, A guide for the use of the function specification method for 2D inverse heat conduction problem, *Rev. Générale de la thermique*, **37**, 17 (1998).

14. T. Loulou, E. Artiykhin and J.P. Bardon, Estimation of the time dependent thermal contact resistance at the mold-casting interface, *The 2th Int. Conf. On Inverse Problems in Eng.*, Le Croisic, France (1996).

15. K. Bourouni, *Etude d'un évaporateur pour le dessalement d'eau par le procédé d'aero-évapo-condensation : expérimentation et modélisation*. Thése de l'université de provence, Aix-Marseille I, France, 1998.

16. J.C. Han, L.S. Fletcher, Influence of concurrent steam shear velocity on falling film evaporation and boiling on a horizontal tube, *Proc. of the joint ASME-JSME Thermal Engineering conference*, 527, 1987.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

TWO DIMENSIONAL COMPUTATIONAL ESTIMATION OF TRANSIENT BOUNDARY CONDITIONS FOR A FLAT SPECIMEN

Said Abboudi
CREST, UTBM, site de Sévenans,
B.P. 449 - 90010 Belfort Cedex, France
said.abboudi@utbm.fr

Eugène Artioukhine
CREST, Institut de Génie Energétique,
Université de Franche-Comté
2, Avenue Jean Moulin, 90000 Belfort Cedex,
France
artyukh@ige.univ-fcomte.fr

ABSTRACT

The problem studied in this work is the estimation of the heat flux density at the boundary of a specimen in the two dimensional case. To solve this inverse heat conduction problem, two the most popular methods are analyzed: the iterative regularization method (IRM) which is the whole time domain estimation, and the sequential function specification method (FSM).
These two methods are first verified and then compared between them by using simulated numerically data. The influence of the time step (delta Fourier number) as well as the number of approximation parameters of the heat flux spline-approximation is analyzed.
A comparison of results obtained by using the two above methods are demonstrated.

NOMENCLATURE

C	specific heat
M_y, M_t	number of approximation parameters
N_t	number of time steps
N	number of temperature sensors
T, T_a, T_0	temperature,ambient, initial temperature
Q_i^j	sensitivity coefficient at time t_j and position X_i
a, b	dimensions of the specimen
f_i	measured temperature at position X_i
h	heat transfer coefficient
t , t_f	time, final time
$q(y,t)$	heat flux density
y	adjoint variable
J	temperature variation
r	density
l	thermal conductivity
a	thermal diffusivity
d^2	estimated error or criterion
Δt	time step
ΔFo	delta Fourier number
Δq	heat flux increment

INTRODUCTION

The ill-posed inverse problem of estimating the surface heat flux from transient temperature histories measured in a heat conducting solid is constantly of a great interest during three last decades. A literature review and a presentation of different methods is given, for example, in Tikhonov and Arsenin [1], Beck et al. [2,3,4], Hensel [5], Murio [6], Alifanov et al. [7,8]. Different applications of various methods are presented, in particular, in Zabaras et al. [9] et Delaunay et al. [10]. And Woodbury et al [11]

In this paper, results of numerical analysis are reported, the goal of which is to estimate the heat flux absorbed by a flat specimen cooled at the back side and insulated at its lateral surface.

We use the iterative regularization method (IRM) (Alivanov et al., [8]) and the function specification method (FSM) (Beck et al., [2]) to solve this inverse heat conduction problem. The first numerical algorithm is based on the minimization of the residual functional which is the integrated difference between temperature histories measured and those calculated at the

sensor locations by solving the direct problem. The conjugate gradient method is used to solve the inverse problem. The residual functional gradient is computed by solving the adjoint problem and the optimal descent parameter is calculated by solving the problem for temperature variations. The heat flux evolution is approximated by cubic B-splines (Alivanov et al [7]). The second numerical algorithm is based on the minimization of the discrete least squares criterion by taking into account a few future time steps. These two methods are first verified and compared between them by using simulated numerically data. The influence of the delta Fourier number is analyzed.

INVERSE PROBLEM FORMULATION

The specimen is heated by a heat flux of unknown density at the active surface and cooled by a forced convection flow at the opposite surface.

The following hypotheses have been taken into account:

- thermophysical properties are supposed to be constant,
- heat transfer is two-dimensional.
- heat transfer coefficient is constant at the cooled surface.

Under these conditions, the heat transfer process in the specimen can be described by the following system of equations :

$$\frac{\partial T}{\partial t} = a\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right), \quad 0 < x < a, 0 < y < b, 0 < t \le t_f \qquad (1)$$

$$-\lambda \frac{\partial T(0,y,t)}{\partial x} = q(y,t) \qquad (2)$$

$$-\lambda \frac{\partial T(a,y,t)}{\partial x} = h[T(a,y,t) - T_a] \qquad (3)$$

$$-\lambda \frac{\partial T(x,0,t)}{\partial y} = 0 \qquad (4)$$

$$-\lambda \frac{\partial T(x,b,t)}{\partial y} = 0 \qquad (5)$$

$$T(x,y,0) = T_0, \; 0 < x < a, \; 0 < y < b \qquad (6)$$

In the model (1)-(6), the heat flux density $q(y,t)$ is unknown. To get additional information about the temperature distribution in the specimen, temperature histories are measured in the specimen at a certain number of points N with coordinates $(x,y) = (X_n, Y_n)$, $n = 1,2,...,N$:

$$T_{meas}(X_n, Y_n, t) = f_n(t), \; n = 1,2,...,N \qquad (7)$$

This information, together with the model (1)-(6), is used to solve the inverse problem.

ITERATIVE REGULARIZATION METHOD (IRM)

To build a computational algorithm, we use the variational formulation of the inverse problem of interest. The problem is to find an unknown function $q(y,t)$ for which temperature histories computed from the mathematical model (1) to (6) at the sensor locations would be close to measured histories. That leads to the problem of minimizing the residual functional:

$$J(q) = \sum_{n=1}^{N} \int_0^{t_f} [T(X_n, Y_n, t; q) - f_n(t)]^2 dt \qquad (8)$$

where $T(X_n, Y_n, t; q)$, $n = 1,2,...,N$, are temperature histories computed at the sensor locations with a given heat flux density $q(y,t)$.

The unknown function is parametrized in the form of a cubic B-spline :

$$q(y,t) = \sum_{n=1}^{M_y} \sum_{m=1}^{M_t} p_{nm} \beta_n(y) \beta_m(t) \qquad (9)$$

where p_{nm}, $n = 1,2,...,M_y$, $m = 1,2,...,M_t$, are unknown parameters, $\beta_n(y)$, $\beta_m(t)$, $n = 1,2,...,M_y$, $m = 1,2,...,M_t$, are given basis cubic B-splines. The numbers of approximation parameters M_y and M_t are usually fixed a priori. As a result, the inverse problem is reduced to the estimation of a vector p of parameters p_{nm}, $n = 1,2,...,M_y$, $m = 1,2,...,M_t$.

The unknown function is considered as an element of the function space $L_2([0,b] \times [0,t_f])$ of parametrized functions. We use the unconstrained conjugate gradient method of

optimization. The residual functional gradient as well as the descent direction in L_2 space have the form :

$$J'_q(y,t) = \sum_{n=1}^{M_y}\sum_{m=1}^{M_t} g_{nm}\varphi_n(y)\varphi_m(t) \qquad (10)$$

$$D(y,t) = \sum_{n=1}^{M_y}\sum_{m=1}^{M_t} d_{nm}\varphi_n(y)\varphi_m(t) \qquad (11)$$

So, the gradient is characterized by the vector $g=(g_{nm})$ and the descent direction by the vector $d=(d_{nm}), n=1,2,\ldots M_y, m=1,2,\ldots,M_t$. It is easy to show that the residual functional minimization with respect to desired parametrized functions is reduced to those with respect to unknown parameters. The successive improvements of desired parameters are constructed as follows :

$$p^{s+1} = p^s + g^s d^s, \quad s=0,1,\ldots \qquad (12)$$

where *s* is the iteration index, g^s is the descent parameter, p^0 is an initial guess for the vector of unknowns parameters given a priori. The vector d^s is computed as follows :

$$d^s = -g^s + b^s d^{s-1}, \quad s=0,1,\ldots, \qquad (13)$$

where $b^0=0$, $b^{(s)} = \dfrac{(g^{(s)}-g^{(s-1)}, g^{(s)})}{\left\| g^{(s)} \right\|^2}$

The realization of the iterative procedure (12) is based on computing the vector *g* at each iteration. This vector is determined by the relationship for the residual functional variation :

$$dJ(q,dq) = (J'_p, dp)_{R^M} = (J'_q, dq)_{L_2} \qquad (14)$$

where J'_p is the residual functional gradient in R^M space of approximation parameters and J'_q is the gradient in L_2 space of parametrized functions, $(\cdot,\cdot)$ is the scalar product.

By using the parametric form (7), it can be shown that the vector g is computed as follows (Alivanov et al., [1]):

$$J'_p = Gg \qquad (15)$$

where *G* is the Gram's matrix for basis functions.

The most effective method for calculating the gradient J'_p in $R^{M_y} \times R^{M_t}$ space is based on introducing an adjoint problem. The following expression for the gradient components can be derived :

$$J'_{p_{nm}} = \int_0^b \int_0^{t_f} y(0,y,t)\varphi_n(t)\varphi_m(t)dtdy,$$
$$m=1,2,\ldots,M_t \ , \ n=1,2,\ldots,M_y \qquad (16)$$

where $y(x,y,t)$ is the solution of the adjoint problem [2]).

A linear approximation is used to estimate the parameter g^s:

$$g^s = -\frac{\sum_{n=1}^{N}\int_0^{t_f}[T(X_n,Y_n,t)-f_n(t)]J(X_n,Y_n,t)dt}{\sum_{n=1}^{N}\int_0^{t_f}[J(X_n,X_n,t)]^2 dt} \qquad (13)$$

(17)

where $J(x,y,t)$ is the solution of the following boundary-value problem for temperature variations [2]:

To obtain stable solutions of the inverse problem under consideration, the iterative regularization is used [2]. The main idea is to terminate the iterative procedure with the residual criterion :

$$J(q^{s^*}) \approx d^2 \qquad (18)$$

where d^2 is the total (integrated) measurement error defined by :

$$d^2 = \sum_{n=1}^{N} \int_0^{t_f} s_n^2(t)dt \qquad (19)$$

$s_n^2(t)$ is an estimate of the time-dependent standard deviation for the *n*th measured temperature history. This procedure gives the most stable solution. The number *s** of the last iteration is the regularization parameter of the method.

It is necessary to note that the numbers of approximation parameters M_y and M_t should be correctly chosen for the desired function. These numbers have to be chosen so that the residual criterion would be verified.

One iteration of the numerical algorithm includes the following steps :

· solution of the direct problem and computation of the residual functional,

· verification of the residual criterion,

· solution of the adjoint problem and computation of the residual functional gradient in L_2 space,

· computation of the descent direction,

·.solution of the problem for temperature variations and computation of the optimal descent parameter,

· calculation of the heat flux approximation.

FUNCTION SPECIFICATION METHOD (FSM)

To estimate transient heat flux, the sum of squares function (Beck et al. [3]) is minimized with respect to q^{n+1}:

$$J(q^{n+1}) = \sum_{i=1}^{N}\sum_{j=1}^{r} [T(X_i, Y_i, t_{n+j}; q^{n+1}) - f_i(t_{n+j})]^2 + R \qquad (20)$$

where:

$$R = \sum_{k=0}^{2} r_k \int_0^b \left[\frac{\partial^k q}{\partial y^k}\right]^2 dy \qquad (21)$$

The first subscript refers to the space where *N* is the sensor number and the second to the time where *r* is the number of future time steps.

$q^{n+1} = [\, q_1^{n+1}\ q_2^{n+1}\ ....\ q_M^{n+1}\,]$ is the unknown heat flux.

The space stabilization, for the function specification method, is obtained by the regularization term R witch is a function of the regularization parameters r_0, r_1 and r_2 of the zeroth, first and second order respectively. The stabilization in time is obtained through the temporary assymption that flux is constant over the r futures time steps:

$$q^{n+1} = q^{n+2} = ... = q^{n+r} = q^n + \Delta q^{n+1} \qquad (22)$$

Then the minimization of the sum $J(q^{n+1})$ with respect to: Dq_l^{n+1}, $l = 1, 2, ..., M$, leads to the following system:

$$(K + K_R)Dq^{n+1} = F + F_R \qquad (23)$$

where the coefficients of the matrix *K* are:

$$K_{ml} = \sum_{i=1}^{N}\sum_{j=1}^{r} \left[Q_{i,m}^{n+j} \,.\, Q_{i,l}^{n+j}\right]$$

and of the vector *F* are :

$$F_m = \sum_{i=1}^{N}\sum_{j=1}^{r} Q_{i,m}^{n+j}\left[f_i(t_{n+j}) - T(X_i, Y_i, t_{n+j}; q^{n+1})\right] \qquad (24)$$

where: $Q_{i,m}^{n+j} = \left.\frac{\partial T_i}{\partial q_m}\right|_{q_1^{n+1},..., q_M^{n+1}}^{n+j}$ is the sensitivity coefficient at the time t_{n+j} and the sensor locations X_i.

the coefficients of the matrix K_R and the vector F_R are function of the regularization parameters and the heat flux density at the nth time step [12,13].

One iteration of the numerical algorithm includes the following steps :

· Computation of the sensitivity matrix K,

· Computation of the temperature T_i^{n+j} with the heat flux $q^{n+j} = q^n$,

· Computation of elements of the matrix K_R and vectors F and F_R.

· Computation of the heat flux increment Δq^n and the heat flux q^{n+1}.

· Computation of the temperature T_i^{n+1} with the heat flux $q^{n+1} = q^n + Dq^{n+1}$,

NUMERICAL SIMULATION AND COMPARISON OF THE TWO METHODS

To simulate the numerical solution and to compare the two methods, we supposed in the problem (1) to (6) that $l = 1$ *W/(m.°C).*., $a = 0{,}1$ *m,* $b = 0{,}1$ *m,* $h = 10$ *W/(m².°C),* $T_0 = 0$ *°C,* $T_a = 10$ *°C,*

The simulated heat flux evolution is :

$q(y,t) = q_0 f(y) g(t)$

where $f(y) = exp(-y^2 / b^2)$

and $g(t) = q_0(1 + sin(wt))exp(-mt)$

$0 < y < b$, $0 < t \le t_f$, $q_0 = 1000$ *W/m²,*

$w = 15.7$ rd/s, $b = 20$ m, $m = 0.05$, $\Delta t = 0{,}01$ s,

$t_f = 1$ *s.*

rC is computed from the delta Fourier number by the relation: $rC = \dfrac{Dtl}{DFo\, L_c^2}$ where L_C is the distance between the heated surface to the sensors locations.

The measured temperature histories were simulated numerically on the line $x = 0.1\,a$ of the specimen, at nine uniformly distributed space nodes $y_j = jb/10$, $j = 1,2,...,9$, with a random noise 4% of the maximal temperature value.

For the IRM, we used $M_y = 21$ parameters in the *Oy* direction and $M_t = 51$ parameters in time to estimate the unknown heat flux evolution.

For the FSM, the sensitivity coefficients are computed by solving the sensitivity boundary problem associated to the system (1) to (6). The number of the future time steps $r = 10$ is fixed and the regularization parameters r_0, r_1 and r_2 are determined for each case.

The alternative direction implicit (ADI) method is used for to solve different boundary problems..

In figures 1a, 1b and 1c, we show, respectively, the exact heat flux $q_{ex}(y,t)$ and the exact temperature evolutions $T(x,t)$ at $y = b/2$ and $T(y,t)$ at $x = a/2$ computed with the analyzed heat flux .

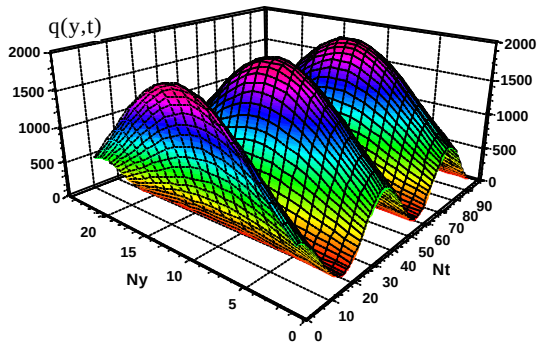


Figure 1a. Exact heat flux evolution.

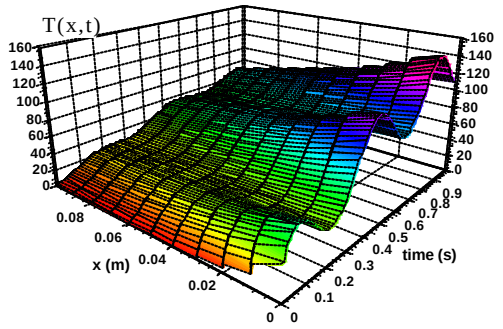


Figure 1b. Exact temperature evolution $T(x,t)$ at $y = b/2$

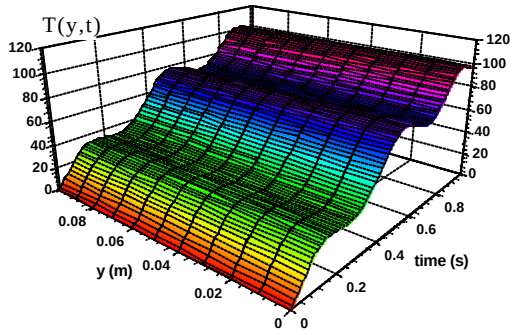


Figure 1c. Exact temperature evolution $T(y,t)$ at $x = a/2$.

For the test heat flux evolution studied in this work, the comparison between the two methods (IRM and FSM) show as in the case of the one-

dimensional study [14], a good agreement between them for delta Fourier numbers higher than 0.5.

In what follows, we show results of the comparison between the two methods for three values of delta Fourier number ΔFo *=1*; *0.1* and *0.01*. If ΔFo is less than *0.01*, to conclude on their performance, it is necessary to optimize the set of parameters for each method.

In figures 2a and 2b we present, for ΔFo *=0.1*, the results of estimating the heat flux *q(y,t)* for IRM and FSM methods respectively. We note that the general form of these surfaces in function of space coordinate *y* and time *t* is in a good agreement with studied exact heat flux, figure 1a.

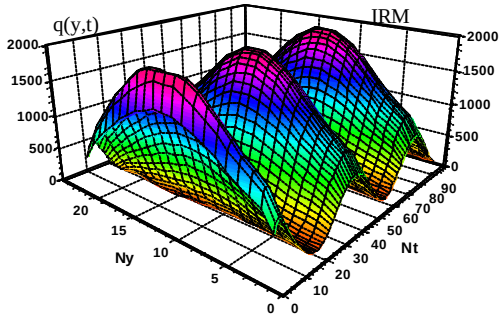


Figure 2a. Estimated heat flux evolution.

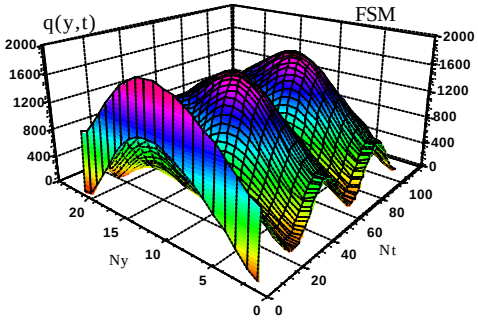


Figure 2b. Estimated heat flux evolution.

For the three studied delta Fourier numbers, we present a serie of curves in which, we compare the two methods through the results of the estimated heat flux, figures 3a, 3b and 3c but also through the temperature evolutions (at a fixed sensor location), figures 4a, 4b and 4c.

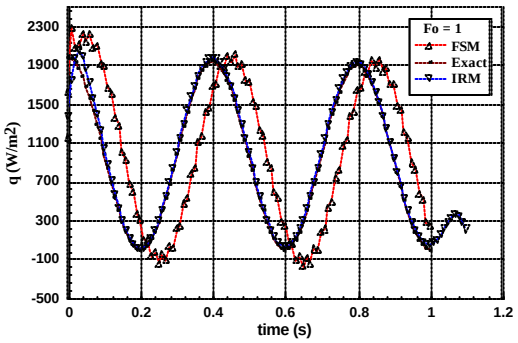


Figure 3a. Estimated and exact heat flux evolution ΔFo *=1*.

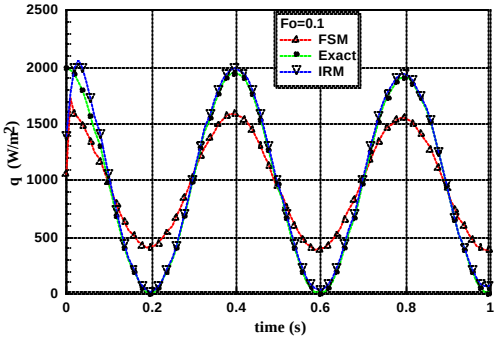


Figure 3b. Estimated and exact heat flux evolution ΔFo *=0.1*.

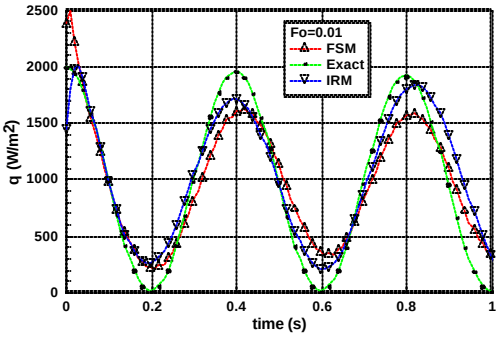


Figure 3c. Estimated and exact heat flux evolution ΔFo *=0.01*.

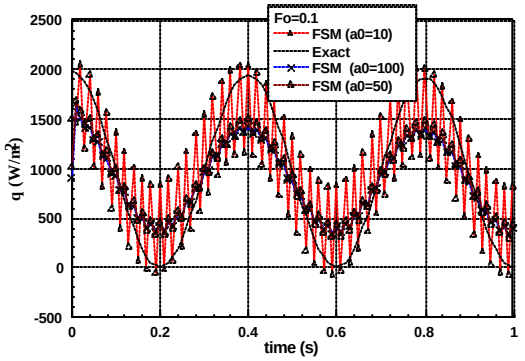


Figure 3d. Estimated and exact heat flux evolution $\Delta Fo = 0.1$, $r_0 = a_0\, 10^{-8}$.

These curves show generally, the presence, for the two methods, of a difference between the estimated and exact results which tends to increase when the delta Fourier number decrease. This variation reaches its maximum value when one approaches the high values of exact heat flux. The two methods present different behaviors.

For FSM method, we note the presence of fluctuations (or instabilities) around the exact value of the heat flux analyzed. The amplitude of these fluctuations increases when the delta Fourier number falls and decrease according to the regularization parameters of the method.

Let as note that the choice of these parameters must correspond with a good conditioning of the matrix system (32); for a fixed small condition number, the interval domain of these parameters is very large [13].

The numerical tests show that the estimated results are very sensitive to the parameter of order zero r_0 which, in addition, gives the best result of the estimate heat flux. In figure 4d, we present, for three values of r_0, a comparison between estimated (for $r_0 = 10^{-7}$, $5\ 10^{-7}$ and 10^{-6}) and exact heat flux. We can notice that these fluctuations clearly decreased for $r_0 = 10^{-6}$ but on the other hand, the difference between the estimated and exact results slightly increased.

For the IRM method, the results are more stable and approach exact heat flux a little better than FSM method in particular for $DFo = 0.01$. This is a direct consequence of the use of the regularizing residual criterion (19). For all considered cases, the residual values d^2 were computed by using the formula (18). These values are:

ΔFo	d^2
1	81.5817
0.1	20.6836
0.01	0.5476

Lastly, we find these same remarks on the curves relating to the estimated temperatures at, for example, the point of measurement $(0.05,0.02)$, figures 4a and 4c.

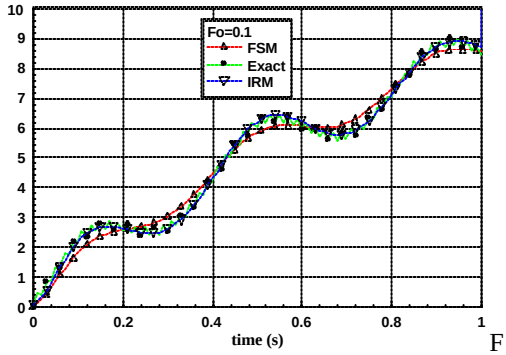


Figure 4b. Estimated and exact temperatures evolution $\Delta Fo = 0.1$

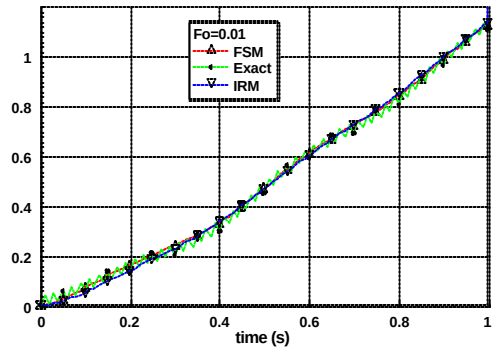


Figure 4c. Estimated and exact temperatures evolution $\Delta Fo = 0.01$.

We note that the comparis on results obtained in this paper differ from those presented for example, in Beck, [3]. An explanation is that we use another realization of the IRM method based on a spline-approximation of the unknown heat flux evolution. In particular, the residual functional

gradient is not equal to zero at the final time with this approximation.

CONCLUSIONS

To estimate a heat flux evolution from measured internal temperature histories, the IRM and FSM was studied. A comparison between these two methods was carried out for different dimensionless ΔFo numbers.

This comparative study between the two methods in two-dimensional linear heat conduction problem studied shows that there is a good agreement between the estimated and exact heat flux for delta Fourier numbers higher than approximately *0.5*.

For smaller ΔFo numbers a more detailed analysis of these two methods is necessary. Indeed, for IRM method, the iterative regularization criterion allows to improve a little better the estimated results by an analysis of the numbers of approximation parameters used in the relation (9).

For SFM method, the estimated results become very sensitive to the various parameters of the methods: the sensor locations, the number of future time steps and the regularization parameters. An optimization of these parameters is needed.

It should be underlined that the realized comparison of the two methods was rather restricted. It is necessary to continue this analysis to establish the domains where each of the methods is more effective.

We note finally that the IRM method is more time consuming that the FSM method (about five to six time slower) .

REFERENCES

[1] Tikhonov, A., Arsenin V., 1977, *Solution of Ill-posed Problems*, Wiley, New York.

[2] Beck, J. V., Blackwell, B., St. Clair, C. R., 1985, *Inverse Heat Condition : Ill-Posed Problems*, Wiley Intersc., New York.

[3] Beck, J. V., 1993, *Comparison of the Iterative Regularization and Function Specification Algorithms for the Inverse Heat Conduction Problem*, in Zabaras et al., 1993.

[4] Beck, J. V., Blackwell, B., and Hajisheikh, A., 1996, *Comparison of Some Inverse Heat Conduction Methods Using Experimental Data*, International Journal of Heat and Mass Transfer, Vol. 39, No. 17, pp. 3649-3657.

[5] Hensel, E., 1991, *Inverse Theory and Applications for Engineers*, Prentice Hall, Englewood Cliffs, New Jersey.

[6] Murio, D., 1993, *The Mollification Method and the Numerical Solution of Ill-Posed Problems*, John Wiley and Sons, Inc., New York.

[7] Alifanov, O. M., Artyukhin, E. A., and Nenarokomov, A. V., 1987, Spline-Approximation of the Solution of the Inverse Heat Conduction Problem, Taking Account of the Smoothness of the Desired Function, *High Temperature*, Vol. 25, No.5, pp. 520-526.

[8] Alifanov, O. M., Artyukhin, E. A., and Rumyantsev, S. V., 1995, *Extreme Methods for Solving Ill Posed Problems with Applications to Inverse Heat Transfer Problems*, Begell House, Inc., New York.

[9] Zabaras, N., Woodbury, K., and Raynaud, M., (Eds), 1993, Proceedings of the First International Conference on *Inverse Problems in Engineering: Theory and Practice*, Palm Coast, Florida, USA, June 13-18 1993, ASME, 1993.

[10] Delaunay, D., Jarny,Y., and Woodbury, K., (Eds), 1996, Proceedings of the 2nd International Conference on *Inverse Problems in Engineering: Theory and Practice*, Le Croisic, France, June 9-14 1996, ASME, 1998.

[11] K. A. Woodbury (ed), Proceedings of the 3rd International Conference on *Inverse Problems Engineering: Theory and Practice*, June 13-18, 1999, Port Ladlow, Washington, USA, ASME 2000.

[12] Blanc G, Raynaud M. and Chau T; H., *A guide for the use of the function specification method for 2D inverse heat conduction problems*. Rev Gén. Therm, 1998, 37, 17-30.

[13] Blanc G, *Problème inverse multimdiemnsionnel de conduction de la chaleur Optimisation du positionnement des capteurs et utilisation des mesures de thermo-deformations*, Thèse de Doctorat de l'INSA de Lyon, Janvier 1996.

[14] Abboudi S., Artioukhine, E., Riad H., *Computational and experimental estimation* of boundary conditions for a flat specimen. Proceedings of the Third International Conference on *Inverse Problems in Engineering: Theory and Practice*, Port Ladlow, Seattle, USA, June 13-18 1999, ASME.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

Parameter Estimation in Solid-Liquid Adsorption with a Stochastic Global Optimization Method

João Flávio V. de Vasconcellos[1] and Antônio J. Silva Neto[1,2]
[1]Department of Mechanical Engineering and Energy, Instituto Politécnico, IPRJ, Universidade do Estado do Rio de Janeiro, UERJ, P.O. Box 97282, 28601-970, Nova Friburgo, Rio de Janeiro, RJ, Brazil
jflavio@iprj.uerj.br
ajsneto@iprj.uerj.br
[2]Nuclear Engineering Program – PEN/COPPE, Universidade Federal do Rio de Janeiro, UFRJ, Caixa Postal 68509, CEP 21945-970, Rio de Janeiro, RJ, Brazil
ajsneto@lmn.con.ufrj.br

Cesar Costapinto Santana
School of Chemical Engineering
Universidade Estadual de Campinas- UNICAMP
CP 6066, CEP 13081-970, Campinas, SP, Brazil
santana@feq.unicamp.br

Francisco J. C. P. Soeiro
Mechanical Engineering Department
College of Engineering
Universidade do Estado do Rio de Janeiro
CEP 20550-013, Rio de Janeiro, RJ, Brazil
soeiro@uerj.br

ABSTRACT

In the present work and inverse mass transfer problem involving biomolecules adsorption is solved. The mass transfer coefficients and the coefficients of the adsorption isotherm are estimated using an implicit formulation with the cost function of the squared residues, between calculated and measured concentrations, being minimized with a Genetic Algorithm. One configuration was considered: the transient experimental data on the solute concentration is obtained in a stirred-bath system. In the implicit formulation considered here, the direct problem is solved several times along the iterative procedure used for the solution of the inverse problem. The solution of the direct problem was obtained using a combination of finite volume method and difference finite method, and was validated with real experimental data of the concentration for the protein Bovine Serum Albumin (BSA) on Accell Plus QMA® macroporous adsorbent resin using Tris-HCL buffer solution.

INTRODUCTION

The solution of inverse problems has several relevant applications in engineering. As an example, a lot of attention has been devoted to the estimation of boundary and initial conditions in heat conduction problems as well as thermal properties and heat source intensities in such diffusive processes. On the other hand, despite its relevance in mechanical and chemical engineering, there is not a sufficient number of published results on inverse mass transfer or heat convection problems. Recently Denisov [1] has considered the estimation of an isotherm of absorption, and Lugon *et alii* [2] have investigated the determination of adsorption isotherms in a gas-liquid adsorption system with applications in the food and pharmaceutical industry.

In this work the mass transfer coefficients and the coefficients of the adsorption isotherm are estimated using an implicit formulation for the inverse mass transfer problem, with the cost function of the squared residues, between calculated and measured concentrations, being minimized with a Genetic Algorithm. The analysis and solution of inverse mass transfer problems are used herein to estimate parameters in a rather important problem in the modern biotechnology industry: the protein adsorption in chromatography systems. Chromatography systems involve the movement of macromolecules from a liquid solution to the active sites located in the interior of the pores of the adsorbent. Therefore the phenomenon is controlled by mass transfer mechanisms. These mechanisms consist of external diffusion, the effective diffusion in the pores and the velocity of the adsorption on the active site of the adsorbent.

Proteins can be separated by some form of chromatography or selective adsorption, and this process may be carried out in a packed column or a stirred tank using a specially chosen adsorbent [3]. To have a better understanding of the

processes involved, as well as to allow the design and scale up of adsorption equipment, it is necessary the determination of the physicochemical properties used in the mathematical and numerical models developed for the system analysis. In order to achieve this objective, inverse mass transfer problems may be used as important tools.

A useful literature related to numerical and experimental studies on protein adsorption kinetics is available. Horstmann and Chase [4] developed a model describing the kinetics of adsorption of proteins to adsorbents. This model includes the effects of external film mass transfer and pore diffusion. Numerical solution of the governing differential equations was carried out using a finite difference method. Liapis and Rippin [5] described a multicomponent adsorption equation from a finite bath onto adsorbent particles. This model includes the external film resistance effect and diffusional resistance within the particle. Orthogonal collocation was used to solve the equations for two component adsorption. Estimates were made of pore and solid diffusion within carbon particles by using superposition of model prediction onto experimental results. Firouztale *et alii* [6] employed theoretical models and nonlinear regression analysis to determine the values of liquid film mass transfer and effective pore diffusion coefficient. These values seem to be consistent with values estimated from empirical relations. Conder and Hayek [3] used a two-equation model to simulate the protein adsorption on a rigid, silica-based medium. The differential equations were solved using an approach equivalent to that employed by Horstmann and Chase [4]. The mass transfer coefficient and pore effective diffusivity were evaluated for Bovine Serum Albumin (BSA) adsorption on Accell Plus QMA® macroporous adsorbent resin by Silva *et alii* [7]. The model solution was obtained through the orthogonal collocation method. A nonlinear multivariable optimization method was used to fit the model to experimental results.

The objectives of the present study are: (i) present the finite volume / finite difference formulation for a set of differential equations that models the mass transfer phenomena; (ii) compare the numerical results with experimental data; and (iii) use a global optimization method (Genetic Algorithm) with the finite volume / finite difference formulation and real experimental data to estimate parameters related to adsorption of biomolecules in a stirred bath.

DIRECT PROBLEM FORMULATION

A model that describes the protein adsorption on macro porous solids was presented by Horstmann and Chase [4]. This model includes the effects of external film mass transfer and pore diffusion as well as an expression for the rate of surface reaction. The assumptions used as the basis of the model can be seen in Horstmann and Chase [4] and will be briefly listed here to the completeness of the present work: (i) The adsorbent is made of a porous material, into which the solute must diffuse, in a manner described by an effective diffusivity, D_{eff} ; (ii) Mass transfer to the surface of the adsorbent is governed by a film model characterized by a mass transfer coefficient, k_s ; (iii) The bed is homogeneous, i. e., adsorbent particles are spherical, with uniform size, i. e., radius *R*, and density ρ; (iv) The protein is distributed evenly throughout the interior of the particle; (v) Adsorption is isothermal and its equilibrium behavior can be represented by the Langmuir equation; and (vi) Axial dispersion is negligible in this simulation.

We now present the mathematical model developed. The mass balance over a solid particle leads to the differential equation that describes the solute (protein) diffusion in the particle's pores (adsorbent resin) [8]

$$\frac{\partial C_i}{\partial t} = \frac{\Psi}{r^2}\frac{\partial}{\partial r}\left[r^2 \frac{\partial C_i}{\partial r} \right] \tag{1a}$$

$$\Psi = \frac{D_{eff}}{1 + \dfrac{\rho\, q_m k_d}{\varepsilon_p (k_d + C_i)^2}} \tag{1b}$$

where C_i is the protein concentration in the liquid phase in the interior of the particles pores, D_{eff} is the effective diffusion coefficient, ρ is the density of the adsorbent particles, ε_p is the particle porosity, q_m is the Langmuir isotherm equilibrium constant and k_d is the maximum binding capacity of adsorbent and *t* and *r* are the temporal and spatial (radial) variables respectively.

The initial condition is the following

$$C_i(r,t=0)=0 \text{ for } 0 \le r \le R \quad (2)$$

and the boundary conditions are

$$\frac{\partial C_i}{\partial r}=0 \text{ for } t>0 \text{ and } r=0 \quad (3)$$

$$\varepsilon_p D_{eff} \frac{\partial C_i}{\partial r} = k_s(C_b - C_i) \text{ for } t>0 \text{ and } r=R \quad (4)$$

where C_b is the bulk concentration of protein in the liquid phase, R is the particle radius and k_s is the external mass transfer coefficient.

The mass balance in the bulk liquid phase with respect to the protein concentration can be written as [8]

$$\frac{\partial C_b}{\partial t} = -\frac{3}{R}\frac{1-\varepsilon_b}{\varepsilon_b} k_s\left(C_b - C_i\big|_{r=R}\right) \quad (5)$$

where ε_b is the bed porosity. Eq. (5) is subjected to the following initial condition

$$C_b = C_0 \text{ when } t = 0 \quad (6)$$

When the geometry, physicochemical properties, process parameters and initial and boundary conditions are known, the solute concentration in both liquid and solid phases can be calculated. This is the direct problem. In the next section we show the solution of the direct mass transfer problem under consideration using a finite volume approximation for the solid phase, and a finite difference approximation for the liquid phase.

FINITE VOLUME FORMULATION

The complexity of the analysis of adsorption systems depends markedly on the type of isotherm involved and on the degree of mass-transfer resistance encountered. Analytical solutions can be obtained in the case of negligible mass transfer combined with rapid equilibrium [8]. However, these conditions are not commonly found in practical situations. In this work, the coupling of equations (1) and (5), with respective initial and boundary conditions precludes the determination of an analytical solution. Therefore, the finite volume method is used to solve the problem in the solid phase governed by Eq. (1) while the finite difference method is used to solve the problem in the liquid phase governed by Eq. (5). The simplicity of Eq. (5), with no spatical dependence, was the main reason to use the finite difference method to solve this equation.

To the best of the authors' knowledge, no previous studies of adsorption in a stirred bath using finite volume method have been reported by other authors. Finite difference method was used to solve all differential equations in Refs. [2-4] while orthogonal collocation method on finite difference method was used in Refs. [5-7].

SOLID PHASE EQUATIONS

The basic step in the derivation of a finite volume method is to integrate the conservation equation in divergence form over the control volume represented in Fig. 1.

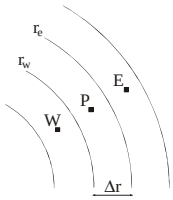


Figure 1 - Elemental control volume used to integrate Eq. (1).

The integration of the transient term in Eq. (1) both in time as well as over the control volumes yields

$$\int_{\Delta r}\left[\int_{\Delta t}\frac{\partial C_i}{\partial t}dt\right]4\pi r^2 dr = \frac{4\pi}{3}\left(C_{iP} - C_{iP}^{\,o}\right)\left(r_e^3 - r_w^3\right) \quad (7)$$

where the superscript "*o*" means quantities evaluated at the previous time level. Integration of the right-hand side term of Eq. (1) gives

$$\int_{\Delta r}\int_{\Delta t}\left[\frac{\Psi}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial C_i}{\partial r}\right]\right]dt\,4\pi r^2 dr = 4\pi\Psi_P\left[r_e^2\frac{C_{iE}-C_{iP}}{\Delta r} - r_w^2\frac{C_{iP}-C_{iW}}{\Delta r}\right]\Delta t \quad (8)$$

where

$$\Psi_P = \frac{D_{eff}}{1+\dfrac{\rho\, q_m k_d}{\varepsilon_p\left(k_d + C_{iP}\right)^2}} \quad (9)$$

Introducing the results of these integrations into Eq. (1), we obtain

$$A_P C_{i\,P} = A_E C_{i\,E} + A_W C_{iW} + S_P C_{i\,P}^{o} \qquad (10)$$

where

$$A_E = \Psi_P \frac{r_e^2}{\Delta r} \;, \quad A_W = \Psi_P \frac{r_w^2}{\Delta r} \qquad (11a,b)$$

$$S_P = \frac{1}{3\Delta t}\left(r_e^3 - r_w^3\right), \; A_P = A_E + A_W + S_P \qquad (11c,d)$$

For each interior volume there will be a corresponding discretized equation as represented by Eqs. (10) and (11). For a near-boundary control volume, the discretization equation should contain the boundary condition. Therefore, two new equations must be written, one for the first volume, where Eq. (3) should be satisfied, and the other for the last volume, where Eq. (4) should be satisfied.

The integration of the transient term in Eq. (1) remains unchanged for the first and the last volume, but the integration of the right-hand side term of Eq. (1) gives for the first volume

$$\int_{\Delta r}\int_{\Delta t}\left[\frac{\Psi}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial C_i}{\partial r}\right]\right]dt\,4\pi r^2 dr = 4\pi\Psi_P\left[r_e^2\frac{C_{iE}-C_{iP}}{\Delta r}\right]\Delta t \qquad (12)$$

Introducing the results of these integrations into Eq. (1), it can be written in a form equivalent to Eq. (10) with

$$A_E = \Psi_P \frac{r_e^2}{\Delta r} \;, \; A_W = 0 \qquad (13a,b)$$

$$S_P = \frac{r_e^3}{3\Delta t} \;, \; A_P = A_E + S_P \qquad (13c,d)$$

The integration of the right-hand side term of Eq. (1) gives the finite volume equation for the last volume

$$\int_{\Delta r}\int_{\Delta t}\left[\frac{\Psi}{r^2}\frac{\partial}{\partial r}\left[r^2\frac{\partial C_i}{\partial r}\right]\right]dt\,4\pi r^2 dr = 4\pi\Psi_P\left[R^2\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R} - r_w^2\frac{C_{i\,P}-C_{iW}}{\Delta r}\right]\Delta t \qquad (14)$$

Introducing the results of integration into Eq. (1), leads to

$$A_P C_{i\,P} = A_W C_{iW} + S_P \qquad (15)$$

where

$$A_W = \Psi_P \frac{r_w^2}{\Delta r} \;, \; A_P = A_W + \frac{1}{3\Delta t}\left(r_e^3 - r_w^3\right) \qquad (16a,b)$$

$$S_P = \frac{C_{i\,P}^{o}}{3\Delta t}\left(r_e^3 - r_w^3\right) + \Psi_P R^2 \left.\frac{\Delta C_i}{\Delta r}\right|_{r=R} \qquad (16c)$$

Special care is required to obtain an approximation for $\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R}$ in Eq. (16b). Using Eq. (4) this term can be evaluated as

$$\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R} = \frac{k_s}{\varepsilon_p D_{eff}}\left(C_b - C_i\big|_{r=R}\right) \qquad (17)$$

Since the values for C_i are stored at the center of each volume, as represented in Fig. 1, the value of $C_i\big|_{r=R}$ is not available to be used in Eq. (17). This value needs to be carefully determined or the solution will change appreciably. In this work a Taylor series expansion is used to determine the value of $\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R}$ yielding

$$C_i\big|_{r=R} = \frac{9C_{iN} - C_{iN-1} + 3\dfrac{\Delta r\, k_s}{\varepsilon_p D_{eff}} C_b}{8 + 3\dfrac{\Delta r\, k_s}{\varepsilon_p D_{eff}}} \qquad (18)$$

With Eqs. (10) and (15) a system of algebraic linear equations is assembled for a given time level for the unknowns C_{ij} with *j=1,2,...,N*. In the present work this system is solved using the GMRES method [9] with a preconditioning procedure based on a incomplete LU factorization [10]. We have used the routines developed by Balay *et alii* [11].

LIQUID PHASE EQUATION

For the solution of the mass transfer problem in the liquid phase, a numerical approximation to Eq. (5) is constructed based on the following third-order accurate finite difference in time [12]:

$$\frac{f^{t+\Delta t}-f^{t}}{\Delta t}=\frac{1}{2}\left[\frac{\partial f}{\partial t}-\frac{\Delta t}{6}\frac{\partial^2 f}{\partial t^2}\right]^{t+\Delta t}+\frac{1}{2}\left[\frac{\partial f}{\partial t}+\frac{\Delta t}{6}\frac{\partial^2 f}{\partial t^2}\right]^{t} \tag{19}$$

where the superscripts t and $t + \Delta t$ indicate, respectively, evaluation at the old and new time level in the numerical solution. Using C_b in place of f in Eq. (19) gives

$$\frac{C_b^{\,t+\Delta t}-C_b^{\,t}}{\Delta t}=\frac{1}{2}\left[\frac{\partial C_b}{\partial t}-\frac{\Delta t}{6}\frac{\partial^2 C_b}{\partial t^2}\right]^{t+\Delta t}+\frac{1}{2}\left[\frac{\partial C_b}{\partial t}+\frac{\Delta t}{6}\frac{\partial^2 C_b}{\partial t^2}\right]^{t} \tag{20}$$

In order to simplify Eq. (20), we assume that the rate of change of the bulk liquid concentration is equal to the rate of change of the concentration in the pore liquid at the surface of the particle, i. e.,

$$\frac{\partial C_b}{\partial t}\cong\left.\frac{\partial C_i}{\partial t}\right|_{r=R} \tag{21}$$

Therefore, it can be proved that the second derivative of concentration with respect to time is negligible. This simplification gives the equation used to calculate the bulk liquid concentration

$$C_b=C_b^o-\frac{\Delta t}{2}\frac{1-\varepsilon_b}{\varepsilon_b}\frac{3}{R}D_{eff}\,\varepsilon_p\left\{\left[\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R}\right]^{t+\Delta t}+\left[\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R}\right]^{\Delta t}\right\} \tag{22}$$

Inspecting Eq. (22) one recognizes that the protein concentration in the liquid phase in the interior of particles pores, C_i at $r = R$, is also required to calculate the protein concentration in the liquid. Inspecting Eq. (18) one recognizes that the protein concentration $C_i\big|_{r=R}$ is a function of C_b and C_i in the interior of the particle. Therefore, an iterative procedure must be used in order to update the values of C_b and C_i in Eqs. (18) and (22).

The solution procedure for the determination of C_b and C_i is as following:

1) Set the initial values for C_b and C_{ij} with *j=1,2,...,N*;

2) Calculate $C_i\big|_{r=R}$ using Eq. (18) with the most recent values for C_b and C_i;

3) Calculate $\left.\frac{\Delta C_i}{\Delta r}\right|_{r=R}$ using Eq. (17);

4) Calculate C_b using Eq. (22);

5) Assemble the system of linear algebraic equations using Eqs. (10), (11), (13), (15) and (16);

6) Solve the system of linear algebraic equations for the determination of C_{ij} with *j=1,2,...,N*;

7) Return to step 2 until a prescribed convergence criteria is achieved; and

8) Return to step 2 for the next time step.

GENETIC ALGORITHM

We use here an implicit formulation in which the inverse problem is written as an optimization problem[13]. Genetic Algorithms belong to the general category of stochastic global optimization methods. They have their philosophical basis in a process found in nature related to the evolution of the different species. Darwin's theory of survival of the fittest gives the main idea of the method. A set of feasible designs constitutes a generation that has a fixed number of individuals. A set of better designs is derived from the previous generation where the individuals are allowed to reproduce and cross among themselves with bias allocated to the fittest members. Combinations of the most favorable characteristics of the mating members of the population results in a new generation that is more fit than the previous one. In a genetic algorithm, a binary string represents each set of variables. Its natural counterpart is the chromosomal string of an individual in a biological population. This chromosomal structure represents the memory of a generation which is changed when the members of the population reproduce.

A fitness function is created indicating how "good" a member is in his generation. For an unconstrained maximization problem, the cost function can be used as the fitness function. For a minimization problem, the cost function has to be altered. The inverse of the cost function, or the difference between a large number and the cost function for each member of the population, can be used as the fitness function in a minimization problem. For constrained problems, an exterior

penalty function approach can be used to transform the problem into an unconstrained one.

Based on the ideas described above a genetic algorithm is implemented with three basic operations: reproduction, crossover and mutation. In the reproduction process individuals are selected based on their fitness. One simplistic way to selecting members from an initial population is to assign a probability of selection $f_i/\Sigma f_i$ to the *ith* member where f_i is the fitness. A new population of the same size as the original is created with a higher average fitness value. After reproduction, more copies of the dominant or fit designs are present in the population. The crossover allows for an exchange of characteristics among members of the population with the intent to improve the fitness of the next generation. Crossover is executed by selecting two mating parents, randomly choosing sites on the genetic strings and swapping strings of 0's and 1's between these chosen sites. In the present work the uniform type of crossover was used. In this case, a mask of 0's and 1's is randomly generated to indicate the positions in the chromosome strings where the parents' bits will be exchanged. For example, a mask 00101 indicates that the digits in the third and in the fifth positions will be swapped between the mating parents. The remaining digits will be kept unchanged. The crossover operation will be performed with a user defined probability p_c. Mutation is a safeguard against losses of useful genetic material during reproduction and crossover. The process of mutation is simply to choose few members from the population based on a probability of mutation p_m and switch a 0 to 1 and vice-versa at a randomly selected site on the selected member.

The genetic algorithm code used in this work was developed by Wall [13].

RESULTS AND DISCUSSION

VALIDATION OF THE DIRECT PROBLEM

All the solutions shown herein were performed for Bovine Serum Albumin (BSA, molecular weight = 67,000) on Accell Plus QMA®macroporous adsorbent resin using Tris-HCL buffer solution (0.02 M pH 7.0). This resin has a global specific mass of 1.16 g/cm3, an average diameter of 46 μm, an average particle porosity of 0.60, and a bed porosity of 0.992 [7].

Values of D_{eff} and k_s were obtained by Silva *et alii* [7] using a nonlinear regression for a given initial concentration of adsorbate. They also obtained the equilibrium isotherm of BSA using the measurement technique briefly described in that work. The values obtained in their work were: $q_m = 302.66$ mg/g and $k_d = 0.0051$ mg/ml. These values were used in most of the simulations performed in this work.

The same experimental data and parameter values presented in Ref. [7] were used in the present work to check how the numerical model compares with the experimental results. This comparison is presented in Fig. 2 and it can be seen that the finite volume / finite difference model fits fairly well the experimental data.

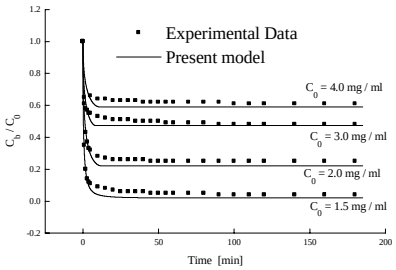


Figure 2 - Experimental and calculated kinetic curves.

SENSITIVITY ANALYSIS

An important requirement in parameter estimation is that the sensitivity coefficients should not be of small magnitude, and when more than one parameter are estimated simultaneously, their sensitivity coefficients must be linearly independent over the experimental time domain [14]. Similar shapes (time dependence) of sensitivity coefficients for two different parameters indicate that their effects on the model response are similar, being impossible therefore to tell them apart. Larger sensitivity coefficients are related to better chances of obtaining a good estimate.

In the present work we analyzed the scaled sensitivity coefficients which are defined as

$$X_s(t) = \frac{\partial C_b(t)}{\partial \beta_s}\beta_s \quad , \text{s= } 1, 2, ..., Np \qquad (23)$$

where β_s are the unknowns, that in the present work may be one of these parameters: $\{D_{eff}, k_s, q_m, k_d, \varepsilon_p\}$. As it can be observed in Eq. (23) the scaled sensitivity coefficients have all the

same units of protein concentration, and a direct comparison is then possible.

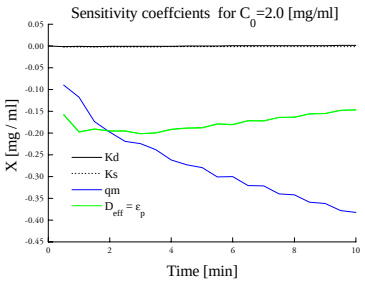


Figure 3 - Sensitivity coefficients for C_o = 2.0 [mg/ml].

In order to calculate the normalized sensitivity coefficients, approximate values of β_s are needed. For five initial BSA concentration values, C_0 = 1.5, 2.0, 3.0, 3.5, 4.0 mg/ml, and using parameters available in [7], we plotted the scaled sensitivity coefficients. Fig. 3 displays the scaled sensitivity coefficients for C_0 = 2.0 mg/ml. It is shown in this figure that only $\{D_{eff}, q_m\}$ can be simultaneously estimated because the sensitivity coefficients for $\{k_s, k_d\}$ are of small magnitude and nearly equal, therefore both parameters cannot be accurately estimated. Also q_m and ε_p are correlated.

ESTIMATION OF THE BSA PROPERTIES

Three tables are shown with the estimation of the BSA properties for three different values for the initial concentration, because as observed by Silva *et alii* [7] some of the properties depend on BSA initial concentration. On these tables are shown three different test cases. Tests 1 were performed just to determine the value of the residual cost function, using parameters provided by Silva *et alii* [7]. In this test, we did not estimate any property. Properties were determined in test-case 2, D_{eff}, and in test-case 3, D_{eff}, q_m. The residual cost function was also calculated for these cases.

$$Q = Q(\vec{Z}) = \sum_{l=1}^{K} [C_{bl} - Y_l]^2 \qquad (24)$$

where Y_l are the experimental data.

Tables 2 to 4 summarize the estimated coefficients for $C_0 = 1.5$, $C_0 = 2.0$ and $C_0 = 3.0$ mg/ml, respectively. As previously mentioned k_s and k_d cannot be estimated because of the small magnitude of the sensitivity coefficients. Therefore, we have used the values for k_s shown in Table 1 and for all test cases, $k_d = 0.0051$ mg/ml [7].

Table 1 - Values of for different initial concentration [7].

Initial BSA Concentration [mg/ml]	k_s x 10^2 [cm/s]
1,50	20,17
2,00	6,88
3,00	4,05

As can be seen in tables 2 to 4, values of the residual cost function are lower using the estimated coefficients calculated by the formulation presented herein than that estimated by Silva *et alii* [7].

Table 2 - Estimated parameters for $C_0 = 1.5$ mg/ml

Test Case	Fixed Parameters	Estimated Parameters	Cost Function Q[mg/ml]^2
1	D_{eff} = 15.37 x 10^{-7} q_m = 302.66		4.5247 x 10^{-3}
2	q_m = 302.66	D_{eff} = 25.37 x 10^{-7}	5.4720 x 10^{-4}
3		D_{eff} = 25.10 x 10^{-7} q_m = 296.65	5.4459 x 10^{-4}

Table 3 - Estimated parameters for $C_0 = 2.0$ mg/ml

Test Case	Fixed Parameters	Estimated Parameters	Cost Function Q[mg/ml]^2
1	D_{eff} = 6.79 x 10^{-7} q_m = 302.66		8.8426 x 10^{-3}
2	q_m = 302.66	D_{eff} = 6.80 x 10^{-7}	8.7909 x 10^{-3}
3		D_{eff} = 6.79 x 10^{-7} q_m = 302.69	8.7495 x 10^{-3}

CONCLUSION

The agreement between the experimental observations and the numerical results suggests that the method of finite volume / finite difference was used successfully to solve the coupled non-

linear equations associated with adsorption of BSA in a stirred bath. The method of Genetic Algorithm was also used successfully to determine parameters of these non-linear equations. It was possible to obtain, using the combination of these methods and experimental data, the pore effective diffusivity, D_{eff} , and the Langmuir isotherm equilibrium constant, q_m .

Table 4 - Estimated parameters for $C_0 = 3.0$ *mg/ml*

Test Case	Fixed Parameters	Estimated Parameters	Cost Function Q[mg/ml]^2
1	D_{eff} = 3.72 x 10^{-7} q_m = 302.66		9.8418 x 10^{-3}
2	q_m = 302.66	D_{eff} = 3.70 x 10^{-7}	9.0883 x 10^{-3}
3		D_{eff} = 4.06 x 10^{-7} q_m = 275.79	7.7862 x 10^{-3}

ACKNOWLEDGMENTS

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

REFERENCES

[1] Denisov, A M., “Inverse problems of absorption dynamics”, Minisymposium on Inverse Problems In Medicine, Engineering and Geophysics, Proc. XXIII Brazilian Congress on Applied and Computational Mathematics, Invited Lecture, Santos, Brazil, pp. 24-32, 2000.

[2] Lugon Jr., J., Silva Neto, A. J. and Santana, C. S., “An inverse mass transfer problem in gas-liquid adsorption in columns of bubble and foam”, Inverse Problems in Engineering Symposium, College Station, Texas, USA, 2001.

[3] Conder, J. R. and Hayek, B. O., “Adsorption kinetics and equilibria of bovine serum albumin on rigid ion-exchange and hydrophobic interaction chromatography matrices in a stirred cell”, *Biochemical Engineering Journal*, Vol. 6, pp. 215-223, 2000.

[4] Horstmann, B. J. and Chase, H. A., “Modelling the affinity adsorption of immunoglobulin G to Protein A Immobilized to Agarose Matrices”, *Chem. Eng. Res. Des.*, Vol. 67, 1989.

[5] Liapis, A. I. and Rippin, D. W. T., “A general model for the simulation of multicomponent adsorption from a finite bath”, *Chemical Engineering Science*, Vol. 32, pp. 619-627, 1977.

[6] Firouztale, E., Scott, A. P., Dalvie, S. K. and von Blohn, G. M., “Experimental and theoretical study of key parameters of adsorption on reverse phase macroporous resins”, AIChE Symposium Series, Vol. 88, n^o 290, pp.191, 1992.

[7] Silva, F. R. C., Pereira, J. A. M., Araújo, M. O. D. and Santana, C. C., “Mass transfer parameters evaluation in protein adsorption on macroporous resin”, *Hungarian Journal of Industrial Chemistry*, Vol. 27, pp. 183-187, 1999.

[8] Blanch, H. W. and Clark, D. S., Biochemical Engineering, Marcel Dekker Inc., 1997.

[9] Saad, Y. and Schultz, M. H., “GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems”, *SIAM J. Sci. Stat. Comput.*, Vol. 3, pp. 856-869, 1986.

[10] Venkatakrishnan, V., “Preconditioned conjugate gradient methods for the compressible Navier-Stokes Equations”, AIAA Journal, Vol. 29, n^o 7, pp. 1092-2000, 1991.

[11] Balay, S., Buschelman, K., Gropp, W. D., Kaushik, D., McInnes, L. C. and Smith, B. F., ”PETSc homepage”, “http://www.mcs.anl.gov/petsc”, 2001.

[12] van Genuchten, M. T. and Gray, W. G.,”Analysis of some dispersion corrected numerical schemes for solution of the transport equation”, *International Journal for Numerical Methods in Engineering*, Vol. 12, pp. 387-404, 1978.

[13] Wall, Matthew., ”GAlib: A C++ Library of Genethic Algorithm Components”, *Documentation Revision B*, 1996.

[14] Beck, J. V., Blackwell, B. and St. Clair Jr., C. R., Inverse Heat Conduction, Wiley, New York, 1985.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

A ONE-DIMENSIONAL INVERSE RADIATIVE TRANSFER PROBLEM WITH TIME-DEPENDENT BOUNDARY CONDITIONS

Nancy I. Alvarez Acevedo
Nilson C. Roberty
Antônio J. Silva Neto[2,1]
[1] Nuclear Engineering Program, COPPE,
Universidade Federal de Rio de Janeiro,
P.O. Box 68509, 21945-470
Rio de Janeiro, RJ, Brazil
nancy@lmn.con.ufrj.br,nilson@lmn.con.ufrj.br,ajsneto@lmn.con.ufrj
[2]Department of Mechanical Engineering and Energy
Instituto Politécnico,IPRJ, Universidade do
Estado do Rio de Janeiro, UERJ,
P.O. Box 97282, 28601-970, Nova Friburgo,
RJ, Brazil
ajsneto@iprj.uerj.br

ABSTRACT

In the present work we present the formulation and the solution of three inverse radiative transfer problems in one-dimensional homogeneous participating media with time-dependente boundary conditions. In the first one a energy balance for the incoming and exit radiation allows the direct determination of the absorption coefficient. The second and third inverse problems consider the simultaneous estimation of the total extinction and scattering coefficients using the source-detector and Levenberg-Marquardt methods respectively. Test case results are presented.

INTRODUCTION

The analysis of inverse problems involving the interaction of particles or radiation, such as neutrons, gamma rays or photons, with a participating medium, in which absorption, emission and scattering takes place, has been used in engineering for non destructive testing and in medicine for diagnosis and therapy.

The mathematical modelling of radiative transfer, in which the phenomena of absorption, emission and scattering are taken into account is usually made using the linearized Boltzmann equation, that is also known as the transport equation. The corresponding inverse problems are called inverse radiative transfer problems.

These inverse problems are basically devoted to the estimation of radiative properties or sources using experimental data related to the radiation intensity at the boundaries and/or the interior of the medium under investigation. Carita Montero et al. [1,2] estimated the absorption coefficient in heterogeneous two dimensional absorbing media using the so called q-ART algorithm. Kauati et al. [2] estimated the scattering and total extinction coefficients in homogeneous one dimensional media using the source-detector methodology. In a recente conference McCormick [9] presented a review of different approaches for solving inverse problems involving particles transport.

Inverse problems are formulated either explicity [2], sometimes not requiring the solution of the direct problem, [6] or implicity [4], usually involving the solution of an optimization problem. One fundamental result in the context of the explicit methods was published by Siewert [7].

Most of the work published so far considers the steady-state transport equation, but as an example of time dependent problems, Duracz and McCormick [8] investigated an inverse radiative transfer problem with collimated pulsed illumination of a slab target.

In the present work we tackle the inverse radiative transfer problem in a one-dimensional homogeneous media whose boundary surfaces are subjected to the incidence of external time dependent radiation.

We first consider the determination of the absorption coefficient using an energy balance for the incoming and exit radiation. Then we present one explicit formulation for the simultaneous estimation of the total extinction and the scattering coefficients.

The explicit method is the source-detector methodology [2].

MATHEMATICAL FORMULATION OF THE DIRECT PROBLEM

A plane-parallel anisotropically scattering slab of thickness L, with transparent boundaries, is subjected to an external time dependent irradiation (see Fig.1).

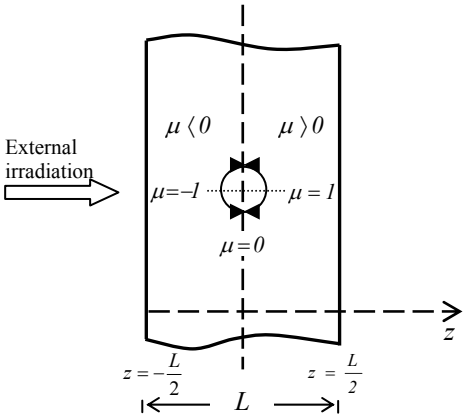


Figure 1. Plane-parallel participating medium subjected to external time dependent irradiation.

The mathematical formulation of the direct radiation problem is given by the linearized Boltzmann equation [9]

$$\frac{1}{c}\frac{\partial\phi(z,\mu,t)}{\partial t}+\mu\frac{\partial\phi(z,\mu,t)}{\partial z}+\sigma_t(z)\phi(z,\mu,t)=\frac{1}{2}\int_{-1}^{1}\sigma_s(z,\mu',\mu)\phi(z,\mu',t)d\mu'+S(z,\mu,t)$$

in $-\frac{L}{2}\langle z\langle\frac{L}{2}$ and $-1\le\mu\ge 1$ (1a)

with boundary conditions

$$\phi(-\frac{L}{2},\mu,t)=\delta(t-t_0)\delta(\mu-\mu_0),\ \ \mu\rangle 0 \quad (1b)$$

$$\phi(\frac{L}{2},\mu,t)=0,\ \ \mu\langle 0 \quad (1c)$$

and initial condition

$$\phi(z,\mu,0)=0$$

in $-\frac{L}{2}\langle z\langle\frac{L}{2}$ and $-1\le\mu\le 1$ for $t=0$ (1d)

where c is the velocity of propagation of the radiation, ϕ is the transient radiation intensity, z is the spatial variable, μ is the direction cosine of the radiation beam with the positive z axis, σ_t is the total extinction coefficient, σ_s is the scattering coefficient, and S is an internal radiation source.

The boundary condition (1b) implies on a short pulse of energy released at time $t=t_0$, and angular direction correspondent to μ_0.

When the geometry, the initial and boundary conditions, the source term, and the radiative properties are known, the radiative transfer problem(1) can be solved for the determination of the radiation intensity at any location $-\frac{L}{2}\le z\le\frac{L}{2}$, for any angular direction $-1\le\mu\le 1$, and at any time, $t\rangle 0$. This is the direct problem.

SOLUTION OF THE DIRECT PROBLEM

Applying the Laplace transform

$$\tilde{\phi}(z,\mu,s)=\Lambda\{\phi(z,\mu,t)\}=\int_0^\infty\phi(z,\mu,t)e^{-st}dt \quad (2)$$

in all terms of Eq. (1a) one obtains

$$\mu\frac{\partial\tilde{\phi}(z,\mu,s)}{\partial z}+\tilde{\sigma}_t(z,s)\tilde{\phi}(z,\mu,s)=\int_{-1}^{1}\sigma_s(z,\mu,\mu')\tilde{\phi}(z,\mu',s)d\mu'+\tilde{S}(z,\mu,s) \quad (3a)$$

where

$$\tilde{\sigma}_t(z,s)=\sigma_s(z)+\sigma_a(z)+\frac{s}{c} \quad (3b)$$

$$\tilde{S}(z,\mu,s)=\int_0^\infty S(z,\mu,t)e^{-st}dt \quad (3c)$$

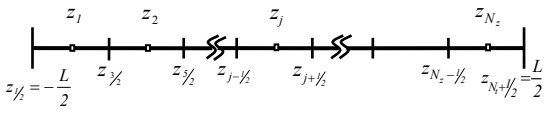
In Eq. (3b) σ_a is the absorption coefficient.

The Laplace transform of the boundary conditions (1b-c) yields

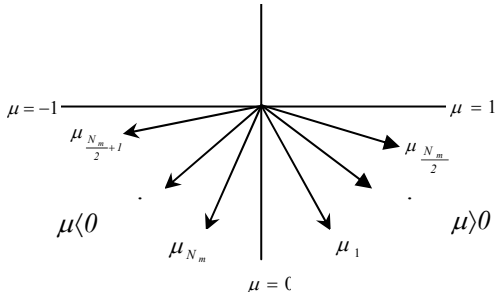
$$\tilde{\phi}(-\frac{L}{2},\mu,s)=\int_0^\infty e^{-st}\phi(-\frac{L}{2},\mu,t)dt=\delta(\mu-\mu_1),\ \ \mu\rangle 0 \quad (3d)$$

$$\tilde{\phi}(\frac{L}{2},\mu,s)=\int_0^\infty e^{-st}\phi(\frac{L}{2},\mu,t)dt=0,\ \ \mu\langle 0 \quad (3e)$$

For the solution of problem (3) we use the discrete ordinates method in which the integral term on the right hand side of Eq. (3a) is replaced by a numerical quadrature. The integro-differential equation (3a) is then replaced by a set of coupled ordinary differential equations (ODE). For each direction of the now discretized angular domain (see Fig. 2b) there is one ODE. To deal wih the spatial dependence of the radiation intensity, the spatial physical domain is replaced by a computational grid, as shown in Fig. 2a, and the derivative on the left hand side of Eq. (3a) is replaced by a finite difference approximation.



(a) Discretization of the spatial domain



(b) Discretization of the angular domain

Figura 2. Computational grid

The scattering coefficient $\sigma_s(z,\mu,\mu')$ can be written as

$$\sigma_s(z,\mu,\mu') = \sigma_s(z)p(\mu,\mu') \tag{4}$$

where the phase function of anisotropic scattering is usually represented in terms of a series of Legendre polynomials as

$$p(\mu,\mu') = \sum_{l=0}^{N_l}(2l+1)f_l P_l(\mu)P_l(\mu_n) \tag{5}$$

and f_l are the expansion coefficients. Using a numerical quadrature for the integral term on the right hand side of Eq. (3a), and using Eqs. (4) and (5), one obtains

$$\frac{1}{2}\int_{-1}^{1}\sigma_s(z,\mu,\mu')\tilde{\phi}(z,\mu',s)d\mu'$$
$$=\frac{1}{2}\sum_{n=1}^{N_m}w_n\sum_{l=0}^{N_l}\sigma_{s_l}(z)(2l+1)P_l(\mu)P_l(\mu_n)\tilde{\phi}(z,\mu_n,s) \tag{6}$$

where w_n are the weights of the quadrature.

Using

$$\left.\frac{\partial\tilde{\phi}(z,\mu,s)}{\partial z}\right|_{j,m} = \frac{\tilde{\phi}(z_{j+1/2},\mu_m,s)-\tilde{\phi}(z_{j-1/2},\mu_m,s)}{(z_{j+1/2}-z_{j-1/2})}$$
$$= \frac{\tilde{\phi}_m^{\,j+1/2}-\tilde{\phi}_m^{\,j-1/2}}{\Delta z_j} \tag{7}$$

$$\left.\tilde{\phi}(z,\mu,s)\right|_{j,m} = \frac{\tilde{\phi}(z_{j+1/2},\mu_m,s)+\tilde{\phi}(z_{j-1/2},\mu_m,s)}{2}$$
$$= \frac{\tilde{\phi}_m^{\,j+1/2}+\tilde{\phi}_m^{\,j-1/2}}{2} = \tilde{\phi}_m^{\,j} \tag{8}$$

where j and m are respectively the indices of the nodes of the spatial and angular grids, with $j=1,2,.....,N_z$ and $m=1,2,...,N_m$, and defining

$$\left.\tilde{q}(z,\mu,s)\right|_{j,m} = \tilde{q}(z_j,\mu_m,s)$$
$$= \frac{1}{2}\sum_{l=0}^{N_l}\sigma_{sl}(z_j)(2l+1)P_l(\mu_m)$$
$$\times\sum_{n=1}^{N_m}w_n P_l(\mu_n)\tilde{\phi}(z_j,\mu_n,s)+\tilde{S}(z_j,\mu_m,s)=\tilde{q}_m^{\,j} \tag{9}$$

Eq. (3a) is replaced by the approximation

$$\mu_m\left(\frac{\tilde{\phi}_m^{\,j+1/2}-\tilde{\phi}_m^{\,j-1/2}}{\Delta z_j}\right)+\tilde{\sigma}_t^{\,j}\left(\frac{\tilde{\phi}_m^{\,j+1/2}+\tilde{\phi}_m^{\,j-1/2}}{2}\right)=\tilde{q}_m^{\,j}$$
$$\text{for } m=1,2,...,N_m \tag{10}$$

To obtain the solution of problem (3) we start at the left boundary of the domain, i.e. $z=-\frac{L}{2}$, using boundary condition (3d). Then we make a sweep forward from $z=-\frac{L}{2}$ towards $z=\frac{L}{2}$ using Eq. (10) written as

$$\tilde{\phi}_m^{\,j+1/2} = \left(\frac{1-\frac{\tilde{\sigma}_t^{\,j}\Delta z_j}{2\mu_m}}{1+\frac{\tilde{\sigma}_t^{\,j}\Delta z_j}{2\mu_m}}\right)\tilde{\phi}_m^{\,j+1/2}+\frac{\tilde{q}_m^{\,j}}{\left(\frac{\mu_m}{\Delta z_j}+\frac{\tilde{\sigma}_t^{\,j}}{2}\right)}$$

for $j=1,2,...,N_z$ and $m=1,2,...,\frac{N_m}{2}$ $(\mu\rangle 0)$ (11)

Note that in the forward sweep we calculate the radiation intensity only for $\mu\rangle 0$.

When the forward sweep is completed we capture the information given by the boundary condition (3e) and we make a backward sweep, i.e. from $z=\frac{L}{2}$ to $z=-\frac{L}{2}$, using Eq.(10) written as

$$\tilde{\phi}_m^{j-1/2}=\left(\frac{1+\frac{\tilde{\sigma}_t^j \Delta z_j}{2\mu_m}}{1-\frac{\tilde{\sigma}_t^j \Delta z_j}{2\mu_m}}\right)\tilde{\phi}_m^{j+1/2}+\frac{\tilde{q}_m^j}{\left(\frac{\tilde{\sigma}_t^j}{2}-\frac{\mu_m}{\Delta z_j}\right)}$$

for $j=N_z,N_z-1,..1$

and $m=\frac{N_m}{2}+1,\frac{N_m}{2}+2,...,N_m$ $(\mu\langle 0)$ (12)

After each cycle of one forward and one backward sweep is completed the value of $\tilde{q}_m^{\,j}$, with $j=1,2,...,N_z$, and $m=1,2,...,N_m$, is updated. The forward and backward sweeps are continued until a convergence criteria is satisfied such as

$$\left|\frac{\tilde{\phi}_m^{j,k+1}-\tilde{\phi}_m^{j,k}}{\tilde{\phi}_m^{j,k}}\right|\langle\varepsilon \quad (13)$$

where k is the iteration counter and ε is a tolerance.

FORMULATION AND SOLUTION OF THE INVERSE RADIATIVE TRANSFER PROBLEMS

We consider here the different formulations for inverse radiative transfer problems. In the first one we determine the absorption coefficient by integrating the original equation in the spatial and angular domains, and by choosing properly the Laplace transform parameter s. The second one is the source-detector formulation [2].

First Inverse Problem – Absorption Coefficient Determination

To keep the presentation simple let's consider Eq. (3a) without the internal source, and then we integrate in the spatial and angular domains,

$$\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-1}^{1}\mu\frac{\partial\tilde{\phi}(z,\mu,s)}{\partial z}d\mu dz=\int_{-\frac{L}{2}}^{\frac{L}{2}}\tilde{\sigma}_t(z,s)\int_{-1}^{1}\tilde{\phi}(z,\mu,s)d\mu dz$$
$$+\frac{1}{2}\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-1}^{1}\left\{\int_{-1}^{1}\sum_{l=0}^{N_l}\sigma_{s_l}(2l+1)P_l(\mu)P_l(\mu')\tilde{\phi}(z,\mu',s)d\mu'\right\}d\mu dz \quad (14)$$

The lef hand side of the Eq. (14) results in

$$LHS=\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-1}^{1}\mu\frac{\partial\tilde{\phi}(z,\mu,s)}{\partial z}d\mu dz$$
$$=\int_{-1}^{1}\mu\left[\tilde{\phi}(\frac{L}{2}\mu,s)-\tilde{\phi}(-\frac{L}{2}\mu,s)\right]d\mu \quad (15)$$

and the right hand side gives

$$RHS=-\int_{-L/2}^{L/2}\left\{\left(\sigma_s+\sigma_a+\frac{s}{c}\right)\int_{-1}^{1}\tilde{\phi}(z,\mu,s)d\mu\right\}dz$$
$$+\int_{-L/2}^{L/2}\left\{\frac{1}{2}\sum_{l=0}^{N_l}(2l+1)\sigma_{s_l}\int_{-1}^{1}P_l(\mu)d\mu\int_{-1}^{1}P_l(\mu')\tilde{\phi}(z,\mu',s)d\mu'\right\}dz$$
$$=-\int_{-L/2}^{L/2}\left\{\left(\sigma_s+\sigma_a+\frac{s}{c}\right)\int_{-1}^{1}\tilde{\phi}(z,\mu,s)d\mu\right\}dz$$
$$+\int_{-L/2}^{L/2}\left\{\frac{\sigma_{s_0}}{2}\int_{-1}^{1}P_0(\mu)d\mu\int_{-1}^{1}P_0(\mu')\tilde{\phi}(z,\mu',s)d\mu'\right\}dz$$
$$+\int_{-L/2}^{L/2}\left\{\frac{1}{2}\sum_{l=1}^{N_l}(2l+1)\sigma_{s_l}\int_{-1}^{1}P_l(\mu)d\mu\int_{-1}^{1}P_l(\mu')\tilde{\phi}(z,\mu',s)d\mu'\right\}dz \quad (16)$$

Knowing that

$$\sigma_s=\sigma_{s_0} \quad (17)$$

$$\int_{-1}^{1}P_l(\mu)d\mu=0 \text{ for } l\geq 1 \quad (18)$$

$$P_0(\mu)=1,\ -1\leq\mu\leq 1 \quad (19)$$

we obtain from Eq. (16), for a homogeneus medium

$$RHS=-\left[\sigma_a+\frac{s}{c}\right]\int_{-\frac{L}{2}}^{\frac{L}{2}}\tilde{\phi}(z,\mu,s)d\mu dz \quad (20)$$

Introducing Eqs. (15) and (20) into Eq. (14), and considering that the medium is homogeneous, i.e. the radiative properties are constant, results

$$-\left[\sigma_a+\frac{s}{c}\right]\int_{-\frac{L}{2}}^{\frac{L}{2}}\tilde{\phi}(z,\mu,s)\,d\mu dz$$
$$=\int_{-1}^{0}\mu\tilde{\phi}(\frac{L}{2},\mu,s)d\mu-\int_{-1}^{0}\mu\tilde{\phi}(-\frac{L}{2},\mu,s)d\mu+$$
$$+\int_{0}^{1}\mu\tilde{\phi}(\frac{L}{2},\mu,s)d\mu-\int_{0}^{1}\mu\tilde{\phi}(-\frac{L}{2},\mu,s)d\mu \qquad (21)$$

The first and the fourth term on the right hand side of Eq. (21) correspond to the incoming heat radiative flux that is considered given as boundary condition, and the second and third terms correspond to the exit radiative heat flux than can be measured.

If we choose properly the Laplace transform parameter, i.e

$$s=-c\sigma_a \qquad (22)$$

we obtain from Eq. (21)

$$\overbrace{\int_{-1}^{0}|\mu|\tilde{\phi}(\frac{L}{2},\mu,s)-\int_{0}^{1}\mu\tilde{\phi}(-\frac{L}{2},\mu,s)d\mu}^{incoming\ radiative\ heat\ flux}$$
$$=\underbrace{\int_{-1}^{0}|\mu|\tilde{\phi}(-\frac{L}{2},\mu,s)-\int_{0}^{1}\mu\tilde{\phi}(\frac{L}{2},\mu,s)d\mu}_{exit\ radiative\ heat\ flux} \qquad (23)$$

Therefore, to obtain one estimate for the absorption coefficient σ_a we have just to solve the transformed radiative transfer problem (3) using different values for the Laplace transform parameter s, and we calculate the exit radiative heat flux. When the exit radiative heat flux becomes equal to the incoming radiative heat fllux as show in Eq. (23) it means that the relation (22) is valid, and as s and c are known, an estimate for σ_a is obtained.

Second Inverse Problem - Explicit Formulation

Here we use the source-detector methodology originally proposed by Roberty [3] which consists on using the original transport equation and an adjoint equation, using reference values for the unknowns we want to determine, in order to assemble a system of nonlinear equations with the unknowns explicitily represented.

In this section we will also focus on the estimation of the total extinction and scattering coefficients.

The Source Problem. To keep the presentation simple we consider the radiative transfer problem (3) without the internal source term, and a homogeneous medium. The source problem is then written as

$$\frac{\partial\tilde{\phi}(z,\mu,s)}{\partial z}+\tilde{\sigma}_t\tilde{\phi}(z,\mu,s)$$
$$=\frac{1}{2}\sum_{l=0}^{N_l}\sigma_{sl}(2l+1)P_l(\mu)\int_{-1}^{1}P_l(\mu')\tilde{\phi}(z,\mu',s)d\mu' \qquad (24a)$$

$$\tilde{\phi}(-\frac{L}{2},\mu,s)=f(\mu),\quad \mu\rangle 0 \qquad (24b)$$

$$\tilde{\phi}(\frac{L}{2},\mu,s)=0,\quad \mu\langle 0 \qquad (24c)$$

where $f(\mu)$ represents the boundary condition at $z=-\frac{L}{2}$, for example $f(\mu)=\delta(\mu-\mu_0)$. Therefore a external collimated source can be located around the boundary surface in different positions.

Detectors can also be placed in different positions around the boundary surface.

The Detector Problem. For each location where a detector is positioned, an adjoint problem is formulated. This formulation is obtained from the source problem by reversing the direction of radiative transfer, i.e. by replacing μ by $-\mu$,

$$-\mu\frac{\partial\tilde{\phi}^*(z,\mu,s)}{\partial z}+\tilde{\sigma}_t^R\tilde{\phi}^*(z,\mu,s)$$
$$=\frac{1}{2}\sum_{l=0}^{N_l}\sigma_{sl}^R(2l+1)P_l(\mu)\int_{-1}^{1}P_l(\mu')\tilde{\phi}^*(z,\mu',s)d\mu' \qquad (25a)$$

$$\tilde{\phi}^*(-\frac{L}{2},\mu,s)=g(\mu),\quad \mu\rangle 0 \qquad (25b)$$

$$\tilde{\phi}^*(\frac{L}{2},\mu,s)=0,\quad \mu\langle 0 \qquad (25c)$$

where $\tilde{\sigma}_t^R$ and σ_{sl}^R are reference values for the unknowns $\tilde{\sigma}_t$ and σ_{sl}, respectively.

By imposing the coincidence of the location of the detector, with those for the source, function $g(\mu)$ represents the measurement that would be obtained by the detector for the strength of the source located at that position,

$$g(\mu) = \eta f(-\mu) \tag{26}$$

where η is the efficiency of the detector.

The Auxiliary Problem. Reversing again the direction of radiation transfer, and defining an auxiliary function

$$\tilde{\phi}^0(z,\mu,s) = \tilde{\phi}^*(z,-\mu,s) \tag{27}$$

we obtain from the detector problem the following auxiliary problem

$$\mu\frac{\partial\tilde{\phi}^0(z,\mu,s)}{\partial z} + \tilde{\sigma}_t^R\tilde{\phi}^0(z,\mu,s)$$
$$= \frac{1}{2}\sum_{l=0}^{N_l}\sigma_{sl}^R(2l+1)P_l(\mu)\int_{-1}^{1}P_l(\mu')\tilde{\phi}^0(z,\mu',s)d\mu' \tag{28a}$$

$$\tilde{\phi}^0(-\frac{L}{2},\mu,s) = \eta f(\mu) \tag{28b}$$

$$\tilde{\phi}^0(\frac{L}{2},\mu,s) = 0 \tag{28c}$$

The Inverse Transport Equation. In the first step on the ITE construction we multiply Eq. (24a) by the adjoint function $\tilde{\phi}^*(z,\mu,s)$, and integrate over the spatial and angular domains, $z = \left[-\frac{L}{2},\frac{L}{2}\right]$ and $\mu = [-1,1]$, respectively. Integrating by parts the resulting equation, introducing Eq. (25a), and replacing the adjoint function $\tilde{\phi}^*(z,\mu,s)$ by the auxiliary function $\tilde{\phi}^0(z,\mu,s)$, we get

$$\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-1}^{1}\left[\tilde{\sigma}_t - \tilde{\sigma}_t^R\right]\tilde{\phi}(z,\mu,s)\tilde{\phi}^0(z,-\mu,s)d\mu dz$$

$$-\frac{1}{2}\sum_{l=0}^{N_l}(2l+1)\left[\sigma_{sl} - \sigma_{sl}^R\right]\int_{-\frac{L}{2}}^{\frac{L}{2}}\int_{-1}^{1}P_l(\mu)\tilde{\phi}(z,\mu,s)$$

$$\times\int_{-1}^{1}P_l(\mu')\tilde{\phi}^0(z,-\mu',s)d\mu' d\mu dz =$$

$$= -\int_{-1}^{1}\mu\tilde{\phi}(\frac{L}{2},\mu,s)\tilde{\phi}^0(\frac{L}{2},-\mu,s)d\mu +$$
$$+\int_{-1}^{1}\mu\tilde{\phi}(-\frac{L}{2},\mu,s)\tilde{\phi}^0(-\frac{L}{2},-\mu,s)d\mu \tag{29}$$

Equation (29) is actually a system of nonlinear equations in which the unknowns $\tilde{\sigma}_t$ and σ_{sl} appear explicitily. For each pair source-detector there is one equation.

On the right hand side of Eq. (29) the boundary conditions come into place, as well as the measured values for the exit radiation intensity (Laplace transformed).

The system of nonlinear equations (29) can be solved using a Gauss-Newton linearization scheme or a more robust algorithm such as the MART (Multiplicative Algebraic Reconstruction Technique) that was developed for tomographic image reconstruction.

RESULTS

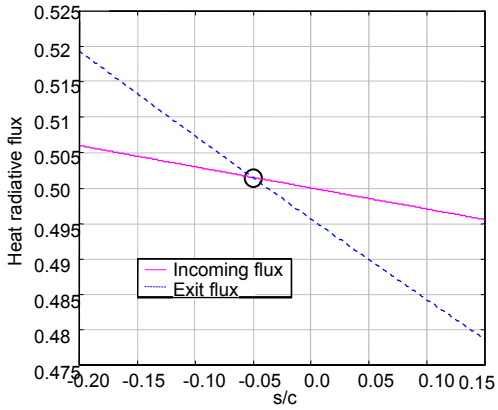
We present in this section results for the two inverse problems described in the previous section.

First Inverse Problems.- Absorption Coefficient Estimation

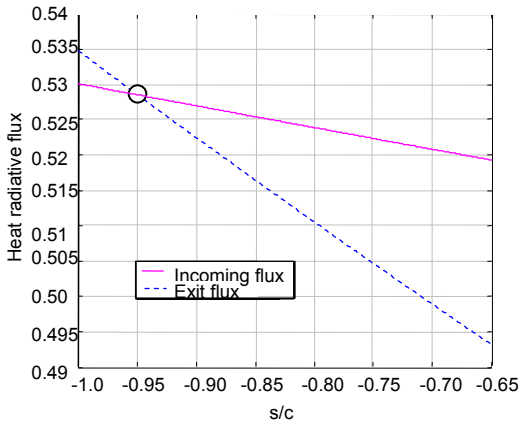
According to Eqs. (22) and (23) we look for the value of the Laplace parameter for which the incoming radiative heat flux is equal to the exit radiative heat flux. Knowing then the value of s/c we obtain from Eq. (22) the value of the absorption coefficiente. Observe that c is given.

In Fig. 3 we show two test case results. Several runs were made considering the radiative properties in the range $0{,}05m^{-1} \le \sigma_a$, $\sigma_s \le 0{,}95m^{-1}$, and slabs with thickness $0{,}1m \le L \le 0{,}5m$, but due to space limitations we are presenting the results for just on case. Here we have a slab of thickness $L = 0{,}1m^{-1}$. The exact value of the absorption coefficient is $\sigma_a = 0{,}05m^{-1}$ ifor the case presented Fig. 3a and $\sigma_a = 0{,}95m^{-1}$ in Fig. 3b.

Plotting the incoming and exit radiative heat fluxes, and finding the value for which they equal each other, the exact values for σ_a are recovered. Observe that here we are considering only noiseless data.



(a) $\sigma_a = 0{,}05 m^{-1}$



(b) $\sigma_a = 0{,}95 m^{-1}$

Figura 3. Absorption coefficient determination. $\sigma_s = 0{,}1 m^{-1}$ and $L = 0{,}1 m^{-1}$

Second Inverse Problem. Results for the source-detector methodology

To demonstrate the feasibility of the solution of the inverse problem with the source-detector methodology, we present in Table 1 the results obtained for the simultaneous estimation of the total extinction and scattering coefficients. The reference values are $\tilde{\sigma}_t^R = \sigma_s^R = 0$, and the exact values are $\sigma_t^{exact} = 15 m^{-1}$ and $\sigma_s^{exact} = 10 m^{-1}$. The initial guess values are taken equal to the reference values, i.e. $\tilde{\sigma}_t^0 = \sigma_t^R$ .and $\tilde{\sigma}_t^0 = \sigma_s^R$

As reported by Kauati et al. [2] when the reference values are far from the exact values it is necessary to restart the algotihm using the reference values taken as the estimated values obtained in the previous cycle of iterations, i.e. $\tilde{\sigma}_t^{R^{k+1}}$ is $\tilde{\sigma}_t^k$ and $\sigma_s^{R^{k+1}} = \sigma_s^k$, where k is the counter of the cycles of iterations. As shown in Table 1, at the end of the first cycle of iterations one gets $\tilde{\sigma}_t^1 = 6{,}5150 m^{-1}$ and $\sigma_s^1 = 6{,}8004 m^{-1}$. These values are then used as the reference values, being obtained with the souce-detector methodology the estimates $\tilde{\sigma}_t^2 = 10{,}659 m^{-1}$ and $\sigma_s^2 = 10{,}390 m^{-1}$. The procedure of sequentially obtaining new estimates by updating the reference values is continued until no change, within a tolerance, is observed.

The results shown in Table 1 were obtained for $s/c = -1 m^{-1}$.

Table 1. Estimated values for σ_t and σ_s at each iteration $s/c = -1 m^{-1}$

Cycle k	Reference values	Iter.	$\sigma_t \ [m^{-1}]$	$\sigma_s \ [m^{-1}]$
1	$\sigma_t^R = 0$	1	1,9508	1,3248
	$\sigma_s^R = 0$	2	6,5150	6,8004
2	$\sigma_t^R = 6{,}5150$	1	10,6586	10,3904
	$\sigma_s^R = 6{,}8004$			
3	$\sigma_t^R = 10{,}658$	1	13,4931	10,9328
	$\sigma_s^R = 10{,}390$			
9	$\sigma_t^R = 14{,}999$	1	15,0004	10,0005
	$\sigma_s^R = 10{,}000$	2	15,0004	9,9999
10	$\sigma_t^R = 15{,}000$	1	14,9999	9,99999
	$\sigma_s^R = 10{,}000$	2	14,9999	10,00000

In Fig. 5 are shown the results for $s/c = 0$ $1m^{-1}$.

In the test cases presented here we have considered synthetic experimental data without noise. In fact there is some level noise present in the data due to the numerial approximation used for the solution of the direct problem and round-off errors.

CONCLUSIONS

The test case results presented demonstrate the feasibility of the three techniques used for solving inverse radiative transfer problems in a one-dimensional homogeneous participating medium using time dependent boundary conditions.

The next step on this research consists on considering noise in the experimental data.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by CAPES–Comissão de Aperfeiçoamento de Pessoal de Nível Superior, CNPq-Consellho Nacional de Desenvolvimento Cientifico e Tecnológico and FAPERJ-Fundação Carlos Chagas Filho de Amparo a Pesquisa do Estado de Río de Janeiro.

REFERENCES

1. R.F. Carita Montero, N.C. Roberty, and A.J. Silva Neto, Absorption coefficients estimation in heterogeneus media using a domain partition consistent with divergent beams, *Inverse Problems in Engineering,* to appear (2002).

2. A.T.Kauati, A.J. Silva Neto and N.C. Roberty, A source-detector methodology for the construction and solution of the one dimensional inverse transport equation. *Inverse Problems in Engineering,* **9**, pp. 45-66 (2001).

3. M.P Mengüç, J. Manickavasagan and D.A. D'Sa, Determination of radiative properties of pulverized coal particles from experiments, *Fuel*, **73** , pp. 513-625 (1994).

4. A.J. Silva Neto and M.N. Özisik, An inverse problem of simultaneous estimation of radiation phase function, albedo and optical thickness, *J. Quant. Spectrosc. Radiat. Transfer,* **53**, pp. 397-409 (1995).

5. N.J. McCormick, Inverse particle transport: Review of recent analysis and applications, *Proc. Mathematics and Computation, Reactor Physics and Environmental Analysis in Nuclear Applications*, Madrid, Spain, **1**, pp. 467-473 (1999).

6. A.J. Silva Neto and N.J. McCormick, An explicit formulation for the inverse radiative transfer problem of single scattering albedo and diffuse reflectivities estimation, *Inverse Problems in Engineering Symposium, Texas*, USA (2001).

7.C.E. Siewert, On a possible experiment to evaluate the one-speed or constant cross-section model of the neutron-transport equation, J. Math. Phys., **19**, pp. 1587-1588 (1978).

8. T. Duracz and N.J. McCormick, Numerical study of the time-dependent radiative-transfer inverse problem, *J. Opt. Soc. Am. A.,* **4**, pp. 1849-1854 (1987).

9. M.N. Özisik, Radiative Transfer and Interactions with Conduction and Convection, John Wiley and Sons. (1973)

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

Estimation of Initial Condition in Heat Conduction by Neural Network

Élcio H. Shiguemori **José Demísio S. da Silva** **Haroldo F. de Campos Velho**

Laboratory for Computing and Applied Mathematics – LAC
National Institute for Space Research – INPE
São José dos Campos, SP, Brazil
[elcio, haroldo, demisio]@lac.inpe.br

ABSTRACT

This paper describes a methodology for using neural networks in an inverse heat conduction problem. Three neural network (NN) models are used to determine the initial temperature profile on a slab with adiabatic boundary condition, given a transient temperature distribution at a given time. This is an ill-posed 1D parabolic inverse problem, where the initial condition has to be estimated. Three neural network models addressed the problem: a feedforward network with backpropagation, radial basis functions (RBF), and cascade correlation. The input for the NN is the temperature profile obtained from a set of probes equally spaced in the one-dimensional domain. The NNs were trainned considering a 5% of noise in the experimental data. The training was performed considering 500 similar test-functions and 500 different test-functions. Good reconstructions have been obtained with the proposed methodology.

NOMENCLATURE

ASE	Average square error
b_k	Bias employed in the NNs
CasCor	Cascate correlation NN
$f(x)$	Unknown nitial condition
$g(x)$	Activation function
w_{ji}	Conection weight of a NN
$N(b_m)$	Norm of the eigenfunction
NN	Neural network
RBF	Radial base function NN
$T(x,t)$	Temperature calculate
$\tilde{T}(x,t)$	Experimental temperature
$X(b_m x)$	eigenfunction
a	Regularization parameter
b_m	Eigenvalue in Eq. (2)
h	Learning rate
m	Random variable
s	Standard deviation
Ω	Space domain
$\Re^+$	Positive real number set

INTRODUCTION

Neural networks have emerged as a new technique to solve inverse problems. This approch was used to identify initial conditions in inverse heat conduction problem on a slab with adiabatic boundary conditions, from transient temperature distribution, obtained at a given time. Three neural networks architectures have been proposed to address the problem: the multilayer perceptron with backpropagation, radial basis functions (RBF), both trained with the whole temperature history mapping, and cascade correlation.

The results are compared with those obtained with non-linear least square approach and standard regularization schemes [1, 2].

Preliminar results using backpropagation and radial basis function neural networks were obtained using whole time history, but with only three different test functions for the learning process [3, 4]. The reconstructions obtained were worse than those identified with regularization techniques. In that strategy two NNs were coupled: the first NN was used for determining the time-period to get the observational data, and another one to find the initial condition itself. That strategy constituted in a novelty in the field, but propably the poor set of test functions for learning step did not permit a good reconstruction. In order to overcome this constrain, 500 functions were used for the learning process in this work. In addition, two groups of test functions were used. In the first group 500 completly different test functions were used, while for the second group 500 similar test-functions were used.

Numerical experiments were carried out with synthetic data with 5% of noise was used to simulate experimental data.

DIRECT HEAT TRANSFER PROBLEM

The direct problem under consideration consists of a transient heat conduction problem in a slab with adiabatic boundary condition, with an initial temperature profile denoted by $f(x)$. Mathematically, the problem can be modeled by the following heat equation

$$\begin{aligned} &\frac{\partial T(x,t)}{\partial t} = \frac{\partial^2 T(x,t)}{\partial x^2} && (x,t) \in \Omega \times \Re^+ \\ &\frac{\partial T(x,t)}{\partial x} = 0 && (x,t) \in \partial\Omega \times \Re^+ \\ &T(x,0) = f(x) && (x,t) \in \Omega \times \{0\} \end{aligned} \qquad (1)$$

where x represents space (the distance between a point in the slab and one of its endpoints), t is the time, $f(x)$ is the initial condition, $T(x,t)$ represents the temporal evolution of the temperature at each point of the slab, and $\partial\Omega$ represents the boundaries of domain Ω. All of these terms are dimensionless quantities and $\Omega = (0,1)$ is the 1D space domain.

The direct problem solution, for a given initial condition $f(x)$ is explicitly obtained using separation of variables, for $(x,t) \in \Omega \times R^+$:

$$T(x,t) = \sum_{m=0}^{+\infty} e^{-b_m^2 t} \frac{1}{N(b_m)} X(b_m, x) \int_0^1 X(b_m, x') f(x') dx' \qquad (2)$$

where $X(b_m,x) = \cos(b_m x)$ are the *eigenfunctions* associated to the problem, $b_m = m\pi$ are the *eigenvalues* and $N(b_m) = \int_\Omega X(b_m, x') f(x') dx'$ represents the *integral normalization* (or the *norm*) [5].

The inverse problem consists in estimating the initial temperature profile $f(x)$ for a given transient temperature distribution $T(x,t)$ at a time t [1].

NEURAL NETWORK ARCHITECTURES

Artificial neural networks (ANN) are made of arrangements of processing elements (*neurons*). The artificial neuron model basically consists of a linear combiner followed by an activation function. Arrangements of such units form the ANNs that are characterized by:

1. Very simple neuron-like processing elements;
2. Weighted connections between the processing elements (where knowledge is stored);
3. Highly parallel processing and distributed control;
4. Automatic learning of internal representations.

ANNs aim to explore the massively parallel network of simple elements in order to yield a result in a very short time slice and, at the same time, with insensitivity to loss and failure of some of the elements of the network. These properties make artificial neural networks appropriate for application in pattern recognition, signal processing, image processing, financing, computer vision, engineering, etc. [6-9].

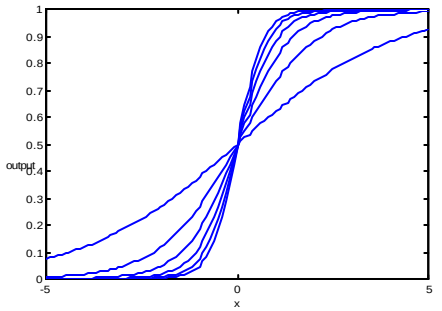
The simplest ANN model is the single-layer Perceptron with a hard limiter activation function, which is appropriate for solving linear problems. This fact prevented neural networks of being massively used in the 1970s [6]. In the 1980s they reemerged due to Hopfield´s paper on recurrent networks and the publication of the two volumes on parallel distributed processing (PDP) by Rumelhart and McClelland [6].

There exist ANN different architectures that are dependent upon the learning strategy adopted. This paper briefly describes the three ANNs used in our simulations: the multilayer Perceptron with backpropagation learning, radial basis functions (RBF), and cacade correlation. Detailed introduction on ANNs can be found in [6] and [9].

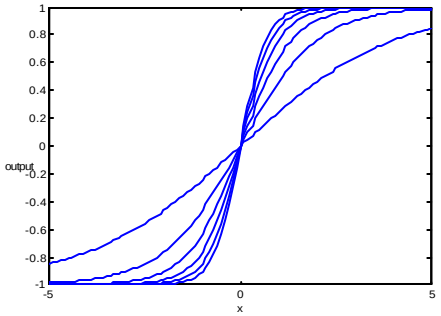
Multilayer perceptrons with backpropagation learning algorithm, commonly referred to as backpropagation neural networks are feedforward networks composed of an input layer, an output layer, and a number of hidden layers, whose aim is to extract high order statistics from the input data [4]. Figure 2 depicts a backpropagation neural network with a hidden layer. Functions g and f provide the activation for the hidden layer and the output layer neurons, respectively. Neural networks will solve nonlinear problems, if nonlinear activation functions are used for the hidden and/or the output layers. Figure 1 shows examples of such functions.

A feedforward network can input vectors of real values onto output vector of real values. The connections among the several neurons (Figure 2) have associated weights that are adjusted during the learning process, thus changing the performance of the network. Two distinct phases can be devised while using an ANN: the training phase (learning process) and the run phase (activation of the network). The training phase consists of adjusting the weights for the best performance of the network in establishing the mapping of many input/output vector pairs. Once

trained, the weights are fixed and the network can be presented to new inputs for which it calculates the corresponding outputs, based on what it has learned.



(a)



(b)

Figure 1: Two activation functions: (a) sigmoid $g(x) = (1+e^{-x})^{-1}$; (b) $g(x) = \tanh(x) = (1-e^{-x})(1+e^{-x})^{-1}$.

The backpropagation training is a supervised learning algorithm that requires both input and output (desired) data. Such pairs permit the calculation of the error of the network as the difference between the calculated output and the desired vector. The weight adjustments are conducted by backpropagating such error to the network, governed by a change rule. The weights are changed by an amount proportional to the error at that unit, times the output of the unit feeding into the weight. Equation 3 shows the general weight correction according to the so-called the delta rule

$$\Delta w_{ji} = \eta d_j y_i \qquad (3)$$

where, d_j is the local gradient, y_i is the input signal of neuron j, and η is the learning rate parameter that controls the strength of change.

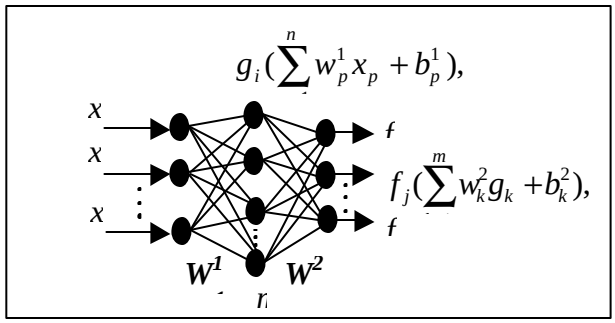


Figure2: The backpropagation neural network with one hidden layer.

Radial basis function networks are feedforward networks with only one hidden layer. They have been developed for data interpolation in multidimensional space. RBF nets can also learn arbitrary mappings. The primary difference between a backpropagation with one hidden layer and an RBF network is in the hidden layer units. RBF hidden layer units have a receptive field, which has a center, that is, a particular input value at which they have a maximal output. Their output tails off as the input moves away from this point. The most used function in an RBF network is a Gaussian (Figure 3).

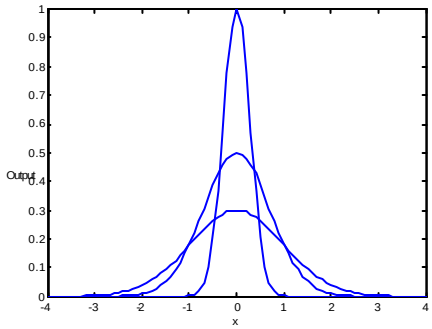


Figure 3: Gaussian for three differents variances.

RBF networks require the determination of the number of hidden units, the centers, and the sharpness (standard deviation) of their Gaussians. Generally, the centers and standard deviations are decided on first by examining the vectors in the training data. The output layer weights are then trained using the Delta rule.

The training of RBF networks can be conducted: (1) on classification data (each output representing one class), and then used directly as classifiers of new data; and (2) on pair of points $(x, f(x))$ of an unknown function f, and then used

to interpolate. The main advantage of RBF networks relies on the fact that one can add extra units with centers near elements of the set of input data, which are difficult to classify.

Like Backpropagation networks, RBF networks can be used for processing time-varying data and many other applications.

The third ANN used in the present paper is the cascade correlation. This NN permits dynamically to find out the appropriated number of neurons, begining with just the input and output layers, with all the neurons fully interconnected (there is no hidden layer). The weights on these connections are determined using a conventional learning. Next, new neurons are considered sequentially, and weights between the candidate units and the inputs are selected to maximize the correlation between the activation of the neuron(s) and the residual error of the net. Once a neuron is selected, its weights on the inputs are frozen, and are not subsequently changed when considering new neurons. Additional neurons are applied until a specified small error is reached.

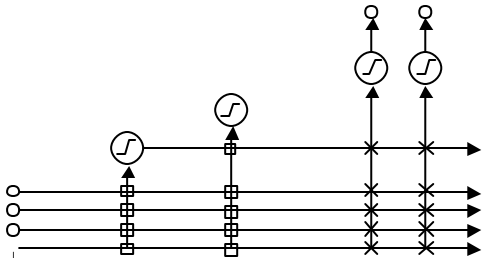


Figure 4: Casccade correlation network with 2 hidden layers. The symbol ⊘ denotes a neuron.

Figure 4 shows a cascate correlation (CasCor) network into which two candidate neurons have been implemented. These neurons use a conventional activation function, as shown in Figure 2. Each open box in the figure represents a weight that is trained only once (when the neuron is a candidate) and then is frozen. But the cross marks represent weights that are repeatedly changed as the network evolves. Note that the structure of the network is such that the inputs remain directly connected to the outputs, but also some information is filtered through the neurons. The direct input to output connection can handle the linear portion of the mapping, while the non-linearities are addressed by the neurons.

NEURAL NETWORK FOR DETERMINING THE INITIAL CONDITION

Artificial neural networks have two stages in their application, firstly the learning and activation steps. During the learning step, the weights and bias corresponding at each connection are adjusted to some reference examples. For activation, the output is obtained based on the weights and bias computed in the learning phase. A supervised learning was used for all NN architectures.

The numerical experiment for inverse problem is based on two test functions, the triangular function

$$f(x)=\begin{cases}2x & x\in[0,0.5]\\ 2(1-x) & x\in(0.5,1]\end{cases} \tag{6}$$

and semi-triangular function

$$f(x)=\begin{cases}0.55 & 0\le x\le 0.2\\ 8/3\,x & 0.2<x<0.5\\ -28/5\,x+23/5 & 0.5<x<0.75\\ 2/9 & 0<x\le 1\end{cases}. \tag{7}$$

The experimental data (measured temperatures at a time $\tau>0$), which intrinsically contains errors in the real world, is obtained by adding a random perturbation to the exact solution of the direct problem, such that

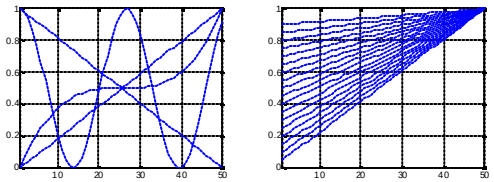
$$\tilde{T}=T_{\text{exact}}+sm \tag{8}$$

where s is the standard deviation of the errors and m is a random variable taken from a Gaussian distribution, with zero mean and unitary variance. Twin numerical experiments were performed. In the first one, a noiseless *observational* data were employed (s=0). For the second numerical experiment was carried out using 5% of noise (s=0.05).

For the NNs, the training sets are constituted by synthetic data obtained from the forward model, i.e., profile of a *measure* points from probes spread in the space domain. Two different data sets were used. The first data set is the profiles obtained from 500 similar functions (see examples in Figure 5a). The second one is that obtained with 500 no-similar functions (Figure 5b). Similar functions are those belonging to the same class (linear function class, trigonometric function class, such as sine functions with

different amplitude and/or phase, and so on). No-similar functions are those complete different, in which each one belonging a distinct class.

Figure 5 shows a set of functions used in the leraning stage, appling non-similar (Fig. 5a) and similar functions (Fig. 5b).



(a) (b)

Figure 5: Sample of test functions for training: (a) non-similar functions; (b) similar functions.

The *activaction* is a regular test used for checking out the NN performance, where a function belonging to the test function set is applied *to activate* (to run) the NN. Good activations were obtained for all three NNs for observational data with noise and noiseless data, for similar and non-similar test function sets (not shown). In the activation test the NN trainned with similar data were systematicaly better than the training with non-similar functions (not shown too), with and without noise in the data. A summary of the training results for three NNs is presented in Table 1.

Table 1: Training results for the neural networks used for initial condition reconstruction.

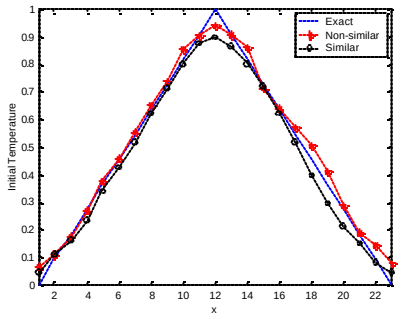
Multi-layer perceptron				
Data	Noise	Hidden Neurons	Training Epochs	ASE
Non-similar	0%	25	150000	0,0487
Similar	0%	20	50000	0,0127
Non-similar	5%	20	300000	0,0694
Similar	5%	20	50000	0,0144
Radial base function				
Data	Noise	Hidden Neurons	Training Epochs	ASE
Non-similar	0%	20	50000	0,0576
Similar	0%	20	50000	0,0095
Non-similar	5%	20	300000	0,0873
Similar	5%	20	50000	0,0123
Cascate correlation				
Data	Noise	Hidden Neurons	Training Epochs	ASE
Non-similar	0%	10	300000	0,0746
Similar	0%	05	63000	0,0230
Non-similar	5%	02	2000	0,1389
Similar	5%	05	63000	0,0318

Nevertheless, the activation test is an important procedure, indicating the permormance of a NN, the effective test is defined using a function (initial condition) that did not belong to the training function set. This action is called the *generalization* of the NN. Functions as expressed by Eqs. (6) and (7) did not belong to the function set in the traning step.

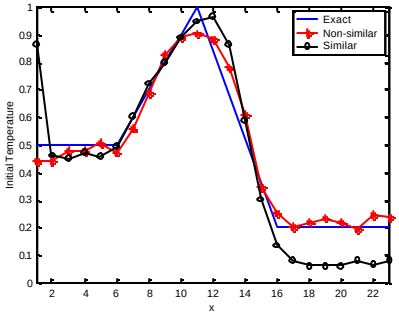
Figures 6, 7, and 8 show the initial condition reconstruction for noiseless experimental data, and Table 2 presents the Average Square Error (ASE) for three NNs used in this paper. Differently from the results for the activation test, reconstruction using non-similar functions were better than estimation with similar functions.

Table 2: Activation results for the noiseless experimental data.

Multi-layer perceptron		
***f*(*x*)**	Data	ASE
Triangular	Non-similar	0.0136
Triangular	Similar	0.0139
Semi-triangular	Non-similar	0.0246
Semi-triangular	Similar	0.1599
Radial base function		
***f*(*x*)**	Data	ASE
Triangular	Non-similar	0.0065
Triangular	Similar	0.0079
Semi-triangular	Non-similar	0.0275
Semi-triangular	Similar	0.0498
Cascate correlation		
***f*(*x*)**	Data	ASE
Triangular	Non-similar	0.0253
Triangular	Similar	0.0845
Semi-triangular	Non-similar	0.0471
Semi-triangular	Similar	0.1462



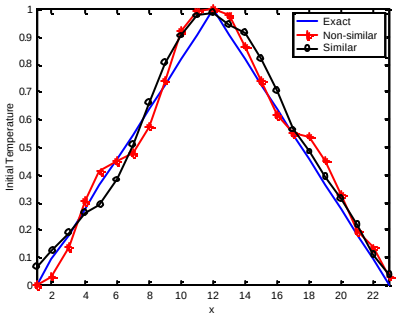
(a)



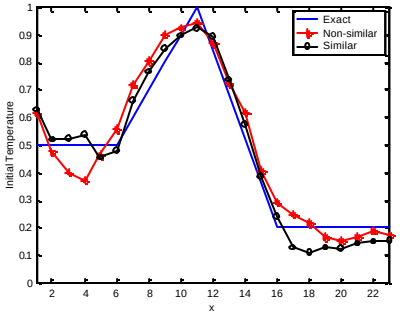
(b)

Figure 6: Reconstruction using multi-layer perceptron NN with noiseless data.

The worse reconstructions for noiseless data were obtained using CasCor-NN (see table 2 and Figures 6, 7, and 8), and the best identifications were obtained using RBF-NN. However, good initial condition identifications were gotten with three NN architectures.

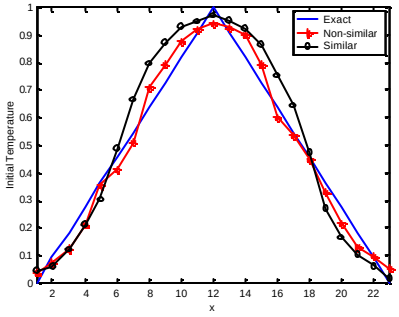


(a)

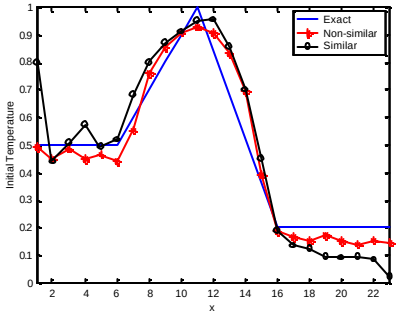


(b)

Figure 7: Reconstruction using radial base function NN with noiseless data.



(a)



(b)

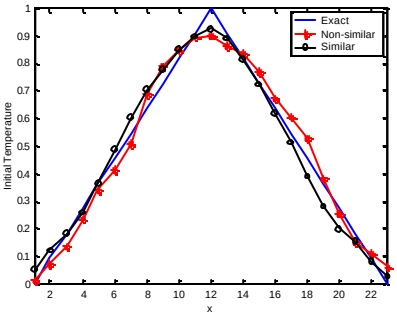
Figure 8: Reconstruction using cascate correlation NN with noiseless data.

Table 3: Activation results for the experimental data with 5% of noise.

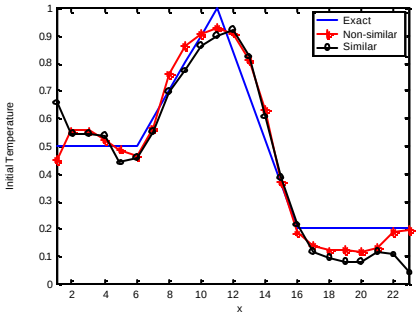
Multi-layer perceptron		
***f*(*x*)**	Data	ASE
Triangular	Non-similar	0.0227
Triangular	Similar	0.0210
Semi-triangular	Non-similar	0.0621
Semi-triangular	Similar	0.0786
Radial base function		
***f*(*x*)**	Data	ASE
Triangular	Non-similar	0.0308
Triangular	Similar	0.0331
Semi-triangular	Non-similar	0.0563
Semi-triangular	Similar	0.0396
Cascate correlation		
***f*(*x*)**	Data	ASE
Triangular	Non-similar	0.0384
Triangular	Similar	0.0947
Semi-triangular	Non-similar	0.0486
Semi-triangular	Similar	0.1294

Real tests for inverse problems must be performed using some level of noise in the synthetic experimental data. As it was mentioned, the real experimental data were simulated corrupting the output data from direct problem with Gaussian white noise, see Eq. (8).

As with our numerical experiment with noiseless data, the identification of the initial condition was effective for all NNs used here.

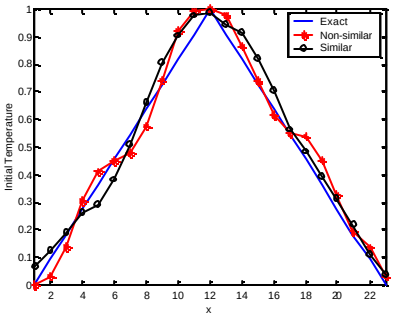


(a)

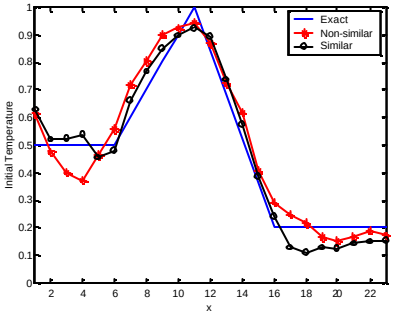


(b)

Figure 9: Reconstruction using multi-layer perceptron NN with 5% of noise.



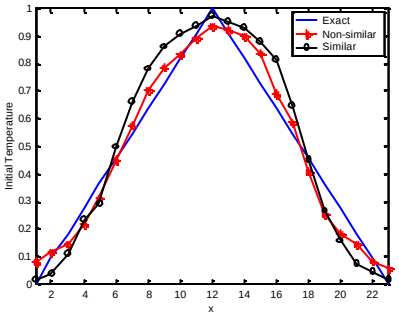
(a)



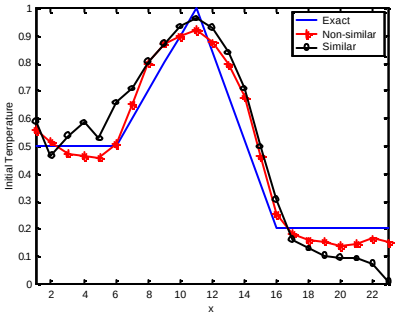
(b)

Figure 10: Reconstruction using radial base function NN with 5% of noise.

Figures 9, 10 and 11 show the reconstructions for multi-layer perceptron, RBF and CaCor NNs. Table 3 presents the ASE for two test function in the generalization. As expected the reconstruction with data contamined with noise was worse than noiseless data. But, the NNs were robust in the identification with noise in the experimental data.



(a)



(b)

Figure 11: Reconstruction using cascate correlation NN with 5% of noise.

FINAL REMARKS

Three architectures of neural networks were studied in the reconstruction of the initial condition of a heat conduction problem. All of NNs were effective for solving this inverse problem. Different from previous results [3, 10], reconstructions are comparable with those obtained with regularization methods [2], even for data containing noise. However, the NN do not remove the inherent ill-posedness of the inverse problem.

The initial condition estimation problem is a harder inverse problem than the identification of boundary condition in heat transfer [11-13].

An interesting remark is the result for the activation test, where the training with similar functions produced better identification than non-similar function. However, reconstructions using non-similar functions were systematicly better for the generalization, except in only one case: the estimation of semi-triangular function by RBF-NN with 5% of noise (Table 3).

The worse estimation was obtained with CasCor-NN. A future work could be done using the strategy adopted by Hidalgo and Gómez-Treviño [14]. To accommodate large amounts of noise, they added a regularization term to the least squares objective function of the neural network.

Processing with NNs is a two step process: training and activation. After the training phase, the inversion with NNs is much faster than the regularization methods, and the NNs do not need a mathematical model to simulate the forward model. In addition, NNs is an intrinsicly parallel algorithm. Finally, NNs can be implemented in hardware devices, the neurocomputers, becoming the inversion processing faster than NNs emulated by software.

REFERENCES

1. W.B. Muniz, H.F. de Campos Velho and F.M. Ramos (1999): A comparison of some inverse methods for estimating the initial condition of the heat equation, *J. Comp. Appl. Math*., **103**, 145 (1999).

2. W.B. Muniz, F.M. Ramos and H.F. de Campos Velho, Entropy- and Tikhonov-based regularization techniques applied to the backwards heat equation, *Comp. Math. Appl*., **40**, 1071 (2000).

3. E. Issamoto, F.T. Miki, J.I. da Luz, J.D. da Silva, P.B. de Oliveira, H.F. de Campos Velho, An Inverse Initial Condition Problem in Heat Conductions: A Neural Network Approach, *Braz. Cong. Mech. Eng.* (COBEM), Proc. in CD-ROM - paper code AAAGHA, 238 (1999), Unicamp, Campinas (SP), Brasil.

4. F.T. Miki, E. Issamoto, J.I. da Luz, P.B. de Oliveira, H. F. de Campos Velho, J.D. da Silva, A Neural Network Approach in a Backward Heat Conduction Problem, *Braz. Conf. Neural Networks*, Proc. in CD-ROM - paper code 0008, 019 (1999), São José dos Campos (SP), Brasil.

5. M.N. Özisik, *Heat Conduction*, Wiley Interscience, 1980.

6. S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan. New York, 1994.

7. C-T Lin and G. Lee, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall, New Jersey, 1996.

8. M. Nadler and E.P. Smith, *Pattern Recognition Engineering*, John Wiley & Sons, New York, 1993.

9. L.H. Tsoukalas and R.E. Uhrig, *Fuzzy and Neural Approaches in Engineering*, John Wiley and Sons, New York, 1997.

10. F.T. Miki, E. Issamoto, J.I. da Luz, P.B. de Oliveira, H. F. de Campos Velho, J.D. da Silva, An inverse heat conduction problem solution with a neural network approach, *Bulletim of the Braz. Soc. for Comp. Appl. Math*. (SBMAC), 2000, available in the internet: www.sbmac.org.br/publicacoes .

11. J. Krejsa, K.A. Woodbury, J.D. Ratliff, M. Raudensky (1999): Assessment of strategies and potential for neural networks in the IHCP, *Inverse Probl. Eng*., **7,** 197 (1999).

12. K.A. Woodbury, Neural networks and genetic algorithms in the solution of inverse problems, *Bulletim of the Braz. Soc. for Comp. Appl. Math*. (SBMAC), 2000, available in the internet: www.sbmac.org.br/publicacoes .

13. E.H. Shiguemori, F.P. Harter, H.F. de Campos Velho, J.D.S. da Silva, Estimation of boundary conditions in heat transfer by neural networks, *Braz. Cong. on Comp. and Appl. Math*., Belo Horizonte (MG), Brazil, 559 (2001).

14. H. Hidalgo, E. Gómez-Treviño, Application of Constructive Learning Algorithms to the Inverse Problem, *IEEE T. Geosci. Remote*, **34**, 874 (1996).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

A Parametric Study of a New Regularization Operator: the Non-extensive Entropy

Élcio H. Shiguemori **Haroldo F. de Campos Velho**
José Demísio S. da Silva **Fernando M. Ramos**
Laboratory for Computing and Applied Mathematics – LAC
National Institute for Space Research – INPE
São José dos Campos, SP, Brazil
[elcio, haroldo, demisio, fernando]@lac.inpe.br

ABSTRACT

Tsallis' non-extensive entropy is used as a regularization operator. The parameter "*q*" has a central role in Tsallis' thermostatiscs formalism. Here, several values of q are investigated in an inverse problem, using $q<1$ and $q>1$ ($q=1$ is the standard Boltzmann-Gibbs-Shannon formulation). Estimation of initial condition in heat conduction problem is employed as the inverse test problem. Two methods are studied for determining the regularization parameter, the maximum curvature for the L-curve, and the Morozov's discrepancy principle.

NOMENCLATURE

$E(.)$	Square error in L_2-norm
$f(x)$	Unknown nitial condition
$J_a[\tilde{T}, f]$	Objective function
q	Non-extensivity parameter
$R(.)$	Residue in L_2-norm
S_q	Non-extensive entropy
$T(x,t)$	Temperature calculate
$\tilde{T}(x,t)$	Experimental temperature
a	Regularization parameter
m	Random variable
s	Standard deviation
Λ	Space domain
$W(f)$	Regularization function
$\Re^+$	Positive real number set

INTRODUCTION

Inverse problems belongs the class of *ill-posed* problems. In these type of problems existence, uniqueness and stability of their solutions cannot be ensured.

The regularization theory appeared in the 60's as a general method for solving inverse problems [1]. In this approach, the non-linear least square problem is associated with a regularization term (*a priori* or additional information), in order to obtain a well-posed problem. A well-known regularization technique was proposed by Tikhonov [1], where smoothness of the unknown function is searched. Similarly to Tikhonov's regularization, the maximum entropy formalism searchs for *global* regularity and yields the smoothest reconstructions which are consistent with the available data. The maximum entropy principle was first proposed as a general inference procedure by Jaynes [2] on the basis of Shannon's axiomatic characterization of the amount of information [3]. This principle has successfully been applied to a variety of fields [4].

A non-extensive statistics has been proposed for entropy [5, 6]. Recently, the non-extensive entropic form (S_q) was used as a new regularization operator [7], using only $q=0.5$. The q parameter plays a central role in the Tsallis' thermostatistics, in which $q=1$ the Boltzmann-Gibbs-Shannon's entropy is recovered. In the present study several values for q were used for the *non-extensive* entropic regularization term.

This new regularization operator was tested for estimating initial condition in heat conduction problem [8, 9]. Synthetic data with Gaussian white noise corruption were used to simulate experimental data.

Two methods were investigated for determining the regularization parameter: the Morozov's discrepancy principle [10], and the maximum curvature scheme of the curve relating smoothness versus fidelity, inspired in Hansen's geometrical criterion [11].

NON-EXTENSIVE ENTROPY AS A NEW REGULARIZATION OPERATOR

A non-extensive form of entropy has been proposed by Tsallis [5] by the expression

$$S_q(\boldsymbol{p}) = \frac{k}{q-1}\left(1 - \sum_{i=1}^{N_p} p_i^q\right) \qquad (1)$$

where p_i is a probability, and q is a free parameter. In thermodynamics the parameter k is known as the Boltzmann's constant. Similarly as in the mathematical theory of information, k=1 is considered in the regularization theory. Tsallis' entropy reduces to the the usual Boltzmann-Gibbs-Shanon formula

$$S(p) = -k\sum_{i=1}^{N_p} p_i \ln p_i \qquad (2)$$

in the limit $q \to 1$.

As for extensive form of entropy, the equiprobability condition produces the maximum for the non-extensive entropy function, and this condition is expressed as

$$\left(S_q\right)_{\max} = \frac{1}{1-q}\left(N_p^{1-q} - 1\right) \quad (N_p \geq 1) \qquad (3)$$

where, in the limit $N_p \to \infty$, S_q diverges if $q \leq 1$, and saturates at $1/(q\text{-}1)$ if q>1 [6].

The equiprobability condition leads the regularization function defined by the operator $S_q(\boldsymbol{p})$, given by Eq. (1), to search the smoothest solution for the of the unknown vector $\boldsymbol{p}$.

The parameter q has a central role in Tsallis' thermostatiscs, and it is called the *non-extensivity parameter* – see Eqs. (A.5) and (A.6) in the Appendix. For q<5/3, the standard central limit theorem applies, implying that if p_i is written as a sum of M random independent variables, in the limit case $M \to \infty$, the probability density function for p_i in the distribution space is the *normal* (Gaussian) distribution [6]. However, for $5/3 < q < 3$ the Levy-Gnedenko's central limit theorem applies, resulting for $M \to \infty$ the Levy distribution as the probability density function for the random variable p_i . The index in such Levy distribution is $g=(3\text{-}q)/(q\text{-}1)$ [6].

The non-extensive approach has been used in many different applications, such as in a certain type of anomalous diffusion process [6], as well as the statistical model for data from turbulent flow [12, 13] and from financial market [12]. According to Plastino and Plastino [14], the first experimental confirmation of Tsallis' non-extensive formalism is the Boghosian's approach of the two dimensional pure electron plasma [15]. Some properties of the thermostatistics formalism are described in the Appendix.

In the present study the maximum non-extensive entropy principle of zeroth-order is applied as a regularization function for estimating the initial condition in conductive heat transfer. The forward problem is presented in the next Section.

FORWARD HEAT CONDUCTION PROBLEM

The direct (forward) problem consists of a transient heat conduction problem in a slab with adiabatic boundary condition and initially at a temperature denoted by $f(x)$. The mathematical formulation of this problem is given by the following heat equation

$$\begin{aligned} &\frac{\partial T(x,t)}{\partial t} = \frac{\partial^2 T(x,t)}{\partial x^2} && (x,t) \in \Lambda \times \Re^+ \\ &\frac{\partial T(x,t)}{\partial x} = 0 && (x,t) \in \partial\Lambda \times \Re^+ \\ &T(x,0) = f(x) && (x,t) \in \Lambda \times \{0\} \end{aligned} \qquad (4)$$

where $T(x,t)$ (temperature), $f(x)$ (initial condition), x (spatial variable) and t (time variable) are dimensionless quantities and Λ=[0,1]. The set of partial differential equations is solved by using a central finite difference approximation for space variable $O(Dx^2)$, and explit Euler method for numerical time integration $O(Dt)$ [14].

This problem has been used for testing different methodologies in inverse problems [8, 9, 17-19], and it is badly conditioned problem [8].

INVERSE ANALYSIS

Typically, for (inverse) ill-posed problems, that existence, uniqueness and stability of their solutions cannot be ensured. In some sense, a solution can be formulated in such a way that existence and uniqueness can be relaxed, but this solution can still be unstable under the presence of noise in the experimental data. Hence, it requires some regularization technique, i.e., the incorporation in the inversion procedure of some available information about the true solution. Following the Tikhonov's approach [1], a regularized solution is obtained by choosing the function f^* that minimizes the following functional

$$J_a[\tilde{T}, f] = \left\| \tilde{T} - T(f) \right\|_2^2 + a\Omega[f] \qquad (5)$$

where $\tilde{T} = \tilde{T}(x,t)$ is the experimental data ($t = t$), $T(f)$ is the temperature computed from the forward model at time t, $\Omega[f]$ denotes the regularization term given by expression (1) with k=1, a is the regularization parameter, and $\|\cdot\|_2$ is the 2-norm.

The regularization parameter a is chosen by two methods: numerically, assuming that a bound d (or the 'statistics') of the measurement error is known, i.e., $\left\| T_{\text{exact}} - \tilde{T} \right\|_2 \le d$ - this numerical procedure is based on Morozov's discrepancy principle [10]; graphically, finding out the point of maximum curvature in the curve $\Omega[f_a] \times \left\| \tilde{T} - T(f_a) \right\|_2^2$, a type of L-curve [11, 20].

Optimization Algorithm

The optimization problem is iteratively solved by the quasi-newtonian optimizer routine from the NAG Fortran Library [21], with variable metrics. This algorithm is designed to minimize an arbitrary smooth function subject to constraints (simple bound, linear or nonlinear constraints), using a sequential programming method.

This routine has been successfully used in several previous works: in geophysics, hydrologic optics, and meteorology.

ESTIMATING INITIAL CONDITION BY THE NON-EXTENSIVE ENTROPY

The numerical experiment with the non-extensive entropy is based on two test functions, the triangular function

$$f(x) = \begin{cases} 2x & x \in [0, 0.5] \\ 2(1-x) & x \in (0.5, 1] \end{cases} \qquad (6)$$

and semi-triangular function

$$f(x) = \begin{cases} 0.55 & 0 \le x \le 0.2 \\ 8/3\,x & 0.2 < x < 0.5 \\ -28/5\,x + 23/5 & 0.5 < x < 0.75 \\ 2/9 & 0 < x \le 1 \end{cases} . \qquad (7)$$

The experimental data (measured temperatures at a time τ>0), which intrinsically contains errors, is obtained by adding a random perturbation to the exact solution of the direct problem, such that

$$\tilde{T} = T_{\text{exact}} + sm \qquad (8)$$

where s is the standard deviation of the errors and m is a random variable taken from a Gaussian distribution, with zero mean and unitary variance. All tests were carried out using 5% of noise (s=0.05).

It is important to observe that the spatial grid consists of 101 points (N_x=101), and the time-integration is performed up to t=0.01. The residue $R(f_a)$ and the error $E(f_a)$ are defined by

$$R(f_a) = \left\| \tilde{T} - T(f_a) \right\|_2^2 \qquad (9)$$

$$E(f_a) = \left\| f_a - f_{\text{exact}} \right\|_2^2 \qquad (10)$$

If we effectively want to apply some kind of regularization, which means a>0 in Eq. (5), then the discrepancy principle - an *a-posteriori* parameter choice rule - implies that a suitable regularized solution can be obtained. Since the spatial resolution is N_x=101, the *optimum* a is reached for $R(f_*) \cong N_x s^2 = 0.2525$. Table 1 shows the square diference term $R(f_*)$ obtained for different values of a, and the *optimum* value is pointed out for each case in bold font.

A set of tables (Tables 1, 2, 3 and 4) presents the least squares (or residual) term $R(f_\alpha)$ and the error $E(f_a)$ between the approximated (or calculated) solution f_a and the exact solution f_{exact} obtained for two values of q=0.5 and 2.0, from a family of regularization non-extensive entropy functions S_q, for different values of a. The value of a satisfying the discrepancy principle is pointed out for each test functions by bold font. Regularized solutions are presented in Figures 1 and 2.

The parameter vector was always subjected to the following simple bounds: $1.2 \ge f_k \ge -0.2$ for the triangular test function, and $1.2 \ge f_k \ge 0$ for the semi-triangular test function, with $k = 1, 2, \ldots, N_x$.

Table 1: Determining a by Morozov's criterion: q=0.5, for triangular test function.

a	$R(f_a)$	$E(f_a)$
0.0001	0.1853	2.7298
0.0003	0.1856	0.5388
0.0010	0.1861	0.3443
0.0285	**0.2525**	**0.3994**
0.0999	0.7728	0.8684

Table 2: Determining a by Morozov's criterion: q=0.5, for semi-triangular test function.

a	$R(f_a)$	$E(f_a)$
0.0001	0.1851	4.9359
0.0003	0.1854	0.6740
0.0010	0.1856	0.2113
0.0346	**0.2525**	**0.2400**
0.0999	0.6687	0.7207

Table 3: Determining a by Morozov's criterion: q=2.0, for triangular test function.

a	$R(f_a)$	$E(f_a)$
0.0001	0.1852	2.3979
0.0003	0.1854	0.3578
0.0010	0.1856	0.1383
0.0414	**0.2525**	**0.2370**
0.0999	0.5174	0.5977

Table 4: Determining a by Morozov's criterion: q=2.0, for semi-triangular test function.

a	$R(f_a)$	$E(f_a)$
0.0001	0.1851	4.0920
0.0003	0.1854	0.4455
0.0010	0.1856	0.1807
0.0419	**0.2525**	**0.4379**
0.0999	0.5131	0.8651

Figures 1a-1d show the estimation of four q-values for triangular initial condition, where the regularization parameter was computed by Morozov's principle. The best reconstruction was found by q=2.5, but good reconstructions were obtained for other values of q too. Figures 2a-2d depict the reconstructions for the semi-triangular test function, showing good reconstructions for all values of q.

Another criterion for finding the regularization parameter was also investigated, and it is based on the maximum curvature in the L-curve [11]. Figures 3a-3b show the L-curve for triangular test function using q=0.5 and 2.0, respectively. The L-curve for semi-triangular test function is displayed in Figures 4a-4b. The regularization parameter a is chosen at the corner of the L-curve.

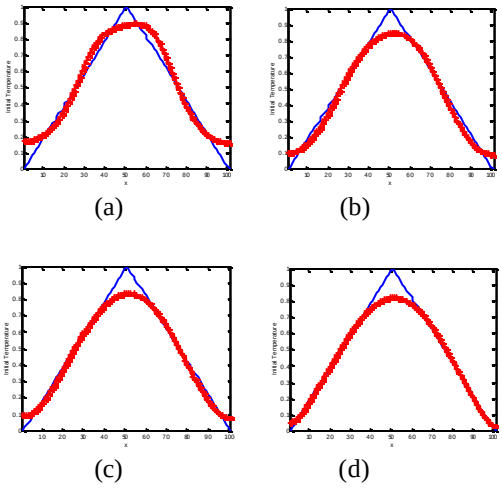


Figure 1: Reconstructions with 5% of noise, with a determined by Morozov's principle: (a) q=0.5; (b) q=1.5; (c) q=2.0; (d) q=2.5.

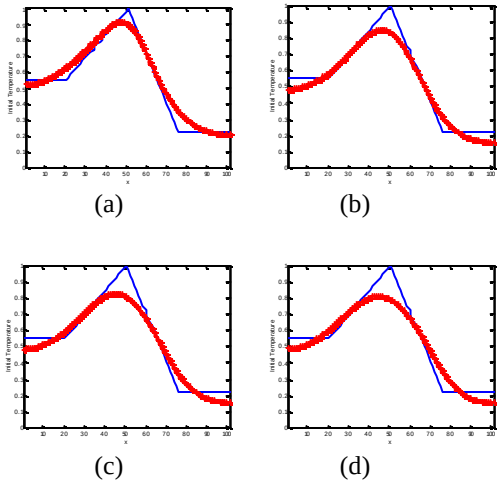
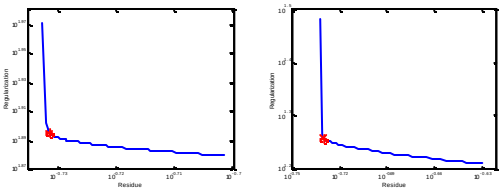


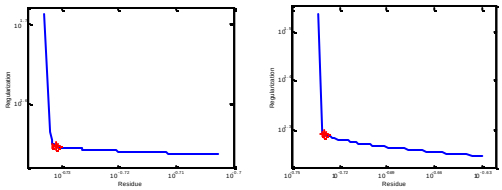
Figure 2: Reconstructions with 5% of noise, with a determined by Morozov's principle: (a) q=0.5; (b) q=1.5; (c) q=2.0; (d) q=2.5.

Reconstructions using a as computed by Hansen's criterion [11] are shown in Figures 5a-5d for triangular test function, and Figures 6a-6d for semi-triangular test function. The best reconstruction was obtained using q=2.5, and the worst for q=0.5.



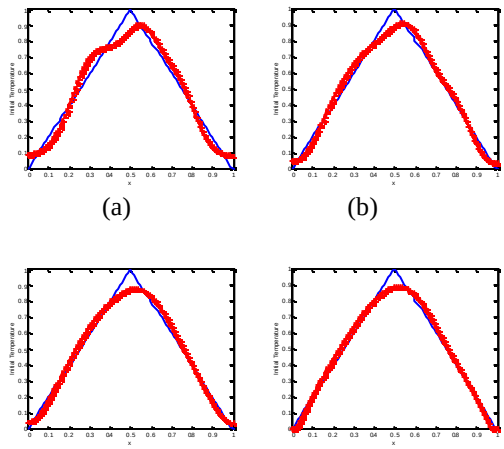
(a) (b)

Figures 3: L-curve for triangular test function. (a) q=1.5; (b) q=2.5.



(a) (b)

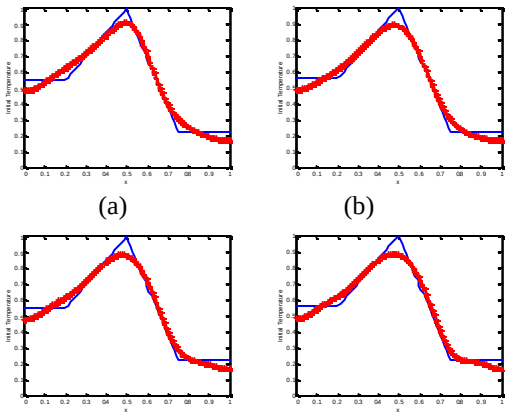
Figures 4: L-curve for semi-triangular test function. (a) q=1.5; (b) q=2.5.



(a) (b)

(c) (d)

Figure 5: Reconstructions for triangular test function, with a determined by Hansen's criterion: (a) q=0.5; (b) q=1.5; (c) q=2.0; (d) q=2.5.



(a) (b)

(c) (d)

Figure 6: Reconstructions for semi-triangular test function, with a determined by Hansen's criterion: (a) q=0.5; (b) q=1.5; (c) q=2.0; (d) q=2.5.

CONCLUSION

The implict strategy and the regularization techniques adopted in this work yield good results in reconstructing the initial condition of the heat equation. The Morozov's discrepancy principle was efficient to estimate the regularization parameter in the analyzed cases. The Hansen's criterion also produced good estimates for the Lagrange multiplier a. According to the Tables 1-4, the value of regularization parameter calculated by the Morozov's principle tends to over-estimate the value of a. However, looking at Table 5, it is possible to realize that the a computed by the Hansen's criterion is closer to the $a_{optimum}$. Nevertheless, sometimes is hard to obtain the L-curve. One case specially difficult was found to q=2.5, for some values of a was not possible to obtain a solution (no convergence). A possible solution to convergence would be to change the deterministic optimizer by a stochastic one.

The new regularization technique used in this work, namely the maximum non-extensive entropy of zeroth order, worked very well for the backwards heat equation for all parameter q tested. The choice q=0.5 made by Rebollo-Neira et al. [7] was linked to a previous result of physical relevance related to the relaxation of two-dimensional turbulence [15]. There is no reason to restrict the regularization operator S_q at

q=0.5. Actually, the worse reconstructions for the triangular test function were obtained using $S_{0.5}$!

Table 5: Regularization parameter computed by test functions using Morozov's principle and Hansen's criterion.

Triangular Test Function				
	q=0.5	q=1.5	q=2.0	q=2.5
α_{Morozov}	0.0285	0.0231	0.0414	0.0579
α_{Hansen}	0.0011	0.0008	0.0040	0.0040

Semi-triangular Test Function				
	q=0.5	q=1.5	q=2.0	q=2.5
α_{Morozov}	0.0346	0.0231	0.0419	0.0583
α_{Hansen}	0.0040	0.0010	0.0040	0.0050

Table 6 shows that an appropriated choice of the non-extentive parameter q can improve the reconstruction. The worked example is a linear model. However, the regularization theory can also be applied to a non-linear problems [1].

Hence, one can say that the regularization technique worked effectively for this extremly ill-posed problem and, in this sense, the exploration of new entropy-based methodologies seems to be necessary and promising. Higher order of the maximum non-extensive entropy deserves to be examinated [22-24] too, as well as other criteria for determining the regularization parameter [25].

Table 6: Estimation error computed with the regularization parameter found by Morozov's principle and Hansen's criterion.

Triangular Test Function		
	$E(f_{\alpha-\text{Morozov}})$	$E(f_{\alpha-\text{Hansen}})$
q=0.5	0.3994	0.3490
q=1.0	0.3151	0.1959
q=1.5	0.2599	0.1697
q=2.0	0.2370	0.1530
q=2.5	0.2561	0.1302

Semi-triangular Test Function		
	$E(f_{\alpha-\text{Morozov}})$	$E(f_{\alpha-\text{Hansen}})$
q=0.5	0.2400	0.1676
q=1.0	0.4056	0.1958
q=1.5	0.3205	0.1745
q=2.0	0.4379	0.1800
q=2.5	0.5959	0.2041

REFERENCES

1. A.N. Tikhonov and V.I. Arsenin, *Solutions of Ill-posed Problems*, John Wiley & Sons, 1977.

2. E.T. Jaynes, Information theory and statistical mechanics, *Phys. Rev.*, **106**, 620 (1957).

3. C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Univ. of Illinois Press, 1949.

4. C. R. Smith and W. T. Grandy (Eds.), *Maximum-Entropy and Bayesian Methods in Inverse Problems*, in Fundamental Theories of Physics, Reidel, Dordrecht, 1985.

5. C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, *J. Statistical Physics*, **52**, 479 (1988).

6. C. Tsallis, Nonextensive statistics: theoretical, experimental and computational evidences and connections, *Braz. J. Phys.*, **29**, 1 (1999).

7. L. Rebollo-Neira, J. Fernandez-Rubio and A. Plastino, A non-extensive maximum entropy based regularization method for bad conditioned inverse problems, *Phys. A*, **261**, 555 (1998).

8. W.B. Muniz, F.M. Ramos and H.F. de Campos Velho, Entropy- and Tikhonov-based regularization techniques applied to the backwards heat equation, *Comp. Math. Appl.*, **40**, 1071 (2000).

9. W.B. Muniz, H.F. de Campos Velho and F.M. Ramos (1999): A comparison of some inverse methods for estimating the initial condition of the heat equation, *J. Comp. Appl. Math.*, **103**, 145 (1999).

10. V.A. Morozov, On the solution of functional equations by the method of regularization, *Soviet Math. Dokl.*, **7**, 414 (1966).

11. P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev.*, **34**, 561 (1992).

12. F.M. Ramos, C. Rodrigues Neto, R.R. Rosa, L.D. Abreu Sa, M.J.A. Bolzan, Generalized thermostatistical description of intermittency and nonextensivity in turbulence and financial markets, *Nonlinear Anal.-Theor.*, **23**, 3521 (2001).

13. F.M. Ramos, R.R. Rosa, C. Rodrigues Neto, M.J.A. Bolzan, L.D.A. Sa and H.F. de Campos Velho, Nonextensive statistics and three-dimensional fully developed turbulence, *Phys. A*, **295**, 250 (2001).

14. A. Plastino and A.R. Plastino, Tsallis's entropy and Jaynes' information theory formalism, *Braz. J. Phys.*, **29**, 50 (1999).

15. B.M.R. Boghosian, Thermodynamic description of two-dimensional Euler turbulence using Tsallis' statistics, *Phys. Rev.*, **E 53**, 4754 (1996).

16. J.D. Hoffmann, *Numerical Methods for Engineers and Scientists*, McGraw-Hill, 1993.

17. E. Issamoto, F.T. Miki, J.I. da Luz, J.D. da Silva, P.B. de Oliveira, H.F. de Campos Velho, An inverse initial condition problem in heat conduction: a neural network approach, *Braz. Cong. on Mechanical Eng.* (COBEM), Campinas (SP), Brazil, Proc. in CD-ROM - paper code AAAGHA, page 238 in the Abstract Book (1999).

18. F.T. Miki, E. Issamoto, J.I. da Luz, P.B. de Oliveira, H. F. de Campos Velho, J.D. da Silva (1999): A neural network approach in a backward heat conduction problem, *Braz. Conference on Neural Networks*, Proc. in CD-ROM, paper code 0008, pages 19-24, São José dos Campos (SP), Brazil (1999).

19. E.H. Shiguemori, H.F. de Campos Velho, J.D.S. da Silva, Estimation of initial condition in heat conduction by neural network, in this Proceedings - ICIPE-2002, May 26-31, Angra dos Reis (RJ), Brazil.

20. T. Reginska, A regularization parameter in discrete ill-posed problem, *SIAM J. Sci. Comput.*, **17**, 740 (1996).

21. E04UCF routine, NAG Fortran Library Mark 17, Oxford, UK (1995).

22. F.M. Ramos and H.F. de Campos Velho, Reconstruction of geoelectric conductivity distributions using a minimum first-order entropy technique, *2nd International Conference on Inverse Problems on Engineering*, Le Croisic, France, **vol. 2** , 199 (1996).

23. H.F. de Campos Velho and F.M. Ramos, Numerical inversion of two-dimensional geoelectric conductivity distributions from eletromagnetic ground data, *Braz. J. Geophys.*, **15**, 133 (1997).

24. F.M. Ramos, H.F. de Campos Velho, J.C. Carvalho and N.J. Ferreira, Novel approaches on entropic regularization, *Inverse Problems*, **15**, 1139 (1999).

24. M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics, 1999.

APPENDIX – SOME PROPERTIES FOR NON-EXTENSIVE THERMOSTATISCS

For N_p micro-states with propabilities $p_i \geq 0$, $i=1, \ldots, N_p$:

A1: Non-extensive entropy:

$$S_q(p) = \frac{k}{q-1}\left(1 - \sum_{i=1}^{N_p} p_i^q\right) \tag{A.1}$$

A2: q-expectation of an observable:

$$O_q \equiv \langle O \rangle_q = \sum_{i=1}^{N_p} p_i^q \, o_i \tag{A.2}$$

Properties

1. If $q \to 1$:

$$S_1 = k \sum_{i=1}^{N_p} p_i \ln p_i \tag{A.3}$$

$$O_1 = \sum_{i=1}^{N_p} p_i \, O_i \tag{A.4}$$

2. Non-extensive entropy is positive: $S_q \geq 0$.

3. Non-extensivity

$$S_q(A+B) = S_q(A) + S_q(B) + (1-q) S_q(A) S_q(B) \tag{A.5}$$

$$O_q(A+B) = O_q(A) + O_q(B) + (1-q)\left[O_q(A) S_q(B) + O_q(B) S_q(A)\right] \tag{A.6}$$

4. Max S_q under constrain $O_q = \sum_i p_i^q \, e_i$ (canonical ensemble):

$$p_i = \frac{1}{Z_q}\left[1 - b(1-q) e_i\right]^{1/(1-q)} \tag{A.7}$$

where the e_i is the energy of state i, $O_q = U_q$ is the non-extensive form to the internal energy, and the normalization factor Z_q (partition function), for $1<q<3$, is given by

$$Z_q = \left[\frac{p}{b(1-q)}\right]^{1/2} \frac{\Gamma\left[(3-q)/2(q-1)\right]}{\Gamma\left[1/(q-1)\right]} \tag{A.8}$$

For $q = 1$ yields

$$p_i = e^{-b e_i} / Z_1 \quad . \tag{A.9}$$

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSE RADIATIVE TRANSFER PROBLEMS IN TWO-DIMENSIONAL PARTICIPATING MEDIA

Mariella J. Berrocal Tito[1,2]
Nilson Costa Roberty[1]
[1]Nuclear Engineering Program, PEN /COPPE, Universidade Federal do Rio de Janeiro, UFRJ, P.O. Box 68509, 21945-970, Rio de Janeiro, RJ, Brazil,
nilson@lmn.con.ufrj.br
[2]Facultad de Ciencias, Universidad Nacional de Ingeniería, Av. Tupac Amaru s/n, Lima - Perú

Antônio J. Silva Neto[3,1]
[3]Department of Mechanical Engineering and Energy, InstitutoPolitécnico, IPRJ, Universidade do Estado do Rio de Janeiro, UERJ. P.O. Box 97282, 28601-970, Nova Friburgo, Brazil, ajsneto@lmn.con.ufrj.br, ajsneto@iprj.uerj.br
Jorge Bravo Cabrejo[4]
[4]Facultad de Física, UNMSM, Lima - Perú. Av. Venezuela s/n, d220008@unmsm.edu.pe

ABSTRACT

In the present work inverse radiative transfer problems in two-dimensional heterogeneous participating media are considered. An implicit formulation is used in which the cost functional of the squared residues between calculated and measured exit radiation intensities is minimized. The Levenberg - Marquardt method, a gradient-based minimization algorithm, is used, and therefore at every iteration of the iterative procedure the solution of the direct problem is required. For the solution of the direct problem, which is modeled by the linearized Boltzmann equation, the discrete ordinates method with a finite difference approximation for the spatial derivatives is used. The inverse problems considered are related to the estimation of the absorption and anisotropic scattering coefficients, as well as the estimation of source terms.

INTRODUCTION

The analysis of inverse radiative transfer problems has attracted the attention of many researchers due to several relevant applications in different areas such as oil industry [1], computerized tomography [2], biomedical engineering [3], and ocean optics [4] among others.

Most of the work published on inverse radiative transfer problems involves the determination of radiative properties [5,6], or intensity and/or location of internally distributed sources [7], and are usually related to one dimensional Cartesian geometry.

These inverse problems are either formulated explicitly [6], sometimes not requiring the solution of the direct problem, or implicitly, usually involving the solution of an optimization problem [5].

The mathematical modeling of the interaction of radiation with participating media is usually done with the linearized Boltzmann equation, also known as transport equation. Several methods have been developed for the solution of the transport equation [8], and in recent years an increasing interest has been observed towards the discrete ordinates method [9], that was originally proposed by Chandrasekhar [10], and has been intensively applied in radiative heat transfer by Fiveland [11] and Truelove [12].

In the present work we consider inverse radiative transfer problems in two-dimensional participating media. An implicit formulation is used, and the inverse problem is solved as an optimization one, in which the standard squared residues functional is minimized with a gradient based method. The direct problem of radiative transfer is solved with the discrete ordinates methods to deal with the angular dependence of the radiation intensity, and a finite difference approximation for the spatial derivatives.

Test case results are presented for homogeneous and heterogeneous two-dimensional media. The estimation of a constant internal source, asymmetry factor, as well as total extinction and scattering coefficients are considered. Synthetic experimental noisy, and noiseless data were used in the simulations. A sensitivity analysis is made, and then it is shown that better estimates are obtained when only detectors located at the region of highest sensitivity are used.

MATHEMATICAL FORMULATION OF THE DIRECT PROBLEM

The linearized Boltzmann equation whichthat is used in the mathematical modeling of the interaction of radiation with a participating medium, i.e a medium in which absorption, emission and scattering takes place, is written for a gray medium, in the steady state, and no spectral dependency case as

$$\underset{\sim}{\Omega}\cdot\nabla I(\underset{\sim}{r},\underset{\sim}{\Omega}) + s_t(\underset{\sim}{r})I(\underset{\sim}{r},\underset{\sim}{\Omega}) = \int_{S^2} s_s(\underset{\sim}{r},\underset{\sim}{\Omega}'\cdot\underset{\sim}{\Omega})I(\underset{\sim}{r},\underset{\sim}{\Omega}')d\underset{\sim}{\Omega}' + S(\underset{\sim}{r},\underset{\sim}{\Omega}) \qquad (1)$$

where I it is the intensity of the radiation, $\underset{\sim}{r}$ represents a point in the domain, $\underset{\sim}{\Omega}$ is the direction of propagation of the radiation, s_t is the coefficient of total extinction (absorption + out scattering), s_a is the absorption coefficient, s_s is the scattering coefficient, $\underset{\sim}{\Omega}'$ is the direction of incident radiation that is scattered at point $\underset{\sim}{r}$ in the direction $\underset{\sim}{\Omega}$, and S is the source term. The integral term on the right hand side of Eq. (1) is known as the in scattering term. In the discrete ordinates method the integral in the in scattering term is replaced by a numerical quadrature, and the integro - differential equation (1) is replaced by a set of coupled ordinary differential equations (ODE). For each direction of the now discretized angular domain there is one ODE.

In a two-dimensional cartesian domain

$$\underset{\sim}{r} = x\cdot\hat{e}_x + y\cdot\hat{e}_y \qquad (2)$$

$$\underset{\sim}{\Omega} = m_l\cdot\hat{e}_x + x_l\cdot\hat{e}_y \qquad (3)$$

$$\nabla = \frac{\partial}{\partial x}\hat{e}_x + \frac{\partial}{\partial y}\hat{e}_y \qquad (4)$$

where *l=1,2..Lo* represents the directions in the discretized angular domain, and m_l and x_l are direction cosines with respect to the axis represented by the unit vectors $\hat{e}_x$ and $\hat{e}_y$, respectively. In a rectangular domain, $0 \le x \le Lx$, $0 \le y \le Ly$ an approximation for Eq. (1) obtained with the discrete ordinates method can be written as [8]

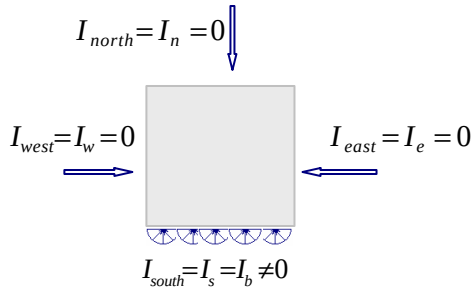
$$m_l\frac{\partial I_l}{\partial x} + x_l\frac{\partial I_l}{\partial y} + s_t I_l = \sum_{m=1}^{Lo}\frac{s_s}{4p}w_m f_{lm} I_m + S_l$$
$$\text{for } l = 1,2,..,Lo,\quad -1 \le m \le 1\quad -1 \le x \le 1 \qquad (5)$$

where w_m are the weights of the quadrature, and f_{lm} represents the phase function of anisotropic scattering. In the present work we consider only isotropic scattering or the linear anisotropic scattering in which cases the phase function can be represented by [11]

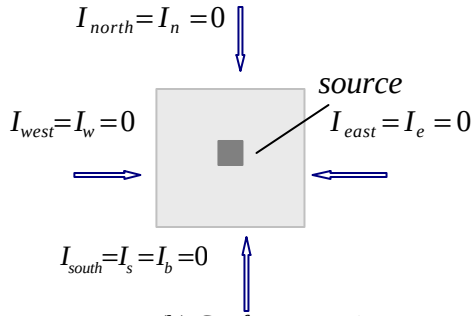
$$f_{lm} = 1 + A_1[m_l m_m + x_l x_m] \qquad (6)$$

where A_1 is the asymmetry factor, and it varies in the range; $-1 \le A_1 \le 1$. Isotropic scattering is modeled by A_1=0, and forward and backward scattering are given by positive and negative values respectively.

Several sets of boundary conditions can be used for the two-dimensional rectangular domain considered here, but in the present work we will focus only on the two cases represented in Fig. 1.



(a) Configuration 1



(b) Configuration 2

Figure 1. Boundary Conditions.

In configuration 1, Fig 1a, the lower surface of the medium, y = 0, is subjected to external isotropic incident radiation. In configuration 2, Fig 1b , there is no incidence of external radiation on any of the boundaries of the domain, but there is a radiation source inside the domain. When the geometry, boundary conditions, intensity and location of internal sources, and radiative properties are known the radiative transfer problem (5) can be solved and the radiation intensity can be determined for any point *(x,y)* and direction $\underline{\Omega}_l$, l= 1,2,..,Lo. This is known as the direct problem.

SOLUTION OF THE DIRECT PROBLEM

In order to solve problem (5) with the boundary conditions represented in Fig. 1 we have considered a finite difference method in which the spatial physical domain is replaced by a computational grid, Fig. 2, and the derivatives on the left hand side of Eq. (5) are replaced by finite differences,

$$m_l \frac{\Delta I_{l(i,j)}}{\Delta x} + x_l \frac{\Delta I_{l(i,j)}}{\Delta y} + s_{t(i,j)} I_{l(i,j)} = \\ = \sum_{m=1}^{Lo} \frac{s_{s(i,j)}}{4p} w_m f_{lm(i,j)} I_{m(i,j)} + S_{l(i,j)} \\ l = 1,2,..,Lo \ , \qquad (7)$$

with

$$\Delta I_{l(i,j)x} = Ie_{l(i,j)} - Iw_{l(i,j)} \qquad (8a)$$

$$\Delta I_{l(i,j)y} = In_{l(i,j)} - Is_{l(i,j)} \qquad (8b)$$

where *e,w,n*, and *s* represent respectively the east, west, north, and south boundaries, of the computational cell represented by the grid node *(i,j)*. The values of $I_{l(i,j)}$ can be approximated by, [8]

$$I_{l(i,j)} = g \ In_{l(i,j)} + (1- g \ ) Is_{l(i,j)} \\ = g \ Ie_{l(i,j)} + (1- g \ ) Iw_{l(i,j)} \qquad (9)$$

being usually used g = 0,5.

Following the signs of the direction cosines, the angular domain can be divided in four quadrants as follows: (I) m_l >0 and x_l >0; (II) m_l <0 and x_l >0; (III) m_l >0 and x_l <0; and (IV) m_l <0 and x_l <0.

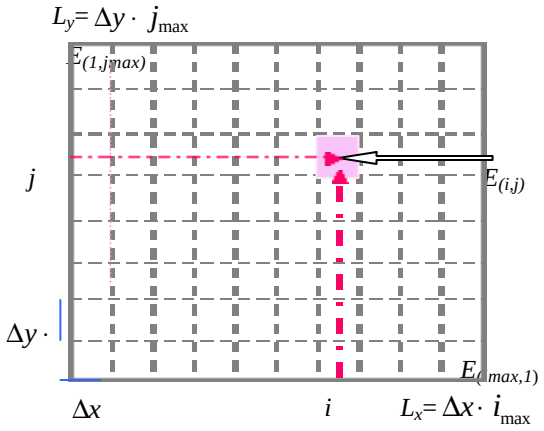


Figure 2 Computational grid with a regular spacing, $\Delta x = \Delta y$.

To obtain the solution for the radiative transfer problem in the rectangular domain sweeps across the grid nodes are made. First if we consider a sweep from $x = 0$ to $x = L_x$ and $y = 0$ to $y = L_y$, we are able to capture the values of the boundary conditions at $x = 0$ and $y = 0$, and propagate them inside the domain. This is made considering the first quadrant (see Fig. 3). A second sweep is made from x = L_x to x = 0 and y = 0 to y = L_y, capturing the boundary conditions at $x = L_x$ and $y = 0$, and propagating them inside the domain. This is related to the second quadrant A similar procedure is then used for the third and fourth quadrants. In summary, to complete once cycle four sweeps are necessary in order to capture the boundary conditions at the four boundary surfaces and cover the four quadrants.

For the first quadrant, m_l >0 and x_l >0, it is convenient to write from Eq. (9),

$$In_{l(i,j)} = \frac{1}{g} [I_{l(i,j)} - (1- g \ ) Is_{l(i,j)}] \qquad (10a)$$

$$Ie_{l(i,j)} = \frac{1}{g} [I_{l(i,j)} - (1- g \ ) Iw_{l(i,j)}] \qquad (10b)$$

From Eqs. (7), (8) and (10) one obtains

$$I_{l(i,j)} = \frac{\Delta y \cdot m_l \cdot Iw_{l(i,j)} + \Delta x \cdot x_l \cdot Is_{l(i,j)} + \Delta x \cdot \Delta y \cdot g \cdot Q_{l(i,j)}}{\Delta y \cdot m_l + \Delta x \cdot x_l + s_{t(i,j)} \cdot \Delta x \cdot \Delta y \cdot g} \qquad (11)$$

where

$$Q_{l(i,j)} = \sum_{m=1}^{Lo} \frac{s_{s(i,j)}}{4p} w_m f_{lm(i,j)} I_{m(i,j)}$$

with $i=1,2,..,i_{max}$, $j=1,2,..,j_{max}$. (12)

For the second quadrant $m_l<0$ and $x_l>0$. In this case it is convenient to write from Eq. (9)

$$Iw_{l(i,j)} = \frac{1}{(1-g)}\left[I_{l(i,j)} - g\ Ie_{l(i,j)}\right] \quad (13)$$

and use also Eq. (10a) for $In_{l(i,j)}$. From Eqs. (7), (8), (10a) and (13) one obtains

$$I_{l(i,j)} = \frac{\frac{-g}{(1-g)}\cdot\Delta y\cdot m_l\cdot Ie_{l(i,j)} + \Delta x\cdot x_l\cdot Is_{l(i,j)} + \Delta x\cdot\Delta y\cdot g\cdot Q_{(i,j)}}{\frac{-g}{(1-g)}\cdot\Delta y\cdot m_l + \Delta x\cdot x_l + s_{t(i,j)}\cdot\Delta x\cdot\Delta y\cdot g}$$

with $i=i_{max},i_{max}-1,..,1$; and $j=1,2,..j_{max}$ (14)
where $Q_{l(i,j)}$ is given by Eq. (12).

The equations to be used for the other two sweeps, considering quadrants 3 and 4, can be derived in a similar way, yielding

Quadrant III : $m_l>0$ and $x_l<0$

$$I_{l(i,j)} = \frac{\Delta y\cdot m_l\cdot Iw_{l(i,j)} - \frac{g}{(1-g)}\cdot\Delta x\cdot x_l\cdot In_{l(i,j)} + \Delta x\cdot\Delta y\cdot g\cdot Q_{(i,j)}}{\frac{-g}{(1-g)}\cdot\Delta y\cdot m_l + \Delta x\cdot x_l + s_{t(i,j)}\cdot\Delta x\cdot\Delta y\cdot g} \quad (15)$$

Quadrant IV: $m_l<0$ and $x_l<0$

$$I_{l(i,j)} = \frac{\frac{-g}{(1-g)}\cdot\Delta y\cdot m_l\cdot Iw_{l(i,j)} - \frac{g}{(1-g)}\cdot\Delta x\cdot x_l\cdot In_{l(i,j)} + \Delta x\cdot\Delta y\cdot g\cdot Q_{(i,j)}}{\frac{-g}{(1-g)}\cdot\Delta y\cdot m_l - \frac{g}{(1-g)}\cdot\Delta x\cdot x_l + s_{t(i,j)}\cdot\Delta x\cdot\Delta y\cdot g} \quad (16)$$

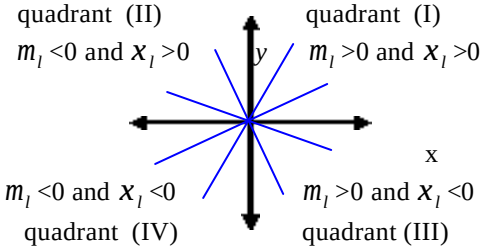


Figure 3. Four quadrants of angular domain.

MATHEMATICAL FORMULATION AND SOLUTION OF THE INVERSE PROBLEM

When the geometry, the boundary conditions the intensity and/or location of internal sources, or radiative properties, are unknown, but experimental data related to the radiation intensity inside and/or at the boundary of the domain are available one can attempt to obtain an estimate for the unknown quantities. This is known as the inverse problem.

In the present work we consider the inverse problem of radiative properties and internal source intensity estimation, using measurements of the radiative heat flux as shown schematically in Fig.4. Detectors $D1^k$ in system 1, Fig. 4a, are therefore able to measure the positive half range heat flux

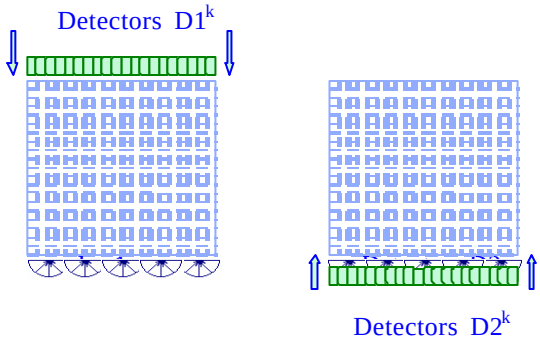
$$q^{+}_{(k,j\,max)} = D1^k = \sum_{\substack{l=1 \\ x>0}}^{Lo} w_l\cdot In_{l(i,j\,max)}\cdot x_l$$

$k=1,2,..,i_{max}$ (17a)

and detectors $D2^k$ in system 2, Fig 4b, are able to measure the negative half range heat flux

$$q^{-}_{(k,1)} = D2^k = \sum_{\substack{l=1 \\ x<0}}^{Lo} w_l\cdot Is_{l(i,1)}\cdot|x_l|$$

$k=1,2,..,i_{max}$ (17b)



(a) system 1 north boundary (b) system 2, south boundary

Figure 4. Location of detectors.

For the solution of the inverse radiative transfer problems under analysis here we use an implicit formulation that consists on the minimization of the squared residues functional

$$R = \sum_{k=1}^{K}\left[q^k_{cal}(\underline{z}) - q^k_{meas}\right]^2 = \vec{F}^T\vec{F} \quad (18)$$

where q^k_{cal} and q^k_{meas} are respectively the calculated and measured values of the radiative heat flux, $\underline{z}$ is the vector of unknowns, K is the total number of experimental data available, k indicates the position of the detectors and the

relationship of parity of the solution q^k_{cal} and q^k_{meas}, and

$$F^k(\underline{z}) = q^k_{cal}(\underline{z}) - q^k_{meas} \qquad (19)$$

Due to space limits we will not make presentation of the formulation of the Levenberg Marquardt method. Details can be found in [5].

RESULTS

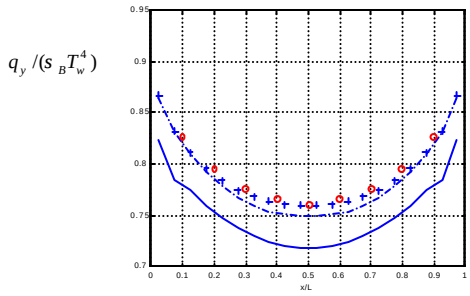
Validation of the Direct Problem Solution

The algorithm developed for the computation of the radiation intensities was validated making a comparison of the results obtained with those published for a radiative transfer problem [8,11].

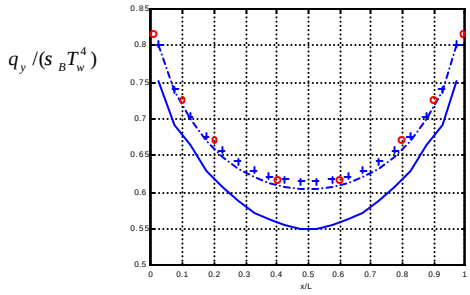
Consider a two-dimensional absorbing and emitting square medium whose inferior (south) boundary is kept at temperature T_w, and the other boundaries are kept at T =0. In Fig. 5 are represented the calculated values of heat transfer from inferior boundary, $q_y/(\sigma_B T_w^4)$, for different values of the absorption coefficient, where σ_B is the Stefan - Boltzmann Constant, $q_y = \pi I_{bw} - q^-_{(k,1)}$ and $q^-_{(k,1)}$ is given by Eq. (17b).

The solid and dashed lines represent, respectively, the results for S_2 and S_4 quadratures. The crosses represent the solution for an S_6 quadrature. The circles represent the values calculated by Razzaque et al. [13] using a finite element method.

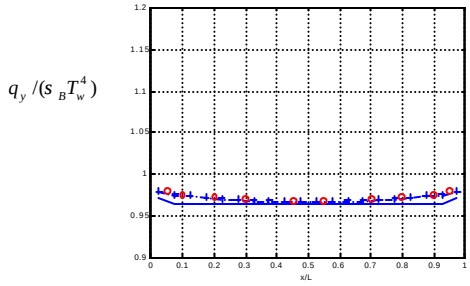
There for all test cases used for the analysis of the inverse problem are have used S_8 and S_{10} quadratures.



(a) σ_a = 0,1



(b) σ_a = 1,0



(c) σ_a = 2,0

Figure 5. Validation of the direct problem solution.

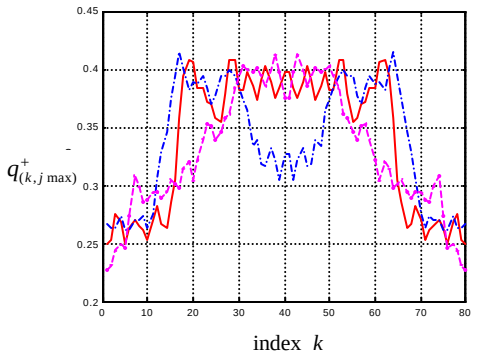
Inverse Problem Test Case Results

Considering the configuration 1 shown in Fig 1, we have observed that the solution of the direct problem for the values of $q^+_{(k,j\max)}$, Eq. (17a), that could he measured using the detectors shown in Fig. 4a, was affected by the ray effect. The solutions for $q^-_{(k,1)}$, given by Eq. (17b), that could be measured using the detectors shown in Fig. 4b where much smoother. In Fig. 6 are shown the results obtained for $q^+_{(k,j\max)}$ and $q^-_{(k,1)}$ for a computational grid with, i_{max}=j_{max}=80, and using different quadratures: LSN S_8 (Level Symmetric N), LSH S_{10} (Level Symmetric Hybrid) and EWO S_{10} (Equal Weight Odd) [11,12].

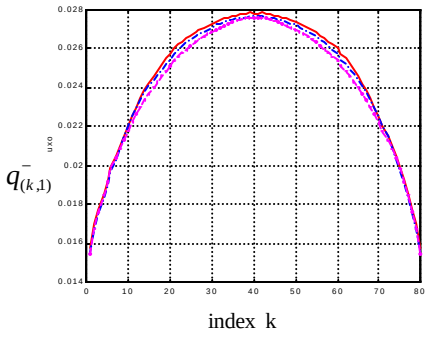
Even though the intensities for the $q^-_{(k,1)}$ (reflected by the medium) where smaller than the intensities for $q^+_{(k,j\max)}$ (transmitted through the medium) we decided to use only the detectors according to system 2 (located at the south boundary of the domain), as shown in Fig. 4b.

Homogeneous Media without Internal Sources

Using "exact data" we have solved the inverse problem for the estimation of σ_a and σ_s in the range [0,1.0] using as initial guesses $\sigma_a^o = \sigma_s^o = A_1^o$ =0,0. Convergence was obtained without regularization in almost all cases with exception



(a) $q^{+}_{(k,j\,max)}$



(b) $q^{-}_{(k,1)}$

LSN LSH EWO

Figure 7. Calculated values of the radiative heat fluxes $q^{+}_{(i,j_{max})}$ and $q^{-}_{(i,1)}$. $\sigma_a = \sigma_s = A_1 = 0.1$

for $0{,}9 \le \sigma_s \le 1$ and $0{,}5 \le \sigma_a \le 0{,}8$. Using regularization (a nonzero damping factor in the Levenber-Marquardt algorithm), convergence was obtained for all cases.

Homogeneous Media with Internal Sources

Using also "exact data" we estimated σ_a, σ_s and A_1 for a medium with an internal source of known constant strength S=1. The square domain has a side of length 1, i.e. $Lx = Ly = 1$ and the source occupies a square region inside it with $0{,}35 \le x, y \le 0{,}65$. We considered the test case of $\sigma_a = \sigma_s = A_1 = 0.1$ and initial guesses $\sigma_a^{\,o} = \sigma_s^{\,o} = A_1^{\,o} = 0$. Convergence was obtained within 6 iterations with a convergence criterion of $|\Delta\sigma_a| + |\Delta\sigma_s| + |\Delta A_1| < 10^{-6}$.

We then solved the inverse problem considering known $\sigma_a = \sigma_s = A_1 = 0.1$ and looking for the estimation of the intensity of the source $S = 1$, using the initial guess $S = 0$. Convergence was obtained in one iteration.

Heterogeneous Media

In Fig 7 are shown the shapes of the heterogeneous media considered. They are composed of two different materials, referred to as 1 and 2.

First we consider known σ_{a1}, σ_{s1} and A_{11}, and we estimate σ_{a2}, σ_{s2} and A_{12}. Then we consider known σ_{a2}, σ_{s2} and A_{12}, and we estimate σ_{a1}, σ_{s1} and A_{11}. Following that we considered the simultaneous estimation of σ_{a1}, σ_{s1}, A_{11}, σ_{a2}, σ_{s2} and A_{12}. Several test cases were run but we selected just two examples.

In Table 1 are shown the results for the estimation of σ_a, σ_s and A_1 in a homogeneous medium, with an internal source of constant strength and external detectors, as shown in Fig 8. The simulation considered a computational grid of 20x20 nodes. As expected deviations are observed when the number of detectors is too small. Later we present a simulation where the results improve even when a reduced number of detectors is used. In Table 2 are shown the results for a heterogeneous medium without internal sources. Here we have also considered only external detectors.

In all test cases zero values are used as initial guesses for the unknowns.

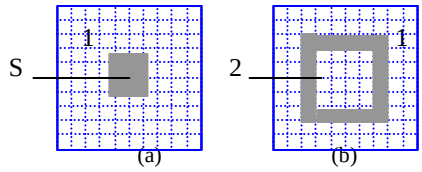
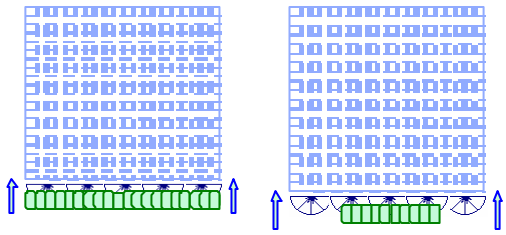


Figure 7. Heterogeneous media composed of two different materials



(a) external detectors (b) reduced number of detectors

Figure 8. Location of the detectors used for the solution of the inverse problem

Table 1. Estimates of the properties in the homogeneous medium with source. S = 1

Exact values		s_a=0,1	s_s=0,1	A_1=0,1
# Detectors		20	8	6
Estimates	s_a	0,100	0,1000	0,1002
	s_s	0,100	0,1000	0,9977
	A_1	0,100	0,1000	0,8082
	S	1,000	1,0000	1,0001

Table 2. Estimate of the properties in a heterogeneous medium (see fig. 7b)

Unknowns		s_{a1}=0,1	s_{s1}=1,0	A_{11}=0,1
		s_{a2}=2,0	s_{s2}=0,1	A_{12}=0,0
# Detectors		20	8	6
Estimates	s_{a1}	0,100	0,100	0,10579
	s_{s1}	0,999	0,999	0,87577
	A_{11}	0,100	0,100	-0,2149
	s_{a2}	1,999	1,999	1,958
	s_{s2}	0,100	0,100	0,16487
	A_{12}	1,32e-12	-1,0e-09	-0,1211

Experimental Data with Noise

In all cases presented before we considered noiseless data. As real data was not available, we consider synthetic data by adding pseudo random noise generated by the computer to the exact calculated values of the radiative heat flux

$$q^{-}_{(k,1)\exp} = q^{-}_{(k,1)exato} + r_k \, s \qquad (29)$$

where r_k is a random number and s simulates the standard deviation of measurement errors.

In Fig. 9 are presented the estimates obtained for the exact values $s_a = s_s = A_1 =$ 0,5 in a homogeneous media, considering an error of 1,25% in the maximum value measured for $q^{-}_{(k,1)}$. For smaller values of $q^{-}_{(k,1)}$ the errors are larger. In Fig 9 the different runs are related to different sets of random noise, simulating different experiments.

Sensitivity Coefficients

In Fig. 10 are presented the sensitivity coefficients for s_a, s_s and A_1 in a homogeneous media, considering the detector system shown in Fig. 4b. We observe that higher values are obtained for the detectors located close to the center of the south boundary of the domain. Using the detectors located in the position for which the sensitivity is higher we obtained better estimates for the unknowns as shown in Fig. 11. Compare the results shown in Figs. 9 and 11. In both cases we have 1,25% error in the synthetic experimental data. We stress that in the former all detectors were considered, while in the latter just a smaller number located close to the center of the south boundary with higher sensitivity

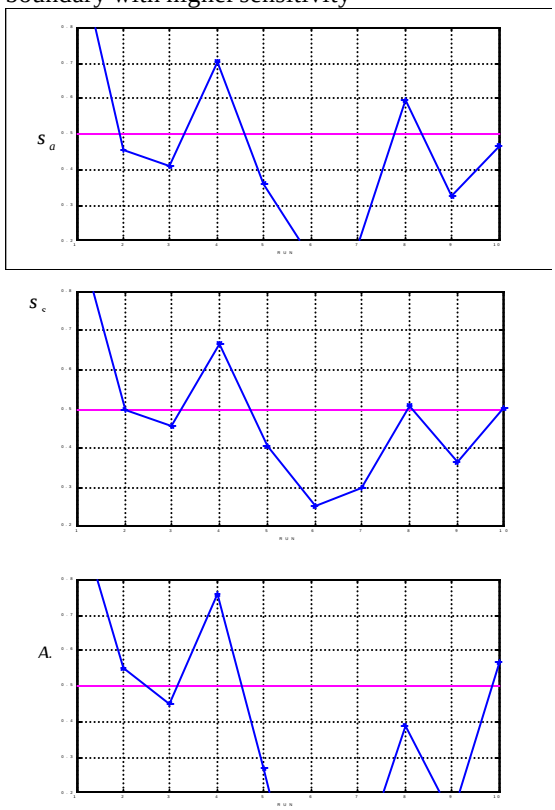


Figure 9. Estimates for s_a, s_s and A_1 in a homogeneous medium considering synthetic data with 1,25% error.

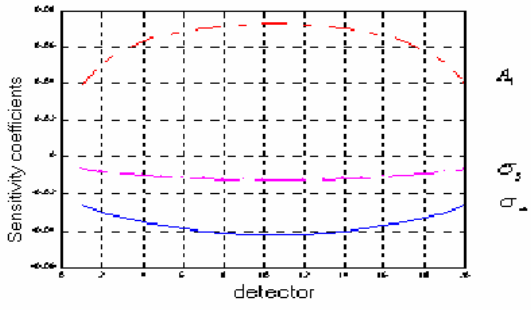


Figure 10. Sensitivity coefficients for $s_a = s_s = A_1 = 0.1$

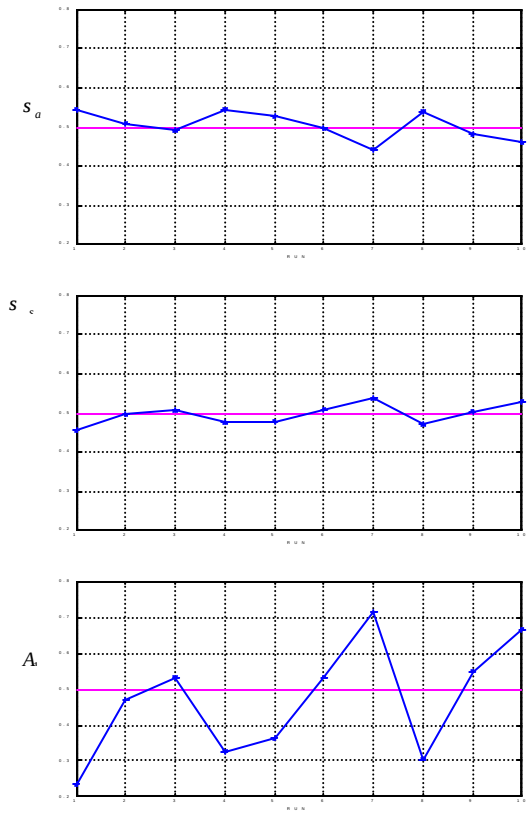


Fig 11. Estimates for s_a, s_s and A_1 in a homogeneous medium considering synthetic data with 1.25% error. Only the detectors located at positions of higher sensitivity were considered

CONCLUSIONS

The test case results presented demonstrate the feasibility of solving inverse radiative transfer problems in two-dimensional media using external detectors, measures the radiative heat flux. Both homogeneous and heterogeneous media are considered, and the estimates are related to radiative properties and internal sources. Following the results of the sensitivity analysis results of the inverse problems improved when detectors located only at the regions of higher sensitivity were used.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPERJ – Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, and CAPES – Comissão de Aperfeiçoamento de Pessoal de Nível Superior.

REFERENCES

1. Kedrow, N. V. Popov , A. S. Khimchenko and A. S. Tsybin, Neutron – gamma well logging apparatus with more sensitivity for chlorine, *Applied Radiation and Isotopes*, **48**, pp. 1647-1648 (1997).

2. G. T. Herman, Image Reconstruction from Projections – The Fundamentals of Computerized Tomography, Academic Press (1980).

3. N. J. McCormick, Inverse photon transport methods for biomedical applications, *Proc. 1st International Conference on Inverse Problems in Engineering Theory and Practice*, ASME. Florida, USA, pp. 253-258. (1993).

4. R. A. Leathers, C. S. Roesler and N. J. McCormick, Ocean inherent optical property determination from in-water light field measurements, *Applied Optics*, **38**, pp. 5096-5103 (1999).

5. A. J. Silva Neto and M. N. Özisik, An inverse problem of simultaneous estimation of radiation phase function, albedo and optical thickness, *J. Quant. Spectrosc. Radiat. Transfer*, **53**, pp. 397-409, (1995).

6. A. T. Kauati, A. J. Silva Neto and N. C. Roberty, A Source-detector methodology for the construction and solution of the one dimensional inverse transport equation, *Inverse Problems in Engineering*, **9**, pp. 25-66 (2001).

7. C. E. Siewert, An inverse source problem in radiative transfer, *J. Quant Spectrosc. Radiat. Transfer*, **50**, pp 603- 609 (1993).

8. F. M Modest, Radiative Heat Transfer, McGraw-Hill, Inc. (1993).

9. M. Sakami, A. Charette, and V. Le Dez, Radiative heat transfer in three - dimensional enclosures of complex geometry by using the discrete-ordinate method, *J. Quant. Spectrosc. Radiat. Transfer*, **59**, pp. 117-136. (1998).

10. S. Chandrasekhar, Radiative Transfer, Dover Publications. (1960).

11. W. A. Fiveland, Discrete – ordinates solutions of the radiative transport equation for rectangular enclosures, *Journal of Heat transfer*, **106**, pp 699-706 (1984).

12. J. S. Truelove, Three- dimensional radiation in absorbing-emitting-scattering media using the discrete-ordinates approximation. *J. Quant. Spectrosc. Radiat. Transfer*, **39**, pp. 27-31 (1988).

13. M. N. Razzaque, D. E. Klein, and J. R. Howell, Finite element solution of radiative transfer in a two–dimensional rectangular enclosure with gray participating media, *Journal of Heat Transfer*, **105**, pp. 933-936 (1983).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

RECONSTRUCTION OF ABSORPTION AND SCATTERING COEFFICIENTS WITH A DISCRETE ORDINATES METHOD CONSISTENT WITH THE SOURCE-DETECTOR SYSTEM

Raúl F. Carita Montero[1]
Nilson C. Roberty[1]
Antônio J. Silva Neto[2,1]
[1]Nuclear Engineering Program – PEN/COPPE, Universidade Federal do Rio de Janeiro, UFRJ P.O. Box 68509, 21945-970, Rio de Janeiro, RJ, Brazil,
carita@lmn.con.ufrj.br, nilson@lmn.con.ufrj.br, ajsneto@lmn.con.ufrj.br
[2]Department of Mechanical Engineering and Energy, Instituto Politécnico, IPRJ, Universidade do Estado do Rio de Janeiro, UERJ, P. O. Box 97282, 28601-970, Nova Friburgo, RJ, Brazil,
ajsneto@iprj.uerj.br

ABSTRACT

In the present work the absorption and scattering coefficients in heterogeneous two-dimensional media are estimated using the source-detector methodology and a discrete ordinates method whose directions of radiation propagation are taken in a way consistent with the source-detector system for parallel beams of radiation.

The domain partition, the solution of the direct problem and the source detector methodology are briefly described, and preliminary test case results are presented.

INTRODUCTION

The well known transmission tomography is based on a simplification of the linearized Boltzmann equation in which the scattering phenomenon is not taken into account [1-3]. But, in some practical applications scattering cannot be neglected [4]. In such cases the mathematical model for the radiation propagation may be made with an approximation for the transport equation based on discrete ordinates, being therefore given by a set of first order partial differential equations [5].

In the present work the domain partition is constructed to be consistent with the directions of propagation of the parallel beams of radiation. These directions are therefore used for the discrete ordinates scheme, allowing the solution of the direct problem with a discontinuous Galerkin method, marching along each of the discrete ordinates directions. The formulation of the inverse problem with the source-detector methodology [4, 6] leads to a system of non-linear integral equations. The solution of the inverse problem is done with *q*-ART type algorithms [1], based on the method of projection into convex sets.

Test case results are presented for isotropic scattering media without internal source and transparent boundaries. Refractive index are neglected.

MATHEMATICAL FORMULATION OF THE DIRECT PROBLEM

Consider an absorbing and scattering two-dimensional medium with no internal radiation sources, subjected to externally generated parallel beams of radiation.

For the steady-state situation, with no spectral dependency, the following mathematical formulation is obtained from the linearized Boltzmann equation [7] which is used for the modeling of the interaction of the radiation with the participating medium

$$\vec{\Omega}\cdot\nabla f(\vec{x},\vec{\Omega}) + s_t(\vec{x}) f(\vec{x},\vec{\Omega}) =$$

$$= \int_{2p} s_s(\vec{x}, \vec{\Omega}' \cdot \vec{\Omega}) f(\vec{x}, \vec{\Omega}') d\vec{\Omega}' \quad \text{in domain } D \qquad (1a)$$

where f is the intensity of the radiation, $\vec{x}$ represents a point in the domain, $\vec{\Omega}$ is the direction of propagation of the radiation, s_t is the total extinction coefficient (absorption + out scattering), s_s is the scattering coefficient, and $\vec{\Omega}'$ is the direction of incident radiation that is scattered at point $\vec{x}$ to the direction $\vec{\Omega}$.

In order to complete the mathematical formulation of the direct problem of radiative transfer represented by Eq. (1) boundary conditions are necessary. Considering that the intensity of the externally incident radiation source is known, f^{in}, we write

$$f(\vec{x}, \vec{U}) = f^{in}(\vec{x}, \vec{U}) \quad \text{for } \vec{U} \cdot \vec{n}(\vec{x}) < 0 \text{ at } \vec{x} \in \partial D^- \qquad (1b)$$

where ∂D^- is the boundary of domain D through which the external radiation comes into the medium, and $\vec{n}(\vec{x})$ is the outward normal at point $\vec{x}$ of the boundary.

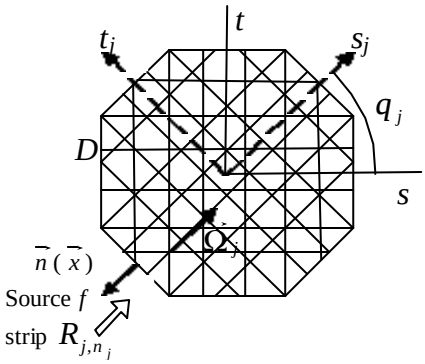


Figure 1- Domain partition consistent with parallel beams of radiation

Using a domain partition consistent with the parallel beams of radiation as shown in Fig. 1, and considering a particular direction $\vec{\Omega}_j$ with a coordinate system given by (s_j, t_j) that is rotated in accordance to the particular direction $\vec{\Omega}_j$, Eq. (1a) can be written for each strip R_{j,n_j} in the following way

$$\vec{\Omega}_j \cdot \nabla f_j(\vec{x}) + s_t(\vec{x}) f_j(\vec{x}) = \sum_{k=1}^{2J} w(\vec{\Omega}_k \cdot \vec{\Omega}_j) s_s(\vec{x}, \vec{\Omega}_k \cdot \vec{\Omega}_j) f_k(\vec{x})$$

for $j=1,\ldots,2J$; and $\vec{x} \in R_{j,n_j} \subset D$, $n_j=1,\ldots,2M$ (2a)

where $2M$ is the total number of parallel strips that compose the domain partition for each set of parallel beams of radiation that propagates along the direction $\vec{\Omega}_j$, R_{j,n_j} represents each strip, w are the corresponding weights due to the quadrature and f_k is the intensity of the radiation incident along direction $\vec{\Omega}_k$ that is scattered at position $\vec{x} \in R_{j,n_j}$ into direction $\vec{\Omega}_j$.

The boundary condition (1b) is written as

$$f_j(\vec{x}) = f_j^{in}(\vec{x}) \quad \text{for } \vec{\Omega}_j \cdot \vec{n}(\vec{x}) < 0 \quad \text{at } \vec{x} \in \partial R_{j,n_j}^- \qquad (2b)$$

where $\partial R_{j,n_j}^-$ represents the boundary of the strip through which the external radiation comes into the strip.

Rotated Discrete Ordinates Approximation

The formulation of the radiative transfer problem given by Eqs. (2) consists on a discrete ordinates formulation with the discretization of the angular domain being made in a way that is consistent with parallel beams of radiation originated at external sources.

Adopting the coordinate system that rotates in accordance to each particular direction $\vec{\Omega}_j$, the first term in the left hand side of Eq. (2a) can be written as

$$\vec{\Omega}_j \cdot \nabla f_j(\vec{x}) = \frac{df_j(s_j(\vec{x}), t_j(\vec{x}))}{\partial s_j} \qquad (3)$$

for $\vec{x} \in R_{j,n_j}$ $j=1,\ldots,2J$, $n_j 1,\ldots,2M$

Given the geometry, the radiative properties, and the boundary conditions, the values of f_j, $j=1,\ldots,2J$, at each rectangle $m_j=1,\ldots,2M$ of each strip R_{j,n_j} can be calculated. This is the direct problem. For each strip R_{j,n_j} there is an ordinary differential equation to be solved, therefore the

2Jx2M equations form a system of coupled equations to the solved.

Rearranging the summation terms in the right hand side of Eq. (2a), and using Eq. (3), we obtain the following matricial form

$$\frac{df_j(s_j(\dot{x}),t_j(\dot{x}))}{\partial s_j}+\sum_{k=1}^{2J} s_{kj} f(s_k(\dot{x}),t_k(\dot{x}))=0$$

for $j=1,\ldots,2J$ (4)

where the matrix elements of $\boldsymbol{s}$ are given by the combination of the scattering and absorption coefficients

$$s(l,l,\vec{x})=s_a(\vec{x})+\sum_{\substack{k=1\\k\neq l}}^{K} w(\vec{\Omega}_k,\vec{\Omega}_l)s_s(\vec{x},\vec{\Omega}_k\cdot\vec{\Omega}_l), \quad (5a)$$

$$s(l,m,\vec{x})=-w(\vec{\Omega}_l,\vec{\Omega}_m)s_s(\vec{x},\vec{\Omega}_l\cdot\vec{\Omega}_m), \quad (5b)$$

for $l,m=1,2,\ldots,2J$ e $l\neq m$.

This configuration is similar to that employed in computerized transmission tomography, with exception of the position of the detectors where in the case considered here they measure not only the transmitted but also the scattered radiation.

Considering the rotational symmetry of the system source-detector, the matrix $\boldsymbol{s}$ is symmetrical and cyclical, being written then in the form

$$s=\left[\begin{array}{ccccc:ccccc}
s_{1,1} & s_{1,2} & s_{1,3} & \cdots & s_{1,J} & s_{1,J+1} & s_{1,J} & s_{1,J-1} & \cdots & s_{1,2}\\
s_{1,2} & s_{1,1} & s_{1,2} & \cdots & s_{1,J-1} & s_{1,J} & s_{1,J+1} & s_{1,J} & \cdots & s_{1,3}\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots\\
s_{1,J} & s_{1,J-1} & s_{1,J-2} & \cdots & s_{1,1} & s_{1,2} & s_{1,3} & s_{1,4} & \cdots & s_{1,J+1}\\
\hdashline
s_{1,J+1} & s_{1,J} & s_{1,J-1} & \cdots & s_{1,2} & s_{1,1} & s_{1,2} & s_{1,3} & \cdots & s_{1,J}\\
s_{1,J} & s_{1,J+1} & s_{1,J} & \cdots & s_{1,3} & s_{1,2} & s_{1,1} & s_{1,2} & \cdots & s_{1,J-1}\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots\\
s_{1,2} & s_{1,3} & s_{1,4} & \cdots & s_{1,J+1} & s_{1,J} & s_{1,J-1} & s_{1,J-2} & \cdots & s_{1,1}
\end{array}\right] \quad (6)$$

One can observe that every row of the matrix $\boldsymbol{s}$ can be obtained from the previous one by just making the shift of the coefficients.

Shifting the incidence direction by an angle

$$q_j=(j-1)\frac{p}{J} \quad (7)$$

we can write without loss of generality that

$$s_{j,j}=s_{1,1} \quad , j=1,2,\ldots,2J \quad (8)$$

and the terms out of the diagonal are taken symmetrically to the diagonal.

For the solution of the problem given by Eqs. (4-8) we now march along the discrete directions $\vec{\Omega}_j$, *j=1,2... 2J*, following each one of the strips, and for each direction the corresponding coordinate system is used.

To the best of our knowledge this is a new procedure in the context of the discrete ordinates method [8].

Domain Partition

We have *2J* regular square meshes, called here pixel meshes, each one compatible with one angular flux transport direction. In each of these rotated coordinate systems the transport problem can be written as a one dimensional problem. The complete intersection of all pixel meshes are taken into account for the domain partition in the polygons p_e with $e=1,\ldots,E$, where E is the total number of polygons.

First, the intersections of all lines are calculated (observe that each strip has two limiting lines). The points where the intersections take place are the vertices of the polygons p_e.

Now a major task has to be performed, that is, from a set of vertices (x_1, y_1), (x_2, y_2),…, (x_L, y_L) where L is the total number of intersection points, we have to identify those that form a polygon. An algorithm has been developed, in which for each rectangular pixel, we sequentially investigate for all intersection points contained in the interior or at the boundary of the pixel, all possible elements that can be created around each intersection point. Obviously each intersection point has to be a vertex of all polygons created around it. The polygons constructed in this way are characterized by the indices that relate the intersection points with the indices of all strips that gave origin to them. Afterwards, in the polygons numbering step, if a polygon is created identical to the polygon created by another of its vertices, the element counter is not increased by one. In this way multiplicity is avoided in the process of polygons numbering, and at the same time the vertices that belong to each polygon are determined.

The output of the computer code written for the construction of the natural base of simple Lebesgue functions consists on the ordered listing of all polygons, their areas, optionally their vertices, and its pertinence matrix (incidence matrix) for its localization in each one of the oriented pixel. The pertinence matrix is defined by two sets of indices, i.e. $n(e,j)$ and $m(e,j)$ that

locates each particular element *e* with respect to the rotated coordinate system (s_j,t_j), $j=1,\ldots,2J$.

A computer code for the automatic mesh generation was written in the MATLAB environment. More information about this kind of mesh generation can be found in Reis and Roberty [8] and Carita, Roberty and Silva Neto [1-3]. The set characterization of the oriented pixels are $K(n_j,m_j,j)\in V_M^J$ where

$$K(n_j,m_j,j)=$$
$$\{\vec{x}=(x,y)\in\Re^2 = n_j < x\sin\theta_j + y\cos\theta_j < n_j+1$$
$$\text{and } m_j < -x\cos\theta_j + y\sin\theta_j < m_j+1\} \quad (9)$$

and V_M^J is the domain partition composed by the polygons $e=1,..,E$.

SOLUTION OF THE DIRECT PROBLEM

We have solved the direct problem and formulated the source detector methodology equations, to be used in the inverse problem solution [4, 6], within the context of the Galerkin discontinuous method formulated by Lesaint and Raviart [9-11]. The combination of the discontinuous finite element formulation with a domain partition constructed according to the discrete ordinate oriented pixel meshes, produces a simplified and robust algorithm for solving the discrete problem, and to formulate the respective reconstruction inverse problem.

The system (4) can be written for each strip R_{j,n_j} as

$$\frac{\partial f^{f}_{j,n_j}(s_j,t_j)}{\partial s_j} = -s\,(j,j,e(s_j,t_j))f^{f}_{j,n_j}(s_j,t_j) + \sum_{\substack{k=1\\k\neq j}}^{2J} s\,(k,j,e(s_j,t_j))f^{f}_{k,n_k}(s_k,t_k) \quad (10a)$$

$$f^{f}_{j,n_j}(1,n_j) = f^{in,f}_{j,n_j} \text{ for } j=1,\ldots,2J;\ n_j=1,\ldots,2M \quad (10b)$$

where superscript *f* indicates a specific source for the direct problem.

We have adopted the following representation for the angular flux

$$f_j(\vec{x}) = \sum_{n_j=1}^{2M}\sum_{m_j=1}^{2M} p(s_j)\,c(n_j,m_j,j)(\vec{x}) \qquad j=1,..,2J \quad (11a)$$

where

$$c(n_j,m_j,j)(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x}\in K(n_j,m_j,j) \\ 0 & \text{otherwise} \end{cases} \quad (11b)$$

and $p(s_j)\in P_r(K(n_j,m_j,j))$ where P_r is the set of polynomials of order *r*.

For the combination of absorption and scattering coefficients we have used

$$s\,(k,j,\vec{x}) = \sum_{e=1}^{E} s\,(k,j,e)\,c_e(\vec{x}) \quad (12)$$

where the characteristic function c_e assumes the following values

$$c_e(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x}\in \bigcap_{j=1}^{2J} K(n(e,j),m(e,j),j) \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

The approximation space for this discontinuous flux representation is

$$W_M^J = \{v\in L_2(D): v|_{K(n_j,m_j,j)}\in P_r(K(n_j,m_j,j))\} \qquad j=1,...,2J \quad (14)$$

The variational formulation for Eqs.(10) is: for $K_j\equiv K(n_j,m_j,j)\in V_M^J$, given $f^{J-}_{M,j}$ on $\partial K_j^- \equiv \partial K^-(n_j,m_j,j)$, find $f^J_{M,j}\in W_M^J$ such that:

$$(\vec{\Omega}_j\cdot\nabla f^J_{M,j} + s_t f^J_{M,j}, v)_{K_j} - \int_{\partial K_j^-}[f^J_{M,j}]\,v^+\vec{\Omega}_j\cdot\vec{n}\,dS = (g,v)_{K_j} \quad \forall v\in P_r(K_j),\ \forall K(n_j,m_j,j) \quad (15)$$

where $[f^J_{M,j}] = f^{J+}_{M,j} - f^{J-}_{M,j}$ is the jump at the influx boundary and *g* corresponds to the last term in the right hand side of Eq. (10a). If we adopt a constant polynomial inside the rotated pixel i.e. $p(s_j)$=constant we obtain the following discretized equation

$$[f(n_j,1+m_j,j) - f(n_j,m_j,j)] + f(n_j,1+m_j,j)\sum_{e\in K(n_j,m_j,j)} s\,(j,j,e)a(e) = \sum_{e\in K(n_j,m_j,j)}\sum_{\substack{k=1\\k\neq j}}^{2J} s_e(k,j,e)a(e)\times f(n(e,k),1+m(e,k),k)$$
$$\forall\ j=1,...,2J,\ n_j=1,\ldots,2M, \text{ and } m_j=1\ldots,2M \quad (16a)$$

where $a(e)$ is the area of element e, and

$$f^{+}(n_j, m_j, j) = f^{-}(n_j, m_j + 1, j) = f(n_j, m_j + 1, j) \quad (16b)$$

as represented in Fig. 2.

When f^{f}_{j,n_j} in Eq. (10a) assumes only non zero values, for $j=1,\ldots,2J$, $n_j=1,\ldots,2M$, and $m_j=1,\ldots,2M$, we can write a convenient equation for the logarithm of the flux

$$\frac{\partial \ln f^{f}_{j,n_j}(s_j,t_j)}{\partial s_j} = -s\,(j, j, e(s_j,t_j)) + \\ + \sum_{\substack{k=1 \\ k\neq j}}^{2J} s\,(k, j, e(s_j,t_j)) f^{f}_{k,n_k}(s_k,t_k) / f^{f}_{j,n_j}(s_j,t_j) \quad (17a)$$

$$\ln f^{f}_{j,n_j}(1,n_j) = \ln f^{in,f}_{j,n_j} \\ \text{for } j=1,\ldots,2J;\ nj=1,\ldots,2M \quad (17b)$$

A discretized version of this equation can be derived in a similar way to that used to obtain Eq. (16a) from Eq. (10a).

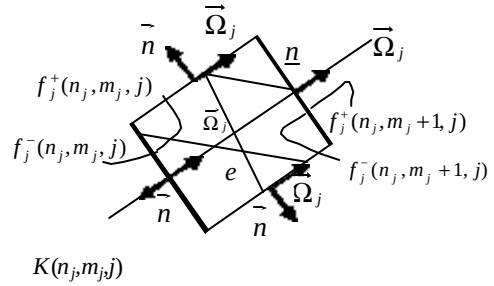


Figure 2.- Rotated pixel (n_j,m_j) in direction $\vec{\Omega}_j$ for discontinuous Galerkin finite element.

MATHEMATICAL FORMULATION AND SOLUTION OF THE INVERSE PROBLEM

Here we describe the formulation of the radiative transfer problem using the source-detector methodology [4, 6]. In all previous works we have considered only applications in one-dimensional media. In the present work we consider a two-dimensional absorbing and scattering media.

The source-detector methodology consists on a explicit formulation for the inverse radiative transfer problem in participating media. For each source-detector pair a non-linear equation is derived using a convolution of the direct problem (source problem) with the solution of the adjoint problem (detector problem). In both problems the modeling of scattering, absorption and emission (when included in the formulation) is done with the linearized Boltzmann equation, but in the latter reference values are available for the unknowns that we want to determine, being therefore included in the formulation. This approach results in a system of non-linear equations with the unknowns explicitly represented. This system of equations is called Inverse Transport Equation (ITE).

The basic steps of the source-detector methodology are therefore the formulation and solution of the following problems: (i) source problem; (ii) detector and auxiliary problems, and (iii) Inverse Transport Equation (ITE).

The Source Problem

The source problem is given by Eqs. (10a-b), or in the discretized version given by Eqs. (16a-b).

The Detector and Auxiliary Problems

The detector problem is obtained with an adjoint problem formulated by reversing the direction of radiation transfer, i.e. by replacing m by $-m$ in Eq. (10a),

$$-\frac{\partial f^{*d}_{j,n_j}(s_j,t_j)}{\partial s_j} = -s^{\,d}(j, j, e(s_j,t_j)) f^{*d}_{j,n_j}(s_j,t_j) \\ + \sum_{\substack{k=1 \\ k\neq j}}^{2J} s^{\,d}(k, j, e(s_j,t_j)) f^{*d}_{k,n_k}(s_k,t_k) \quad (18a)$$

$$f^{*d}_{j,n_j}(2M+1, m_j) = f^{out,*d}_{j,n_j} \\ \forall\ j=1,\ldots,2J \text{ and } n_j=1,\ldots,2M \quad (18b)$$

where $s^{\,d}$ represents a reference value for the unknown s that contains the primary unknowns s_a and s_s as shown in Eqs. (5a-b).

By reversing once more the direction of radiation propagation we obtain from problem (18) the following auxiliary problem

$$\frac{\partial f^{d}_{j,n_j}(s_j,t_j)}{\partial s_j} = -s^{\,d}(j, j, e(s_j,t_j)) f^{d}_{j,n_j}(s_j,t_j) \\ + \sum_{\substack{k=1 \\ k\neq j}}^{2J} s^{\,d}(k, j, e(s_j,t_j)) f^{d}_{k,n_k}(s_j,t_j) \quad (19a)$$

$$f^{d}_{j,n_j}(1,n_j) = f^{out,*d}_{j+J,2M+1-n_j}$$
$$\forall\ j=1,\ldots,2J \text{ and } n_j=1,\ldots,2M \qquad (19b)$$

Deriving the product of the radiation intensities $f^{f}_{j,n_j} f^{*d}_{j,n_j}$, and introducing Eqs (10a) and (19a), we obtain

$$\frac{\partial f^{f}_{j,n_j} f^{*d}_{j,n_j}}{\partial s_j} = -[s(j,j,e(s_j,t_j)) - s^{d}(j,j,e(s_j,t_j))] f^{f}_{j,n_j} f^{*d}_{j,n_j} + \sum_{\substack{k=1 \\ k\neq j}}^{2J} [s(k,j,e(s_j,t_j)) f^{f}_{k,n_k} f^{*d}_{j,n_j} - s^{d}(k,j,e(s_j,t_j)) f^{*d}_{k,n_k} f^{f}_{j,n_j}]$$
$$\forall\ j=1,\ldots,2J \text{ and } n_j=1,\ldots,2M \qquad (20)$$

Due to the symmetric nature of the system of rotated equations, the adjoint detector problem is equivalent to an auxiliary problem with the same radiative properties of the detector problem.

Observe that the auxiliary problem is equivalent to a source problem, but for the radiative properties reference values are used.

In terms of the discrete values of the angular flux, this equivalence can be stated

$$f^{*d}_{j,n_j}(s_j,t_j) = f^{d}_{j+J,2M+1-n_j}(2M+2-s_j,2M+1-t_j) \qquad (21)$$

The Discretized Inverse Transport Equation (DITE)

To obtain the DITE we integrate the source detector equation (20) along a specified strip R_{j,n_j}

$$[f^{f}(n_j,2M+1,j) f^{d}(n_j,1,j+J) - f^{f}(n_j,1,j) \times f^{d}(n_j,2M+1,j+1)]\Delta t_j =$$
$$= \sum_{m_j=1}^{2M} \sum_{e\in K(n_j,m_j,j)} [s(j,j,e) - s^{d}(j,j,e)] a(e) f^{f}(n_j,1+m_j,j) \times f^{d}(n_j,2M+1-m_j,j+J) -$$
$$- \sum_{\substack{m_j=1 \\ k\neq j}}^{2M} \sum_{k=1}^{2J} \sum_{e\in K(n_j,m_j,j)} [s(k,j,e) f^{f}(n(e,k),1+m(e,k),k) \times f^{d}(n_j,2M+2-m_j,j+J) -$$
$$- s^{d}(k,j,e) f^{d}(n(e,k),2M+1-m(e,k),k+J) \times f^{f}(n_j,m_j,j)] a(e) \qquad (22)$$

where $s(k,j,e)$ has been defined in Eq. (12) and $f(n_j,m_j,j)$ are rotated pixel values for the discontinuous flux as shown in Eqs. (11a, 16a, 16b).

Alternative Formulation of the Inverse Problem

When $f^{f}_{j,n_j}(s_j,t_j) \neq 0$ an alternative formulation to DITE can be obtained using the logarithmic formulation of the radiative transfer problem, Eqs. (17a-b), yielding

$$\ln \frac{f^{f}(n_j,2M+1,j)}{f^{f}(n_j,1,j)} = \sum_{m_j=1}^{2M} \left(\sum_{e\in K(n_j,m_j,j)} s(j,j,e) a(e) \right) - \sum_{\substack{m_j=1 \\ k\neq j}}^{2M} \sum_{k=1}^{2J} \left(\sum_{e\in K(n_j,m_j,j)} s(k,j,e) \frac{f^{f}(n(e,k),m(e,k),k)}{f^{f}(n_j,m_j,j)} \right) \qquad (23)$$

SOLUTION OF THE INVERSE PROBLEM

As mentioned before the solution of a direct problem, such as the source and auxiliary problems, is done by an algorithm that marches along each strip. The scattering term in the equation is calculated using the values for the fluxes in the previous iteration.

For the solution of the inverse problem, Eqs. (22) or (23), we must distinguish two situations. When the system is consistent a solution can be obtained with the q-ART algorithm [1]. The special case $q \to 0$ yields the MART maximum entropy algorithm used in image reconstruction [1, 8]. Even with the use of noisy data good results may be obtained with this methodology when a natural base representation of the radiative properties is used. The other situation is when the system is inconsistent and it may occur because of the large number of equations derived from all pairs source-detectors. In this case two approaches can be used: (i) a Tikhonov regularization can be applied; or (ii) only the equations with highest sensitivity are considered.

PRELIMINARY RESULTS

To present the results of the inverse problem solution we have considered two test cases. For the domain D the radiative properties are $\sigma_a = 2/(MJ)$ and $\sigma_s = 1/(4MJ)$. These values were used to allow that a significant part of the radiation goes through the medium, and measurements can be made at the point the radiation leaves the medium.

In the first test case shown in Fig. 3 both the absorption and scattering coefficients in a

embedded rectangular region inside domain *D* were amplified by a factor of 10.

In Fig. 5 are shown the reconstructed values of the radiative properties for test case 1.

In the second test case shown in Fig. 4 there are two rectangular regions with different properties from domain *D*. The highest peak represents an amplification by the factor 10 of the absorption coefficient only, and the second peak represents an amplification by the same factor of the scattering coefficient. In Fig. 6 are shown the reconstructed values of the radiative properties for test case 2.

The results presented here were obtained using the logarithmic formulation of the inverse problem given by Eq. (23) which has been intensively used for the transmission tomography, but here we consider a discrete ordinates domain partition consistent with the parallel beams of incident external radiation.

Even though it is possible to simultaneously estimate the absorption and scattering coefficients, the results shown in Figs. 5 and 6 take into account only the out-scattering term (see Eq. 5(a)), and the off-diagonal terms are neglected (in-scattering term). Therefore only the diagonal terms in matrix are estimated.

A total number of *Jx2M* exit radiation intensities were taken as simulated experimental data. The synthetic data was produced using the exact values of the radiative properties including the in-scattering term. As the inverse problem is solved neglecting this term, the in-scattering effect can be considering as noise in the data.

We are now working in the full implementation of the inverse radiative transfer problem in which the DITE given by Eq. (22) is used for the scattered radiation and the logarithmic formulation is used for the strip related to the incoming and transmitted radiation.

CONCLUSIONS

The methodology proposed for the solution of the direct problem of the radiation transport, consistent with the source-detector system, is convenient for the computational implementation of the inverse problem solution for the estimation of the scattering and absorption coefficients, generalizing the results previously obtained by Carita Montero, Roberty and Silva Neto [1-3]. The marches along the space coordinate are made following radiation parallel beams, and the term of in-scattering is approximated by a quadrature where the quadrature points are determined by the partition of the domain consistent with the geometry of the source-detector system. Therefore this approach is expected to minimize the errors due to the spatial and angular domains discretization in the inverse problem solution considered here.

With the use of sources and detectors located in different positions (as in the computerized tomography, CT), the method of Galerkin and the source detector methodology were able to locate the defect and to give an estimate for the extinction coefficient.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPERJ - Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, and CAPES – Comissão de Aperfeiçoamento de Pessoal de Nível Superior.

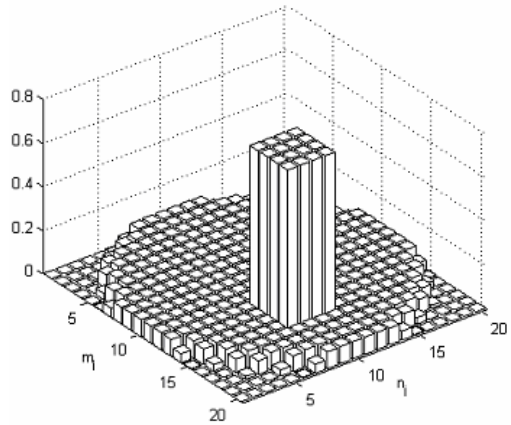


Figure 3- The extinction coefficient for configuration 1.

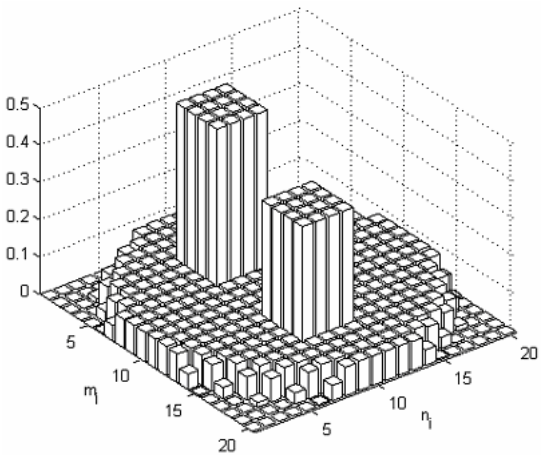


Figure 4- The extinction coefficient for configuration 2.

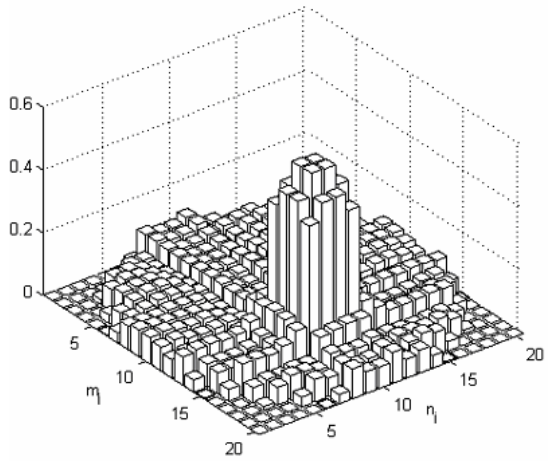


Figure 5- The reconstructed extinction coefficient for configuration 1.

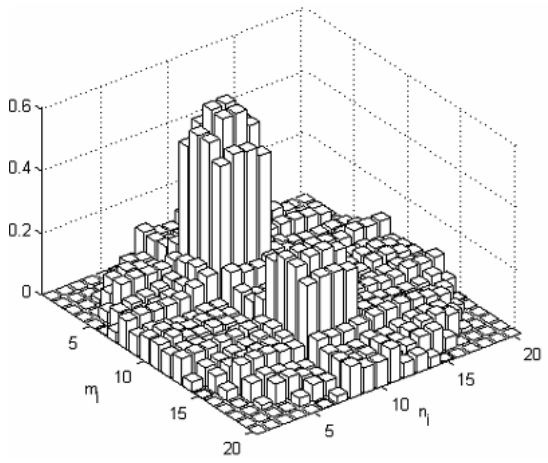


Figure 10- The reconstructed extinction coefficient for configuration 2.

REFERENCES

1. R. F. Carita Montero, N. C. Roberty, and A. J. Silva Neto, Absorption coefficient estimation in heterogeneous participating media using a domain partition consistent with divergent beams, *Inverse Problems in Engineering*, to appear (2002).

2. R. F. Carita Montero, N. C. Roberty, and A. J. Silva Neto, Absorption coefficient estimation in two-dimensional participating media using a base constructed with divergent beams, *Proc 34th National Heat Transfer Conference*, Pittsburgh, Pennsylvania, USA (2000).

3. R. F. Carita Montero, N. C. Roberty, and A. J. Silva Neto, Natural base construction for absorption coefficient estimation in heterogeneous participating media with divergent beams, *Proc. 3rd International Conference on Inverse Problems in Engineering: Theory and Practice*, Port Ludlow, Washington, USA pp.389-395 (1999).

4. A. T. Kauati, A. J. Silva Neto, and N. C. Roberty, Use of the source-detector technique in tomography with anisotropic scattering, *Proc. 17th Brazilian Congress of Biomedical Engineering*, Santa Catarina, pp. 1009-1014, in Portuguese (2000).

5. S. Chandrasekhar, Radiative Transfer, Dover Publications (1960).

6. A. T. Kauati, A. J. Silva Neto, and N. C. Roberty, A source-detector methodology for the construction and solution of the one-dimensional inverse transport equation, *Inverse Problems in Engineering*, **9**, 45-66 (2001).

7. M. N. Özisik, *Radiative Transfer and Interactions with Conduction and Convection*, John Wiley & Sons (1973).

8. M. L. Reis, and N. C. Roberty, Maximum entropy algorithms for image reconstruction from projections, *Inverse Problems*, **8**, pp. 623-644, (1992).

9. C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, New York, 1990.

10. P. Lesaint, Finite element methods for symmetric hyperbolic systems, *Num. Math.*,**21** pp. 244-255, (1973).

11. P. Lesaint, and P.A. Raviart, On a finite element method for solving the neutron transport equation *in Mathematical Aspects of Finite Elements in Patial Differential Equations*, C. de Boor ed., Academic Press, New York (1974).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

METAL/MOLD HEAT TRANSFER COEFFICIENTS DURING HORIZONTAL AND VERTICAL UNSTEADY-STATE SOLIDIFICATION OF Al-Cu AND Sn-Pb ALLOYS

Carlos A. dos Santos

Cláudio A. Siqueira

Amauri Garcia
Department of Materials Engineering
State University of Campinas, UNICAMP
Campinas, S.P., Brazil
alex@fem.unicamp.br

José M.V. Quaresma
Federal University of Pará, UFPA
Belém,P.A., Brazil

Jaime A. Spim
Federal University of Rio Grande do Sul, UFRGS
Porto Alegre,R.S., Brazil

ABSTRACT

In the present work, the metal/mold heat transfer coefficients (h_i) are determined during unidirectional solidification of Al-Cu and Sn-Pb alloys. The effects of casting assembly (horizontal and vertical), alloy composition, material and thickness of the mold and melt superheat are investigated. By using measured temperatures in both casting and metal, together with numerical solutions of the solidification problem, metal/mold heat transfer coefficients are quantified based on solution of the inverse heat conduction problem. Experimental temperatures are compared with simulations furnished by an explicit finite difference numerical model, and an automatic search selects the best theoretical-experimental fitting from a range of values of hi. Experiments were conducted to analyze the evolution of h_i during solidification of Al-2,4.5,5,8,10,15,33 wt% Cu alloys and Sn-5,10,15,20,30,39 wt% Pb in horizontal and vertical steel chills. The results permitted the establishment of expressions as a power function of time, for different alloy compositions, casting assembly material and thickness of the mold and melt superheat.

NOMENCLATURE

A : area [m^2]
c : specific heat [J/kg.K]
e : mold wall thickness [m]
h_i : metal/mold heat transfer coefficient [$W/m^2.K$]
h_o: mold/environment heat transfer coefficient
h_g : overall heat transfer coefficient
h_w : mold/coolant heat transfer coefficient
k : thermal conductivity [W/m.K]
k_o : partition coefficient
L : latent heat of fusion [J/kg]
q : heat flux [W]
t : time [s]
T : temperature [^oC]
T_F : fusion temperature
T_E : eutectic temperature
T_L : liquidus temperature
T_P : pouring temperature
T_o : cooling water temperature

Greek symbols

ε : emissivity
ρ : density [kg/m^3]
α : thermal diffusivity [m^2/s]

Subscripts

I : interface
S : solid
L : liquid
M: mold
IC : casting surface
IM : mold surface
w : water

INTRODUCTION

For the purpose of accurate mathematical modeling of solidification processes, it is essential that correct boundary conditions be established. The heat transfer at the metal/mold interface is one of these boundary conditions, which is of central importance when considering the magnitude of heat transfer during the early stages of solidification. The way heat flows across the casting and the mold surfaces directly affects the evolution of solidification and plays a notable role in determine the freezing conditions within the casting, mainly in foundry systems of high thermal diffusivity like chill casting. Gravity or

pressure die casting, continuous casting and squeeze castings are some of the process where product soundness is more directly affected by heat transfer at the metal/mold interface. The loss of heat when a metal first comes into contact with the mold is regulated not only by the heat storage capacity of the mold material, but also by the heat transfer conditions within the metal itself and particularly at the metal/mold interface. The solid bodies are only in contact at isolated points and the actual area of contact is only a small fraction of the nominal area, as shown in Figure 1.

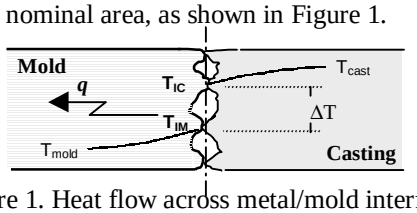


Figure 1. Heat flow across metal/mold interface.

Part of the heat flow follows the paths of actual contact, but the remainder must pass through the gaseous and non-gaseous interstitial media between the surface peaks. The interstices are limited in size, so that convection can be neglected. If temperature differences are not extreme, radiation does not play a significant role and most of the energy passes by conduction across the areas of actual physical contact.

The heat flow across a casting/massive mold interface, as shown by the schematic representation of Figure 1, can be characterized by a macroscopic average metal/mold interfacial heat transfer coefficient (h_i), given by:

$$h_i = \frac{q}{A(T_{IC} - T_{IM})}, \qquad (1)$$

where q is the average heat flux across the interface [W], T_{IC} and T_{IM} are, respectively casting and mold surface temperatures [K] and A is area. In water-cooled molds the overall heat flow is affected by a series of thermal resistances, as shown in Figure 2.

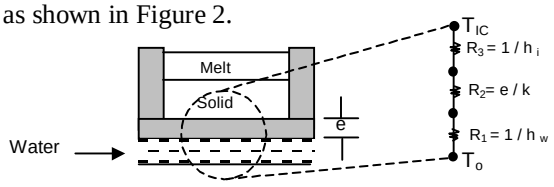


Figure 2. Thermal resistances (cooled mold).

The interfacial resistance between the casting and the mold surface is generally the largest, and the overall thermal resistance ($1/h_g$) is given by:

$$\frac{1}{h_g} = \frac{1}{h_w} + \frac{e}{k_M} + \frac{1}{h_i} \qquad (2)$$

where h_g is the overall heat transfer coefficient between casting surface and the coolant fluid, e is the wall thickness of the bottom of the mold and h_w is the mold/coolant heat transfer coefficient. The average heat flux from casting surface to the cooling water is then given by:

$$q = h_g (T_{IC} - T_o), \qquad (3)$$

where T_o is the temperature of the cooling water.

The present study describes a method for obtaining interfacial heat transfer coefficients as a function of time, from experimental data concerning the solidification of Al-Cu and Sn-Pb alloys in massive and cooled molds. Experimental temperatures in the casting and in the mold are compared with simulations furnished by a numerical model, and an automatic search selects the best fitting from a range of values of interfacial heat transfer coefficients. The effects of alloy composition, material and thickness of the mold and melt superheat are also investigated.

HEAT TRANSFER COEFFICIENT

Several researches have attempted to quantify the transient interfacial heat transfer during solidification in terms of a heat transfer coefficient [1-5]. These studies have highlighted the different factors affecting heat flow across on interface during solidification. These factors include the thermophysical properties of the contacting materials, the casting and mold geometry, the roughness of the mold contacting surface, mold coatings, contact pressure, mold superheat, initial temperature of the mold, etc. The heat transfer coefficient shows a high value in the initial stage of solidification and then decline to a low steady value because the casting contracts from the mold surface, creating an interfacial gap. Most of the methods of calculation of h_i existing in the literature are based on temperature histories at interior points of the casting or mold together with mathematical models of heat transfer during solidification. Among these methods, those based on the solution of the inverse conduction problem have been widely used in the quantification of the transient interfacial heat transfer [6-9].

In the present work, a similar procedure determines the values of h_i, which minimizes an objective function defined by:

$$F(h) = \sum_{i=1}^{n} (T_{est} - T_{exp})^2, \qquad (4)$$

where T_{est} and T_{exp} are, respectively the estimated and experimentally measured temperatures at

various thermocouples locations and times, and n is the iteration stage. The procedure has been detailed by some of the present authors in a previous publication [9].

EXPERIMENTAL PROCEDURE

The casting assemblies used in both horizontal and vertical solidification experiments are shown in Figure 3. Horizontal solidification experiments were performed with Sn, Pb, Al and Al-Cu alloys (4.5; 15; 33wt% Cu) and Sn-Pb alloys (5; 10; 15 and 39wt% Pb), including short and long freezing range alloys, as well as eutectic compositions. The casting and chill materials selected for experimentation, and the employed thermophysical properties are summarized in Table 1. The main design criteria were to ensure a dominant unidirectional heat flow during solidification. This objective was achieved by adequate insulation of the chill and casting chamber. Copper and low carbon steel chills were used, with the heat-extracting surfaces being polished. In order to investigate the influence of chill thickness on heat flow, five different thicknesses of chills were used (X = 6, 17, 28, 39 and 60 mm). Two chromel-alumel thermocouples were introduced in the chill; one near the chill/casting interface and the other at the outer surface, and a third one was placed in the casting and located at 20 mm from the interface, as indicated in Figure 3. Each alloy was melted in an electric resistance-type furnace until the molten metal reached a predetermined temperature. It was then stirred until the temperature was brought to a specified value and poured into the casting chamber. The aluminum alloys were degassed with hexachloroethane tablets before pouring. The effect of liquid metal superheat on heat transfer coefficient was also investigated, by using a Sn-10 wt% Pb alloy, a 60 mm thick carbon steel chill and different degrees of superheat: 20°C, 40°C, 70°C and 100°C above liquidus temperature. The vertical solidification apparatus was designed in such a way that the heat was extracted only through the water-cooled bottom, promoting upward directional solidification. The use of such experimental configuration permits natural convection to be minimized, as well as solute convection due to buoyancy forces if the rejected solute has a higher density than the alloy melt. A stainless steel mold was used having an internal diameter of 50mm, height 110mm and a wall thickness of 5mm. The inner vertical surface was covered with a layer of insulating alumina to minimize radial heat losses, and a top cover made of an insulating material was used to reduce heat losses from the metal/air surface. The bottom part of the mold was closed with a thin (3mm) disc of carbon steel. The alloys were melted in situ and the lateral electric heaters had their power controlled in order to permit a desired superheat to be achieved. To begin solidification, the electric heaters were disconnected and at the same time the water flow was initiated. Experiments were carried out with Al-Cu alloys (2; 5; 8 and 10wt% Cu) and Sn-Pb alloys (5, 10, 15, 20 and 30wt% Pb) at various superheats. Temperatures in the casting were monitored during solidification via the output of a bank of type K thermocouples (1.6 mm diameter) accurately positioned with respect to the heat-extracting surface.

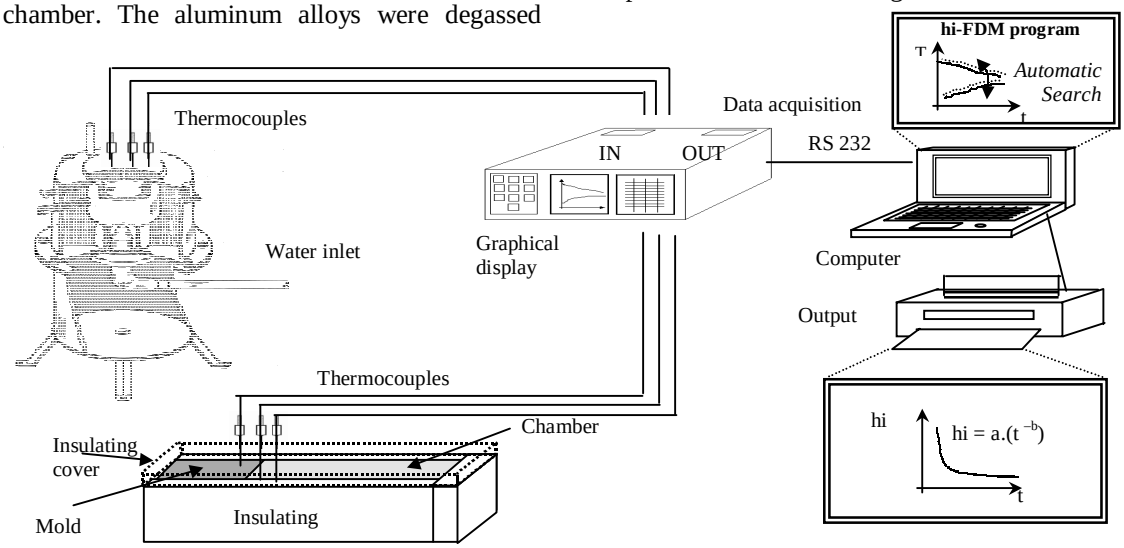


Figure 3. Casting arrangement and position of the thermocouples in mold wall and metal.

Table 1. Casting and chill materials used for experimentation and the corresponding thermophysical properties [9-12].

	Al	Al-2%Cu	Al-4.5%Cu	Al-5%Cu	Al-8%Cu	Al-15%Cu	Al-33%Cu	Steel SAE 1010	Pb	Sn-39%Pb	Sn-20%Pb	Sn-10%Pb	Sn-5%Pb	Sn	Copper
k_S	222	209	193	192.4	188.4	179	155	46	34.7	54.7	59	63	64	67	372
k_L	92	90.7	85	89	87	80	71		29.7	31.7	32	33	33	33	
c_S	1123	1109	1092	1090	1088	1080	1070	527	129.8	186.2	200	209	221	221	419
c_L	1086	1074	1059	1057	1039	999	895		138.2	212.9	231	243	259	259	
ρ_S	2550	2588	2650	2667	2746	2910	3410	7860	11340	8840	8250	7840	7720	7300	8960
ρ_L	2380	2423	2480	2501	2580	2760	3240		10678	8400	7860	7480	7380	7000	
α_S (10^{-5})	7.75	7.28	6.67	6.62	6.31	5.67	4.25		2.37	3.35	3.58	3.84	3.91	4.15	
α_L (10^{-5})	3.36	3.49	3.24	3.36	3.24	2.90	2.45		2.04	1.79	1.76	1.81	1.82	1.82	
L	385000	383566	381900	381415	379264	274270	350000		26205	47560	52580	56140	57120	60710	
T_F	660	660	660	660	660	660	660		327	232	232	232	232	232	
T_E		548	548	548	548	548	548			183	183	183	183		
T_L		654	645	643	633	618					202	216	220		
ε								0.8							0.023
k_o		0.17	0.17	0.17	0.17	0.17					0.0656	0.0656	0.0656		

The thermocouples were calibrated at the melting points of aluminum (for Al-Cu alloys) and Tin (for Sn-Pb alloys), exhibiting fluctuations of about 1.0 ^oC and 0.4 ^oC respectively. The experimental profiles plotted are the averages of three thermocouple reading at each location in chill and casting. Results from repeated experiments have shown differences not greater than 4^oC. All of the thermocouples were connected by coaxial cables to a data logger interfaced with a computer, and the temperature data were acquired automatically. The temperature files were used in a finite-difference heat flow program to estimate the transient heat transfer coefficients.

RESULTS AND DICUSSION

Horizontal Casting:

Effect of Alloy Composition: solidification simulation of each test casting was performed by adopting two different approaches for the liberation of the latent heat of fusion. For eutectic alloys and pure metals, the latent heat (L) was transformed into equivalent number of degrees by considering a temperature accumulation factor (λ) related to L by the specific heat (λ = L/c). For short or long freezing range alloys, the latent heat evolution was taken into account by using Scheil's equation until the remaining liquid reached the eutectic composition. Temperature was experimentally measured in two locations: in the chill at 3 mm from the metal/ mold interface and in the casting at 20 mm from this interface. In Figure 4 and Figure 5 typical experimental thermal responses are compared to those numerically simulated by using the transient h_i profile which provides the best curve fitting for Al-Cu and Sn-Pb alloys.

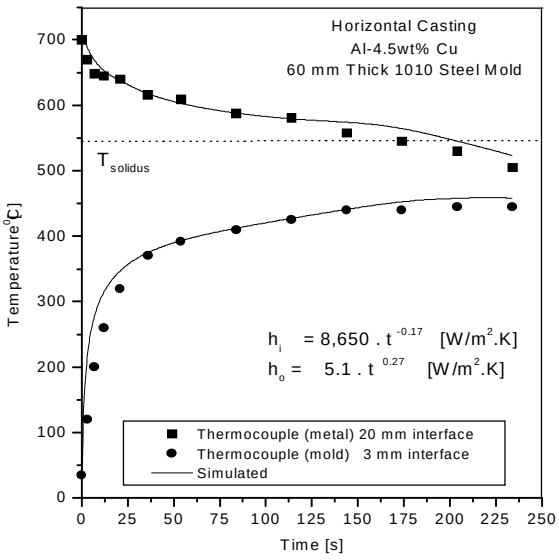


Figure 4. Experimental and simulated temperatures responses at two locations in casting and chill: 60 mm thick steel chill and a superheat $\Delta T = 0.1\ T_L$ (10% of liquidus temperatures) Al-4.5wt% Cu

Figure 6 shows the metal/mold heat transfer coefficients profiles as a function of time for the case of Al-Cu alloys solidifying against a 60 mm thick carbon steel chill.

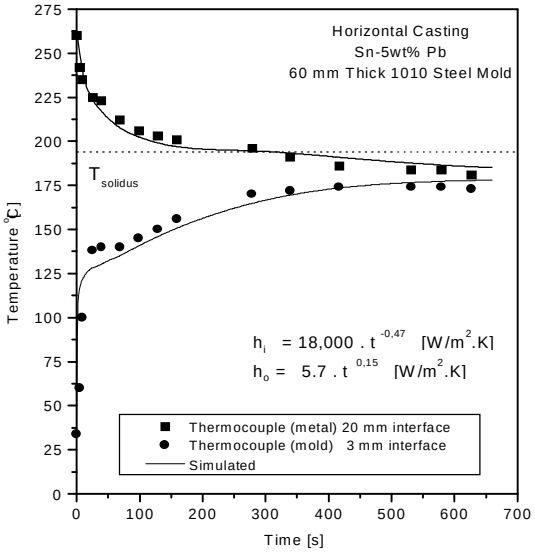


Figure 5. Experimental and simulated temperatures responses at two locations in casting and chill: 60 mm thick steel chill and a superheat $\Delta T = 0.1\ T_L$, Sn-5wt% Pb.

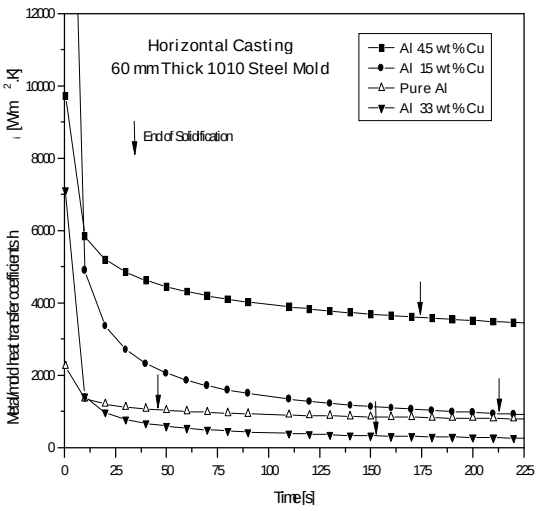


Figure 6. Evolution the metal/mold interfacial heat transfer coefficients as a function of alloy composition: Al-Cu system, 60 mm thick steel chill and $\Delta T = 0.1\ T_L$.

The observed differences in the h_i profiles between the pure metals and the other alloys examined, can be explained by the total shrinkage accompanying solidification, the extent of the solidification range and the wetting of the mold by the melt. It can be seen that the h_i profile increases with increasing mushy zone length. For longer mushy zones, the interdendritic liquid can feed better the solidification contraction causing a continued presence of liquid at the interface, leading to higher values of h_i . This can be taken as a general trend, but care should be exercised when applying this conclusion to the early beginning of solidification. As can be seen in Figure 6, the Al-15wt% Cu alloy exhibits initial h_i values higher than those corresponding to the Al 4.5wt% Cu alloy, which has a longer mushy zone. At the initial stage of solidification the wetting of the mold by the melt seems to be the dominant factor controlling heat transfer coefficient. Anyway a more complex experimental set-up would be necessary for an accurate characterization of initial heat transfer coefficients, as fluid flow in the cast alloys, with its associated heat transfer, would be at its strongest.

Effects of Chill Material and Chill Thickness: the effects of chill material and chill thickness on heat transfer coefficient are shown in Figures 7 and 8, for the cases of a Sn 10wt % Pb alloy solidifying respectively against a carbon steel and copper chills at a superheat of $\Delta T = 0.1\ T_L$.

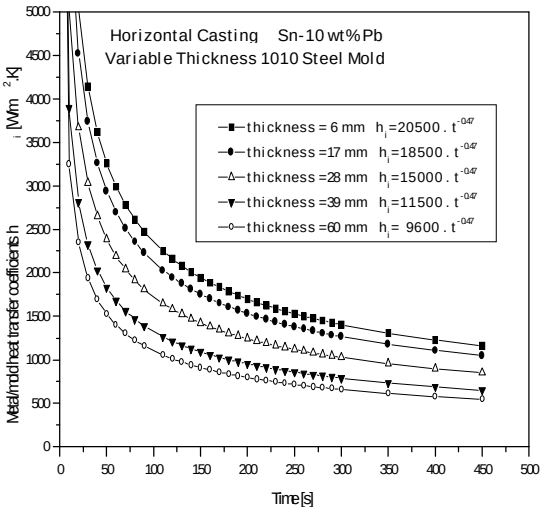


Figure 7. Evolution of the metal/mold heat transfer coefficients as a function of chill thickness: Sn-10 % wt Pb alloy, steel chill and $\Delta T = 0.1\ T_L$.

As can be seen by comparing Figures 7 and 8, the heat transfer coefficient profiles increase with increasing thermal diffusivity of the chill material.

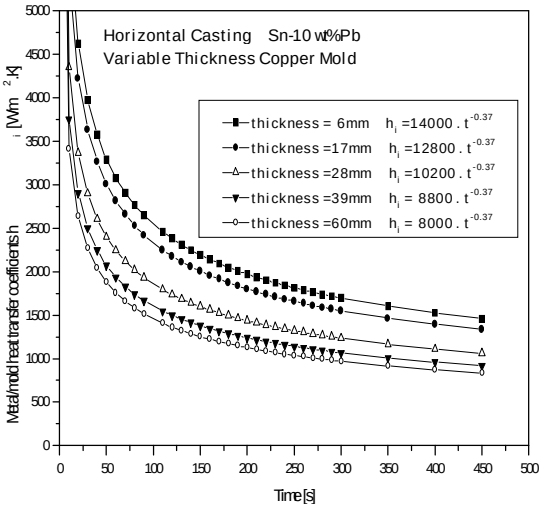


Figure 8. Evolution of the metal/mold heat transfer coefficients as a function of chill thickness: Sn-10 % wt Pb alloy, cooper chill and $\Delta T = 0.1\ T_L$.

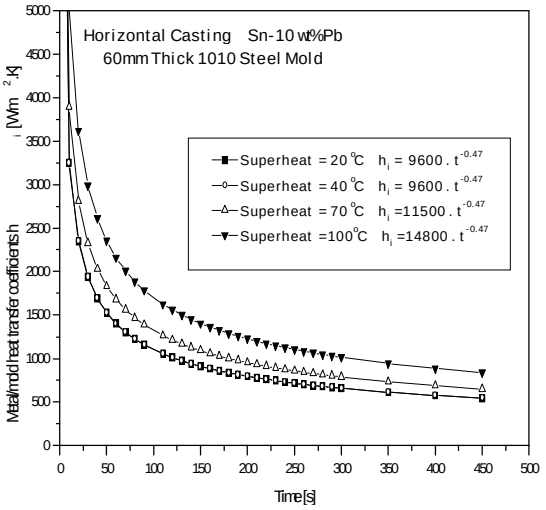


Figure 9. Evolution of the metal/mold heat transfer coefficients as a function of superheat: Sn 10 % wt Pb alloy, 60 mm thick steel chill.

These results are in agreement with other studies in the literature [3, 7]. It can also be seen that the h_i profiles increase with decreasing chill thickness. The chill temperature rises more rapidly from beginning of solidification with decreasing chill thickness. As a consequence, mold expansion favors the thermal contact between metal and chill surface and as the solidified shell is not so thick as for thicker chills, this translates to lower contraction away from the chill. Both factors will contribute to an increase in h_i values.

Effect of Superheat: the heat transfer coefficient increases with increasing values of superheat, as can be seen in Figure 9 for a Sn 10wt% Pb alloy solidifying against a 60 mm thick carbon steel chill. The fluidity of molten alloys increase with increasing superheat, favoring the wetting of the chill by the melt [9,13]. Some results reported in the literature indicate that the surface of solidified shell becomes smoother as the superheat increases for the same chill microgeometry, thus increasing the interfacial contact [14]. Some differences can be observed for values higher than 40° C (Figure 9). This is not the case for Al, where fluidity plays a more significant role. The initial value of h_i rises from about 1500 W/m^{2}K to 6000 W/m^{2}K if the superheat is increased from 10% of the melting point (T_f) to 20% of T_f [9].

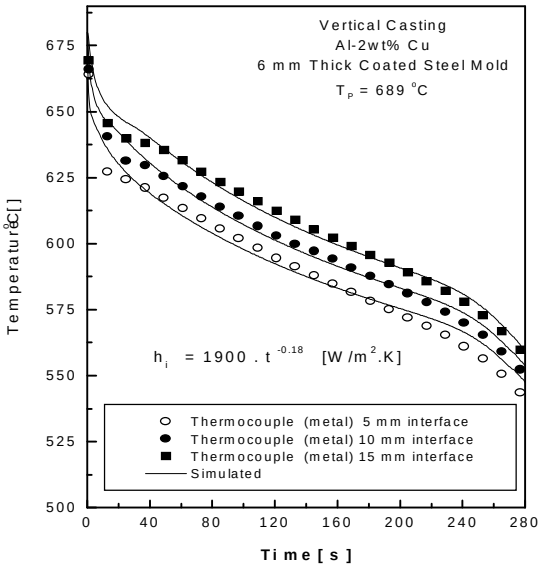
Table 2 summarizes all the values of h_i, expressed as a power function of time, determined during the present experimental investigation for Al-Cu and Sn-Pb alloys under different conditions.

Table 2. Metal/mold heat transfer coefficients during the horizontal directional solidification.

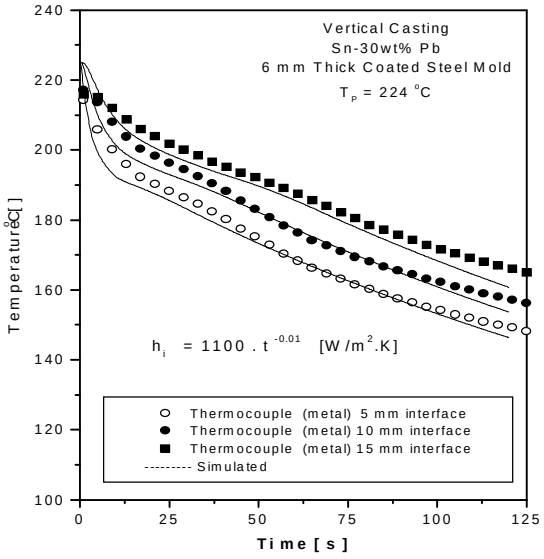
Alloy	*Pouring Temperature*	*Mold condition*		*hi (W/m².K), t (s)*
Al Pure	*726 ºC* *792 ºC*	*1010 Steel mold*		*2000 (t)$^{-0,17}$* *12500 (t)$^{-0,17}$*
Al4.5wt%Cu	*645 ºC*	*1010 Steel mold*		*8650 (t)$^{-0,17}$*
Al15wt%Cu	*618 ºC*	*1010 Steel mold*		*17000 (t)$^{-0,54}$*
Al33wt%Cu	*548 ºC*	*1010 Steel mold*		*4900 (t)$^{-0,54}$*
Sn Pure	*232 ºC*	*1010 Steel mold*		*6800 (t)$^{-0,47}$*
Sn5wt%Pb	*220 ºC*	*1010 Steel mold*		*18000 (t)$^{-0,47}$*
Sn10wt%Pb	*210 ºC*	*1010 Steel mold*		*9600 (t)$^{-0,47}$*
Sn10wt%Pb	*231 ºC*	*1010 Steel mold*	*thickness*	-
			6 mm	*20500 (t)$^{-0,47}$*
			17 mm	*18500 (t)$^{-0,47}$*
			28 mm	*15000 (t)$^{-0,47}$*
			39 mm	*11500 (t)$^{-0,47}$*
			60 mm	*9600 (t)$^{-0,47}$*
		Copper mold	*6 mm*	*14000 (t)$^{-0,37}$*
			17 mm	*12800 (t)$^{-0,37}$*
			28 mm	*10200 (t)$^{-0,37}$*
			39 mm	*8800 (t)$^{-0,37}$*
			60 mm	*8000 (t)$^{-0,37}$*
Sn10wt%Pb	*230 ºC* *250 ºC* *280 ºC* *310 ºC*	*1010 Steel mold*		*9600 (t)$^{-0,47}$* *9600 (t)$^{-0,47}$* *11500 (t)$^{-0,47}$* *14800 (t)$^{-0,47}$*
Sn20wt%Pb	*202 ºC*	*1010 Steel mold*		*8400 (t)$^{-0,47}$*
Sn39wt%Pb	*183 ºC*	*1010 Steel mold*		*7800 (t)$^{-0,47}$*

Vertical Casting:

Figure 10 shows typical examples of the evolution of thermal profiles during the course of different experiments of upward directional solidification of an Al-Cu alloy and a Sn-Pb alloy.



(A)



(B)

Figure 10. Typical experimental temperatures responses at three locations in casting: Al-Cu alloys and Sn-Pb alloys (coated mold): (A) Al-2wt% Cu, (B) Sn-30wt% Pb.

Figure 11 shows a typical example of the evolution of metal/mold heat transfer coefficients (h_i) during the course of different experiments of upward directional solidification of a Al-5wt%Cu alloy.

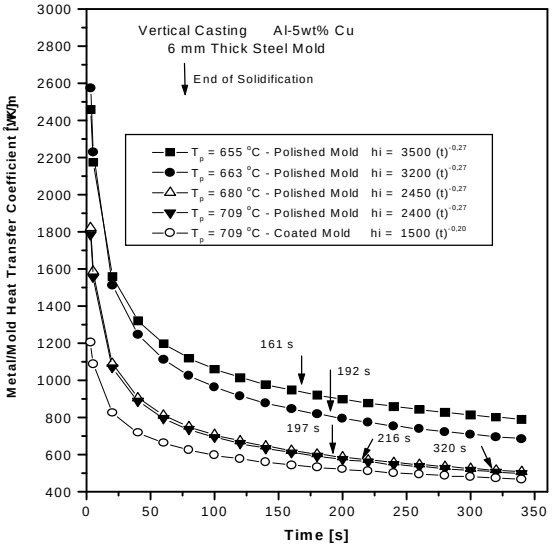


Figure 11. Evolution of the metal/mold heat transfer coefficients as a function of superheat: Al-5wt% Cu, water-cooled steel mold.

Table 3 summarizes all the values of h_i, expressed as a power function of time, determined during the present experimental investigation for Al-Cu and Sn-Pb alloys under different conditions.

Table 3. Metal/mold heat transfer coefficients during the vertical directional solidification.

Alloy	Pouring Temperature T_p (K) and Mold Surface Condition	h_i (W/m²K), t (s)
Al 2wt% Cu	T_p = 660°C – polished mold	$h_i = 2600\ (t)^{-0.14}$
	T_p = 674°C – polished mold	$h_i = 2100\ (t)^{-0.14}$
	T_p = 689°C – polished mold	$h_i = 2000\ (t)^{-0.14}$
	T_p = 709°C – polished mold	$h_i = 1850\ (t)^{-0.14}$
	T_p = 689°C – coated mold	$h_i = 1900\ (t)^{-0.18}$
Al 5wt%Cu	T_p = 655°C – polished mold	$h_i = 3500\ (t)^{-0.27}$
	T_p = 663°C – polished mold	$h_i = 3200\ (t)^{-0.27}$
	T_p = 680°C – polished mold	$h_i = 2450\ (t)^{-0.27}$
	T_p = 709°C – polished mold	$h_i = 2400\ (t)^{-0.27}$
	T_p = 709°C – coated mold	$h_i = 1500\ (t)^{-0.20}$
Al 8wt%Cu	T_p = 640°C – polished mold	$h_i = 5100\ (t)^{-0.33}$
	T_p = 652°C – polished mold	$h_i = 5100\ (t)^{-0.33}$
	T_p = 690°C – polished mold	$h_i = 5100\ (t)^{-0.33}$
		$h_i = 1000\ (t)^{-0.065}$

	T_p = 652°C – coated mold	
Al 10wt%Cu	T_p = 700°C – polished mold	$h_i = 5700\ (t)^{-0.33}$
	T_p = 705°C – polished mold	$h_i = 5700\ (t)^{-0.33}$
	T_p = 705°C – coated mold	$h_i = 1000\ (t)^{-0.075}$
Sn 10wt%Pb	T_p = 220°C – coated mold	$h_i = 1200\ (t)^{-0.01}$
Sn 15wt%Pb	T_p = 234°C – coated mold	$h_i = 1200\ (t)^{-0.1}$
Sn 30wt%Pb	T_p = 224°C – coated mold	$h_i = 1100\ (t)^{-0.01}$
Sn 5wt%Pb	T_p = 225°C – inox mold	$h_i = 650\ (t)^{-0.07}$
Sn 10wt%Pb	T_p = 220°C – inox mold	$h_i = 1300\ (t)^{-0.07}$
Sn 20wt%Pb	T_p = 207°C – inox mold	$h_i = 900\ (t)^{-0.07}$
Sn 30wt%Pb	T_p = 197°C – inox mold	$h_i = 400\ (t)^{-0.07}$

It can be seen both in Table 3 and Figure 11 that as expected, under the same condition of melt superheat, the heat transfer coefficient decreases when the mold surface is coated with an insulating layer of alumina. It can also be observed that the increase in melt superheat decreases the heat transfer coefficient. This is, apparently, in contradiction with results obtained in a previous article concerning the horizontal directional solidification of Al-Cu and Sn-Pb alloys [9] and can be explained by the differences on the physical configuration of the two experimental set-ups. In both cases the superheat delays the solidification evolution. In the horizontal solidification this will translate to a higher h_i profile for higher melt superheats, since the contraction of metal from the mold wall will also be delayed. In the upward solidification the casting weight will contribute to a better metal/mold thermal contact if lateral contraction is effective (permitting the ingot to be gradually detached from lateral walls). This will happen sooner for solidification without superheat, and as a consequence a higher hi profile will be provided with decreasing melt superheat.

CONCLUSIONS

From the experimental work on horizontal and vertical directional solidification of Al-Cu and Sn-Pb alloys, the following conclusions can be made:

- for both the horizontal and vertical casting, a rapid drop in interfacial heat transfer coefficient occurs during the initial stages of solidification;
- the metal/mold (h_i) or the overall heat transfer coefficient (h_g) can be expressed as a power function of time, of the form $h = C_1.(t)^{-m}$, where h [$W/m^2.K$], t [s] and C_1 and m are constants, which depend on alloy composition, chill material and superheat.
- the transient heat transfer coefficients profiles increase with increasing melt superheat for horizontal directional solidification. A reverse situation has been observed for vertical upward directional solidification, where the h_g profiles decrease with increasing melt superheat.

Acknowledgements

The authors would like to acknowledge financial support provided by FAPESP (The Scientific Research Foundation of the State of São Paulo, Brazil) and CNPq (The Brazilian Research Council).

REFERENCES

[1] M. Prates and H. Biloni, Variables Affecting the nature of chill-zone. *Metallurgical Transactions*, v.3A, p.1501, (1972).

[2] K. Ho and R.D. Pehlke, Mechanisms of heat transfer at a metal-mold interface. *AFS Transactions*, v.92, p.587, (1984).

[3] K. Ho and R.D. Pehlke, Metal-mold interfacial heat transfer. *Metallurgical Transactions*, v.16B, p.585, (1985).

[4] A.V. Reddy and C. Beckermann, Measurements of metal-mold interfacial heat transfer coefficients during solidification of Sn and Sn-Pb alloys. *Experimental Heat Transfer*, v.6, p.111, (1993).

[5] J.F. Evans, D.H. Kirkwood and J. Beech, The determination of metal-mold interfacial heat transfer coefficients and the prediction of gross shrinkage cavities in chill mold castings. *The Minerals, Metals & Materials Society*, (1991).

[6] J.V. Beck, Combined parameter and function estimation in heat transfer with application to contact conductance. *Transactions of the ASME*, v.110, p. 1046, (1988).

[7] M. Krishnan and D.G.R. Sharma, Determination of the interfacial heat transfer coefficients h in unidirectional heat flow by Beck's nonlinear estimation procedure. *International Communications on Heat Mass Transfer*, v.23, p.203, (1996).

[8] M.A. Martorano and J.D.T. Capocchi, Heat transfer coefficient at the metal mould interface in the unidirectional solidification of Cu-8% Sn alloys. *International Journal of Heat and Mass Transfer*, v.43, p.2541, (2000)

[7] M.A. Taha, N.A. El-Mahallawy; A.W.M. Assar and R.M. Hammouda, Effect of melt superheat and chill material on interfacial heat transfer coefficient in end-chill Al and Al-Cu alloys casting. *Journal of Materials Science*, v.27, p. 3467, (1992).

[9] C.A. Santos, J.M.V. Quaresma and A. Garcia, Determination of transient interfacial heat transfer coefficients in chill mold castings. *Journal of Alloys and Compounds*, v.319, n.1-2, p.174, (2001).

[10] L.F. Mondolfo, Effect of interfacial energies on heterogeneous nucleation. *Materials Science Technology*, v.5, p.118, (1976).

[11] R.D. Pehlke et alli, Sumary of Thermal Properties for Casting Alloys and Mold Materials. *University of Michigan*, (1982).

[12] Y.S. Toloukian et alli, Thermophysical Properties of Matter. *IFI/Plenum, N.Y.*, v.1, (1970).

[13] T.S.P. Kumar and K.N. Prabhu, Heat flux transients at the casting/chill interface during solidification of aluminum base alloys. *Metallurgical Transactions*, v. 22B, p.717, (1991).

[14] M.C. Flemings, Solidification Processing. *Mc-Graw Hill*, (1974).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

TRANSIENT AND STEADY STATE FREE CONVECTION FROM A HORIZONTAL CYLINDER

A. F. Emery
University of Washington
Seattle, WA 98195-2600
emery@u.washington.edu

ABSTRACT

Estimating parameters in steady state, homogeneous systems is a well developed process, usually involving the least squares technique or the Bayesian statistical approach. When the parameters vary with time or space, the process is more difficult. The Forward Variable method was developed to estimate time varying surface fluxes for inverse heat conduction problems. Electrical and systems engineers tend to utilize the Kalman filter approach in such situations. The two approaches have much in common and the Kalman filter can often be applied to conduction/convection problems. It is particularly useful when multiple parameters vary with time. In addition, it lends itself conveniently to situations in which the measured data are noisy and/or correlated. This paper describes the application of both methods to the problem of transient and steady state free convection from a horizontal cylinder. The problem involves noisy data and an exceptionally strong early time variation of the heat transfer coefficient. Because of the inherent smoothing effect of the Kalman filter, it is possible to estimate the heat flux with greater precision. The paper discusses both approaches and compares the resulting estimates of the parameters.

NOMENCLATURE

A	Surface area
B	Matrix
$E[]$	Expected value
h	Convective Coefficient
$\bar{h}$	Mean value of h
h_{ss}	Steady state h
k	Thermal Conductivity of the air
Nu	Nusselt Number (hD/k)
Q	Heat supplied by heater
Ra	Rayleigh Number
t	time
T_a	Ambient temperature
T_s	Surface temperature
x	Column vector
x^T	Transpose of x
y	State variable vector
$\hat{y}$	Estimated value of y
$\hat{y}_{i\|i-1}$	Conditional expectation
$\sigma()$	Standard deviation
$\mathbf{W}$	Covariance matrix of noise
$\mathbf{\Sigma}$	Covariance matrix of the errors in the state variables

INTRODUCTION

The prediction of parameters in a model of a real process by an inverse technique is not complete unless some estimate of the precision of the estimated parameter can be provided. For parameters which are constants, this is usually obtained through the classical least squares approach. For time varying parameters the precision is usually assessed by comparing the simulated response to the measured response. Unfortunately, frequently one may find that a good agreement between the responses does not imply an acceptable level of precision in the estimated parameter. In addition, most tests of inverse techniques are based upon a simulated noisy response obtained by adding noise to the predicted response under the assumption that the noise is stationary and white. In reality, the measured response often contains non-stationary and colored noise. These effects may be eliminated when estimating constant parameters, but they can cause serious difficulty when treating time or spatially varying parameters. This paper compares the

use of the forward variable approach [1] with the extended Kalman filter method in the estimation of the convective heat transfer coefficient for free convection from a horizontal cylinder.

THE EXPERIMENT

Consider the transient convection from a horizontal cylinder which has a substantial thermal capacitance and whose surface flux thus varies significantly with time. The experiment takes place in a setting with varying air temperature, local drafts, and a strong radiation component. At steady state the heat transfer will be a function of the Rayleigh number (which is proportional to $T_s - T_a$) and the surface emissivity. In addition to estimating the steady state behavior, we are interested in how well such correlations predict the time varying convective flux when the cylinder is heating. Experiments were performed at different heating levels to produce a range of steady state surface temperatures, and thus Rayleigh numbers, to investigate the validity of the steady state correlations reported in the literature [2, 3] and to determine if they were applicable to transient heating.

Table 1
Experimental Conditions

Measurement	Range	σ
Heat Flux, Q/A	2100W/m^2	3W/m^2
T_s	295-417K	0.1K
$T_s - T_a$	2-120K	0.2K

The experiment consisted of measuring the temperature history of a hollow copper cylinder, 33.5 mm OD, 9.33 mm ID and 33.6 cm long. Twelve thermocouples were staked slightly below the surface around and along the cylinder and an electrical heater inserted into the center. The cylinder was suspended in room air and surrounded by a wire mesh, located 30 cm horizontally away from the cylinder to dampen any air flow perturbations. Local air temperature was measured 6 cm horizontally from the cylinder axis. The cylinder surface was covered with a very thin dense layer of soot so that it radiated as a black surface. After the cylinder was determined to be in equilibrium with the ambient air, a constant heater power was applied and the power and temperatures were measured with a data acquisition system at 5 second with a data acquisition system at 5 second intervals. A statistical analysis of the measurements gave the ranges and standard deviations (σ) shown in Table 1.

COMPARISON OF MEASURED AND MODELED RESPONSES

Because a conjugate cfd simulation of the system using FIDAP [4] showed that the temperature in the cylinder was isothermal to within 0.1C, the cylinder was modeled as a lumped parameter transient conduction problem with a time varying convective heat transfer coefficient. Figure 1 compares the measured temperature and surface flux histories for a typical case.

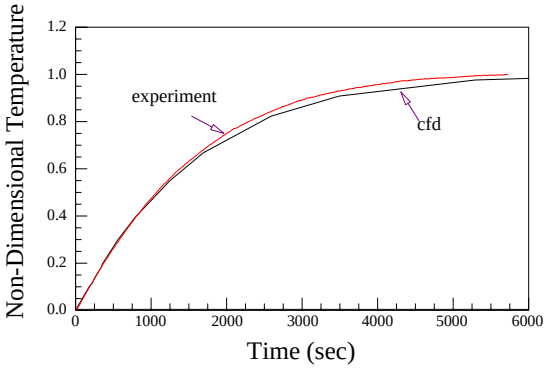


Figure 1a: Temperature History

While the figures suggest that there is good agreement between the measured and the cfd simulation, a closer examination reveals that there are important differences, particularly at early times. These effects are better illustrated by comparing the values of the convective heat transfer coefficient, given in terms of the Nusselt number Nu, as predicted by the forward variable method with the cfd predictions as shown in Figure 2.

We note that the steady state values are in good agreement, but the transient values differ substantially at early times. The good agreement of the cfd results with those obtained from the correlation $Nu = 0.48Ra^{0.25}$ using the measured temperatures suggests that the flow was able to adapt within seconds to the heating and that the correlation would adequately predict the transient behavior except at times less than a few seconds. The question

is then what is the cause of the different behavior of the values obtained from the inverse

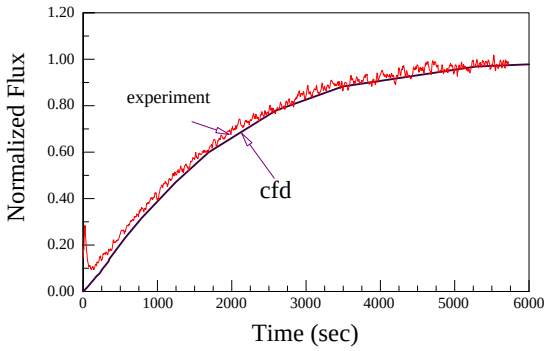


Figure 1b Heat Flux History

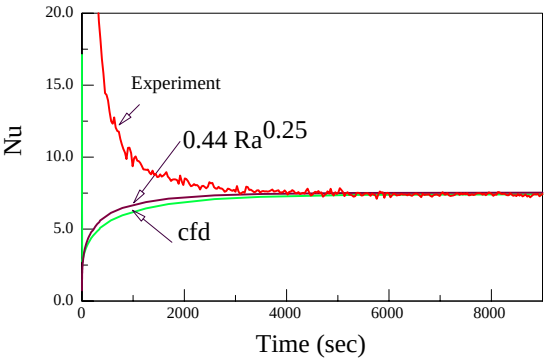


Figure 2: Nusselt Numbers estimated by the forward variable method and from the cfd simulation

solution and whether these differences reflect a basic fault in the model of the process.

WHITE AND COLORED NOISE

Before describing the inverse model in detail, it is important to realize that in almost all studies of the inverse method the noise is assumed to be stationary and white. That is, the mean and standard deviation of the noise are constant with time and the errors are statistically independent. Rarely do the the studies examine the consequences of violating these assumptions. Although the measurements of heater power and temperature were noisy, the most important and dominant noise was in the measured ambient temperature. Figure 3a shows a characteristic history. A spectral analysis [5, 6] of this history showed that the noise was stationary, but decidedly not white. The effect that this 'coloration' of the noise has on the estimated precision is quite dramatic.

We can appreciate this effect by considering how the precision of the heat transfer coefficient at steady state is related to the noise in the measured temperatures. Since h is defined by

$$h_{ss} = \frac{Q_{ss}/A}{T_s - T_a} \tag{1}$$

the precision of the average value $\overline{h}_{ss}$ is given by the law for propagation of errors [7] by

$$\sigma^2(\overline{h}_{ss}) = (\frac{\partial h}{\partial Q})^2\sigma^2(\overline{Q}_{ss}) + (\frac{\partial h}{\partial T_s})^2\sigma^2(\overline{T}_s) + (\frac{\partial h}{\partial T_a})^2\sigma^2(\overline{T}_a) \tag{2}$$

where we have assumed that the errors in Q, T_s, and T_a are independent. Now using the usual least squares approach to determine the mean value of T_a we have

$$\overline{T}_a = \sum_{i=1}^{i=N} T_a(i) \tag{3a}$$

$$\sigma^2(\overline{T}_a) = One^T \mathbf{W}^{-1} One \tag{3b}$$

where One is a vector of N ones, One^T is its transpose, and $\mathbf{W}$ is the covariance matrix of the noise. If the noise is white and stationary, $\mathbf{W}$ is a diagonal matrix with diagonal elements having the constant value of σ^2_{noise} and the standard deviation of the average ambient temperature is given by the usual expression [8]

$$\sigma^2(\overline{T}_a) = \frac{\sigma^2_{noise}}{N} \tag{4}$$

However, if the noise is colored, then $\mathbf{W}$ is a full matrix and the degrading effect of coloration is a significant increase in $\sigma^2(\overline{T}_a)$ from that which would be computed by assuming that the noise is white, i.e. $\mathbf{W}$ is diagonal. Figure 3b shows autocorrelation coefficient of the ambient temperature shown in Figure 3a. Values of ρ less than 0.2 in absolute magnitude are usually taken to mean that there is minimal autocorrelation [9]. From the figure, it is apparent that a significant autocorrelation, i.e., coloring, exists. An analysis of the information content shows that it is equivalent to that contained in white noise sampled at $1/5^{th}$ the rate. The lower curve in Figure 3b shows

the effect of this lower sampling rate. Table 2 gives the respective values of the standard deviations of $\overline{T}_a$ computed using the full

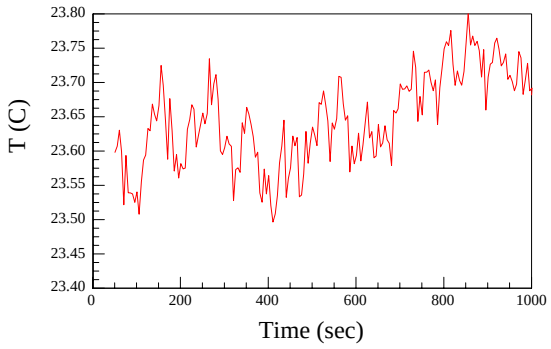


Figure 3a: Typical Time History of the Ambient Temperature

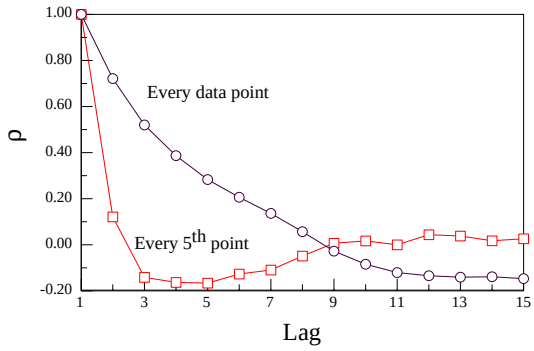


Figure 3b: Autocorrelation Coefficient for the Ambient Temperature shown in Figure 3a

matrix **W** (i.e. correlated noise) and that computed by assuming that the errors were uncorrelated (a diagonal **W**).

Table 2
Standard Deviation of the Average T_a

sampling rate	full **W**	diagonal **W**	ratio
Every point	0.1266	0.0567	2.23
$1/5^{th}$	0.0520	0.0513	1.01

Assuming that the noise is white substantially underestimates $\sigma(\overline{T}_a)$ when every data point is used because of the high degree of correlation. On the other hand, when a reduced sampling rate is used, the noise is essentially uncorrelated. The consequence of ignoring the effects of correlation is that $\sigma(h_{ss})$ would be underestimated by over a factor of 2.

As described in the following section, the least squares solution for h requires a knowledge of **W**. Because it adds significantly to the computational costs of the least squares method and because it is neither easy nor convenient to evaluate [10, 11], every effort should be made to made to ensure that the errors are not correlated. Fortunately, reducing the sampling rate is often an effective way to eliminate the correlation as shown above. Of course this means then that rapidly varying parameters may not be well estimated.

STEADY STATE RESULTS

The history of Nu shown in Figure 2 was determined with the forward variable method which is designed to estimate time varying parameters. Instead of choosing $h(t_i)$ to match the measured response at time t_i to the model solution, the value of $h(t_i)$ is chosen to minimize the variance of the error in $\hat{h}$ found by a least squares solution using the responses measured at $t_{i-r_b}, t_{i-r_b+1}, ...t_i, t_{i+1}, ...t_{i+r_f}$

Although the points used are usually at future times, $r_b = 0, r_f > 1$, one could just as easily use a combination of forward and backward points. r is an adjustable parameter and its optimal value often varies throughout the entire history of $h(t)$ but it is generally more convenient to use a constant value [1]. Since the noise is assumed to be stationary and white, the procedure can be applied without a knowledge of $\sigma(noise)$, and in fact the solution is independent of the noise which only enters into the final estimate of the standard deviation of the estimated parameter. The method is a form of data smoothing. The steady state values of Nu were best fit with $Nu = 0.44Ra^{0.25}$ which differs from the published correlation by $\approx 10\%$ and from the correlation of Chu and Churchill [2] by less than 5%.

Figure 4 compares the standard deviations of h_{ss} as a function of the number of points, r, sampled in the forward variable approach. Although the forward variable method was designed for varying parameters, it proves to be effective when there is noise and, in our case, when the ambient temperature is constantly changing throughout the experiment. Increasing r reduces the uncertainty, but of course increases the computational cost.

The reduction in $\sigma(h_{ss})$ is dramatic. The method as originally proposed assumes that the noise is white, i.e., all points are equally weighted, which we know from Figure 2b is not the case. Using every 5^{th} point as suggested by Figure 2b, gives the results shown by the lower curve on Figure 4 and the significant improvement in the precision of the estimate is clearly seen. Not only is the uncertainty in h_{ss} reduced, but the computational cost is much less. Note that the reduction in $\sigma(h_{ss})$ is in line with the results given in Table 2.

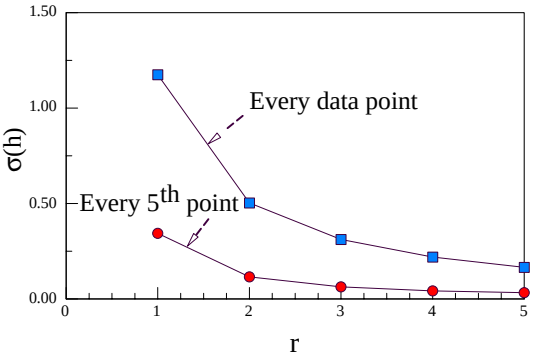


Figure 4: $\sigma(h_{ss})$ as a function of the forward variable parameter r

TRANSIENT RESULTS

Using either the values of h derived from the cfd simulation or assuming a constant value over a short time, it can be shown that the sensitivity, $\partial T/\partial h$, is zero at time zero and slowly increases to a maximum at steady state. Because of the very small sensitivity at short times, the experiment is not suitable for estimating $h(t)$ at the start of heating. However, by 100 seconds, the sensitivity has increased to $0.1C/(watt - C/m^2)$ which with the value of $\sigma(T_s - T_a)$ of 0.2C gives a value of $\sigma(Nu) \approx 2$, and increasingly smaller values as time increases. The imprecision is of the order of the oscillations shown in the time history of h in Figure 2, giving a range of h not much different than that shown in the figure. This being so, one would conclude that the experimentally determined values of h are statistically different from the cfd simulation results. The obvious question is why? A possible approach to answering this is through the Kalman filter method in which the history of $\sigma(h)$ is tracked and in which multiple parameters or other possible effects are easily incorporated. The idea behind the Kalman filter is to estimate a value of the parameter at time t_i based upon all of the data up to and including that at time t_j. If $j < i$ the process is termed *prediction*, if $j = i$ *filtering*, and if $j > i$ *smoothing*. In essence the forward variable method is really the Kalman smoothing filter approach in which the covariance matrix is assumed to be diagonal with constant elements and the past data are downgraded in importance.

Consider the determination of a constant parameter using the least squares approach with N data points. Usually, the data is treated in a batch mode, i.e., all N data points are considered at once. Assuming that the temperature can be expressed as a function of the convective coefficient, h, as

$$T_i = F(\overline{h}) + \frac{\partial F}{\partial h}|_{\overline{h}}(h - \overline{h}) \tag{5}$$

$$\hat{h} = \overline{h} + (A^T \mathbf{W}^{-1} A)^{-1} A^T \mathbf{W}^{-1} \{T - F(\overline{h}\} \tag{6a}$$

$$\sigma(\hat{h}) = (A^T \mathbf{W}^{-1} A)^{-1} \tag{6b}$$

where A is a N dimensional vector with components of $\frac{\partial F}{\partial h}|_{\overline{h}}$, $\{T - F(\overline{h}\}$ is the vector of measured temperatures and $\mathbf{W}$ is the covariance matrix of the noise. While the solution of Eq. 6 gives the estimate of h and of $\sigma(h)$ based on all N data points, the solution involves substantial computational expense when N is large and it is not easy to monitor how the error behaves as new data points are added. The recursive least squares approach was designed to allow a new datum point to be considered without inverting the entire matrix. It also permits one to monitor the behavior of $\sigma(h)$ as N increases.

The Kalman filter is an extension of the usual recursive least squares approach based upon state variables which has been found in its usual formulation to handle slowing varying parameters. Details are given in a number of excellent texts [10, 11]. Let the state of the system at time t_i be defined as the m component vector, y_i. In this study, $y_i^T = [T_i, h_i]$ where T_i is the temperature and h_i is the convective heat transfer coefficient at time t_i. y_i is given in terms of y_{i-1} by the equation

$$\{y_{i+1}\} = \begin{pmatrix} T_{i+1} \\ h_{i+1} \end{pmatrix} = \begin{pmatrix} T_i + \frac{dT}{dt}|_i \Delta t \\ h_i \end{pmatrix} = \mathbf{A}_i\{y_i\} + \mathbf{B}_i f_i + \mathbf{C}_i w_i \tag{7a}$$

where f_i is a forcing function (in this case the heater power) and w_i is the process noise with a covariance matrix of $\mathbf{Q}_i$. The measured response, z_i, is given by

$$z_i = \mathbf{D}\ y_i + n_i \qquad (7b)$$

where n_i is the measurement noise which has a covariance of $\mathbf{R}_i$. Both noises are assumed to be white i.e., $\mathbf{Q}$ and $\mathbf{R}$ are diagonal matrices, and to be uncorrelated with each other. Let $\hat{y}_{i|i-1}$ represent our estimate of the system state at time i based upon measurements up to time *i-1*. The procedure consists of the following steps

a) a prediction of the response $\hat{y}_{i|i-1}$ based upon $\hat{y}_{i-1|i-1}$.

$$\hat{y}_{i|i-1} = \mathbf{A}_{i-1}\hat{y}_{i-1|i-1} + \mathbf{B}_{i-1}f_{i-1} + \mathbf{C}_{i-1}w_{i-1} \qquad (8a)$$

b) and a correction

$$\hat{y}_{i|i} = \hat{y}_{i|i-1} + \mathbf{G}_i(z_i - \mathbf{D}\hat{y}_{i|i-1}) \qquad (8b)$$

The gain matrix, $\mathbf{G}_i$ is chosen to minimize the mean square errors of the state variables and is recursively computed from

$$\begin{aligned} \mathbf{P}_{i|i-1} &= \mathbf{A}_{i-1}\mathbf{P}_{i-1|i-1}\mathbf{A}_{i-1}^T + \mathbf{C}_{i-1}\mathbf{Q}_{i-1}\mathbf{C}_{i-1}^T \\ \mathbf{G}_i &= \mathbf{P}_{i|i-1}\mathbf{D}_i^T(\mathbf{D}_i\mathbf{P}_{i|i-1}\mathbf{D}_i^T + \mathbf{R}_i)^{-1} \\ \mathbf{P}_{i|i} &= (\mathbf{I} - \mathbf{G}_i\mathbf{D}_i)\mathbf{P}_{i|i-1} \end{aligned} \qquad (9)$$

where $\mathbf{P}_{i|i}$ is the conditional covariance matrix of the error in the estimated state variable

$$\mathbf{P}_{i|i} = E(\{y_i - \hat{y}_{i|i}\}\{y_i - \hat{y}_{i|i}\}^T) \qquad (10)$$

Figure 5a illustrates the histories of $\sigma(T)$ and $\sigma(h)$ corresponding to the Nu shown in Figure 2. Both the estimated temperatures and the heat transfer coefficients show acceptably low standard deviations. Based upon these results, one would conclude that the estimated history of Nu was correct. However, this estimated history is at variance with the cfd simulations and requires an explanation.

The most probable cause of the differences in h is an inadequacy in the model to represent the early time behavior. Even if the early time value of h were of the order of 10 times the steady state, the Biot number would still be of the order of 0.01, validating the use of the lumped capacitance model to estimate h. Since the value of $h(t)$ is directly proportional to the

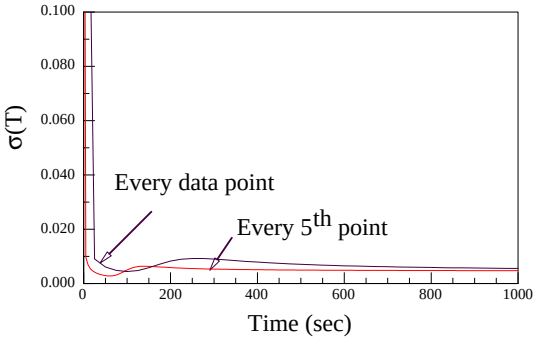


Figure 5a: Time history of $\sigma(T)$

The most probable cause of the differences in h is an inadequacy in the model to represent the early time behavior. Even if the early time value of h were of the order of 10 times the steady state, the Biot number would still be of the order of 0.01, validating the use of the lumped capacitance model to estimate h. Since the value of $h(t)$ is directly proportional to the surface flux which is the difference between the heat supplied to the cylinder by the heater, Q, and that stored in the cylinder, an error in estimating either quantity would give erroneous values of h.

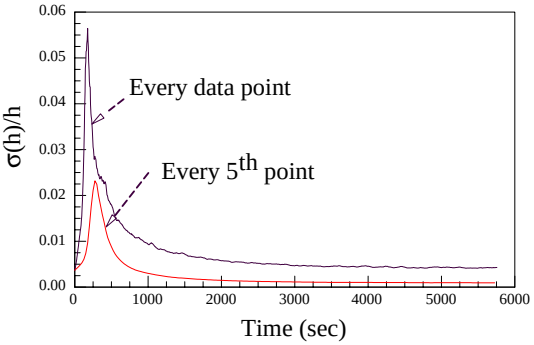


Figure 5b: Time history of $\sigma(h)$

One of the convenient features of the Kalman filter approach is the ease with with additional state variables can be included. In our case, an augmented state variable, $y_i^T = [T_i, h_i, L_i, S_i]$, where L_i represents a heat loss and S_I represents an error in the heat stored was used. Figure 6 compares the estimated surface heat flux based upon the original solution and that including either L or S. We note how much better the fluxes agree when a loss variable is included, suggesting that the heat supplied

to the cylinder was in error especially at early times.

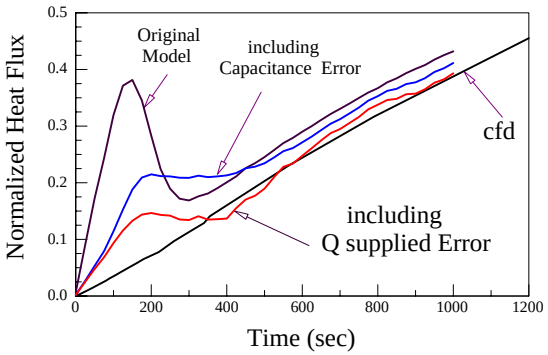


Figure 6: Estimated Surface Heat Flux showing the effect of inaccuracies in the heat supplied and stored

Since there is a contact resistance between the heater and the cylinder, it is possible that at early times the measured heater power does not represent the heat into the cylinder. Instead, the heater may be significanlty hotter than the cylinder, thus storing heat in itself and losing heat out through the ends.

Both effects had been anticipated and hopefully mitigated by insulating the ends of the cylinder and using a thermal grease at the interface between the heater and cylinder. Thermal grease usually contains metallic particles (often copper) and is of very high viscosity, thus making it difficult to insert the heater when copiously used. A new cylinder was constructed and sufficient grease applied that a hydraulic press was required to insert the heater. In addition, thermocouples were installed at the heater/cylinder interface to more accurately characterize the temperature associated with the stored energy.

Figure 7 depicts the estimated heat flux and the time history of $\sigma(h)$ for the new system. The experimental results are now in much better agreement with the cfd simulation, compare with Figure 2, except at very early times. Comparable results were found using the forward variable method. As expected, the standard deviations are reduced when the sampling rate is reduced, but of course doing so reduces the ability to represent the early transient effects.

Could better results be obtained at the earlier times to validate the cfd simulations? Theoretically yes, but realistically no. At very early times the sensitivity to h is small and the noise in the difference, $T_s - T_a$, is large compared to the difference, leading to very high values of $\sigma(h)$. Furthermore, to get values of h at times in the range of 10 seconds would require a very high sampling rate, leading to a high degree of correlation and a further increase in the uncertainty.

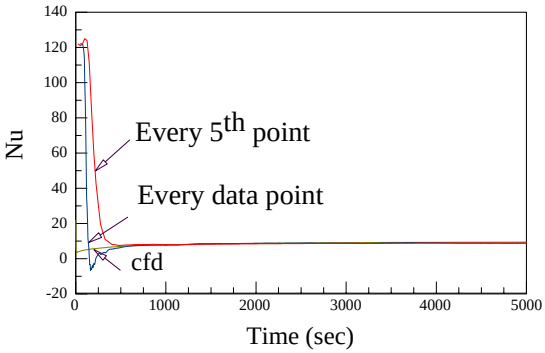


Figure 7a: History of h for the new system computed using the Kalman Filter

The Kalman filter can account for colored noise, but only with a high computational cost for analyzing the noise and including it in such a way that the filtering estimator yields a stable computation.

Furthermore, unless a strong fading memory is used, this approach reflects the total past history leading to a transient response which would be too slow for this specific problem. In addition, one would be required to very accurately measure the time history of the heater power and to eliminate, or at least to accurately determine, the heat losses associated with the momentarily high heater temperatures.

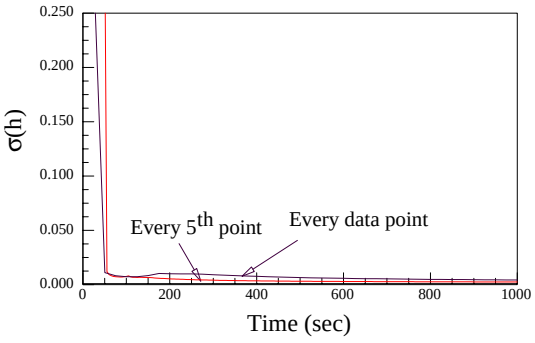


Figure 7b: History $\sigma(h)$ for the new system computed using the Kalman Filter

CONCLUSIONS

The use of the Kalman filter made it possible to track the history of the uncertainties and to determine the times when the heat transfer coefficient was most poorly estimated. By then expanding the state variables to include losses and imprecision in the estimate of the stored energy we were able to recognize that the difficulty was most probably due to interfacial resistance. When the resistance was reduced, substantially better results were obtained.

The Kalman filter approach is not a total panacea for problems like this. While it allows for the inclusion of fading memory, variations in how the terms in the alogrithm, Eq. 9, are evaluated, and including varying degrees of smoothing, the method is very much ad-hoc in nature. Since all of these are choices made by the analyst, usually based upon a subjective interpretation of the results, there are no well defined procedures or criteria for making them.

In analyzing the results and predicting the history of h, both the forward variable and Kalman filter approaches estimated the temperature history of the cylinder very well as evidenced by the very small values of σT, being in the order of 0.05C. Looking at the time history of $\sigma(T)$ and $\sigma(h)$, one would conclude that the inverse estimation of h was accurate unless a reason to suspect the results was known. In this case, the results of the cfd simulation and the correlation, makes it clear that the inverse estimation of the first experiment, and even that of the second at early times, are significnatly in error. Once again reinforcing the idea that one must know much about the problem and must design the experiment carefully before conducting it.

ACKNOWLEDGEMENT

This work was conducted under the sponsorship of the Sandia National Laboratories, Albuquerque, NM. The author wishes to acknowledge the support of the laboratory and the helpful assistance of Drs. S. Kempka, B. Blackwell, and K. Dowding.

REFERENCES

1 Blanc, G., Beck, J. V., Raynaud, M., "Solution of the Inverse Heat Conduction Problem with a Time-Variable Number of Future Temperatures," *Numerical Heat Transfer, B,* pp 437:451, 1997

2 Incropera, F. and DeWitt, D., Introduction to Heat Transfer, J. Wiley and Sons, Publ., New York, NY, 2001

3 Kakac, S., Shah, R. K. and Aung, W., Handbook of Single- Phase Convective Heat Transfer, J. Wiley and Sons, Publ., New York, NY, 1987

4 FIDAP-8, Fluent Corporation, NH, 1988

5 Shiavi, R., *Introduction to Applied Statistical Signal Analysis,* Academic Press, N.Y., NY, 1999

6 Bendat, J. S. and Piersol, A. G., *Random Data: Analysis and Measurement Processes,* J. Wiley and Sons, N.Y., NY, 2000

7 Coleman, H. W. and Steele Jr., W. G., *Experimentation and Uncertainty Analysis for Engineers,* J. Wiley and Sons, N. Y., NY, 1989

8 Stark, H. and Woods, J. W., *Probability, Random Processes, and Estimation Theory for Engineers,* Prentice Hall, N.Y., NY, 1994

9 Johnston, J., *Econometric Methods,*, McGraw-Hill, N.Y., NY, 1972

10 Chui, C. K. and Chen, G., *Kalman Filtering with Real Time Applications,* Springer Publ., N.Y., NY, 1999

11 Anderson, B. D. O. and Moore, J. B., 1979, *Optimal Filtering,* Prentice-Hall Publ. N.Y., NY

4^{th} *International Conference on Inverse Problems in Engineering*
Angra dos Reis, Brazil, 2002

THE NUMERICAL SOLUTION OF THE BACKWARD HEAT CONDUCTION EQUATION USING ARBITRARY PRECISION ARITHMETIC

Kentaro Iijima
Graduate School of Science and Engineering,
Ibaraki University, Mito 310-8512, Japan
nm0101a@mcs.ipc.ibaraki.ac.jp

Keisuke Hayashi
Graduate School of Science and Engineering,
Ibaraki University, Mito 310-8512, Japan
ghayashi@mcs.ibaraki.ac.jp

Kazuei Onishi
Department of Mathematical Sciences,
Ibaraki University, Mito 310-8512, Japan
onishi@mito.ipc.ibaraki.ac.jp

ABSTRACT

We consider a numerical method to find a temperature $u(\boldsymbol{x},t)$ of the backward heat conduction equation

$$\frac{\partial u}{\partial t} = \Delta u$$

in a bounded two-dimensional domain D and for $\bar{T} \le t < 0$ from given temperature $u(\boldsymbol{x},0)$ at the past time $\bar{T} < 0$ with prescribed temperature on the boundary $u(\boldsymbol{x},t)$ for $\boldsymbol{x} \in \Gamma = \partial D$ and $t \in [\bar{T},0]$. This is a typical severely ill-posed problem.

We apply the finite difference method for the approximation of the backward heat equation and we use arbitrary precision arithmetic in numerical computations in order to solve this backward heat conduction problem numerically. By the term *arbitrary precision arithmetic*, we mean a computer arithmetic system in which numbers of figures as large as required are taken for numerical computation in order to keep the significant digits in the arithmetic.

By using a few numerical experiments on simple examples, we observed the relationship between the number of figures and the error of numerical solutions. We examined the dependence of the rate of growth of errors on the number of figures used in the arithmetic.

We conclude that the problem is solved over a longer time interval when the number of figures is increased in the arbitrary precision arithmetic.

INTRODUCTION

Let $D \subset \boldsymbol{R}^2$ be a bounded domain with boundary $\Gamma = \partial D$ of piecewise C^2-class. Let $\bar{T} < 0$ be a negative constant. Our problem consists of finding a function $u \in C(\bar{D} \times [\bar{T},0])$ that is twice continuously differentiable with respect to the space variable $\boldsymbol{x}$ and continuously differentiable with respect to the time variable t in $D \times [\bar{T},0)$ and that satisfies

$$\frac{\partial u}{\partial t} = \Delta u \quad \text{in} \quad D \times [\bar{T},0) \tag{1}$$

$$u(\cdot,0) = w \quad \text{in} \quad D \tag{2}$$

$$u = f \quad \text{on} \quad \Gamma \times [\bar{T},0]. \tag{3}$$

Here, $w \in C(\bar{D})$ and $f \in C(\Gamma \times [\bar{T},0])$ are given continuous functions subject to the compatibility condition

$$f(\cdot,0) = w \ \text{ on } \ \Gamma.$$

This problem is called a backward heat conduction problem. It is known that the backward heat conduction problem is unstable [4], [5]. That is, the change of solution u may be very large, although the change of given w is small. Because of this instability, when we solve the backward heat conduction problem numerically, round-off error and discretization error have great influence on the numerical solution. In order to solve unstable problems, investigations of regularization schemes have been done up to present [3]. Recently, new investigations were started by using infinite precision numerical simulation [1], whose method

reduces the round-off error by arbitrary precision arithmetic, and reduces the discretization error by spectral collocation method arbitrarily. In the paper [2] discretization error and data error of the solution of the backward heat conduction problem are considered.

In order to reduce round-off error, it is effective that we use arbitrary precision arithmetic as a numerical method to solve this problem. In problems where no discretization error is involved, this is especially true.

INSTABILITY OF THE PROBLEM

The following theorem will describe the instability of the solution of the backward heat conduction problem.

Theorem 1 ([4])
Assume that the initial data w and boundary data $f = 0$ are given. If u solves the backward heat conduction problem (1) – (3), then it satisfies

$$\|u(\cdot,t)\| \geq \|w\| \exp\left(\frac{\int_D |\nabla w(\boldsymbol{x})|^2 d\boldsymbol{x}}{\int_D |w(\boldsymbol{x})|^2 d\boldsymbol{x}}(-t)\right), \quad t \in [\bar{T}, 0]$$

where $\|\cdot\|$ is L^2-norm in D.

From this theorem, we mention two remarks:

1. In the backward heat conduction problem (1)–(3), the solution u does not depend continuously on the data w. For example, we set $D := (0,1) \times (0,1)$ and

   $$w_n(x,y) := \frac{1}{n}\sin(n\pi x)\sin(n\pi y), \ \ n \in \boldsymbol{N}.$$

   When initial data is given as w_n and boundary data is given as $f = 0$, we can see that u_n is a unique solution of the problem (1)–(3). By Theorem 1, we obtain

   $$\|u(\cdot,t)\| \geq e^{2(n\pi)^2(-t)}\|w\|. \qquad (4)$$

   From the estimate (4), for all $C > 0$ there exists $n_0 \in \boldsymbol{N}$ such that

   $$\begin{aligned} \|u_{n_0}(\cdot,t)\| &\geq e^{2(n_0\pi)^2(-t)}\|w_{n_0}\| \\ &> C\|w_{n_0}\|, \quad t \in [\bar{T}, 0]. \end{aligned}$$

   Therefore u_n ($n \in \boldsymbol{N}$) do not depend on w_n continuously.

2. The solution u is increasing exponentially with respect to $-t$.

DISCRETIZATION

In this section, in order to solve the backward heat conduction problem numerically, we discretize the backward heat conduction problem using the finite difference method. We obtain an explicit difference scheme, due to the discretization by backward difference, with respect to time direction. In numerical computation by the explicit difference scheme, we do not need to solve simultaneous linear equations. Therefore, we are able to reduce the great workload needed in computation when using arbitrary precision arithmetic.

After this section, let D be the square domain $(0,1) \times (0,1)$ which can be easily discretized by the finite difference method.

Discretization by Backward Difference

Let D be the square domain $(0,1) \times (0,1)$. We divide the domain D into n^2 equal parts and define lattice points (x_i, y_j). Then step size is $h = \frac{1}{n}$ and lattice points (x_i, y_j) are given by $(x_0, y_0) := (0,0)$, $x_{i+1} := x_i + h$, $y_{j+1} := y_j + h$, where $i,\ j = 0, 1, \cdots, n-1$. Let Δt be positive time step size. Then we set the k-th time step t_k as $t_0 := 0$, $t_{k+1} := t_k - \Delta t$, where $k = 0, 1, 2, \cdots$. Let u_{ij}^k be the value of an approximation to the exact solution u of the problem (1)–(3) at (x_i, y_j, t_k). Here, for a discretization we use the backward difference with respect to time direction, and the central difference with respect to both x direction and y direction. Then the difference scheme is described by

$$\begin{aligned} \frac{u_{ij}^k - u_{ij}^{k+1}}{\Delta t} &= \frac{u_{i+1j}^k - 2u_{ij}^k + u_{i-1j}^k}{h^2} \\ &\quad + \frac{u_{ij+1}^k - 2u_{ij}^k + u_{ij-1}^k}{h^2}, \end{aligned} \qquad (5)$$

where $i,\ j = 1, 2, \cdots, n-1$; $k = 0, 1, 2, \cdots$. Moreover, when the equation (5) is solved with respect to u_{ij}^{k+1}, we can get explicit difference scheme

$$\begin{aligned} u_{ij}^{k+1} &= -\rho u_{ij-1}^k - \rho u_{i-1j}^k + (1+4\rho)u_{ij}^k \\ &\quad -\rho u_{i+1j}^k - \rho u_{ij+1}^k, \end{aligned} \qquad (6)$$

where $i,\ j = 1, 2, \cdots, n-1;\ k = 0, 1, 2, \cdots$, and $\rho = \dfrac{\Delta t}{h^2}$.

Compatibility of the Difference Scheme

A difference scheme must tend to the original differential equation which should be approximated, when step size Δt and h tend to 0. The above requirement is given as the compatibility of the difference scheme.

Definition 1 (Compatibility)
Let L be a linear differential operator. Suppose that a partial differential equation is given by

$$\frac{\partial u}{\partial t} = Lu, \tag{7}$$

and the corresponding difference scheme is given by

$$u^{k+1}(x, y) = S(h, \Delta t)u^k(x, y). \tag{8}$$

Between step sizes Δt and h, we set a relation $h = h(\Delta t)$, where $h(\Delta t) \to 0$ as $\Delta t \to 0$. Let a solution $u(x, y, t)$ of the differential equation (7) satisfy

$$u(x, y, t + \Delta t) - S(h(\Delta t), \Delta t)u(x, y, t) = o(\Delta t)$$

$$\Delta t \to 0.$$

Then we define that a difference scheme

$$u^{k+1}(x, y) = S(h(\Delta t), \Delta t)u^k(x, y)$$

is compatible with the differential equation (7).

Let u be a solution of our backward heat equation, and $\rho = \frac{\Delta t}{h^2}$ be a positive constant. For each $t \in [\bar{T}, 0)$, we set

$$\begin{aligned} & S(h, \Delta t)u(x, y, t) \\ := \ & -\rho u(x, y - h, t) - \rho u(x - h, y, t) \\ & +(1 + 4\rho)u(x, y, t) - \rho u(x + h, y, t) \\ & -\rho u(x, y + h, t), \end{aligned} \tag{9}$$

and $u_{ij}^k := u^k(x_i, y_j)$ by a solution u^k of the difference scheme

$$u^{k+1}(x, y) = S(h, \Delta t)u^k(x, y).$$

Then u_{ij}^k satisfies the difference scheme (6). We show that

$$\begin{aligned} & \frac{u(x, y, t) - u(x, y, t - \Delta t)}{\Delta t} \\ & - \frac{u(x - h, y, t) - 2u(x, y, t) + u(x + h, y, t)}{h^2} \\ & - \frac{u(x, y - h, t) - 2u(x, y, t) + u(x, y + h, t)}{h^2} \\ = \ & \frac{\partial u}{\partial t}(x, y, t) + o(1) - \frac{\partial^2 u}{\partial x^2}(x, y, t) + o(1) \\ & - \frac{\partial^2 u}{\partial y^2}(x, y, t) + o(1) \\ = \ & \frac{\partial u}{\partial t} - \Delta u(x, y, t) + o(1) \\ = \ & o(1), \qquad \Delta t \to 0,\ t \in [\bar{T}, 0]. \end{aligned}$$

When this expression is solved with respect to $u(x, y, t - \Delta t)$, we get

$$u(x, y, t - \Delta t) \ = \ S(h, \Delta t)u(x, y, t) + o(\Delta t).$$

Therefore the difference scheme (6) is compatible with the heat equation (1).

Instability of the Difference Scheme

Although a difference scheme is compatible with original differential equation, conditions are also desired where a solution of a difference scheme does not diverge as $\Delta t \to 0,\ h \to 0$.

Definition 2 (Stability)
We call that a difference scheme (8) is stable if there exists a constant $C > 0$ such that $\|S^k\| \le C$ for all k and Δt which satisfy $\bar{T} \le -k\Delta t \le 0$. Here, for a bounded operator $A : C(\bar{D}) \to C(\bar{D})$ the operator norm $\|\cdot\|$ is defined by

$$\|A\| := \sup_{\varphi \in C(\bar{D}) \setminus \{0\}} \frac{\|A\varphi\|_\infty}{\|\varphi\|_\infty}$$

using the maximum norm $\|\cdot\|_\infty$. We define that a difference scheme is unstable if the difference scheme is not stable.

Application of the Backward Difference. We show that the difference scheme (6) using the backward difference is unstable. We denote the imaginary unit by i. For all $\xi,\ \eta \in \boldsymbol{R}$ there exists $g \in \boldsymbol{C}$ such that

$$\tilde{u}(ph, qh, k\Delta t) := g^k e^{i\xi ph} e^{i\eta qh} \tag{10}$$

satisfies the difference scheme (6) for any $p,\ q,\ k \in \boldsymbol{N}$. The value g is called amplification factor. The amplification factor g means

magnification ratio of $\tilde{u}$ per a time step. In fact, when $\tilde{u}$ is substituted in (6), we have

$$\begin{aligned} & g^{k+1}e^{i\xi ph}e^{i\eta qh} \\ = \; & -\rho g^k e^{i\xi ph}e^{i\eta(q-1)h} - \rho g^k e^{i\xi(p-1)h}e^{i\eta qh} \\ & +(1+4\rho)g^k e^{i\xi ph}e^{i\eta qh} \\ & -\rho g^k e^{i\xi(p+1)h}e^{i\eta qh} - \rho g^k e^{i\xi ph}e^{i\eta(q+1)h}. \end{aligned}$$

Consequently,

$$\begin{aligned} g &= -\rho e^{-i\eta h} - \rho e^{-i\xi h} + 1 + 4\rho \\ & \quad -\rho e^{i\xi h} - \rho e^{i\eta h} \\ &= -2\rho\cos(\xi h) - 2\rho\cos(\eta h) + 1 + 4\rho. \quad (11) \end{aligned}$$

Therefore by setting g for all $\xi,\ \eta \in \boldsymbol{R}$ as (11), $\tilde{u}$ satisfies the difference scheme (6). From (11) we get

$$1 \le g \le 1 + 8\rho, \quad \xi, \eta \in \boldsymbol{R}. \qquad (12)$$

Stability of a difference scheme can translate as follows. There exists a constant $C > 0$ such that $|g|^k < C$ for all k and Δt which satisfy $\bar{T} \le -k\Delta t \le 0$. But from (11), for all $L \in (1, 1+8\rho)$ there exists $\xi_L,\ \eta_L \in \boldsymbol{R}$ such that an amplification factor g_L for $\xi_L,\ \eta_L$ satisfies $g_L \ge L$. Then we can take L as constant, because ρ is constant. By setting $k_{\bar{T}} := [-\frac{\bar{T}}{\Delta t}]$, we have $\bar{T} \le -k_{\bar{T}}\Delta t \le 0$ and $k_{\bar{T}} \to \infty$ as $\Delta t \to 0$. Therefore $|g_L|^{k_{\bar{T}}} \to \infty$ as $\Delta t \to 0$, and we get that the difference scheme (6) is unstable with constant ρ.

Application of the Forward Difference. We also have that the difference scheme using the forward difference

$$\begin{aligned} \frac{u_{ij}^k - u_{ij}^{k+1}}{\Delta t} &= \frac{u_{i+1j}^{k+1} - 2u_{ij}^{k+1} + u_{i-1j}^{k+1}}{h^2} \\ & \quad + \frac{u_{ij+1}^{k+1} - 2u_{ij}^{k+1} + u_{ij-1}^{k+1}}{h^2} \qquad (13) \end{aligned}$$

where $i, j = 1, 2, \cdots, n-1;\ k = 0, 1, 2, \cdots$ is unstable. Analogously when using the backward difference by substituting $\tilde{u}(x_i, y_j, t_k)$ for u_{ij}^k in (13), we get instability of (13).

Convergence of Solutions of the Difference Scheme

In the previous subsection, we showed that difference schemes (6) and (13) are unstable. When an initial–boundary value problem of differential equation is well-posed, we suppose using unstable scheme for the problem. Then we can conclude that a solution of the difference scheme does not converge to a solution of the problem by the following Lax's equivalence theorem.

Theorem 2 (Lax's equivalence theorem[6]) *If a difference scheme is compatible with a initial value problem, then stability of the difference scheme is equivalent to convergence of the difference scheme.*

ARBITRARY PRECISION ARITHMETIC

When the conventional numerical calculation using the unstable difference scheme (6) is carried, solution of the difference scheme diverges as k increases. The cause of divergence is due to severe increase in numerical errors. A numerical error consists of a round-off error and a discretization error. In this section, we consider to reduce accumulation of round-off error by using arbitrary precision arithmetic.

Implementation of Arbitrary Precision Arithmetic

We represent an arbitrary precision number x by radix 10000 fixed point. For each component of an integer type array which has p components, we set a non-negative integer a_i which has 4 figures by radix 10, where $i = 0, 1, \cdots, p-1$. For an argument sgn which represents signum of x, we set $sgn = 1$ if x is non-negative, $sgn = -1$ otherwise. Let f be the number of components in the array which includes the decimal part of x. Then the number of figures in the decimal part of x is $frac = 4f$. Then x is represented by

$$x = sgn \sum_{i=0}^{p-1} a_i \times 10000^{i-f}$$

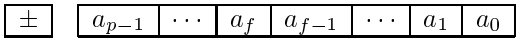
with

$$p \ge 1, \quad 0 \le a_i \le 9999,\ i = 0, 1, \cdots, p-1,$$

$$sgn = \pm 1, \quad 0 \le f \le p-1.$$

The number x is represented in the following schema.

Signum Integer type array



We use the C programming language. For elementary four arithmetic operations we use algorithms similar to hand calculation.

NUMERICAL EXAMPLES

In the square domain $D := (0,1) \times (0,1)$, a function $\tilde{u}$ defined by

$$\tilde{u}(x,y,t) := \frac{1}{3}\{6(x-0.5)t + (x-0.5)^3 + 6(y-0.5)t + (y-0.5)^3\}$$

satisfies the heat equation (1). Then in the backward heat conduction problem (1)–(3), we set data w and f as

$$w = \tilde{u}(\cdot, 0) \quad \text{in} \quad D,$$
$$f = \tilde{u} \quad \text{on} \quad \Gamma \times [\bar{T}, 0].$$

We solve the problem numerically by the difference scheme (6) using the arbitrary precision arithmetic. We observe relation between the numerical solution u_{ij}^k and the exact solution $u(x_i, y_j, t_k)$. A error is measured by

$$err_k := \max_{i,j=0,1,\cdots,n} |u_{ij}^k - u(x_i, y_j, t_k)|. \quad (14)$$

As a remark, $\tilde{u}$ satisfies the difference scheme (6) exactly. Therefore discretization error is not involved, and numerical error consists only of round-off error.

The Influence of the Number of Figures

We set the number of space division $n = 20$, the space step size $h = 0.05$, the time step size $\Delta t = 0.25 \times 10^{-4}$, and the coefficient $\rho = 0.01$. We obtain a numerical solution at k just before a step when err_k attains to 10^{-4}.

In Case the Number of Figures is 28. We experiment numerically with setting the number of figures in decimal part as $frac = 28$. Then we obtained the numerical solution which contains error $\mathrm{err}_k = 0.99681 \times 10^{-4}$ at $k = 773$, $t = -0.019325$. We show the exact solution and obtained numerical solution in Figure 1.

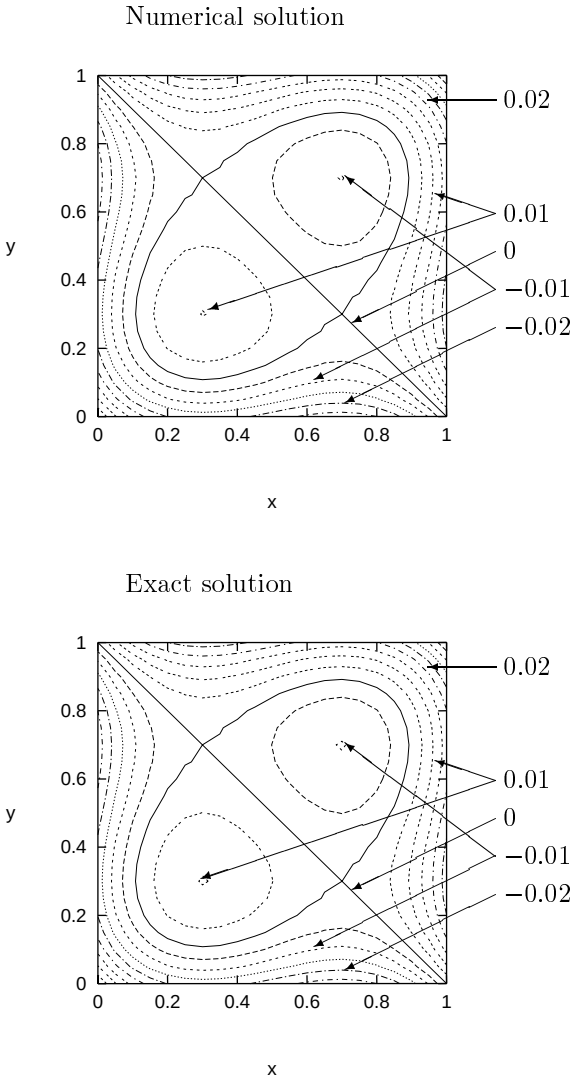


Figure 1: Solutions at $t = -0.019325$ using 28 figures in decimal part

In Case the Number of Figures is 88. We experiment numerically with setting the number of figures in decimal part as $frac = 88$. Then we obtained the numerical solution which contains error $\mathrm{err}_k = 0.97985 \times 10^{-4}$ at $k = 2594$, $t = -0.06485$. We show the exact solution and obtained numerical solution in Figure 2.

If we had not used the arbitrary precision arithmetic, a numerical solution would have already overflown before the final time step k. We see that the numerical solution is close to the exact solution in Figure 1 and 2.

The Relation between Time Steps and Errors for the Number of Figures

We derive the relation between $frac$, k, and err_k by plotting these data.

In Case the Number of Division of the Domain is 10^2. When the number of division in each axis direction is 10, we plotted $frac$, k, and err_k. In Figure 3, graphs are shown for k and err_k as $frac = 28$, 60, 88, and 120. Then we obtain the relation

$$\log_{10} err_k \approx 0.0314k - frac \qquad (15)$$

from these graphs by fitting a regression line. In fact, lines which satisfy (15) fit in well with data of experiments.

In Case the Number of Division of the Domain is 20^2. When the number of division in each axis direction is 20, we plotted $frac$, k, and err_k. In Figure 4, graphs are shown for k and err_k as $frac = 28$, 60, and 88. Then we obtain the relation

$$\log_{10} err_k \approx 0.0318k - frac \qquad (16)$$

from these graphs by fitting a regression line. In fact, lines which satisfy (16) fit in well with data of experiments. In relations (15), (16), we see that a change in the coefficient of k is small, although n changes from 10 to 20.

The Relation between Time Steps and Errors for the Coefficient ρ

As $frac = 60$, we plotted err_k against k for $\rho = 0.005$, 0.01, 0.02, and 0.04 in Figure 5. For each ρ, we obtain a linear relation between $\log_{10} err_k$ and k. Fitting a regression line for

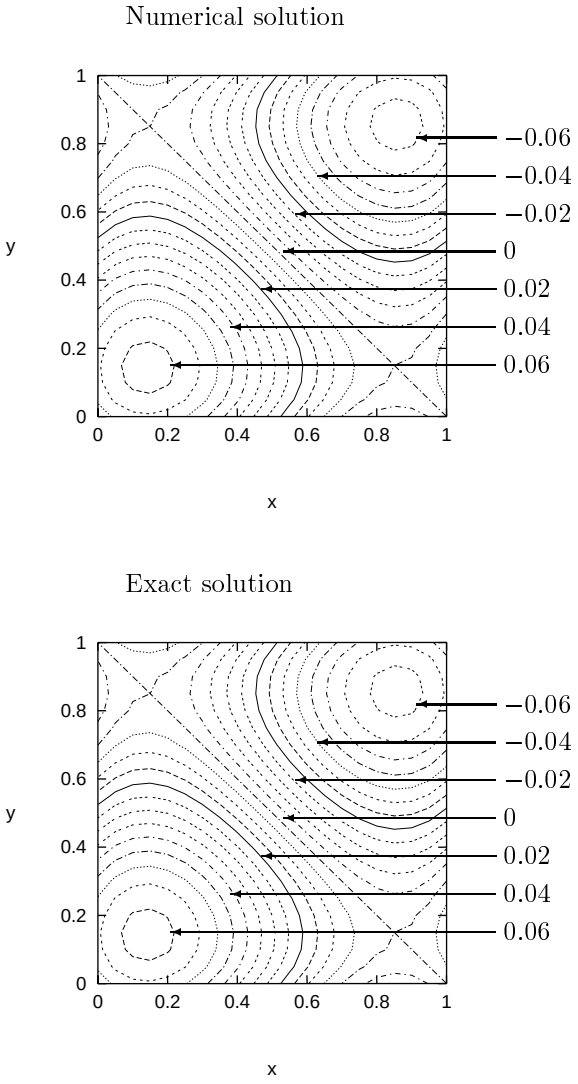


Figure 2: Solutions at $t = -0.06485$ using 88 figures in decimal part

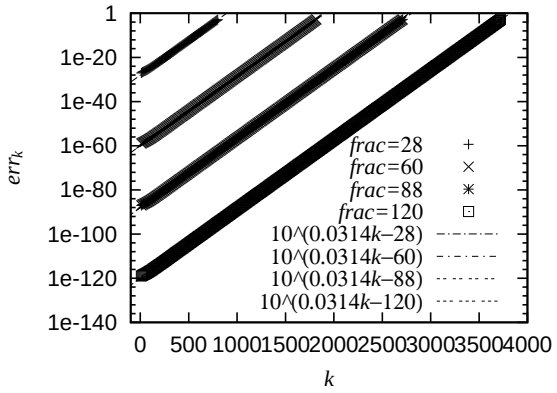


Figure 3: The relation between time steps and errors for the number of figures ($n = 10$)

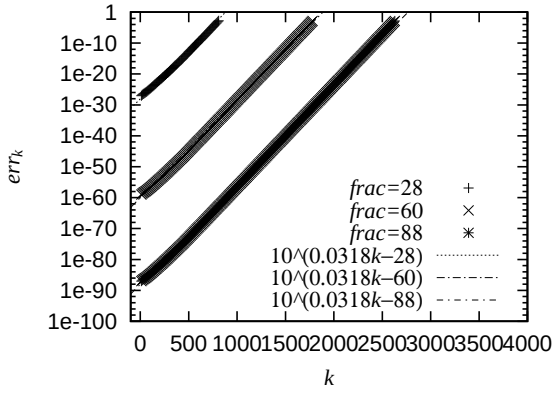


Figure 4: The Relation Between Time Steps and Errors for the Number of Figures ($n = 20$)

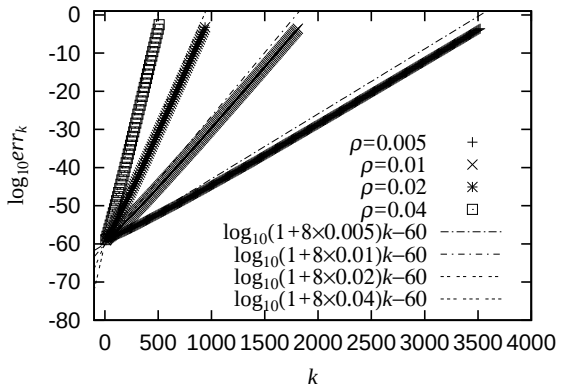


Figure 5: The relation between time steps and errors for ρ

$\log_{10} err_k$, we see that the gradient of this line decreases as ρ tends to 0.

From the relation (12), we can estimate that err_k will become $1+8\rho$ times par time step. By considering (15) and (16), we expect a relation

$$\log_{10} err_k \approx k \log_{10}(1 + 8\rho) - frac \qquad (17)$$

between ρ, k, and err_k. In fact in Figure 5, behavior of data in experiments follows the relation (17). Moreover we observe that data of experiments do not exceed a line which satisfies the relation (17) when k is sufficiently large. As an explanation of this result we guess that right hand side of the relation (17) represents logarithm of accumulation error which has an overestimated amplification factor from the inequality (12).

CONCLUSIONS

1. We solve the backward heat conduction problem by finite difference method using arbitrary precision arithmetic.

2. The amplification factor g satisfies

$$1 \leq g \leq 1 + 8\rho.$$

The adopted difference scheme is unstable because a magnification ratio of $\tilde{u}$ per time step is larger than 1.

3. When discretization error is zero, we showed that $frac$ and ρ satisfy

$$err_k \approx (1 + 8\rho)^k \times 10^{-frac},$$

where $frac$ is the number of figures in decimal part, and ρ is the coefficient in difference scheme. Therefore, we can estimate how many figures are necessary when using arbitrary precision arithmetic in order to obtain numerical solution whose error is smaller than given limit of error.

4. The next problem we must consider will be the efficiency of arbitrary precision arithmetic in problems which include discretization error.

REFERENCES

[1] Hitoshi Imai, Toshiki Takeuchi, and Masahiro Kushida, On numerical simulation of partial differential equations in infinite precision, *Advances in Mathematical Sciences and Applications*, **9** (1999), No. 2, pp. 1007–1016.

[2] Fritz John, Numerical solution of the equation of heat conduction for preceding times, *Annali di Matematica Pura ed Applicata. Series IV*, **40** (1955), pp. 129–142.

[3] Rainer Kress, Linear Integral Equations, Second Edition, Springer-Verlag, New York (1998).

[4] Jijun Liu, Determination of temperature field for backward heat transfer, *Communications of the Korean Mathematical Society*, **16** (2001), No. 3, pp. 385–397.

[5] L. E. Payne, Improperly posed problems in partial differential equations, *Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, (1975).

[6] Robert D. Richtmyer and K. W. Morton, Difference Methods for Initial-value Problems, Second Edition, Interscience, Publishers, New York · London · Sydney (1967).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ESTIMATION OF DRYING FOOD THERMOPHYSICAL PROPERTIES BY USING TEMPERATURE MEASUREMENTS

Ljubica P. Kanevce
Faculty of Technical Sciences
St. Kliment Ohridski University
Bitola, Macedonia
kanevce@osi.net.mk

Gligor H. Kanevce
Macedonian Academy of
Sciences and Arts
Skopje, Macedonia
kanevce@osi.net.mk

George S. Dulikravich
Department of Mechanical and Aerospace Engineering
Multidisciplinary Analysis, Inverse Design and Optimization (MAIDO) Program
The University of Texas at Arlington, Arlington, Texas 76019, U.S.A
gsd@mae.uta.edu

ABSTRACT

In this paper an inverse problem of simultaneous estimation of the thermophysical properties, together with the heat and mass transfer coefficients of a drying food sample is analyzed. The solution method is based on using only temperature measurements.

The Levenberg-Marquardt procedure of minimization of the least-squares norm is applied for the solution of the presented parameter estimation problem. As a representative vegetable material, a potato has been chosen. Numerical experiments have been conducted to investigate applicability of the method to the foods. In order to simulate real measurements, a normally distributed error was added to the numerical temperature response. An analysis of the influence of the drying parameters needed for the design of the proper experiment is presented as well. In order to perform this analysis, the sensitivity coefficients and the sensitivity matrix determinant were examined.

NOMENCLATURE

a = water activity
c = heat capacity (dry basis), J/K/kg db
C = concentration of water vapor, kg/m^3
D = moisture diffusivity, m^2/s
D_0 = Arhenius factor for m. diffusion, m^2/s
E_0 = activation energy for m. diff., kJ/mol
h = heat transfer coefficient, W/m^2/K
h_D = mass transfer coefficient, m/s
ΔH = latent heat of vaporization, J/kg
$\mathbf{I}$ = identity matrix
j_m = mass flux, kg/m^2/s
j_q = heat flux, W/m^2
$\mathbf{J}$ = sensitivity matrix
κ = thermal conductivity, W/m/K
L = flat plate thickness, m
p_s = saturation pressure , Pa
$\mathbf{P}$ = vector of unknown parameters
R = ideal gas constant, kJ/mol/K
t = time, s
T = temperature, ^{0}C
T_k = absolute temperature, K
$\mathbf{T}$ = vector of estimated temperatures, ^{0}C
v = velocity, m/s
x = distance from the mid-plane, m
X = moisture content (dry basis), kg/kg db
$\mathbf{Y}$ = vector of measured temperatures, ^{0}C
δ = thermo-gradient coefficient, 1/K
ε = phase conversion factor
ε = relative error
σ = standard deviation
μ = damping parameter
ρ = density, kg/m^3
φ = relative humidity

Subscripts

a = drying air
s = dry solid
w = water

INTRODUCTION

There are several methods for describing the direct problem of complex simultaneous heat and moisture transport processes within drying

material. In the approach proposed by Luikov [1] the moisture and temperature fields in the drying body are expressed by a system of two coupled partial differential equations. The system of equations incorporates coefficients that must be determined experimentally.

For many drying processes, including the drying processes considered in this paper, the influence of the thermal diffusion is small and can be ignored. In this case, Luikov's moisture transport equation is the same as Fick's second law equation, where concentration has been converted to moisture content on a dry basis. The moisture diffusivity has the same meaning in both of these approaches. It accounts for various types of possible drying processes including molecular (liquid) diffusion, vapor diffusion, surface diffusion, hydrodynamic flow, Knudsen flow, and other considerations. An effective moisture diffusivity, which lumps all possible moisture transport mechanisms into a single measurable parameter, is often used to characterize the drying behavior regardless of the dominating mechanism. This moisture diffusivity depends on moisture content and temperature. This effect cannot be ignored in the drying process calculation for the majority of practical cases.

All the coefficients except for the moisture diffusivity can be relatively easily determined by experiments [2, 3]. A number of methods for the experimental determination of the moisture diffusivity exist such as: sorption kinetics methods, permeation methods, concentration-distance methods, drying methods, radiotracer methods, and methods based on the techniques of electron spin resonance and nuclear magnetic resonance. But, there is no standard method for the experimental determination of the moisture diffusivity. The adoption of a generalized method for moisture diffusivity estimation would be of great importance; however, this does not seem probable in the near future [4].

The application of the moisture diffusivity estimation methods based on the experimental drying curves in relation to the analytical solution of the differential diffusion equation seems to be the most popular experimental practice [5, 6]. Numerical solutions of the Fick's law differential diffusion equation with constant [7] or moisture and temperature dependent [5] diffusivity also have been used for the moisture diffusivity estimation.

Kanevce, Kanevce and Dulikravich [8, 9, 10, 11] and Dantas, Orlande and Cotta [12, 13, 14] recently analyzed a method of the moisture diffusivity estimation by temperature response of a drying body. The main idea of this method is to make use of the interrelation between the heat and mass (moisture) transport processes within the drying body and from its surface to the surroundings. Then, the moisture diffusivity can be estimated on the basis of an accurate and easy to perform single thermocouple temperature measurement by using an inverse approach.

The objective of this paper is an analysis of the possibilities of simultaneous estimation of the moisture diffusivity together with other thermophysical properties of the vegetables as well as the heat and mass transfer coefficients. The method requires a single drying experiment and requires a single temperature measurement probe. As a representative drying vegetable product, a thin slice of a potato has been chosen.

The present parameter estimation problem is solved by using the Levenberg-Marquardt method of minimization of the least-squares norm. Instead of actual temperature measurements, the temperature response during convective drying is obtained from the numerical solution of the non-linear one-dimensional Luikov's equations. In order to simulate real measurements, a normally distributed error was added to the numerical temperature response.

An analysis of the influence of the drying air velocity, temperature and relative humidity, drying body dimensions and drying time on the moisture diffusivity estimation that enables the design of the proper experiment is conducted as well. In order to realize this analysis, the sensitivity coefficients and the sensitivity matrix determinant were calculated for the characteristic drying regimes and drying body dimensions.

A MATHEMATICAL MODEL OF DRYING

In the case of an infinite flat plate of thickness $2L$, if the shrinkage of the material during drying can be neglected, the resulting system of equations for energy and moisture transport can be expressed as

$$c\rho_s \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x} \right) + \varepsilon \rho_s \Delta H \frac{\partial X}{\partial t} \qquad (1)$$

$$\frac{\partial X}{\partial t} = \frac{\partial}{\partial x}\left(D \frac{\partial X}{\partial x} + D\delta \frac{\partial T}{\partial x} \right) \qquad (2)$$

where $T(x, t)$ is the unsteady temperature field and $X(x, t)$ is the moisture content field.

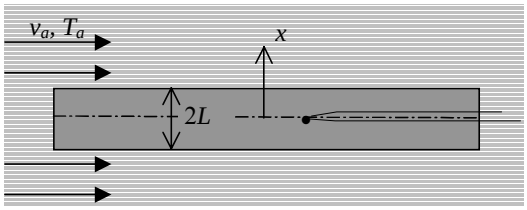


Fig. 1. Scheme of the drying experiment

As initial conditions, uniform temperature and moisture content profiles are assumed

$$t = 0 \quad T(x,0) = T_0\,, \quad X(x,0) = X_0 \tag{3}$$

In the convective drying experiment (Fig. 1) the surfaces of the drying body are in contact with the drying air thus resulting in a convective boundary conditions for both the temperature and the moisture content

$$\begin{aligned} -k\left(\frac{\partial T}{\partial x}\right)_{x=L} + j_q - \Delta H(1-\varepsilon) j_m = 0 \\ D\rho_s\left(\frac{\partial X}{\partial x}\right)_{x=L} + D\delta\rho_s\left(\frac{\partial T}{\partial x}\right)_{x=L} + j_m = 0 \end{aligned} \tag{4}$$

where the convective heat flux, $j_q(t)$, and mass flux, $j_m(t)$, on these surfaces are

$$\begin{aligned} j_q &= h\left(T_a - T_{x=L}\right) \\ j_m &= h_D\left(C_{x=L} - C_a\right) \end{aligned} \tag{5}$$

The water vapor concentration in the drying air, C_a, is calculated by

$$C_a = \varphi\, p_s(T_a) / 461.9 /(T_a + 273) \tag{6}$$

The water vapor concentration of the air in equilibrium with the surface of the body exposed to convection is calculated by

$$C_{x=L} = \frac{a(T_{x=L}, X_{x=L})\; p_s(T_{x=L})}{461.9\,(T_{x=L} + 273)} \tag{7}$$

The water activity, a, or the equilibrium relative humidity of the air in contact with the convection surface at temperature $T_{x=L}$ and moisture content $X_{x=L}$ is calculated from experimental water sorption isotherms.

The problem is symmetrical, and boundary conditions on the mid-plane of the plate ($x = 0$) are

$$\left(\frac{\partial T}{\partial x}\right)_{x=0} = 0, \quad \left(\frac{\partial X}{\partial x}\right)_{x=0} = 0 \tag{8}$$

In order to approximate the solution of Eqs. (1, 2), an explicit procedure has been used [15].

ESTIMATION OF PARAMETERS

The estimation methodology used is based on minimization of the ordinary least square norm

$$\mathbf{E}(\mathbf{P}) = [\mathbf{Y} - \mathbf{T}(\mathbf{P})]^{\mathrm{T}}[\mathbf{Y} - \mathbf{T}(\mathbf{P})] \tag{9}$$

Here, $\mathbf{Y}^{\mathrm{T}} = [Y_1, Y_2, \ldots, Y_{\mathrm{imax}}]$ is the vector of measured temperatures, $\mathbf{T}^{\mathrm{T}} = [T_1(\mathbf{P}), T_2(\mathbf{P}), \ldots T_{\mathrm{imax}}(\mathbf{P})]$ is the vector of estimated temperatures at time t_i (i = 1, 2, …, imax), $\mathbf{P}^{\mathrm{T}} = [P_1, P_2, \ldots P_{\mathrm{N}}]$ is the vector of unknown parameters, imax is the total number of measurements, and N is the total number of unknown parameters (imax ≥ N).

A version of Levenberg-Marquardt method was applied for the solution of the presented parameter estimation problem [16]. This method is quite stable, powerful, and straightforward and has been applied to a variety of inverse problems. It belongs to a general class of damped least square methods [17]. The solution for vector $\mathbf{P}$ is achieved using the following iterative procedure

$$\mathbf{P}^{r+1} = \mathbf{P}^r + [(\mathbf{J}^r)^T\mathbf{J}^r + \mu^r\mathbf{I}]^{-1}(\mathbf{J}^r)^T[\mathbf{Y} - \mathbf{T}(\mathbf{P}^r)] \tag{10}$$

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial T_1}{\partial P_1} & \cdots & \dfrac{\partial T_1}{\partial P_N} \\ . & & \\ \dfrac{\partial T_{i\max}}{\partial P_1} & \cdots & \dfrac{\partial T_{i\max}}{\partial P_N} \end{bmatrix} \tag{11}$$

Near the initial guess, the problem is generally ill-conditioned so that large damping parameter is chosen thus making term $\mu\mathbf{I}$ large as compared to term $\mathbf{J}^{\mathrm{T}}\mathbf{J}$. The term $\mu\mathbf{I}$ damps instabilities due to ill-conditioned character of the problem. So, the matrix $\mathbf{J}^{\mathrm{T}}\mathbf{J}$ is not required to be non-singular at the beginning of iterations and the procedure tends towards a slow-convergent steepest descent

method. As the iteration process approaches the converged solution, the damping parameter decreases, and the Levenberg-Marquardt method tends towards Gauss method. In fact, this method compromises between the steepest descent and Gauss method by choosing μ so as to follow the Gauss method to as large an extent as possible, while retaining a bias towards the steepest descent direction to prevent instabilities. The presented iterative procedure stops if the norm of gradient of **E**(**P**) is sufficiently small, or if the ratio of the norm of gradient of **E**(**P**) to the **E**(**P**) is small enough, or if the changes in the vector of parameters are very small [18].

RESULTS AND DISCUSSION

In this paper, application of the proposed method for the estimation of the thermophysical properties of vegetables has been analyzed. As a representative vegetable product, a single slice of a potato was chosen.

The variation in water activity with change in moisture content of samples at a specified temperature is defined by sorption isotherms. Sorption isotherms of most vegetables are nonlinear and generally sigmoid in shape. There are many different models for describing the sorption isotherms of foods [3, p. 25]. In recent years, the most widely accepted and efficient model for sorption isotherms of foods has been the GAB (Guggenheim-Anderson-de Boer) model. It is a semi-theoretical model and has been considered the best-fit model for many food materials over a wide range of water activity. The GAB isotherm equation can be written as

$$X = \frac{X_m CKa}{(1-Ka)(1-Ka+CKa)} \tag{12}$$

where the water activity, a, represents the relative humidity of the air in equilibrium with the drying object at temperature, T, and moisture content, X. The monolayer moisture, X_m, and the adsorption constants C and K are related as Arrhenius type equations

$$C = C_0 \exp\left(\frac{\Delta H_C}{R_w T_k}\right) \tag{13}$$

$$K = K_0 \exp\left(\frac{\Delta H_K}{R_w T_k}\right) \tag{14}$$

$$X_m = X_{m0} \exp\left(\frac{\Delta H_x}{R_w T_k}\right) \tag{15}$$

GAB model parameters, C_0, ΔH_C, K_0, ΔH_K, X_{m0}, and ΔH_K can be estimated by different regression procedures from experimental isotherm data. The Gane experimental results for potatoes (cited by [3], p. 45) were used in this paper

$C_0 = 6.609 \cdot 10^{-1}$; $\Delta H_c = 528.4$ kJ/kg
$K_0 = 0.606$; $\Delta H_k = 53.33$ kJ/kg
$X_{m0} = 2.489 \cdot 10^{-2}$; $\Delta H_x = 123.6$ kJ/kg

Heat capacity of food materials can be taken as equal to the sum of the heat capacity of solid matter and water absorbed by that solid

$$c = c_s + c_w . X \tag{16}$$

Although the heat capacity of solid matter and water are functions of the temperature, constant values have been most widely used. The following values, proposed in [19] for potato, were used.

$c_s = 1381$ J/kg/K, and
$c_w = 4187$ J/kg/K.

The shrinkage of the material during drying has been neglected in this paper. This leads to a constant dry material density. The value for dried potato [20] used in this paper was

$\rho_s = 1610$ kg/m^3.

From the experimental and numerical examinations of the transient moisture and temperature profiles [15] it was concluded that for practical calculations, the influence of the thermodiffusion, δ, is small and can be ignored. Consequently,

$\delta = 0$

was utilized in this paper.

It was also concluded that the system of two simultaneous partial differential equations (1, 2) could be used by treating the thermal conductivity

as constant. A mean value from the results obtained in [21] for the potato was taken as

$k = 0.40$ W/m/K.

There is very little published data concerning the value of the ratio of water evaporation rate to the reduction rate of the moisture content during the drying of food. For this reason and because the influence of the phase conversion factor on the transient moisture content and temperature profiles is very small, a mean value was used as

$\varepsilon = 0.5$

Moisture diffusivity of foods is often considered as an Arrhenius-type temperature function

$$D = D_0 \exp(-E_0/RT_k) \qquad (17)$$

where, D_0 is the Arrhenius factor, E_0 is the activation energy for moisture diffusion, R is the ideal gas constant, and T_k is the absolute temperature. In this paper the values obtained in [5] for potato from the drying data were used

$D_0 = 8.63 \cdot 10^{-5}$ m^2/s ; $E_0 = 30.6$ kJ/mol.

For the direct problem solution, the system of Equations (1) and (2) with the initial conditions, equation (3), and the boundary conditions, equations (4) and (8), have been solved numerically for a potato slice sample, with the above thermophysical properties.

For the inverse problem investigated here, values of the moisture diffusivity, the heat capacity, the thermal conductivity, the phase conversion factor and the density of the potato as well as the heat and mass transfer coefficients were regarded as unknown. All other quantities appearing in the direct problem formulation were assumed to be known. Thus, the vector of the unknown parameters was

$$\mathbf{P}^{\mathrm{T}} = [D_0, E_0, \rho_s, c_s, k, \varepsilon, h, h_D] \qquad (18)$$

For the estimation of these unknown parameters, the transient reading of a single temperature sensor located at the position $x = 0$, has been considered.

Numerical experiments have been conducted to investigate applicability of the method to the food processing involving drying of thin flat samples. In order to simulate real measurements, a normally distributed error with zero mean and standard deviation, σ, was added to the numerical temperature responses. An analysis of the influence of the drying air parameters, drying time, and dimensions of the drying sample needed for the design of the proper experiment has been conducted as well. In order to perform this analysis, the sensitivity coefficients and the sensitivity matrix determinant have been examined.

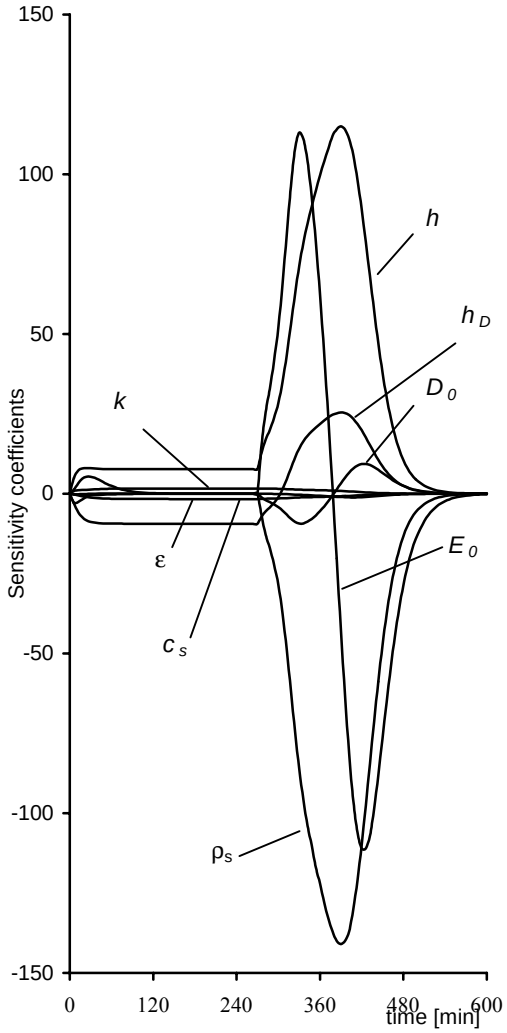


Fig. 2. Relative sensitivity coefficients

The sensitivity coefficients analysis has been carried out for an infinite flat plate model of a

slice of a potato with initial moisture content of $X(x, 0) = 4.00$ kg/kg and initial temperature $T(x, 0) = 20.0\ ^{O}C$. The possibilities of simultaneous estimation of the temperature-dependent moisture diffusivity together with other thermophysical properties of the potato as well as the heat and mass transfer coefficients have been investigated for the variety of boundary conditions and dimensions of the drying sample.

The drying air bulk temperature, T_a, was varied between 30 and 70 ^{O}C, the drying air velocity, v_a, between 1.0 and 4.0 m/s, the relative humidity of the drying air, φ_a, between 0.02 and 0.08 and the potato slice thickness $2L$, between 2 and 8 mm.

The best combination of the relative temperature sensitivity coefficients with respect to all unknown parameters, was obtained with $T_a = 70\ ^{O}C$, $v_a = 3$ m/s, $\varphi_a = 0.05$ and $2L = 4$ mm. Figure 2 shows the relative sensitivity coefficients $P_j \partial T_i / \partial P_j$, i = 1, 2,…, imax, for temperature with respect to all unknown parameters, j = 1, 2,…,8.

It can be seen that the relative sensitivity coefficients with respect to the phase conversion factor, ε, and the thermal conductivity, k, are very small. This indicates that ε and k cannot be estimated in this case. But, this also indicates that the influence of the phase conversion factor and the thermal conductivity on the transient moisture content and temperature profiles is very small in this case. This can be explained by the very small heat transfer Biot number ($Bi = hL/k = 0.156$) and consequently very small temperature gradients inside the body during the drying (Fig. 4). Due to these reasons, the phase conversion factor and the thermal conductivity were taken as known quantities for the examination below.

Heat capacity of wet potato was taken as equal of the sum of the heat capacity of solid matter and absorbed water. Since the heat capacity of the solid matter presents only a few percent of the overall heat capacity (Eq. 16) of the potato, the relative sensitivity coefficients with respect to the heat capacity of solid matter is also very small. Consequently, the value of the heat capacity of the solid matter was also taken as known.

The relative sensitivity coefficients with respect to the dry matter density, ρ_s, and the convection heat transfer coefficient, h, are linearly-dependent except for the first period of drying, when the relative sensitivity coefficients with respect to the dry matter density is nearly equal to zero. This makes it possible to simultaneously estimate ρ_s and h with the "exact" (without noise) temperature data. But, with temperature data with an added simulated noise, local minimums were obtained that depended on the initial guesses. Due to these reasons and the fact that the density of the dry sample material can be relatively easily determined by a separate experiment, the density of the dry sample material was also considered as known.

Thus, it appears to be possible to estimate simultaneously the moisture diffusivity parameters, D_0 and E_0, the convection heat transfer coefficient, h, and the mass transfer coefficient, h_D.

Figure 3 presents transient variation of the sensitivity determinant if five, (D_0, E_0, ρ_s, h, h_D) and four (D_0, E_0, h, h_D) parameters are simultaneously considered as unknown. Elements of this sensitivity determinant were defined [22] for a large, but fixed number of transient temperature measurements (501 in these cases).

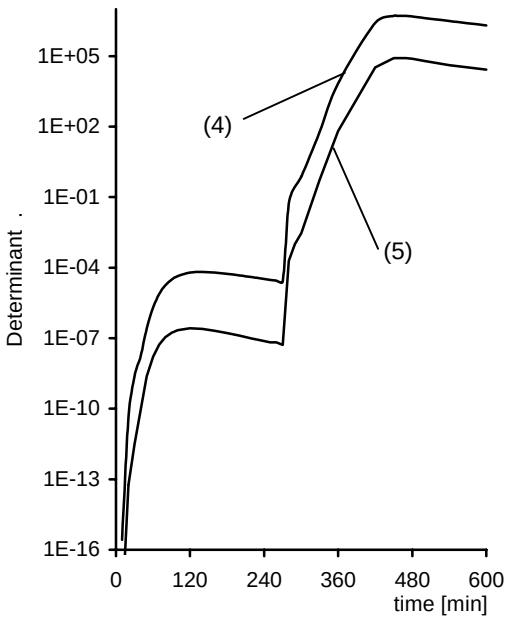
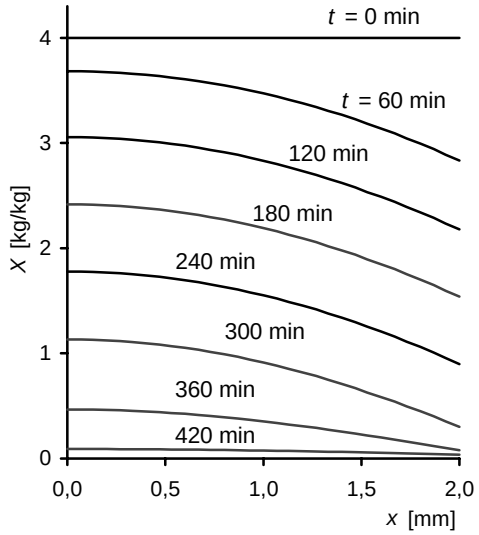


Fig. 3. Sensitivity determinant

The maximum sensitivity determinant value corresponds to the drying time when nearly equilibrium moisture content and temperature profiles have been reached (Fig. 4 and 5).

Table 1. Estimated parameters

Parameters	Exact values	Estimated values			Relative errors for σ = 0.5 [%]
		σ = 0	σ = 0.3 ^{O}C	σ = 0.5 ^{O}C	
$D_0 \cdot 10^5$ [m^2/s]	8.63	8.6306	8.6648	8.7142	0.98
E_0 [kJ/mol]	30.6	30.600	30.611	30.626	0.09
h [W/m^2/K]	31.2	31.200	31.200	31.200	0.00
$h_D \cdot 10^2$ [m/s]	3.42	3.4200	3.4203	3.4205	0.01



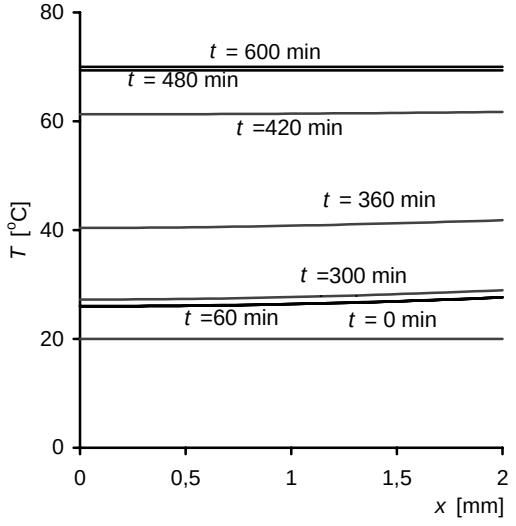


Fig. 4. Transient moisture content and temperature profiles

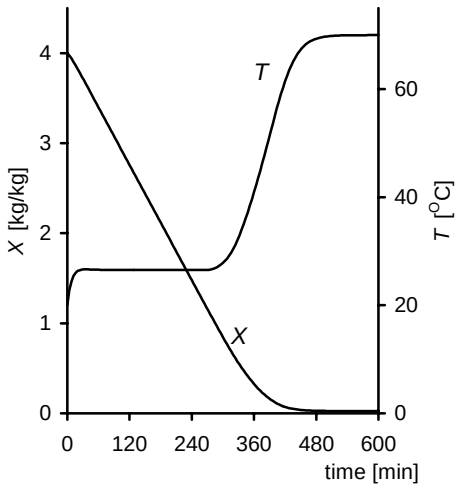


Fig. 5. Volume-averaged moisture content and temperature changes during the drying

Table 1 shows the computationally obtained parameters. For comparison, the values of exact parameters and the values estimated with "exact" (without noise) temperature data are also shown. The obtained results show good agreement between the evaluated and exact values of parameters. The moisture diffusivity and the boundary conditions parameters, (D_0, E_0, h, h_D), can be simultaneously estimated with the relative errors within one percent even in the case of temperature measurement errors with σ = 0.5 ^{O}C.

CONCLUSIONS

The inverse problem of simultaneous estimation of thermophysical properties and the boundary conditions parameters of drying vegetables by using only temperature measurements has been analyzed. As a representative vegetable product, a slice of a potato has been chosen.

It can be concluded that in the convective drying experiment it is possible, based on a single thermocouple temperature response, to estimate simultaneously the two moisture diffusivity parameters, the convection heat transfer coefficient, and the mass transfer coefficient.

REFERENCES

1. A. V. Luikov, *Teplomassoobmen,* Energia, Moscow, Russia, 1972, p. 468.

2. V. T. Karathanos, Z. B. Maroulis, D. Marinos-Kouris, and D. G. Saravacos, Higrothermal and quality properties applicable to drying. Data sources and measurement techniques, *Drying Technology*, **14**, 1403 (1996)

3. S. Rahman, *Food Properties Handbook*, CRC Press, Inc., Boca Raton, New York, 1995

4. N. P. Zogzas, Z. B. Maroulis, and D. Marinos-Kouris, Moisture diffusivity data compilation in food products, *Drying Technology,* **14**, 2225 (1996)

5. N. P. Zogzas, and Z. B. Maroulis, Effective moisture diffusivity estimation from drying data. A comparison between various methods of analysis, *Drying Technology,* **14**, 1543 (1996)

6. H. Feng, J. Tang, R. P. Cavalieri, Combined microwave and spouted bed drying of diced apples: effect of drying conditions on drying kinetics and product temperature, *Drying Technology,* **17**, 1981 (1999)

7. R. W. Daud, M. H. Ibrahim, and Z. M. Talib, Parameter estimation of Fick's law drying equation, *Drying Technology,* **15**, 1637 (1997)

8. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Moisture diffusivity estimation by temperature response of a drying body, *Inverse Problems in Engineering Mechanics II*, *eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam,* 43, (2000)

9. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Influence of boundary conditions on moisture diffusivity estimation by temperature response of a drying body, *Proceedings of NHTC'00, ASME paper NHTC2000-12296, 34th ASME National Heat Transfer Conf., August 20-22, Pittsburgh, PA, U.S.A.,* 12296, (2000)

10. G. H. Kanevce, L. P. Kanevce, V. B. Mitrevski and G. S. Dulikravich, Moisture diffusivity estimation from temperature measurements: influence of measurement accuracy, *Proceedings of the 12th International Drying Symposium, IDS'2000, August 28-31, Noordwijkerhout, The Netherlands*, 337, (2000)

11. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Simultaneous estimation of thermophysical properties and heat and mass transfer coefficients of a drying body, *Inverse Problems in Engineering Mechanics III*, *eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam,* 3, (2002)

12. L. B. Dantas, H. R. B. Orlande, R. M. Cotta, R. De Souza, and P. D. C. Lobo, Inverse analysis in moist capillary porous media, *15th Brazilian Congress of Mechanical Engineering*, *November 7-10, Aguas de Lindola, Sao Paulo, Brazil,* (1999)

13. L. B. Dantas, H. R. B. Orlande, R. M. Cotta and P. D. C. Lobo, Parameter estimation in moist capillary porous media by using temperature measurements, *Inverse Problems in Engineering Mechanics II, eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam,* 53, (2000)

14. L. B. Dantas, H. R. B. Orlande and R. M. Cotta, Effects of lateral heat losses on the parameter estimation problem in moist capillary porous media, *Inverse Problems in Engineering Mechanics III, eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam,* 13, (2002)

15. G. H. Kanevce, Numerical study of drying, *IDS '98, Vol. A, Halkidiki, Greece*, 256, (1998)

16. D. W. Marquardt, An algorithm for east squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math.*, 431 (1963).

17. J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*, John Wiley&Sons, Inc., New York, 1977, p. 370.

18. R. C. Pfafl and B. J. Mitchel, A general method for simultaneous measurement of thermal properties, *AIAA Paper, No. 69-602*, (1969)

19. R. Neisteruh, Changes at thermal properties of fruits and vegetables during drying, *Drying Technology,* **14**, 415, (1996)

20. K. K. Krokida, Z. B. Maroulis, Effect of microwave drying on some quality properties of dehydrated products, *Drying technology*, **17**, 449, (1999)

21. G. Donsi, G. Ferrari, R. Nigro, Experimental determination of thermal conductivity of apple and potato of different moisture contents, *Journal of food engineering*, **30**, 263, (1996)

22. M. N. Ozisik and H. R. B. Orlande, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor and Francis, New York, 2000

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

SOLVING GEOMETRIC INVERSE HEAT TRANSFER PROBLEMS ON ANALOG AND HYBRID COMPUTERS

Yuri M. Matsevity
Institute of Mechanical Engineering Problems at National Academy of Science of Ukraine
Pozharsky 2/10, Kharkov 61046, Ukraine
matsevit@ipmach.kharkov.ua

Andrey O. Kostikov
Institute of Mechanical Engineering Problems at National Academy of Science of Ukraine
Pozharsky 2/10, Kharkov 61046, Ukraine
kostikov@ipmach.kharkov.ua

ABSTRACT

In the paper it is considered the algorithms of search of optimum solutions and analogous devices that realize them. In conclusion it is proved the necessity of creation of hybrid (analogous + digital) computing devices to solve such a class of IHCPs.

INTRODUCTION

One of the most important applications of the methodology of solving the inverse heat transfer problem (IHTP) is investigating the thermal regimes of electronic equipment (EE). Such investigations involve the entire gamut of inverse problems. When solving boundary IHTPs, the boundary conditions of convective heat transfer on the surfaces of boards and electronic units; and contact heat transfer between the components of units, and between the board and the add-on components, etc. are identified. When solving the internal IHTPs, the thermal and physical characteristics of multilayer boards with metallised interconnections, and of other electronic units comprising composite structures and/or a combination of heterogeneous components are defined. When solving retrospective IHTPs, one can determine the state of electronic components in the initial or earlier moment of time, whereas the temporal IHTP allows to register the time when the object or its individual components reach a preset (or critical) temperature.

A special position is held by geometric IHTPs, in which the known information about the object's temperature field allows to determine its geometric parameters. Of the entire gamut of IHTPs, such problems are the least studied and infrequently solved ones. At the same time, in many cases their solution is the necessary condition for investigating the thermal state of thermally loaded systems and structures (in particular, electronic equipment), since the geometric characteristics are part of the uniqueness conditions. Without knowing these conditions, it is impossible to define the object's temperature field.

Solving these problems makes it possible to effectively arrange micro channels in the body of the board being cooled to provide internal and convective-conductance cooling [1], or resistive elements in the heater's body when optimising the thermal regimes of EE [2]. Solving the geometric IHTP allows to determine the domain of permissible allocation of given heat sources (EE components), the specified locations of sources of different type with constraints on the temperature and the temperature gradients in the board.

It is also necessary to solve geometric IHTPs for optimising temperature fields in boards when specifying the sites of electronic components, or minimising the packaging surface (volume) of EE, as well as when searching for sources, or during their successive allocation.

The most effective methods of solving IHTPs are those based on solving the conditional optimisation problem, in particular, the method of automated fitting [3]. In view of the fact that the major share of computing time needed for solving IHTPs by such methods on digital computers is spent on simulating a succession of temperature fields, using array analog processors in such cases is far more effective in solving field theory problems. These processors have a discrete structure enabling the use of straightforward stepping search algorithms, which are practical for hardware implementation of the automated fitting method.

IDENTIFYING THE ZONES OF FEASIBLE ALLOCATION OF THERMAL OBJECTS IN THE PRESENCE OF INHIBITION ZONES

When designing EE, a problem frequently encountered is to define the domain of feasible allocation of thermal sources in an object, with a constraint on the maximum temperature. To solve such a problem, the device [4], whose schematic is shown in Fig. 1, can be used.

The device comprises a source assignment unit *SAU* based on flip-flops *FF* and delay elements *DE*; current lead-in unit *CLU*, based on

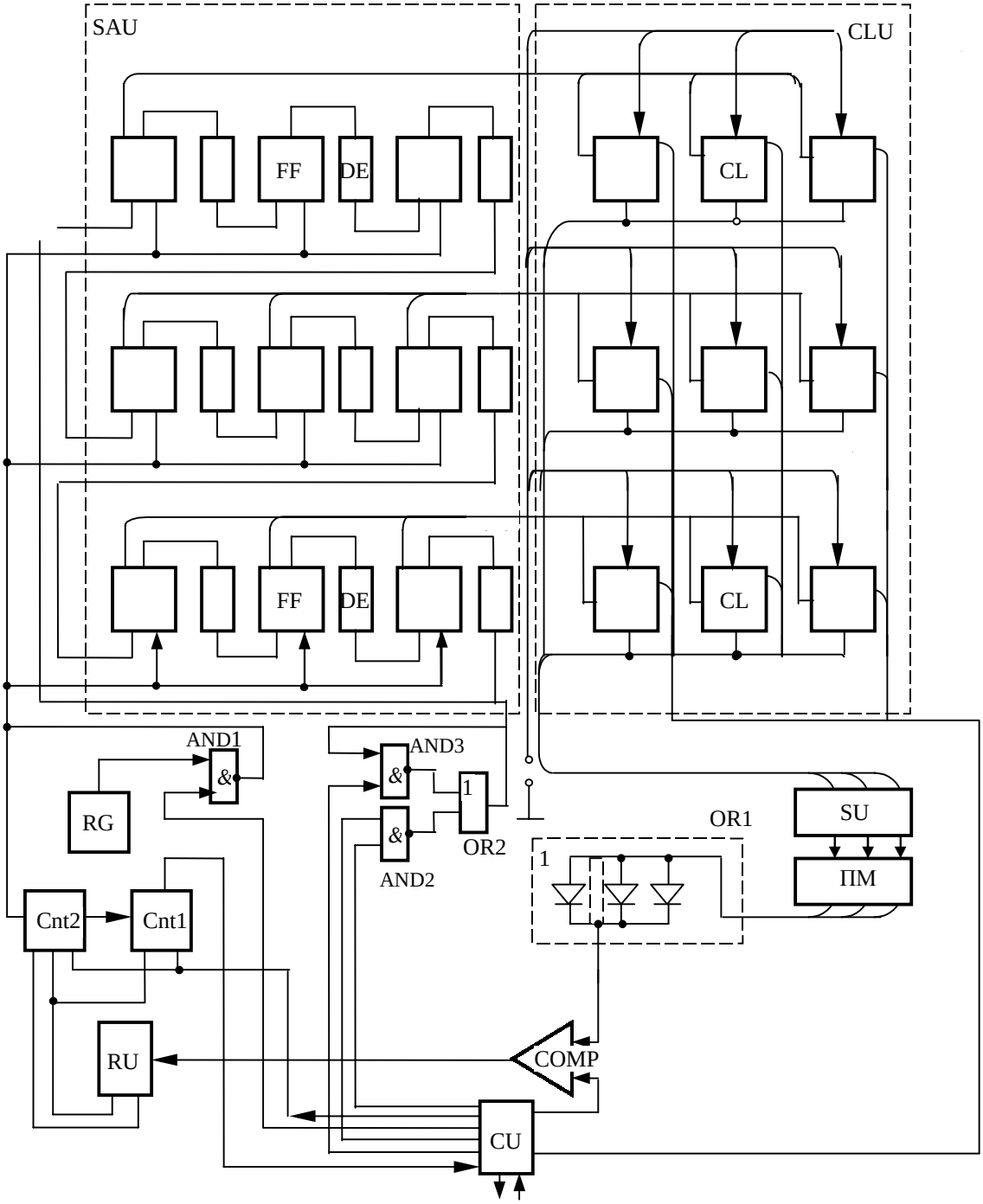


Fig 1. Simulator for allocation of thermal sources

current lead-ins *CL*; switching unit *SU*; control unit *CU*; *OR1* gate; comparison unit *COMP*; recording unit *RU*; counters *Cnt1* and *Cnt2*; pulse generator *PG*; *AND1-AND3* gates and the *OR2* gate. The device operates as follows. In the initial state, all the *SAU* flip-flops are in the "0" state. Gates *AND1-AND3* are disabled and the current lead-ins are in the state corresponding to I = 0, where I is the controlled current-lead-in current. The pulse counters are in the "0" state. Via the *CU*, the digital computer *DC* program enables the inputs of the *AND1* and *AND2* gates. As a result, the *SAU* starts receiving information from the *DC* via *AND2*, and shift pulses from the *PG* via *AND1*. In so doing, the assignment domain is set in the *SAU* flip-flops via *AND2* according to the program in the *DC*. Having set the information in the *SAU*, *AND2* is disabled. According to the data calculated in the *DC*, the control unit outputs voltage to the *COMP* unit and a control signal to the controlled current lead-ins. The *CU* sends a coded signal to the counters to set therein numbers proportional or equal to the co-ordinates x and y of the centre of the assignment domain. The number corresponding to co-ordinate x is set in *Cnt2*, and the number corresponding to the co-ordinate y is set in *Cnt1*. This completes the phase of setting up the device for solving the problem.

Problem solving is initiated by a command from the *DC*, or by depressing a button. Clock pulses from the *PG* are fed to the *SAU*, and the assignment domain starts shifting from left to right and from up to down. The voltage applied from the *CU* to the control input of the current lead-in determines the value of the current in the current lead-ins. Thus, in shifting, the assignment domain will "bypass" the *SAU*, and, hence, the *CLU*. Therefore, the passive model *PM*, connected to the *CU*, will receive current via the current lead-ins, but only to points corresponding to the places of allocation and to the shape of the assignment domain.

The current flow time interval depends on the repetition rate of the shifted pulses. The number of pulses fed to *SAU* is controlled by the pulse counters, resulting in that their content during the entire period of problem solving is actually the co-ordinates x and y of the assignment domain centre.

Since all the points of the model connected to the *CU* are connected to *OR1* and to *COMP*, as soon as the voltage in any point of the model becomes greater than (or equal to) the admissible voltage from *COMP*, the latter outputs a signal to enable the recording unit, which registers the domain co-ordinates.

These data can be input via the *CU* to the *DC*, which will register those points of the assignment domain centre where the *COMP* had triggered.

When the assignment domain has scanned all the flip-flops and taken its initial state, Cnt1 outputs an end-of-solution pulse to the *CU*, and the *DC* will receive a command to issue new, or modified information.

Hence, the device described identifies the domain of feasible allocation of the heat source without violating the constraint on the object's maximum temperature.

In the general case, the problem can be stated as follows. The temperature field is formed by not only the boundary conditions and the source being allocated, but also by the fixed discrete sources found in the object. Different admissible temperature values are specified for some points in the domain. In addition, inhibition domains are specified, i.e. domains wherein it is prohibited to allocate sources due to design, technological and other reasons. It is required to find such feasible heat source allocation zones, which do not to disturb the constraints on the temperatures in the points being controlled. This problem is solved by means of device [5].

SUCCESSIVE ALLOCATION OF HEAT SOURCES

The problem of successive allocation of heat sources in a given domain arises when it is necessary to arrange these sources in such a manner that the temperature field generated therein minimises a functional (target function) under given constraints on the maximum (minimal) temperature.

The search consists in determining the temperature when locating the heat source in all points of the given neighbourhood of its initial allocation. The value of the target function, when locating a source in a successive point, is compared to its value when the source was located in the previous point. Analysing the temperature functional in all points of the given neighbourhood allows to determine the location of a source with a better value of the target function. The centre of the new neighbourhood being investigated is relocated to this point, and the values of the target function are determined in this neighbourhood similar to that for the previous neighbourhood. The search process is terminated if the target function in the source location found

satisfies the given system of constraints, or if there exists no better value of the target function in the source's neighbourhood.

The schematic of device [6], which implements the scheme of searching for a location of a heat source just described, is shown in Fig. 2. In this case, the objective is to maximise the temperature, the constraint not exceeding the given value.

The codes of the co-ordinates of the source's initial location are input to the control unit *CU*. After the device has been enabled, these codes are input to registers *Rg1* and *Rg2*. Then the code from these registers is applied to only one *AND* gate of the network corresponding to these registers. The signal from this *AND* gate is fed to its respective gate *G* of the current lead-ins switch *SW*, thus enabling a current flow from the current source *CS* (which is an analog of the heat source) to the respective gate of the passive model *PM* node.

The current fed to the given *PM* node, in

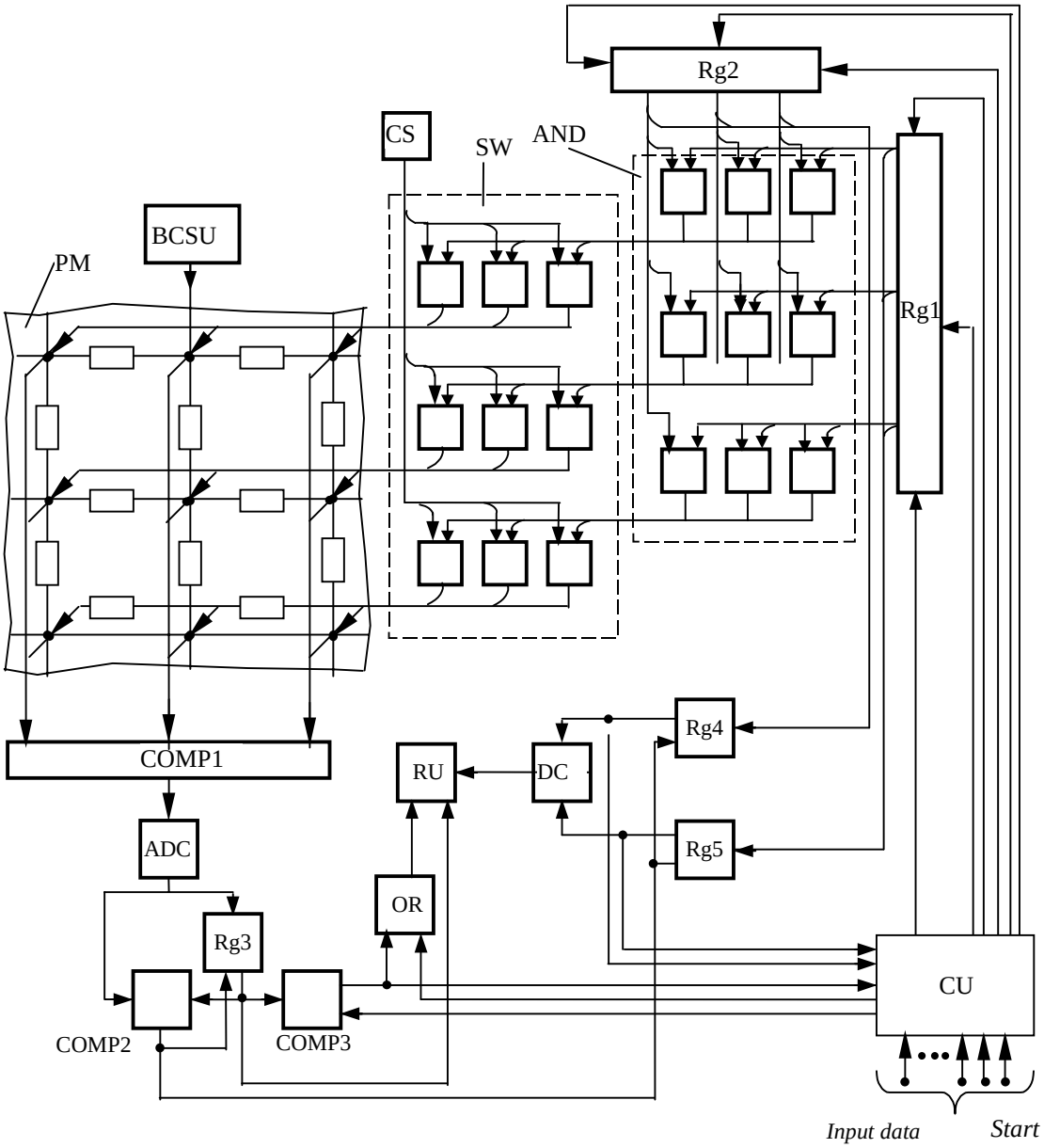


Fig. 2. Device for solving heat transfer problems

combination with the boundary conditions setting unit *BCSU*, forms an electric field corresponding to the temperature field being investigated. The potentials from the controlled PM nodes are applied to the input of the comparison unit *COMP1*. This unit selects the maximum potential and applies it to the input of the analog-to-digital converter *ADC*, which converts it to a code corresponding to the given potential. This code is then applied to the comparison unit *COMP2*, whose other input is at «0» level because register *Rg3* has not yet received a control signal to enable information writing. *COMP2* compares the code corresponding to the maximum potential and the «0» level to deliver a signal to registers *Rg1* - *Rg5*, thus enabling information writing (storage). Register *Rg3* stores a code corresponding to the maximum potential of the source location being analysed. Registers *Rg4* and *Rg5* store codes that characterise its location. The code in *Rg3* is applied to the comparison unit *COMP3*, at whose second input there is present a previously applied code corresponding to the admissible temperature. If this temperature value exceeds the value obtained in the model, *COMP3* will not output a termination signal.

Further, a signal from the *CU* initiates a shift of the "1's" in registers *Rg1* and *Rg2* one bit "up" and to the "left", i.e. to a point, which is the first one among the analysed points that are closest to the initial location of the source. The temperature field corresponding to this source location is analysed similar to that as described above. If the temperature corresponding to the new source location is less than that obtained at the previous step, information is not rewritten into registers *Rg3* to *Rg5*. Otherwise the information is rewritten as described above.

By controlling the location of "1's" in registers *Rg1* and *Rg2*, one can scan all the points in the neighbourhood of the initial location of the source. Having analysed the last point of the neighbourhood being investigated, *Rg3* and decoder *DC* output signals to recording unit *RU*, which registers the maximum potential and the number of the respective *PM* node. Further, the process runs in a similar manner for the new location of the source, etc. until the maximum temperature shall exceed the admissible value at a successive step. In this case, *COMP3* outputs a signal to the *OR* gate, thus enabling *RU* to register the given temperature value and the respective source location (number of *PM* node). At the same time, the *COMP3* output is applied to the input of the flip-flop to disable the flow of clock pulses from the clock pulse generator. This terminates the process of problem solving and disables device operation when allocating one heat source.

The approach described above is suitable for allocating one source. If several sources are to be allocated, then one of the sources is selected and the problem for this source is solved as described above. In so doing, the remaining sources retain their initial positions. Then the same search process is carried out for another source, etc. The search process in this case is considered terminated if neither source can be moved to other points of its neighbourhoods with a improvement on the target function.

ALLOCATING HEAT SOURCES OF DIFFERENT TYPES TO FIXED POSITIONS

In designing engineering systems, there occur problems when, for each source of a set of sources (in the general case of sources of different intensity and different spatial geometry of carriers), it is necessary to specify a location taken from a number of a priori given sites. A partial problem is the case of allocating intensities of sources to fixed positions, i.e. on the assumption that the carriers of sources have the same spatial geometry. An example of this can be the problem of packaging a rack having mounting locations, with electronic units having different heat liberation values, or creating different electromagnetic fields.

The characteristic features of the given class of problems are the a priori specified constraints on the location of sources in the domain and on the resulting temperature field.

For the sake of clarity, let us consider the problem of allocating point heat sources of different intensity, which, together with the boundary conditions, create a stationary field in a 2-dimensional domain. The problem is to define such a rearrangement of sources, which would ensure that the temperature values in the monitored points do not exceed an a priori specified value. Since the sources are point ones, the conditions of their mutual non-intersection and belonging to the domain are not checked.

The mathematical statement of the this problem leads to searching for such rearrangement of sources, which would satisfy the inequality

$$\min_{\pi \in \mathbf{Đ}} \max_{P_k \in \hat{\mathbf{U}}} T(P_k, \pi) \le T^* \quad (k = 1, 2, ..., p), \qquad (1)$$

where T is a function characterising the temperature field in domain Ω; π is the rearrangement of a set of rearrangements Π; Pk (k = 1, 2, …, p) are the temperature monitoring points; and T* is the critical temperature.

Let us consider a feasible hardware implementation [7] of the algorithm of solving problem (1). The input information of the *SAU* (Fig. 3) are the codes corresponding to the intensities of the sources being allocated. These codes are input to the rearrangement enumeration unit *REU* [8] intended for enumeration of n! permutations of n symbols, which are the intensities of sources. Since the *REU* outputs are connected to the digital-to-analog converter *DAC*, the codes corresponding to the intensities of sources are converted to currents-analogs of the sources. The *DAC* output currents flow to the nodes of the passive model *PM* to form, in combination with the boundary conditions setting unit *BCSU*, an electric field. The potentials from the outputs of the R-array (the monitored model nodes), which are proportional to the values of the field in these points, are applied to the inputs of the comparison unit *COMP*. The unit selects the maximum potential applied to the first input of adder *ADD*. The potential corresponding to the admissible temperature is applied to the second input o *ADD* from the voltage divider *VD*. The *ADD* output error signal is fed to the *SAU* input. If the difference between the selected maximum potential value and the admissible value satisfies constraint (1), the problem is considered solved. *SAU* outputs a control signal to the recording unit *RU*, which registers the corresponding rearrangement. This signal also disables the flow of clock pulses from the pulse generator in the *SAU*. If the field formed by using the model does not satisfy the constraint (1), clock pulses continue to flow to *REU*, thus enabling the formation of new rearrangements of sources to be analysed similarly. The device described allows to minimise the maximum temperature of the

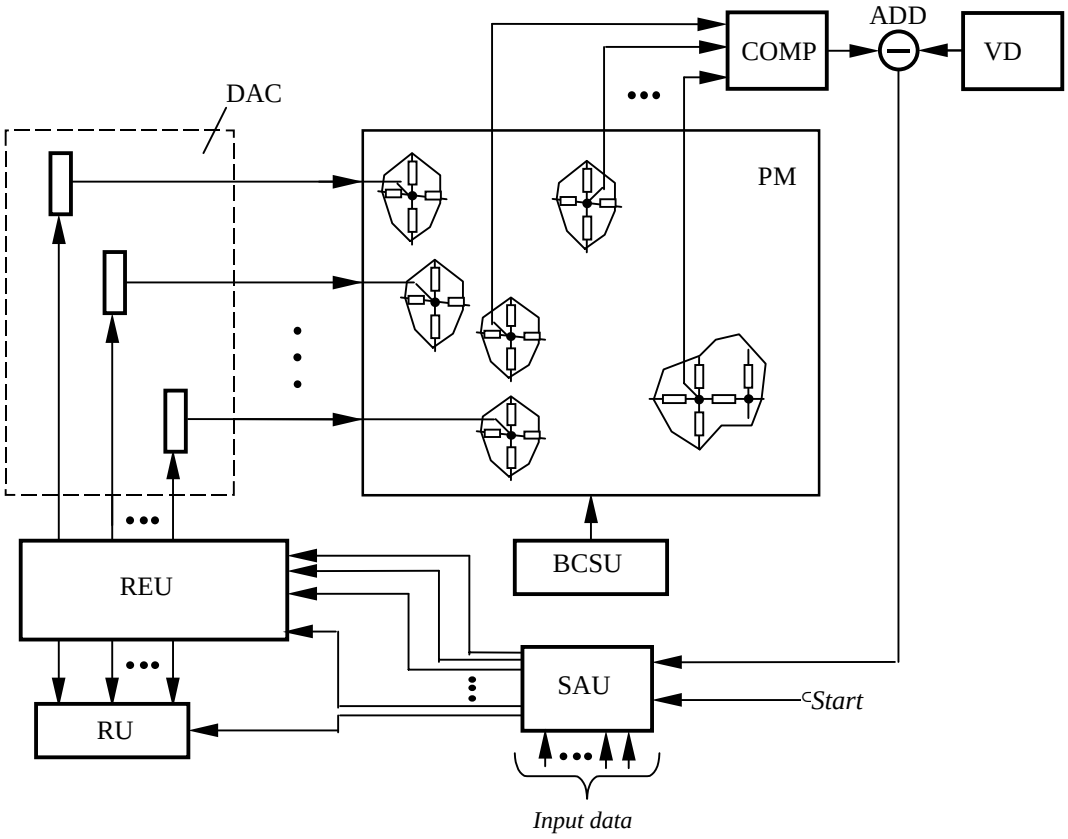


Fig. 3. Device for allocating heat sources to fixed positions

points monitored to an a priori level b enumerating the rearrangement of sources.

Often, when designing engineering systems, the number of sites differs from the number of objects being allocated. For instance, during thermal design of electronic equipment the number of sites that could accommodate electronic components, which are heat sources, is greater, as a rule, than the number of components proper. In this case, it is necessary to enumerate the feasible allocations of sources to select their optimal arrangement based on the criteria of minimising the system's maximum temperature. This problem is solved by means of device [9].

Similar devices [10] can be used for thermal design of electronic equipment, namely, during packaging/heat synthesis of electronic units to minimise their geometric characteristics (overall dimensions, areas of boards and volumes of electronic devices), and consequently, the masses of electronic components and units.

CONCLUSION

All the devices described allow for modifications, making it possible to find multi-choice solutions. This is often necessary in thermal design of electronic equipment for the designer to have the option of choosing such a solution from a set of admissible solutions, which would best satisfy the design, technological and other constraints unaccounted for in the process of solving the basic problem. In this case, one carries out additional investigations with other combinations of allocating sources and constraints. Such investigations are called forth by considerations evolving in the process of developing a streamlined design of an electronic unit.

In conclusion, let us note that in the investigations described in [9, 10, 11 and 12] the authors made an attempt, on the one hand, to overcome the low accuracy of analog calculations and the impossibility of performing logic operations in analog computers. On the other hand, the authors made an attempt to meet cost-performance challenges involving extended digital computer time needed for multi-choice solutions of field theory problems. All this proves the effectiveness of combining both types of computers in problem-oriented hybrid systems. In particular, in paper [11], the authors offer one of the feasible structures of a problem-oriented hybrid computing system intended for optima allocation of physical objects.

REFERENCES

1. Kostikov A. O. and Matsevity Yu. M., Thermal designing of a ceramic printed ci rcuit board with a micro-channel cooling system, *Proc. of 5th Sino–Russian–Ukrainian Symposium on Space Science and Technology*, Harbin, China, June 6–9, 2000, Harbin Institute of Technology, P. 738–743.

2. Kostikov A., Matsevity Yu. and Tsakanyan O., Cermet Heating Elements in the System of Providing the Thermal Regime of Spacecraft Devices, *Proc. 34th Space Congress*, Cocoa Beach, Florida, April 29 – May 2, 1997., Session IIIB, 6

3. Matsevity Yu. M. and Lushpenko S. F., *Identifying the heat and physical properties of solids*, Kiev: Nauk. Dumka, 1990, 216 pp.

4. *Inv. Cert. USSR No. 568953, IPC2 G 06 G 7/56*. Device for simulating the interaction of a moving heat source with a temperature field, Yu. M. Matsevity, Yu. G. Stoyan and V. N. Shapran.– Publ. 15.08.77. Bulletin № 30.

5. *Inv. Cert. USSR No. 1251122, IPC4 G 06 G 7/56*. Device for simulating physical fields, Yu. M. Matsevity, Yu. G. Stoyan and V. P. Putyatin.– Publ. 15.08.86. Bulletin № 30.

6. *Inv. Cert. USSR No. 1278900, IPC4 G 06 G 7/56*. Device for solving heat transfer problems, Yu. M. Matsevity, Yu. G. Stoyan and V. P. Putyatin.– Publ. 23.12.86. Bulletin № 47.

7. *Inv. Cert. USSR No. 1059584, IPC3 G 06 G 7/56* . Device for simulating physical fields, Yu. M. Matsevity, Yu. G. Stoyan, V. P. Putyatin et al. – Publ. 07.12.83. Bulletin № 45.

8. *Inv. Cert. USSR No. 748416, IPC2 G 06 F 15/20* . Device for enumerating rearrangements, S. N. Borisov, O. V. Viktorov, L. N. Minin et al. – Publ. 15.07.80. Bulletin № 26.

9. *Inv. Cert. USSR No. 1246120, IPC4 G 06 G 7/46* . Device for solving inverse field theory problems, Yu. M. Matsevity, Yu. G. Stoyan, V. P. Putyatin et al. – Publ. 23.07.86. Bulletin № 27.

10. *Inv. Cert. USSR No. 1109767, IPC3 G 06 G 7/46* . Device for solving inverse field theory problems, Yu. M. Matsevity, Yu. G. Stoyan and V. P. Putyatin. – Publ. 23.08.84. Bulletin № 31.

11. Stoyan Yu. G. and Putyatin V. P. *Optimising engineering systems with sources of physical fields*, Kiev: Nauk. Dumka, 1988, 192 pp.

12. Matsevity Yu. M. and Kuniesh J. *Hybrid simulation of heat processes*. Kiev: Nauk. Dumka, 1987, 268 pp.

4th International Conference of Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSE ANALYSIS OF HEAT FLOW AT A SOLID - SOLID ELECTRO-THERMAL CONTACT

G. Le Meur, B. Bourouga, Y. Jarny

LTI- UMR CNRS 6607 Ecole polytechnique de l'université de Nantes
rue Christian Pauc BP 90604 - 44306 Nantes cedex 03, France
brahim.bourouga@polytech.univ-nantes.f

ABSTRACT

An experimental methodology and an inverse approach are presented to model the thermal condition at the electro-thermal contact between two solid bodies. The modelling heat conduction equations at the interface involve two contact parameters which are estimated according to steady state and transient analysis. The results of both methods can then be compared.

The transient method is based on the minimization of a least squares criterion by using a Gauss-Newton algorithm. The experimental conditions are designed in order to satisfy the linear assumption of the modelling equations, and thus to simplify the sensitivity analysis.

The inverse analysis shows that the contact parameters to be estimated are highly correlated. the correlation coefficient can be reduced by choosing the time interval of measurements. The optimal extent of this interval is prone to the parameter values to be estimated.

NOMENCLATURE

a_i : Thermal diffusivity of material i (m²/s)
b_i : Thermal effusivity of material i (W/m².K.s$^{1/2}$)
cc : Correlation coefficient
c_{pi} : Specific heat of material i (J/Kg.K)
L_i : Distance between boundary condition thermocouple and interface in material i (m)
P_i : Voluminal dissipated heat flux in material i (W/m^3)
R_{TC} : Thermal contact resistance (m².K/W)
R_{EC} : Electrical contact resistance (Ω.m²)
S_R : Measurement sensitivity to R_{TC} (W/m²)
S_α : Measurement sensitivity to α (K)
t_w : Estimation window (s)
T^i_{mj} : Measured temperature by thermocouple j in material i (°C)
x^i_{mj} : Position of thermocouple j in material i (m)
α : Partition coefficient of generated heat flux at the interface
φ_g : Generated heat flux at the interface (W/m²)
λ_i : Thermal conductivity of material i (W/m.K)
θ_i : Temperature of water box i (°C)
ρ_{ei} : Electrical resistivity of material i (Ω.m)
ρ_i : Density of material i (Kg/m^3)
σ_i : Electrical conductivity of material i
i = 1,2
j = 1,2

INTRODUCTION

In the case of the perfect thermal contact, the boundary condition at the solid-solid interfaces does not pose a particular problem in the sense that the two equations which describe it are well-known : they are the equations of temperature continuity and heat flux conservation. When it is established that the contact remains imperfect, a temperature jump prevails at the interface. In the absence of a generated heat flux, this temperature jump is entirely due to the thermal contact resistance R_{TC}. On the other hand, when there is a thermal dissipation at the interface, heat flux generated φ_g also takes part in the temperature jump; but not in all its integrality. That introduces a second parameter into the expression of the boundary condition. This parameter is defined as a ' partition coefficient of heat flux generated at the interface ' which is noted α ($\alpha < 1$). This type of boundary condition at internal bounderies can be found in several cases such as sliding or rolling contacts and the thermoelectric contacts (connectors, spot welding...).

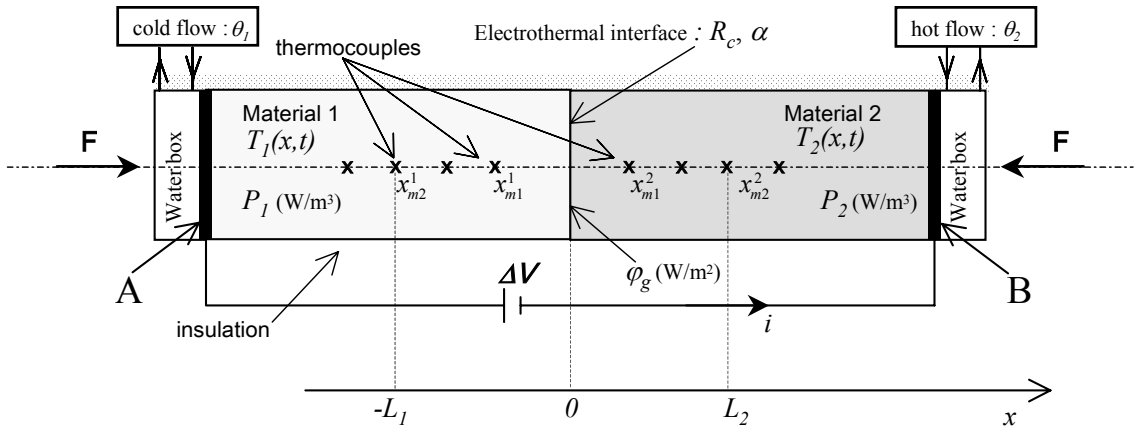


Figure 1 : Scheme of measurement principle

The definition of α knew a significant evolution that one finds primarily in studies on the sliding contacts. L. Mazo and *alii* [1, 2] and Badie-Levet [3] initially defined it in the case of the sliding contact, like a ' partition coefficient of generated heat flux by friction at the interface '. More recently, about the same subject, J P. Bardon [4] introduced a new concept : the one of ' share of the generated heat flux which takes part in the temperature jump '. Always in the case of the sliding contact, this concept, probably intrinsic to the interface, was taken again thereafter by P. Chantrenne and M. Raynaud [5] which consider that heat flux is dissipated in a subjacent volume of the interface. N. Laraqi proposes an approach of the generation of flux by friction based on a Gaussian distribution of the zones of dissipation [6], but the problem involved in the characterization of the intensity of the sources remains posed. Recently, we proposed a microscopic ideal model considering a current/flux tube which crosses a thermoelectric contact [7]. The results show that coefficient α approaches the ratio of the Lorentz numbers defined by the Wiedman - Frantz law. But this results is unsatisfactory. It does not take into account topography of both surfaces in contact.

Lastly, let us note that the bibliography reveals that, for a long time, the electrical contact arouses interest [8, 9] but the thermal boundary condition did not the subject of a particular attention of the electricians.

In this work, we are interested in the thermal characterization of a static thermoelectric contact where the heating is not too high, i.e. where the heat conduction problem remains linear and where parameters R_{TC} and α can be supposed constant. We present the principle of an experimental methodology which leads to a simultaneous estimation of both parameters R_{TC} and α. This text is organized into three sections. In the first, one presents the principle of measurement. The second is devoted to the sensitivity analysis. In the third section we present results of the inverse analysis optimizing the parameters estimation.

PRINCIPLE OF MEASUREMENT OF CONTACT PARAMETERS

We develop here an experimental approach which can be exploited in stationary thermal mode as in transient mode. The two methods of measurement must give the same result but probably not with the same precision. We are interested primarily in the transient method and more particularly with the identifiability problem with a view to study the dynamic contact where the two required parameters are temporal laws.

Principle of Measurement in Transient Model

The principle of measurement of the parameters R_{TC} and α is found on the transient analysis of the response of two cylindrical samples of the same geometry in imperfect contact. It is schematized on figure 1. The two cylinders are assembled aligned and tightened by means of a press which applies a normal effort F to them. They can be crossed by a heat flux and/or an electrical current. The electrical current is provided by a stabilized power supply connected at the two ends of the cylinders (electrodes A and B). The heat flux is ensured by two water boxes supporting two water flows maintained at two different temperatures θ_1 and θ_2. The side surface is thermally insulated.

The contact in $x = 0$ is imperfect. It is thus seat of an electrical contact resistance which opposes

to the current and of a thermal contact resistance which opposes to the heat flux. The electric contact resistance noted R_{EC} is supposed to be known starting from an auxiliary measurement. It determines generated heat flux density φ_g (W/m²) at the interface $x = 0$ when an electrical current crosses the interface. In the same way, the thermophysical characteristics of both materials are assumed well known, in particular the electrical resistivity which determines the voluminal sources P_1 and P_2 (W/m^3) during the current passage.

On such a device, the estimation of the thermal contact resistance R_{TC} and the partition coefficient of generated heat flux can be carried out by means of a 1D transient linear conductive model defined on the domain [$-L_1$; L_2] for $t > 0$ and by the measurement of the temperature field on both sides of the interface along the cylinders axis in quite selected points. Bounderies $x = -L_1$ and $x = L_2$ are equipped with thermocouples which record permanently the respectively noted temperatures $T_{-L1}(t)$ et $T_{L2}(t)$. Thus, we consider boundary conditions of first kind on these two bounderies. The linear system describing the heat transfer in such a device can be formulated as follows:

$$\frac{1}{a_1}\frac{\partial T_1}{\partial t}=\frac{\partial^2 T_1}{\partial x^2}+\frac{P_1}{\lambda_1} \quad -L_1\le x\le 0, \quad t>0 \tag{1a}$$

$$\frac{1}{a_2}\frac{\partial T_2}{\partial t}=\frac{\partial^2 T_2}{\partial x^2}+\frac{P_2}{\lambda_2} \quad 0\le x\le L_2, \quad t>0 \tag{1b}$$

$$T_1(-L_1,t)=T_{-L1}(t) \quad x=-L_1, t>0 \tag{1c}$$

$$\lambda_1\frac{\partial T_1(0^-,t)}{\partial x}=\lambda_2\frac{\partial T_2(0^+,t)}{\partial x}+\varphi_g \tag{1d}$$

$$\lambda_1\frac{\partial T_1(0^-,t)}{\partial x}+\alpha\varphi_g=\frac{T_2(0^+,t)-T_1(0^-,t)}{R_{CT}} \tag{1e}$$

$$T_2(L_2,t)=T_{L2}(t) \quad x=L_2, t>0 \tag{1f}$$

$$T_1(x,0)=T_i^1(x) \quad -L_1\le x\le 0, t=0 \tag{1g}$$

$$T_2(x,0)=T_i^2(x) \quad 0\le x\le L_2, t=0 \tag{1h}$$

$T_i^j(x)$ (where j=1, 2) represent the initial distribution of temperature of the whole of the device before the electrical current passage. The boundary condition at the interface $x = 0$ is composed of two equations : the heat flux conservation equation (1d) and a condition of third kind (1e).

It is the latter which interests us since the temperature jump at the interface is described by means of the two required parameters R_{TC} and α. The principle of estimation R_{TC} and α is based on the resolution of the inverse heat conduction problem whose direct problem is given by the equations system (1). The resolution of the inverse heat conduction problem will be based on the knowledge of the following data :

- Two temperature recordings within the cylinder 1 noted T_{m1}^1 et T_{m2}^1 and two other temperature recordings within the cylinder 2 noted T_{m1}^2 et T_{m2}^2. These measurements are located at x-coordinates : x_{m1}^1, x_{m2}^1 in medium 1 and x_{m1}^2 et x_{m2}^2 in medium 2.
- The voltage recording (current intensity is imposed and well known).
- The electric and thermal characteristics of materials are supposed known, uniform and constant.
- the electric contact resistance at the interface of the cylinder is the result of an auxiliary measurement. It is used to determine φ_g.
- Note : although the problem (1) is linear, one retained the principle of a numerical resolution of the direct problem. The discretisation of the equations uses the finite difference method. We consider the unconditionally stable scheme of Cranck-Nicholson.

Principle of Measurement in Steady State

In steady state, the estimation of both parameters is done in two stages. Initially, one chooses the outside temperatures θ_1 and θ_2 in the water boxes so that a significant thermal gradient

	ρ_e (Ωm)	λ (W/m.K)	C_p (J/Kg.K)	ρ (Kg/m^3)	$a=\frac{\lambda}{\rho C_p}$ (m²/s)	$b=\sqrt{\lambda\rho C_p}$ (W/m².K.s$^{0.5}$)	L_i (mm)	D (mm)	T_{Li} (°C)
Steel	10^{-6}	40	500	7300	$1{,}1\ 10^{-5}$	12083	40	40	0
Copper	$5\ 10^{-8}$	320	381	8500	$0{,}99\ 10^{-4}$	32192	40	40	0

Table 1 : Physical properties

is established according to the axis of the both cylinders. The measurement of the steady state thermal field according to the axis of both cylinders permits to extrapolate the temperature jump at the interface. The knowledge of the thermal gradient determines the heat flow which crosses the interface. The R_{TC} thus is measured.
In the second stage, one makes circulate an electrical current in the device by means of a feeding connected on electrode A and B (figure 1). A flux density φ_g is generated at the interface $x = 0$ seat of an electric contact resistance. Knowing the value of the R_{TC}, it is easy to determine the value of the partition coefficient α.

SENSITIVITY ANALYSIS

The sensitivity coefficient of the thermal contact resistance will be noted S_R and that of the partition coefficient S_α. The sensitivity coefficient being defined like the derivative of the measured parameter compared to the parameter to estimate. One writes that:

$$S_R = \frac{\partial T_j}{\partial R_{tc}} \text{ (W/m²)} \tag{2a}$$

$$S_\alpha = \frac{\partial T_j}{\partial \alpha} \text{ (K)} \tag{2b}$$

S_R is a heat flux density and S_α a temperature. To obtain the formulation of the sensitivity fields in the calculation domain which interests us, one successively derives the system from equations (1) compared to R_{TC} and α in accordance with the relations (2). The equations system describing the sensitivity coefficient field of the R_{TC} is similar to

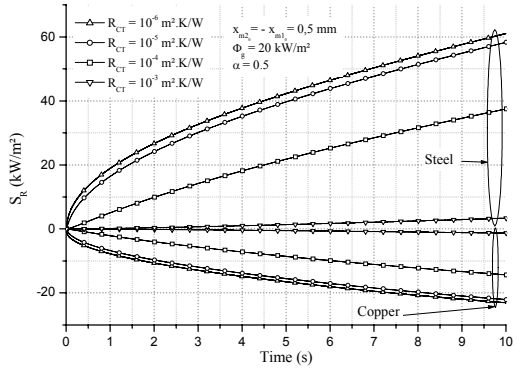


Figure 2 : S_R versus time and R_{TC}.

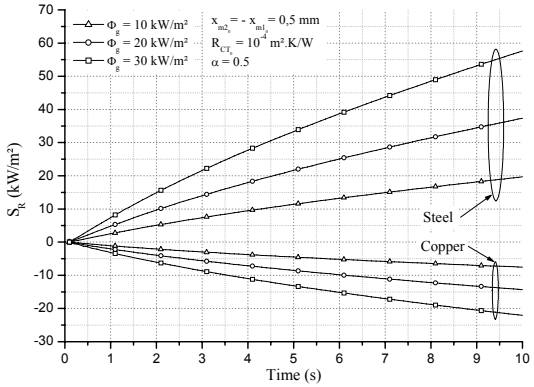


Figure 3 : S_R versus time and φ_g

the system (1) with some differences : all the equations are homogeneous except the derivative of the equation (1e) which, in accordance with (2a), is written as :

$$\lambda_1 \frac{\partial S_R(0,t)}{\partial x} = \frac{1}{R_{TC}} \left(S_{R1}(0,t) - S_{R2}(0,t) + \lambda_1 \frac{\partial T_1(0,t)}{\partial x} - \alpha \varphi_g \right) \tag{3}$$

Thus, for a material couple given, one has :

$$S_R = S_R(\alpha, \varphi_g, R_{TC}, x, t) \tag{4}$$

In the same way, the equations system describing the sensitivity coefficient field of α is similar to the system (1) except equation given by the derivative of (1e) which, in accordance with (2b), is written as :

$$\lambda_1 \frac{\partial S_\alpha(0,t)}{\partial x} + \varphi_g = \frac{S_{\alpha 1}(0,t) - S_{\alpha 2}(0,t)}{R_{TC}} \tag{5}$$

For a material couple given, one thus has :

$$S_\alpha = S_\alpha(\varphi_g, R_{TC}, x, t) \tag{6}$$

To appreciate the variations of S_R and S_α according to the parameters on which they depend, one considers a material couple. The numerical data of table 1 describe the experimental project :

For the R_{TC}, one considers a field of common value in the case of a dry contact metal – metal : $10^{-6} \leq R_{TC} \leq 10^{-3}$ Km²/W.

The field of value of φ_g ranging between 10 and 30 kW/m² is selected so that the heating of the control volume remains ranging between 0 and 50 K (linear problem).

The partition coefficient α varies from 0 to 1.

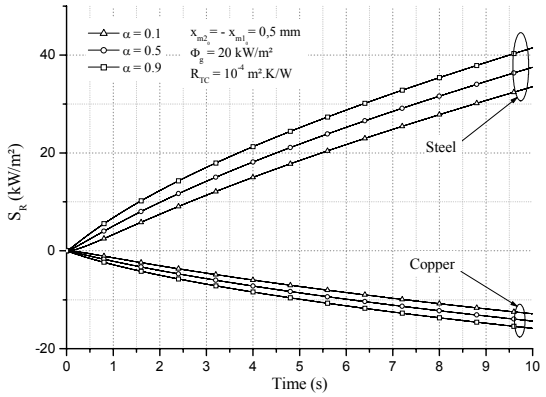


Figure 4 : S_R versus time and α

Sensitivity Coefficient of the R_{TC}. Figure 2 shows the behaviour of S_R according to the time at both measurement points $x^{1}_{m1} = x^{2}_{m2} = 0.5$ mm for various values of the R_{TC}. S_R is monotonous increasing in time and monotonous decreasing according to R_{TC}. The time scale retained in this calculation (10 seconds) is of the same order as the value of the fundamental time-constant of the studied thermal system. One notes that S_R has appreciable levels of value, about several 10 kW/m². For a given value of the R_{TC}, S_R is maximum in steady state. The sensitivity related to the measurement point in copper is definitely lower (in absolute value) than that related to the measurement point in steel.

Figure 3 shows the temporal law of S_R for R_{TC} = 10^{-4} Km²/W and three values of φ_g. It reveals that S_R is monotonous increasing according to φ_g.

Figure 4 shows the temporal law of S_R for R_{TC} = 10^{-4} m².K/W and three values of α. It reveals that

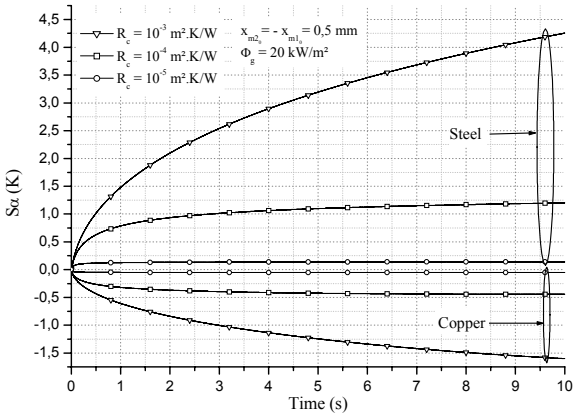


Figure 5 : S_α versus time and R_{TC}

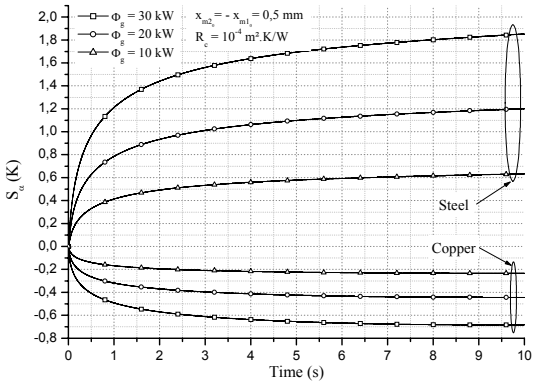


Figure 6 : $S\alpha$ versus time and φ_g

S_R is monotonous increasing according to α.

We also checked that S_R is decreasing when the measurement points are increasingly far away from the interface. But for the positions domain prospected (a few millimetres) these variations are not significant in front of the mean value level.

Sensitivity coefficient S_α. Figure 5 shows the sensitivity coefficient of α on the same scale of time. Like S_R, S_α is monotonous increasing in time and monotonous decreasing according to *RTC*. Figure 6 shows that S_α is monotonous increasing according to φ_g. S_α presents appreciable levels of value, about several Kelvins. There too, it is noted that the sensitivity related to the measurement point in copper is definitely lower (in absolute value) than that related to the measurement point in steel.

ESTIMATION OF THE PARAMETERS R_{TC} AND α

The selected estimation technique is a variant of the Gauss – Newton method.

Principle of the Estimation Method

Let β the vector to be estimated such as : $\beta = (R_{TC}, \alpha)$. The estimation method is based on the minimization algorithm of Gauss-Newton. The criterion of least squares to minimize can be written as:

$$J(\beta) = \frac{1}{2}\sum_{k=1}^{N_t}\left[Y(\beta)^k - \tilde{Y}^k\right]^2 \qquad (7)$$

where Y^k is the temperature vector calculated at time t^k at sensor locations x^1_{m1} and x^2_{m1}. $\tilde{Y}^k$ is the temperature measured at time t^k.

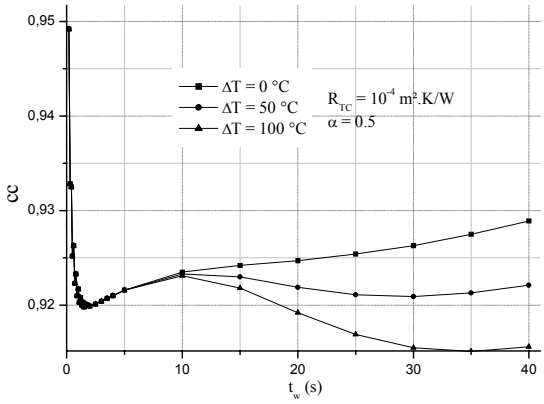


Figure 7 : *cc* versus tw and ΔT

Starting from an initial vector β^0 the algorithm determines at each iteration the vector β^{n+1} (n>0) given by the equation (8) :

$$\beta^{n+1} = \beta^n + \Delta\beta^n \quad (8a)$$

$$\Delta\beta^n = \left[X^{t(n)} X^{(n)} \right]^{-1} X^{t(n)} \left[\widetilde{Y} - Y^n \right] \quad (8b)$$

where X^n is the sensitivity matrix computed at iteration n.

The calculation of $\Delta\beta^n$ according to the equation (8)b poses the numerical problem of the inversion of matrix $\left[X^{t(n)} X^{(n)}\right]$ which is badly conditioned. This problem was regularized by multiplying the diagonal terms of the matrix by a factor $(1+\varepsilon)$, where $\varepsilon \langle\langle 1$. This operation introduces a bias into the final estimation. The value of ε is chosen to ensure an optimum conditioning and minimized bias.

The other improvement made to the method consisted in optimizing the length of the vector displacement of reiterated given by the equation (8b). The operation consists in balancing term $\left[X^{t(n)} X^{(n)}\right]^{-1} X^{t(n)} \left[\widetilde{Y} - Y^n\right]$ by a coefficient μ^n given by:

$$\mu^n = \frac{1}{2} \frac{J(\beta^n)}{J(\beta^0)} \quad (9)$$

The coefficient ½ was obtained empirically. This refinement of the Gauss – Newton method permits to divide the iteration count by a factor seven.

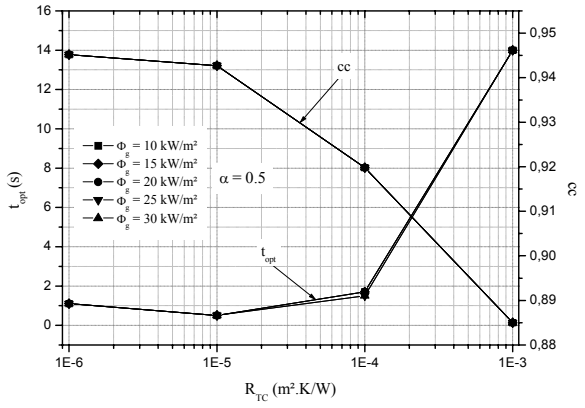


Figure 8 : t_{opt} and *cc* versus R_{TC} and φ_g

Identifiability Problem

The problem of identifiability is significant in this application. Indeed α has sense only if a significant heat flux is generated at the interface. One understands by ' significant ' the fact that generated heat flux is at the origin of a significant component of the heating which, generally in the case of the thermoelectric systems, is due to several internal and external sources. The origin of the generated heat flux at the interface is the electrical contact resistance which draws its substance from the electric constriction phenomenon. And this last is superimposed on the thermal constriction phenomenon. The function of constriction which one notes F is the same one in the two phenomenologies since electric and thermal constriction resistances, respectively R_{EC} et R_{TC}, are written :

$$R_{CE} = \frac{F}{\sigma} \quad ; \quad R_{CT} = \frac{F}{\lambda} \quad (10)$$

Thus this thermoelectric coupling can induce a certain correlation between the thermal contact resistance R_{TC} and the partition coefficient of generated heat flux α. This correlation is defined by a correlation coefficient noted *cc* given by :

$$cc = \frac{(X^T X)_{12}}{\sqrt{(X^T X)_{11} \cdot (X^T X)_{22}}} \quad (11)$$

where $(X^T X)_{ij}$ are the terms of the square matrix 2x2 $X^T X$.

The analysis of the curves of sensitivity for different geometries (in particular values of L_i *(i = 1,2)* shows that it is more interesting to carry

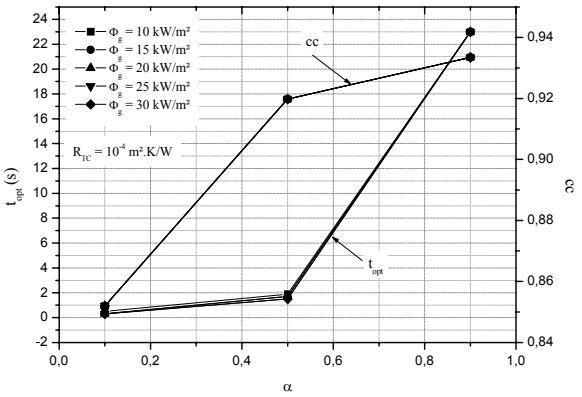


Figure 9 : t_{opt} and cc versus α and φ_g

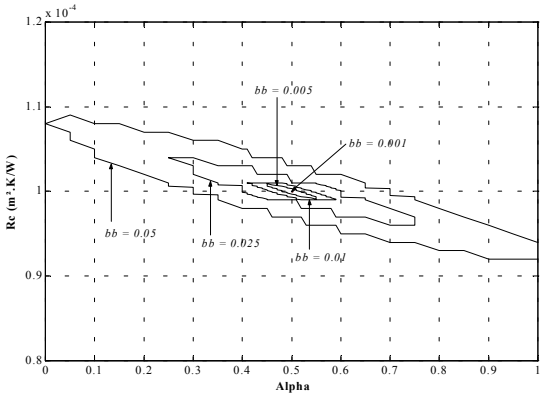


Figure 10 : Imprecision domain of estimation for different values of signal/noise ratio *bb*

out the simultaneous estimation of R_{TC} and α at the beginning of the transient mode. Physically that is explained by the chronology of the thermal events which are the heating components related to each source (internal or external). The heating component which appears in first in the interface region, i.e. very close to the measurement point is that due to the generated heat flux at the interface. It is this argument which led us to consider the behaviour of the correlation coefficient according to the extent of the window of estimation that one will note t_w. We represent this evolution of cc on figure 7. And for showing the chronology of the thermal events well one considered 3 values of the external source $\Delta T = T_{L2} - T_{L1}$. One observes on this figure that, whatever the value of ΔT, cc admits an optimal value 0.92 which always corresponds to the same temporal window covering the interval [0, t_{opt} = 1.1 s]. Beyond the value of t_w = 7 to 8 s approximately, the curves of cc separate but the difference remain very weak. And it is only beyond $\Delta T = 50$ *K* that appears a second minimum slightly more favourable to the estimation (cc = 0.915).

Knowing that the sensitivity coefficients S_R and S_α are functions of the interface parameters R_{TC}, α and φ_g, it is significant to emphasize the behaviour of t_{opt} according to the latter. Figure 8 shows that the optimal estimation window is extended more and more according to the R_{TC}, in particular beyond R_{TC} = 10^{-4} K.m²/W. This increase of the estimation window corresponds to a less and less strong correlation between both parameters to estimate. t_{opt} and cc are indifferent to the intensity of generated heat flux.

Figure 9 reveals that t_{opt} is monotonous increasing according to α. It is clearly more sensitive to α in the domain α > 0.5. The levels of value of t_{opt} in the domain α > 0.5 are an order of magnitude weaker than those observed in the domain α < 0.5. On the other hand cc is strongly sensitive to α in the domain α < 0.5. Let us note here that $\alpha\varphi_g$ is the share of generated heat flux at the interface in the most resistive material. This time the increase of the estimation window is accompanied by an increasingly strong correlation between both parameters to identify. It is there an opposite tendency of that observed on the variations according to R_{TC}.

Finally, the figure 10 presents the confidence region of the estimation according to the signal/noise ratio. It is still checked that the estimation quality is degraded according to an increasing amplitude of the noise. For a white noise of average amplitude of the order of the percent, the estimation of a R_{TC} of 10^{-4} K.m²/W is accompanied by a relative inaccuracy of about 1 %. That goes from pair with an inaccuracy on α of about 16 %. That is related to the geometrical configuration of the iso – $J(\beta)$ in the (RTC, α) plan.

CONCLUSION

The theoretical principle of an experimental methodology was presented to model the thermal condition at the electro-thermal contact between two solid bodies. The modelling heat conduction equations at the interface involve two parameters which are estimated according to both steady state and transient analysis. These two approaches are different but they are based on the same experimental apparatus. For the steady state experiments the estimated values are directly issued from algebraic equations and for the transient experiments, an inverse method is used.

The results of both methods can then be compared.

The transient method is based on the minimization of a L-S criterion by using a Gauss-Newton algorithm. The experimental conditions (cylinder sizes, heating level, sensor locations) have to be carefully designed in order to satisfy the linear assumption of the modelling equations, and thus to simplify the sensitivity analysis.

Numerical experiments have shown that the two contact parameters to be estimated are highly correlated. The inverse analysis shows that the correlation coefficient can be reduced by choosing the time interval of measurements. The optimal extent of this interval is prone to the parameter values to be estimated. Thus for a given set of experimental conditions, the resolution of the inverse problem is performed using an initial guess of the couple (RTC, α). This choice is based on physical considerations which take into account different experimental variables like the nature of the materials, the states of contact surfaces, the contact pressure. The first estimations can be then refined by adjusting the estimation time interval.

REFERENCES

1. L. Mazo, B. Cassagne, D. Badie-Levet, J. P. Bardon, Etude des conditions de liaison thermique dans le cas du frottement sec métal-plastique. Revue Générale de Thermique (204) (1978) 919-933.

2. L. Mazo, Contribution à l'étude thermocinétique du frottement métal-plastique. Thèse de doctorat, Université de Nantes, 1977.

3. D. Badie-Levet, Champs de température engendrés par le frottement de deux anneaux concentriques en rotation relative, dans le cas d'une condition de contact imparfait, Revue Générale de Thermique (275) (1984) 671-679.

4. J. P. Bardon, Bases Physiques des conditions de contact thermique imparfait entre deux milieux en glissement relatif. Revue Générale de Thermique (386) (1994) 85-91.

5. P. Chantrenne Et M. Raynaud, A microscopic thermal model for dry sliding contact, International Journal of Heat and Mass Transfer 40 (5) (1997) 1083-1094.

6. N. Laraqi, Température de contact et coefficient de partage du flux généré par frottement sec entre deux solides. Approche nouvelle de la génération de flux. International Journal of Heat and Mass Transfer 35 (11) (1992) 3131-3139.

7. Le Meur, B. Bourouga et J.P. Bardon, Etude théorique des mécanismes de base du transfert de chaleur à travers une interface siège d'une dissipation par effet joule, Actes du Congrès SFT, Nantes, 2001, Ed. Elsevier, pp 653-658.

8. R. Holm, Electric contacts. Theory and application, 4ème édition, Springer-Verlag, Berlin, 1967.

9. L. Fechant, Le contact électrique : phénomène physique et matériaux, éd. Hermes, collection SEE,1995.

ADAPTIVE MULTISCALE ESTIMATION OF A SPATIALLY DEPENDENT DIFFUSION FUNCTION WITHIN POROUS MEDIA FLOW

Harald Krüger
University of Bergen
Department of physics
Bergen, Norway
Harald.Kruger@fi.uib.no

Alv-Arne Grimstad
RF-Rogaland Research
Bergen, Norway
Alv-Arne.Grimstad@rf.no

Trond Mannseth
RF-Rogaland Research
Bergen, Norway
Trond.Mannseth@rf.no

ABSTRACT

Adaptive multiscale estimation of a spatially dependent diffusion function is considered. When utilizing pressure data, the method has been successful in identifying coarse-scale structures of the diffusion function. In this paper the method is used with pressure and fluid rate data as information sources. The estimate is expected to be improved when additional data is included. However, this expectation is not always met, and in some cases the simultaneous use of these data types in the estimation might even cause problems. The reason for the lack of improvement is investigated together with methods for successful estimation utilizing pressure and fluid rates.

NOMENCLATURE

$\boldsymbol{c}$	Parameter vector
$\boldsymbol{d}$	Measured production data
J	Objective function
$\tilde{J}$	Predicted attainable obj. function
k	Absolute permeability
$\boldsymbol{x}$	Spatial coordinate vector

Greek symbols

σ	Standard deviation
$\tilde{\sigma}$	Approximate standard deviation

INTRODUCTION

When producing a hydrocarbon reservoir, production data - well pressures and fluid well rates - can be used to infer properties of the reservoir. In this paper a porous medium containing two fluid phases is considered. The equations governing fluid flow in the reservoir are of advection-diffusion type. The spatially dependent diffusion function - the permeability - is estimated using multiscale inversion techniques.

With multiscale estimation (see e.g. [1][2][3][4][5][6]), the inverse problem is solved by performing a hierarchical search for the right parameterization while solving a sequence of estimation problems with an increasing parameterization dimension.

With *adaptive* multiscale estimation [7](see [8][9][10] for preliminary version of the methodology and [11] for an alternative methodology), the new parameterization at an arbitrary stage in the estimation sequence is such that new degrees of freedom are not necessarily introduced all over the porous medium. On the contrary, the aim is to introduce new degrees of freedom only where it is warranted by the data.

The parameterization selection is based on a linearization of the model output as a function of the parameters in the permeability expansion. It is therefore important that the linearization assumption is valid for this method to work properly.

Previous work [7] has shown that utilizing the well pressure alone leads to a good estimation of coarse-scale permeability structures. Including also fluid well rates in the data, one would expect to be able to improve on estimates based on well pressures only. However, recent studies show that the expectation of improvement is not met. The main topic of this paper is to find reasons for this and to suggest improvements.

The paper starts by describing the forward model and the adaptive multiscale estimation methodology used to solve the inverse prob-

lem. Then, examples of application of adaptive multiscale estimation are given, utilizing pressure data alone and both pressure and fluid rate data as information sources. The observations from these examples are then discussed. Finally, productive utilization of both pressure and fluid rate data is considered.

FORWARD MODEL

The two conservation equations describing two-phase incompressible, immiscible porous-media flow in a gravity-free environment with isotropic permeability are (see e.g., [12])

$$\phi(\boldsymbol{x})\frac{\partial s_o}{\partial t} - \nabla\cdot(\lambda_o k(\boldsymbol{x})\nabla p_o) = q_o(\boldsymbol{x}), \quad (1)$$

$$\phi(\boldsymbol{x})\frac{\partial s_w}{\partial t} - \nabla\cdot(\lambda_w k(\boldsymbol{x})\nabla p_w) = q_w(\boldsymbol{x}). \quad (2)$$

The subscripts o and w refer to the oil and water phases, respectively, which are assumed to be the only fluid phases present. Furthermore, $\lambda_i = k_{ri}(s_i)\mu_i^{-1}$, s_i is the saturation, p_i the pressure, μ_i the viscosity, q_i the external volumetric flow rate, and $k_{ri}(s_i)$ the relative permeability, each with respect to the ith fluid phase. The functions $\phi(\boldsymbol{x})$ and $k(\boldsymbol{x})$ are the porosity and the absolute permeability (i.e. the function to be estimated), respectively, of the porous medium. In order to close the system a completely saturated medium is assumed and an equation of state defining the capillary pressure, P_c, is needed,

$$s_o + s_w = 1, \quad (3)$$

$$p_o - p_w = P_c(s_w). \quad (4)$$

In addition to suitable initial and boundary conditions, the functions ϕ, k, k_{ri} and P_c have to be specified a priori in order to obtain meaningful solutions for s_i and p_i.

INVERSE PROBLEM

The inverse problem of finding k through analysis of production data, when the other functions are assumed known, is the topic of this paper.

General Methodology

Generally, permeability estimation requires the permeability to be parameterized. In this paper this is obtained by setting $k_N(\boldsymbol{x}) = \boldsymbol{c}_N^T\boldsymbol{\psi}_N(\boldsymbol{x})$, where $\boldsymbol{c}_N \in \Re^N$ is the parameter vector and $\boldsymbol{\psi}_N$ is an N-vector of real-valued piecewise constant basis functions spanning the space within which k is to be estimated. Let $\boldsymbol{d} \in \Re^M$ denote the measured time series of the production data, and let $\boldsymbol{m}(\boldsymbol{c}) \in \Re^M$ denote the corresponding production data calculated using the model equations.

The estimation problem is solved by finding the parameter vector $\boldsymbol{c}_N^{est} \in \Re^N$ which minimizes the objective function (output least squares)

$$J(\boldsymbol{c}_N) = (\boldsymbol{d} - \boldsymbol{m}(\boldsymbol{c}_N))^T \boldsymbol{D}^{-1}(\boldsymbol{d} - \boldsymbol{m}(\boldsymbol{c}_N)), \quad (5)$$

subject to $k > 0$. Here $\boldsymbol{D}^{-1}$ is the inverse of the covariance of the measurement errors. In order to solve this minimization problem the Levenberg-Marquardt algorithm, as implemented in [13], is used.

Multiscale estimation

The question of how to select the resolution, N, for the parameter estimation problem remains. The purpose of multiscale estimation is to avoid guessing at a fixed resolution and subsequently estimate all the corresponding parameters simultaneously. Multiscale estimation solves a sequence of parameter estimation problems with initial low, but increasing resolution.

Several sequential estimations are performed, each (from the second one) starting from the permeability estimate from the previous estimation, but with a more detailed representation of the permeability.

With ordinary multiscale estimation the resolution is gradually increased by the same amount in the entire grid. A risk of over-parameterization is present due to the rapid increase in the number of new parameters.

With adaptive multiscale estimation [7] the resolution is gradually increased only in regions of the grid where it is expected to be productive. The risk of over-parameterization is

therefore reduced. Figure 1 illustrates the difference between ordinary and adaptive multiscale estimation.

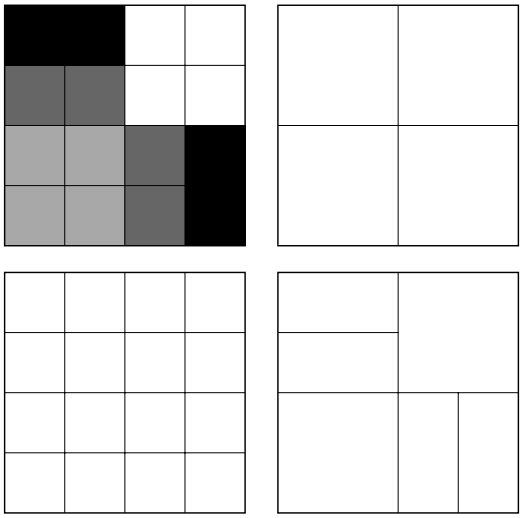


Figure 1: Illustration of excessive parameterization with ordinary multiscale estimation. Going from left to right: The true permeability, ordinary multiscale 4-parameter representation, ordinary multiscale 16-parameter representation, and a sufficiently detailed adaptive multiscale representation.

To perform adaptive multiscale estimation, a way to select parameterization refinement is needed. A description of such a method follows next. A more in-depth description of this method is given in [7]. An alternative method for parameterization selection is presented in [11].

Selection of parameterization

Introduction of new parameters will reduce the objective function value, but also increase computational cost and, in general, increase the parameter uncertainty. The parameterization must be selected in a proper way in order to obtain a successful result.

In order to select the most productive parameterization it would be helpful to obtain the minimum objective function values after introduction of each candidate refinement. This is, however, clearly unattainable due to computational costs.

To avoid this costly computation of the objective function values, the predicted attainable objective function value, $\tilde{J}$, is introduced. $\tilde{J}$ is an approximation of $\min J$ based on a linearization of $\boldsymbol{m}(\boldsymbol{c})$ and is therefore available by analytical calculation.

$$\tilde{J}(P_Q) = \Delta \boldsymbol{d}^T \left(D^{-1} - G\right) \Delta \boldsymbol{d}, \qquad (6)$$

$$G = \left[D^{-1} A_Q \left(A_Q^T D^{-1} A_Q\right)^{-1} A_Q^T D^{-1}\right], \qquad (7)$$

where $\tilde{J}(P_Q)$ is the predicted attainable objective function value with the parameterization P_Q, containing Q parameters. The sensitivity matrix for the parameterization P_Q is denoted A_Q, and $\Delta \boldsymbol{d}$ is the current residual.

In order to find the best parameterization P_Q with Q parameters, $\tilde{J}$ for different parameterizations containing Q parameters are compared. The parameterization with the lowest $\tilde{J}$ is the best choice with Q parameters (local winner).

The local winners for different numbers of parameters are then compared in order to find the parameterization to be used in the next step of the estimation sequence.

Generally, an increase in the number of parameters will decrease $\tilde{J}$. Hence, selecting the parameterization with the lowest $\tilde{J}$ will introduce more parameters than needed. In order to determine the number of new parameters to be introduced it is not optimal to have only $\min \tilde{J}$ as criterion. A measure of the uncertainty in $\tilde{J}$ is needed in order to assess whether the reduction of $\tilde{J}$ is significant.

The measurement errors are random and therefore $\tilde{J}$ is also random. The standard deviation, $\sigma(\tilde{J})$, of $\tilde{J}$ can potentially be used as a measure of the uncertainty of $\tilde{J}$. However, $\sigma(\tilde{J})$ is unknown. Instead we use an approximation [7], $\tilde{\sigma}(\tilde{J})$, to $\sigma(\tilde{J})$:

$$\tilde{\sigma}(\tilde{J}(P_Q)) = \sqrt{4\tilde{J}(P_Q) - 2\,(M - Q)} \qquad (8)$$

(M is the number of measurements of production data).

An increase in the number of parameters will reduce $\tilde{J}$ but if $\tilde{J}(P_{Q+1}) + \tilde{\sigma}(\tilde{J}(P_{Q+1})) \geq \tilde{J}(P_Q)$ the reduction is not considered to be significant and P_Q is the new parameterization. Figure 2 shows an example of choosing the parameterization.

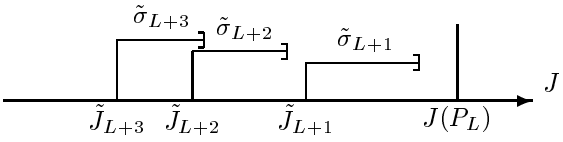


Figure 2: Example of ranking of predictions from refinement analysis. The original number of parameters is L. In this case Q=L+2, since $\tilde{J}_{L+3} + \tilde{\sigma}_{L+3} \geq \tilde{J}_{L+2}$ and $\tilde{J}_{L+2} + \tilde{\sigma}_{L+2} < \tilde{J}_{L+1}$.

The sequence of estimations is terminated when the value of the objective function can be explained by random measurement errors, i.e. the sequence is terminated when $J(P_Q) < (M-Q) + \sqrt{2(M-Q)}$ (see e.g. [14]). Termination is also effectuated whenever the criterion for introduction of new parameters advises against introducing more parameters.

APPLICATION OF ADAPTIVE MULTISCALE ESTIMATION

The numerical examples are performed on a square horizontal reservoir with constant thickness and no-flow outer boundaries (see [7]). For the simulation, the reservoir is divided into 16x16 grid blocks. Water is injected at a constant and equal rate from wells located in grid blocks (2,2),(2,15) and (15,2). Oil and water are produced at constant pressure from a well located in grid block (15,15). The reservoir is initially fully saturated with oil. Table 1 shows the properties of the fluids and the reservoir.

Table 1: Fixed reservoir and fluid properties.

Reservoir dimensions	150m×150m×10m
Water viscosity	$\mu_w = 1.0 \cdot 10^{-3}$Pa·s
Oil viscosity	$\mu_o = 1.3 \cdot 10^{-3}$Pa·s
Water rel. perm.	$k_{rw} = s_w^2$
Oil rel. perm.	$k_{ro} = s_o^2$
Capillary pressure	$P_c(s_w) \equiv 0$kPa

Data for the history matching are generated by adding uncorrelated Gaussian noise to a set of pressure and rate data calculated using a selected permeability profile (the true permeability). The noise level in the data, represented by the standard deviation, is 0.25 per cent of the total production rate for the oil and water rate data, and between 0.3 and 1.7 per cent of the pressure difference between the injection and production wells for the pressure data.

For the estimations the time series of data contain 400 observations of pressure in each injection well, and 400 observations of oil and water rates in the production well when rate data are used. The time of observations ranges from the start of the injection and until 0.625 pore volumes are injected.

Estimating with pressure data

In [7] it was found that adaptive multiscale estimation of coarse-scale permeability structures was successful when utilizing pressure data as the only information source (P-estimation). This is illustrated for the true permeability k_{t1} given in figure 3.

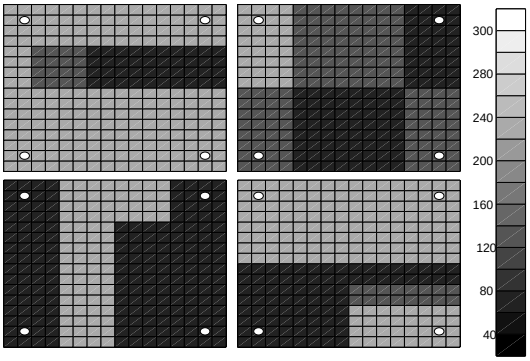


Figure 3: The true permeabilities going from left to right: k_{t1}, k_{t2}, k_{t3} and k_{t4}. The gray scale represents the permeability in mD (the same scale will be used in the rest of the paper.). White dots indicate well positions.

The estimation sequence for k_{t1} is given in figure 4. The permeability is not completely recovered, but the coarse-scale variations are identified.

Examples of other true permeabilities (k_{t2}, k_{t3} and k_{t4}) successfully estimated with P-estimation are given in figure 3. k_{t2} and k_{t3} where completely recovered with P-estimation, while coarse-scale variations in k_{t4} was recovered [7].

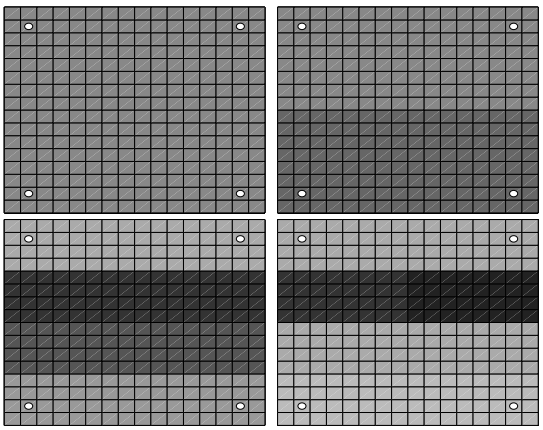


Figure 4: The P-estimation sequence for k_{t1}. In this example the entire estimation sequence is shown, for other examples only the final estimates will be shown.

Estimating with pressure and fluid rate data

In this section pressure data and fluid rate data are both utilized in order to do the estimation (PR-estimation). There are potentially different information content in the different data types. The information quality may vary from data type to data type, but different data types may also contain information about different features of the reservoir. More data points is expected to produce better estimates.

PR-estimation of k_{t1} produces the final estimate given in figure 5. The expectation of

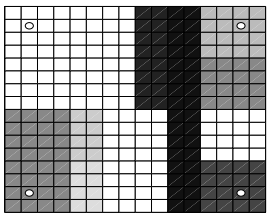


Figure 5: The PR-estimate.

a better estimate is not met. On the contrary, the result is clearly way off, and hence much worse than the estimate obtained by P-estimation.

PR-estimation of k_{t2} through k_{t4} produces results which are fairly good, but in no cases better than results obtained with P-estimation.

The next two sections discuss potential reasons for the lack of improvement in the estimates when performing PR-estimation, and alternative ways of performing PR-estimation.

DISCUSSION OF OBSERVATIONS

We will discuss the observations with respect to model nonlinearity and with respect to data weighting.

Model nonlinearity

The parameterization selection depends on a linearization of the model function so it is a prerequisite that the model function does not have large nonlinearity.

If the nonlinearity of $\boldsymbol{m}(\boldsymbol{c})$ with PR-estimation is systematically higher than the nonlinearity with P-estimation it might explain the lack of improvement.

Curvature measures of nonlinearity,

$$\Gamma^{\boldsymbol{h}} = \|\boldsymbol{m}_{hh}\|/\|\boldsymbol{m}_h\|^2, \tag{9}$$

where $\boldsymbol{m}$ is the model function and subscripts h and hh denotes the first and second order directional derivatives in parameter space, can be used to quantify model nonlinearity (for a more thorough discussion of curvature measures see [15]).

A relative curvature is defined as $\gamma^{\boldsymbol{h}} = \Gamma^{\boldsymbol{h}}\rho$, where ρ is a normalizing factor such that $\gamma^{\boldsymbol{h}} \ll 1$ corresponds to a (close to) linear model.

The most suitable choice of ρ may vary from problem type to problem type. Bates and Watts [15] investigated a different problem type than ours, so their suggestion of normalizing factor can not be used here [16]. We are using $\rho = \|A\Delta\tilde{\boldsymbol{c}}\|^{-1}$, where A is the sensitivity matrix and $\Delta\tilde{\boldsymbol{c}}$ the step corresponding to the "linearized minimum" $\tilde{J}$ [16].

It is also a question whether curvatures in all directions should be taken into account or only curvatures for directions close to $\Delta\tilde{\boldsymbol{c}}$.

The following different curvature measures have been tested:

$$g = \max_{\substack{\boldsymbol{h} \\ \|\boldsymbol{h}\|=1}} \gamma^{\boldsymbol{h}} \tag{10}$$

and

$$g_\Omega = \max_{\substack{\boldsymbol{h}\in\Omega \\ \|\boldsymbol{h}\|=1}} \gamma^{\boldsymbol{h}} \qquad (11)$$

where $\Omega \subset \Re^Q$, such that $\forall \boldsymbol{b} \in \Omega$, $\frac{\Delta\bar{\boldsymbol{c}}^T\boldsymbol{b}}{\|\Delta\bar{\boldsymbol{c}}\|\,\|\boldsymbol{b}\|} \geq f$. (In this paper $f = 0.86$.)

Both these curvature measures show similar trends when applied to the data. Only results with g_Ω are therefore given.

For k_{t1} the range of g_Ω is from 1 to 11 when performing P-estimation. There is an increase in the nonlinearity when the number of parameters increase. Trying to estimate this case by PR-estimation gives a range of curvatures from 3 to 33. The numbers can not be compared directly since they are obtained from different parameter spaces.

For k_{t2} the ranges are 1 to 23 for the P-estimation and 1 to 17 for the PR-estimation. The other cases give similar results for the curvature.

PR-estimation and P-estimation both start with the same parameterization (a single parameter). At some point in the estimation sequence PR-estimation results in a different choice of parameterization than P-estimation. The nonlinearity at this particular point is potentially of special interest, since one may argue that PR-estimation selects a non optimal parameterization exactly there. However, there is not much difference in the curvature at such points. The small difference that exists, however, seems always to be favorable to the P-estimation case. It is difficult to decide whether this is a coincidence or a small effect of nonlinearity problems with PR-estimation at these choices of parameterizations.

Based on the curvature measure values obtained, it can not be concluded with certainty that increased nonlinearity is the reason for the lack of improvement of PR-estimations compared to P-estimations.

Weighting of the different data types

The information content in the different data types may, as mentioned earlier, vary. Weighting the data according to the measurement errors (as in eq. 5) might therefore not be the best choice.

Figure 6 shows an example where production data with different measurement error magnitudes have been used. The only difference between the sets of data are the noise level added to the production data. The result is that for estimation sequences performed on these cases, different relative weighting of the data types is applied since the data are weighted according to the measurement errors. This example shows that the relative weighting is important for the final result. This is also observed in other examples.

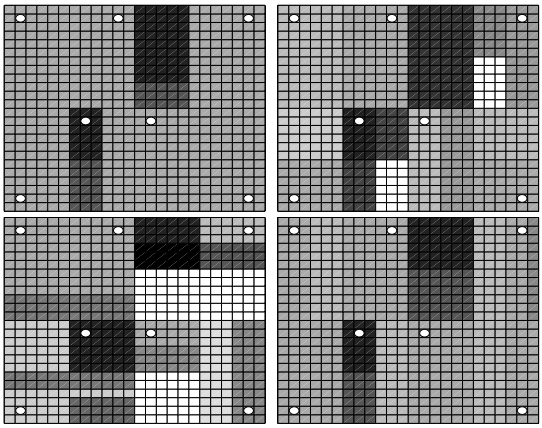


Figure 6: Example showing the importance of the relative weights. Going from left to right: The true permeability, the estimate with high relative weight on pressure, the estimate with high relative weight on fluid rates, and the estimate with the best weighting found (in this case approximately equal weight on pressure and rates) are shown. Wells in grid blocks (2,2), (23,2) and (11,23) are injection wells while the others are production wells.

It is seen that the measurement errors on each data type are not always an indication of the optimal way of weighting the data in the estimation. The problems observed with PR-estimation are potentially explained by the weight on each data type not being optimal.

POTENTIAL REMEDIES

Optimal weighting

Figure 6 showed an example where the relative weighting had been varied by varying the noise added to the real production data. In real life the measurement errors and noise exist in the data, and the assignment of relative weight has

to be done subsequently to the acquisition of the data. Finding the optimal weighting, subsequently to the acquisition of the data, might be a potential remedy.

Utilizing the best relative weight of the different data insures the result to be at least as good as the estimate from either data type alone (in many cases a lot better). There is a catch though: The best weighting has to be found from case to case. It is case dependent and not only dependent on the types of data used.

Sequential integration of pressure and rate data

Instead of introducing the different data types in the estimation sequence simultaneously, they can be introduced sequentially. The fluid rate data are introduced after the P-estimation sequence has ended. A PR-estimation sequence is started at the point where the P-estimation ended to see if pressure and rate data contain additional information that can be utilized.

Potentially, this method will perform at least as good as the P-estimation. If the fluid rate data contain information about variation on finer scale it might improve the estimate.

The results for k_{t1} from such an estimation are given in figure 7. In this sequential PR-

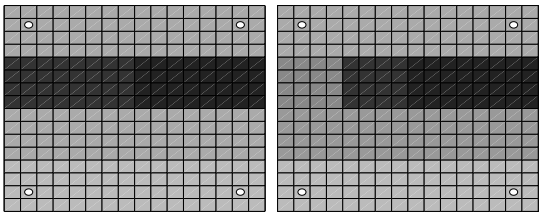


Figure 7: The final estimate with sequential PR-estimation is shown to the right. The result from P-estimation is shown to the left for comparison (see also figure 4).

estimation the final estimate is an improvement compared to the P-estimation, since it has a parameterization that is closer to resolving the true permeability completely. It is, however, to early in the development of sequential PR-estimation to say if this method always is preferable to P-estimation.

SUMMARY

Adaptive multiscale estimation for solving the inverse problem of estimating the permeability from production data has been considered. Earlier [7], the method has proved successful in utilizing pressure data in estimating coarse-scale permeability variations.

Applying the method introducing pressure and fluid rates simultaneously in the estimation sequence did not give the expected improvement in the estimates.

Possible reasons for the lack of improvement have been discussed. It can not be concluded that the model nonlinearity of the combined pressure and fluid rate data is to blame for the lack of improvement. However, the weighting of the different data types may be important for the estimation result.

Potential remedies are suggested in order to apply pressure and fluid rates in a manner giving an improved estimate. The obvious choice of finding the optimal weighting of the data types have some potential, but it is difficult to apply in real life since the optimal weights are case dependent. Sequential introduction of the data types has been tested and found to be a potential practical method for obtaining improved estimate from pressure and fluid rate data.

ACKNOWLEDGMENTS

The research presented in this paper has been financially supported by ENI-Agip SpA, Norsk Hydro ASA, and The Research Council of Norway.

REFERENCES

[1] K. Brusdal and T. Mannseth: *"Effects of basis-element norm rescaling on the convergence speed for a nonlinear parameter estimation problem."*, SIAM J. Sci. Comput., 21(6):2114–2125, 2000.

[2] G. Chavent and J. Liu: *"Multiscale parameterization for the estimation of a diffusion coefficient in elliptic and parabolic problems."*, Fifth IFAC Symposium on Control of Distributed Parameter Systems, Perpignan, France, June 1989.

[3] J. Liu: *"A multiresolution method for distributed parameter estimation."*, SIAM J. Sci. Comput., 14(2):398–405, March 1993.

[4] T. Mannseth: *"Prediction uncertainty assessment through scale splitting."*, In Proc. EAGE/SPE 3rd International Symposium on Petroleum Geostatistics, Toulose, France, April 1999. (extended abstract).

[5] T. Mannseth, K. Brusdal, A.-A. Grimstad, J.-E. Nordtvedt, and G. Nævdal: *"Nonlinearity, scale, and sensitivity for parameter estimation problems: Some implications for estimation algorithms."*, In *Proc. 3rd International Conference on Inverse Problems in Engineering: Theory and Practice*, Port Ludlow, WA, USA, June 1999.

[6] S. Yoon, A. H. Malallah, A. Datta-Gupta, D. W. Vasco, and R. A. Behrens: *"A multiscale approach to production data integration using streamline models."*, In 1999 SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, October 1999. SPE 56653.

[7] A-A. Grimstad, T. Mannseth, G. Nævdal, H. Urkedal: *"Adaptive multiscale permeability estimation"*, Accepted for publication in Computational Geosiences.

[8] A.-A. Grimstad, T. Mannseth, G. Nævdal, and H. Urkedal: *"Scale splitting approach to reservoir characterization."*, In Paper SPE66394 presented at the SPE Reservoir Simulation Symposium, Houston, Texas, February 2001.

[9] A.-A. Grimstad, T. Mannseth, J.-E. Nordtvedt, and G. Nævdal: *"Reservoir characterization through scale splitting."*, In 7th European Conference on the Mathematics of Oil Recovery. European Association of Geoscientists & Engineers, September 2000.

[10] A.-A. Grimstad, T. Mannseth, J.-E. Nordtvedt, and G. Nævdal: *"Scale splitting can reduce cost and complexity of reservoir characterization."*, In Proc. 62nd EAGE Conference and Exhibition, Glasgow, 2000. (extended abstract).

[11] H. Ben Ameur, G. Chavent, J. Jaffré: *"Refinement and coarsening of parameterization for the estimation of hydraulic transmissivity."*, In Proc. 3rd International Conference on Inverse Problems in Engineering: Theory and Practice, Port Ludlow, WA, USA, June 1999.

[12] K. Aziz and A. Settari: *"Petroleum Reservoir Simulation"*, Applied Science, London, 1979.

[13] J.J. Moré: *"The Levenberg-Marquardt algorithm: Implementation and theory."*, In G. A. Watson, editor, Numerical Analysis, Lecture Notes in Mathematics, vol. 630, pages 105-116. Springer Verlag, Berlin, 1977.

[14] A. Sen and M. Srivastava: *"Regression Analysis; Theory, Methods, and Applications"*. Springer-Verlag, 1990.

[15] D.M. Bates, D.G. Watts: *"Relative curvature measures of nonlinearity (with discussion)"*, J. R. Statist. Soc. Ser. B, 42(1):1:25, 1980.

[16] A-A. Grimstad: *"A more correct scaling of curvature measures"*, Unpublished note.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ESTIMATION OF THERMOPHYSICAL PROPERTIES OF MOIST MATERIALS UNDER DIFFERENT DRYING CONDITIONS

Gligor H. Kanevce
Macedonian Academy of Sciences and Arts
Skopje, Macedonia
kanevce@osi.net.mk

Ljubica P. Kanevce
Faculty of Technical Sciences
St. Kliment Ohridski University
Bitola, Macedonia
kanevce@osi.net.mk

George S. Dulikravich
Department of Mechanical and Aerospace Engineering
The University of Texas at Arlington
Arlington, Texas 76019, U.S.A
gsd@mae.uta.edu

Helcio R. B. Orlande
Department of Mechanical Engineering, EE/COPPE
Federal University of Rio de Janeiro, UFRJ
Rio de Janeiro, RJ, Brazil
helcio@serv.com.ufrj.br

ABSTRACT

This paper deals with the solution for the inverse problem of simultaneously estimating moisture content and temperature-dependent moisture diffusivity, together with thermal conductivity, heat capacity, density and phase conversion factor of a drying body, as well as boundary condition coefficients, by using only temperature measurements. Two different physical problems, convective and contact drying, are examined and compared.

The present parameter estimation problem is solved with the Levenberg-Marquardt method of minimization of the least-squares norm, by using simulated experimental data. The temperature responses during drying are obtained with a numerical solution of the non-linear one-dimensional Luikov's equations. As a representative drying body, a mixture of bentonite and quartz sand with known thermophysical properties has been chosen. Analyses of the sensitivity coefficients and of the determinant of the information matrix are presented.

NOMENCLATURE

a = water activity
c = heat capacity, *J/K kg*
C = concentration of water vapor, kg/m^3
D = moisture diffusivity, m^2/s
h = heat transfer coefficient, W/m^2K
h_D = mass transfer coefficient, *m/s*
ΔH = latent heat of vaporization, *J/kg*
$\mathbf{I}$ = identity matrix
j_m = boundary mass flux, $kg/m^2\ s$
j_q = boundary heat flux, W/m^2
$\mathbf{J}$ = sensitivity matrix
k = thermal conductivity, *W/mK*
L = flat plate thickness, *m*
p_s = saturation pressure , *Pa*
$\mathbf{P}$ = vector of unknown parameters
q = applied heat flux, W/m^2
t = time, *s*
T = temperature, ^0C
$\mathbf{T}$ = vector of estimated temperatures, ^0C
V = velocity, *m/s*
x = spatial coordinate, *m*
X = moisture content (dry basis), *kg/kg db*
$\mathbf{Y}$ = vector of measured temperatures, ^0C
d = thermo-gradient coefficient, *1/K*
ε = phase conversion factor
σ = standard deviation
μ = damping parameter
ρ = density, kg/m^3
φ = relative humidity

Subscripts

a = drying air
s = dry solid

INTRODUCTION

Drying of hygroscopic capillary-porous bodies is a complex process of simultaneous heat and moisture transport within the material and from its surface to the surroundings, caused by a number of mechanisms. There are several

different methods of mathematically modeling the drying process. In the approach proposed by Luikov [1], the drying body moisture content and temperature field are expressed by a system of two coupled partial differential equations. The system of equations incorporates coefficients that must be determined experimentally. The main problem is the determination of the moisture diffusivity connected with the difficulty of moisture content measurements. Local moisture content measurements are practically unfeasible especially for small drying objects. Standard drying curves measurements (body mean moisture content during drying) are complex and have low accuracy.

Dantas et al [2, 3, 4] and Kanevce, Kanevce and Dulikravich [5, 6, 7, 8] recently analyzed the application of inverse analysis approaches to the estimation of thermophysical properties of drying bodies. The main idea of the applied methods was to take advantage of the relation between the heat and mass (moisture) transport processes within the drying body and from its surface to the surrounding media. Then, the estimation of the thermophysical properties of the drying body could be performed on the basis of accurate and easy-to-perform thermocouple temperature measurements, by using an inverse analysis approach. Kanevce, Kanevce and Dulikravich [5, 6,7,8] analyzed this idea by using the temperature response of a body exposed to convective drying, while Dantas et al [2, 3, 4] examined contact drying experiments.

The objective of this paper is to compare these two kinds of experiments for the estimation of the thermophysical properties of a drying body. In order to perform this analysis, the sensitivity coefficients and the determinant of the information matrix were calculated.

A MATHEMATICAL MODEL OF DRYING

Two different physical problems, convective and contact drying, are analyzed here. In the convective drying experiment (Fig. 1) the boundaries of the drying body are in contact with the drying air, thus resulting in a convective boundary condition for both temperature and moisture content. In the contact drying experiment (Fig. 2), one of the boundaries of the one-dimensional body is in contact with a heater. That boundary is impervious to moisture transfer. The other boundary is in contact with the dry air, thus resulting in a convective boundary condition.

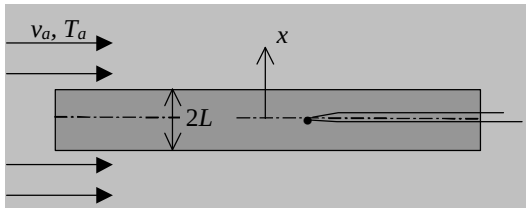


Fig. 1. Scheme of the convective drying experiment

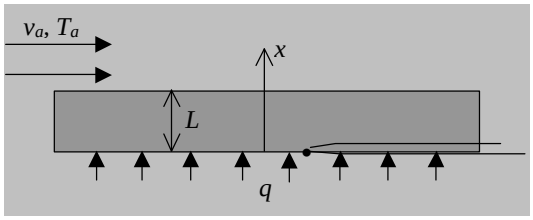


Fig. 2. Scheme of the contact drying experiment

An infinite flat plate of the capillary porous material with negligible shrinkage has been considered in both kinds of experiments examined here.

The system of equations for energy balance and moisture transport can be expressed [1] as

$$c\rho_s \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(k \frac{\partial T}{\partial x} \right) + \varepsilon \rho_s \Delta H \frac{\partial X}{\partial t} \qquad (1)$$

$$\frac{\partial X}{\partial t} = \frac{\partial}{\partial x}\left(D \frac{\partial X}{\partial x} + D\delta \frac{\partial T}{\partial x} \right) \qquad (2)$$

where $T(x,t)$ and $X(x,t)$ are the unsteady temperature and moisture content fields, respectively. From previous experimental and numerical examinations of the transient moisture and temperature fields [9] it was concluded that, for practical calculations, the influence of the thermo-gradient coefficient, δ, is small and can be ignored. It was also concluded that this system of coupled partial differential equations can be used by treating the transport coefficients as constants, except for the moisture diffusivity, D. Consequently, the resulting system of equations for temperature and moisture content prediction becomes

$$\frac{\partial T}{\partial t} = \frac{k}{c\rho_s}\frac{\partial^2 T}{\partial x^2} + \frac{\varepsilon \Delta H}{c}\frac{\partial X}{\partial t} \tag{3}$$

$$\frac{\partial X}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial X}{\partial x} \right) \tag{4}$$

As initial conditions, uniform temperature and moisture content profiles are assumed, that is,

$$T(x,0) = T_0, \; X(x,0) = X_0 \quad , \; for\; t = 0 \tag{5}$$

The boundary conditions on the body surface exposed to convection, at $x = L$, are

$$\begin{aligned} -k\left(\frac{\partial T}{\partial x}\right)_{x=L} + j_q - \Delta H(1-\varepsilon) j_m = 0 \\ D\rho_s\left(\frac{\partial X}{\partial x}\right)_{x=L} + D\delta\rho_s\left(\frac{\partial T}{\partial x}\right)_{x=L} + j_m = 0 \end{aligned} \tag{6}$$

where the convective heat flux, $j_q(t)$, and mass flux, $j_m(t)$, are given respectively by

$$\begin{aligned} j_q &= h\left(T_a - T_{x=L}\right) \\ j_m &= h_D\left(C_{x=L} - C_a\right) \end{aligned} \tag{7}$$

The water vapor concentration in the drying air, C_a, is calculated by

$$C_a = j \; p_s(T_a)/[461.9(T_a + 273)] \tag{8}$$

The water vapor concentration of the air in equilibrium with the surface of the body exposed to convection is calculated by

$$C_{x=L} = \frac{a(T_{x=L}, X_{x=L})\, p_s(T_{x=L})}{461.9(T_{x=L} + 273)} \tag{9}$$

The water activity, a, or the equilibrium relative humidity of the air in contact with the convection surface at temperature $T_{x=L}$ and moisture content $X_{x=L}$, is calculated from experimental water sorption isotherms.

In the case of the convective drying experiment the problem is symmetrical and boundary conditions on the mid-plane of the plate ($x = 0$) are given by:

$$\left(\frac{\partial T}{\partial x}\right)_{x=0} = 0, \quad \left(\frac{\partial X}{\partial x}\right)_{x=0} = 0 \tag{10.a}$$

In the case of the contact drying experiment the boundary conditions at the surface $x = 0$, in contact with the heater that provides the heat flux q, are:

$$-k\left(\frac{\partial T}{\partial x}\right)_{x=0} = q, \quad \left(\frac{\partial X}{\partial x}\right)_{x=0} = 0 \tag{10.b}$$

ESTIMATION OF PARAMETERS

The estimation methodology used is based on minimization of the ordinary least square norm

$$E(\mathbf{P}) = [\mathbf{Y} - \mathbf{T}(\mathbf{P})]^T [\mathbf{Y} - \mathbf{T}(\mathbf{P})] \tag{11}$$

Here, $\mathbf{Y}^T = [Y_1, Y_2, \ldots, Y_{imax}]$ is the vector of measured temperatures, $\mathbf{T}^T(\mathbf{P}) = [T_1(\mathbf{P}), T_2(\mathbf{P}), \ldots T_{imax}(\mathbf{P})]$ is the vector of estimated temperatures at time t_i ($i = 1, 2, \ldots, imax$), $\mathbf{P}^T = [P_1, P_2, \ldots P_N]$ is the vector of unknown parameters, $imax$ is the total number of measurements, and N is the total number of unknown parameters ($imax \geq N$).

A version of Levenberg-Marquardt's method was applied for the solution of the present parameter estimation problem [10]. This method is quite stable, powerful, straightforward and has been applied to a variety of inverse problems. It belongs to a general class of damped least square methods [11]. The solution for the vector $\mathbf{P}$ is achieved with the following iterative procedure

$$\mathbf{P}^{r+1} = \mathbf{P}^r + [(\mathbf{J}^r)^T \mathbf{J}^r + m^r \mathbf{I}]^{-1} (\mathbf{J}^r)^T [\mathbf{Y} - \mathbf{T}(\mathbf{P}^r)] \tag{12}$$

where the superscript r denotes the number of iterations and the sensitivity matrix is given by:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial T_1}{\partial P_1} & \cdots & \frac{\partial T_1}{\partial P_N} \\ \cdot & & \\ \frac{\partial T_{imax}}{\partial P_1} & \cdots & \frac{\partial T_{imax}}{\partial P_N} \end{bmatrix} \tag{13}$$

Near the initial guess, the problem is generally ill conditioned so that a large damping parameter is chosen, thus making the term $\mu\mathbf{I}$ large as compared to the term $\mathbf{J}^T\mathbf{J}$. The term $\mu\mathbf{I}$ damps instabilities due to ill-conditioned character of the problem. Hence, the matrix $\mathbf{J}^T\mathbf{J}$ is not required to be non-singular at the beginning of iterations and the procedure tends towards the slow-convergent

steepest descent method. As the iteration process approaches the converged solution, the damping parameter decreases and the Levenberg-Marquardt method tends towards Gauss' method. The iterative procedure of the Levenberg-Marquardt method is stopped when the ordinary least squares norm, $E(\mathbf{P})$ is sufficiently small, or if the norm of the gradient of $E(\mathbf{P})$ is small enough, or if the changes in the vector of parameters are very small [13].

RESULTS AND DISCUSSION

For the direct problem solution, the system of equations (3) and (4) with initial conditions given by equations (5) and boundary conditions given by equations (6) and (10a) (or 10.b), has been solved numerically for a model material [5], involving a mixture of bentonite and quartz sand, with the following experimentally determined thermophysical properties [9]: $\rho_s = 1738$ kg/m^3, $\Delta H = 2.31\text{x}\,10^6$ J/kg, $c = 1550$ J/K kg db, $k = 2.06$ W/m K and $\varepsilon = 0.5$.

The experimentally obtained desorption isotherms of the model material are given by the following empirical equation [9]:

$$a = 1 - \exp\{-1.5\text{x}10^6 (T+273)^{-0.91} X^{[-0.005(T+273)+3.91]}\} \qquad (14)$$

where the water activity, a, represents the relative humidity of the air in equilibrium with the drying object at temperature, T, and moisture content, X.

The following empirical expression can describe the experimentally obtained relationship for the moisture diffusivity of this material

$$D = D_X X^{-2} \left(\frac{T+273}{303} \right)^{D_T} \qquad (15)$$

where $D_X = 9.0\text{x}10^{-12}$ and $D_T = 10$.

In the inverse problem investigated here, the values of D_X, D_T, r_s, c, k, e, h and h_D are regarded as unknown for the convective drying experiment. For the contact drying experiment there is one additional unknown parameter, the applied heat flux, q. All other quantities appearing in the direct problem formulation were assumed to be exactly known.

For the estimation of these unknown parameters, the transient readings of a single temperature sensor, located at the position $x = 0$, were considered available for the inverse analysis. Simulated experimental data were used in this work. Such data were obtained from the numerical solution of the direct problem presented above, by treating the values and expressions for the material properties as known. In order to simulate real measurements, normally distributed errors, with zero mean and constant standard deviation σ, were added to the numerical temperature responses.

Convective drying experiment

The vector of unknown parameters in the case of the convective drying experiment is

$$\mathbf{P}^{\mathrm{T}} = [D_x, D_T, \rho_s, c, k, \varepsilon, h, h_D] \qquad (16)$$

The possibility of simultaneously estimating the moisture content and temperature dependent moisture diffusivity, together with other thermophysical properties of the model material, as well as the heat and mass transfer coefficients in the convective drying experiment, by using only temperature measurements, was already investigated in [8]. Here, we will outline the main conclusions and results.

Following the conclusions of previous works [6, 7] the selected drying air bulk temperature, speed, and relative humidity were taken as $T_a = 80\ ^{\mathrm{O}}\mathrm{C}$, $V_a = 10$ m/s and $\varphi = 0.12$, respectively.

The analysis of the sensitivity coefficient has been carried out for a plate of thickness $2L = 6$ mm, with initial moisture content of $X(x,0) = 0.20$ kg/kg db and initial temperature $T(x,0) = 20^{\mathrm{O}}\mathrm{C}$. Figure 3 shows the relative sensitivity coefficients $P_j \partial T/\partial P_j$, with respect to all unknown parameters, $j = 1, 2, \ldots, 8$.

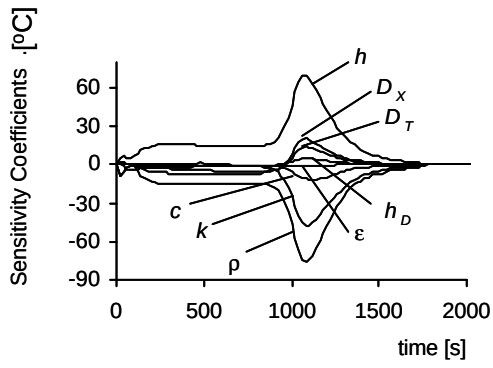


Fig.3. Relative sensitivity coefficients for the convective drying experiment

The temperature sensitivity coefficient with respect to the phase conversion factor, ε, is very

small. This indicates that ε cannot be estimated in this case. The relative sensitivity coefficients with respect to the dry material density, r_s, and the convection heat transfer coefficient, h, are linearly-dependent. This makes it impossible to simultaneously estimate r_s and h. Due to these reasons and to the fact that the density of the dry material can be relatively easily determined by a separate experiment, the density of the dry material was assumed as known for the inverse analysis.

The relative sensitivity coefficient with respect to the thermal conductivity, k, is very small, except for the moment when the body moisture content is nearly equal to its equilibrium value. This is also a moment when a small evaporation rate and a fast body temperature increase occur. Temperature measurements of a single thermocouple do not make it possible to estimate the thermal conductivity, if the initial guess is higher than the exact value of the parameter. For the cases when the initial guess for thermal conductivity is smaller than the exact value, the estimation of the thermal conductivity by a single thermocouple temperature response of a thin drying plate is possible.

An analysis of the determinant of the information matrix $\mathbf{J}^T\mathbf{J}$ with normalized elements confirms the previous conclusions. Figure 4 presents transient variations of the determinant of the information matrix if five, (D_X, D_T, c, h, h_D), six, (D_X, D_T, c, h, h_D, r_s), seven, (D_X, D_T, c, h, h_D, r_s, k), and eight, (D_X, D_T, c, h, h_D, r_s, k, e) parameters are simultaneously considered as unknowns.

Based on the foregoing analyses of the sensitivity coefficients and of the determinant of the information matrix, we now consider as unknown parameters for the inverse problem the moisture diffusivity parameters, D_X and D_T, the specific heat, c, the convection heat transfer coefficient, h, and the convection mass transfer coefficient, h_D. For the solution of such a parameter estimation problem with the Levenberg-Marquardt method, we use simulated measurements of a single thermocouple, with different levels of random errors, including $\sigma = 0$ (errorless measurements), 0.2 and 0.5 ^{O}C. Table 1 shows the parameters estimated for these different levels of random error. For comparison, the exact values for the parameters are also shown in this table. The obtained results show good agreement between the estimated and exact values for the parameters. For measurements with a standard deviation of 0.5 ^{o}C, the maximum relative error between estimated and exact values is of 4.4% for h_D, but for the other parameters such an error is smaller than 1%.

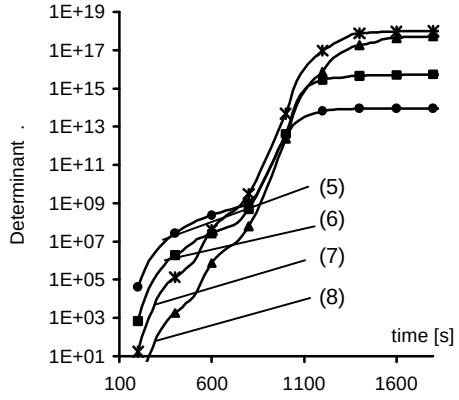


Fig.4. Determinant of the Information Matrix

Contact drying experiment

The vector of unknown parameters in the case of the contact drying experiment is

$$\mathbf{P}^T = [D_x, D_T, \rho_s, c, k, \varepsilon, h, h_D, q] \qquad (17)$$

The analysis of the sensitivity coefficients has been carried out for an infinite flat plate with initial moisture content of $X(x,0) = 0.20$ kg/kg and initial temperature $T(x,0) = 20.0$ ^{O}C. The possibility of simultaneously estimating the moisture content and temperature-dependent moisture diffusivity together with other thermophysical properties of the model material, as well as the heat and mass transfer coefficients and the applied heat flux, has been investigated for a variety of boundary conditions and dimensions of the drying body.

The drying air bulk temperature, T_a, was varied between 20 and 80 ^{O}C, the drying air velocity, V_a, between 3 and 10 m/s, the applied heat flux, q, between 1000 and 5000 W/m^2 and the plate thickness L, between 3 and 6 mm. The relative humidity of the drying air was $\varphi = 0.12$. The best combination of the relative temperature sensitivity coefficients with respect to all unknown parameters, was obtained with T_a = 20 ^{O}C, V_a = 10 m/s, q = 3000 W/m^2 and L = 3 mm.

Table 1. Estimated parameters in the convective drying experiment

Parameters	Exact	Estimated			Relative errors
		σ = 0	σ = 0.2 °C	σ = 0.5 °C	for σ = 0.5 [%]
D_X x 10^{12} [m^2/s]	9.00	8.99	9.04	9.06	0.7
D_T	10.0	10.0	9.999	10.1	1.0
c [J/K kg]	1550	1551	1550	1551	0.1
h [W/m^2 K]	83.1	83.1	83.2	83.3	0.2
h_D x 10^2 [m/s]	9.29	9.29	9.12	8.88	4.4

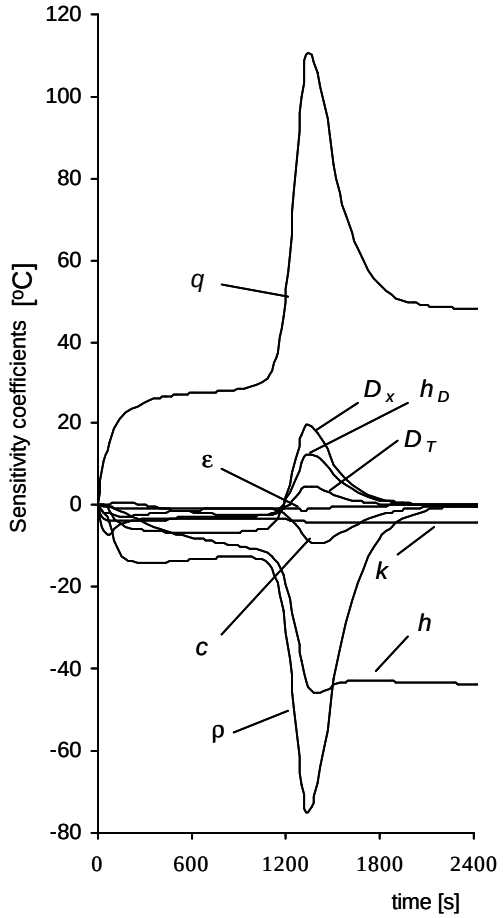


Fig.5. Relative sensitivity coefficients for the contact drying experiment

Figure 5 shows the relative sensitivity coefficients $P_j \P T/\P P_j$, with respect to all unknown parameters. It can be seen that the relative sensitivity coefficients with respect to the applied heat flux, the dry material density and the convection heat transfer coefficient are much larger than the other sensitivity coefficients. Due to the same reasons underlined in the case of convective drying experiment, the phase conversion factor and the dry material density were taken as known quantities for the cases examined below.

Figure 6 presents the transient variation of the determinant of the information matrix if nine, (D_X, D_T, c, r_s, k, e, h, h_D, q), seven, (D_X, D_T, c, k, h, h_D, q), six, (D_X, D_T, c, h, h_D, q) and five (D_X, D_T, h, h_D, q) parameters are simultaneously considered as unknown. Elements of the information matrix were defined for a large, but fixed number of transient temperature measurements (501 in these cases) [13]. The maximum determinant of the information matrix corresponds to the drying time when equilibrium moisture content and temperature profiles have been reached, as can be seen in figures Fig. 7 and 8.

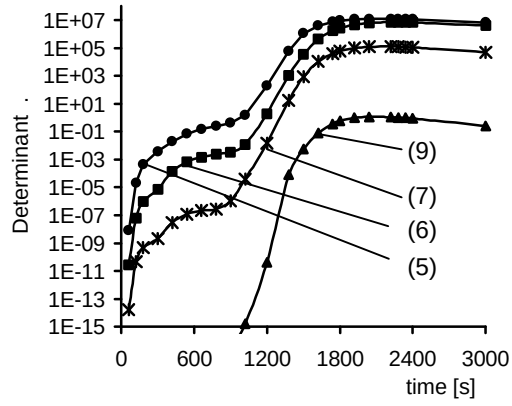


Fig.6. Determinant of the information matrix for the contact drying experiment

Table 2. Estimated parameters in the contact drying experiment

Parameters	Exact values	Initial guesses	Estimated values				Relative errors for σ = 0.5 [%]
			σ = 0	σ = 0.5	σ = 0.5	σ = 0.5	
D_X x10^{12} [m^2/s]	9.00	11.00	9.00	9.049	9.063	9.041	0.5
D_T	10.0	12.0	10.0	9.904	9.806	9.874	1.3
c [J/K kg]	1550	1300	1550	-	1533	1531	1.2
k [W/m K]	2.06	2.70	2.06	-	-	2.12	2.9
h [W/m^2 K]	68.7	80.0	68.70	68.697	68.64	68.59	0.2
h_D·x 10^2 [m/s]	6.94	8.00	6.94	6.90	6.89	6.82	1.7
q [W/m^2]	3000	3500	3000	3000	2997	3003	0.1

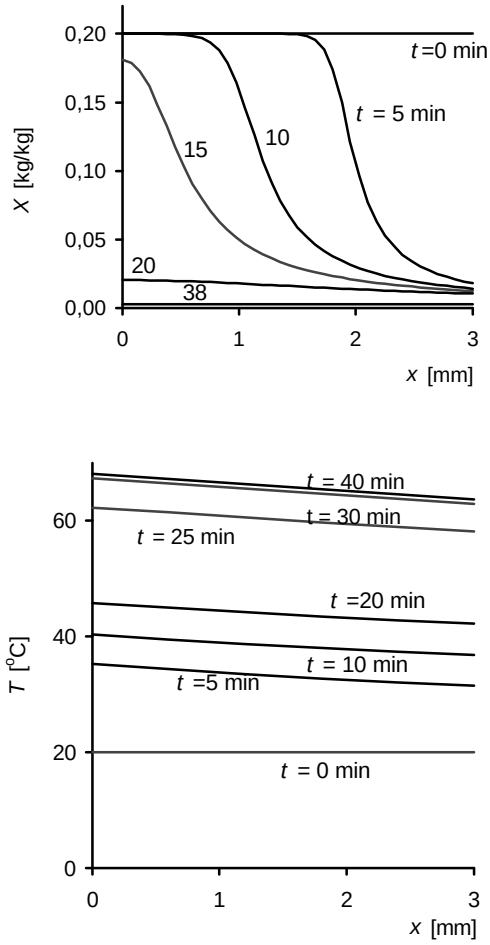


Fig.7. Transient moisture content and temperature profiles in the case of the contact drying experiment

Table 2 shows the estimated parameters for σ=0.5, for 5, 6 and 7 unknown parameters. For comparison, the values of exact parameters and the values estimated with errorless (σ=0) temperature data are shown in this table. Table 2 also shows the initial guesses used for the Levenberg-Marquardt method, as well as the relative errors for the case involving the estimation of 7 unknown parameters. Estimated values of similar accuracy have been obtained with other initial guesses. If the dry material density and the phase conversion factor are considered as known, the remaining seven (D_X, D_T, c, k, h, h_D, q), parameters can be simultaneously estimated with relative errors within 2.9%. The accuracy of computing the parameters in the case when six (D_X, D_T, c, h, h_D, q) parameters were simultaneously estimated was within two percent. In the case of simultaneous estimation of the moisture diffusivity and the boundary conditions parameters, (D_X, D_T, h, h_D, q), the relative errors of the computed parameters were within one percent.

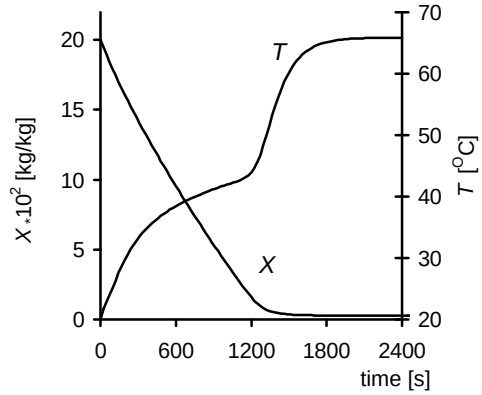


Fig.8. Volume-averaged moisture content and temperature changes during the contact drying

CONCLUSIONS

The use of two types of experiments, convective and contact drying, for the solution of the inverse problem of simultaneous estimation of thermophysical properties of a drying body together with the boundary condition parameters, by using only temperature measurements, has been analyzed in this paper.

Values of two moisture diffusivity parameters, the dry material density, the thermal conductivity, the specific heat, the phase conversion factor, the convection heat transfer coefficient and the mass transfer coefficient were regarded as unknown quantities in the convective drying experiment. In the contact drying experiment an additional unknown parameter, the applied heat flux, was taken into account.

In the convective drying experiment based on a single thermocouple transient response, it is possible to estimate simultaneously five of the eight unknown parameters: the two moisture diffusivity parameters, the specific heat, the convection heat transfer coefficient and the mass transfer coefficient.

In the contact drying experiment it is possible to estimate simultaneously seven of the nine unknown parameters: the two moisture diffusivity parameters, the specific heat, the thermal conductivity, the convection heat transfer coefficient, the mass transfer coefficient and the applied heat flux.

The application of the convective or contact drying experiment for the estimation of the thermophysical properties of the drying body primarily depends on the available experimental setup; but the use of the contact drying experiment allows for the estimation of the thermal conductivity together with the other parameters.

REFERENCES

1. A. V. Luikov, *Teplomassoobmen*, Energia, Moscow, Russia, 1972, p. 468.

2. L. B. Dantas, H. R. B. Orlande, R. M. Cotta, R. De Souza, and P. D. C. Lobo, Inverse analysis in moist capillary porous media, *15th Brazilian Congress of Mechanical Engineering, November 7-10, Aguas de Lindoia, Sao Paulo, Brazil* (1999)

3. L. B. Dantas, H. R. B. Orlande, R. M. Cotta and P. D. C. Lobo, Parameter estimation in moist capillary porous media by using temperature measurements, *Inverse Problems in Engineering Mechanics II*, eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam, 2000, p.53.

4. L. B. Dantas, H. R. B. Orlande and R. M. Cotta, Effects of lateral heat losses on the parameter estimation problem in moist capillary porous media, *Inverse Problems in Engineering Mechanics III*, eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam, 2002, p.13.

5. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Moisture diffusivity estimation by temperature response of a drying body, *Inverse Problems in Engineering Mechanics II*, eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam, 2000, p.43.

6. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Proceedings of NHTC'00, ASME paper NHTC2000-12296, 34th ASME National Heat Transfer Conference, August 20-22, Pittsburgh, PA, U.S.A. , 2000.

7. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Proceedings of the 12th International Drying Symposium, IDS'2000, August 28-31, Noordwijkerhout, The Netherlands, 2000.

8. G. H. Kanevce, L. P. Kanevce and G. S. Dulikravich, Simultaneous estimation of thermophysical properties and heat and mass transfer coefficients of a drying body, *Inverse Problems in Engineering Mechanics III*, eds: M. Tanaka, and G. S. Dulikravich, Elsevier, Amsterdam, 2002, p. 3.

9. G. H. Kanevce, Numerical study of drying, *IDS '98, Vol. A , pp. 256-263, Halkidiki, Greece*, 1998.

10. D. W. Marquardt, An algorithm for least squares estimation of nonlinear parameters, *J. Soc. Ind. Appl. Math*., 431 (1963).

11. J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*, John Wiley&Sons, Inc., New York, 1977.

12. R. C. Pfafl and B. J. Mitchel, A general method for simultaneous measurement of thermal properties, *AIAA Paper, No. 69-602*, 1969.

13. M. N. Ozizik and H. R. B. Orlande, *Inverse Heat Transfer: Fundamentals and Applications*, Taylor and Francis, New York, 2000.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

PARAMETER ESTIMATION OF THERMO-OPTICAL GLASS PROPERTIES FROM EXPERIMENTAL PHASE LAG SIGNALS OBTAINED THROUGH PERIODIC HEAT FLUX EXCITATION

Stéphane André [♠]
LEMTA-UMR CNRS 7563,
2 av de la forêt de Haye, BP160,
54504 Vandœuvre Lés Nancy Cedex , France,
sandre@ensem.inpl-nancy.fr

Jean-Jacques Serra
DGA/DCE/CTA/LOT –
Groupe des Hauts Flux thermiques
BP6, Odeillo, 66125 Font Romeu, France.
serra@imp-odeillo.fr

Norberto Cella
Instituto Politécnico, IPRJ,
Universidade do Estado do Rio de Janeiro,UERJ,
P.O. Box 97282, 28601-Nova Friburgo, RJ, Brazil
cella@iprj.uerj.br

Antônio J. Silva Neto
Instituto Politécnico, IPRJ,
Universidade do Estado do Rio de Janeiro,UERJ,
P.O. Box 97282, 28601-Nova Friburgo, RJ, Brazil
ajsneto@iprj.uerj.br

ABSTRACT

The experimental technique based on the established periodic regime of heat transfer is applied for characterization of semi-transparent materials (STM). The direct model for the time dependent conductive and radiative transfer coupled problem is analytical. The noisy experimental signal consists in phase lag evolutions versus frequency of excitation. They have been obtained through a system based on a modulated heat flux supplied by a laser that impinges on various glasses samples. The identification process is made by minimizing the cost functional of the squared residuals between measured and calculated phase lag, using either direct or stochastic algorithms. The results show that thermal diffusivity estimations match the values obtained with the flash method. Concerning possible estimation of the two additional parameters introduced by the radiative transfer, especially of an "average" extinction coefficient, one shows that even though such goal is accomplished in most cases, it is not feasible for the important case from an experimental point of view of small optical thickness, for which the two parameters are highly correlated. Neither deterministic nor stochastic algorithms are shown to yield trustable results in such ill-posed conditions.

NOMENCLATURE

e	thickness of the sample
f^*	dimensionless frequency = $f\, t_c$
h	heat exchange coefficient
k	thermal conductivity
n	gray index of refraction
N_{Pl}	Planck number = $k\beta/4n^2\sigma T_0^3$
t_c	characteristic time of conduction = e^2/α
T_0	reference temperature of the sample
α	thermal diffusivity
β	spectrally averaged (mean) extinction (absorption) coefficient / parameter vector
ε_i	emissivity of boundary i
φ	phase lag variable (rad)
ϕ	heat flux density (W/m²)
σ	Stefan-Boltzmann constant / standard deviation (Std Dev)
τ_0	optical thickness = βe
ω	pulsation of excitation = $2\pi f$
$\hat{}$ / $\bar{}$	upperscripts for estimated / averaged values

INTRODUCTION

Characterizing a glass in view of heat transfer modeling generally means first to measure the true thermal (or phononic) diffusivity and/or its conductivity in distinct thermal experiments, and secondly to measure the spectrally dependent

♠ *Present position :* *Visiting Professor of the CNPq, Instituto Politécnico, UERJ, Brazil, sandre@iprj.uerj.br*

absorption coefficient and refractive index in optical experiments. In this paper, we are looking for establishing whether or not a single thermal experiment could lead simultaneously to the correct thermal diffusivity as well as to some spectrally averaged or mean absorption coefficient. The interest would be to reduce properties data base of glasses only to temperature dependent mean absorption coefficient that may be used with sufficient accuracy in gray modeling of more complex processes. It may also be used for the qualification of ceramic thermal barriers at high temperatures by comparative measurements. One previous attempt to reach such an objective has been carried out by Schatz and Simmons [1], for optically thick samples and low frequency temperature waves.

It has been shown already [2] that the flash method is able to provide a correct measurement of the single phononic diffusivity. Therefore, the thermal diffusivities measured on pure silica and Zinc Selenide (ZnSe) with flash method will provide reference estimations. A theoretical stochastic analysis of the parameter estimation problem (PEP) [3], has shown that firstly, the modulated technique can lead to a good estimation of the phononic diffusivity and secondly, that retrieving mean absorption coefficient estimations is possible in most cases but very difficult in the important experimental case of small optical thickness. Real experimental measurements obtained at the solar oven of Odeillo (France) have been used for inverse recovering of the parameters of the model and confirm the results previously obtained.

DIRECT MODELS

When PEP is applied to real measurements, the experimentalist will most certainly benefit from developing more than one direct model. The objective reason for this is that generally, the physics of the mechanisms at hand and the hypothesis made on the problem may lead to asymptotic behavior and/or to various idealized conditions respectively. Therefore, the size of the unknown parameter vector can change, the combination of the various parameters can be different (e.g. when dimensioned or dimensionless models are considered) which may change the nature of the PEP by affecting the parameter sensitivities. Parameter or model reduction may then be readily applied as needed in the identification process. This kind of driving thought is applied here and three direct models are presented next.

- The first direct model (DM1) is obtained when considering an opaque material. This model can be useful to handle data obtained on very thin semi-transparent materials having high reflective opaque boundaries (e.g. 10 μm coating of gold). In this case, the radiative transfer is weakened to approach a zero heat flux and heat is transferred by conduction only. As a consequence, only the thermal diffusivity can obviously be estimated.
- The second model (DM2) is obtained when a pure scattering media is considered towards the radiative intensity exchange. In this case, the radiative transfer is uncoupled from the conductive one. It is also the model appropriate for a material layer of weak absorption when a direct radiative exchange may be preponderant between the two frontiers (radiative equilibrium).
- The third model (DM3) is for the more general case of coupled conductive and radiative heat transfer.

All three models are analytical, derived using the quadrupole method, and are briefly described here. For a more detailed description, see [4] for a general overview of the quadrupole method and [3] for the specific details concerning the problem under analysis in the present work. In what follows, the temperature variable θ corresponds to the variations of the absolute temperature $\tilde{T}$ with respect to the reference temperature T_0: $\theta = \tilde{T} - T_0$

❑ In the case of DM1, it can be readily derived that in a periodic steady-state regime, one can look for a solution involving the complex amplitude $\theta^*(z,\omega) = A(z,\omega)\exp(-j\varphi(z,\omega))$ of the complex temperature $T^* = \theta^* \exp(j\omega t)$, which derives from the solution of the transformed one dimensional time dependent heat equation

$$\frac{d^2\theta^*}{dz^2} = \frac{j\omega}{\alpha}\theta^* = \gamma^2\theta^* \qquad (1)$$

A and φ denotes respectively the magnitude (modulus) and phase lag of the complex temperature θ^*. Equation (1) can also be seen as the Laplace transform of the heat equation where the Laplace variable has been substituted by the pure imaginary number $j\omega$. When the knowledge

of entire temperature field is not necessary, one can use the quadrupole method [4] to obtain an intrinsic solution of (1), formulated in terms of a matrix transfer Q_C linking the **i**nput and **o**utput vectors, i.e. temperature and heat flux density.

$$\begin{bmatrix} \theta_i^* \\ \phi_i^* \end{bmatrix} = \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} \theta_o^* \\ \phi_o^* \end{bmatrix} \qquad (2)$$

where $A_c = D_c = \cosh(\gamma e)$, $B_c = (1/k\gamma)\sinh(\gamma e)$, $C_c = k\gamma \sinh(\gamma e)$. In the case where heat losses are considered on both sides of the sample, Ohm's law gives rise to $\theta^* = (1/h)\phi^*$ which leads, with the quadrupole method, to a simple matrix product (quadrupoles in cascade):

$$\begin{bmatrix} \theta_i^* \\ \phi_i^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/h & 1 \end{bmatrix} \begin{bmatrix} A_c & B_c \\ C_c & D_c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1/h & 1 \end{bmatrix} \begin{bmatrix} \theta_o^* \\ \phi_o^* \end{bmatrix} \qquad (3)$$

Finally, specifying the input flux as $\phi_i^* = \phi$, amplitude of the modulated excitating flux whose phase is equal to the reference 0, origin of any phase lag, and the output flux ϕ_o^* as being equal to 0 (no back surface heat source), the model DM1 for the temperature evolution of the rear face is

$$\theta_o^* = \frac{\phi}{C_{eq}(\omega)} \qquad (4)$$

with $C_{eq} = C_c + 2A_c/h + B_c/h^2$, being the coefficient C of the **eq**uivalent quadrupole resulting from the above matrix product.

❑ In the case of DM2, one has just to connect a radiative resistance R_r in parallel to the conductive quadrupole Q_C. The expression of this resistance (in dimensionless form) can be easily derived from any textbook on radiative transfer and looks like

$$R_r = \frac{k/e}{4n^2\sigma T_0^3}\left(\beta e + \frac{1}{\varepsilon_1} + \frac{1}{\varepsilon_2} - 1\right) \qquad (5)$$

This leads to a quadrupole matrix Q_2 whose coefficients are respectively $A_2 = D_2 = (B_c + R_r A_c) / (B_c + R_r)$, $B_2 = R_r B_c / (B_c + R_r)$, $C_2 = (2 (A_c - 1) + R_r C_c) / (B_c + R_r)$. In this case the radiative transfer adds only one additional parameter to the problem. In the case of perfectly reflecting walls, R_r tends to infinity which means that no heat flux can be exchanged by radiation and DM2 reduces to DM1.

❑ In the case of DM3, it has been shown [4] that a quadrupole matrix may be obtained as a transfer function of a STM layer where coupled conductive and radiative transfer takes place, when the two following conditions hold. First, one linearizes the heat transfer about the reference temperature T_0 of the experiment. This means that the emission terms of the blackbody function, in the 4th power of the temperature are linearized through $\theta^4 = 4T_0^3\theta$, so that the radiative flux only involves the 1st order of the temperature. This condition holds when a maximum of 50 K of temperature differences between two different locations inside the STM sample is considered. This is the case in our experiments. The second condition for which this analytical model can be derived is the approximation of the exponential integral functions of the radiative flux equation by pure exponentials of the form $A\exp(-Bx)$. This is named 'kernel substitution' technique and even though it has not been intensively used in engineering calculations it has been shown to be of good precision [4] (semi-transparent medium temperature field calculations within 2% of a numerical 'exact' solution).
A differential equation in the fourth power of the temperature can then be obtained instead of (1), by twice differentiating the heat equation and the quadrupole formulation follows. The expressions for the 4 coefficients of the quadrupole matrix cannot be given in a single formula but are easily calculated analytically. The resulting computational time for this direct modeling are of the same order of calculating basic functions like the hyperbolic cosh and sinh of the conductive quadrupole Q_c. This makes this approach interesting in view of exhaustive iterative estimation procedures. In case of black walls and very transparent material, DM3 reduces to DM2.

Parameter Vector

According to the direct models presented above, we have three possible parameter vectors.
For DM1, the parameter vector is the singleton $\beta = \{\beta_1 = tc = e^2/\alpha\}$. For DM2, the parameter vector is made by both t_c and the dimensionless radiative

resistance R_r. Thus $\beta = \{\beta_1=t_c,\ \beta_2=R_r\}$. For DM3, it is $\beta = \{\beta_1=tc,\ \beta_2=N_{Pl},\ \beta_3=\tau_0\}$. One can note that DM2 introduces the single parameter R_r for the radiative contribution that is made of both the two parameters of DM3, N_{Pl} and τ_0.

These parameter vectors can be extended when considering heat losses. In the case of modulated experiments of heat transfer, this is required at very low frequencies and we won't need to take this parameter into account (i.e. negligible sensitivity) in the present study.

One additional parameter may also be introduced. It is a 'nuisance' parameter that stems from experimental constraints. Due to the electronics of the lock-in detection, the absolute value of the phase lag in the measurements shown experiences an offset. This parameter will be considered as unknown and noted β_0 later on.

Direct simulations

Figure 1 shows theoretical simulations obtained with DM3 for different values of the optical thickness and Planck number, exhibiting how radiative transfer may alter the well known monotonically decreasing phase lag evolution of opaque materials (curve 0). Symbol f_{co} in the figure means that attention has been paid to the fact that at high frequency, the temperature can not be measured anymore (low detectivity limit reached). Simulations are stopped when the amplitude of the signal falls below the technical limit of the detection system used. Curve 1 has been obtained for $N_{Pl} = \tau_0 = 0.5$. One can observe the phase lag blocking occurring for a given high frequency, indicating a single in-phase radiative transfer. Curves 2,3 exhibit all the intermediate behaviors obtained for other sets of (N_{Pl}, τ_0) until the opacity limit (Curve 4, $N_{Pl} = \tau_0 = 5$).

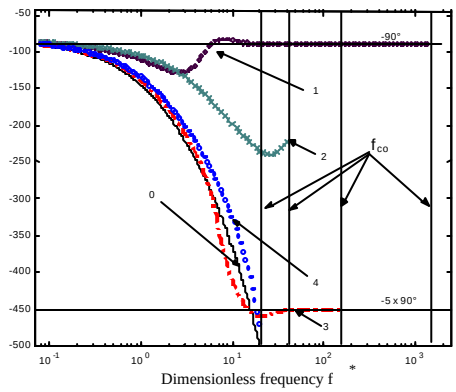


Figure 1 : Phase-lag φ versus frequency f^*

PARAMETER ESTIMATION PROBLEM

Experimental system

The experimental setup is derived from the one described in [5] but adapted to monolayer geometry (Fig 2). The temperature investigations are carried out by setting the sample at the focus of a double reflection 1-m diameter solar furnace. This facility is able to deliver a flux density up to 900 kW/m^2 onto a 25 mm diameter focal zone. It is equipped with a slow moving iris diaphragm that can tune the incident flux, but is unable to modulate this flux at high frequency. So, the modulated flux is delivered by a remote controlled 40 Watts CO_2 laser (Millenium Lasers Ltd., Model ML30/S) the beam of which is colinear to the concentrator axis and spreads over a 15 mm gaussian spot onto the sample surface. The rear face temperature is measured by optical pyrometry. The mean value is determined with a slow response pyrometer (Ultrakust InfraPlus). The temperature oscillations are monitored with an arrangement involving a ZnSe lens and a fast response IR detector (InSb) in order to sight a 2 mm diameter zone.

The control computer is connected to two IEEE488 devices :

- a synthesized function generator (Stanford Research, Model DS345) that controls the laser chopper frequency,
- a lock-in amplifier (Stanford Research, Model SR830) for recording the temperature phase lag: difference between the IR detector output and the function generator command (hardware demodulation).

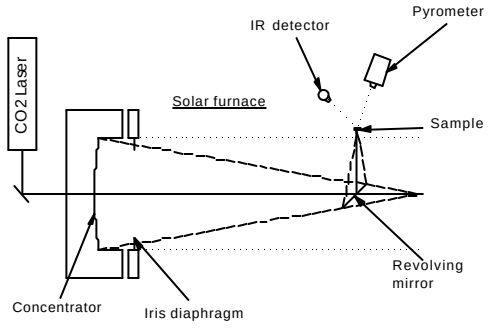


Figure 2 : Experimental set-up.

Measurements

The experimental data consists on the acquisition of the signal phase lag (with respect to

the laser input) versus the frequency of excitation. For each data set sequence programming, the following experimental parameters have to be selected:
- frequency values: linear or logarithmic evolution or values list,
- stabilization time before measurements start,
- measurement recording definition: data number, integration time, duration.

For a given frequency value, the lock-in outputs (phase and modulus) are measured a given number of times. Phase measurement reliability can be evaluated from the observed standard deviation σ_φ and from the modulus evolution. When signal-to-noise ratio decreases, phase inaccuracy increases. In the experimental results shown, only 5 phase measurements are repeated for each data point. Rather than having a statistical mean that could be introduced as a diagonal weighting matrix in deterministic algorithms, the measured standard deviation (Std Dev) provides an estimation of the squared residues cost function that can be expected. In our experiments, one have generally $S = \sum_i (3\sigma_{\varphi_i})^2 \cong 0.075$.

Estimation procedures

❑ It has been shown previously [6] that accounting for the various sources that generates and transforms the noise that corrupts the data, it can be considered as additive, of gaussian type, having zero-mean value and Std Dev proportional to the square root of the band pass integration. When setting the integration time in an appropriated manner for each tested frequency, the Std Dev of the noise can be made constant. All these conditions are included in the model for the noise that has been considered in the stochastic theoretical study.

❑ Estimations are performed in the classical way of minimizing the uni-modal cost functional that consists on the square of the phase lag residuals, i.e. distance between experimental and calculated signals. Four inversion algorithms have been considered in the estimation software :

- the simplex Nelder-Mead and Levenberg-Marquardt algorithms, based on a deterministic search of the minimum, based on the computation of the descending gradients. These methods are very efficient in terms of computational time. In some complicated cases where a good design of the PEP can not be achieved (i.e. in case of correlated parameters), such methods lead to local minima that are dependent upon the initial guesses. This induces an error on all the estimates, according to the influence of the noise on the sensitivity matrix. This problem may be overcome in some situations, when multiplying starting guesses and averaging the resulting estimates. This situation may arise here as a previous theoretical stochastic study [3] shows that in the particular case of small optical thicknesses, a high correlation exists between the two "radiative" parameters τ_0 and N_{Pl}.
- the stochastic optimization algorithms used are the "Simulated Annealing" method (SA), which has been already used in heat transfer problems by Silva Neto and Soeiro [7], and the method due to Berg [8]. They consist basically in a random walk over the cost-functional landscape over the hyper-space domain of the parameter vector. In order to avoid to be trapped in a local minimum, the SA algorithm make use of a probabilistic model for escaping from it, that relies on an Arrhenius (or thermally activated) law where the temperature parameter used here has nothing to do with the physical variable of the problem, but can be considered as a fine tuning parameter for the algorithm. On the other hand, the Berg method retains all the local minima encountered during pure random walks.

The software has been written in MATLAB language. It makes use of the built-in function fminsearch for the Nelder-Mead simplex algorithm, of the function leasqr.m developed by Richard I. Shrager, from the National Institute of Health (Bethesda, MD, USA) and modified by Arthur Jutan and Ray Muzic, available at the Internet Matlab site (matlab42/toolbox/contrib), and functions SimAnneal.m and randcost.m developped by Peter Cervelli from the Dpt of Geophysics (Stanford University).

Results and discussion

In this section, we give examples showing the ability of this method to yield the phononic diffusivity of STM easily, in all experimental configurations. Focusing on pure aspects of PEP, we mainly discuss the problem of estimating radiative parameters. For this purpose, we concentrate our efforts on measurements where an evident correlation exists between parameters N_{Pl} and τ_0 when using DM3.

Identification test case : Figure 3 shows phase lag measurements (star dots) for a sample of 4 mm thickness of pure $Si0_2$, at 600 °C. The

emissivities are considered as equal to 1 (Graphite paint coating). At this temperature, the phononic diffusivity measured with the flash method is $5.7\,10^{-7}$ m²/s. This yields a phononic conductivity k=1.55 W/m/K when considering experimental data of $\rho C_P(T)$ [2]. The expected characteristic time is then t_c = 28.2 s. Eighteen measurement points have been acquired between 0.02 and 1 Hz. The average Std Dev associated to each data point is of the order of σ_φ=0.7° The phase lag blocking predicted by the model at "high" frequency can be noted here and looking back at Fig. 1, one can state that it corresponds to a case of small Planck number N_{Pl} and small optical thickness τ_0.

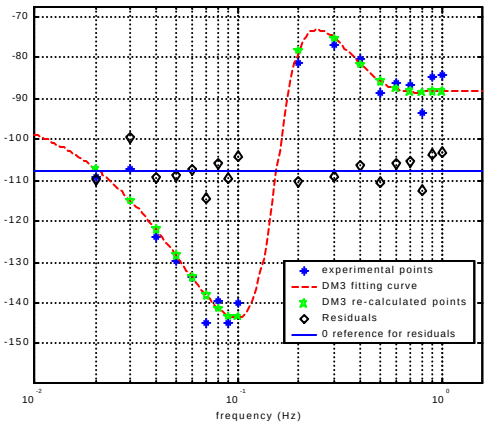


Figure 3 : Example of experimental phase lag measurements along with a model fitting.

In order to allow the discussion that will follow, let's anticipate the identified results and consider DM3 in the case

$$\beta = \{t_c = 28.2,\ N_{Pl} = 0.15,\ \tau_0 = 0.1\}$$

A theoretical stochastic analysis [3], allows to calculate the expected variance-covariance matrix, that is presented below in a modified form where the diagonal terms correspond to the relative error that is expected to affect the parameters, calculated through $\hat{\sigma}_{\beta_i}/\hat{\beta}_i$ and expressed in %, and where the off-diagonal terms are the cross-correlation coefficients. One obtains the following result when considering a noise of $\sigma_\varphi = 0.7°$. The assumptions on the measurement errors are the standard ones denoted 11111111 according to the notation of Beck & Arnold.

$$\mathrm{cov}_{\mathrm{mod}}(\hat{\beta}) = \begin{bmatrix} 0.9 & 0.48 & 0.5 \\ 0.48 & 22.6 & 0.99 \\ 0.5 & 0.99 & 24.3 \end{bmatrix} \qquad (6)$$

One can first conclude that the two radiative parameters are highly correlated in this region of the parameter space and that it is impossible to pretend identifying both within a good confidence interval. The second conclusion is that besides this, the identification of the diffusivity can be performed in fairly good conditions : 95% confidence bounds of the order of 2.1% (t-distribution).
Despite this obviously ill-conditionned PEP, we will implement lately stochastic algorithms for performing identification in this hard test case; in order to compare their performance to direct methods.

Use of reduced DM2. We first start estimations with DM2, which is the reduced model that corresponds to the expected situation of small optical thickness. Independently of the algorithms used and of any initial guesses, all algorithms converges onto the following result

$$\alpha = 6.\ 10^{-7}\ \mathrm{m^2/s} \text{ and } R_r = 2.2$$

The diffusivity value is 6% higher than the one found with the flash method. The quality and facility for performing the identification can be confirmed when using the residuals at convergence for calculating the variance-covariance matrix of the parameters. One finds

$$\mathrm{cov}_{\mathrm{mod}}(\hat{\beta}) = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 1.0 \end{bmatrix} \qquad (7)$$

This means that a 0.6% and 1% of relative error are obtained respectively for α and R_r. In this case the value found for the nuisance parameter β_0 was 38°. Considering the value of R_r and the value of the thermal conductivity obtained by the flash method, eq. (5) allows then to calculate both the optical thickness and Planck number that should be estimate later on. One finds $\tau_0 = \beta e \cong 0.93$ and then to $N_{Pl} \cong 1$ (see formula of table 1).

Use of full DM3. The thermal diffusivity value that has been estimated from DM2 will be introduced as initial guess when using DM3. One will then be able to confirm that this value may

not be altered much despite the high correlation between the two radiative parameters. Concerning these latter, starting with the knowledge drawn from direct simulations performed for various parameter vectors, one will retain to initiate estimations, the initial guess of N_{Pl}=0.5 and $\tau_0 = 0.5$ (curve 1 of Fig. 1) and also different values of the parameter vector in the neighborhood of this point : 0.1 and 1 for both N_{Pl} and τ_0. This leads to 9 different starting points for identification.

Concerning deterministic algorithms, three clearly distinct local minima (1,2,3) have been found (Table 1). When using the Simulated Annealing algorithm, one have only to provide lower and upper bounds for the parameter vector. Then the algorithm initiates itself and does random walks. When simply repeating the estimation procedure a certain number of times, this algorithm recovers the previous minima encountered but also finds more minima (4). This is not a surprise as it may explore other regions of the cost-functional topography. Deterministic algorithms directly goes to the local minima closest to the initial guess.

The conclusion that can be stated is that both algorithms are unable to find the good global minimum in this ill-conditioned case.

Table 1. Local Minima

Parameter vector components	**Minima**			
	N°1	**N°2**	**N°3**	**N°4**
α (10^7 m²/s)	5.2	4.9	5.85	4.8
N_{Pl}	0.6	0.74	0.16	0.82
τ_0	0.38	0.5	0.07_6	0.56
β_0	42°	43°	38°	44°
$R_r = \frac{N_{Pl}}{\tau_0}(\tau_0 + 1)$	2.18	2.22	2.26	2.28

Thanks to the reduced direct model DM2 used before, one can try to go further on, searching for the more correct or likely minimum. Looking at Table 1, one can see that only for the minima n°3, the good diffusivity value is found : 5.85 10^{-7} m²/s. Among all other values found, this is the closest one to the value estimated with DM2 and to the value obtained by the flash method. As proved by the covariance matrix of the estimated parameters calculated in (6), the correct diffusivity can be found with DM3 that is not corrupted by the large error made on the correlated parameters. Furthermore, this minimum may also be considered as the good one as only in this case, the value of the nuisance parameter β_0 found is the same as the one found with DM2. This parameter also has a unique value for a given set of data that has not to depend on the estimation procedure.

Retaining minimum 3 as the most probable one, one can proceed to estimations in conditions where β_0 and $\beta_1 = t_c$ (i.e. α) are considered as known parameters. One have then a cost function that depends only on the two radiative parameters $\beta_2 = N_{Pl}$ and $\beta_3 = \tau_0$. One can then focus on a small domain of variations of both parameters, centered on the values given by minimum 2: $\beta_c = \{N_{Pl} \cong 0.16\,,\, \tau_0 \cong 0.07\}$. The algorithm of Berg allows to perform random walks and retaining all the minima encountered on them. Then a statistics for the estimations can be obtained and studied. In the same manner, if N distinct minima are found, N identifications are repeated with different initial guesses selected in the prescribed constraint domain using deterministic algorithms.

4000 random walks (106000 function evaluations) have been performed in a parameter domain centered on β_c, and of 40% span for the parameter movements around β_c. 200 distinct minima have been found. The statistics show a distribution of the parameter estimates that is rather flat, not of gaussian type, centered about the initial values $\bar{\hat{\beta}} = \{N_{Pl} = 0.14_6\,,\, \tau_0 = 0.07\}$. If one calculates a standard deviation of the parameter estimates distribution about their mean, it is found $\sigma_{\hat{\beta}} \big/ \bar{\hat{\beta}} = \{19.3\,\%, 20.6\%\}$ which is the value drawn from the stochastic analysis given in Eq.(6). If one proceed with 200 repetitions of deterministic algorithms for different initial guesses in the same domain, exactly the same result is found thus confirming published results on the comparison of both approaches in parameter estimation. Two conclusions can be drawn at this point. The non gaussian distribution of the estimates shall lead to the conclusion that the minimum found is not the good one. Furthermore, using the values of $\bar{\hat{\beta}}$ along with the definition of N_{Pl} and τ_0, one finds a conductivity of 2.8 W/m/K which is far from the known value. Supposing that the good minimum (calculated previously with the estimated

parameter R_r of DM2) when introducing the knowledge of the thermal conductivity) can be found, it will be obtained with a very large error. These two conclusions shall lead to this last one : it is impossible to identify both parameters in the case of small optical thicknesses and therefore use has to be made of model DM2.

This result can be illustrated by a contour plot of the cost functional in the parameter space defined above (Fig 4). Some minima found during the random walks or repeated deterministic estimations are plotted also (big dots). The reason of this very ill conditioned situation can be understand when looking at the very narrow and non convex shape of the minimum valley. The minimum cost loci in the vicinity of β_c follow a line that gives the constant ratio $N_{Pl}/\tau_0 = 2$, nearly the value of the radiative resistance of DM2. This was also the case for the other minima found far from β_c but still in the valley (see Table 1). One can then understand how the physics rules the PEP. One conclusion is also that it may be dangerous to retain as best estimates, those giving the lowest cost (S=0.063 in our case) as it is done in stochastic algorithms. It is worthwhile to note that the averaged cost function value found in the valley is of $\overline{S} = 0.07$, which is a little below the optimal value predicted by computing it from the measured Std Dev of the experimental data. (S=0.075).

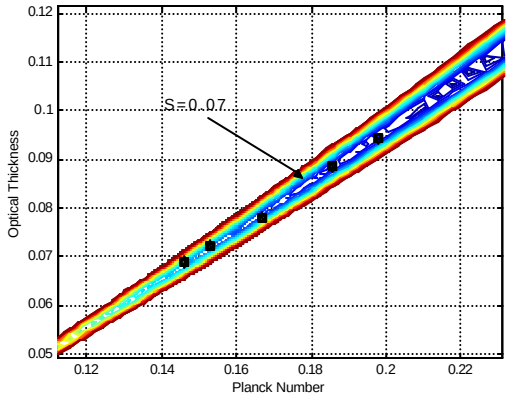


Figure 4: Contour plot of the cost function.

CONCLUSIONS

From both paper [3] and the present results with real experimental data, one can drawn several conclusions :

The periodic method is a good method for measuring the true diffusivity of STM, using either DM1 (results not shown here for small optical thickness samples having gold coated boundaries), or DM2 and DM3, depending on the case. Averaged gray optical parameters (mean absorption coefficient) can be measured with good accuracy only in the approximate range $0.5 \le \tau_0 \le 5$ (results not shown here), which seriously compromises the objective of being able to obtain through this method trustable average optical coefficients.

Acknowledgements. The authors acknowledge the financial support provided by FAPERJ and CNPq through research grants and a Visiting Professor scholarship.

REFERENCES

1. J. F. Schatz and G. Simmons, Method of simultaneous measurement of radiative and lattice thermal conductivity , *J. Appl. Phys.*, **43**, no. 6, 2586-2594 (1972).

2. S. Andre and A. Degiovanni, Experimental measurements of the phonic diffusivity of semi-transparent materials up to 800K, *Int. J. Glass Science & Tech.*, **66**, no.11, 291-298, (1993).

3. S. Andre, D. Maillet and A. Degiovanni, Periodic Method for Characterization of Glass by use of an Analytical Model, *Proceedings of the 3rd Int Symp. Rad. Transfer*, ICHMT, Turkey, 8 pages, (June 2001).

4. D. Maillet, A. Degiovanni, J.C. Batsale, C. Moyne and S. Andre, Thermal Quadrupoles : Solving the Heat Equation through Integral Transforms. *John Wiley & Sons*, (October 2000).

5. E. Ranchy, J-J. Serra and Ph. Hervé, Determination of heat transfer modes in Y_2O_3-ZrO_2 plasma sprayed coatings, *High Temp. Chem. Processes*, **3**, 233-238 (1994).

6. C. Gervaise, Caractérisation thermique multi-échelles de revêtements réfractaires – Comportement sous excitation périodique et identification par méthodes inverses, *Thèse de Doctorat de 3ème cycle* (1999).

7. A. J. Silva Neto and F. J.C.P Soeiro, A combination of gradient based and global optimization methods for the solution of inverse heat transfer problems, *Inverse Problems in Engineeering*, accepted for publication (2002).

8. B. Berg, Locating global minima in optimization problems by a random-cost approach, Nature, 361, 708-710 (1993).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

AN EXPLICIT FORMULATION BASED ON THE MOMENTS OF THE EXIT RADIATION INTENSITY FOR THE ONE-DIMENSIONAL INVERSE RADIATIVE TRANSFER PROBLEM

Antônio J. Silva Neto
Department of Mechanical Engineering and Energy, Instituto Politécnico, IPRJ, Universidade do Estado do Rio de Janeiro, UERJ, P.O. Box 97282, 28601-970, Nova Friburgo, Rio de Janeiro, RJ, Brazil
ajsneto@iprj.uerj.br
Nuclear Engineering Program – PEN/COPPE, Universidade Federal do Rio de Janeiro, UFRJ, Caixa Postal 68509, CEP 21945-970, Rio de Janeiro, RJ, Brazil
ajsneto@lmn.con.ufrj.br

Norman J. McCormick
Department of Mechanical Engineering, University of Washington, P.O. Box 352600, 98195-2600, Seattle, Washington, USA
mccor@u.washington.edu

ABSTRACT

An explicit formulation is developed for the simultaneous estimation of the single scattering albedo and diffuse reflectivities of a semi-transparent slab using moments of the exit radiation intensity. Results from preliminary numerical tests are presented.

INTRODUCTION

The procedure usually employed for solving inverse problems of radiative transfer involves a numerical iterative implicit solution in which a cost functional (*e.g.* squared residues between calculated and measured quantities) is minimized when estimating the unknowns [1-4].

For some problems in very simple geometries, analytically based procedures can be developed that provide an alternative to the implicit method [5-7]. The unknowns to be recovered with an analytically based system of equations appear explicitly, although in some cases the resulting equations from the explicit formulation are nonlinear. In such cases an iterative solution is required for the equations that differs from that with an iterative implicit formulation.

The explicit algorithms developed in Refs. [5-7] do not require the solution of the direct problem. In Ref. [8] is presented the source-detector methodology, which is also an explicit formulation but it requires the solution of the direct problem in the analytical formulation of the inverse problem.

In this work two formally exact equations are used for estimating the single scattering albedo and the diffuse reflectivities, at the inner boundary surfaces of the medium, of an isotropically scattering, homogeneous slab medium that is illuminated from the outside. The explicit solution of this inverse radiative transfer problem for the three unknowns requires only experimental data of the ingoing and exiting radiation intensity moments acquired with external sensors. Numerical tests are presented for the case in which two equations come from an experiment with isotropic external illuminations and a third equation comes from an experiment with a constant heat flux illumination. The tests are presented for media with different optical thicknesses.

The formulation also can be easily used to give two formally exact equations for a collimated incident illumination. The three unknowns then could be obtained from two experiments with different incident illumination directions.

MATHEMATICAL FORMULATION OF THE DIRECT PROBLEM

As schematically represented in Fig. 1, a plane-parallel, gray, isotropically scattering slab of optical thickness τ_0, with diffusely reflecting boundaries, is subjected to external isotropic irradiation at both boundaries, $\tau = 0$ and $\tau = \tau_0$.

It is assumed that the emission of radiation by the medium due to its temperature is negligible in comparison to the external radiation.

The mathematical formulation of the direct radiation problem is given by [9]

$$\mu\frac{\partial I(\tau,\mu)}{\partial\tau}+I(\tau,\mu)=\frac{\omega}{2}\int_{-1}^{1}I(\tau,\mu')d\mu',$$

$$0\le\tau\le\tau_0,\quad -1\le\mu\le 1 \qquad (1)$$

where $I(\tau,\mu)$ is the dimensionless radiation intensity, τ is the optical distance, μ is the direction cosine of the radiation beam with the positive τ axis, and ω is the unknown single scattering albedo. The boundary conditions for isotropic external illuminations are given by

$$I(0^+,\mu)=A_1+2\rho_1\int_0^1 I(0^+,-\mu')\mu'd\mu' \quad ,\mu>0 \quad (2a)$$

$$I(\tau_0^-,-\mu)=A_2+2\rho_2\int_0^1 I(\tau_0,\mu')\mu'd\mu' \quad ,\mu>0 \quad (2b)$$

and the boundary conditions for constant heat flux illuminations are

$$I\left(0^+,\mu\right)=(C_1/\mu)+2\rho_1\int_0^1 I\left(0^+,-\mu'\right)\mu'd\mu' \ ,\mu>0 \qquad (2c)$$

$$I\left(\tau_0^-,-\mu\right)=(C_2/\mu)+2\rho_2\int_0^1 I\left(\tau_0^-,\mu'\right)\mu'd\mu',\mu>0 \qquad (2d)$$

Here ρ_1 and ρ_2 are the unknown diffuse reflectivities, and A_1, A_2, C_1, and C_2 are the strengths of the external irradiations just inside the boundaries at $\tau=0^+$ and $\tau=\tau_0^-$, respectively. Index of refraction effects at the boundary have been neglected.

MATHEMATICAL FORMULATION OF THE INVERSE PROBLEM

We wish to solve an inverse problem of radiation governed by Eq. (1). Initially we show that the two diffuse reflectivities ρ_1 and ρ_2 and the single scattering albedo ω can be estimated by solving

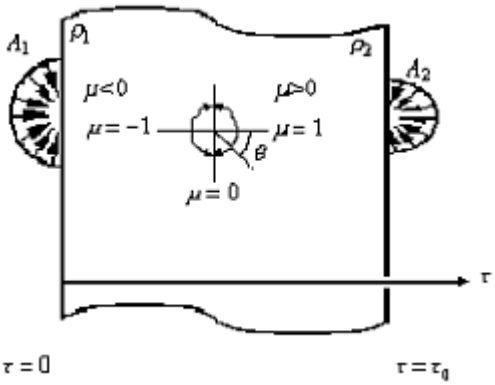


Figure 1 – Plane-parallel medium with diffusely reflecting boundaries subjected to isotropic external irradiations.

three coupled equations for appropriately selected conditions on A_1, A_2, C_1, and C_2. Two equations come from an experiment with isotropic external illuminations, as given by Eqs. (2a,b), and a third equation comes from an experiment with constant external heat flux illuminations as in Eqs. (2c,d). The procedure to derive the equations utilizes ideas developed by Siewert [5,10] and McCormick [11].

General Inverse Equations

We will develop two general equations before searching for the appropriate boundary illuminations to impose. Both general equations are derived with the use of a change from μ to $-\mu$ in Eq. (1),

$$-\mu\frac{\partial I(\tau,-\mu)}{\partial\tau}+I(\tau,-\mu)=\frac{\omega}{2}\int_{-1}^{1}I(\tau,\mu')d\mu',$$

$$0\le\tau\le\tau_0,\quad -1\le\mu\le 1 \qquad (3)$$

Multiplication of Eq. (1) by $\partial I(\tau,-\mu)/\partial\tau$ and Eq. (3) by $\partial I(\tau,\mu)/\partial\tau$, integration of each equation over $-1 \le \mu \le 1$, and then adding the results we obtain

$$\frac{d}{d\tau}\int_{-1}^{1}I(\tau,\mu)I(\tau,-\mu)d\mu=\frac{\omega}{2}\frac{d}{d\tau}\left[\int_{-1}^{1}I(\tau,\mu)d\mu\right]^2 \qquad (4)$$

Integration of Eq. (4) over the spatial domain, $0^+\le\tau\le\tau_0^-$, and use of symmetry on the left hand side of the equation gives

$$4\left[\int_0^1 I(\tau_0^-,\mu)I(\tau_0^-,-\mu)d\mu - \int_0^1 I(0^+,\mu)I(0^+,-\mu)d\mu\right] =$$

$$= \omega\left\{\left[\int_{-1}^1 I(\tau_0^-,\mu)d\mu\right]^2 - \left[\int_{-1}^1 I(0^+,\mu)d\mu\right]^2\right\} \quad (5)$$

Equation (5) is the first general inverse equation for this type of problem for a homogeneous medium.

The second general inverse equation is derived by first multiplying Eq. (1) by $\mu I(\tau,-\mu)$ and Eq. (3) by $\mu I(\tau,\mu)$, and then integrating each equation over $-1 \le \mu \le 1$. After subtracting the results we find

$$\frac{d}{d\tau}\int_{-1}^1 \mu^2 I(\tau,\mu)I(\tau,-\mu)d\mu =$$

$$= -\omega\left[\int_{-1}^1 \mu I(\tau,\mu)d\mu\right]\left[\int_{-1}^1 I(\tau,\mu')d\mu'\right] \quad (6)$$

Then we integrate Eq. (1) over $-1 \le \mu \le 1$ to obtain the energy conservation equation

$$\int_{-1}^1 I(\tau,\mu)d\mu = -(1-\omega)^{-1}\frac{d}{d\tau}\int_{-1}^1 \mu I(\tau,\mu)d\mu \quad (7)$$

From the symmetry and the definition of $\tilde{\omega}$,

$$\int_{-1}^1 \mu^2 I(\tau,\mu)I(\tau,-\mu)d\mu = 2\int_0^1 \mu^2 I(\tau,\mu)I(\tau,-\mu)d\mu \quad (8)$$

$$\tilde{\omega} = \frac{\omega}{1-\omega} \quad (9)$$

we obtain from Eqs. (6) and (7),

$$4\frac{d}{d\tau}\int_0^1 \mu^2 I(\tau,\mu)I(\tau,-\mu)d\mu = \tilde{\omega}\frac{d}{d\tau}\left[\int_{-1}^1 \mu I(\tau,\mu)d\mu\right]^2 \quad (10)$$

Integration of Eq. (10) over $0^+ \le \tau \le \tau_0^-$ leads to the final result for the second equation,

$$4\left[\int_0^1 \mu^2 I(\tau_0^-,\mu)I(\tau_0^-,-\mu)d\mu - \int_0^1 \mu^2 I(0^+,\mu)I(0^+,-\mu)d\mu\right] =$$

$$= \tilde{\omega}\left\{\left[\int_{-1}^1 \mu I(\tau_0^-,\mu)d\mu\right]^2 - \left[\int_{-1}^1 \mu I(0^+,\mu)d\mu\right]^2\right\} \quad (11)$$

Isotropic and Constant Heat Flux Illuminations

We wish to incorporate the boundary conditions, given by Eqs. (2a,b) into Eqs. (5) and (11), which is done by breaking up the integrals over –1 to 1 into separate integrals from –1 to 0 and from 0 to 1. To do this we define the moments of the radiation intensities $I(0^+,-\mu)$ and $I(\tau_0^-,\mu)$, for $\mu > 0$, inside the surfaces $\tau = 0$ and $\tau = \tau_0$, as

$$E_n(0) = \int_0^1 \mu^n I(0^+,-\mu)d\mu, \quad n = 0,1,2 \quad (12a)$$

$$E_n(\tau_0) = \int_0^1 \mu^n I(\tau_0^-,\mu)d\mu, \quad n = 0,1,2 \quad (12b)$$

The moments correspond to the outward portions of the energy density (n = 0) and the heat flux $(n=1)$. Measurement of the emerging radiation quantities for n = 2 would be a problem because they do not correspond to any common measurement system, but one can measure the radiation intensity at different polar angles, and then use Eqs. (12a,b) to obtain the moments.

The inverse Eqs. (5) and (11) are written for radiation intensities inside the surfaces, *i.e.* at $\tau = 0^+$ and $\tau = \tau_0^-$. Therefore it is necessary to write Eqs. (12a,b) using quantities that can be measured using only external detectors, as for example the exit radiation intensities $I(0^-,-\mu)$ and $I(\tau_0^+,\mu)$, for $\mu > 0$. A schematic representation is given in Fig. 2.

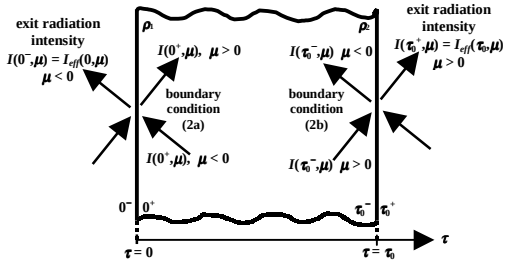


Figure 2 – Schematic representation of the radiation intensities at the boundaries of the medium.

The number of photons interacting with the surfaces can be preserved with the constraints

$$\int_{-1}^{1} \mu\left[I(0^+,\mu) - I(0^-,\mu)\right]d\mu = 0 \qquad (13a)$$

$$\int_{-1}^{1} \mu\left[I(\tau_0^+,\mu) - I(\tau_0^-,\mu)\right]d\mu = 0 \qquad (13b)$$

which can be satisfied by the equations

$$I(0^+,-\mu) = \frac{I(0^-,-\mu)}{(1-\rho_1)} = \frac{I_{eff}(0,-\mu)}{(1-\rho_1)},\ 0 \le \mu \le 1 \quad (14a)$$

$$I(\tau_0^-,\mu) = \frac{I(\tau_0^+,\mu)}{(1-\rho_2)} = \frac{I_{eff}(\tau_0,\mu)}{(1-\rho_2)},\ 0 \le \mu \le 1 \quad (14b)$$

where the symbols $I_{eff}(0,-\mu)$ and $I_{eff}(\tau_0,\mu)$ for $\mu > 0$ are used to represent quantities that can be effectively measured using only external detectors.

With the definitions

$$F_n(0) = \int_0^1 \mu^n I_{eff}(0,-\mu)d\mu, \qquad n = 0,1,2 \qquad (15a)$$

$$F_n(\tau_0) = \int_0^1 \mu^n I_{eff}(\tau_0,\mu)d\mu, \qquad n = 0,1,2 \qquad (15b)$$

Eqs. (12a,b) can be written as

$$E_n(0) = \frac{F_n(0)}{(1-\rho_1)}, \qquad n = 0,1,2 \qquad (16a)$$

$$E_n(\tau_0) = \frac{F_n(\tau_0)}{(1-\rho_2)}, \qquad n = 0,1,2 \qquad (16b)$$

so the boundary conditions of Eqs. (2a-d) now can be written as

$$I(0^+,\mu) = A_1 + 2\rho_1 E_1(0), \quad \mu > 0 \qquad (17a)$$

$$I(\tau_0^-,-\mu) = A_2 + 2\rho_2 E_1(\tau_0), \quad \mu > 0 \qquad (17b)$$

$$I(0^+,\mu) = (C_1/\mu) + 2\rho_1 E_1^*(0), \quad \mu > 0 \qquad (17c)$$

$$I(\tau_0^-,-\mu) = (C_2/\mu) + 2\rho_2 E_1^*(\tau_0), \quad \mu > 0 \qquad (17d)$$

where the superscript * distinguishes the constant heat flux experimental results from those for the isotropic illumination experiment.

By recognizing in Eqs (5) and (11) the terms that correspond to the boundary conditions and those that relate to the moments of the exit radiation intensities, and by introducing Eqs. (12), (15), (16) and (17a,b), two inverse equations for an isotropic illumination are obtained,

$$\omega = \frac{N}{D}, \qquad \tilde{\omega} = \frac{\tilde{N}}{\tilde{D}} \qquad (18, 19)$$

where

$$N = 4\{E_0(\tau_0)[2\rho_2 E_1(\tau_0) + A_2] - E_0(0)[2\rho_1 E_1(0) + A_1]\} \qquad (20a)$$

$$D = [2\rho_2 E_1(\tau_0) + E_0(\tau_0) + A_2]^2 - [2\rho_1 E_1(0) + E_0(0) + A_1]^2 \qquad (20b)$$

$$\tilde{N} = 4\{E_2(\tau_0)[2\rho_2 E_1(\tau_0) + A_2] - E_2(0)[2\rho_1 E_1(0) + A_1]\} \qquad (20c)$$

$$\tilde{D} = \left[E_1(\tau_0)(1-\rho_2) - \frac{A_2}{2}\right]^2 - \left[E_1(0)(1-\rho_1) - \frac{A_1}{2}\right]^2 = \left[F_1(\tau_0) - A_2/2\right]^2 - \left[F_1(0) - A_1/2\right]^2 \qquad (20d)$$

To obtain a third equation we utilize the boundary conditions given by Eqs. (17c,d) with the simplification that $C_2 = 0$. If we were to attempt to use these boundary conditions with Eq. (5) then an essential singularity corresponding to integrating $1/\mu$ from 0 to 1 would occur, so we focus on use of Eqs. (17c,d) with Eq. (11) to obtain

$$\tilde{\omega} = \tilde{N}^*/\tilde{D}^* \qquad (21)$$

where

$$\tilde{N}^* = 4\left\{2\rho_2 E_1^*(\tau_0)E_2^*(\tau_0) - E_1^*(0)\left[2\rho_1 E_2^*(0) + C_1\right]\right\} \qquad (22a)$$

$$\tilde{D}^* = \left[E_1^*(\tau_0)(1-\rho_2)\right]^2 - \left[E_1^*(0)(1-\rho_1) - C_1\right]^2 = F_1^*(\tau_0)^2 - \left[F_1^*(0) - C_1\right]^2 \qquad (22b)$$

Equations (18), (19) and (21) are sufficient, in principle, to determine the three unknowns ω ρ_1, and ρ_2 provided two of the equations do not become degenerate. Note that the equations involve a highly nonlinear search, however, even though they appear formally in closed form. From Eqs. (16) we see that the moments of the exit radiation intensities depend on the diffuse reflectivities, and in Eqs. (20) there are several

products of those quantities. Also, one can observe that the optical thickness τ_0 cannot be determined with the time-independent experiments as proposed here because integrations were performed in the optical variable in the range $[0,\tau_0]$, thus eliminating the optical thickness from the final formulation of the inverse problem.

A disadvantage of these equations is that the experiments would be difficult to perform, so we also develop additional equations from which all three parameters can be obtained simultaneously from experiments done with collimated incident illuminations.

Because the direct measurement of $E_n(0)$, $E_n(\tau_0)$, $E_n^*(0)$ and $E_n^*(\tau_0)$ for $n=0$ and 2 would be difficult, however, a more practical approach is to change the surface illumination conditions.

Collimated Illumination

More and potentially better equations can be obtained if the external illuminations previously considered for the isotropic and constant heat flux cases are replaced by a set of collimated external illuminations. We will now set $A_2 = 0$ in the boundary condition given by Eq. (2b) and assume that the set of measurements are made so that the boundary conditions to be satisfied are

$$I_j(0^+,\mu) = A_j\delta(\mu-\mu_j) + 2\rho_1\int_0^1 I_j(0^+,-\mu')\mu' d\mu', \quad \mu>0 \tag{23a}$$

$$I_j(\tau_0^-,-\mu) = 2\rho_2\int_0^1 I_j(\tau_0^-,\mu')\mu' d\mu', \quad \mu>0 \tag{23b}$$

with j = 1 to J denoting each experiment performed. Clearly J must be such that the use of of the collimated boundary conditions in Eqs. (5) and (11) gives at least three equations, but more equations could be used.

By redefining the moments of the exit radiation given by Eqs. (15) and (16), including the index j to indicate each experiment performed, with $j = 1, 2, \ldots, J$, the boundary conditions (23) are written as

$$I_j(0,\mu) = A_j\delta(\mu-\mu_j) + 2\rho_1 E_{1j}(0), \quad \mu>0 \tag{24a}$$

$$I_j(\tau_0,-\mu) = 2\rho_2 E_{1j}(\tau_0), \quad \mu>0 \tag{24b}$$

for j = 1 to J. With the same procedure as before, boundary conditions (24) lead to the following system of equations

$$\omega = N_j/D_j, \qquad j = 1, 2, \ldots, J \tag{25}$$

$$\tilde{\omega} = \tilde{N}_j/\tilde{D}_j, \qquad j = 1, 2, \ldots, J \tag{26}$$

where the right hand sides of Eqs. (25) and (26) are given by Eqs. (20a-d) with $A_2 = 0$, A_1 replaced by Aj, and $E_n(0)$ and $E_n(\tau_0)$ replaced by $E_{nj}(0)$ and $E_{nj}(\tau_0)$.

SOLUTION OF THE INVERSE PROBLEM

For the vector of unknowns

$$\vec{z} = \{\omega, \rho_1, \rho_2\}^T \tag{27}$$

the solution of the inverse problem is obtained by solving a system of nonlinear equations

$$\vec{H}(\vec{z}) = 0 \tag{28}$$

For the first formulation with an isotropic illumination and constant heat flux, we obtain from Eqs. (9), (18), (19) and (21)

$$H_1(\vec{z}) = D\omega - N = 0 \tag{29a}$$

$$H_2(\vec{z}) = \tilde{D}\omega - \tilde{N}(1-\omega) = 0 \tag{29b}$$

$$H_3(\vec{z}) = \tilde{D}^*\omega - \tilde{N}^*(1-\omega) = 0 \tag{29c}$$

A similar set of equations corresponding to Eqs. (25) and (26) follows for the second formulation involving the set of collimated illuminations.

From a Taylor series expansion, by keeping only the terms up to the first order,

$$H(\vec{z}^{k+1}) = H(\vec{z}^k) + J^k\Delta\vec{z}^k = 0 \tag{30}$$

where k is an iteration counter, J is the Jacobian matrix, and $\Delta\vec{z}^k$ are corrections for the unknowns to allow for new estimates

$$\vec{z}^{k+1} = \vec{z}^k + \Delta\vec{z}^k \tag{31}$$

The elements of the Jacobian matrix are given by

$$J_{ij} = \frac{\partial H_i}{\partial z_j}, \qquad j = 1,2,3 \tag{32}$$

where $i = 1, 2, 3$ for the first inverse problem formulation, and $i = 1, 2, ..., 2J$ for the second inverse problem formulation.

After multiplying Eq. (32) by the transpose of the Jacobian matrix, the vector of corrections can be calculated using

$$J^{T^k} J^k \Delta \vec{z}^k = -J^{T^k} H\left(\vec{z}^k\right) \tag{33}$$

Therefore, to obtain the solution of the system of nonlinear equations (29), an iterative procedure is constructed by starting with an initial guess $\vec{z}^0$, and obtaining in a sequential way corrections and new estimates for the unknowns using respectively Eqs. (33) and (31).

The iterative procedure is continued until a convergence criteria such as

$$\left|\Delta z_j^k / z_j^k\right| < \varepsilon, \qquad j = 1, 2, 3 \tag{34}$$

is satisfied, where ε is a small number. This tolerance should be related to the level of noise present in the experimental data. We used 10^{-5} for our numerical tests.

RESULTS

We used a Newton multivariable method to examine the performance of the explicit inverse method of analysis presented here for the estimation of ω, ρ_1, and ρ_2. Several test cases have been studied. We considered participating media with optical thicknesses varying in the range $0.1 \le \tau_0 \le 5.0$, single scattering albedo values in the range $0.1 \le \omega \le 0.9$ and diffuse reflectivities in the range $0.05 \le \rho_1, \rho_2 \le 0.9$.

In Tables 1 to 4 are shown the results for just a few cases in media with different optical thicknesses. These tables are self contained but a few remarks are made. In Table 1 for the cases with $\omega = 0.9$ we observe the importance of properly choosing the strength of the external sources. For $A_1 = C_1 = 1$ and $A_2 = 0$ convergence was not achieved, but good results are obtained with $A_1 = C_1 = 1$ and $A_2 = 5$. This happened because in the first situation the importance of ρ_1 is felt only after two internal reflections take place and the exit radiation at $\tau = 0$ and $\tau = \tau_0$ is measured.

Table 1 – Test results for a medium with optical thickness $\tau_0 = 0.1$. Noiseless data.

Case	Control Parameters	Initial Guess	Estimates
$\omega = 0.9$ $\rho_1 = 0.9$ $\rho_2 = 0.1$	$A_1 = C_1 = 1$ $A_2 = 0$	$\omega^0 = 0.4$ $\rho_1^0 = 0.4$ $\rho_2^0 = 0.4$	did not converge
$\omega = 0.9$ $\rho_1 = 0.9$ $\rho_2 = 0.1$	$A_1 = C_1 = 1$ $A_2 = 5$	$\omega^0 = 0.4$ $\rho_1^0 = 0.4$ $\rho_2^0 = 0.4$	$\omega = 0.8999$ $\rho_1 = 0.8999$ $\rho_2 = 0.0999$

In Table 2, for $\omega = 0.5$ and 0.9 poor results were obtained with $A_2 = 0$, but very good results were obtained with $A_2 = 5$. The importance of finding proper values for the strength of the external illumination can also be seen in Table 3 for $\omega = 0.9$.

Table 2 – Test results for a medium with optical thickness $\tau_0 = 1$. Noiseless data.

Case	Control Parameters	Initial Guess	Estimates
$\omega = 0.5$ $\rho_1 = 0.1$ $\rho_2 = 0.9$	$A_1 = C_1 = 1$ $A_2 = 0$	$\omega^0 = 0.4$ $\rho_1^0 = 0.4$ $\rho_2^0 = 0.4$	$\omega = 0.9858$ $\rho_1 = 0.9331$ $\rho_2 = 0.9683$
$\omega = 0.5$ $\rho_1 = 0.1$ $\rho_2 = 0.9$	$A_1 = C_1 = 1$ $A_2 = 5$	$\omega^0 = 0.4$ $\rho_1^0 = 0.4$ $\rho_2^0 = 0.4$	$\omega = 0.4999$ $\rho_1 = 0.0999$ $\rho_2 = 0.8999$
$\omega = 0.9$ $\rho_1 = 0.1$ $\rho_2 = 0.9$	$A_1 = C_1 = 10$ $A_2 = 0$	$\omega^0 = 0.4$ $\rho_1^0 = 0.4$ $\rho_2^0 = 0.4$	$\omega = 0.9949$ $\rho_1 = 0.7057$ $\rho_2 = 0.8701$
$\omega = 0.9$ $\rho_1 = 0.1$ $\rho_2 = 0.9$	$A_1 = C_1 = 10$ $A_2 = 5$	$\omega^0 = 0.4$ $\rho_1^0 = 0.4$ $\rho_2^0 = 0.4$	$\omega = 0.8999$ $\rho_1 = 0.0999$ $\rho_2 = 0.8999$

In Table 3 we observe for higher values of the diffuse reflectivities in a highly scattering medium with a large optical thickness that convergence could not be achieved. We have tried different combinations for the strength of external irradiation.

Table 3 – Test results for a medium with optical thickness τ_0 = 5. Noiseless data.

Case	Control Parameters	Initial Guess	Estimates
$\omega = 0.5$ $\rho_1 = 0.5$ $\rho_2 = 0.5$	$A_1 = C_1 = 1$ $A_2 = 0.8$	$\omega^0 = 0.8$ $\rho_1^0 = 0.8$ $\rho_2^0 = 0.8$	$\omega = 0.4998$ $\rho_1 = 0.5000$ $\rho_2 = 0.5002$
$\omega = 0.9$ $\rho_1 = 0.2$ $\rho_2 = 0.2$	$A_1 = C_1 = 10$ $A_2 = 5$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.8999$ $\rho_1 = 0.1999$ $\rho_2 = 0.2011$
$\omega = 0.9$ $\rho_1 = 0.8$ $\rho_2 = 0.2$	$A_1 = C_1 = 10$ $A_2 = 5$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	did not converge
$\omega = 0.9$ $\rho_1 = 0.8$ $\rho_2 = 0.2$	$A_1 = C_1 = 1$ $A_2 = 5$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.8999$ $\rho_1 = 0.8001$ $\rho_2 = 0.1999$
$\omega = 0.9$ $\rho_1 = 0.9$ $\rho_2 = 0.9$	$A_1 = C_1 = 10$ $A_2 = 0$ or $A_1 = C_1 = 10$ $A_2 = 20$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	did not converge

In Table 4 for $\omega = 0.5$ and 0.9 we observe that an under-relaxation is necessary to achieve convergence. In Eq. (31) instead of using the full correction $\Delta\vec{z}^k$ we used $\gamma\Delta\vec{z}^k$, where γ is a damping factor with $0 < \gamma < 1$. We also used damping factors on the test cases in Table 3.

All results presented in Tables 1 to 4 were obtained using simulated noiseless data. In Tables 5 and 6, data with different levels of noise are shown. The largest deviations were obtained for the estimation of ρ_1, which is consistent with a sensitivity analysis that demonstrated that the sensitivity to ρ_1 is lower than the sensitivity to ω and ρ_2.

CONCLUSIONS

The results demonstrate the applicability and accuracy of an explicit method developed here for the solution of an inverse radiative transfer problem with reflecting boundaries. The computational implementation of the formulation with collimated external irradiation is in progress. We note that an alternative approach to this problem has been developed in Ref. [12].

The solution obtained with the explicit formulation developed here also has been used as the initial guess for a more complicated inverse radiative transfer problem solved with a Levenberg-Marquardt implicit formulation. Although the problem involved a larger number of unknowns, use of the initial guess enabled convergence of the implicit formulation in a situation in which convergence otherwise did not occur[13].

Table 4 – Test results for a medium with optical thickness τ_0 = 0.5. Noiseless data.

Case	Control Parameters	Initial Guess	Estimates
$\omega = 0.5$ $\rho_1 = 0.8$ $\rho_2 = 0.8$	$A_1 = C_1 = 10$ $A_2 = 0$ $\gamma = 1.0$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.9916$ $\rho_1 = 0.9940$ $\rho_2 = 0.9852$
$\omega = 0.5$ $\rho_1 = 0.8$ $\rho_2 = 0.8$	$A_1 = C_1 = 10$ $A_2 = 0$ $\gamma = 0.05$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.5000$ $\rho_1 = 0.8001$ $\rho_2 = 0.8000$
$\omega = 0.9$ $\rho_1 = 0.9$ $\rho_2 = 0.9$	$A_1 = C_1 = 1$ $A_2 = 3$ $\gamma = 1.0$	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	did not converge
$\omega = 0.9$ $\rho_1 = 0.9$ $\rho_2 = 0.9$	$A_1 = C_1 = 5$ $A_2 = 1$ $\gamma = 1.0$	$\omega^0 = 0.7$ $\rho_1^0 = 0.7$ $\rho_2^0 = 0.7$	$\omega = 0.9989$ $\rho_1 = 0.9927$ $\rho_2 = 0.9809$
$\omega = 0.9$ $\rho_1 = 0.9$ $\rho_2 = 0.9$	$A_1 = C_1 = 5$ $A_2 = 1$ $\gamma = 0.1$	$\omega^0 = 0.7$ $\rho_1^0 = 0.7$ $\rho_2^0 = 0.7$	$\omega = 0.8993$ $\rho_1 = 0.8995$ $\rho_2 = 0.8995$

Table 5 – Test results for a medium with optical thickness τ_0 = 0.5. Data with noise.

Case	Control Parameters	Initial Guess	Estimates
$\omega = 0.1$ $\rho_1 = 0.1$ $\rho_2 = 0.9$	$A_1 = C_1 = 1$ $A_2 = 0$ 1% error	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.0922$ $\rho_1 = 0.0384$ $\rho_2 = 0.8961$
$\omega = 0.1$ $\rho_1 = 0.1$ $\rho_2 = 0.9$	$A_1 = C_1 = 1$ $A_2 = 0$ 5% error	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.0953$ $\rho_1 = 0.0182$ $\rho_2 = 0.8945$
$\omega = 0.5$ $\rho_1 = 0.8$ $\rho_2 = 0.8$	$A_1 = C_1 = 10$ $A_2 = 0$ 2% error	$\omega^0 = 0.5$ $\rho_1^0 = 0.5$ $\rho_2^0 = 0.5$	$\omega = 0.5005$ $\rho_1 = 0.8122$ $\rho_2 = 0.8112$

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by CNPq, FAPERJ and CAPES. A.J. Silva Neto acknowledges Prof. Nilson C. Roberty for useful discussions, and the doctorate student Rosana P. F. Pinheiro for the aid in the computational implementation.

Table 6 – Test results for a medium with optical thickness $\tau_0 = 1$. Data with noise.

Case	Control Parameters	Initial Guess	Estimates
$\omega = 0.5$	$A_1 = C_1 = 1$	$\omega^0 = 0.4$	$\omega = 0.4998$
$\rho_1 = 0.1$	$A_2 = 5$	$\rho_1^0 = 0.4$	$\rho_1 = 0.0984$
$\rho_2 = 0.9$	1% error	$\rho_2^0 = 0.4$	$\rho_2 = 0.8998$
$\omega = 0.5$	$A_1 = C_1 = 1$	$\omega^0 = 0.4$	$\omega = 0.5056$
$\rho_1 = 0.1$	$A_2 = 5$	$\rho_1^0 = 0.4$	$\rho_1 = 0.1266$
$\rho_2 = 0.9$	5% error	$\rho_2^0 = 0.4$	$\rho_2 = 0.9029$
$\omega = 0.5$	$A_1 = C_1 = 1$	$\omega^0 = 0.4$	$\omega = 0.4995$
$\rho_1 = 0.1$	$A_2 = 5$	$\rho_1^0 = 0.4$	$\rho_1 = 0.0883$
$\rho_2 = 0.9$	8% error	$\rho_2^0 = 0.4$	$\rho_2 = 0.8982$
$\omega = 0.9$	$A_1 = C_1 = 10$	$\omega^0 = 0.4$	$\omega = 0.8984$
$\rho_1 = 0.1$	$A_2 = 5$	$\rho_1^0 = 0.4$	$\rho_1 = 0.0959$
$\rho_2 = 0.9$	0.5% error	$\rho_2^0 = 0.4$	$\rho_2 = 0.8994$
$\omega = 0.9$	$A_1 = C_1 = 10$	$\omega^0 = 0.4$	$\omega = 0.8944$
$\rho_1 = 0.1$	$A_2 = 5$	$\rho_1^0 = 0.4$	$\rho_1 = 0.0506$
$\rho_2 = 0.9$	1% error	$\rho_2^0 = 0.4$	$\rho_2 = 0.8943$

REFERENCES

1. K. Kamiuto, Application of the emerging-intensity fitting method for inverse scattering problems to a system bounded by transparent plates, *J. Quant. Spectrosc. Radiat. Transfer*, **46**, 159-164 (1991).

2. A. J. Silva Neto and M. N. Özisik, An inverse problem of estimating thermal conductivity, optical thickness, and single scattering albedo of semi-transparent medium, *Proc. 1st International Conference on Inverse Problems in Engineering: Theory and Practice*, ASME, Florida, USA, 267-273 (1993).

3. A. J. Silva Neto and M. N. Özisik, An inverse problem of simultaneous estimation of radiation phase function, albedo and optical thickness, *J. Quant. Spectrosc. Radiat. Transfer*, **53**, 397-409 (1995).

4. A. J. Silva Neto and F. J. C. P. Soeiro, A combination of gradient based and global optimization methods for the solution of inverse heat transfer problems, *Inverse Problems in Engineering*, to appear (2002).

5. C. E. Siewert, On the inverse problem for a three-term phase function, *J. Quant. Spectrosc. Radiat. Transfer*, **22**, 441-446 (1979).

6. N. J. McCormick, Unified approach to analytical solutions of three inverse transport problems, *Progress in Nuclear Energy*, **34**, 425-430 (1999).

7. A. J. Silva Neto and N. J. McCormick, An explicit formulation for the inverse radiative transfer problem of single scattering albedo and diffuse reflectivities estimation, *Inverse Problems in Engineering Symposium*, Texas, USA (2001).

8. A. T. Kauati, A. J. Silva Neto and N. C. Roberty, A source-detector methodology for the construction and solution of the one-dimensional inverse transporte equation, *Inverse Problems in Engineering*, **9**, 45-66 (2001).

9. M. N. Özisik, *Radiative Transfer and Interactions with Conduction and Convection*, Wiley (1973).

10. C. E. Siewert, On a possible experiment to evaluate the one-speed or constant cross section model of the neutron-transport equation, *J. Math. Phys.*, **19**, 1587-1588 (1978).

11. N. J. McCormick, Transport scattering coefficients from reflection and transmission measurements, *J. Math. Phys.*, **20**, 1504-1507 (1979).

12. C. E. Siewert, Inverse solutions to radiative-transfer problems with partially transparent boundaries and diffuse reflection, *J. Quant. Spectrosc. Radiat. Transfer*, **72**, 299-313 (2002).

13. A. J. Silva Neto, Explicit and implicit formulations for inverse radiative transfer problems, *Fifth World Congress on Computational Mechanics, Minisymposium 125 - Computational Treatment of Inverse Problems in Mechanics*, Vienna, Austria (2002).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

MULTIOBJECTIVE PARAMETER ESTIMATION PROBLEMS: AN IMPROVED STRATEGY

Cláudia M. Silva, Domingos F. S. Souza, Evaristo C. Biscaia Jr.
Department of Chemical Engineering, PEQ/COPPE
Federal University of Rio de Janeiro, UFRJ - Rio de Janeiro, RJ, Brazil
cmartins@peq.coppe.ufrj.br, evaristo@peq.coppe.ufrj.br

ABSTRACT

A multiobjective optimization approach has been applied to solve parameter estimation problems. An improved genetic algorithm has been proposed to optimize mathematical model parameters. The algorithm makes use of a new concept of fitness function, based on an extended ranking procedure and a new class of operators, which enhance the algorithm performance.

Two processes have been analyzed: a grain cooling process and a grain drying process. In order to estimate the coefficient of heat transfer and the drying rate parameters of these models, minimization of the sum of the least squares of temperature and equilibrium moisture content have been conducted. Experimental data obtained from the soybean cooling in a continuous cross-flow moving bed heat exchanger and the corn drying in a fixed bed dryer have been used to evaluate the estimated parameters. The simulated results demonstrated the algorithm efficiency to perform parameter estimation. The validated model consistently fits the experimental data.

NOMENCLATURE

Cp	specific heat (kJ/kgK)
Dp	particle diameter (m)
fm	drying rate
F	objective function
F_k	fitness function
G	mass velocity of drying gas (kg/m^2s)
h	heat transfer coefficient (kJ/m^3s)
H	specific enthalpy (kJ/kg)
N	number of experimental data
S	solid superficial area/bed volume (m^{-1})
t	time(s)
T	temperature (K)
UR	relative humidity
x	gas flow direction
X	dimensionless parameter
Y	moisture content of the grain (-)
z	bed height coordinate
Z	dimensionless parameter

Greek symbols

ε	void fraction in bed of solids
ϕ	particle shape factor
λ	latent heat of evaporation (kJ/kg)
ρ	density (kg/m^3)

Subscripts and Superscripts

exp	experimental
g	gas
gt	total gas
l	liquid
R	reference
s	solid
se	solid in equilibrium
v	vapor

INTRODUCTION

Accurate modeling is an important prerequisite for optimization, control and design of engineering processes. Mathematical models are set of expressions that simulate the behavior of a system, describing quantitatively the mechanisms of the process. The generality of a model depends upon the process complexity as well as the available information concerning the system. In order to validate the model, it is often necessary to estimate the model parameters of nonlinear algebraic or differential equations.

The parameter estimation procedure is based on the minimization of the total sum of every observed error. Such procedure is often conducted using optimization algorithms. Traditionally, the estimation methods are "hill-climbing" methods that converge to only one stationary point. The objective functions based upon a nonlinear model and experimental data frequently present more than one optimum. The hill-climbing algorithms are easily trapped in local optimum, and are

unable to exploit the existing optima when the eigenvalue ratio becomes large [1]. Heuristic optimization techniques have been successfully used to solve parameter estimation problems [1,2,3,4].

Parameter estimation constitutes a multiobjective optimization problem, as it entails opposing requirements to be satisfied simultaneously. As the targets are often in conflict, a compromise solution has to be sought. The estimation problem has been handled by single objective approaches, in which all targets are combined into a single objective function by using weighting factors. A meaningful optimization approach implies the use of the multiobjective concept, which generates a family of equally good solutions, called Pareto optimal set, instead of a global optimum. The choice of the optimal results requires additional knowledge of the process, in order to select the best compromise solution between the competing goals.

This contribution is focused on the development of a multiobjective optimization algorithm to solve parameter estimation problems. An improved approach based on the genetic algorithm technique is proposed. A new class of operators consisting of a Pareto-set filter, an elitism operator and a niche operator is introduced. These operators enhance the performance of the algorithm and are considerably simple to implement. A new ranking strategy is proposed to treat multidimensional problems. The new strategy does not require any previous knowledge of the relative importance of individual objectives. A fitness function based on each rank population size and rank level is also provided to determine the reproduction ratio.

In order to evaluate the algorithm performance, two different processes have been studied: a soybean grain cooling process and a corn grain drying process. The mathematical model of each process was employed to generate the dependent variable values under the same conditions of the experiments. The minimization of every discrepancy between the measured and the predicted variables guides the search for the best-fit parameters. The heat and moisture transfer coefficients of these processes are estimated by the temperature and equilibrium moisture content least squares. Experimental data obtained from a continuous cross-flow moving bed heat exchanger and a fixed bed dryer have been used to evaluate the estimated parameters.

The simulated results demonstrated the ability of the method to handle multiobjective optimization problems. The proposed algorithm can successfully seek tradeoff surface regions and find the Pareto optimal set. It has been shown that the validated model fits the experimental data satisfactorily, indicating to be a consistent strategy to solve parameter estimation problems.

MULTIOBJECTIVE GENETIC ALGORITHM

The genetic algorithm technique simulates a natural evolution process: the fittest species survive and propagate while the less successful tend to disappear. The GA optimization procedure consists of a search for nondominated solutions. The concept of nondominance refers to the solutions for which no objective can be improved without worsening at least one of the other objectives. The progress strategy is guided by the fitness evaluation, and consists of performing the population with genetic operators to generate the next population. Different adaptations of the original GA are presented in the literature [5,6,7]. A detailed background on the GA theory is reported in Goldberg [8] and Busacca *et al*. [9].

The Proposed Algorithm

The multiobjective optimization algorithm developed is an improved version of a Pareto genetic algorithm proposed by Cheng and Li [5]. The standard ranking procedure, based on the concept of nondominance, is extended to treat multidimensional problems. New operators are introduced to enhance the algorithm performance: (a) a niche operator, which prevents genetic drift and maintains a uniformly distributed population along the optimal set; (b) a Pareto-set filter, which avoids missing optimal points during the evolutionary process and (c) an elitism operator, which insures the propagation of the best result of each individual objective function. These operators reduce the necessary number of generations and are computationally feasible for even complicated problems [10]. A new concept of fitness function is adopted in order to provide the evolutionary process with robust and stable convergence. The computational code developed operates in a continuous variable space, which is computational fast and stable in converging to the Pareto optimal set.

The proposed GA procedure works through the following steps:

(a) creation of a random initial population;
(b) evaluation of the individuals;

(c) ranking of the individuals, calculation of the fitness, registration of the best individuals;
(d) registration of all nondominated individuals in the Pareto set filter operator;
(e) selection of pairs of individuals as parents;
(f) crossover of the parents to generate the children;
(g) replacement of the individuals using the niche operator;
(h) genetic mutation;
(i) replacement of the individuals using the elitism operator.

The improvements made to the original method are discussed in the next sections.

Ranking procedure. The ranking procedure consists of the classification of the individuals into categories according to the concept of Pareto set dominance. First, all nondominated individuals of the population are identified and assigned rank 1. These individuals are virtually removed from the population and a new evaluation is conducted on the remaining individuals. The next set of nondominated points are identified and assigned rank 2. This procedure continues until all the individuals are classified. Figure 1 illustrates the ranking procedure for a maximization problem.

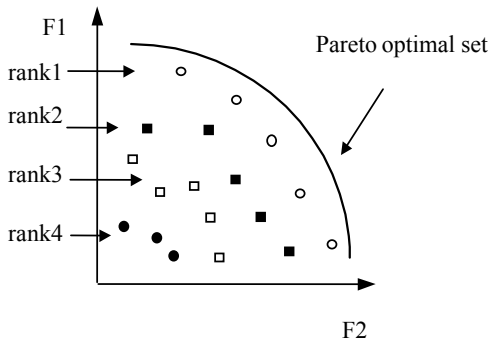


Figure 1 - Population ranking for a two-dimensional maximization problem

Mathematically, the classification method proposed by the authors is conducted in four basic steps:

(a) the points are sorted according to the evaluation of an objective function, randomly chosen as a reference function;
(b) all points that produce opposite effects on the function values are selected as potential candidates to the rank, i.e., the sequence of points that simultaneously produces an increase on the reference function and a decrease on the other functions, or vice versa, will be selected;
(c) repeat steps (a) and (b) until all the objective functions have been chosen as the reference function;
(d) all points selected at least (n-1) times in the step (b) will be assigned to the rank.

These steps are repeated for each rank, until all points are classified.

Fitness function. After being ranked, all individuals in the population are evaluated by a fitness function. The fitness value represents a measure of each individual performance, which will be used on the selection procedure. Every individual belonging to the same rank class is considered equivalent and has the same fitness value. Thus, it has the same probability of being selected for reproduction. The fitness function, F_k, is determined to each individual of the same rank k by the following equations, as proposed by Cheng and Li [5]:

$$SS = \frac{1}{P_s}\sum_{k=1}^{N_r}(N_r - k + 1)P_{sk} \qquad (1)$$

$$F_k = \frac{N_r - k + 1}{SS} \qquad (2)$$

where Nr is the highest rank of the population, P_s is the population size and P_{sk} is the population size of rank k. According to this fitness definition, the larger the population size at a rank is, the smaller the fitness of a point. Hence, the reproduction ratio of individuals at each rank depends on both the rank level and population size.

Niche operator. The niche technique consists of a replacement procedure of individuals. This operator has been applied to determine which individuals will go to the next population. The fitness of each child originated by the crossover procedure is calculated in the domain of its parental population. The replacement of the parents only occurs if the child fitness exceeds the parents' superior fitness. Otherwise, both parents will go to the next generation [5]. This operator helps to avoid the phenomenon known as genetic drift, which makes a population become clustered at certain regions.

The genetic drift results from stochastic errors associated with the genetic operators and the use of a population of finite size. Hence, this technique has been adopted with the purpose of maintaining the appropriate diversity of the population.

Elitism Operator. The GA's evolutionary process, as mentioned before, is guided by the ranking of the population. In the ranking procedure, points are categorized into groups according to their nondominance levels. The closer to the Pareto set a group of points is, the higher its probability to propagate to the next generation. The concept of nondominance refers to a characteristic behavior presented by a sequence of points sorted according to their objective function values. Hence, the classification of a group of points depends on their best objective function values.

The elitism procedure consists of the propagation of the best solutions to the next generation. For this purpose, it maintains an elitism file where the best solution of each individual objective function is registered. At each generation, this file is updated: if a better solution was generated it replaces the one stored. The individuals selected to the next population by the crossover and mutation procedures are submitted to the elitism operator. The points registered in the elitism file will randomly replace some of the individuals selected as candidates to the next generation. Since the elitism operator stores the best solutions obtained till the current iteration, it guarantees the propagation of best solutions during the evolutionary search procedure. As a consequence, it increases the convergence of the optimization process as well as the robustness of the algorithm.

Pareto-set filter operator. The Pareto set filter operator works during the whole optimization process. All points assigned rank 1 at each generation are registered in a Pareto set file. This file is dynamically updated by using a filter operator. In this filter, the nondominated solutions of the current population are compared with those already stored in the file, from the previous generations. A new evaluation is conducted in all points in the filter, according to the following rules:

(a) all points in the filter reassigned rank 1, i.e., points identified as nondominated, are recorded in the Pareto set file. The dominated ones are discarded;
(b) if the number of points in the file is inferior to the population size, the new nondominated points are stored. Otherwise, if the file is full, the most similar points in the Pareto set file are replaced.

The concept introduced to measure the similarity between two points in rule (b) is based on the euclidean distance between the values of the objective functions. Points with minimum distance relative to the others are removed from the file. Such procedure maintains an even distribution of the points in the file, which helps to provide the optimal set with the best solutions. At the end of the optimization process, the file itself comprises the Pareto optimal set and constitutes the result of the optimization.

The proposed algorithm is detailed in a previous work [10]. The computational load associated to the algorithm relies on the number of objective functions of the problem. At each generation, the objective functions are evaluated just once for each point. Therefore, the total number of the objective evaluations is a function of the size of the population, the number of generations and the number of objective functions optimized.

PARAMETER ESTIMATION PROBLEM

Grain Cooling Process

The first case study is the estimation of the overall heat transfer coefficient in a grain cooling process. The process is conducted in a continuous cross-flow moving bed heat exchanger, fed with warm soybean grains and dry air. The model formulation considers the convection on the particle surface as the dominant mechanism in the gas-solid heat transfer [11]. Some assumptions were made to simplify the mathematical formulation of the problem:

(a) gas and solid velocity profiles are linear;
(b) the temperature profiles in both phase feeding regions are considered linear;
(c) heat losses through the equipment walls are negligible;
(d) the heat transfer perpendicular to the flow is negligible;
(e) solid and gas flows are unidirectional;
(f) the physical and chemical properties are constant throughout the bed.

A hyperbolic system of differential equations are provided to describe the steady state heat transfer mechanism as follows:

Gas phase:

$$Cp_g G_g \frac{\partial T_g}{\partial x} + h\, S\,(T_g - T_s) = 0 \qquad (3)$$

Solid phase:

$$Cp_s G_s \frac{\partial T_s}{\partial z} + h\, S\,(T_g - T_s) = 0 \qquad (4)$$

where $S = \dfrac{6(1-\varepsilon)}{\phi Dp}$

subject to the initial conditions

$$T_g(0,Z) = T_{g0} \text{ and } T_s(X,0) = T_{s0} \qquad (5)$$

After integration, the equations (3) and (4) can be expressed in dimensionless form as:

$$v(X,Z) = e^{-Z} \int_0^X e^{-\xi} I_0\left(2\sqrt{Z\xi}\right) d\xi \qquad (6)$$

$$u(X,Z) = 1 - e^{-X} \int_0^Z e^{-\eta} I_0\left(2\sqrt{X\eta}\right) d\eta \qquad (7)$$

where I_0 is the modified Bessel function of the first kind of order zero, and

$$v = \frac{T_g - T_{g0}}{T_{s0} - T_{g0}} \qquad u = \frac{T_s - T_{g0}}{T_{s0} - T_{g0}} \qquad (8)$$

$$X = \frac{x\,h\,S}{Cp_g G_g} \qquad Z = \frac{z\,h\,S}{Cp_s G_s} \qquad (9)$$

The optimization problem involves the minimization of the least squares of the measured temperature and the one predicted by the model. The heat transfer coefficient is the decision variable. The parameter estimation problem is formulated as follows:

$$F_1 = \frac{1}{N}\sum_{k=1}^{N}\left(T_{g\exp k}(x,z) - T_{gk}(X,Z)\right)^2 \qquad (10)$$

$$F_2 = \frac{1}{N}\sum_{k=1}^{N}\left(T_{s\exp k}(x,z) - T_{sk}(X,Z)\right)^2 \qquad (11)$$

Experimental results obtained in a pilot scale heat exchanger have been used to estimate the heat transfer coefficient. The physical properties and parameters of the problem are presented in [11].

In order to investigate the multiobjective behavior of the formulated problem, the decision variable is plotted in Figure 2 as a function of the objective functions.

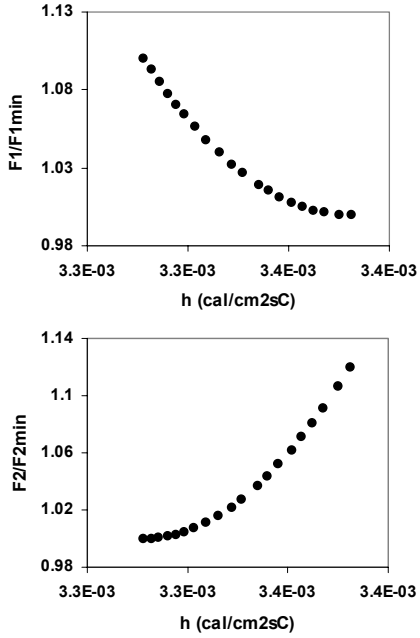


Figure 2 – Heat transfer coefficient of the grain cooling process as a function of the objective functions F_1 and F_2.

The objective functions are affected in opposing ways by changes in the decision variable. Therefore, the estimation parameter of the soybean cooling process constitutes a multiobjective optimization problem, and is suitable to evaluate the algorithm performance.

Grain Drying Process

The second problem deals with the estimation of the heat transfer coefficient and the drying rate parameters of a corn drying process, conducted in a fixed bed dryer. The drying model constitutes a set of differential equations describing the temperature and moisture content profiles in the gas and solid phases. An algebraic-differential approach is adopted to express the energy balances in term of the enthalpies. The hypotheses considered in the model formulation involve [12]:

(a) a uniform plug flow of air through the bed;
(b) adiabatic walls, the heat losses are negligible;
(c) the temperature gradients within the individual particles are negligible;
(d) the heat capacities are constant;
(e) the packing density of the grain in the bed is uniform;
(f) the evaporating water extracts the heat of desorption from the grain and enters the air stream as water vapor at the grain temperature.

The differential equations describing the heat and mass transfer mechanisms are presented:

Solid phase mass balance

$$(1-\varepsilon)\rho_s \frac{\partial Y_s}{\partial t} = -fm \tag{12}$$

Solid phase energy balance

$$(1-\varepsilon)\rho_s \frac{\partial H_{sl}}{\partial t} = h(T_g - T_s) - fm\langle H\rangle \tag{13}$$

$$H_{sl} = H_s + Y_s H_l \tag{14}$$

Gas phase mass balance

$$\varepsilon\rho_g \frac{\partial Y_{gt}}{\partial t} + G_g \frac{\partial Y_{gt}}{\partial z} = fm \tag{15}$$

Gas phase energy balance

$$\varepsilon\rho_g \frac{\partial H_{gv}}{\partial t} + G_g \frac{\partial H_{gv}}{\partial z} = -h(T_g - T_s) + fm\langle H\rangle \tag{16}$$

$$H_{gv} = H_g + Y_{gt} H_v \tag{17}$$

The enthalpies are related to the temperatures as:

$$H_s = Cp_s(T_s - T_R) \tag{18.a}$$

$$H_l = Cp_l(T_s - T_R) \tag{18.b}$$

$$H_g = Cp_g(T_g - T_R) \tag{18.c}$$

$$H_v = \lambda_{T_R} + Cp_v(T_g - T_R) \tag{18.d}$$

$$\langle H\rangle = \lambda_{T_R} + Cp_v(T_g - T_R) \tag{18.e}$$

The drying rate expression for corn grains is described as

$$fm = (b_1 Y_s - b_2)(Y_s - Y_{se}) \tag{19}$$

where b_1 and b_2 are the parameters to be estimated in order to determine the drying rate. The sorption isotherms, Y_{se}, for corn grains ($T_s \geq 40^0$C) are expressed by:

$$Y_{se} = \frac{1}{100}\sqrt{\frac{\ln(1-UR)}{d_1(d_2 T_g - d_3)}} \tag{20}$$

where d_1, d_2 and d_3 are parameters present in the literature [13].

The partial differential and algebraic equation system is discretized in the spatial domain into 50 points, using the implicit Euler form. The resulting system consists of 300 equations, which were integrated using the solver DASSL [14].

The formulation of the multiobjective parameter optimization problem comprises the minimization of the following objective functions:

$$F_1 = \frac{1}{N}\sum_{k=1}^{N}\left(\frac{T_{g\exp k}(t,z) - T_{gk}(t,z)}{T_{g\exp k}(t,z)}\right)^2 \tag{21}$$

$$F_2 = \frac{1}{N}\sum_{k=1}^{N}\left(\frac{Y_{s\exp k}(t,z) - Y_{sk}(t,z)}{Y_{s\exp k}(t,z)}\right)^2 \tag{22}$$

Three decision variables are selected: the heat transfer coefficient, h, and the drying rate parameters of equation (19), b_1 and b_2. The experimental data obtained by Calçada [13] in a soybean drying process were used to estimate the parameters above. The physical properties and parameters related to the process are presented in the same work.

RESULTS AND DISCUSSION

The proposed algorithm has been applied to estimate the model parameters for the two cases formulated. The optimizations have been carried out using the GA parameters summarized in Table 1, established according to a sensitivity study of a previous work [10].

Table 1. Genetic algorithm parameters

Population size	20 individuals
Crossover probability	75 %
Mutation probability	5 %
Number of children /crossover	1

Grain Cooling Process

Figure 3 shows the Pareto set obtained in the soybean cooling process. The number of

generations required to optimize the heat transfer coefficient was 32 iterations.

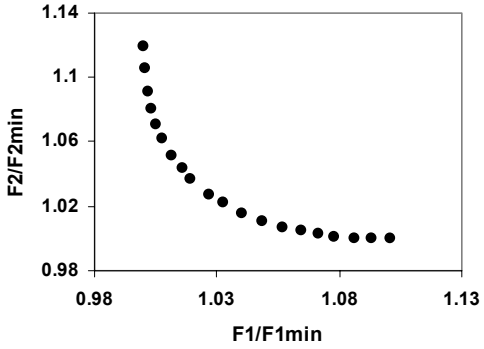


Figure 3 – Pareto set for the soybean cooling process.

The optimal values of the heat transfer coefficient corresponding to both extremes of the Pareto set were used to verify the parameter estimation. The simulations of the model illustrated in Figures 4 and 5 show that the estimated parameter satisfactory fits the experimental data.

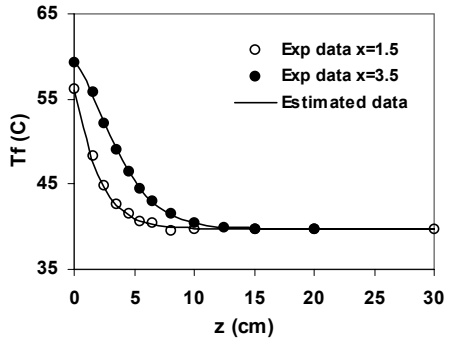


Figure 4 - Air temperature profiles as a function of the position in the heat exchanger.

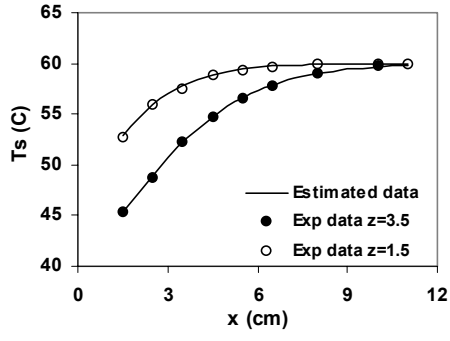


Figure 5 - Soybean temperature profiles as a function of the position in the heat exchanger.

Grain Drying Process

The Pareto set obtained for the corn drying process is depicted in Fig. 6. Sixty-four iterations were required to generate the solution set.

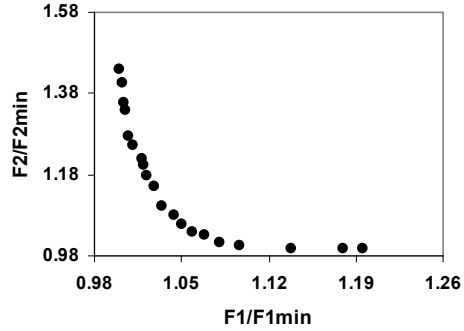


Figure 6 – Pareto set for the corn drying process.

Simulations of the model using the best set of results for each objective function are illustrated in Fig 7 and 8. The influence of the multiobjective behavior can be observed in both cases by the difference on the moisture content profiles.

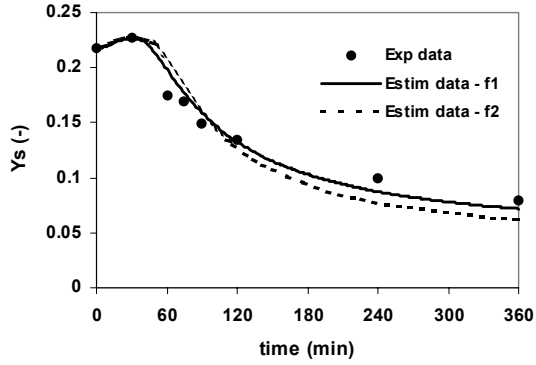


Figure 7 – Moisture content profile as a function of time using the best set of results for F_1 and F_2.

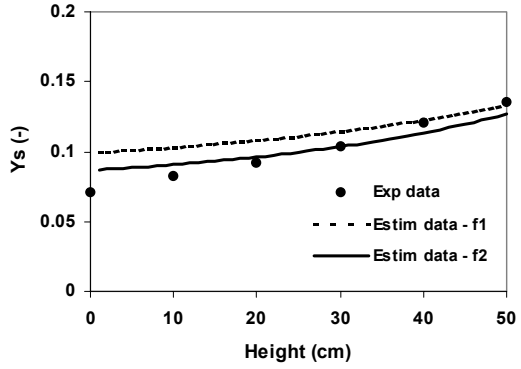


Figure 8 – Moisture content profile at different heights using the best set of results of F_1 and F_2.

In Figure 7, the parameter estimation based on F_1 provides the best description of the moisture content. In Figure 8, the moisture content profile is better simulated using the results based on F_2. The selection of the best compromise solution will depend on the judgment of the engineer, based on additional knowledge of the system.

For the sake of comparison, Figure 9 presents the estimated correlation and the ones proposed in the literature, by Lewis [15] and Mancini [16]. The estimated results are the ones that best fit the experimental data in all tested cases.

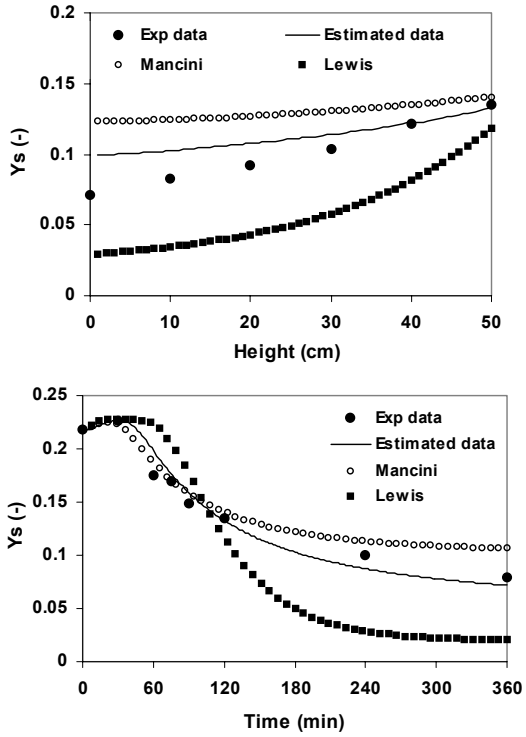


Figure 9 – Corn moisture profiles as a function of the height (t = 120 min), and the time (z = 50 cm).

CONCLUSIONS

Our focus in this contribution has been to evaluate the performance of the proposed algorithm in conducting multiobjective parameter estimation. The algorithm has been applied to a grain cooling process and a grain drying process in order to determine the heat and moisture transfer coefficients.

The model was validated against the experimental data. Simulations were carried out to verify the estimated parameters. The model predictions have been shown to match the measured values within experimental error.

ACKNOWLEDGEMENTS

We would like to thank the CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico – for the scholarship provided.

REFERENCES

[1] Z. S. H. Chan, H. W. Ngan, Y.F. Fung and A. B. Rad, An Advanced Evolutionary Algorithm for Parameter Estimation of the Discrete Kalman Filter, *Comp. Phys. Comm.*, **142**, 248 (2001).

[2] J. Nougués, M.Grau and L.Puigjaner, Parameter Estimation with Genetic Algorithm in Control of Fed-Batch Reactors, *Chem. Engng. Proc.*, **41**, 303 (2002).

[3] F. S. Wang and J. W. Sheu, Multiobjective Parameter Estimation Problems of Fermentation Processes Using a High Ethanol Tolerance Yeast, *Chem. Engng. Sci.*, **55**, 3685 (2000).

[4] T. Park and G. Froment, A Hybrid Genetic Algorithm for the Estimation of Parameters in Detailed Kinetic Models, *Comp. Chem. Engng*, **22**, S103 (1998).

[5] F. Y. Cheng and D. Li,. Genetic Algorithm Development for Multiobjective Optimization of Structures, *AIAA Journal*, **36**, 6, 1105 (1998).

[6] V.G Toshinsky, H. Sekimoto and G. Toshinsky, A Method to Improve Multiobjective Genetic Algorithm optimization of a Self Fuel Providing LMFBR by Niche Induction among Nondominated Solutions, *Proc. Nucl. Energy*, **27**, 397 (2000).

[7] K.Wang, Y. Qian, Y. Yuan and P. Yao, Synthesis and Optimization of Heat Integrated Distillation Systems using an Improved Genetic Algorithm, *Comp. Chem. Engng.*, **23**, 125 (1998).

[8] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading MA, 1989.

[9] P.G. Busacca, M. Marseguerra and E. Zio, Multiobjective Optimization by Genetic Algorithms: Application to Safety Systems, *Reliab. Engng. Syst. Saf.*, **72**, 59 (2001).

[10] C. M. Silva and E. C. Biscaia, Genetic Algorithm Development for Multiobjective Optimization of Polymerization Reactors, submitted to *Comp. Chem. Engng*.

[11] D. J. M. Sartori, *Transferência de Calor em Leito Deslizante*, PhD Thesis, COPPE/UFRJ, Rio de Janeiro, 1986.

[12] Y. Sun, C.Pantelides, Z.Chalabi, Mathematical Modeling and Simulation of Near-ambient Grain Drying, *Comp. Eletr. Agric.*, **13**, 243 (1995).

[13] L. A. Calçada, *Secagem de Materiais Granulares Porosos*, PhD Thesis, COPPE/UFRJ, Rio de Janeiro, 1998.

[14] L.R. Petzold, *DASSL: A Differential-Algebraic System Solver*, Lawrence Livermore National Laboratory, Livermore, CA, USA (1989).

[15] W. K. Lewis, The Rate of Drying of Solid Materials, *Ind. Engng. Chem.*, **13**, 427 (1921).

[16] M. C. Mancini, *Transferência de Massa em Secadores de Grãos*, PhD Thesis, COPPE/UFRJ, Rio de Janeiro, 1996.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

ON-SITE CALIBRATION OF A PHASE FRACTION METER BY AN INVERSE TECHNIQUE

Vanessa P. Rolnik
Núcleo de Engenharia Térmica e Fluidos
Escola de Engenharia de São Carlos
Universidade de São Paulo
Av. Trabalhador Sãocarlense, 400
13566-590 São Carlos - SP, Brasil
vanessa@sc.usp.br

Paulo Seleghim Jr.
Núcleo de Engenharia Térmica e Fluidos
Escola de Engenharia de São Carlos
Universidade de São Paulo
Av. Trabalhador Sãocarlense, 400
13566-590 São Carlos - SP, Brasil
seleghim@sc.usp.br

ABSTRACT

The formal calibration procedure of a phase fraction meter is based on registering the outputs resulting from imposed phase fractions at known flow regimes. This can be straightforwardly done in laboratory conditions, but is rarely the case in industrial conditions, and particularly for on-site applications. Thus, there is a clear need for less restrictive calibration methods regarding to the prior knowledge of the complete set of inlet conditions. A new procedure is proposed in this work for the on-site construction of the calibration curve from total flow mass values of the homogeneous dispersed phase. The solution is obtained by minimizing a convenient error functional, assembled with data from redundant tests to handle the intrinsic ill-conditioned nature of the problem. Numerical simulations performed for increasing error levels demonstrate that acceptable calibration curves can be reconstructed, even from total mass measured within a precision of up to 2%. Consequently, the method can readily be applied, especially in on-site calibration problems in which classical procedures fail due to the impossibility of having a strict control of all the input/output parameters.

NOMENCLATURE

α	volumetric solid fraction
$\dot{m}_s$	solids (particulate) mass flow rate
Q_s	solids volumetric flow rate
ρ_s	solids density
C	inter-electrode capacitance
M_s	total solids mass
$\phi(\cdot)$	calibration curve to be reconstructed
a_i	expansion coefficients of $\phi(\cdot)$
e	error functional

INTRODUCTION

The continuous measurement of physical parameters in multiphase flows is of great interest, not only for the monitoring and control of industrial equipment, but also to obtain phenomenological insight in research applications. In particular, the development of phase fraction sensors is a subject very frequently found in the related scientific literature, considering that this parameter is one of the most adequate when describing and analyzing two-phase flows.

A good illustration of this is the determination of composition and flow rates in the petroleum industry (Fischer, 1994). In the absence of an acceptable on-line measurement device, the most common procedure is to adopt a strategy based on the continuous separation of the fluid constituents and subsequent measurement by conventional single-phase techniques followed by recombination. Despite the extensive number of applications based on this approach, its efficiency and economic feasibility are not satisfactory. The large size of conventional separation equipment, such as hydro-cyclones and other centrifugal separators, is frequently restrictive and even prohibitive in offshore production systems for instance. In addition to this, in some situations, special thermal or chemical methods are required to deal with emulsion formation. Also, the flashing of dissolved gases from the liquid phase requires a more elaborate temperature and pressure control system implying in more complex operation conditions (Rajan *et al.*, 1993). Thus, the development of simple, robust and inexpensive non-intrusive multiphase flow sensors suited for industrial applications is still an open problem.

In this context, electrical sensing techniques are particularly attractive due to its capability of resolving fast changes in the flow structure,

besides being simple to implement and not expensive as well. Its general principle of operation is based on differences or contrasts in the electrical properties of the phases of the mixture, and also on the assumption that the electromagnetic sensing field is instantaneously modulated by the flow (Chang and Watson, 1994). In general, problems associated with electrical sensing are related to electrochemical effects (Hemp, 1994), electrostatic stray charges (Green and Thorn, 1998), nonlinearly due to the influence of flow regimes (Andreusi *et al.,* 1988), and electrode erosion or coating. One way of overcoming these problems is to design the sensor and the sensing strategy specifically for the desired application, therefore the need for on-site calibration methods (Duncan and Trabold, 1997).

The formal calibration procedure of a measurement device is based on the construction of the output/input relation by imposing known inputs and registering the corresponding outputs. In the case of a phase fraction probe this is not sufficient because, in addition to being correlated to the phase fraction, the probe's response will be also strongly correlated to the flow regime. In other words, the same phase fraction may result in different responses depending on the topological organization of the constituent phases within the sensing volume (see for instance the work of Andreussi *et al.*, 1988). Attempts have been made aiming to minimize the influence of the flow regime, mostly relied on the optimization of the electrodes (Seleghim and Hervieu, 1998) or on the sensing strategy (Klug and Mayinger, 1994). A very promising approach is based on fuzzy and neural signal processing techniques and is implemented so to previously identify the flow regime and subsequently take the correct calibration curve (Tsoukalas *et al.* 1997; Mi *et al.* 1998; Crivelaro and Seleghim, 1999).

Thus, in a strict sense, the calibration of a phase fraction meter requires the ability to impose known phase fractions at known flow regimes, which can be straightforwardly done in laboratory conditions. In industrial conditions this is rarely the case. For instance, if it is necessary to calibrate a fraction meter placed on a pneumatic conveying line, the control of the flow conditions would require the addition of auxiliary equipment that would produce significant disturbances in the operating conditions and, consequently, compromising the final result. Still, the majority of industrial scale pneumatic conveyors are designed to operate at nonpermanent flow conditions, which may take the form of discrete structures such as solid plugs and rolling dunes, or alternating flow regimes associated with varying inlet mass flow rates. This justifies the need for less restrictive calibration methods, in particular with regard to the prior knowledge of the complete set of inlet conditions.

The purpose of this paper is to contribute in this direction by proposing a new inverse procedure for the on-site construction of the calibration curve of a phase fraction meter from less restrictive data, i.e. not the instantaneous flow rate signal but its integral value. The problem will be precisely stated in section 2 and a numerical simulation will be presented in section 3 (in which the consequences and a solution for the problems associated with the inverse nature of the formulation will be shown). A final conclusion and the references are presented respectively in section 4 and 5.

STATEMENT OF THE PROBLEM

Consider the homogeneous flow of a two-phase mixture trough a capacitive fraction meter installed on a light phase pneumatic transport system (Fig. 1). The volumetric solid fraction (α) is defined as the ratio between the volumes occupied by the solids (V_s) and the total sensing volume (V_s+V_a), which, according to the one-dimensional one-velocity model (Bergels *et al.*, 1981), can be written as

$$\alpha = \frac{V_s}{V_a + V_s} = \frac{Q_s}{Q_a + Q_s} = \frac{\dot{m}_s}{\rho_s Q_a + \dot{m}_s} \qquad (1)$$

in which the volumetric flow rates (Q_s and Q_a), solid density ρ_s and mass flow rate $\dot{m}_s$ is defined in Fig. 1.

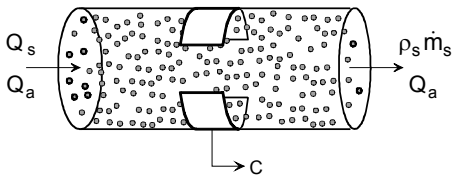


Figure 1: Homogeneous flow of a gas-solid mixture trough a capacitive fraction meter

The capacitance (C) measured between the sensing electrodes depends on how the electrical field traverses the sensing volume, which is

related to the permittivity of the medium and also to the geometric organization of the different phases. Thus, the capacitance is strongly correlated to the solid fraction and to the flow regime (homogeneous by assumption), and the formal relation between these variables, i.e.

$$\alpha = \phi_{flow}(C) \tag{2}$$

is known as the meter's calibration curve as mentioned before. Substituting equation (2) into (1) and isolating the solid mass flow rate yields:

$$\dot{m}_s = \frac{\rho_s Q_a}{\frac{1}{\phi(C)} - 1} \tag{3}$$

In an industrial pneumatic conveying system the instantaneous values of the air volumetric flow rate (the continuous phase) can be readily determined, for instance with an orifice plate or simply from the blower's performance curve by measuring its rotation and pressure rise. However, the measurement of the instantaneous values of the solid flow rate is quite complex without modifying the piping to install auxiliary equipment. In addition, the solid flow rate must be measured at the fraction meter's section since it can vary significantly along the transport line as well as in time (Ostrowski *et al.*, 1999). A more convenient variable to measure would be the total solid mass (M_s) conveyed in a given time interval:

$$M_s = \int_{\Delta t} \dot{m}_s \, dt \tag{4}$$

This expression can be obtained from the integration of equation (3), with the additional assumption (for simplicity and without loss of generality) of a constant volumetric air flow rate. It will then result

$$\frac{M_s}{\rho_s Q_a} = \int_{\Delta t} \frac{1}{\frac{1}{\phi(C)} - 1} dt \tag{5}$$

To obtain the calibration curve $\alpha = \phi(C)$ it is necessary to solve the integral equation (5) on the input data M_s, ρ_s, Q_a and $C = C(t)$, the instantaneous capacitance values delivered by the fraction meter. Equation (5) can also be seen as a special case of an inhomogeneous Fredholm equation of the first kind, which is known to be ill conditioned. Consequently, as it will be shown on the sequel, if specific methods to deal with the ill-conditioned nature of the problem are not employed, the calibration curve will be extremely sensitive to small changes in the input parameters, which in fact are likely to happen due to intrinsic experimental errors. This is so because the integration of the unknown function in (5) causes an information loss. A proper method to deal with this must, in some way, restore the lost information from some prior knowledge or from redundant measurements.

NUMERICAL SIMULATION AND MANAGING THE ILL-CONDITIONED NATURE OF THE INVERSE PROBLEM

In order to demonstrate the statements above consider the following numerical simulations. First, suppose that the calibration curve is given by the following representative formula (taken from previous laboratory experiment – Hervieu, 1999):

$$\alpha = \phi(C) = 0.06 - \sqrt{0.0036 - 0.12\,C} \tag{6}$$

This equation is not known a priori and will have to be reconstructed from experimental data. Suppose now that, due to a specific operating condition, the instantaneous measured solids fraction values follow the equation

$$\alpha(t) = 0.025\left[1 + \sin(\pi t)\right] \tag{7}$$

This being, the substitution of (7) into (6) yields the instantaneous capacitance values delivered by the probe:

$$C(t) = \left[1 + \sin(\pi t)\right]\left[0.025 - 0.0052\left[1 + \sin(\pi t)\right]\right] \tag{8}$$

The resulting instantaneous solids mass flow rate for homogeneous flow can be determined by introducing (7) into expression (3), which yields

$$\dot{m}_s = \frac{\rho_s Q_a}{\left(0.025\left[1 + \sin(\pi t)\right]\right)^{-1} - 1} \tag{9}$$

The following figure illustrates the behavior of these curves for $\rho_s = 3000\ kg/m^3$ and $Q_a = 0.01\ m^3/s$

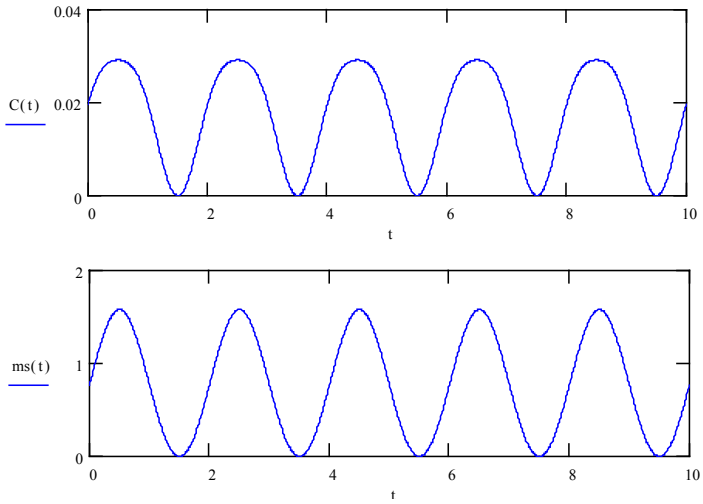


Figure 3: Capacitance in pF (equation (8)) and mass flow rate in kg/s (equation (9)) as a function of time in sec, resulting from (6) and (7) in an homogeneous gas-solid flow.

The inverse problem, such as formulated above, consists in reconstructing the calibration curve (2) (which was imposed to follow (6)) from measurements of the total mass M_s and the instantaneous capacitance values (8). To do this we can start by expressing (2) according to

$$\alpha = \phi(C) \cong \sum_{0}^{N} a_i \varphi_i(C) \qquad (10)$$

where $\{\varphi_i(C)\}$ is a convenient set of known functions. Subsequently, substituting (8) into (10) and the result in (5) we obtain a formal relation with the capacitance signal, the approximated calibration curve and the measured total mass, that is (equation (11)):

$$\frac{M_s}{\rho_s Q_a} \cong \int_{\Delta t} \left\{ \left[\sum_{0}^{N} a_i \varphi_i \left([1+\sin(\pi t)] \times \right. \right. \right.$$
$$\left. \left. \left. \times [0.025 - 0.0052(1+\sin(\pi t))] \right) \right]^{-1} - 1 \right\}^{-1} dt \qquad (11)$$

The difference between both sides of the expression above constitutes an error functional expressing how well is the approximation given by (10). The influence of experimental errors in the measurements can be introduced, for instance, by randomly perturbing M_s. and C. We thus define (equation (12)):

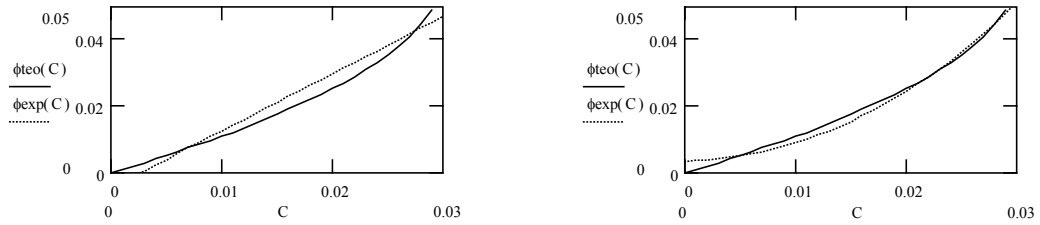
$$e = \left(\frac{M_s(1+\delta)}{\rho_s Q_a} - \int_{\Delta t} \left\{ \left[\sum_{0}^{N} a_i \varphi_i \left([1+\sin(\pi t)][0.025 + \right. \right. \right.$$
$$\left. \left. \left. - 0.0052[1+\sin(\pi t)]] + \varepsilon(t) \right) \right]^{-1} - 1 \right\}^{-1} dt \right)^2 \qquad (12)$$

where δ and $\varepsilon(t)$ are centered uniform random variables. The reconstruction of the calibration curve can now be achieved by searching the minimum of equation (12).

The steepest descent method is employed in this work, with increments calculated from an arbitrated step Δa, according to the following rule:

$$\Delta a \left(\frac{\partial e}{\partial a_0} \right)^{-1} = \ldots = \Delta a \left(\frac{\partial e}{\partial a_N} \right)^{-1} \qquad (13)$$

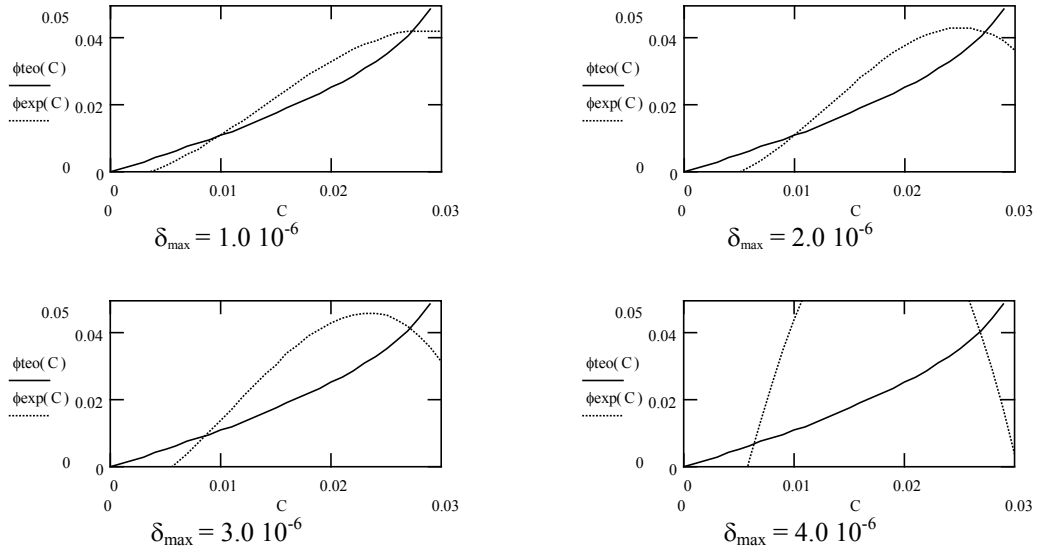
Consider $\delta = \varepsilon(t) = 0$ to illustrate the ill-posed nature of the problem. In this case, the minimum of (12) can be calculated by solving equations (13) directly, which produces progressively better approximations as the polynomial order N in (10) is increased. This is as shown in Figure 4 for a linear and a third order approximation.



4.a) $\phi_{exp}(C) = a_0+a_1C$ 4.b) $\phi_{exp}(C) = a_0+a_1C+a_2C^2+a_3C^3$

Figure 4: Reconstruction of the calibration curve from simulated experimental data b searching the minimum of (12) for increasing approximation order in (10)

As it can be seen, the reconstruction works adequately under the assumption of no experimental errors. In a real situation however, these errors must be taken into account and, due to the inverse nature of the problem, this will produce an extremely negative affect. This can be illustrated by performing the same reconstruction as shown in Figure 4.b (third order approximation) and considering only a centered experimental error in the measurement of M_s, i.e. $\varepsilon(t) = 0$ while δ varies randomly between $\pm\delta_{max}$. The following figure shows the influence of increasing values of δ_{max} in the reconstructed calibration curve. Although neglecting the error in the measurement of the instantaneous capacitance values, even unrealistic experimental errors the order of $\delta_{max} = 10^{-6}$ have a disastrous effect.



$\delta_{max} = 1.0\ 10^{-6}$ $\delta_{max} = 2.0\ 10^{-6}$

$\delta_{max} = 3.0\ 10^{-6}$ $\delta_{max} = 4.0\ 10^{-6}$

Figure 5: The influence of experimental errors in the reconstruction of the calibration curve (third order polynomial) from the minimum of (12).

As mentioned before, this problem is probably triggered by the integration of the unknown function $\phi(C)$ in (5) rendering the problem extremely ill conditioned. To deal with this there are some mathematical methods, mostly based on the construction of regularizing operators from a priori information. Another way of overcoming the problem, which in fact is more natural in the case we are dealing with, consists in reintroducing the lost information by redundant measurements.

Consider then a set of measurements of the total mass $M_s^{(k)}$ and the corresponding

capacitance historic $C^{(k)}(t)$, performed over different time intervals $\Delta t^{(k)}$ ($k = 1,2,\ldots M$). For the purposes of this numerical simulation, these data can be generated by randomly varying δ and $\varepsilon(t)$, i.e.

$$M_s^{(k)} = M_s\left(1+\delta^{(k)}\right) \tag{14}$$

$$C^{(k)}(t) = \varepsilon^{(k)}(t) + \left[1+\sin(\pi t)\right]\times \\ \times\left[0.025 - 0.0052\left[1+\sin(\pi t)\right]\right] \tag{15}$$

in which

$$\delta^{(k)} \le \left|\delta_{max}\right| \quad \text{and} \quad \varepsilon^{(k)}(t) \le \left|\varepsilon_{max}\right| \tag{16}$$

Within these definitions, the error associated with each measurement can be quantified by the following

$$e = \left(\frac{M_s(1+\delta^{(k)})}{\rho_s Q_a} - \int_{\Delta t^{(k)}} \left\{ \left[\sum_0^N a_i \varphi_i \left([1+\sin(\pi t)][0.025 + \right.\right.\right. \\ \left.\left.\left. - 0.0052[1+\sin(\pi t)]] + \varepsilon^{(k)}(t) \right) \right]^{-1} - 1 \right\}^{-1} dt \right)^2 \tag{17}$$

and a global error function can be defined by calculating the Euclidean norm of $\{e^{(k)}\}$:

$$E = E(a_0, a_1, \ldots a_N) = \sqrt{\sum_{k=1}^{M} e^{(k)}} \tag{18}$$

The problem can now be solved by searching for the minimum of (18) instead of (12). Figure 6 shows the results for increasing error in the measurement of the total mass (δ_{max}) and time intervals varying randomly between 10 and 30s according to a uniform law. One can readily conclude that the admissible error have been significantly increased.

To illustrate how the ill-conditioned nature of the problem manifests itself at the formulation proposed here some numerical simulations were carried out with a reduced number of parameters. The graphic in Fig. 7 shows (18) with N = 2, i.e. $E = E(a_0, a_1)$, to

allow visualization, and M = 100 (that is 100 evaluations of (17) with different values of δ). It is possible to observe that there exists a pronounced minimum at the point (0.0002, 1.4489) which corresponds to the calibration curve

$$\alpha = \phi(C) = 0.0002 + 1.4489\,C \tag{19}$$

Equation (19) corresponds to the projection of the actual calibration curve (6) onto the subspace of all possible linear calibration curves. It is also important to stress that the topology of the (18) have particular features which will induce severe difficulties regarding the convergence of the minimization procedure (a tradeoff between setting very small iteration steps and intense oscillatory behavior of the global error function).

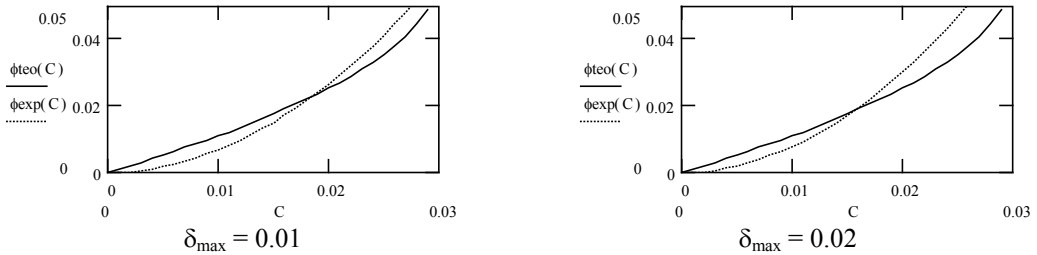


Figure 6: The influence of experimental errors in the reconstruction of the calibration curve (third order polynomial) from the minimization of (18)

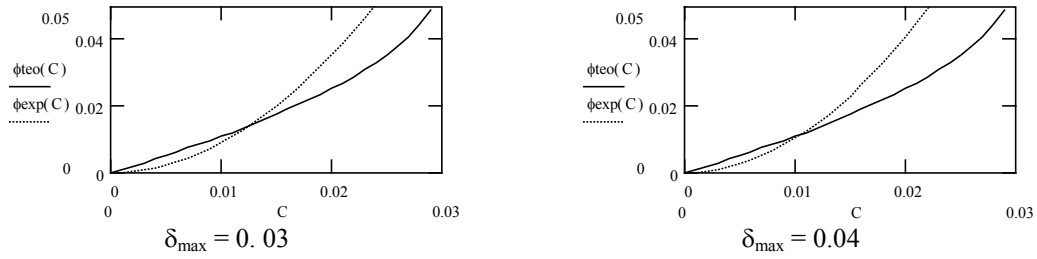


Figure 6: (continued)

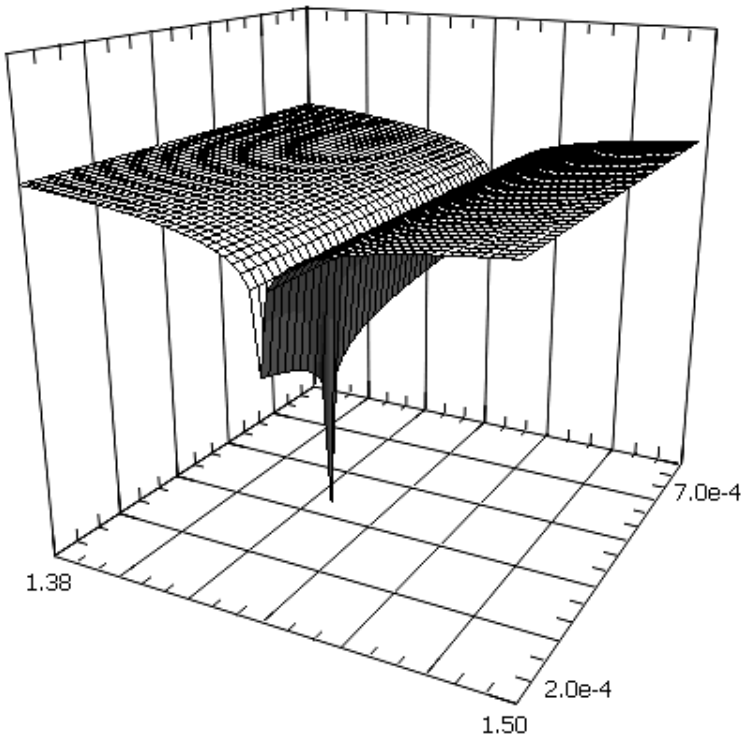


Figure 7: Error surface constructed for a two-parameter (linear) approximation of the calibration curve

CONCLUSION

A new inverse procedure is proposed in this work for the on-site reconstruction of the calibration curve from total flow mass values of the homogeneous dispersed phase. The problem is formulated as an intrinsically ill conditioned integral equation solved by setting an appropriate approximation for the calibration curve (equation (10), figure 4). Under these circumstances, the method cannot be applied in practical situations due to the severe influence of very small experimental errors (illustrated in figure 5).

The proposed solution to this is based on minimizing an error functional constructed from a set of redundant measurements, which restores the lost information associated to the integration of the instantaneous mass flow rate in the one-dimensional one-velocity flow model in equation (5). Numerical simulations performed for increasing errors demonstrate that acceptable calibration curves can be reconstructed, even from total mass measurements within a precision of up to 2%. Thus, the method can readily be applied, especially in on-site calibration problems

in which classical procedures fail due to the impossibility of having a strict control of all the input/output parameters.

REFERENCES

1. P. Andreussi, DI Donfrancesco, A. & Messia, M., An impedance method for the measurement of liquid hold-up in two-phase flow, International Journal on Multiphase Flow, Vol. 14, No. 6, pp.777-785, 1988.

2. A.E. Bergles, Collier J.G., Delhaye J.M., Hewitt G.F. and Mayinger F., Two-phase flow and heat transfer in the power and process industries. Hemisphere Publishing Corporation, 707p, 1981.

3. J.S. Chang and Watson A, Electro-magnetic hydrodynamics. IEEE Transactions on Dielectrics and Electrical Insulation, Vol.1, No.15, pp.871-895, 1994.

4. O. K. Crivelaro and Seleghim Jr. P, Identifiação de regimes de escoamento bifásico horizontal com auxílio de redes neurais. XV Congresso Brasileiro de Engenharia Mecânica, Águas de Lindoia - SP, 22-26 novembro, 1999.

5. D. Duncan and Trabold D.A., Cross-duct electric field perturbation void fraction probes. International Communications in Heat and Mass Transfer, Vol. 24, No.2, pp.285-294, 1997.

6. C. Fischer, Development of a metering system for total mass flow and composition measurement of multiphase/multicomponent flows such as oil/water/air mixtures. Flow Measurement and Instrumentation., Vol. 5, No. 1, pp.31-42, 1994.

7. R.G. Green and Thorn R., Sensor systems for lightly loaded pneumatic conveyors. Powder Technology, Vol. 95, No. 1, pp. 79-92, 1998.

8. J. Hemp, Error in potential measurements due to nonuniform contact impedance of electrodes. Quarterly Journal of Mechanics and Applied Mathematics, Vol. 47, Part 1, pp. 175-182, 1994.

9. E. Hervieu, Calibration of a segmented ring electrode void fraction meter under different flow regimes. Private communication, Comissariat à l'Energie Atomique – Grenoble – France, 1999.

10. F. Klug and Mayinger F., Impedance Based Flow Reconstruction - A Novel Flow Composition Measuring Technique for Multi-Phase-Flows. Nuclear Engineering and Design, Vol. 146, No.1-3, pp.35-42, 1994.

11. Y. Mi, Ishii M. and Tsoukalas L..H., Vertical Two-Phase Flow Identification Using Advanced Instrumentation and Neural Networks. Nuclear Engineering and Design, Vol. 184, No. 2-3, pp.409-420, 1998.

12. K. Ostrowski, Luke S.P., Bennett M.A. and Williams R.A., Real time visualization and analysis of dense phase powder conveying. Powder Technology, Vol.102, pp.1-13, 1999.

13. V.S.V. Rajan, Ridley R. K. and Rafa K. G., Multiphase flow measurement - A review. J. Energy Research, Vol. 115, No. 3, pp.115-161, 1993.

14. P. Seleghim Jr. and Hervieu, E, Direct imaging of horizontal gas-liquid flows. Measurement Science & Technology, Vol.9 No.8 (1998), pp.1492-1500, 1998.

15. L.H. Tsoukalas, Ishii M. and Mi Y., A Neurofuzzy Methodology for Impedance-Based Multiphase Flow Identification. Engineering Applications of Artificial Intelligence, Vol.10, No.6, pp.545-555, 1997.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support provided by FAPESP and CNPq through grants 98/12921-1, 99/02821-2 and 520.723/97-0.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSE GEOMETRY PROBLEM OF ESTIMATING THE PHASE FRONT MOTION OF ICE IN A THERMAL STORAGE SYSTEM

Dong-Seong Kwag, In-Son Park
Graduate School of Mechanical Engineering, Hanyang University, Seoul, Korea,
mmkkds@templab.hanyang.ac.kr

Woo-Seung Kim
Department of Mechanical Engineering, Hanyang University, Ansan, Kyunggi-do, Korea,
wskim@hanyang.ac.kr

ABSTRACT

The moving phase change interface in the latent heat energy storage system is estimated by applying a two-dimensional inverse geometry problem. The energy storage system of the vertical tube type is considered for the present inverse geometry problem. To solve the phase change problem, the boundary element method is adopted. The moving phase change interface is estimated by using the conjugate gradient method. Estimation of the phase front motion is verified by conducting the inverse analysis for an assumed phase front motion. An inverse analysis for the desired temperature distributions is executed to investigate the possibility of desired front motion control or ice monitoring. The effects of the noise levels and the thermocouple spacing on the inverse solutions are also examined.

NOMENCLATURE

C_P	constant-pressure specific heat
J	functional defined by Eq. (5)
J'	gradient of functional
k	thermal conductivity
La	latent heat
M	total number of thermocouples
T	dimensionless temperature
ΔT	sensitivity function
Ste	Stefan number
Y	dimensionless measured temperature

Greek Symbols

b	search step size
g	conjugate coefficient
Ω	computational domain
Γ	boundary of computational domain
l	Lagrange multiplier of adjoint problem
w	random variable in measured data
s	standard deviation of the measurement errors

Subscripts

L	liquid region
m	location of thermocouple
S	solid region
∞	working fluid

Superscripts

$\hat{}$	estimated variables
n	number of iteration
$*$	Green function

INTRODUCTION

By using inverse analysis, engineers can obtain the thermal quantities that are difficult to measure or analyze. When the shape of the problem is unknown, the inverse geometry problem can be applied to the shape estimation.

Thermal energy storage is becoming popular as an energy saving method to solve the imbalance between day and night energy demand patterns. Especially, latent heat energy storage system has advantages of low volume/energy ratio together with its small temperature deviation as compared with sensible heat storage system.

In the previous work of thermal storage system by Ismail et. al. [1,2,3,4], a direct phase change problem on the solidification of phase change material was studied numerically.

Zabaras et. al. [5] estimated the thermal boundary conditions for desired phase front motion. Huang and Chen [6] carried out the inverse geometry problem successfully by using boundary element method and conjugate gradient method. Huang and Hsiung [7] also conducted the optimal design problem of the cooling passage inside a turbine blade by using inverse geometry problem.

The objective of this work is to extend the inverse geometry problem to the phase change

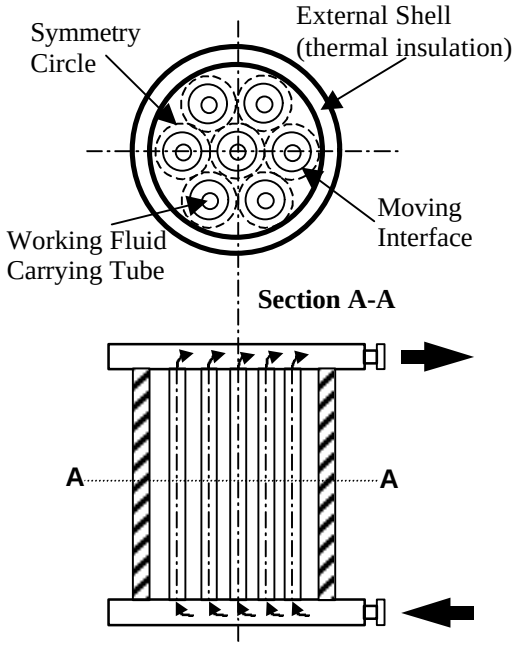


FIGURE 1 Thermal storage unit and tube arrangement inside a storage tank.

problem in estimating the ice growth process in a vertical tube type energy storage system.

DIRECT PROBLEM

Two-Dimensional Phase Change Problem

In the thermal storage system shown in Figure 1, the liquid water is contained inside the external shell, while the working fluid flows into a set of tubes inside the shell. When the temperature of the working fluid is lowered below the phase change temperature, the solid-liquid interface is growing on the outside surface of the inner tube.

The mathematical modeling is conducted using only one inner tube of the thermal storage system as shown in Figure 2. The thermal resistance of the inner tube is assumed to be negligible. To transform the formulation to the dimensionless form, the following dimensionless quantities are defined.

$$T = \frac{\overline{T} - \overline{T}_i}{\overline{T}_m - \overline{T}_0}\ ,\quad x = \frac{\overline{x}}{\overline{L}}\ ,\quad y = \frac{\overline{y}}{\overline{L}}\ ,\quad t = \frac{\overline{k}_S \overline{t}}{\overline{r}_S \overline{C}_{PS} \overline{L}^2}\ ,$$

$$\mathrm{Bi} = \frac{\overline{h}\overline{L}}{\overline{k}_S}\ ,\ \mathrm{Ste} = \frac{\overline{C}_{PS}\left(\overline{T}_m - \overline{T}_0\right)}{\overline{La}}\ ,\ a = \frac{\overline{k}_S \overline{r}_L \overline{C}_{PL}}{\overline{k}_L \overline{r}_S \overline{C}_{PS}}$$

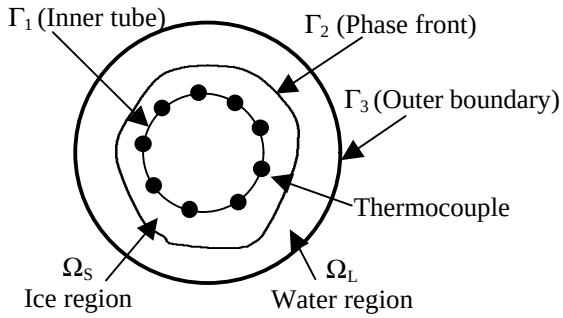


FIGURE 2 Solid and liquid domain with boundary shape.

where the '¯' denotes dimensional quantities, $\overline{La}$ means the latent heat, *x*, *y* are dimensionless geometrical coordinates, and *á* is dimensionless thermal diffusivity. $\overline{L}$, $\overline{T}_0$ are reference length and initial temperature, respectively. Bi is the Biot number that means heat transfer between tube surface and working fluid, and Ste is the Stefan number that means the effect of latent heat. In the formulation, it is assumed that the phase change is dominated by heat conduction. Considering the above assumptions, the dimensionless governing equation for a two-dimensional phase change problem is given as:

$$\left(\frac{\partial^2 T_S}{\partial x^2} + \frac{\partial^2 T_S}{\partial y^2}\right) = \frac{\partial T_S}{\partial t} \qquad \text{in } \grave{U}_S \quad (1a)$$

$$\frac{\partial T_S}{\partial n} = Bi\left(T_S - T_\infty\right) \qquad \text{along } \tilde{A}_1 (1b)$$

$$\left(\frac{\partial^2 T_L}{\partial x^2} + \frac{\partial^2 T_L}{\partial y^2}\right) = a\,\frac{\partial T_L}{\partial t} \qquad \text{in } \grave{U}_L \quad (1c)$$

$$\frac{\partial T_L}{\partial n} = 0 \qquad \text{along } \tilde{A}_3 (1d)$$

$$T_L = 0 \qquad \text{at } t = 0 \quad (1e)$$

$$\frac{\partial T_S}{\partial n} - \frac{\overline{k}_L}{\overline{k}_S}\frac{\partial T_L}{\partial n} = \frac{1}{\mathrm{Ste}} v_n \qquad \text{along } \tilde{A}_2 (1f)$$

$$T_S = T_L = T_m \qquad \text{along } \tilde{A}_2 (1g)$$

where *n* is normal outward direction to the boundary, v_n is the *n*-direction velocity of moving interface, and T_m is the phase change temperature.

Boundary Element Method

Since the BEM [8] can readily accommodate the change of unknown boundary, the BEM is adopted for the estimation of interface shape.

After multiplying conduction equation by the Green function, integrating the result with respect to all space and time domain, and applying Green's second identity, the following equation is obtained.

$$cT_N(t_J) + K\int_0^{t_J}\int_\Gamma Tq^* d\Gamma dt = K\int_0^{t_J}\int_\Gamma qT^* d\Gamma dt + \int_\Omega T_0 T^* d\Omega \quad (2)$$

$$T^* = \frac{1}{4pKt}\exp\left[-\frac{r^2}{4Kt}\right] \quad (3)$$

$$q^* = \frac{d}{8pK^2t^2}\exp\left[-\frac{r^2}{4Kt}\right] \quad (4)$$

where K is the thermal diffusivity, t_J denotes the final time step, and N is a point to be analyzed. T^* is Green function, and q^* is the derivative of T^* with respect to the normal direction. r is the length to a point on the boundary from N, and $\hat{o} = t_J - t$. d is the partial derivative of r respect to normal direction.

Since the initial temperature is zero, the last integral term of right-hand side of Eq. (2) is eliminated. Reformulating the result, the boundary element equation is obtained. By solving the boundary element equation, the unknown temperatures or heat fluxes can be obtained.

INVERSE PROBLEM

The present inverse estimation of the phase change interface is to be performed in such a way that the following functional is minimized.

$$J\left[\Gamma_2(x,y,t)\right] = \int_{t=0}^{t_J}\sum_{m=1}^{M}\left[T_{S,m}(t) - Y_m(t)\right]^2 dt \quad (5)$$

Here $T_{S,m}(t)$ is the calculated temperatures by the estimated interface, and $Y_m(t)$ is the measured temperatures.

Conjugate Gradient Method

The CGM iterative process is used for the estimation by minimizing the functional. The phase change interface is given as

$$\hat{\Gamma}_2^{n+1}(x,y,t) = \hat{\Gamma}_2^n(x,y,t) - b^n P^n(x,y,t) \quad (6a)$$

$$\hat{x}_2^{n+1} = \hat{x}_2^n - b^n P^n(x,y,t)\times\cos f \quad (6b)$$

$$\hat{y}_2^{n+1} = \hat{y}_2^n - b^n P^n(x,y,t)\times\sin f \quad (6c)$$

where f is the angle between the horizontal and outward normal direction of the unknown boundary. $P(x,y,t)$ is the direction of descent (i.e., search direction) given by

$$P^n(x,y,t) = J'^n(x,y,t) + g^n P^{n-1}(x,y,t) \quad (7)$$

where g^n is the conjugate coefficient defined as

$$g^n = \frac{\int_{t=0}^{t_J}\int_{\Gamma_2}\left(J'^n\right)^2 d\Gamma_2 dt}{\int_{t=0}^{t_J}\int_{\Gamma_2}\left(J'^{n-1}\right)^2 d\Gamma_2 dt}, \quad g^0 = 0 \quad (8)$$

Sensitivity Problem and Search Step Size

The sensitivity problem is obtained by replacing $\tilde{A}_2$ by $\tilde{A}_2 + \ddot{A}\tilde{A}_2$ and T by $T+\ddot{A}T$ in the direct problem, then subtracting the direct problem from the resultant expression, and neglecting the second-order terms.

$$\left(\frac{\partial^2\Delta T_S}{\partial x^2} + \frac{\partial^2\Delta T_S}{\partial y^2}\right) = \frac{\partial\Delta T_S}{\partial t} \quad \text{in } \grave{U}_S \quad (9a)$$

$$\frac{\partial\Delta T_S}{\partial n} = \mathrm{Bi}(\Delta T_S) \quad \text{along } \tilde{A}_1 (9b)$$

$$\Delta T_L = 0 \quad \text{at } t=0 \quad (9c)$$

$$\frac{\partial\Delta T_S}{\partial n} - \frac{\bar{k}_L}{\bar{k}_S}\frac{\partial\Delta T_L}{\partial n} = \frac{1}{\mathrm{Ste}}\Delta v_n \quad \text{along } \tilde{A}_2 (9d)$$

$$\Delta T_S = (\Delta\Gamma_2)\frac{\partial T_S}{\partial n} \quad \text{along } \tilde{A}_2 (9e)$$

Only one of the two boundary conditions is applied to unknown interface, since the measured temperatures are used additionally. If Eq. (9d) is used for the boundary condition, the iteration cannot be performed because of the fluctuation resulting from the condition including a 'v_n' term coupled with an unknown interface that is changed with iterations. Therefore, Eq. (9e) is used for the interface condition in this study.

The functional can be rewritten as follows.

$$J\left[\hat{\Gamma}_2^{n+1}(x,y,t)\right] = \int_{t=0}^{t_J}\sum_{m=1}^{M}\left[T_{S,m}\left(\hat{\Gamma}_2^n\right) - b^n \Delta T_{S,m}\left(P^n\right) - Y_m\right]^2 dt \quad (10)$$

The following expression on the search step size is obtained by differentiating Eq. (10) with respect to $\hat{a}_n$ and letting by zero.

$$b^n = \frac{\int_{t=0}^{t_J}\sum_{m=1}^{M}\left[T_{S,m}\left(\hat{\Gamma}_2^n\right) - Y_m\right]\Delta T_{S,m}(P^n)dt}{\int_{t=0}^{t_J}\sum_{m=1}^{M}\left[\Delta T_{S,m}\left(P^n\right)^2\right]dt} \quad (11)$$

The sensitivity function $\Delta T_{S,m}(P^n)$ can be obtained from the solutions of Eq. (9) by letting $\Delta\Gamma_2 = -P_n$.

Adjoint Problem and Gradient Equation

To obtain the adjoint problem, Eq. (1) is multiplied by the $l_S(x,y,t)$, then integrated over all space and time domain, and added the functional. In the resulting equation, the Ä*J* is obtained by perturbing $\tilde{A}_2$ by Ä$\tilde{A}_2$ and *T* by Ä*T*, then subtracting the original equation, and neglecting the second-order terms. This gives

$$\Delta J = \int_{t=0}^{t_J}\int_{\Gamma_1} 2(T_S - Y)\Delta T d(x - x_m) d(y - y_m) d\Gamma_1 dt + \int_{t=0}^{t_J}\int_{\Omega_S} l_S\left[\frac{\partial^2 \Delta T_S}{\partial x^2} + \frac{\partial^2 \Delta T_S}{\partial y^2} - \frac{\partial \Delta T_S}{\partial t}\right] d\Omega dt \quad (12)$$

where $d(\bullet)$ is the Dirac delta function. In Eq. (12), the domain integral term is reformulated based on Green's second identity, the boundary and the initial conditions of the sensitivity problem are utilized, and the integral including Ä*T* becomes zero. Then the following adjoint problem is obtained

$$\left(\frac{\partial^2 l_S}{\partial x^2} + \frac{\partial^2 l_S}{\partial y^2}\right) = -\frac{\partial l_S}{\partial t} \quad \text{in } \grave{U}_S \quad (13a)$$

$$\frac{\partial l_S}{\partial n} = \mathrm{Bi}\, l_S - 2(T_S - Y)d(x - x_m)d(y - y_m) \quad \text{along } \tilde{A}_1 (13b)$$

$$l_S = 0 \quad \text{at } t = t_f \quad (13d)$$

$$l_S = 0 \quad \text{along } \tilde{A}_2 (13c)$$

After the introduction of the adjoint problem, the following integral term is left.

$$\Delta J = \int_{t=0}^{t_J}\int_{\Gamma_2} -\frac{\partial l_S}{\partial n}\frac{\partial T_S}{\partial n}\Delta\Gamma_2 d\Gamma dt \quad (14)$$

By definition [6], the functional increment can be represented as

$$\Delta J = \int_{t=0}^{t_J}\int_{\Gamma_2} J'(x,y,t)\Delta\Gamma_2 d\Gamma dt \quad (15)$$

By comparing Eq. (14) with Eq. (15), the gradient of functional $J'(x,y,t)$ is obtained as

$$J'(x,y,t) = -\frac{\partial l_S}{\partial n}\frac{\partial T_S}{\partial n}\bigg|_{\Gamma_2} \quad (16)$$

In Eq. (16), the estimation cannot be performed since the initial and final values are always equal to zero. Therefore the following assumptions are introduced.

$$\frac{\partial l_S(x,y,t_J)}{\partial n} = \frac{\partial l_S(x,y,t_{J-1})}{\partial n} \quad (17)$$

$$\frac{\partial T_S(x,y,0)}{\partial n} = \frac{\partial T_S(x,y,\Delta t)}{\partial n} \quad (18)$$

where Ä*t* is the numerical time increment.

Stopping Criterion

If the problem contains no measurement errors, the traditional check condition is specified as

$$J[\hat{\Gamma}_2^{n+1}(x,y,t)] < e \quad (19)$$

where *å* is a small value. However, the measured temperature data may contain measurement errors. Therefore the stopping criteria *å* is obtained by using the discrepancy principle [9].

$$e = M s^2 t_J \quad (20)$$

where *ó* is the standard deviation of the measurements, which is assumed to be a constant.

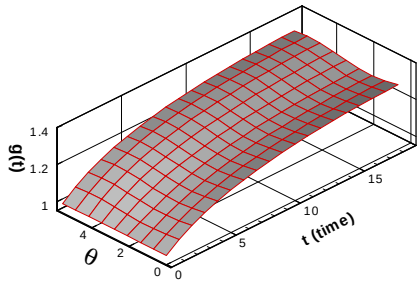


FIGURE 3 Exact interface shape of ice.

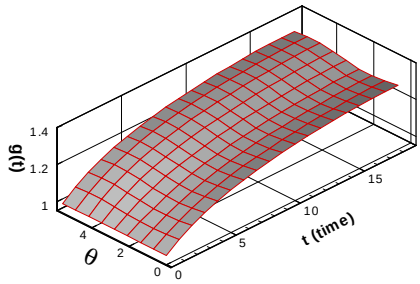


FIGURE 4 Estimated interface shape of ice using ó = 0.0 and M = 9.

Computational Procedure

When $\hat{\Gamma}_2^n(x,y,t)$ is available, the calculation for the iteration n+1 is summarized as below.

Step 1. Solving the direct problem given by Eq. (1), and compute the functional of Eq. (5)

Step 2. Examine the stopping criteria given by the Eq. (19) and Eq. (20).

Step 3. Solve the adjoint problem of Eq. (13).

Step 4. Compute $J'(x,y,t)$ by Eq. (16).

Step 5. Compute g^n and P^n from Eq. (7) and Eq. (8), respectively.

Step 6. Solve the sensitivity problem of Eq. (9).

Step 7. Compute the search step size of Eq. (11).

Step 8. Estimate the new $\hat{\Gamma}_2^{n+1}(x,y,t)$ from Eq. (6), and return to step 1.

RESULTS

We consider one specific inverse problem whose phase change interface is assumed to be the known function to verify the methodology considered in the present work. The specified moving velocity of the interface and the relational equation between the moving velocity and the phase change interface are used for assuming the known function of the interface.

The temperature readings for this inverse problem are produced numerically by adding numerical noises to the exact solution of the direct problem. The numerically produced measurement data can be expressed as follows.

$$Y = Y_{exact} + ws \tag{21}$$

where Y_{exact} is the exact solution of the direct problem, *ó* is the standard deviation of the measurements, and *ù* is a random variable generated by IMSL subroutine DRNNOR. *ù* has a value within –2.576 and 2.576 for the 99% confidence bounds.

The reference length representing the radius of inner tube is given as unity. The thermal properties of ice in the solid and liquid regions are given as:

$\bar{r}_L = 999.8\ kg/m^3$, $\bar{C}_{PL} = 4.225\ kJ/kg\cdot°C$, $\bar{k}_L = 0.566\ W/m\cdot°C$, $\bar{r}_S = 920\ kg/m^3$, $\bar{C}_{PS} = 2.11\ kJ/kg\cdot°C$, $\bar{k}_S = 2.2\ W/m\cdot°C$, $\overline{La} = 333.4\ kJ/kg$

In order to specify an arbitrary phase change configuration, the moving velocity of a phase change interface is assumed as

$$v_n(q,t) = 0.04 - 0.0005\times t + (0.04 - 0.01\times\frac{\sin q}{(t+1)})\times\exp[-4\times\frac{t}{t_J}] \tag{22}$$

The explicit relationships between the moving velocity and the phase change interface are given as

$$x(q,t) = x(q,t-1) + v_n(q,t-1)\times dt\times\cos f \tag{23a}$$

$$y(q,t) = y(q,t-1) + v_n(q,t-1)\times dt\times\sin f \tag{23b}$$

From the interface velocity given by Eq. (22) and the explicit relationships given by Eq. (23), the phase change interface for the validation is specified. The interface configuration specified by Eq. (22) and Eq. (23) is shown in Figure 3.

The initial guessed interface configuration for the iteration is assumed to be the surface shape of inner tube, $\Gamma_1(x,y,t)$.

The difference between the exact and the estimated interface is expressed by the following shape error.

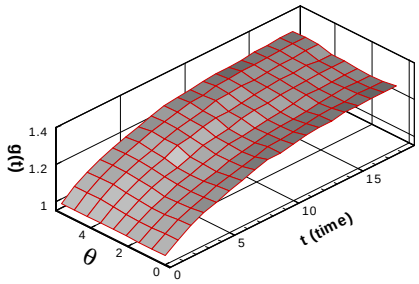


FIGURE 5 Estimated interface shape of ice using ó = 0.05 and M = 9.

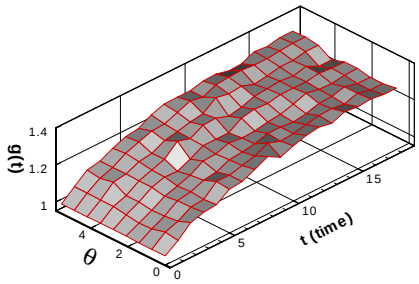


FIGURE 6 Estimated interface shape of ice using ó = 0.2 and M = 9.

$$shape\ \ error(\%) = \sum_{i=1}^{I}\sum_{j=0}^{J}\left|\frac{\hat{\Gamma}_2(x_i,y_i,t_j)-\Gamma_2(x_i,y_i,t_j)}{\Gamma_2(x_i,y_i,t_j)\times[I\times(J+1)]}\right|\times 100\% \quad (24)$$

where I and J represent the number of elements discretized by space and time, respectively.

The relative quantity of measurement errors is calculated by the following measurement error.

$$measurement\ error(\%) = \frac{2.576s}{Y_{avg}}\times 100 \quad (25)$$

Here Y_{avg} is the average of measured data.

Validation Case Using 9-Thermocouples

The inverse analysis using 9-thermocouples, denoted by the dots in Figure 2, is performed. Without measurement errors ($ó$ = 0), the given stopping criteria $å$ = 0.002 is satisfied after the estimation with 33-iterations. The estimated result of the phase change interface is shown in Figure 4. In this case, the functional is converged to 0.0018 from 1900, the shape error to 0.02% from 11.0%. From Figures 3 and 4, it can be seen that the phase change interface of ice is obtained accurately in the case of no measurement errors.

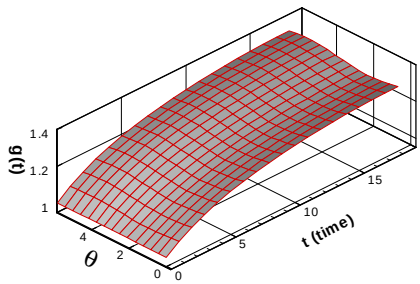


FIGURE 7 Estimated interface shape of ice using ó = 0.0 and M = 5.

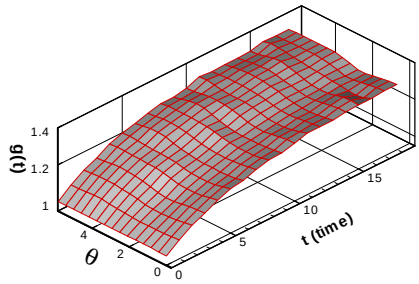


FIGURE 8 Estimated interface shape of ice using ó = 0.05 and M = 5.

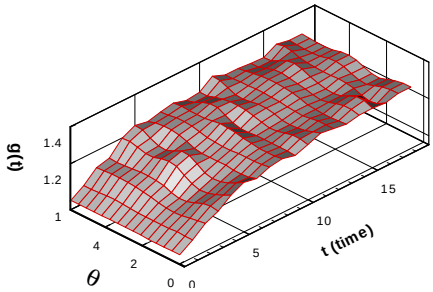
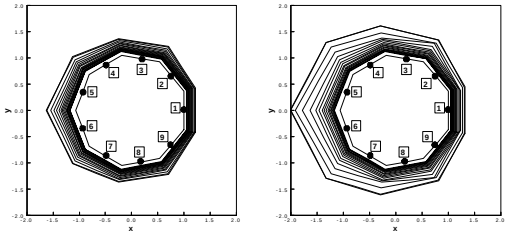


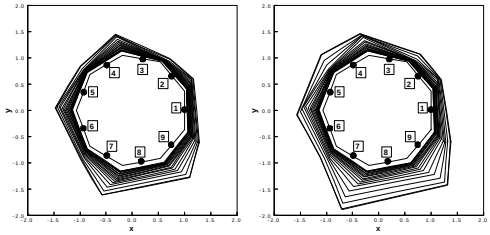
FIGURE 9 Estimated interface shape of ice using ó = 0.2 and M = 5.

Since the measurement errors are always introduced in any real measurement, it is required to observe the influence of measurement errors on the inverse estimation. Figure 5 shows the estimated interface after 13-iterations in the case of $ó$ = 0.05. In this case, the measurement error is calculated as 1.40%. The functional is converged to 0.37 from 1900, and the shape error to 0.27% from 11.0%. In Figure 6, the estimated interface converged after 8-iterations is shown in the case of ó = 0.2. Here the measurement error is



(a) lower temperature increment (b) higher temperature increment

FIGURE 10 Estimated ice shape for higher desired temperatures at 1 and lower at 5, 6.



(a) lower temperature increment (b) higher temperature increment

FIGURE 11 Estimated ice shape for higher desired temperatures along $\overline{1,2,3}$, medium along $\overline{4,5,6}$ and lower along $\overline{7,8,9}$.

calculated as 5.70%. The functional is converged to 2.2 from 1900, and the shape error on 0.75% from 11.0%. The above results indicate that the phase change interface estimation is accurately performed even though measured temperatures contain measurement errors.

Validation Case Using 5-Thermocouples

In the most of real systems, there are some difficulties in installing enough sensors. Therefore, it is necessary to verify the effect of the less number of sensors. To examine the influence of the number of sensors, the inverse analysis is performed using 5-thermocouples.

In the case with 5-thermocouples, only five-spatial information can be obtained. But the five-spatial information is not enough to express the interface for the inverse calculation. Therefore the phase change interface is expressed more closely by interpolation between each thermocouple.

Figure 7 shows the estimated result of phase change interface in the case with no measurement error. It satisfies the given stopping criteria $\mathring{a}$ = 0.002 after 33-iterations. The functional is converged to 0.0018 from 1900, the shape error to 0.14% from 11.0%. Using 5-thermocouples without measurement errors, the phase change interface of ice is estimated accurately.

In Figure 8, the estimated interface after 18-iterations is shown in the case of $\acute{o}$ = 0.05. The measurement error is 1.40%, the functional is converged to 0.19 from 1100, and the shape error to 0.35% from 11.0%. As expected, the estimated interface with measurement errors is worse than that with no measurement error. But the estimated interface shows a good agreement with the exact interface.

In Figure 9, the estimated interface after 12-iterations is shown in the case of $\acute{o}$ = 0.2. Here the measurement error is 5.80%, the functional is converged to 3.4 from 1100, and the shape error to 1.1% from 11.0%. Though small number of thermocouple and large measurement error, the interface can also be estimated with a good agreement.

Results on the Desired Temperature Distributions

To illustrate the ability of the present inverse estimation in operating the practical thermal storage system and monitoring the generated ice, we consider arbitrary temperature distributions on the thermocouple location. The estimations for the arbitrary temperature distributions are performed.

Figure 10(a) shows the estimated interface configuration when imposing relatively higher temperature distribution on the first thermocouple and lower temperature distributions on the fifth and sixth. It can be seen that a thinner ice layer is generated around the first thermocouple, and a thicker ice layer around the fifth and sixth. Figure 10(b) is the estimated result for the similar temperature distribution as in Figure 10(a), but temperature increases more with time compared to that of Figure 10(a). Figure 10(b) shows that the ice growth is more active than that of Figure 10(a) because of increasing temperature differences.

Figure 11(a) illustrates the estimated interface configuration when imposing relatively lower temperature on the ninth point and increasing the temperature clockwise from that point. It can be seen that the thick ice layer is generated around the ninth thermocouple, and the thickness of ice

layer decreases clockwise from that location. Figure 11(b) shows the estimated result with the similar temperature distributions as in Figure 10(a), but temperature increases more with time compared to that of Figure 11(a). It is observed that thickness of ice layer in Figure 11(b) is thicker than that of Figure 11(a).

CONCLUSION

In this work, the phase change interface is estimated using the inverse geometry technique by applying the conjugate gradient method (CGM) along with the boundary element method (BEM). Several cases for different numbers of sensors and measurement errors are successfully performed. Physical validity for an arbitrary temperature distribution imposed on the thermocouples is also examined. The present results show that the phase change interface can be estimated with a good accuracy by using the method considered in this work. The present work may also be used for controlling and monitoring the thermal storage system.

ACKNOWLEDGEMENTS

This work was supported by the research fund of Hanyang University (HY-2001).

REFERENCES

1. K. A. R. Ismail and M. M. Abugderah, Performance of a thermal storage system of the vertical tube type, *Energy Conversion & Management*, **41**, pp. 1165-1190 (2000)

2. K. A. R. Ismail, C. L. F. Alves and M.S. Modesto, Numerical and experimental study on the solidification of PCM around a vertical axially finned isothermal cylinder, *Applied Thermal Engineering*, Vol. **21**, pp. 53-77 (2001)

3. K. A. R. Ismail, J. R. Henriquez, L. F. M. Moura and M. M. Ganzarolli, Ice formation around isothermal radial finned tubes, *Energy Conversion & Management*, Vol. **41**, pp. 585-605 (2000)

4. K. A. R. Ismail and M. M. Goncalves, Thermal performance of a pcm storage unit, *Energy Conversion & Management*, Vol. **40**, pp. 115-138 (1999)

5. N. Zabaras, Y. Ruan and O. Richmond, Design of two - dimensional Stefan processes with desired freezing front motions, *Numerical Heat Transfer, B*, **21**, pp. 307-325 (2000)

6. C. H. Huang and H. M. Chen, Inverse geometry problem of identifying growth of boundary shapes in a multiple region domain, *Numerical Heat Transfer, A*, **35**, pp. 435-450 (1999)

7. C. H. Huang and T. Y. Hsiung, An inverse geometry problem of estimating optimal shape of cooling passages in turbine blades, *International Journal of Heat and Mass Transfer*, **42**, pp. 4307-4319 (1999)

8. C. A. Brebbia and J. Dominguez, Boundary elements; an introductory course, 2nd. Ed., McGraw-Hill., 1992.

9. O. M. Alifanov, Solution of an inverse problem of heat conduction by iteration methods, *Journal of Engineering Physics*, **26**, No. 4, pp. 471-476. (1974)

IMAGING AND INVERSE SCATTERING

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil 2002

COMPUTATION OF MAGNETIC FIELD SOURCES FROM MEASUREMENTS USING ITERATIVE REGULARIZATION

S. Bégot
P. Hiebel, J.M. Kauffmann
L2ES LRE T 31
Université de Technologie de Belfort-Montbéliard
Université de Franche-Comté
IGE 2 avenue Jean Moulin
90000 Belfort France
sylvie.begot@univ-fcomte.fr

E.A. Artioukhine

CREST UMR6000
Université de Franche-Comté
IGE 2 avenue Jean Moulin
90000 Belfort France

ABSTRACT

This paper deals with the resolution of inverse problems in magnetostatics. The cases the authors have considered are related to the determination of the current density or positions of air coils on the basis of magnetic field measurements. These problems are ill-posed, so regularization techniques are needed. The authors discuss the main iterative regularization principles and D-optimal experimental design. Numerical and experimental results are presented.

NOMENCLATURE

A: Matrix of the contribution of each coil to each field value.
B: Magnetic field values.
e: Maximum relative error on the estimated parameters.
E: Quadratic error between the exact and computed solution.
J: Coil current densities.
M_F: Information matrix.
R: Quadratic residue between the computed and measured field.
Z: Coil axial positions.
δ^2: Level of experimental error.

INTRODUCTION

In superconducting magnet technology, coils are located in cryostatic equipment. Therefore, a numerical tool calculating the coil parameters from magnetic field measurement is interesting for faulty operating conditions determination. This leads us to solve two different inverse problems: a linear inverse problem to compute the current densities on the basis of field measurements, and a nonlinear inverse problem for the computation of the coil positions.

Several aspects across electromagnetism raise inverse problems. Inverse problems for Maxwell equations in general are discussed in [1]. In the case of superconducting magnet diagnosis, this subject has been discussed by [2] who uses the Levenberg-Marquard algorithm. Reference [3] employs the Truncated Singular Value Decomposition for identification of the plasma magnetic contour from external measurement by means of equivalent currents. Magnetostatic inverse problems have also been discussed in biomagnetism by several authors. Reference [4] gives an overview of the different methods used.

In this paper, the authors present a different approach based on iterative regularization [5,6]. This method was chosen because it applies to linear and nonlinear problems. We reconstruct the current densities then the coil axial positions of Magnetic Resonance Imaging (MRI) magnets. The main contribution of this paper is the experimental validation of the numerical method presented in [7].

The first part of this paper describes the problem statement. In the second part, the resolution of the linear inverse problem is detailed. Numerical and experimental results are presented, a D-optimal experimental design is calculated. In the third part, the resolution of the nonlinear inverse problem with numerical and experimental results is described.

PROBLEM STATEMENT

Description of MRI magnets

The magnets we are considering are magnets for MRI. These magnets produce a high and uniform magnetic field in a Volume Of Interest (VOI), which is a sphere at the center of the magnet. Typically, the field values range from 0.5 T to 1.5 T and the field homogeneity is 10 ppm (part per million). To create such strong magnetic field, superconducting technology is required. Therefore the magnet coils are made of superconducting wire and the whole magnet is in a cryostat. The magnet length is about 1.5 m and the bore diameter 0.9 m (Fig. 1). The field homogeneity is checked by performing measurements at the surface of the VOI.

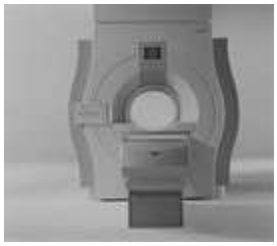


Figure 1. Example of a MRI magnet

Magnetostatics equations

We assume that the magnetic field is static and that the equipment is in vacuum. The Biot-Savart law gives the magnetic flux density $\boldsymbol{B}$ at the point P produced by a current distribution $\boldsymbol{j}$ at the point M [8]:

$$\mathbf{B}(P) = \frac{\mu_0}{4\pi} \iiint_{\text{Coil volume}} \frac{\boldsymbol{j}(M) \times \mathbf{MP}}{\|\mathbf{MP}\|^3} d\tau \qquad (1)$$

In a MRI magnet, the principal magnetic field is parallel to the magnet axis. Thus, we consider only the axial component of the field B_z. The system possesses cylindrical symmetry, the current density is only azimuthal (Fig. 2).

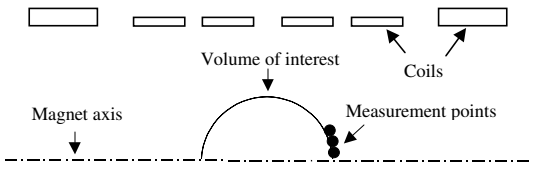


Figure 2. Cross section of the magnet geometry

Inverse problems statement

We will discuss two types of problems.

Linear inverse problem. The coil geometry and positions are known, and we reconstruct the current density. This leads us to solve a linear integral equation of Fredholm of the 1st kind. The unknowns are the current density values. We denote $\mathbf{J}$ the vector of all the densities. The spatial discretrization and collocation of the Biot-Savart law lead to a singular integral discretization. It can be stated as the matrix relation:

$$A\mathbf{J} = \mathbf{B} \qquad (2)$$

where $\mathbf{B}$ denotes the magnetic field values, and A is the matrix of the contributions of each coil to the field values in each point.

Nonlinear inverse problem. The current density, coil geometry, and coil radius are known, and we reconstruct the coil axial positions. The integral equation to be solved is then nonlinear. The purpose is to evaluate a displacement from the normal position, which causes a loss in the field homogeneity. The unknowns are the coils axial positions $\mathbf{Z}$. As this problem is nonlinear, the discretization of (1) leads to a formulation of the type:

$$A(\mathbf{Z}) = \mathbf{B} \qquad (3)$$

where A depends on the coil positions $\mathbf{Z}$.

RESOLUTION OF THE LINEAR INVERSE PROBLEM

The authors use iterative regularization [5-7]. According to this method principle, we formulate the inverse problem as an optimization problem, then solve it by an iterative method coupled to a stopping rule. The number of iterations in the minimization is the regularization parameter.

Objective function

The objective function to minimize is the quadratic residue R between the computed magnetic field and the measured field.

$$R(\mathbf{J}) = \|A\mathbf{J} - \mathbf{B}\|^2 \qquad (4)$$

The direct problem i.e. $A\mathbf{J}$ is computed by integration of the Biot-Savart law.

Minimization procedure

The minimization is performed with the conjugate gradient method [9].
The residue gradient is calculated by:

$$G_i = 2(A\mathbf{J} - \mathbf{B})^T A_i \qquad (5)$$

where A_i denotes the ith column of the matrix A. The descent parameter is calculated by making a one variable minimization along the descent direction **D**. As the problem is linear, this minimization has an analytical solution:

$$\gamma = -\frac{(A\mathbf{D})^T (A\mathbf{J} - \mathbf{B})}{(A\mathbf{D})^T (A\mathbf{D})} \qquad (6)$$

Choice of the regularization parameter

The authors employ two different ways to choose the regularization parameter.

Discrepancy principle [5]. The authors match the residue and the error level δ in the input data. We define the error level by:

$$\delta^2 = \left\| \mathbf{B}_{measured} - \mathbf{B}_{exact} \right\|^2 \qquad (7)$$

Thus, the final iteration is determined when:

$$R \approx \delta^2 \qquad (8)$$

L-curve [10]. The value of the residue is plotted versus the solution norm for different regularization parameters. This curve has an L-shape. The iteration number is chosen by selecting the value corresponding to the corner of the L.

Numerical example

We consider a magnet (Fig. 3) divided into 24 coils. Our goal is to compute the current density values from 101 measurement points: 1 at the VOI center and 100 on the sphere.

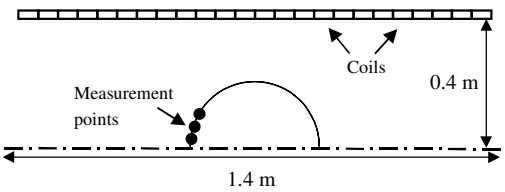


Figure 3. Geometry of the numerical example

We have simulated experimental data by adding a gaussian random noise of 10 ppm to the exact data. For this noisy data, the exact value of δ^2 is $2.7.10^{-9}$.

The minimization procedure starts with all values set to zero. The application of the discrepancy principle (Fig. 4) or the use of the L-Curve (Fig. 5) indicates the 23rd iteration as the final iteration to be taken.

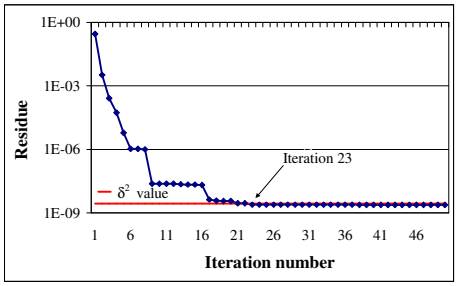


Figure 4. Application of the discrepancy principle

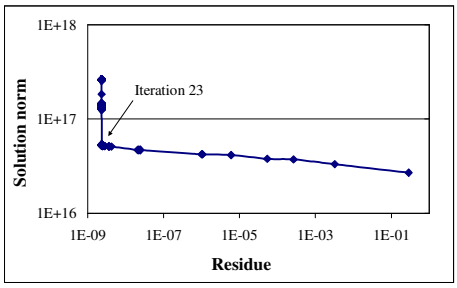


Figure 5. L-curve.

In order to check the optimality of this iteration number, we plot the quadratic error:

$$E = \left\| \mathbf{J}_{calulated} - \mathbf{J}_{exact} \right\|^2 \qquad (9)$$

versus the iteration number. We observe that the error E decreases until the 23rd iteration then increases (Fig. 6). So, the final iteration selected is the optimal iteration. Comparing the computed solution for 23 iterations and the exact solution, we observe a good agreement (Fig. 7). The maximum relative error on the estimated densities:

$$e = \max_{coil} \frac{\left| j_{exact} - j_{calculated} \right|}{j_{exact}} \qquad (10)$$

is 9%. At the end of the minimization, the computed solution becomes unstable (Fig. 8).

This numerical test shows that the proposed method enables us to find the current density with measured field values as input data.

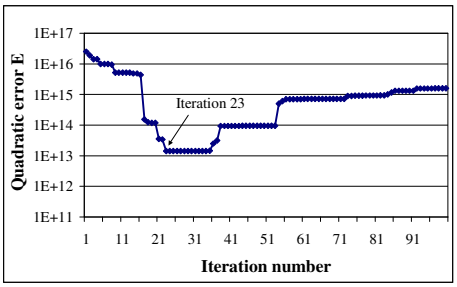


Figure 6. Quadratic error versus iteration number

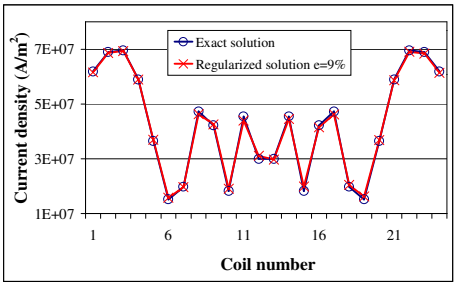


Figure 7. Exact and calculated solutions 23 iterations

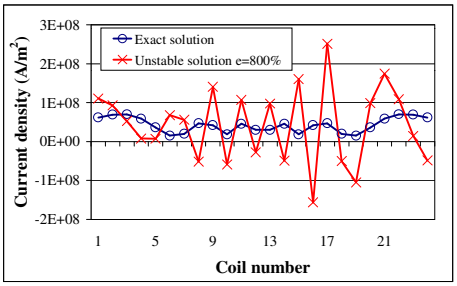


Figure 8. Exact solution and calculated solutions 500 iterations

Experimental validation

The aim of this experimental validation is to show the instability of real magnetostatic inverse problems and to test the proposed method on these problems.

Experimental scheme. We use resistive circular air coils of rectangular cross section. The experimental scheme is presented in figure 9. The power supply is current controlled. The magnetic field strength is measured by a Hall effect probe. For the measured field (about 10^{-2} T), this equipment has an absolute accuracy of 3.10^{-4} T. For each point, we perform 100 acquisitions of the field values and we take the mean of these acquisitions. As a result, the variance of the measurement error is reduced by a factor of 10 [11]. The probe position is controlled by displacement tables.

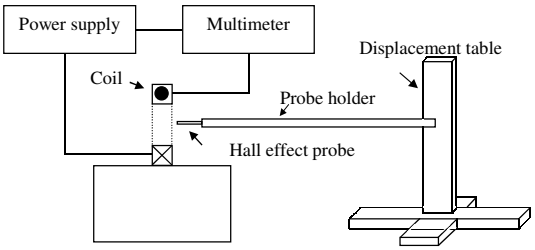


Figure 9. Experimental scheme

Inverse problem under analysis. We assume the coil is divided into 5 sections. The magnetic field is measured in 5 points regularly distributed in 20 mm on the coil axis (Fig. 10). The value of δ^2 is $4.5.10^{-9}$. The exact current density that we search is constant and equals to $4.45.10^6$ A/m^2.

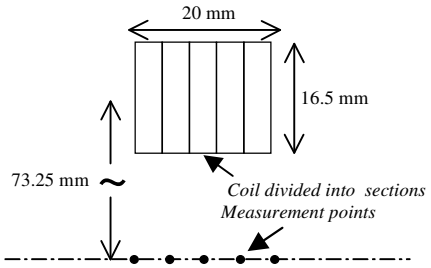


Figure 10. Experimental inverse problem

Resolution by iterative regularization.. Applying the discrepancy principle, we select the first iteration as the final one (Fig. 11). The use of the L-curve selects iteration 3 (Fig. 12). By plotting the quadratic error versus the number of iterations (Fig. 13), we observe that the discrepancy principle application is efficient to determine the right iteration number. The computed solution is in good agreement with the exact solution (Fig. 14). The use of the L-curve,

in that case, produces a less accurate solution.

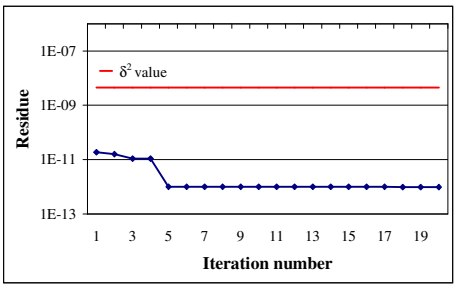


Figure 11. Application of the discrepancy principle

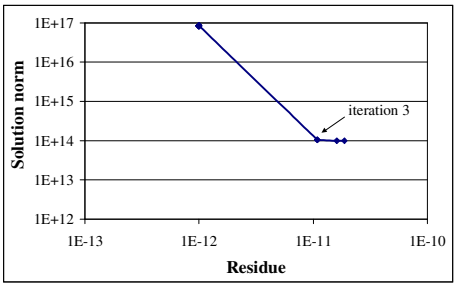


Figure 12. L-curve

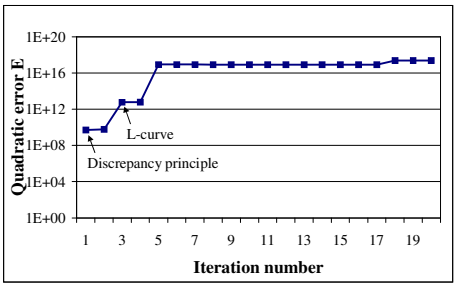


Figure 13. Quadratic error versus iteration number

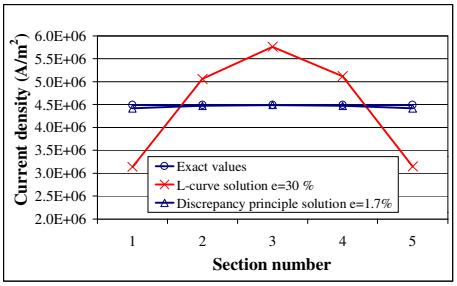


Figure 14. Computed values

As expected, at the end of the minimization, the computed solution is unstable (Fig. 15).

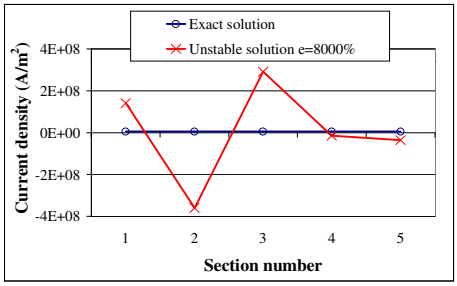


Figure 15. Unstable solution

This experimental tests show that the proposed method enables us to compute stable solutions for the problem considered.

D-optimal experimental design

The aim of this section is to present the computation of a D-optimal experimental design [12,13] and to illustrate that this plan increases the accuracy of the inverse problem solution.

Theoretical results. We assume additive, independent errors of zero mean and constant variance. With the least squares estimator, the estimation of the parameters **J** is given by:

$$\mathbf{J} = (A^T A)^{-1} A^T \mathbf{B} \qquad (11)$$

The matrix $M_F = A^T A$ is called the information matrix. It has been demonstrated that:

$$V(\mathbf{J}) = \sigma^2 (A^T A)^{-1} \qquad (12)$$

with V variance-covariance matrix of the parameters, σ variance of the measurement error. So the variance of the estimated parameters depends not only on the variance of the measurement errors but also on the experimental design through the model matrix A. A D-optimal design consists in maximizing the determinant of the information matrix. This ensures that the generalized variances of the parameters are minimized, and that the volume of the confidence region of the estimated parameters is minimal.

Computations. We choose to take as many measurement points as unknown currents in order to minimize the number of measurements. In this case, A is a square matrix and:

$$det(M_F) = det^2(A) \qquad (13)$$

So we minimize:

$$d = -\log\left(\det\left(M_F\right)\right) \qquad (14)$$

To take into account the physical limitations of experiments, we minimize (14) under constraints. We use the projected conjugate gradient algorithm. The gradient in a $\boldsymbol{u}$ direction is calculated by:

$$G_i = -2\frac{\partial\, det(A)}{\partial u_i} / \left(ln\,10\, det\left(A\right)\right) \qquad (15)$$

The derivative of det*(A)* is the sum of the determinants obtained by derivation of each row of *A*. The descent parameter is numerically computed by a golden section algorithm.

Numerical Results. We consider the superconducting magnet of figure 16 made of 6 coils. The classic choice for the experimental design would be to take the measurements on the surface of VOI.

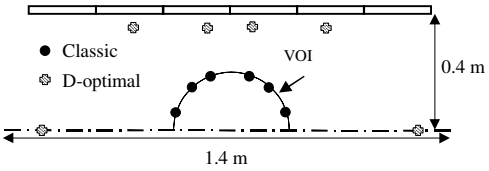


Figure 16. Magnet shape and experimental designs

We compute the D-optimal experimental points following the method presented in the preceding section. As the measurements are not possible in the cryostat, we limit the y-coordinate to 0.35 m. The computed positions are presented in figure 16, the determinant of the information matrix and the condition number of *A* are given in Table 1. The increase in the determinant goes along with an improvement in the condition number of *A*. As a result, the inverse problem is more stable. The comparison between the accuracy of the computed current densities for both designs shows a significant improvement (Fig. 17).

Table 1. Determinant and condition number

	Classic design	D-optimal design
-log(det(M_F))	111	100
Condition number	503	5

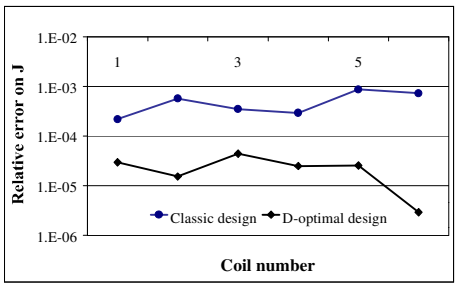


Figure 17. Accuracy obtained for both designs

So, The construction of a D-optimal experiment for an inverse problem improves the accuracy of solutions.

RESOLUTION OF THE NONLINEAR INVERSE PROBLEM

Computations

The computations of the coil geometric dimensions $\mathbf{Z}$ are performed with the same principles as in the linear case. The objective function is expressed as:

$$R(\mathbf{Z}) = \left\| A(\mathbf{Z}) - \mathbf{B} \right\|^2 \qquad (16)$$

In the minimization procedure, the gradient is computed as:

$$G_i = 2\left(A(\mathbf{Z}) - \mathbf{B}\right)^T \left(\frac{\partial A(\mathbf{Z})}{\partial \mathbf{Z_i}} \right) \qquad (17)$$

The derivative of *A* with respect to the coil axial positions is computed by derivation of (1) with respect to the integral bounds in *z*.

The descent parameter is chosen as the linear estimation of the solution of the one variable minimization.

$$\gamma = -\frac{\left(\frac{\partial A}{\partial \mathbf{Z}}.\mathbf{D}\right)^T \left(A(\mathbf{Z}) - \mathbf{B}\right)}{\left(\frac{\partial A}{\partial \mathbf{Z}}.\mathbf{D}\right)^T \left(\frac{\partial A}{\partial \mathbf{Z}}.\mathbf{D}\right)} \qquad (18)$$

This enables us to avoid the use of a numerical line search.

Numerical example

The problem we propose to solve is the diagnosis of a magnet that presents a low homogeneity due to a manufacturing error. Some

of its coils have an incorrect axial position. As a test case, we use the magnet represented schematically in figure 18.

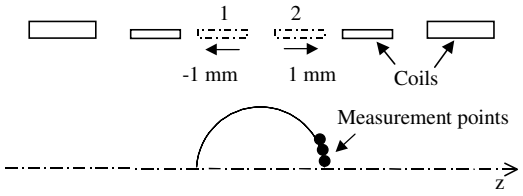


Figure 18. Geometry of the nonlinear numerical example

A defect is simulated by a 1 mm displacement of the central coils. We simulate experimental measurements by adding a uniformly distributed random noise of 10 ppm to the exact field values. We use 29 measurement points on the VOI surface. The starting points of the minimization are the coil theoretical values. All coils are allowed to move. Selecting the final iteration using exact δ^2 leads to stopping the minimization at the 6th iteration (Fig. 19). For this iteration, we observe that the quadratic error is minimal (Fig. 20).

By comparing the exact values and the regularized solution (Table 2), we observe that we are able to compute the simulated defects.

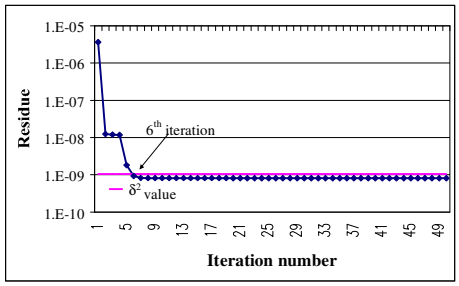


Figure 19. Application of the discrepancy principle.

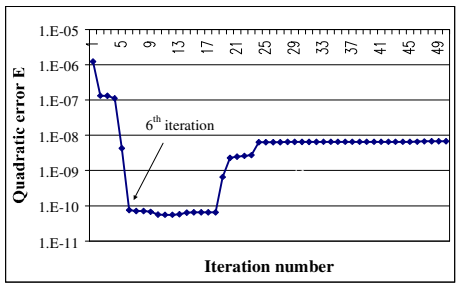


Figure 20. Quadratic error versus iteration number

Table 2. Computed defects for the regularized solution

	Exact Displacement (mm)	Computed Displacement (mm)	Relative Error (%)
Coil 1	**-1**	**$-9.94.10^{-1}$**	**0.6**
Coil 2	**1**	**$9.92.10^{-1}$**	**0.8**

At the end of the minimization, the accuracy of the computed solution is lower, but large oscillations in the computed values are not observed (Table 3). This fact can be explained by the physical nature of the inverse problem to be solved. Large instability in the solution occurs when the unknown values increase in magnitude and take different signs to produce exactly the measured field. In this problem, a large increase in the coil axial positions cannot produce a similar field to the one measured. Thus, the instability is limited.

Table 3. Computed defects for the unstable solution

	Exact displacement (mm)	Computed displacement (mm)	Relative error (%)
Coil 1	**-1**	**$-8.92.10^{-1}$**	**10.8**
Coil 2	**1**	**$1.10.10^{-1}$**	**10**

Experimental validation

We use two Helmholtz air coils (Fig. 21) with the same experimental scheme presented in figure 9. These coils, which are separated by a distance equal to their radius, produce a homogeneous field at the center of the system [8]. We compute their axial positions from two field measurement points at the coil center. We apply the iterative regularization with the discrepancy principle (Fig. 22). We observe that the computed positions match the exact values (Table 4).

So, the numerical and experimental tests illustrate the efficiency of the proposed method to reconstruct the coil axial positions.

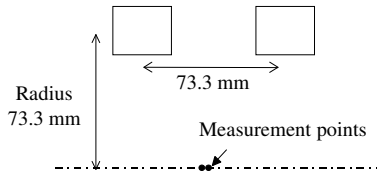


Figure 21. Helmholtz coils under analysis

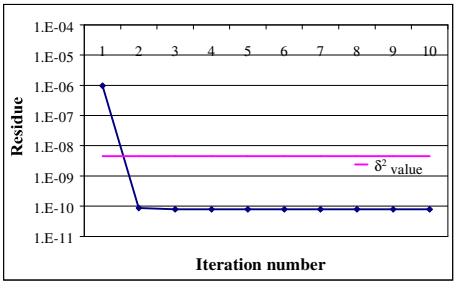


Figure 22. Application of the discrepancy principle

Table 4. Exact and computed positions

	Exact position (mm)	Computed position (mm)	Relative error (%)
Coil 1	**-36.65**	**-36.77**	**0.3**
Coil 2	**36.65**	**36.77**	**0.3**

CONCLUSION

The authors have presented iterative regularization applied to a linear and nonlinear inverse problem of magnetostatics. The numerical and experimental tests show that in a magnet it is possible to find the current density or a faulty coil axial position with measured field values as input data. The method used always produces better results than the usual least squares solution. For unstable problems, this improvement is significant, the regularized solutions are close to the exact ones, while the least square solutions are strongly distorted. The authors have analyzed the parameters that influence the stability of the inverse problems: measurement noise level, number of unknowns and number of measurements, and the unknown solution itself in the nonlinear case.

REFERENCES

1. V.G. Romanov, S.I. Kabanikhin, *Inverse problems for Maxwell's Equations*, VSP, 1994.

2. S. Russenschuck *et al*., Integrated design of superconducting accelerator magnets. A case study of the main quadrupole. *Eur. Phys. J. AP,* **1**, 93 (1998).

3. P. Bettini *et al.*, Identification of the plasma magnetic contour from external magnetic measurements by means of equivalent currents, *Eur. Phys. J. AP*, **13,** 51 (2001).

4. J.T. Nenonen, Solving the inverse problem in Magnetocardiography, *IEEE Engineering in medicine and biology* Sept 1994, 487.

5. O.M. Alifanov, E.A. Artyukhin, S.V. Rumyantsev, *Extreme methods for solving ill-posed problems with applications to inverse heat transfer problems,* Begell house, 1995

6. S.F. Gilyazov, N.L. Gol'dman, *Regularization of ill-posed problems by iteration methods*, Kluwer academic publishers, 2000

7. S. Bégot, E. Voisin, P. Hiebel, J.M. Kauffmann, E. Artioukhine, Resolution of linear magnetostatic inverse problem using iterative regularization, *Eur. Phys. J. AP*, **12** n°2, 123 (2000).

8. J.D. Jackson, *Classical electrodynamics*, 3rd ed., Wiley: New York, 1999.

9. P.E. Gill, W. Murray, M.H Wright, *Practical optimization*, Academic Press: New York, 1981.

10. P.C. Hansen, The use of the L-curve in the regularization of discrete ill-posed problems; *SIAM J. Sci. Comput* , **14** n°6, 1487, (1993).

11. R. Barlow, *Statistics*, Wiley 1989.

12. V.V. Fedorov, *Theory of optimal experiments*, Academic Press 1972.

13. S. Bégot, E. Voisin, P. Hiebel, J.M. Kauffmann, E. Artioukhine, D-optimal experimental design applied to a linear magnetostatic inverse problem, *IEEE Trans. Magn.* accepted for publication.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSION OF SPECTROSCOPIC DATA OF CO_2 RADIATION

P. Alkhoury
Laboratoire d'Energétique et d'Economie d'Energie,
Université Paris X.
1chemin Desvallières.
92410 Ville d'Avray, FRANCE.
philippe.alkhoury@cva.u-paris10.fr

F. Clément
INRIA-Rocquencourt
BP 105
78153 Le Chesnay Cedex, FRANCE.
Francois.Clement@inria.fr

P. Hervé
philippe.herve@cva.u-paris10.fr

G. Chavent
Guy.Chavent@inria.fr

ABSTRACT

This paper deals with inversion of spectroscopic measurements in order to obtain temperature and CO_2 concentration distributions in a flame. This is a non-intrusive method, where a CO_2 high-resolution spectrum, characterized by its strong dependence with temperature, is observed. Modeling is performed by a one-dimensional Radiative Transfer Equation (RTE).

Then minimization of the misfit between computed and observed spectra gives temperature and concentration profiles. The RTE is solved as a differential equation, and an adjoint-state method is used to compute a descent direction for the misfit. In order to analyze the information content of data with some noise level, for some concentration and temperature distribution, a Singular Value Decomposition is used; it enables to estimate the number of retrievable parameters. Then Gauss-Newton optimization method is applied to perform inversion and to evaluate parameters. Finally, we compare inversion results with intrusive measurements on real data from turbojets and gas turbines.

NOMENCLATURE

C	CO_2 Concentration	%
h	Length of a subsegment	cm
J_Σ	Objective function	
k	Boltzmann constant $k = 1.38e\text{-}23$	$J.K^{-1}$
k_1	Constant in Planck's Law $k_1 = 1.19e\text{-}8$	$W.cm^3$
k_2	Constant in Planck's Law $k_2 = 1.438$	$cm.K$
K_σ	Absorption Coefficient for σ	cm^{-1}
L_0	Background Radiation	$W.m^{-2}.sr^{-1}.cm^{-1}$
L_σ	Radiative Intensity function for σ	$W.m^{-2}.sr^{-1}.cm^{-1}$
L_σ^{mes}	Real measurement for σ	$W.m^{-2}.sr^{-1}.cm^{-1}$
L_Σ^{mes}	Real measurements for Σ	
L_σ^0	Local Plank function for σ	$W.m^{-2}.sr^{-1}.cm^{-1}$
N	Number of subsegments	
P	Pressure	bar
T	Temperature distribution	K
u	State variable	
$\bar{u}$	Adjoint-state variable	
Φ_σ	Lagrangian function for σ	
σ	Wavenumber	cm^{-1}
Σ	A set of wavenumbers	

INTRODUCTION

There are a lot of studies on the inversion of the Radiative Transfer Equation (RTE) in order to construct the temperature distribution, to find some medium characteristics, or to calculate some source properties for industrial or ecological purposes. A lot of work has been reported on the reconstruction of flame temperature. Determining gas distribution is an additional requirement to be pulled out from spectroscopic data. Yousefian *et al*. [1] [2] have used infrared emission to retrieve temperature and concentration profiles for a moderately noisy measurements in axisymetric flame by Abel Inversion. L. H. Liu *et al*. [3] [4] worked on the estimation of temperature and spectral absorption coefficient in a non-grey

medium, and on the identification of source term distribution. Measuring temperature in order to retrieve some radiative parameters is the subject of some works like H. M. Park and T. Y. Yoon [5] [6].

The accumulation of high-resolution spectroscopic data motivates the use of more generic methods that could be applied for different wavenumbers and different instrument resolutions. So we have expressed the absorption coefficient in a general form using HITEMP database [7]. Then we have used the derivative formula instead of the integrated one because it realizes less time execution. Still a very important question rarely answered is what are for a considered problem the limits of this method, in other words the number of parameters we hope retrieve by fitting spectroscopic data. We propose to analyze the quantity of information contained in a spectrum by using SVD of Jacobian matrix [8] [9] [10] [11], then we will compare this a priori result with profile identification.

Some hypotheses were made for this model resulting in a first approximation of the real radiative behavior of a gas. The main hypothesis is the local thermal stability and the establishment of temperature and concentration profiles. Then we neglect the particles emission-absorption (soot's particles for example), the reflection at the air-gas interface, and the scattering effect compared with the emission-absorption one. Finally, the gas real index is always supposed equal to one.

The spectrum is observed between 2379 cm^{-1} and 2397cm^{-1}, which is in the range of high sensitivity with respect to the temperature, so we hope pull out a good estimation of temperature's parameters at least. For each wavenumber σ we measure, by a detector placed at $x = d$, the resulting signal of absorption and emission radiation of a quantity of CO_2 situated between 0 and d (see the figure 1).

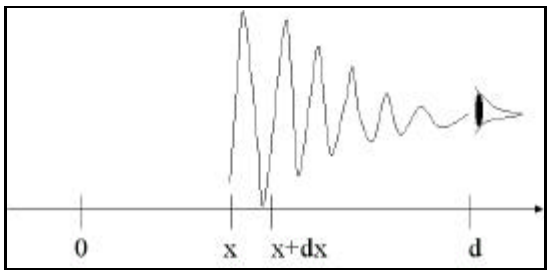


Figure 1: The observation

DIRECT AND INVERSE PROBLEM

These hypotheses are assumed for one-dimensional model and enable us to express the radiation intensity $L_\sigma(x)$ of a gas measured at x emitted by the quantity of the gas situated at the left of the point x. This radiation intensity $L_\sigma(x)$ depends on the temperature distribution T and the CO_2 concentration profile C between 0 and x. That dependence is given by either of the two following expressions :

- the differential formula :

$$\frac{dL_\sigma}{dx}(x) = K_\sigma(x)\left[L^0_\sigma(x) - L_\sigma(x)\right], \qquad L_\sigma(0) = L_0; \tag{1}$$

- the integral formula frequently sees in the RTE literature :

$$L_\sigma(x) = L_0 e^{-\int_0^x K_\sigma(v)dv} + \int_0^x K_\sigma(u) L^0_\sigma(u) e^{-\int_u^x K_\sigma(v)dv} du \tag{2}$$

Where K_σ is the monochromatic absorption coefficient. This coefficient is expressed in the form:

$$K_\sigma(T) = \left[a_0 + a_1 T + a_2 T^2 + a_3 T^3\right]\frac{PC}{10kT}. \tag{3}$$

Where P is the pressure, the coefficients (a_0, a_1, a_2, a_3) depend on the wavenumber σ, and are obtained from numerical study over a high-resolution spectroscopic database (HITEMP).

L^0_σ is the black body emission expressed by Planck's law:

$$L^0_\sigma(T) = \frac{k_1 \sigma^3}{\exp(\frac{k_2 \sigma}{T}) - 1}. \tag{4}$$

The measurement device is situated in the position d (see figure 1), it supplies an array of radiation measurements $L^{mes}_\Sigma = \left(L^{mes}_\sigma\right)_{\sigma\in\Sigma}$ for some set of wavenumbers Σ. We define the function:

$$F_\Sigma(T,C) = (f_\sigma(T,C))_{\sigma\in\Sigma} = (L_\sigma(d))_{\sigma\in\Sigma}. \quad (5)$$

So we have to solve the following inverse problem:

> Find the two profiles of temperature and CO_2 concentration (T,C) that satisfy
> $$F_\Sigma(T,C) = L_\Sigma^{mes}. \quad (6)$$

The least squares formulation of this problem is to find $m = (T, C)$ solution of:

$$\min_m \left\| F_\Sigma(m) - L_\Sigma^{mes} \right\|_{W,2}. \quad (7)$$

W is a weight vector balancing each measurement. This is an ill-posed problem for which existence, uniqueness and stability of the solution are difficult to prove. This problem is studied by application of the non-linear least squares theory. The study of the linearized operator, by the sensitivity analysis, gives an idea of the behavior of the differential operator. First we have to compute the function $F_\Sigma(T,C)$ as it is defined in (5).

DISCRETISATION OF THE RTE AND SENSITIVITY ANALYSIS

For a wavenumber $\sigma\in\Sigma$ we want to evaluate $f_\sigma(T,C)$. Let $(x_i)_{0\le i\le N}$ be a uniform subdivision of the segment $[0,1]$ with step $h = \frac{1}{N}$. We can approximate each function defined on $[0,1]$ by a piecewise constant function. So we associate a vector of N components to all of these functions: temperature $T = (T_i)_{1\le i\le N}$, CO_2 concentration $C = (C_i)_{1\le i\le N}$. We consider the following notations, where we distinguish by the subscript m vectors associated to T and C profiles:

$$\begin{aligned}
&m = (T, C).\\
&K_m = (K_{m,i}) = (K_\sigma(T_i, C_i))_{1\le i\le N}.\\
&L_m = (L_{m,i}) = (L^0_\sigma(T_i))_{1\le i\le N}.\\
&R_m = (R_{m,i})_{0\le i\le N} = (L_\sigma(x_i))_{0\le i\le N}.\\
&u_m = (K_m, L_m, R_m). \quad (8)
\end{aligned}$$

And we note:

$$\begin{aligned}
&u = (K, L, R) = ((K_i)_{1\le i\le N}, (L_i)_{1\le i\le N}, (R_i)_{0\le i\le N}).\\
&\bar{u} = (\bar{K}, \bar{L}, \bar{R}). \quad (9)
\end{aligned}$$

Notice that

$$f_\sigma(T,C) = L_\sigma(d) = R_{m,N}. \quad (10)$$

The differential formula gives a recurrent relation and enables to compute $f_\sigma(T,C)$ efficiently.

$$\begin{aligned}
&R_{m,0} = L_0 = L_\sigma(0).\\
&R_{m,i} = (1 - hK_{m,i})R_{m,i-1} + hK_{m,i}L_{m,i}, i = 1..N. \quad (11)
\end{aligned}$$

We define the Lagrangian function:

$$\begin{aligned}
\Phi_\sigma(u, \bar{u}; m) = R_N &+ \sum_{i=1}^{N} (K_i - K_{m,i})\overline{K_i}\\
&+ \sum_{i=1}^{N} (L_i - L_{m,i})\overline{L_i}\\
&+ \sum_{i=1}^{N} (R_i - (1 - hK_i)R_{i-1} - hK_iL_i)\overline{R_i}.\\
&+ (R_0 - L_0)\overline{R_0}. \quad (12)
\end{aligned}$$

We notice that for all $\bar{u}$:

$$\Phi_\sigma(u_m, \bar{u}; m) = f_\sigma(T,C) \quad (13)$$

Where $u_m = (K_m, L_m, R_m)$ is the solution of the state-equations (8) and (11).

Let $\bar{u}$ be the vector that:

$$\frac{\partial \Phi_\sigma}{\partial u}(u, \bar{u}; m) = 0 \quad (14)$$

That gives the following adjoint-state equations:

$$\begin{aligned}
&\overline{R}_N = -1,\\
&\overline{R}_{i-1} = \overline{R}_i(1 - hK_i) && i = N,...,1.\\
&\overline{K}_i = h\overline{R}_i(L_i - R_{i-1}) && i = N,...,1.\\
&\overline{L}_i = hK_i\overline{R}_i && i = N,...,1. \quad (15)
\end{aligned}$$

This implies for $u = u_m$:

$$\frac{\partial \Phi_\sigma}{\partial T_i}(u_m, \bar{u}; m) = \frac{\partial f_\sigma}{\partial T_i}(T, C),$$
$$\frac{\partial \Phi_\sigma}{\partial C_i}(u_m, \bar{u}; m) = \frac{\partial f_\sigma}{\partial C_i}(T, C). \qquad (16)$$

Those two equations lead up to the following relations:

$$\frac{\partial f_\sigma}{\partial T_i}(T, C) = -\left[\overline{K_i}\,\frac{\partial K_\sigma}{\partial T_i}(T_i, C_i) + \overline{L_i}\,\frac{dL_\sigma}{dT_i}\right],$$

$$\frac{\partial f_\sigma}{\partial C_i}(T, C) = -\overline{K_i}\,\frac{\partial K_\sigma}{\partial C_i}(T_i, C_i). \qquad (17)$$

Finally we will obtain the following formula:

$$\frac{\partial f_\sigma}{\partial T_i}(T, C) = \frac{K_\sigma(T_i)}{T_i}\overline{K_i} - \left(a_1 + 2a_2 T_i + 3a_3 T_i^2\right)\frac{PC_i\,\overline{K_i}}{10kT_i} - \frac{k_2\sigma}{T_i^2}\overline{L_i}L_\sigma^0(T_i)\exp\left(\frac{k_2\sigma}{T_i}\right), \qquad (18)$$

$$\frac{\partial f_\sigma}{\partial C_i}(T, C) = -\left(a_0 + a_1 T_i + a_2 T_i^2 + a_3 T_i^3\right)\frac{P\overline{K_i}}{10kT_i}.$$

Therefore the derivative of the function is:

$$Df_\sigma(T, C) = \left[\left(\frac{\partial f_\sigma}{\partial T_i}\right)_{1\le i\le N}\left(\frac{\partial f_\sigma}{\partial C_i}\right)_{1\le i\le N}\right]. \qquad (19)$$

Then for a set of wavenumbers Σ, the evaluation of $F_\Sigma(T,C)$ and its derivative $DF_\Sigma(T,C)$ is an iterative procedure. $DF_\Sigma(T,C)$ is the sensitivity matrix [10] [11], its Singular Value Decomposition (SVD) gives:

$$DF_\Sigma(T,C) = U\Lambda V^t. \qquad (20)$$

V contains the right singular vectors, U the left ones, and Λ contains the singular values. We can read the sensitivity for parameters of different types on the components of these vectors. Sensitivity increases when we have high singular values. In order to avoid problems of weak sensitivity due to some scale differences, we normalize the different types of parameters. For Gaussian distribution of temperature and concentration (see figure 2), we plot (see figure 3), from left to right from top to bottom, the first eight right singular vectors, each one has the first half of its components in temperature parameters space, and the others in the concentration parameters one. We see that the look of the first right singular vector is nearly the same as temperature and concentration profiles, so Gaussian parameterizations of profiles is proper to use here.

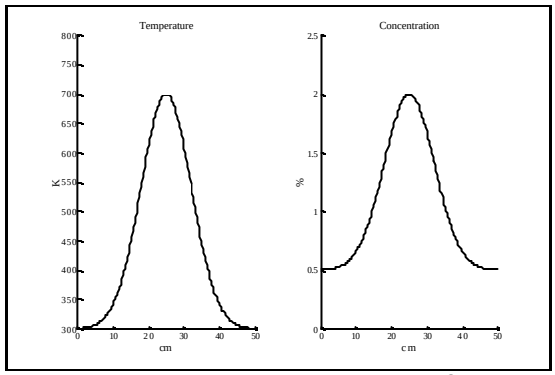


Figure 2: Gaussian distribution profiles.

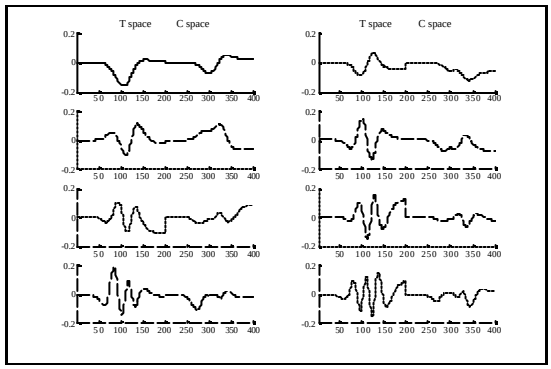


Figure 3: Right singular vectors for normalized parameters.

NOISE EFFECT

We note that for every temperature we have:

$$F_\Sigma(T,0) = 0. \qquad (21)$$

By a first order approximation we have :

$$F_\Sigma(T,C) = DF_\Sigma(T,C).[(T,C)-(T_1,0)]. \qquad (22)$$

Let T_1 tend to zero, by taking vector norms, we have:

$$\|F_\Sigma(T,C)\|_{R^I} \le \|DF_\Sigma(T,C)\|_{L(R^{2N},R^I)} \|(T,C)\|_{R^{2N}}. \qquad (23)$$

Let

$$m_0 = (T_0, C_0),$$
$$R^0_\Sigma = F_\Sigma(m_0). \qquad (24)$$

We note the greatest singular value:

$$\lambda_1(m_0) = \|DF_\Sigma(T,C)\|_{L(R^{2N},R^I)}. \qquad (25)$$

For some perturbation on measurements ε, we have to search the perturbation on the parameters δm, which minimizes the following function $J(\delta m)$:

$$\frac{1}{2}\left\langle \left(R^0_\Sigma + \varepsilon\right) - F_\Sigma(m_0+\delta m), \left(R^0_\Sigma + \varepsilon\right) - F_\Sigma(m_0+\delta m)\right\rangle. \qquad (26)$$

Its associated linearized problem is:

$$J_{lin}(\delta m) = \frac{1}{2}\left\langle \varepsilon - DF_\Sigma(m_0).\delta m, \varepsilon - DF_\Sigma(m_0).\delta m\right\rangle. \qquad (27)$$

The decomposition of δm on the base of right singular vectors and the decomposition of ε on the left singular vectors give:

$$\delta m = \sum_{i=1}^{2N} \delta a_i v_i$$
$$\varepsilon = \sum_{i=1}^{I} \varepsilon_i u_i. \qquad (28)$$

The solution of the linearised problem satisfies:

$$\varepsilon_i = \lambda_i \delta a_i,\ i = 1,..., I \qquad (29)$$

This equality combined with inequality (23) gives the following estimation:

$$S/N(a_i) \ge S/N(R^0_{\sigma i}) + (\lambda_i/\lambda_1)_{dB},\ i = 1,..,I. \qquad (30)$$

S/N signifies signal to noise ratio, a_i is the value of i-th parameter, $R^0_{\sigma i}$ is the value of i-th measurement, and λ_i is the i-th singular value. When the number $S/N(a_i)$ is positive a significant value of the i-th parameter is expected. So we can define the quantity of information non-hidden by the noise as the number i_N of parameters which still have a significant value:

$$i_N = \text{card}\left\{ i/\ -(\lambda_i/\lambda_1)_{dB} < \min_j S/N(R^0_{\sigma j}) \right\}. \qquad (31)$$

This number depends on temperature and concentration distributions, and it differs from a wavenumber set to another.
We consider Gaussian distributions for both temperature and concentration (see figure 2).

We will show that some subsets of wavenumbers could contain almost all the information. So we consider the following four subsets of wavenumbers between 2379 cm^{-1} and 2397 cm^{-1} (see figure 4): Peaks (29 wavenumbers) corresponding to local maxima or 'turned back maxima' of spectrum intensity, Troughs (32 wavenumbers) corresponding to local minima of spectrum intensity, Peaks and Troughs (61 wavenumbers), and all the wave numbers (901 wavenumbers).

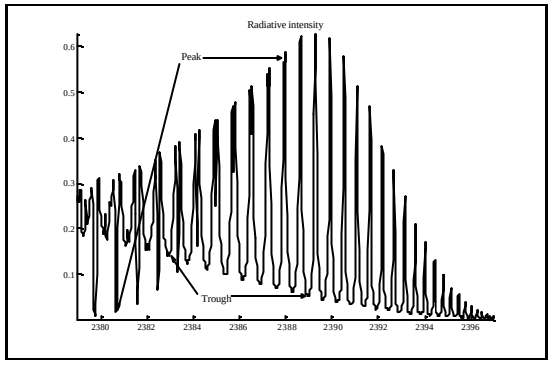


Figure 4: Definition of Peaks and Troughs .

In figure 5 we plot $-(\lambda_i/\lambda_1)_{dB}$ for the greatest eight singular values, and for each one of

the previous sets of wavenumbers, the number i_N, corresponding to a noise level of 1%, is evaluated in order to compare these sets. Peaks and Troughs set realizes i_N=6, which is nearly the same number corresponding to the set of all wavenumber. Consequently the set of Peaks and Troughs contains nearly almost all of the total information.

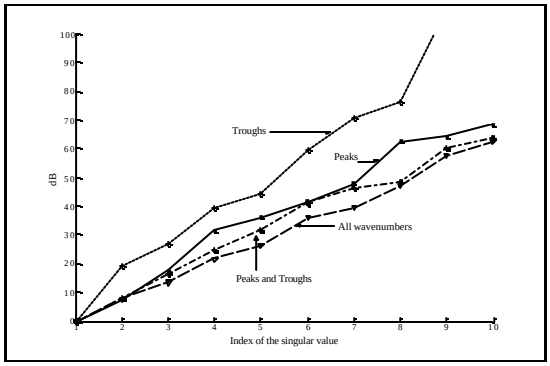


Figure 5: Comparaison of the four sets of wavenumbers for a Gaussian distribution.

PARAMETRIZATION

In order to discretize the differential equation, we use about 200 geometric points. This yields 200x2 parameters in order to reconstruct profiles, which is impossible to identify. Regularization is necessary to reduce the difficulty of this problem

We have used several parameterizations or a priori profiles as a regularization method. The main example here is the parameterization of a Gaussian profile, it is characterized by four parameters per profile :

$$T(x) = T_\infty + (T_{max} - T_\infty)\exp\left[-\left(\frac{x - x_T}{\alpha_T}\right)^2\right]$$

$$C(x) = C_\infty + (C_{max} - C_\infty)\exp\left[-\left(\frac{x - x_C}{\alpha_C}\right)^2\right] \quad (32)$$

When we have a good estimation of some parameters, such as T_∞ and C_∞, a sub-parameterization is considered, then identification is made on the reduced set of parameters.

For the new parameterization a sensitivity analysis is done to find the maximum number of estimable parameters associated to noise level of 1%. For Gaussian parameterization, there are at most 4 retrievable parameters for a noise level of 1%, see figure 6.

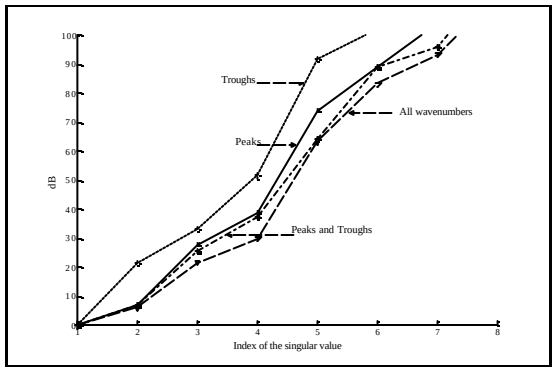


Figure 6: Number of retrievable parameters for Gaussian parameterization.

The precedent study shows that the set of Peaks and Troughs contains lot enough of information relatively to the set of all wavenumbers. So it will be used for further tests.

For Gaussian parameterization, we have to choose a subparameterization with only four parameters to identify. To guide the choice, we plot the first right singular vectors (see figure 7), corresponding to the greatest singular values. The first one indicates clearly to the first parameter T_{max}, so it has more chance to be retrieved. With maximal values T_{max} and C_{max}, widths α_T and α_C are chosen to be retrieved. An appropriate estimation of other parameters is used; for example boundary values give an estimation of T_∞ and C_∞.

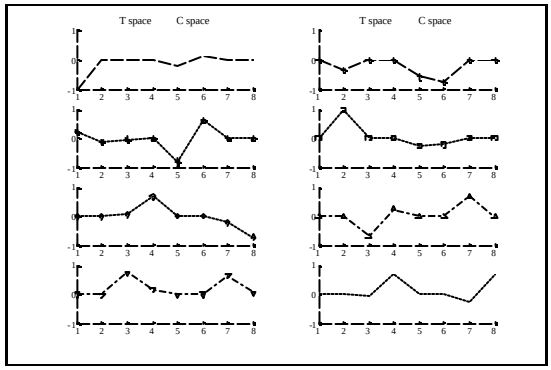


Figure 7: Right singular vectors for Gaussian parameterization.

For this subparameterization, similar study for a noise level of 1% shows that we hope to identify 2 parameters.

OPTIMIZATION AND IDENTIFICATION

Identification is performed by optimizing the following objective function:

$$J_\Sigma(p) = \frac{1}{2}\sum_{\sigma\in\Sigma} W_\sigma\left(L_\sigma^{mes} - L_\sigma(p)\right)^2 \qquad (33)$$

In this formula, W_σ is the weight coefficient corresponding to σ, and p refers to the parameterization.

A trust region reflective method was used where Gauss-Newton and Conjugate Gradient algorithms are the main tools used for each iteration.

In the following identification tests on synthetic data with noise, we try to retrieve profiles and see the effect of the noise on parameters values.

Noise is generated as uniform random distribution with maximal amplitude equals to γ percent of the maximal spectral measurement.

The first identification result on synthetic data relative to a noise level of 1% affirms the sensitivity analyzing conclusion, (see table 1). So we have retrieved T_{max} with a good precision, then comes α_T with some error of 3.79%. The values of the other parameters have less precision. With a higher noise level of 10%, T_{max} retrieved value still has a good precision of 0.325%.

Table 1: synthetic test of noise effect .

Noise level γ	T_{max} K	C_{max} %	α_T cm	α_C cm
0	700	2	28	28
1%	699.40	2.097	26.940	26.673
10%	697.73	1.654	33.899	32.636

The previous study is effected for other types of profiles for example double humps profiles and piecewise profiles, and every type of profile has its proper results So sensitivity analysis for double humps profiles shows that at most eight parameters of 400 parameters could be retrieved, but with parameterization of 10 parameters 5 parameters could be retrieved which is relatively better.

The following real case identification could conclude this study.

REAL CASE

Measurements are effected on the exhaust of Rolls-Royce Spey at DeRA, in order to develop non-intrusive method.

Chosen for being not very far from the reality, a Gaussian a priori profile is adopted to do inversion. So we hope to identify at most 4 parameters. These are chosen to be the maximal value and the width of each profile.

Compared with intrusive measurements which are direct measurements, we have attained reasonable results (see figures 8 and 9).

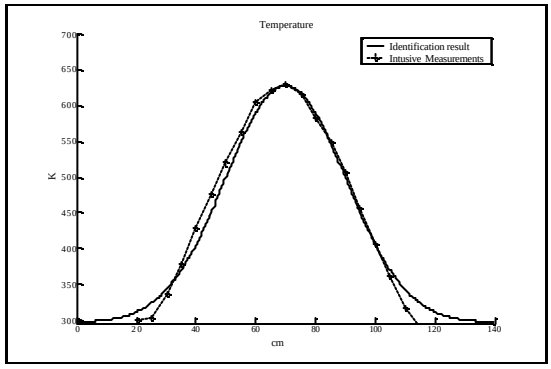


Figure 8: Identification result and intrusive measurement for temperature profile.

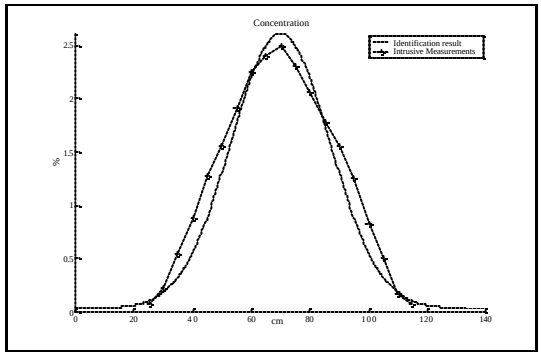


Figure 9: Identification result and intrusive measurement for concentration profile.

CONCLUSION

Inversion of spectroscopic data of CO_2 emission gives good results for retrieving temperature and CO_2 concentration profiles, when an a priori profile is used as a regularization way for the ill-posed problem. Adjoint state method accelerates Jacobian computation required by Gauss-Newton method, where conjugate gradient is used to minimize objective function. For higher noise level temperature profile is steadier than concentration one, and still be close to the solution. With real data, the identification result is too close to the intrusive measurements, that assures good confidence and accuracy in our inversion algorithme.

KEYWORDS

Inversion of spectroscopic data, SVD, Parameterization, Gauss-Newton, Conjugate Gradient.

ACKNOWLEDGEMENTS

We acknowledge here DERA for their results of intrusive measurements, and the company AUXITROL for their supports.

REFERENCES

[1] F. Yousefian, M. Lallemand, 1998. *Inverse radiative analysis of high-resolution infrared emission data for temperature and species profiles recoveries in axisymetric semi-transparent media. J. Quant. Spectrosc. Radiat. Transfer* Vol. 60, No. 6, pp. 921-931, 1998.

[2] F. Yousefian, M. Sakami, M. Lallemand, 1999. *Recovery of temperature and species concentration profiles in flames using low-resolution infrared spectroscopy. J. Heat Transfer* Vol. 121, pp. 268-279, May 1999.

[3] L. H. Liu, 2000. *Simultaneous identification of temperature profile and absorption coefficient in one-dimensional semitransparent meduim by inverse radiation analysis. J. Heat and Mass Transfer* Vol. 27, No. 5, pp. 635-643, 2000.

[4] L. H. Liu, H. P. Tan, 2001. *Inverse radiation problem in three-dimensional complicated geometric systems with opaque boundaries. J. Quant. Spectrosc. Radiat. Transfer* Vol 68, pp. 559-573, 2001.

[5] H. M Park, T. Y. Yoon, 1999. *Solution of inverse radiative problems using the Karhunen-Loève Galerkin procedure. J. Quant. Spectrosc. Radiat. Transfer* Vol 68, pp. 489-506, 2001.

[6] H. M. Park, T. Y. Yoon, 2000. *Solution of the inverse radiation problem using a conjugate gradient method. J. Heat and Mass Transfer* Vol 43, pp. 1767-1776, 2000.

[7] L. S. Rothman *et al*. 1998. *The HITRAN molecular spectroscopic database and HAWKS (HITRAN Atmospheric Workstation): 1998 edition*. Papar appeared in *J. Quant. Spectrosc. Radiat. Transfer* Vol. 60, No. 5, pp. 665-710, 1998.

[8] B. Lavaud, N. Kabir, G. Chavent 1999. *Pushing AVO inversion beyond linearised approximation*. Paper to appear in Geophysical Journal International.

[9] G.Chavent 1991. *On the theory and practice of non-linear least-squares* Water Resources, 1991, Vol. 14, No. 2, Page 55.

[10] G. H. Golub, C. F. Van Loan, 1989. *Matrix computation, second edition*. The John Hopkins University Press, Baltimore.

[11] A. Bjorck, 1996. *Numerical methods for Least Squares problems*. SIAM, Philadelphia, PA, U.S.A 1996.

A NOVEL INVERSE PROBLEM IN MEDICAL EMISSION IMAGING

Mai K. Nguyen
Equipe Traitement des Images et du Signal
CNRS/E.N.S.E.A/Université de Cergy-Pontoise
Cergy-Pontoise. France
nguyen@ensea.fr

T.T. Truong
Laboratoire Physique Théorique et Modélisation
CNRS/Université de Cergy-Pontoise
Cergy-Pontoise. France
tuong.truong@ptm.u-cergy.fr

H.D. Bui
Laboratoire de Mécanique des Solides
Ecole Polytechnique
Palaiseau, France
bui@lms.polytechnique.fr

ABSTRACT

This paper presents a recent result on the feasibility of reconstruction of the radioactivity distribution of an object from a sequence of Compton-scattered radiation data in emission imaging. This may be regarded as a novel inverse principle as opposed to the traditional one in which the object is reconstructed only from the non-scattered rays collected at different incident directions. The new inversion procedure is described by a invertible linear integral transform which may be viewed as generalized Radon transform and has several advantages over the old one. It improves significantly the contrast of the reconstructed image. The required data for reconstruction is easily acquired from an energy and space measuring gamma camera under the form of scattered distribution images classified by their Compton-scattering angle. The motion of the camera in standard tomographic procedure is no longer necessary for data taking.

NOMENCLATURE

$f(\mathbf{V})$	: object density function.
$\tilde{g}(\mathbf{D},\theta)$ *or* $g(\mathbf{D},t)$	: image density function.
n_e	: electron density in medium.
$p(\mathbf{D},t\|\mathbf{V})$	: Point Spread Function.
t	$= tan\theta$.
$\mathcal{T}$	: integral transformation.
$Y(s)$	: Heaviside unit step function.
μ	: absorption coefficient.
θ	: Compton scattering angle .

INTRODUCTION

In gamma-ray emission imaging the goal is to reconstruct the object under study (e.g. body organ) from photon intensity distributions measured on a gamma camera, hereafter called images. However, this image is considerably affected by the unavoidable Compton-scatter effect. Up until now, most image reconstruction methods discard the scattered photons and operate only with primary (or non-scattered) ones. Nevertheless, such restored images remain unsatisfactory, due to the very small number of detected primary photons (the noise in the image increases unacceptably with scatter elimination) and statistical fluctuations. The solution to the scatter problem remains a major technical challenge in gamma-ray imaging.

In the present work, we propose an alternative approach by taking advantage of the properties of scattered photons and show that these photons can be used to improve the image quality. In fact, we establish a new relation between the object radioactivity density and a series of images formed by the *single*-scattered photons and *parameterized* by various scatter angles. This relation appears as a weighted sum of conical linear integral transforms. It may be viewed as a generalized Radon Transform (integral on cone surfaces instead of planes). We name the kernel of the new transform SPSF (Scatter Point Spread Function). Moreover and most importantly, we prove that this new transformation is *invertible*. These theoretical results allow us to reconstruct a three-dimensional object from a series of secondary angle-parameterized images. This possibility demonstrates a new imaging principle which is based on Compton-scattering properties and in which the motion (displacement or rotation) of the detector is no longer necessary.

The paper is organized as follows. Based on a

physical analysis of Compton scattering we first establish a so-called imaging equation which expresses the secondary image density function as a linear conical integral transform of the object density function and explicitly evaluate the kernel (SPSF) of this mapping. We next present the most important property of this mapping which is the existence of an inverse mapping with explicit kernel. Conclusions and comments are given in the last section.

SECONDARY IMAGE FORMATION BY COMPTON-SCATTERED PHOTONS

Compton scattering

Our point of view in this work is to focus on the emitted photons which undergo at least one Compton scattering and study how they may turn out to be relevant for image formation process [1].

First, we recall the relation between energy and scattering angle for the Compton-scattered photon [2]:

$$E = E_0 \frac{1}{1+\varepsilon(1-\cos\theta)} \tag{1}$$

where θ is the scattering angle as measured from the incident photon direction; E_0, the photon initial energy; $\varepsilon = \frac{E_0}{mc^2}$ and mc^2 the rest energy of the electron . Eq. (1) shows that single-scattered photons have a continuous energy spectrum: $0 \leq E \leq E_0$, which is related to the scattering angle θ. But Compton scattering is also a quantum phenomena: the emergence of the scattered photon has a probability of occurrence given by the Compton differential cross-section :

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2} P(\theta) \tag{2}$$

where r_e is the classical radius of the electron and $P(\theta)$ the so-called Klein-Nishina probability for deflection by an angle θ:

$$P(\theta) = \frac{1}{[1+\varepsilon(1-\cos\theta)]^2}$$

$$\left[1+\cos^2\theta + \frac{\varepsilon^2(1-\cos\theta)^2}{1+\varepsilon(1-\cos\theta)}\right]. \tag{3}$$

As a result of Compton scattering, photons leaving an emitting point source can enter the detector along the collimator's direction after one or more scattering. However, since single-scattered photons dominate the process, we shall limit ourselves to single-scattered photons.

The imaging equation

Thus at given angle θ, let $\tilde{g}(\mathbf{D},\theta)$ be the photon flux density at detector site $\mathbf{D}$. This quantity

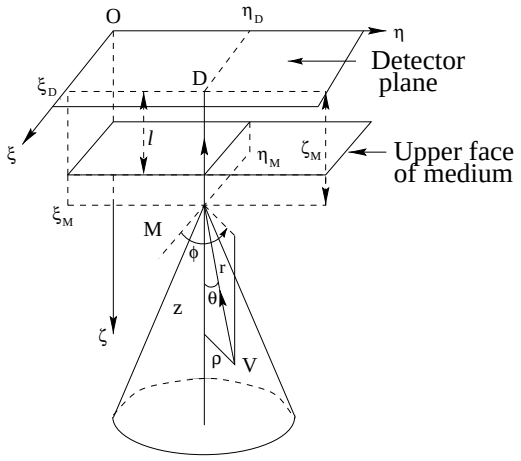


Figure 1: Coordinate system for the calculation of $\mathcal{T}$

describes essentially a *secondary* emission imaging process since it is based on secondary emission sites formed by the free electrons of the surrounding medium. In conventional image processing for γ-ray camera, they are discarded by filtering or by other correction techniques. We study now the secondary image formation process by such scattered photons.

Let $\mathbf{V}$ denote an object voxel of coordinates (ξ_V, η_V, ζ_V) and $f(\mathbf{V})$ be the object density function, defined as $f(\mathbf{V}) = f(\xi_V, \eta_V, \zeta_V)$. This is also the number of photons emitted per unit time and per unit object (or source) volume, uniformly distributed around the 4π solid angle at site $\mathbf{V}$. Thus in one given direction, making an angle θ with the vertical downward direction $O\zeta$ (see Fig. 1), the number of photons emitted in a small solid angle $d\Omega_M$ around site $\mathbf{M}$ by an elemental source volume $d\mathbf{V}$, per unit time is:

$$\frac{1}{4\pi} f(\mathbf{V}) d\mathbf{V} d\Omega_M. \tag{4}$$

Now, by definition of the solid angle we have $d\Omega_M = \frac{d\sigma_M}{MV^2}$, where $d\sigma_M$ is the area element around site $\mathbf{M}$ normal to the direction VM. Consequently the flux of photons arriving at $\mathbf{M}$ in the VM direction is:

$$\frac{f(\mathbf{V})d\mathbf{V}}{4\pi} \frac{1}{MV^2} \exp(-\mu r) \tag{5}$$

where $r = VM$ and μ is the linear mean coefficient of absorption in biological tissues (i.e. organ and surrounding medium).

But site $\mathbf{M}$ is located either in the studied organ or in the surrounding biological medium, and

there are $n_e(\mathbf{M})d\mathbf{M}$ free electrons in a small volume element $d\mathbf{M}$. Note that in general the density of electrons is not uniform. Now since $(\frac{d\sigma}{d\Omega})d\Omega_D$ is the differential cross-section of photons scattering in the θ-direction, and since $d\Omega_D = \frac{d\sigma_D}{MD^2}$, the number of photons reaching a unit detector surface at $\mathbf{D}$ per unit time (after division by the area $d\sigma_D$) is the flux density recorded by the detector at site $\mathbf{D}$:

$$\frac{f(\mathbf{V})d\mathbf{V}}{4\pi}\,\frac{e^{-\mu r}}{MV^2}\,n_e(\mathbf{M})\,d\mathbf{M}\,\frac{r_e^2}{2}\,P(\theta)\,\frac{e^{-\mu(\zeta_M-l)}}{MD^2}, \tag{6}$$

which depends on $d\mathbf{V}$ and $d\mathbf{M}$.

Consequently, the number of photons recorded per unit time and unit detector area at site $\mathbf{D} = (\xi_D, \eta_D)$, $\tilde{g}(\mathbf{D}, \theta)$, is due to all emitting point sources $\mathbf{V}$ situated on a cone with opening angle θ, axis parallel to $O\zeta$ and with apex the scattering site on the vertical line MD:

$$\tilde{g}(\mathbf{D},\theta) = \int d\xi_M d\eta_M \frac{d\zeta_M}{\zeta_M^2}$$

$$\delta(\xi_D - \xi_M)\delta(\eta_D - \eta_M) n_e(\mathbf{M}) e^{-\mu(\zeta_M - l)}$$

$$\int \frac{f(\mathbf{V})d\mathbf{V}}{4\pi}\,\frac{\delta(Cone)\exp(-\mu r)}{MV^2}\,\frac{r_e^2}{2}P(\theta), \tag{7}$$

where $\delta(Cone)$ restricts the integration over $\mathbf{V}$ to the circular cone . If one uses the local spherical coordinates centered at $\mathbf{M}$ in figure 1, we have $d\mathbf{V} = r^2 dr \sin\alpha d\alpha d\phi$ then $\delta(Cone) = \frac{1}{r}\delta(\theta - \alpha)$. Clearly $\tilde{g}(\mathbf{D}, \theta)$ has the dimension of a photon flux through a plane (flux density).

Thus the whole process defines the integral mapping:

$$\mathcal{T} : f(\mathbf{V}) \longmapsto \tilde{g}(\mathbf{D}, \theta). \tag{8}$$

Note that $\mathcal{T}$ maps a function of three variables $f(\mathbf{V})$ to $\tilde{g}(\mathbf{D}, \theta)$ which is also a function of three scalar variables, since $\mathbf{D} \in \mathcal{R}^2$ and $\theta \in \mathcal{R}$.

Explicit expressions of $\mathcal{T}$

In order to focus only on Compton effect features, other image degradations (absorption, imperfect detector) are now neglected and we set $n_e =$ constant and $\mu = 0$. This is a plausible hypothesis since soft biological tissues have a mass density close to water [2].

In Fig. 1, the measuring apparatus formed by the collimator, the detector and the photomultiplier bank is collapsed into a rectangle in the horizontal plane $O\xi\eta$. Moreover, we assume that the organ under study lies below the plane $\zeta = l$.

The coordinates of $\mathbf{V}$ in this system are:

$$\begin{aligned} \xi_V &= \xi_M + r\sin\theta\cos\phi, \\ \eta_V &= \eta_M + r\sin\theta\sin\phi, \\ \zeta_V &= \zeta_M + r\cos\theta, \end{aligned}$$

where θ is the scattering angle and ϕ the azimuthal angle of $\mathbf{V}$ with respect to the cone axis DM. The integration measure on the cone is $r\sin\theta d\phi dr$. Equation (7) after integration over ξ_M and η_M becomes our basic *imaging equation*:

$$\tilde{g}(\mathbf{D},\theta) = \tilde{K}(\theta)\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}\int_0^{2\pi} d\phi \int_{0_+}^\infty \frac{dr}{r}$$

$$f(\xi_D + r\sin\theta\cos\phi, \eta_D + r\sin\theta\sin\phi, \zeta_M + r\cos\theta), \tag{9}$$

where the factor $K(\theta)$ contains terms dependent on θ:

$$\tilde{K}(\theta) = \frac{n_e}{4\pi}\frac{r_e^2}{2}P(\theta)\sin\theta. \tag{10}$$

In this expression of $\tilde{g}(\mathbf{D}, \theta)$, the integral on r is formally divergent near the origin since $f(\mathbf{V})$ is usually a bounded function (the activity density is everywhere finite). Mathematically, the regularization of the integral in equation (9) is achieved by assigning a cut-off for the integration in r, here the lower integration bound is noted 0_+. In fact in practical numerical calculations this cut-off can be easily realized by appropriate discretization of the integration range .

As $f(\mathbf{V})$ is of compact support, this integral can be understood as the sum, along a vertical line, of integrals of $f(\mathbf{V})$ on cones of axis parallel to $O\zeta$, apex $\mathbf{M}$ and opening angle θ. The image of $f(\mathbf{V})$ is thus represented by a set of *conical integrals* instead of planar integrals. In this sense one may view this transformation $\mathcal{T}$ as a *generalized* Radon transform.

We can rewrite the transformation $\mathcal{T}$ in a more suggestive form using $t = \tan\theta$, $\tilde{K}(\theta) = K(t)$ and $\tilde{g}(\mathbf{D}, \theta) = g(\mathbf{D}, t)$, as :

$$g(\mathbf{D},t) = \int d\mathbf{V}\, p(\mathbf{D}, t|\mathbf{V})\, f(\mathbf{V}). \tag{11}$$

where $p(\mathbf{D}, t|\mathbf{V})$ stands for the kernel of the transformation $\mathcal{T}$:

$$p(\mathbf{D},t|\mathbf{V}) = K(t)\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}\int_0^\infty \frac{dr}{r}\int_0^{2\pi} d\phi$$

$$\delta(\xi - \xi_D - r\sin\theta\cos\phi)\,\delta(\eta - \eta_D - r\sin\theta\sin\phi)$$

$$\delta(\zeta - \zeta_M - r\cos\theta). \tag{12}$$

Physically this kernel is also called the Point Spread Function (PSF), or image of a point source at site $\mathbf{V} = (\xi, \eta, \zeta)$. Nevertheless this image is due to *single scattered* photons rather than primary photons. This is why we call this kernel SPSF (Scatter PSF).

Computation of the SPSF kernel

Using the Fourier representation of the δ-function, and computing the integral in ζ_M as follows:

$$\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}\delta(\zeta-\zeta_M-r\cos\theta)=\frac{Y(\zeta-l-r\cos\theta)}{(\zeta-r\cos\theta)^2}, \tag{13}$$

with $Y(\zeta)$ as the Heaviside unit step function. We can now represent the kernel $p(\mathbf{D},t|\mathbf{V})$ by its Fourier transform $\mathcal{P}(u,v|\zeta,t)$:

$$p(\mathbf{D},t|\mathbf{V})=\int\int dudv\,\mathcal{P}(u,v|\zeta,t)$$
$$\exp 2i\pi[(\xi-\xi_D)u+(\eta-\eta_D)v],$$

with:

$$\mathcal{P}(u,v|\zeta,t)=2\pi K(t)\int_{0_+}^\infty \frac{dr}{r}J_0(2\pi r\sin\theta\sqrt{u^2+v^2})$$
$$\frac{Y(\zeta-l-r\cos\theta)}{(\zeta-r\cos\theta)^2}. \tag{14}$$

This result show that $\mathcal{P}(u,v|\zeta,t)$ is a function of the sole variable $\sqrt{u^2+v^2}$. We can now reobtain the SPSF by Fourier transformation. But in view of the rotational symmetry in Fourier space, the Fourier transform is equivalent to a Hankel transform of order zero [3]. If we introduce the following polar coordinates $(\xi-\xi_D)=\rho\cos\alpha$, $(\eta-\eta_D)=\rho\sin\alpha$ and $u=\sigma\cos\psi$, $v=\sigma\sin\psi$, then :

$$p(\mathbf{D},t|\mathbf{V})=2\pi\int_0^\infty \sigma\,d\sigma\,J_0(2\pi\rho\sigma)\,\mathcal{P}(\sigma|\zeta,t). \tag{15}$$

Now observing that $z=r\cos\theta=\rho/t$ where $t=\tan\theta$ and rearranging the expression of $\mathcal{P}(u,v|\zeta,t)$ as Hankel transform of order zero of the function (or distribution):

$$\frac{1}{z^2}\frac{Y(\zeta-z-l)}{(\zeta-z)^2},$$

we have:

$$p(\mathbf{D},t|\mathbf{V})=2\pi\int_0^\infty \sigma\,d\sigma\,J_0(2\pi\rho\sigma)\frac{K(t)}{t^2}\,2\pi$$
$$\int_0^\infty \rho\,d\rho\,J_0(2\pi\rho\sigma)\frac{1}{z^2}\frac{Y(\zeta-z-l)}{(\zeta-z)^2}. \tag{16}$$

Now Hankel's identity [4]:

$$(2\pi)^2\int \sigma\,d\sigma\,J_0(2\pi\rho'\sigma)\int \rho\,d\rho\,J_0(2\pi\rho\sigma)f(\rho)=f(\rho'),$$

allows to conclude that for scattering angles $0<\theta<\pi/2$, $(t>0)$, the SPSF has the expression:

$$\begin{aligned}p(\mathbf{D},t|\mathbf{V}) &= p(\xi_D,\eta_D,t|\xi,\eta\zeta)\\ &= t^2K(\theta)\tfrac{Y(t(\zeta-l)-\rho)}{\rho^2(t\zeta-\rho)^2}.\end{aligned} \tag{17}$$

The SPSF diverges as $\rho^{-2}\to\infty$ near the origin at constant θ, when the scattering site $\mathbf{M}$ coincides with the point source $\mathbf{V}$. It also diverges when the scattering site $\mathbf{M}$ comes to the detection site $\mathbf{D}$, or when $t\zeta=\rho$, but this will not occur because of the limit $\zeta>l$. This kernel is also translationally invariant in the detector plane:

$$p(\mathbf{D},t|\mathbf{V})=p((\xi-\xi_D),(\eta-\eta_D)|t,\zeta). \tag{18}$$

Now for scattering angles θ in $[\pi/2,\pi[$ or $(t<0)$, the SPSF is:

$$p(\mathbf{D},t|\mathbf{V})=t^2K(\theta)\frac{1}{\rho^2(|t|\zeta+\rho)^2}. \tag{19}$$

There is no need to insert the Heaviside function because the scattering point lies always lower than the point source. To sum up we have the following expression for the SPSF:

$$p(\mathbf{D},t|\mathbf{V})=Y(t)t^2K(\theta)\frac{Y(t(\zeta-l)-\rho)}{\rho^2(t\zeta-\rho)^2}$$
$$+Y(-t)t^2K(\theta)\frac{1}{\rho^2(|t|\zeta+\rho)^2}, \tag{20}$$

valid in the whole range of t values.

It is noted that the SPSF has the general shape of a "Mexican hat" for $0<\theta<\pi/2$ and of a simple peak at the origin for other angles. Moreover, it has rotational symmetry around the projection of the point source on the detector (horizontal) plane (see Fig. 2).

As examples illustrating our idea of taking advantage of Compton-scattered rays, we present in Fig. 3 a three dimensional object (cylinder) and in Fig. 4 some of its scattered-ray images at various deflection angles ($\theta=36,53$, and 90 degrees, which correspond to energy losses of -5%, -10% and -22% of the incident photon energy). It is clear that the scattered-ray images carry useful information on the object structure and deserve to be used in image reconstruction. Considering them as a pure noise and discarding them as usually done is certainly not the most appropriate solution for the improvement of image quality.

INVERSION OF THE TRANSFORM $\mathcal{T}$

Computation of the inverse kernel

Having explicitly computed the kernel of the transformation $\mathcal{T}$, we show now that this transformation is *invertible* [5] and compute also the expression of the inverse kernel. Introducing the two-dimensional Fourier transform $\tilde{F}(u,v,\zeta_M;\theta)$ of $f(\mathbf{V})$ under the form:

$$\tilde{F}(u,v,\zeta_M;\theta)=\int\int d\xi_D d\eta_D\,f(\xi_D,\eta_D,\zeta_M;\theta)$$

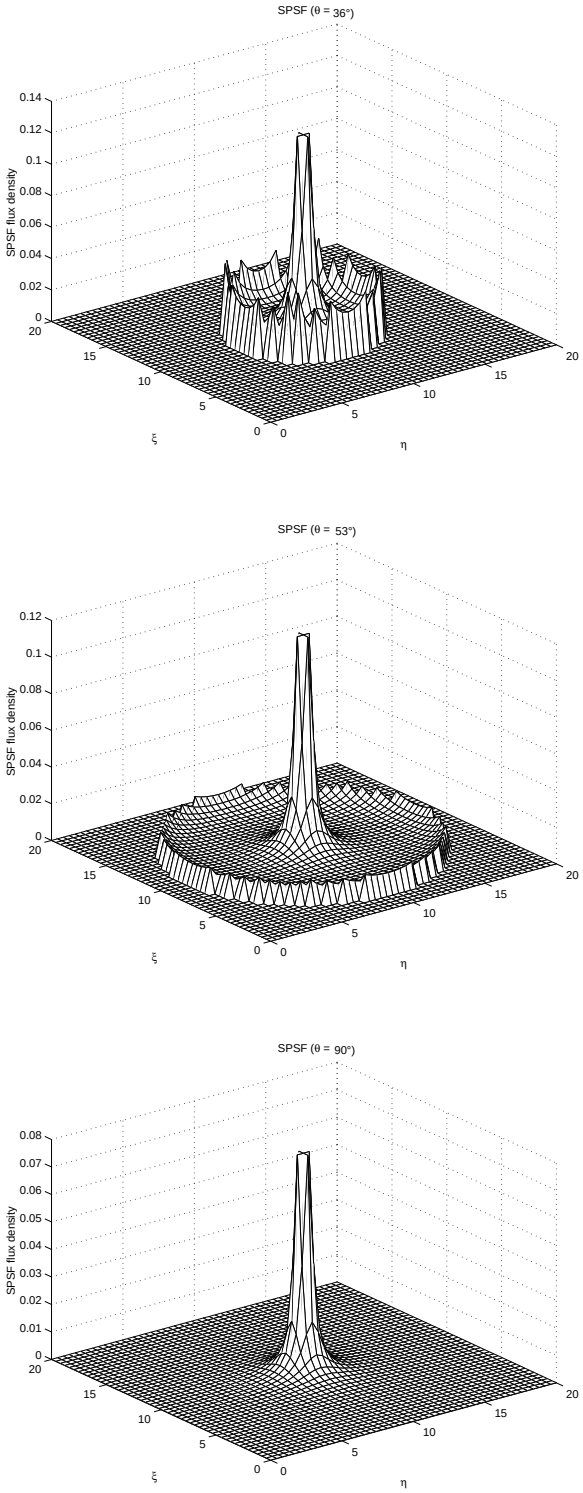


Figure 2: SPSF of the transform $\mathcal{T}$ at various scattering angles ($\theta = 36, 53$ and 90 degrees respectively) .

$$\exp -2i\pi(u\xi_D + v\eta_D), \tag{21}$$

and similarly $\tilde{G}(u,v;\theta)$ for $\tilde{g}(\mathbf{D},\theta)$, the imaging equation 9 reads now:

$$\tilde{G}(u,v;\theta) = 2\pi\tilde{K}(\theta)\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}\int_{0_+}^\infty \frac{dr}{r}$$

$$\tilde{F}(u,v,\zeta_M + r\cos\theta)J_0(2\pi r\sin\theta\sqrt{u^2+v^2}). \tag{22}$$

Now using the variables t and z defined earlier, we can cast this last equation as a Hankel transform of order zero (of course with the usual care for the lower bound of the integration):

$$G(u,v;t) = 2\pi K(t)\int_{0_+}^\infty z\,dz\,J_0(2\pi zt\sqrt{u^2+v^2})$$

$$\frac{1}{z^2}\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}F(u,v,\zeta_M+z). \tag{23}$$

The fact we integrate z over $]0,\infty[$, means that we consider here only scattering angles $\theta \in]0,\pi/2]$ or $t>0$ and $z>0$.

Thus inverting "formally" the Hankel transform (the lower integration bound is considered as to be 0), one obtains:

$$\frac{1}{z^2}\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}F(u,v,\zeta_M+z) = 2\pi(u^2+v^2)$$

$$\int_{0_+}^\infty t\,dt\,J_0(2\pi zt\sqrt{u^2+v^2})\frac{G(u,v;t)}{K(t)}. \tag{24}$$

In the other range of θ, i.e. $\theta \in [\pi/2,\pi[$ we have $t<0$ and $z<0$, the previous relation becomes:

$$\frac{1}{z'^2}\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}F(u,v,\zeta_M-z') = 2\pi(u^2+v^2)$$

$$\int_{0_+}^\infty t'\,dt'\,J_0(2\pi z't'\sqrt{u^2+v^2})\frac{G(u,v;-t')}{K(-t')}, \tag{25}$$

where $z=-z'$ and $t=-t'$ with $z'>0$ and $t'>0$. We can put the two inverse cases together in a single formula, valid for all t, as follows:

$$\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}F(u,v,\zeta_M+z) = 2\pi(u^2+v^2)\,z^2$$

$$\int_{0_+}^\infty t\,dt\,J_0(2\pi|z|t\sqrt{u^2+v^2})$$

$$\left[Y(z)\frac{G(u,v;t)}{K(t)} + Y(-z)\frac{G(u,v;-t)}{K(-t)}\right]. \tag{26}$$

Now the left hand side of the previous equation can be put under an alternative form using Fourier analysis:

$$\int_l^\infty \frac{d\zeta_M}{\zeta_M^2}F(u,v,\zeta_M+z) =$$

$$\int_{-\infty}^{\infty} dw\, \mathcal{J}_l(w) \bar{F}(u,v,w) \exp[-2i\pi\, w(l+z)], \quad (27)$$

where $\bar{F}(u,v,w)$ is the three dimensional Fourier transform of $f(\mathbf{V})$, $\mathcal{J}_l$ the Fourier transform of $Y(s)(s+l)^{-2}$ is given in Fourier tables as:

$$\mathcal{J}_l(w) = 2i\pi w$$
$$\{e^{2i\pi lw}\,[Ci(2\pi l|w|) - i\,\epsilon(w)\,Si(2\pi l|w|)] - \frac{i}{2\pi wl}\}, \quad (28)$$

here the variable s is related to ζ_M by $\zeta_M = (s+l)$ and $\epsilon(w)$ is the sign function of w. The variable $(z+l)$ may take now any value in $\mathcal{R}$.

To extract $\bar{F}(u,v,w)$, we make an inverse Fourier transform on the last form of the imaging equation (26):

$$\bar{F}(u,v,w) = \frac{1}{\mathcal{J}_l(w)} \int_{-\infty}^{\infty} dz\, e^{2i\pi w(z+l)}\,(u^2+v^2)\, z^2$$
$$2\pi \int_{0_+}^{\infty} t\, dt\, J_0(2\pi|z|t\sqrt{u^2+v^2})$$
$$\left[Y(z)\frac{G(u,v;t)}{K(t)} + Y(-z)\frac{G(u,v;-t)}{K(-t)}\right]. \quad (29)$$

Finally, performing the three dimensional Fourier transform on the previous equation (29) and reexpressing $G(u,v;t)$ as inverse two dimensional Fourier transform of $g(\mathbf{D},t)$, the inverse transform $\mathcal{T}^{-1}$ can be put as:

$$f(\mathbf{V}) = \int\int_{-\infty}^{\infty} d\mathbf{D} \int_0^{\infty} dt\, p^{-1}(\mathbf{V}|\mathbf{D},t)\, g(\mathbf{D},t), \quad (30)$$

with the inverse kernel:

$$p^{-1}(\mathbf{V}|\mathbf{D},t) =$$
$$\int\int_{-\infty}^{\infty} du\, dv\, e^{2i\pi[u(\xi-\xi_D)+v(\eta-\eta_D)]}(u^2+v^2)$$
$$2\pi \int_{-\infty}^{\infty} z\, dz\, J_0(2\pi|z|t\sqrt{u^2+v^2})\, z\, H_l(\zeta+z+l)\, t\mathcal{K}_t, \quad (31)$$

where:

$$\int_{-\infty}^{\infty} dw\, \exp[2i\pi(\zeta+z+l)w]\,\frac{1}{\mathcal{J}_l(w)} = H_l(\zeta+z+l),$$

and $\mathcal{K}_t$ is a discrete linear operator defined by:

$$\mathcal{K}_t\, g(\mathbf{D},t) = \left[Y(z)\frac{g(\mathbf{D},t)}{K(t)} + Y(-z)\frac{g(\mathbf{D},-t)}{K(-t)}\right].$$

Now by splitting up the integration range of z into two parts: $]-\infty,0]$ and $[0,\infty[$ and taking care of the action of $\mathcal{K}_t$, we obtain a new form of the inversion formula:

$$f(\mathbf{V}) = \int_{-\infty}^{\infty} d\xi_D \int_{-\infty}^{\infty} d\eta_D \int_0^{\infty} dt$$
$$\int\int_{-\infty}^{\infty} du\, dv\, e^{2i\pi[u(\xi-\xi_D)+v(\eta-\eta_D)]}(u^2+v^2)$$
$$2\pi \int_0^{\infty} z\, dz\, J_0(2\pi|z|t\sqrt{u^2+v^2})\, z$$
$$\left[H_l(\zeta+z+l)\frac{tg(\mathbf{D},t)}{K(t)} + H_l(\zeta-z+l)\frac{tg(\mathbf{D},-t)}{K(-t)}\right]. \quad (32)$$

The integration on z in Eq. (32) may be viewed as a Hankel transform of the function $z\,H_l(\zeta \pm z + l)$:

$$2\pi \int_0^{\infty} z\, dz\, J_0(2\pi|z|t\sqrt{u^2+v^2})\, z\, H_l(\zeta \pm z + l) =$$
$$\mathcal{H}^{\pm}(\zeta+l, t\sqrt{u^2+v^2}). \quad (33)$$

Then the two dimensional Fourier transform of the kernel $p^{-1}(\mathbf{V}|\mathbf{D},t)$ reads simply:

$$\mathcal{P}^{-1}(u,v|\zeta,t) = (u^2+v^2)\mathcal{K}_t(u^2+v^2,\zeta), \quad (34)$$

where $\mathcal{K}_t(u^2+v^2,\zeta)$ is a linear discrete operator defined by:

$$\mathcal{K}_t(u^2+v^2,\zeta)g(\mathbf{D},t) =$$
$$[\frac{t\mathcal{H}^+(\zeta+l,t\sqrt{u^2+v^2})}{K(t)}g(\mathbf{D},t) +$$
$$\frac{t\mathcal{H}^-(\zeta+l,t\sqrt{u^2+v^2})}{K(-t)}g(\mathbf{D},-t)]. \quad (35)$$

The integration over t in Eq. (32) has the meaning that one must use a *series of images labelled by the scattering angle* θ to reconstruct the object. Thus this angle θ plays the role of a scanning parameter (spatial rotation angle) in a standard tomographic imaging procedure. This inverse kernel is obviously translationally invariant in the detector plane. Let $\mathbf{S}$ be the projection of $\mathbf{V}$ on the detector plane, then we may write:

$$p^{-1}(\mathbf{V}|\mathbf{D},t) = p^{-1}(\mathbf{S},\zeta|\mathbf{D},t) = p^{-1}(\mathbf{DS}|\zeta,t).$$
$$= \int\int_{-\infty}^{\infty} du\, dv\, e^{2i\pi[u(\xi-\xi_D)+v(\eta-\eta_D)]}$$
$$(u^2+v^2)\mathcal{K}_t(u^2+v^2,\zeta). \quad (36)$$

Proof of the inversion relation

In this section we shall give the proof of:

$$\int d\mathbf{D}\, dt\, p^{-1}(\mathbf{V}|\mathbf{D},t)p^*(\mathbf{D},t|\mathbf{V}') = \delta(\mathbf{V}-\mathbf{V}') \quad (37)$$

where the star means complex conjugate. By going to the two dimensional Fourier transform, the left hand side can be written as:

$$\int_0^{\infty} dt \int\int_{-\infty}^{\infty} du\, dv\, e^{2i\pi[u(\xi-\xi')+v(\eta-\eta')]}$$

$$\mathcal{P}^{-1}(u,v|\zeta,t)\mathcal{P}(u,v|\zeta',t). \qquad (38)$$

In order to perform this calculation an alternative form for $\mathcal{P}(u,v|\zeta',t)$ using a Fourier representation is needed. We first express:

$$\frac{Y(\zeta'-z'-l)}{(\zeta'-z')^2} = \int_{-\infty}^{\infty} dw'\, \mathcal{J}_l(w') e^{2i\pi(\zeta'-z'-l)w'}, \qquad (39)$$

and insert it in:

$$\mathcal{P}(u,v|\zeta,t) = 2\pi K(t) \int_{0_+}^{\infty} \frac{dz'}{z'} J_0(2\pi z' t\sqrt{u^2+v^2})$$

$$\int_{-\infty}^{\infty} dw'\, \mathcal{J}_l(w') e^{2i\pi(\zeta'-z'-l)w'}. \qquad (40)$$

Now using the expression of $\mathcal{P}^{-1}(u,v|\zeta,t)$ in Eq. (34) and this representation for $\mathcal{P}(u,v|\zeta',t)$ we can perform the t integration:

$$\int_0^{\infty} dt \mathcal{P}^{-1}(u,v|\zeta,t)\mathcal{P}(u,v|\zeta',t) =$$

$$2\pi \int_0^{\infty} dt\, K(t) \int_{0_+}^{\infty} \frac{dz'}{z'} J_0(2\pi z' t\sqrt{u^2+v^2})$$

$$\int_{-\infty}^{\infty} dw'\, \mathcal{J}_l(w') e^{2i\pi(\zeta'-z'-l)w'} 2\pi \frac{t}{K(t)}(u^2+v^2)$$

$$\int_{-\infty}^{\infty} z\, dz\, J_0(2\pi|z|t\sqrt{u^2+v^2})\, z \int_{-\infty}^{\infty} dw \frac{e^{2i\pi(\zeta+z+l)w}}{\mathcal{J}_l(w)}, \qquad (41)$$

and we see that the $K(t)$ cancel out. The integration on t can be carried out using the identity:

$$\int_0^{\infty} \sigma\, d\sigma\, J_0(z'\sigma) J_0(z\sigma) = \frac{1}{z'}\delta(z-z'). \qquad (42)$$

Thus after integration over dz for $t>0$ and also $z>0$, we get:

$$\int_0^{\infty} dt \mathcal{P}^{-1}(u,v|\zeta,t)\mathcal{P}(u,v|\zeta',t) =$$

$$\int_0^{\infty} \frac{dz'}{z'} \int_{-\infty}^{\infty} dw' \mathcal{J}_l(w') e^{2i\pi(\zeta'-z'-l)w'}$$

$$\frac{1}{z'}\delta(z-z') \int_{-\infty}^{\infty} z^2\, dz \int_{-\infty}^{\infty} dw\, \frac{1}{\mathcal{J}_l(w)} e^{2i\pi(\zeta+z+l)w}. \qquad (43)$$

The integration with respect to z' removes all the z factors, and we are left with:

$$\int_0^{\infty} dt \mathcal{P}^{-1}(u,v|\zeta,t)\mathcal{P}(u,v|\zeta',t) =$$

$$\int_{-\infty}^{\infty} dz\, e^{2i\pi(w-w')} \int_{-\infty}^{\infty} dw'\, \mathcal{J}_l(w') e^{2i\pi w'(\zeta'-l)}$$

$$\int_{-\infty}^{\infty} dw\, \frac{1}{\mathcal{J}_l(w)} e^{2i\pi w(\zeta+l)}$$

$$= \int\int_{-\infty}^{\infty} dw\, dw' \mathcal{J}_l(w') \frac{\delta(w-w')}{\mathcal{J}_l(w)}$$

$$e^{2i\pi w'(\zeta'-l)}\, e^{2i\pi w(\zeta+l)}$$

$$= \int_{-\infty}^{\infty} dw\, e^{2i\pi(\zeta'-\zeta)w} = \delta(\zeta'-\zeta). \qquad (44)$$

The remaining integrations over du and dv yield obviously $\delta(\xi'-\xi)\,\delta(\eta'-\eta)$. The product of the kernels $p^{-1}(\mathbf{V}|\mathbf{D},t)$ with the SPSF kernel $p(\mathbf{D},t|\mathbf{V})$ is thus the identity.

Further mathematical analysis of the inverse problem in Hadamard sense (existence, uniqueness, stability of the solution) goes beyond the scope of this paper and shall not included in the present work. Numerical calculations lead to approximatively the same numerical problems as in known Radon transform, but they are more involved because of the conical geometry. They depend on concrete applications, instrumental constraints, acquisition configurations, number of measures, choice of regularization methods, optimization algorithm and discrete implementation. Thus they are highly technical and inappropriate for inclusion here.

CONCLUSION AND COMMENTS

In this work we have given a theoretical result showing the feasibility of object reconstruction using Compton-scattered rays in gamma emission imaging. The new linear integral transform as well as its kernel (SPSF) are derived directly from a Compton-scattering physical analysis [5], which has been strongly recommended to be included in quantitative imaging by experts of the field [6]. The new imaging principle proposed here takes advantage of properties of scattered rays instead of rejecting them as usually done in most actual scatter correction methods. It will improve the signal-to-noise ratio, and consequently the contrast. Moreover, the fact that the scattering angle is a free parameter is used here as a disposable parameter to collect data, the multi-views of the object are obtained without the need of moving the detector. This possibility is particularly interesting in the applications where the number of measures is very limited, e.g. in non destructive control. Several perspectives of the proposed approach are possible: for example extension of this result obtained in emission imaging to transmission imaging and to higher order Compton scattering.

References

[1] Mai K. Nguyen, C. Fay, L. Eglin and T. T. Truong, Apparent image formation by Compton scattered photons in gamma-ray imaging

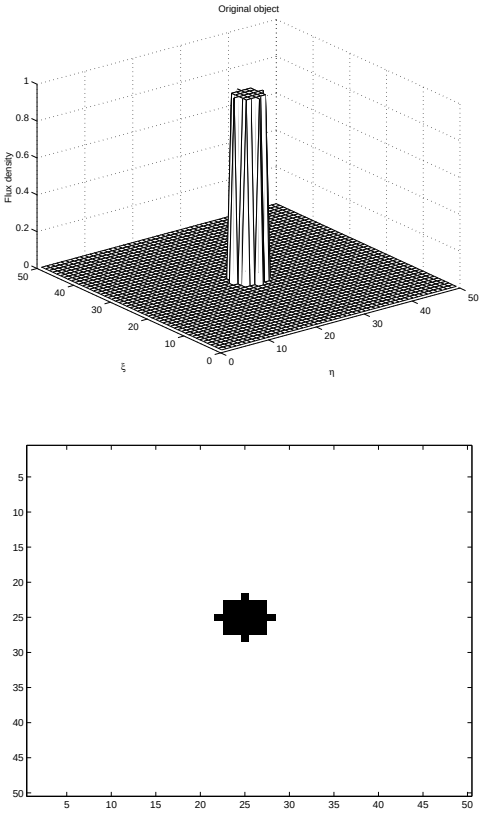


Figure 3: A cylinder original object and its cross-section.

IEEE Signal Processing Letters, **8** (9), pp. 248-251, 2001.

[2] H. H. Barrett and W. Swindell, *Radiological Imaging* I and II, Academic Press, New York, (1981).

[3] B. Davies, *Integral Transforms and their Applications*, Springer, New York, (2001).

[4] N. N. Lebedev, *Special Functions and their Applications* Dover, New York, (1955)

[5] Mai K. Nguyen and T. T. Truong, On an integral transform and its inverse in nuclear imaging, *Inverse Problems*,**18** (1),pp. 265-277, 2002.

[6] M. S. Rosenthal, J. Cullom, W. Hawkins, S. C. Moore, B. M. W. Tsui and M. Yester, "Quantitative SPECT imaging : A review and recommendations by the focus committee of the society of nuclear medicine computer and instrumentation council," *J. Nucl. Med.*, Vol. 36, pp. 1489-1513, 1995.

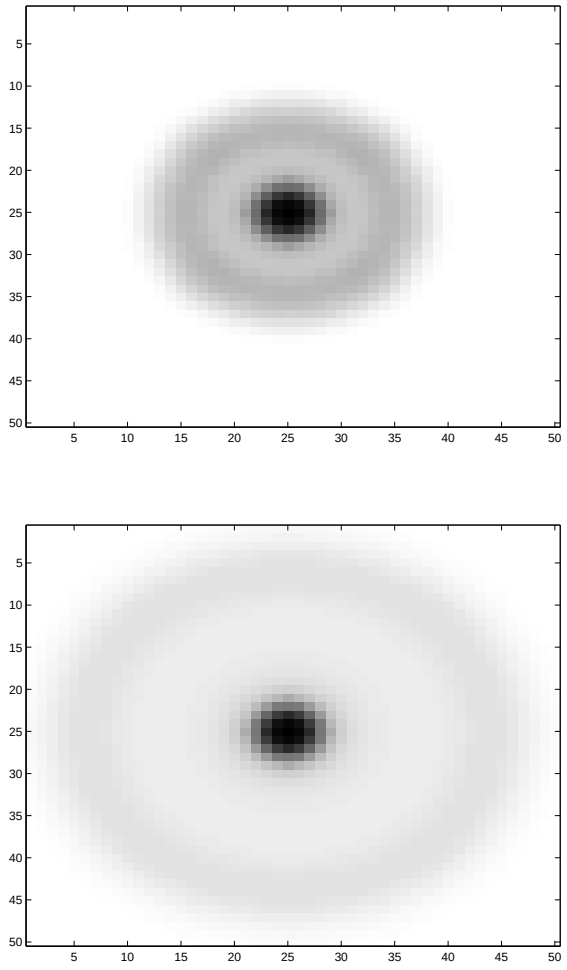


Figure 4: Series of secondary images parameterized by scattering angles ($\theta = 36, 53$, and 90 degrees respectively) .

DETERMINATION OF THE ELECTRONIC DENSITY IN A MEDIUM BY AN INVERSE METHOD BASED ON DOUBLE-COMPTON SCATTERING IN TRANSMISSION IMAGING.

T.T. Truong
Laboratoire de Physique Théorique et Modélisation.
Université de Cergy-Pontoise
Cergy-Pontoise. France
nguyen@ensea.fr

Mai K. Nguyen
Equipe Traitement des Images et du Signal
ENSEA/Université de Cergy-Pontoise
Cergy-Pontoise. France
tuong.truong@ptm.u-cergy.fr

H. D. Bui
Laboratoire de Mécanique des Solides.
Ecole Polytechnique
Palaiseau, France
bui@lms.palaiseau.fr

C. Daveau
Laboratoire de Mathématiques, Equipe d'Analyse numérique
Université Paris XI, Orsay
Orsay, France
Christian.Daveau@math.u-psud.fr

ABSTRACT

We propose a new method for reconstructing defects in bulk materials, via the recovery of the electronic volume density n_e in the material, using doubly Compton-scattered gamma photons in transmission imaging modality. We establish an integral relationship between the photon flux density after a double scattering and the product of electronic densities at two different sites. This photon flux density at different scattering angles can be measured by an energy-position detector in a fixed configuration. Thus the reconstruction of n_e can be formulated as an inverse problem of the integral relation, which is bilinear in the electronic densities. This theoretical result opens the way for a new imaging principle, which exploits scattered radiation rather than discarding it as in most existing imaging procedures. In this new procedure, the motion of the detector is no longer necessary as it is the case of conventional tomography.

NOMENCLATURE

$g(\mathbf{D}, \mathbf{S}\|\tau)$	: flux density at detection site $\mathbf{D}$ due to incoming pencil source beam having a trace at $\mathbf{S}$ on the detector.
$G(\mathbf{D}, \tau)$	: flux density at site $\mathbf{D}$ for given τ.
l	: distance slab-detector.
L	: thickness of slab.
$n_e(\mathbf{M})$	: electronic density at site $\mathbf{M}$. also noted $n_e(\mathbf{D}, \tau)$.
n_e^0, n_e^1	: constant electronic density in surrounding medium and in defect.
Φ_0	: constant incident flux density.
ρ	: distance $\|\mathbf{DS}\|$.
τ	: cotangent of the scattering angle θ.

INTRODUCTION

The goal in non-destructive control is to obtain information (location, form, representative parameters) about defects in bulk materials. This is up to now done by X-ray or gamma ray transmission imaging. In these methods of investigation, the defect is represented by its linear attenuation distribution μ. The data consists of line integrals of the attenuation density along lines joining source points to detector points. A three dimensional reconstruction of the volume defect can be realized if a complete set of data is obtained with various directions of the incident source beam. However this may not be always realizable in practical situations where the geometry of the piece of material imposes severe restrictions on the motion of the radiation source.

In this paper, we describe an alternative method of investigation of defects. Instead of representing the material under study by its absorption function, as usually done, we shall describe it alternatively by its electronic density n_e. Indeed at the location of defects (inclusions, voids or cracks) n_e will change drastically and will exhibit discontinuous jumps in values. The distribution of scattered photons will change accordingly. The determination of n_e from single scattered photons (in-

stead of non-scattered ones) has been introduced in Compton Scatter Tomography [6]. But no analytical inversion procedure is available. In the present work we propose an analytical inversion method for obtaining n_e from the series of doubly Compton-scattered photon distributions *at various scattering angles*. Moreover we show that in this procedure the motion of the detector is no longer necessary and consequently may be more interesting from the point of view of operational level [2, 3].

In order to focus solely on the scattering aspect, the attenuation mainly due to photoelectric effect shall be neglected here. This hypothesis is acceptable in the low energy range of gamma photons. In a transmission imaging system with collimated detector, if the material slab is of thickness L and limited by parallel planes, for incoming gamma photons perpendicularly to the slab to emerge parallel to the incident direction, there must be at least two Compton collisions. This is why we propose to exploit the properties of double Compton scattering for determining n_e, since higher order collisions have much weaker probability of occurrence.

The paper is organized as follows. In the next section, we establish the basic relation connecting the flux density of photons $G(\mathbf{D}, \tau)$ - image measured on the detector plane - to $n_e(\mathbf{r})$ - the electronic distribution in the bulk material. This relation will be called also the *"imaging equation"*. Next, we introduce the concept of *Transmission Pencil Source Function* (TPSF), which plays the role of the well-known PSF in emission imaging. Then images of cracks of simple form are computed as illustrations with the use of the TPSF. In the following section, the *inverse problem* of determining the volume extent of the crack using the measured flux density $G(\mathbf{D}, \tau)$, is considered. It is shown that an *analytical* solution exists if one makes use of the whole *set of data collected at all scattering angles* θ. This theoretical result opens the way for a new transmission imaging principle in which one takes advantage of the double Compton scattering whereby the complete data is obtained without moving the detector. Conclusions and outlook are given in the last section.

IMAGING EQUATION

We consider an incident gamma-ray beam of initial energy E_0 and of constant flux density Φ_0 on a slab of material of thickness L, in which the electronic density is $n_e(\mathbf{r})$. Normally the gamma rays at this energy will go through the object, if no attenuation is assumed. They are collected on the plane of a collimated gamma camera, which has an axis parallel to the incident beam, and located at a distance l from the slab face.

However part of the gamma rays will undergo

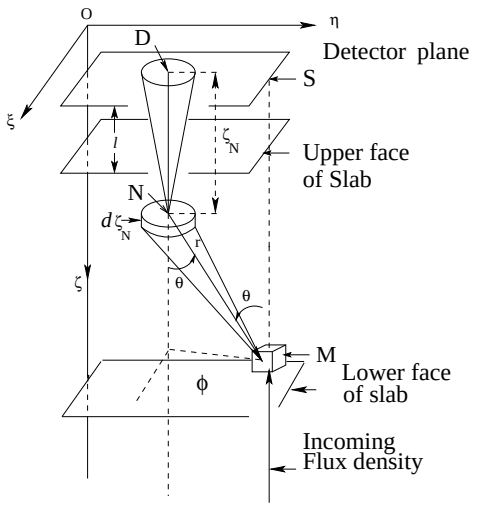


Figure 1: Schematic representation of the functioning principle

Compton scattering, and actually two scattering if they are to emerge parallel to the incident direction (see figure 1). They form images *parameterized by the scattering angle* θ, or alternatively by the outgoing photon energy E. The photon energy recorded after two collisions with the same scattering angle θ is given by the Compton formula:

$$E = \frac{E_0}{1 + 2\epsilon(1 - \cos\theta)},$$

where $\epsilon = \frac{E_0}{mc^2}$ is the ratio of the initial photon energy to the rest energy of the electron. This formula has the same form as the formula for one collision, but with 2ϵ instead of ϵ.

Let Φ_0 be the constant incident flux density. Now the flux density of photons $\Phi(\mathbf{M})$, scattered at a site $\mathbf{M}\,(\xi_M, \eta_M, \zeta_M)$ (first collision site) in the bulk material into a solid angle $d\Omega_N$ is given by the expression:

$$\Phi_0 r_e^2 P(\theta) d\Omega_N n_e(\mathbf{M}) d\mathbf{M} = \Phi(M) d\sigma_N$$

where $d\Omega_N = d\sigma_N / MN^2$, $d\sigma_N$ being the elementary surface subtended at site $\mathbf{N}\,(\xi_N, \eta_N, \zeta_N)$(second collision site). The differential Compton cross section is the term $r_e^2 P(\theta)$ where r_e is the classical electron radius and $P(\theta)$, the so-called Klein-Nishina probability function [1].

$$P(\theta) = \frac{1}{2[1 + \varepsilon(1 - \cos\theta)]^2}$$
$$\left[1 + \cos^2\theta + \frac{\varepsilon^2(1 - \cos\theta)^2}{1 + \varepsilon(1 - \cos\theta)}\right]. \qquad (1)$$

The flux density of photons arriving at site $\mathbf{N}$ is then obtained by dividing out $d\sigma_N$. This flux will be scattered next at site $\mathbf{N}$ into the solid angle $d\Omega_D$. Finally the flux density photons collected at detection site $\mathbf{D}$ $((\xi_D, \eta_D, 0)$ reads:

$$\Phi_0 r_e^2 P(\theta) \frac{n_e(\mathbf{M})}{MN^2} d\mathbf{M} r_e^2 P(\theta) \frac{n_e(\mathbf{N})}{ND^2} d\mathbf{N}.$$

In figure 1, we have given also the coordinate system to be used in the calculations to come. The scattering site $\mathbf{M}$ has coordinates:

$$\begin{aligned} \xi_M &= \xi_D + r \sin\theta' \cos\phi, \\ \eta_M &= \eta_D + r \sin\theta' \sin\phi, \\ \zeta_M &= \zeta_N + r \cos\theta'. \end{aligned}$$

where (r, θ', ϕ) are the local spherical coordinates centered at $\mathbf{N}$.

The total flux density $G(\mathbf{D}, \tau)$ on the detector is the summation on contributions of:

• for given $\mathbf{N}$, all sites $\mathbf{M}$ in the medium on a cone of apex $\mathbf{N}$, opening angle θ and axis $O\zeta$,

• all sites $\mathbf{N}$ on a vertical at site $\mathbf{D}$, verifying $l < \zeta_N < \infty$, since the slab is put at a distance l from the detector.

Thus we introduce the corresponding integration measures:

$$\begin{aligned} \delta(Cone) &= r^{-1}\delta(\theta' - \theta) \\ \delta(Line) &= \delta(\xi_N - \xi_D)\delta(\eta_N - \eta_D). \end{aligned}$$

Consequently the photon flux density at $\mathbf{D}$ is given by:

$$G(\mathbf{D}, \tau) = [r_e^2 P(\theta)]^2 \int_l^\infty \delta(Line)\, d\xi_N d\eta_N d\zeta_N$$

$$\int \frac{1}{r}\delta(\theta' - \theta) r^2 dr d\theta' \sin\theta' d\phi \frac{1}{r^2} n_e(\mathbf{M}) \frac{1}{\zeta^2} n_e(\mathbf{N}) \Phi_0. \tag{2}$$

(Recall that Φ_0 is the incident flux density of photons.)

Some integrations are now performed in order to eliminate the δ-functions. The result may be expressed in terms of the constant:

$$K(\tau) = [r_e^2 P(\theta)]^2 \sin\theta.$$

where $\tau = \cot\theta$ and reads as:

$$\frac{G(\mathbf{D}, \tau)}{K(\tau)} = \int_l^\infty \frac{d\zeta_N}{\zeta_N^2} n_e(\mathbf{D}, \zeta_N) \int_{0_+}^\infty \frac{dr}{r} \int_0^{2\pi} d\phi$$

$$n_e(\xi_D + r\sin\theta\cos\phi, \eta_D + r\sin\theta\sin\phi, \zeta_N + \cos\theta)\Phi_0.$$

Introducing now the projection of site $\mathbf{M}$ on the detector plane $\mathbf{S}$, we observe that $\xi_S = \xi_D + r\sin\theta\cos\phi$, $\eta_S = \eta_D + r\sin\theta\sin\phi$ and $\zeta_S = 0$. We may put the result as:

$$G(\mathbf{D}, \tau) = K(\tau) \int_{0_+}^\infty \frac{dr}{r} \int_0^{2\pi} d\phi$$

$$\int_l^\infty \frac{d\zeta_N}{\zeta_N^2} n_e(\mathbf{D}, \zeta_N) n_e(\mathbf{S}, \zeta_N + r\cos\theta)\Phi_0, \tag{3}$$

here this expression is evidently bilinear in the electron densities. Let us call $\rho = r\sin\theta$, the distance between sites $\mathbf{D}$ and $\mathbf{S}$,i.e. $|\mathbf{D} - \mathbf{S}| = \rho$, then we can rewrite conveniently this expression as:

$$G(\mathbf{D}, \tau) = K(\tau) \int_{0_+}^\infty \frac{\rho d\rho}{\rho^2} \int_0^{2\pi} d\phi$$

$$\int_l^\infty \frac{d\zeta_N}{\zeta_N^2} n_e(\mathbf{D}, \zeta_N) n_e(\mathbf{S}, \zeta_N + \rho\tau)\Phi_0,$$

this can be rearranged as an integral over $\mathbf{S}$ as:

$$G(\mathbf{D}, \tau) = \int_{beamsection} d\mathbf{S} g(\mathbf{S}, \mathbf{D}|\tau)\Phi_0, \tag{4}$$

where $g(\mathbf{S}, \mathbf{D}|\tau)$, is defined as *the Transmission Pencil Source Function* (TPSF) of the problem. *The TPSF represents the transmission image by double Compton scattering of a unit flux density pencil source at infinity incoming normally on the detector plane.* This function will play the role of the PSF function in emission imaging. It is given by:

$$g(\mathbf{S}, \mathbf{D}|\tau) = \frac{K(\tau)}{\rho^2} \int_l^\infty \frac{d\zeta_N}{\zeta_N^2} n_e(\mathbf{D}, \zeta_N) n_e(\mathbf{S}, \zeta_M). \tag{5}$$

where $\zeta_M = (\zeta_N + \rho\tau)$. This function describes thus the illumination of the material slab by an incoming pencil beam and serves to explore the structure of defects in the bulk.

For an arbitrary incoming beam with finite section, e.g. from an extended source far away, $\phi_0 = \phi_0(\mathbf{S})$ and this must be taken care of in the integration standing in equation (4).

IMAGES OF SIMPLE SYSTEMS

The purpose of this section is to illustrate the imaging mechanism of the TPSF on simple systems. As already mentioned, gamma rays can be used to reconstruct variations of the electronic density at very small (atomic) scales. But for the localization of macroscopic defects in materials, these details are irrelevant and homogeneous medium can be considered to have a constant electronic density equal to its mean macroscopic value [4]. Defects may appear as cracks or voids where the electronic density drops brutally to zero or filled volumes with a different value of the electronic density. And this is precisely what we wish to detect.

1-Image of a homogeneous material slab

As the electronic density is assumed to be constant and equal to n_e^0, the TPSF can be explicitly

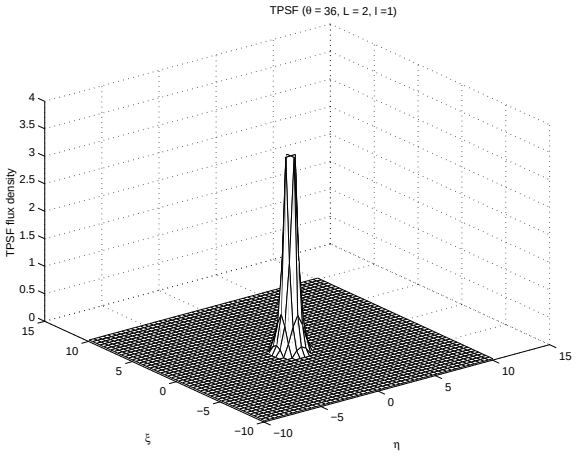


Figure 2: TPSF Flux density at $\theta = 36$ degrees

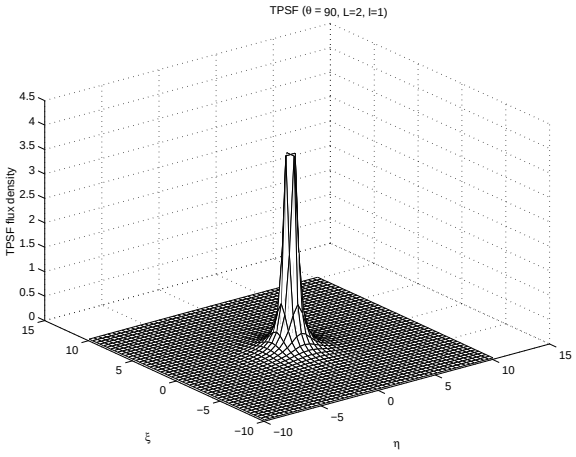


Figure 3: TPSF Flux density at $\theta = 90$ degrees

evaluated:

$$\frac{g(\mathbf{S},\mathbf{D}|\tau)}{K(\tau)(\frac{n_e^0}{\rho})^2} = \begin{cases} Y(L-\rho\tau)\frac{(L-\rho\tau)}{l(l+L-\rho\tau)} & \text{if } \tau > 0 \\ Y(L+\rho\tau)\frac{(L+l+\rho\tau)}{(l+L)(-\rho\tau)} & \text{if } \tau < 0 \\ \frac{L}{l(l+L)} & \text{if } \tau = 0. \end{cases} \quad (6)$$

Here $Y(x)$ stands for the Heaviside unit step function. Rotational symmetry around the incident pencil beam is reflected by the sole dependence on ρ. This image will serve as basis for detecting defects.Fig 2 and Fig.3 give a representation of the TPSF at two scattering angles.

At 90 degrees scattering angle, the TPSF has a very simple behavior, i.e. its varies as ρ^{-2}.

Finally the form of the peak changes according to the ratio l/L.

2-Image of a linear defect

Consider a linear defect situated on a line perpendicular to the detector plane at a site $\mathbf{S}_0$ on the detector plane. The defect has a length $h << L$ and its middle point is at a distance ζ_0 from $\mathbf{S}_0$. On the defect there is an electronic density n_e^1, different from the slab electronic density n_e^0 . The question is now how would the TPSF detect this defect? Two situations are possible:

• if the defect is not exactly positioned on the direction of the incoming beam and the detection site $\mathbf{D}$ different from $\mathbf{S}_0$, then the measured flux density $g(\mathbf{S},\mathbf{D}|\tau)$ is the same as in the case of the homogeneous slab (see equation 6). But if $\mathbf{D}$ is precisely at $\mathbf{S}_0$ then the measured flux density is, for $\tau > 0$, given by:

$$g(\mathbf{S},\mathbf{D}|\tau) = \frac{K(\tau)}{\rho_0^2}$$

$$\left\{ \frac{(n_e^0)^2(L-\rho_0\tau)}{l(l+L-\rho_0\tau)} - \frac{[(n_e^0)^2-(n_e^1)^2]h}{\zeta_0^2 - \frac{1}{4}h^2} \right\}, \quad (7)$$

where $\rho_0 = |\mathbf{S}_0\mathbf{D}|$. When $n_e^1 \to 0$, the defect becomes a crack, and when $h \to 0$ we get the limit of the point defect. Then it is clear that the TPSF cannot "see" a point defect since we recover the flux density of equation 6.

• if the probing pencil beam falls directly on the defect, then one would measure, also for $\tau > 0$ or the scattering angle $0 < \theta < \pi/2$:

$$g(\mathbf{S},\mathbf{D}|\tau) = \frac{K(\tau)}{\rho^2}$$

$$\left\{ \frac{(n_e^0)^2(L-\rho\tau)}{l(l+L-\rho\tau)} - \frac{[(n_e^0)^2-(n_e^1)^2]h}{(\zeta_0-\rho\tau)^2 - \frac{1}{4}h^2} \right\}. \quad (8)$$

Thus a linear vertical defect would cause a jump in the recorded flux density as compared to the pattern of a homogeneous slab. For $\tau < 0$ or a scattering angle larger than $\pi/2$, the main features of the pictures remains the same but the expressions are slightly modified.

3-Image of a rectangular defect

Having studied in detail the case of a linear perpendicular defect in the previous section, it is now easy to treat many other cases. The simplest one which can be generated is the case of the rectangle perpendicular to the detector with center at (ρ_0, ζ_0), with sides (h', h) respectively parallel and orthogonal to the detector.

• Despite the non-zero width h', the behavior of the TPSF with respect to this kind of defect is exactly the same as in the case of the linear defect of the previous subsection, if the incident pencil beam does not cross the defect and if the detection line (the line perpendicular to the detector at $\mathbf{D}$)

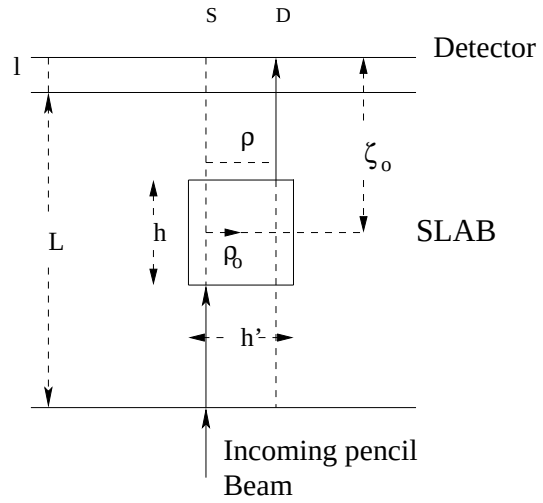


Figure 4: Rectangular defect plane with **S** and **D** inside the plane.

crosses the defect. To see the spread of the defect parallel to the detector one should move **D** along the detector plane so that $(\rho_0 - h'/2) < \rho < (\rho_0 + h'/2)$.

• The interesting case occurs when the lines perpendicular to the detector at **S** and at **D** are inside the rectangle and at a distance $\rho < h'$.

Then there are two values of the flux density depending on τ :

a) if $\rho\tau \leq h$ and $\rho < h'$ we have:

$$g(\mathbf{S},\mathbf{D}|\tau) = \frac{K(\tau)}{\rho^2}\,\{$$

$$2n_e^0 n_e^1\left(\frac{(\zeta_0-\rho\tau)}{(\zeta_0-\rho\tau)^2-\frac{1}{4}h^2}-\frac{\zeta_0}{\zeta_0^2-\frac{1}{4}h^2}\right)+$$

$$(n_e^1)^2\left(\frac{(h-\rho\tau)}{(\zeta_0-\frac{1}{2}h)(\zeta_0-\rho\tau+\frac{1}{2}h)}+\right.$$

$$\left.(n_e^0)^2\left(\frac{(L-\rho\tau)}{l(l+L-\rho\tau)}-\frac{(h+\rho\tau)}{(\zeta_0+\frac{1}{2}h)(\zeta_0-\rho\tau-\frac{1}{2}h)}\right)\right\}$$

b) but if $\rho\tau > h$ and $\rho < h'$ we have:

$$g(\mathbf{S},\mathbf{D}|\tau) = \frac{K(\tau)}{\rho^2}\left\{(n_e^0)^2\left(\frac{(L-\rho\tau)}{l(l+L-\rho\tau)}\right.\right.$$

$$\left.-\frac{h}{(\zeta_0-\rho\tau)^2-\frac{1}{4}h^2}-\frac{h}{\zeta_0^2-\frac{1}{4}h^2}\right)$$

$$\left.+n_e^0 n_e^1\left(\frac{h}{(\zeta_0-\rho\tau)^2-\frac{1}{4}h^2}+\frac{h}{\zeta_0^2-\frac{1}{4}h^2}\right)\right\} \qquad (9)$$

From these results one can draw several conclusions:

- for $h \to 0$, we get the limit of an infinitely thin horizontal defect: the flux density is that of the homogeneous medium, in other words it is "invisible" by the TPSF. Also it is easy to generalize to a horizontal disk defect or to any horizontal planar defect: they cannot be detected by the TPSF.
- for $n_e^1 \to 0$, we get the limit of a rectangular crack. This crack does present a non-zero "thickness" h with respect to the direction of the incoming probing radiation.
- There are similar results for $\tau < 0$.

4-Image of an arbitrary defect

From the previous considerations, we can make general observations on how images are formed by the TPSF. First of all, it must present a certain "thickness" with respect to the incoming pencil radiation. Then we can decompose the problem into planar imaging problem by considering the intersection of the plane (π) which contains the two parallel lines perpendicular to the detector at **D** and at **S** and the the slab of material to be investigated.

If these lines do not intersect the section of a defect (Σ) in (π), then the flux density recorded on the detector is that of a homogeneous medium. Now if the perpendicular line at **D** intersects (Σ), then the measured flux density has the pattern of a linear perpendicular defect studied in the previous subsection. Note that the height of such a defect varies as site **D** moves on the detector plane, and in the previous subsection we have looked at a particular section (Σ), which is a simple rectangle. Of course, the interesting case would be the case where both perpendicular lines at **D** and **S** intersect (Σ), but this times the intersection lengths are no longer equal. The general effect remains similar.

So in principle, a general image obtained of an arbitrary defect of nonzero measure in the vertical direction by the TPSF, can be constructed from the TPSF image of a finite linear perpendicular defect. The question is now how the defect in the bulk can be reconstructed in space when a set of images, labelled by τ, or equivalently by the scattering angle θ has been collected beforehand.

THE INVERSE PROBLEM

This is the problem of reconstructing the electronic density $n_e(\mathbf{r})$ from measurements made on the detector. In this section, we show that this problem has a solution, provided that certain working conditions are assumed. More precisely, we show that the TPSF images can be used effectively to reconstruct the electronic density as follows.

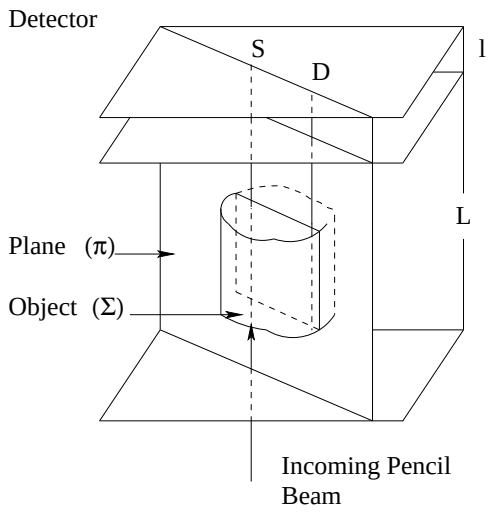


Figure 5: Arbitrary object analysis

Introducing the one-dimensional Fourier representation of the electronic densities:

$$n_e(\mathbf{D},\zeta_N) = \int_{-\infty}^{\infty} dw\, e^{2i\pi w\zeta_N}\, \tilde{n}_e(\mathbf{D},w),$$

$$n_e(\mathbf{S},\zeta) = \int_{-\infty}^{\infty} dw'\, e^{2i\pi w'\zeta}\, \tilde{n}_e(\mathbf{S},w'),$$

in equation (5), the TPSF can be recast under the form:

$$g(\mathbf{S},\mathbf{D}|\tau) = \frac{K(\tau)}{\rho^2} \int\int dw\, dw'\, \tilde{n}_e(\mathbf{D},w)\, \tilde{n}_e(\mathbf{S},w')$$

$$\mathcal{J}_l(w+w')\, \exp[2i\pi(lw + lw' + w'\rho\tau)], \qquad (10)$$

where:

$$\mathcal{J}_l(w) = \int_l^{\infty} \frac{d\zeta_N}{\zeta_N^2}\, e^{2i\pi(w+w')\zeta_N} = 2i\pi w$$

$$\{e^{2i\pi lw}\, [Ci(2\pi l|w|) - i\,\varepsilon(w)\, Si(2\pi l|w|)] - \frac{i}{2\pi wl}\},$$

$\varepsilon(w)$ being the sign function of w [5]. Now since $\tau \in \mathcal{R}$, we could use it as Fourier variable to invert equation 10, and obtain:

$$\int_{-\infty}^{\infty} d\tau\, e^{-2i\pi\nu\tau}\, \frac{g(\mathbf{S},\mathbf{D}|\tau)}{K(\tau)} = \frac{1}{\rho^2}\, \tilde{n}_e(\mathbf{S},\frac{\nu}{\rho}) e^{2i\pi l\frac{\nu}{\rho}}$$

$$\int_{-\infty}^{\infty} dw\, \tilde{n}_e(\mathbf{D},w)\, \mathcal{J}_l(w+\frac{\nu}{\rho}) e^{2i\pi lw}.$$

Here the left hand side appears as *a superposition of TPSF images taken at various scattering angles*, however the right hand side remains a bilinear construct in the electronic densities at two different sites. To get a better insight, we transform the right hand side by back Fourier transform into:

$$\int_{-\infty}^{\infty} dw\, \tilde{n}_e(\mathbf{D},w)\, \mathcal{J}_l(w+\frac{\nu}{\rho}) e^{2i\pi lw} =$$

$$\int_l^{\infty} \frac{d\zeta}{\zeta^2} n_e(\mathbf{D},\zeta) e^{2i\pi\frac{\nu}{\rho}(\zeta-l)},$$

and end up, after another Fourier inversion with a final form:

$$\int_{-\infty}^{\infty} d\tau\, e^{-2i\pi\nu\tau}\, \frac{g(\mathbf{S},\mathbf{D}|\tau)}{K(\tau)} =$$

$$\frac{1}{\rho^2}\, \tilde{n}_e(\mathbf{S},\frac{\nu}{\rho}) \int_l^{\infty} \frac{d\zeta}{\zeta^2} n_e(\mathbf{D},\zeta)\, e^{2i\pi\frac{\nu}{\rho}\zeta}. \qquad (11)$$

Remark. One could exchange the roles of $\mathbf{D}$ and $\mathbf{S}$: the incoming pencil beam falls perpendicularly on the detector at site $\mathbf{D}$ and $\mathbf{S}$ becomes the detection site. Then one has, with the same separation ρ:

$$\int_{-\infty}^{\infty} d\tau\, e^{-2i\pi\nu\tau}\, \frac{g(\mathbf{D},\mathbf{S}|\tau)}{K(\tau)} =$$

$$\frac{1}{\rho^2}\, \tilde{n}_e(\mathbf{D},\frac{\nu}{\rho}) \int_l^{\infty} \frac{d\zeta}{\zeta^2} n_e(\mathbf{S},\zeta)\, e^{2i\pi\frac{\nu}{\rho}\zeta}.$$

Inversion is in fact based on one of these equations. In general, we have thus a pair of integral equations which are quadratic in the unknown functions $n_e(\mathbf{r})$. To our knowledge a general solution does not exist. However under certain practical hypotheses, an inversion procedure may be constructed.

Let us consider equation (11), which describes the imaging of the medium by the TPSF centered at $\mathbf{S}$, the measurement of flux density being made at site $\mathbf{D}$. If $\mathbf{S}$ is chosen in such a way that $\tilde{n}_e(\mathbf{S},\frac{\nu}{\rho})$ is *known*, one may locate in the sample a vertical line to the detector along which the material is homogeneous, i.e. free of defects. Then we have:

$$n_e(\mathbf{S},\zeta) = \begin{cases} n_e^0 & \text{if } l < \zeta < (L+l) \\ 0 & \text{otherwise.} \end{cases}$$

Then $\tilde{n}_e(\mathbf{S},\frac{\nu}{\rho})$ may be exactly evaluated:

$$\tilde{n}_e(\mathbf{S},\frac{\nu}{\rho}) = \frac{n_e^0\rho}{\pi\nu}\, e^{-i\pi\frac{\nu}{\rho}(2l+L)}\, \sin\pi\frac{\nu}{\rho}L.$$

Hence

$$\int_l^{\infty} \frac{d\zeta}{\zeta^2} n_e(\mathbf{D},\zeta)\, e^{2i\pi\frac{\nu}{\rho}\zeta} = \frac{\pi\nu\rho}{n_e^0}$$

$$\frac{e^{i\pi\frac{\nu}{\rho}(2l+L)}}{\sin\pi\frac{\nu}{\rho}L} \int_{-\infty}^{\infty} d\tau\, e^{-2i\pi\nu\tau}\, \frac{g(\mathbf{S},\mathbf{D}|\tau)}{K(\tau)}. \qquad (12)$$

To extract now the electronic density at **D**, we perform inverse Fourier transformation by multiplying equation (12) on both sides by

$$\int_{-\infty}^{\infty} d(\frac{\nu}{\rho})\, e^{2i\pi\frac{\nu}{\rho}\zeta'},$$

and integrate over $(\frac{\nu}{\rho})$ to get the expression of the electronic density at site $(\mathbf{D}, \zeta')$:

$$n_e(\mathbf{D},\zeta') = \zeta'^2 \int_{-\infty}^{\infty} d(\frac{\nu}{\rho})\, e^{2i\pi\frac{\nu}{\rho}(\frac{2l+L}{2}-\zeta')}\, \frac{\pi\nu\rho}{n_e^0}$$

$$\frac{1}{\sin\pi\frac{\nu}{\rho}L} \int_{-\infty}^{\infty} d\tau\, e^{-2i\pi\nu\tau}\, \frac{g(\mathbf{D},\mathbf{S}|\tau)}{K(\tau)}. \qquad (13)$$

An alternative way of inverting consists of choosing a fixed detection site **D** with a perpendicular line along which the electronic density is constant. Then one can calculate immediately:

$$\int_{l}^{\infty} \frac{d\zeta}{\zeta^2} n_e(\mathbf{D},\zeta)\, e^{2i\pi\frac{\nu}{\rho}\zeta} = n_e^0 \mathcal{J}_l(\frac{\nu}{\rho}).$$

Hence one deduces the Fourier transform $\tilde{n}_e(\mathbf{S}, \frac{\nu}{\rho})$ of $n_e(\mathbf{S}, \tau)$ and consequently:

$$n_e(\mathbf{S},\zeta') = \int_{-\infty}^{\infty} d(\frac{\nu}{\rho})\, e^{2i\pi\frac{\nu}{\rho}\zeta'}$$

$$\frac{\rho^2}{n_e^0 \mathcal{J}_l(\frac{\nu}{\rho})} \int_{-\infty}^{\infty} d\tau\, e^{-2i\pi\nu\tau}\, \frac{g(\mathbf{D},\mathbf{S}|\tau)}{K(\tau)}. \qquad (14)$$

This time we keep the detection site fixed, and move around the incoming pencil beam to collect the data before computing the reconstruction of the defect. This is an equivalent procedure and the choice between the two may depend on the practicability of the measures at hand.

CONCLUSIONS AND OUTLOOK

In this work, we propose an *analytical* inverse method to determine the electronic density from double Compton scattering in transmission imaging. This result is used for the detection of defects in homogeneous medium. This detection procedure does not require the motion of neither the incident radiation source nor the material under investigation. This represents a real advantage in some non-destructive controls in which the number of views in conventional tomography is very limited.

References

[1] H. H. Barrett and W. Swindell, *Radiological Imaging* I and II, Academic Press, New York, (1981).

[2] Mai K. Nguyen, C. Fay, L. Eglin and T. T. Truong, Apparent image formation by Compton scattered photons in gamma-ray imaging *IEEE Signal Processing Letters*, **8** (9), pp. 248-251, 2001.

[3] Mai K. Nguyen and T. T. Truong, On an integral transform and its inverse in nuclear imaging, *Inverse Problems*, **18** (1), pp. 265-277, 2002.

[4] R. Cesaro, A.L. Hanson, G.E. Gigante, L.J. Pedraza and S.Q.G. Mathtaboally, *Interactions of keV Photons with Matter and New Applications*, Physics Reports **213** (3), pp. 117-178, 1992.

[5] J. Lavoine, *Transformation de Fourier des Pseudo-fonctions avec Tables de Nouvelles Transformées* , CNRS, Paris, (1963).

[6] E. M. A. Hussein, Compton Scatter Imaging Systems, in *Bioinstrumentation: research, developments and applications*, D. L. Wise Editor, chap.35, pp. 1053-1086, 1990.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

DYNAMIC ELECTRICAL IMPEDANCE IMAGING OF BINARY-MIXTURE FIELDS WITH EXTERNAL AND INTERNAL ELECTRODES

Kyung Youn Kim, Sook In Kang
Department of Electrical & Electronic Engineering
Min Chan Kim
Department of Chemical Engineering
Junghoon Lee
Department of Computer Science and Statistics
Cheju National University, Cheju 690-756, Korea

Sin Kim
Department of Nuclear and Energy Engineering
E-mail:sinkim@cheju.ac.kr
Chang Ik Kang
Department of Marine Instrumentation Engineering
Cheju National University, Cheju 690-756, Korea

ABSTRACT

In the conventional electrical impedance tomography (EIT), the internal impedivity distribution, that is mixture distribution, is reconstructed based on the physical relationship between the known sets of injected currents through the electrodes and induced voltages on the surface of the domain of interest under the assumption that the domain is stationary during the measurements. This study considers a dynamic electrical impedance imaging to binary-mixture systems with known internal structures to which additional electrodes can be attached. We attempt the utilization of the additional electrodes to enhance the recontruction. Also, we assume the domain is undergoing rapid transient, so the resistivity distribution in the domain of interest changes rapidly within the time taken to acquire a full set of independent measurement data. The dyamic EIT problem is treated as the nonlinear state estimation problem and the unknown state (resistivity) is estimated with the aid of extended Kalman filter in a minimum mean square error sense. Computer simulations for the two-dimensional object with abrupt changing resistivity distribution are provided to illustrate the reconstruction performance of the proposed algorithm.

NOMENCLATURE

C	time-updated error covariance matrix
e_l	l^{th} electrode
F	state transition matrix at time k
I_l	injected current through e_l
J	Jacobian matrix
L	number of electrodes
N	number of elements
R^*	regularization matrix
U	measured boundary potential
u	potential distribution
v	measurement error
V	calculated bounday potential
w	process error
y	pseudo-measurement

Greek symbols

α	regularization parameter
Γ	covariance
ρ	resistivity

INTRODUCTION

Binary-mixture systems are frequently encountered in many engineering applications including heat exchangers, oil or natural gas pumping systems, fluidized beds and so on. For the optimal operation as well as for the design and performance analysis, the knowledge on mixture distribution is of importance. Many imaging techniques using radiological, optical, sonic or electromagnetic characteristics have been proposed and applied to visualize mixture systems. Among these, electrical impedance tomography (EIT), a relatively new imaging modality, is considered in this work.

Over the past few decades, EIT techniques have received much attention from both theoretical and experimental points of view since they can be used as an alternative imaging modality for monitoring tool in many engineering fields. This is mainly due to the relatively cheap electronic hardware requirements, noninvasive measurement sensing, and relatively good temporal resolution [1-3]. In EIT, the internal impedivity (inverse of admittivity) distribution corresponding to the mixture distribution is

reconstructed with the known sets of injected currents through the electrodes and induced voltages on the surface of the domain of interest (the object) based on the relationship between electrical current and potential distribution. EIT includes both electrical capacitance tomography (ECT) and electrical resistance tomography (ERT). However, more frequently in EIT it is assumed that the resistive part of the impedivity dominates and only the resistivity (inverse of conductivity) distribution inside the object is estimated.

A basic idea of the conventional EIT is that the distribution of internal electrical properties is stationary during acquisition of a full set of EIT data required for the image reconstruction. However, the characteristic time of the systems under severe transient could be shorter than the time interval that would be required for the EIT measurement. We introduce a dynamic EIT imaging technique for the case where the resistivity distribution inside the object changes rapidly within the time taken to acquire a full set of independent measurement data. In this work, the inverse problem is treated as the nonlinear state estimation problem and the unknown state (resistivity) is estimated with the aid of extended Kalman filter in a minimum mean square error sense.

In particular, this study considers the systems with internal structures where internal electrodes could be attached. Such systems are often possible in heat exchangers with tube banks or mixers with rotors. Additional electrodes attached to the internal structure would be expected to enhance the reconstruction performance. Several simulated mixture distributions are reconstructed with and without internal electrodes and comparisons are made to illustrate the relative reconstruction performance. Several numerical results will be given to corroborate that the internal electrodes improve the spatial and temporal resolution of reconstructed images.

MATHEMATICAL MODEL FOR DYNAMIC EIT PROBLEMS

Problem Definition

In the present EIT problem, the quantity to be estimated is actually the resistivity distibution, which should satisfy the physical relationship between the internal resistivity distribution ρ and the potential distribution u expressed as a partial differential equation (Laplace equation) over the domain of interest Ω with appropriate boundary conditions:

$$\nabla \cdot \left(\rho^{-1} \nabla u\right) = 0\,, \quad x \in \Omega \tag{1}$$

subject to

$$\begin{aligned} &\int_{e_l} \frac{1}{\rho}\frac{\partial u}{\partial \nu} dS = I_l\,, \quad x \in e_l\,, \quad l = 1,2,\cdots,L \\ &\frac{1}{\rho}\frac{\partial u}{\partial \nu} = 0, \quad x \in \partial\Omega \setminus \cup_{l=1}^{L} e_l \end{aligned}\,, \tag{2}$$

where I_l denotes the injected current through the l^{th} electrode e_l and L is the number of the electrodes mounted on the surface. Hence, the present problem is to find the resistivity distribution both satisfying the above constraints and minimizing the difference between the calculated bounday potential V and the measured boundary potential U:

$$\min_{\rho} \Phi = \frac{1}{2}\left[U - V(\rho)\right]^T \left[U - V(\rho)\right] \tag{3}$$

Owing to the complexity of this relationship, it is impossible to obtain a closed-form solution for arbitrary resistivity distribution. Hence, various reconstruction algorithms have been developed in the literature to estimate the internal resistivity distribution of the object. This work adopts the finite element method (FEM) to discretize the domain and to find the potential field induced by the applied currents in a predetermined manner. Also, as a solver to the present dynamic EIT problem the extended Kalman filter is introduced.

Most of the reconstruction algorithms presented so far are mainly focused on the case where the internal resistivity of the object is time-invariant within the time taken to acquire a full set of independent measurement data. As is well known, the conventional EIT imaging techniques such as backprojection or modified Newton-Raphson (mNR) algorithm use a full set of voltage measurements for each image [4,5]. However, in some real applications such as mixture fields in multi-phase heat transfer media and chemical process lines undergoing severe transient, this stationary approximation seems to be inadequate to obtain satisfactory temporal

resolution for the reconstructed images due to the rapid changes in resistivity distribution.

More recently, dynamic imaging techniques have been developed to enhance the temporal resolution of the reconstructed images in the situations where the resistivity distribution inside the object changes rapidly in time. In most of these techniques, the inverse reconstruction problem is treated as state estimation problem and the time-varying state is estimated with the aid of linearized Kalman filter (LKF)[6-10] or extended Kalman filter (EKF)[11,12].

Quite often in real situations there are partially known fixed internal structures inside the object. These internal structures can be, for example, an impeller drive shaft or a mixing paddle in process vessels and an assembly of fuel rods in nuclear reactor. The internal structures inside the object may results in difficulties in the image reconstruction by EIT especially when the high resistive region is near the conductive internal structure or vice versa [13,14]. The so-called masking effect in the reconstructed image may be significant for the high-contrast case. There are two ways to get around these difficulties; the one is to use the internal structure as additional electrodes [13,15,16] and the other is to take into account it as a priori information in the inverse procedure [14]. However, all of the above approaches are for the case where the resistivity distribution inside the object is time-invariant for one classical frame.

The purpose of the present work is to develop a dynamic EIT reconstruction algorithm for the visualization of transient binary-mixture systems whose internal structures are known *a priori*. For the utilization of the prior information as well as the enhancement of the image reconstruction, we consider the case where additional electrodes can be attached to the known internal structure.

The inverse problem is treated as the state estimation problem and the resistivity distribution is estimated with EKF. In other to deal with the well-known ill-posedness of the EIT inverse problem, a smoothness assumption is made and the generalized Tikhonov regularization technique is also introduced in the cost functional.

We carried out extensive computer simulations with synthetic data to illustrate the reconstruction performance, and to investigate the effects of additional internal electrodes on the spatial and temporal resolution of the reconstructed images.

Extended Kalman Filter Approach

In case where the resistivity distribution inside the object changes rapidly within the time taken to acquire a full set of independent measurement data, the conventional static imaging techniques which need a full set of voltage measurements for each image often fail to obtain satisfactory temporal information on the resistivity distribution. We consider the underlying inverse problem as a state estimation problem to estimate rapidly time-varying distribution of the resistivity. In the state estimation problem, we need so-called the dynamic model which consists of the state equation, i.e., for the temporal evolution of the resistivity and the observation equation, i.e., for the relationship between the resistivity and boundary voltage.

In general, the temporal evolution of the resistivity distribution ρ_k in the object Ω is related by the nonlinear mapping. Here, the state equation is assumed to be of the linear form, of which the modeling uncertainty is compensated by the process noise

$$\rho_{k+1} = F_k \rho_k + w_k \tag{4}$$

where $F_k \in R^{N\times N}$ is the state transition matrix at time k and N is the number of finite elements in the FEM (finite element method). In particular, we take $F_k = I_N$ where $I_N \in R^{N\times N}$ is an identity matrix, to obtain the so-called random-walk model. It is assumed that the process error, w_k is white Gaussian noise with the following covariance which determines the rate of changes in resistivity distribution

$$\Gamma_k^w = \mathrm{E}[w_k w_k^T] \tag{5}$$

Let $U_k \in R^L$, defined as

$$U_k \equiv [U_k^1 U_k^2 ... U_k^L]^T \tag{6}$$

be the surface measurement voltages induced by the k^{th} current pattern. Then the observation equation can be described as the following nonlinear mapping with measurement error

$$U_k = V_k(\rho_k) + v_k \tag{7}$$

where the measurement error v_k is also assumed to be white Gaussian noise with covariance

$$\Gamma_k^v = \mathrm{E}[v_k v_k^T] \tag{8}$$

Linearizing Eq. (7) about the current predicted state $\rho_{k|k-1}$ we obtain

$$\begin{aligned} U_k &= V_k(\rho_{k|k-1}) + J_k(\rho_{k|k-1})\cdot(\rho_k - \rho_{k|k-1}) \\ &\quad + H.O.T + v_k \end{aligned} \tag{9}$$

where H.O.T represents the higher-order terms which will be considered as additional noise, and $J_k(\rho_{k|k-1}) \in R^{L\times N}$ is the Jacobian matrix defined by

$$J_k(\rho_{k|k-1}) \equiv \left.\frac{\partial V_k}{\partial \rho}\right|_{\rho=\rho_{k|k-1}} \tag{10}$$

Define the pseudo-measurement as

$$y_k \equiv U_k - V_k(\rho_{k|k-1}) + J_k(\rho_{k|k-1})\cdot\rho_{k|k-1} \tag{11}$$

Then we obtain the linearized observation equation by considering the *H.O.T.* in (9) as additional noise

$$y_k = J_k(\rho_{k|k-1})\cdot\rho_k + \bar{v}_k \tag{12}$$

where $\bar{v}_k$ is composed of measurement error and linearization error and also assumed to be white Gaussian noise with covariance as

$$\bar{\Gamma}_k = \mathrm{E}[\bar{v}_k \bar{v}_k^T] \tag{13}$$

In Kalman filtering we estimate the state ρ_k based on all the measurements taken up to the time k. With the Gaussian assumptions the required estimate is obtained by minimizing the cost functional which is formulated based on the above state and observation equations (4) and (12), respectively. The cost functional for the conventional Kalman filter is of the form

$$\begin{aligned} G^a(\rho_k) = \tfrac{1}{2}\{&\|\rho_k - \rho_{k|k-1}\|_{C_{k|k-1}^{-1}} \\ &+ \|y_k - J_k(\rho_{k|k-1})\cdot\rho_k\|_{(\bar{\Gamma}_k)^{-1}}\} \end{aligned} \tag{14}$$

where $C_{k|k-1} \in R^{N\times N}$ is the time-updated error covariance matrix, which is defined by

$$C_{k|k-1} \equiv E[(\rho_k - \rho_{k|k-1})(\rho_k - \rho_{k|k-1})^T] \tag{15}$$

In order to mitigate the inherent ill-conditioned nature of the EIT inverse problem, additional constraint is included in the cost functional

$$\begin{aligned} G^b(\rho_k) = \tfrac{1}{2}\{&\|\rho_k - \rho_{k|k-1}\|_{C_{k|k-1}^{-1}} \\ &+ \|y_k - J_k(\rho_{k|k-1})\cdot\rho_k\|_{(\bar{\Gamma}_k)^{-1}} \\ &+ \alpha\|R^*\rho_k\|\} \end{aligned} \tag{16}$$

where α is regularization parameter and R^* is modified regularization matrix. There are two approaches for selecting regularization parameter α that have gained wide acceptance. The one is composite residual and smoothing operator (CRESO) and the other is the L-curve approach [17]. In this paper, α is chosen *a posteriori* by simple trial-and-error method. One popular conventional method for the choice for the regularization matrix R^* is a difference-type matrix on the basis of the generalized Tikhonov regularization technique [6] by the smoothness assumptions in resistivity distributions. In this method, the resistivity distribution is parameterized such that

$$\rho = \sum_{n=1}^{N} \rho_n \chi_n \tag{17}$$

where χ_n is the characteristic function of the n^{th} finite element. Recalling that the problem domain is discretized into triangular elements, the i^{th} row of R^* is

$$R_i^* = (0,0,...,0,-1,0,...,0,-1,0,...,0,3,0,...,0,-1,0,...0) \tag{18}$$

where 3 is located at the i^{th} column and –1 is placed in the columns corresponding to elements having common edge with the i^{th} element. Sometimes in real situations, there are partially known internal structures in which additional electrodes can be attached. This work considers

the utilization of the prior information on the interior structure. In this case, the regularization matrix can be obtained from Eq. (18) by removing the -1 that corresponds to element having common edge with the known internal structure. Also, the number 3 in Eq. (18) should be replaced by 2 since the smoothness assumption is violated between the known element and background.

Define the augmented pseudo-measurement, $\bar{y}_k \in R^{(L+N)\times 1}$ and pseudo-measurement matrix, $H_k \in R^{(L+N)\times N}$ as

$$\bar{y}_k \equiv \begin{pmatrix} y_k \\ 0 \end{pmatrix} \tag{19}$$

$$H_k \equiv \begin{pmatrix} J_k \\ \sqrt{\alpha} R^* \end{pmatrix} \tag{20}$$

Then the cost functional, Eq. (16) can be rearranged as

$$G(\rho_k) = \frac{1}{2}\{\| \rho_k - \rho_{k|k-1} \|_{C_{k|k-1}^{-1}} + \| \bar{y}_k - H_k \rho_k \|_{(\Gamma_k)^{-1}}\} \tag{21}$$

where the augmented covariance matrix, $\Gamma_k \in R^{(L+N)\times(L+N)}$ is defined by

$$\Gamma_k \equiv \text{Blockdiag}[\bar{\Gamma}_k, I_N] \tag{22}$$

Minimizing the cost functional, Eq. (21), and solving for the updates of the associated covariance matrices we obtain the recursive extended Kalman filter algorithm which consists of the following two steps [18]:

(i) Measurement Updating Step (Filtering)

$$G_k = C_{k|k-1} H_k^T [H_k C_{k|k-1} H_k^T + \Gamma_k]^{-1} \tag{23}$$

$$C_{k|k} = (I - G_k H_k) C_{k|k-1} \tag{24}$$

$$\rho_{k|k} = \rho_{k|k-1} + G_k [\bar{y}_k - H_k \cdot \rho_{k|k-1}] \tag{25}$$

(ii) Time Updating Step (Prediction)

$$C_{k+1|k} = F_k C_{k|k} F_k^T + \Gamma_k^w \tag{26}$$

$$\rho_{k+1|k} = F_k \rho_{k|k} \tag{27}$$

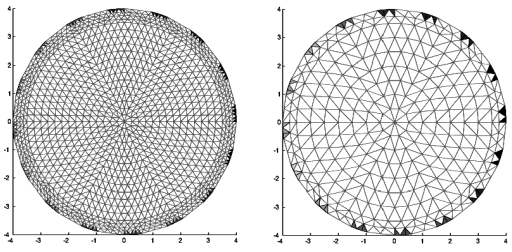
Hence, we can find the estimated state $\rho_{k|k}$ for the true state ρ_k in a recursive minimum mean square error sense for $k = 1,2,\ldots,rK$, where K is the number of the independent current patterns and r is the number of the classical frames.

COMPUTER SIMULATIONS

We carried out extensive computer simulations with synthetic data to evaluate the reconstruction performance of the proposed algorithm.

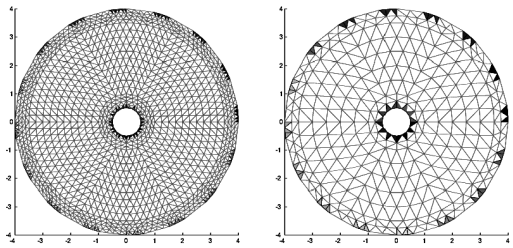
The FEM meshes without internal electrodes used for the forward and inverse solvers are shown in Fig. 1 (a) and (b), respectively. In the forward computations we used the FEM with a mesh of 2400 elements and 1281 nodes. In the inverse computations, we used the FEM with a mesh of 600 elements and 341 nodes to reduce the computational burden. For the current injection and corresponding voltage measurement, traditional adjacent method [1] was employed through 16 boundary electrodes (L) so that the total measurement voltage data were 256 (16×16).

The FEM meshes with a single internal electrode used for the forward and inverse solvers are shown in Fig. 2 (a) and (b), respectively. In the forward computations we used the FEM with a mesh of 2480 elements and 1336 nodes. In the inverse computations, we used the FEM with a mesh of 620 elements and 358 nodes. We injected electrical current between 16 boundary electrodes and the internal electrode and measured the corresponding voltage on the 17 electrodes (L) so that the total measurement voltage data were 272 (17×16).



(a) (b)

Fig. 1. FEM meshes without internal electrodes used for
(a) forward solver and (b) inverse solver.



(a) (b)

Fig. 2. FEM meshes with an internal electrode used for
(a) forward solver and (b) inverse solver.

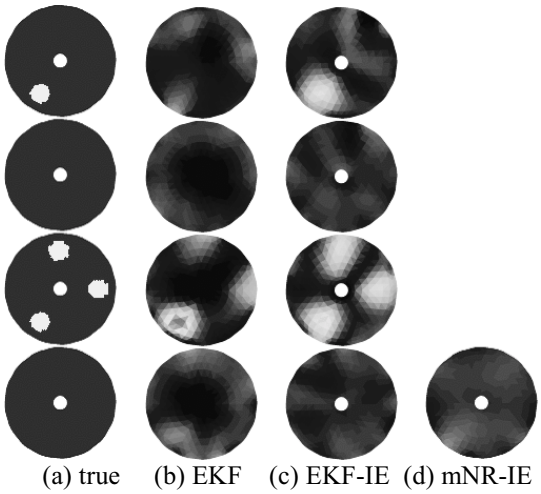
To compare the reconstruction performance, we used the static algorithm based on the modified Newton-Raphson algorithm with internal electrodes (mNR-IE)[16], dynamic algorithm based on the extended Kalman Filter (EKF)[12], and the Extended Kalman Filter with internal electrodes (EKF-IE) described in this paper. The parameters used for the three methods are as follows. The regularization parameter(α) is set to 0.5 in both simulations. The initial resistivity value is set to the same as the background value in all cases. For simplicity, it is assumed that the covariance matrices for all the EKFs are diagonal and time-invariant. The covariance matrix for process noise(Γ_k^w) is $10\mathrm{I_N}$, the covariance matrix for measurement noise(Γ_k^v) is 0.0001 I_L and the initial value for the state error covariance matrix ($C_{1|0}$) is I_N in both simulations.

The First Simulation

For the verification of the proposed model, we simulate binary-mixture distributions in a circular domain (8 cm in diameter). It is assumed that there is a known conductive circular structure (about 1 cm in diameter, 0.0001Ωcm in resistivity), which is located at the center and is used as an internal electrode. This configuration could be found in many engineering applications like mixer with rotor, shell-and-tube type heat exchanger and so on.

Consider the following scenario of binary-mixture field undergoing transient. A single anomaly emerges and disappears. And then, three anomalies emerge and disappear. This transient is assumed to take place within the time interval required to measure a full set of EIT data (1 frame) in the context of the static EIT image reconstruction. The stationary approximation could not be effective over the entire time interval required for the injection of all the predetermined 16 current patterns and the measurement of the induced boundary potentials. Hence, we assume each transient stage to take the time interval equivalent to 4 current injections evenly. The background resistivity and the anomaly resistivity are set to 300Ωcm and 600Ωcm, respectively. The simulated evolution of the mixture distribution is depicted in the first column of Fig. 3. In order to make the simulation more realistic, we assume that the measured boundary voltages are contaminated by the measurement error of 1%.

The transient mixture field described above is visualized with different EIT algorithms; EKF, EKF-IE and mNR-IE. Results from each algorithm are presented in Fig. 3(b), (c) and (d). As can be expected, Fig. 3(d) shows that mNR-IE totally fails to reconstruct the internal mixture field since the internal resistivity distribution is not stationary. The reproduced images by EKF without internal electrode seem to be rather vague although they are much better than the images by mNR-IE. However, the reconstructed images by EKF-IE in Fig. 3(c) show that the suggested algorithm gives significant improvement in the spatial resolution.



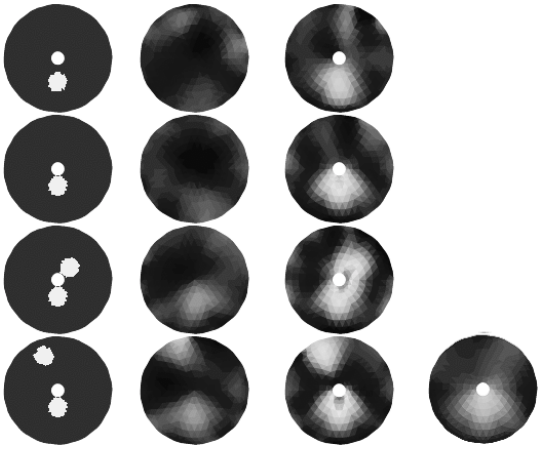
(a) true (b) EKF (c) EKF-IE (d) mNR-IE

Fig. 3. The first simulation.

The Second Simulation

In the second simulation, we assume the same condition for the internal structure and the resistivity values as in the first simulation. In the measurement of boundary voltages, as in the first simulation, random error of 1% is included. However, a different evolution of mixture distribution is considered as shown in Fig. 4(a).

As much similar to the results from the first simulation, mNR-IE cannot extract any useful information from the EIT image reconstruction. Even with EKF, the reconstruction seems to be more deteriorated compared to the first simulation. It should be noted that in this simulation we intentionally place two anomalies and the conductive internal structure close together. Then, resistive anomalies impede electrical current flowing to the conductive internal structure and the spatial resolution is expected to substantially worsen. However, the utilization of the internal electrode improves the spatial resolution in the binary mixture field as can be seen in Fig. 4(c). Also, the temporal information for the abruptly changing mixture fields is severely lost in the reconstructed images by the conventional EKF.



(a) true (b) EKF (c) EKF-IE (d) mNR-IE

Fig. 4. The second simulation.

CONCLUSIONS

In usual static EIT techniques, the internal resistivity distribution is assumed to be stationary during the measurement of the required full set of independent EIT data. Quite often in real situations, however, the object to be visualized may undergo severe transient. This study considers the application of the EIT technique to the visualization of binary-mixture systems under rapid transient. Also, we propose an EIT image reconstruction algorithm that utilizes the internal electrode attached to the known internal structure in addition to the external electrodes. Actually, in many engineering applications, there are known internal structures where the use of additional electrodes can be possible.

For the dynamic electrical impedance imaging, EIT inverse problem is formulated as a state estimation problem and the state (resistivity distribution) is estimated with the aid of the extended Kalman filter along with the inclusion of the internal electrode mounted on the known internal structure. Through the comparison with the dynamic algorithm using extended Kalman filter as well as the static algorithm like modified Newton-Raphson method, it is shown that the proposed method improves reconstruction performance significantly in the sense of the spatial and temporal resolution.

ACKNOWLEDGMENT

This work was supported by the Nuclear Academic Research Program of the Ministry of Science and Technology (MOST), Korea.

REFERENCES

1. J. G. Webster, *Electrical Impedance Tomography*, Adam Hilger, Bristol, 1990.

2. R. A. Williams and M. S. Beck, *Process Tomography: Principles, Techniques and Applications*, Butterworth-Heinemann, Oxford, 1995.

3. M. S. Beck and R. A. Williams, Process tomography: a European innovation and its applications, *Meas. Sci. Technol.*, **7**, 215 (1996).

4. D. C. Barber and B. H. Brown, Applied potential tomography, *J. Phys. E*, **17**, 723 (1984).

5. T. J. Yorkey, J. G. Webster, and W. J. Tompkins, Comparing reconstruction algorithms for electrical resistance tomography, *IEEE Trans. Biomed. Eng.*, **34**, 843 (1987).

6. M. Vauhkonen, P. A. Karjalainen, and J. P. Kaipio, A Kalman filter approach to track fast impedance changes in electrical impedance tomography, *IEEE Trans. Biomed. Eng.*, **45**, 486 (1998).

7. J. P. Kaipio, P. A. Karjalainen, E. Somersalo, and M. Vauhkonen, State estimation

in time-varying electrical impedance tomography, *Ann. NY Acad. Sci.*, **873**, 430 (1999).

8. P. J. Vauhkonen, M. Vauhkonen, T. Mäkinen, P. A. Karjalainen, and J. P. Kaipio, Dynamic electrical impedance tomography - phantom studies, *Inverse Prob. Eng.*, **8**, 495 (2000).

9. M. Vauhkonen, *Electrical Impedance Tomography and Prior Information*, Doctoral Dissertation, Dept. of Applied Physics, University Kuopio, 1997.

10. A. Seppanen, M. Vauhkonen, P. J. Vauhkonen, E.Somersalo, and J. P. Kaipio, State estimation with fluid dynamical evolution models in process tomography - An application with impedance tomography, *Inverse Prob.*, **17**, 467 (2001).

11. K. Y. Kim, B. S. Kim, M. C. Kim and Y. J. Lee, "On-line image reconstruction in dynamic electrical impedance tomography based on the extended Kalman filter," Proc. of Compumag-Evian (France), Vol. IV, pp. 70-71, 2001.

12. K. Y. Kim, B. S. Kim, M. C. Kim, Y. J. Lee, and M.Vauhkonen, Image reconstruction in time-varying electrical impedance tomography based on the extended Kalman filter, *Meas. Sci. Technol.*, **12**, 1032 (2001).

13. R. A. Williams, X. Jia, and S. L. McKee, Development of slurry mixing models using resistance tomography, *Powder Technol.*, **87**, 21 (1996).

14. L. M. Heikkinen, M. Vauhkonen, T. Savolainen, K. Leinonen, and J. P. Kaipio, Electrical process tomography with known internal structures and resistivities, *Inverse Prob. Eng*., **9**, 431, 2001.

15. G. M. Lyon and J. P. Oakley, "A simulation study of sensitivity in stirred vessel electrical impedance tomography," In M. S. Beck, E. Campogrande, M. Morris, R. A. Williams, and R. C. Waterfall, editors, Tomography Techniques for Process Design and Operation, pp. 137-146, Southampton UK, 1993, Computational Mechanics Publications.

16. L. M. Heikkinen, M. Vauhkonen, T. Savolainen, and J. P. Kaipio, Modelling of internal structures and electrodes in electrical process tomography, *Meas. Sci. Technol*., **12**, 1012 (2001).

17. P. R. Johnston and R. M. Gulrajani, Selecting the corner in the L-curve approach to Tikhonov regularization, *IEEE Trans. Biomed. Eng*., **47**, 1293 (2000).

18. M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*, Prentice Hall, Englewood Cliffs, New Jersey, 1993.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

RETRIEVAL OF TEMPERATURE AND MOISTURE PROFILES OVER BRAZIL USING SATELLITE DATA

João C. Carvalho, Nelson J. Ferreira, Fernando M. Ramos
Instituto Nacional de Pesquisas Espaciais
São José dos Campos, SP, Brazil.
jcarlos@cptec.inpe.br

Lydie Lavanant
Centre de Météorologie Spatiale, Météo-France
Lannion, Cotês d'Armor, France.

ABSTRACT

In this paper we investigate the problem of reconstructing temperature and humidity profiles in a tropical atmosphere using the Inversion Coupled with Image (ICI) model. The analyses were done using the Advanced TIROS Operational Vertical Sounder (ATOVS) data during February of 2000. The results are compared with numerical weather prediction data. Emphasis is given in analysing the role of the channels combination used in the retrieval process. Different surface types (sea and land) and atmospheric conditions (clear and cloud sky) were also considered.

INTRODUCTION

The vertical structure of temperature and water vapor plays an important role in the dynamics and radiation exchange of the atmosphere. For years the radiosonde observations have been the primary source of data for monitoring the tropospheric conditions. Despite the importance of this observing system, there is still a lack of observation in several regions of the Earth. In this sense, retrievals of temperature and humidity profiles from satellite are important for a number of applications such as weather analysis and data assimilation in numerical weather prediction models.

Interpretation of satellite radiances in terms of meteorological parameters requires the inversion of the radiative transference equation (RTE) where measurements of radiation performed in different frequencies are related to the energy from different atmospheric regions. The obtained solution is highly indeterminate for a set of observed radiances. The degree of indetermination is associated with the spectral resolution and the number of spectral channels. Moreover, usually this solution is very unstable with regard to noises in the measuring process [1,2]. Currently there is an effort to improve the spectral and spatial resolution of the sounders on board of meteorological satellites. Also, several methodologies and models have been developed to improve the satellite data processing. Due to the difficulty of obtaining correct RTE solutions, several approaches and methods were developed to extract information from satellite data [3-9].

Since May 1998 a new generation of radiometers for satellite retrieval of vertical temperature and humidity profiles became available. The ATOVS system consist basically of three instruments: the HIRS, composed of 20 spectral channels in the infrared spectral range, the AMSU-A with 15 microwave channels and AMSU-B, composed of 5 channels also localized in the microwave spectral range. Table 1 shows the main characteristics of this radiometer and its meteorological applications. A complete description of the ATOVS system can be found in [10].

The main purpose of this work is to investigate the problem of reconstructing temperature and humidity profiles in a tropical atmosphere using data remotely collected by the ATOVS sounder. To accomplish this task, we chose the Inversion Coupled with the Image (ICI) model [11] developed at the Centre de Météorologie Spatiale (CMS), where it has been operational since 1996.

THE ICI INVERSION SYSTEM

The ICI was chosen because its easy implementation and simple conception to make experiments. The main ICI modules are: the guess profiles library search, the inversion module and the tuning module, responsible for model periodic calibration [11]. Figure 1 shows a schematic diagram of the ICI working system.

Table 1. Characteristics of ATOVS system: channels, spectral range (wavelength or frequency) and main meteorological applications.

HIRS		
Chan.	*Wavelength (μm)*	*Meteor. Application*
1-7	14.9-13.4	Temperature Sounding
8	11.1	Surface Temperature
9	9.6	Total Ozone
10	12.4	Moisture Sounding
11,12	6.5; 7.3	Moisture Sounding
13-17	4.13-4.57	Temperature Sounding
18-19	3.76-4.13	Surface Temperature
20	0.7	Cloud Detection
AMSU-A		
Chan.	*Frequency (GHz)*	*Meteor. Application*
21,22	23.8;31.4	Total Water Content and Surface Characteristics
23	50.3	
24-34	52.8-57.3	Temperature Sounding
35	89.0	Surface Characteristics
AMSU-B		
Chan.	*Frequency (GHz)*	*Meteor. Application*
36,37	89.0;150.0	Surface Characteristics
38-40	183.0	Moisture Sounding

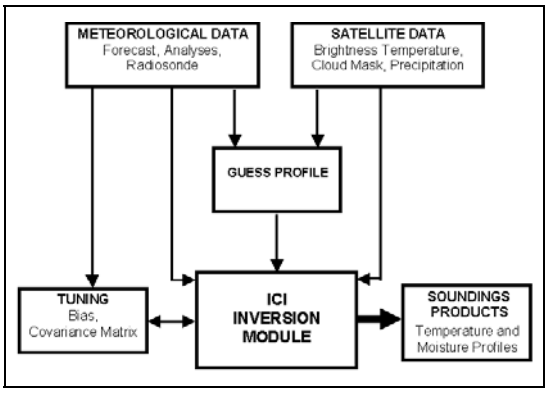


Figure 1. Description simplified of the ICI inversion processing system.

Pre-Processing and Data Sets

In this study 15 days of NOAA-15 satellite data over Brazil were used for reconstructing temperature profiles during the period 23/02/2000 to 08/03/2000. The meteorological data used in the initialization and validation process (radiosondage and numerical weather forecasting data) was supplied by the Brazilian Centre for Weather and Climate Forecast (CPTEC). Since the satellite raw data is received in the HRPT (High Resolution Picture Transmission) format, it is necessary to process it before the retrieval process. The ATOVS and AVHRR Processing Package (AAPP) model was used to perform the ingestion and pre-processing of the HRPT data [12,13]. This procedure supplies calibrated data of brightness temperature for all ATOVS channels, located in the terrestrial coordinates (latitude and longitude) and mapped for a common grid resolution.

In order to use the infrared spectral channels in the retrieval process, it was necessary to perform cloud cover detection [14] in the study area. The microwave spectral channels generally are not affected by clouds, but in some frequencies (specially in the AMSU-B channels), the presence of rain and ice hydrometeors particles becomes important due to scattering and absorption processes. Thus a technique based in the 21, 22 and 35 AMSU-A channels brightness temperature differences was used to perform the precipitation and scattering identification over sea and land locations [15]. The areas identified as contaminated by precipitation or scatterings are not used in the current study.

The Forward Model

The interaction between the radiation that arrive in the satellite sensor and the atmosphere is described by the RTE. For a non-scattering atmosphere in local thermodynamic equilibrium, the RTE is written as:

$$R(\nu) = \tau_s(\nu)\varepsilon_s(\nu)B_s(\nu,T) + \int_{\tau_s}^{1} B(\nu,T)d\tau(\nu) + (1-\varepsilon_s(\nu)\tau_s^2(\nu)\int_{\tau_s}^{1}\frac{B(\nu,T)}{\tau^2(\nu)}d\tau(\nu) \quad (1)$$

where R is the spectral radiance; ν is the channel frequency; B is the Planck function which is a function of the temperature T and pressure p; ε the surface emissivity; τ the layer to space atmospheric transmittance function and the subscript s denotes surface [9,16].

During the retrieval process, the atmospheric radiative transfer model is used many times; thus, the forward model need to be fast enough to work in an operational inversion scheme but sufficiently accurate to maintain the retrieval quality. The ICI system uses the RTTOV-6 model, a fast radiative transference code [17,18], where the transmittance is not calculated directly

but rather through an approximation using a multi-regression method.

If the satellite observed brightness temperature R of each channel is known, then R can be considered a nonlinear function of the atmospheric temperature profile (T), water vapour mixing ratio profile (q), surface skin temperature (T_s), surface emissivity (ε_s), ozone profile, uniformly mixed gases profiles, etc. That is, $R = R(T,q,T_s,q_s,\ldots)$. The uniformly mixed gases are supplied by internal coefficients and a standard ozone profile is utilized (US standard atmosphere 1976). For the infrared spectral channels over sea, the emissivity is given according to the scanning secant, the average wind speed and the surface temperature, using values described in [19]. Over land the emissivity is considered as a constant and set to 1. In the microwave spectral region over sea, the emissivity is calculated using the algorithm FASTEM [20], which takes into account variables such as wind speed, skin surface temperature, etc. Over land, in the current study, the microwave emissivity information is provided through an atlas of emissivity and is dependent on the frequencies and scanning angles [21,22]. Thus, in general we have:

$$\mathbf{R}_i = \mathbf{R}_i(\mathbf{x}), \quad i = 1,\ldots n \tag{2}$$

where n is the number of brightness temperatures and the vector $\boldsymbol{x}$ contains the 57 parameters to be estimated: 40 atmospherics temperatures (that correspond to 40 pressure levels in the atmosphere, from 1000 to 0.1 hPa), 15 atmospheric water vapour mixing ratio levels (from 1000 to 300 hPa), one surface skin temperature and one surface water vapour mixing ratio.

Inversion Scheme of the ICI Model

One can separate the ICI retrieval process in two different steps; the first one is related to the guess selection using a least square optimization in the brightness temperature space, and in the second step the inversion process uses an Bayesian approach to retrieve the temperature and moisture vector.

The selection of the guess profile is performed through a search in the library of temperature and moisture profiles. For each profile i, a distance d_i is computed as:

$$d_i = (\mathbf{r}^m - \mathbf{r}(\mathbf{x}_i))^T \mathbf{B}^{-1} (\mathbf{r}^m - \mathbf{r}(\mathbf{x}_i)) \tag{3}$$

where $\boldsymbol{r}^m$ and $\boldsymbol{r}_i$ represent the vector of brightness temperature of the observed profile and the vector of brightness computed from the candidate to guess profile, respectively; and $\boldsymbol{B}$ is the brightness temperature covariance matrix that is calculated in agreement with the type of prevailing cloud coverage (clear, partially clear and cloudy) and air mass class (polar, middle latitude, tropical). The guess profile is defined as the average of the 10 profiles that minimize the computed distances d_i. The library of guess profile is built sampling the radiosoundings and analyses profiles from the 15 previous days provided in the local acquisition zone. The library contains approximately 2000 profiles (temperature and moisture).

The ICI inversion process consists of finding the most probable atmospheric profile $\boldsymbol{x}$ given the measurements $\mathbf{r}^m$, i.e. of maximizing the conditional probability of $\mathbf{x}$ given $\mathbf{r}^m$: $max\ P(\mathbf{x}|\mathbf{r}^m)$. According to the Bayes theorem, in the case of Gaussian error distributions, the most probable solution is that which minimizes the objective function:

$$\begin{aligned}\mathbf{J}(\mathbf{x}) &= (\mathbf{x}-\mathbf{x}^b)^T \mathbf{C}^{-1} (\mathbf{x}-\mathbf{x}^b) \\ &+ \{\mathbf{r}^m - \mathbf{r}(\mathbf{x})\}^T (\mathbf{E})^{-1} \{\mathbf{r}^m - \mathbf{r}(\mathbf{x})\}\end{aligned} \tag{4}$$

where $\mathbf{x}^b$ is the background profile; $\mathbf{C}$ is the error covariance matrix of the background; $\mathbf{r}(\mathbf{x})$ is the radiative transfer model or "forward model"; $\boldsymbol{E}$ is error covariance matrix of the combined measurement and forward model errors.

An approach to obtain a minimum for Equation (4) is to use the Newtonian iteration method, which employs the following updating rule:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{J}''(\mathbf{x}_n)^{-1} \mathbf{J}'(\mathbf{x}_n)\,, \tag{5}$$

where $\mathbf{J}'(\mathbf{x})$ the gradient of $\mathbf{J}(\mathbf{x})$ with respect to $\boldsymbol{x}$ and $\mathbf{J}''(\mathbf{x})$ is their second derivative of the Equation (4). By matrix manipulation, one can find the following estimation scheme for this problem:

$$\begin{aligned}\mathbf{x}_{n+1} &= \mathbf{x}_n + (\mathbf{x}^b - \mathbf{x}_n) + \mathbf{W}_n \{\mathbf{r}^m - \mathbf{r}(\mathbf{x}_n) \\ &- \mathbf{K}_n (\mathbf{x}^b - \mathbf{x}_n)\} \\ \mathbf{W}_n &= \mathbf{C}\mathbf{K}_n^T (\mathbf{K}_n \mathbf{C} \mathbf{K}_n^T + \mathbf{E})^{-1}\,,\end{aligned} \tag{6}$$

where **K**(**x**) contains the partial first-derivatives of **r**(**x**) with respect to the elements of **x** [23,24]. Iteration in Equation (6) ends when the increment $(\mathbf{x}_{n+1} - \mathbf{x}_n)$ is small enough and when $\{(\mathbf{r}^m - \mathbf{r}(\mathbf{x}_{n+1})\}$ is of the order of the measurement error in all channels.

RESULTS

In practice an important factor that contribute to the retrieval accuracy is the correct selection of the spectral channels that will be used in the inversion process. Channels with large noise and those which cannot be properly simulated by the forward model generally downgrade the quality of the solutions. However, the mere exclusion these channels may also eliminate useful information.

Figure 2 shows the error statistics concerning the mean (bias) and the standard deviation (stdev) of the difference between the simulated and measured brightness temperature. In practice, these errors are caused by various sources such as forward models approximations, measurement errors, among others [25]. As expected, the error levels vary considerably from channel to channel. The accuracy also depends on the surface type (sea and land). In general, surface channels yield the worst result due to their high sensitivity to surface parameters such as the emissivity and the surface temperature. The channels associated with the water vapor absorption band (moisture channels) also show a larger error (stdev). We can also see that the values of standard deviation are higher over land, especially in the AMSU channels. This occurs because the surface parameters have lower variability and easier to be estimated over sea. However, one can notice that the forward model underestimates the brightness temperature, althought a bias correction scheme was applied in the measured brightness temperature prior to its use in the inversion process.

Table 2. Channels selection used in the ICI inversion process.

Case	Channel Selection	
	Clear Sky	Cloud Sky
T1	2-8; 13-16; 25-32	2-3; 25-32
T2	2-8; 11-16; 25-32; 38-40	2-3; 25-32; 38-40
T3	2-8; 10-16; 24-32; 37-40	2-3; 24-32; 37-40
T4	2-8; 11-16; 25-32; 38-40 + TPW	2-3; 25-32; 38-40 + TPW

To study the performance of ICI viz-a-viz the selection of spectral channels used in the retrieval, we considered four different strategies (see Table 2). For the T1 case, we used only temperature sounding channels. In the T2 case, a combination of temperature and moisture channels were used, but the surface channels were not included. In the T3 case, some surface channels were added, and, finally, for the T4 case, instead of surface channels, additional information was used in the guess choice process. This additional information takes the form of a "pseudo channel" constraining the integrated atmospheric water vapor content (i.e., the total precipitable water-

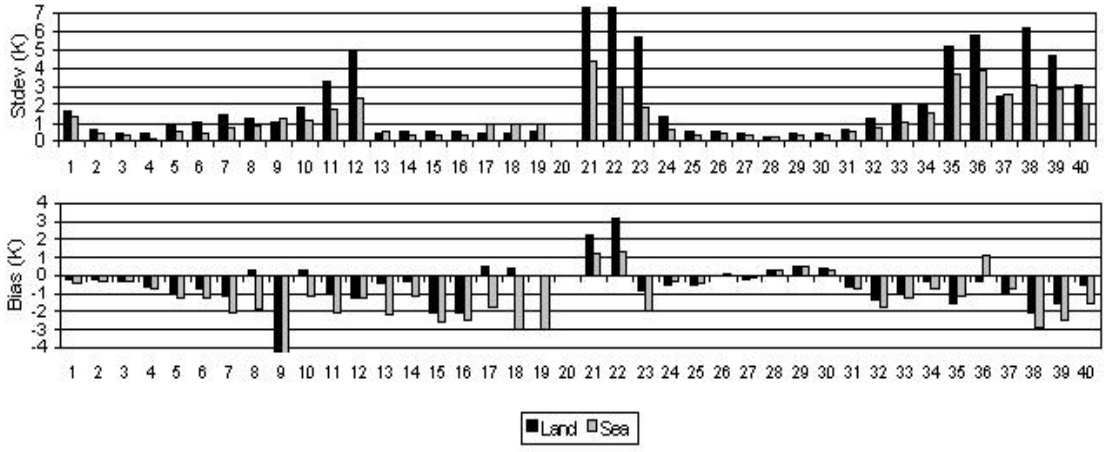


Figure 2. Error statistics concerning the mean (bias) and the standard deviation (stdev) of the difference between the simulated and measured brightness temperature. Statistics performed for the NOAA-15 satellite from 28/02/2000 to 08/03/2000.

TPW). The TPW over sea is estimated from channels 21,22 e 23 [15], while over land a multiple-regression algorithm using various channels is applied.

Temperature Retrievals

Figures 3 and 4 show the temperature mean RMS error for retrievals over, respectively, land and sea, for clear and cloudy sky conditions, and test cases T1, T2 and T3. As expected, since the channels errors (stdev) are smaller over sea (Figure 2), we found that at the lower surface levels (up to 700 hPa), the temperature estimates are less accurate over land than over sea. On the other hand, the impact of the cloud cover type (cloudy or clear) was not significant, which means that the information supplied by the infrared channels produces a small improvement on the retrieved profile.

Concerning the different channel selection strategies, we observed that T1 yielded the worst results for all situations while T2 and T3 showed the smallest error level, respectively, over land and sea. This result suggests that the use of surface channels (as in T3 case) is only useful over sea although other cases need to be evaluated in order to obtain a statistically significant result.

Table 3 shows the values of RMS error (from 1000 to 10 hPa) for the guess and retrieval profiles (T3 case). We notice a considerable error decrease as the result of the inversion process. The largest improvement in the temperature profile is observed over land for clear sky conditions.

Table 3. RMS error for guess and inversion profile for the case T3 for the follow variables: temperature (from 1000 to 10 hPa) and moisture (from 1000 to 500 hPa).

Surf. type	Temp. (K)		Moist. (g/kg)	
	Guess	Invers.	Guess	Invers.
Sea clear	1.03	0.82	1.00	0.91
Land clear	1.18	0.88	1.32	1.29
Sea cloud	1.09	0.86	0.99	0.92
Land cloud	1.07	0.87	1.21	1.20

The previous analysis concern the statistics of the mean error, but for meteorological purposes it is also important to know the horizontal structure of the retrieved fields. Our results showed that the retrieved temperature fields for different pressure levels were similar to the observation (true) fields at the same levels, with a RMS error not exceeding 2 K, except in a few locations near the surface.

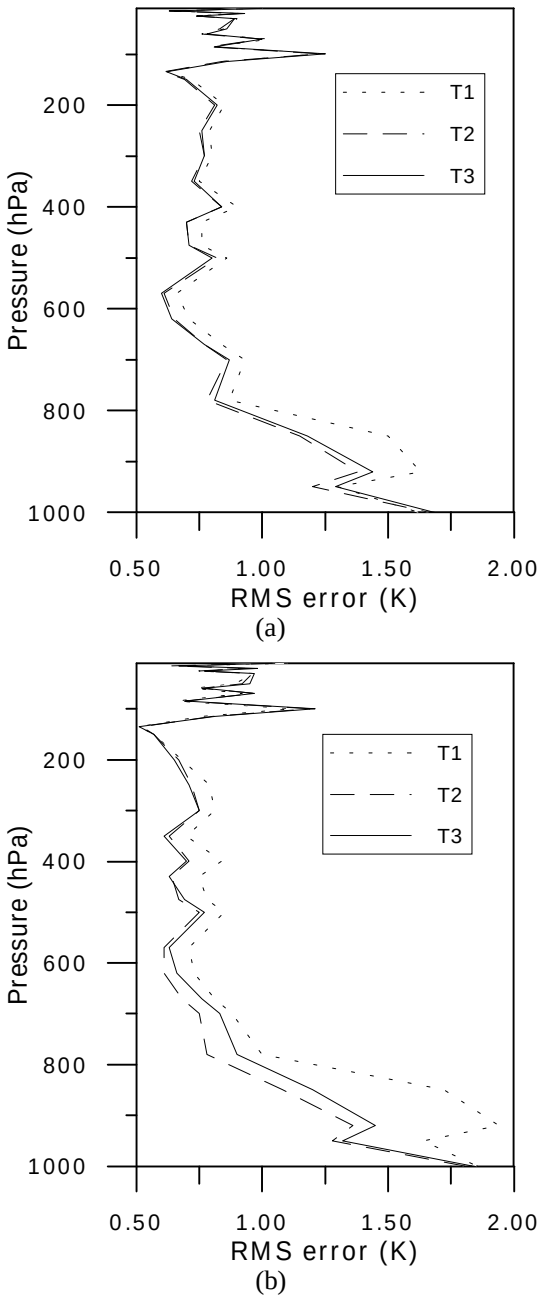


Figure 3. Mean RMS error for temperature profile over land (T1, T2 and T3 channel configurations): a) clear sky condition (1292 profiles) and b) cloud sky conditions (1664 profiles).

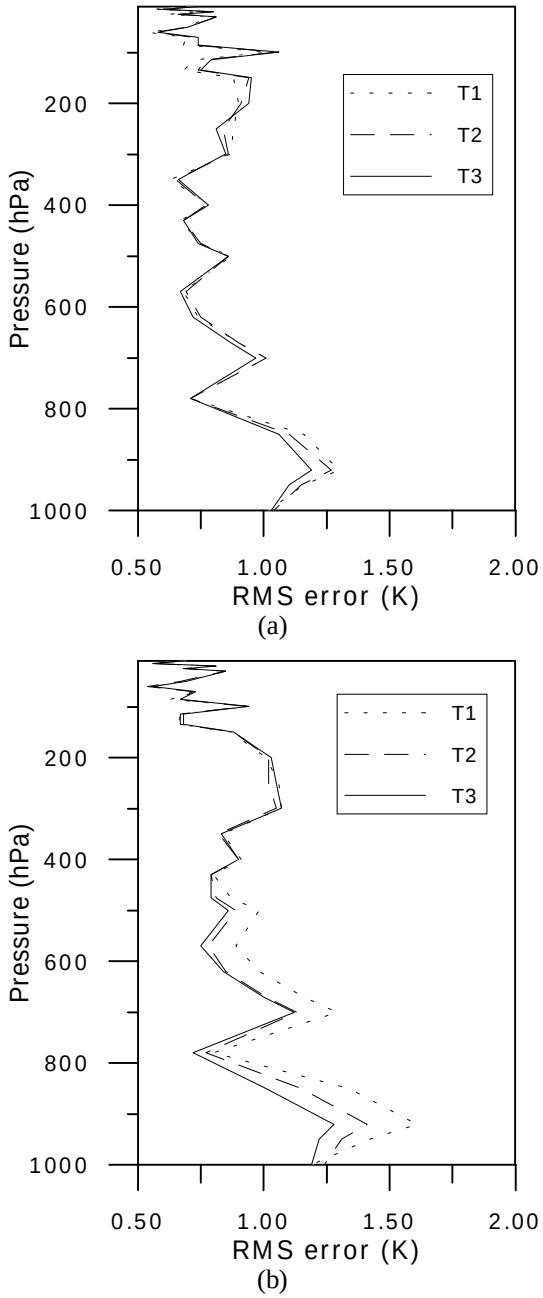


Figure 4. Mean RMS error for temperature profiles over sea (T1, T2 and T3 channel configurations): a) clear sky condition (1279 profiles) and b) cloud sky conditions (847 profiles).

Moisture Retrievals

In addition to the problems already mentioned for retrieving the temperature profiles, the moisture inversion process involves additional difficulties. The moisture fields usually are subject to a large spatial and temporal variability, especially in tropical regions, which makes difficult the comparison with observed data, necessary for validation purposes. Moreover the water vapor channels present a larger noise level.

Figures 5 and 6 show the mean RMS error for moisture retrievals over, respectively, land and sea, for four different channels configurations. As in the case of temperature retrieval, we found that at the lower surface levels (up to 700 hPa), the temperature estimates are less accurate over land than over sea. Also, the error levels found for different cloud coverages were similar. As expected, when the moisture sensitive channels were not used (T1 case), the error levels are larger in all situations. Overall, the best results were obtained with the T4 channel configuration, except over land for cloudy condition, because the TPW channel was not used.

Differently from the temperature retrieval cases, the impact caused by the moisture inversion process in the mean error level decrease is small for almost all situations (Table 3). However, a qualitative analysis of individual horizontal fields show that the ICI inversion system identifies properly the structure of the moisture field. The areas with higher and lower water vapor contents are correctly represented, mainly in over sea. On the other hand, there are some regions where the error became significant, especially in soundings over land.

CONCLUDING REMARKS

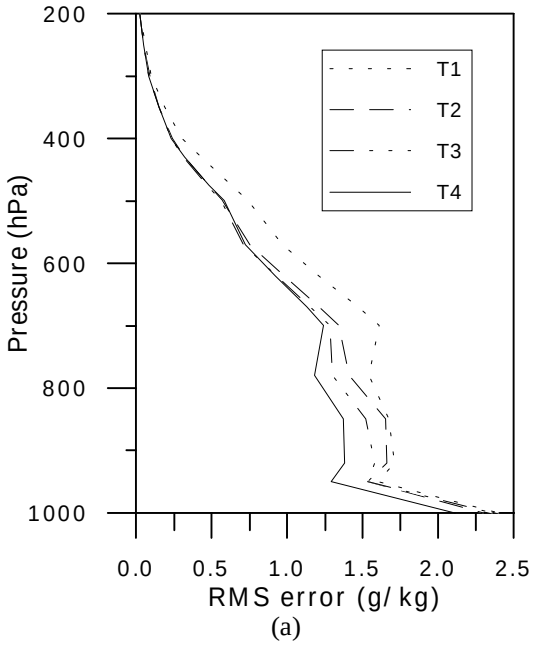
In this paper we investigated the problem of retrieving atmospheric temperature and moisture profiles from satellite data. Emphasis was given in the analysis the role of different channel combination used in the retrieval process. Different surface type (sea and land) and atmospheric conditions (clear and cloud sky) were also considered.

The results showed that it is easier to retrieve temperature and moisture profiles over sea than over land. This occurs because over land, the forward model is less accurate due to the difficulty of estimating surface parameters such as the emissivity and the surface temperature. Surprisingly, we found the atmospheric conditions do not affect significantly the accuracy of the inversion process. Also even without most of the infrared channels (cloudy sky situations) it is possible to obtain accurate retrievals. We also

observed that the use of TPW as additional source of information improve considerably the error statistics of moisture soundings over sea. Finally, we stress that the present results and conclusions are representative of the period and region in study and, therefore, further studies are necessary to obtain more general conclusions.

ACKNOWLEDGEMENTS

This work was supported by FAPESP-Brazil. The first author thanks the CNPq-Brazil by the financial support given through of the PhD program 142400/1998-0. NJF and FMR also acknowledge the support given by CNPq thought the research Grant 300486/96-0 and 300171/97-8, respectively.



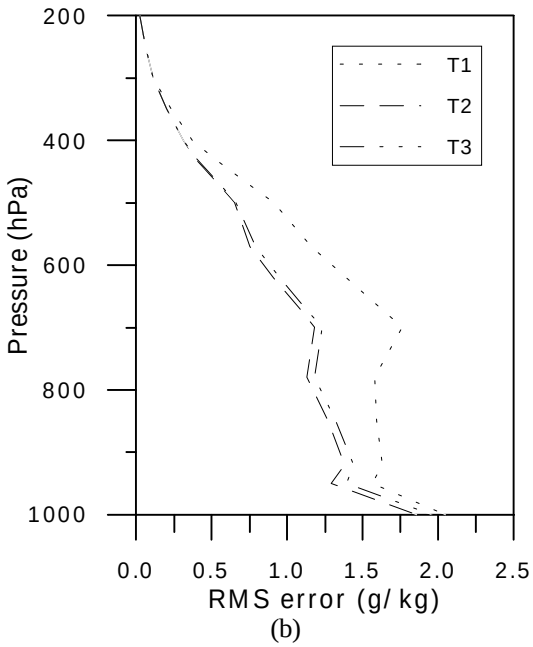
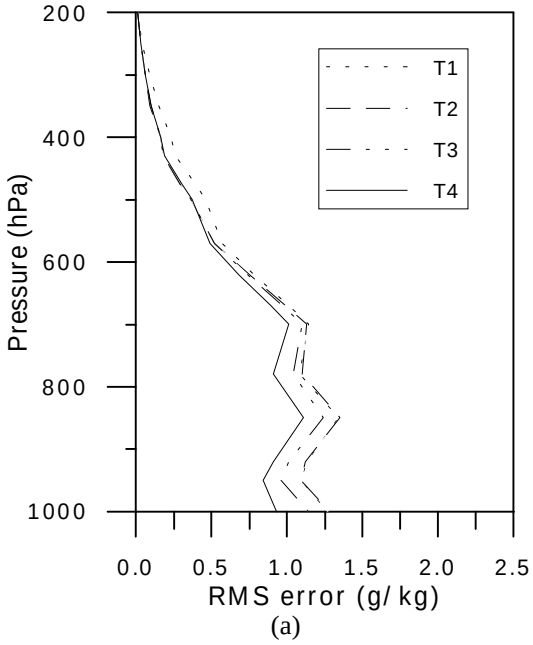


Figure 5. Same as Fig. 3 but for moisture profile.



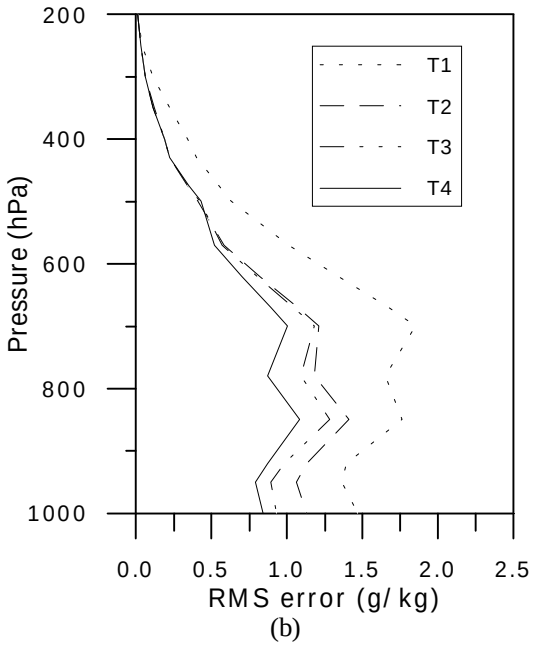


Figure 6. Same as Fig. 4 but for moisture profile.

REFERENCES

1. C. D. Rodgers, Retrieval of the atmospheric temperature and composition from remote measurements of thermal radiation, Rev. Geophys. Space Phys., **14**, 609-624 (1976).

2. S. Twomey, Introduction to the mathematics of inversion in remote sensing and interative measurements, Amsterdam, Elsevier Scientific, 1977.

3. M. T. Chahine, Inverse problems in radiative transfer: determination of atmospheric parameters, Jour. Atmos. Sci., **27**, 960 (1970).

4. W. L. Smith, H. M. Woolf, A. J. Schriener, Simultaneous retrieval of surface and atmospheric parameters: a physical analytically direct approach, Adv. in Rem. Sens., **7** (1985).

5. L. Lavanant, P. Brunel, G. Rochard, T. Labrot, D. Pochic, Current status for the ICI retrieval scheme, Tech. Proc. of the IX-ITSC (1997).

6. J.-P. Chaboureau, A. Chédin and N. A. Scott, Remote sensing of the vertical distribution of the atmospheric water vapor from the TOVS observations: Method and validation, Jour. of Geophys. Res., **103**, 8743-8752 (1998).

7. J. C. Carvalho, F. M. Ramos, N. J. Ferreira, H. Campos Velho, Retrieval of vertical temperature profiles in the atmosphere, 3rd International Conference on Inverse Problems in Engineering, Port Ludlow, USA (1999).

8. F. M. Ramos, H. F. Campos Velho, J. C. Carvalho, N. J. Ferreira, Novel approaches on entropic regularization. Inverse Problems, **15**, n.5, 1139-1148 (1999).

9. J. Li, W. W. Wolf, W. P. Menzel, W. Zhang, H.-L. Huang, T. H. Achtor, Global soundings of the atmosphere from ATOVS measurements: the algorithm and validation, Jour. of App. Met. **39**, 1948-1968 (2000).

10. G. Goodrun, K. B. Kidwell, W. Winstons, NOAA KLM user's guide (1998).

11. L. Lavanant, P. Brunel, G. Rochard, T. Labrot, NOAA15 Soundings profiles retrieved with the ICI scheme, Tec. Proc. of the X-ITSC (1999).

12. K. D. Klaes, ATOVS data processing in Europe, Tec. Proc. of the IX-ITSC (1997).

13. K. D. Klaes, R. Schraidt, The European ATOVS and AVHRR Processing Package (AAPP) Development, Tec. Proc. of the X-ITSC (1999).

14. L. Lavanant, H. LeGleau, M. Derrien, S. Levasseur, G. Monnier, L. Ardouin, P. Brunel, B. Bellec, AVHRR cloud mask for sounding applications, Tec. Proc. of the X-ITSC (1999).

15. N. Grody, F. Weng, R. Ferraro, Application of AMSU for obtaining water vapour, cloud liquid water, precipitation, snow cover and sea ice concentration. Tec. Proc. of the X-ITSC (1999).

16. K. N. Liou, An introduction to atmospheric radiation, academic press, Orlando, 1982.

17. J. R. Eyre, A fast radiative transfer model for satellite soundings system, ECMWF Technical Memorandum, **176** (1991).

18. R. Sanders, M. Matricardi, P. Brunel, An improved fast radiative transfer model for assimilation of radiance observations, Quart. Jour. R. Met. Soc., **102**, 1407-1425 (1998).

19. K. Masuda, T. Takashima, Y. Takayama, Emissivity of pure and sea water for the model sea surface in the infrared windows regions, Rem. Sens. of Envir. (1998).

20. S. J. English, T. J. Takashima, A fast generic millimetre-wave emissivity model, Microwave Remote Sensing of the Atmosphere Environment, SPIE, 3503 (1998).

21. C. Pringent, E. Mathews, R. Rossow, Microwave land surface emissivities estimated from SSM/I observations, Jour. Geophys. Res., **102**, 21867-21890 (1997).

22. C. Pringent, J. P. Wigneron, W. B. Rossow, J. P. Pardo-Carrion, Frequency and angular variations of land surface microwave emissivities: Can we estimate SSM/T and AMSU emissivities from SSM/I emissivities?, IEEE Trans. on Geos. and Rem. Sens., **20** (1999).

23. ICI Documentation, Description Document for the ICI sounding software related to meteorological NOAA orbiting satellites, Scientific Part (1999).

24. J. R. Eyre, Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory ans Simulation for TOVS, Q. Jour. R. Met. Soc., **115**, 1001-1026 (1989).

25. M. J. Uddstrom and L. M. McMillin, Extraction of atmospheric signals from radiance measurements: some limitations, High Spectral Resolution Infrared Remote Sensing for Earth's Weather and Climate Studies, Paris, Springer-Verlag (1993).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INVERSION OF ELASTIC LIGHT SCATTERING MEASUREMENTS TO DETERMINE REFRACTIVE INDEX AND PARTICLE SIZE DISTRIBUTION OF POLYMERIC EMULSIONS

Gloria L. Frontini
Instituto de Investigación en Ciencia y Tecnología de Materiales (INTEMA)
Univ. Nac. de Mar del Plata
Mar del Plata, Argentina
gfrontin@fi.mdp.edu.ar

Elena M. Fernández Berdaguer
Instituto de Cálculo FCEyN , Fac. Ingeniería.
Univ. Nac. de Buenos Aires

ABSTRACT

Polymeric emulsions are well characterized by the knowledge of their particle size distributions (PSD). Elastic-light-scattering (ELS) measurements can be inverted to estimate the PSD in the range from 50 nm. up to several micrometers. The relative refractive index of the particles is required in computation procedures to obtain the PSD. Small differences in the assumed refractive index may cause significant differences in the resulting PSD. From the scattering data, the refractive index can be determined. In this paper we present the corresponding technique. We consider polymeric emulsions where the non-absorption assumption is reasonable.

We propose a methodology based on Tikhonov regularization applied only to the distribution. However, we solve the minimization problem simultaneously with respect to the distribution and the refractive index. To select the regularization parameter, we include the Generalized Cross Validation technique. From simulated ELS measurements we show that the problem is solved successfully.

NOMENCLATURE

A matrix form of operator T
D particle diameter
f particle size distribution (PSD)
f vector form of PSD
g noisy ELS measurement
H matrix form of L
I light scattering intensity
J functional
q smoothing restriction on f
m relative particle refractive index
n_1 particle refractive index
n_2 solvent refractive index
S Lorentz-Mie scattering function
T Operator that represents the integral equation
Greeks
ε experimental error
γ regularization parameter
θ scattering angle
Superscript
$\sim$ estimated values
$*$ adjoint

INTRODUCTION

Emulsion polymerization systems are typically composed by three coexisting phases: a continuous (usually aqueous) phase, monomer droplets, and polymer particles. In particular, the size of the polymer particles greatly affects the properties of the final material. So, the knowledge of the emulsion particle size distribution (PSD) is important to characterize the material.

The most popular technique for the determination of PSDs is elastic light scattering (ELS) [1], since it is an easy to perform and non-destructive experimental technique. The emulsion sample is illuminated by a monochromatic beam and the scattered light, averaged over time, is measured as a function of the scattering angle. This angular intensity, having the same wavelength as the incident beam, is influenced by the size, the shape and the optical contrast of the particles (Mie theory [2]).The difficulty associated with this technique is related to the inversion of the measurements to extract the desired information from the data. Size, shape, and optical contrast can be described by a large number of parameters, which cannot completely be extracted from the intensity spectrum stem from a light scattering experiment. The reasons of this fact are the ill-posed nature of the problem and the statistical noise of experimental data. The ill-posed nature of the problem is due to the lack of identifiability of the parameter. This problem is mentioned in the literature as existence of singularities [3]. In our case the refractive index seems to be identifiable.

The determination of the PSD from ELS measurements, assuming that all parameters in the model are known exactly has been studied by several authors [4,5]. Since the inverse problem can be stated as linear, several

regularization methods can be applied [6]. There is a large variety of inversion techniques for light scattering data [7,8]. In general, these techniques are classified as either analytical or empirical. Analytical techniques involve formal solutions of integral equations that describe the light scattering process, and require the use of *a priori* information regarding the distribution function because of the ill-posed nature of the inverse problem. Empirical inversion techniques generally require that a parametric model of the light scattering process be developed. The parameters are then adjusted within physically realistic bounds so that a least squares fit of the measured data is obtained.

We are interested in the problem of determining PSD and relative refractive index of polydisperse colloidal particles. In this case the inverse problem is non-linear. Some previous publications have considered this problem. Schnablegger and Glatter [3], reported a methodology to retrieve PSD and refractive index from ELS in the presence of scattered light reflected from the walls of the sample holder. They represented the sought distribution by a series of β-spline functions and imposed to the solution smoothing and positive constraints, and determined the regularization parameter by means of a sensitivity plot constructed from the residuals. Jones et. al. [10], combined analytical and empirical inversion techniques to obtain optical parameters and PSD. They chose an orthonormal base to expand the solution and followed a procedure consisting in sequential steps. First, the retrieval of the refractive index through the use of the unconstrained solution, then, the retrieval of the PSD through the use of the constrained solution, and finally the retrieval of the absorption index by matching the measured and calculated scattering patterns. They concluded from simulated experiments that this procedure gives the best results for narrow distributions.

Our goal is to develop a more general and analytical methodology to estimate simultaneously PSD and refractive index from ELS measurements. Our approach involves techniques for the automatic determination of the regularization parameter, and yields good results with fewer empirical considerations.

THE DIRECT PROBLEM

Consider a given particle of a particular shape, size and composition, which is illuminated by a light beam of a specific wavelength, intensity and polarization. The determination of the resulting electromagnetic field from the scattering process, based on the solution of Maxwell equations is called the 'direct problem'. The most important problem in the theory of absorption and scattering by small particles that can be solved using a mathematical closed formulae is that by a sphere of arbitrary radius and refractive index. This formulae was derived by Mie[11] and other authors [12]. The angular scattered intensity due to a polydisperse system can be regarded as a linear combination of Lorenz-Mie form factors $S(\theta,D,m)$, where θ is the scattered angle, D is the particle diameter, and m is the relative complex refractive index of the particles in the solvent. If we proceed and regard the coefficients of the linear combination as a continuous function $f(D)$ that represents the emulsion PSD , we obtain the integral expression that has to be evaluated to find the angular-dependent scattering curve of the polydisperse system, as

$$I(\theta,m) = \int_{D_{\min}}^{D_{\max}} S(\theta,D,m)\, f(D)\, dD\,,$$

$$\theta \in [\theta_1,\theta_2] = \Theta \subset \Re \qquad (1)$$

For the calculations of the direct problem we use the computer programs reported in the literature [13]. A similar program is also given in [14].

The complex relative refractive index m is defined as the ratio of the complex refractive index of the particle n_1 to that of the solvent n_2 . Because of the small value of the imaginary part of the refractive index (i. e. , absorption coefficient), less than 0.01 for polymeric emulsions, m is approximated by its real value (i. e., scattering coefficient). Scattering and absorption coefficients are usually referred as optical constants. We use this approximation since for the range of wavelengths of the incident beam commonly used in the experimental equipment, the absorbed light by the polymer particles can be neglected for the materials we are considering.

As a concluding remark, it can be said that the availability of reliable optical constants is critical for the use of Mie theory.

THE INVERSE PROBLEM

In practice, it often occurs that the particles responsible for the scattering cannot be analyzed directly. From a study of the scattered field, we then have to determine the characteristics of the particles.

In this article, we consider that the particle characteristics to be determined from the scattered intensity spectrum, $I(\theta,m)$, are $f(D)$ and m. All other parameters are assumed known. It is obvious from Eq. (1) that the relation between $I(\theta,m)$ and $f(D)$ is linear. Contrarily, m appears in the kernel in a non-linear way, therefore $I(\theta,m)$ is non-linear in m. Thus, the determination of all the unknowns needs the solution of a non-linear inverse problem.

There are two aspects to be taken into account to solve this class of non-linear inverse problem. The first one, the relations described in last paragraph, will allow us to derive a quasi-analytical solution. The second is that the problem ill posedness is only with respect to $f(D)$. This property can be described by the fact that small perturbations on m produces large differences in $I(\theta,m)$ when one evaluates the direct problem, behavior opposite to that observed for ill posedness. On the other hand, the decay rate of the singular values of the kernel of the integral equation, which can be used as a measure

of the degree of ill-posedness [5], is practically not affected by the value of *m*.

Regularization of the Inverse Problem

The general formulation of the inverse problem addressed in last paragraphs is the following.

Consider the integral equation

$$T(m)[f] = \int_{D_{\min}}^{D_{\max}} S(\theta, D, m) f(D) dD = I(\theta, m), \qquad \theta \in \Theta \tag{2}$$

where the distribution *f(D)* and the parameter *m* are the unknowns. The set of admissible parameters is

$$\mathcal{P} = \{(m,f): m \in [M_0, M_1], f \in L^2([D_{min}, D_{max}]; \Re)\}.$$

Noisy measurements $g(\theta)$ are available, with experimental error $\varepsilon(\theta)$, i.e. :

$$g(\theta) = I(\theta, m^o) + \varepsilon(\theta) \tag{3}$$

where m^o is the actual value of parameter *m*.

We propose to find the solution of the problem optimizing the following functional

$$\underset{\mathcal{P}}{Min}\, J(m,f) = \left\| T(m)[f] - g \right\|_{L^2}^2 + \gamma \left\| q(f) \right\|_{L^2}^2 \tag{4}$$

Eq. (4) shows that we apply Tikhonov's regularization [15] only to *f(D)*. *q(f)* is the smoothing restriction included as the a-priori information about the sought distribution and γ the regularization parameter which weighs this inclusion.

The possible local minima are found solving the equations:

$$\frac{\partial}{\partial f} J(m,f) = 0\,, \qquad \frac{\partial}{\partial m} J(m,f) = 0\,.$$

In the first equation, *m* is fixed, and the derivative must be interpreted as a functional derivative because f is a function (see Appendix A). The solution of this equation is the classical for linear inverse problems, i.e.

$$f = \left(T(m) * T(m) + \gamma q\right)^{-1} T(m) * g\,. \tag{5}$$

In the second equation, for fixed *f*, *J(m,f)* is a function of the real variable *m*. We proceed as follows to solve it. Writing *J(m,f)* in terms of the scalar product we obtain

$$J(m,f) = \left(T(m)[f] - g, T(m)[f] - g\right)_{L^2} + \left(q(f), q(f)\right)_{L^2}.$$

Differentiating with respect to *m* we have

$$\frac{\partial J}{\partial m}(m,f)\delta m = 2\left(\frac{\partial}{\partial m} T(m)[f]\delta m, T(m)[f] - g\right)_{L^2}$$
$$= 2\left(\delta m, \left(\frac{\partial}{\partial m} T(m)[f]\right)^* \left(T(m)[f] - g\right)\right)_{L^2}.$$

Since the above equation holds for every $\delta m \in \Re$, we obtain the formula

$$\left(\frac{\partial}{\partial m} T(m)[f]\right) * \left(T(m)[f] - g\right) = 0 \tag{6}$$

Substitution of Eq. (5) into Eq. (6) yields:

$$\left(\frac{\partial}{\partial m} T(m)[\left(T(m) * T(m) + \gamma q\right)^{-1} T(m) * g]\right) *.$$
$$.\left(T(m)[\left(T(m) * T(m) + \gamma q\right)^{-1} T(m) * g] - g\right) = 0\,. \tag{7}$$

In fact, the solution of this single equation [Eq. (7)] in terms of *m*, for a specific value of the regularization parameter γ, gives an estimation of the parameter $\tilde{m}$. Numerically this is calculated easily since the unknown *m* is one-dimensional. Evaluation of Eq. (5) using $\tilde{m}$ gives the estimation of the PSD, $\tilde{f}$.

Although γ, *m*, and *f* must be determined to solve the problem, a prior selection of the regularization parameter must be performed, since γ is constant in the minimization problem (Eq. 4), as usual in Tikhonov's regularization. The relation between the selection of the regularization parameter and the quality of the estimations is analyzed in the next section.

Selection of the Regularization Parameter

We follow two different approaches for the selection of the regularization parameter. The first one can be applied in any situation since no estimation on noise level is needed. It involves an iterative procedure which behavior is robust and gives good estimations of the problem unknowns. The second approach is more analytical and needs no iterations. However, it needs one to assume bounds on the perturbation of the sought parameter and on the measurement noise which greatly affects the quality of the solution. Thus, its application in real experiments may be more difficult.

Iterative method. For any γ Eq. (7) can be solved and the value of m determined. It has been noticed that regardless the γ value, the solution $\tilde{m}$ is nearly the same. In fact, this is true as long as γ is in a range where multiple solutions in Eq. (7) are avoided. Contrarily, the estimated distribution $\tilde{f}(D)$ obtained evaluating Eq. (6) for each γ may differ greatly. It is necessary to obtain the optimal value of the regularization parameter to retrieve the correct distribution. The Generalized Cross Validation (GCV) technique developed by Golub et. al.

[16] for the selection of the regularization parameter for linear inverse problems could be used. This is possible since once *m* is determined, the inverse problem stated by Eq.(2) becomes linear.

We propose an iterative procedure for the selection of the regularization parameter, γ, to find the estimation of the unknown parameter in the kernel, $\tilde{m}$, and the unknown distribution $\tilde{f}$, that can be summarized as follows

1. Select an initial value of the regularization parameter, γ_0.
2. Find $\tilde{m}_1$ such that Eq. (7) is fulfilled.
3. Find a new value of the regularization parameter γ_{GCV_1} applying the GCV technique, for the linear problem stated using $\tilde{m}_1$.
4. Repeat steps 2 and 3 finding $\tilde{m}_i$ and γ_{GCV_i} for *i=1, 2 . . .* , until the parameters values stay invariant. Call $\tilde{m}$ and γ the final values.
5. Find $\tilde{f}$ evaluating Eq. (6) .

The application of this iterative process gives, in few steps, good estimations of the solution of the inverse problem.

One step method. We analyzed an alternative method based on Neubauer's technique [17] to select the optimal regularization parameter that can be used when good estimations of the modeling error and the noise level present in the measurement are known.

Let m^o be the exact value of parameter m . As a consequence of the uncertain value of parameter *m* in the kernel, there is an error *h* that can be considered as a modeling error, given by

$$\left\| T(m) - T(m^o) \right\| = h(m) \le \tilde{h} \ . \qquad (8)$$

Let also $\delta = \|\varepsilon\|$ be the noise level present in the measurement. Neubauer [17] derived a method for choosing the regularization parameter that takes into account this situation. (See the appendix). The constant value of γ in Eqs. (4), (6) and (7) should be changed by $\gamma(m)$, the optimal regularization parameter that corresponds to the problem defined to each value of *m*. Notice that an additional term containing $d\gamma(m)/dm$ will appear in Eqs. (5) and (7). The simultaneous solution of Eq. (7) and Eq. (A-3) gives, in one step, the values of $\tilde{m}$ and of the regularization parameter $\gamma_{Neub} = \gamma(\tilde{m})$. Finally, as before, $\tilde{f}$ is found from Eq. (6). It should be said that Neubauer's method is derived for $\|L(f)\|_{L^2}^2 = \|f\|_{L^2}^2$ in Eq. (4).

In practice, *h(m)* can never be known exactly, since $T(m^o)$ is unknown, then a superior bound $\tilde{h}$ should be estimated. In this case the obtained solution my differ greatly form the optimal.

COMPUTATIONS AND RESULTS

To illustrate the validity of the regularization method to obtain the refractive index and the PSD of a polymeric emulsion from the knowledge of the scattered intensity spectrum, we consider two examples.

In order to compare the results for situations involving random measurement errors, we assume normally distributed uncorrelated errors with zero mean and constant standard deviation. The simulated noisy measurements can be expressed, in the discrete domain, as

$$g_{\varepsilon i} = g_i + \varepsilon_i \qquad (9)$$

where $g_i = I(\theta_i)$ is the exact solution of the direct problem for the exact value of the parameter m^o and the exact distribution **f,** corresponding to a particular scattered angle θ_i, and ε_i the noise added at that angle. Let us write the discrete version of the inverse problem stated by Eq. (4) as

$$\underset{\mathbf{f},m}{Min} \quad \mathbf{J} = \left\| \mathbf{g}_\varepsilon - \mathbf{A}(m)\mathbf{f} \right\|^2 + \gamma \mathbf{f}^{\mathbf{T}} \mathbf{H} \mathbf{f} \qquad (10)$$

where vector $\mathbf{g}_\varepsilon$ represents the scattered intensitiy for all measured angles, matrix $\mathbf{A}(m)$ the discrete form of the operator *T(m)* , $\mathbf{H}=\mathbf{K}^{\mathbf{T}}\mathbf{K}$ and $\mathbf{Kf}$ is the discrete form of *q(f)*. We consider in all cases *q(f)* as the second derivative of *f*.

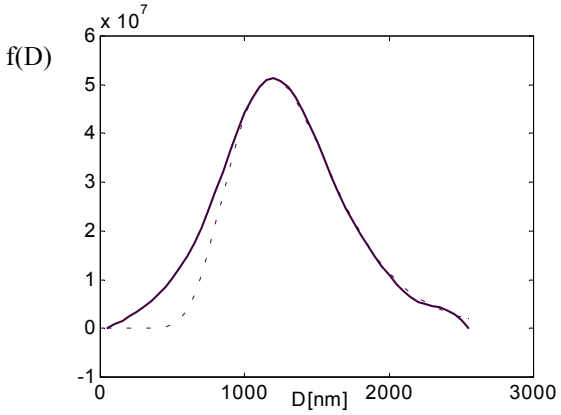


FIGURE 1.*Real (- - -) and Estimated (—) PSD for Example I*

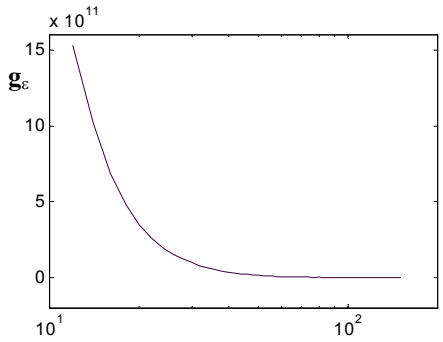


FIGURE 2. *Light Scattering Spectrum for Example I*

The first example corresponds to a polystyrene emulsion in water (m^o=1.1867) having spherical particles with diameters in the range of 50 nm to 2550 nm distributed as a broad number PSD, as the one shown in Figure 1. The simulated spectrum $\mathbf{g}_\varepsilon$ (Figure 2) corresponds to 70 equally spaced scattered angles from 12º to 150º, where the noise standard deviation σ_ε is equal to 1% of the mean value of the measurements. We use this data to retrieve m^o and f, the solution of the inverse problem**.**

The results obtained following the iterative method are $\tilde{m}$=1.1867 and $\tilde{\mathbf{f}}$ as in Figure 1 (solid line), for γ= 4.0E5. The iterative process is shown for different values of γ_o in Table 1. For the $\tilde{m}$ value obtained for each step we applied the GCV technique to obtain a new value of the regularization parameter, γ_{GCV}. The last two columns show that the convergence of the method is achieved from any initial value.

In order to help understanding the behavior of the methodology we plot the evolution of a function that we call $J_\gamma(m)$ in Figure 3. $J_\gamma(m)$ is the functional $J(m,f)$ given by Eq. (4) in which Eq. (5) was introduced. For each γ we find a curve $J_\gamma(m)$ vs. m, which minimum occurs at $\tilde{m}$. This value is in fact the same found solving Eq. (7).

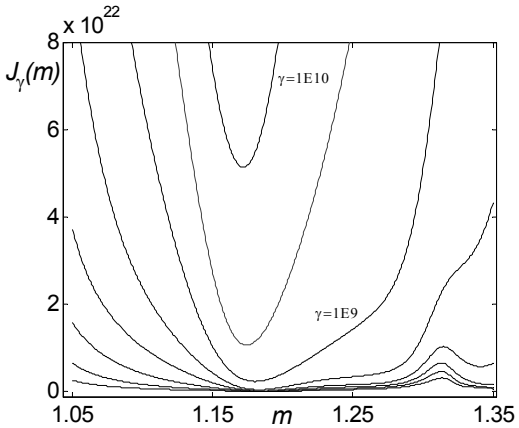


FIGURE 3 *Evolution of the functional for Example I*

Table 1. *Results of the iterative method for Example I*

γ_0	$\tilde{m}_1$	γ_{GCV_1}	$\tilde{m}_2$	γ_{GCV_2}	$\tilde{m}_3$	γ	$\tilde{m}$
1.0 E-1	undetermined						
1.0 E 0	1.1895	1. E 5	1.1875	3.2E5	1.1869	4.0E5	1.1867
1.0 E 1	1.1890	1.6 E	1.1875	3.2E5	1.1869	4.0E5	1.1867
1.0 E 2	1.1885	2. E 5	1.1870	4.0E5	1.1867	4.0E5	1.1867
1.0 E 3	1.1885	2. E 5	1.1870	4.0E5	1.1867	4.0E5	1.1867
1.0 E 4	1.1850	3.2E 5	1.1870	4.0E5	1.1867	4.0E5	1.1867
1.0 E 5	1.1875	3.2E 5	1.1870	4.0E5	1.1867	4.0E5	1.1867
1.0 E 6	1.1860	5.0E 5	1.1865	5.0E5	1.1866	5.0E5	1.1866
1.0 E 7	1.1845	1.0E 4	1.1850	3.2E5	1.1869	4.0E5	1.1867
1.0 E 8	1.1830	5.0E 3	1.1885	2.0E5	1.1872	4.0E5	1.1867
1.0 E 9	1.1805	2.5E 3	1.1870	4.0E5	1.1867	4.0E5	1.1867
1.0E 10	1.1745	8.0E 2	1.1885	2.0E5	1.1872	4.0E5	1.1867

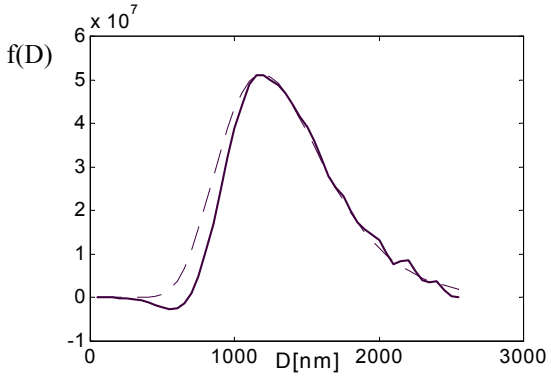


FIGURE 4. *Real (- - -) and Estimated (—) PSD for Example I obtained by the one step method*

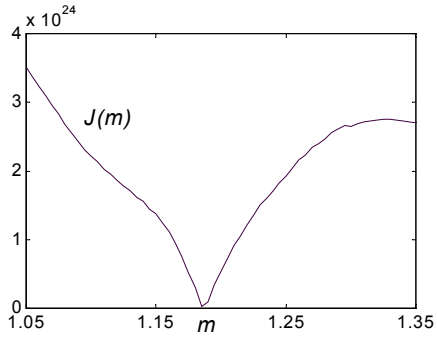


FIGURE 5. *Evolution of the functional for Example I obtained by the one step method*

The results stem from applying the one step method are $\gamma_{Neub} = 1.1\,E4$, $\tilde{m}$=1.1867 and $\tilde{\mathbf{f}}$ as in Figure 4. This solution was obtained using the exact norm of the model-

ing error *h(m),* not available in real experiments. We tried for several values of $\tilde{h}$ with no success, showing that the one step method is not appropriate for this case. The difference between the retrieved PSDs in Figures 1 and 4 is due to the fact that in the one step method $\|q(f)\|_{L_2}^2 = \|f\|_{L_2}^2$, as we said in the previous section.

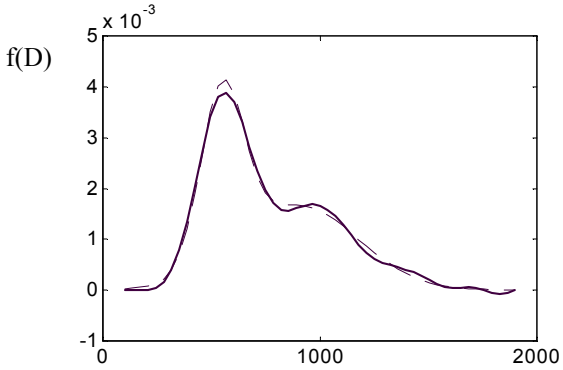


FIGURE 6. *Real (- - -) and Estimated (—) volumePSD for Example II obtained by the iterative method.*

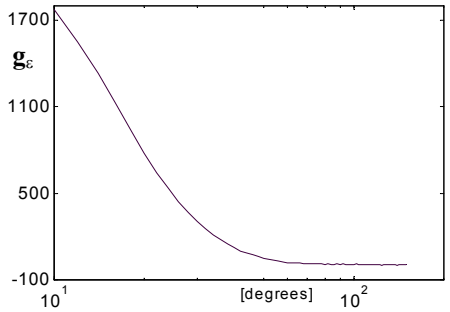


FIGURE 7. *Light Scattering Spectrum for Example II*

The second example was taken from the literature [3]: a bimodal gaussian distribution shown in Figure 6 represents the volume distribution of an emulsion. Its actual relative refractive index is m^o=1.25. The simulated measurements, shown in Figure 7, were generated with additive noise with statistical parameters as in the first example.

For this case we show the results obtained using the iterative procedure. The result for the steps found for two different starting points are shown in Table 2. Figure 8 shows the functional behavior for some specific values of the regularization parameters.

Table 2. *Results of the iterative method for Example II*

γ_0	$\tilde{m}_1$	γ_{GCV_1}	$\tilde{m}_2$	γ_{GCV_2}	$\tilde{m}_3$	γ	$\tilde{m}$
1.0 E 6	1.242	2.5E 7	1.250	3.1E5	1.252	3.1E5	1.252
1.0E 11	1.278	5.6E 7	1.254	3.5E5	1.252	3.1E5	1.252

Thus, the solution obtained is $\tilde{m}$ =1.252, γ=3.1E5 and the estimated PSD as in Figure 6 in solid line.

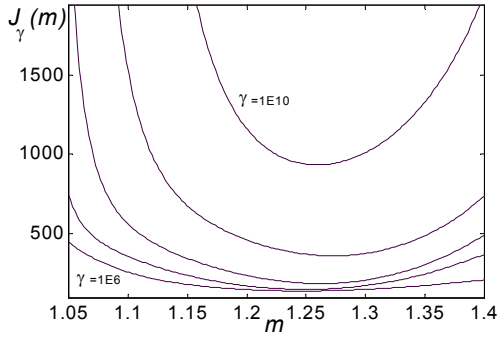


FIGURE 8. *Evolution of the functional for Example II*

CONCLUSIONS

An inversion technique that retrieves the particle size distribution and the relative refractive index of non-absorbing spherical particles from simulated measurements of elastic light scattering has been developed. The general formulation of the problem takes the form of a non-linear inverse problem since the Fredholm equation representing the light scattered has an unknown parameter in its kernel.

The solution is obtained by means of an iterative procedure that improves the values of the refractive index and the regularization parameter, based on the generalized cross validation technique. Attempts to use a non-iterative method based on Neubauer's approach [17] were less successful.

Because of the ill-posed nature of the inverse light scattering problem, *a-priori* information regarding the PSD was used as in Tikhonov regularization. No *a-priori* information about the refractive index was necessary. High accuracy of the resulting refractive index for scattering data with typical error level was observed.

The methodology proposed seems to be more general and simpler to be implemented than those presented in Refs [3] and [10] . We consider that it is not convenient to include the non-negativity restrictions in the optimization process as Ref. [3] does. If this would be done, the non-linear formulation changes and the estimation of the refractive index could not be achieved with the same accuracy. For cases were the measurement error is higher, non-negativity restriction could be added for a second optimization problem considered after determining the value of the refractive index.

APPENDIX A

As *f* is a function the derivative must be interpreted as a functional derivative, for example the Frèchet derivative. For the sake of completeness, we recall the definition.

Let X and Y be Banach spaces, and D an open set in X . The mapping $F:D\subset X\rightarrow Y$ is Frèchet differentiable at x εD if there is a linear operator T from X to Y such that

$$\lim_{h\to 0}\frac{\|F(x+h)-F(x)-Th\|}{\|h\|}=0$$

APPENDIX B

Generalized Cross Validation (GCV)

One of the most popular methods for the determination of the regularization parameter is the GCV due to Golub et. al. [16].

For the regularized linear inverse problem in discrete form given by

$$\underset{\mathbf{f}}{Min}\quad \|\mathbf{Af}-\mathbf{g}_\varepsilon\|^2+\gamma\mathbf{f}^{\mathbf{T}}\mathbf{Hf}\,, \qquad \text{(A-1)}$$

the GCV criterion for the selection of the regularization parameter γ is to minimize the function $V(\gamma)$

$$\underset{\gamma}{Min}\,V(\gamma)=n\frac{\|\mathbf{Af}-\mathbf{g}\|^2}{\left(trace\left(\mathbf{I}-\tilde{\mathbf{A}}(\gamma)\right)\right)^2}=n\frac{\sum_{i=1}^{n}\left(\frac{\gamma}{\lambda_i+\gamma}\right)^2 z_i^2}{\left(\sum_{i=1}^{p}\frac{\gamma}{\lambda_i+\gamma}+n-p\right)^2} \qquad \text{(A-2)}$$

where

$[z_1 z_2 ... z_n]^T = \mathbf{U}^{\mathrm{T}}\mathbf{g}$, λ_i (i=1,..., p) are the eigenvalues of $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, $\lambda_i=0$, $i>p$. $\mathbf{U}$ is defined by the singular value decomposition (SVD) of $\mathbf{X}\rightarrow\mathbf{X}=\mathbf{UDV}^{\mathrm{T}}$ where $\mathbf{X}=\mathbf{AK}^{-1}$, $\mathbf{H}=\mathbf{K}^{\mathbf{T}}\mathbf{K}$, and $\tilde{\mathbf{A}}(\gamma)=\mathbf{A}\left(\mathbf{A}^{\mathbf{T}}\mathbf{A}+\gamma\mathbf{K}^{\mathbf{T}}\mathbf{K}\right)^{-1}\mathbf{A}^{\mathbf{T}}$.

Neubauer's Method

For the regularized linear inverse problem in discrete form given by

$$\underset{\mathbf{f}}{Min}\quad \|\mathbf{A}(m)\mathbf{f}-\mathbf{g}_\varepsilon\|^2+\gamma\mathbf{f}^{\mathbf{T}}\mathbf{If}\,,$$

where $\|\mathbf{A}(m)-\mathbf{A}\|\le h$ and $\|\mathbf{g}_\varepsilon-\mathbf{g}\|\le\delta$ defines the existence of errors in the model $(\mathbf{A}(m))$ and the data $(\mathbf{g}_\varepsilon)$, the regularization parameter γ is selected as the value that satisfies ([14]):

$$\mathcal{M}(\gamma,\mathbf{g}_\varepsilon,\mathbf{A}(m))=\gamma^3\left\langle(\mathbf{A}(m)\mathbf{A}(m)^T+\gamma\mathbf{I})^{-3}\mathbf{g}_\varepsilon,\mathbf{g}_\varepsilon\right\rangle=$$

$$=\left(Dh\|\mathbf{f}_\gamma\|+\delta\right)^2 \qquad \text{(A-3)}$$

where D=1.1 and $\mathbf{f}_\gamma=\left(\mathbf{A}(m)^T.\mathbf{A}(m)+\gamma\mathbf{I}\right)^{-1}\mathbf{A}(m)^T\mathbf{g}_\varepsilon$.

Acknowledgements

We want to thank the financial support from CONICET, Universidad Nacional de Mar del Plata and Universidad Nacional de Buenos Aires (Argentina).

References

1. O. Glatter, J. Siebere and H. Schnablegger, A Comparative Study on Different Scattering Techniques and Data Evaluation Methods for Sizing of Colloidal Systems Using Light Scattering, *Part. Part. Syst. Charact.* ***8****, 274,* 1991

2. M. Kerker, *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press, New York, 1969

3. H. Schnablegger and Otto Glatter. Simultaneous Determination of Size Distribution and Refractive Index of Colloidal Particles from Static Light -Scattering Experiments. *J. Of Colloid and Interface Science*, ***158****, 228*, 1993

4. G. E. Eliçabe and G. L. Frontini, Determination of the Particle Size Distribution of Latex Using a Combination of Elastic Light Scattering and Turbidimetry: A Simulation Study, *J. of Colloid and Interface Science* ***181****, 669*, 1996

5.O. Glatter, M. Hofer, C. Jorde and W. Eigner,.Interpretation of Elastic Light-Scattering Data in Real Space, *J. of Colloid and Interface Science* ***105****, 577*,1985

6. P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems. Numerical Aspects of Linear Inversion*, Siam, Philadelphia, (1998)

7. O. Glatter, M. Hofer, C. Jorde and W. Eigner. Optical Sizing of Small Colloidal Particles: an optimized Regularization Technique, *Applied Optics* ***30****, 33*, 1991

8. S. W. Provencher, A constrained regularization method for inverting data represented by linear algebraic or integral equations, *Comput. Phys. Commun,* ***27****, 213,* 1982

9. H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, 1996

10. M. R. Jones, B. P. Curry, M. Q. Brewster and K. H. Leong. Inversion of light-scattering measurements for particle size and optical constants: theoretical study, *Applied Optics,* ***33****, 18*, 1994

11. G. Mie, *Ann. Phys*., **25**,377, 1908

12. H. C. Van de Hulst, *Light Scattering by Small Particles*, Dover Publications, New York, 1981

13. P. W. Barber and S. C. Hill, *Light Scattering by Particles: Computational Methods*, World Scientific Pub, Singapore, 1998

14. C. Bohren and D. Huffman, *Absorption and Scattering of Light by Small Particles*. John Willey & Sons, 1983

15. A. N. Tikhonov and V. Arsenin, *Solution of Ill-posed Problems,* Wiley, New York, 1977

16. G. H. Golub, M. Heath and G. Wahba, Generalized Cross Validation as a Method for choosing a good Parameter, *Technometrics***, 21**, 215, 1979.

17. A. Neubauer, An a posteriori parameter choice for Tikhonov Regularization in the Presence of Modeling Error. *App. Num. Math*., ***4*** *, 507*, 1988.

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

SOME COMMENTS CONCERNING INVERSE PROBLEMS IN PARTICLE TRANSPORT THEORY

Liliane Basso Barichello

Instituto de Matemática
Universidade Federal do Rio Grande do Sul
91509 900 Porto Alegre, RS, Brazil
lbaric@mat.ufrgs.br

ABSTRACT

Inverse problems in different applications of the particle transport theory are briefly mentioned and a short review is presented regarding to approaches based on analytical solutions of the direct problem, manipulations of the equation itself and the so-called "explicit methods", for radiative-transfer problems. In particular, numerical results obtained for boundary conditions, source term, albedo, phase function, reflection coefficients and optical thickness estimation are discussed.

INTRODUCTION

Investigations on inverse problems in transport theory can be dated of many years ago. In a Case's paper [1] the question of getting an expression for the scattering kernel, for example, was already discussed. According to McCormick [2], the research on inverse problems governed by the linear integrodifferential transport equation dates back at least to the late 1960s. Very good reviews of the development in this field can be found in some of the McCormick's works [2, 3, 4, 5], who has made many contributions to the field of the inverse problems in transport theory.

Since the transport equation is the basis for modeling a wide class of problems, the techniques developed can be associated to different applications: radiative transfer, neutral and charged particles, time-dependent and time-independent models, and so on. In this sense, we note, for example, the works of Badruzzaman [6] in nuclear geophysics, more specifically nuclear oil well logging; inverse time-dependent problems related to photon transport in an interstellar cloud [7]; estimation of the optical properties of sea water [8] and several other works related to radiative transfer applications, as listed later on in this paper.

In our opinion, an important issue to be emphasized is the considerable number of explicit expressions and techniques based on manipulations of the transport equation itself or of analytic solutions to direct problems [9], which have been obtained in trying to solve these problems. We can mention, for example, an important result obtained by Siewert [10] who presented a closed-form expression to estimate the scattering-to-total cross-section ratio for a semi-infinite homogeneous medium, considering isotropic scattering, using only external angular flux measurements. Of course, due to the complexity of the general model, the techniques may be associated to approaches based on numerical schemes [11, 12], however, even to models which include multidimensional applications [13]; multigroup equations [14]; polarization effects [15, 16, 17] and different geometries [18], the so-called explicit techniques have been successfully applied.

In this work, without seeking to do a complete review, but in order to give instead a general idea of the use of semi-analytical techniques for solving inverse problems in transport theory, we restrict ourselves to the classical radiative transfer problem and we review some interesting results obtained in reconstructing boundary conditions, source terms, scattering coefficients and optical properties of the medium.

FORMULATION

We consider the equation of transfer [19] for the radiation intensity $I(\tau, \mu)$, written as

$$\mu\frac{\partial}{\partial\tau}I(\tau,\mu)+I(\tau,\mu)=\frac{\varpi}{2}\sum_{l=0}^{L}\beta_l P_l(\mu)\times\int_{-1}^{1}P_l(\mu')I(\tau,\mu')d\mu'+S(\tau)\quad(1)$$

where $\tau\in(0,\tau_0)$ is the (dimensionless) optical variable, $\mu\in[-1,1]$ is the cosine of the polar angle (as measured from the *positive* τ axis) used to describe the direction of propagation of the radiation and ϖ is the albedo for single scattering. In addition, the β_l are the coefficients in a Legendre polynomial expansion of the scattering law. For direct problems in radiative transfer, we normally supplement Eq. (1) with boundary conditions of the form

$$I(0,\mu)=F_1(\mu)+\rho_1^s I(0,-\mu)+2\rho_1^d\int_0^1 I(0,-\mu')\mu'd\mu'\quad(2a)$$

and

$$I(\tau_0,-\mu)=F_2(\mu)+\rho_2^s I(\tau_0,\mu)+2\rho_2^d\int_0^1 I(\tau_0,\mu')\mu'd\mu'\quad(2b)$$

for $\mu\in(0,1]$. Here $F_1(\mu)$ and $F_2(\mu)$ are considered given, $S(\tau)$ is an inhomogeneous source term and ρ_β^s and ρ_β^d, for $\beta=1$ and 2, are coefficients for specular and diffuse reflection.

An Inverse Boundary Condition Problem

From the field of radiation therapy, where one may want to predict the angular shape and strength of an exposing radiation beam incident on the surface of the body that will give rise to a desired internal dose, can come a motivation for solving an inverse boundary-condition problem. For the problem considered, we suppose that the radiation density

$$\Phi(\tau)=\int_{-1}^{1}I(\tau,\mu)d\mu\quad(3)$$

is known at certain positions $\{\tau_i\}$ within the medium, and we then seek to determine what functions $F_1(\mu)$ and $F_2(\mu)$ for $\mu\in(0,1]$ can induce such a radiation density. Based on specific features of the method, the homogeneous version of this problem ($S(\tau)=0$) was first solved by the LTS_N method [20] and later a more general approach was proposed [21], based on spectral methods, where the desired boundary data were computed from

$$F_1(\mu)=\sum_{k=0}^{K_1}a_k H_k(\mu)\quad(4a)$$

and

$$F_2(\mu)=\sum_{k=0}^{K_2}b_k H_k(\mu)\quad(4b)$$

for $\mu\in(0,1]$. Here $\{H_k(\mu)\}$ denotes a set of basis functions that has to be specified. The constants $\{a_k\}$ and $\{b_k\}$ are to be found simply by constraining the approximate form of the radiation density to match the desired radiation density $\Phi(\tau)$ at K_1+K_2+2 positions given by $\{\tau_i\}$:

$$\sum_{k=0}^{K_1}a_k\Xi_k^1(\tau_i)+\sum_{k=0}^{K_2}b_k\Xi_k^2(\tau_i)=\Phi(\tau_i).\quad(5)$$

In Eq. (5), for $i=1,2,\ldots,K_1+K_2+2$ and $k=0,1,2,\ldots,K_\alpha$, $\alpha=1$ and 2,

$$\Xi_k^\alpha(\tau)=\int_{-1}^{1}\Upsilon_k^\alpha(\tau,\mu)d\mu,\quad(6)$$

where $\Upsilon_k^\alpha(\tau,\mu)$ satisfy an the homogeneous version of Eq. (1) with boundary conditions such that

$$\Upsilon_k^\alpha(0,\mu)=\delta_{1,\alpha}H_k(\mu)+\rho_1^s\Upsilon_k^\alpha(0,-\mu)+2\rho_1^d\int_0^1\Upsilon_k^\alpha(0,-\mu')\mu'd\mu'\quad(7a)$$

and

$$\Upsilon_k^\alpha(\tau_0,-\mu)=\delta_{2,\alpha}H_k(\mu)+\rho_2^s\Upsilon_k^\alpha(\tau_0,\mu)+2\rho_2^d\int_0^1\Upsilon_k^\alpha(\tau_0,\mu')\mu'd\mu'\quad(7b)$$

for $\mu\in(0,1]$. We note that, even though methods for solving direct problems are very important in this kind of approach (since we may need to solve K_1+K_2+2 times the direct problem), the solution given by Eqs. (4) is independent of the method used to solve the basic problems which define the functions $\Upsilon_k^\alpha(\tau,\mu)$. In Ref. [21] the spherical harmonics method

was used to get numerical results for test cases defined by variations of the function

$$F_1(\mu) = \sqrt{(1-\mu^2)} + \frac{1}{1+\mu^2} + \mu \log(\mu) e^{-\mu} \tag{8}$$

and $F_2(\mu) = 0$ along with $\varpi = 0.99$, $\tau_0 = 10.0$, $L = 99$, $\rho_1^s = 0.1$, $\rho_1^d = 0.2$, $\rho_2^s = 0.3$, and $\rho_2^d = 0.4$. In addition

$$H_k(\mu) = P_{2k}(\mu) \tag{9}$$

for $k = 0, 1, \ldots, K_1$ and various values of K_1. In Table 1, some results obtained [21] with a $N = 199$ approximation of the spherical harmonics method, are showed.

Table 1: Direct and Inverse Results for Boundary Data

ξ	$F_1(\xi)$	$\widehat{F_1}(\xi)$
0.00	2.00	1.88
0.10	1.78	1.80
0.20	1.68	1.67
0.30	1.60	1.60
0.40	1.53	1.54
0.50	1.46	1.45
0.60	1.37	1.37
0.70	1.26	1.26
0.80	1.13	1.13
0.90	0.95	0.94
1.00	0.50	0.47

The complete analysis of the results of this problem is presented in Ref. [21], where it can be seen that the radiation density estimates are much better than the boundary conditions estimates. From there, we can also note that more accurate results for the boundary conditions were obtained for simpler expressions, based on Eq. (8), of the incoming intensity. An important tool used to check the ill-conditioning was the condition number estimates of the linear system involved, given by Eq. (5). Based on that, either the number or the position of the measurements were changed. In addition, for some test cases, non-physical results were obtained, pointing out the important issue of existence and uniqueness analysis of solutions.

Similar approach was used in order to estimate boundary conditions in hydrologic optics [22].

An Inverse Source Problem

If the boundary functions are now considered known, one can seek to determine the inhomogeneous source term $S(\tau)$, in Eq. (1), given that we know the results $I(0, -\mu)$ and $I(\tau_0, \mu)$, for $\mu \in (0, 1]$. In this case, Siewert [23] chose two sets – (i) Hermite cubic splines and (ii) the Legendre polynomials $P_k(2\xi - 1)$ – of basis functions $H_k(\xi)$, defined on the interval $[0, 1]$, to write

$$S(\tau) = \sum_{k=0}^{K_1} a_k H_k(\tau). \tag{10}$$

In solving this problem, Siewert [23] was able to deduce the unknown source term for three considered test problems

$$S_1(\tau) = [1 + (\tau/\tau_0)^2]^4, \tag{11}$$

$$S_2(\tau) = 1 + \sin(\pi\tau/\tau_0) \tag{12}$$

and

$$S_3(\tau) = 2 + \sin(2\pi\tau/\tau_0) \tag{13}$$

where it was allowed radiation incident on both sides of the layer and both specular and diffuse reflection on each of the surfaces. Anisotropic scattering ($L = 24$) was also considered. It was reported [23] that for those test cases, it was possible to reduce the accuracy of the exiting (known) distributions to just two figures and still obtain meaningful results for the desired source term. In addition, good results were obtained for the source term by using only 10 points on each of the two boundaries where the exiting distributions were known. Finally, it was also possible to deduce the source term from the exiting intensity known only in one of the two surfaces.

We note that, in regard to inverse problems in radiative transfer, to the best of our knowledge, the inverse source problem is the one to which some existence and uniqueness results are available [24], even though it is regarding to the case of a semi-infinite plane-parallel layer. In fact, it was showed that the problem is ill-posed and under certain constraints (bounded by an exponential function) related to the intensity and the source the problem has at most one solution, which involves the analytic continuation of exiting intensity [24].

The optical thickness

We consider now the problem [25] defined by the homogeneous version of Eq. (1), subject to boundary conditions as Eqs. (2) for the particular case where the coefficients $\rho_\beta^s = 0$ and $\rho_\beta^d = 0$, for β= 1 and 2, and we assume we know ϖ, the order L, the coefficients β_l that define the scattering law and (from experimental data) the exiting intensities

$$I(0, -\mu) = G_1(\mu) \tag{14a}$$

and

$$I(\tau_0, \mu) = G_2(\mu) \tag{14b}$$

for $\mu \in (0, 1]$.

Different approaches can be referenced in regard to estimation of optical thickness [26, 27]. However, recently, based on the method of elementary solutions [28], an explicit expression for the optical thickness τ_0 was found by Siewert [25], for this anisotropic case, in terms of the incoming and exiting intensities and the discrete eigenvalues ν_j, written as

$$\tau_0 = \nu_j \ln\left\{M(0, \nu_j)/M(\tau_0, \nu_j)\right\} \tag{15}$$

where, for $j = 0, 1, \ldots, K_1$

$$M(0, \pm\nu_j) = \int_0^1 \phi(\pm\nu_j, \mu) F_1(\mu)\mu \mathrm{d}\mu - \int_0^1 \phi(\mp\nu_j, \mu) G_1(\mu)\mu \mathrm{d}\mu \tag{16}$$

and

$$M(\tau_0, \pm\nu_j) = \int_0^1 \phi(\pm\nu_j, \mu) G_2(\mu)\mu \mathrm{d}\mu - \int_0^1 \phi(\mp\nu_j, \mu) F_2(\mu)\mu \mathrm{d}\mu. \tag{17}$$

Here

$$\phi(\pm\nu_j, \mu) = \frac{\varpi\nu_j}{2} \sum_{l=0}^{L} \beta_l P_l(\mu) g_l(\pm\nu_j) \frac{1}{\nu_j \mp \mu}, \tag{18}$$

and $g_l(\xi)$ denotes the Chandrasekhar polynomials [19] defined by the three-term recursion formula

$$(2l + 1 - \varpi\beta_l)\xi g_l(\xi) = (l + 1) g_{l+1}(\xi) + l g_{l-1}(\xi) \tag{19}$$

and the starting values

$$g_0(\xi) = 1 \quad \text{and} \quad g_1(\xi) = (1 - \varpi)\xi. \tag{20a,b}$$

In addition, the discrete eigenvalues can be computed as zeros of the dispersion function [25]

$$\Lambda(z) = 1 + \int_{-1}^{1} \Psi(\xi) \frac{\mathrm{d}\xi}{\xi - z}, \quad z \notin [-1, 1], \tag{21}$$

where the characteristic function is

$$\Psi(\xi) = \frac{\varpi}{2} \sum_{l=0}^{L} \beta_l P_l(\xi) g_l(\xi). \tag{22}$$

According to Ref. [25], without adding noise to the solution of the direct problem, by using Eq. (15), it was possible to get numerical results for the optical thickness with five figures of accuracy, what can be considered a good test for the correctness of derived expressions. Of course, by adding 5% of random noise to the solution of the direct problems this accuracy is lost, but still meaningful results are obtained. In addition, simpler results than the ones given by Eqs. (16) and (17), were found for the conservative case $\varpi = 1$.

The scattering coefficients

Following a previous work [29] and using simple iterative procedures along with some results derived by McCormick [30], such that, for $\varpi \in (0, 1)$,

$$4 \int_0^1 [G_2(\mu)F_2(\mu) - G_1(\mu)F_1(\mu)]\mathrm{d}\mu = \varpi \sum_{l=0}^{L} (-1)^l \beta_l [I_l^2(\tau_0) - I_l^2(0)] \tag{23}$$

and

$$4 \int_0^1 [G_2(\mu)F_2(\mu) - G_1(\mu)F_1(\mu)]\mu^2 \mathrm{d}\mu = \varpi \sum_{l=0}^{L} (-1)^l (2l + 1)(\beta_l/h_l)[J_l^2(\tau_0) - J_l^2(0)], \tag{24}$$

where

$$I_l(\tau) = \int_{-1}^{1} I(\tau, \mu) P_l(\mu) \mathrm{d}\mu, \tag{25a}$$

$$J_l(\tau) = \int_{-1}^{1} I(\tau,\mu) P_l(\mu) \mu \mathrm{d}\mu, \tag{25b}$$

and

$$h_l = 2l + 1 - \varpi\beta_l, \tag{26}$$

Siewert [25] worked on the determination of the coefficients of the scattering law. Two cases were considered: (i) the binomial scattering law

$$p(\cos\Theta) = \frac{K+1}{2^K}\big(1+\cos\Theta\big)^K \tag{27}$$

which can be rewritten in terms of Legendre polynomials as

$$p(\cos\Theta) = \sum_{l=0}^{L} \beta_l P_l(\cos\Theta), \tag{28}$$

and where the β coefficients can be computed as [31]

$$\beta_l = \Big(\frac{2l+1}{2l-1}\Big)\Big(\frac{K+1-l}{K+1+l}\Big)\beta_{l-1}, \tag{29}$$

for $l = 1, 2, ...,$ with $\beta_0 = 1$; (ii) the Henyey-Greenstein law

$$p(\cos\Theta) = \sum_{l=0}^{L} (2l+1) g^l P_l(\cos\Theta). \tag{30}$$

It should be noted that while Eqs. (23) and (24) require the intensity on the two boundaries of the medium to be known, the optical thickness τ_0 is not needed. Again, numerical results were obtained with five figures of accuracy, without adding noise to the solution of the direct problem. By adding 5% of random noise to the solution of the direct problem this accuracy is lost, but still meaningful results are obtained. In these problems, the approach used for solving the direct problem is a version of the discrete-ordinates method [32, 33]. Finally, the conservative case $\varpi = 1$ was also treated.

We note that simultaneous estimation of radiation phase function and albedo has been also investigated, for example, for Chalhoub and Campos Velho [34].

The albedo and the reflection coefficients

Looking back again to Eq. (1), we consider now the particular case of isotropic scattering

$$\mu\frac{\partial}{\partial\tau} I(\tau,\mu) + I(\tau,\mu) = \frac{\varpi}{2}\int_{-1}^{1} I(\tau,\mu')\mathrm{d}\mu', \tag{31}$$

for $\tau \in (0, \tau_0)$ and $\mu \in [-1, 1]$, and boundary conditions written as

$$I(0,\mu) = F_1(\mu) + 2\rho_1^d \int_0^1 I(0,-\mu')\mu'\mathrm{d}\mu' \tag{32a}$$

and

$$I(\tau_0,-\mu) = 2\rho_2^d \int_0^1 I(\tau_0,\mu')\mu'\mathrm{d}\mu' \tag{32b}$$

for $\mu \in (0, 1]$. If we then assume we know the function $F_1(\mu)$ and the surface results

$$R(\mu) = (1-\rho_1^d) I(0,\mu) \tag{33a}$$

and

$$T(\mu) = (1-\rho_2^d) I(\tau_0,\mu), \tag{33b}$$

for $\mu \in (0, 1]$, we seek to determine the single-scattering albedo ϖ and the two coefficients for diffuse reflection ρ_1^d and ρ_2^d [35].

This problem has been also solved [35], based on previous results [29], in a collection of exact expressions and the use of Newton's method. Although numerical results were discussed [35] for half-space and finite-media applications, equal and unequal coefficients, in particular we chose to comment here, the case of equal reflection coefficients, where the solution involves to solve a cubic (algebraic) equation for ρ, *viz.*

$$(c+\widehat{d}\rho)(\widehat{\alpha}\rho^2 + \widehat{\beta}\rho + \widehat{\gamma}) = (\delta + \widehat{\epsilon}\rho) \times [(a+b\rho)^2 - (\widehat{a}+\widehat{b}\rho)^2] \tag{34}$$

and then compute the albedo from either of equations

$$\varpi[(a+b\rho)^2 - (\widehat{a}+\widehat{b}\rho)^2] = c + \widehat{d}\rho \tag{35a}$$

and

$$\varpi(\widehat{\alpha}\rho^2 + \widehat{\beta}\rho + \widehat{\gamma}) = \delta + \widehat{\epsilon}\rho, \tag{35b}$$

where the constant coefficients presented in Eqs. (34) and (35) are clearly defined in Ref. [35].

The determination of the reflection coefficients is of interest to the knowledge of the radiative properties of a material and it could be cited as a boundary condition inverse problem, in the sense of "parameter estimates"; differently of the "function estimates" proposed

in Eqs. (4). Same kind of application is also of interest in rarefied gas problems, when one would like to determinate the accommodation coefficients, which are used to describe the gas-surface interaction [36].

CONCLUDING COMMENTS

Some inverse problems in radiative transfer were commented here, where, in general, the optical thickness, the albedo, the coefficients of the scattering law, the reflection coefficients, the boundary conditions and the source term are to be determined. We chose to comment on solutions based on either explicit formulas or explicit methods, where the solution of the inverse problem is constructed in terms of the direct problem or manipulations of the equation itself. We intended to emphasize some aspects: (i) the fundamental point of having good solutions for direct problems – easy to implement, accurate and fast solutions. In this regard we could expect that some of the results here showed, may be improved once methods more recently developed for dealing with the direct problems [32] could be used in a more efficient way. For the boundary condition problem, for example, a new analytical version of the discrete-ordinates method [32], based on "half-range" quadrature schemes, seems more appropriate to deal with the known discontinuity present in the boundary conditions of the problem. (ii) explicit results can always be used as starting points when it comes to the solution of more complex models, mostly based on iterative procedures; (iii) as mentioned in the beginning of this paper, since different applications in transport theory are based on the same model, we believe the results commented here can be used in some other fields of the general area of transport theory.

Acknowledgment – This work is supported in part by FAPERGS and CNPq of Brasil.

REFERENCES

1. K. M. Case. Inverse problem in transport theory. *Phys. of Fluids,* **16**, 10, 1607 (1973).

2. N. J. McCormick. Inverse radiative transfer problems:a review. *Nucl. Sci. Eng.,* **112**, 185 (1992).

3. N. J. McCormick. Recent developments in inverse scattering transport methods. *Trans. Theory Stat. Phys.,* **13**, 15 (1984).

4. N. J. McCormick. Methods for solving inverse problems for radiation transport-anupdate. *Trans. Theory Stat. Phys.,* **15**, 759 (1986).

5. N. J. McCormick. Inverse problems:methods and applications. *TEMA–Proc. XXIII Nac. Meeting on Appl. and Comp. Math., Brasil,* **2**, 1 (2001).

6. A. Badruzzaman. Computational methods in nuclear geophysics. *Prog. Nucl. Energy,* **25**, 265 (1991).

7. M. Lisi and S. Totaro. Inverse problems related to photon transport in an interstellar cloud. *Proceedings of the 17th International Conference on Transport Theory,* **London, UK**, 64 (2001).

8. E. S. Chaloub and H. F. C. Velho. Estimation of the optical properties of sea water from measurements of exit radiance. *J. Quant. Spectrosc. Radiat. Transfer,* **in press**, (2001).

9. C. E. Siewert. A new approach to the inverse problem. *J. Math. Phys.,* **19**, 2619 (1978).

10. C. E. Siewert. On a possible experiment to evaluate the validity of the one-speed or constant cross section model of the neutron-transport equation. *J. Math. Phys.,* **19**, 1587 (1978).

11. A. T. Kauati, A. J. S. Neto and N. C. Roberty. A source-detector methodology for the construction and solution of the one-dimensional inverse transport equation. *Inv. Probl. Eng.,* **9**, 45 (2001).

12. R. F. C. Monteiro, N. C. Roberty and A. J. S. Neto. Absorption coefficient estimation in heterogeneous participating media using a domain partition consistent with divergente beams. *Inv. Probl. Eng.,* **in press**, (2001).

13. C. E. Siewert and W. L. Dunn. On inverse problems for plane-parallel media with nonuniform surface ilumination. *J. Math. Phys.,* **23**, 1376 (1982).

14. C. E. Siewert and M. N. Ozisik. On the inverse problem for multigroup neutron transport theory. *Nucl. Sci. Eng.,* **63**, 95 (1977).

15. C. E. Siewert. Determination of the single-scattering albedo from polarization measurements of a rayleigh atmosphere. *Astrop. Spac. Sci.,* **60**, 237 (1979).

16. C. E. Siewert and J. R. Maiorino. The inverse problem for a finite rayleigh-scattering atmosphere. *J. Appl. Math. Phys. (ZAMP),* **31**, 767 (1980).

17. C. E. Siewert. Solutions to an inverse problem in radiative transfer with polarization - I. *J. Quant. Spectrosc. Radiat. Transfer,* **30**, 523 (1983).

18. C. E. Siewert. A radiative-transfer inverse-source problem for a sphere. *J. Quant. Spectrosc. Radiat. Transfer,* **52**, 157 (1994).

19. S. Chandraselhar. *Radiative transfer.* Oxford University Press, London, 1950.

20. L. B. Barichello and M. T. Vilhena. An inverse problem in neutron and radiation transport (in portuguese). *Proceedings of the IX ENFIR, Brasil,* **1**, 22 (1993).

21. L. B. Barichello, R. D. M. Garcia and C. E. Siewert. On inverse boundary-conditions problems in radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer,* **57**, 405 (1997).

22. H. F. C. Velho, M. R. Retamoso and M. T. Vilhena. Estimation of boundary condition in hydrologic optics. *Appl. Num. Math.,* **in press**, (2001).

23. C. E. Siewert. An inverse source problem in radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer,* **50**, 603 (1993).

24. E. W. Larsen. The inverse source problem in radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer,* **15**, 1 (1975).

25. C. E. Siewert. Inverse solutions to radiative-transfer problems based on the binomial or the Henyey-Greenstein scattering law. *J. Quant. Spectrosc. Radiat. Transfer,* **72**, 827 (2002).

26. R. Sanchez, N. J. McCormick and H. C. Yi. Iterative inverse radiative transfer method to estimate optical thickness and surface albedo. *Trans. Theory Stat. Phys.,* **19**, 357 (1990).

27. H. C. Yi, N. J. McCormick and R. Sanchez. Cloud optical thickness estimation from irradiance measurements. *J. Atmos. Sci.,* **47**, 2567 (1990).

28. K. M. Case. Elementary solutions of the transport equation and their applications. *Annals Phys. (NY),* **9**, 1 (1960).

29. C. E. Siewert. On the inverse problem for a three-term phase function. *J. Quant. Spectrosc. Radiat. Transfer,* **22**, 441 (1979).

30. N. J. McCormick. Transport scattering coefficients from reflection and transmission measurements. *J. Math. Phys.,* **20**, 1504 (1979).

31. N. J. McCormick and R. Sanchez. Inverse problem transport calculations for anisotropic sacttering coefficients. *J. Math. Phys.,* **22**, 199 (1981).

32. L. B. Barichello and C. E. Siewert. A discrete-ordinates solution for a non-grey model with complete frequency redistribution. *J. Quant. Spectrosc. Radiat. Transfer,* **62**, 665 (1999).

33. C. E. Siewert. A concise and accurate solution to chandrasekhar's basic problem in radiative transfer. *J. Quant. Spectrosc. Radiat. Transfer,* **64**, 109 (2000).

34. E. S. Chaloub and H. F. C. Velho. Simultaneous estimation of radiation phase function and albedo in natural waters. *J. Quant. Spectrosc. Radiat. Transfer,* **69**, 137 (2001).

35. C. E. Siewert. Inverse solutions to radiative-transfer problems with partially transparent boundaries and diffuse reflection. *J. Quant. Spectrosc. Radiat. Transfer,* **72**, 299 (2002).

36. M. M. R. Williams. A review of the rarefied gas dynamics theory associated with some classical problems in flow and heat transfer. *Z. Angew. Math. Phys.,* **52**, 500 (2001).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

THE SOURCE-DETECTOR METHODOLOGY FOR APPLICATIONS WITH ANISOTROPIC SCATTERING

Adriana T. Kauati
Biomedical Engineering Program –PEB/COPPE
Universidade Federal do Rio de Janeiro,
P.O. Box 68510, 21945-970, Rio de Janeiro, RJ, Brazil
kauati@peb.ufrj.br

Antônio J. Silva Neto[1,2]
Nilson C. Roberty[2]
[1]*Department of Mechanical Engineering and Energy, Instituto Politécnico, IPRJ,*
Universidade do Estado do Rio de Janeiro, UERJ,
P.O. Box 97282, 28601-970, Nova Friburgo, RJ, Brazil
ajsneto@iprj.uerj.br

[2]*Nuclear Engineering Program – PEN/COPPE*
Universidade Federal do Rio de Janeiro,
P.O. Box 68509, 21945-970, Rio de Janeiro, RJ, Brazil
ajsneto@lmn.con.ufrj.br, nilson@lmn.con.ufrj.br

ABSTRACT

In the present work the source-detector methodology is used for the estimation of the scattering and total extinction (absorption + scattering) coefficients as well as the first coefficient of anisotropic scattering.

The source-detector methodology is an explicit formulation for inverse radiative transfer problems that consists on formulating and solving a nonlinear system of equations derived from the direct problem and an adjoint problem.

The mathematical formulation and test case results for homogeneous media are presented.

NOMENCLATURE

a	surface in which the source is located ($a = 0$ or $a = L$)
A	amplitude of the strength of the external collimated radiation source
b	surface for which the boundary condition is written ($b = 0$ or $b = L$).
$g^{b,k_{a'}}(m)$	measurement obtained by the detector
q*	reference value for $S(x,\mu)$
$S(x,\mu)$	internally distributed source
x	spatial coordinate

Greek letters

f	Radiation Intensity
μ	direction cosine of the radiation beam with the positive x axis
$\sigma_s(x,\mu',\mu)$	scattering coefficient
$\sigma_s^R(x, \mu', \mu)$	reference value for $\sigma_s(x)$
$\sigma_t(x)$	total extinction coefficient (absorption + scattering)
$\sigma_t^R(x)$	reference value for $\sigma_t(x)$

INTRODUCTION

Several outstanding examples of the application of inverse radiative transfer problems can be found in different areas of human activity such as biomedical engineering [1-3] and environmental engineering [4-7].

Most of the work on inverse radiative transfer problems published so far involves the estimation of radiative properties [8-14], internal sources [15-17], and boundary conditions or properties [18-21], and they are based on experimental data of the radiation intensity, or some related quantity such as radiative heat flux. In some cases detectors can even be placed inside the medium under analysis [22,23].

A lot of effort has been devoted to the estimation of anisotropic scattering phase function [24-28], with a particular interest on a

special case in which the phase function is based on one parameter, the asymmetry factor [29-34].

In the present work we use the source-detector methodology for the estimation of the extinction and scattering coefficients.

The source-detector methodology consists on a explicit formulation for the inverse radiative transfer problem in participating media. For each pair source-detector a non-linear equation is derived using a convolution of the direct problem (source problem) with the solution of the adjoint problem (detector problem). In both problems the modeling of scattering, absorption and emission phenomena is done with the linearized Boltzmann equation, but in the latter reference values are available for the unknowns that we want to determine, being therefore included in the formulation. This approach results in a system of non-linear equations with the unknowns explicitly represented. This system of equations is called Inverse Transport Equation (ITE).

In a previous work [35] the particular case of isotropic scattering was considered. Here the formulation is extended to the case of anisotropic scattering. Test case results are presented for the estimation of one coefficient of the expansion of the phase function in Legendre polynomials.

THE SOURCE-DETECTOR METHODOLOGY

The source-detector methodology for the analysis of inverse radiative transfer problems involves the formulation and solution of the following problems [35-38]: (I) source problem; (ii) detector and auxiliary problems, and (iii) inverse transport equation (ITE).

The Source Problem

Consider a plane-parallel, gray, anisotropic scattering slab of thickness L, with transparent boundaries subjected to an external collimated radiation source that may be positioned in different locations around the medium. These locations are represented by the index $k_o=1, 2, \ldots I_o$ for the boundary at $x=0$, and by $k_L=1, 2, \ldots I_L$ for the boundary at $x=L$. In a compact form we write $k_a=1, 2, \ldots I_a$, where $a=0$ or $a=L$.

The mathematical formulation of the one-dimensional steady-state radiative transfer problem in a participating medium, i.e. a medium in which absorption, emission, and scattering takes place, is given by the linearized Boltzmann equation, that for the case of azimuthal symmetry is given by [39]

$$\mu \frac{\partial f^{a,k_a}(x,\mu)}{\partial x} + \sigma_t(x)\, f^{a,k_a}(x,\mu) - \frac{1}{2}\int_{-1}^{1} \sigma_s(x,\mu',\mu)\, f^{a,k_a}(x,\mu')\, d\mu' = S(x,\mu)$$

$$0 < x < L, \quad -1 \le \mu \le 1 \tag{1a}$$

$$f^{a,k_a}(b,\mu) = A\,\delta_{ab}\, f^{b,k_a}(\mu) \tag{1b}$$

where $f^{a,k_a}(x,\mu)$ is the radiation intensity, x is the spatial coordinate, μ is the direction cosine of the radiation beam with the positive x axis, $\sigma_t(x)$ is the total extinction coefficient (absorption + scattering), $\sigma_s(x,\mu',\mu)$ is the scattering coefficient, $S(x,\mu)$ is an internally distributed source, A is the amplitude of the strength of the external collimated radiation source and $f^{b,k_a}(\mu)$ represents its dependence with the polar angle. The indices a and b represent, respectively, the surface in which the source is located ($a = 0$ or $a = L$) and the surface for which the boundary condition is being written ($b = 0$ or $b = L$).

In the present work we have used the discrete ordinates method for the solution of the source problem.

The Detector and Auxiliary Problems

Consider that for a given position of the external radiation source, detectors are placed at positions around the medium represented by the index $k_{0'} = 1, 2, \ldots I_{0'}$ for the boundary at $x = 0$ and $k_{L'} = 1, 2, \ldots I_{L'}$ for the boundary at $x = L$. In a compact form we write $k_{a'} = 1, 2, \ldots, I_{a'}$, where $a' = 0$ or $a' = L$.

Now an adjoint problem is formulated by reversing the direction of radiation transfer, i.e. by replacing μ by $-\mu$ in problem (1),

$$-\mu \frac{\partial f^{*a',k_a'}(x,-\mu)}{\partial x} + \sigma_t^R(x)\, f^{*a',k_a'}(x,-\mu) - \frac{1}{2}\int_{-1}^{1} \sigma_s^R(x,\mu',\mu)\, f^{*a',k_a'}(x,-\mu')\, d\mu' = q^*(x,\mu)$$

$$0 < x < L, \quad -1 \le \mu \le 1 \tag{2a}$$

$$f^{*a'k_a'}(b,\mu) = \delta_{a'b}\, g^{b,k_a'}(\mu) \tag{2b}$$

where $\sigma_t^R(x)$, $\sigma_s^R(x,\mu',\mu)$ and q^* are reference functions for the unknowns $\sigma_t(x)$, $\sigma_s(x)$ and $S(x,\mu)$ respectively.

By imposing the coincidence of the location of the detector, $k_{a'} = 1, 2, \ldots, I_{a'}$, with those for the source, $k_a = 1, 2, \ldots, I_a$, function $g^{b,k_{a'}}(m)$ represents the measurement that would be obtained by the detector for the strength of the source located at that position. In fact, for a given source problem in which the external source is located at position k_a, the calculated values for the intensities of the exit radiation, $f^{a,k_a}(0, m_{k_{0'}})$ with $\mu<0$, and $f^{a,k_a}(L, m_{k_{L'}})$ with $\mu>0$, are used as the values for $g^{b,k_{a'}}(m_{k_{a'}})$ in Eq. (2b).

Problem (2) is the detector problem. By reversing again the direction of radiation transfer, and defining an auxiliary function

$$f^{0}(x,m) = f^{*}(x,-m) \qquad (3)$$

one obtains from the detector problem the following auxiliary problem

$$m\frac{\partial f^{0a',k_{a'}}(x,m)}{\partial x} + s_t^R(x) f^{0a',k_{a'}}(x,m) - \frac{1}{2}\int_{-1}^{1} s_s^R(x,m',m) f^{0a',k_{a'}}(x,m')\,dm' = q^0(x,m)$$

$$0 < x < L, \quad -1 \le m \le 1 \qquad (4a)$$

$$f^{0a',k_{a'}}(b,m) = \delta_{a'b}\delta_{k_{a'}k_a} f^{b,k_a}(a',m) \qquad (4b)$$

Observe that for each source problem there is a set of detector problems, and consequently auxiliary problems. Therefore, the function $f^{0a',k_{a'}}$ has also a dependence on the indices (a, k_a) that represent a particular source problem. Just to keep the presentation simple we have omitted the representation of the dependence of the solution of the detector, and auxiliary problems on a specific source problem.

The formulations of the source and auxiliary problems, given respectively by Eqs (1a) and (4a), are very similar. In the former are used the unknown radiative properties and source, while in the latter reference values for these unknown functions are employed. Therefore, the same computational programs developed for the solution of the source problem can be used for the solution of the auxiliary problem.

The Inverse Transport Equation (ITE)

The first step on the derivation of the inverse transport equation consists on multiplying Eq. (1a), that models the source problem, by the solution of the correspondent adjoint problem, i.e. the detector problem, $f^{*a',k_{a'}}(x,m)$, and then integrating the resulting expression both in the spatial and angular domains, i.e. x=[0,L] and μ=[-1,1], respectively, yielding

$$\int_{-1}^{1}\int_{0}^{L}\left[m\frac{\partial f^{a,k_a}(x,m)}{\partial x} f^{*a',k_{a'}}(x,m) + s_t(x) f^{a,k_a}(x,m) f^{*a',k_{a'}}(x,m)\right]dx\,dm - \frac{1}{2}\int_{-1}^{1}\int_{0}^{L}\left[f^{*a',k_{a'}}(x,m)\int_{-1}^{1} s_s(x,m',m) f^{a,k_a}(x,m')\,dm'\right]dx\,dm$$

$$= \int_{-1}^{1}\int_{0}^{L}\left[S(x,\mu)\phi^{*a',k'_a}(x,\mu)\right]dx\,d\mu \qquad (5)$$

Note that for each pair source detector, represented by the location of the source at the position given by the indices (a, k_a) and by the location of the detector at the position given by the indices $(a', k_{a'})$, there is one equation corresponding to Eq.(5). Therefore, in fact Eq. (5) represents a system of nonlinear equations. Taking into account all possible combinations of source and detector locations there is a total of I_t nonlinear equations where $I_t = (I_0 + I_L)\times(I_{0'} + I_{L'})$.

The nonlinearity of the equations results from the fact that the unknowns s_t, s_s and S appear in Eq. (5) multiplying the solutions of the source and detector problems that themselves depend on the value of the unknowns.

Integrating by parts the first term on the left hand side of Eq. (5), and introducing the formulation of the detector problem given by Eq. (2a) one obtains.

$$\int_{-1}^{1}\int_{0}^{L}\left[s_t(x) - s_t^R(x)\right] f^{a,k_a}(x,m) f^{*a',k_{a'}}(x,m)\,dx\,dm - \frac{1}{2}\int_{-1}^{1}\int_{0}^{L}\int_{-1}^{1}\left[s_s(x,m',m) - s_s^R(x,m',m)\right] f^{a,k_a}(x,m) f^{*a',k_{a'}}(x,m')\,dm'\,dx\,dm - \int_{-1}^{1}\int_{0}^{L}\left[S(x,m) f^{*a',k'_a}(x,m) - q^*(x,m) f^{a,k_a}(x,m)\right]dx\,dm$$

$$\int_{-1}^{1} m\left[f^{*a',k_{a'}}(x,m) f^{a,k_a}(x,m)\right]_0^L dm \qquad (6)$$

Defining the deviations from the reference values

$$\Delta\sigma_t(x) = \sigma_t(x) - \sigma_t^R(x) \qquad (7a)$$

$$\Delta\sigma_s(x,\mu',\mu) = \sigma_s(x,\mu',\mu) - \sigma_s^R(x,\mu',\mu) \qquad (7b)$$

and using the definition of the auxiliary function, Eq. (3), we obtain from Eq. (6),

$$\int_{-1}^{1}\int_{0}^{L} \Delta s_t(x) f^{a,k_a}(x,m) f^{\circ a',k_{a'}}(x,-m)\,dx\,dm$$

$$-\tfrac{1}{2}\int_{-1}^{1}\int_{0}^{L}\int_{-1}^{1} \Delta s_s(x,m',m) f^{a,k_a}(x,m) f^{\circ a',k_{a'}}(x,-m)\,dm'\,dx\,dm$$

$$-\int_{-1}^{1}\int_{0}^{L}\left[S(x,m) f^{\circ a',k_{a'}}(x,-m) - q^{\circ}(x,-m) f^{a,k_a}(x,m)\right] dx\,dm$$

$$= -\int_{-1}^{1} m\left[f^{a,k_a}(x,m) f^{\circ a',k_{a'}}(x,-m)\right]_0^L dm \qquad (8)$$

We call the system of equations represented by Eq. (8) as the Inverse Transport Equation (ITE). Solving this system of equations we obtain estimates for the unknowns. Observe that on the right hand side of Eq. (8) we use the experimental data on the exit radiation intensity.

The formulation described here for the inverse radiative transfer problem of estimating the scattering and total extinction coefficients along with the internal radiation source is an explicit formulation. As one can see from Eq.(8) the unknowns explicitly appear in the formulation of the inverse problem.

Solution of the ITE

In the present work we are interested only on the estimation of the coefficients of total extinction σ_t and scattering σ_s. Therefore, we consider $S(x,\mu)=0$.

Considering a homogeneous medium and a representation of the anisotropic scattering coefficient in terms of Legendre polynomials

$$s_s(m',m) = \sum_{l=0}^{N_l} (2l+1) s_{s_l} P_l(m) P_l(m') \qquad (9)$$

Eq. (8) is written as

$$\int_{-1}^{1}\int_{0}^{L} (s_t - s_t^R) f^{a,k_a}(x,m) f^{\circ a',k_{a'}}(x,-m)\,dx\,dm$$

$$-\tfrac{1}{2}\sum_{l=0}^{N_l}(2l+1)(s_{sl} - s_{sl}^R)\int_{0}^{L}\int_{-1}^{1} P_l(m) f^{\circ a',k_{a'}}(x,-m)$$

$$\times\int_{-1}^{1} P_l(m') f^{a,k_a}(x,m')\,dm'\,dm\,dx$$

$$= -\int_{-1}^{1} m f^{a,k_a}(L,m) f^{\circ a',k_{a'}}(L,-m)\,dm$$

$$+\int_{-1}^{1} m f^{a,k_a}(0,m) f^{\circ a',k_{a'}}(0,-m)\,dm \qquad (10)$$

The system of nonlinear equations (10) can be solved using a Gauss-Newton linearization scheme or a more robust algorithm such as the MART (Multiplicative Algebraic Reconstruction Technique) that was developed for tomographic image reconstruction. More details on the solution of the ITE can be seen in Refs. [35-38].

RESULTS

In the present work we have considered the linear anisotropic scattering phase function, with Eq. (9) being written as

$$s_s(m',m) = s_s(A_1 + A_2 m' m) \qquad (11)$$

with $A_1 = 1$ for the isotropic scattering term. Therefore A_1 is considered known

In Table 1 are presented the estimates for the scattering and total extinction (absorption + scattering) coefficients as well as the estimated values of the anisotropic scattering phase function coefficient A_2.

All computational implementation has been done with Borland C++ Builder and each simulation has taken approximately six hours of CPU time on a IBM compatible personal computer with 200 MHz Pentium processor.

In Table 1 the experimental values correspond to the exact values of the unknown properties that we want to estimate. In all test cases we have as initial guesses

$$s_t^0 = s_t^R, \quad s_s^0 = s_s^R \text{ and } A_2^0 = A_2^R \qquad (12)$$

As reported by Kauati et al. [35] to achieve convergence we have used sequentially cycles of properties estimation with the source–detector methodology. Considering k as the cycle counter, as soon as new estimates are obtained, σ_t^k, σ_s^k and A_2^k, they are used as the reference values for a new cycle of iterations of the source detector methodology, i.e

$$\sigma_t^{R^{k+1}} = \sigma_t^k, \sigma_s^{R^{k+1}} = \sigma_s^k \text{ and } A_2^{R^{k+1}} = A_2^k \quad (13)$$

Table 1 – Test Case Results

Coeff.	Reference	Experimental	Estimate
σ_s	0.9000	0.9000	0.896066
σ_t	1.0000	1.0000	1.000079
A_2	0.48520	0.4000	0.424091
σ_s	0.9000	0.9000	0.898902
σ_t	1.0000	1.0000	1.000006
A_2	0.48520	0.1000	0.105832
σ_s	0.9000	0.7000	0.689847
σ_t	1.0000	1.0000	1.000371
A_2	0.48520	0.1000	0.159318
σ_s	0.7000	0.7000	0.7006618
σ_t	1.0000	1.0000	0.999828
A_2	0.1000	0.48520	0.442454
σ_s	0.5000	0.7000	0.695907
σ_t	1.0000	0.9000	0.900088
A_2	0.48520	0.3000	0.326594
σ_s	0.1000	0.1000	0.100260
σ_t	0.3000	0.5000	0.499996
A_2	0.48520	0.5000	0.491172

The cycles of the source-detector methodology are continued until the values for the estimated parameters do not vary within a tolerance.

Even though we have considered only noiseless data (except for the noise due to the numerical approximation for the solution of the direct problem and round – off errors) it can be observed from Table 1 that deviations can be observed in the estimation of A_2. We have tried to estimate higher order terms of anisotropic scattering but results were poor. A more in-depth investigation is necessary.

CONCLUSIONS

In the present work we demonstrate the feasibility of using the source-detector methodology for simultaneously estimating the scattering and total extinction coefficients along with the first coefficient of anisotropic scattering, when the anisotropic phase function is represented in a series of Legendre polynomials.

In the next step of this research the effects of the level of noise in the experimental data will be investigated.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico, FAPERJ – Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, and CAPES – Comissão de Aperfeiçoamento de Pessoal de Nível Superior.

REFERENCES

1. M. S. Patterson, Noninvasive measurements of tissue optical properties: current status and future prospects, *Comments Mol. Cell. Biophys*, **8**, pp. 387-417 (1995).

2. N. J. McCormick, Inverse photon transport methods for biomedical applications, *Proc. 1st International Conference on Inverse Problems in Engineering: Theory and Practice*, Florida, USA, pp.253-258 (1993).

3. L. Fukshansky, N. Fukshansky-Kazarinova and A. M. Remisowsky, Estimation of optical parameters in a living tissue by solving the inverse problem of the multiflux radiative transfer, *Appl. Optics*, **30**, No.22, pp.3145-3153 (1991).

4. R. A. Leathers, C. S. Roesler and N. J. McCormick, Ocean inherent optical property determination from in-water light field measurements, *Applied Optics*, **38**, pp.5096-5103 (1999).

5. N. J. McCormick, R. Sanchez and H.C. Yi, Cloud optical thickness estimation from ground-level measurements, *Atmosf. Oceanic Opt.*, **10**, pp.702-708 (1997).

6. L. J. Holl and N. J. McCormick, Ocean optical property estimation with the Zaneveld-Wells algorithm, *Applied Optics*, **34**, pp.5433-5441 (1995).

7. H. C. Yi, N. J. McCormick and R. Sanchez, Cloud optical thickness estimation from

irradiance measurements, *Journal of the Atmospheric Sciences*, **47**, pp.2567-2579 (1990).

8. V. P. Nicolau, M. Raynaud and J. F. Sacadura, Spectral radiative properties identification of fiber insulating materials, *Int. J. Heat Mass Transfer*, **37**, pp.311-324 (1994).

9. M. P. Mengüç, S. Manickavasagam and D. A. D'Sa, Determination of radiative properties of pulverized coal particles from experiments, *Fuel*, **73**, pp.613-625 (1994).

10. V. I. Gryn, The uniqueness of the determination of the absorption coefficient in the spherically symmetric case, *Comp. Maths Math. Phys*., **33**, p. 1353-1364 (1993).

11. P. M. Koleniskov, Inverse problems of radiative heat transfer in polydis persed media, *J. Engng. Phys*., **56**, pp.503-509 (1989).

12. K. Kamiuto, M. Sato and M. Iwamoto, Determination of the radiative properties of high-porosity materials by use of the emerging-intensity fitting method, *J. Quant. Spectrosc. Radiat. Transfer*, **42**, pp. 477-482 (1989).

13. C. – H. Ho and M. N. Özisik, An inverse radiation problem, *Int J. Heat Mass Tansfer*, **32**, pp.335-341 (1989).

14. F. G. Bakirov, G. N. Zverev, R. S. Kashapov and Z. G. Shaikhutdinov, Z. G., Solution of inverse problems of radiative transport by soot particles of complex shapes, *J. Engng. Phys.*, **49**, pp.921-925 (1986).

15. H. Y. Li, An inverse source problem in radiative transfer for spherical media, *Numerical Heat Transfer, Part B*, **31**, pp.251-260 (1997).

16. C. E. Siewert, An inverse source problem in radiative transfer, *J. Quant. Spectrosc. Radiat. Transfer*, **50**, pp.603-609 (1993).

17. H. Y. Li and M. N. Özisik, Estimation of the radiation source term with a conjugate-gradient method of inverse analysis, *J. Quant Spectrosc. Radiat. Transfer*, **48**, pp.237-244 (1992).

18. L. H. Liu, H. P. Tan and Q. Z. Yu, Simultaneous identification of temperature profile and wall emissitivities in one-dimensional semitransparent medium by inverse radiation analysis, *Numerical Heat Transfer, Part A*, **36**, pp.511-525 (1999).

19. L. B. Barichello, R. D. M. Garcia and C. E. Siewert, On inverse boundary-condition problems in radiative transfer, *J. Quant. Spectrosc. Radiat. Transfer*, **57**, pp.405-410 (1997).

20. H. Y. Li and M. N. Özisik, Inverse radiation problem for simultaneous estimation of temperature profile and surface reflectivity, *Journal of Thermophyics and Heat Transfer*, **7**, pp.88-93 (1993).

21. R. Sanchez, N. J. McCormick and H. C. Yi, Iterative inverse radiative transfer method to estimative optical thickness and surface albedo, *Transport Theory and Statistical Physics*, **19**, pp357-385 (1990).

22. N. Fukshansky-Kazarinova, L. Fukshansky, M. Kühl and B. B. Jorgensen, Solution of the inverse problem of radiative transfer on the basis of measured internal fluxes, *J. Quant. Spectrosc. Radiat. Transfer*, **59**, pp.77-89 (1998).

23. H. C. Yi, R. Sanchez and N. J. McCormick, Bioluminescence estimation from ocean in situ irradiances, *Applied Optics*, **31**, pp. 822-830 (1992).

24. H. Y. Li and C. Y. Yang, A genetic algorithm for inverse radiation problems, *Int. J. Heat Mass Transfer*, **40**, pp.1545-1549 (1997).

25. A. J. Silva Neto, A. J. and M. N. Özisik, An inverse problem of simultaneous estimation of radiation phase function, albedo and optical thickness, *J. Quant. Spectrosc. Radiat. Transfer*, **53**, pp. 397-409 (1995).

26. B. M. Agarwal and M. P. Mengüç, Forward and inverse analysis of single and multiple scattering of collimated radiation in an axisymmetric system, *Int. J. Heat Mass Transfer*, **34**, pp. 633-647 (1991).

27. M. P. Mengüç and S. Subramaniam, A step-function approximation for the experimental determination of the effective scattering phase function of particles, *J. Quant. Spectrosc. Radiat. Transfer*, **43**, pp.253-265 (1990).

28. C. E. Siewert, On the inverse problem for a three-term phase function, *J. Quant. Spectrosc. Radiat. Transfer*, **22**, pp.441-446 (1979).

29. J.-H. Tsai, Inverse scattering problem with two-flux methods, *Int. Comm. Heat Mass Transfer*, **20**, pp.585-596 (1993).

30. M. P. Mengüç and S. Manickavasagam, Inverse radiation problem in axisymmetric cylindrical scattering media, *Journal of Thermophysics and Heat Transfer*, **7**, pp.479-486 (1993).

31. S. Subramaniam and M. P. Mengüç, Solution of the inverse radiation problem for inhomogeneous and anisotropically scattering media using a Monte Carlo technique, *Int. J. Heat Mass Transfer*, **34**, pp.253-266 (1991).

32. K. Kamiuto, A constrained least-squares method for limited inverse scattering problems, *J.*

Quant. Spectrosc. Radiat. Transfer, **40**, pp.47-50 (1988).

33. K. Kamiuto and J. Seki, Study of the P_1 approximation in an inverse scattering problem, *J. Quant. Spectrosc. Radiat. Transfer*, **37**, pp.455-459 (1987).

34. J. F. Sacadura, G. Uny and A. Venet, Models and experiments for radiation parameter estimation of absorbing, emitting and anisotropically scattering media, *Proc. 8th International Heat Transfer Conference*, California, USA, pp.565-570 (1986).

35. A. T. Kauati, A. J Silva Neto and N. C. Roberty, A source-detector methodology for the construction and solution of the one-dimensional inverse transport equation, *Inverse Problems in Engineering*, **9**, pp. 45-66 (2001).

36. A. T. Kauati, A. J Silva Neto and N. C. Roberty, Use of the source-detector technique in tomography with anisotropic scattering, *Proc. 17th Brazilian Congress of Biomedical Engineering*, Santa Catarina, pp. 1009-1014, in Portuguese (2000).

37. A. J. Silva Neto and N. C. Roberty, The source-detector approach for the solution of inverse problems in participating media, *Proc. 5th N-NE Brazilian Congress of Mechanical Engineering*, Ceará, Brazil, **III**, pp.340-347, in Portuguese (1998).

38. A. J. Silva Neto and N. C. Roberty, Sacttering tomography modelling as an inverse problem, *Proc. .4th Brazilian National Forum on Science and Technology in Health*, Paraná, Brazil, pp.317-318, in Portuguese (1998).

39. M. N. Özisik, Radiative Transfer and Interactions with Conduction and Convection, John Wiley & Sons (1973).

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

INFERENCE OF VELOCITY FIELDS BASED ON TOMOGRAPHIC MEASUREMENTS IN PROCESS INDUSTRY

Aku Seppänen
Marko Vauhkonen
Jari P. Kaipio
Department of Applied Physics
University of Kuopio
Kuopio, Finland
Aku.Seppanen@uku.fi

Erkki Somersalo
Institute of Mathematics
Helsinki University of Technology
Helsinki, Finland

ABSTRACT

In this paper we consider tomographic imaging of moving fluids with electrical impedance tomography (EIT). In EIT the conductivity distribution is reconstructed based on electrical measurements from the boundary of the object. In order to reconstruct time-varying targets with EIT we use the state estimation approach which allows the change of the target during the observations. The use of state-space formalism in EIT also enables the inference of underlying physical parameters in the target evolution. In this paper we propose a novel approach for estimating velocity fields based on EIT measurements. The velocity field is parametrized and the parameters of the velocity are estimated simultaneously with the conductivity distribution. The numerical studies show that it is possible to estimate the velocity based on EIT data - at least with some accuracy.

INTRODUCTION

Process tomography is a non-intrusive imaging technique used in process industry. Typical applications in which tomographic imaging is used are fluid flow in pipelines and mixing and foaming processes. In process tomography the target is reconstructed based on boundary measurements. One of the modalities used in process tomography is the electrical impedance tomography (EIT). In EIT we apply electric currents to the electrodes on the surface of the object and measure the corresponding voltages on the electrodes. The conductivity distribution is reconstructed based on the boundary voltage data.

However, in the case of moving fluids the final goal of the imaging procedure is not usually the conductivity distribution. Instead, parameters associated with the process dynamics are often in the focus of interest. Such parameters are for example the velocity field and the total mass flow rate. The traditional method for reconstructing velocity fields based on EIT is called the cross-correlation method [1, 2, 3], which is based on post processing the reconstructed conductivity images.

In this paper we propose a novel approach for estimating velocity fields in process tomography. In our approach we consider the velocity primarily as an unknown function and reconstruct it together with the conductivity distribution. For this end the EIT is formulated in the state-space formalism.

We have previously used the state-space formalism in EIT in the case that the velocity field is known *a priori* [4]. The use of state estimation approach in EIT have been shown to be advantageous, especially in the case that the rate of the change of the target is high in comparison to rate of the measurements [4, 5, 6, 7, 8].

In the state estimation paradigm the object is modeled as a stochastic process. The physical model that approximates the target evolution is written in the operator form and the discrepancies due to the approximation are included in the model as a stochastic process. The observation model with the evolution model constitute the state-space representation and the reconstruction problem can be written as the state estimation problem.

The structure of this paper is following. First, we review the observation model of EIT. Next, we formulate the convection-diffusion (CD) model which we use as the evolution model in the case of moving fluid. We consider both the case that the velocity is known and the case that the velocity is unknown. Next, we consider the associated nonlinear state estimation problem. Finally, we show two numerical examples.

OBSERVATION MODEL

In EIT alternating currents I_ℓ are applied to electrodes on the surface of the object and the resulting voltages U_ℓ on these electrodes are measured. The conductivity distribution σ is reconstructed based on the voltage measurements. In *stationary* EIT the conductivity distribution $\sigma = \sigma(x)$ is assumed to be a function of the spatial variable only, so that when different measurement sets are obtained, they correspond to the same target. The observation model of EIT is conventionally written as

$$V = U(\sigma) + v, \tag{1}$$

where V is a vector including the RMS-values of the measured voltages, U is a model between the conductivity σ and the measurements and v is additive observation noise. In order to reconstruct the conductivity σ based on the voltage data V the conductivity distribution and the model $U = U(\sigma)$ must be discretized. In the following we identify the actual infinite dimensional distribution σ with the finite dimensional vector of coordinates of the approximation of σ with respect to some basis. We denote the number of basis functions by N. The EIT reconstruction problem is ill-posed and spatial regularization is required in practically all cases in stationary reconstruction. In the Tikhonov regularization approach the conductivity is computed by minimizing the functional

$$\Xi(\sigma) = \|V - U(\sigma)\|^2 + \alpha^2 \|R(\sigma - \sigma_*)\|^2 \tag{2}$$

with respect to σ. In (2) α is the regularization parameter, R is the regularization matrix and σ_* is a prior guess for σ.

In order to reconstruct the conductivity distribution σ based on voltage observations we need to fix the forward model of EIT. There are several models, which are referred to as electrode models. In this paper we use the most accurate known model, the complete electrode model [9, 10]

$$\nabla \cdot (\sigma \nabla u) = 0, \quad x \in \Omega \tag{3}$$

$$u + z_\ell \sigma \frac{\partial u}{\partial n} = U_\ell, \quad x \in e_\ell,\ \ell = 1, \ldots, L \tag{4}$$

$$\int_{e_\ell} \sigma \frac{\partial u}{\partial n} \mathrm{d}S = I_\ell, \quad x \in e_\ell,\ \ell = 1, \ldots, L \tag{5}$$

$$\sigma \frac{\partial u}{\partial n} = 0, \quad x \in \partial\Omega \backslash \cup_{\ell=1}^{L} e_\ell \tag{6}$$

where $u = u(x)$ is the electric potential, e_ℓ is the ℓ^{th}electrode, z_ℓ is the contact impedance between the ℓ^{th}electrode and the object, U_ℓ is the potential on ℓ^{th}electrode, I_ℓ is the injected current and n is outward unit normal. In addition, to ensure the existence and uniqueness of the solution the following conditions are required

$$\sum_{\ell=1}^{L} I_\ell = 0\ , \ \sum_{\ell=1}^{L} U_\ell = 0. \tag{7}$$

The forward problem of EIT is to compute the potential $u = u(x)$ and the voltages U_ℓ, given the conductivity distribution σ, the contact impedances z_ℓ and the current pattern $I = (I_1, \ldots, I_L)$. The forward problem can be approximated using the finite element method (FEM). The weak form of the complete electrode model (3–7) was given in [10] and its FEM implementation for example in [11]. Usually the conductivity distribution is approximated by a piecewise constant function and an associated basis. However, we use the piecewise linear approximation as in [4].

In stationary EIT the voltage measurements V correnspond to set of several different current patterns. The set of current patterns used for one stationary reconstruction is called the *frame*. In process tomography, however, the target can change very rapidly. Consequently, the assumption of target being (even approximately) non-varying during each frame is not necessarily valid. In state estimation approach, or *dynamical EIT*, we take into account the change of the object. We approximate that the conductivity distribution does not change during each current injection, but varies between consecutive current patterns. That is, we consider the observation model (1) in the case that the measurements correspond to a single current pattern at time t instead of several current patterns. Thus, we write the time dependent observation model of EIT

$$V_t = U_t(\sigma_t) + v_t, \tag{8}$$

where V_t, U_t, σ_t and v_t are the voltage measurements, the forward model, the discretized conductivity distribution and the observation noise at time t, correspoindingly.

It is important to notice that the time-dependent observation model (22) is *highly underdeterministic*, because the number of unknown parameters is considerably higher than the number of observations at time t. Thus, it is unlikely to obtain a feasible reconstruction for σ_t by applying spatial regularization to the case of time-dependent observation model – unless the spatial prior information

is very specific. Our approach to the problem is following: We add temporal prior information to the reconstruction. In other words, we link the states σ_t and σ_{t+1} corresponding to times of consecutive current injections. This is performed by modeling the time dependence of the conductivity distribution. In the next section we model the evolution of the conductivity distribution in the case of moving fluid by using the stochastic convection-diffusion equation. When we have the observation model and the stochastic evolution model for the conductivity we can reconstruct the conductivities σ_t by using state estimation. Furhermore, in the following sections we show that the state estimation approach also enables the estimation of parameters associated with the transition from the state σ_t to σ_{t+1}, that is, estimation of the velocity field.

EVOLUTION MODEL

In this section we model for the conductivity distribution in the case of moving fluid. First we review the convection-diffusion model in the case that the velocity field is know. Next, we introduce an approach for modeling the process in the case that the velocity field is unknown.

Evolution model in the case of a priori known velocity field

Consider a system with moving fluid. We model the time-varying concentration distribution $c = c(x,t)$ by the stochastic convection-diffusion equation

$$dc = -\bar{v} \cdot \nabla c \, dt + \nabla \cdot \kappa \nabla c \, dt \; + dB(t), \quad (9)$$

where $\bar{v}$ is the velocity field, κ is the diffusion coefficient and $B(t)$ is Brownian motion [12]. The stochastic process $dB(t)$ represents the discrepancies between the model and the actual system. In this section we assume that the velocity field $\bar{v}$ is known *a priori*.

We also set the following initial and boundary conditions

$$c(x,0) = c_0(x), \;\; x \in \Omega \quad (10)$$

$$c(x,t) = c_{\text{in}}(x,t), \;\; x \in \partial\Omega_{\text{in}} \quad (11)$$

$$\frac{\partial c}{\partial n}(x,t) = 0, \;\; x \in (\partial\Omega \backslash \partial\Omega_{\text{in}}) \, , \quad (12)$$

where $c_{\text{in}}(x,t)$ is the time-varying concentration on the input boundary $\partial\Omega_{\text{in}}$ and n is the outward unit normal. Although the input concentration $c_{\text{in}}(x,t)$ is written as a Dirichlet condition, it is the primary unknown variable in the model. Thus we consider $c_{\text{in}}(x,t)$ as a stochastic process and write

$$c_{\text{in}}(x,t) = \bar{c}_{\text{in}}(x,t) + \eta(x,t), \;\; x \in \partial\Omega_{\text{in}} \, , \quad (13)$$

where $\bar{c}_{\text{in}}(x,t)$ represents the known part of the input that we can either measure (indirectly) or approximate, and $\eta(x,t)$ is a stochastic process. Usually $\bar{c}_{\text{in}}(x,t)$ is the spatial/temporal average mean of $c_{\text{in}}(x,t)$ or its estimate.

In [4, 13] we solved the initial value/boundary value problem (9-12) numerically by using the FEM. We do not show the details here but we note that the procedure is straightforward. Following the FEM scheme the numerical solution of the CD problem is obtained in the recursive form

$$c_{t+1} = \bar{F}_t c_t + \bar{s}_{t+1} + \psi_{t+1} \quad (14)$$

where $c_t \in \mathbb{R}^N$ includes the discretized values of the concentration at time t, $\bar{F}_t \in \mathbb{R}^{N \times N}$ is the evolution matrix, $\bar{s}_t \in \mathbb{R}^N$ includes contribution of the known part of the input of the system and $\psi_t \in \mathbb{R}^N$ is the state noise process in this representation. At this point we remark that $\bar{F}_t, \bar{s}_{t+1}$ and ψ_{t+1} all depend on velocity field and diffusion coefficient.

In the case of a multi-phase system the target possesses also contrast with respect to conductivity. The relation between the conductivity σ and the concentration c depends on the process and it is usually nonlinear. In this paper we assume, for simplicity, that the contrast in concentration is small enough to justify an affine approximation between the conductivity and the concentration. Furthermore, since the associated constant of this approximation can be absorbed into the observations, we write $\sigma = \lambda c$. In that case we obtain an approximation for the state equation in terms of the conductivity

$$\sigma_{t+1} = F_t \sigma_t + s_{t+1} + w_{t+1}, \quad (15)$$

where $F_t = \bar{F}_t$, $s_{t+1} = \lambda \bar{s}_{t+1}$ and $w_{t+1} = \lambda \psi_{t+1}$.

Evolution model with unknown velocity field

Next we consider the CD model in the case that the velocity is an unknown function. First, we write a parametric form for velocity field $\bar{v}$. That is we represent $\bar{v}$ as a linear combination of K basic functions

$$\bar{v}(x,t) = \sum_{k=1}^{K} \beta_{t,k} \bar{v}_k(x), \quad (16)$$

where $\bar{v}_k(x)$ is the k^{th}basic function for velocity and $\beta_{t,k}$ is the corresponding (unknown) coefficient at time t. For further use we define the vector $\beta_t = [\beta_{t,1}, \ldots, \beta_{t,K}]^T$. Since the coefficients $\beta_{t,k}$ are unknown we model them as random variables. In this paper we use the *random walk model*, that is $\beta_{t+1,k} = \beta_{t,k} + \epsilon_{t,k}$, or

$$\beta_{t+1} = \beta_t + \epsilon_t, \tag{17}$$

where $\epsilon_t = [\epsilon_{t,1}, \ldots, \epsilon_{t,K}]^T$ is Gaussian noise vector.

Next we substitute the representation (16) into the CD equation (9). As in the previous section we solve the CD equation numerically by using the FEM. If we again assume the linear dependence between the conductivity σ and the concentration c we obtain – equivivalently to (15) – the evolution equation for the conductivity

$$\sigma_{t+1} = F(\beta_t, t)\sigma_t + s(\beta_t, t) + w(\beta_t, t), \tag{18}$$

where the evolution matrix F, input term s and the state noise w are now functions of the random variable β_t.

Next we introduce a new state variable θ_t defined as

$$\theta_t := \begin{bmatrix} \beta_t \\ \sigma_t \end{bmatrix} \in \mathbb{R}^{(K+N)\times 1} \quad . \tag{19}$$

Furthermore, by stacking the equations (17) and (18), we obtain

$$\begin{bmatrix} \beta_{t+1} \\ \sigma_{t+1} \end{bmatrix} = \begin{bmatrix} \beta_t \\ F(\beta_t, t)\sigma_t + s(\beta_t, t) \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ w(\beta_t, t) \end{bmatrix} \tag{20}$$

which is clearly the evolution model for the variable θ_t. For convenience we write the equation (20) in the form

$$\theta_{t+1} = f(\theta_t) + \varpi_t(\theta_t), \tag{21}$$

where the nonlinear function f and state-dependent noise ϖ are defined as the first and second term on the right hand side in the equation (20), correspondingly.

By definition (19) the variable θ_t includes the conductivity distribution and the velocity parameters at time t. In the next section we consider the estimation of θ_t based on EIT observations.

STATE ESTIMATION

In this section we consider estimation of the state variable θ_t introduced in the previous section. For this end, we first define the model $U_t^*(\theta_t)$, that maps the variable θ_t to voltage observations. Since the measured voltages do not explicitely depend on the velocity of the fluid – or the coefficients β_t – the observation model can simply be written in the form

$$V_t = U_t^*(\theta_t) + v_t, \tag{22}$$

where $U_t^*(\theta_t) = U(\sigma_t)$, that is, the function $U_t^*(\theta_t)$ does not depend on β_t.

The evolution model (21) and the observation model (22) constitute the so-called state-space representation, based on which the state variable θ_t can be estimated given the voltage data $V_t, t = 1 \ldots, T$. Both of the evolution model and the observation model are nonlinear and thus, nonlinear state estimation is required.

In the case of linear models the state estimation problem is to find the conditional expectation of θ_t based on a set of observations $V_j,\ j \in \mathcal{J}$ where $\mathcal{J}$ denotes a subset of the available/usable observations. We denote the expectation of θ_t based on observations $V_j,\ j \in \mathcal{J},\ \mathcal{J} = \{1, \ldots, k\}$, by $\hat{\theta}_{t|k}$. Some of the estimates can be computed recursively. For example, the linear on-line estimator $\hat{\theta}_{t|t}$ is given by Kalman filter algorithm. Similarily, the so-called fixed-lag Kalman smoother gives an on-line – but not real time – estimate $\hat{\theta}_{t|t+q}$, where the lag $q > 0$ defines the number of future observations taken into account in computation of the estimate of θ_t. Equivalently, we can write $\hat{\theta}_{t-q|t}$ which means that at the (present) time t we obtain an estimate for the variable $\hat{\theta}_{t-q}$. For reference of linear filtering and smoothing theory, see for example [14, 15].

In the nonlinear case the conditional expectations can not be computed recursively. However there is a variety of suboptimal algorithms which have been used for nonlinear state estimation. The simplest algorithm is the linearized Kalman filter where both the evolution model and the observation model are linearized in advance and the Kalman filter is applied to the linearized models. A more advanced way is to use the so-called extended Kalman filter (EKF), where the models are linearized in each time step. The EKF algorithm is of the form [15]

$$\hat{\theta}_{t|t-1} = f(\hat{\theta}_{t-1|t-1}) \tag{23}$$

$$C_{\tilde{\theta}_{t|t-1}} = J_{f_t} C_{\tilde{\theta}_{t-1|t-1}} J_{f_t}^T + C_{\varpi_{t-1}} \quad (24)$$

$$K_t = C_{\tilde{\theta}_{t|t-1}} J_{U_t^*}^T (J_{U_t^*} C_{\tilde{\theta}_{t|t-1}} J_{U_t^*}^T + C_{v_t})^{-1} \quad (25)$$

$$C_{\tilde{\theta}_{t|t}} = (I - K_t J_{U_t^*}) C_{\tilde{\theta}_{t|t-1}} \quad (26)$$

$$\hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + K_t (V_t - U^*(\hat{\theta}_{t|t-1})) \quad (27)$$

In equations (23–27) $C_{(\cdot)}$ are covariance matrices. Especially, C_{v_t} is the covariance of the observation noise and C_{ϖ_t} is the covariance of the state noise. The matrices J_{f_t} and $J_{U_t^*}$ are the Jacobians of nonlinear functions f and U_t^*, correspondingly, evaluated at the predictor $\hat{\theta}_{t|t-1}$, that is,

$$J_{f_t} = \left.\frac{\partial f_t}{\partial \theta}\right|_{\theta=\hat{\theta}_{t|t-1}}, J_{U_t^*} = \left.\frac{\partial U_t^*}{\partial \theta}\right|_{\theta=\hat{\theta}_{t|t-1}} \quad (28)$$

Furthermore, since in our case the state noise ϖ is state-dependent, the covariance of C_{ϖ_t} is also evaluated in the predictor in each time step. The construction of the Jacobian J_{f_t} is straightforward but tedious work. Thus, we don't show the derivation here. The computation of the Jacobian $J_{U_t^*}$ is considered e.g. in [11].

In the linear case the fixed-lag smoother have been shown to give essentially better results than the Kalman filter since the smoother also takes into account the observations of the future. We refer to the following algorithm as the extended fixed-lag smoother (EFLS), because it is a modification of the fixed-lag smoother where the predictor $\hat{\theta}_{t|t-1}$ is used as the linearization point – similarly as in EKF. The EFLS estimates $\hat{\theta}_{t-q|t}$ are given by recursions

$$\hat{\theta}_{t-i|t} = \hat{\theta}_{t-i|t-1} + K_{t-i}(V_t - U^*(\hat{\theta}_{t|t-1}) \quad (29)$$

$$K_{t-i} = C^{(i,0)}_{\tilde{\theta}_{t|t-1}} J_{U_t^*}^T (J_{U_t^*} C_{\tilde{\theta}_{t|t-1}} J_{U_t^*}^T + C_{v_t})^{-1} \quad (30)$$

$$C^{(i+1,0)}_{\tilde{\theta}_{t+1|t}} = C^{(i,0)}_{\tilde{\theta}_{t|t-1}} (I - K_t J_{U_t^*})^T J_{f_t}^T \,, \quad (31)$$

where the inner iteration $i = 0, \ldots, q$ is performed for all t and where the covariance matrix

$$C^{(0,0)}_{\tilde{\theta}_{t|t-1}} = C_{\tilde{\theta}_{t|t-1}} \quad (32)$$

is obtained from the filter equations (24) and (26).

We still remind that in equations (23-27) and (29-29) the estimates $\hat{\theta}_{t|j}$ are not conditional expectations for θ_t as in the linear case.

NUMERICAL STUDIES

In this section show two numerical examples of the simultaneous estimation of conductivity and velocity based on EIT measurements. In the first example we form the target evolution by using parabolic flow field in the pipe. When solving the inverse problem the velocity field is known to be parabolic, that is, the number of basic functions for velocity equals to one. In the second example we have a non-parabolic velocity field, and in the inverse problem the velocity is represented with three basic functions. In both cases the true velocity field is stationary.

In order to make the computations in feasible time, we study two-dimensional cases in this paper. However, the extension to three dimensions is straightforward. We have already studied the state estimation approach in 3D [16].

Case 1: Parabolic velocity field

The first task was to form the target evolution. We assumed that the velocity field is parabolic. The parabolic flow at x-direction is of the form

$$v_x(x,y) = \frac{3}{2} v_{x,\mathrm{mean}} \left[1 - \left(\frac{|y - y_0|}{R}\right)^2\right] \quad (33)$$

where $v_{x,\mathrm{mean}}$ is the spatial average of the velocity, y_0 is the y-coordinate of the center of the pipe and R is the radius of the pipe. We choosed that the average velocity is $v_{x,\mathrm{mean}} = 450\mathrm{cm/s}$. It is worth to notice that with this flow rate the flow is actually turbulent and the velocity profile is not parabolic.

Next we computed the time-varying conductivity distribution. We assumed that the conductivity in the input layer $\sigma_{\mathrm{in}}(x,t) = \lambda c_{\mathrm{in}}(x,t)$ consists of poorly conducting "bubbles" in a homogeneous background. The occurrence of the bubbles was such that the distribution of the bubble centers across the input boundary was uniform and the time difference between consecutive occurences of the bubbles was narrow Gaussian distribution. The conductivity distributions were computed at time intervals $t = 1, \ldots, 64$ so that the actual time step Δt between t and $t+1$ was $5\mathrm{ms}$.

After the time-varying conductivity distribution was constructed, the voltage observations were computed from the FEM approximation of (3–7). The electrical potential u inside the pipe was approximated with second order basis functions in the FEM solution [17]. The placement of the electrodes is shown in Figure 1. The currents were injected from opposite electrode pairs and for each current injection the voltages were measured on all opposite electrode pairs. The order of the current patterns was randomized and

the time step between consecutive measurements was $5\mathrm{ms}$. All the voltage measurements corresponding to each current injection were assumed to be made simultanously and instantly. Thus, the conductivity was non-varying during each current injection. We computed the voltages corresponding to 64 current injections. Zero-mean Gaussian observation noise was added to the computed voltages. The observation noise consisted of two parts. First, we added to each observation noise with std 1% of the value of the corresponding observation. Secondly, we added to all observations noise with std 0.1% of the maximum voltage.

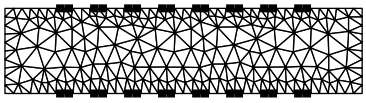


Figure 1: The FEM mesh of the EIT computations. The thick lines represent the placement of the electrodes.

Next we solved the inverse problem. That is, we estimated the velocity profile and the conductivity distribution based on the noisy observations. We assumed that the velocity is known to be of the form (33), but the average velocity $v_{x,\mathrm{mean}}$ was unknown. Thus, we had one unknown parameter for the velocity field. We computed the EFLS estimates for the states θ_t as described above. In the evolution model the diffusion coefficient was assumed to be known. The input of the system σ_{in} was unknown. Thus, we used the best homogeneous estimate σ_0 for computing "known source term" s_t. The fluctuations of the input were considered as a stochastic process as in equation (13) and the estimated variances of the input fluctuations were included in the state noise covariance matrix C_{ϖ_t} as in [4]. We assumed that the velocity field is stationary, that is, we set the variance of the noise ϵ_t in equation (17) equal to zero. As an initial quess for $v_{x,\mathrm{mean}}$ we used $200\mathrm{cm/s}$. The variance of the inital state was set to be "large". The covariance matrix C_{v_t} was constructed knowing the variances of the measurement noise. In EFLS algorithm we chose the lag q to be 6.

Figure 2 represents the true and the reconstructed conductivity distribution at some time steps. The reconstructions of the conductivities are quite feasible. Furthermore, Figure 3 represents the profile of the reconstructed velocity at x-direction at the final time step $t = 64$. The stationary parabolic field is reconstructed very accurately. Figure 4 shows the time evolution of the estimated mean velocity. As the time evolves the fluctuations of the estimate get smaller.

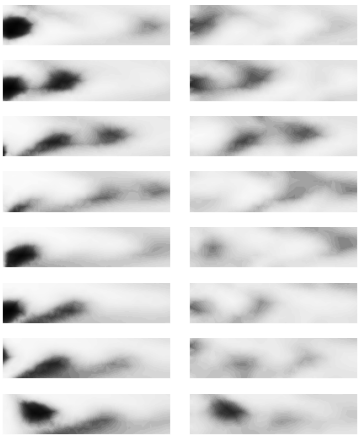


Figure 2: Left) True conductivity distribution, Right) the estimated conductivity distribution at time intervals $t = 16, 20, \ldots, 44$.

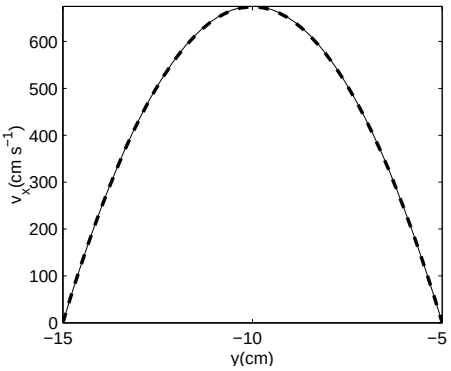


Figure 3: Profile of the true velocity at x-direction (solid line) and the Profile of the estimated velocity at x-direction (dashed line).

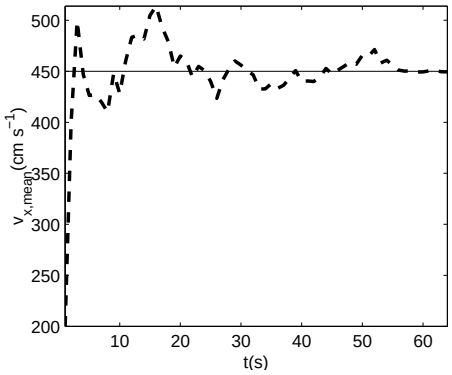


Figure 4: Time evolution of the estimate for the mean velocity (dashed line). The solid line represents the true mean velocity.

Case 2: Estimation of three velocity parameters

In this example the velocity field was set be non-symmetric with respect to y-axis of the pipe. The velocity at y-direcition was zero, and the velocity at x-direcition was independent of x-coordinate. The velocity profile was of the form of the function plotted with dashed line in Figure 5. In the inverse problem we had three basic functions for $\bar{v}$. The profiles of the basic functions are shown in figure 5. In other respects we proceeded the exactly as in the first example.

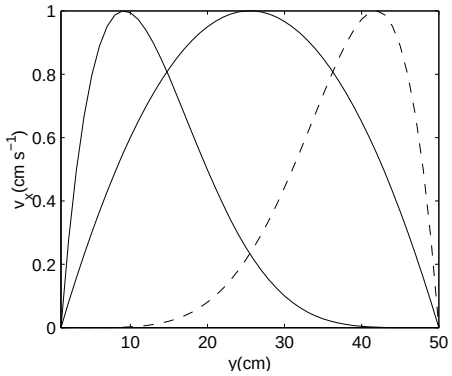


Figure 5: The profiles of the three basic functions for the velocity field

Figure 6 represents the true and the reconstructed conductivity distribution at certain time intervals. Again, the conductivities are tracked relatively well.

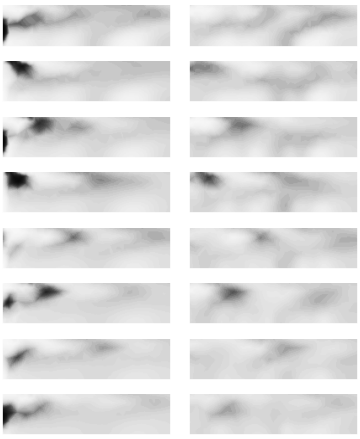


Figure 6: Left) True conductivity distribution, Right) the estimated conductivity distribution at time intervals $t = 16, 20, \ldots, 44$.

Figure 7 represents the profile of the reconstructed velocity at x-direction at the final time step $t = 64$. The estimated profile is quite feasible. The reason for the bias in the estimate is due to the suboptimality of the nonlinear state estimation algorithm – and perhaps partly due to nonuniqueness of the solution.

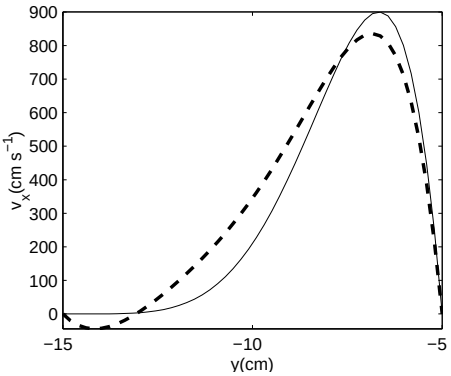


Figure 7: Profile of the true velocity at x-direction (solid line) and the Profile of the estimated velocity at x-direction (dashed line).

CONCLUSIONS

This paper is the first step towards simultaneous estimation of the velocity fields and conductivity distributions in based on EIT measurements. The main aim of this study was to find out if the direct estimation of velocity fields is possible. Based on the numerical studies our answer is: Yes – at least whith some accuracy. Evidently, there are still several issues that should be improved in the mod-

eling and the reconstruction. However, some of the extensions are relatively straightforward and their implementation only necessitates small changes to the overall estimation scheme.

In the numerical studies we considered only cases where the velocity was stationary. It is yet to be investigated is it possible to track velocities also in the nonstationary case. Furthermore, in our examples we only had a small number of basic functions for velocity. It is evident that if the number of basic functions is increased too much the estimation problem has not unique solution. However, as in stationary estimation, the spatial regularization can also be used in state estimation, for reference, see [6, 18]. That is, if we have also spatial prior information about the target we can take it into account in order to obtain a unique solution. Moreover, it is not obvious how to choose the basic functions for velocity. Our suggest is to use computational fluid dynamics (CFD) for constructing the basis. The use of CFD in this sense can be interpretated as the use of more specific prior information in the inverse problem.

Finally, we note that our nonlinear state estimation algorithm, EFLS, is far from optimal and it can be improved.

REFERENCES

1. T. Dyakowski. Process tomography applied to multi-phase flow measurement. *Meas. Sci. Tech.*, **7**, 343 (1996)

2. E. O. Etuke and R. T. Bonnecaze. Measurement of angular velocities using electrical impedance tomography. *Flow. Meas. Instr.*, **9**, 159 (1998)

3. G. P. Lucas, J. Cory, R. C. Waterfall, W. W. Loh, and F. J. Dickin. Measurement of the solids volume fraction and velocity distribution in solids-liquid flows using dual-plane electrical resistance tomography. *Flow. Meas. Instr.*, **10**, 249 (1999)

4. A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation with fluid dynamical evolution models in process tomography – an application to impedance tomography. *Inv. Probl.*, **17**, 467 (2001)

5. M. Vauhkonen, P.A. Karjalainen, and J.P. Kaipio. A Kalman filter approach to track fast impedance changes in electrical impedance tomography. *IEEE Trans. Biomed. Eng.*, **45**, 486 (1998)

6. J.P. Kaipio, P.A. Karjalainen, E. Somersalo, and M. Vauhkonen. State estimation in time-varying electrical impedance tomography. *Ann. New York Acad. Sci.*, **873**, 430 (1999)

7. P.J. Vauhkonen, M. Vauhkonen, and J.P. Kaipio. Fixed-lag smoothing and state estimation in dynamic electrical impedance tomography. *Int. J. Numer. Methods Eng.*, **50**, 2195 (2001)

8. P.J. Vauhkonen, M. Vauhkonen, T. Mäkinen, P.A. Karjalainen, and J.P. Kaipio. Dynamic electrical impedance tomography – phantom studies. *Inv. Prob. Eng.*, **8**, 495 (2000)

9. K.-S. Cheng, D. Isaacson, J.C. Newell, and D.G. Gisser. Electrode models for electric current computed tomography. *IEEE Trans. Biomed. Eng.*, **36**, 918 (1989)

10. E. Somersalo, M. Cheney, and D. Isaacson. Existence and uniqueness for electrode models for electric current computed tomography. *SIAM J. Appl. Math.*, **52**, 1023 (1992)

11. M. Vauhkonen. *Electrical Impedance Tomography and Prior Information*. PhD thesis, University of Kuopio, Kuopio, Finland, 1997.

12. B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, 1998.

13. A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation with fluid dynamical evolution models in process tomography – EIT application. Technical Report 6/2000, University of Kuopio, Department of Applied Physics, 2000. http://venda.uku.fi/research/IP/publications /deptrep/fluiddyneit.pdf.

14. J.L. Melsa and D.L. Cohn. *Decision and Estimation Theory*. McGraw-Hill, 1978.

15. B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice Hall, 1979.

16. A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation in three dimensional impedance imaging – Use of fluid dynamical evolution models. In *Proceedings of 2nd World Congress on Industrial Process Tomography*, pages 198–206, 2001.

17. P.J. Vauhkonen, M. Vauhkonen, T. Savolainen, and J.P. Kaipio. Three-dimensional electrical impedance tomography based on the complete electrode model. *IEEE Trans. Biomed. Eng.*, **46**, 1150, (1999)

18. D. Baroudi, J.P. Kaipio, and E. Somersalo. Dynamical electric wire tomography: Time series approach. *Inv. Probl.*, **14**, 799 (1998)

4th International Conference on Inverse Problems in Engineering
Rio de Janeiro, Brazil, 2002

AN INVERSION ALGORITHM FOR CAPACITANCE TOMOGRAPHY IMAGING OF TWO-PHASE FLOW REGIMES

Andres Fraguela
Facultad de Ciencias Fisico Matematicas
Benemerita Universidad Autonoma de Puebla
Puebla, Pue., Mexico
fraguela@fcfm.buap.mx

Carlos Gamio
Programa de Ductos
Instituto Mexicano del Petroleo
Mexico, D.F., Mexico
jgamio@imp.mx

Doris Hinestroza
Departamento de Matematicas
Universidad del Valle
Cali, Colombia
dhinest@imp.mx

ABSTRACT

Electrical-capacitance tomography (ECT) is a novel technique for the non-invasive internal visualization of industrial processes involving electrically non-conducting mixtures. One of the most promising potential applications of ECT is the measurement of gas-oil-water multiphase flow through pipes. ECT is based on the inversion of mutual capacitance measurements taken between a number of electrodes, placed around a particular cross-section of the pipe (which must have insulating walls in this area), in order to find the electrical permittivity distribution (reflecting the mixture composition). A new inversion algorithm based on a set of integral equations (equivalent to the differential model formulation) is presented here. Using these integral equations and expanding the unknown permittivity with respect to an orthonormal basis conveniently chosen in a Sobolev space of functions defined in the sensor cross-section, the inverse problem of determining the permittivity reduces to solving a quadratic optimization problem with quadratic constraints.

NOMENCLATURE

$c_{i,j}$	modified data values
$C_{i,j}$	inter-electrode mutual capacitances
V	electric potential
R_i	radius of the region boundaries
S_i	circular arc corresponding to electrode i
L_2, H_1	function spaces

Greek symbols

Δ	Laplace operator
∇	gradient operator
ε	electrical permittivity

INTRODUCTION

Oil wells typically produce not just oil, but a complex multiphase mixture having variable amounts of oil, gas and water. The determination of the quantity of each component actual being produced by each specific well is of the greatest importance for the efficient exploitation of the reservoirs. The conventional way of doing this is by separating the mixture and measuring each individual component using single-phase flowmeters. However, the three-phase separators needed are excessively bulky and expensive.

Multiphase flow measurement techniques that do not require mixture separation have emerged in the last decade [8]. Among the most promising approaches currently under investigation is one based on multiphase flow visualization using tomographic methods [3], in particular electrical-capacitance tomography (ECT). The main advantage of the tomographic methods lies in their inherently non-invasive and flow-regime independent operation.

ECT is an emerging technique aimed at the non-invasive internal visualization of electrically non-conducting mixtures in industrial processes like mixing, separation and multiphase flow [10,11]. Only the application to flow imaging will be considered here. The basic principle of this method is to place a sensor containing an array of between 8 and 16 contiguous sensing electrodes around the pipe carrying the process fluids, at the

cross-section to be investigated (see fig. 1). The pipe wall should be electrically non-conducting in the zone of the electrodes. The electrodes are typically 10 cm long. The sensor also has a an outer cylindrical metallic screen covering the whole thing, which is always kept at an electric potential of zero volts. The sensing electrodes are connected to an apparatus that allows all the mutual capacitances between the different electrode pairs to be measured, and from this set of measurements the electrical permittivity distribution inside the sensor is obtained using a suitable inversion algorithm. The permittivity distribution reflects the spatial arrangement of the phases in the flow. Image reconstruction can thus be regarded as an inverse permittivity problem.

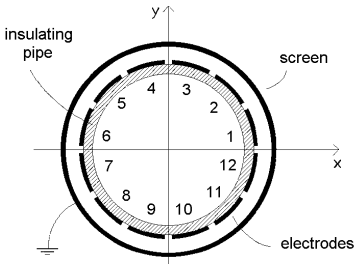


Fig.1 Cross-section of sensor

For 2-phase flows like gas-oil, the permittivity distribution directly determines the distribution of each phase, whereas for 3-phase ones like gas-oil-water, the distribution of an additional parameter (i.e. conductivity, etc.) must be obtained first through a different tomographic modality [1] in order to resolve each phase distribution.

To measure the mutual capacitance $C_{i,j}$ between electrodes i and j, a sinusoidal voltage of magnitude V is applied to electrode i (source) keeping the rest at zero potential, and the electric charge Q induced on electrode j (detector) is measured. $C_{i,j}$ is then given by

$$C_{i,j} = \frac{Q_j}{V_i} \qquad (1)$$

Because $C_{i,j} = C_{j,i}$, there are only $n = ½\,N\,(N - 1)$ independent mutual capacitance values, where N is the number of electrodes. For a 12-electrode sensor $n = 66$.

The use of the cylindrical guards at the ends of the sensing electrodes (and assuming that the phase distribution changes slowly in the axial direction) allows the sensor to be represented by a two-dimensional (2-D) model [9]. Assuming that the flow changes negligibly during the time required for one set of 66 measurements, and that the frequency of the excitation voltage is so small that the corresponding wavelength is much larger that the sensor dimensions, a static model can be considered.

PROBLEM FORMULATION

The forward problem is to determine the mutual capacitances $C_{i,j}$ given a known permittivity distribution inside the pipe. The inverse problem involves finding an unknown permittivity distribution inside the pipe based on the knowledge of all the $½\,N\,(N - 1)$ mutual capacitances $C_{i,j}$ (note that in the pipe itself and in the area between the pipe and the screen the permittivity distribution is known).

Mathematical Model

The sensor is modeled (fig.2) as a dielectric region Ω made of three subregions: a circle Ω_1 and two adjacent concentric rings Ω_2 and Ω_3, representing the interior of the pipe, the pipe wall and the area between the pipe and the screen, respectively. The electrodes, being very thin and having very small gaps between them, are modeled as equipotential surfaces (lines in the 2-D model) covering the entire boundary between Ω_2 and Ω_3. The screen is modeled as an equipotential line on the outer perimeter of Ω_3.

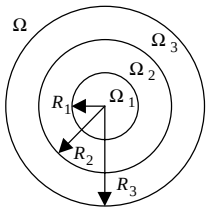


Fig.2 Sensor model

Let $z = (x,y)$ and $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$ where $\Omega_1 = \{z : |z| < R_1\}$, $\Omega_2 = \{z : R_1 < |z| < R_2\}$ and $\Omega_3 = \{z : R_2 < |z| < R_3\}$.

Each one of the subregions Ω_i ($i = 1,2,3$) has an electrical permittivity $\varepsilon_i(z)$. For $i = 2,3$ the function $\varepsilon_i(z)$ has a known constant value ε_i, whereas it is unknown for $i = 1$.

We shall consider that electrode i lies on the arch S_i defined by

$$S_i = \{z: |z| = R_2,\ 2\pi(i\text{-}1)/N \le \arg(z) \le 2\pi i/N\ \}$$

We shall denote by $V^{(i)}(z)$, $i = 1,2,..,N$, the potential produced in Ω when the potential in electrode i is set to 1 and that of the remaining electrodes is set to zero.

We shall consider that $V_j^{(i)}(z) = V^{(i)}(z)$ and $\varepsilon(z) = \varepsilon_j(z)$ if $z \in \Omega_j$ (j = 1,2,3). $\varepsilon_j(z)$ is constant ε_j for j = 2,3.

The electric potential $V^{(i)}(z)$, i = 1,2,..,N in the model must satisfy the following:

$$\nabla(\varepsilon_k(z)\nabla V_k^{(i)}) = 0 \text{ in } z \in \Omega_k \ (k = 1,2,3) \quad (2)$$

and the boundary conditions

$$V^{(i)}(z) = \begin{cases} 0 \text{ in } |z| = R_3 \\ 1 \text{ in } z \in S_i \\ 0 \text{ in } |z| = R_2, \ z \notin S_i \end{cases} \quad (3)$$

$$V_1^{(i)}(z) = V_2^{(i)}(z), \text{ in } |z| = R_1 \quad (4)$$

$$\varepsilon_1(z)\frac{\partial V_1^{(i)}}{\partial n_1}(z) = \varepsilon_2 \frac{V_2^{(i)}}{\partial n_1}(z), \text{ in } |z| = R_1 \quad (5)$$

where n_1 is the exterior normal unitary vector to the circle $|z| = R_1$.

The Inverse Problem

In what follows, our model will be determined by the boundary value problem (2)-(5). We can describe the inverse problem as follows:

Given ½ N (N - 1) values $C_{i,j}$ (i, j = 1,2,…,N) ($i < j$), of the mutual capacitances between the electrodes S_i and S_j, determine approximately the value of $\varepsilon_1(z)$ using the model (2)-(5).

The direct problem consists in calculating the solution $V^{(i)}(z)$ of the boundary value problem (2)-(5) and the mutual capacitance values $C_{i,j}$, for a known permittivity $\varepsilon_1(z)$.

The direct problem is a well posed problem in the Hadamard sense and is numerically stable [6], but the inverse problem is an ill-posed problem in the sense that small errors in the data may produce a large error in the reconstructed solution. On the other hand, the numerical discretization of the problem can also lead to significant deviations in the calculation of $\varepsilon_1(z)$, and in this case regularized algorithms [5] have to be used. We will propose an algorithm that will take into account the ill-posedness and the numerical instability in the solution of the inverse problem.

The mutual capacitance values are given by

$$C_{i,j} = K\int_{S_j} \varepsilon(z)\frac{\partial V^{(i)}(z)}{\partial n} ds \quad (6)$$

where n_2 is the exterior normal unitary vector to the circle $|z| = R_2$ and K is a constant with units of inverse potential and $i \neq j$.

We shall observe that is possible to obtain from (6) the relation

$$\int_{S_j^-} \frac{\partial V_2^{(i)}}{\partial n_2} ds = \frac{1}{\varepsilon_2}\left\{\varepsilon_3 \int_{S_j^+} \frac{\partial V_3^{(i)}}{\partial n_2} ds - \frac{C_{i,j}}{K}\right\} \quad (7)$$

where S_j^- and S_j^+ denote the arc S_j obtained as a limit set of points from Ω_3 and Ω_2, respectively.

The fact that the function $V_3^{(i)}$ is decoupled in the model (2)-(5) and is an harmonic function in Ω_3, with boundary conditions given in (3), means that we can calculate $V_3^{(i)}$ independently, which allows us to obtain the value of the right-hand side of (7) from the knowledge of the mutual capacitance values measured experimentally $C_{i,j}$. For that reason, if we denote

$$c_{i,j} = \int_{S_j^-} \frac{\partial V_2^{(i)}}{\partial n_2} ds \quad (8)$$

then, the new equivalent problem that we will consider is the following:

Given ½ N (N - 1) values $c_{i,j}$ (i, j = 1,2,…,N) ($i < j$), obtained using (7) *and* (8), *determine approximately the value of $\varepsilon_1(z)$ using the model* (2)-(5).

Uniqueness

In order to give sense to the uniqueness of the inverse problem solution, we shall consider the auxiliary boundary value problem:

$$\nabla(\varepsilon(z)\nabla V) = 0 \text{ in } \Omega_1 \cup \Omega_2 \quad (9)$$

under the continuity conditions defined in (4)-(5) on $|z| = R_1$, and

$$V(z) = \phi(z) \quad (10)$$

for $|z| = R_2$.

To every fixed $\varepsilon_1(z)$ corresponds a Dirichlet-Neumann type operator

$$T_\varepsilon : H^{1/2}(|z| = R_2) \to H^{-1/2}(|z| = R_2)$$

that associates the value of the normal derivative

$$T_\varepsilon(\phi) = \frac{\partial V}{\partial n}(z) = \Psi(z) \quad (11)$$

where $V \in H^1(\Omega_1 \cup \Omega_2)$ and $|z| = R_2$.

Theorem 1. *If the Dirichlet-Neumann operators corresponding to two continuous functions $\varepsilon_1(z)$ and $\varepsilon_1^*(z)$ coincide for all possible boundary conditions* (10), *then $\varepsilon_1(z) = \varepsilon_1^*(z)$ in $|z| \le R_1$.*

Proof. Suppose that in addition to the system (4)-(5) and (9)-(11) we have another analogous system denoted by (4')-(5') and (9')-(11'), obtained by substituting ε_1^* and V^* for ε_1 and V, respectively. If we subtract (9), (10) and (11) from (9'), (10') and (11'), respectively, then for the harmonic function $W_2 = V_2^* - V_2$ in $R_1 < |z| < R_2$, we have that $W_2(z) = \dfrac{\partial W_2}{\partial n_2}(z) = 0$ in $|z| = R_2$.

As a consequence of the theorem about analytic continuation for harmonic functions, it follows that $W_2(z) \equiv 0$, and therefore $V_2^*(z) = V_2(z)$ in $R_1 < |z| < R_2$.

Then, from (4), (5) and (4'), (5'), we can conclude that, in $|z| = R_1$:

$$V_1(z) = V_1^*(z) \tag{12}$$

$$\varepsilon_1(z)\frac{\partial V_1}{\partial n_1}(z) = \varepsilon_1^*(z)\ \frac{V_1^*}{\partial n_1}(z) \tag{13}$$

Using (13) and following the line of reasoning of reference [2], the relations given below can be obtained

$$\varepsilon_1(z)\nabla V_1(z) = \varepsilon_1^*(z)\nabla V_1^*(z) =$$

$$= -\frac{1}{\pi}\int_{|\xi|=R_1} \varepsilon_1(\xi)\frac{\partial V_1}{\partial n_1}(\xi)\nabla_z \ln|\xi - z|\, ds_\xi \tag{14}$$

Applying Green's formula to the functions V_1 and V_1^*, and to the Green's function $d(z,\xi)$ corresponding to the Dirichlet problem for the Laplace operator in $|z| < R_1$, the following relation is obtained [2]

$$V(z) = \int_{|\xi|<R_1} (1-\varepsilon(\xi))\,\nabla V(\xi)\cdot\nabla_\xi\, d(z,\xi)d\xi$$

$$- \int_{|\xi|=R_1} V(\xi)\frac{\partial d}{\partial n_1}(z,\xi)ds_\xi \tag{15}$$

which is valid for the pairs (V_1, ξ_1) and (V_1^*, ξ_1^*). Using this two relations (one for each pair) and (14), the following equality is obtained

$$W_1(z) = \int_{|\xi|<R_1} \nabla W_1(\xi)\cdot\nabla_\xi d(z,\xi)d\xi \tag{16}$$

for $|z| < R_1$, where $W_1 = V_1^* - V_1$.

From (16), we conclude that $W_1(z)$ is harmonic in $|z| < R_1$, and, since $W_1(z) = 0$ in $|z| = R_1$ then $W_1(z) \equiv 0$ in $|z| < R_1$, which takes us to the conclusion that $V_1 = V_1^*$.

But then, subtracting conveniently equations (9) and (9'), as well as the boundary conditions (5) and (5'), we get

$$\nabla\left(|\varepsilon_1(z) - \varepsilon_1^*(z)|\nabla V_1\right) = 0 \text{ in } |z|<R_1 \tag{17}$$

$$|\varepsilon_1(z) - \varepsilon_1^*(z)|\frac{\partial V_1}{\partial n_1}(z) = 0 \text{ in } |z|=R_1 \tag{18}$$

By Green's formula, we conclude that $|\varepsilon_1(z) - \varepsilon_1^*(z)|\,|\nabla V_1|^2 = 0$ in $|z| < R_1$. Thus, $V_1(z)$ must be constant in the open set ω of $|z| < R_1$ composed by the points where $\varepsilon_1(z) \neq \varepsilon_1^*(z)$, regardless the value of the boundary condition $\phi(z)$ in (10). It is clear that this cannot happen except if ω is the empty set, i.e. $\varepsilon_1 \equiv \varepsilon_1^*$. □

In our case, where we have as data of the problem only some averages of the normal derivative $\dfrac{\partial V_2}{\partial n_2}(z)$ over arcs in $|z| = R_2$, when the function V_2 is the solution of the problem (4), (5), (9), (10) corresponding to a very limited set of boundary conditions (10), obviously there cannot be uniqueness in the recovery of $\varepsilon_1(z)$. For other issues related with this result of uniqueness we refer to [1,4].

PROBLEM SOLUTION

From now on we will consider that the solution $V_3^{(i)}$ is known. In the scheme of the solution we will introduce several auxiliary functions whose analytical or numerical calculation can be constructed independently.

Fundamental Relations

To construct the algorithm for the solution of the inverse problem we have first to find a set a relations between the functions $V_1^{(i)}$, $V_2^{(i)}$ and ε_1, that allow us to obtain ε_1 approximately from the data $\{c_{i,j}\}$, $i, j = 1,2,\ldots,N$, given by (7) and (8).

We will classify the relations in two different types: (*a*) those that allow us to obtain $V_2^{(i)}$ and $\dfrac{\partial V_2^{(i)}}{\partial n_2}$ in $|\xi| = R_1$ from the data $\{c_{i,j}\}$ and also

ensure that $V_2^{(i)}$ and $\dfrac{\partial V_2^{(i)}}{\partial n_2}$ in $|\xi| = R_1$ are the restriction and the normal derivative of a harmonic function in Ω_2 that satisfies the boundary conditions (3) in $|\xi| = R_2$; and (*b*) those that allow us to obtain simultaneously $V_1^{(i)}$, $\nabla V_1^{(i)}$ and ε_1 approximately from the minimization of an objective functional, in such way as to satisfy equation (2) in Ω_1 and the boundary conditions (4) and (5), taking into account the a priori information that we have about $\varepsilon_1(z)$ in terms of bounds.

If we apply Green's formula to the Laplace operator in Ω_2 and the functions $V_2^{(i)}(\xi)$ and $N(z,\xi)$, where

$$\Delta N(z,\xi) = \delta(z-\xi);\quad z,\xi \in \Omega_2 \tag{19}$$

$$\frac{\partial N}{\partial n_z}(z,\xi) = 0 \ \text{ in } \ |z| = R_1, \quad \xi \in \Omega_2 \tag{20}$$

$$N(z,\xi) = 0 \ \text{ in } \ |z| = R_2, \quad \xi \in \Omega_2 \tag{21}$$

$\delta(z\text{-}x)$ being Dirac's delta function, and we take the limit when $z \to z_o$ with $|z_o| = R_1$, then we obtain for $i = 1,2,\ldots,N$

$$V_2^{(i)}(z_o) = \int_{|\xi|=R_1} N(z_o,\xi)\frac{\partial V_2^{(i)}}{\partial n_2}(\xi)\,ds_\xi + \int_{S_i} \frac{\partial N}{\partial n_2}(z_o,\xi_1)\,ds_{\xi_1} \tag{22}$$

where $|\xi| = R_1$ and $|\xi_1| = R_2$, which express the relation that must exist between the boundary values of a harmonic function in Ω_2 and its normal derivative in $|\xi| = R_1$, for the boundary condition (3) to be satisfied in $|\xi_1| = R_2$.

Now, in order to find a relation between $\varepsilon_1(z)$, $\nabla V_1^{(i)}(z)$ and $\dfrac{\partial V_2^{(i)}}{\partial n_1}(z)$ we consider the boundary problem

$$\nabla(\varepsilon_1(x,y)\nabla V_1^{(i)}) = 0 \ \text{ in } \ |z| < R_1$$

$$\varepsilon_1(x,y)\frac{\partial V_1^{(i)}}{\partial n_1}(z) = \Psi \ \text{ in } \ |z| = R_1$$

with $\int_{|\xi|=R_1} \Psi\,ds = 0$.

Making the change of variables

$$\varepsilon_1(x,y)\nabla V_1^{(i)} = \nabla u^{(i)} \tag{23}$$

then $u^{(i)}$ satisfies the boundary problem

$$\Delta u^{(i)} = 0 \ \text{ in } \ |z| < R_1$$

$$\frac{\partial u^{(i)}}{\partial n_1}(z) = \Psi \ \text{ in } \ |z| = R_1$$

whose solution is given by

$$u^{(i)}(z) = -\frac{1}{\pi}\int_{|\xi|=R_1} \Psi(\xi)\ln|\xi - z|\,ds_\xi + const \tag{24}$$

with $|z| < R_1$.

From (23) and (24) and the boundary conditions (5) we obtain

$$\varepsilon_1(z)\nabla V_1^{(i)}(z) = -\frac{\varepsilon_2}{\pi}\int_{|\xi|=R_1} \frac{\partial V_2^{(i)}}{\partial n_1}(\xi)\nabla_z \ln|\xi - z|\,ds_\xi \tag{25}$$

If we apply Green's formula to the function $V_1^{(i)}$, and the Green's functions $d(z,\xi)$ and $n(z,\xi)$ for the Dirichlet and Neumann problems in $|z| < R_1$, respectively, we have

$$V_1^{(i)}(z) = \int_{|\xi|<R_1} \nabla V_1^{(i)}(\xi)\cdot\nabla_\xi d(z,\xi)\,d\xi + \int_{|\varphi|=R_1} F(z,\varphi)\frac{\partial V_2^{(i)}}{\partial n_1}(\varphi)\,ds_\varphi + h_i(z) \tag{26}$$

$$V_1^{(i)}(z) = \int_{|\xi|<R_1} \nabla V_1^{(i)}(\xi)\cdot\nabla_\xi n(z,\xi)\,d\xi + \int_{|\varphi|=R_1} G(z,\varphi)\frac{\partial V_2^{(i)}}{\partial n_1}(\varphi)\,ds_\varphi + a_i \tag{27}$$

where

$$F(z,\varphi) = \frac{\varepsilon_2}{\pi}\int_{|\xi|<R_1} \nabla_\xi \ln|\xi - z|\cdot\nabla_\xi d(z,\xi)\,d\xi - \int_{|\psi|=R_1} N(\psi,\varphi)\frac{\partial d}{\partial n_\psi}(z,\psi)\,ds_\psi, \tag{28}$$

$$G(z,\varphi)=\frac{\varepsilon_2}{\pi}\int_{|\xi|<R_1}\nabla_\xi \ln|\xi-z|\cdot\nabla_\xi n(z,\xi)d\xi$$
$$+\varepsilon_2 n(z,\varphi)-\frac{1}{2\pi R_1}\int_{|\psi|=R_1}N(\psi,\varphi)ds_\psi, \qquad (29)$$

$$h_{(i)}(z)=-\int_{S_i}\left\{\int_{|\psi|=R_1}\frac{\partial d}{\partial n_\psi}(z,\psi)\frac{\partial N}{\partial n_{\xi_1}}(\psi,\xi_1)ds_\psi\right\}ds_{\xi_1} \qquad (30)$$

and

$$a_i=-\frac{1}{2\pi R_1}\int_{S_i}\left\{\int_{|\psi|=R_1}\frac{\partial N}{\partial n_{\xi_1}}(\psi,\xi_1)ds_\psi\right\}ds_{\xi_1} \qquad (31)$$

Relation (22) is of type (*a*), whereas (25)-(27) are relations of type (*b*).

Theorem 2*. For every strictly positive and continuously differentiable function* $\varepsilon_1(z)$ *in* $|z|\le R_1$*, the solution to the boundary value problem* (2)-(5) *is unique and coincides , for each* $i = 1,...,N$*, with the solution of the system of integral relations* (22), (25)-(27).

Proof. Using Green's formula in a convenient way [2], integral relations (22), (25)-(27) are obtained for the solution of the boundary value problem (2)-(5). It is easy to see that the solution to this problem is unique for each $i = 1,...,N$.

Suppose now that $V^{(i)}$ satisfy the integral relations (22), (25)-(27). Obviously, (2) is obtained from (22) and (25).

If we suppose that $V_2^{(i)}(z) = g(z)$ in $|z| = R_2$, then from Green´s formula and from (25) we get

$$\int_{|\xi_1|=R_2}\left[g(\xi_1)-\chi_i(\xi_1)\right]\frac{\partial N}{\partial n_{\xi_1}}(z_2,\xi_1)\,ds_{\xi_1}=0$$
$$\text{in} \quad |z| = R_2 \qquad (32)$$

where χ_i is the characteristic function of the arc S_i.

If we consider the harmonic function $W_2^{(i)}$ in $R_1 < |z| < R_2$ that satisfies the Dirichlet condition $g - \chi_i$ in $|z| = R_2$ and the null Neumann condition in $|z| = R_1$, then, applying Green's formula to $W_2^{(i)}$ and to the Green's function $N(z,\xi)$ in the ring $R_1 < |z| < R_2$, and using (32), it is deduced that $W_2^{(i)}(z) \equiv 0$ and, therefore, $g \equiv \chi_i$, from where the boundary condition (3) in $|z| = R_1$ is obtained.

In order to obtain the boundary conditions (4) and (5), equalities (22) and (25) are substituted into (26) and (27), resulting in the following relations

$$V_1^{(i)}(z)=\int_{|\xi|<R_1}(1-\varepsilon_1)\nabla V_1^{(i)}(\xi)\nabla_\xi d(z,\xi)d\xi$$
$$-\int_{|\xi|=R_1}V_2^{(i)}(\xi)\frac{\partial d}{\partial n_1}(z,\xi)ds_\xi \qquad (33)$$

$$V_1^{(i)}(z)=\int_{|\xi|<R_1}(1-\varepsilon_1)\nabla V_1^{(i)}(\xi)\nabla_\xi n(z,\xi)d\xi$$
$$+\varepsilon_2\int_{|\xi|=R_1}n(z,\xi)\frac{\partial V_2^{(i)}}{\partial n_1}(\xi)ds_\xi \qquad (34)$$
$$-\frac{1}{2\pi R_1}\int_{|\xi|=R_1}V_2^{(i)}(\xi)ds_\xi$$

From (33) and (34), and applying conveniently Green's formula to the function $V_1^{(i)}$, the following equalities are deduced in $|z| < R_1$:

$$\int_{|\xi|=R_1}\left[V_2^{(i)}(\xi)-V_1^{(i)}(\xi)\right]\frac{\partial d}{\partial n_1}(z,\xi)\,ds_\xi=0, \qquad (35)$$

$$\int_{|\xi|=R_1}\left[\varepsilon_1(\xi)\frac{\partial V_1^{(i)}}{\partial n_1}(\xi)-\varepsilon_2\frac{\partial V_2^{(i)}}{\partial n_1}(\xi)\right]\cdot n(z,\xi)\,ds_\xi$$
$$-\frac{1}{2\pi R_1}\int_{|\xi|=R_1}\left[V_1^{(i)}(\xi)-V_2^{(i)}(\xi)\right]ds_\xi=0 \qquad (36)$$

Repeating the same line of thought used with (32), (4) and (5) can be obtained from (35) and (36).□

Scheme of the Algorithm for the Solution of the Inverse Problem

We will explain the process of reconstruction of the permittivity coefficient in several steps.

If we look for the function $V_2^{(i)}$ as

$$V_2^{(i)}(r,\theta) = a_0^{(i)} + b_0^{(i)}\ln r$$
$$+\sum_{n=1}^{\infty}(a_n^{(i)}r^n + b_n^{(i)}r^{-n})\cos n\theta \qquad (37)$$
$$+\sum_{n=1}^{\infty}(c_n^{(i)}r^n + d_n^{(i)}r^{-n})\sin n\theta$$

and using the boundary condition (3) in $|z| = R_2$ and considering the experimental data $C_{i,j}$, then, solving a system of linear equations we can find the coefficients $a_n^{(i)}$, $b_n^{(i)}$, $c_n^{(i)}$ and $d_n^{(i)}$, for $0 \le n \le ½(N\text{-}2)$. Once we find these coefficients that determine $V_2^{(i)}$, we are able to calculate both $V_1^{(i)}$ and $\nabla V_1^{(i)}$ aproximately using equalities (25)-(27). We can note that in these equalities appear in explicit form the coefficients of (37). We will find $V_1^{(i)}$ in the form

$$V_1^{(i)}(z) = \sum_{k=1}^{M} A_k^{(i)} W_k(z) \tag{38}$$

where $\{W_k(z)\}_{k=1}^{\infty}$ is an orthonormal basis in the Sobolev space $H^1(\Omega_1)$ and the number M should be chosen appropriately [2] and it is assumed that (38) is an approximation to $V_1^{(i)}$ in $H^1(\Omega_1)$.

To find the coefficients $A_k^{(i)}$ of (45) we will minimize the functional

$$l(\vec{V}_1) = \sum_{i=1}^{N} \left\{ \left\| f_i(z) \frac{\partial V_1^{(i)}}{\partial y}(z) - g_i(z) \frac{\partial V_1^{(i)}}{\partial x}(z) \right\|^2 + \left\| V_1^{(i)}(z) - \int_{|\xi|<R_1} \nabla_\xi V_1^{(i)}(\xi) \cdot \nabla_\xi d(z,\xi) d\xi - F_i(z) \right\|^2 + \left\| V_1^{(i)}(z) - \int_{|\xi|<R_1} \nabla_\xi V_1^{(i)}(\xi) \cdot \nabla_\xi n(z,\xi) d\xi - G_i(z) \right\|^2 \right\} \tag{39}$$

where the norm in (39) is taken in the square-integrable function space $L_2(\Omega_1)$, $\vec{V}_1 = (V_1^{(i)},.., V_1^{(N)})$, and for $z = (x,y)$ and $\xi = R_1 e^{i\tau}$, the functions

$$f_i(z) = \int_{|\xi|=R_1} \frac{\partial V_2^{(i)}}{\partial n_1}(\xi) \frac{x - R_1 \cos\tau}{|\xi - z|^2} ds_\xi \tag{40}$$

$$g_i(z) = \int_{|\xi|=R_1} \frac{\partial V_2^{(i)}}{\partial n_1}(\xi) \frac{y - R_1 \sin\tau}{|\xi - z|^2} ds_\xi \tag{41}$$

$$F_i(z) = \int_{|\varphi|=R_1} F(z,\varphi) \frac{\partial V_2^{(i)}}{\partial n_1}(\varphi) ds_\varphi + h_i(z) \tag{42}$$

$$G_i(z) = \int_{|\varphi|=R_1} G(z,\varphi) \frac{\partial V_2^{(i)}}{\partial n_1}(\varphi) ds_\varphi + a_i \tag{43}$$

Note that with the minimization of $l(\vec{V}_1)$ we are ensuring that relations (25)-(27) are approximately met.

The minimization problem is solved under the restrictions:

$$0 \le \varepsilon_{min} \le \varepsilon_1(z) \le \varepsilon_{min} \tag{44}$$

If we choose $\{W_k(z)\}_{k=1}^{\infty}$ as an orthonormal basis in $H^1(\Omega_1)$ such that $\{\gamma_k W_k(z)\}_{k=1}^{\infty}$ is an orthonormal basis in $L_2(\Omega_1)$ for some adequate coefficients γ_k [6], then besides avoiding the instability of the numerical differentiation, we have that the problem of minimizing the functional $l(\vec{V}_1)$ subject to the restriction (44) is equivalent [2] to solving N quadratic optimization problems:

$$min \left| F_i \vec{a}_i \right|_M^2 + \left| D\vec{a}_i - \vec{b}_i \right|_M^2 + \left| N\vec{a}_i - c_i \right|_M^2 \tag{45}$$

subject to the quadratic constraints

$$\frac{\varepsilon_2 R_1}{\sqrt{2\pi}\varepsilon_{\max}} \le \left| A_i \vec{a}_i \right|_M \le \frac{\varepsilon_2 R_1}{\sqrt{2\pi}\varepsilon_{\min}} \tag{46}$$

$$\frac{\varepsilon_2 R_1}{\sqrt{2\pi}\varepsilon_{\max}} \le \left| B_i \vec{a}_i \right|_M \le \frac{\varepsilon_2 R_1}{\sqrt{2\pi}\varepsilon_{\min}} \tag{47}$$

The matrices F_i, A_i, B_i, D and N are described in detail in [2]. We can observe that the coefficients in these matrices depend of the values of the coefficients in (38) and on the Fourier coefficients of the fuctions (40)-(43) with respect to $\{W_k(z)\}$. Once we obtain the coefficients of $V_1^{(i)}$, we can use (25) to find an approximation of the permittivity of the form

$$\varepsilon_1(z) = \frac{\varepsilon_2}{\pi N} \sum_{i=1}^{N} \frac{\left| \int_{|\xi|=R_1} \frac{\partial V_2^{(i)}}{\partial n_1}(\xi) \nabla_z \ln|z - \xi| ds_\xi \right|}{\left| \nabla V_1^{(i)}(z) \right|} \tag{49}$$

If we have the partition $(\omega_k)_{k=1}^{\tilde{M}}$ of $|z| < R_1$, then expression (48) can be used to obtain an average value $\varepsilon_1^{\,k}$ of $\varepsilon_1(z)$ for each element of the partition. It is possible to prove that we can take $M = \tilde{M}$ and equal to the maximum number of

elements of a partition in which each component ω_k is such that the change of ${\varepsilon_1}^k$ in ω_k can be detected by the instrument used to measure the mutual capacitances.

CONCLUSION

Obviously the function ${V_1}^{(i)}$ depends implicitly in a non-linear way on the unknown permittivity $\varepsilon_1(z)$, and, therefore, obtaining $\varepsilon_1(z)$ from the measurements (6) consists formally in solving the nonlinear equation

$$A(\varepsilon_1) = \vec{C}$$

where the operator A acts on a space of functions defined in Ω_1 and takes its values on $\Re^{N(N-1)/2}$.

The traditional general method to solve this kind of problem is to discretize the operator A into an operator $\tilde{A}$ acting in $\Re^{N(N-1)/2}$ on itself, and after that minimizing the functional

$$\left\| \tilde{A}(\tilde{\varepsilon}_1) - \vec{C}_{obs} \right\| \qquad (49)$$

on a set vectors $\tilde{\varepsilon}_1$ that correspond to an approximate representation of the function ε_1, and where $\vec{C}_{obs}$ corresponds to the vector of data obtained as a result of the measurement of the mutual capacitances and the formulas (7) and (8) [1, 7-10].

The problem with considering the functional (55) is that it is a nonlinear functional that can have multiple extreme points, and, according to (6), (7) and (8), in the definition of $\tilde{A}$ appears implicitly an ill-posedness due to the derivatives of ${V_2}^{(i)}$ involved in the definition of $\{c_{ij}\}$, and it is well known that the process of derivation is an ill-posed problem and numerically unstable.

To avoid these difficulties we have developed a new algorithm with the following advantages over the traditional method:

- The traditional method based on minimizing (49) is fully iterative and the forward problem must be solved on each iteration, while in the proposed direct method we have to do many of the calculations only one time, which should result in a more efficient algorithm.
- The proposed procedure allows calculating the accumulated error in each step more efficiently. This error is composed basically of the truncation error that we have when we approximate the series by a partial sum, plus the error coming from the measured data.
- In the non-linear functional case (49) the regularization techniques employed have an empiric character, because there are no theoretical results applicable. On the other hand, in the new method proposed we end up with a quadratic optimization problem with quadratic constraints, whose solution has been more thoroughly studied.

REFERENCES

1. M. Cheney, D. Isaacson, An overview of inversion algorithms for impedance imaging, *Contemporary Mathematics*, **122**, 29 (1991)

2. A. Fraguela, C. Gamio and D. Hinestroza, The inverse problem of electrical capacitance tomography and its application to gas-oil 2-phase flow imaging, *WSEAS Transactions on Systems*, **1**(2), 130 (2002)

3. E. A. Hammer and G. A. Johansen, Process tomography in the oil industry: state of the art and future possibilities, *Measurement and Control*, **30**, 212 (1997)

4. V. Isakov, On the uniqueness in the inverse conductivity problem with one boundary measurement, *Lectures in Applied Mathematics* (AMS), **30**, 105 (1994)

5. A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, AMS 120, Springer, New York, 1996

6. V. P. Mijailov, *Ecuaciones Diferenciales en Derivadas Parciales*, Mir-Moscow, 1982

7. D. M. Spink and J. M. Noras, Recent developments in the solution of the forward problem in capacitance tomography and implications for iterative reconstruction, *GPA*, India, 115 (1988).

8. R. Thorn, G. A. Johansen and E. A. Hammer, Recent developments in three-phase flow measurement, *Measurement Science and Technology*, **8**, 691 (1997)

9. C. G. Xie, A. Plaskoswki and M. S. Beck, 8-electrode capacitance system for two-component flow identification. Part 1: Tomographic flow imaging, *IEE Proceedings Pt. A*, **136**, 173 (1989)

10. W. Q. Yang, M. S. Beck and M. Byars, Electrical capacitance tomography: from design to applications, *Measurement and Control*, **28**, 261 (1995)

11. W. Q. Yang, A. L. Stott, M. S. Beck and C. G. Xie, Development of capacitance tomographic imaging systems for oil pipeline measurements, *Review of Scientific Instruments*, **66**, 4326 (1995)