ICIPE 2014

8th international conference on inverse problems in engineering

Kraków, Poland, 12–15 May, 2014

Conference Proceedings

Edited by:
Ireneusz Szczygieł, Andrzej J. Nowak, Marek Rojczyk

8th International Conference on Inverse Problems in Engineering

Conference Proceedings

Edited by: **Ireneusz Szczygieł, Andrzej J. Nowak, Marek Rojczyk**

Institute of Thermal Technology
Silesian University of Technology

May 12–15, 2014
Gliwice–Krakow, Poland

Edited by:
Ireneusz Szczygieł, Andrzej J. Nowak, Marek Rojczyk
Institute of Thermal Technology
Silesian University of Technology
ul. Konarskiego 22
44-100 Gliwice, Poland
e-mail: Ireneusz.Szczygiel@polsl.pl
e-mail: Andrzej.J.Nowak@polsl.pl
e-mail: Marek.Rojczyk@polsl.pl

Organized by:

Institute of Thermal Technology
Silesian University of Technology
ul. Konarskiego 22
44-100 Gliwice, Poland
http://www.itc.polsl.pl

Under auspices of:

The Committee of Thermodynamics and Combustion
Polish Academy of Sciences
http://www.ktis.pan.pl

Cover design: Vero Graphic Design Weronika Hluchnik http://www.vgd.com.pl
Computer typesetting: Ziemowit Ostrowski (LaTeX2e) based on the camera-ready contributions provided by authors.

ISBN 978-83-61506-23-2

Seminar Chairmen:

Irenusz Szczygieł – Conference Chair, Silesian University of Technology, Gliwice (Poland)
Andrzej J. Nowak – Conference Co-Chair, Silesian University of Technology, Gliwice (Poland)
Keith A. Woodbury – ICIPE Steering Committee, The University of Alabama, Tuscaloosa, AL (USA)

Scientific Committee:

Oleg Alifanov (Russia), Eugene Artiyoukhine (France), Czél Balázs (Hungary), Jean Luc Battaglia (France), James Beck (USA), Ryszard Białecki (Poland), Marc Bonnet (France), Cara Brooks (USA), Tadeusz Burczynski (Poland), John Cannon (USA), Alan Miquel Cerrolaza (Venezuela), Jeongeung Choi (USA), Marcelo Colaco (Brazil), Kevin Cole (USA), Andrei Constantinescu (France), Renato Cotta (Brazil), Fillipo de Monte (Italy), Brian Dennis (USA), Ingham Derek (UK), Eduardo Divo (USA), Kirk Dolan (USA), Kevin Dowding (USA), George Dulikravich (USA), Ashley Emery (USA), Adam Fic (Poland), Colin Fox (New Zeland), Jay Frankel (USA), Brizeida Gamez (Venezuela), Gilmar Guimaraes (Brazil), Bojan Guzina (USA), Alemdar Hasanov (Turkey), Cheng-Hung Huang (Taiwan), Yvon Jarny (France), Alain Kassab (USA), Alexander Katsevich (USA), Patricia Lamm (USA), German Larrazabal (Venezuela), Antonio Leitao (Brazil), Daniel Lesnic (UK), Denis Maillet (France), Toshiro Matsumoto (Japan), Robert McMasters (USA), Faissal Moslehy (USA), Zuhair Nashed (USA), A.V. Nenarokomov (Russia), Antonio Silva Neto (Brazil), Andrzej J. Nowak (Poland), David Ojeda (Venezuela), Helcio R. B. Orlande (Brazil), Ziemowit Ostrowski (Poland), Darrell Pepper (USA), Nilson Roberty (Brasil), Keith A. Woodbury (USA), Luiz Wrobel (UK), Nicholas Zabaras (USA).

Organizing Committee:

Irenusz Szczygieł – Conference Chair (Poland)
Andrzej J. Nowak – Conference Co-Chair (Poland)
Keith A. Woodbury – ICIPE Steering Committee (USA)
Marek Rojczyk – Conference Secretary (Poland)
Ryszard Białecki – Member (Poland)
Andrzej Szlęk – Member (Poland)
Ziemowit Ostrowski – Member (Poland)
Gabriel Węcel – Member (Poland)

History of the Conferences

During the years that the international conference (*International Conference on Inverse Problems in Engineering - ICIPE*) does not take place, informal two-day seminars (*Inverse Problem Symposia - IPS*) have been held at various universities in the USA. The sites and associated organizers are listed for purposes of archiving the historical continuity of these meetings.

Previous international meetings (International Conferences on Inverse Problems in Science and Engineering – ICIPE) were held at Palm Coast, FL, in 1993, at Le Croisic, France, in 1996, at Port Ludlow, WA, in 1999, at Angra dos Reis, Brazil, in 2002, at Claire College, Cambridge, England, in 2005, at Dourdan, France, in 2008 and at Orlando, FL, in 2011.

These international conferences grew out of several informal meetings held in the years prior to the Palm Coast conference, which focused mainly on inverse problems in heat transfer. Nicholas Zabaras broadened the scope of the Conference in the first international conference in 1993, especially into the field of solid mechanics. This conference was successful in attracting participation from the mechanics group.

Sites and organizers for the International Conferences and Symposia

Year	Venue	Chair
1988 – 91	Michigan State University, East Lansing, Michigan, USA	James V. Beck
1992	Michigan State University, East Lansing, Michigan, USA	Keith A. Woodbury
1993	*Palm Coast, Florida, USA*	*Nicholas Zabaras*
1994	University of Cincinnati, Cincinnati, Ohio, USA	Diego Murio
1995	Ohio State University, Columbus, Ohio, USA	Hank Busby
1996	*Le Croisic, France*	*Didier Delaunay*
1997	Rensselaer Polytechnic Institute, Troy, New York, USA	Antoinette Maniatty
1998	Ball State University, Muncie, Indiana, USA	Lija Guo
1999	*Port Ludlow, Washington, USA*	*Keith A. Woodbury*
2000	Texas A&M University, College Station, Texas, USA	Ted Watson
2001	University of Texas – Arlington, Arlington, Texas, USA	A. Hadji-Sheikh
2002	*Angra dos Reis, Brazil*	*Helcio R.B. Orlande*
2003	The University of Alabama, Tuscaloosa, AL, USA	Keith A. Woodbury
2004	University of Cincinnati, Cincinnati, Ohio, USA	Diego Murio
2005	*Claire College, Cambridge, England*	*Daniel Lesnic*
2006	Iowa State University, Ames, Iowa, USA	Mark Bryden
2007	Michigan State University, East Lansing, Michigan, USA	Neil T. Wright
2008	*Dourdan, France*	*Denis Maillet*
2009	Michigan State University, East Lansing, Michigan, USA	Kirk Dolan
2011	*University of Central Florida, USA*	*Alain Kassab*
2014	*Silesian University of Technology, Poland*	*Ireneusz Szczygieł*

(International Conference on Inverse Problems in Engineering – ICIPE are *italicized*.)

Preface

This is the 25th in the series of national and international meetings on Inverse Problems that were initiated at Michigan State University in 1988. The 2014 edition is the 8th International Conference. Both the Scientific Committee Members and the Participants come from many countries all over the world: USA, Brazil, Russia, Poland, France among others.

The primary purpose of the Conference is to provide a forum of scientist and graduate students in sciences and engineering to present recent results of the inverse problems.

The very important objective is also to exchange the knowledge and experience to strengthen existing but also start new international scientific cooperation.

The problems coped with involve theory, applications and, of course, the combination of both. The topics of the Conference cover inverse problems in all branches of science and engineering including thermal sciences, structure mechanics, fluid flows, medical applications and many other. In particular, the covered fields of interest can be summarized as:

- Statistical and Probabilistic Methods
- Identification in Nonlinear Differential Equations
- Regularization Techniques
- Design and Shape Optimization
- Inverse Scattering and Time Reversal
- Determination of Boundary and Initial Conditions
- Computational Methods
- Identifiability Concepts

ICIPE 2014 lasted four days: from May, 12th to May 15th in the former capital of Poland: Krakow. Krakow is the second largest and one of the oldest cities in Poland. Situated on the Vistula River in the Lesser Poland region, the city dates back to the 7th century. Krakow has traditionally been one of the leading centers of Polish academic, cultural, and artistic life and is one of Poland's the most important economic hubs. It was the capital of Poland from 1038 to 1569.

The editors would like to thank the distinguished scientists who accepted the invitation to deliver Plenary Lectures. The list of Keynote Lectures (in the order of presentations) is given below:

1. Prof. James V. Beck, Prof. Em., Michigan State University, East Lansing, MI, USA, *Inverse Heat Conduction Problem: Insights from scaled sensitivity coefficients, digital filter coefficients and intrinsic verification*
2. Prof. Oleg M. Alifanov, Russia, *Inverse Problems in Identification and Modeling of Thermal Processes: Theory and Practice*
3. Prof. Helcio R. B. Orlande, Ph.D., Federal University of Rio de Janeiro, Brazil, *Application of particle filters to some inverse problems in biomedical engineering*

4. Prof. Jan Taler, Ph.D., DSc, Krakow University of Technology, Poland, *Monitoring of thermal stresses in pressure components using inverse heat conduction method*

5. Prof. Assad Oberai, Ph.D., Rensselaer Polytechnic Institute, NY, USA, *Biomechanical Imaging: Theory, Practice and Applications*

6. Prof. Aleksey V. Nenarokomov, Ph.D., Moscow Aviation Institute University of Aerospace Technology, Russia, *Optimal experiment design with application to inverse problems*

We also would like to thank all members of Scientific Committee for the effort with the valuable paper reviews as well as thank the local organising committee for the hard work which eventually led to a successful and fruitful meetings.

Ireneusz Szczygieł, Andrzej J. Nowak and Marek Rojczyk,

Gliwice – Krakow, Poland, May 2014

Contents

Full length papers

Full length papers

ESTIMATING OF EXTERNAL HEAT FLUX FOR ABLATIVE THERMAL PROTECTION OF SPACECRAFT BY INVERSE PROBLEMS TECHNIQUE

O.M. Alifanov[1], S. A. Budnik[1], A.V. Nenarokomov[1], A.V. Netelev[1], V.M.Ydin[2]
[1] *Moscow Aviation Institute, Dept. of Space System Engineering*
4 Volokolamskoe Hgw., Moscow, 125993, Russia
aleksey.nenarokomov@mai.ru
[2] *Central Aerohydrodynamic Institute*
1 Zhukovsky Street, TsAGI Zhukovsky, Moscow Region, 140180,Russian Federation

Key words: inverse problems, thermokinetic, heat shield, ablative thermal protection, diagnostic heat flux

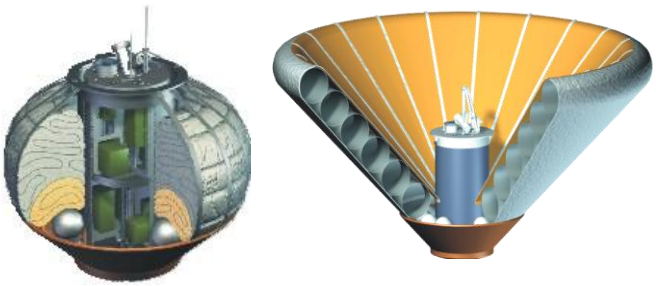
Abstract
The objective of this paper is to estimate heat fluxes on the surface of advanced materials with known thermal and thermokinetic properties using the approach based on inverse methods. In many practical situations it is impossible to measure directly heat fluxes on the surfaces of analyzed composite structures (in particular a thermal protection systems (TPS) of spacecraft) especially in the case of thermokinetic processes inside materials. Presently, heat and mass transfer in ablative materials are being studied intensively. This interest is associated with the important practical applications of the results of these investigations in aerospace technology, chemical technology, and other fields. But traditional measurements of the heat flux made to solve this problem are currently based on the use of a calorimetric method or the thin-wall method. However, the basic deficiency of these methods in the considered conditions is the requirement of conservation of the calorimeter mass or the wall thickness in the course of the experiment. This significantly limits the measurement time, since with the prolonged action of a flow on a calorimeter or a thin wall ablation breakdown begins. As a result, these methods cannot easily take account of the influence on the heat transfer of factors such as internal decomposing of the materials. The sufficiently effective way that can be used to overcome these difficulties is the solution of boundary inverse heat transfer problems. By solving such inverse problems, the boundary conditions and unsteady temperature field are reconstructed from interior temperature measurements in solids. Such problems are ill-posed in mathematical sense and their main feature shows itself in the solution instabilities. The general method of iterative regularization is concerned with application to the estimation of external heat flux.

1. Introduction

Moving in the planets atmosphere re-entry vehicles are influenced by large force and thermal loadings caused by gas approach flow. Frontal aerodynamic shields are used for RV protection from these loadings and its effective braking; a level of loadings depends on shields lateral dimensions. A trade-off between a decrease of a level of force and thermal loadings on RV by increasing RV's lateral dimension and limitations of launch vehicle payload fairing cross section could be a frontal shield of opening construction.

One of the options of this aerodynamic shield could be an inflatable shield (Fig. 1). In general this shield is a closed air-tight casing (or several casings) forming selected aerodynamic shape after filling it with a gas. This casing connects to a rigid frontal element of the aerodynamic shield which forms a RV's inflatable brake mechanism (IRDT). The object is placed at the rigid frontal element designed for descending in the atmosphere. At the orbital injection phase IRDT is placed under the

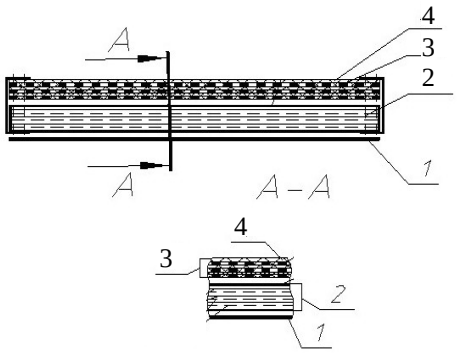
fairing in the compact volume folded position (Fig. 1a), and right before the aerodynamic deceleration phase, i.e. before the entry to the atmosphere, is brought into a state of operating unfolded position (Fig. 1b). Flexible thermal protective coating (FTPC) is used for the protection of air-tight inflatable casing, with its material destruction temperature around 500K. Principle scheme of FTPC is shown at the Fig. 2. Coating consists of two multilayered packages – one external thermal protective and one internal thermal insulting. The external layer consists of silica fabric with sublimating polymer material.



а) IRDT
in a folded position

б) IRDT
in an unfolded position

Figure 1: RV with inflatable brake mechanism



1 – air-tight casing,
2 – heat insulting package,
3 – thermal protective package,
4 – sublimating material layer.

Figure 2: Scheme of IRDT's flexible thermal protective coating

To provide mathematical simulation of IRDT the following initial data are nesessary: heat flux $q_{wi}(\tau)$ and/or external temperature variation $T_{wi}(\tau)$ at selected points at the surface. These measurements could be done at different TPC development phases during laboratory, benchmark, full-scale (large-scale) and flight experiments and tests.

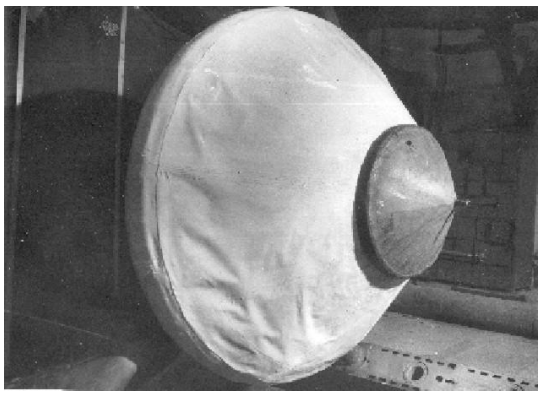
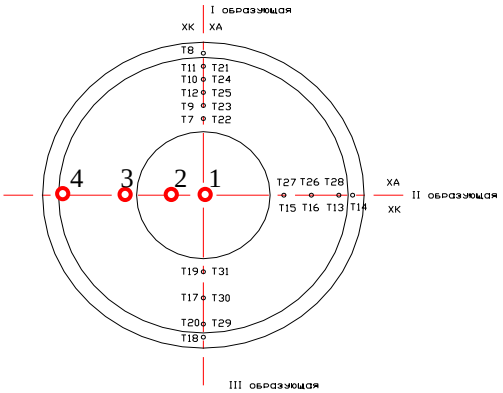


Figure3 RV's mockup with the IRDT at the thermal stand



- selected points, - measurement points,

Figure 4 The scheme of suggested measurement points placing at the IRDT surface

Photo of RV's mockup with the IRDT installed at the test stand is presented in Fig. 3 and the scheme of suggested measurement points placing at the IRDT surface during benchmark tests is shown at the Fig 4.

Results of simulation (at the points presented in Fig. 4) of a heat flux at the RV's surface with the mass 138.5 kg and IRDT diameter 2.3 m for nominal model of the Earth atmosphere are presented in Fig. 5.

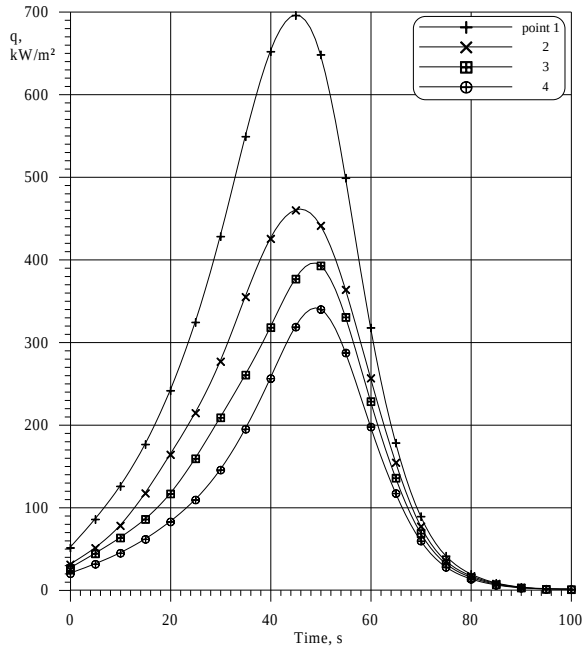


Figure 5: Heat flux at selected points at the RV's surface

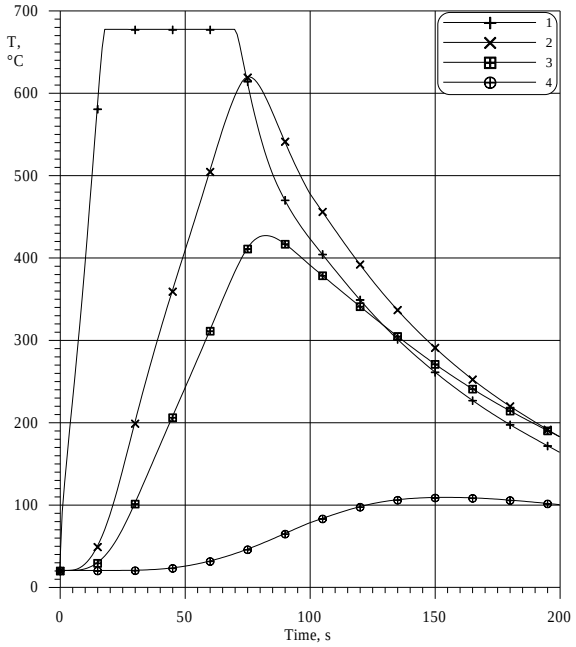
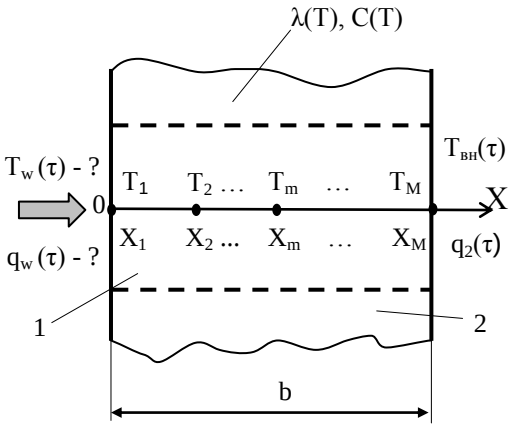


Figure 6: Temperature at selected points on the thermal protection thickness for the point 3at the RV's surface

Results of calculation of FTPC heating at the Point 3at the RV's surface at selected points on coating thickness (Fig. 2) are presented in Fig. 6.

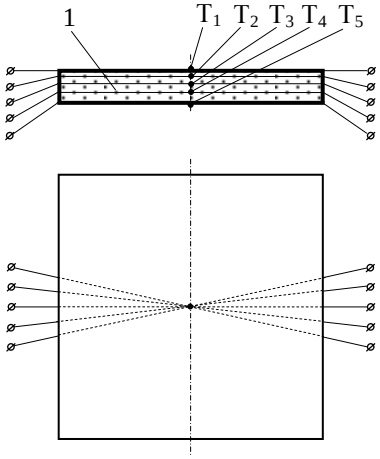
2. Heat flux sensor development

The experimental specimen of thermal protection coating is proposed as multilayer slab in the rectangular parallelepiped shape made of the regular material (Fig. 7). The choice of specimen size is depended on hardware opportunities. The important parameter is the ratio of specimen thickness to its length and width, which provides he realization accuracy of one-dimensional temperature field in the line of axis. That's why, for example, the ratio 1:10 – 1:15 is recommended for materials with low thermal conductivity. The scheme of thermal testing of the specimen with a sensor is shown at the Fig. 7. The prototype of sensors for IRDT based on sublimating polymer material were developed and manufactured for experimental verification of sensor structure and temperature measurement processing methods at the installation TVS-2M (MAI). Sensor scheme is shown in Fig. 8.



1 – sensor, 2 – experimental specimen, b – specimen thickness

Figure 7: Heat flux sensor



1 –sensor, $T_1 \ldots T_5$ - thermocouples

Figure 8: Scheme of sensor prototype for coating based on sublimating polymer

Special device was developed at MAI for installation of internal thermocouples X_2, X_3 and X_4 into sensors during the process of thermocouples manufacturing. Thermocouples at the heated (X_0) and reverse (X_5) surfaces of prototypes of sensors installed during the preparation of thermal tests during the integration of experimental sample in the experimental module.Photos of some prototypes of sensors with installed inside internal thermocouples are presented in Fig. 9 –11.

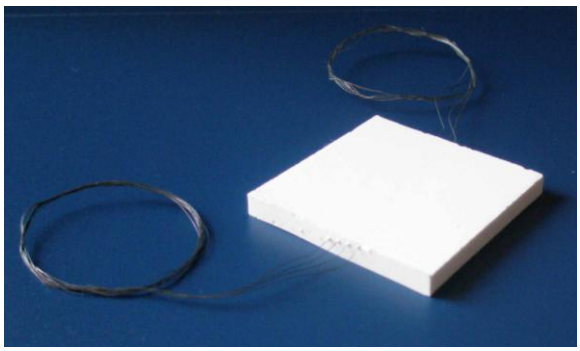


Figure 9: Prototype of heat flux sensor for IRDT based on sublimating polymer

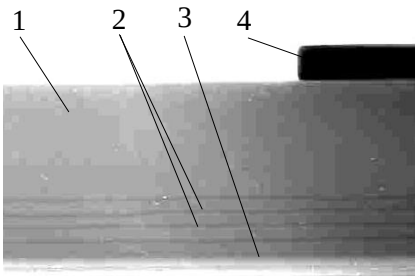


Figure 10: – X-ray picture of central part of prototype of sensor

The symmetric scheme of contact heating of two samples [11] was used in tests with the help of flat heating element (HE). The scheme is shown in Fig. 11. In this scheme the protective frames of sensor samples A and B made of sublimating polymer simulate the thermal-protective covering in which the sensors are installed. The stainless steel foil was used as of HE material with size (length×width×thickness) 120×80×0,1mm. The size of operating zone of HE equals to 80×60mm. The back surfaces of sensors and thermal protective covering are covered by thermal-isolated slab made from material which thermal conductivity coefficient are much more less than investigated sensors

have. It allows to model the presence of thermal-isolating under the external covering layer in consistent of real multilayered flexible thermal protective covering.

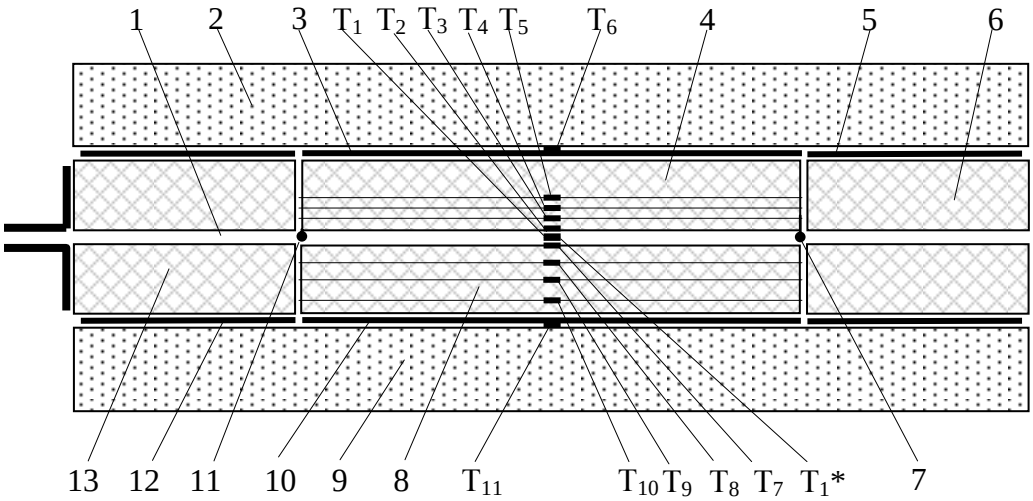


Fig. 11**:** The scheme of thermal test: 1 – heating element (HE), 2 – thermal-isolated slab, 3 – thin copper slab on the sample A, 4 – sample A (upper), 5 – frame of upper copper slab, 6 – protective frame of sample A, 7 – point of voltage measurement, 8 – sample B (lower), 9 – thermal-isolated slab, 10 – thin copper slab on the sample B, 11- point of voltage measurement, 12 – fram of lower thin copper slab, 13 – protective frame of sample B. T_1 - thermocouple on the HE, T_1* - the reserve thermocouple on the HE, T_2 and T_7 - thermocouples on the heating surfaces of A and B samples respectively, $T_3 \div T_5$ – inner thermocouples of sample A, T_6 and T_{11} – thermocouples on thin copper slabs, $T_8 \div T_{10}$ –inner thermocouples of sample B.

The construction and technological peculiarities of flux thermal sensor prototype designed for flexible Thermal Protection based on sublimating polymer as well as samples of sensors prototypes designed for thermal tests and methodology of thermal tests executing were detailed considered in [14]. In this paper the results of sensor prototype study in condition of thermal tests with vacuum heating (1×10^{-3}bar) and heating rate 32,7K/sec within temperature range 300-925K are presented. The vacuum installation TVS-2M developed for destructive materials of Thermal Protection was employed. The prototype of heat flux sensors were placed in the experiment module EM-3D (Fig. 12-13).

The two couples of experimental samples of sensors prototypes D_{1A}, D_{1B} and D_{2A}, D_{2B} based on sublimating polymer were manufactured for execution of thermal tests. The two first sensors were used for pilot tests execution and two second sensors were used for optional tests execution. The sensors had form of rectangular parallelepiped with size (length×width×thickness) 60×60×6mm (Fig 11).

Table 1: The real coordinates of measurements in sensors D_{2A} and D_{2B}.

Sensor	x_1,[mm]	x_2, [mm]	x_3, [mm]	x_4, [mm]	x_5, [mm]
D_{2A}	0,0	$1,10^{\pm0,1}$	$1,70^{\pm0,05}$	$2,20^{\pm0,05}$	$6,20^{-0,1}$
D_{2B}	0,0	$2,30^{-0,1}$	$3,60^{-0,2}$	$5,00^{-0,2}$	$6,10^{\pm0,05}$

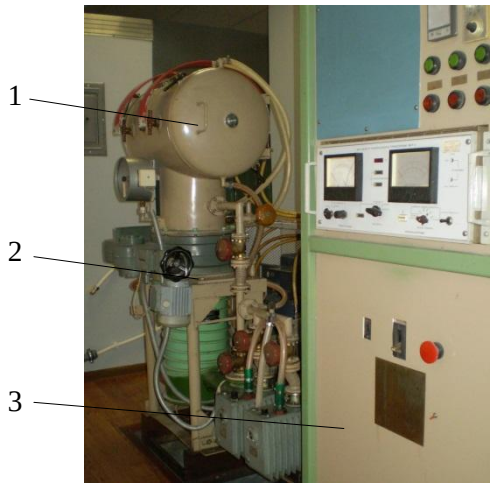


Figure 12: The general view of stand TVS-2M: 1 – vacuum chamber, 3 – aggregates of vacuum system, 2 – manipulating and control rack.

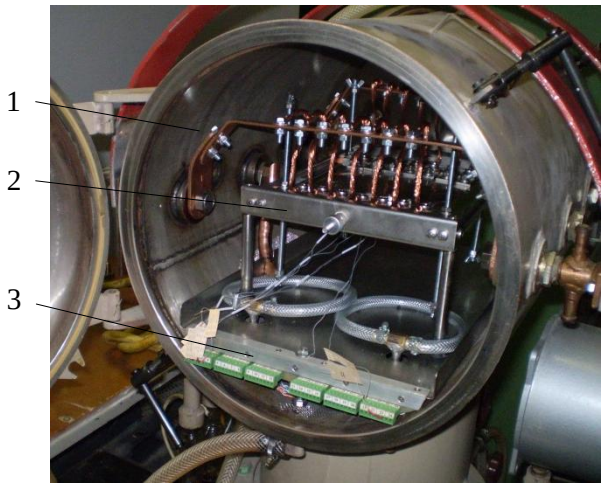


Figure 13: Experimental module EM-3D installed in vacuum chamber of stand TVS-2M: 1 – vacuum chamber, 2 – module EM-3D, 3 – connector blocks of thermocouples

3. Innverse problem algorithm

The mathematical model of heat transfer in considered layer of Thermal Protection with thremokinetics is

$$C\left(T(\tau,x)\right)\rho\frac{\partial T(\tau,x)}{\partial\tau}=\frac{\partial}{\partial x}\left(\lambda\left(T(\tau,x)\right)\frac{\partial T(\tau,x)}{\partial x}\right)+$$
$$+C_g\left(T(\tau,x)\right)\int_{x_н}^{x}\frac{\partial\rho(\xi,\tau)}{\partial\tau}d\xi\,\frac{\partial T(\tau,x)}{\partial x}+H\left(T(\tau,x)\right)\frac{\partial\rho(x,\tau)}{\partial\tau} \quad (1)$$

$$(x,\tau)\in Q=(0,L)\times(0,\tau_m]$$

$$T(0,x)=T_0(x),\ 0<x<L, \quad (2)$$

$$-\alpha_1\lambda\left(T(0,\tau)\right)\frac{\partial T(0,\tau)}{\partial x}+\beta_1 T(0,\tau)=q_1\left(T(0,\tau),\tau\right) \quad (3)$$

$$-\alpha_2\lambda\left(T(L,\tau)\right)\frac{\partial T(L,\tau)}{\partial x}+\beta_2 T(L,\tau)=q_2\left(T(L,\tau),\tau\right) \quad (4)$$

$$\rho(x,\tau_r)=\rho_0 \quad (5)$$

$$\rho(x,\tau_c)=\rho_c \quad (6)$$

$$\frac{\partial\rho(x,\tau)}{\partial\tau}=\begin{cases}0, T(x,\tau)<T_r\\ -\rho^n A\exp\left(\dfrac{-E}{RT(x,\tau)}\right), \rho(x,\tau)>\rho_c, T(x,\tau)\ge T_r\\ 0, \rho(x,\tau)\le\rho_c\end{cases} \quad (7)$$

where heat flux $q_1(\tau)$ is desired function of time. The experimental equipment and the method described below could be applied for estimating $q_1(\tau)$.

The results of temperature measurements inside the specimen are assigned as necessary additional information to solve an inverse problem

$$T^{\exp}(x_m, \tau) = f_m(\tau), \quad \mathrm{m} = \overline{1, \mathrm{M}} \tag{8}$$

where *M* is the number of installed thermocouples.

Let's introduce in the consideration a least-square discrepancy of the calculated (by model (1)-(7)) and experimental (8) temperature values in points of thermocouples positioning

$$J(q_1(\tau)) = J(\bar{p}) = \sum_{m=1}^{M} \int_{\tau_{\min}}^{\tau_{\max}} \left(T(x_m, \tau) - f_m(\tau)\right)^2 d\tau \tag{9}$$

It is assumed here that the conditions of uniqueness of the inverse problem solving are satisfied. An iterative regularization method stated here, according to general definition of Tikhonov's regularizing operators [9], gives rise to a regularizing set of operators, in which a regularization parameter is the number of the last iteration. For linear ill-posed problems the iterative regularization method has received a rigorous mathematical proves and practical verification through data of mathematical modeling. 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 [9-11], the unknown function can be determined through minimization of functional (9) by gradient methods of the first order prior to a fulfillment of the condition

$$\mathrm{J}(\bar{\mathrm{p}}) \le \delta_{\mathrm{f}} \tag{10}$$

where $\delta_f = \sum_{m=1}^{M} \int_{\tau_{\min}}^{\tau_{\max}} \sigma_m(\tau)\, d\tau$ - integral error of temperature measurements $f_m(\tau)$, $m=1,M$, and $\sigma_m(\tau)$ - 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 $\overline{p^0}$ is set

2) a value of the unknown vector at the next iteration are calculated

$$\begin{gathered} \overline{p}^{s+1} = \overline{p}^{s} + \gamma^{s} \overline{g}^{s} \\ \overline{g}^{s} = -\overline{J'}^{s} + \beta^{s} \overline{g}^{s-1} \\ \beta^0 = 0, \\ \beta^{\mathrm{s}} = \left\langle \left(\overline{J}_p'^{(s)} - \overline{J}_p'^{(s-1)}\right), \overline{J}_p'^{(s)} \right\rangle_{R^{N_P}} \Big/ \left\| \overline{J}_p'^{(s)} \right\|_{R^{N_P}} \end{gathered} \tag{11}$$

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 has been obtained

$$J'_{q_2} = -\psi_1(x_0, \tau) \tag{12}$$

where $\psi(x,\tau)$ - solution of a boundary-value problem adjoint to a linearized form of the initial problem (1)-(8).
To calculate the descent step a linear estimation [10, 12] is used:

$$\gamma^s = \left(\sum_{m=1}^{M} \int_{\tau_{min}}^{\tau_{max}} (T(x_m,\tau) - f_m(\tau)) \Delta T(x_m,\tau) d\tau \right) / \left(\sum_{m=1}^{M} \int_{\tau_{min}}^{\tau_{max}} \Delta T^2(x_m,\tau) d\tau \right) \tag{13}$$

where the boundary-value problem for the Freshet differential of $T(x,\tau)$ at the point $\{ q_1(\tau) \}$ (noted as $\Delta T(x,\tau)$).

4. Testing esting of heat flux sensor

The tests were carried out in the air in condition of decreased pressure 1×10^{-3} bar into vacuum chamber. During the realization of heating manipulation the indicated value of thermocouple T_1, installed on HE were used in capacity of feedback coupling. The pilot tests of sensor samples D_{1A} and D_{1B} were carried out for verification of heating manipulation subsystem work in real test condition, measurement and data collection and for selection of heating mode as well as the verification of chosen temperature measurement scheme. The optional tests of sensor samples D_{2A} (A - upper) and D_{2B} (B – lower) were prepared and carried out with taking into account results of pilot test. The heating program T_{pr} includes several stages:
primary stage with duration of 4 seconds and constant temperature value T_o = 298K, which equals to temperature into the vacuum chamber in the moment of tests beginning, which is necessary for manipulation system becoming on operative mode;
the first operating stage is heating by linear behavior from set temperature T_o to maximum temperature T_{max} = 950K with the given heating rate 32,7K/sec;
the second operating stage with constant temperature T_{max} = 950K during the 32sec;
the operating stage is decreasing of temperature from T_{max} = 950K to 350K.
The graph of heating program $T_{pr}(\tau)$ is presented in Fig. 14.

The measurements results, which were obtained during the thermal tests $T_m(\tau), 0 \le \tau \le \tau_e$ after primary data handling are given in Fig. 6-8.
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 the electric power released on HE can be written as

$$Q_{electr} = U * I \tag{14}$$

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 [11]

$$q_1(\tau) = Q_{electr} / (2A) = U * I / (2A) \tag{15}$$

where A - heater's operating surface area.
In the hole the carried out test approved efficiency and workability of developed technology and investigated sensor prototype for the solving thermal diagnostic problem (determining of

non-stationary temperature $T(x,\tau)$ and heat flux $q_1(\tau)$ of flexible thermal-protective covering based on sublimating polymer). The experimental data obtained from testing presented in previous subsection have been used to determine the heat flux on the heating surface of analyzed sensors. For this, the results of tests on the time uniform grid should be submitted as required by the software of inverse problem solution.

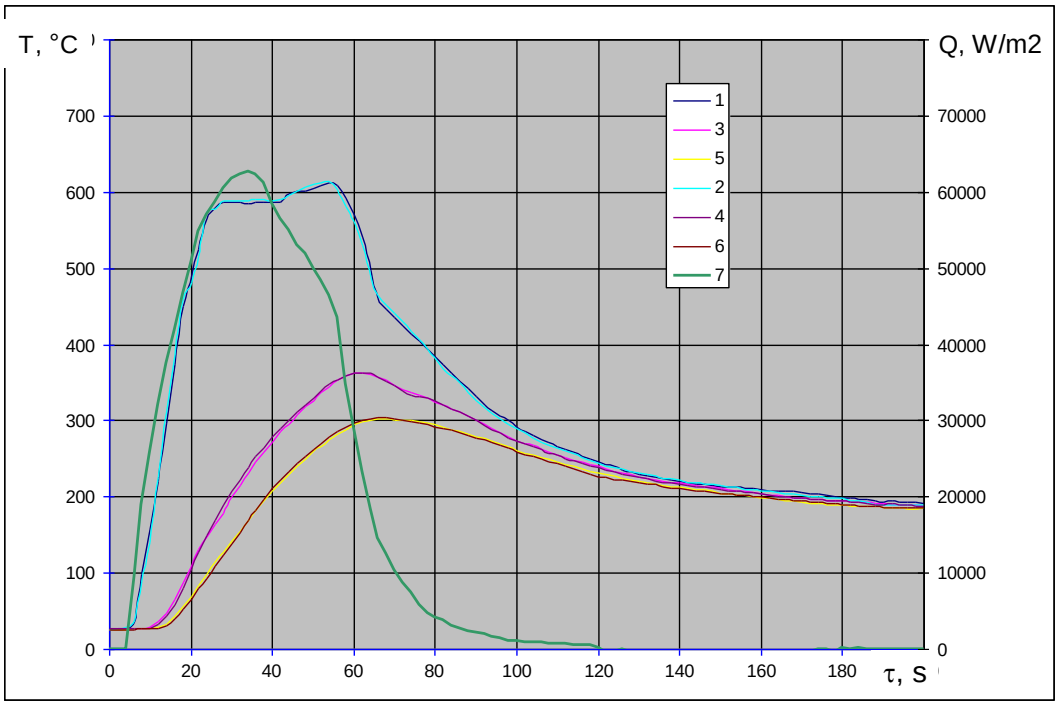


Figure. 15: Comparing of calculated in positions of T_2 (1), T_3 (3) and T_4 (5) and measured T_2 (2), T_3 (4) and T_4 (6) temperatures (sensor D_{2A}) and estimated values if heat flux (7).

To solve inverse problem in the pointed versions there was used software developed and presented above. All input data are presented in previous sections. So the comparing of calculated and measured temperatures on specimens' surfaces for sensors are presented in Fig. 14. Tab. 2 includes the obtained values of least square and maximum deviation of calculated temperatures from that measured in the experiments.

Table 2: The deviation of the calculated temperatures and measured temperatures for heat flux sensors

Sensors	Least-squares temperature deviation (K)	Maximum temperature deviation (K)
D_{2A}	12.7	19.1
D_{2B}	9.98	18.8

6. Conclusions

The paper seeks to describe the algorithm developed to process the data of unsteady-state thermal experiments for determining of the external heat flux. The algorithm suggested is based on the solving of the nonlinear inverse heat transfer problem in an extreme formulation. The following main factors

have an influence on the accuracy of the boundary inverse heat conduction problem (in sequence of significance): the errors in coordinates of thermosensor positions; the errors in values of a-priori known thermal properties of the sensor; the errors in estimating the residual level. For partially decomposed materials the model of heat transfer with temperature-dependent thermal characteristics and one-stage thermal kinetics is approximate, and characteristics are effective, since the heat transfer in such material is provided by heat conduction and a few different transformation processes depended on conditions of heating.

Acknowledgments

This work was done with the financial support of Russian Government (grant # NSh- 6343.2014.8) and RFBR.

References

[1] Hansel, J.G. and McAlevy, R.F., *Energetics and chemical kinetics of polystyrene surface degradation in inert and chemically reactive environments*., AIAA Journal, 4(5), pp. 1055-1066 ,(1966)
[2] Lundell, J.H., Dickey, R.R. and Jones J.W., *Performance of charring ablative materials in the diffusion - controlled Surface Combustion regime*, AIAA Journal, 6(6), pp. 1115-1124 , (1968)
[3] Pfal, R.C. Jr. and Mitchel, B.J., *Simultaneous measurement of six thermal properties of a charring ablator*, Int. Journal Heat and Mass Transfer, 13(2), pp.275-281.(1970)
[4] Potts, R.L., *Hybrid integral/quasi-steady solution of charring ablation*, AIAA Paper, 1667, 20p., (1990)
[5] Reinikka, E.A. and Wells, P.B., *Charring ablators in lifting reentry*, AIAA Paper. No.63-181, 15p.
[6] Vojvodich N.S. and Pope R.B. *Effect of gas composition on the ablation behavior of a charring material*, AIAA Journal, 2(3), pp.536-542.,(1964)
[7] Adarkar, D.B. and Hartsook, L.B., *An integral approach to transient charring ablator problems*, AIAA Journal, 4(12), pp.2246-2248, (1966)
[8] Clark, B.L., *A parametric study of the transient ablation of Teflon*, Trans. ASME. Journal of Heat Transfer, 94,pp.347-354, (1972)
[9] Alifanov, O.M., *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin \ Heidelberg, p. 348, (1994)
[10] Alifanov, O.M., Artyukhin, E.A. and Rumyantsev, S.V., *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Problems*, Begell House, New York\Wallinford(UK), p. 306., (1995)
[11] Artyukhin, E.A., Ivanov, G.A. and Nenarokomov, A.V., *Determination of a complex of materials thermophysical properties through data of nonstationary temperature measurements*, High Temperature, 31, pp. 235, (1993).
[12] Artyukhin, E.A. and Nenarokomov, A.V., *Coefficient Inverse heat conduction problem*, Journal of Engineering Physics, 53, pp. 1085, (1987).
[13] Alifanov, O.M., Budnik, S. A., Nenarokomov, A.V. and Netelev A.V., *Estimation of thermal properties of materials with application for inflatable spacecraft structures testing*, Inverse Problems in Science and Engineering, 20, pp.650-661, (2012)

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Experimental-computational system for non-contact diagnostics of elastic materials

Oleg M. Alifanov[1], Aleksey V. Nenarokomov[1], Kirill A. Nenarokomov[1], Anna V.Terentieva[1], Dmitry M. Titov[1], Valery S.Finchenko[2]

[1] Moscow Aviation Institute, Moscow, Russia
[2] Lavochkin Association, Khimki, Russia

KEY WORDS: Nonlinear Acoustics, Non-destructive Diagnostics of Defects, Parametric Identification, Conjugate Gradient Method.

Abstract
The purpose of this research is to develop experimental technique and setup for remote (non-contact) diagnostics of structural defects in elastic materials. The performance of equipment is based on nonlinear interaction of two acoustic beams of finite amplitude in investigated material. The most promising direction in further development of non-destructive diagnostics (research methods) for the elastic composite materials is to use the procedure of inverse problems [1,2]. Such problems are of great practical importance in the study of properties of composite materials used as non-destructive elastic surface coating in objects of space technology, power engineering etc The experimental equipment and the method developed could be applied for determination of material's properties; the availability of corresponded experimental facilities allows us to provide uniqueness of the solution. Elastic properties of flaws differ from ones of base medium. Gradient of elastics properties on the border of flaws results in the appearance of non-classic nonlinearity in the region. Such nonlinearity exceeds significantly the physical nonlinearity of a medium under diagnostics. Determination of spatial distribution of structural acoustic nonlinearity in the sample under investigation allows spotting the defects. In the paper the experimental technique for remote diagnostics of subsurface flaws in elastic materials is considered.

1. INTRODUCTION

The surface of re-entry vehicles is interacted with high-enthalpy gas flows, and its protection as well as breaking in the atmosphere is provided by aerodynamic screens of Thermal Protection Systems (TPS), and external heat fluxes determined by the diameters of such screens. Therefore it is necessary to find the compromise between the maximal diameters of TPS screens for re-entry stage and constrains of launch vehicles. At last time the transformed structures of re-entry space vehicles become more and more popular. And one of such recently developed approach is inflatable systems (Inflatable Re-entry Demonstrator Technology – IRDT, Fig.1**).**

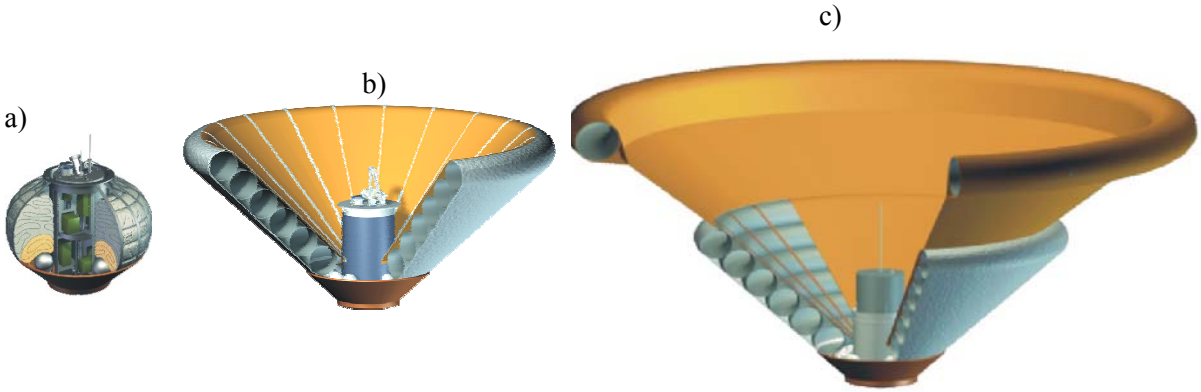


Figure 1: Re-entry vehicle with inflatable sructure: a - initial position, b - 1st stage, c – 2nd stage

There are a lot of possible applications of IRDT, and just one of them is presented in Fig.2.

The realization of IRDT concept should be provided by solving of some none-traditional technical problems:

- development of structure of flexible TPS for inflatable screen, provided its keeping in special container and transformation in functioning position (Fig.3);
- computational-experimental analisys of heat transfer in frexible TPS;
- development of supercharge system for inflatable screen;
- computational-experimental analysis of dynamics of transformation for re-entry flight;;
- analysis of stress and stability of IRDT;
- in-site diagnostics of covert defects of flexible TPS (delaminations, cracks, breakage, friction refinements, gapes, etc.) after testing (especially vibration testing (Fig.4).

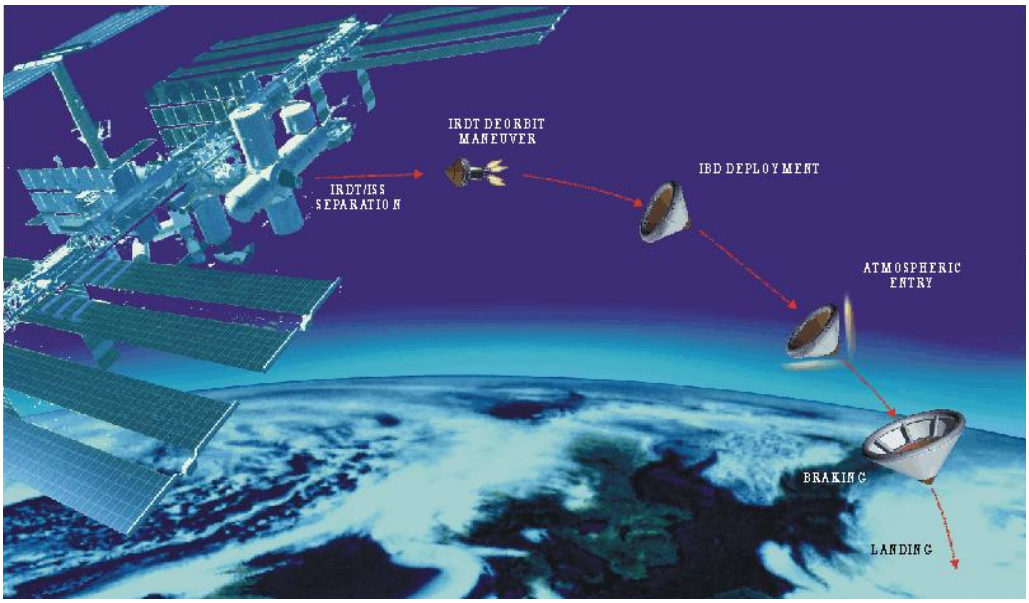


Figure 2: The Application of IRDT

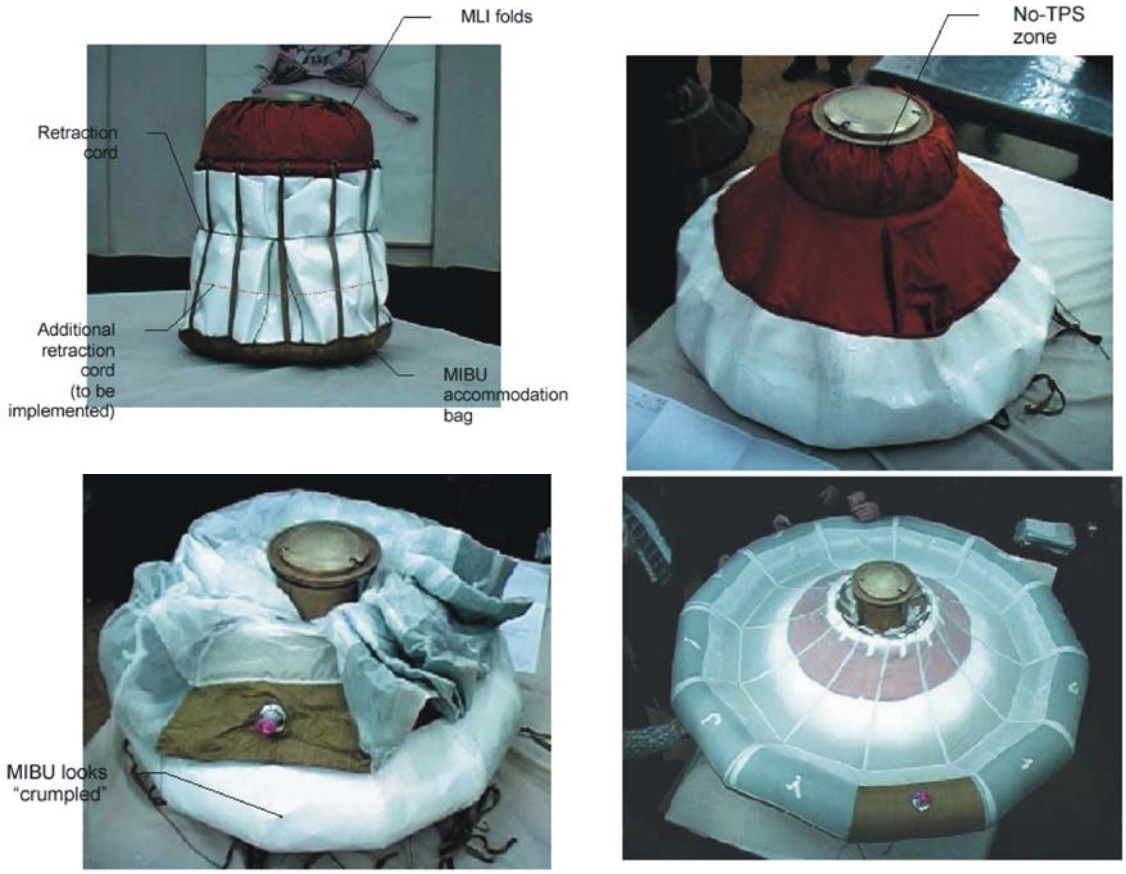


Figure 3: TPS structure testing a – initial position; b – transformation 1; c – transformation 2; d – final position

Figure 4: Vibtation tests

This paper is devoted to the last item in the list above. Enough sensitive nonlinear acoustics diagnostics methods are suitable for non-destructive diagnostics of defects. Sometimes the position of defects is required. The developed technique costrains the wave field so that defects can be located.
This paper has focused on development and analysis the nonlinear acoustics methods in order to make them suitable for the diagnostics of flexible TPS.

2. MATHEMATICAL MODEL

In terms of acoustics pressure the mathematical model of wave propagation in specimen of flexible TPS in general case is [3]

$$\varepsilon_l(p,x)\frac{\partial^2 p_l}{\partial \tau^2} = c_l(p,x)\frac{\partial^2 p_l}{\partial x^2} + a_l(p,x)\frac{\partial p_l}{\partial \tau} + b_l(p,x)\frac{\partial p_l}{\partial x} + S_l(p,x)$$

$$X_{l-1} < x < X_l, l = \overline{1,L},\ \ 0 < \tau \le \tau_{max} \tag{1}$$

$$p_l(x,0) = p_l^0(x),\ X_{l-1} \le x \le X_l,\ l = \overline{1,L} \tag{2}$$

$$\frac{\partial p_l}{\partial \tau}(x,0) = 0,\ X_{l-1} \le x \le X_l,\ l = \overline{1,L} \tag{3}$$

$$\alpha_1 \frac{\partial p_1(X_0,\tau)}{\partial x} + \beta_1 p_1(X_0,\tau) = v_1(p_1(X_0,\tau),\tau),\ 0 < \tau \le \tau_{\max} \tag{4}$$

$$w_l \frac{\partial p_l(X_l,\tau)}{\partial x} = w_{l+1}\frac{\partial p_{l+1}(X_l,\tau)}{\partial x},\ l = \overline{1,L-1},\ 0 < \tau \le \tau_{\max} \tag{5}$$

$$p_l(X_l,\tau) = p_{l+1}(X_l,\tau),\ l = \overline{1,L-1},\ 0 < \tau \le \tau_{\max} \tag{6}$$

$$\alpha_2 \frac{\partial p_L(X_L,\tau)}{\partial x} + \beta_2 p_L(X_L,\tau) = v_2(p_L(X_L,\tau),\tau),\ ,\ 0 < \tau \le \tau_{\max} \tag{7}$$

where coefficients $\varepsilon_l(p,x), a_l(p,x), b_l(p,x), c_l(p,x), S_l(p,x),\ l = \overline{1,L}$ can be unknown.

Let's introduce the following notation X_m is the coordinate of m^{th} pressure sensor installed on external boundary of system. And the results of pressure measurements in the specimen are assigned as necessary additional information to solve an inverse problem

$$p_{\exp}(X_m,\tau) = f_m(\tau),\ m = 1,L \tag{8}$$

With the presented statement of inverse problem, the data gained in one experiment are not sufficient for simultaneous recovery of all unknown coefficients. But bellow we presented the universal algorithm without analysis of uniqueness.

3. INVERSE PROBLEM ALGORITHM

There are a few approaches to realize inverse problems methods with application to acoustics [4-11]. Here the extreme method is used. Thus, the identification problem is formulated as funding the function $p(X_m,\tau)$ which provides a minimum of least-square objective function (residual functional of the calculated and experimental pressure in points of sensors positioning:

$$J = \sum_{m=1}^{L}\int_0^{\tau_m}\left[p(X_m,\tau) - f_m(\tau)\right]^2 d\tau \tag{9}$$

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 $\overline{u^0}$ is set

2) a value of the unknown vector at the next iteration are calculated

$$\begin{aligned} \overline{u}^{s+1} &= \overline{u}^{s} + \gamma^{s}\overline{g}^{s}, \quad s = 1,\ldots,s^{*}, \\ \overline{g}^{s} &= -J_{u}^{'} + \beta^{s}\overline{g}^{s-1}, \quad g^{0} = 0, \\ \beta^{0} &= 0, \\ \beta^{s} &= \left\langle \left(\overline{J}_{u}^{\prime(s)} - \overline{J}_{u}^{\prime(s-1)}\right), \overline{J}_{u}^{\prime(s)} \right\rangle_{U} \Big/ \left\| \overline{J}_{\varepsilon}^{\prime(s)} \right\|_{U}^{2} \end{aligned} \tag{10}$$

where $\overline{J}_{u}^{\prime(s)}$ - value of the functional gradient at current iteration.

To solve the minimization problem for residual functional a reformulation based on the following Lagrangian is necessary [2]

$$\begin{aligned} \Im = J &+ \sum_{l=1}^{L}\int_{0}^{\tau}\int_{x_l}^{x_{l+1}} \psi_l(x,\tau)\left[\varepsilon_l(p)\frac{\partial^2 p_l}{\partial\tau^2} + c_l(p)\frac{\partial^2 p_l}{\partial x^2} + a_l(p)\frac{\partial p_l}{\partial\tau} + b_l(p)\frac{\partial p_l}{\partial x} + S_l(p)\right]dx d\tau - \\ &+ \sum_{l=1}^{L}\int_{x_l}^{x_{l+1}} \eta_l(x,0)\left[p_l(x,0) - p_l^{0}(x)\right]dx + \sum_{l=1}^{L}\int_{x_l}^{x_{l+1}} \xi_l(x,0)\left[\frac{\partial p_l}{\partial\tau}(x,0)\right]dx + \\ &+ \int_{0}^{\tau_{\max}} \eta(X_0,\tau)\left[\alpha_1(p_1(X_0,\tau))\frac{\partial p_1(X_0,\tau)}{\partial x} + \beta_1 p_1(X_0,\tau) - q_1(p_1(X_0,\tau),\tau)\right]d\tau + \\ &+ \sum_{l=1}^{L}\int_{0}^{\tau_{\max}} \mu(X_l,\tau)\left[p_l(X_l,\tau) - p_{l+1}(X_l,\tau)\right]d\tau + \\ &+ \sum_{l=1}^{L-1}\int_{0}^{\tau_{\max}} \eta(X_l,\tau)\left[w_l\frac{\partial p_l(X_l,\tau)}{\partial x} - w_{l+1}\frac{\partial p_{l+1}(X_l,\tau)}{\partial x}\right]d\tau + \\ &+ \int_{0}^{\tau_{\max}} \eta(X_L,\tau)\left[\alpha_1\frac{\partial p_1(X_0,\tau)}{\partial x} + \beta_1 p_1(X_0,\tau) - v_1(p_1(X_0,\tau),\tau)\right]d\tau, \end{aligned} \tag{11}$$

where $\psi_l(x,\tau), \eta_l(x),\ l=\overline{1,L},\ \eta(X_l,\tau), l=\overline{0,L}, \mu(X_l,\tau), l=\overline{1,L-1}$ are the Lagrange multipliers. If the functions $\varepsilon_n(p,x), a_n(p,x), b_n(p,x), c_n(p,x), S_n(p,x)$ have received variations $\Delta\varepsilon_n, \Delta a_n, \Delta b_n, \Delta c_n, \Delta S_l$. Than pressure $p_l(x,\tau)$ will have variations $\vartheta_l(x,\tau),\ l=\overline{1,L}$, which are covered the following problem

$$\begin{aligned} &\varepsilon_l\frac{\partial^2\vartheta_l}{\partial\tau^2} = c_l\frac{\partial^2\vartheta_l}{\partial x^2} + a_l\frac{\partial\vartheta_l}{\partial\tau} + b_l\frac{\partial\vartheta_l}{\partial x} + S_l + \\ &+ \left[\frac{dc_l}{dP}\frac{\partial^2\vartheta_l}{\partial x^2} + \frac{da_l}{dP}\frac{\partial\vartheta_l}{\partial\tau} + \frac{db_l}{dP}\frac{\partial\vartheta_l}{\partial x} + \frac{dS_l}{dP} - \frac{\partial^2 P_l}{\partial\tau^2}\frac{d\varepsilon_l}{dP}\right]\vartheta_l + \\ &+ \delta_n^{l}\left[-\Delta\varepsilon_l(p)\frac{\partial^2 p_l}{\partial\tau^2} + \Delta c_l(p)\frac{\partial^2 p_l}{\partial x^2} + \Delta a_l(p)\frac{\partial p_l}{\partial\tau} + \Delta b_l(p)\frac{\partial p_l}{\partial x} + \Delta S_l(p)\right], \\ &X_{l-1} < x < X_l,\ 0 < \tau \le \tau_{\max},\ l=\overline{1,L}, \end{aligned} \tag{12}$$

$$\vartheta_l(x,0)=0,\ X_{l-1}\le x\le X_l,\ l=\overline{1,L}, \tag{13}$$

$$\frac{\partial\vartheta_l}{\partial t}(x,0)=0,\ X_{l-1}\le x\le X_l,\ l=\overline{1,L} \tag{14}$$

$$\alpha_1\frac{\partial\vartheta_1(X_0,\tau)}{\partial x}+\left(\beta_1-\frac{\partial v_1(p_1(X_0,\tau),\tau)}{\partial P}\right)\vartheta_1(X_0,\tau)=0,,\ 0<\tau\le\tau_{\max} \tag{15}$$

$$w_l\frac{\partial\vartheta_l(X_l,\tau)}{\partial x}=w_{l+1}\frac{\partial\vartheta_{l+1}(X_l,\tau)}{\partial x},\ l=\overline{1,L-1},\ ,\ 0<\tau\le\tau_{max} \tag{16}$$

$$\vartheta_l(X_l,\tau)=\vartheta_{l+1}(X_l,\tau),\ l=\overline{1,L-1},\ ,\ 0<\tau\le\tau_{max} \tag{17}$$

$$\alpha_2\frac{\partial\vartheta_L(X_L,\tau)}{\partial x}+\left(\beta_2-\frac{\partial v_2(p_L(X_L,\tau),\tau)}{\partial p}\right)\vartheta_L(X_L,\tau)=0,\ ,\ 0<\tau\le\tau_{\max} \tag{18}$$

where $\delta_n^l=\begin{cases}0,\ l\ne n\\ 1,\ l=n\end{cases}$.

Introducing the adjoint problem as:

$$\varepsilon_l\frac{\partial^2\psi_l}{\partial\tau^2}=c_l\frac{\partial^2\psi_l}{\partial x^2}+2\frac{\partial\psi_l}{\partial x}\frac{dc_l}{dp}\frac{\partial p_l}{\partial x}+2\frac{dc_l}{dp}\frac{\partial^2 p_l}{\partial x^2}\psi_l+\psi_l\frac{d^2c_l}{dp^2}\left(\frac{\partial p_l}{\partial x}\right)^2-2\frac{d\varepsilon_l}{dp}\frac{\partial^2 p_l}{\partial\tau^2}\psi_l$$
$$-2\frac{\partial\psi_l}{\partial\tau}\frac{d\varepsilon_l}{dp}\frac{\partial p_l}{\partial\tau}--\psi_l\frac{d^2\varepsilon_l}{dp^2}\left(\frac{\partial p_l}{\partial\tau}\right)^2-a_l\frac{\partial\psi_l}{\partial\tau}-b_l\frac{\partial\psi_l}{\partial\tau}+\frac{\partial S_l}{\partial p}\psi_l, \tag{19}$$

$X_{l-1}<x<X_l,\ 0<\tau\le\tau_{\max},\ l=\overline{1,L},$

$$\psi_l(x,\tau_{max})=0, \tag{20}$$
$$X_{l-1}\le x\le X_l,\ l=\overline{1,L}$$

$$\frac{\partial\psi}{\partial t}(x,\tau_{max})=0, \tag{21}$$
$$X_{l-1}\le x\le X_l,\ l=\overline{1,L}$$

$$\alpha_1c_1\frac{\partial\psi_1(X_0,\tau)}{\partial x}+\left[\beta_1c_l-\frac{\partial v_1(p_1(X_0,\tau),\tau)}{\partial p}c_1+\alpha_1\frac{dc_l}{dp}\frac{\partial p_1}{\partial x}-\alpha_1b_l\right]\psi_1(X_0,\tau)- \tag{22}$$
$$-2\alpha_1[p_1(X_0,\tau)-f_1(\tau)]=0,\ 0<\tau\le\tau_{\max}$$

$$\frac{c_l}{w_l}\psi_l(X_l,\tau)=\frac{c_{l+1}}{w_{l+1}}\psi_{l+1}(X_l,\tau),\ l=\overline{1,L-1},\ 0<\tau\le\tau_{\max} \tag{23}$$

$$c_l\frac{\partial\psi_l(X_l,\tau)}{\partial x}-c_{l+1}\frac{\partial\psi_{l+1}(X_l,\tau)}{\partial x}+\left[\frac{dc_l}{dp}\frac{\partial p_l}{\partial x}-b_l\right]\psi_l(X_l,\tau)- \tag{24}$$
$$-\left[\frac{dc_{l+1}}{dp}\frac{\partial p_{l+1}}{\partial x}-b_{l+1}\right]\psi_{l+1}(X_l,\tau)=0,\ l=\overline{1,L-1},\ 0<\tau\le\tau_{\max}$$

$$\alpha_2c_L\frac{\partial\psi_L(X_L,\tau)}{\partial x}+\left[\beta_2c_L-\frac{\partial v_2(p_L(X_L,\tau),\tau)}{\partial p}c_L+\alpha_2\frac{dc_L}{dp}\frac{\partial p_L}{\partial x}-\alpha_2b_L\right]\psi_L(X_L,\tau)- \tag{25}$$
$$-2\alpha_2[p_L(X_L,\tau)-f_L(\tau)]=0,\ 0<\tau\le\tau_{\max}$$

we receive the variation of residual functional as

$$\delta J = \int_0^{\tau_{\max}} \sum_{l=1}^{L} \int_{X_{l-1}}^{X_l} \psi_{l,m}(x,\tau)\vartheta_n^l \left[-\Delta\varepsilon_l(p)\frac{\partial^2 p_l}{\partial \tau^2} + \Delta c_l(p)\frac{\partial^2 p_l}{\partial x^2} + \Delta a_l(p)\frac{\partial p_l}{\partial \tau} + \Delta b_l(p)\frac{\partial p_l}{\partial x} + \Delta S_l(p) \right] dx d\tau \quad (26)$$

This formula can be considered as differential of residual functional, and in the case when unknown functions depend only on spatial coordinate x or/and time τ it can be transformed to gradient. If unknown functions depends on pressure (non-linear problem) it is convenient to introduce parameterization of unknown functions

$$\varepsilon_n(p) = \sum_{k=1}^{N_\varepsilon} \varepsilon_k \varphi_k(p) \quad a_n(p) = \sum_{k=1}^{N_a} a_k \varphi_k(p) \quad b_n(p) = \sum_{k=1}^{N_b} b_k \varphi_k(p)$$

$$c_n(p) = \sum_{k=1}^{N_c} c_k \varphi_k(p) \quad S_n(p) = \sum_{k=1}^{N_S} S_k \varphi_k(p) \quad (27)$$

where $\varepsilon_k, k = \overline{1, N_\varepsilon}; a_k, k = \overline{1, N_a}; b_k, k = \overline{1, N_b};\ c_k,\ k = \overline{1, N_c}; S_k, k = \overline{1, N_S}$ are unknown parameters, $\varphi_i(P),\ i = \overline{1,5}$ are systems of basic functions. Applying variations $\Delta u_i(p) = \sum_{k=1}^{N_i} \Delta u_{ik} \varphi_k(p),\ i = \overline{1,5}$ the gradient of residual functional is thus given by

$$\frac{\partial J}{\partial c_k} \equiv J'_{c_k} = \int_0^{\tau_{\max}} \int_{X_{n-1}}^{X_n} \psi_l(x,\tau)\frac{\partial^2 p_n}{\partial x^2}\varphi_k(p)\,dx\,d\tau$$

$$k = \overline{1, N_c}$$

$$\frac{\partial J}{\partial \varepsilon_k} \equiv J'_{\varepsilon_k} = -\int_0^{\tau_{\max}} \int_{X_{n-1}}^{X_n} \psi_n(x,\tau)\frac{\partial^2 p_n}{\partial \tau^2}\varphi_k(p)\,dx\,d\tau \quad (28)$$

$$k = \overline{1, N_\varepsilon}$$

The descent parameter γ^s is determined from the condition

$$\gamma_s = Arg \min_{\gamma \in R^1} J\left(u^s - \gamma g\left(J'^{(s)}\right)\right) \quad (29)$$

A linear estimation is used for determination of the descent step. It can be calculated as

$$\gamma_s = -\frac{\sum_{m=1,L} \int_0^{\tau_m} \left[p(X_m,\tau) - f_m(\tau)\right]\vartheta_m(X_m,\tau)\gamma_s g\left(J'^{(s)}\right) d\tau}{\sum_{m=1,L} \int_0^{\tau_m} \left[\vartheta_m(X_m,\tau)\gamma_s g\left(J'^{(s)}\right)\right]^2 d\tau} \quad (30)$$

where $\vartheta_{l,m}(x,\tau)$ is the Freshet derivative of $p(x,t)$ and is the solution of the

$$\varepsilon_l \frac{\partial^2 \vartheta_l}{\partial \tau^2} = c_l \frac{\partial^2 \vartheta_l}{\partial x^2} + a_l \frac{\partial \vartheta_l}{\partial \tau} + b_l \frac{\partial \vartheta_l}{\partial x} + S_l +$$
$$+ \left[\frac{dc_l}{dp} \frac{\partial^2 \vartheta_l}{\partial x^2} + \frac{da_l}{dp} \frac{\partial \vartheta_l}{\partial \tau} + \frac{db_l}{dp} \frac{\partial \vartheta_l}{\partial x} + \frac{dS_l}{dp} - \frac{\partial^2 p_l}{\partial \tau^2} \frac{d\varepsilon_l}{dp} \right] \vartheta_l +$$
$$+ \delta_n^l \left[- \frac{\partial^2 p_l}{\partial \tau^2} \gamma_s \sum_{k=1}^{N_e} G_k \varphi_k(p) + \frac{\partial^2 p_l}{\partial x^2} \gamma_s \sum_{k=1}^{N_c} G_k \varphi_k(p) + \frac{\partial p_l}{\partial \tau} \gamma_s \sum_{k=1}^{N_a} G_k \varphi_k(p) + \right. \quad (31)$$
$$\left. + \frac{\partial p_l}{\partial x} \gamma_s \sum_{k=1}^{N_b} G_k \varphi_k(p) + \gamma_s \sum_{k=1}^{N_S} G_k \varphi_k(p) \right],$$
$$X_{l-1} < x < X_l,\ 0 < \tau \le \tau_{\max},\ l = \overline{1, L},$$

$$\vartheta_l(x,0) = 0,\ X_{l-1} \le x \le X_l,\ l = \overline{1, L}, \quad (32)$$

$$\frac{\partial \vartheta_l}{\partial t}(x,0) = 0,\ X_{l-1} \le x \le X_l,\ l = \overline{1, L} \quad (33)$$

$$\alpha_1 \frac{\partial \vartheta_1(X_0, \tau)}{\partial x} + \left(\beta_1 - \frac{\partial v_1(p_1(X_0, \tau), \tau)}{\partial p} \right) \vartheta_1(X_0, \tau) = 0,\ 0 < \tau \le \tau_{\max} \quad (34)$$

$$w_l \frac{\partial \vartheta_l(X_l, \tau)}{\partial x} = w_{l+1} \frac{\partial \vartheta_{l+1}(X_l, \tau)}{\partial x},\ l = \overline{1, L-1},\ 0 < \tau \le \tau_{max} \quad (35)$$

$$\vartheta_l(X_l, \tau) = \vartheta_{l+1}(X_l, \tau),\ l = \overline{1, L-1},\ 0 < \tau \le \tau_{max} \quad (36)$$

$$\alpha_2 \frac{\partial \vartheta_L(X_L, \tau)}{\partial x} + \left(\beta_2 - \frac{\partial v_2(p_L(X_L, \tau), \tau)}{\partial p} \right) \vartheta_L(X_L, \tau) = 0,\ 0 < \tau \le \tau_{\max} \quad (37)$$

4. EXPERIMENTAL APPROVES

Nonlinearity of a material with damaged structure is the result of elastic properties changes on the border of inhomogenuity and surrounding medium. Knowing the value of signal amplitude on heterodyne frequency during scanning makes it possible to restore the spatial distribution of structural damages in the investigated material [4]. The scheme for the computer-controlled experimental setup is on Fig. 5.

The measurement of nonlinear response value dependence on coordinate allows to define spatial distribution of nonlinear parameter in the sample and to localize the damages of structure. Background noise measured on an object without structural damages is used as a zero signal value in all of the experiments. For testing measurements the sample of regular TPS was made. Its density was 800 kg/m^3. The sizes of the sample were 100x100x3 mm. The subsurface structural inclusions (inner cutting of different depth) were maden at internal surface of the sample and imitated the structural damage of a composite material after testing. Som results of corresponded inverse problem solving are presented in Fig.6.

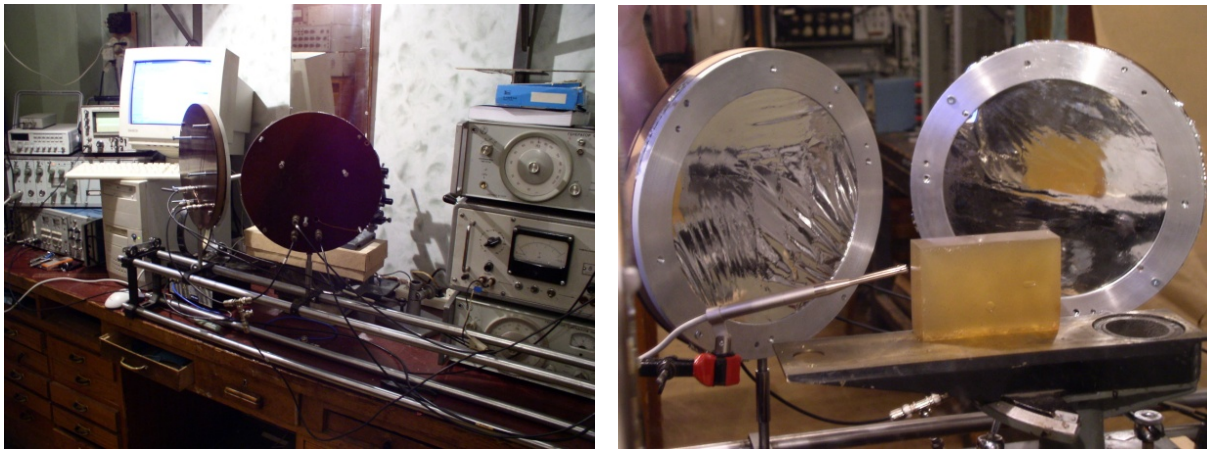


Figure 5: The experimental facilities

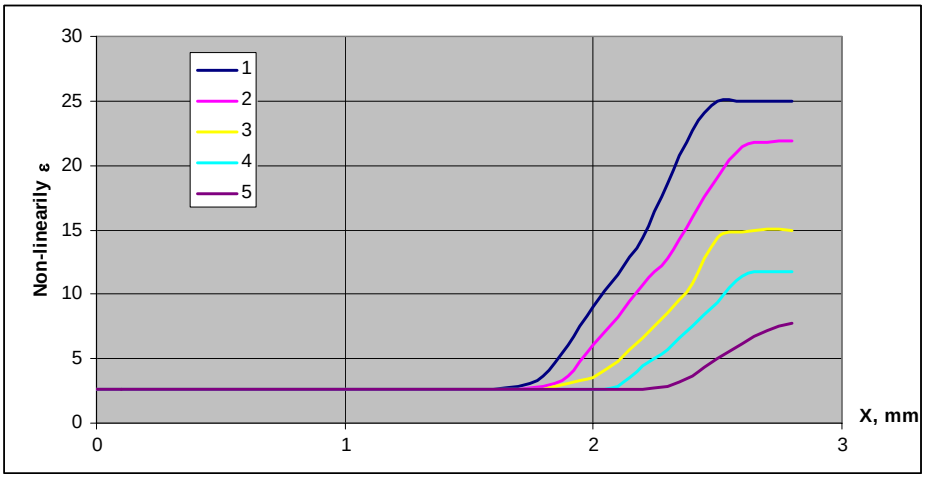


Figure 6: The non-linearity of the samples as function of cutting depth: 1 – depth is 1,5 mm, 2 – depth is 1.2 mm, 3 - depth is 1 mm, 4 – depth is 0.8 mm, 5 – depth is 0.6 mm

5. CONCLUSIONS

The experimental facility and corresponded software were developed, allowing with the high accuracy to define the defects in elastic TPS materials. The deviations of the calculated pressure (using $\varepsilon(x)$ estimations) from the pressure measured in the experiments are insignificant showing sufficient accuracy in the estimations of $\varepsilon(x)$ of the analyzed materials.

ACKNOWLEDGMENTS

The authors are grateful to the Russian Foundation for Basic Research for the financial support of this study. This work was also supported by the Russian president grant for Scientific Team NSh-6343.2014.8.

REFERENCES

[1] Alifanov, O.M., *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin, 1994.

[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, 1995.

[3] O.V. Rudenko, S.N. Soluyan, *Theoretical basics of nonlinear acoustics*. Nauka, 1975 (In Russian).

[4] Norton S.J. *Iterative Inverse Scattering Algorithms: Methods of Computing Frechet Derivatives*.- Journal of Acoustical Society of America, (1999), Vol.106, No. 5, pp. 2653-2660.

[5] Karim G.Sabra. Using Cross Correlations of Turbulent *Flow-induced Ambient Vibration to Estimate the Structural Impulse Response*. Application to Structural Health Monitoring.- Journal of Acoustical Society of America, (2007), Vol. 121, No. 4, pp.1987-1995

[6] B.M.Gibbs, R.Cookson, and N. Qi. *Vibration Activity and Mobility of Structurae-borne Sound Sources by a Reception Plate Method*.- Journal of Acoustical Society of America, (2008), Vol. 123, No. 6, pp. 4199-4209

[7] Anton I. Lavrentyev and Stanislav I. Rokhlin. *Determination of Elastic Module Density, Attenuation, and Thickness of a Layer Using Ultrasonic Spectroscopy at Two Angles*.- Journal of Acoustical Society of America, (1997), Vol. 102, No. 6, pp.3467-3477.

[8] Nicolas P. Valdivia, Earl G. Williams, and Peter C. Herdic. *Approximations of Inverse Boundary Element Methods with Partial Measurements of the Pressure Field*.- Journal of Acoustical Society of America, (2008), Vol. 123, No. 1, pp. 109-120.

[9] Daniel E. Astral and Mats G. Gustafsson, *Optimal Detection of Crack Echo Families in Elastic Solides*.- Journal of Acoustical Society of America, (2003), Vol. 113, No. 5, pp. 2732-2741.

[10] F. Simonetti and P. Cawley. *A Guided Wave Technique for the Characterization of Highly Attenuative Viscoelastic Materials*.- Journal of Acoustical Society of America, (2003), Vol. 114, No. 1, pp.158-165.

[11] J. Bucaro, A.J. Romano, *Detection and Localization of inclusions in Plates Using Inversion of Point Actuated Surface Displacement*.- Journal of Acoustical Society of America, (2004), Vol. 115, No. 1, pp. 201-206.

[12] Anthony J. Romano, Joseph A. Bucaro, Joseph F. Vignola, and Phillip B. Abraham, Detection and Localization of Rib Detachment in Thin Metal and Composite Plates by Inversion of Laser Doppler Vibrometry Scans.- Journal of Acoustical Society of America, (2007), Vol. 121, No. 5, pp. 2667-2672.

Numerical calculation of tertiary air duct in the cement kiln installation

Borsuk G., Wydrych J., Dobrowolski B.

Department of Thermal Engineering and Industrial Facilities
Opole University of Technology, Opole, Poland
e-mail: g.borsuk@po.opole.pl, j.wydrych@po.opole.pl, b.dobrowolski@po.opole.pl

Key words: pneumatic transport, particle concentration, tertiary air, rotary kiln

Abstract
Rotary kiln installation as a very complicate system of many elements influences on the cement production. During operation of tertiary air installation some problems with settling the particles are possible. This paper presents results of numerical calculations of the gas-particle flow in tertiary air duct installation. Gas-particle flow was calculated using Euler method for air motion and Lagrange method for the particles motion. This paper shows that it is possible to focus on studies of the tertiary air installation without the analysis of the other processes in the rotary kiln. Solving the inverse problem it gave the possibility to analyse several variants of modification the tertiary air installation.

1 Introduction

Because of the clinker production process, portion of the fuel can be burned in preheater, it is necessary to supply additional air (tertiary air), which should be taken from the clinker cooler [1]. Trouble at the tertiary air conveying is that clinker particles which are entrained from the cooler may be settle in the duct [2]. For optimization of working conditions, it is necessary to understand the detailed process which takes place in the kiln. The specific requirements of the process are designing the burner system and provide a burner design that will deliver efficient and reliable combustion. Yet, extensive use of this type of transport may bring about some problems that should be eliminated.

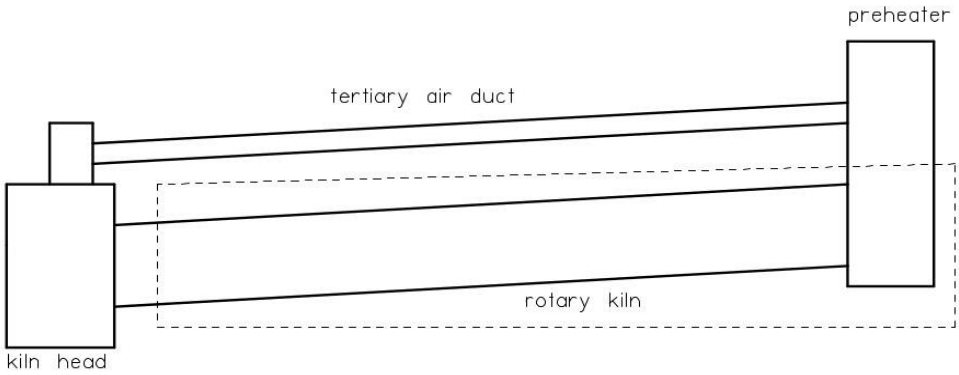


Figure 1 Rotary kiln general view

For describing the complex process of cement production a numerical calculation domain was built. For subsequent analyses only flow distribution through the kiln head was important, authors decided to build the cement kiln installation only partially. Fig. 1 shows general view of the rotary kiln system

with connection between kiln head and preheater. Part of the installation inside a dotted line has been omitted and then there was a possibility to build a calculation domain for tertiary air duct only. Fig. 2 shows tertiary air duct with inlet and outlet. In Fig. 3 head of rotary kiln connected with "basic" tertiary air duct is presented. Flow parameters for the system are based on the following assumptions: constant velocity in the inlet cross-section depends on the gas flow rate to the kiln head, at the outlet of the kiln and pipeline constant pressure value was determined for the "basic" variant for the appropriate distribution of gas stream to the kiln and tertiary air duct.

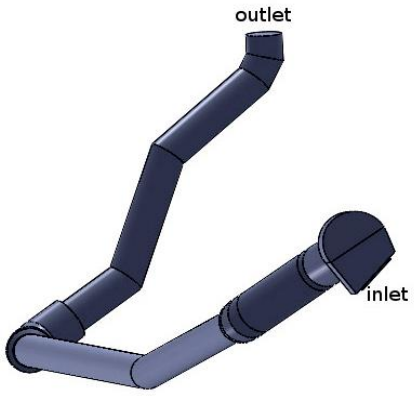


Figure 2 Tertiary air duct

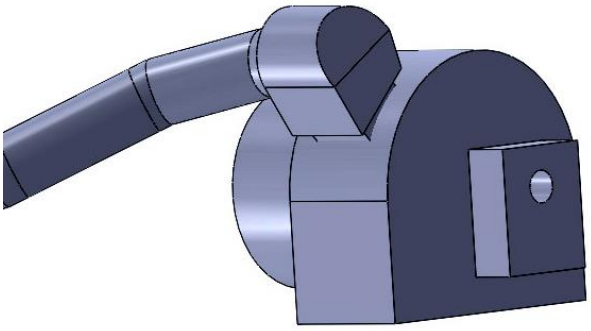


Figure 3 Rotary kiln head for "basic" conditions

2 Mathematical model

In order to realize numerical calculations, the mathematical model containing equations of motion for the gaseous phase and coal dust particles was formulated. The air motion was described with the Euler method, and the particle motion – by the Lagrange method. To analyse motion of the gas-particle polydispersive mixture, in this paper the PSI-Cell method was applied[3].

Numerical calculations are based on the following assumptions: considered flow is isothermal, stationary and without phase changes, both phases are incompressible, Gas motion is described in the uniform, generalized conservative form, isolating convection, diffusion and source components. In a consequence we obtain

$$\frac{\partial(\rho\phi)}{\partial t}+\frac{\partial\left(\rho U_i\phi\right)}{\partial x_i}=\frac{\partial}{\partial x_i}\left(\Gamma_\phi\frac{\partial\phi}{\partial x_i}\right)+S_\phi+S_{\phi p}, \tag{1}$$

where ϕ is a generalized dependent variable, Γ_ϕ is the coefficient of diffusion transport, and the source term S_ϕ contains all the remaining components of the differential equations[2]. The coefficients Γ_ϕ and S_ϕ are dependent on the variable ϕ. In the PSI-Cell method it is assumed that particles of the disintegrated phase are the sources of mass, momentum and energy occurring as additional components $S_{\phi p}$ in equations of the continuous (gaseous) phase.

The system of equations is accompanied by suitable boundary and initial conditions. The above system of partial differential equations is non-linear. Particular equations are coupled, so they have to be solved with special numerical techniques.

In order to calculate turbulence model k-ε was used. The standard k-ε model is a semi-empirical model based on model transport equations for the turbulence kinetic energy k and its dissipation rate ε. The model transport equation for k is derived from the exact equation, while the model transport equation for ε was obtained using physical reasoning and bears little resemblance to its mathematically exact counterpart [4].

In the derivation of the k-ε model, it was assumed that the flow was fully turbulent, and the effects of molecular viscosity were negligible. The standard k-ε model is therefore valid only for fully turbulent flows [5]. The turbulence kinetic energy, k and its rate of dissipation ε, are calculated from the following transport equations [6]:

$$\frac{\partial}{\partial t}(\rho k)+\frac{\partial}{\partial x_i}(\rho k u_i)=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right]+G_k+G_b-\rho\varepsilon-Y_M+S_k, \tag{2}$$

$$\frac{\partial}{\partial t}(\rho \varepsilon)+\frac{\partial}{\partial x_i}(\rho \varepsilon u_i)=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_j}\right]+C_{1\varepsilon}\frac{\varepsilon}{k}(G_k+C_{3\varepsilon}G_b)-C_{2\varepsilon}\rho\frac{\varepsilon^2}{k}+S_\varepsilon, \tag{3}$$

In these equations, G_k represents the generation of turbulence kinetic energy due to the mean velocity gradients. G_b is the generation of turbulence kinetic energy due to buoyancy. Y_M represents the contribution of the fluctuating dilatation in compressible turbulence to the overall dissipation rate. $C_{1\varepsilon}$, $C_{2\varepsilon}$, and $C_{3\varepsilon}$ are constants. σ_k and σ_ε are the turbulent Prandtl numbers for k and ε, respectively. S_k and S_ε are user-defined source terms. The turbulent (or eddy) viscosity μ_t is computed by combining k and ε as follows:

$$\mu_t=\rho C_\mu\frac{k^2}{\varepsilon}, \tag{4}$$

The model constants $C_{1\varepsilon}$, $C_{2\varepsilon}$, C_μ, σ_k and σ_ε have the following default values $C_{1\varepsilon}=1.44$, $C_{2\varepsilon}=1.92$, $C_\mu=0.09$, $\sigma_k=1.0$ and $\sigma_\varepsilon=1.3$.

The particle trajectory should be known during calculation of the mentioned above source components of differential equations. The particle trajectory is calculated according to its equation of motion. If the phase density difference is big, the equation of particle motion can be written as [11][12]:

$$\frac{d\vec{u}_p}{dt}=\frac{1}{\tau_p}(\vec{u}-\vec{u}_p)+g+3{,}084\left(\frac{\mu}{\rho}\right)^{1/2}\left(\frac{1}{|\vec{\omega}|}\right)^{1/2}\{\vec{\omega}\times(\vec{u}-\vec{u}_p)\}f(Re_p,Re_s), \tag{5}$$

3 Pressure drop optimization

Information about flow rate of incoming air to the head of rotary kiln was received from the monitoring system of the cement factory. Having information about the area of the inlet velocity was set to 5,538 m/s. Pressure loss between inlet to head and outlet from tertiary air pipe was treated as a base information for calculation proper distribution of gas. The primary step of setting the numerical calculation was the pressure at the outlet to the kiln, as the flow distribution is the same as at the start of the analysis. Over 70% of the incoming flow to the head is transported through the kiln. Minor rest is transported through the tertiary air installation.

Fig. 4 presents velocity profile for the “basic” conditions. In the tertiary air duct velocity is going to be in range from 0 to 20 m/s. In Fig. 5 there are results of calculations particles concentration in selected cross sections of the head of kiln.

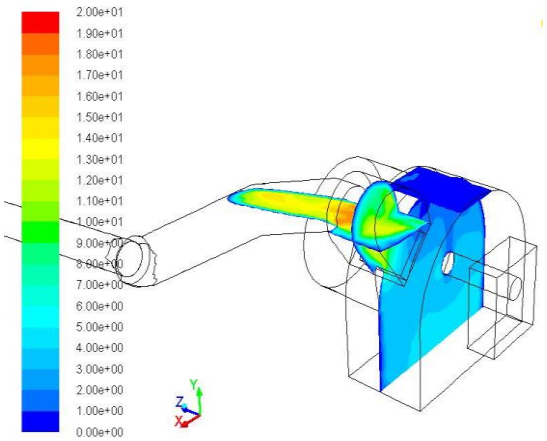


Fig. 4 Velocity profile [m/s] – variant V0

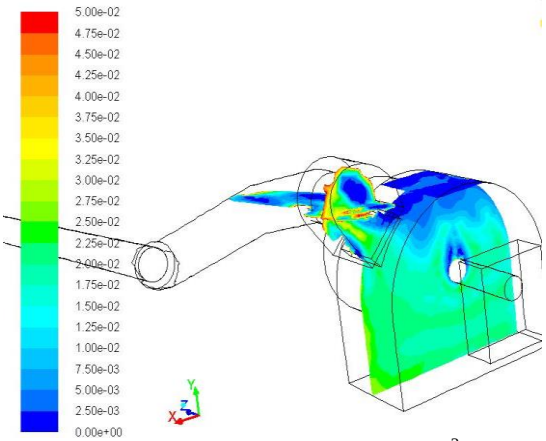


Fig. 5 Particle concentration [kg/m^3] – variant V0

For the particle calculation Rosin-Rammler-Sperling distribution was used and 10 fractions of particle with the range from 15 to 600 μm were determined. Tab. 1 presents the particle percentage of each fraction determined as a part of the mass flow for one fraction. In the calculations was assumed the same number of intake point for all particle fractions. Such adopted boundary conditions gave the results of the pressure drop calculation as was assumed. Calculations were made by using ANSYS Package [13]. Fig. 6 shows particle trajectories in the head of rotary kiln.

Table 1. Particle percentage

d[μm]	Percentage
15	7.51
35	9.39
55	10.32
75	14.12
110	17.16
150	18.66
220	16.67
350	5.15
450	0.95
600	0.07

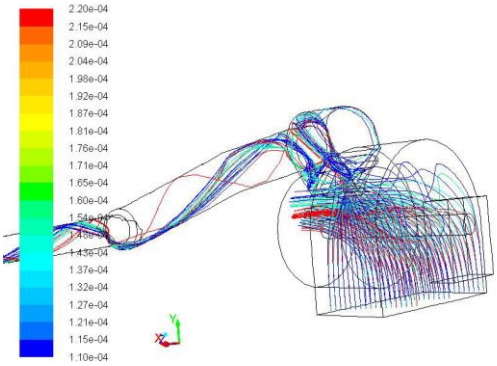


Figure 6 Particle trajectories (d110, d150, d220 μm) – variant V0

4 Tertiary air installation with solid particle settler.

For the "basic" conditions obtained from the initial calculations (variant V0), special settler just after the outlet of the kiln head as one of the tertiary air installation sections was used. Fig. 7 shows selected variants of the modification. In variants V1 and V2 inlet and outlet in the settler are situated under 90° angle, in variants V3 and V4 inlet and outlet are in the same axis but another relative to the circular base of the settler. Selected particle trajectories are presented for each variant. In variants V1 and V3 settler is situated back to the rotary kiln, in variants V2 and V4 – closer to the kiln. In each variant some particles are transported with the gas through the tertiary air duct, especially in variant V2 a "rope" of particles is visible. For this set, the largest number of particles is transported in the tertiary air duct, so in these variant working conditions are the worst of the possible.

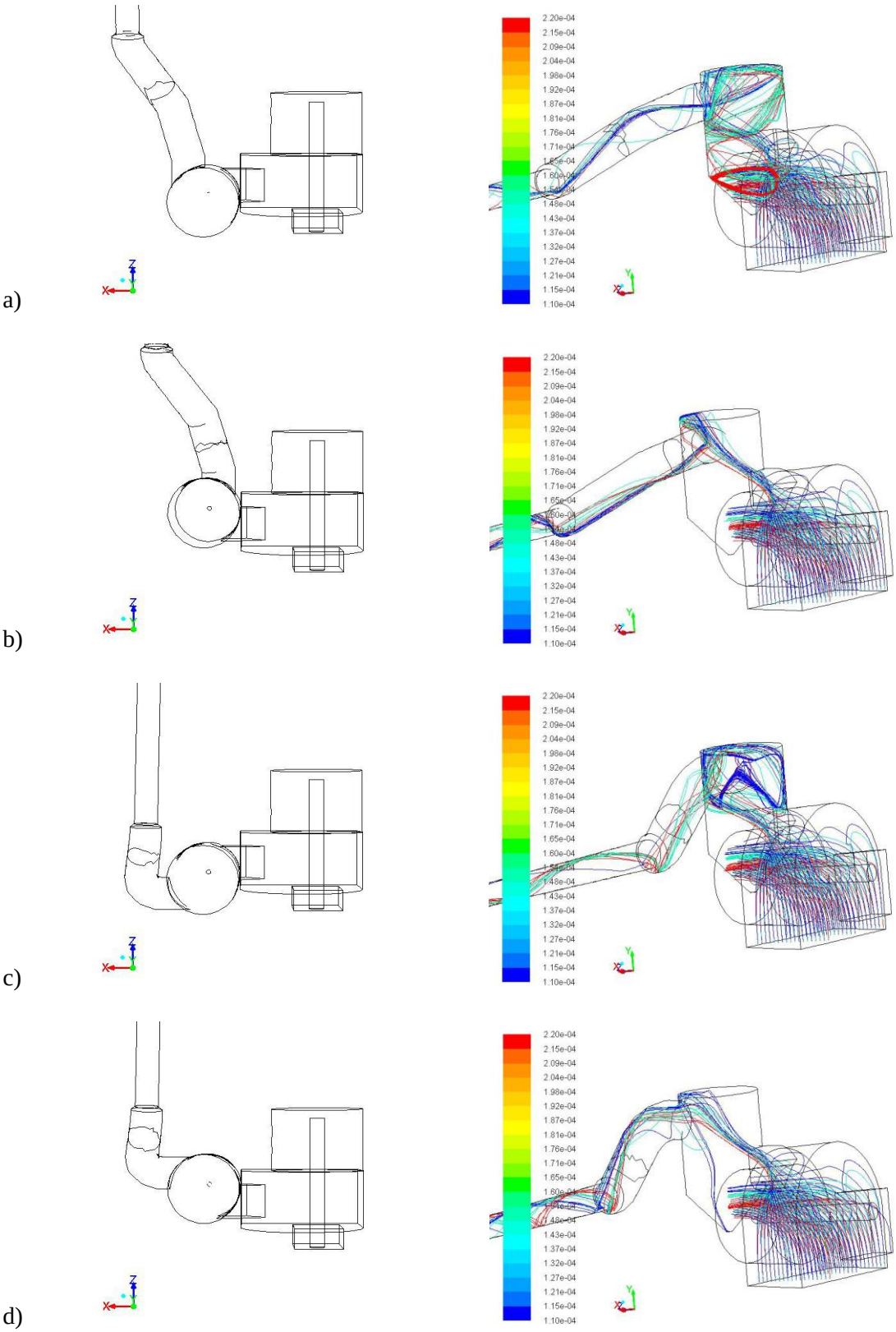


Figure 7 Considered modifications of the settler and selected particle trajectories (d110, d150, d220 μm) for: a) variant V1, b) variant V2, c) variant V3, d) variant V4

4.1 Particle conveying.

The next step was the opportunity of counting particles transported in tertiary air installation. Tab. 2 presents number of particles at the outlet from tertiary air installation for "basic" conditions and next 4 variants.

Table 2 Number of particles at the outflow from tertiary air installation

diameter	V0	V1	V2	V3	V4
d015	809	797	805	704	736
d035	780	745	737	683	721
d055	749	737	730	417	578
d075	678	705	705	570	564
d110	665	651	667	304	614
d150	497	307	619	610	605
d220	481	257	498	496	499
d350	80	2	24	2	21
d450	0	0	0	0	0
d600	0	0	0	0	0

For better explanation of the problem, next step there were calculations in order to obtain information on mass particle flow for each of the variants. Tab. 3 shows the results of the calculation.

Table 3 Mass flow rate [t/h]

diameter	V0	V1	V2	V3	V4
d015	0,1110	0,1094	0,1105	0,0966	0,1010
d035	0,1338	0,1278	0,1265	0,1172	0,1237
d055	0,1413	0,1390	0,1377	0,0786	0,1090
d075	0,1750	0,1819	0,1819	0,1471	0,1455
d110	0,2085	0,2042	0,2092	0,0953	0,1925
d150	0,1695	0,1047	0,2111	0,2080	0,2063
d220	0,1465	0,0783	0,1517	0,1511	0,1520
d350	0,0075	0,0002	0,0023	0,0002	0,0020
d450	0,0000	0,0000	0,0000	0,0000	0,0000
d600	0,0000	0,0000	0,0000	0,0000	0,0000
sum	**1,0932**	**0,9455**	**1,1308**	**0,8942**	**1,0321**

Tab. 3 shows that the best results in terms of reducing the number of particles being transported in the tertiary air installation is obtained in variant V1, where inlet and outlet are positioned relative to the settler at an angle of 90°. In variants V3 and V4 mass flow rate is smaller than the "basic" system, but the ratio of particles having diameters exceeding 150 μm is greater. This is probably connected with results of the formation of a „rope" of particles passing through the settler. This phenomena is even more visible in the variant V2, where most of the particles exiting from the head after being distributed to the pipeline, avoiding the middle part of the settler.

5 Summary

From the results of numerical calculations, the following conclusions can be drawn:

- It is possible to develop the inverse problem in order to focus on studies of the tertiary air installation without the analysis of the other processes in a rotary kiln,
- Initial conditions were selected on the basis of a predetermined pressure drop between kiln head and preheater,
- The pressure at the outlet to the rotary kiln was selected as long, as the distribution of the gas flow rate is consistent with the real installation,
- Solving the inverse problem, it was possible to analyse several variants of modification the tertiary air installation

References

[1] H. Lederer: A new rotary kiln burner technology, World cement (1996)
[2] G. Borsuk, B. Dobrowolski, J. Wydrych: Gas - solids mixture flow through a two - bend system. -Chemical and Process Engineering vol. 27, nr 3/1, (2006), pp. 645-656
[3] J. Wydrych: Comparative analysis of the methods of simulation of flow in boiler dust systems, Chemical and Process Engineering, vol. 31, no. 4, (2010), pp. 603-656
[4] J. Wang., S. Shirazi: A CFD based correlation for mass transfer coefficient in elbows, International Journal of Heat and Mass Transfer, 44, (2001), pp. 1817-1822
[5] H. Akili, E.K. Levy, B. Sahin: Gas-solid flow behavior in a horizontal pipe after a 90° vertical-to-horizontal elbow, Powder Technology 116, (2001), pp. 43-52
[6] B. Kuan, W. Yang, M.P. Schwarz: Dilute gas–solid two-phase flows in a curved 90° duct bend: CFD simulation with experimental validation, Chemical Engineering Science, 62, (2007), pp. 2068–2088
[7] D.K. Fidaros, C.A. Baxevanou, C.D. Dritselis, N.S. Vlachos, Numerical modelling of flow and transport processes in a calciner for cement production, Powder Technology, 171, (2007), pp. 81-95
[8] Z. Hu, J. Lu, L. Huang, S. Wang, Numerical simulation study on gas-solid two-phase flow in pre-calciner, Communications in Nonlinear Science and Numerical Simulation, 11, (2006), pp.440-451
[9] Z. Zhou, H. Zhu, B. Wright, A. Yu, P. Zulli: Gas-solid flow in an ironmaking blast furnace-II: Discrete particle simulation, Powder Technology, Vol. 208, Issue 1, (2011), pp. 72-85
[10] R. Saidura, M.S. Hossaina, M.R. Islama, H. Fayazb, H.A. Mohammed, A review on kiln system modeling, Renewable and Sustainable Energy Reviews, 15, (2011), pp. 2487-2500.
[11] S. Laín, M. Sommerfeld, Numerical calculation of pneumatic conveying in horizontal channels and pipes: Detailed analysis of conveying behavior, International Journal of Multiphase Flow, no. 39, (2012), pp. 105–120
[12] H. Bilirgen, E.K. Levy: The effect of a band on the particle cross-section concentration and segregation in pneumatic conveying systems, Powder Technology, 98, (1998), pp. 95-103
[13] ANSYS Fluent, Release 14.0, ANSYS Inc.,

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

ESTIMATION OF THE LOCAL HEAT TRANSFER COEFFICIENT IN COILED TUBES: COMPARISON BETWEEN TIKHONOV REGULARIZATION METHOD AND GAUSSIAN FILTERING TECHNIQUE

F. Bozzoli[1*], L. Cattani[1], S. Rainieri[1], F.S.V. Bazán[2], L.S. Borges[3]

[1]Department of Industrial Engineering,
University of Parma, Parco Area delle Scienze 181/A I-43124 Parma, Italy
e-mail: fabio.bozzoli@unipr.it

[2]Department of Mathematics
Federal University of Santa Catarina, CEP 88040-900, Florianópolis/SC, Brazil.

[3]Department of Mathematics
Federal University of Technology, CEP 84016-210, Ponta Grossa/PR, Brazil.

Key words: coil, Tikhonov regularization, Gaussian filter

Abstract
Most of passive techniques for heat transfer enhancement (e.g. rough surfaces, swirl-flow devices and coiled tubes) origins an irregular distribution of the heat transfer coefficient at the fluid wall-interface along the wall perimeter. This irregular distribution could be critical in some industrial applications but most of the available research papers, mainly due to the practical difficulty of measuring heat flux on internal wall surface of a pipe, presents the results only in terms of Nusselt number averaged along the wall circumference. The application of inverse heat conduction problem solution techniques overcomes this limitation because they enable to estimate the local convective heat transfer coefficient starting from the temperature distribution acquired on the external wall surface. In this work, two different techniques were considered and compared: Tikhonov Regularization Method (TRM) and Gaussian Filtering Technique (GFT). They were tested considering a particular problem within passive heat transfer enhancement techniques: estimating the local convective heat transfer coefficient in coiled tubes. The comparison, performed both by synthetic and experimental data, revealed that the TRM performs better than the GFT if the signal is affected by low noise level while, for higher noise value, their efficiency is comparable.

1 Introduction

Savings in materials and energy use provide strong motivation for adopting techniques of heat transfer enhancement in the design of commercial heat exchangers. Enhancement techniques can be separated into two categories: passive and active. Passive methods require no direct application of external power and they usually employ special surface geometries, which cause heat transfer enhancement. On the other hand, active schemes (e.g. electromagnetic fields and surface vibration) do require external power for operation, [1].

Passive techniques are commercially more attractive because no power is required to facilitate the enhancement and among them treated surfaces, rough surfaces, displaced enhancement devices, swirl-flow devices, surface-tension devices, coiled tubes, or flow additives are found. Most of them origins an irregular distribution of the velocity field over cross-section of the tube and this could lead to a significant variation of the heat transfer coefficient at the fluid-wall interface along the circumferential angular coordinate, [2,3]. Although this irregular distribution could be critical in some industrial applications, such as in the ones that involve a thermal process, most of the papers available in the scientific literature did not investigate this aspect, mainly due to the practical difficulty of measuring heat flux on internal wall surface of a pipe, and they presented the results only in terms of the Nusselt number averaged along the wall circumference.

The application of the Inverse Heat Conduction Problem (IHCP) solution techniques overcomes the experimental problem of monitoring heat transfer on the internal wall of the tube because they enable to estimate the local convective heat transfer coefficient starting from the temperature distribution acquired on the external wall surface. However, this approach presents some complications due to the fact that IHCP is ill-posed and, consequently, it is very sensitive to small perturbations in the input data. In order to bypass the ill-posedness of inverse problems, many techniques based on the processing of the experimental data have been suggested and validated in literature. In this work, two different techniques were considered and compared: Tikhonov Regularization Method (TRM) and Gaussian Filtering Technique (GFT).

Filtering approach is one of the most promising ones when input data are represented by spatially highly resolved temperature maps, such as the ones obtained by new generation infrared cameras. This approach is based on filtering out the unwanted noise from the raw temperature data in order to make feasible the direct calculation of its Laplacian, embedded in the inverse heat conduction problem's formulation [4]. In the present investigation the Gaussian filter, already successfully applied in literature for solving IHCPs, was adopted. The choice of the cut-off frequency of the filter, a crucial problem of this approach, was performed according to the discrepancy principle [5].

On the other hand, TRM is one of the most famous approach for this kind of problem [6]; it solves the original ill-posed IHCP by the minimization of an objective function expressed by the sum of the squared difference between measured and estimated temperature discrete data and of a regularization parameter times a term that expresses the smoothness of the unknown quantity. The success of this approach relies on a proper choice of the regularization parameter and this is not an easy task. In the present paper, in order to make the comparison with the filtering technique more reliable, the discrepancy principle was adopted for determining a proper value of the regularization [5]. It has to be reminded that, if for TRM other approaches (e.g. Fixed-point, L-curve) are available, for filtering approaches the discrepancy principle is the only reliable one.

In this paper, these two methods were tested considering a particular application: starting from temperature distribution on the external wall surface of a coiled tube, the local convective heat transfer coefficient at the internal wall surface has to be estimated. This application, from the industrial point of view, is particularly interesting because helical coil is a widely adopted technique of heat transfer enhancement in the thermal processing of highly viscous fluids. Its effectiveness relies on the fact that wall curvature gives origin to the centrifugal force inducing local maxima in the fluid velocity distribution, locally increasing the temperature gradients at the wall and, as a result, maximizing the

heat transfer [7-10]. The uneven distribution of the velocity field leads to a significant variation of the heat transfer coefficient along the wall perimeter. For all these reasons this tube geometry is an ideal benchmark for testing the IHCP solution techniques. In this paper the comparison is performed both with synthetic and experimental data.
Finally, it must be stressed that, as observed in [11], the use of data filtering approach to the solution of the IHCP is still not well-established, even though these filters are available in many engineering data processing environments with very moderate computational costs. Therefore, it is desirable that some further research efforts will be devoted to this research topic and the present paper has been developed in this direction.

2 Problem's definition

A typical condition in which curved tubes are tested is the one in which the convective heat transfer within the fluid that flows inside the tube occurs under the uniform heat flux boundary conditions, such as the one considered by Rainieri et al. [10,12] in their experimental investigations where a heat flux was dissipated by Joule effect directly within the tube wall.
In order to evaluate the local actual value of the convective heat transfer coefficient at the fluid internal wall interface on a given cross section (as highlighted in figure 1), the temperature distribution is acquired on the external wall surface and then the IHCP in the wall domain is solved by considering the convective heat transfer coefficient distribution on the internal wall surface to be unknown.

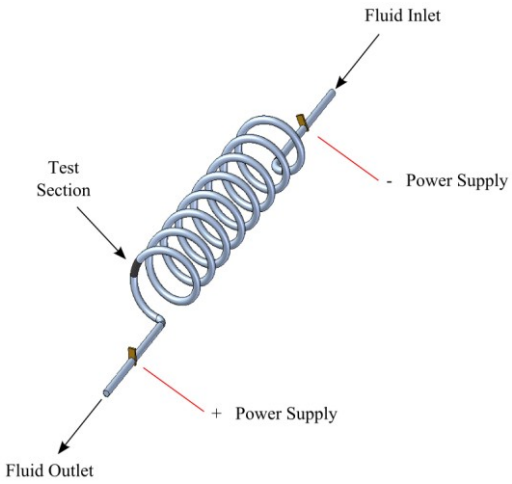


Figure 1: Coiled tube.

The test section (sketched in Fig. 2) is efficiently modelled as a 2-D solid domain by assuming that along the axis of the tube the temperature gradient is almost negligible.
In this domain the steady state energy balance equation is expressed in the form:

$$k\nabla^2 T + q_g = 0 \tag{1}$$

where q_g is the heat generated per unit volume and k is the tube wall thermal conductivity. The following two boundary conditions completed the energy balance equation:

$$k\frac{\partial T}{\partial r} = \frac{(T_{env} - T)}{R_{env}} \tag{2}$$

that is applied on surface S_1 and where R_{env} is the overall heat transfer resistance between the tube wall and the surrounding environment with the temperature T_{env} ;

$$-k\frac{\partial T}{\partial r} = h_{int} \cdot (T_b - T) \tag{3}$$

that is applied on surface S_2 and where T_b is the bulk fluid temperature and h_{int} is the local convective heat transfer coefficient at the fluid- internal wall interface.

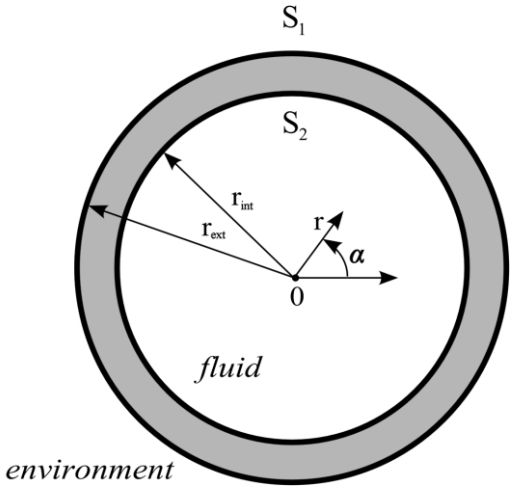


Figure 2: Geometrical domain with coordinate system.

For a given experimental set-up the bulk fluid temperature can be easily calculated by imposing that the energy balance is satisfied [10,12].

3 Solution techniques

In this paragraph the two techniques employed to solve the IHCP and to estimate the local convective heat transfer coefficient are synthetically presented. An exhaustive discussion of these approaches can be find, for the TRM, in [12] and for GFT in [4].

3.1 Gaussian filtering technique

The filtering approach aims to filter out the unwanted noise from the raw temperature data in order to make feasible the direct calculation of its Laplacian, embedded in IHCP's formulation. The effectiveness of the Gaussian kernel in this kind of inverse problem was experimented by Murio et al. [13], Delpueyo et all [14] and Bozzoli et al. [4].
In order to implement the above procedure, that can be ultimately classified as a function estimation procedure, a thin wall model of the test section was considered. Under this approach, the convective heat transfer coefficient at the fluid-coiled wall interface follows form the steady state local energy balance:

$$h_{int} = \frac{k\, ln\left(\frac{r_{ext}}{r_{int}}\right)\cdot\frac{d^2Y_f}{d\alpha^2} - r_{ext}\cdot\frac{\left(Y_f - T_{env}\right)}{R_{env}} + \frac{q_g}{2}\cdot\left(r_{ext}^2 - r_{int}^2\right)}{r_{int}\cdot\left(Y_f - T_b\right)}, \quad (4)$$

where Y_f is the measured temperature distribution on the external tube's wall, processed by the Gaussian filter .

A Gaussian filter is a filter whose impulse response is a Gaussian function. Since the Fourier transform of a Gaussian is again a Gaussian function, the application of a Gaussian filter has the effect of reducing the data high-frequency components. It is widely adopted in the enhancing of images' quality within graphics software. The transfer function in the frequency domain, of this kind of filter can be expressed as follows:

$$H(u,v) = e^{-(u+v)^2/2u_c{}^2}, \quad (5)$$

where u_c is the cutoff frequency. Since in real application the optimal cutoff frequency value for each kind of filter is not known a priori, a criterion to choose it must be selected in order to make successful the regularization procedure. This choice is similar to the selection of optimal regularization parameters when dealing with regularization methods for inverse estimation [5,6,15]. In the present analysis the criterion provided by the discrepancy principle, originally formulated by Morozov [16], was adopted.

According to this principle, the inverse problem solution is regarded to be sufficiently accurate when the difference between filtered and measured temperatures is close to the standard deviation of the raw measurements. The cutoff frequency was determined as the frequency at which the condition expressed in Eq. (6) is satisfied.

$$\frac{\left\|\mathbf{Y}_f - \mathbf{Y}\right\|_2}{\sqrt{N}} = \sigma_Y \quad (6)$$

where $\|\ \|_2$ stands for the 2-norm of a vector, N is the size of the vector $\mathbf{Y}$ and σ_Y is the standard deviation of the measurement error.

3.2 Tikhonov regularization method

It is helpful to reformulate the direct problem making explicit the local convective heat flux at the fluid-internal wall interface; under this approach the boundary condition expressed by Eq.(3) becomes:

$$-k\frac{\partial T}{\partial r} = q(\alpha), \quad (7)$$

which is applied on surface S_2 and where q is the local convective heat flux at the fluid-internal wall interface, assumed to be varying with the angular coordinate α.

To express the problem in the discrete domain, the convective heat flux distribution can be simplified by considering that it is described by a continuous, piecewise linear function. Proceeding this way the direct problem becomes linear with respect to the heat flux q(α) and its discrete version can be described as follows:

$$\mathbf{T} = \mathbf{Xq} + \mathbf{T_{q=0}}, \quad (8)$$

where **T** is the vector of the discrete temperature data at the external coil surface, **q** is the heat flux vector at the fluid-internal wall interface, $\mathbf{T_{q=0}}$ is a constant term and **X** is the sensitivity matrix.
The direct formulation of the problem is concerned with the determination of the temperature distribution on the tube external wall when the convective heat flux vector **q** is known. In the inverse formulation considered here, **q** is instead regarded as being unknown, whereas the surface temperature **Y** is measured.
As the inverse problem is ill-posed, in order to cope with the presence of noise in the measured temperature some type of regularisation is required. The TRM, successfully applied in the inverse heat-transfer literature [17,18], makes it possible to reformulate the original problem as a well-posed problem that consists of minimising the following objective function:

$$J(\mathbf{q}) = \left\| \mathbf{Y} - \mathbf{Xq} - \mathbf{T_{q=0}} \right\|_2^2 + \lambda^2 \left\| \mathbf{Lq} \right\|_2^2, \ \lambda>0, \tag{9}$$

where $\left\| \cdot \right\|_2^2$ stands for the 2-norm, λ is the regularisation parameter, **L** is a discrete derivative operator and **T** is the distribution of the external surface temperature derived from a direct numerical solution of the problem obtained by imposing a given convective heat flux distribution on the internal wall side **q**. Often, **L** is the zero, first or second derivative operator: in this work the second-order derivative formulation was chosen to preserve the local variation in the heat-flux distribution.
The function expressed in Eq. (9) represents a trade-off between the fidelity of the fit and the stability of the solution. Thus, an appropriate choice of λ is a crucial point to find a reliable approximation of the wanted solution.
In the present paper, in order to make the comparison with the filtering technique more consistent, the discrepancy principle was adopted for determining a proper value of the regularization parameter [5,18]. The discrepancy principle suggests computing λ in such a way that the corresponding Tikhonov solution, that is the minimum of the Eq. (9), satisfies the non linear equation:

$$\frac{\left\| \mathbf{Y} - \mathbf{Xq} - \mathbf{T_{q=0}} \right\|_2}{\sqrt{N}} = \sigma_Y \tag{10}$$

where N is the size of the vector **Y** and σ_Y is the standard deviation of the measurement error.
Once the heat-flux distribution at the fluid-wall interface compatible with the experimental temperature data has been determined through the strategy described above, the local convective heat-transfer coefficient can be determined, as follows:

$$h_{\text{int}}(\alpha) = \frac{q_\lambda(\alpha)}{T(\alpha, r = r_{\text{int}}) - T_b}, \tag{11}$$

where $q_\lambda(\alpha)$ is the heat flux distribution estimated under the solution approach based on the Tikhonov regularisation method and $T(\alpha,r=r_{int})$ is temperature distribution on the tube internal wall efficiently estimated by numerically solving the direct problem expressed by Eq. (8) by imposing a convective heat flux equal to $q_\lambda(\alpha)$.

4 Numerical comparison

The comparison of the two above described techniques was first performed by adopting synthetic data. By imposing a known distribution of h_{int} and by solving the governing Eq. (1-3), a synthetic temperature distribution on the external wall surface was obtained. The physical and geometrical parameters used in this work correspond to a stainless steel tube with an internal radius of 7 mm and an external radius of 8 mm. For the local convective heat transfer coefficient, a distribution according to the data of Jayakumar et al. [8], and shown in Fig. 3, was adopted. This distribution, corresponding to the thermally fully developed region, was derived by numerical simulations performed under the turbulent flow regime and it shows a significant variation along the curvilinear coordinate.

Then the synthetic temperature distribution on the external wall surface, deliberately spoiled by random noise, was used as the input data of the inverse problem. In particular, a white noise characterized by a standard deviation ranging from 0.01 K - 5 K was considered.

In Fig. 3 the local heat transfer coefficient distributions restored by TRM and GFT for two different noise level are compared with the exact one.

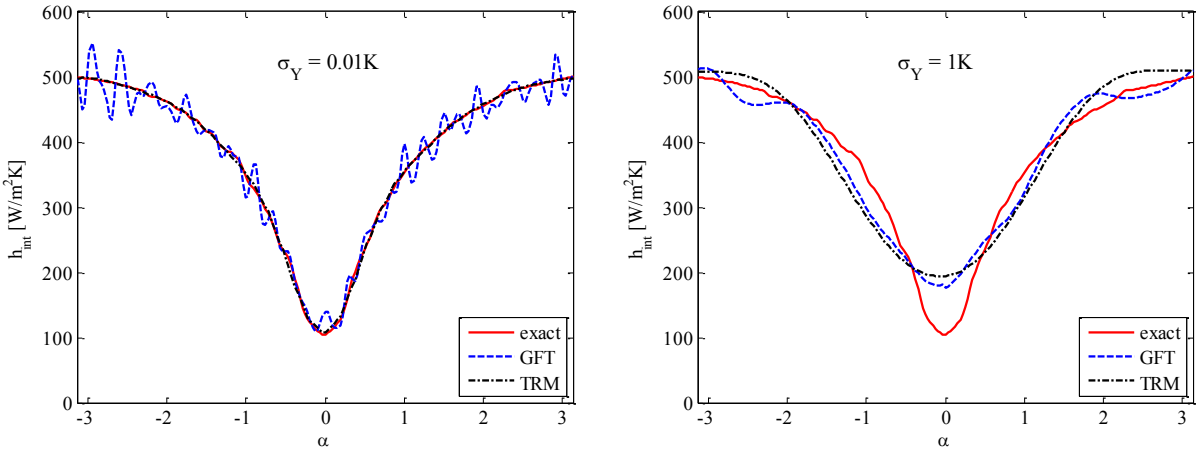


Figure 3: Exact and reconstructed h_{int} distribution obtained by Tikhonov Regularization Method (TRM) and Gaussian Filtering Technique (GFT) for two different noise level.

In order to quantify the effectiveness of the two approaches at different signal to noise level a residual analysis could be performed by plotting the estimation error, defined as follows:

$$E = \frac{\|(h_{int})_{restored} - (h_{int})_{exact}\|_2^2}{\|(h_{int})_{exact}\|_2^2} \tag{12}$$

versus the standard deviation of the measurement error.

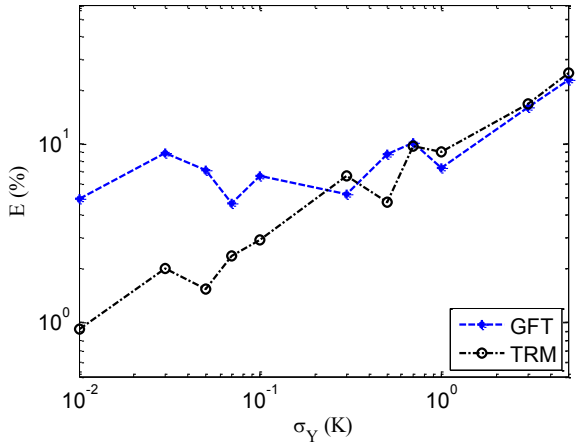


Figure 4: Estimation error on h_{int} for Tikhonov Regularization Method (TRM) and Gaussian Filtering Technique (GFT) at different noise level.

This comparison highlights that, for the problem here investigated, the TRM performs better than the GFT if the noise is lower than 0.1 K while, for higher noise value, their efficiency is almost comparable.

5 Experimental comparison

The two estimation techniques were compared trough their application to a set of experimental data obtained in [12]: a stainless steel coiled tube was tested under the prescribed condition of uniform heating generated by Joule effect in the wall. Ethylene Glycol was used as working fluid and laminar flow regime was investigated. The temperature distribution of the tube external wall was obtained by moving the infrared camera around the section and merging together the acquired multiple images. Further details on the experimental facilities and procedures can be found in [12] .

In order to apply both the estimation procedures, the noise level in the acquired data has to be estimated; it was done by measuring the surface temperature distribution while maintaining the coil wall under isothermal conditions.

For a representative Reynolds number value, the distributions of the convective heat transfer coefficient restored by TRM and GFT are compared in Fig. 5: this plot underlines that, for the case here investigated, the two approaches give equivalent results although the implementation of the regularization approach through Gaussian filter requires a lower computational cost if compared to Tikhonov regularization approach.

The residuals between the experimental and the reconstructed temperature, plotted in Fig. 6, gives deeper insight into the performances of two estimation procedures: the residuals produce by TRM are more randomly distributed than the ones by GFT. This observation suggests that, in this case, even if no appreciable differences are noticeable in terms of estimated local convective heat transfer coefficient, TRM works slightly better than GFT which tends to excessively smoothen the temperature distribution .

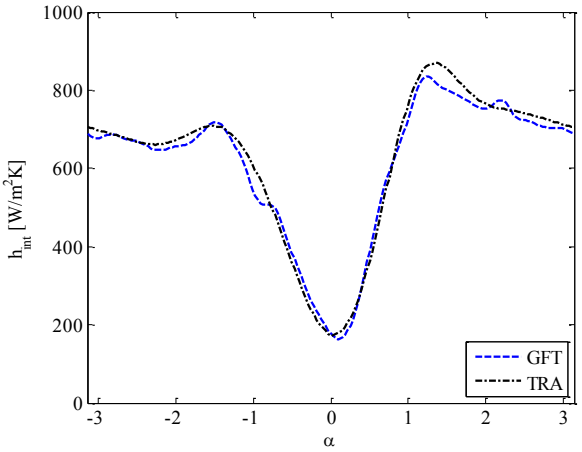


Figure 5: Convective heat transfer coefficient distribution at the fluid-internal wall interface obtained by Tikhonov Regularization Method (TRM) and Gaussian Filtering Technique (GFA), Re = 665.

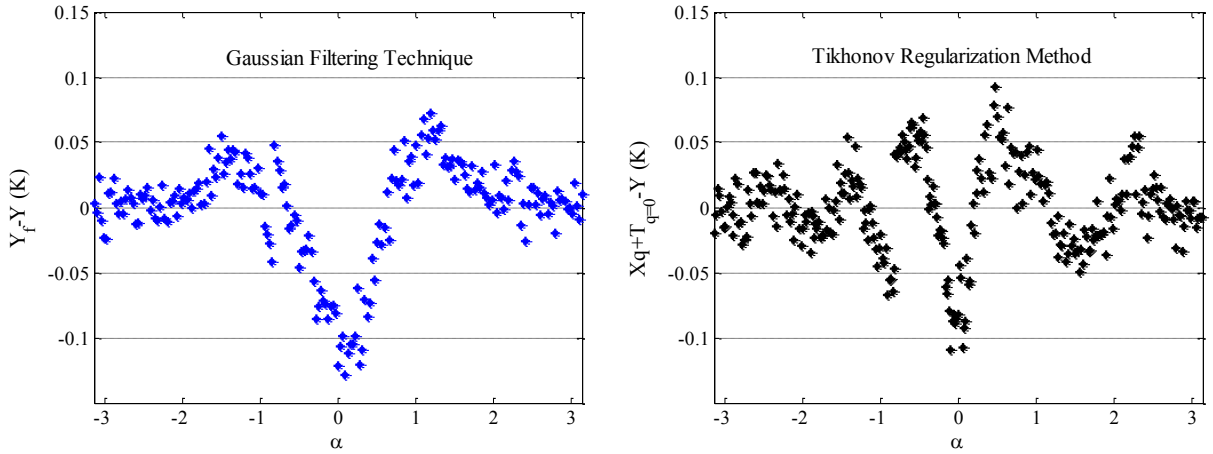


Figure 6: Residuals between the experimental and the reconstructed temperature obtained by GFT (left) and TRM (right).

6 Conclusions

Two different Inverse Heat Conduction Problem solution techniques were considered and compared: Tikhonov Regularization Method (TRM) and Gaussian Filtering Technique (GFT). They were tested with regard to the estimation of the local convective heat transfer coefficient in coiled tubes. In order to make the evaluation more robust, the discrepancy principle was adopted for the both techniques.
The comparison, firstly performed through their application to synthetic data, revealed that the TRM performs better than the GFT if the signal is affected by low noise level while, for higher noise value, their efficiency is comparable.
The following application to experimental data established that, for the experimental facilities here adopted, the two approaches give equivalent results although the implementation of the regularization approach through GFA requires a lower computational cost if compared to TRM.

References

[1] R. Webb, N.H. Kim, *Principles of Enhanced Heat Transfer*, second ed., Taylor & Francis, (2005).

[2] A.E. Bergles, *Techniques to enhance heat transfer* in: *Handbook of Heat Transfer*, McGraw-Hill, New-York, (1998).

[3] S. Rainieri, F. Bozzoli, L. Schiavi, G. Pagliarini, *Numerical analysis of convective heat transfer enhancement in swirl tubes*, International Journal of Numerical Methods for Heat and Fluid Flow, 21(5), (2011), pp. 559-571.

[4] F. Bozzoli, G. Pagliarini, S. Rainieri, *Experimental validation of the filtering technique approach applied to the restoration of the heat source field*, Experimental Thermal and Fluid Science, 44, (2013), pp. 858-867.

[5] J.V. Beck, B. Balckwell, C.R.St. Clair, *Inverse Heat Conduction – Ill-posed Problems*, Wiley-Interscience, New York, USA, (1985).

[6] A.N. Tikhonov, V.Y. Arsenin, *Solution of Ill-Posed Problems*, Winston & Sons, Washington DC, USA, (1977).

[7] G. Yang, F. Dong, M.A. Ebadian, *Laminar forced convection in a helicoidal pipe with finite pitch*, International Journal of Heat and Mass Transfer, 38(5), (1995), pp. 853-862.

[8] J.S. Jayakumar, S.M. Mahajani, J.C. Mandal, N.I. Kannan, P.K. Vijayan, *CFD analysis of single-phase flows inside helically coiled tubes*, Computer and Chemical Engineering, 34, (2010), pp. 430-446.

[9] P. Naphon, S. Wongwises, *A review of flow and heat transfer characteristics in curved tubes*, Renewable & Sustainable Energy Reviews, 10, (2006), pp. 463–490.

[10] S. Rainieri, F. Bozzoli, L. Cattani, G. Pagliarini, *Compound convective heat transfer enhancement in helically coiled wall corrugated tubes*, International Journal of Heat and Mass Transfer, 59, (2013), pp. 353–362.

[11] N. Renault, S. André, D. Maillet, C. Cunat, *A two-step regularized inverse solution for 2-D heat source reconstruction*, International Journal of Thermal Sciences, 47 (7), (2008), pp. 834–847.

[12] F. Bozzoli, L. Cattani, S. Rainieri, F.S.V. Bazán, L.S. Borges, *Estimation of the local heat-transfer coefficient in the laminar flow regime in coiled tubes by the Tikhonov regularisation method*, International Journal of Heat and Mass Transfer, 72, (2014), pp. 352-361.

[13] D.A. Murio, *The Mollification Method and the Numerical Solution of Ill-Posed Problems*, John Wiley and Sons, New York, (1993).

[14] D. Delpueyo, X.Balandraud, M. Grédiac, *Heat source reconstruction from noisy temperature fields using an optimised derivative Gaussian filter*, Infrared Physics & Technology, 60, (2013), pp. 312–322.

[15] F.S.V. Bazan, L.S. Borges, *GKB-FP: an algorithm for large-scale discrete ill-posed problems*, Bit Numerical Mathematics, 50(3), (2010), pp. 481-507.

[16] V.A. Morozov, *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, (1984).

[17] P.C. Hansen, Rank-*deficient and discrete ill-posed problems*, SIAM, Philadelphia, (1998).

[18] H.R.B. Orlande, F. Olivier, D. Maillet, R.M. Cotta, *Thermal measurements and inverse techniques*, Taylor and Francis, New York , (2011).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

The Use of COMSOL and Inverse Problem Technique to Estimate the Heat Flux on a Cutting Tool

R. F. Brito[a], S. R. Carvalho[b], S. M. M. Lima e Silva*[a]

[a]Instituto de Engenharia Mecânica, Universidade Federal de Itajubá – UNIFEI, Itajubá, Minas Gerais, Brasil.
[b]Faculdade de Engenharia Mecânica, Universidade Federal de Uberlândia – UFU, Uberlândia, Minas Gerais, Brasil.
e-mail: rogbrito@unifei.edu.br, srcarvalho@mecanica.ufu.br, metrevel@unifei.edu.br

Keywords: COMSOL, inverse problems, machining process, heat transfer.

Abstract
This work proposes the use of inverse problem techniques in connection with COMSOL to estimate the heat flux and the temperature field on a turning cutting tool in transient regime. The main purpose of the present work is to present the improvements performed in relation to the authors' previous work to develop the complex geometry of a machining process. Specification function, which is an inverse problem technique, was implemented in a program to estimate the heat flux applied on the tool, from the experimental temperature records. Once the heat flux is known, COMSOL is again utilized to obtain the temperature field on the cutting tool. The validation of the methodology is carried out by comparing the numerical and experimental results of temperature.

1 Introduction

Several engineering processes have their performance and quality affected by high temperature values. A typical example is the machining process in which cutting tool temperatures may be higher than 900ºC [1]. High temperatures change the microstructure and physical properties of the tool during machining, thus reducing their capacity to resist mechanical stress [2]. The direct consequence of these alterations is the reduction of their lifespan and performance. This leads to high operation costs and reduction of the end product quality. The right knowledge of the temperature values and applied heat flux in this kind of process, results in advantages like the development of more efficient cooling techniques as well as better specifications of the cutting parameters in machining processes. These temperatures have a controlling influence on the wear rate of the cutting tool as well as on the friction between the chip and the tool. However, the direct measurement of the temperature in a machining process is difficult to accomplish due to the movement of the piece as well as the presence of chips. Thus, the use of inverse heat conduction techniques conveys a good alternative to obtain these temperatures, since these techniques allow the use of experimental data obtained from accessible regions. Inverse problems consist of obtaining the value of a variable through the measurement of

another variable measured directly [3]. These techniques often use optimization algorithm in order to minimize the error between the calculated and real value of the variable in question. Nowadays, several researchers have proposed the combination of inverse techniques and numerical heat transfer solutions to analyze the thermal fields during machining processes.

Conveying a greater availability of computational resources, the use of numerical methods gained terrain, and it did not take long before they started being used, along with experimental methods in the studies of temperature fields on cutting tools. A three-dimensional finite difference-based model to predict temperature in machining processes was presented in [4]. The FDM based model proposed in this paper offered very rapid and reasonably accurate solutions. The simulated results were validated with infrared thermal measurements which were determined from the machining of AISI 1050 and AISI H13 materials under various cutting conditions. In the study of [5] an analytical and numerical model for cutting temperature prediction of 316L stainless steel was developed. The simulation model was set up in commercial FEM software of Abaqus6.8, which is good at nonlinear dynamic calculation. An ALE finite element model, which combines the advantages of both Lagrangian and Eulerian techniques, was used. The Johnson-Cook plasticity model was used to model the workpiece material. The analytical modeling and FEM modeling results match very well. In [6] the temperature distribution of the micro-cutter in the micro-end-milling process was investigated by numerical simulations and experimental approach. Micro-end-milling processes were modeled by the three-dimensional finite element method coupling thermal–mechanical effects. The micro-cutter cutting temperature distribution, the effect of various tool edge radii on cutting force, and the effective stress during micro-end-milling of aluminum alloy Al2024-T6 using a tungsten-carbide micro-cutter were investigated on. The simulation results showed that with the increase of the tool edge radius the cutting force increases, while the effective stress and mean cutting temperature decreases slightly.

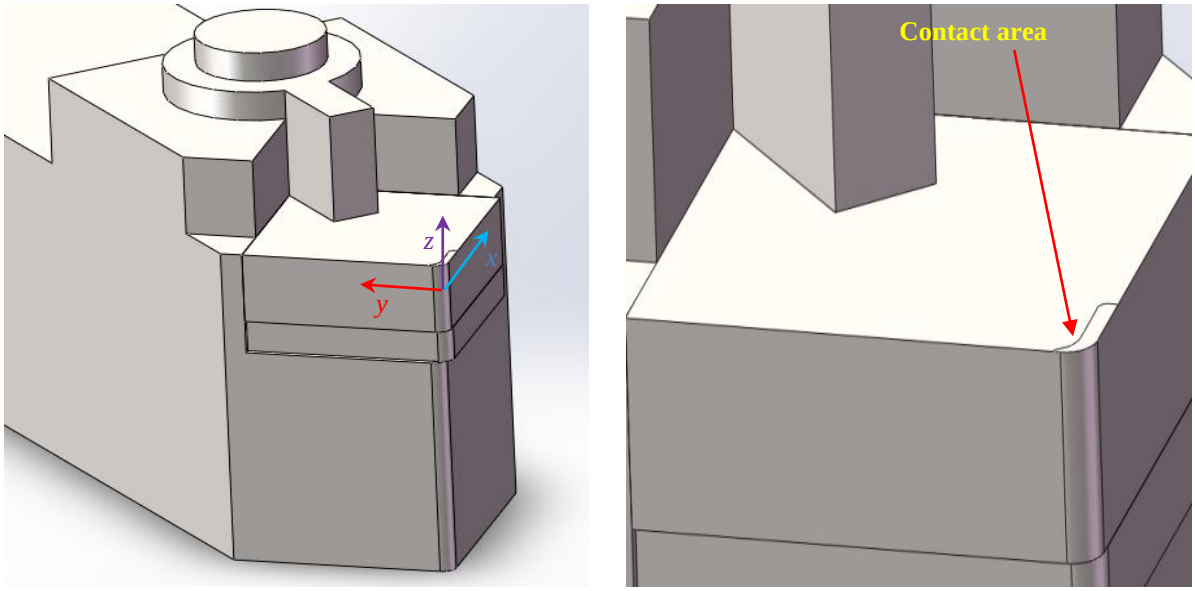
Inverse techniques have already been used to study temperature fields on a cutting tool. The solution of a three dimensional inverse heat conduction problem using an Evolutionary Algorithm (EA) was demonstrated in [7]. The heat flux on the tool during the turning process was determined by using evolutionary operations combined with measured temperatures on the tool surface. The three-dimensional conduction in the tool and tool holder was simulated using FLUENT. In [8] an inverse method was proposed to estimate the heat sources in the transient two-dimensional heat conduction problem in a rectangular domain with convective bounders. The non homogeneous partial differential equation (PDE) is solved using the Integral Transform Method. The test function for the heat generation term was obtained by the chip geometry and thermomechanical cutting. Then the heat generation term was estimated by the conjugated gradient method (CGM) with adjoint problem. The sequential function specification method was used to estimate the transient heat flux imposed on the rake face of a cutting tool during the cutting operation with two different assumptions [9]. In one of them the thermal conductivity is assumed to be constant, and in the other one it varies with the temperature. The cutting tool was modeled as a three dimensional object. Simulated temperature data was used to recover the heat flux at the cutting tool surface using linear as well as nonlinear solutions.

This work proposes the use of inverse problem techniques with the commercial software COMSOL® 4.3 to estimate the heat flux and the temperature field in the contact area under transient regime, in a turning cutting tool. A Matlab program, with Specification Function technique, was developed to estimate the heat flux applied on the cutting tool, using experimental temperature records in a determined point. The validation of the proposed methodology was accomplished in controlled experiments in laboratory.

2 Theoretical Formulation

2.1 Temperature model

The problem dealt with in this work is represented by Fig. 1a, which represents a set consisting of a cutting tool, a hard metal, a wedge positioned under the cutting tool between the tool and the tool holder. There is also a staple and a bolt to fix the set. In Figure 1a the schematic model for the thermal problem of machining is presented. The heat generation during the machining process is indicated by a distribution of unknown heat flux $q''(x,y,t)$, over the arbitrary area by x and y. A blown up view of the set is shown in Fig. 1b.



a) b)

Figure 1: a) Thermal problem scheme and b) Detail of the contact interface between the tool and the workpiece.

The heat diffusion equation ruling this problem may be given as:

$$\frac{\partial^2 T}{\partial x^2}(x,y,z,t)+\frac{\partial^2 T}{\partial y^2}(x,y,z,t)+\frac{\partial^2 T}{\partial z^2}(x,y,z,t)=\frac{1}{\alpha}\frac{\partial T}{\partial t}(x,y,z,t), \tag{1}$$

subject to the following boundary conditions

$$-k\frac{\partial T}{\partial z}(x,y,0,t)=q_0'', \text{ on the contact interface with the workpiece (Fig. 1b)} \tag{2}$$

and

$$-k\frac{\partial T}{\partial \eta}=h(T-T_\infty), \text{in the remaining regions of the set} \tag{3}$$

and having the following as the initial condition

$$T(x,y,z,t)=T_0, \text{ at } t = 0 \tag{4}$$

The direct problem consists in solving the heat diffusion equation according to the boundary conditions (Eqs. 1 to 4). The COMSOL® 4.3 program, which solves thermal problems by using the finite element method, is used for this purpose. The use of COMSOL for the numerical resolutions of differential equations that rule the physical phenomenon investigated should be highlighted. Also, COMSOL allows adjusting any boundary conditions, as well as modeling the geometry so as to faithfully represent the system investigated as presented in Fig. 1a.

2.2 The Inverse Problem

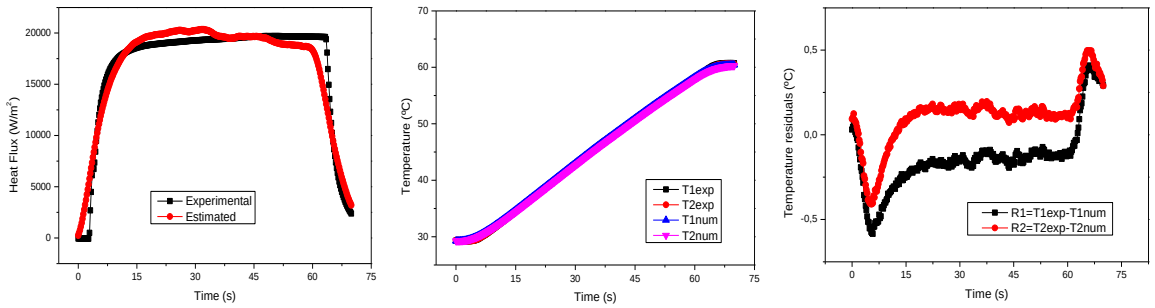
The inverse technique adopted in this work is the Specification Function [3]. This technique requires the calculation of the sensitivity coefficient which is done numerically from Duhamel Theorem [10]. The sensitivity coefficient is then obtained with the use of a numerical probe which follows the temperature changes in the point equivalent to where the thermocouples were placed in the experiments. Once the sensitivity coefficient is at hand, the heat flux is estimated with the use of a Matlab language program. Another important parameter is the value of future time steps *r*. In the Specification Function technique, a determined value of future time steps *r* is used to estimate the heat flux at present instant. In the resolution of the inverse problem, the Specification Function searches for a heat flux value that minimizes the objective function given in Eq. 5, for each time step

$$F = \sum_{p=1}^{r} \sum_{j=1}^{ns} \left(Y_{j,M+p-1} - T_{j,M+p-1} \right)^2 \tag{5}$$

3 Validation of the Methodology Proposed

A great difficulty in the solution of inverse heat conduction problems is the validation of the technique used. This difficulty is inherent to the problem, once the validation of the estimated heat flux requires the previous knowledge of the experimental heat flux. It is observed that in real inverse problems, as in machining process, the experimental heat flux is not known. Thus, an alternative for the validation of the inverse technique is to carry out a controlled experiment, in which the heat flux and the temperature are measured at the cut tool. In this sense, before the analysis of the real machining process, a cemented carbide tool with dimensions of 0.0127 x 0.0127 x 0.0047 m was used. A heat flux transducer and two thermocouples previously calibrated and a kapton electric heater were used on this tool. This heater was connected to a digital power supply (MCE). The heat flux transducer was located between the heater and the tool, in order to measure the heat supplied to the tool. The temperatures at the tool were measured with two thermocouples. The heat flux and temperatures signals were acquired by a HP Series 75000 data acquisition system, controlled by a PC. Temperatures were measured using type K thermocouples (30AWG) welded by capacitive discharge and calibrated by using a bath temperature calibrator ERTCO with a stability of ± 0.01 ºC.

The solution of the three-dimensional heat diffusion equation is obtained with the use of the finite element method, through the commercial software COMSOL® 4.3. For this, a computational thermal model was used to faithfully represent the experimental model of the sample. This model was discretized in a computational tetrahedral mesh. The validation results are presented in Figs. 2a, 2b and 2c. Figure 2a presents a comparison between the experimental and estimated flux, whereas Fig. 2b compares the experimental and numerical temperatures. Figure 2c presents the deviation between the experimental and numerical temperatures. The Specification Function method for *r* equal to 10 future time steps was used in Fig. 2a.

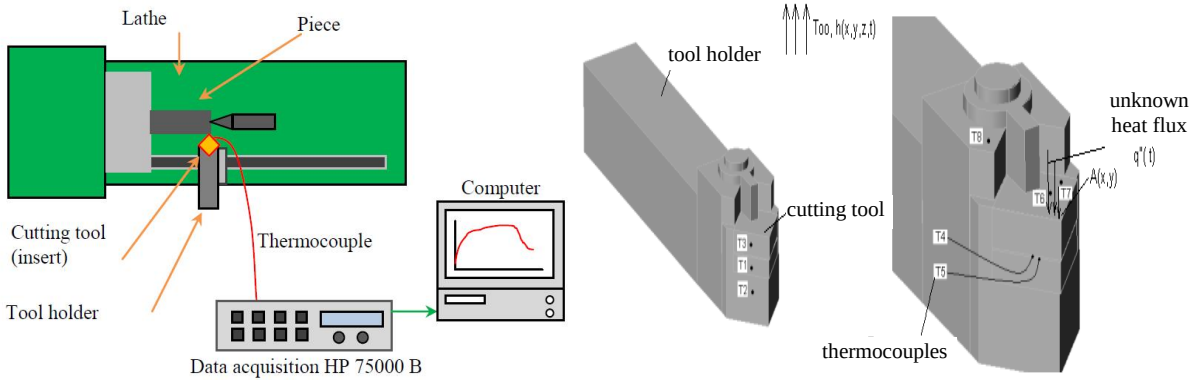


a) b) c)

Figure 2. a) Experimental and estimated heat flux, b) experimental and calculated temperatures c) temperature residuals.

4 Experimental Assembly in a Real Machining Process

The machining test was carried out in a conventional lathe IMOR MAXI–II–520–6CV without coolant. The material used in the experimental test was a cylindrical gray cast iron bar FC 20 EB 126 ABNT of 77 mm in external diameter. The insert and tool holder used were cemented ISO SNUN12040408 K20/Brassinter and ISO CSBNR 20K12/SANDVIK COROMAT, respectively. The temperatures were measured on accessible locations of the insert, the shim and the tool holder by using K type thermocouples and a data acquisition system HP 75000 Series B controlled by a PC (Fig. 3a). Table 1 presents the location of the thermocouples shown in Fig. 3b.



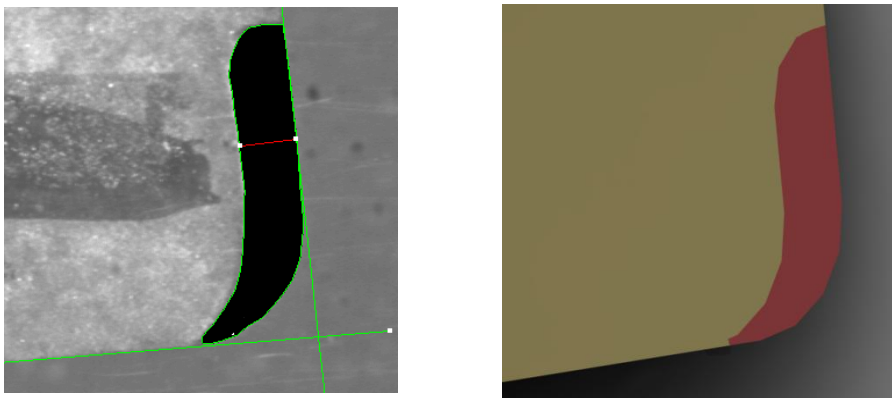
a) b)

Figure 3. a) Experimental apparatus used to acquire the temperature signals in the tool during machining and b) detail of the position of the thermocouple welded to the tool.

Table 1: Locations of the thermocouples shown in Fig. 3b.

Position/Thermocouples	1	2	3	4	5	6	7	8
x [mm]	0.0	0.0	0.0	4.490	6.528	7.222	9.512	5.300
y [mm]	6.45	7.25	3.950	4.116	6.579	4.740	1.715	14.55
z [mm]	-6.55	-11.65	-2.12	-4.83	-4.83	0.0	0.0	9.4

The chip-tool contact area determination represents one of most important and delicate aspects among the main sources of errors in the solution of the thermal model problem. Some methods to identify this area can be found in the literature as, for example, the use of image analyzer software [11] or the application of coatings [12]. In both processes, the area is measured after cutting. This procedure is also used here. However, in this work, the interface contact areas were obtained from the three tests carried out with the same cutting condition. In order to measure the contact area an image system program with video camera Hitachi CCD, KP-110 model, an AMD PC- K6 450 MHz and the GLOBAL LAB image software were used. A typical contact area is presented in Figs. 4a and 4b. The contact area value was 1.41mm^2, obtained for feed rate of 0.138mm/rot, cutting speed of 135.47m/min and depth of cut of 5.0mm.



a) b)

Figure 4. a) Image treatment of the contact area and b) contact area on the computational model.

Many tests were carried out to observe the influence of cutting speed, feed rate and depth of cut in the temperature distribution. However, due to the limitation of pages in this paper the results are presented for only two tests. The test identifications with the cutting conditions are presented in Tab. 2. Each cutting condition was repeated three times to observe the repeatability. In each experiment the total number of measurement of each thermocouple was nt = 180 with a time sample of 0.5s. The thermal conductivity and diffusivity of the tool are respectively, λ = 43.1 Wm/K and α = 14.8 x 10^{-06}m^2/s [13].

Table 2: Cutting conditions.

Cutting parameters	Test 1	Test 2
Feed rate	0.138 mm/rev	0.138 mm/rev
Cutting speed	135.47 m/min	135.47 m/min
Depth of cut	5.0 mm	1.0 mm
Final diameter	72 mm	76 mm

The tool holder is AISI 1045 steel and its thermal properties are λ = 49.8 Wm/K and α = 13.05 x 10^{-06} m^2/s [14]. The support below the tool has the same thermal properties as the tool. All the faces, except the chip-tool interface, were submitted to a constant convection heat transfer coefficient, h = 20 W/m^{2}K. Another important error source that must be taken into account is the thermal contact resistance existing among the tool, the shim and the tool holder. The thermal contact depends on many parameters and conditions such as contact nature, surface properties, pressures, etc. Much effort has been spent in the model and identification technique. Although, this problem is not treated here, the

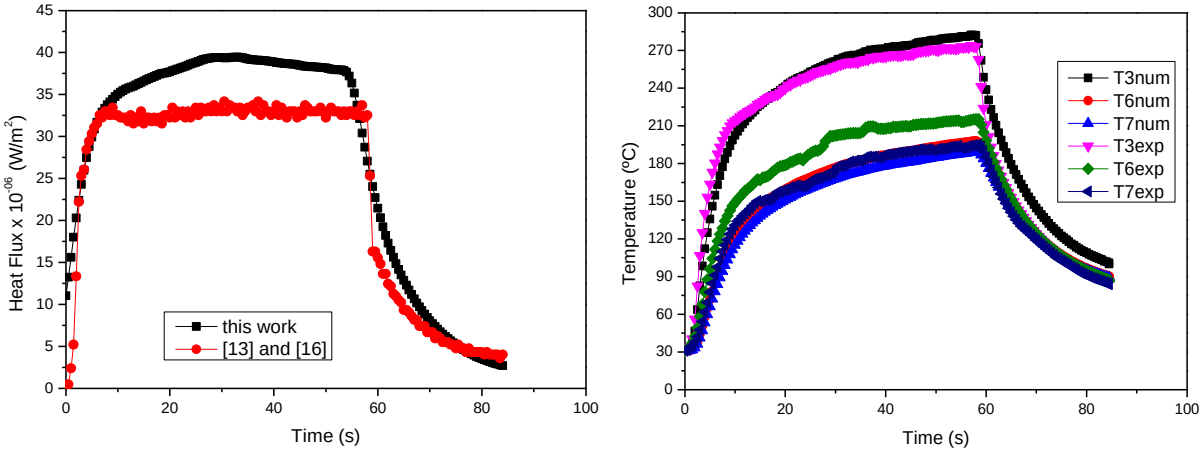
effect of this thermal contact is simulated as a gap of 10μm thickness between the materials involved with the air properties at 300 K. The air thermal properties used are λ = 0.026 Wm/K and α = 22.5 x $10^{-06}m^2/s$ [15].

5 Results Analysis

In this section, the results for the estimation of the heat flux and temperature calculations by using inverse problem technique Specification Function with software COMSOL 4.3 are presented. As aforementioned, for the study of the temperature field in the cutting tool, 3 experiments were carried out with no alterations in the assembly conditions or operations. Each experiment lasted 90s, with temperature reading at every 0.5s, totaling 180 temperature values. It is worth mentioning that in the beginning of the experiment, there is no contact between the tool and the workpiece, therefore the tool is found at uniform room temperature. The cutting time happened between the initial time and 60s. After the 60th second, the cutting is stopped and the tool moves off the workpiece. And it is during the cutting time that heat flux is applied on the tool.

The sensitivity coefficient was calculated numerically with the use of the COMSOL® 4.3 software, as the direct problem, using boundary conditions of heat flux of 1W/m² and initial temperature at 0ºC and an average convection coefficient of 20W/m^{2}K. Many simulations were performed, as in [13], to analyse the influence of the value h = 20W/m^{2}K.

In Figure 5a, the heat flux was estimated for Test 1 (Tab. 2) by using the Specification Function Technique for the future time steps parameter, r = 10. In this figure, a comparison with the heat flux estimated in [13] and [16] is also presented. Tests were carried out with higher and lesser values of future time steps to confirm this value of r. According to the graph, the heat flux is applied from the beginning of the machining process up to approximately 60s. After this time interval, the applied heat flux is null, that is, no machining occurs on the material. In the time interval between 0 and 60s, the average applied heat flux was approximately 35MW/m². The computational time to estimate the heat flux by using the Function Specification technique was 9 minutes and 24 seconds using an Intel Core I7, 6 GB, Windows 7 Ultimate 64 bits. In Figure 5b, a comparison between experimental and calculated temperatures in positions T_3, T_6 and T_7 is presented. In this figure, good results can be seen when comparing estimated and experimental temperatures, especially for thermocouple T_3. Only the results for the highest values of temperature are presented.



a) b)

Figure 5. a) Estimated heat flux and b) Comparison between experimental and estimated temperatures for thermocouples T_3, T_6 and T_7.

Figure 6a presents the estimated heat flux for Test 2 (Tab.2) by using the Specification Function for r = 10. A comparison between experimental and calculated temperatures in positions T_3, T_6 and T_7 is presented in Fig 6b. There is also a good agreement when comparing with authors' previous work (Fig. 6a) and with the experimental temperatures (Fig. 6b). In addition a comparison between the calculated temperatures in positions T_3, T_6 and T_7 for Tests 1 and 2 is presented in Fig. 6c. It may be noticed that the temperatures increase for a higher depth of cut.

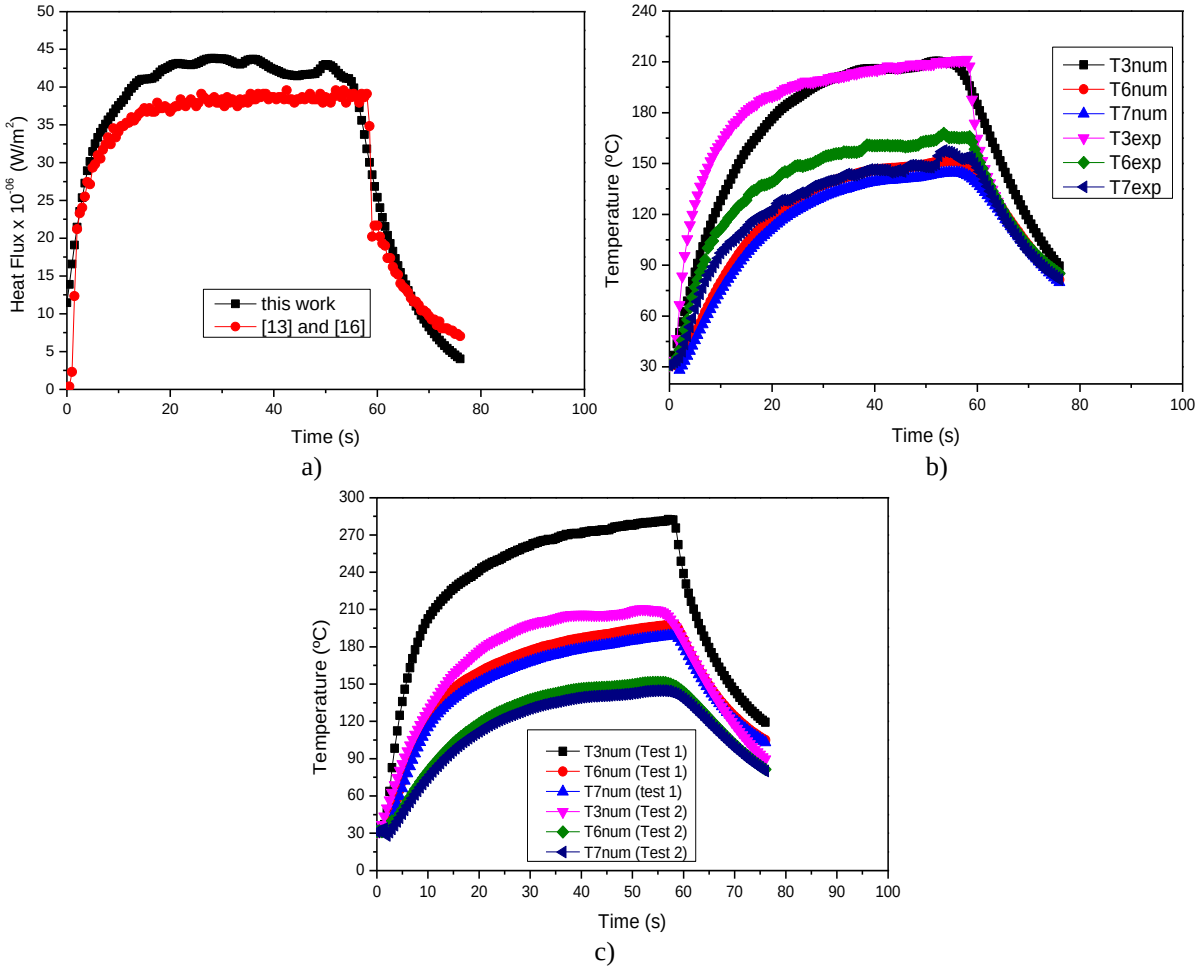
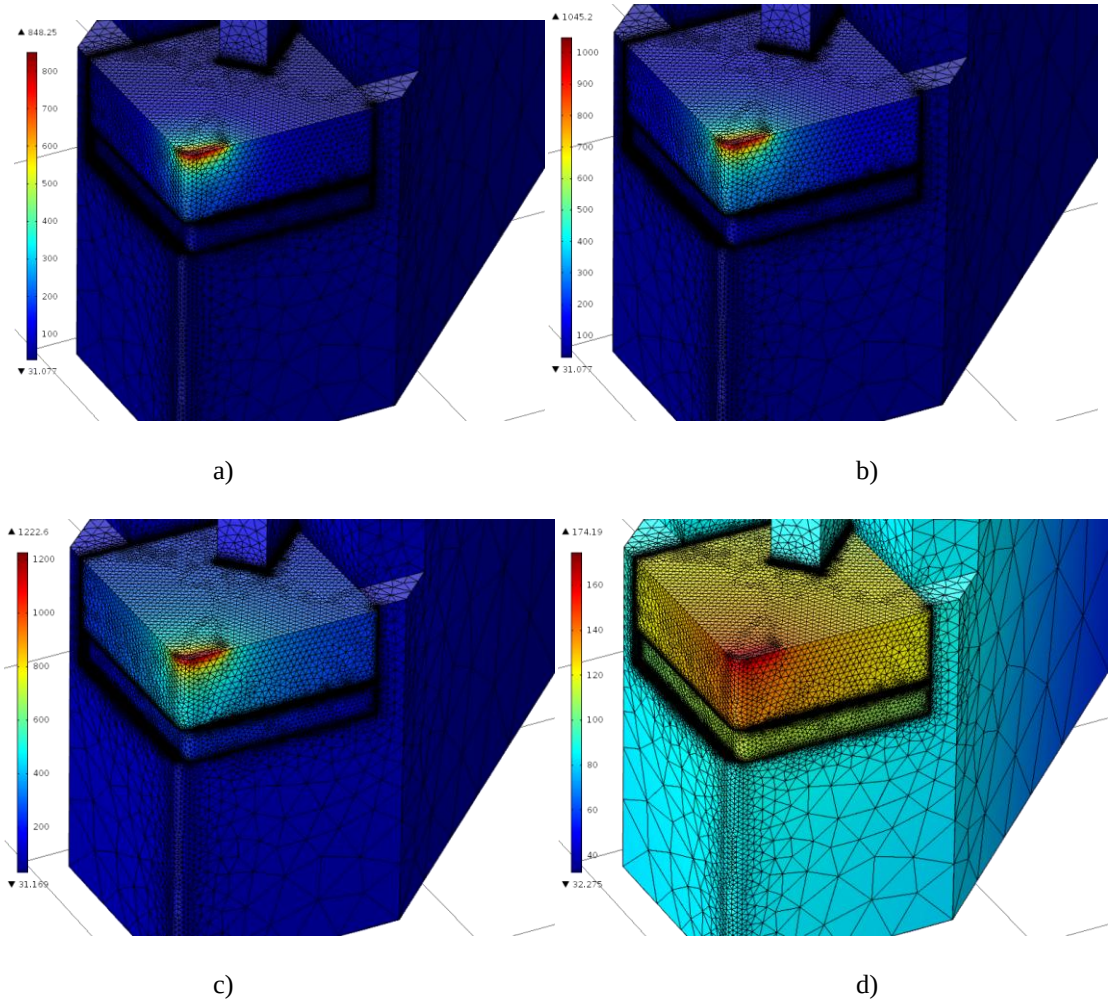


Figure 6. a) Estimated heat flux, b) Comparison between experimental and estimated temperatures for thermocouples T_3, T_6 and T_7 and c) comparison between estimated temperatures in positions T_3, T_6 and T_7 for Tests 1 and 2.

To complete, Figures 7a, 7b, 7c and 7d show a representation of the temperature field in the tool set for Test 1 according to the COMSOL® program at time 5s, 10s, 50s and 80s, respectively. According to Fig. 7c there is a high temperature gradient in the insert.



a) b)

c) d)

Figure 7: Temperature field in the set at instants a) t = 5s, b) 10s, c) 50s and b) t = 80s.

6 Conclusions

The temperature field in any region of the tool set (insert, shim and tool-holder) is calculated from the heat flux estimation at the cutting interface. A significant improvement in the technique to estimate heat flux and temperatures in a machining process was presented in this work. For this, the Inverse Problem Specification Function Method and the COMSOL® 4.3 were joined. In addition, several cutting tests using cemented carbide tools were performed in order to check the model and to verify the influence of the cutting parameters on the temperature field. The use of commercial packages for the numerical resolutions of differential equations that rule the physical phenomenon investigated should be highlighted, for these programs allow adjusting any boundary conditions, as well as modeling the geometry so as to faithfully represent the system investigated.

Acknowledgements

The authors would like to thank CNPq, FAPEMIG and CAPES for their financial support, in special Process 470971/2011-8 sponsored by CNPq.

References

[1] E. M. Trent, P. K. Wright, *Metal cutting*, Butterworth Heinemann, 4th ed., Woburn, United States (2000).

[2] M. Gostimirovic, P. Kovac, M. Sekulic, *An inverse heat transfer problem for optimization of the thermal process in machining*, Sādhanā, 36, (2011), 489-504.

[3] J. V. Beck, B. Blackwell, C. St. Clair, Inverse Heat Conduction: Ill-posed Problems, Wiley-Interscience Publication, New York, (1985).

[4] D. Ulutan, I. Lazoglu, C. Dinc, *Three-dimensional temperature predictions in machining process using finite difference method*, Journal of Materials Processing Technology, 209, (2009), 1111-1121.

[5] L. Wang, Z. L. Sun, X. K. Wang, S. C. Guo, *Numerical and analytical modelling of temperature rise on the machined stainless steel 316L*, International Conference on Mechanic Automation and Control Engineering : (MACE 2010), (2010).

[6] K. Yang,Y.-C. Liang, K.-n. Zheng, Q.-s. Bai, W.-q. Chen, *Tool edge radius effect on cutting temperature in micro-end-milling process*, Int. J. Adv. Manuf. Technol., 52, (2011), 905–912.

[7] K. A. Woodbury, S. Duvvuri, Y. K. Chou, J. Liu, Use of evolutionary algorithms to determine tool heat fluxes in a machining operation. Proceedings of the Inverse Problems Design and Optimization Symposium, IPDO, Miami Beach, Florida, USA, (2007).

[8] V. M. Luchesi, R. T. Coelho, *An inverse method to estimate the moving heat source in machining process*, Applied Thermal Engineering, 45-46, (2012), pp. 64-78.

[9] F. Samadi, F. Kowsary, A. Sarchami, *Estimation of heat flux imposed on the rake face of a cutting tool: A nonlinear, complex geometry inverse heat conduction case study*, International Communication on Heat and Mass Transfer, 39, (2012), 298-303.

[10] C. A. Corrêa Ribeiro, *Uso combinado do software comercial CFX e técnicas de problemas inversos em transferência de calor*, Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Itajubá, Itajubá, Minas Gerais, Brasil, (2012).

[11] T. C. Jen, G. Gutierrez, *Numerical heat transfer analysis in transient cutting tool temperatures*, Proceedings of 34th National Heat Transfer Conference, Pittsburgh, Pennsylvania, August, (2000), pp.20-22.

[12] D. W. Yen, P. K. Wright, *A remote temperature sensing technique for estimating the cutting interface temperature distribution*, Journal of Engineering for Industry, 108, (1986), pp.252-263.

[13] S. R. Carvalho, S. M. M. Lima e Silva, A. R. Machado, G. Guimarães, *Temperature determination at the chip-tool interface using an inverse thermal model considering the tool holder*", Journal of Materials Processing Technology, 179, (2006), pp. 97-104.

[14] W. Grzesik, P. Niesłony, M. Bartoszuk, Modelling of the cutting process analytical and simulation methods, Advances in Manufacturing Science and Technology, 33, (2009), pp. 5-29.

[15] F. P. Incropera, D. P. DeWitt, T. L. Bergman, A. S. Lavine, *Fundamentals of heat and mass transfer*, 6th ed., John Wiley & Sons, USA, (2007).

[16] S. R. Carvalho, M. R. Santos, P. F. B. Souza, G. Guimarães, S. M. M. Lima e Silva, Comparison of inverse methods in the determination of heat flux and temperature in cutting tools during a machining process, High Temperatures High Pressures, 38, (2009), pp. 119-136.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Improvements and Error Analysis on the Different Heat Flux Intensity Method to Estimate Thermal Properties

G. V. V. Campanaro, L. F. S. Carollo, A. L. F. Lima e Silva, S. M. M. de Lima e Silva*

Instituto de Engenharia Mecânica – IEM, Laboratório de Transferência de Calor – LabTC, Universidade Federal de Itajubá - UNIFEI, Av. BPS, 1303, Bairro Pinheirinho, CEP 37500-903, Caixa Postal 50, Itajubá, MG, Brasil
e-mail: guilhermemustang@hotmail.com, felipecarollo@yahoo.com.br, alfsilva@unifei.edu.br, metrevel@unifei.edu.br

Keywords: thermophysical properties, heat transfer, experimental technique, error analysis.

Abstract
In this work, thermal conductivity, λ, and volumetric heat capacity, ρc_p are estimated for a titanium sample. Two different intensities of heat flux were used in the same experiment with the purpose of increasing significantly the sensitivity for the thermophysical property estimation. The purpose of the present work is to present the improvements performed in relation to previous work of the authors. This paper also presents a study on the thermal contact resistance and an analysis of the uncertainties that occur in the simultaneous estimation of λ and ρc_p. The thermal contact resistance is calculated considering the distance between the resistive heater and the sample and a kapton layer of the resistive heater. Satisfactory results are obtained for this analysis because these influences result in a temperature difference of around 0.1 ° C, which is equal to the uncertainty of the thermocouple. The uncertainty analysis is based on the propagation of uncertainty by taking the experimental thermal contact resistance and numerical errors into account. The uncertainty analysis is considered satisfactory because the obtained result was lower than 5%.

1 Introduction

Nowadays, due to globalization, more and more new techniques are required to quickly, reliably and accurately determine the thermophysical properties of materials. Another important aspect is the economic issue, because the lower the cost to determine the parameters, ensuring reliability, the greater the chance to compete in the national and international markets. The technique proposed in this paper can be used, for example, to correctly choose, under the point of view thermal properties, the materials to be used in the manufacture of a heat exchanger. This choice is made by taking into account the values of thermophysical properties, which should be ideal to yield a saving that is directly linked to energy and environmental issues, widely discussed in the current global circumstances. Another example can be a machining process which great part of the heat generated by friction between the workpiece and the cutting tool must be transferred to the tool holder, as the tool wear is directly linked to temperature increase. Thus, the right tool for the process can be chosen through the

knowledge of its thermal conductivity, since this property determines the range of the working temperature of the material. From these needs, researchers have developed many techniques which are being improved continuously [1,2].
There are several techniques to determine the thermophysical properties of diverse materials [3,4,5]. These techniques may determine these properties separately or isolatedly, moreover, most of these estimations occur rapidly, safely and precisely. Hence, an experimental assembly is necessary to perform this estimation. Therefore, when experiments are performed, several uncertainties occur related to the thermal contact resistance as well as in the equipment used to measure the temperature, the heat flux, among others. These uncertainties are intrinsic to the process and cannot be avoided. Therefore, the correct procedure is to perform a controlled experiment to quantify these uncertainties. However, authors like [6,7,8], and others have performed experiments without mentioning these uncertainties. Although these uncertainties are mentioned and quantified in [3,4,5], the method and the process to measure these uncertainties are not described. A method to estimate simultaneously thermal conductivity, λ and volumetric heat capacity, ρc_p, of titanium ASTM B265 is presented in this work.
An uncertainty analysis is also presented in this work as well a study of the contact resistance through the experiment aforementioned to estimate the thermal properties simultaneously. The study on the contact resistance displays the influence of the Kapton layer on the resistive heater and the distance between the heater and the sample. The uncertainty analysis is done by taking into account the influence of the thermal contact resistance and the numerical and experimental temperature errors. In addition, the main purpose of the present work is to present the improvements performed in relation to [9]. The heater, now, is completely symmetrical and the metal samples were rectified.

2 Theory

2.1 Thermal Model and Sensitivity Coefficient

Figure 1 shows the proposed one-dimensional thermal model, which consists of a sample located between a resistive heater and an insulator. The sample has much smaller thickness than its others dimensions and all the surfaces, except the heated ($x = 0$), are isolated to ensure the one-direction heat flux.

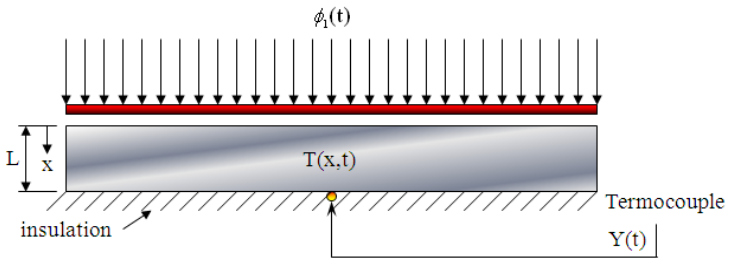


Figure 1: One-dimensional thermal model.

The heat diffusion equation for the problem presented in Figure 1 can be written as:

$$\frac{\partial^2 T(x,t)}{\partial x^2} = \frac{\rho c_p}{\lambda} \frac{\partial T(x,t)}{\partial t}, \tag{1}$$

subject to the boundary conditions:

$$-\lambda \frac{\partial T(x,t)}{\partial x} = \phi_1(t) \text{ at } x = 0, \tag{2}$$

$$\frac{\partial T(x,t)}{\partial x} = 0 \text{ at } x = L, \tag{3}$$

and the initial condition:

$$T(x,t) = T_0 \text{ at } t = 0, \tag{4}$$

where x is the Cartesian coordinate, t the time, ϕ_1 the prescribed heat flux, T_0 the initial temperature of the sample and L the thickness.

Studies of the sensitivity coefficient for each sample are performed in this work in order to determine the ideal region to estimate the properties and the best configuration of the experimental setup. This study provides information such as: the correct positioning of the thermocouples, the experimental time, and the time interval of the applied heat flux incidence. The higher the coefficients value, the better the chance of obtaining the properties reliably. The sensitivity coefficient is defined by the first partial derivative of the temperature in relation to the parameter to be analyzed (λ or ρc_p), being written as follows:

$$X_{ij} = P_i \frac{\partial T_j}{\partial P_i}, \tag{5}$$

where T is the numerical temperature, P the parameter to be analyzed (λ or ρc_p), i the index of parameter, and j the index of points. As in this work, only two properties will be analyzed, $i = 1$ for λ and $i = 2$ for ρc_p.

2.2 Thermal Conductivity and Volumetric Heat Capacity Simultaneous Estimation and Heat Flux Analysis

To estimate the two properties it is necessary to use an objective function. Usually, the objective function is simply the square difference between the temperatures [10,11]. However, since the thermal contact resistance is a systematic error, this influence needs to be considered in the analysis, because this value is constant and permanent. So, this influence was included in the objective function with the purpose of considering an initial error, therefore, the objective function will never be equal to zero. Thus, the objective function used in this work is based on the square difference between the experimental and numerical temperatures plus the influence of the thermal contact resistance. This equation can be written as:

$$F = \left[\left(R''_{c,1} + R''_{c,2}\right)\phi_a\right]^2 + \sum_{j=1}^{m}\left(Y_j - T_j\right)^2, \tag{6}$$

To obtain the values for λ and ρc_p in each experiment, the BFGS (Broydon Fletcher Goldfarb Shanno) sequential optimization technique, presented in [12], was used. This technique is a particularity of the variable metric methods. The advantages of this technique are its fast convergence and readiness for working with many design variables. Because it is a first order method, it is necessary to know the gradient of the objective function.

3 Experimental Apparatus

The experimental apparatus used to determine the properties of ASTM B265 Grade 2 Titanium is shown in Fig. 2. The ASTM B265 Grade 2 Titanium plate is 50.00 x 49.98 x 8.84 mm in dimension.

The 44.5 x 44.5 x 0.25 mm resistive kapton heater has a resistance of 15 Ω and was used due to its small thinness, allowing faster overall warming. This heater was connected to a digital power supply Instrutemp ST – 305D-II to provide the necessary heat flux. In this work, different intensities of heat flux were used in the same experiment as an attempt to achieve the best condition to estimate the properties simultaneously in accordance to the analyses of the sensitivity coefficients. To achieve this heat flux condition, the digital power supply has a configuration that allows working with parallel or series connection. Then, the series condition was used to provide the highest heat flux for the first period of the experiment, and the parallel condition to supply the lowest heat flux for the second part. A symmetrical assembly was used to minimize the errors in the measurement of the heat flux to be generated on the sample surface. In addition, the applied current value was measured by the calibrated multimeter Minipa ET-2042C and weights were used on top of the isolated sample-heater set to improve the contact between the components. To ensure a one-direction flux and minimize the effect of convection caused by the air circulating in the environment, the sample-heater set was isolated with polystyrene plates. Temperatures were measured using type K thermocouples (30AWG) welded by capacitive discharge in the middle of the bottom surface (25.0 x 25.0 x 8.84 mm) and calibrated by using a bath temperature calibrator Marconi MA 184 with a resolution of ± 0.01 °C. These thermocouples were connected to Agilent 34980A data acquisition set controlled by a microcomputer. In order to obtain better results, all experiments were performed in controlled room temperature.

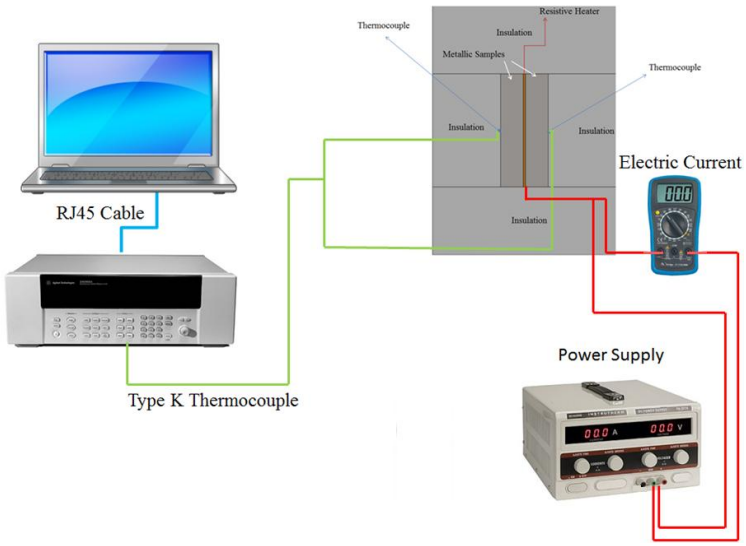


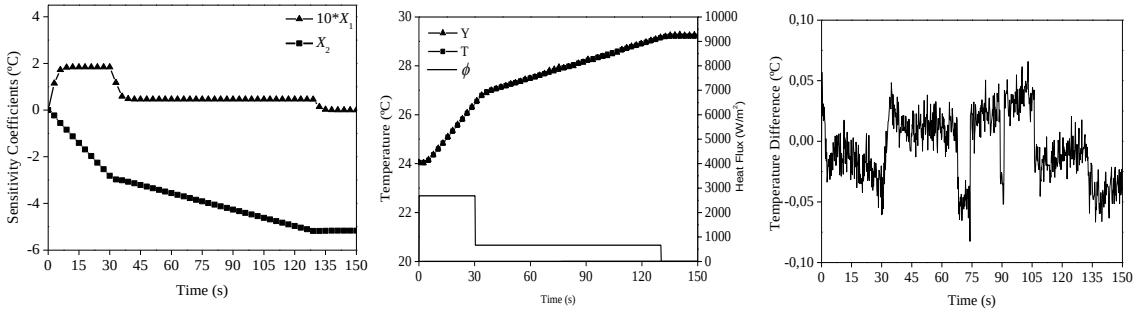
Figure 2: Sketch of experimental apparatus used to determine the properties.

4 Results and Discussion

Twenty experiments were performed to simultaneously estimate the thermal conductivity and the volumetric heat capacity of a titanium sample. This number of experiments was done in order to obtain reliable estimates of standard deviation and average of the data. According to the literature this number has to be at least 20 experiments [13]. Each experiment lasted 150 s, and the heat flux was imposed from 0 to 130 s. From 0 to 30 s, the imposed heat flux was approximately 2682 Wm^{-2}; from 30 to 130 s, the imposed heat flux was around 664 Wm^{-2}. The time interval used to monitor the temperature was 0.2 s. To guarantee the hypotheses of constant thermal properties, this configuration for the heat fluxes was chosen to keep the temperature difference lower than 5 K. This temperature difference is based on

the difference between the final and initial temperatures which are measured having the thermocouple on the same position of the sample. The sensitivity analyses were performed to determine the best region to estimate the properties. These analyses were performed by using the values of λ and ρc_p obtained from [14] and the parameters described above. In addition, several analyses of the objective function (Eq. 6) with the sensitivity coefficient analyses were performed to determine the properties in the selected region [9]. This selected region corresponds to a set of points which provides accurate thermal properties estimation. Since this estimation presents an accurate result, it can be said that this region of points presents enough influence to determine these properties. Figure 3a shows the sensitivity coefficients at $x = L$ for λ and ρc_p.

X_1 represents the sensitivity coefficient for λ and X_2 represents the sensitivity coefficient for ρc_p, both on the isolated surface. The former is multiplied by a factor to improve the visualization of the curve. In Figure 3a, one can see that X_1 increases in the first 10 s and remains constant thereafter until the change of heat flux; and X_2 increases proportionally with the temperature. Because of this behavior, the highest heat flux was imposed in the first period of time, resulting in high sensitivity for λ. The lowest heat flux was imposed in the second part to increase the sensitivity for ρc_p and maintain the sensitivity for λ. Figure 3b presents the distribution of experimental and numerical temperatures for the sample at $x = L$ and the imposed heat flux at $x = 0$. The numerical temperature is achieved by applying the values of the estimated properties, λ and ρc_p, from one of the accomplished experiments. These temperatures present good agreement, which can be proved by analyzing the temperature residuals shown in Fig. 3c. These residuals are calculated by the difference between the experimental and numerical temperatures. The maximum value found was lower than 0.08 ºC, which is much lower than the thermocouple uncertainty. This difference will be considered as the imperfection of the thermal insulation in the uncertainty analysis.



a) b) c)

Figure 3: a) Sensitivity coefficients as a function of time for titanium sample, b) experimental (Y) and numerical (T) temperatures with heat flux (ϕ) as a function of time for titanium sample and c) Temperature difference $Y - T$ for titanium sample.

A significant improvement obtained in this work concerns the difference between the measured temperatures in $x = L$ on both samples used for the assembly to guarantee the symmetrical heat flux (Fig. 2). In [9] this difference presented an average error of approximately 0.2 ºC, with a maximum difference of 0.4 ºC, as may be seen in Figs. 4a and 4b. Thus, three significant changes were made to minimize this difference. The first concerns the asymmetry presented as a manufacturing error of the resistive heater. All Omega heaters previously used displayed this small asymmetrical heat flux when the experiment is performed under unsteady; this problem did not happen under steady condition. For the solution of this problem, resistive heaters purchased from Laboratório de Meios Porosos e Propriedades Termofísicas da Universidade Federal de Santa Catarina were used. Several tests were performed to verify whether the problem of asymmetry had been solved. One of these results will be described in Fig. 5a and 5b.

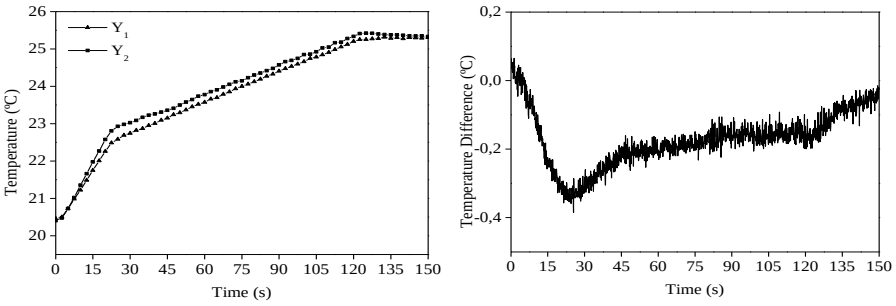


Figure 4: a) Experimental temperatures comparison in $x = L$ for titanium sample and b) temperature difference $Y_1 - Y_2$ for titanium sample [9].

All the samples used for the measurements were rectified aiming to minimize the thermal contact resistance between the heater and the samples. Therefore, Silver Artic thermal paste used in [9] was unnecessary. This procedure reduced the great uncertainty added to the value of the estimated thermal properties, mainly for the low values displayed by this thermal paste. Finally, the assembly of the samples was made in the vertical to avoid the influence one sample weight on the other. Figure 5a presents a comparison of the experimental temperature at $x = L$ for the titanium as shown in Fig. 4a. As already mentioned above, the purpose of these procedures was to solve the asymmetry of the temperatures measured at $x = L$. The effective proof of this significant improvement may be verified in Fig. 5b. It may be observed from this figure that the average of the temperature differences is close to 0 ºC and as expected, the difference between these temperatures vary randomly; whereas in Fig. 5b, the difference shows a systematic error, especially in the region of highest heat.

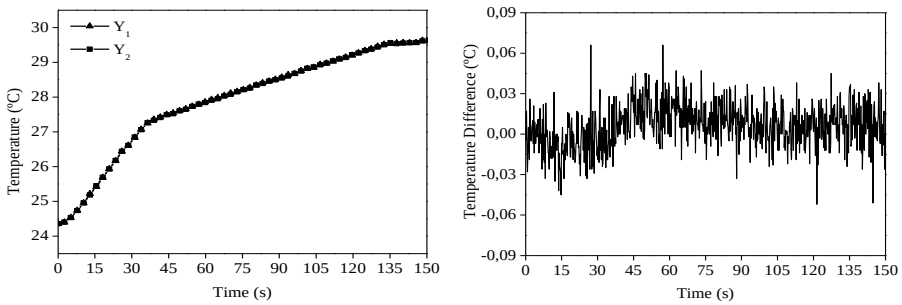


Figure 5: a) Experimental temperatures comparison in $x = L$ and b) temperature difference $Y_1 - Y_2$ for titanium sample.

Table 1 displays the estimated property results (λ and ρc_p) for all the 15 experiments carried out on titanium ASTM B265 Grade 2. To prove the symmetric assembly, these properties were estimated for the two temperatures signals measured on the surface of the two samples used at $x = L$. In this table, Th_1 stands for Thermocouple 1 and Th_2 for thermocouple 2. It may be seen that in several cases the results of both properties stand very close to each other, mainly for ρc_p. Table 2 presents the average value, the standard deviation and the percent difference between the estimated average value and the literature values for λ and ρc_p on titanium ASTM B265 Grade 2. It may be observed that the average values for λ and ρc_p for both thermocouples were practically the same. One may see that the results are in accordance with the work found in literature due to the low standard deviation and small percent difference (Tab. 2).

Table 1: Obtained values for ASTM B265 Grade 2 Titanium samples.

Experiment	λ (W/mK)		ρc_px10^{-6} (Ws/m^{3}K)	
	Th_1	Th_2	Th_2	Th_2
1	17.71	17.71	2.65	2.64
2	17.71	17.71	2.64	2.64
3	17.00	16.39	2.60	2.60
4	17.71	17.71	2.64	2.65
5	17.71	17.71	2.65	2.64
6	16.98	16.58	2.67	2.66
7	17.10	17.71	2.60	2.61
8	17.71	17.71	2.63	2.63
9	17.71	17.71	2.66	2.65
10	17.71	17.71	2.64	2.64
11	17.11	18.93	2.67	2.68
12	17.71	16.81	2.66	2.67
13	19.47	18.73	2.69	2.68
14	17.71	17.18	2.66	2.67
15	16.44	17.71	2.60	2.61
16	17.71	17.71	2.62	2.61
17	17.71	17.72	2.65	2.61
18	17.71	17.72	2.64	2.66
19	17.72	20.61	2.64	2.66
20	17.72	17.72	2.65	2.67

Table 2: Statistic values obtained for the ASTM B265 Grade 2 Titanium.

Thermocouple	Property	Present work	[14]	S. D.	Difference (%)
1	ρc_px10^{-6} (Ws/m^{3}K)	2.64	2.66	± 0.024	0.75
	λ (W/mK)	17.60	18.06	± 0.57	2.54
2	ρc_px10^{-6} (Ws/m^{3}K)	2.64	2.66	± 0.025	0.75
	λ (W/mK)	17.77	18.06	± 0.88	1.63

5 Thermal Contact Resistance and Uncertainty Analysis

The thermal contact resistance was analyzed with the purpose to find out if there is a significant influence during the temperature measurements. This study was divided into two parts: the first part considered the thermal resistance caused by the applied thermal compound; the second part took into account the influence of the kapton layer present in the resistance heater. In [9] the effect of the thermal resistance contact due to the thermal paste ($R''_{c,1}$) was added as a systematic error to the objective function as presented in Eq. (6). However, since the samples used in this work were rectified and the thermal paste was not used, this effect was disregard in Eq. (6); only the effect of the kapton layer of the resistive heater was considered. The influence of the kapton layer was analyzed. An

Optical Microscope Jenavert Zeiss (2000x) with the Image Analyzer Olympus Model TVO.5XC-3 was used to measure the thickness of the kapton on the heater. The kapton layer presents a thickness of 10.64 x10-6 m since the thermal conductivity of the kapton is 0.12 $Wm^{-1}{}^{o}C^{-1}$ [15]; a thermal contact resistance equal to $88.67x10^{-6}$ $m^{2o}CW^{-1}$ was obtained. By considering the applied heat flux average, 1117 Wm^{-2}, this thermal contact resistance corresponds to a temperature difference of 0.10 ºC.
An uncertainty analysis was done to verify whether the obtained results were reliable. This analysis was based on the uncertainty propagation procedure. In this procedure, it is necessary to decide which errors will be analyzed. In this work the procedure to determine the uncertainty in the estimation of λ and ρc_p is based on linear propagation of uncertainties of the variables: temperature measurement, the imposed heat flux, measurement instruments, thermal contact resistance, imperfection of thermal insulation, and the numerical errors (BFGS and finite difference method. As the thermocouple was located in the middle of the bottom surface (opposite the heater), the uncertainty imposed by the position of the thermocouple was disregarded. The hypothesis of linear propagation is used because the objective function is based on the difference between the experimental and numerical temperatures. This analysis is in accordance with the theory of error propagation extracted from [16]. The uncertainty of the estimation can be calculated based on the objective function, which is composed of the experimental and numerical temperature, and the BFGS method. Thus:

$$\begin{aligned} U_{final}^2 &= U_Y^2 + U_T^2 + U_{BFGS}^2 \\ U_{final}^2 &= U_{aquisitiondata}^2 + U_{thermocouple}^2 + U_{contactresistance}^2 + U_{thermalinsulation}^2 \\ &+ U_{currentmultimeter}^2 + U_{voltagemultimeter}^2 + U_{FDM}^2 + U_{BFGS}^2 \end{aligned} \quad (7)$$

After the final uncertainty has been defined, it is necessary to quantify the partial uncertainty. To define these values, the authors decided to consider each uncertainty divided by the mean value of the analyzed parameter. As the variations of the uncertainties considered in this work were very small, the uncertainties were considered constant. The data acquisition uncertainty is calculated from the resolution of the equipment which is 0.01 ºC and the maximum difference of the temperature which is approximately 5.00 ºC. This temperature difference is based on the difference between the initial and the highest temperatures which are measured by having the thermocouple on the same position on the sample. Thus, this uncertainty is:

$$U_{aquisitiondata} = \frac{0.01}{5.00} = 0.2\,\%, \quad (8)$$

Now the uncertainty of the thermocouple is calculated by considering the a oscillation of 0.1 ºC and the same difference of average temperature of 5.0 ºC. Thus:

$$U_{thermocouple} = \frac{0.10}{5.00} = 2.0\,\%, \quad (9)$$

The uncertainty due to the thermal contact resistance of 0.1 ºC was estimated in Section 6. Therefore:

$$U_{contactresistance} = \frac{0.10}{5.00} = 2.0\,\%, \quad (10)$$

As aforementioned, the maximum difference between the numerical and experimental temperatures of 0.08 °C presented in Fig. 6 was considered as uncertainty due to the imperfection of the thermal isolation; hence:

$$U_{thermalisnsulation} = \frac{0.08}{5.00} = 1.60\,\%, \quad (11)$$

For the calculation of the multimeter uncertainty, the resolution of the digital device divided by the average current and the average resistance is used; the values are 0.67 A and 15.0 Ω. It is worth highlighting that this multimeter was used in the measurement of the current value, with a resolution of 0.01 A and the resistance with a resolution of 0.1 Ω. In this manner, the following equations are achieved:

$$U_{currentmultimeter} = \frac{0.01}{0.67} = 1.49\,\%, \tag{12}$$

$$U_{resistancemultimeter} = \frac{0.1}{15.0} = 0.67\%, \tag{13}$$

Finally, the uncertainty of the mathematical methods used must be quantified. For the BFGS method the error of 0.01 ºC was adopted and for the FDM 0.05 ºC as base values for the definition of the uncertainty; thus the following is obtained:

$$U_{BFGS} = \frac{0.01}{5.00} = 0.2\,\%, \tag{14}$$

$$U_{FDM} = \frac{0.05}{5.00} = 1.0\,\%, \tag{15}$$

Once all the partial uncertainties have been calculated, it is possible to determine the uncertainty of the thermal properties by substituting Eqs. (8) to (15) in Eq. (7).

$$\begin{aligned} U_{final}^2 &= 0.2^2 + 2.00^2 + 2.00^2 + 1.60^2 + 1.49^2 + 0.67^2 + 0.20^2 + 1.0^2, \\ U_{final} &= 3.78\,\%, \end{aligned} \tag{16}$$

As it can be seen, the uncertainty value is in accordance with that found in the literature, because the presented value is lower than 5 %.

6 Conclusions

This paper presented a significant improvement in the technique that uses different intensities of heat flux in the same experiment to estimate the thermal conductivity and the volumetric heat capacity of metal samples simultaneously. The estimated properties are in good agreement with the literature. This affirmation can be proved due to the small difference between the literature and estimated values, and the low standard deviation. In addition, an error analysis based on thermal contact resistance and uncertainty analysis in the estimation of λ and ρc_p was presented. Other contributions of this work are: the low cost of the experimental apparatus in terms of the equipment used, the facility to assemble this apparatus, the short duration of the experiment, the small size of the samples, and the study of the global minimum value along with the sensitivity.

Acknowledgements

The authors would like to thank CNPq, FAPEMIG and CAPES for their financial support, in special Process APQ-1940-11 sponsored by FAPEMIG and Process 470971/2011-8 sponsored by CNPq.

References

[1] R. F. Brito, S. R. Carvalho, S. M. M. Lima e Silva, J. R. Ferreira, *Thermal analysis in coated cutting tools*, Int. Commun. Heat Mass Transfer, 36, (2009), 314-321.

[2] S. R. Carvalho, M. R. Santos, P. F. B. Souza, G. Guimarães, S. M. M. Lima e Silva, *Comparison of inverse methods in the determination of heat flux and temperature in cutting tools during a machining process*, High Temperatures-High Pressures, 38, (2009), pp. 119-136.

[3] Y. Jannot, A. Degiovanni, G. Payet, 2009, *Thermal conductivity measurement of insulating materials with a three layers device,* International Journal of Heat and Mass Transfer, 52, (2009), pp. 1105-1111.

[4] J. Xamán, L. Lira, J. Arce, *Analysis of the temperature distribution in a guarded hot plate apparatus for measuring thermal conductivity*, Applied Thermal Engineering, 29, (2009), pp. 617-623.

[5] M. Thomas, N. Boyard, N. Lefèvre, Y. Jarny, D. Delaunay, *An experimental device for the simultaneous estimation of the thermal conductivity 3-D tensor and the specific heat of orthotropic composite materials*, International Journal of Heat and Mass Transfer, 53, (2010), pp. 5487-5498.

[6] R. Le Goff, D. Delaunay, N. Boyard, Y. Jarny, T. Jurkowski, R. Deterre, *On-line temperature measurements for polymer thermal conductivity estimation under injection molding conditions*, International Journal of Heat and Mass Transfer, 52, (2009), pp. 1443-1450.

[7] E. D. Kravvaritis, K. A. Antonopoulos, C. Tzivanidis, C., *Experimental determination of the effective thermal capacity function and other thermal properties for various phase change materials using the thermal delay method*, Applied Energy, 88, (2011), pp. 4459-4469.

[8] C. S. Sanjaya, T. Wee, T. Tamilselvan, *Regression analysis estimation of thermal conductivity using guarded-hot-plate apparatus*, Applied Thermal Engineering, 31, (2011), pp. 1566-1575.

[9] L. F. S. Carollo, A. L. F. Lima e Silva, S. M. M. Lima e Silva, *Applying different heat flux intensities to simultaneously estimate the thermal properties of metallic materials*, Measurement Science & Technology, 23, (2012), pp. 01-11.

[10] M. H. Adjali, M. Laurent, *Thermal conductivity estimation in non-linear problems*, Int. J. Heat Mass Transfer, 50, (2007), 4623-4628.

[11] V. L. Borges, P. F. B. Souza, G. Guimarães, *Experimental determination of thermal conductivity and diffusivity using a partially heated surface method without heat flux transducer,* Inverse Problems in Science and Engineering, 16, (2008), pp. 1047-1067.

[12] G. N. Vanderplaats, *Numerical optimization techniques for engineering design*, 4th ed., McGraw Hill, USA, (2005).

[13] J. P. Holman, *Experimental methods for engineers*, 7th ed., McGraw-Hill Book Company, New York, (2001).

[14] GMTTitanium, ASTM B265 Grade 2 Titanium, Avaliable at: www.gmttitanium.com, Access date: February 22, (2010).

[15] Matweb, Kapton, Available at: www.matweb.com, Access date: February 16, 2012.

[16] J. R. Taylor, *An introduction to error analysis: the study of uncertainties in physical measurements*, 2nd edition, University Science Books.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

A Backward Reciprocity Function to Estimate Spatial and Transient Unknown Heat Fluxes in Multilayered Composites using Non-Intrusive Measurements

Marcelo J. Colaço[1*], Carlos J. S. Alves[2]

[1]Department of Mechanical Engineering
Federal University of Rio de Janeiro, PEM/COPPE, Cx. Postal 68503, CEP 21941-972, Rio de Janeiro, RJ, Brazil
e-mail: colaco@ufrj.br

[2]Department of Mathematics
Technical University of Lisbon, Av. Rovisco Pais, 1, Lisbon, 1049-001, Portugal
e-mail: carlos.alves@math.ist.utl.pt

Key words: reciprocity function, double layered composites, transient and spatial contact failures, non-intrusive method.

Abstract
This paper presents a computationally fast methodology to estimate spatial and transient unknown heat fluxes between sandwiched materials, using non-intrusive measurements. The methodology is formulated in such a way that the spatial and temporal variation of the unknown quantity is obtained simply by solving a linear system, whose solution vector is composed of some integrals containing the measured temperatures and applied heat flux at an external boundary. Good estimates are obtained, even for functions containing discontinuities. Simulated measurements with and without errors are considered, showing promising good results.

1 Introduction

The non-destructive evaluation of contact failures in composite materials is important in several applications [1,2]. In fact, different areas, such as electronic packaging, nuclear reactors, aerospace and biomedicine, among others, rely on efficient heat removal mechanisms, which are strongly dependent on the perfect contact among the layers of a system. When there is a total contact failure between two layers of a material, the heat flux at such interface will have a very small value, whereas for a perfect contact, this heat flux will be large.

There are different types of damages in laminated composite materials, which are generally classified as [2]: delamination or disbound, debound, and kissing bounds. Delamination or disbound refers to the

lack of contact between two layers of a laminated composite. When such failure is corrected and the defect appears once again, it is called debond. Finally, kissing bounds are failures related to the lack of adhesion between the layers. Kissing bounds are more difficult to be detected, since the layers are still in contact. The technique described in this paper is related to the first two types of defects, where the two layers might have an intermittent contact along their interface.

Some previous studies were related to the estimative of a time and/or spatial dependent variation of the thermal contact conductance at the contacting interface between two materials. Thermal contact conductance is defined as the ratio between the heat flux and the temperature jump at the interface between two materials. Fieberg and Kneer [3], motivated by the study of temperature distribution in internal combustion engine components, solved an inverse heat conduction problem to estimate the heat flux at the interface between two solids and used temperature measurements at the interface to estimate the thermal contact resistance (defined as the reciprocal of the thermal contact conductance). In their work, the measurements were taken by an infrared camera pointed at the interface between the two materials, which required access to such location. As a result, only a time dependent global contact resistance with constant spatial distribution could be estimated. Also motivated by automotive applications, Huang and Ju [4] estimated the transient periodic thermal contact conductance between the exhaust valve and its seat in an internal combustion engine, using intrusive measurements. Yang [5] also used an inverse heat conduction problem to estimate a time dependent contact resistance in single-coated optical fibers. Although good results were obtained, intrusive measurements were required. No spatial variation was considered. Also based on intrusive methods, Ozisik *et al*. [6] and Orlande and Ozisik [7] used the conjugate gradient method of function estimation with adjoint problem formulation in order to estimate the contact conductance during a solidification process and for periodically contacting surfaces, respectively.

Shojaeefard and Goudarzi [8] estimated the thermal contact resistance for various levels of pressure applied to composite materials using intrusive measurements. The same authors mentioned that values of thermal contact resistances presented in the literature differ significantly, probably due to the fact that they derive from different experimental conditions. Shojaeefard *et al*. [9,10] also estimated a time dependent thermal contact conductance, using the conjugate gradient method with adjoint problem, but again using intrusive measurements.

Gill *et al*. [11] solved an inverse heat conduction problem to estimate the spatial distribution of the thermal contact resistance. The authors mentioned that several models consider such resistance constant in space, although it actually varies spatially. The results obtained by the authors were very sensitive to measurement errors and required temperatures measurements very close to the interface. However, the main contribution was to estimate the spatial variation of the thermal contact resistance instead of using a constant value.

Other authors [12-15] also obtained the overall thermal contact resistance between different materials, using intrusive measurements.

In this paper we describe a non-intrusive method for determining the heat flux at the interface between two materials. The method is based on the reciprocity functional approach [16-19] and does not involve any iterative procedure. The method starts solving some auxiliary problems, which do not depend on the unknown quantity. Once these problems are solved, several different functional forms of the heat flux at the inaccessible interface can be determined solving a linear system, whose solution vector is composed of some integrals of the imposed heat flux and measured temperatures at an outer surface.

This paper is related to our previous works [17-19]. In [17,19] we presented a non-intrusive method for estimating thermal contact resistances in steady-state problems and in [18] we presented an

extension for transient problems. This paper is a completely new approach to transient problems, since it does not uses any iterative procedure, as we did in [18].

2 Physical Problem

Lets consider a generic domain Ω, divided in three parts so that $\Omega = \Omega_1 U \Gamma U \Omega_2$, where Ω_1 is a domain with a thermal conductivity K_1 and thermal diffusivity α_1, Ω_2 is a domain, with a thermal conductivity K_2 and thermal diffusivity α_2 that is in contact with Ω_1, and Γ is the contact interface between them. In this paper, it is considered that thermal conductivities and diffusivities are known constants. The boundary of Ω_1 is $\partial\Omega_1 = \Gamma_0 U \Gamma_1 U \Gamma$, where the surface Γ_0 is subjected to a prescribed heat flux q and its temperature is measured. Γ_1 is the lateral surface of Ω_1 and Γ is the contact surface between Ω_1 and Ω_2. Similarly, the boundary of Ω_2 is $\partial\Omega_2 = \Gamma_{00} U \Gamma_2 U \Gamma$, where Γ_{00} is the lower surface, Γ_2 is the lateral surface of Ω_2 and Γ is the contact surface. Fig. 1 shows the geometry for a three-dimensional case.

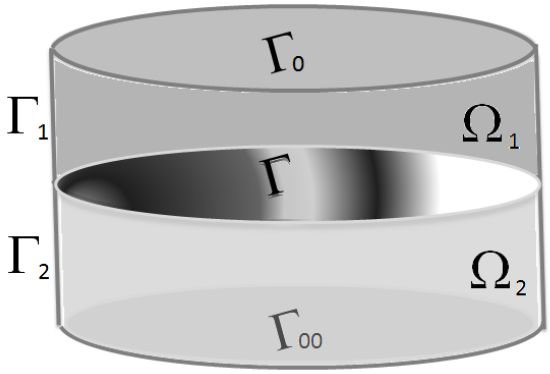


Figure 1: Geometry for a three-dimensional case

The lateral surfaces $\Gamma_1 U \Gamma_2$ are assumed to be thermally insulated while the lower surface Γ_{00} is subjected to a prescribed temperature. The measurement surface Γ_0 is assumed to have a prescribed heat flux q imposed on it. The contact surface Γ is assumed to have a Robin boundary condition, i.e., $-K_1 \partial T_1/\partial \mathbf{n} = h(T_1 - T_2) = K_2 \partial T_2/\partial \mathbf{n}$, where $\mathbf{n}$ is the normal derivative outward the boundary, K_1 is the thermal conductivity of region 1, T_1 and T_2 are the temperatures at the interface of domains Ω_1 and Ω_2 , respectively, and h is the thermal contact conductance.

The statement of the interface transient heat transfer problem for constant conductivities K_1 and K_2 can be formulated as the following direct problem:

$$\frac{1}{\alpha_1}\frac{\partial T_1}{\partial t} = \nabla^2 T_1 \qquad \text{in } \Omega_1, \text{ for } t>0 \qquad (1.a)$$

$$-K_1 \frac{\partial T_1}{\partial \mathbf{n}_1} = q \qquad \text{at } \Gamma_0, \text{ for } t>0 \qquad (1.b)$$

$$\frac{\partial T_1}{\partial \mathbf{n}_1} = 0 \qquad \text{at } \Gamma_1, \text{ for } t>0 \qquad (1.c)$$

$$-K_1 \frac{\partial T_1}{\partial \mathbf{n}_1} = h\left(T_1 - T_2\right) \qquad \text{at } \Gamma, \text{ for } t>0 \qquad (1.d)$$

$$\frac{1}{\alpha_2}\frac{\partial T_2}{\partial t} = \nabla^2 T_2 \qquad \text{in } \Omega_2, \text{ for } t>0 \qquad (1.e)$$

$$\frac{\partial T_2}{\partial \mathbf{n}_2} = 0 \qquad \text{at } \Gamma_2, \text{ for } t>0 \qquad (1.f)$$

$$T_2 = 0 \qquad \text{at } \Gamma_{00}, \text{ for } t>0 \qquad (1.g)$$

$$K_2 \frac{\partial T_2}{\partial \mathbf{n}_2} = -K_1 \frac{\partial T_1}{\partial \mathbf{n}_1} \qquad \text{at } \Gamma, \text{ for } t>0 \qquad (1.h)$$

$$T_1 = T_2 = 0 \qquad \text{for } t=0 \text{ in } \Omega_1 \cup \Omega_2 \qquad (1.i)$$

The *inverse problem* under analysis in this paper consists of estimating the normal derivative $\partial T_1/\partial \mathbf{n}$ at the inaccessible contact surface Γ, by using some extra transient temperature measurements Y at the boundary Γ_0. For the inverse problem all the other quantities appearing in the formulation of the physical problem are assumed as deterministically known, except h, T_1 and T_2 at the boundary Γ.

In this paper we will present a formulation that allows the determination of such unknown function using a backward reciprocity function. This formulation is based on the idea of the reciprocity functional [16-19], extended to a transient case. The approach used here is different from our previous paper [18] in the sense that a non-iterative procedure is used in this paper. The sought function can then be obtained simply solving a linear system, whose solution vector is composed of some integrals in space and time of a function that depends on Y and q. The method firstly solves some auxiliary problems that do not depend on the unknown function. Once this pre-processing is done, such result can be used to estimate different variations of the unknown function in a post-processing stage, that does not require any iterative procedures and can recover the entire field of interest (in time and space) in short CPU time. Notice that the pre-processing stage can be done in advance and the results stored for different estimates.

3 Backward Reciprocity Function

Let us write the following *auxiliary problem* for some functions $G_{1,j}$, for domain Ω_1:

$$\frac{1}{\alpha_1}\frac{\partial G_{1,j}}{\partial t} = -\nabla^2 G_{1,j} \qquad \text{in } \Omega_1, \text{ for } 0<t<t_f \qquad (2.a)$$

$$G_{1,j} = \psi_j \qquad \text{at } \Gamma_0, \text{ for } 0<t<t_f \qquad (2.b)$$

$$\frac{\partial G_{1,j}}{\partial \mathbf{n}_1} = 0 \qquad \text{at } \Gamma_1, \text{ for } 0<t<t_f \qquad (2.c)$$

$$\frac{\partial G_{1,j}}{\partial \mathbf{n}_1} = 0 \qquad \text{at } \Gamma, \text{ for } 0<t<t_f \qquad (2.d)$$

$$G_{1,j} = 0 \qquad \text{for } t=t_f \text{ in } \Omega_1 \cup \Omega_2 \qquad (2.e)$$

Notice that this problem includes (2.e) a final condition (at $t=t_f$) instead of an initial condition, and it must be solved backwards. Also, the boundary condition at Γ_0 is an imposed function $\psi_j(x,t)$, which is taken as an element of an orthonormal basis function, on the spatial component.

Let us now write the following identity, concerning the boundary Ω_1, where Eqs. (1.a) and (2.a) are used:

$$\int_{\Omega_1}\left(G_{1,j}\nabla^2 T_1 - T_1\nabla^2 G_{1,j}\right)d\Omega_1 = \int_{\Omega_1}\left(G_{1,j}\frac{1}{\alpha_1}\frac{\partial T_1}{\partial t} + T_1\frac{1}{\alpha_1}\frac{\partial G_{1,j}}{\partial t}\right)d\Omega_1 \tag{3}$$

Using Green's second identity, we can then obtain:

$$\int_{\Gamma_0\cup\Gamma_1\cup\Gamma}\left(G_{1,j}\frac{\partial T_1}{\partial \mathbf{n}_1} - T_1\frac{\partial G_{1,j}}{\partial \mathbf{n}_1}\right)d\left(\Gamma_0\cup\Gamma_1\cup\Gamma\right) = \int_{\Omega_1}\left(G_{1,j}\frac{1}{\alpha_1}\frac{\partial T_1}{\partial t} + T_1\frac{1}{\alpha_1}\frac{\partial G_{1,j}}{\partial t}\right)d\Omega_1 \tag{4}$$

or, using the chain rule:

$$\int_{\Gamma_0\cup\Gamma_1\cup\Gamma}\left(G_{1,j}\frac{\partial T_1}{\partial \mathbf{n}_1} - T_1\frac{\partial G_{1,j}}{\partial \mathbf{n}_1}\right)d\left(\Gamma_0\cup\Gamma_1\cup\Gamma\right) = \int_{\Omega_1}\left[\frac{1}{\alpha_1}\frac{\partial\left(G_{1,j}T_1\right)}{\partial t}\right]d\Omega_1 \tag{5}$$

Using now Eqs. (1.b), (1.c), (2.c), (2.d), and the fact that some measurements $T_1=Y$ are available at boundary Γ_0, we obtain:

$$\int_{\Gamma_0}\left(-G_{1,j}\frac{q}{K_1} - Y\frac{\partial G_{1,j}}{\partial \mathbf{n}_1}\right)d\Gamma_0 + \int_{\Gamma}\left(G_{1,j}\frac{\partial T_1}{\partial \mathbf{n}_1}\right)d\Gamma = \int_{\Omega_1}\left[\frac{1}{\alpha_1}\frac{\partial\left(G_{1,j}T_1\right)}{\partial t}\right]d\Omega_1 \tag{6}$$

Integrating now both sides in time and using (1.i) and (2.e), we obtain:

$$\int_t\int_{\Gamma_0}\left(-G_{1,j}\frac{q}{K_1} - Y\frac{\partial G_{1,j}}{\partial \mathbf{n}_1}\right)d\Gamma_0\,dt = -\int_t\int_{\Gamma}\left(G_{1,j}\frac{\partial T_1}{\partial \mathbf{n}_1}\right)d\Gamma\,dt \tag{7}$$

After solving Eqs. (2.a)-(2.e), we can obtain the value of $G_{1,j}$ at interface Γ and call it β_j. Notice that since those equations are solved for an orthonormal basis function Ψ_j, the value of β_j is also a basis function. Then, approximating the normal derivative of T_1 at interface Γ as:

$$\left.\frac{\partial T_1}{\partial \mathbf{n}_1}\right|_{\Gamma} = \sum_i \beta_i\alpha_i \tag{8}$$

we can then write Eq. (7) as:

$$\int_t \int_{\Gamma_0} \left(-G_{1,j} \frac{q}{K_1} - Y \frac{\partial G_{1,j}}{\partial \mathbf{n}_1} \right) d\Gamma_0 \, dt = -\left\langle \beta_j, \beta_i \right\rangle_{L^2\left((0,t_f)\times t_f\right)} \alpha_i \tag{9}$$

Then, after solving the linear system given by Eq. (9) and obtaining the values of α_i, the normal derivative of temperature at interface Γ can be obtained using Eq. (8). Notice that since β_j is a function of time and space, the estimated function also varies with these two variables. Also, the auxiliary problems need to be evaluated only once, since they do not depend on the unknown quantity. Thus, once these problems are solved, different heat fluxes can be obtained simply solving the linear system given by Eq. (9), whose solution vector comprehends a series of integrals of the auxiliary variables multiplied by the imposed heat flux and measure temperatures. Notice that the coefficients of such linear system are given by the double integral (in space and time) of the coefficients β_i and β_j and can be stored in advance to estimate different functional forms of the sought function.

4 Results

In this paper we considered the simulation of a problem that can be reduced to a simplified two-dimensional setting, using the same material for both domains (AISI 1050 steel) with length equals to 0.04 m and a total thickness equal to 0.02 m (0.01 m for each layer). The imposed heat flux q was set to 100,000 W/m^2, which resulted in a maximum increment achieved for the temperature of 270 ^{o}C after 10 minutes of simulated experiment. Lower applied heat fluxes were also simulated and the results were similar to the ones presented here. Six different profiles for the unknown heat flux at the connecting interface were considered. These profiles were obtained by using a constant thermal contact resistance, according to Table 1. Notice that a time dependent thermal contact conductance could also be used, but to simplify the direct problem simulation, a constant value was used in this paper.

Table 1: Functions used for the contact conductance

Profile	***h* [W/(m^{2o}C)]**
1	1000 for $x < L/4$ and $x > 3L/4$ 0 for $L/4 < x <$ L/2
2	1000 for ($x < L/4$) and ($L/2 < x < 3L/4$) 0 for ($L/4 < x < L/2$) and ($x > 3L/4$)
3	1000 sin(πx)
4	1000 sin($2\pi x$)
5	1000
6	1000 for ($x < L/4$) and ($L/2 < x < 3L/4$) 500 for ($L/4 < x < L/2$) 0 for ($x > 3L/4$)

Simulated measurements were used in this paper, which were obtained by solving the direct problem given by Eqs. (1.a)-(1.i) for the six different thermal contact conductances given in Table 1. The solution was obtained by using the finite difference technique after a proper grid convergence analysis has been conducted. Since the direct and inverse problems rely on different techniques, the so-called inverse crime was avoided.

To verify the stability of the inverse problem solution, simulated experimental errors were introduced in the measured temperatures according to:

$$Y = T\left(\Gamma_0\right) + \varepsilon\sigma \tag{10}$$

where ε is a random variable with a Gaussian distribution and σ is the standard deviation of the measurements. In order to generate the Gaussian random variable ε, with zero mean and unit variance, the Box-Muller transformation was used

$$\varepsilon = \sqrt{-2\ln(u)}\cos(2\pi v) \tag{11}$$

where u and v are two uniformly distributed random numbers (between 0 and 1) generated by the Fortran intrinsic function random_number().

As discussed before, the auxiliary problem was solved for a set of orthonormal functions. In this study, for the two-dimensional setting, a combination of sine and cosine-wave functions (Fourier series) was chosen. As it will be shown, this choice does not provide the best estimate for discontinuous functions, mainly due to the Gibbs phenomenon. The number of test functions used was set to forty-one in time and forty-one in space. Notice that the auxiliary problems must be solved only one time. Once their solution is obtained, different heat fluxes can be obtained.

The linear system was solved using the preconditioned bi-conjugate gradient method, where a Jacobi preconditioner was employed, for errorless measurements. When noise was introduced, a truncated SVD solver was used, where the minimum singular value allowed was 0.001. The double integrals appearing on the left and right side of Eq. (9) were calculated using the IMSL DBS2IG subroutine, where a second order B-Spline was used. The frequency of measurements used was set to 0.96 Hz.

Figure 2 shows the estimate of the heat flux at the contact interface resulting from the six different thermal contact conductances presented in Table 1, for measurements with and without noise. In this figure, x coordinate is aligned to the Γ boundary presented in Fig. 1. Considering the non-intrusive character of such method, and the very little computational effort to post-process the results (once the temperatures are measured, the entire estimate is performed in a few seconds), the results are very good, although some oscillations are present, mainly due to the Gibbs phenomena. It is interesting also to notice that even in the case where experimental errors were introduced in the measurements, the estimate was reasonable good.

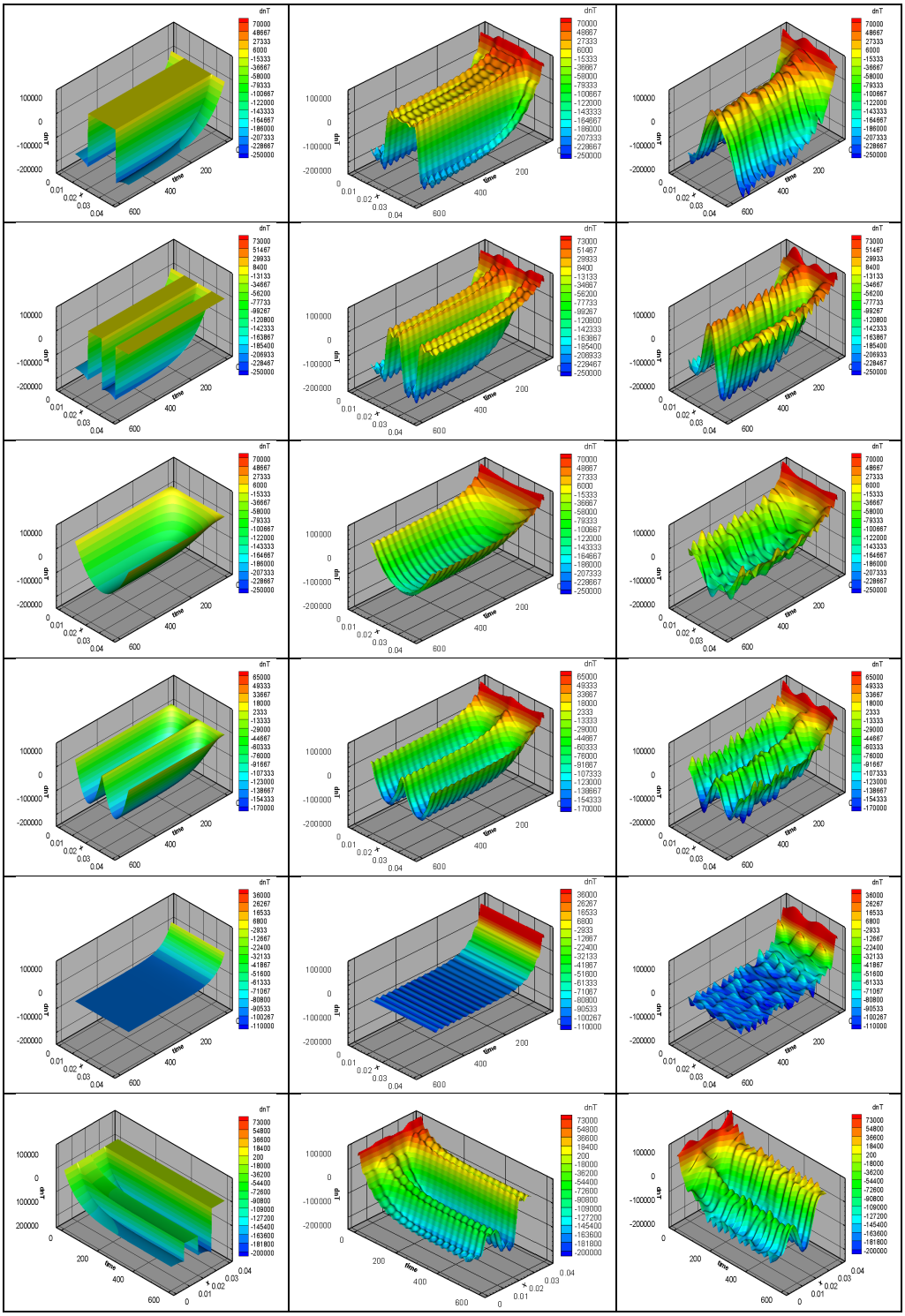


Figure 2: Exact (left) and estimated (middle and right) variation of $\partial T_1/\partial \mathbf{n}$ at Γ for errorless measurements (middle) and for measurements containing noise with σ=0.5 ^{o}C (right).

5 Conclusions

In this paper we presented a non-intrusive methodology to estimate unknown heat fluxes located in the interface between two materials. The method was capable to capture the spatial and temporal distribution of the unknown quantity in a fast way, only solving a linear system, whose solution vector was composed of some integrals of the imposed heat flux and measured temperatures at an outer interface. Measurements with and without errors were considered, showing a good estimate of the function. For errorless measurements, a preconditioned bi-conjugate gradient method was used, whereas for measurement containing errors a truncated SVD solver was employed. Future works shall investigate the use of more general orthonormal basis functions, as well as other regularization techniques.

Acknowledgements

The authors would like to thank the agencies CNPq, CAPES, FAPERJ, ANP/PRH37, and FCT for fostering science and for financial support for this work. This work was part of a collaboration project between Brazil and Portugal (CAPES/FCT 305).

References

[1] J. Schöntag, D. Willemann, J.R. Albertazzi, *Depth assessment of defects in composite plates combining shearography and vibration excitation*, Proc. SPIE, 7387:1, (2010).

[2] F. Gay, S. Hoa, S. Tsai, *Composite Materials Design and application*, CRC Press, (2003).

[3] C. Fieberg, R. Kneer, *Determination of Thermal Contact Resistance from Transient Temperature Measurements*, International Journal of Heat and Mass Transfer, 51, (2008), pp. 1017-1023.

[4] C.-H. Huang, T.-M. Ju, *An Inverse Problem of Simultaneously Estimating Contact Conductance and Heat Transfer Coefficient of Exhaust Gases Between Engine's Exhaust Valve and Seat*, International Journal for Numerical Methods in Engineering, 38:5, (1995), pp. 735-754.

[5] Y.-C. Yang, *Estimation of Thermal Contact Resistance and Thermally Induced Optical Effects in Single-Coated Optical Fibers*, Optics Communications, 278, (2007), pp. 81-89.

[6] M.N. Ozisik, H.R.B. Orlande, L.G. Hector, P.N. Anyalebechi, *Inverse Problem Of Estimating Interface Conductance During Solidification Via Conjugate Gradient Method*, Journal of Materials Processing and Manufacturing Science, 1, (1992), pp. 213-225.

[7] H.R.B. Orlande, M.N. Ozisik, *Inverse Problem Of Estimating Interface Conductance Between Periodically Contacting Solids*, Journal of Thermophysics and Heat Transfer, 7, (1993), pp. 319-325.

[8] M.H. Shojaeefard, K. Goudarzi, *The Numerical Estimation of Thermal Contact Resistance in Contacting Surfaces*, American Journal of Applied Sciences, 5:11, (2008), pp. 1566-1571.

[9] M.H. Shojaeefard, K. Goudarzi, M. Sh. Mazidi, *Inverse Heat Transfer Problem of Thermal Contact Conductance Estimation in Periodically Contacting Surfaces*, Journal of Thermal Science, 18:2, (2009), pp. 150-159.

[10] M.H. Shojaeefard, M. Sh. Mazidi, V. Kh. Mousapour, *The Estimation of Time-Varying Thermal Contact Conductance Between Two Fixed Contacting Surfaces*, Mechanika, 19:2, (2013), pp. 167-171.

[11] J. Gill, E. Divo, A.J. Kassab, *Estimating Thermal Contact Resistance Using Sensitivity Analysis and Regularization*, Engineering Analysis with Boundary Elements, 33, (2009), pp. 54-62.

[12] P.J. Halliday, R. Parker, R.B. Piggott, A.C. Smith, D.C. Steer, *Estimation of the Thermal Contact Resistance Between Potato Granules and Steel*, Journal of Food Engineering, 28, (1996), pp. 261-270.

[13] X. Zhang, P. Cong, S. Fujiwara, M. Fujii, *A New Method for Numerical Simulation of Thermal Contact Resistance in Cylindrical Coordinates*, International Journal of Heat and Mass Transfer, 47, (2004), pp. 1091-1098.

[14] X. Luo, H. Feng, J. Liu, M.L. Liu, S. Liu, *An Experimental Investigation on Thermal Contact Resistance Across Metal Contact Interfaces*, Proc. of 12th International Conference on Electronic Packaging Technology & High Density Packaging, (2011).

[15] J. Liu, H. Feng, X. Luo, R. Hu, S. Liu, *A Simple Setup to Test Thermal Contact Resistance Between Interfaces of Two Contacted Solid Materiais*, Proc. of 11th International Conference on Electronic Packaging Technology & High Density Packaging, (2010).

[16] S. Andireux, A.B. Abda, *The Reciprocity Gap: A General Concept for Flaw Identification Problems*, Mechanics Research Communications, 20:5, (1993), pp. 415-420.

[17] M.J. Colaço, C.J.S. Alves, *A Fast Non-Intrusive Method for Estimating Spatial Thermal Conductance by Means of the Reciprocity Functional Approach and the Method of Fundamental Solutions*, International Journal of Heat and Mass Transfer, 60, (2013), pp. 653-663.

[18] M.J. Colaço, C.J.S. Alves, H.R.B. Orlande, *Non-Intrusive Transient Method for Estimating Spatial Thermal Contact Resistance Using the Reciprocity Functional Approach*, Proc. of 4th Inverse Problems, Design and Optimization Symposium, June 26-28, Albi, France, (2013).

[19] L.A.S. Abreu, C.J.S. Alves, M.J. Colaço, H.R.B. Orlande, *A Non-Intrusive Inverse Problem Technique for the Identification of Contact Failures in Double-Layered Composites*, Proc. of 15th International Heat Transfer Conference, August 10-15, Kyoto, Japan, (2014).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Identification of aerodynamic coefficients of a projectile and reconstruction of its trajectory from partial flight data

V. Condaminet[ab], **F. Delvare**[a], **D. Choï**[a], **H. Demailly**[c], **C. Grignon**[c], **S. Heddadj**[b]

[a]Laboratoire LMNO, Université de Caen, Boulevard Maréchal Juin, 14032 Caen, France
[b]Nexter Munitions, 7 Route de Guerry, 18023 Bourges Cedex, France
[c]DGA Land Systems, rocade Est - Echangeur de Guerry, 18000 Bourges, France
e-mail: vincent.condaminet@unicaen.fr

Key words: aerodynamic coefficients identification, non-linear optimization, free flight, regularization

Abstract

A method is proposed to identify the aerodynamic coefficients and to reconstruct the trajectory of a fin-stabilized projectile from partial flight data. A reduced model for the flight mechanic equations is used instead of the more general six-degree-of-freedom model. An optimization problem is introduced to find the set of aerodynamic coefficients minimizing the difference between the measured and calculated state parameters. The use of Lagrange multipliers instead of a penalization term is preferred to take into account the flight mechanics equations and therefore to denoise data. An iterative method is also developed and tested on real flight data to reconstruct the trajectory of the projectile where measurements are not available.

1 Introduction

We have a good representation of the aerodynamic forces and moments acting on the projectile during its flight and each aerodynamic forces and moments are connected to an aerodynamic coefficient. Accurate knowledge of the aerodynamic coefficients of a projectile is then essential in understanding, controlling and predicting its flight attitude, in terms of precision and stability. Different tools can be used to identify these coefficients, such as aerodynamic numerical codes [1] or wind tunnel tests. The coefficients obtained by numerical codes are sometimes incompatible with the technical stakes in question. The tests performed in wind tunnel are subjects to a perturbation in the flow due to the sting used to hold the mock-up in the flow vein, altering the quality of the determination of the drag coefficient. The firing test, allowing real conditions to be reproduced, remains the most reliable method to study the aerodynamics of a projectile. Our aim is to develop a technique to identify the aerodynamic coefficients using measurements from a flight tests.

Several identification techniques have been previously developed [2, 3, 4, 5]. Some of them are based on least squares optimization so as to minimize the deviation between the measured and computed trajectories [2, 3]. Other methods consist, for example, in employing a Kalman filter with the aim to reduce the inaccuracy of the estimations due to noise related to the instrumentation [4, 5].

In this paper, we first recall the method developed by Demailly et al [6] for the identification of the aerodynamic coefficients using a reduced model of the projectile flight attitude. This model takes into account the state parameters the most representative of the motion of a projectile, namely axial velocity v_i and roll rate ω_c. This identification method consists in a least square minimization composed of two terms. The first term characterises a deviation between the measured data and the calculated quantities. The second term is a penalization term expressing the flight mechanics equations in discrete form.

In section 2, the forces and moments applied on the projectile and the reduced model introduced by Demailly et al [6] are presented. In section 3, the penalization method introduced to take into account the flight mechanic equations is compared with the use of equality constraints. Finally, in the section 4, an iterative method is proposed to reconstruct the trajectory of the projectile in taking into account the partial measured data and the iterative method is tested on real flight data taking from cards and RADAR equipments.

2 Reduced system of equations of motion

The identification of aerodynamic coefficients by inverse method requires first of all to establish the most representative direct model of the attitude of the projectile in flight. On this purpose, the inventory of the aerodynamic forces and moments must be done. Several authors have already been interested in the inventory of forces and moments applied on the projectile during its flight as Fleck [7] or Lieske and Mackenzie [3].

2.1 General formulation

Aerodynamic forces and aerodynamic moments are each associated with aerodynamic coefficients. Forces like gravity $\overrightarrow{g}$ and Coriolis $\overrightarrow{Cor}$ forces can be distinguished from aerodynamic forces (Figure 1) such as the drag force $\overrightarrow{D}(C_x)$, the lift force $\overrightarrow{L}(C_z)$ and the Magnus force $\overrightarrow{K}(C_y)$ which are induced by the air flow around the projectile. The projectile is also subject to aerodynamic moments (Figure 1) as the rolling moment $\overrightarrow{M_E}(C_{l0})$, the roll damping moment $\overrightarrow{M_R}(C_{lp})$, the pitching moment $\overrightarrow{M_A}(C_m)$, the pitch damping moment $\overrightarrow{M_D}(C_{mq})$ and the Magnus moment $\overrightarrow{M_M}(C_{np})$.

According to the basic principle of the dynamic, we obtain a system of non-linear differential equations. These equations are representative of the projectile's attitude in a six degrees of freedom model which gives the position (three parameters) and orientation (three angles) of the projectile.

2.2 Formulation of a reduced model for a fin-stabilized projectile

Demailly et al [6] showed that, for a fin-stabilized projectile, a reduction of this model is possible. A sensitivity study of the state parameters with respect to a variation of each aerodynamic coefficients shows that the axial velocity component (v_i) depends only on the drag coefficient C_x and the roll rate parameter ω_c is influenced only by the coefficients C_x, C_{lp} and C_{l0}. The equations governing the evolution of v_i

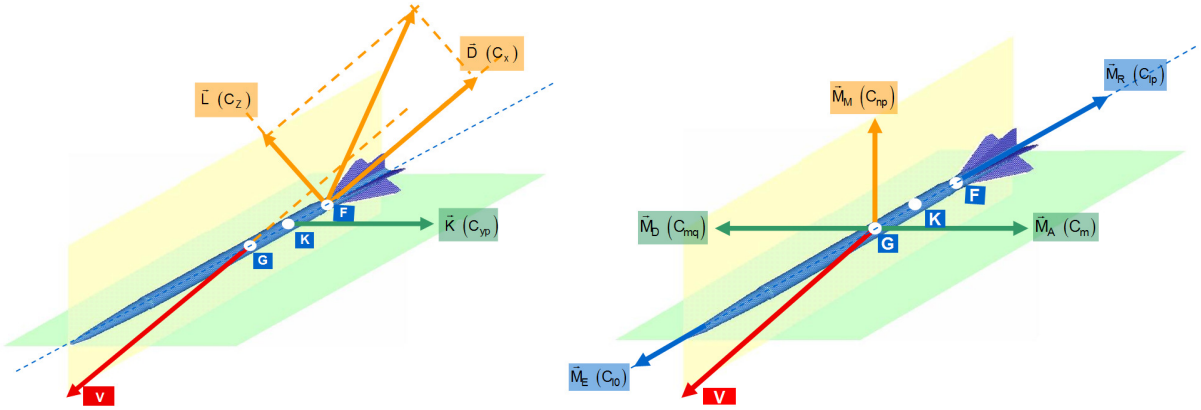


Figure 1: left: Inventory of the forces; right: Inventory of the moments [6]

and ω_c can be decoupled from the other equations. This amounts to consider a model composed only of the drag force $\overrightarrow{D}$, the rolling moment $\overrightarrow{M_E}$ and the roll damping moment $\overrightarrow{M_R}$

$$\overrightarrow{D} = -\frac{1}{2}\rho S C_x v_i \overrightarrow{v_i} \qquad \begin{aligned} \overrightarrow{M_E} &= \frac{1}{2}\rho S D C_{l0} v_i \overrightarrow{v_i} \\ \overrightarrow{M_R} &= -\frac{1}{2}\rho S D^2 C_{l0} \omega_c \overrightarrow{v_i} \end{aligned}$$

where ρ is the density of the air, D is the projectile's calibre, S is the frontal area of the projectile, m is the projectile's mass and I_1 the projectile's longitudinal inertia. The principle of the dynamic reduces then to the following system of two equations :

$$\begin{cases} \dot{v_i} = -\dfrac{\rho S}{2m} v_i^2 C_x \\ \dot{\omega_c} = -\dfrac{\rho S}{2I_1}[D^2 C_{lp} \omega_c v_i - D C_{l0} v_i^2] \end{cases} \tag{1}$$

For the sake of simplicity, the state parameters, denoted $U = (v_i, \omega_c)$, will be associated with the aerodynamic coefficients $C = (C_x, C_{lp}, C_{l0})$ in the form of the following dynamical system :

$$\dot{U} - f(U, C, t) = 0 \tag{2}$$

3 Identification procedures of the aerodynamic coefficients

During a flight, discrete and partial measurements of the state parameters v_i and ω_c are recorded. The principle of the identification method is to look for the set of aerodynamic coefficients (C_x, C_{lp} and C_{l0}) associated with the couple (v_i^*, ω_c^*) to be as close as possible to the measurements (ϕ_d) and satisfying the flight mechanics equations (in a discrete form) at all times. As an example, if the Euler explicit scheme is employed, the discrete flight mechanics equations are:

$$A_i^m(U^{m+1}, U^m, C) = U^{m+1} - U^m - f(U^m, C, t^m)(t^{m+1} - t^m) = 0 \ \ \forall t^m \in [t^0 \dots \ t^{N-1}] \tag{3}$$

where U^m is the vector composed of the state parameters vector value at time t^m.

3.1 Non linear optimization using a penalization term

Demailly et al [6] have proposed a non-linear optimization problem which minimizes the distance between measured and calculated data using the least-squares technique.

$$\begin{cases} \text{Find } \Psi = (U^*, C^*) \text{ such as} \\ J(\Psi^*) \leq J(\Psi) \ \forall \ \Psi = (U, C) \\ J(\Psi) = ||U - \phi_d||^2 + ||R(\Psi)||^2 \end{cases} \tag{4}$$

A penalization term $||R(\Psi)||$ is introduced to take into account the flight mechanics equations in a discrete form (equation (3)). As an example, if we use the Euler explicit scheme, the penalization term is:

$$||R(\Psi)||^2 = \sum_{m=1}^{N-1} [U^m - U^{m+1} - f(U^m, C, t^m)(t^{m+1} - t^m)]^2$$

The Newton-Raphson technique is then used to solve the system of non-linear optimality equations. It appears, see Figure 2, that a penalization term is not appropriate for noisy data, numerical solution obtained with this method is not correctly denoised.

3.2 Non linear optimization under equality constraints

We propose another technique, to take into account the flight mechanics in their discrete form. We take into account the flight mechanics by equality constraints which lead to the following constrained minimization problem:

$$\begin{cases} \text{Find } \Psi^* = (U^*, C^*) \text{ which minimizes} \\ J(\Psi) = ||U - \phi_d||^2 \ \forall \ \Psi = (U, C) \\ \text{under the equality constraints (3)} \end{cases} \tag{5}$$

Obviously, the matrix system size ($4N \times 4N$) is greater compared to the use of a penalization term ($(2N + 2) \times (2N + 2)$) because we need to introduce Lagrange multipliers to take into account the equality constraints. The calculation time is consequently greater.

3.3 Numerical tests

We use simulated data which have been generated numerically and noised to compare both identification methods. We remark that the use of equality constraints enables not only to identify the aerodynamic coefficients, but also to denoise the experimental measurements, in comparison with the use of a penalization term. For the sake of confidentiality, the curve presented in Figure 2 has been normalized ($\omega_c^* = \frac{\omega_c}{\omega_{c_{max}}}$ and $t^* = \frac{t}{t_{max}}$).

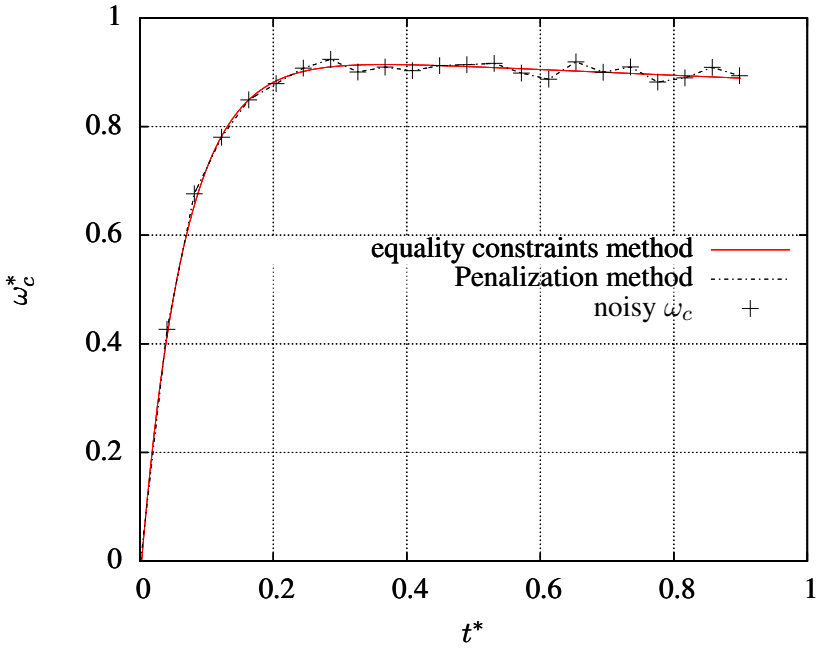


Figure 2: Comparison between the penalization method and the equality constraints method

In addition, Table 1 (which corresponds to the results of Figure 2) shows that the identification of the aerodynamic coefficients is better when using equality constraints rather than using a penalization term. The relative error is calculated from the reference solution.

	Relative error (%)			
	C_x	C_{lp}	C_{l0}	CPU Time (s)
Penalization method	0.26	8.12	8	11
Equality constraints method	0.26	1.78	2	245

Table 1: Comparison of the relative error on the coefficients identification obtained by both identification techniques

4 An iterative identification method from partial flight measurements : application to real data

4.1 Iterative regularization method

In this section, we are interested in the reconstruction of the entire flight from partial experimental data as the measurements from instrumentation are not available on the entire flight. In addition, the various state parameters often have different acquisition frequencies. An iterative regularization method inspired from [8, 9] introduced by Cimetière et al has been developed to determine the missing information. The functional J^k contains two terms. The first one characterizes the validity given to measures and the second one characterizes the distance to a priori known information. The solution of this optimization problem is the limit of the solutions of a sequence of minimization problems where the a priori information is updated at each step.

$$\begin{cases} \text{Find } \Psi^{k+1} = (U^{k+1}, C^{k+1}) \text{ which minimizes} \\ J_c^k(\psi) = ||U - \phi_d||^2 + c||U - U^k||^2 \quad \text{with } \psi = (U, C) \\ \text{under the equality constraints (3)} \end{cases} \tag{6}$$

4.2 Application to real flight data

The effectiveness of this iterative method is now illustrated by comparing its results with those obtained by Demailly et al [6] in the case of real flight data. We only have information on the two state parameters which are axial velocity v_i and roll rate ω_c. The roll rate measurements comes from two different sources. Cards were positioned on the beginning of the flight to determine the angular attitude of the projectile between each of them. The roll rate parameter is then deduced from these measurements. The second source of measurements for the roll rate parameter is captured from RADAR (associated with reflectors) on another part of the trajectory. There is therefore a gap where no measures are available, between the cards and RADAR measures. In addition, cards and RADAR measures have different acquisition frequencies. We compare the efficiency of the two methods : the penalization method [6] and the iterative regularization method. For the iterative regularization method, the interpolated data are updated at each iteration, i.e. we trust only the measured data.

4.2.1 Comparison between the penalization method and the iterative method to reconstruct the projectile attitude

On the Figure 3, the reconstruction with both methods using only the 9^{th} card data and all the RADAR data is shown. It illustrates the efficiency of the iterative regularization method, in comparison with the penalization method to reconstruct the evolution of the state parameter ω_c over the entire flight. In addition, the iterative method enables to reconstruct the initial value of ω_c. We observe (Table 2) that the coefficients identified with the regularization method are closer to those identified with all data than the coefficients identified using a penalization term (the relative error is calculated from a identified ω_c curve which is passing through all the measured data).

	Relative error (%)			
	C_x	C_{lp}	C_{l0}	CPU Time (s)
Penalization method	2	40	39	35
Iterative regularization method	1.5	11	11	780
Reference (all data)	0	0	0	

Table 2: Comparison of the relative errors on the coefficients identification

4.2.2 Influence of the cards data on the reconstruction of the roll rate parameter

We also wanted to explore the possibility to dispense cards measures to identify properly the coefficients. A first identification test have been realized without card data, a second with the first card data and a last test was realized with the 9^{th} card data. The reconstructions associated to this test are shown in Figure 4 and the corresponding relative errors on the identified coefficients are calculated relatively to the aerodynamic coefficients identified with all data. We can observe, Figure 4 and Table 3, the importance of having data at the beginning of the flight to identify with accuracy the aerodynamic coefficients and to reconstruct properly the projectile attitude.

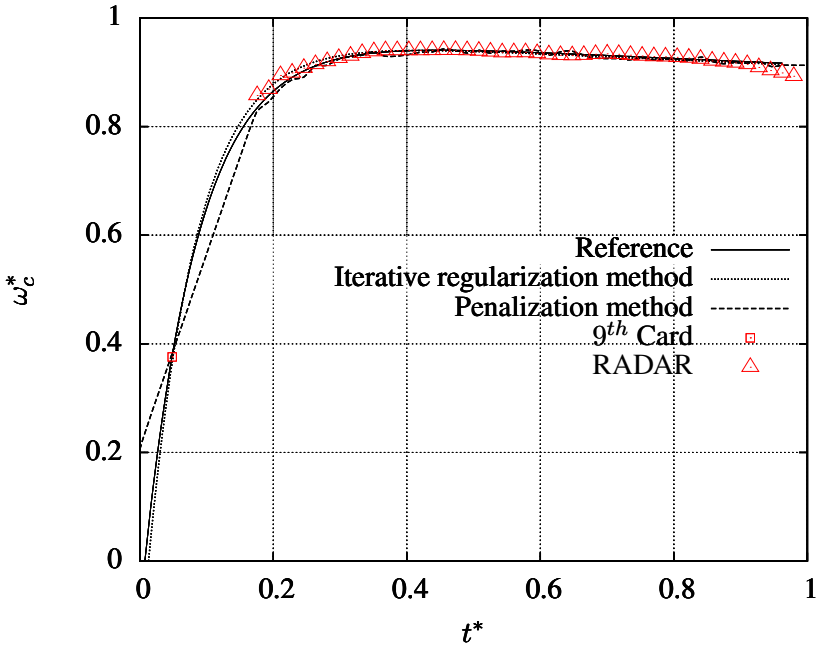


Figure 3: Comparison of the reconstructions of the adimensional roll rate ω_c^* using the iterative regularization method and the penalization method

	Relative error (%)		
	C_x	C_{lp}	C_{l0}
Without card data	1	11	11
With the 1^{st} card data	1.5	7	7
With all data	0	0	0

Table 3: Influence of the number and the positions of cards data on the identification of the aerodynamic coefficients

5 Conclusion

We have developed an optimization method with equality constraints which has proved to be most effective to denoise the flight measurements than the use of penalization term to take into account the flight mechanics. Then we have introduced optimization techniques which enables, from partial flight data, to identify with accuracy the aerodynamic coefficients at simultanously to reconstruct the entire trajectory of a projectile. This method has been tested with a reduced model developed by Demailly et al [6] and validated using real flight data. The technique developed opens a large field of applications such as to determine the initial conditions of a firing test or to study the influence of the instrumentation position along the flight.

Acknowledgements

We would like to acknowledge the French Ministry of Defence and the company Nexter Munitions for the financial support.

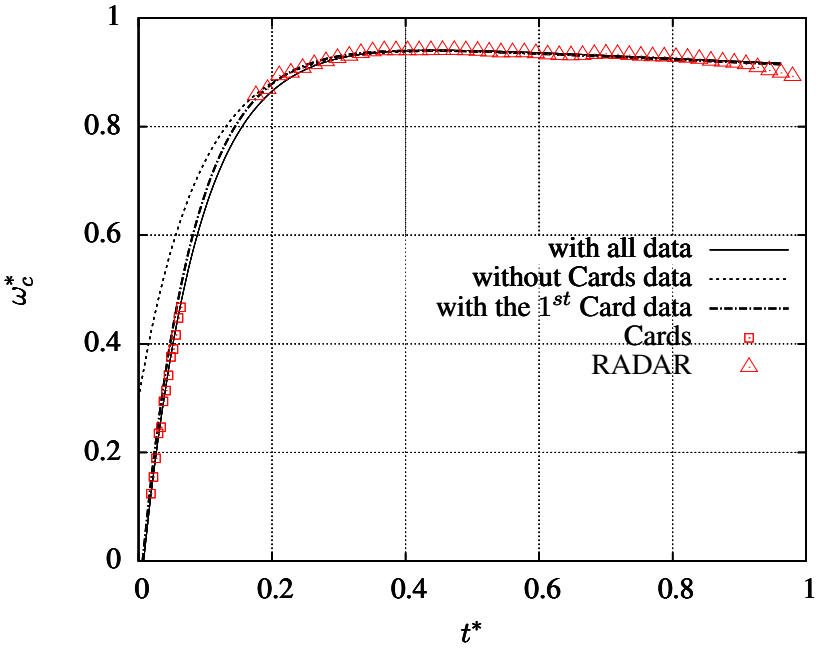


Figure 4: Effect of the cards data on the reconstruction of the roll rate using the iterative method

References

[1] P. Rebuffet, *L'aérodynamique Expérimentale*, 3, Vol. 1, Dunod, Paris, 1969.

[2] G.T. Chapman, D. Kirk *A method for extracting aerodynamic coefficients from free-flight data*, AIAA J. 8, 753-758, 1970.

[3] R.F. Lieske, A.M. Mackenzie *Determination of aerodynamic drag from radar data*, Aberdeen Proving Ground Tech. Rep., Memorandum rept no. 0465057, Institut Saint-Louis, 1982.

[4] G. Mahmutyazicioglu, B.E. Platin *Aerodynamic Data Identification Using Linear Aeroballistic Theory*, System Identification for Integrated A/C Development and Flight Testing, RTO Symposium, RTO MP-11, Madrid, Spain May, Research and Technology Organization, 1998.

[5] J. Quanwei, C. Qiongkang *Dynamic model for real-time estimation of aerodynamic characteristics*, J. Aircraft, 26, 315-321, 1986.

[6] H. Demailly, F. Delvare, C. Grignon, S. Heddadj, P. Bailly, *Identification of aerodynamic coefficients of a kinetic energy projectile from flight data*, Inverse Problems in Science and Engineering, 21, 1, 63-83, 2013.

[7] V. Fleck, *Introduction à la balistique extérieure avec exercices, Cours de Balistique extérieure*, Coetquidant, France, 2005.

[8] A. Cimetière, F. Delvare, M. Jaoua, F. Pons, *Solution of the Cauchy problem using iterated Tikhonov regularization*, Inverse Problems, 17, 553-570, 2001.

[9] F. Delvare, A. Cimetière, J.L. Hanus, P. Bailly, *An iterative method for the Cauchy problem in linear elasticity with fading regularization effect*, Computer Methods in Applied Mechanics and Engineering, 199, 49-52, 3336-3344, 2010.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Reconstruction of Continuous Mechanical Deformations in Power Transformer Windings

Mariana Dalarsson*, **Martin Norgren, Hadi Emadi**

Department of Electromagnetic Engineering, School of Electrical Engineering,
KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden
e-mail: mariana.dalarsson@ee.kth.se

Key words: Inverse Problems, Power Transformers, Continuous Deformations, Microwave Radiation

Abstract

We present an online method to detect continuous mechanical deformations of windings in a power transformer. We employ an approximate cylindrical continuous model of an outer winding surrounded by the transformer tank wall, where the power transformer winding can be viewed as a continuous metallic structure with deformations described by means of suitable continuous functions. In such a model, the perturbation theory for waveguide conductor boundaries can be used and the computation complexity can be reduced as compared to the previously studied discrete conductor models. The idea is to insert antennas inside the transformer tank above and below the winding to radiate and measure microwave fields that interact with the metallic structure bound by the outer winding and the transformer tank wall. The responses from the radiated waves are assumed to be sensitive to the continuous mechanical deformations on the outer surface of the outer winding. The goal is to reconstruct the continuous deformation function from measurements of the scattered fields at both ends. The inverse problem of reconstructing sinusoidal boundary deformations in a coaxial waveguide is solved by using a first order perturbation method applied to the dominant TEM-mode. The reconstruction results from using reflection data are presented and they indicate an agreement between the reconstructed and true continuous functions that describe the sinusoidal mechanical deformations of transformer windings.

1 Introduction

Local deformations of the power transformer windings, due to the heavy mechanical forces from short circuit currents, are known to increase the risk of failures which may cause serious electrical power outages. It is therefore of interest to develop suitable on-line methods for early detection of such winding deformations. Some consequences of various degradation phenomena in power transformers can be diagnosed by frequency response analysis (FRA) method, applicable only when the transformer is disconnected from the power grid. In e.g. [1] FRA has been used for detection of winding deformations.

However, in order to reduce risks of power outages, the online methods - applicable while the transformer is in operation - are desirable. In [2] and [3] we studied an online method based on positioning microwave antennas inside the transformer tank and using the microwaves to reconstruct the set of all radial positions of the individual winding segments or turns along the axial winding dimension. In these papers, we used an approximate model and applied a full wave method that takes into account the details of the winding geometry. The locations of the individual conductors are then reconstructed using an optimization method. In order to reduce the computational complexity in [2] and [3] we studied cases with up to ten winding segments or turns. However, in a more realistic case the number of winding turns may be prohibitively large to allow a full-scale numerical model used in the optimization procedure. Our primary interest is, however, to detect small winding deformations and it is therefore of interest to study whether inversion methods, based on weak scattering, can be utilized for the reconstruction of transformer winding deformations. In a recent paper [4] a first step towards such a method was taken, where each winding was not considered in detail, but instead modeled as an equivalent outer boundary surface, which shape is to be reconstructed. In [4] a specific piecewise linear shape of the winding deformation in a parallel-plate waveguide model of the power transformer winding structure was assumed. In the present paper we pursue this investigation a step further by considering the continuous axially symmetric sinusoidal deformation in a coaxial waveguide model of the power transformer winding structure. We present a simple and computationally efficient first order perturbation method to solve the inverse problem of reconstructing deformations in the lower coaxial waveguide boundary. The model is tested using synthetic measurement data from simulation of the structure in the commercial FEM program HFSS.

2 Problem formulation and scattering analysis

We consider an axially symmetric scattering configuration with a coaxial waveguide oriented along the z-axis, shown in Fig. 1. The inner boundary cylinder is located at $r = r_0$ while the outer boundary cylinder is located at $r = r_0 + a$ such that the radial width of the unperturbed cavity is equal to a. In the context of a power transformer, the outer boundary models the wall of the transformer tank while the lower boundary models the outermost layer of the winding structure. Hence, we describe the winding as an equivalent PEC surface. In a more realistic treatment it can be described by e.g. an anisotropic boundary condition [5]. Although a realistic transformer is filled with oil, we here assume that the medium inside the waveguide is air (or vacuum). At the inner boundary cylinder along the section $z_1 < z < z_2$ there is a local deformation described by the equation

$$r = ag(z) \quad \text{with} \quad \max|g(z)| \ll 1 \quad , \quad g(z_1) = g(z_2) = 0 \tag{1}$$

The inverse problem at hand is to reconstruct $g(z)$ in the *estimation region* $z_1 < z < z_2$ using scattering data obtained when the waveguide is excited from both ends. In order to focus the present study on the primary scattering mechanism, due to the local deformation of the lower boundary, we assume that there are no reflections from the ends of the waveguide, i.e. the waveguide is infinitely long.

2.1 The direct scattering problem

Following [3] we assume that the radial width of the unperturbed cavity is small compared to the mean radius of the coaxial wave guide, i.e. that $a \ll 2r_0$. As r_0 represents the radius of the outermost layer of the winding structure and a represents its distance to the tank wall, such assumption is generally justified for the realistic power transformers where a is typically about 10-15 percent of the outer diameter of

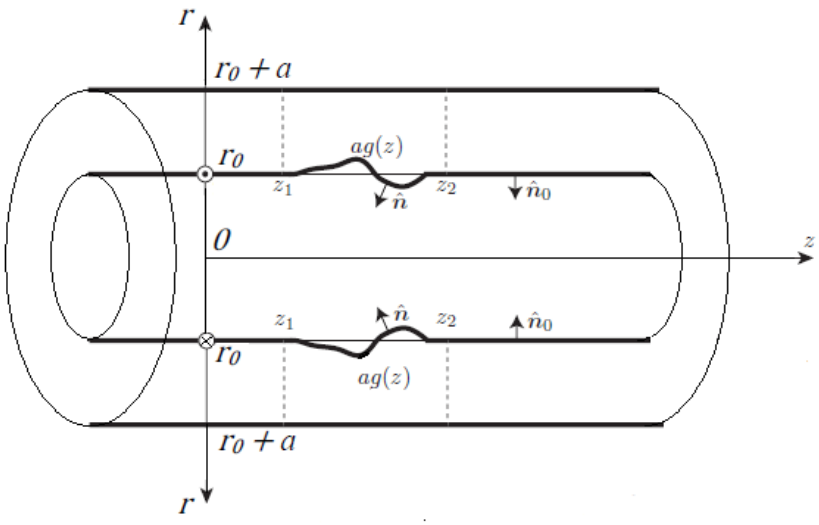


Figure 1: Geometry of the coaxial waveguide model.

the winding package $d_0 = 2r_0$. Thus the model based on the above assumption can be considered as reasonably good for the present analysis, given the fact that the very model of coaxial cylindrical waveguide is approximate in the first place. In the present investigations, the cavity size a = 1 m is used for simplicity. Since we are mainly concerned with investigating the reconstruction principles, the dimensions used are not intended to mimic a realistic power transformer. Furthermore, we restrict our present analysis to mechanical deformations that possess cylindrical symmetry (i.e. are independent on the polar angle φ) and consider the TM-modes only. In such a case we can use a new variable $\rho = r - R_0$ and use the approximate solution for the TM-modes as follows

$$\boldsymbol{E}_m = E_{mr}\,\mathbf{e}_r \propto \cos(\frac{m\pi\rho}{a}) \quad , \quad \boldsymbol{H}_m = H_{m\varphi}\,\mathbf{e}_\varphi \propto \cos(\frac{m\pi\rho}{a}) \quad , \quad m = 0, 1, 2, \ldots \tag{2}$$

and especially the dominant mode ($m = 0$) that propagates at all frequencies. Furthermore, we note that the longitudinal component of the electric field E_z is non-zero only for $m \geq 1$.

2.1.1 Perturbation method

The inversion scheme (Sect. 2.2) is based on solving the direct scattering problem by means of a boundary perturbation method for waveguides, similar to the one used in [6]. For the dominant mode the corrections to the transmission parameters are second order in the perturbation $g(z)$:

$$S_{12} = 1 + \mathcal{O}(\max|g^2|) \quad , \quad S_{21} = 1 + \mathcal{O}(\max|g^2|) \tag{3}$$

Thereby we consider transmission data to be too sensitive for measurement errors, and consequently it will not be included in the subsequent analysis. Locating the measurement planes to the boundaries of the *estimation region* $z_1 < z < z_2$, the first order approximation of the reflection parameters become

$$S_{11}(k) \approx -\mathrm{j}ke^{\mathrm{j}2kz_1}G^*(k) \quad , \quad S_{22}(k) \approx -\mathrm{j}ke^{-\mathrm{j}2kz_2}G(k) \tag{4}$$

where $k = \omega\sqrt{\epsilon_0\mu_0}$ is the vacuum wavenumber and we introduce a "k-transform" $G(k)$ of $g(z)$ and its complex conjugate as follows

$$G(k) = \int_{z_1}^{z_2} g(z)e^{\mathrm{j}2kz}dz \quad , \quad G^*(k) = \int_{z_1}^{z_2} g(z)e^{-\mathrm{j}2kz}dz \tag{5}$$

2.1.2 HFSS model

As generator of synthetic measurement data, we use a full-wave FEM model implemented in the commercial program HFSS. This model takes into account all TM modes in the waveguide. The output is the set of complex S-matrices solved in our case for the frequency range 20 MHz to 150 MHz with a step size of 5 MHz. The reason for choosing the frequencies up to 150 MHz is that one must be cautious about the existence of higher order modes that can propagate freely or be trapped in the coaxial waveguide structure. The unperturbed waveguide cavity size a = 1 m implies the cutoff frequencies at multiples of 150 MHz. In other words, in the unperturbed region, the first higher order mode starts to propagate at approximately 150 MHz and the present method, based on the contributions from the ground mode only, is not expected to give reasonably accurate results at higher frequencies.

2.1.3 Transmission Line (TL) model

For comparison purposes, we have also modeled the TEM-mode scattering with a transmission line (TL) model, in which the L and C parameters are computed from the local value $a(1-g(z))$ of the waveguide width. From [7], it follows that e.g. S_{11} obeys the Riccati equation

$$\frac{dS_{11}}{dz} = \mathrm{j}k[\frac{1}{1-g} + 1 - g]S_{11} + \frac{\mathrm{j}k}{2}[\frac{1}{1-g} - 1 + g](1 + S_{11}^2) \quad , \quad S_{11}(z_2) = 0 \tag{6}$$

Note that, like the perturbation method, this model does not take into account any higher order TM-modes. Approximating the equation (6) to the first order in g, it is easily verified that it implies the solution for S_{11} given in the first of the equations (4). Thus, to the first order in g the TL-model and the waveguide perturbation method are equivalent to each other.

2.2 The inverse scattering problem

Let us now consider the perturbation function $g(z)$ with the properties given in (1). Since we require that $g(z_1) = g(z_2) = 0$, any such function can be expanded into the Fourier Sine series of the form

$$g(z) = \sum_{n=1}^{\infty} p_n\phi_n(z) \quad , \quad \phi_n(z) = \sin(n\pi\frac{z - z_1}{z_2 - z_1}) \tag{7}$$

where we use the infinite set of mutually orthogonal sine functions $\phi_n(z)$ satisfying the required conditions $\phi_n(z_1) = \phi_n(z_2) = 0$. The coefficients p_n in (7) are real numbers and they are not functions of z. If we then apply the "k-transform", as defined in (5), to both sides of the equation (7), we obtain

$$G(k) = \sum_{n=1}^{\infty} p_n\Phi_n(k) \quad , \quad G^*(k) = \sum_{n=1}^{\infty} p_n\Phi_n^*(k) \tag{8}$$

where

$$\Phi_n(k) = \int_{z_1}^{z_2} \phi_n(z)e^{\mathrm{j}2kz}dz = \frac{n\pi(z_2 - z_1)(e^{\mathrm{j}2kz_1} - (-1)^n e^{\mathrm{j}2kz_2})}{n^2\pi^2 - 4k^2(z_2 - z_1)^2} \tag{9}$$

and $G(k)$ is given by (5). In order to be able to perform the numerical optimizations, we approximate the continuous inverse problem with a discrete inverse problem, where the deformation $g(z)$ is expanded into a finite set of functions, whereby we truncate the infinite series in (8) to a finite number of terms denoted by N, as follows

$$G(k) = \sum_{n=1}^{N} p_n \Phi_n(k) \quad , \quad G^*(k) = \sum_{n=1}^{N} p_n \Phi_n^*(k) \tag{10}$$

Since $\{p_n\}_{n=1}^{N}$ are real-valued, it is convenient to treat the real and imaginary parts of (10) as separate equations

$$G_R(k) = \sum_{n=1}^{N} p_n \Phi_{nR}(k) \quad , \quad G_I(k) = \sum_{n=1}^{N} p_n \Phi_{nI}(k) \tag{11}$$

where $G(k) = G_R(k) + \mathrm{j}G_I(k)$ and $\Phi_n(k) = \Phi_{nR}(k) + \mathrm{j}\Phi_{nI}(k)$, with $G_R(k)$, $G_I(k)$, $\Phi_{nR}(k)$ and $\Phi_{nI}(k)$ being real-valued functions of k. If we substitute $G(k)$ and $G^*(k)$ from (4) into (10), we obtain

$$k \sum_{n=1}^{N} p_n \Phi_n^*(k) = \mathrm{j}S_{11}(k) e^{-\mathrm{j}2kz_1} \quad , \quad k \sum_{n=1}^{N} p_n \Phi_n(k) = \mathrm{j}S_{22}(k) e^{\mathrm{j}2kz_2} \tag{12}$$

where we see that

$$G^*(k) = \frac{\mathrm{j}S_{11}(k)}{k} e^{-\mathrm{j}2kz_1} \quad \Rightarrow \quad G(k) = \frac{-\mathrm{j}S_{11}^*(k)}{k} e^{\mathrm{j}2kz_1} \quad , \quad G(k) = \frac{\mathrm{j}S_{22}(k)}{k} e^{\mathrm{j}2kz_2} \tag{13}$$

From the two results for $G(k)$, given in (13), we see that in theory measurements of S_{11} and S_{22} give the same $G(k)$ and consequently the same equation (10). Thus it is in principle possible to perform the reconstruction of the expansion coefficients $\{p_n\}_{n=1}^{N}$, using the measurements of either S_{11} or S_{22} (one-sided reflection data) or both S_{11} and S_{22} (two-sided reflection data).

In practice however, the measurements of S_{11} and S_{22} may provide different values of $G(k)$ due to the contributions of higher-order modes, potential measurement errors, the truncation of the infinite series and other approximations used in the present model. Thus if we use the measured S_{11} and S_{22}, by means of equations (13) we obtain the "measured" functions $G_{1M}(k)$ and $G_{2M}(k)$ as follows

$$G_{1M}(k) = \frac{-\mathrm{j}S_{11}^*(k)}{k} e^{\mathrm{j}2kz_1} \quad , \quad G_{2M}(k) = \frac{\mathrm{j}S_{22}(k)}{k} e^{\mathrm{j}2kz_2} \tag{14}$$

In the present paper we perform the reconstructions using contributions from both S_{11} and S_{22} in such a way that we calculate a simple arithmetic mean value of their respective contributions. In other words, as measured $G_M(k)$, we use

$$G_M(k) = \frac{1}{2} \left[G_{1M}(k) + G_{2M}(k) \right] \tag{15}$$

As we have pointed out before, in theory, the two complex functions $G_{1M}(k)$ and $G_{2M}(k)$ are exactly equal to each other. Although the actual complex functions $G_{1M}(k)$ and $G_{2M}(k)$ from the synthetic measurements differ from each other, a simple analysis of their real and imaginary parts indicates that their deviations from the real and imaginary parts of the exact analytical function $G(k)$, obtained in the case of some known deformation functions $g(z)$, are approximately equal in size with opposite signs. Thus, an arithmetic mean of two functions $G_{1M}(k)$ and $G_{2M}(k)$ is expected to be closer to the exact analytical function $G(k)$ than either of the two functions $G_{1M}(k)$ and $G_{2M}(k)$ by themselves. We have

therefore chosen to use (15) in the present paper. However, there are in general no restrictions to use either of the functions $G_{1M}(k)$ and $G_{2M}(k)$ or any weighted mean value between the two in the present model.

The coefficients $\{p_n\}_{n=1}^{N}$ are collected into the vector $\mathbf{p}$. From measurements of $G_M(k)$ at several values of k, i.e. frequencies, the right hand side of (10) is collected into the vector $\mathbf{G} = [\mathbf{G}_{MR}\, \mathbf{G}_{MI}]^T$, while the evaluations of $\Phi_n(k)$ are collected into the matrix with elements $\Phi_{nk} = \Phi_n(k)$, such that

$$\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi_R} \\ \boldsymbol{\Phi_I} \end{bmatrix} \quad , \quad \boldsymbol{G} = \begin{bmatrix} \boldsymbol{G_R} \\ \boldsymbol{G_I} \end{bmatrix} \tag{16}$$

The vector $\mathbf{p}$ is determined as the solution of the optimization problem

$$\mathbf{min}\, ||\boldsymbol{\Phi}\mathbf{p} - \mathbf{G}||_{\mathbf{2}}^{\mathbf{2}} \tag{17}$$

with respect to the optimization parameter α to be defined below. Since for each value of k, i.e. for each frequency, we have two real valued equations (11), in the present paper we choose exactly $N/2$ different k-values which gives us the $N \times N$ square matrix $\boldsymbol{\Phi}$ and the vector $\mathbf{G}$ of even length N. The main advantage of our choice is that it makes the matrix $\boldsymbol{\Phi}$ analytically invertible, while a potential disadvantage is a minor loss of generality of the model. This choice is in line with our initial exact analytical investigation based on the solution of the linearized Ricatti equation. In principle, there are no restrictions in the present model to choose any number of modes N with more than $N/2$ different k-values. This amounts to having more data points than expansion coefficients, in which case the vector $\mathbf{p}$ could be determined as the solution of the least square problem (17) using standard numerical methods. Although we choose the even number of expansion coefficients N and exactly $N/2$ different k-values, we expect to study more general cases of overdetermined solutions for $\mathbf{p}$ in our coming studies.

For optimization purposes, we employ the analogue of the variational principle of quantum mechanics [8] whereby we modify the equations (12) by introducing a variational parameter α which defines the effective size of the estimation region, as follows

$$z_1 \rightarrow z_1 - \alpha \quad , \quad z_2 \rightarrow z_2 + \alpha \tag{18}$$

From the family of all possible reconstruction curves for different values of α, by means of the variational method, we determine the value of $\alpha = \alpha_{\text{opt}}$ which gives the best estimate of the reconstructed deformation function $g(z)$, by minimizing the norm of $\mathbf{p}$. It should also be noted that the variational parameter α according to (18), is introduced on both sides of the estimation region in a symmetric fashion.

2.3 The sinusoidal perturbation function

In [4] a specific piecewise linear shape of the winding deformation in a parallel-plate waveguide model of the power transformer winding structure was assumed. In the present paper we consider the axially symmetric sinusoidal deformation in a coaxial waveguide model of the power transformer winding structure. The function $r = g(z)$ is given by

$$r = g(z) = \left\{ \begin{array}{ccc} 0 & , & z_1 \leq z \leq 0 \\ D\sin(\frac{n\pi z}{b}) & , & 0 \leq z \leq b \\ 0 & , & b \leq z \leq z_2 \end{array} \right\} \tag{19}$$

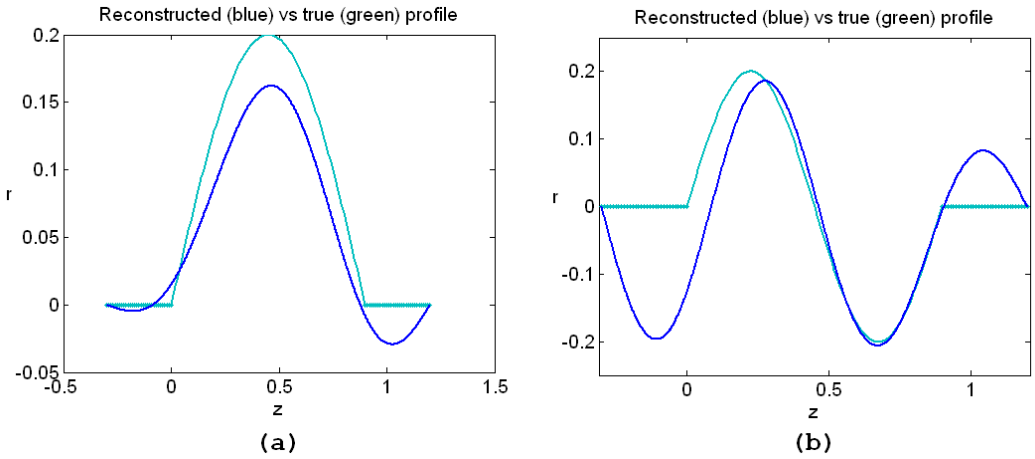


Figure 2: Reconstructions using HFSS data for $N = 4$, $D = 0.2$ (20 % intrusion and extrusion), dimensions $z_1 = -0.3$ m, $b = 0.9$ m and $z_2 = 1.2$ m and (a) $n = 1$ and (b) n = 2. In both graphs, the true shape is represented by the green line and the reconstructed shape by the blue line.

3 Reconstruction results

The general geometry of our transformer winding model is shown in 1. This geometry with $r = g(z)$ given by (19) is used to generate synthetic measurement data using HFSS. The reconstructions are then performed using HFSS data as synthetic measurement data for S_{11} and S_{22}. Fig. 2 shows two examples of reconstructions of the function $g(z)$ for $N = 4$ and dimensions $z_1 = -0.3$ m, $b = 0.9$ m and $z_2 = 1.2$ m. Fig. 2(a) shows the reconstruction results for $D = 0.2$ (20 % intrusion), $n = 1$, and $\alpha_{\text{opt}} = 1.2$ m. Fig. 2(b) shows the reconstruction results for $D = 0.2$ (20 % intrusion and extrusion), $n = 2$ and $\alpha_{\text{opt}} = 0.565$ m. From Figs. 2(a) and 2(b) we see that the reconstructions from the HFSS data are qualitatively good but display certain deviations from the actual deformation shapes. This inaccuracy may depend on a number of factors, including the simplifications in the present model as well as the relatively simple variational technique based on the single parameter α. As the present simple method is based on the contributions from the ground mode only, in the reconstructions shown in Fig. 2, we have chosen to use the frequencies below 100 MHz which are well below the expected cutoff frequencies of any modes other than the ground mode.

4 Conclusions

We have presented a simple and computationally efficient first order perturbation method to the inverse problem of reconstructing deformations in a lower coaxial waveguide boundary. The present investigation is a step towards a general diagnostic technique for detecting deformations in power transformers, that we are currently developing. The reconstruction results so far indicate that the method has potential usefulness. If proven sufficiently accurate, the present method can easily be generalized to reconstruction of arbitrary deformations by means of the Fourier series of sinusoidal terms of the kind investigated in e.g. Fig. 2(a) and Fig. 2(b). Otherwise, a natural improvement of the variational technique is to employ some more commonly used methods of regularization of ill-posed problems, such as for example Tikhonov regularization. This will be the objective of our continued efforts.

Acknowledgements

This work was funded by the Swedish Energy Agency, Project Nr. 34146-1.

References

[1] L. Yanming, L. Gang, Z. Linhai, Z. Longjun, and L. Zhiming, *Transformer winding deformation diagnosis using middle band frequency response analysis*, in Proceedings of 2007 International Conference on Solid Dielectrics, Winchester, UK, July 8-13, (2007), pp. 677-680.

[2] M. Dalarsson, A. Motevasselian, and M. Norgren, *Online power transformer diagnostics using multiple modes of microwave radiation to reconstruct winding conductor locations*, Inverse Problems in Science and Engineering, Volume 22, Number 1, (2014), pp. 84-95(12).

[3] M. Dalarsson, A. Motevasselian, and M. Norgren, *Using multiple modes to reconstruct conductor locations in a cylindrical model of a power transformer winding*, International Journal of Applied Electromagnetics and Mechanics, Volume 41, (2013), pp. 279-291.

[4] M. Norgren, M. Dalarsson and A. Motevasselian, *Reconstruction of boundary perturbations in a waveguide*, in Proceedings of 2013 URSI International Symposium on Electromagnetic Theory (EMTS), Hiroshima, Japan, May 20-24, (2013), pp. 934-937.

[5] N. Kumar, S. K. Srivastava, and S. P. Ojha, *A theoretical analysis of the propagation characteristics of an annular circular waveguide with a helical winding as the inner cladding*, Microwave and Optical Technology Letters, Volume 37, Number 1, (2003), pp. 69-74.

[6] H.-S. Tuan, *The radiation and reflection of surface waves at a discontinuity*, IEEE Transactions of Antennas and Propagation, Volume 21, Number 3, (1973), pp. 351-356.

[7] M. Norgren and S. He, *An optimization approach to the frequency domain inverse problem for a nonuniform lcrg transmission line*, IEEE Transactions of Microwave Theory and Technology, Volume 44, Number 8, (1996), pp. 1503-1507.

[8] G. Fonte, *On the Variational Method in Quantum Mechanics*, Il Nuovo Cimento, Volume 49B, Number 2, (1979), pp. 200-220.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Flatness-based Reentry Guidance Using Inverse Problem Approach

R. Esmaelzadeh

Space Research Institute,MUT, Tehran, Iran
esmaelzadeh@aut.ac.ir

Key words: Reentry, Explicit Guidance, Differential Flatness, Inverse Dynamics

Abstract
An explicit guidance law is developed for a reentry vehicle. Motion is constrained to a three-dimensional Bezier curve. Acceleration commands are derived by solving an inverse problem that combined with differential flatness approach. Trajectory is related to Bezier parameters. A comparison with pure proportional navigation shows the same accuracy, but a higher capability for optimal trajectory to some degree. Other advantages such as trajectory representation with minimum parameters, applicability to any reentry vehicle configuration and any control scheme, and Time-to-Go independency make this guidance approach more favorable.

1 Introduction

Generally, the design of guidance algorithms may be defined loosely as the art of finding the correct acceleration commands to move between two given points. Many different techniques have been suggested for the design of guidance algorithms. These range from the earliest algorithms derived using physical insight (e.g., pursuit, proportional navigation (PN) and their variants) to those derived from a systematic application of mathematical techniques. Most current guidance design methods may be classified into two main categories [1]: (1) nominal trajectory-based techniques and (2) on-line trajectory generation, reshaping and prediction schemes. In the first approach, an (optimal) reference trajectory is defined prior to the mission, and during the flight, a controller keeps the vehicle close to the nominal trajectory. The predictive and/or reshaping approaches propagate the future trajectory based on current flight state by

means of onboard numerical integration to calculate the control input during the remaining flight.
Explicit guidance methods are good examples of the second category. A review of literature [2] shows that they have many advantages over other approaches. These methods, which use preset external trajectories, give a huge calculation advantage and can provide a near optimal solution with any desired accuracy. These are applicable to systems that have linear acceleration and aim at constructing guidance algorithms with specified desired dynamics (i.e. solving an inverse problem.). Although some authors [3, 4] have considered the inverse problem as a direct method because of the implicitly parameterized control, it is better that this approach is examined within a different class. In a direct method, we are asked to predict the trajectory of the vehicle if the initial conditions and the time history of the controls are given, meaning Cauchy task, whereas in an inverse problem, we are asked to predict the controls that are compatible with a desired trajectory [5]. Inverse methods are of great interest in the context of synthesizing nonlinear autopilots [6-8] and guidance algorithms [9-13]. A survey about the inverse problem approach in optimal trajectory generation, both in Russia and in the United States, can be found in Yakimenko's paper [3]. In guidance applications, the variable guidance gains are correlated with the shape of the trajectory that will follow and satisfy particular terminal constraints. Although, with an extension of Cameron [14], Page [15], and Taranenko's [13] methods, the use of this approach in guidance algorithm design has been developed by Hough [9] and Yakimenko [3], it still suffers from serious flaws: a relatively large number of optimization parameters (Ops) (Taranenko, 20; Mortazavi [16], 12; and Hough, 8) depending on the vehicle's velocity vector, relatively difficult numerical calculations, accuracy dependence on the number of segments used in the approximation, and offline application.
In this paper, the author extends the previous work done on maximizing terminal velocity [17]. The inverse problem approach is used to develop an explicit guidance law for guiding a hypersonic unthrusted reentry vehicle (RV) to a fixed point on the ground. The geometrical trajectory shape is specified by expressing the altitude and cross range as functions of the range using Bezier curves [18]. The guidance law is based on the normal and side accelerations. The present paper deals with a new, to some extent, simplified method that provides spatial trajectories being presented analytically and completely defined by minimum parameters. This method combines a number of advantages over methods presented by Taranenko and Hough. Although the guidance law is designed for an RV, it can also be applied to any vehicle at any phase.

2 Problem definition

Assuming a spherical, nonrotating Earth (this assumption was made for simplicity, but a similar guidance law can be derived based on more precise equations of motion including terms due to earth rotation) and a gravitational field with $g=\mu/r2$, we have three-dimensional point mass equations of motion for the RV (Fig. 1):

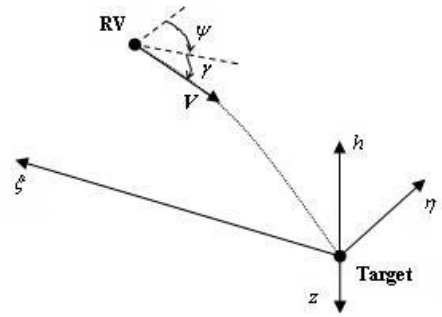


Figure 1: Geometrical definitions.

$$\begin{aligned}
\dot{V} &= \mathrm{d}V/dt = (mg\sin\gamma - D)/m \\
\dot{\gamma} &= \mathrm{d}\gamma/dt = (a_{\mathrm{vc}} + (g - V^2/r)\cos\gamma)/V \\
\dot{\psi} &= \mathrm{d}\psi/dt = a_{\mathrm{hc}}/V\cos\gamma \\
\dot{\xi} &= d\xi/dt = -V\cos\gamma\cos\psi \\
\dot{\eta} &= d\eta/dt = -V\cos\gamma\sin\psi \\
\dot{h} &= dh/dt = -V\sin\gamma
\end{aligned} \tag{1}$$

For the bank-to-turn control configuration (BTT):

$$a_{\mathrm{vc}} = L\cos\sigma/m, \qquad a_{\mathrm{hc}} = L\sin\sigma/m$$

The guidance problem is to find acceleration commands (or equivalently α, and σ for BTT), which steer the vehicle to its target, subject to the state equations (1); known initial conditions, V_0, γ_0, ψ_0, ξ_0, η_0, and h_0; and known final conditions, ξ_f, η_f, and h_f (equivalent to a fixed target position). The solution must satisfy the following constraint:

$$a_{\mathrm{c}} = \sqrt{a_{\mathrm{vc}}^2 + a_{\mathrm{hc}}^2} \leq a_{\max} \tag{2}$$

The $a_{\max}$ can be related to the limitations of angle of attack, dynamic pressure, heat transfer, loading, etc.

3 Differential Flatness and Inverse problem

Differential flatness was first introduced by Fliess and el al [19] in a differential algebraic context. The important property of flat systems is that we can find a set of variables (equal in number to the number of inputs) such that all states and inputs can be expressed in terms of those outputs and a finite number of their time derivatives without any integration procedure. More precisely, we consider the dynamical system of the general form [20]:

$$\dot{x}(t) = f(x(t), u(t))$$
$$y(t) = h(x(t), u(t))$$

where t is the time variable, x is the n-dimensional state vector, u is the m-dimensional input vector, y is the m-dimensional tracking output vector, $f(.)$ and $h(.)$ are a nonlinear

functions. The system is differentially flat if we can find a set of variables $z(t)\in\Re^m$ which are differentially independent, called flat outputs, of the form

$$z(t)=\Phi(x(t),u(t),\dot{u}(t),...,u^{(a)}(t))$$

such that

$$x(t)=\Psi_x(z(t),\dot{z}(t),...,z^{(b)}(t)),\qquad u(t)=\Psi_u(z(t),\dot{z}(t),...,z^{(b+1)}(t))$$

where Φ and Ψ are smooth functions, $z^{(a)}(t)$ and $z^{(b)}(t)$ are respectively the a and b order time derivative of $z(t)$.

In situations where explicit trajectory generation is required, differential flatness can be very useful: since the behavior of flat systems is determined by the flat outputs only, the trajectories can be planned in output space and then mapped to the appropriate inputs. Many authors have been used differential flatness approach to reentry problem guidance [20-23]; As it is proved in the study of Neckel and et al [24], the nonlinear model (1) is not flat if h, ξ and η are considered as flat outputs. To get around this problem, all studies have kept the longitudinal and lateral motions uncoupled. Therefore, only the longitudinal dynamics are inverted using altitude and curvilinear abscissa as flat outputs, the lateral guidance being ensured via a typical roll reversal technique [e.g. 21]. Decoupling has its limitations. For overcoming these limitations, choosing ψ and γ as flat outputs and solving problem by inverse approach is proposed.

To apply this concept, the independent variable is changed from t to ξ in the system equations (1). (The independent variable may be any monotonous variable; Archer's study [25] would be useful for independent variable selection in RV guidance.) After that, we solve for acceleration commands:

$$\begin{aligned} a_{\mathrm{vc}} &= (V^2/r-g)\cos\gamma - V^2\cos\gamma\cos\psi\,\gamma' \\ a_{\mathrm{hc}} &= -V^2\cos^2\gamma\cos\psi\,\psi' \end{aligned} \tag{3}$$

On the other hand, with geometrical considerations, γ and ψ are obtained:

$$\tan\gamma=\cos\psi\,\mathrm{d}h/\mathrm{d}\xi,\qquad \tan\psi=\mathrm{d}\eta/\mathrm{d}\xi \tag{4}$$

In Eq. (3), the desired trajectory shape enters through the curvature terms γ' and ψ', obtained by the implicit differentiation of Eq. (4) with respect to ξ:

$$\begin{aligned} \gamma' &= (h''\cos\psi - h'\psi'\sin\psi)\cos^2\gamma \\ \psi' &= \cos^2\psi\,\eta'' \end{aligned} \tag{5}$$

These functions introduce second derivative terms h'' and η''. Therefore, the guidance commands are related to the shape of the trajectory. An admissible trajectory must satisfy the relation, Eq. (2).

Actual acceleration $\boldsymbol{a}$ lags the acceleration command $\boldsymbol{a}_{\mathbf{c}}$, whose components are specified by Eqs. (3) and (5). For three degree-of-freedom (3DOF) simulations, noninstantaneous response could be modeled by a first-order lag:

$$\mathrm{d}\boldsymbol{a}/dt+\boldsymbol{a}/\tau=\boldsymbol{a}_{\mathbf{c}}/\tau,$$

where the time constant τ approximates the dominant closed loop pole of autopilot and actuator. In the sequel, instantaneous response ($\tau\to 0$) assumed, and it follows that the acceleration commands of Eq. (3) are the actual acceleration components ($\boldsymbol{a}=\boldsymbol{a}_{\mathbf{c}}$).

4 Trajectory generation

Many methods have been used for trajectory generation [3, 9, 26, 27]; all of them having many parameters and requiring specific conditions. In this paper, the Bezier curve is suggested for trajectory generation.

In view of its properties, this curve has been used in various fields of study such as computer graphics [18], robotic guidance [28, 29], airfoil design [30, 31], and trajectory optimization [32]. Mathematically, a parametric Bezier curve of order *n* is defined by

$$P(u) = \sum_{i=0}^{n} B_i J_{n,i}(u) \tag{6}$$

where the Bezier or Bernstein basis or blending function is

$$J_{n,i}(u) = C_N^i u^i (1-u)^{n-i}, \quad C_N^i = \frac{n!}{i!(n-i)!}$$

and u denotes the parameter of the curve taking values in [0, 1]. So, as seen from Eq. (6), the Bezier curve is completely determined by Cartesian coordinates of the control points. The derivative of order *r* of a Bezier curve can be derived as:

$$\frac{d^r}{du^r} P(u) = \frac{n!}{(n-r)!} \sum_{i=0}^{n-r} \Delta^r B_i J_{n-r,i} \tag{7}$$

where, for $i = 0, \ldots, n$,

$$\Delta^0 B_i = B_i,$$
$$\Delta^k B_i = \Delta^{k-1} B_{i+1} - \Delta^{k-1} B_i \quad for \quad k = 0, \ldots, r$$

It is clear that the derivative of order *r* of a Bezier curve at one of its end points only depends on the *r*+1 control points nearest (and including) that end point. It follows that, at *u=0*:

$$\begin{aligned} P'(0) &= n(B_1 - B_0) \\ P''(0) &= n(n-1)(B_2 - 2B_1 + B_0) \end{aligned} \tag{8}$$

For this problem, the parameter u is equal to the normalized range $\bar{\xi} = (\xi - \xi_0)/(\xi_f - \xi_0)$, and the Bezier approximation of the trajectory is determined by coordinates (h_i, η_i) of the control points B_i. With the allowable assumption n=3 for reentry trajectories, the first point $B_0 = (h_0, \eta_0)$ and last point $B_3 = (h_f, \eta_f)$ will be fixed. Now, we have to determine the middle control points $B_1 = (h_1, \eta_1)$ and $B_2 = (h_2, \eta_2)$. In the beginning of trajectory, the second control point, B_1, can be set using Eqs. (4) and (8):

$$\begin{aligned} \eta_1 &= \eta_0 + \lambda \tan\psi_0 / 3 \\ h_1 &= h_0 + \lambda \tan\gamma_0 \sec\psi_0 / 3 \end{aligned} \tag{9}$$

where $\lambda = (\xi_f - \xi_0)$.

On the other hand, from Eqs. (3), (5), and (8), we have:

$$\begin{aligned} a_{hc} &= f_1(\eta_2, \eta_1, \eta_0, X_0) \\ a_{vc} &= f_2(h_2, h_1, h_0, \eta_2, \eta_1, \eta_0, X_0) \end{aligned} \tag{10}$$

where

$$f_1 = -6V_0^2 \cos^2\gamma_0 \cos^3\psi_0(\eta_2 - 2\eta_1 + \eta_0)/\lambda^2$$
$$f_2 = (V_0^2/r_0 - g_0)\cos\gamma_0 - V_0^2 \cos^3\gamma_0 \cos^2\psi_0$$
$$*[6(h_2 - 2h_1 + h_0) - 9\sin 2\psi_0(\eta_2 - 2\eta_1 + \eta_0)(h_1 - h_0)/\lambda]/\lambda^2$$

We know,

$$|a_{hc}| \le a_{max},$$

therefore from Eqs. (3), (5), and (8), we get:

$$g_1(\eta_1, \eta_0, X_0) \le \eta_2 \le g_2(\eta_1, \eta_0, X_0) \tag{11}$$

Where

$$g_1 = -\frac{a_{max}\lambda^2}{6V_0^2 \cos^2\gamma_0 \cos^3\psi_0} + 2\eta_1 - \eta_0, \qquad g_2 = \frac{a_{max}\lambda^2}{6V_0^2 \cos^2\gamma_0 \cos^3\psi_0} + 2\eta_1 - \eta_0$$

With a value in this boundary, a_{hc} may be determined from Eqs. (2) and (10):

$$g_3(h_1, h_0, \eta_2, \eta_1, \eta_0, X_0) \le h_2 \le g_4(h_1, h_0, \eta_2, \eta_1, \eta_0, X_0) \tag{12}$$

where

$$g_3 = [\frac{-\sqrt{(a_{max}^2 - a_{hc}^2)} + (V_0^2/r_0 - g_0)\cos\gamma_0}{V_0^2 \cos^3\gamma_0 \cos^2\psi_0}$$
$$+ 9\sin 2\psi_0(\eta_2 - 2\eta_1 + \eta_0)(h_1 - h_0)/\lambda^3]\lambda^2/6 + 2h_1 - h_0,$$
$$g_4 = [\frac{\sqrt{(a_{max}^2 - a_{hc}^2)} + (V_0^2/r_0 - g_0)\cos\gamma_0}{V_0^2 \cos^3\gamma_0 \cos^2\psi_0}$$
$$+ 9\sin 2\psi_0(\eta_2 - 2\eta_1 + \eta_0)(h_1 - h_0)/\lambda^3]\lambda^2/6 + 2h_1 - h_0$$

By selecting the third control point, B_2, the initial trajectory will be generated, and the RV will follow it so long as the constraints are satisfied. When the acceleration commands of Eq. (6) exceed the maximum allowable acceleration, acceleration command saturation causes the actual values of h, η, γ, and ψ to deviate from the desired values along the initial trajectory. Holding the terminal conditions fixed, Bezier control points should be continuously updated with instantaneous λ values. Using C_0 (position), C_1 (angle), and C_2 (acceleration) continuity conditions, the new trajectory's control points can be obtained automatically. Therefore, for guiding the RV, the only necessary task is to select the third control point, B_2, for the initial Bezier trajectory. It must be noted that all choices in the boundaries of Eqs. (11) and (12) guarantee that the RV reaches the target while satisfying the constraints.

In the case that the final velocity orientation is constrained, a fourth-order Bezier curve is suggested (e.g., if the final velocity vector is constrained to γ_f, and ψ_f, the B_0, B_1, B_2, and B_4 control points would be treated the same way). The fourth control point, B_3, can be set like B_1:

$$\eta_3 = \eta_4 - \lambda\tan\psi_f/3, \qquad h_3 = h_4 - \lambda\tan\gamma_f \sec\psi_f/3$$

5 Simulation results and discussion

To demonstrate the effectiveness of this guidance law, it has been used in a 3DOF (point mass) simulation containing a standard atmosphere, and aerodynamic coefficients as functions of Mach number, angle of attack, and Reynolds number in tabulated form for RV model based on [33].

It is assumed that $\eta_2=2\eta_1-\eta_0$ in Eq. (11) and $h_2=g_4$ in Eq. (12) (i.e., selection of η_2 and h_2 in such a manner that $a_{hc}=0$ and $a_{vc}=a_{max}=g$ at the beginning of flight (For BTT: $L_c=L_{max}$, and $\sigma=180^{\circ}$)). Results of this assumption (EXP) are compared with the performance of pure proportional navigation (PPN) with N=3 [34] and shown in Figs. 2-7 for a sample period of $\Delta t=0.01$ sec.

A fourth-order, fixed-step, Runge-Kutta integrator is used in all simulations. The trajectory boundary conditions for the example problem are shown in Table 1. Both methods have nearly the same impact accuracy, with different behaviors.

Figures 2 and 3 display the simulation flight path profiles. PPN turns quickly to line up with the target, whereas EXP (with those assumptions for selecting η_2 and h_2) shifts the majority of flight time to higher altitudes, where drag is low. Therefore, differences in horizontal paths and acceleration command (Fig. 4) have a very small effect on the velocity profile (Fig. 6).

Table 1: Trajectory boundary conditions.

	Initial	Final
ξ, km	80	0
η, km	2	0
h, km	30	0
V, m s^{-1}	4000	maximum
γ, deg	20	unconstrained
ψ, deg	0	unconstrained

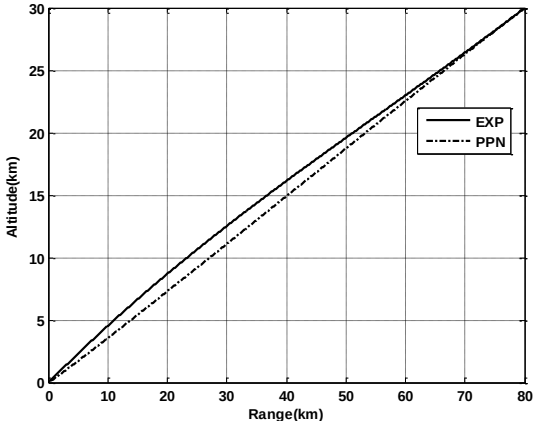


Figure 2: Altitude comparison.

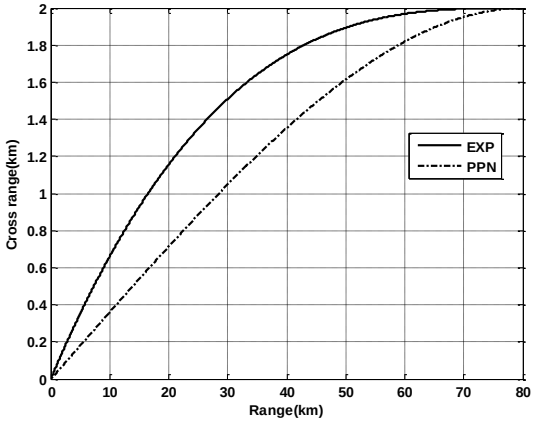


Figure 3: Cross range comparison.

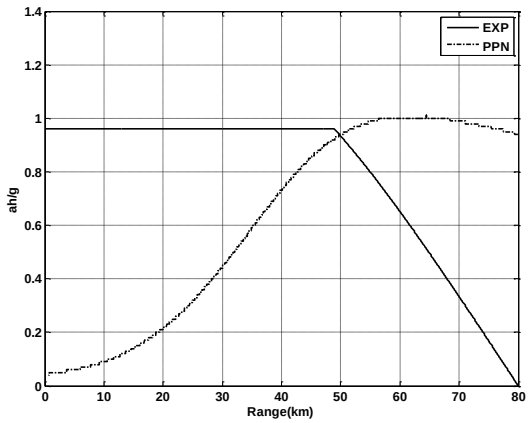


Figure 4: Horizontal acceleration command.

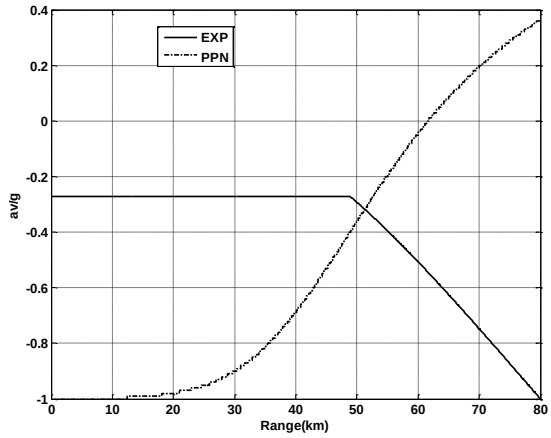


Figure 5: Vertical acceleration command.

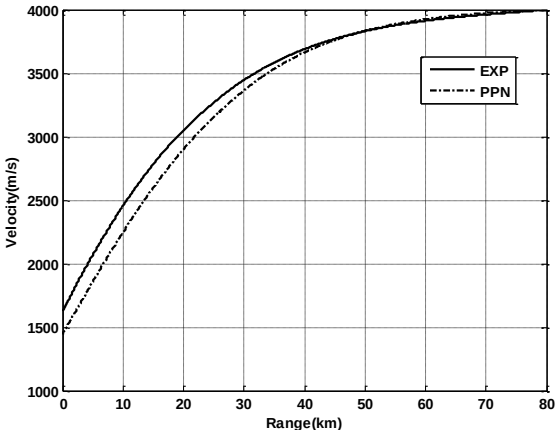


Figure 6: Velocity comparison.

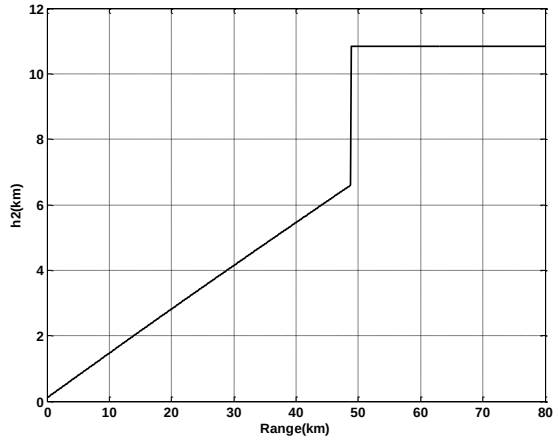


Figure 7: Variation of h_2.

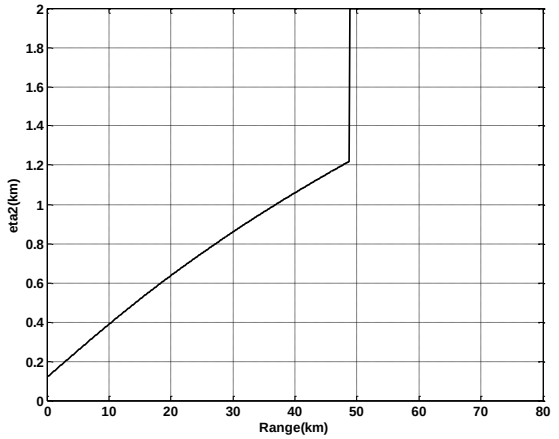


Figure 8: Variation of η_2.

Velocity-range profiles in Fig. 6 are grouped in a similar fashion to the flight profiles. The final velocity produced by the EXP scheme is 1620 m s^{-1}, while the PPN solution is over 13% less (1410 m s^{-1}). This advantage in kinetic energy loss may be useful in some missions. It will be useful if the third control point, B_2, can be optimized with any numerical method to create a performance index such as final velocity, control effort, dynamic pressure, heat transfer, etc.

6 Concluding remarks

An explicit guidance method was devised to obtain descent trajectories for a prescribed destination using differential flatness and inverse problem combination. The guidance commands are related to shape of trajectory, specified by a Bezier curve, to reach the target. During periods of command saturation, the instantaneous Bezier control points vary until sufficient control is available to follow the trajectory. Bezier control points can be determined using any assumption or parameter optimization method. The proposed method is characterized by the following advantages: 1) a priori satisfaction of the

boundary conditions; 2) an absence of "wild" trajectories during path generation; 3) an analytical (parametric) representation of reference trajectory with minimum parameters; 4) applicability to any RV configuration, regardless of its lift-to-drag ratio or range of flight Mach number regime; 5) applicability to any control schemes (bank-to-turn or skid-to-turn), and 6) offline nominal trajectory and Time-to-Go independence. Results compared to pure proportional navigation for terminal velocity were excellent.

7 References

[1] M.H. Gräβlin, J. Telaar, U.M. Schottle, *Ascent and reentry guidance concept based on NLP-methods*, Acta Astronautica, 55, (2004), pp. 461-471.

[2] S.K. Shrivastava, M.S. Bhat, S.K. Sinha, *Closed-loop guidance of satellite launch vehicle- an overview*, J. of the Institute of Engineers, 66, (1986), pp. 62-76.

[3] O.A. Yakimenko, *Direct method for rapid prototyping of near-optimal aircraft trajectories*, J. of Guidance, Control, and Dynamics, 23 (5), (2000), pp. 865-875.

[4] P. Lu, *Inverse dynamics approach to trajectory optimization for an aerospace plane*, J. of Guidance, Control and Dynamics, 16 (4), (1993), pp. 726-731.

[5] M. Borri, C.L. Bottasso, F. Montelaghi, *Numerical approach to flight dynamics*, J. of Guidance, Control, and Dynamics, 20 (4), (1997), pp. 742-747.

[6] O. Kato, I. Sugiura, *An interpretation of airplane general motion and control as inverse problem*, J. of Guidance, Control, and Dynamics, 9 (2), (1986), pp. 198-204.

[7] S.H. Lane, R.S. Stengel, *Flight control design using non-linear inverse dynamics*, Automatica, 24 (4), (1988), pp. 471-483.

[8] E. Sentoh, A.E. Bryson, *Inverse and optimal control for desired outputs*, J. of Guidance, Control and Dynamics, 15 (4), (1992), pp. 687-691.

[9] M.E. Hough, *Explicit guidance along an optimal space curve*, J. of Guidance, Control, and Dynamics, 12 (4), (1982), pp. 495-504.

[10] G. Leng, *Guidance algorithm design: a nonlinear inverse approach*, J. of Guidance, Control, and Dynamics, 21 (5), (1998), pp. 742-746.

[11] C.F. Lin, L.L. Tsai, *Analytical solution of optimal trajectory-shaping guidance*, J. of Guidance, Control, and Dynamics, 10 (1), (1987), pp. 61-66.

[12] S.K. Sinha, S.K. Shrivastava, *Optimal explicit guidance of multistage launch vehicle along three-dimentional trajectory*, J. of Guidance, Control and Dynamics, 13(2),(1990), pp394-403.

[13] V.T. Taranenko, V.G. Momdzhj, *Direct method of modification in flight dynamic problems*, Mashinostroenie Press, Moscow, Russia (in Russian) (1986).

[14] J.D.M. Cameron, *Explicit guidance equations for maneuvering re-entry vehicles*, Proceedings of the IEEE Conference on Decision and Control, New York, USA: Inst. of Electrical and Electronics Engineers, (1977), pp. 670-678.

[15] J. Page, R. Rogers, *Guidance and control of maneuvering reentry vehicles*, Proceedings of the IEEE Conference on Decision and Control, New York, USA: Inst. of Electrical and Electronics Engineers, (1977), pp. 659-664.

[16] M. Mortazavi, *Trajectory determination using inverse dynamics and reentry trajectory optimization*, PhD Thesis, Moscow Aviation Institute, (in Russian) (2000).

[17] R. Esmaelzadeh, A. Naghash, M. Mortazavi, *Explicit Reentry Guidance Law Using Bezier Curves*, Transactions of the Japan Society for Aeronautical and Space Sciences, 50(170), (2008), pp. 225-230.

[18]D.F. Rogers, J.A. Adams, *Mathematical elements for computer graphics*. McGraw-Hill, New York, USA, (1990),.
[19]M. Fliess, J. L´evine, P. Martin, P.Rouchon, *Flatness and defect of nonlinearsystems: introduction theory and examples*, Int. Journal of Control, 61, (1995), pp. 1327–1361.
[20]M. Zerar, F. Cazaurang, A. Zolghadri, *LPV Modeling of Atmospheric Re-entry Demonstrator for Guidance Reentry Problem*, Proceedings of the 44th IEEE Conference on Decision and Control, Seville, Spain, December 12-15, 2005.
[21]V. Morio, F. Cazaurang, P. Vernis, *Flatness-based hypersonic reentry guidance of a lifting-body vehicle*, Control Engineering Practice, 17, (2009), pp. 588-596.
[22]H. Sira-Ramirez, *Soft Landing on a Planet: A Trajectory Planning Approach for the Liouvillian Model*, Proceedings of the American Control Conference, San Diego, 1999.
[23]L.G.W. Sun, Z. Zheng, *Trajectory optimization for guided bombs based on differential flatness*, Proceedings of the Chinese Control and Decision Conference, 2009.
[24]T. Neckel, C. Talbot, N. Petit, *Collocation and inversion for a reentry optimal control problem*, Proceedings of the 5th International Conference on Launcher Technology, 2003.
[25]S.M. Archer, D.D. Sworder, *Selection of the guidance variable for a re-entry vehicle*, J. of Guidance, Control and Dynamics, 2 (2), (1979), pp. 130-138.
[26]K.B. Judd, T.W. McLain, *Spline based path planning for unmanned air vehicles*, AIAA Guidance, Navigation, and Control Conference and Exhibit, Montreal, Canada, (2001).
[27]Z. Shen, *On-board three-dimensional constrained entry flight trajectory generation*, PhD Thesis, Iowa State University, (2002).
[28]K. Nagatani, I. Yosuke, T. Yutaka, *Sensor-based navigation for car-like mobile robots based on a generalized voronoi graph,* Advanced Robotics, 17(5), (2003), pp. 385-401.
[29]J. Zhang, J. Raczkowsky, A. Herp, *Emulation of spline curves and its applications in robot motion control*, IEEE Conf. on Fuzzy Systems, Orlando, FL, USA, (1994), pp. 831-836.
[30]H.S. Myong, J.L Kyu, *Bézier curve application in the shape optimization for transonic airfoils*, 18TH AIAA Applied Aerodynamics Conference, Denver, USA, (2000), pp 884-894.
[31]P. Venkataraman, *Unique solution for optimal airfoil design*, 15th AIAA Applied Aerodynamics Conference, Atlanta , GA, USA, (1997), pp. 205-215.
[32]J.A. Desideri, S. Peigin, S. Timchenko, *Application of genetic algorithms to the solution of the space vehicle reentry trajectory optimization problem*, INRIA Sophia Antipolis, Research Report No. 3843, (1999).
[33]T.F. Lin, and et al., *Novel approach for maneuvering reentry vehicle design*, J. of Spacecraft and Rocket, 40 (5), (2003), pp. 605-614.
[34]N.A. Shneydor, *Missile guidance and pursuit; kinematics, dynamics, and control*, Horwood Publishing Ltd, Chichester, (1998).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Application of inverse analysis in electromagnetic grinding of brown coal for obtaining an optimal particle size distribution – a heuristic approach.

M. Gandor, K. Sławiński, K. Knaś, B. Balt, W. Nowak

Czestochowa University of Technology
Institute of Advanced Energy Technologies
Dabrowskiego 73, 42-201 Czestochowa

mgandor@fluid.is.pcz.pl; kslawinski@fluid.is.pcz.pl; kknas@fluid.is.pcz.pl; bbalt@fluid.is.pcz.pl; wnowak@is.pcz.czest.pl

Key words: milling, brown coal, particle size distribution, optimization, electromagnetic mill

Abstract

This paper presents the research results of milling process optimization in the electromagnetic mill due to obtain the predetermined particle size distribution of the brown coal. Because of the high relevance of brown coal to Polish energy industry (power plants produce 9433 MW of electrical power from brown coal, which corresponds to about 34% share in total fuel usage structure in energy industry of Poland – II qtr. 2013 [1]), there is a great need to foresee highly efficient methods of its mining, valorisation and low emission combustion alongside with CO_2 capture. This paper proposes one of the concepts of adapting low-rank coal to being utilized in modernized and newly built plants, is a simultaneous grinding and drying process in an electromagnetic mill system. In addition, this solution is energy efficient and what is more reduces the space required for its adaptation, thanks electromagnetic mill's compact installation design. It is essential to achieve the desired characteristics of the product through the adequate control of the processes. Major concern of this case study was set on determination of optimal grinding parameters in the electromagnetic mill in order to obtain two products of a desired size distribution (1 – 6,3 mm for the application in fluidized bed boilers and 0 – 315 μm for boiler burners). Authors have presented some theoretical deliberations of the mechanisms and physical phenomena occurring while solid particles fragmentation as well as a literature review of the subject. The process complexity level, taking place in the active area of an electromagnetic mill, involves the influence between particle – milling rod, particle – particle interactions, volume of milling rods degree or coal particle residence time on the size distribution of the product, account for nonlinearity of the problem and make the conditions difficult to rescale. A heuristic approach to inverse problem, hence was chosen to analyse the differences between the desired and obtained particle size distribution. The examinations concerned grinding parameters, which were: total amount of rods (volume-based) and rod sizes (single and multi-size combinations of milling elements). As an investigated material were chosen equivalent samples of Polish brown coal with a particle diameter size ranging of 0 to 10 mm. Influence of the total volume of rods was examined using three values: 100 ml, 150 ml and 200 ml. Two grinding aid element sizes were chosen in a form of ferromagnetic rods: fine rods of the size of 10 × 1 mm and coarse rods of the size of 20 × 2 mm.

1 Introduction

Poland is among countries with significant brown coal deposits. In 2012, the production of brown coal reached almost 80×10^6 Mg, of which $67{,}5 \times 10^6$ Mg constituted for an energetic coal. It accounts for one third of present basic fuels. With the rising demand for and electric energy production (38 GW of power installed in 2013 and a predicted 60 GW in 2030) the significance of brown coal will remain at high level [1, 2]. An increasing interest in this fossil is also justified by the expected period of its mining, which is considered to be about 300 years, in comparison with 200 years for a primary energy carrier - hard coal [2]. In many countries, including in Poland, research is conducted on a large scale related to the effective utilization of brown coal in energy sector and also on its valorization and protection of its deposits [3].

Several problems however accompany the utilization of brown coal associated with the characteristics of this fuel, that need to be taken into account when developing the old and creating new technologies. From the technical point of view, the biggest distinction between coals of different ranks is the amount of moisture and ash content, which has a significant impact on the energy efficiency of the boiler as well as some maintenance problems that may occur during the preparation, transport and combustion. Elevated moisture and ash content cause a decrease in energy density of the fuel, which in turn reduces or excludes the possibility of an economically justified long distance transport. Furthermore, an increased flue gas volume accompanying the combustion of brown coal has negative ecological consequences. Adding brown coal to higher-rank coals, allows to lower the NO_x emissions by introducing volatile components in the vicinity of combustion zone [4]. Research is also conducted on coal gasification in pressure reactors and on underground gasification in situ.

Coal milling plays an important role in its preparation for utilization in various technologies. Grinding and especially selective grinding is one of the basic methods for solid fuel adaptation in a "clean coal technology" program. The assumption of this program is to diminish the negative impact of coal combustion on the environment. Four main methods are favourable: pre-combustion (modification of fuel prior to its combustion), advanced combustion (modification of combustion process, for example with innovative boiler construction), advanced post-combustion (flue gas treatment) and conversion (gasification, pyrolisis etc.) [5]. For pre-combustion the usage of coal blends of desired properties and selective grinding are considered in order to clean the fuel from high contents of sulphur, which occurs in coal as pyrite. Deep coal processing is performed by fine-milling and removing small particles of waste rock. One of the methods of advanced combustion is using a blend of low and high-rank coals in order to obtain lower NO_x emissions. Air dispersed finely ground brown coal also seems to be an interesting cheaper alternative for fuels used in boiler burners, such as fuel oil or mazout. Fine milling of coal allows for lower loss on ignition – smaller particles burn faster and fuel conversion is greater than that of coarse particles, which in turn allows more compact and effective boiler design. For the methods based on conversion it is also important to obtain a desired size of particles, for example for utilization in pressure fluidized gasification, one of the technologies developed under the Strategic Research Program "Advanced technologies for energy generation", in which the authors actively participate. Research has shown, that parameters like volatile matter initiation temperature, peak reactivity value, temperature of char burnout, total burnout time or temperature of self-heating are influenced not only by the size distribution of particles but also by the grinding method used to obtain those particles [4]. Comminution is a highly energy intensive process, the power consumption according to literature can reach from 1 to about 15 kWh/Mg [6, 7]. The range of values is wide and it is affected by parameters like coal grindability, level of moisture or energy required for drying of feed material and product separation. Total power consumption of grinding process is higher, when efficiency of the machines is considered.

The operation of a prototype electromagnetic mill installation for grinding and drying of solid fuels proposed by authors is based on a rapid movement of small milling rods, which is forced with an alternating electromagnetic field. In this technology a high electric energy into mechanical energy conversion occurs and the heat generated by the milling rods is used to reduce moisture in the material. The milling rods move chaotically inside the chamber, hitting the fuel particles, each other and also the walls of the milling chamber. So far no accurate model exist for the phenomena accompanying electromagnetic grinding and a vast majority of parameters are obtained

experimentally. Therefore it is necessary to conduct an inverse analysis of grinding in an electromagnetic mill in order to determine the parameters allowing for an optimal particle size distribution of the product.

2 . Comminution

Comminution is a process of reducing the size of solid bodies to produce smaller elements with externally applied destructive forces. The source of the force can be either mechanical, or more rarely – chemical. Mechanical forces acting on the particles of fuel are dependent on the technology used, in most cases several comminution mechanisms occur simultaneously. Because grinding is the effect of a combination of many processes, it is hard to perform a modelling. As in the case of many disperse material technologies, the analysis of a path of a single particle in the milling area is difficult and the characteristics that are true for a particle can be false for the material in a macro scale. Therefore it is common to make an estimation and to average those parameters.

One of the most essential parameters in milling coal is its grindability. It is a result of coals hardness, strength, crystalline structure. Those features are affected by the rank of coal, petrography and mineral matter content. There is no single reliable method of classification of coals on the basis of their grindability. One of the most common ways to measure coals grindability is the Hardgrove test. The value of Hardgove index is determined via conducting a laboratory scale experiment, thus it doesn't reflect the milling process in an industrial full-scale installation. The value of this index arranges different coals according to the "easiness" with which they are fragmented, however it doesn't give the information on the amount of energy needed to comminute the material. Many authors claim that the value of Hardgrove index is non-additive and it shouldn't be used for estimating grindability of coal blends made from coals of significantly different Hardgrove index.

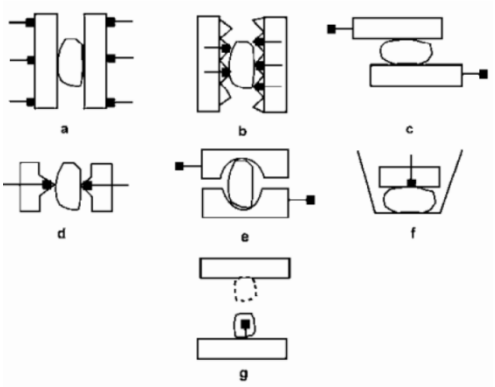


Fig. 1 Different mechanisms of comminution: a – crushing, b – breaking, c – attrition, d – lamination, e – shearing, f – mastication, g – hitting [8].

There are several theories linking grinding with power consumption, they however have some limitations and do not apply in all cases. Kicks theory specifies the energy needed for the fragmentation of particles with the sizes exceeding 1 mm, Bonds theory specifies the approximate energy needed for the fragmentation of particles in the size range of tens to hundreds of μm and Rittingers theory concerns the particles of ultra-small particles. Generally the amount of energy required for fragmentation rises with decreasing of particle size and is proportional to the newly-created surface area. Some improvement of the comminution theory is the Shi-Kojovic model from 2007, which allows to estimate the energy required for grinding in a broad range of sizes (Hardgrove test is made for particles of the size 0,6 – 1,18 mm). Research on the influence of particle size and the energy of comminution showed, that for low values of specific fragmentation energy, with the normalization of the Shi-Kojovic equation, there is a linear correlation between experimentally obtained Hardgrove index and specific energy of comminution. For deeper milling there is a need for larger energies, which can be estimated by the Shi-Kojovic model [9].

An important problem in case of grinding is the loss of energy for grinding of the particles which sizes are already in a desired range. It is essentially relevant in a case when a mill has a limited possibility for particle separation in the process. In such situations it is crucial to separate particles of different granulation prior to milling or to modify the installation [7]. Additional problems are caused by the moisture content in fuel, which lowers the ability for grinding, causes the particles to stick to the grinding surfaces and increases the energy losses due to the deformation in non-elastic bouncing.

Lack of an absolute information on the processes taking place inside the chamber of the electromagnetic mill (the influence of the magnetic field, interaction between milling rods and the material, the number of collisions, influence of the moisture loss, effect of the moisture in the feed material on coals grindability) predestinates the problem for being solved with the use of inverse analysis. Additionally, with so many unknown parameters, a heuristic approach for the problem was selected in order to omit the initial estimation of the fragmentation model, and by analyzing the initial conditions and the obtained results. The goal of the analysis was to describe the processes inside the mill that determine the particle size distribution of the product.

3 A conception of an electromagnetic mill for grinding and activation of the material

Electromagnetic mill is a device for grinding of particulate matter of a maximal size of couple of millimeters with the use of small ferromagnetic grinding aid. Grinding rods are pushed with the force exerted on them by an alternating electromagnetic field produced by the inductor. The field is created by the salient poles of the inductor with a three-phase alternating current from a 50 Hz power grid. The main element of the electromagnetic device for grinding and drying of coal is the mill, which consist of two basic parts: milling chamber with grinding aid and a stator with salient field poles. The milling chamber is a non-ferromagnetic tube in which small ferromagnetic milling rods are rotating suspended in magnetic field. Those two elements create work area in which the feed material is subjected to mechanical, thermal and magnetic treatment. With the increasing magnetic induction inside the chamber, small grinding rods are forced to rotate and move faster and more chaotically. A ferromagnetic rod suspended in magnetic field becomes a magnetic dipole and is attracted by the field with certain force. Thanks to the small size of milling rods and optimized dimension proportions it is possible to gain high acceleration and to quickly obtain maximum velocity of the grinding aid. Because of the very high rotation speed, there is a very large amount of consecutive collisions of rods and fuel particles, which allows for much quicker process when compared to the mills in which the period between impacts is extended.

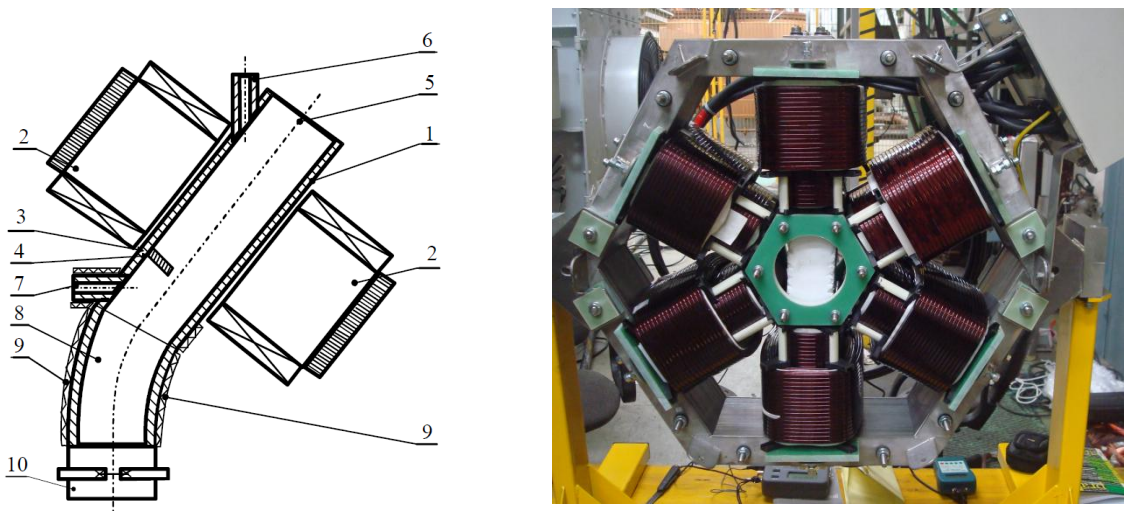


Fig. 2 Design scheme of the electromagnetic mill (left). 1 – milling chamber, 2 – inductor poles, 3 – stop, 4 – vapour transport holes, 5 – material inlet, 6 – grinding aid/gas inlet, 7 – vapour collection vent, 8 – product outlet, 9 – thermal insulation, 10 – fastener for a tight connection of the milling chamber with product container. On the right – physical model of the inductor.

Parameters influencing the operation of inductor, involving the geometry of the inductor, milling chamber as well as dimensions of grinding rods have been determined by mathematical modelling and were presented in a form of design methodology [10].

Research was conducted on a model with a milling chamber diameter of 100 mm.

4 Problem identification

The goal of this research was to establish the main parameters determining particle size distribution of the product (100 – 315 μm range for utilization as fuel in new generation coal dust burners and 1 – 6,3 mm used as fuel for fluidized bed boilers). Since at this stage of research no accurate model exists of the comminution process in an electromagnetic mill, very little information is available on the optimal amounts and sizes of milling rods and volume/weight ratios of grinding aid to coal. In theory, the usage of small milling rods should result in smaller particle sizes of the product. Electromagnetic mill is therefore especially effective in case of fine milling. The force exerted on grinding aid by magnetic field is dependent on the second power of the field induction magnitude in the vicinity of the milling rod. The induction in the neighbourhood of the rod in its axis is smaller when compared to milling rod – free region, so extensive amount of rods and their high density results in weaker induction in the milling chamber. The influence of the differently sized milling rod blends on the obtained particle size distribution of the product remained unexplored and it was also a subject of this research.

5 Experimental

During the experiments, two sizes of milling rods were used: 10 × 1 mm and 20 × 2 mm (length × diameter). Aspect ratio 10 : 1 of length to diameter are dictated by calculations, which showed that for such dimensions of both the magnitude of the induced magnetic field as well as the efficiency of grinding are optimal. The amount of grinding aid in each test was 100 ml, 150 ml and 200 ml.

The feed material used in the tests was brown coal, characterized by the following parameters:

Table 1. Physico-chemical parameters characterizing the sample of brown coal used in the study

Parameter	Value
Received basis:	
Total moisture W_t^r [%]	51,14
Ash kontent A^r [%]	10,74
Total sulphur S_t^r [%]	0,91
Lower heating value Q_i^r [kJ/kg]	8555
Analytical basis:	
Moisture W^a [%]	9,15
Ash A^a [%]	20,04
Volatiles V^a [%]	38,28
Higher heating value Q_s^a [kJ/kg]	23330
Lower heating value Q_i^a [kJ/kg]	21091
Carbon C^a [%]	43,80
Hydrogen H^a [%]	4,65
Nitrogen N^a [%]	0,53
Total sulphur S_t^a [%]	1,77
Dry Basis:	
Ash A_d [%]	23,08
Total sulphur S_t^d [%]	1,82
Volatiles V^{daf} [%]	45,93
Higher heating value Q_s^{daf} [kJ/kg]	19000

For each test a constant volume of feed material was used. Flow of the material was forced by gravity.

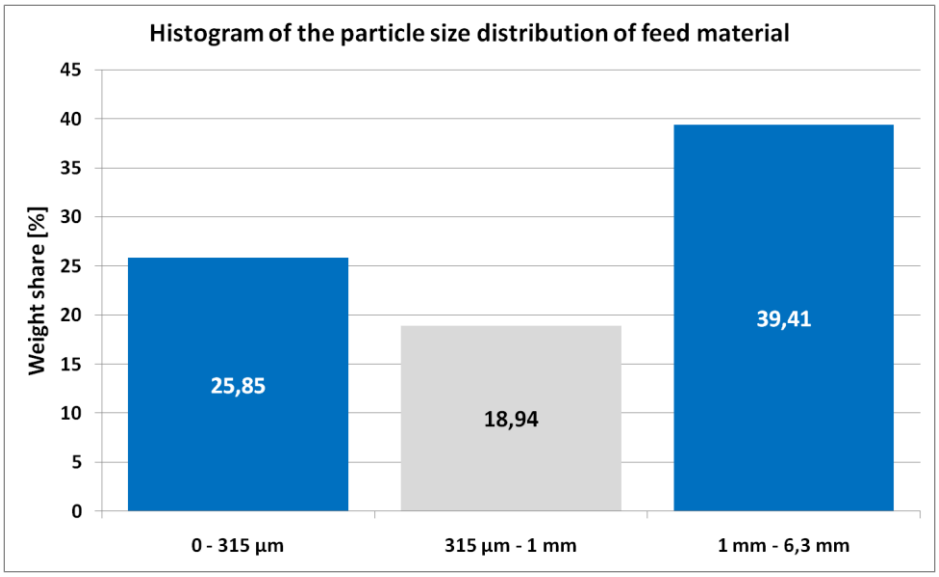


Fig. 3 Histogram of the particle size distribution of feed material, blue colour - desired fraction, gray colour - the intermediate fraction.

Tests were carried out using different volumes (100 - 200 ml), and different sizes of milling rods (10 × 1 and 20 × 2 mm) in order to compare the particle size distribution of the product with the particle size distribution of the feed material. This has resulted in a series of histograms which were used to determine the effect of the grinding aid on the quality of the final product in a form of fractions 0 - 315 µm and 1 - 6.3 mm.

6 Results and discussion

The study begun with the use of 20 × 2 mm milling rods.

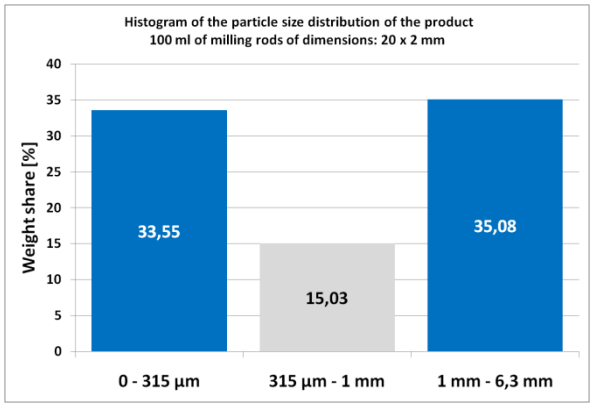


Fig. 4 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray colour - the intermediate fraction. Grinding with 100 ml of milling rods of dimensions 20 × 2 mm.

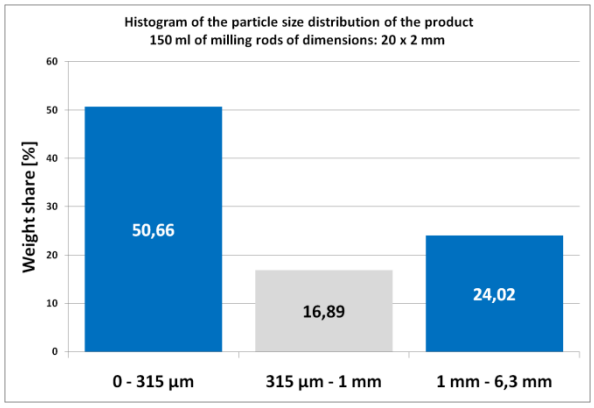


Fig. 5 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray colour - the intermediate fraction. Grinding with 150 ml of milling rods of dimensions 20 × 2 mm.

Particle size distribution analysis of the product showed a slight reduction in the grinding of the coarse fraction and the intermediate, respectively, 4,33% and 3,91% and the increase in the share of fine fraction from 25,85% to 33,55% by weight. It was found that a low efficiency of grinding carried out in this manner can result from too low volume of the grinding aid with respect to the feed material and milling chamber volume.

Increasing the amount of grinding media should result in further reduction of coarser fractions and dispensing the weight loss between the finer fractions. After increasing the volume of grinding media from 100 to 150 ml, a clear decrease in the coarse fraction mass was observed. Part of the lost mass was found within the intermediate fraction, while the vast majority transferred to the fine fraction. As a result no satisfactory reduction of the intermediate fraction was obtained, and it was found that milling rods of the size of 20 × 2 mm may interact mostly with coal particles larger than 1 mm. It was acknowledged that a more favourable effect can not be achieved by further increasing the amount of grinding aid and it is necessary to utilize finer milling rods for the stronger impact on sub 1 mm particles.

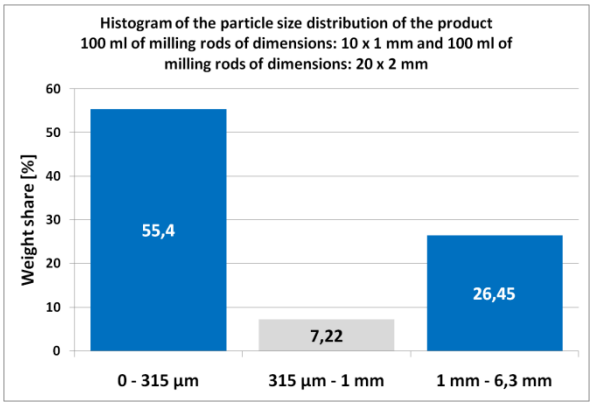


Fig. 6 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray color - the intermediate fraction. Grinding with 100 ml of milling rods of dimensions 10 × 1 mm, and 100 ml of milling rods of dimensions 20 × 1 mm.

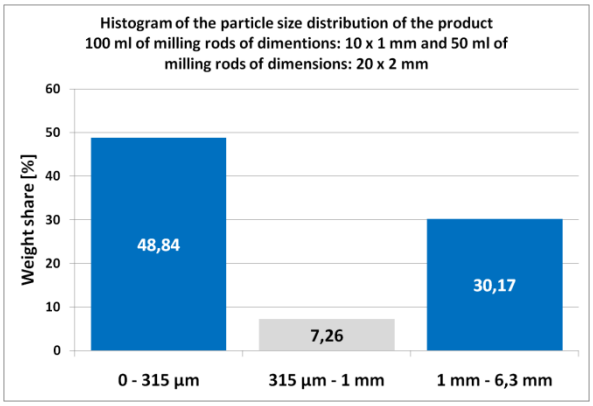


Fig. 7 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray color - the intermediate fraction. Grinding with 100 ml of milling rods of dimensions 10 × 1 mm, and 50 ml of milling rods of dimensions 20 × 2 mm.

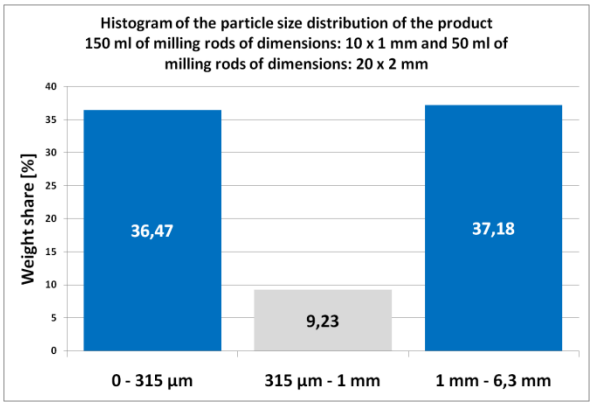


Fig. 8 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray colour - the intermediate fraction. Grinding with 150 ml of milling rods of dimensions 10 × 1 mm, and 50 ml of milling rods of dimensions 20 × 2 mm.

After the change of grinding aid to a blend of 50 : 50 by volume of milling rods with dimensions of 20 × 2 mm , and 10 × 1 mm with a total volume of 200 ml , there was a significant decrease in the share of the intermediate fraction from 16,89 % in the case of using only coarse milling rods to 7,22% in case of using a 50 : 50 blend. As expected , rods of the size of 10 × 1 mm more effectively comminuted particles with the size less than 1 mm. Additionally, a significant share of ultra - fine milling occurred in the process, resulting in more than 55 % content of fraction 0 - 315µm. To obtain information on mutual interactions of different size of milling rods in the grinding process, further studies were performed with the increasing share of the 10 × 1 mm milling rods for both the total volume of 150 and 200 ml. The reduction of the amount of 20 × 2 mm milling rods allowed to keep a larger share of fraction 1 - 6.3 mm. Furthermore, with the increase in the amount of 10 × 1 mm

milling rods up to 150 ml, a minor effect on the reduction of the coarse fraction was registered with a small reduction of the intermediate fraction. The effect of 10 × 1 mm milling rods on milling process was seen for each ratio of the mixture, changing only the ratio between fine and coarse fractions of the product. At this stage , it was found that the presence of small rods (10 × 1 mm) is necessary to reduce the 0 – 315 µm fraction. Effect of the amount of 20 × 2 mm rods on the reduction of 0 – 315 µm fraction is small , and their impact on grinding process is mostly visualized for particles greater than 1 mm. To further reduce the share of the intermediate fraction, it was decided to carry out further studies using different amounts of milling rods with dimensions of 10 × 1 mm.

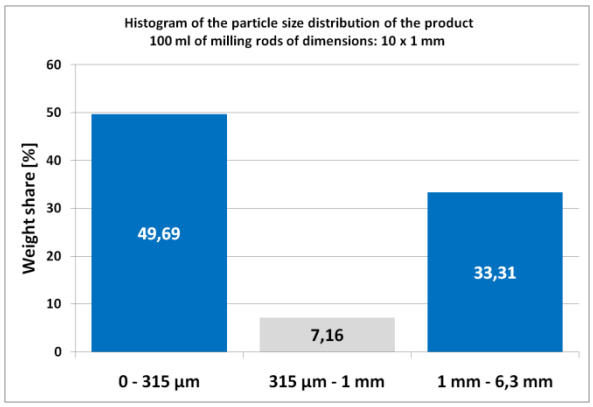


Fig. 9 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray colour - the intermediate fraction. Grinding with 100 ml of milling rods of dimensions 10 × 1 mm.

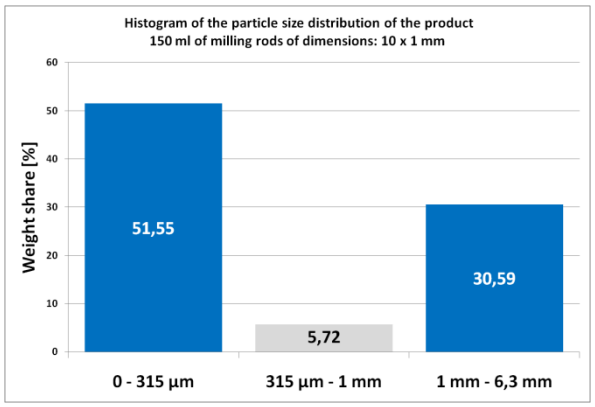


Fig. 10 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray colour - the intermediate fraction. Grinding with 150 ml of milling rods of dimensions 10 × 1 mm.

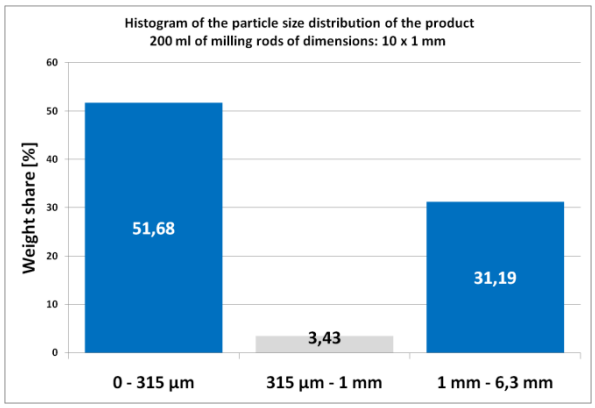


Fig. 11 Histogram of the particle size distribution of the product, the blue colour - desired fraction, gray colour - the intermediate fraction. Grinding with 200 ml of milling rods of dimensions 10 × 1 mm.

As presumed, in the case of milling with 10 × 1 mm rods, the lowest share of the intermediate fraction (from 7.16 to 3.43 %) was obtained. In addition the share decreased with an increasing volume of grinding aid with respect to the volume of the chamber. In each case, similar proportions between fractions were obtained, namely: about 50% of the 0 - 315 µm fraction, about 30% of the 1 - 6.3 mm fraction and 3 to 7 % of the intermediate fraction – which should be re-ground. This result with using 200 ml of 10 × 1 mm grinding aid is the most preferred since it significantly lowers the energy required to re-fragmentation of the intermediate fraction in the next grinding cycle. Different is also the nature of the influence of fine grinding aid on the coarse fraction, when compared to 20 × 2 mm rods. In the case of coarser grinding aid, its total volume influences the fragmentation of the coarse fraction, and the effect increases significantly above 100 ml of grinding aid. For the 10 × 1 mm rods, no such significant impact was observed, whereby it can be concluded, that this type of grinding aid interacts weaker with coarse particles and a few percent reduction in coarse fraction may be due to

crushing occurring in the screw feeder and due to the interactions between the particles of fuel - the collision and mutual attrition.

7 Conclusions

During a research to obtain an optimum particle size distribution of electromagnetically ground lignite, a series of tests were conducted, using different size (10 × 1 mm and 20 × 2 mm) and various volumes of grinding aid (100, 150 and 200 ml). Studies have shown that it is possible to grind the feed in such manner that it is possible to obtain a minimum amount of intermediate fraction, which in turn allows for energy savings from less material requiring re-grinding. It has been found that grinding aid of the size of 20 × 2 mm strongly influences the > 1 mm fraction and by controlling their volume it is possible to obtain greater amounts of 1 - 6.3 mm fraction in the product. In addition, the use of milling rods of that size allows to obtain a significant share of 0 – 315 μm fraction. Still not fully explained is the relatively weak effect on reducing the share of the intermediate fraction. The increased amount of the intermediate fraction as compared to that obtained with finer grinding aid may result from the transfer of weight of the > 1 mm particles into smaller fractions as a result of their fragmentation. Smaller grinding aid can effectively grind particles in the range of 315 μm - 1 mm and therefore reduce the amount of the intermediate fraction. In each case an ultra-fine milling was observed and the size of grinding aid defined the upper limit of the size of affected coal particles. The use of finer grinding aid can produce more homogeneous, fine-grained product, while the coarser milling rods are capable of fragmenting larger particles, resulting in both a linear and a more uniform particle size distribution of the product. A suitable combination of two or more sizes of grinding aid in a total volume considered optimal (200 ml), allows to achieve the desired product properties.

The results presented in this paper were obtained with the research financed by the National Centre for Research and Development under contract SP/E/1/67484/10 – Strategic Research Programme – Advanced technologies for energy production: Development of technologies for highly efficient "zero-emission" coal-fired units with integrated CO_2 capture from flue gas.

8 References

[1] A. Pietraszewski, *Polskie górnictwo węgla brunatnego w pierwszym półroczu 2013 roku*, Węgiel Brunatny 3/84, 2013.
[2] Z. Kasztelewicz, *Węgiel brunatny optymalnym paliwem dla polskiej energetyki w I połowie XXI wieku*, 4 lutego 2010, Warszawa.
[3] J. R. Kasiński, *Zasoby węgla brunatnego w Polsce – stan rozpoznania i podstawowe problemy*, Państwowy Instytut Geologiczny, www.pgi.gov.pl, 2014.
[4] F. Rubiera, A. Arenillas, E. Fuente, N. Miles, J. J. Pis, *Effect of the grinding behaviour of coal blends on coal utilisation for combustion*, Powder Technology 105 (1999) 351 – 356,.
[5] W. Blashke, R. Nycz, *Clean coal-preparation barriers in Poland*, Applied Energy 74 (2003) 343-348.
[6] Lytle J, Choi N, Prisbrey K., *Influence of preheating on grindability of coal*, Int J Miner Process 1992;36:107–12.
[7] V. I. Zasel'skii, G. L. Zaitsev, T. A. Zasel'skaya, *Influence of the granulometric composition of coal batch on the energy requirements of a Hammer Mill*, Coke and Chemistry, Vol. 55, No. 7, pp 282 – 285, 2012.
[8] J. Drzymała, Podstawy mineralurgii. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2009.
[9] Z. Weiran, Z. Yuemin, H. Jaqun, S. Fengnian, D. Chenlong, *Relationship between coal size reduction and energy input in Hardgrove mill*, International Journal of Mining Science and Technology, 22 (2012) 121 – 124.
[10] R. Sosiński, *Opracowanie metodyki projektowania trójfazowych wzbudników z biegunami jawnymi pola wirującego do młynów elektromagnetycznych*, rozprawa doktorska, Politechnika Częstochowska, 2006.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Heuristic approach to inverse analysis of brown coal drying in the electromagnetic mill.

M. Gandor[1], K. Sławiński[1], B. Balt[1], K. Knaś[1], W. Nowak[2]

[1]Czestochowa University of Technology
Institute of Advanced Energy Technologies
Dabrowskiego 73, 42-201 Czestochowa

[2]AGH University of Science and Technology
Al. Mickiewicza 30, 30-059 Kraków
Poland

mgandor@fluid.is.pcz.pl; kslawinski@fluid.is.pcz.pl; bbalt@fluid.is.pcz.pl; kknas@fluid.is.pcz.pl; wnowak@is.pcz.czest.pl

Key words: brown coal, drying, moisture, electromagnetic mill, process optimization.

Abstract

This paper presents an inverse problem solution in the research of brown coal drying in the electromagnetic mill. Representative samples of lignite were chosen from a brown coal mine. Lignite is classified as a low-rank coal due to its low caloric value and high moisture content, reaching approximately 40 - 50%.

Examined coal grains were in the size range of 0 - 80 mm, while the total moisture content in the feed was from 45% to 38%. The mill presented by authors is used for simultaneous grinding and drying of particulate matter. Due to the complexity of the processes occurring in the working chamber and lack of a precise mathematical model, this paper will focus on an inverse analysis throughout the process of drying brown coal. In order to maximize the process of reducing the moisture in fuel, a set of crucial parameters, such as ambient air humidity and temperature must be found. Case study of the moisture content in coal samples, before and after carrying out the drying process, shows that the microstructure of dried brown coal is highly reactive and hygroscopic. An expected reduction of total moisture of the raw lignite is set to be over 20 percentage points in relation to the input material. Analysis of the moisture content in the feed and the product was performed with a drier method at 110 °C.

The research takes into account a number of key parameters, i.e. the moisture content in the feed material, humidity of the ambient air, the temperature of the environment and the final moisture of the product. All parameters have a significant impact on the physical and chemical properties of dried coal. The drying process was carried out in a prototype of an electromagnetic mill. During the laboratory scale research, the determined objective – the reduction of the total moisture over 20 percentage points – was acquired. Furthermore, the optimal set of parameters for electromagnetic mill drying was established based on the inverse analysis.

1. Introduction

Lignite is classified as low-rank coal due to its low calorific value mainly resulting from high moisture content. Power generation in Poland is based mainly on fossil fuels. The combustion of brown coal meets roughly 34% of the electricity and heat demand [1]. Despite the intended increase of natural gas and renewable energy sources in the market, it is predicted that a significant portion of electricity production will still be based on lignite-fired power plants in decades to come [2]. Brown coal is currently an inexpensive source of energy. The energy produced from it is about half of the price of that generated from hard coal. Advantageous geological conditions, large deposits and the simplicity of mining mean that in terms of calorific value, brown coal is one of the cheapest source of energy in Poland, and it will remain so in the near future [3].

2. Drying process.

In drying process, brown coal improves its calorific value by reducing the moisture contained in it. Drying is a thermal process in which water is evaporated from solid materials or solutions. It can be seen that the moisture loss of the material takes place in stages (Figure 1). Initially, the moisture is released from the surface of the lignite particles, after that an internal water transport takes place from the centre of the particle to its surface. The driving force is heat conduction and as a consequence of the energy transfer the breakage of the hydroxyl bonds together with vapour diffusion occurs. As a result of forced convection, mixing and movement of the components take place from the surface of the particles so that the drying agent enriched in moisture is moved away from the surface and is replaced by fresh portion of drying gas. The drying rate is determined by the transport of water through the lignite particles, which eventually alters the size and the porosity of the material after drying. Water from the solid is removed during certain processes, firstly are emptied macropores, mesopores, and then capillars (50 nm in diameter), as a consequence their disintegration and crosslinking occurs. Lignites porous structure during the thermal process is partially disrupted and destroyed, which leads to reduction of the inner surface, in comparison to raw coal surface. Drying process in terms of kinetic approach assumes changes not only in the average moisture content but also the average temperature over time. Furthermore, with this data one can determine the amount of moisture evaporated from the material or the heat inside the material and the amount of energy consumed. Moreover the mass and heat transfer inside the material can be exchanged between the surface of the material and the drying medium. As a result, the moisture content changes in the material subjected to drying [4]. The drying process applies to a variety of wet substances with different physic - chemical and biochemical properties. These substances can be divided into two groups: the colloidal compounds, in the drying process the molecule changes its dimensions in capillary-porous bodies, the drying process do not change their linear dimensions. Drying speed is defined as the amount of moisture evaporated from the dried material per unit time and area of the dried surface [5]:

$$\vartheta = \frac{m \cdot \Delta X}{A \cdot \Delta t} \quad \left[\frac{kg}{s \cdot m^2}\right] \tag{1}$$

where:

- A - surface [m^2]
- t - drying time [s]
- X - absolute humidity of the material [kg/kg]
- m - mass of dry material [kg]

During the drying process the material moisture bounded mechanically and physico-chemically is removed. The drying process is characterized by [4]:

- drying curve [X = X(t)],
- drying rate curve [J = J (x)],
- temperature curve [T = T(x)].

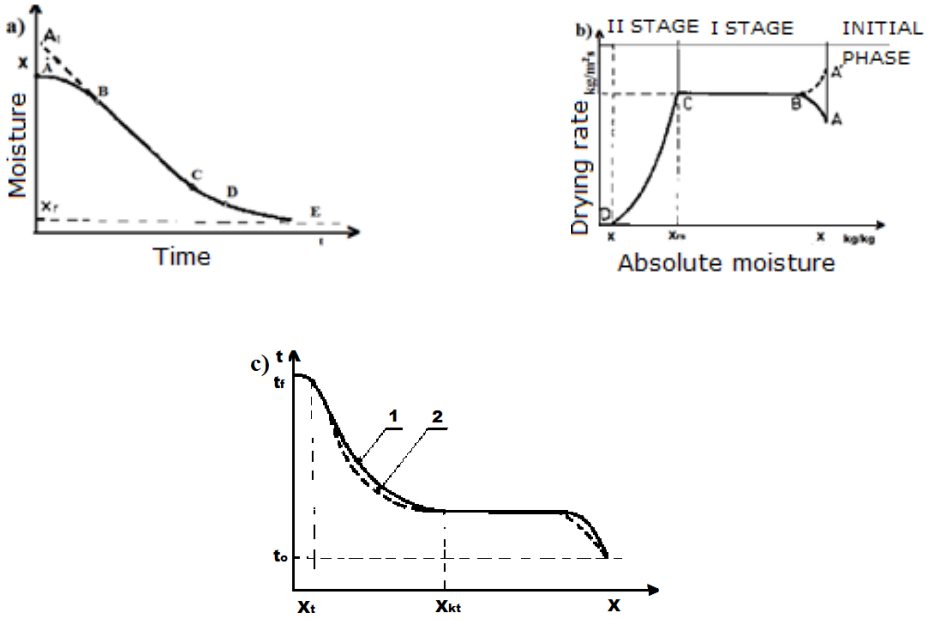


Figure 1: The graphs show curve changes characterizing the drying process: and - drying curve, b - drying rate curve, c - temperature curve [5].

One of the coal drying technologies may be the electromagnetic mill technology, which permits simultaneous drying and grinding of fuel. This method also allows to obtain the desired physical and chemical properties of the product through the appropriate control of the processes in the working chamber of the electromagnetic mill [6].

3. Electromagnetic mill and its implementation for coal drying

Electromagnetic mill, is a novel technology used for drying and grinding of the chosen material. Both processes occur simultaneously, so that there is a considerable saving of time and energy [6]. With the collision of grinding elements with the fuel in the working chamber appears a process of grinding together with the initial release of moisture. Inside the chamber, spin many small rod-shaped steel grinding elements (Fig. 2b). Furthermore, their low weight and high magnitude of the magnetic field, allows them to obtain a very high acceleration. The temperature risen by collisions of grinding elements with the material and grinding elements with each other is significant, what is shown in thermograms appearing in Figure 3 and Figure 4. Electromagnetic mill consists of a cylindrical working chamber, in which both drying and grinding appears, and the salient pole inductor powered by a three-phase current, which consists of six radially arranged poles (Fig. 2a.) that produce a rotating magnetic field. The force acting on grinding elements comes from the magnetic field and depends on the square of the induction field in that area. Operation of the electromagnetic mill does not require large energy inputs (960 kWh/day for the performance of the mill equal to 20 Mg/h [7]), since the grinding elements of small mass easily acquire a great kinetic energy.

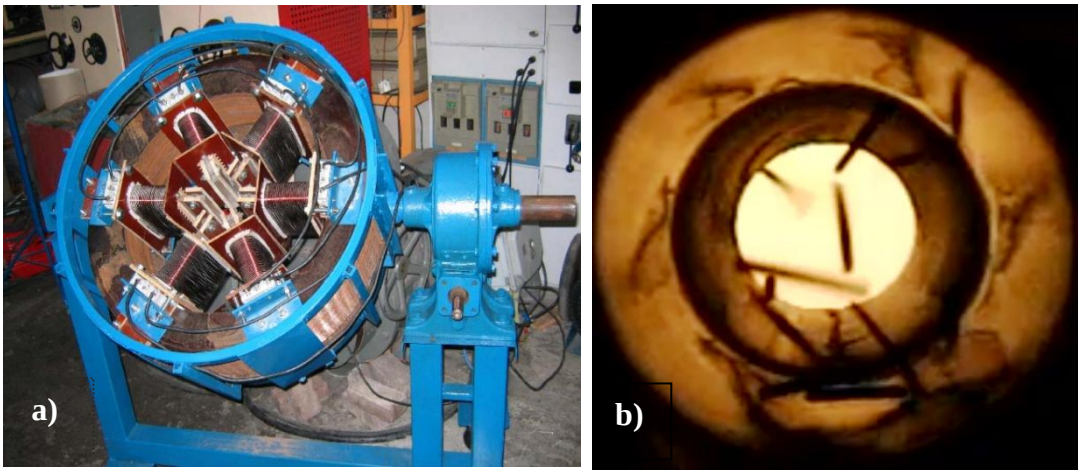


Figure 2: Electromagnetic Mill (a)
Movement of grinding elements in the working chamber (b)

The rise of the temperature of the working chamber from the collisions of the grinding elements and the eddy currents which may occur in chambers walls are shown in Fig. 3, and the temperature of the grinding elements after exiting the chamber shows Fig. 4.

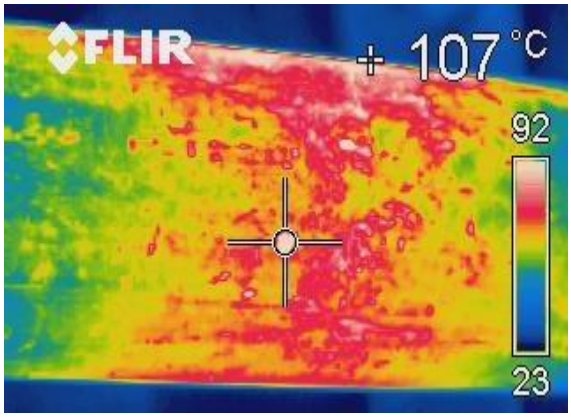


Figure 3: The thermogram of the working chamber with the participation of grinding elements.

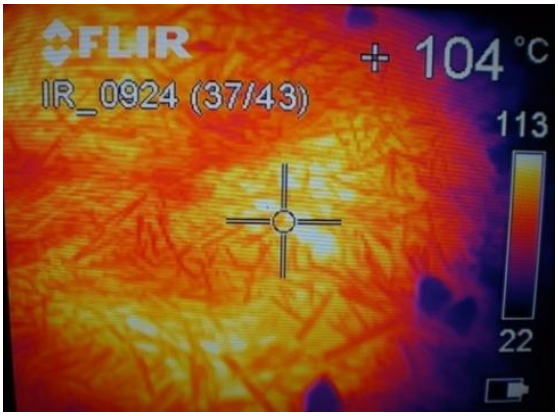


Figure 4: Thermogram of grinding elements after exiting the chamber.

4. Identification of the problem

The aim of the study was to determine the main parameters governing the reduction of moisture in the product of more than 20%. This amount is on the one hand sufficient to effectively increase the calorific value of brown coal, on the other hand, the quality gain is not associated with excessive loss of efficiency in the whole process and does not cause problems of a technological nature with using such prepared fuel [8, 9]. Because of the absence of a precise model describing the drying process occurring in the electromagnetic mill, there is no information about the effects of moisture in the environment and the effects of the drying medium temperature on the degree of drying of the feed. The amount of evaporated moisture is affected by the partial pressure of water vapor in the drying medium - determining the possibility of diffusion and mass transfer of water between the fuel particle and the environment, and the ambient temperature. It should be remembered that within the mill, the conditions are different than in its environment - on the one hand the feed loses heat by evaporation of moisture, on the other hand it obtains heat from hot grinding elements and the walls of the chamber. The goal of this research is also to determine whether it is possible to reduce the amount of water in

the feed using ambient air. In case of the lack of such opportunity or its substantial limitation it is expected to obtain outlines on the conditions necessary to provide for the effective reduction of moisture in the material.

5. The research material and methodology

In this studies, small ferromagnetic grinding elements were used. The feed material was brown coal, characterized by the parameters listed in Table 1. Moreover in all tests a constant volume of feed material was used. In Figure 5 is shown the histogram of size of lignite particles in the feed. Moisture content of the feed was in a range from 35% to 47% depending on the research environment conditions. The effect of the temperature and humidity of the environment on the drying process in the electromagnetic mill were analyzed during 5 tests described in next chapter. Tests were conducted at a constant and optimal quantity of grinding elements.

Table 1: Physico-chemical parameters characterizing the samples of brown coal.

Parameter	Value
Received basis:	
Total moisture W_t^r [%]	51,14
Ash kontent A^r [%]	10,74
Total sulphur S_t^r [%]	0,91
Lower heating value Q_i^r [kJ/kg]	8555
Analytical basis:	
Moisture W^a [%]	9,15
Ash A^a [%]	20,04
Volatiles V^a [%]	38,28
Higher heating value Q_s^a [kJ/kg]	23330
Lower heating value Q_i^a [kJ/kg]	21091
Carbon C^a [%]	43,80
Hydrogen H^a [%]	4,65
Nitrogen N^a [%]	0,53
Total sulphur S_t^a [%]	1,77
Dry basis:	
Ash A^d [%]	23,08
Total sulphur S_t^d [%]	1,82
Volatiles V^{daf} [%]	45,93
Higher heating value Q_s^{daf} [kJ/kg]	19000

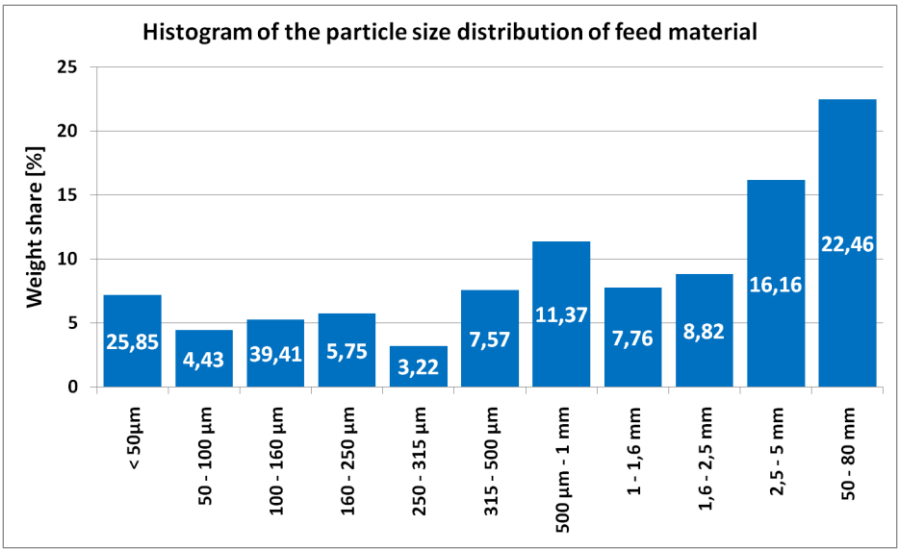


Figure 5: Histogram of grains size variety.

6. Research results

After drying process in the electromagnetic mill, changes in coal properties were revealed. Through the inverse analysis of the drying process, the influence of key parameters of the installation and their impact on the degree of drying and properties of the product were established. In this study five tests were made, with results presented in Figures 6 – 10.

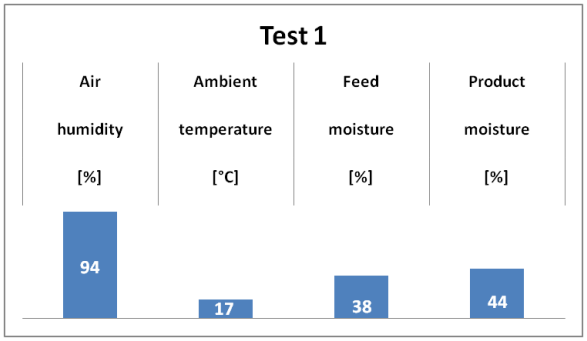


Figure 6: Ambient parameters and properties of the feed and product in test 1.

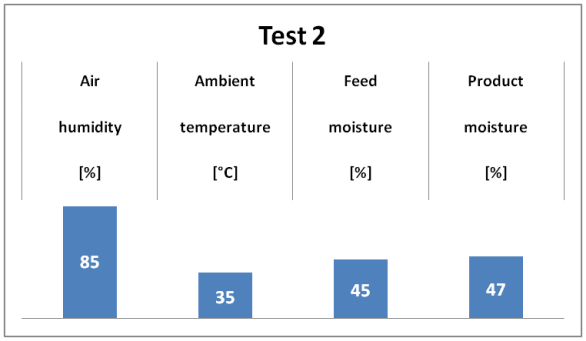


Figure 7: Ambient parameters and properties of the feed and product in test 2.

In case of performing the process with the parameters indicated in Figure 6, one could notice an increase of moisture content in the product. The main factor was a substantial content of moisture in the air, which has impeded or prevented diffusion of water from the fuel particles into the environment. Additionally, further absorption of water from the surroundings by the product has increased the amount of water in the product relative to the feed, which shows a significant increase in surface area of the material after passing the working area and the increase of the reactivity of the material as a result of hygroscopic properties and its activation. It can be concluded that the electromagnetic mill significantly changes physical properties of the material in terms of releasing and absorbing moisture. Another test (Fig. 7) shows that despite a significant increase in the ambient temperature (from 17 °C to 35 °C), it is impossible to dry the material with ambient air humidity exceeding 80%. It must be noted that in both cases, despite the substantial difference in moisture content of the feed (38% and 45%), drying resulted in a similar final moisture (44% and 47%) of the product. This can therefore be inferred to a potential problem appearing while drying in the electromagnetic mill, which was increasing hygroscopic properties in the treated material. It is important for the design of installation and forces the need to isolate the product from the environment when the conditions are unfavorable (high humidity). Further, tests were carried out at a lower humidity.

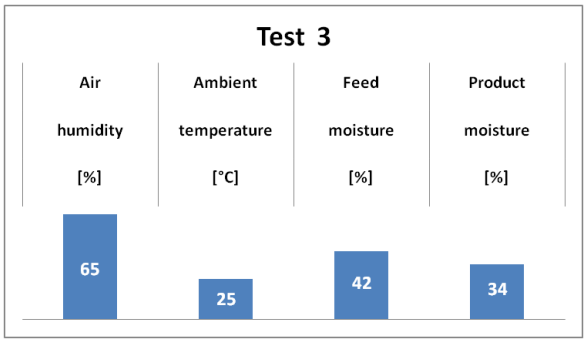


Figure 8: Ambient parameters and properties of the feed and product in test 3.

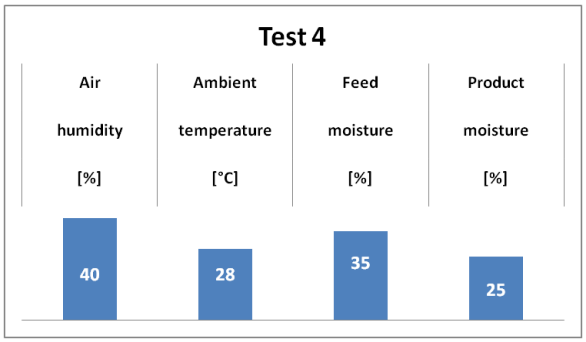


Figure 9: Ambient parameters and properties of the feed and product in test 4.

Tests 3 and 4 were carried out at a similar temperature of air (25 °C and 28 °C) and lower ambient humidity than in previous tests. In both cases, one can observe a reduction in the moisture of the dried lignite by 8% and 10%. A clear relationship can be seen between the obtained degree of moisture reduction and moisture content in the drying medium, while this relationship correlates poorly with the ambient temperature, which may be regarded as the parameter of less importance. From this

observation it was decided to conduct a test in the conditions of low temperature (-15 °C) which provides very low air humidity (18%).

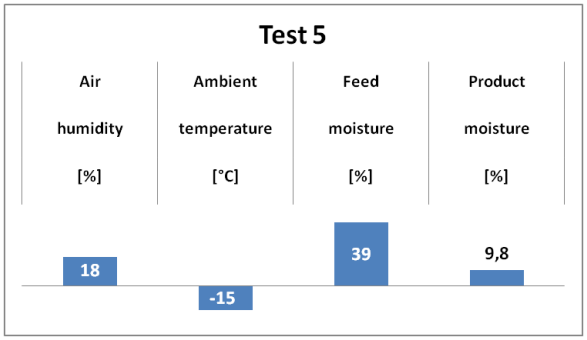


Figure 10: Ambient parameters and properties of the feed and product in test 5.

High moisture reduction (more than 29 percentage points) is ensured by an increased rate of diffusion of moisture from the fuel to low-humidity air. Moreover, such low air moisture content prevents or significantly reduces the re-absorption of water by coal. As stated earlier the temperature of the drying medium is of less importance if the product is isolated from the humid air immediately after the process, or under sealed conditions shall be fed directly into the combustion zone. The results demonstrate the feasibility of simultaneous grinding and drying with the use of an electromagnetic mill, but particular attention should be paid to the collection of the dried material and the selection of the drying medium.

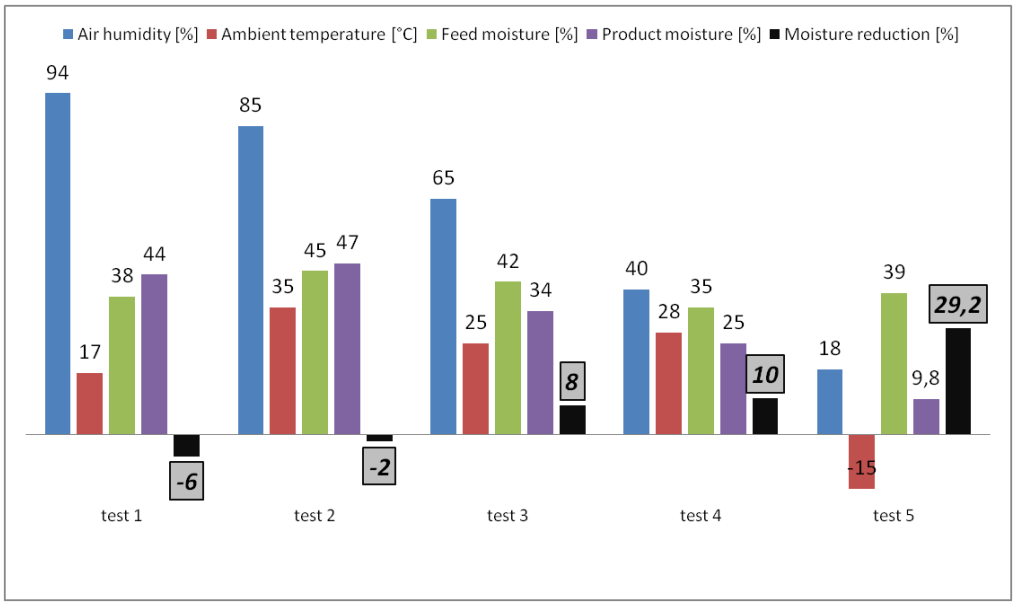


Figure 11: Summary chart of the results (negative values of reduction mean an increase in moisture content, positive values – a reduction of moisture content).

7. Conclusion

This studies have shown the ability to improve the calorific value of lignite by simultaneous grinding and drying in an electromagnetic mill installation. In practice, it is possible to reduce moisture by more than 20 percentage points using dry gas medium. Satisfactory results were obtained even for negative temperatures, which demonstrates the secondary importance of this parameter in the drying process . In the case of using a hot drying agent, the obtained results would be even more advantageous. As a result of the inverse analysis, a significant difference in product properties were determined. The product was hygroscopic and reactive. Large concentrations of highly reactive coal dust can induce an explosion hazard - especially when using hot drying medium with significant content of oxygen. In practice, for safety reasons, drying of brown coal in hot air should be eliminated. A preferred drying

media should be incapable of initiating an explosion due to lack of oxygen, which can be ensured with such gases as carbon dioxide or nitrogen [10]. An alternative is to use a gas mixture low in oxygen, for example – dehumidified flue gases from combustion for energy purposes. Low oxygen content (in the limits of 6 %) reduces the risk of an explosion and high exhaust gas temperature can be effectively used for drying while reducing exergy losses in process and making the valorization of lignite cost-effective. Due to the inclination of the working chamber at an angle relative to the plane of the surface which allows for the gravity transport of feed, a drying medium in a continuous process should be introduced concurrently above the workspace – capturing the released moisture and preventing the absorption of water by the material above the drying area. Particularly important for the dried and activated lignite, is the way of material collecting and storage. It is required to isolate the product from adverse environmental conditions. For this purpose, an additional module of an evaporation chamber was proposed, with the vibrating transporter at the outlet of the mill. In addition, due to changes in products structure, it can be assumed that further release of moisture is distributed in time and takes place outside the mill. Evaporation chamber should have a sufficient length to allow for evaporation of the remaining moisture to the environment and reaching an equilibrium state between the fuel particles and the drying medium. The use of vibrating transporter due to the additional excitation of the particles speeds up this process, and one can reduce the length of the chamber needed to achieve the desired properties of the product. To maintain the low moisture content of the product in the environment it is also necessary to install the fans in the ceiling of the chamber to separate the moisture generated by the process from the product. Given these guidelines the authors have proposed a conceptual installation for drying with the use of electromagnetic mill shown in Figure 12.

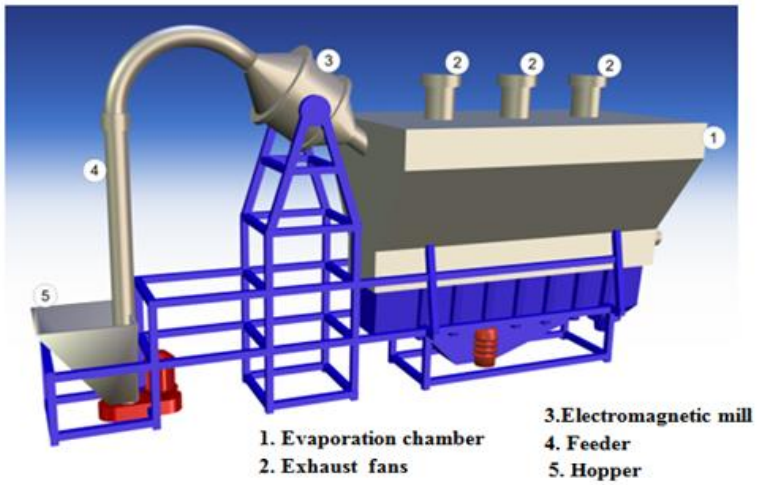


Figure 12: Conceptual facility model for drying and grinding
with the application of electromagnetic mill.

Acknowledgements

This research project is obtained from work co-financed by the National Centre of Research and Development in the framework of Contract SP/E/1/67484/10 – Strategic Research Programme – Advanced technologies for energy generation : Development of a technology for highly efficient zero emission coal-fired power units integrated with CO_2 capture.

The authors acknowledge with gratitude helpful comments from prof. zw. dr hab. inż. Jacek Leszczyński.

References

[1]A. Pietraszewski, Polskie górnictwo węgla brunatnego w pierwszym półroczu 2013 roku. Węgiel Brunatny, 3/84, 2013, PPWB 2013.

[2] Z. Kasztelewicz, *Węgiel brunatny optymalnym paliwem dla polskiej energetyki w I połowie XXI wieku*, 4 lutego 2010, Warszawa.

[3] J. R. Kasiński, *Zasoby węgla brunatnego w Polsce – stan rozpoznania i podstawowe problemy*, Państwowy Instytut Geologiczny, www.pgi.gov.pl, 2014.

[4] C. Strumiło, *Podstawy teorii i techniki suszenia*, WNT, Warszawa, 1983.

[5] B.Staniszewski, *Wymiana ciepła podstawy teoretyczne*. PWN. Warszawa 1983.

[6] R. Sosiński, P. Szymanek, W. Nowak, *Aspekty wykorzystania młyna elektromagnetycznego do aktywowania popiołów lotnych*, Międzynarodowa Konferencja Euro Coal Ash (2008), pp. 175-187.

[7] R. Sosiński, W. Nowak, J. Szymańska, *Problematyka stosowania młynów elektromagnetycznych w wysokowydajnych procesach zagospodarowania popiołów lotnych*, Międzynarodowa Konferencja Euro Coal Ash (2008), pp. 299.

[8] M. Michalski, *Wpływ wstępnego suszenia Węgla brunatnego na pracę kotłów energetycznych typu BB-1150 i CFB-672*, Materiały konferencyjne, Rynek Energii Cieplnej, 2012.

[9] J. Kotowicz, A. Dryjańska, A. Balicki, *Wpływ wybranych parametrów na sprawność kotła CFB typu oxy*, Rynek Energii Nr 2(99), 2012, pp. 120 – 126.

[10] Ł. Bartela, J. Kotowicz, *Analiza wykorzystania azotu jako medium suszącego węgiel spalany w kotle oxy*, Rynek Energii, Nr 2(93), 2011, pp. 49 – 55.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Variational Bayesian inversion for microwave imaging applied to breast cancer detection

L. Gharsalli[1], **H. Ayasso**[2], **B. Duchêne**[1], **A. Mohammad-Djafari**[1]

[1] Laboratoire des Signaux et Systèmes (L2S), (UMR8506: CNRS-SUPELEC-Univ Paris-Sud),
3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France
[2] GIPSA-LAB, Département Image Signal, (CNRS-Univ Grenoble),
BP 46 - 38402, Saint Martin d'Hères, France
e-mail: leila.gharsalli@lss.supelec.fr

Key words: Inverse scattering, microwave imaging, breast cancer detection, Gauss-Markov-Potts prior, Variational Bayesian Approximation

Abstract

In this work, microwave imaging is considered as a nonlinear inverse scattering problem and tackled within a Bayesian estimation framework. The object under test (breast affected by a tumor) is supposed to be composed of compact regions made of a restricted number of different homogeneous materials. This *a priori* knowledge is appropriately translated by a Gauss-Markov-Potts prior. First, we express the *a posteriori* probability laws of all the unknowns and then the Variational Bayesian Approximation (VBA) used to compute the posterior estimators and reconstruct both permittivity and conductivity maps. This approximation consists in the best separable probability law that approximates the true posterior probability law in the Kullback-Leibler sense. This leads to an implicit parametric optimization scheme which is solved iteratively. Some preliminary results, obtained by applying the proposed method to synthetic data, are presented and compared to those obtained by means of the classical contrast source inversion method.

1 Introduction

In the last few decades, microwave scattered imaging has received an increasing interest for medical applications such as breast cancer detection [1]. In addition to the non-ionizing nature of microwaves, one of the motivation for developing a microwave imaging technique for detecting breast cancer is the significant contrast that exists at microwave frequencies between the dielectric properties of normal and malignant breast tissues. All this makes microwave imaging a better alternative, in terms of cost and harmlessness, than X-ray mammography which is the most current breast cancer detection technique.

Hence, measurements of the scattered fields resulting from the interaction between a known interrogating wave in the microwave frequency range and the breast can be used to retrieve a contrast function representative of the dielectric properties (permittivity and conductivity) of the latter. This leads to a non linear ill-posed inverse scattering problem solved, herein, in a variational Bayesian framework. The associated forward problem consists in modeling the wave-breast interaction through a domain integral representation of the electric field in a 2-D configuration in a transverse magnetic polarization case.

The Bayesian framework allows us to take easily into account *a priori* information on the sought solution. Herein, we would like to account for the fact that the breast is composed of a finite number of different tissues distributed in compact regions, meaning that the sought image is composed of a finite number of homogeneous area. This *a priori* is introduced via a Gauss-Markov fields with hidden Potts label fields [2]. Afterwards, the variational Bayesian approximation (VBA) [3] is applied to obtain an estimator of the posterior law. It can be noted that a semi-supervised context is considered herein where the number of different tissues is supposed to be known, while all the unknowns and hyper-parameters of the model are estimated simultaneously through a joint posterior law. The purpose of VBA is to approximate the latter by a free form distribution that minimizes the Kullback-Leibler divergence. This distribution is chosen as a separable law. Then, thanks to the latter method, the initial inverse problem turns into an optimization problem and an analytical approximation of the posterior is obtained. Its use in microwave imaging and optical diffraction tomography has already been treated and results have shown its performances with respect to the computation time and simplicity, compared to other methods such as Monte-Carlo Markov Chain (MCMC) [4, 5].

The main contribution of this work is the application of VBA to breast imaging where the sought contrast is complex valued, contrarily to the case treated in [5], and both permittivity and conductivity maps have to be retrieved. Herein, we discuss the results obtained by means of this approach from synthetic data generated in different configurations involving two different numerical breast phantoms: a simple model made of two homogeneous media and a more sophisticated one built up from a MRI scan of a real breast. Then we present results compared to those obtained by means of the deterministic contrast source inversion method (CSI, [13]).

The paper is organized as follows: section 2 is about the experimental configuration and the forward modeling. The VBA approach and Bayesian computations are discussed in section 3. In section 4, the method is applied to synthetic data and is compared to CSI. Finally, some conclusions and perspectives are given in section 5.

2 The forward modelling

2.1 The experimental configuration

We consider a 2-D configuration in a transverse magnetic polarization case where the object under test is supposed to be cylindrical, of infinite extension along the z axis and illuminated by a line source whose location can be varied and that operates at several discrete frequencies. This source generates an incident electric field E^{inc} polarized along the z axis with an $\exp(-i\omega t)$ implicit time dependence and illuminates the breast from 64 various angular positions uniformly distributed around a 7.5-cm-radius circle centered at the origin and at 6 different frequencies in the band 0.5 - 3 GHz. For each frequency and illumination angle, 64 measurements of the scattered field are performed at angular positions uniformly distributed around the same circle. The breast (domain $\mathcal{D}_2$) is immersed in a background medium (domain $\mathcal{D}_1$) and is

supposed to be contained in a test domain ($\mathcal{D}$). The different media are characterized by their propagation constant $k(\boldsymbol{r})$ such that $k(\boldsymbol{r})^2 = \omega^2\epsilon_0\epsilon_r(\boldsymbol{r})\mu_0 + i\omega\mu_0\sigma(\boldsymbol{r})$, where ω is the angular frequency, ϵ_0 and μ_0 are the permittivity and the permeability of free space, respectively, $\boldsymbol{r} \in \mathcal{D}$ is an observation point and $\epsilon_r(\boldsymbol{r})$ and $\sigma(\boldsymbol{r})$ are the relative permittivity and conductivity of the medium.

Two models of breast are considered herein (Figure 1). Both of them are supposed to be affected by a tumor (domain $\mathcal{D}_3$) with a 2-cm-diameter circular cross-section and of electromagnetic parameters $\epsilon_r = 55.3$ and $\sigma = 1.57$ Sm^{-1}. The first model (Model-1) is rather simple: it consists in an homogeneous breast, with a 9.6-cm-diameter circular cross-section and relative dielectric permittivity and conductivity respectively equal to $\epsilon_r = 6.12$ and $\sigma = 0.11$ Sm^{-1}, immersed in a background medium of electromagnetic parameters $\epsilon_r = 10$ and $\sigma = 0.5$ Sm^{-1}. The second model (Model-2) is more sophisticated. It is built up from a MRI scan of a real breast. Hence, the breast is also supposed to be of circular cross-section with a diameter of 9.2 cm but it is now made of a very heterogeneous medium with parameters varying in the ranges $2.46 \leq \epsilon_r \leq 60.6$ and 0.01 Sm$^{-1} \leq \sigma \leq 2.28$ Sm^{-1}, surrounded by a skin with electromagnetic parameters $\epsilon_r = 35.7$ and $\sigma = 0.32$ Sm^{-1} and immersed in a background medium whose relative dielectric permittivity and conductivity are respectively equal to $\epsilon_r = 35$ and $\sigma = 0.5$ Sm^{-1}.

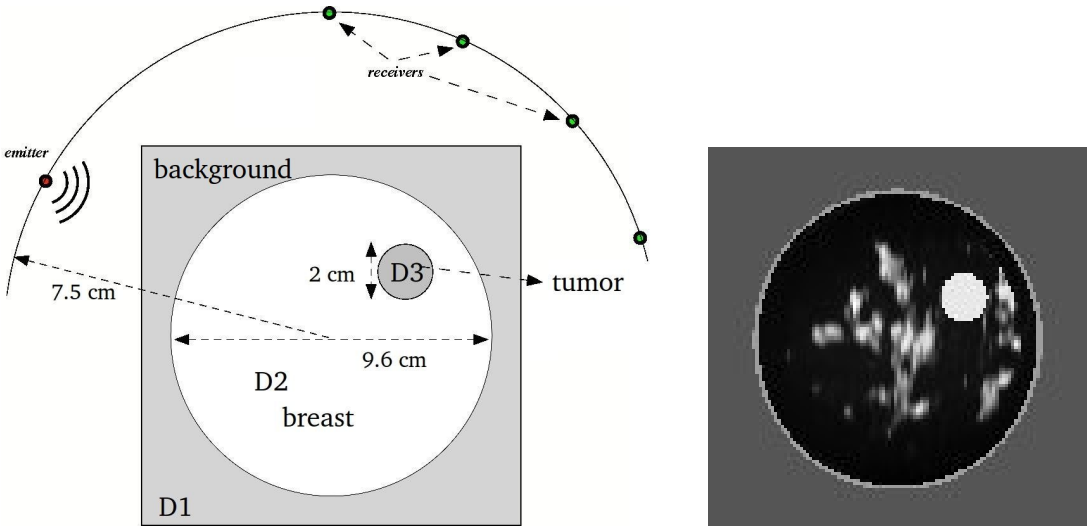


Figure 1: The measurement configuration and the two models of breast (left: Model-1, right: Model-2).

2.2 The problem formulation

The modeling is based upon domain integral representations obtained by applying Green's theorem to the Helmholtz wave equations satisfied by the fields and by accounting for continuity and radiation conditions [8]. The forward model is described by two coupled integral equations. The first one, denoted as the observation equation, is a Fredholm first kind integral equation that relates the scattered field to Huygens-type sources induced within the target by the incident wave, whereas the second one, denoted as the coupling (or state) equation, relates the total field to the induced sources [9, 4]. The forward problem is solved from discrete counterparts of these integral equations obtained by means of a method of moments whith pulse basis and point matching [10], which results in partitioning the test domain $\mathcal{D}$ into $N_{\mathcal{D}}$ elementary pixels small enough to permit considering both the field and the contrast as constant

over each of them. Let us now consider a contrast function χ, defined in $\mathcal{D}$ and null outside the object, such that $\chi(\boldsymbol{r}) = (k(\boldsymbol{r})^2 - k_1^2)/k_1^2$, where k_1 is the propagation constant of the background medium, and define the Huygens-type sources $w(\boldsymbol{r})$ as $w(\boldsymbol{r}) = \chi(\boldsymbol{r})E(\boldsymbol{r})$, where $E(\boldsymbol{r})$ is the total field in the target. The above-mentioned discrete counterparts then read:

$$\boldsymbol{y} = \boldsymbol{G}^o \boldsymbol{w} + \boldsymbol{\epsilon} \tag{1}$$

$$\boldsymbol{w} = \boldsymbol{X}\, \boldsymbol{E}^{inc} + \boldsymbol{X}\, \boldsymbol{G}^c \boldsymbol{w} + \boldsymbol{\xi}, \tag{2}$$

where $\boldsymbol{X} = \mathrm{diag}(\boldsymbol{\chi})$, $\boldsymbol{E}$, $\boldsymbol{\chi}$ and $\boldsymbol{w}$ are vectors that contain the values of $E(\boldsymbol{r}')$, $\chi(\boldsymbol{r}')$ and $w(\boldsymbol{r}')$ at the centers $\boldsymbol{r}'$ of the pixels ($\boldsymbol{r}' \in \mathcal{D}$), $\boldsymbol{y}$ is the vector containing the values of the scattered field $y(\boldsymbol{r})$ at the measurement points $\boldsymbol{r}$, $\boldsymbol{G}^o$ and $\boldsymbol{G}^c$ are huge matrices whose elements result from the integration of the Green's function over the elementary pixels [5] and $\boldsymbol{\epsilon}$ and $\boldsymbol{\xi}$ are two variables that account for the model and measurement errors and that are supposed to be centered and white and to satisfy Gaussian laws (i.e., $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{\epsilon}|\boldsymbol{0}, v_\epsilon \mathbf{I})$ and $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\xi}|\boldsymbol{0}, v_\xi \mathbf{I})$).

Now, the forward problem consists in first solving equation (2) for the induced sources $\boldsymbol{w}$, knowing the contrast $\boldsymbol{\chi}$ and the incident field $\boldsymbol{E}^{inc}$, and then solving equation (1) for the scattered field $\boldsymbol{y}$. At this point it can be noted that the synthetic data of the inverse problem are generated in this way. However, in the case of Model-1, in order to avoid committing an inverse crime which would consist in testing the inversion algorithm on data obtained by means of a model closely related to that used in the inversion, the data are computed by benefiting from the circular symmetry that exists in the absence of the tumor. Hence, the data are computed by means of a model (the data model) where only the domain $\mathcal{D}_3$ occupied by the tumor is discretized, whereas the breast and the background medium are considered as a cylindrically stratified embedding medium and the Green's function is modified consequently. Figure 2 displays the scattered fields obtained by means of the data model on configuration Model-1 for an illumination angle of $45°$ and at two operating frequencies: 1.5 GHz and 3 GHz, compared to that obtained by means of the forward model used for inversion where the test domain $\mathcal{D}$ is a 12.16 cm sided square partitioned into 64×64 square pixels with side $\delta = 1.9$ mm. It can be observed that the results fit relatively well.

3 Bayesian inversion approach

3.1 Hierarchical prior model

The inverse problem consists in retrieving the unknown contrast $\boldsymbol{\chi}$, or more precisely the relative permittivity ϵ_r and the conductivity σ, from the scattered field $\boldsymbol{y}$, given the incident field $\boldsymbol{E}^{inc}$. It can be noted that the induced sources $\boldsymbol{w}$ being also unknown, they must be retrieved at the same time as $\boldsymbol{\chi}$. Hence, assuming that their relation to the contrast is given by the state equation (2), we define their *a priori* probability law as:

$$p(\boldsymbol{w}|\boldsymbol{\chi}) = \exp\left\{-\frac{1}{2v_\epsilon}||\boldsymbol{w} - \boldsymbol{X}\boldsymbol{E}^{inc} - \boldsymbol{X}\boldsymbol{G}^c\boldsymbol{w}||\right\}. \tag{3}$$

Now, let us introduce *a priori* information on the sought solution required in order to counteract the ill-posedness of the inverse problem. It consists in the fact that the sought object is composed of a restricted number K of homogeneous materials distributed in compact regions. This prior information is introduced by means of a hidden variable (or classification label) $z(\boldsymbol{r})$ associated with each pixel $\boldsymbol{r}$, which represents a segmentation of the unknown object. The finite number of homogeneous materials can then be accounted for through the following conditional distribution:

$$p(\chi(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_k, v_k),\ k = 1, \ldots, K, \tag{4}$$

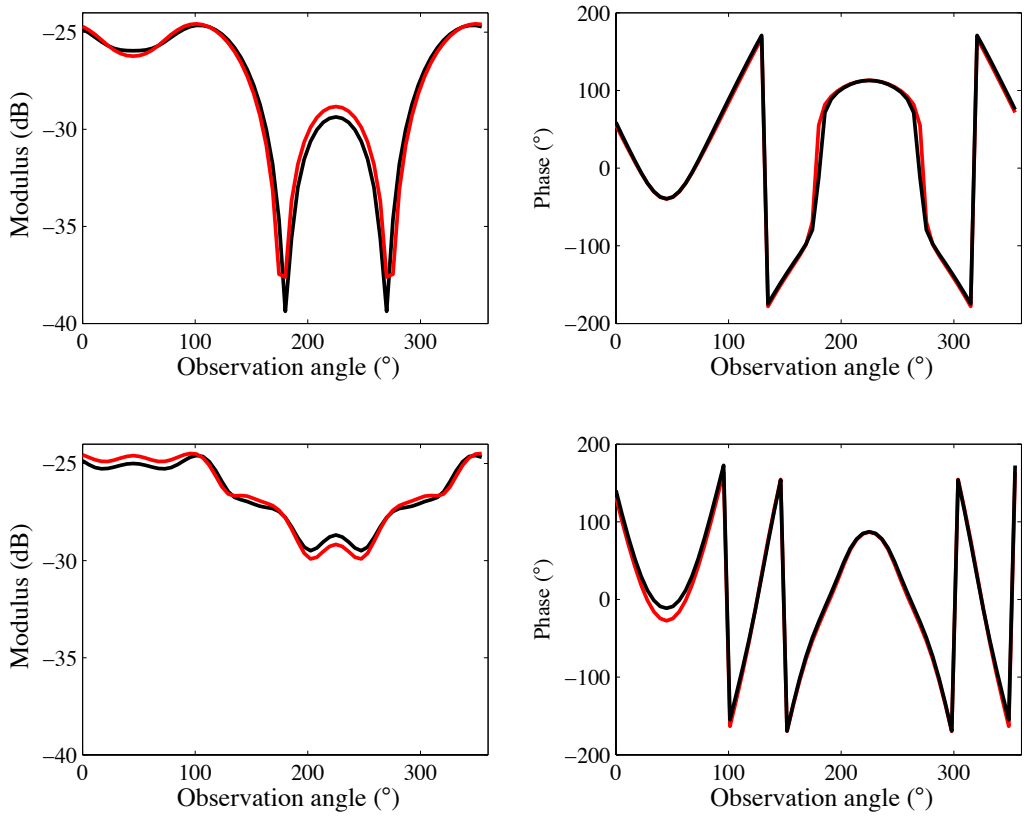


Figure 2: Amplitude (left) and phase (right) of the scattered fields computed by means of the data model (black) and by means of the forward model used for inversion (red) at 1.5 GHz (up) and 3 GHz (down).

which means that all the pixels with the same label ($z(\boldsymbol{r}) = k,\ k \in \{1, ..., K\}$) correspond to the same material with a mean value m_k and a variance v_k.

The compactness of the different regions can be accounted for by relating, in a probabilistic way, the classification label $z(\boldsymbol{r})$ of a pixel $\boldsymbol{r}$ to that of its neighbors. This is done *via* a Potts-Markov model on $\boldsymbol{z}$:

$$p(\boldsymbol{z}|\lambda) = \frac{1}{T(\lambda)} \exp\left\{\lambda \sum_{\boldsymbol{r}\in\mathcal{D}} \sum_{\boldsymbol{r}'\in V_{\boldsymbol{r}}} \delta\left(z(\boldsymbol{r}) - z(\boldsymbol{r}')\right)\right\}, \tag{5}$$

where λ is a parameter that determines the correlation between neighbors (herein $\lambda = 1$), $T(\lambda)$ is a normalization factor and $V_{\boldsymbol{r}}$ is a neighborhood of $\boldsymbol{r}$, herein made of the four nearest pixels.

Now we have all the components necessary to find the expression of the joint posterior law of all the unknowns $(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi})$ with $\boldsymbol{\psi} = \{\boldsymbol{m}, \boldsymbol{v}, v_\epsilon, v_\xi\}$. The latter is obtained by applying the Bayes formula:

$$\begin{aligned} p\left(\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}|\boldsymbol{y}\right) \quad &\propto \quad p\left(\boldsymbol{y}|\boldsymbol{w}, v_\epsilon\right) p\left(\boldsymbol{w}|\boldsymbol{\chi}, v_\xi\right) p\left(\boldsymbol{\chi}|\boldsymbol{z}, \boldsymbol{m}, \boldsymbol{v}\right) \\ &\times \quad p\left(\boldsymbol{z}|\lambda\right) p\left(m_k|\mu_0, \tau_0\right) p\left(v_k|\eta_0, \phi_0\right) \\ &\times \quad p\left(v_\epsilon|\eta_\epsilon, \phi_\epsilon\right) p\left(v_\xi|\eta_\xi, \phi_\xi\right). \end{aligned} \tag{6}$$

Expressions of $p(\boldsymbol{y}|\boldsymbol{w}, v_\epsilon)$, $p(\boldsymbol{w}|\boldsymbol{\chi}, v_\xi)$ and $p(\boldsymbol{z}|\lambda)$ are derived respectively from equations (1), (3) and (5), whereas conjugate priors laws are assigned to the hyper-parameters:

$$\begin{aligned} p(m_k) &= \mathcal{N}(m_k|\mu_0, \tau_0), \quad p(v_k) = \mathcal{IG}(v_k|\eta_0, \phi_0) \\ p(v_\epsilon) &= \mathcal{IG}(v_\epsilon|\eta_\epsilon, \phi_\epsilon), \quad p(v_\xi) = \mathcal{IG}(v_\xi|\eta_\xi, \phi_\xi), \end{aligned} \tag{7}$$

where $\mathcal{N}(m|\mu, \tau)$ and $\mathcal{IG}(v|\eta, \phi)$ stand for Gaussian and inverse-gamma distributions, respectively, and μ_0, τ_0, η_0, ϕ_0, η_ϵ, ϕ_ϵ, η_ξ and ϕ_ξ are meta-hyper-parameters appropriately set to have non-informative priors, i.e. flat prior distributions.

3.2 The Variational Bayesian Approach

All the right-hand side expressions of equation (6) are known, which allows us to obtain the left hand side, i.e. the joint posterior law of all the unknowns, up to a normalizing constant. However, the complexity of its expression makes it very hard to obtain in a tractable form for conventional estimators, such as JMAP or PM, and an approximation is then required. Hence, we opt for an analytical approximation based upon the Variational Bayesian Approach (VBA, [11]) which aims in approximating the true posterior distribution (6) by a simpler separable law $q(\boldsymbol{u}) = \prod_i q(\boldsymbol{x}_i)$ with $\boldsymbol{u} = \{\boldsymbol{\chi}, \boldsymbol{w}, \boldsymbol{z}, \boldsymbol{\psi}\}$, that minimizes the Kullback-Leibler divergence $\mathrm{KL}(q||p) = \int q \ln(q/p)$. We define the separable law as:

$$q(\boldsymbol{u}) = q(v_\epsilon)q(v_\xi) \prod_i q(\chi_i)q(w_i)q(z_i) \prod_k q(m_k)q(v_k). \tag{8}$$

Then we look for the optimal form of q that minimizes the Kullback divergence. This leads to the following parametric distributions:

$$\begin{aligned} q(\boldsymbol{w}) = \mathcal{N}(\widetilde{\boldsymbol{m}}_w, \widetilde{\boldsymbol{V}}_w), \quad q(\boldsymbol{\chi}) = \mathcal{N}(\widetilde{\boldsymbol{m}}_\chi, \widetilde{\boldsymbol{V}}_\chi), \quad q(\boldsymbol{z}) = \textstyle\prod_{\boldsymbol{r}} \widetilde{\boldsymbol{\zeta}}_k(\boldsymbol{r}) \\ q(m_k) = \mathcal{N}(\widetilde{\mu}_k, \widetilde{\tau}_k), \quad q(v_k) = \mathcal{IG}(\widetilde{\eta}_k, \widetilde{\phi}_k) \\ q(v_\epsilon) = \mathcal{IG}(\widetilde{\eta}_\epsilon, \widetilde{\phi}_\epsilon), \quad q(v_\xi) = \mathcal{IG}(\widetilde{\eta}_\xi, \widetilde{\phi}_\xi), \end{aligned} \tag{9}$$

where the tilded variables are mutually dependent and are computed in an iterative way [5].

3.3 Initialization and convergence of the algorithm

The initial number of classes K used for segmentation, it is set to $K = 3$ for Model-1 and $K = 8$ for Model-2, whereas the initial values of the unknowns $\boldsymbol{\chi}^{(0)}$ and $\boldsymbol{w}^{(0)}$ are obtained by backpropagating the scattered field data from the measurement circle onto the test domain $\mathcal{D}$. From $\boldsymbol{\chi}^{(0)}$ and $\boldsymbol{w}^{(0)}$, the classification $\boldsymbol{z}$ and the hyper-parameters (means and variances) can be initialized by means of *K-means* clustering [12]. Here, given the fact that the contrast is complex valued, first the real part is segmented and, then, the same segmentation is used to initialize the imaginary part. Concerning the convergence, the shaping parameters of the equation (9) are iterated until the convergence is reached. The latter is estimated empirically by looking to the evolution of contrast and hyper-parameters in the course of iterations (figure (3)).

4 Results

Figure (4) displays the results obtained for the conductivity of Model-1 after 500 iterations. These results are compared to those obtained, after the same number of iterations, by means of the contrast source

inversion method (CSI, [7]), an iterative deterministic method which consists in minimizing a cost functional, that accounts for both observation and coupling equations, by alternately updating $\boldsymbol{w}$ and $\boldsymbol{\chi}$ with a gradient-based method. In general, the algorithm succeeds in retrieving homogeneous regions that correspond to the background, the breast and the tumor. The results are, however, more accurate than those

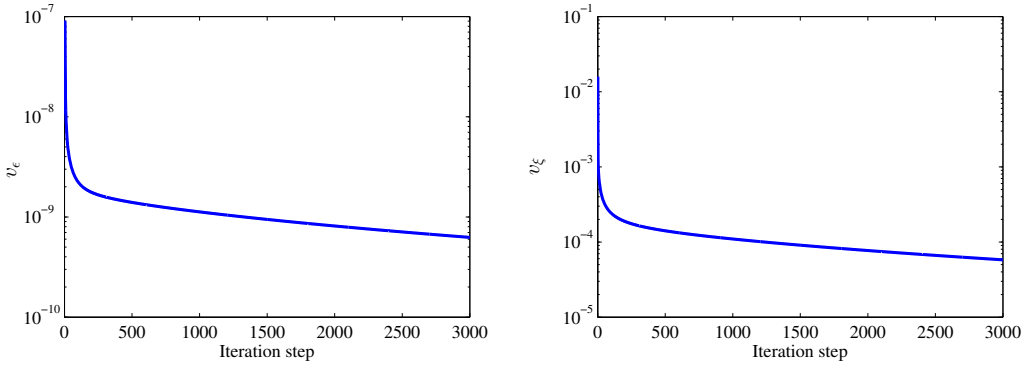


Figure 3: The evolution of the observation v_ϵ (left-) and coupling v_ξ (right) error variances as functions of the iteration step for Model-1.

obtained by means of CSI. This is confirmed by the profiles reconstructed with both methods along an horizontal line crossing the center of the tumor (figure 4 - (d)). Besides, the interest of VBA is that, in addition to an image of the sought object, it yields also an estimate of the hyper-parameters, the segmentation and the variances of the estimators.

Figure (5) displays the results obtained on Model-2, with a test domain $\mathcal{D}$ partitioned into 120×120 1-mm-side square pixels, after 2000 iterations. This figure highlights the observations made on Model-1 about the effectiveness of the proposed VBA technique. The glandular areas in the breast are not exactly found, which is normal as they do not correspond to compact homogeneous regions, but, on the contrary, the tumor is more apparent with VBA than with CSI, particularly in the conductivity map.

5 Conclusion

We consider microwave imaging as an inverse obstacle scattering problem which is known to be ill-posed. This means that a regularization of the problem is required prior to its resolution, and this regularization generally consists in introducing *a priori* information on the sought solution. Herein, an important knowledge about the object under test is that it is composed of a restricted number of homogeneous materials distributed in compact regions. This is tackled in a Bayesian inversion framework via a Gauss-Markov-Potts model. Application to synthetic data shows a good improvement in the reconstruction quality compared to a CSI deterministic approach. However, several points still need to be investigated. Particularly, concerning the convergence fastness, a gradient-like variational Bayesian method [14] is under investigation. This technique has already shown its effectiveness compared to the classical VBA, especially in terms of convergence fastness, in other applications. Furthermore, the use of Gauss-Markov mixtures [4] as prior model for the unknown object may be more adapted, especially when the latter is very heterogeneous, since with this model the links between pixels of the same class are strengthened by means of a Markov field, while the independence between pixels of different classes is kept in order to preserve the contours. An other point is under investigation concerning the number

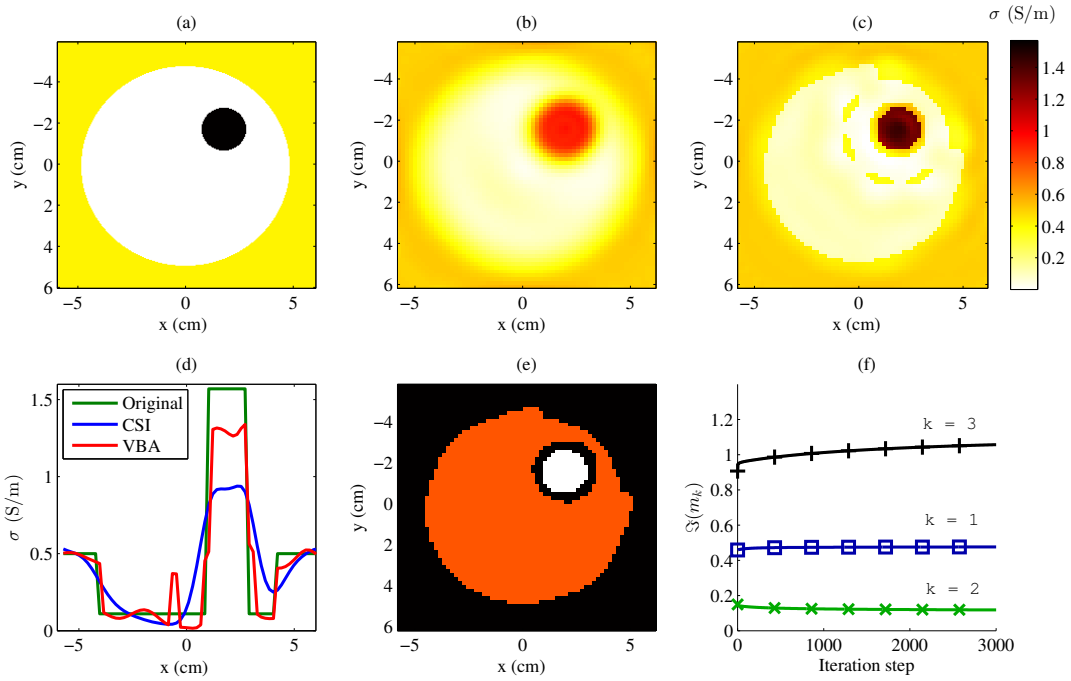


Figure 4: The results obtained for Model-1: maps of conductivity (first row) reconstructed by means of CSI (b) and VBA (c) compared to the real object (a) and (second row) the conductivity profiles reconstructed by means of CSI and VBA compared to the real profile (d), the hidden field (e) and the mean m_k of the imaginary part of the contrast (f) for the 3 classes.

of classes: the latter should converge to an optimal number, if initialized with a higher number, because, during iterations, the number of pixels that belong to some of the classes decreases, which means that the latter should disappear. Alternatively, the estimation of the number of classes, if the latter is unknown, can be tackled by means of a non-parametric approach.

References

[1] T. Henriksson, *Contribution to quantitative microwave imaging techniques for biomedical applications*, Ph.D. thesis, Université Paris-Sud 11 - Mälardalen University, (2009).

[2] A. Mohammad-Djafari, *Gauss-Markov-Potts priors for images in computer tomography resulting to joint optimal reconstruction and segmentation*, International Journal of Tomography and Statistics, 11: W09, (2008), pp. 76–92.

[3] D.J.C. Mackay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, (2003).

[4] H. Ayasso, *Une approche bayésienne de l'inversion. Application à l'imagerie de diffraction dans les domaines micro-onde et optique*, Ph.D. thesis, Université Paris sud 11, (2010).

[5] H. Ayasso, B. Duchêne, and A. Mohammad-Djafari, *Optical diffraction tomography within a variational Bayesian framework*, Inverse Problems in Science and Engineering, 20, no. 1, (2012), pp. 59–73.

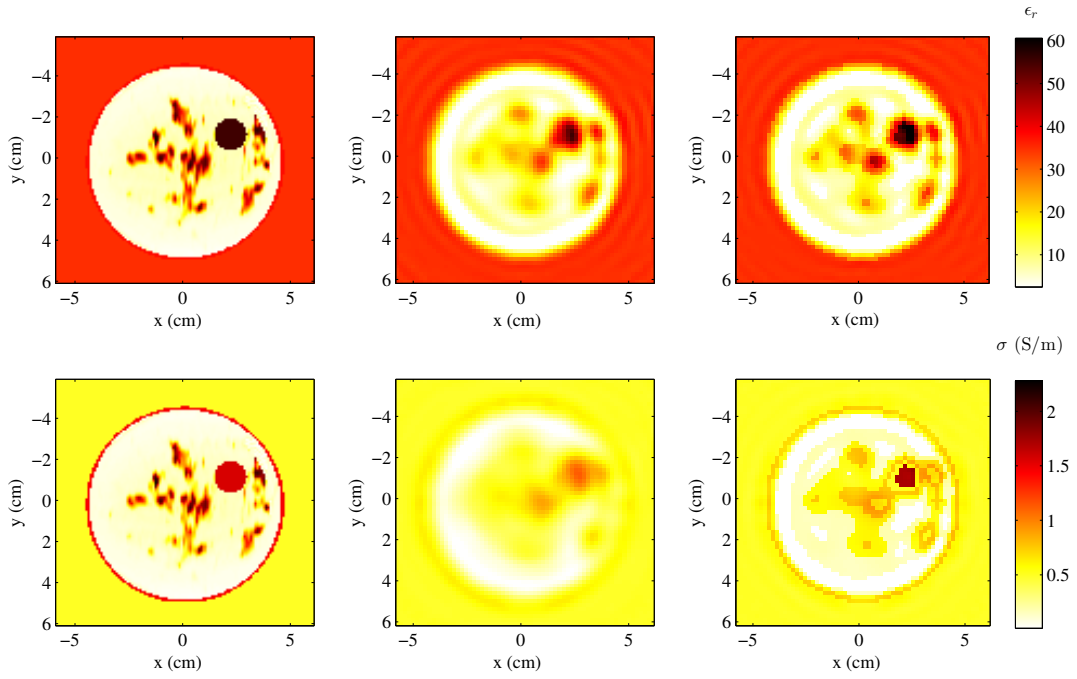


Figure 5: The results obtained for Model-2: maps of permittivity (up) and conductivity (down) reconstructed by means of CSI (second column) and VBA (third column) compared to the real object (first column).

[6] L. Gharsalli, H. Ayasso, B. Duchêne, and A. Mohammad-Djafari, *Microwave tomography for breast cancer detection within a Variational Bayesian Approach*, IEEE European Signal Processing Conference, EUSIPCO, Marrakech (2013).

[7] P.M. van den Berg, A. van Broekhoven, and A. Abubakar, *Extented contrast source inversion*, Inverse problems, 15, (1999), pp. 1325–1344.

[8] W.C. Chew, *Waves and Fields in Inhomogeneous Media*, IEEE Computer Society Press, (1999).

[9] O. Féron, B. Duchêne, and A. Mohammad-Djafari, *Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data*, Inverse Problems, 21, no. 6, (2005), pp. S95–S115.

[10] W.C. Gibson, *The Method of Moments in Electromagnetics*, Chapman & Hall/CRC, Boca Raton, (2008).

[11] V. Smídl, A. Quinn, *The Variational Bayes Method in Signal Processing*, Springer Verlag, Berlin, (2006).

[12] J.B. MacQueen, *Some Methods for Classification and Analysis of Multivariate Observations*, Univ. California Press, vol. 1, (1967), pp. 281–297.

[13] P.M. van den Berg, R.E. Kleinman, *A contrast source inversion method*, Inverse Problems, 13, no. 6, (1997), pp. 1607–1620.

[14] A. Fraysse, T. Rodet, *A gradient-like variational Bayesian algorithm*, Statistical Signal Processing Workshop (SSP), (2011), pp. 605–608.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

A time-dependent coefficient identification problem for the bioheat equation

A. Hazanee* **and D. Lesnic**

Department of Applied Mathematics,
University of Leeds, Leeds, LS2 9JT, UK
e-mail: mmah@leeds.ac.uk

Key words: Inverse problem, Heat source, Bioheat equation, Boundary element method, Regularization

Abstract

This work investigates the inverse problem of finding a time-dependent coefficient - the blood perfusion coefficient - in the bioheat equation with nonlocal boundary conditions and integral mass/energy specification. Conditions which ensure that the problem is well-posed are established and notably, the global continuous dependence on the data is proved. Furthermore, the numerical solution is obtained using the boundary element method together with the Tikhonov regularization. The generalized cross-validation is employed as a possible choice for the regularization parameter. A numerical example is presented to verify the accuracy and stability of the solution.

1 Introduction

Time-dependent coefficient identification problems with nonlocal boundary and/or integral overdetermination conditions have recently attracted revitalising interest, e.g. the reconstruction of a time-dependent diffusivity [1], a blood perfusion coefficient [2], or a heat source [3, 4].

In this paper, the determination of an unknown time-dependent blood perfusion coefficient for the bioheat equation is sought under nonlocal boundary and integral additional conditions. A simple transformation is used to reduce the bioheat equation to the classical heat equation. This inverse problem has already been proved to be uniquely solvable in [5], but no numerical reconstruction has been attempted. Therefore, the purpose of this study is to devise a numerical stable method for obtaining the solution of the inverse problem.

2 Mathematical Formulation

Consider the one-dimensional time-dependent bioheat equation, see e.g. [6],

$$u_t(x,t) = u_{xx}(x,t) - P(t)u(x,t) + f(x,t), \quad (x,t) \in D_T := (0,1) \times (0,T], \tag{1}$$

in a finite slab geometry $x \in (0,1)$, where $T > 0$ is a final time of interest, u is the temperature, P is the blood perfusion coefficient and f is a heat source. Equation (1) has to be solved subject to the initial condition

$$u(x,0) = \varphi(x), \quad x \in [0,1], \tag{2}$$

the convective Robin boundary condition at the left end $x = 0$, namely,

$$-u_x(0,t) = \alpha u(0,t), \quad t \in [0,T], \tag{3}$$

where α is a constant heat transfer coefficient, the nonlocal boundary condition

$$u(0,t) = u(1,t), \quad t \in [0,T], \tag{4}$$

and the integral (nonclassical) mass/energy specification/measurement

$$\int_0^1 u(x,t)dx = E(t), \quad t \in [0,T]. \tag{5}$$

We mention that the nonlocal boundary condition (4) is encountered in biological applications, see [7], whilst the mass/energy specification models processes related to particle diffusion in turbulent plasma, see [8], or heat conduction, see [9]. We are interested to find the pair solution $(P(t), u(x,t)) \in C[0,T] \times (C^{2,1}(D_T) \cap C^{1,0}(\overline{D}_T))$ satisfying the problem (1)–(5).

In [5], the unique solvability and local continuous dependence of the solution upon the data of the inverse problem (1)–(5) have been established. In this section, we establish the global continuous dependence based on a Gronwall's-type inequality. Note that the case $\alpha = 0$ has been dealt with in [3]. Herein we consider the case $\alpha \neq 0$ treated theoretically in [5].

First let us introduce some notations and assumptions as in [5]:

$$Y_0(x) = \begin{cases} \dfrac{2\mu_0}{\alpha(\mu_0^2+\alpha^2-2\alpha)}\left[(\mu_0^2-\alpha^2)\sin(\mu_0 x) + 2\alpha\mu_0\cos(\mu_0 x)\right], & \text{if } \alpha < 0, \\ -\dfrac{s_0+\alpha}{2\alpha(s_0^2-\alpha^2+2\alpha)}\left[(s_0-\alpha)^2 e^{s_0 x} - (s_0+\alpha)^2 e^{-s_0 x}\right], & \text{if } \alpha > 0, \end{cases}$$

$$Y_{2n}(x) = -\frac{4\pi n}{\alpha}\sin(2\pi n x),$$

$$Y_{2n-1}(x) = \frac{2\mu_n}{\alpha(\mu_n^2+\alpha^2-2\alpha)}\left[(\mu_n^2-\alpha^2)\sin(\mu_n x) + 2\alpha\mu_n\cos(\mu_n x)\right], \quad n \geq 1,$$

where s_0 is the unique positive solution of the equation $e^s = 1 + \dfrac{2\alpha}{s-\alpha}$ with $\alpha > 0$, $\mu_n = 2\pi n + O(1) \in (2\pi n, (2n+1)\pi)$, $n \in \mathbb{N}$ and $\mu_n = 2\pi n + O(1) \in ((2n-1)\pi, 2\pi n)$, $n = \mathbb{N}^*$ are monotone increasing positive solutions of the equation $\mu\sin(\mu/2) + \alpha\cos(\mu/2) = 0$ in $\alpha < 0$ and $\alpha > 0$, respectively. Denote by

$$\varphi_n = \int_0^1 \varphi(x) Y_n(x)\,dx, \quad f_n(t) = \int_0^1 f(x,t) Y_n(x)\,dx, \quad n \in \mathbb{N}.$$

We make the following assumptions on the input data φ, E and f:

(A$_1$) $\varphi(x) \in C^2[0,1], -\varphi'(0) = \alpha\varphi(0), \varphi(0) = \varphi(1)$,

(i) $\varphi_0 > 0, \varphi_{2n-1} \geq 0$, for $n \geq 1$, when $\alpha < 0$;

(ii) $\varphi_1 < 0, \varphi_{2n-1} \le 0$, for $n \ge 2$, when $\alpha > 0$.

(A_2) $E(t) \in C^1[0,T], E(0) = \int_0^1 \varphi(x)\,dx, E(t) > 0$ for $t \in [0,T]$.

(A_3) $f(x,t) \in C(\overline{D}_T), f(\cdot,t) \in C^2[0,1]$ for $t \in [0,T]$, $-f_x(0,t) = \alpha f(0,t), f(0,t) = f(1,t)$,

(i) $f_0 \ge 0, f_{2n-1} \ge 0$, for $n \ge 1$, when $\alpha < 0$;

(ii) $f_{2n-1} \le 0$, for $n \ge 2$, when $\alpha > 0$.

Kerimov and Ismailov [5] showed that if assumptions (A_1)–(A_3) are satisfied then the inverse problem (1)–(5) has a unique solution. Moreover, for arbitrary $P \in C[0,T]$ the following representation for $u(x,t)$ holds:

$$u(x,t) = u_0(t)X_0(x) + \sum_{n=1}^{\infty} \left[u_{2n-1}(t)X_{2n-1}(x) + u_{2n}(t)X_{2n}(x)\right], \tag{6}$$

where

$$X_0(x) = \begin{cases} \cos(\mu_0 x) - \frac{\alpha}{\mu_0}\sin(\mu_0 x), & \text{if } \alpha < 0, \\ \dfrac{s_0 - \alpha}{s_0 + \alpha}e^{s_0 x} + e^{-s_0 x}, & \text{if } \alpha > 0, \end{cases} \tag{7}$$

$$X_{2n}(x) = \cos(2\pi n x) - \frac{\alpha}{2\pi n}\sin(2\pi n x), \quad X_{2n-1}(x) = \cos(\mu_n x) - \frac{\alpha}{\mu_n}\sin(\mu_n x), \quad n \ge 1,$$

$$u_k(t) = \frac{\varphi_k e^{-\lambda_k t}}{r(t)} + \int_0^t \frac{f_k(\tau)e^{-\lambda_k(t-\tau)}}{r(t) - r(\tau)}\,d\tau, \quad k \ge 0,$$

$$\lambda_0 = \begin{cases} \mu_0^2, & \text{if } \alpha < 0, \\ -s_0^2, & \text{if } \alpha > 0, \end{cases} \qquad \lambda_{2n} = (2\pi n)^2, \quad \lambda_{2n-1} = \mu_n^2, \quad \text{for } n \ge 1,$$

and

$$r(t) = \exp\left(\int_0^t P(\xi)\,d\xi\right). \tag{8}$$

From (8) we remark that $P(t) = \dfrac{r'(t)}{r(t)}$, $r(t) > 0$, and $r(0) = 1$. Applying the overdetermination condition (5) we obtain that r satisfies the Volterra integral equation of the second kind

$$r(t) = F(t) + \int_0^t K(t,\tau)r(\tau)\,d\tau, \tag{9}$$

where

$$F(t) = \frac{1}{E(t)}\left(\varphi_0 e^{-\lambda_0 t}\int_0^1 X_0(x)\,dx - 2\alpha\sum_{n=1}^{\infty}\frac{\varphi_{2n-1}e^{-\mu_n^2 t}}{\mu_n^2}\right), \tag{10}$$

$$K(t,\tau) = \frac{1}{E(t)}\left(f_0(\tau)e^{-\lambda_0(t-\tau)}\int_0^1 X_0(x)\,dx - 2\alpha\sum_{n=1}^{\infty}\frac{f_{2n-1}(\tau)e^{-\mu_n^2(t-\tau)}}{\mu_n^2}\right). \tag{11}$$

Then we have the following theorem which gives the continuous dependence upon the data of the solution of the inverse problem (1)–(5).

Theorem 1. *Let $\mathcal{J}$ be the class of triplets of the form $\phi = (f, \varphi, E)$ which satisfy the assumptions (A_1)–(A_3) and*

$$\begin{cases} \|f\|_{C^{2,0}(\overline{D}_T)} \leq N_0, \quad \|\varphi\|_{C^2[0,1]} \leq N_1, \quad \|E\|_{C^1[0,T]} \leq N_2, \\ 0 < N_3 \leq \min\{\min\limits_{t\in[0,T]} |E(t)|, \min\limits_{t\in[0,T]} |E'(t)|\}, \end{cases} \tag{12}$$

for some positive constants N_i for $i = \overline{0,3}$. Then the solution pair $(P(t), u(x,t))$ of the inverse problem (1)–(5) *depends continuously upon the data in $\mathcal{J}$.*

Proof: Let $\phi = (f, \varphi, E)$ and $\tilde{\phi} = (\tilde{f}, \tilde{\varphi}, \tilde{E})$ be two sets of data in $\mathcal{J}$. Let (P, u) and $(\tilde{P}, \tilde{u})$ the corresponding solutions. Let $\tilde{r}$, $\tilde{F}$, and $\tilde{K}$ be the analogous quantities corresponding to (9)–(11). We can then estimate the difference

$$r(t) - \tilde{r}(t) = F(t) - \tilde{F}(t) + \int_0^t (K(t,\tau) - \tilde{K}(t,\tau))r(\tau)\,d\tau + \int_0^t \tilde{K}(t,\tau)(r(\tau) - \tilde{r}(\tau))\,d\tau. \tag{13}$$

According to the proof of Theorem 2 of [5] we have the following estimates

$$\|\tilde{K}\|_{C([0,T]^2)} \leq \frac{C_1 N_0}{N_3}, \quad \|K - \tilde{K}\|_{C([0,T]^2)} \leq C_1\|\phi - \tilde{\phi}\|, \quad \|F - \tilde{F}\|_{C([0,T])} \leq C_3\|\phi - \tilde{\phi}\|, \tag{14}$$

for some positive constants C_1, C_2 and C_3, and we denote

$$\|\phi - \tilde{\phi}\| := \|E - \tilde{E}\|_{C^1[0,T]} + \|\varphi - \tilde{\varphi}\|_{C^2[0,1]} + \|f - \tilde{f}\|_{C^2(\overline{D}_T)}. \tag{15}$$

Then, as in [4], using a Gronwall's-type inequality, see Theorem 16 of [10], we can obtain that

$$\|r - \tilde{r}\|_{C[0,T]} \leq C_4\|\phi - \tilde{\phi}\|, \tag{16}$$

for some positive constant C_4. This means that r is continuously dependent upon the input data ϕ. Similarly, one can prove that r' and u also depends continuously upon the data. Finally, as $P(t) = \frac{r'(t)}{r(t)}$, this implies that P also depends continuously upon the data. This concludes the stability proof.

□

Consider now the following transformation, see [11],

$$v(x,t) = r(t)u(x,t). \tag{17}$$

Then the inverse problem (1)–(5) becomes

$$v_t = v_{xx} + r(t)f(x,t), \quad (x,t) \in D_T, \tag{18}$$

$$v(x,0) = \varphi(x), \quad x \in [0,t], \tag{19}$$

$$v(0,t) = v(1,t), \quad v_x(0,t) + \alpha v(0,t) = 0, \quad t \in [0,T], \tag{20}$$

with the transformed integral condition

$$\int_0^1 v(x,t)dx = E(t)r(t), \quad t \in [0,T]. \tag{21}$$

We also have that $r \in C^1[0,T]$, $r(0) = 1$, $r(t) > 0$, for $t \in [0,T]$. Solving the inverse problem (18)–(21) for the solution pair $(r(t), v(x,t))$ yields afterwards the solution pair $(P(t), u(x,t))$ for the inverse problem (1)–(5) as given by

$$P(t) = \frac{r'(t)}{r(t)} \quad \text{and} \quad u(x,t) = \frac{v(x,t)}{r(t)}, \quad (x,t) \in D_T. \tag{22}$$

3 Numerical Method

The numerical discretization of the inverse problem (1)–(5) is based on the boundary element method (BEM), [4, 12, 13]. This reduces the problem to solving a linear system of $N \times N$ equations, generically written as

$$X\underline{\mathrm{r}} = \underline{\mathrm{y}}, \tag{23}$$

where N is the number of constant boundary elements discretising uniformly in time $[0, T]$ each of the boundaries $x = 0$ and $x = L$. In (23), the right-hand side vector $\underline{\mathrm{y}}$ depends only on the initial temperature $\varphi(x)$ and the coefficient α, whilst the matrix X depends on α and the energy measurement $E(t)$.

It is interesting to note that the measurement data (5) is usually contaminated with noise. This is modelled by perturbing (5) with random noise as

$$\underline{\mathrm{e}}^p = \underline{\mathrm{e}} + random('\text{Normal}', 0, \sigma, 1, N), \tag{24}$$

where $\underline{\mathrm{e}} = \left[e_i\right]_N := \left[E(\tilde{t}_i)\right]_N$, $i = \overline{1, N}$, and $random('\text{Normal}', 0, \sigma, 1, N)$ is a command in MATLAB for computing a vector of N random variables generates from a normal distribution with the zero mean and the standard deviation σ given by

$$\sigma = p \times \max_{t\in[0,T]} |E(t)|, \tag{25}$$

where p is the percentage of noise.

Hence, it means that the system matrix X is contaminated with noise and

$$\epsilon \approx \|X - X^\epsilon\|, \tag{26}$$

represents the level of noise. Thus, instead of (23), we have to solve

$$X^\epsilon \underline{\mathrm{r}} = \underline{\mathrm{y}}. \tag{27}$$

For the solution of this system we employ the second-order Tikhonov regularization, [4, 13], which yields the regularization solution

$$\underline{\mathrm{r}}_\lambda = \left((X^\epsilon)^{tr} X^\epsilon + \lambda R^{tr} R\right)^{-1} (X^\epsilon)^{tr} \underline{\mathrm{y}}, \tag{28}$$

where $\lambda \geq 0$ is a regularization parameter to be prescribed and $R^{tr}R$ is a second-order derivative regularization matrix given by, [14],

$$R^{tr}R = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & . & . & . \\ -2 & 5 & -4 & 1 & 0 & . & . & . \\ 1 & -4 & 6 & -4 & 1 & 0 & . & . \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & . \\ . & . & . & . & . & . & . & . \end{bmatrix}.$$

In (28), the regularization parameter can be selected according to the generalized cross validation (GCV) criterion which chooses $\lambda > 0$ as the minimum of the GCV function, see e.g. [15],

$$GCV(\lambda) = \frac{\|X^\epsilon \underline{\mathrm{r}}_\lambda - \underline{\mathrm{y}}\|^2}{[\text{trace}(I - X^\epsilon((X^\epsilon)^{tr} X^\epsilon + \lambda R^{tr} R)^{-1}(X^\epsilon)^{tr})]^2}. \tag{29}$$

The solution of the original inverse problem (1)–(5) can be obtained by substituting all approximate solutions v, r, and r' into (22). In order to obtain the solution $P(t)$, we also need to find the derivative function $r'(t)$ which can be approximated using finite differences as

$$r'(\tilde{t}_1) = \frac{r_\lambda(\tilde{t}_1) - 1}{T/(2N)}, \quad r'(\tilde{t}_i) = \frac{r_\lambda(\tilde{t}_i) - r_\lambda(\tilde{t}_{i-1})}{T/N}, \quad i = \overline{2, N}. \tag{30}$$

4 Numerical Example and Discussion

In this section, we consider a benchmark test example with the input $L = T = 1$, $\alpha = -1$, $\varphi(x) = 1 + x - x^2$, $f(x,t) = (3 + x - x^2)e^{-t}$, and $E(t) = 7e^{-t}/6$. Then the analytical solution of the problem (1)–(5) is given by

$$P(t) = 2, \quad u(x,t) = (1 + x - x^2)e^{-t}, \tag{31}$$

whilst the analytical solution for the transformed problem (18)–(21) is

$$r(t) = e^{2t}, \quad v(x,t) = (1 + x - x^2)e^{t}. \tag{32}$$

We present the numerical results obtained with a BEM mesh of $N = N_0 = 40$, where N_0 is the number of cells discretising uniformly the space interval $[0, L]$.

We start first with the case of exact data $E(t) = 7e^{-t}/6$, i.e. $p = 0$ in (24) and (25). The numerical results for the unknowns $r(t)$, $u(0,t)$, $r'(t)$, and $P(t)$ obtained using the straightforward inversion $\underline{\mathrm{r}} = X^{-1}\underline{\mathrm{y}}$, i.e. without regularization $\lambda = 0$ in (28), are compared with their corresponding analytic values e^{2t}, e^{-t}, $2e^{2t}$, and 2, in Figures 1(a)–1(d), respectively. From Figure 1 it can be seen that all the quantities of interest are accurate.

Next we investigate the stability of the numerical solution with respect to some $p = 1\%$ noise included in the input energy data, as in (24) and (25). The corresponding numerical results to Figure 1 (for exact data) are presented in Figure 2 (for noisy data). In Figures 2(a) and 2(b) the numerical results obtained for $r(t)$ and $u(0,t)$, respectively, are relatively accurate. However, the numerical results obtained for $r'(t)$ and $P(t) = r'(t)/r(t)$ shown in Figures 2(c) and 2(d), respectively, are highly unstable. This is because the differentiation of the noisy function $r(t)$ shown in Figure 2(a) with dashed line using the finite differences (30) is an unstable procedure. In order to deal with this instability one can employ the smoothing spline regularization of [16], but this requires the knowledge of the discrepancy between the exact and numerical values of $r(t)$ in Figure 2(a), which is not available if the exact solution is not available. Alternatively, we employ the second-order Tikhonov regularized solution (28) with the choice of the regularization parameter given by the minimum point of the GCV function (29). This is plotted in Figure 3 for $p = 1\%$ noise and the minimum yields the value $\lambda_{GCV} = 32$. With this value of λ, the solution (28) for $r(t)$ is plotted in Figure 4(a). By comparing with the previous unregularized solution shown in Figure 2(a), one can now see that the obtained solution for $r(t)$ is indeed smooth. Then the process of numerical differentiation (30) is permitted and a stable approximation can be obtained, as shown in Figures 4(c). There are some inaccuracies manifested at the end points $t = 0$ and 1, but this is commonly observed elsewhere when using other stabilising techniques such as the mollification method or, the Tikhonov reguarization for Fredholm integral equation of the first kind presented in detail in [17].

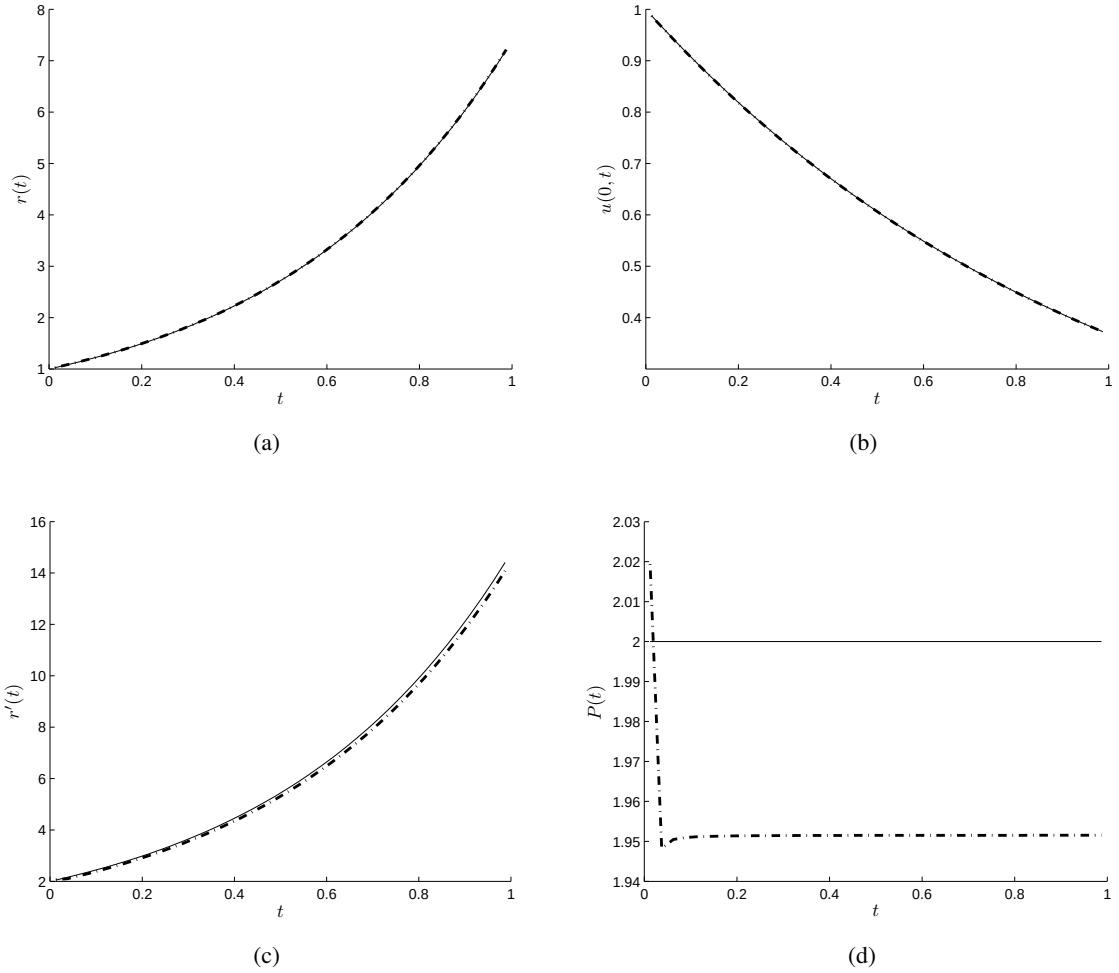


Figure 1: The analytical (——) and numerical $(- - -)$ results of (a) $r(t)$, (b) $u(0,t)$, (c) $r'(t)$, and (d) $P(t)$ obtained using no regularization, i.e. $\lambda = 0$, for exact data.

5 Conclusions

The inverse problem of finding a time-dependent blood perfusion coefficient for the bioheat equation with nonlocal boundary conditions and mass/energy specification has been investigated. The inverse problem has been transformed to an inverse heat source problem with an unknown source present in the integral over-determination condition. The numerical discretization was based on the BEM together with the Tikhonov regularization. The choice of the regularization parameter was based on the GCV. For a typical benchmark test example, accurate and stable numerical solutions have been obtained.

Acknowledgement

A. Hazanee would like to acknowledge the financial support received from the Ministry of Science and Technology of Thailand, and Prince of Songkla University Thailand, for pursuing her PhD at the University of Leeds.

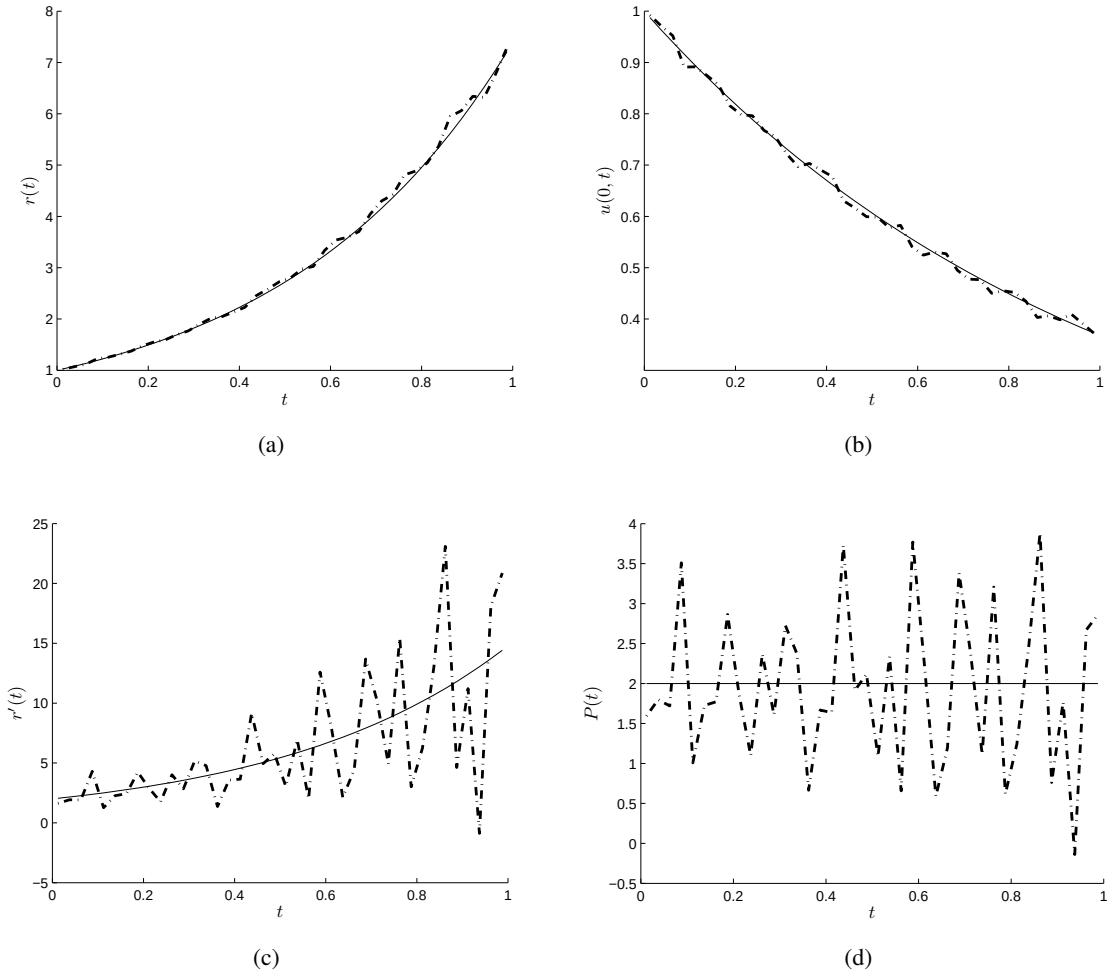


Figure 2: The analytical (——) and numerical $(- - -)$ results of (a) $r(t)$, (b) $u(0,t)$, (c) $r'(t)$, and (d) $P(t)$ obtained using no regularization, i.e. $\lambda = 0$, for $p = 1\%$ noise.

References

[1] N. I. Ivanchov, *Inverse problems for the heat-conduction equation with nonlocal boundary conditions*, Ukrainian Mathematical Journal, 45, (1993), pp. 1186–1192.

[2] M. I. Ismailov, F. Kanca, *An inverse coefficient problem for a parabolic equation in the case of nonlocal bounadry and overdetermination conditions*, Mathematical Methods in the Applied Sciences, 34, (2011), pp. 692–702.

[3] M. I. Ismailov, F. Kanca, D. Lesnic, *Determination of a time-dependent heat source under nonlocal boundary and integral overdetermination conditions*, Applied Mathematics and Computation, 218, (2011), pp. 4138–4146.

[4] A. Hazanee, M. I. Ismailov, D. Lesnic, N. B. Kerimov, *An inverse time-dependent source problem for the heat equation*, Applied Numerical Mathematics, 69, (2013), pp. 13–33.

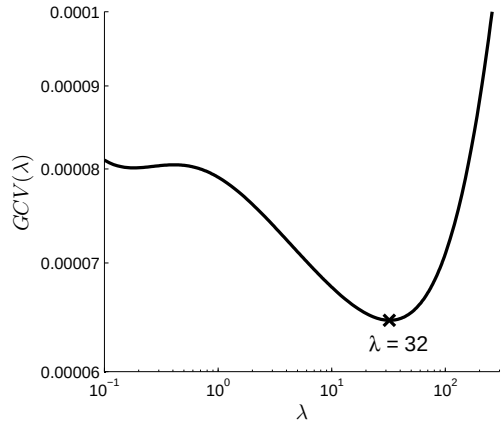


Figure 3: The GCV function, obtained using the second-order Tikhonov regularization for $p = 1\%$ noise.

[5] N. B. Kerimov, M. I. Ismailov, *An inverse coefficient problem for the heat equation in the case of nonlocal boundary conditions*, Journal of Mathematical Analysis and Applications, 396, (2012), pp. 546–554.

[6] D. Trucu, D. B. Ingham, D. Lesnic, *Inverse time-dependent perfusion coefficient identification*, Journal of Physics: Conference Series, 124, (2008), pp. 012050 (28 pages).

[7] A. M. Nakhushev, *Equations of Mathematical Biology*, Vysshaya Shkola, Moscow, (in Russian), (1995).

[8] N. I. Ionkin, *Solution of a boundary-value problem in heat conduction with a non-classical boundary condition*, Differential Equations, 13, (1977), pp. 204–211.

[9] J. R. Cannon, J. Hoek, *Diffusion subject to the specification of mass*, Journal of Mathematical Analysis and Applications, 115, (1986), pp. 517–529.

[10] S. S. Dragomir, *Some Gronwall Type Inequalities and Applications*, RGMiA Monographs, Victoria University, Australia, (2002).

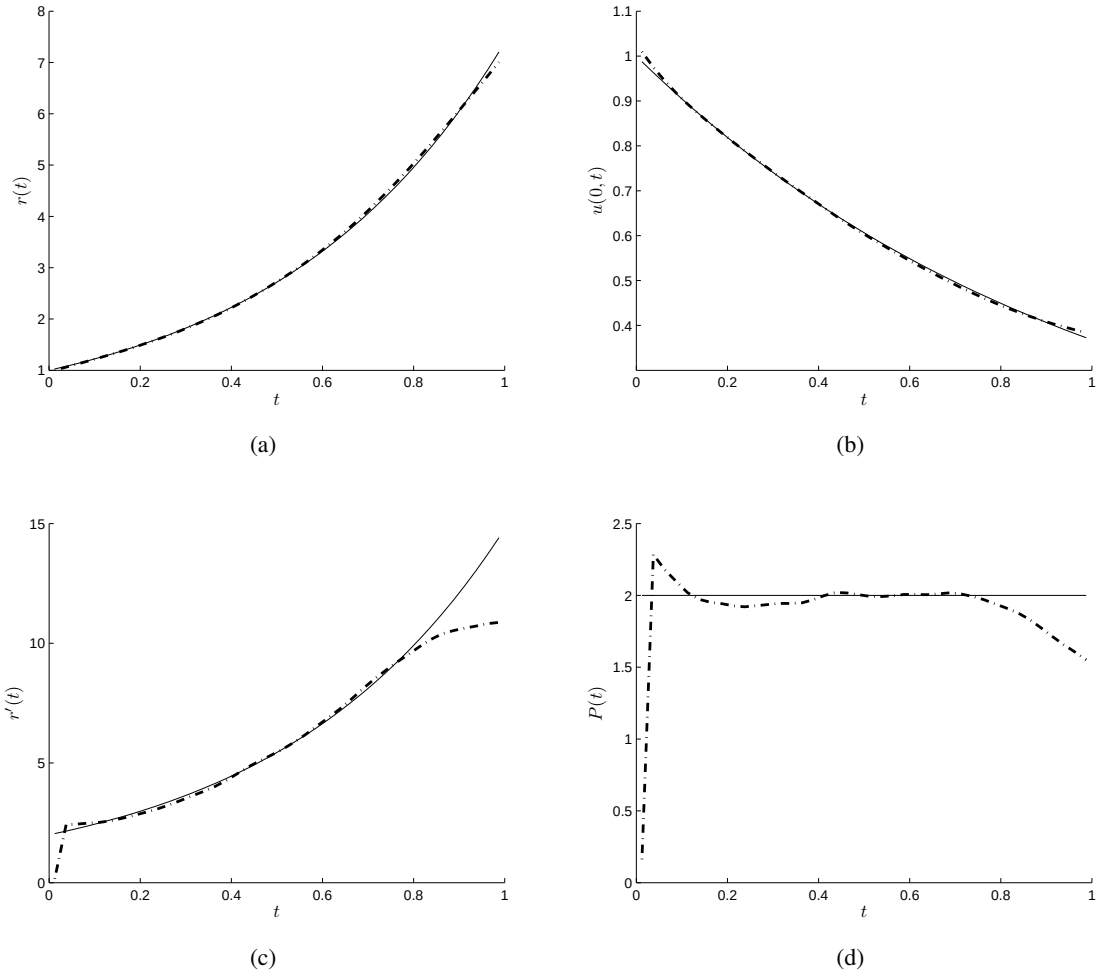
[11] J. R. Cannon, Y. Lin, S. Wang, *Determination of a control parameter in a parabolic partial differential equation*, Journal of the Australian Mathematical Society, Series B, 33, (1991), pp. 149–163.

[12] A. Farcas, D. Lesnic, *The boundary element method for the determination of a heat source dependent on one variable*, Journal of Engineering Mathematics, 54, (2006), pp. 375–388.

[13] A. Hazanee, D. Lesnic, *Determination of a time-dependent heat source from nonlocal boundary conditions*, Engineering Analysis with Boundary Elements, 37, (2013), pp. 936–956.

[14] S. Twomey, *On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature*, Journal of the Association for Computing Machinery, 10, (1963), pp. 97–101.

[15] L. Yan, C. L. Fu, F. L. Yang, *The method of fundamental solutions for the inverse heat source problem*, Engineering Analysis with Boundary Elements, 32, (2008), pp. 216–222.



(a) (b)

(c) (d)

Figure 4: The analytical (——) and numerical $(- - -)$ results of (a) $r(t)$, (b) $u(0,t)$, (c) $r'(t)$, and (d) $P(t)$ obtained using the second-order Tikhonov regularization with λ_{GCV}=32, for $p = 1\%$ noise.

[16] T. Wei, M. Li, *High order numerical derivatives for one-dimensional scatter noisy data*, Applied Mathematics and Computation, 175, (2006), pp. 1744–1759.

[17] D. Trucu, *Inverse Problems for Blood Perfusion Identification*, PhD Thesis, University of Leeds, (2009).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Trefftz functions for the modelling of two-dimensional temperature fields in flow boiling in a minichannel

Sylwia Hożejowska[1], **Beata Maciejewska**[2], **Mieczysław E. Poniewski** [3]

[1,2] Kielce University of Technology, Al. 1000-lecia PP 7, 25-314 Kielce, Poland
e-mail: [1]ztpsf@tu.kielce.pl; [2] beatam@tu.kielce.pl
[3] Warsaw University of Technology, Plock Campus, e-mail: meponiewski@pw.plock.pl

Key words: heat transfer coefficient, flow boiling, minichannel, Trefftz functions, FEM

Abstract
This paper focuses on the numerical determination of two-dimensional temperature fields in flow boiling in an asymmetrically heated rectangular and vertical minichannel. One of the minichannel walls is made of the DC supplied heating foil. The opposite wall is made of glass, which makes it possible to observe the two-phase structure and void fraction. The external side of the heating foil, coated with a layer of liquid crystals, is protected with a pane of glass. Thermosensitive liquid crystals were used to detect two-dimensional temperature distribution of the heating foil. Direct and inverse heat transfer problems that occur in flow boiling in a minichannel were solved using the methods based on Trefftz functions. The selection of particular functions was dependent on the form of the differential equation describing the heat transfer in solids (glass pane, heating foil) and in liquids. These functions were used globally (in the classical Trefftz method) and as shape functions in the FEM (FEMT) to find the two-dimensional temperature fields of the glass pane and the heating foil. Solving these functions led to computing a direct heat transfer problem in the glass pane and an inverse problem in the heating foil and the boiling liquid. Known foil and liquid temperature distributions and gradients enabled the heat transfer coefficient on the foil-liquid contact to be calculated from the Robin boundary condition. The results obtained using the numerical methods described for single-phase flows ending in boiling incipience and for two-phase flows have been summarized, compared and found to be similar.

1 Introduction

Technological advances evolving towards miniaturization impose a search for new heat transfer devices and a modification of the existing ones. The boiling heat transfer process is accompanied by a change of phase, which enhances the efficiency of the device. A body of research is being carried out with the aim of developing miniature heat exchangers that use the change of phase of refrigerant to obtain high heat flux at a small temperature difference between the heating surface and the heated

working fluid. Experimental part of this issue, briefly discussed further in this article, has been described in detail in [6] and [9]. The basic element of the experimental stand, which is the test section with a minichannel and FC-72 refrigerant flowing through it, is shown in Fig.1.

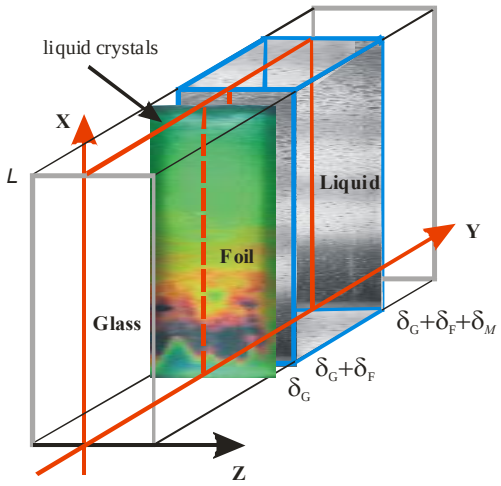


Figure 1: Test section of the minichannel. Temperature distribution at the foil-glass interface was determined using the liquid crystals technique (pictorial view, not to scale).

2 Mathematical Model

For convenience, in mathematical model only two dimensions were taken into account: dimension x along the flow direction and dimension y (perpendicular to the flow direction) relating to the thickness of the protecting glass (δ_G) and the foil (δ_F), and to the depth of the channel (δ_M), Fig. 1. Our considerations focus on the central part of the measurement module (along its height) so that the physical phenomena on the side edges do not affect thermodynamic parameters within the investigated segment, Fig 1. Further in this work the subscript is used to denote the parameters relating to: G - the protecting glass, F - the heating foil, L - the liquid and M - the minichannel. In the presented model the system is assumed to be in the steady state and temperature changes in the glass pane, heating foil and the liquid along the minichannel width are slight. Because Re <2000, the liquid flow was assumed to have a laminar character. Stationary two-dimensional temperatures of the glass pane T_G and the heating foil T_F satisfy the equations [11]

$$\nabla^2 T_G = 0 \quad \text{for} \quad (x,y) \in \Omega_G = \left\{(x,y) \in R^2 : 0 < x < L, \quad 0 < y < \delta_G\right\} \tag{1}$$

$$\nabla^2 T_F = -\frac{q_V}{\lambda_F} \quad \text{for} \quad (x,y) \in \Omega_F = \left\{(x,y) \in R^2 : 0 < x < L, \quad \delta_G < y < \delta_G + \delta_F\right\} \tag{2}$$

The following boundary conditions were adopted at the foil – glass interface

$$T_F(x_{mp},\delta_G) = T_G(x_{mp},\delta_G) = T_{mp} \quad \text{for} \quad mp = 1,2,...,MP \tag{3}$$

$$T_F(x,\delta_G) = T_G(x,\delta_G) \quad \text{for} \quad 0 \le x \le L \tag{4}$$

$$\lambda_F \frac{\partial T_F}{\partial y} = \lambda_G \frac{\partial T_G}{\partial y} \quad \text{for} = \delta_G \text{ i } 0 \le x \le L \tag{5}$$

Insulation was provided on the remaining boundaries of the protecting glass and the heating foil (except the boundary at the foil– liquid interface). The liquid temperature T_L in Ω_L, where

$$\Omega_L = \left\{(x,y) \in R^2 : 0 < x < L, \quad \delta_G + \delta_F < y < \delta_G + \delta_F + \delta_M \right\} \tag{6}$$

is calculated in different ways depending on whether in the minichannel the predominant flow is that of a single phase convection with boiling incipience or that of two-phase convection and extended bubbly and bubbly-slug boiling [4]. The former will be referred to as the single-phase flow with boiling incipience, the latter simply as the two-phase flow.
It is assumed that for the single-phase flow, when the void fraction is very low, liquid temperature T_L changes linearly from the minichannel inlet temperature T_{in} to the outlet temperature T_{out}

$$T_L(x,y) = T_{in} + \frac{T_{out} - T_{in}}{L} x \tag{7}$$

When the bubbly or bubbly-slug boiling occurs in the considerable part of the minichannel, we assume, after [1,4] that
1. the liquid flow is laminar and stationary with a constant mass flux density;
2. the velocity of the liquid has only one component w_x parallel to the heating surface

$$w_x(y) = \frac{\Delta p}{2\mu H}\left[-y^2 + (2\delta_G + 2\delta_F + \delta_M)y - (\delta_G + \delta_F + \delta_M)(\delta_G + \delta_F)\right] \tag{8}$$

3. the liquid temperatures at the inlet and outlet are known and equal to

$$T_L(0,y) = T_{in} \quad \text{for} \quad \delta_G + \delta_F \le y \le \delta_G + \delta_F + \delta_M \tag{9}$$

$$T_L(L,y) = T_{out} \quad \text{for} \quad \delta_G + \delta_F \le y \le \delta_G + \delta_F + \delta_M \tag{10}$$

4. the liquid temperature in the contact with the heating foil satisfies the following condition

$$T_L(x, \delta_G + \delta_F) = \begin{cases} T_F(x, \delta_G + \delta_F), & \text{if} \quad T_F(x, \delta_G + \delta_F) < T_{sat}(x) \\ T_{sat}(x), & \text{if} \quad T_F(x, \delta_G + \delta_F) \ge T_{sat}(x) \end{cases} \tag{11}$$

where T_{sat} is the saturation temperature dependent on the pressure $p(x)$;
5. the two-phase mixture per unit volume in the minichannel contains vapor phase and liquid phase in proportion φ and $(1-\varphi)$, respectively. We assume that the same proportions of vapor and liquid phases refer to any cross-sectional area of the minichannel and then to the heat exchange surface. For bubbly and bubbly-slug flows, following [1,4], we assume that the whole heat flux generated in the foil is transferred to the liquid phase in proportion carried over from the void fraction

$$\lambda_L \frac{\partial T_L}{\partial y} = \lambda_F (1 - \varphi(x)) \frac{\partial T_F}{\partial y} \qquad \text{for} \quad y = \delta_G + \delta_F \ \text{ and } \ 0 \le x \le L \tag{12}$$

Based on the adopted assumptions, the energy equation written only for the liquid phase takes the form

$$\lambda_L \nabla^2 T_L = w_x(y)\, c_L\, \rho_L \frac{\partial T_L}{\partial x} \qquad \text{for} \qquad (x,y) \in \Omega_L \tag{13}$$

where domain Ω_L is given by (6). Where the heating foil temperature distribution and the gradient are known, the heat transfer coefficient $\alpha(x)$ at the foil-liquid interface can be determined using the Robin boundary condition

$$\alpha(x) = -\lambda_F \frac{\partial T_F}{\partial y}(x, \delta_G + \delta_F)\left(T_F(x, \delta_G + \delta_F) - T_{ave}(x)\right)^{-1} \tag{14}$$

where T_{ave} is the reference temperature described further by (7) or equation

$$T_{ave}(x) = \begin{cases} \dfrac{1}{\delta_T} \displaystyle\int_{\delta_G+\delta_F}^{\delta_G+\delta_F+\delta_T} T_L(x,y)\,dy & \text{for Trefftz method} \\ \mathrm{T}_{sat}(x) & \text{for FEMT} \end{cases} \tag{15}$$

The thickness of the thermal layer δ_T in (16) is determined from the equation [1]

$$\delta_T = Pr^{-0.33}\,\delta_h \tag{16}$$

where hydraulic layer δ_h is computed as in [5].

3 Numerical Methods

To calculate two-dimensional temperature distributions of the glass pane and the heating foil, Trefftz functions for the Laplace equation (1) were used. Numerical computations were performed using two methods: the Trefftz method and the FEM with the Trefftz functions as shape functions (FEMT). In addition, the temperature of the liquid for the two-phase flow was determined using the Trefftz functions appropriate for the equation of energy (13) and described in [4].

3.1 Trefftz method

In this approach the unknown temperatures of the protecting glass T_G and the heating foil T_F are approximated with a linear combination of Trefftz functions $u_i(x,y)$ for equation (1), [11], that is

$$T_G(x,y) = \sum_{i=1}^{M} a_i u_i(x,y) \tag{17}$$

$$T_F(x,y) = \tilde{u}(x,y) + \sum_{j=1}^{N} b_j u_j(x,y) \tag{18}$$

where function $\tilde{u}(x,y)$ is the particular solution to equation (2) . For the single-phase flow with boiling incipience, the liquid temperature along the minichannel is assumed to change linearly, according to (7). For the two-phase flow, the liquid temperature T_L satisfies equation (13). Then, the liquid temperature T_L is approximated with a linear combination of the Trefftz functions $\theta_k(x,y)$ for equation (13)

$$T_L(x,y)=\sum_{k=1}^{K} d_k \theta_k(x,y) \tag{19}$$

Unknown coefficients a_i , b_j , d_k for i = 1,..., M, j = 1, ..., N, k= 1,..., K are determined using the least squares method. This leads to the minimization of the functionals for each of the functions, T_G, T_F and T_L. These functionals describe the mean squared error between the approximations and the prescribed boundary conditions. When the Trefftz method is used, the temperature of the protecting glass T_G is approximated first, followed by the approximation of the heating foil temperature T_F, and the liquid temperature T_L , [4,5,11].

3.2 The FEM with Trefftz functions (FEMT)

In this section the Trefftz functions for Laplace equation will serve the purpose of developing new functions of shape in the finite elements method (FEM). Using FEMT:

- we may solve problems in domains of complex shape and
- significantly reduce the number of the Trefftz functions used to approximate the solution in the subdomains as compared to method presented in subchapter **3.1**;
- we may adjust the layout of the nodes in the elements to the given values of temperature measurements and, thus exactly satisfy boundary condition (3);
- only the integrals at the edges of the elements are calculated.

In the presented approach shape functions were generated with a use of the Trefftz functions and the Lagrange interpolation [7]. This method is a generalization of the method described in [2]. Both problems in the protecting glass and the heating foil were solved with FEMT as in [3,10].

In order to calculate T_F and T_G , domains Ω_G, $\overline{\Omega}_F$ are divided into rectangular elements Ω_G^j for $j=1,2,\ldots,K_G$ and $\overline{\Omega}_F^j$ for $j=1,2,\ldots,K_F$, where

$$\overline{\Omega}_F=\left\{(x,y)\in R^2 : x_1<x<x_{MP},\quad \delta_G<y<\delta_G+\delta_F\right\} \tag{20}$$

In (20) x_1 is the coordinate of the first measurement point on the boundary $y=\delta_G$ and x_{MP} the coordinate of the last measurement point. In each element the coordinates of the nodes are determined and then functions of shape $f_{jk}(x,y)$ are developed in the form of linear combinations of the Trefftz functions. The functions of shape in nodes (x_i,y_i) of elements Ω_G^j and $\overline{\Omega}_F^j$ have the following property

$$f_{jk}(x_i,y_i)=\delta_{ki} \ , \qquad i=1,2,\ldots,nn \tag{21}$$

where: δ_{ki} denotes the Kronecker delta, j – number of elements, k – number of basis function in j-*th* element, nn – number of nodes in the element. The temperature approximation of the protecting glass in each element Ω_G^j is presented in the form of a linear combination of the basis functions $f_{jk}(x,y)$

$$T_G^j(x,y)=\sum_{k=1}^{n} T_G^n f_{jk}(x,y) \tag{22}$$

where: n – number of node in the entire domain Ω_G, T_G^n – sought-for temperature value in the n-th node of domain Ω_G. Indexes j, k, nn denote the same as in (21). Unknown coefficients T_G^n of linear

combination (22) are calculated by minimizing the functional that expresses a mean square error of the approximate solution on the boundary as well the difference the approximate temperature and the normal derivatives of approximants on the borders of the neighboring elements. The temperature of the heating foil is calculated analogously.

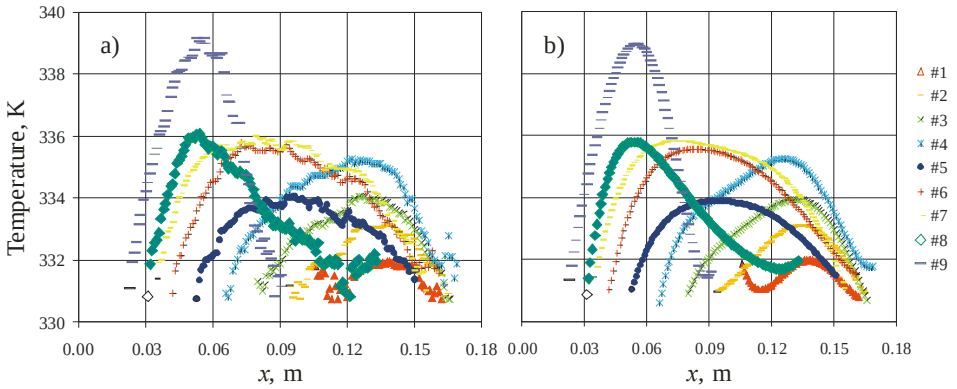


Figure 2: a) Temperature measurements depending on the distance from the inlet to the minichannel obtained through liquid crystal thermography. Experimental data: Reynolds number 867, mass flux 286 kg/(m^{2}s), pressure at the inlet 122 kPa, $q_V = 8.57 \cdot 10^4 \div 1.32 \cdot 10^5$ kW/m^3, b) Temperature measurements dependent on the distance from the inlet to the minichannel, smoothed with the Trefftz functions.

4 Results

Numerical calculations were performed for the data obtained from the experiments described in [9] (for a single phase flow) and in [6] (for a two phase flow). Those experiments pertained to the forced flow of FC-72 in the asymmetrically heated minichannel. The boiling occurred on a flat vertical heating surface made of *Haynes-230* acid-resistant rolled plate. Liquid crystal thermography helped calculate the temperature of the foil in contact with the glass pane. The glass pane isolating the minichannel helped observe the flow structure and determine the void fraction in the manner described by [6]. Other measured values included the temperature of FC-72 at the inlet and outlet of the minichannel, flow velocity, pressure at the minichannel inlet and outlet, voltage drop and the electrical current supplied to the foil. Both methods described in **3.1** and **3.2** involve computing the temperature of the protecting glass prior to computing the temperature of the heating foil. The temperature of the liquid for the two-phase flow is calculated by solving consecutive heat transfer inverse problem when the Trefftz method is used.

4.1 Results of calculations for a single phase flow with boiling incipience

A 360 mm long minichannel, which is 40 mm wide and 1 mm deep constitutes the main part of the experimental set-up. A heating foil used in the experiment was enhanced on one side with unevenly distributed minirecesses of less than 1 μm on average [9]. The focus was placed on the data related to the nucleate boiling incipience in a minichannel. It was assumed that the boiling incipience starts at the point at which the highest temperature of the heating surface was recorded, at the constant heat flux, after which the temperature falls rapidly. The heat flux supplied to the heating surface was

changed gradually in each series. Numerical results were obtained for the experimental data in Fig. 2. First the measurement data were approximated using a linear combination of the Trefftz functions u_i (x,y), as in [3].

Figure. 3 summarizes temperature measurement results for the foil and the temperature distributions of the protecting glass and the heating foil obtained using the Trefftz method. On boundary y= δ_G condition (3) is not fulfilled exactly but the maximum mean squared error calculated for (3) does not exceed 1K. Analogous two-dimensional temperature distribution of the protecting glass and the heating foil was obtained for the FEMT. It has to be emphasized that in the FEMT condition (3) is fulfilled exactly.

The heat transfer coefficient takes on similar values and has the same plot in both methods: it decreases along the minichannel length and rises rapidly with the boiling incipience, Fig. 4.

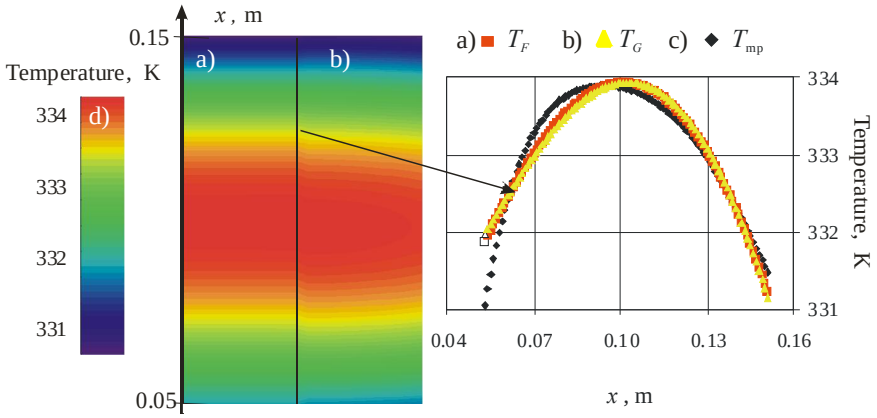


Figure 3: Temperature for setting #5: a) foil (T_F), b) glass (T_G), obtained with the classical Trefftz method, c) measurement temperature (T_{mp}); d) temperature scale.

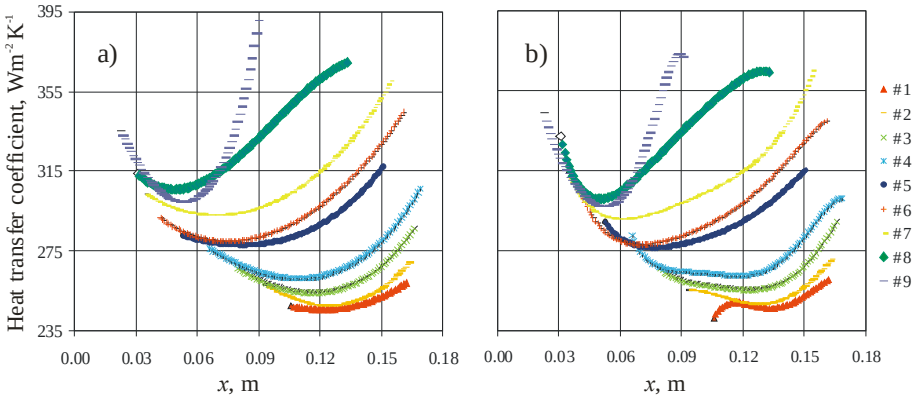


Figure 4: Heat transfer coefficient as a function of distance from the inlet to the minichannel calculated with: a) Trefftz method, b) FEMT.

4.2 Results of calculations for the two-phase flow

Numerical calculations were performed for the data presented in Fig. 5 and obtained from the experiments described [6]. In the vertical minichannel, 360 mm long, 20 mm wide and 1 mm deep,

a smooth heating foil was used. The void fraction calculated in [6] from the minichannel inlet was approximated by a quadratic function, Fig. 5.

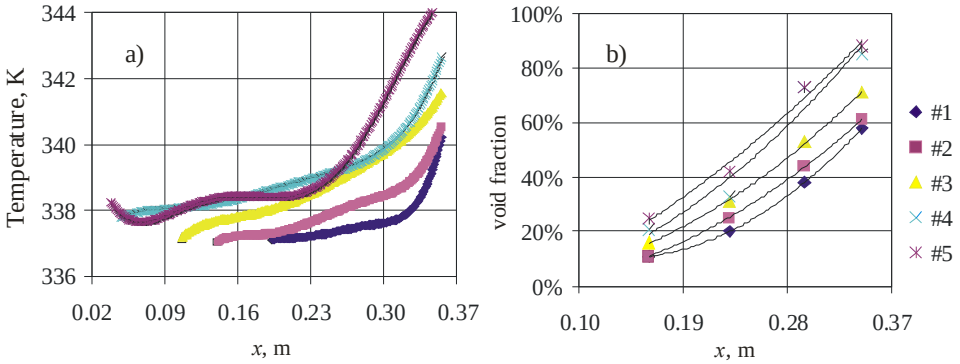


Figure 5: a) Temperature measurements dependent on the distance from the inlet to the minichannel, smoothed with the Trefftz functions, b) void fraction. Experimental data: Reynolds number 1426, mass flux 569 kg/(m^{2}s), pressure at the inlet 126 kPa, $q_V = 9 \cdot 10^5 \div 1.14 \cdot 10^6$ kW/m^3.

Figure 6 presents the two-dimensional distribution of liquid and foil temperature calculated with the Trefftz method. The temperature distribution of liquid, foil and saturation at the foil-liquid interface allowed the comparison of how the calculated temperature satisfied condition (11). The drop in the pressure along the minichannel caused a drop in the saturation and liquid temperatures. Figure 6 indicates the apparent effect of the assumed velocity profile (8) on the liquid temperature distribution.

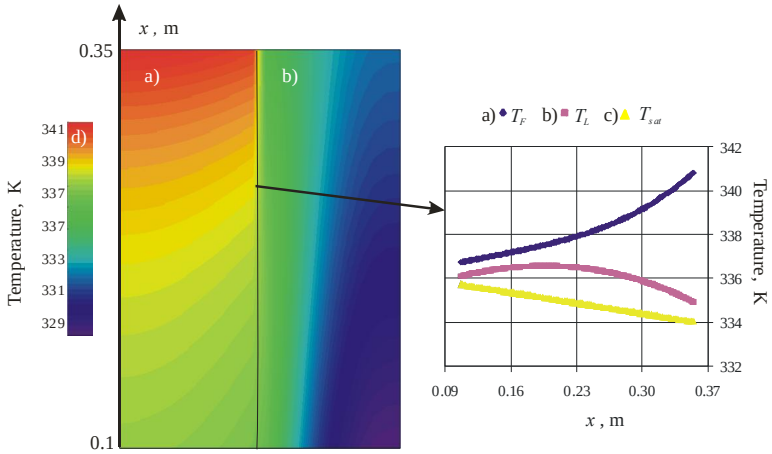


Figure 6: Temperature for setting #3: a) foil, b) liquid obtained with the Trefftz method, c) saturation.

Figure 7 summarizes the changes in heat transfer coefficient as a function of the distance from the minichannel inlet, calculated using the Trefftz method (Fig.7a) and the FEMT (Fig.7b). Heat transfer coefficients calculated using the FEMT have higher values than those calculated with the Trefftz method. The different plot of the heat transfer coefficient, Fig 7a) and Fig 7b), is a result of using a different method of determining the temperature of the liquid at the foil-liquid interface, eq. (15).
For the FEMT, the liquid temperature has the plot similar to that of a linear function, Fig. 6c). Trefftz method-based calculations of the liquid temperature show values lower at first and then higher than the saturation temperature, Fig 6b). The maximum relative differences between the heat transfer

coefficient obtained using the Trefftz method and the heat transfer coefficients obtained with the use of the FEMT were calculated from the formula

$$\Delta_{1,2} = max\left(\sqrt{\frac{\sum_{mp=1}^{MP}\left(\alpha_1\left(x_{mp}\right)-\alpha_2\left(x_{mp}\right)\right)^2}{\sum_{mp=1}^{MP}\left(\alpha_1\left(x_{mp}\right)\right)^2}}, \sqrt{\frac{\sum_{mp=1}^{MP}\left(\alpha_1\left(x_{mp}\right)-\alpha_2\left(x_{mp}\right)\right)^2}{\sum_{mp=1}^{MP}\left(\alpha_2\left(x_{mp}\right)\right)^2}} \right) \tag{23}$$

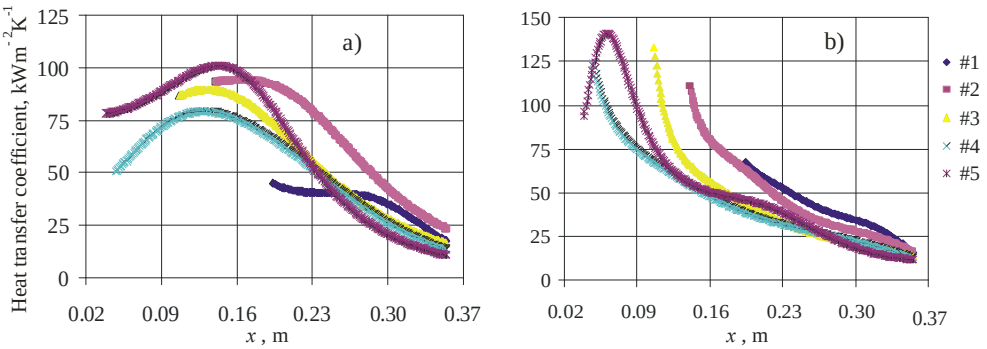


Figure 7: Heat transfer coefficient as a function of the minichannel height calculated by: a) Trefftz method, b) FEMT, for the data from Fig. 5.

Then $\alpha_1(x)$ denotes the heat transfer coefficient obtained with the Trefftz method and $\alpha_2(x)$ the heat transfer coefficient obtained with the FEMT. The results of the calculations are shown in Fig.8.

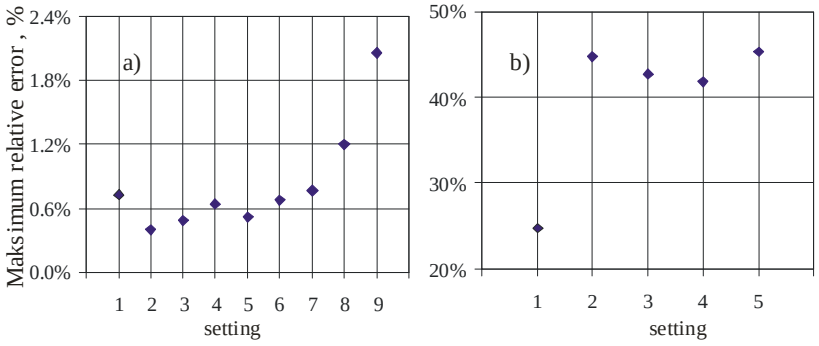


Figure 8. Maximum relative differences (23) calculated for: a) single-phase flow, b) two-phase flow.

5 Conclusions

The presented methods were applied to solve direct and inverse heat transfer problems in flow boiling in a vertical minichannel. These methods are suitable for determining two-dimensional temperature distributions in adjacent domains (protecting glass and heating foil). The two-dimensional liquid temperature distribution in the two-phase flow was calculated with the Trefftz method. Knowledge of the temperature of the heating foil and the liquid at the foil-liquid interface helped calculate the heat

transfer coefficient. Compared with the classical Trefftz method, the application of the FEMT method allowed making calculations using fewer Trefftz functions without compromising the accuracy of the results. In the FEMT, by placing the temperature measurements in the nodes of the finite elements all boundary conditions related to these measurements are satisfied exactly.

Acknowledgements

This paper was partially supported by the Polish Ministry of Higher Education and Science, Grant No. N N512 477339 (contract no. 4773/B/T02/2010/39).

References

[1] T. Bohdal, *Modelling the process of bubble boiling in flows*, Archives of Thermodynamics, 21, (2000), pp.34-75.

[2] M.J. Ciałkowski, A. Frąckowiak, *Solution of the stationary 2D inverse heat conduction problem by Treffetz method* , J. Thermal Science, 11, (2002), pp. 148-162.

[3] K. Grysa, S. Hożejowska, B. Maciejewska, *Compensatory calculus and Trefftz functions applied to local heat transfer coefficient determination in a minichannel,* J. Theoretical and Applied Mechanics, 50, (2012), pp.1087-1096.

[4] S. Hożejowska, R. Kaniowski, M.E. Poniewski, *Application of adjustment calculus to the Trefftz method for calculating temperature field of the boiling liquid flowing in a minichannel*, Int. J. Numerical Methods for Heat and Fluid Flow, 24, (2014), iss.8 – in print.

[5] S. Hożejowska, R. Kaniowski, M.E. Poniewski, *Numerical calculations of two – dimensional temperature field in a vertical minichannel*, 51st European Two-Phase Flow Group Meeting, (2013) p.19, Lyon.

[6] R. Kaniowski, *The analysis of two –phase flow in asymmetrically heated minichannels* (*in Polish*), Ph.D Thesis, Warsaw University of Technology, Plock Campus (2011).

[7] D. Kinkaid, W. Cheney, *Numerical Analysis: Mathematics of Scientific Computing*, 3rd ed. (2002), (Brooks Cole).

[8] B. Maciejewska, M. Piasecka, *Flow boiling in a minichannel: applications of numerical methods in heat transfer coefficient determination***,** Proc. 8th World Congress on Experimental Heat Transfer, Fluid Mechanics and Thermodynamics, (2013), CD, paper No. 119 (in Portugal).

[9] M. Piasecka, *Determination of the temperature field using liquid crystal thermography and analysis of two-phase flow structures in research on boiling heat transfer in a minichannel,* Metrology and Measurement Systems, XX, No. 2, (2013), pp. 205-216.

[10] M. Piasecka, B. Maciejewska, *Enhanced heating surface application in a minichannel flow and use the FEM and Trefftz functions to the solution of inverse heat transfer problem*, Experimental Thermal and Fluid Science, 44, (2013), pp. 23-33.

[11] M. Piasecka, M.E. Poniewski, S. Hożejowska, *Experimental evaluation of flow boiling incipience of subcooled fluid in a narrow channel,* Int. J. of Heat and Fluid Flow, 25, (2004), pp.159-172.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Efficient Markov chain Monte Carlo sampling for electrical impedance tomography

Erfang. Ma*

Mathematics and Physics Center
Xi'an Jiaotong Liverpool University, Renai road 111, Suzhou, China
e-mail: erfang.ma@xjtlu.edu.cn

Key words: electrical impedance tomography, Bayesian inference, Markov chain Monte Carlo

Abstract

This paper studies electrical impedance tomography (EIT) using Bayesian inference [1]. The resulting posterior distribution is sampled by Markov chain Monte Carlo (MCMC) [2]. This paper studies a toy model of EIT as in [3], and focuses on efficient MCMC sampling for this model.

First, this paper analyses the computation of forward map of EIT which is the bottleneck of each MCMC update. The forward map is computed by the finite element method [4]. Here its exact computation has been made up to *five* times more efficient, by updating the Cholesky factor of the stiffness matrix [5]. Since the forward map computation takes up nearly all the cpu-time in each MCMC update, the overall efficiency of MCMC algorithms can be improved almost to the same amount. The forward map can also be computed approximately by local linearisation, and this approximate computation is much more efficient than the exact one. Without loss of efficiency, this approximate computation has been made more accurate here, after a log-transformation has been introduced into the local linearisation process. Later on, this improvement of accuracy will play an important role when the approximate computation of forward map has been employed for devising efficient MCMC algorithms.

Second, the paper presents two novel MCMC algorithms for sampling the posterior distribution in the toy model of EIT. The two algorithms are made within the 'multiple prior update' [6] and the 'delayed-acceptance Metropolis-Hastings' [7] schemes respectively. Both of them have MCMC proposals that are made of localized updates, so that the forward map computation in each MCMC update can be made efficient by updating the Cholesky factor of the stiffness matrix. The algorithms' performances are both compared to that of the standard single site Metropolis [8], which is considered hard to surpass [3]. The algorithm of 'multiple prior update' is found to be *six* times more efficient; while the delayed-acceptance Metropolis-Hastings with single site update is *at least twice* more efficient.

1 Introduction

Electrical impedance tomography (ETI) concerns inferring the distribution of electrical property over a body, with measurement of voltage and current on the boundary of the body [9]. EIT has many applications in various fields, and has been extensively studied during the past years [10].

This paper discusses the reconstruction problem of EIT using Bayesian inference with MCMC. The Bayesian approach to EIT was pioneered by [1, 6]. The Bayesian approach results in a posterior probability distribution about the unknown electrical distribution. With samples of this distribution, we can not only estimate the unknown electrical distribution, but also quantify the error of this estimate. Drawing samples from this posterior distribution is however a non-trivial task and has to rely on the Markov chain Monte Carlo (MCMC) method [2]. Though many MCMC algorithms work for sampling this posterior distribution, few are efficient enough for practical use [3]. In [3], performance of various MCMC algorithms are compared on sampling a posterior distribution for a toy EIT model. They concluded that the classical single-site Metropolis [8] is actually hard to beat for this application, and more efficient MCMC algorithms are required in order for the Bayesian approach to EIT to be used in practice.

This study follows the work in [3], and focuses on efficient MCMC sampling for EIT. We speed up the computation in each MCMC update by updating the Cholesky factor of stiffness matrix in successive calls to the forward map of EIT. We also employ two variants of Metrpolis-Hastings schemes[11], to make novel and efficient MCMC algorithms. The paper is organized as follow. In section 2, we introduce the toy EIT model that this paper focuses on, the posterior probability distribution resulting from the Bayesian approach to this EIT model. MCMC algorithms will be employed for sampling this distribution. In section 3, we speed up the computation of a successive calls to the forward map of EIT in the context of MCMC. This is done by updating the Cholesky factor of the stiffness matrix in the finite element formulation. We also make local linearisation of forward map more accurate, by introducing a log-transformation into this process. In section 4, two novel and efficient MCMC algorithms have been proposed and employed for sampling the posterior distribution in our toy EIT model. Finally, we conclude the findings and discuss the future work.

2 A toy model of EIT

The EIT model discussed in this paper is shown in Figure 1. In this figure, phantom conductivity x is distributed over a square region Ω. Assuming that the region is divided into a 24×24 grid of squares, the true conductivity is constant on each square, and is either three (white colour in the figure) or four (black colour in the figure) in some arbitrary unit. To infer this conductivity distribution, sixteen point electrodes, numbered 1-16, have been evenly placed on the boundary $\partial\Omega$ of the region. Current is injected through one of these electrodes and extracted uniformly through the boundary $\partial\Omega$. The resulting voltage at all of the electrodes are measured. This procedure is then repeated for sixteen times, with each electrode being the injector of current for once. In the end, sixteen sets of measured voltages at all the electrodes are obtained, and they are denoted as u_E.

The measured voltages u_E at all the electrodes depend on the conductivity distribution x over the region. This dependence $u_E(x)$ is the forward map of EIT, given that the injected current is of a fixed pattern.

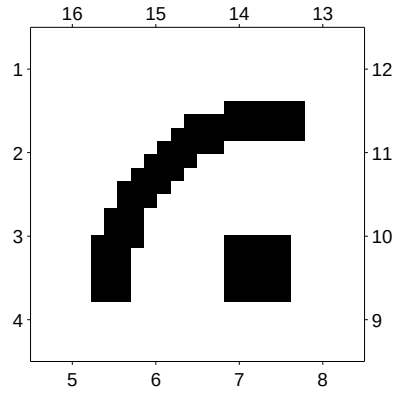


Figure 1: The toy EIT model. Phantom conductivity is distributed over a square region. White and black represent conductivity to be three and four in some unit. Sixteen point electrodes denoted as 1,2,... 16, are evenly placed on the boundary of the region. Through these electrodes current are injected and the resulting voltages are measured.

Evaluating $u_E(x)$ for some x involves solving the following boundary value problem (BVP):

$$-\nabla \cdot (x(s)\nabla u(s)) = 0 \qquad s \in \Omega, \tag{1a}$$

$$-x(s)\frac{\partial u}{\partial n}(s) = j(s) \qquad s \in \partial\Omega, \tag{1b}$$

$$\int_{\partial\Omega} j(s)\mathrm{d}s = 0, \tag{1c}$$

$$\int_{\partial\Omega} u(s)\mathrm{d}s = 0. \tag{1d}$$

$u(s)$ denotes the potential at spatial point s. (1d) represents a way of selecting the potential reference that ensure the solution of this BVP is unique. $j(s)$ is the current pattern through the boundary $\partial\Omega$ of the region. $j(s)$ has to satisfy (1c) because of the current conservation. In this study, $j(s) = \delta(s - e_k) - 1/|\partial\Omega|$. $\delta(s - e_k)$ is a dirac delta function. e_k denotes the location of the k-th point electrode on the boundary $\partial\Omega$. $|\partial\Omega|$ is the length of the boundary. For a conductivity distribution x, we solve this BVP for $u(s)$ which contains u_E. This BVP is solved by the finite element method (FEM). In FEM, the region has been divided into 24×24 squares. So the conductivity x is of 24×24 components. It is undetermined to infer x with only 16×15 measurements. However, the Bayesian approach can overcome this issue by introducing a prior distribution for x.

In the Bayesian approach to our EIT model, the conductivity x is considered to be random. As in [3], an example of the posterior distribution resulting from the Bayesian approach to this model could be:

$$\pi(x|u_E^{\text{noise}}) \propto \exp\left\{\beta \sum_{i\sim j} v(x_i - x_j)\right\} I_{[2.5,4.5]^n}(x) \cdot \exp\left\{-\frac{1}{2\varsigma^2}\|u_E^{\text{noise}} - u_E(x)\|^2\right\}. \tag{2}$$

Here $i \sim j$ is defined to be a horizontal or vertical nearest neighbour as said in [3]. $v(\cdot)$ is the tricube function of [12]:

$$v(t) = \begin{cases} \frac{1}{s_0}\left(1 - |t/s_0|^3\right)^3, & \text{if } |t| < s_0, \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

and $I_{[2.5,4.5]^n}(x)$ is an indicator function restraining every component of x to be within the range $[2.5, 4.5]$. $n = 24^2$ because x is a n-dimensional vector. u_E^{noise} is the voltage measured at the electrodes that is contaminated with additive noise. ς is the level of noise. β is a regularization parameter. Here it is set to be 0.5 as in [3]. Samples from this posterior distribution can be used to produce estimate for the unknown conductivity distribution. To draw samples from this distribution, the MCMC method is employed.

3 Forward map computation

3.1 Exact computation

As said in section 2, evaluating the forward map $u_E(x)$ consists of solving the BVP (1) using the FEM [4]. The FEM results in a sparse and positive system of linear equations:

$$Ku = f. \tag{4}$$

Its solution is the potential u over the region that certainly contains $u_E(x)$. In (4), the stiffness matrix K depends on the conductivity x as follows:

$$K = \sum_{i=1}^{n} x_i K_i + \lambda cc^T. \tag{5}$$

Here K_i for $i = 1, 2 \ldots, n$ are all sparse and symmetric matrices. They are all built on the same local stiffness matrix, and only depend on the geometry of the mesh and the element function employed. c is a n-dimensional column vector with its components either 0 or 1. For each i with $1 \leq i \leq n$, $c_i = 1$ if x_i is on the boundary of the region. Otherwise $c_i = 0$. λcc^T is a penalty term that corresponds to the potential reference (1d). f only depends on the current j through the boundary of the region. In this study, the direct method from [13] is employed for the solution of this system.

In MCMC for EIT, forward map $u_E(x)$ has to be evaluated for a series of $\{x^{(i)}\}$. Normally, the stiffness matrix is formed for each $x^{(i)}$, then the resulting sparse system is solved by the Cholesky factorization. In this study, we obtain the Cholesky factor of the stiffness matrix K^{i+1} for $x^{(i+1)}$ by modifying on that of x^i as said in [5]. Suppose that $x^{(i+1)}$ differs from $x^{(i)}$ at only its k-th component, and $x_k^{(i+1)} - x_k^{(i)} = \delta x$. Then

$$K^{(i+1)} = K^{(i)} + \delta x K_k, \tag{6}$$

according to (5). Recall that K_k is sparse, symmetric and of low rank. So (6) can be written as

$$K^{(i+1)} = K^{(i)} \pm CC^T. \tag{7}$$

for some suitable C. Equation (7) indicates that the method in [5] can be employed to find the Cholesky factor of $K^{(i+1)}$ based on that of $K^{(i)}$. When $x^{(i+1)}$ differs from $x^{(i)}$ at more than one component, $x^{(i+1)}$ can be obtained from a series of successive one-component updates on $x^{(i)}$. Hence the Cholesky factor of $K^{(i+1)}$ can be obtained from a series of updates on that of $K^{(i)}$.

Efficiency of our updating Cholesky factor method was measured on the toy EIT model (figure 1). In FEM, the region was discretized into 24×24 grid of equal size. Bilinear rectangular element was em-

ployed, and the resulting local stiffness matrix was

$$K_4 = \frac{1}{3}\begin{bmatrix} 2 & -0.5 & -1 & -0.5 \\ -0.5 & 2 & -0.5 & -1 \\ -1 & -0.5 & 2 & -0.5 \\ -0.5 & -1 & -0.5 & 2 \end{bmatrix}.$$

Given that $u_E(x)$ had been computed, $u_E(x + \Delta x)$ was calculated by our updating Choleksy factor method as well as in the standard way. All the computation was carried out in MATLAB, and the CPU-time was recorded with the 'profile' command. According to the record, updating Cholesky factor method was found to be about *five* more efficient than the standard method, when Δx had only one non-zero component. The updating Cholesky factor method could still be four and three times more efficient when the non-zero component of Δx was two and four. As Δx had more non-zero components, updating Cholesky factor method got less efficient.

3.2 Approximate computation

In MCMC for EIT, the forward map of EIT is sometimes computed approximately with local linearisation. Approximate computation of forward map can be used to make efficient MCMC algorithms, because it is much cheaper than the exact computation. The approximation of forward map has to be as accurate as possible, in order for MCMC algorithms involving it to achieve maximum efficiency. Because of this, we modify the ordinary approximation of forward map with local linearisation, and make it more accurate.

Normally, the approximation for $u_E(x + \Delta x)$ with local linearisation at x is

$$u_E^*(x + \Delta x) \equiv u_E(x) + \left.\frac{\partial u_E}{\partial x}\right|_x \Delta x. \tag{8}$$

Here $\partial u_E/\partial x$ is the Jacobian of forward map. Computing $u_E^*(x + \Delta x)$ is well known to be much faster than $u_E(x + \Delta x)$. In this study, a new approximation is made by introducing a log-transformation into (8). Define $\rho(x) = \log x$ and $\hat{u}_E(\rho(x)) = u_E(x)$. Let $\Delta\rho = \log(x + \Delta x) - \log(x)$ so that $u_E(x + \Delta x) = \hat{u}_E(\rho + \Delta\rho)$. Then

$$\hat{u}_E^*(\rho + \Delta\rho) \equiv \hat{u}_E(\rho) + \left.\frac{\partial \hat{u}_E}{\partial \rho}\right|_\rho \Delta\rho \tag{9}$$

is our new approximation to $u_E(x + \Delta x)$.

The accuracy of the two approximations (8) and (9) was compared on our toy model (1), in terms of the log-likelihood term: $L(x+\Delta x) = -\frac{1}{2\varsigma^2}\|u_E^{\text{noise}} - u_E(x+\Delta x)\|^2$. The approximated log-likelihood term $L^*(x + \Delta x) = -\frac{1}{2\varsigma^2}\|u_E^{\text{noise}} - \dot{u}_E^*(x + \Delta x)\|^2$ was computed with $\dot{u}_E^*(x + \Delta x)$ being $\hat{u}_E^*(x + \Delta x)$ and $u_E^*(x + \Delta x)$, respectively. The error $\Delta L = \|L^*(x + \Delta x) - L(x + \Delta x)\|$ of approximated likelihood term, associated with approximation (8) and (9) was then compared.

In practice, ΔL was computed for each $\Delta x = (\xi, \xi, \cdots, \xi)$ with ξ taking various values around zero. ΔL was plotted against ξ in figure 2. In this figure, the dashed and continuous curves correspond to the ordinary local linearisation (8) and the local linearisation with log-transformation (9). As seen, the continuous curve is always well below the dashed one. This indicates that our approximation (9) is more accurate.

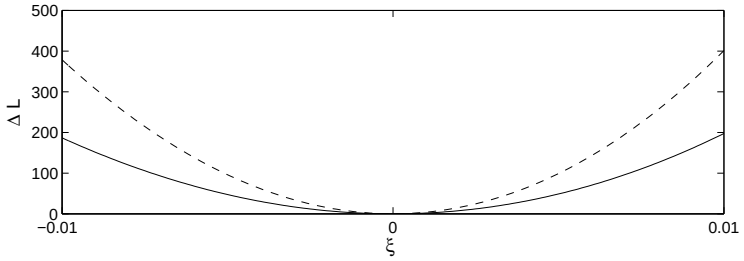


Figure 2: Error ΔL of approximated log-likelihood based on two approximations of forward map $u_E(x+\Delta x)$ when $\Delta x = (\xi, \xi, \cdots, \xi)$. The continuous and dashed curves correspond to the local linearisation with and without log-transformation, respectively.

4 Novel MCMC algorithms

4.1 Multiple prior block update

Our first novel MCMC algorithm is made in the 'multiple prior updates, likelihood acceptance ratio' scheme [6]. This scheme is a special case of the surrogate transition method [2]. Assume $\pi(x) = \pi_p(x)\pi_l(x)$ is the target distribution to be sampled. The multiple prior update scheme is the surrogate transition method with $\pi_p(x)$ as its global approximation $\pi^*(x)$ of $\pi(x)$. Assume that the surrogate Markov chain begins with state x and ends with state y. $S(x, y)$ is the transition function from x to y. Then the probability for the state y to be accepted in the main Markov chain converging to $\pi(x)$ is

$$\min\left\{1, \frac{\pi(y)S(y,x)}{\pi(x)S(x,y)}\right\} = \min\left\{1, \frac{\pi_p(y)\pi_l(y)S(y,x)}{\pi_p(x)\pi_l(x)S(x,y)}\right\} = \min\left\{1, \frac{\pi_l(y)}{\pi_l(x)}\right\}. \tag{10}$$

The second equation in (10) is because

$$\pi_p(y)S(y,x) = \pi_p(x)S(x,y). \tag{11}$$

i.e. the surrogate chain is a reversible Markov chain converging to π_p. Equation (10) indicates that the proposal y almost always gets accepted when $\pi_l(y) \approx \pi_l(x)$.

In this study, the update in the surrogate Markov chain is made on a block of neighboring components altogether, as suggested by [14]. However, the block update here leaves the mean value of state variable invariant. Assume that x and x' are the current and proposal state, respectively. The indices for all the components in the i-th block is $b^i = \left(b^i_1, b^i_2, \cdots, b^i_m\right)$. The update at this i-th block is:

$$x'_{b^i} = x_{b^i} + (I_m - a \cdot a^t)\varepsilon. \tag{12}$$

Here x_{b^i} and x'_{b^i} are the components of current and proposal state that are in the i-th block, respectively. $\varepsilon \sim N(0, \sigma^2 I_m)$ and $a = (1/\sqrt{m}, 1/\sqrt{m}, \cdots, 1/\sqrt{m})$. As seen $(I_m - a \cdot a^t)\varepsilon$ has zero mean, so x' and x has the same mean value. Since x and x' also differs from each other at only few neighbouring components, $u_E(x) \approx u_E(x')$. Hence $\pi_l(x) \approx \pi_l(x')$. So proposal x' will very likely get accepted according to (10).

In practice, the multiple prior block update is mixed with single-site Metropolis ([8]) when sampling the posterior distribution (2). The block update (12) maintains the mean value of the state variable, and the

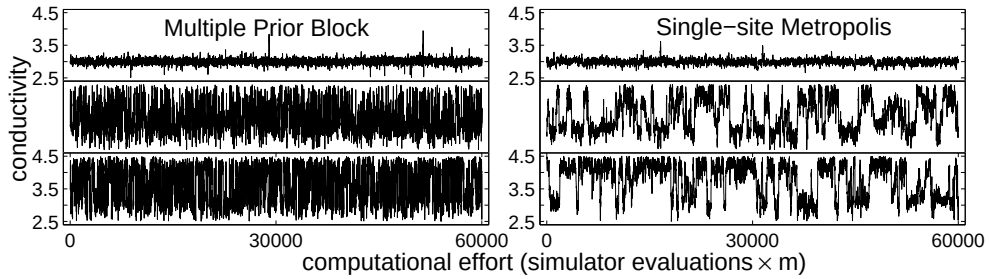


Figure 3: Traces of three representative pixels from Markov chains updated by multiple prior block update and Single-site Metropolis.

Markov chain updated by this proposal alone is reducible. After Inclusion with single-site update, the Markov chain will be desirably irreducible [2]. In practice, 90% of the moves in this mixture are the multiple prior update, so that this mixture update inherit most of its characteristic. The block was chosen at random at a time, and blocks can overlap with each other. This algorithm based on the 2×2 block was found to perform the best. The optimal update in the surrogate chain was found to be about ten.

The performance of our multiple prior block update is compared with that of the single-site Metropolis [8]. As in [3], figure 3 shows the traces of three representative pixels in the conductivity image from these two algorithms. As seen, the three pixels mixed much better in our multiple prior block update. This indicates that our multiple prior block update performs much better. The two algorithms are also compared in terms of the mixing property of the middle curve in the three representative pixels. As said in [3], the middle pixel is on the boundary of conductivity image, and transverse from low to high value during the MCMC run. Its mixing property is an extremely important indicator for the performance of MCMC algorithm.

In case of multiple prior block update, it took $10 \times 10 \times m$ updates for the auto-correlation of the blue pixel to go down to 10%; while it took *10* times more updates in case of single-site Metropolis. On the other hand, the CPU-time for one Multiple Prior block update is about 1.7 times more than that of single-site update. Overall, the multiple prior block update is about $10/1.7 \approx 6$ times more efficient than the single-site Metropolis.

4.2 Delayed Acceptance Metropolis-Hastings

Our second efficient MCMC algorithm is made within the Delayed Acceptance Metropolis-Hastings (DAMH) scheme [7]. As the surrogate transition method, the DAMH also uses an auxiliary probability distribution to make the proposal move. The auxiliary distribution in DAMH, however can be a state-dependent approximation of the target distribution. This makes DAMH distinctive.

Assume that $\pi(.)$ is the target distribution to be sampled. The Markov chain to be designed to converge to π is currently at the state x. Depending on x, $\pi_x^*(.)$ is an approximation of $\pi(.)$. With the terminology from [2], assume that $T_1(x, y)$ is an initial transition function for the Markov chain.

In Metropolis-Hastings (M-H) scheme [11], the proposal $y \neq x$ from $T_1(x, y)$ has to be tested, in order to be accepted as the state for the next step. The test involves an acceptance probability α and is: draw a

sample from $r \sim U(0,1)$. If $r < \alpha$, accept the proposal. Otherwise, reject the proposal and the current state remains to be the state of the next step.

In DAMH, the proposal y has to be subject to *two* steps of the above tests, in order to become the state for the next step. The two acceptance probabilities in the two tests are

$$\alpha_1 = \min\left\{1, \frac{T_1(y,x)\pi_x^*(y)}{T_1(x,y)\pi_x^*(x)}\right\}, \text{ and } \alpha_2 = \min\left\{1, \frac{T_2(y,x)\pi(y)}{T_2(x,y)\pi(x)}\right\}, \tag{13}$$

respectively. Here

$$T_2(x,y) = T_1(x,y)\min\left\{1, \frac{T_1(y,x)\pi_x^*(y)}{T_1(x,y)\pi_x^*(x)}\right\}, \quad y \neq x. \tag{14}$$

The actual transition function based on the initial transition function $T_1(x,y)$ is:

$$A(x,y) = T_1(x,y)\underbrace{\min\left\{1, \frac{T_1(y,x)\pi_x^*(y)}{T_1(x,y)\pi_x^*(x)}\right\}}_{\text{step}I}\underbrace{\min\left\{1, \frac{T_2(y,x)\pi(y)}{T_2(x,y)\pi(x)}\right\}}_{\text{step}II}, \ y \neq x. \tag{15}$$

Assume that

$$\frac{\pi_x^*(y)}{\pi_x^*(x)} \approx \frac{\pi(y)}{\pi(x)} \tag{16}$$

for all x and y that are close to each other. Then

$$\alpha_1 \approx \min\left\{1, \frac{T_1(y,x)\pi(y)}{T_1(x,y)\pi(x)}\right\}, \qquad \alpha_2 \approx 1. \tag{17}$$

Equation (17) indicates that the acceptance probability in the first step of DAMH is approximately the same as the that of M-H given the same initial transition function. The actual transition function in DAMH is effectively almost the same as that of M-H given the same initial transition function. In other words, DAMH produces almost the same amount of statistically independent samples as M-H given the same initial transition function.

When the target distribution $\pi(.) = \pi(.|u_E^{\text{noise}})$ in (2), a candidate of $\pi_x^*(.)$ is

$$\pi_x^*(z|u_E) \propto \exp\left\{\beta\sum_{i\sim j} v(z_i - z_j)\right\} I_{[2.5,4.5]^n}(z) \cdot \exp\left\{-\frac{1}{2\varsigma^2}\|u_E^{\text{noise}} - u_E^*(z)\|^2\right\}. \tag{18}$$

Here $u_E^*(z)$ is the local linearisation of the forward map at state x, as discussed in subsection 3.2. Apparently, (18) satisfies the assumptions (16). What's more, evaluating $\pi_x^*(y)/\pi_x^*(x)$ has negligible cost compared to that of $\pi(y)/\pi(x)$. When computing $\pi(y)/\pi(x)$ takes up almost all the CPU-time for each update in MH and DAMH, the average CPU-time for each DAMH update is α_1 times that of M-H given the same initial transition function. This is because $\pi(.)$ is only required to be evaluated α_1 times on average in each DAMH update; but every time in each M-H update. In DAMH, evaluating $\pi(.)$ only happens in computing α_2, which is only required when the proposal gets accepted in the first step. Recall that α_1 is the probability for the proposal to get accepted in the first step. Therefore, $\pi(.)$ is computed α_1 times on average in each DAMH update. Overall, the same initial transition function in DAMH is expected to be $1/\alpha_1$ times more efficient than in M-H.

In this study, DAMH with single site update was discussed for sampling the posterior distribution (2) of interest.

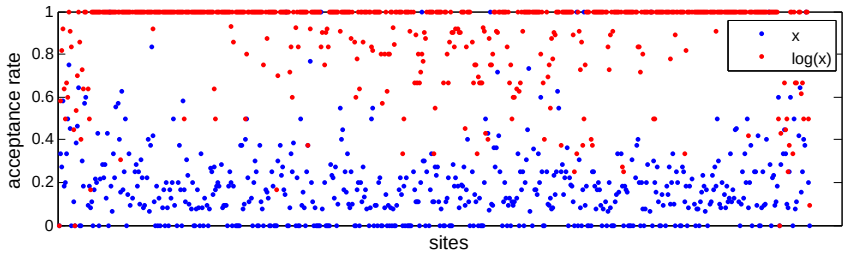


Figure 4: Empirical acceptance rates in the second step of DAMH with two approximations of forward maps.

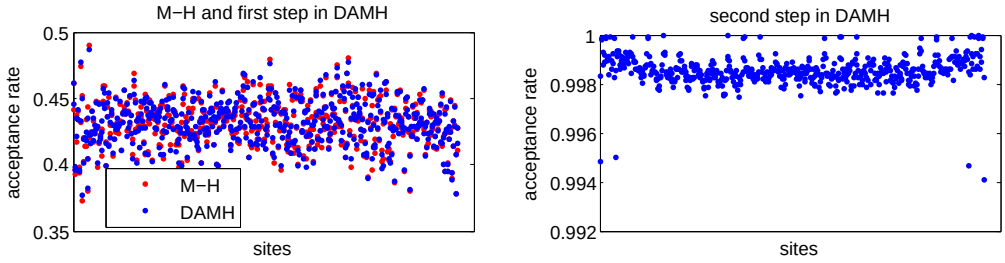


Figure 5: Empirical acceptance rate at every sites in DAMH with single-site update and single-site Metropolis.

First, our DAMH algorithm used two different approximations of forward map, as discussed in subsection 3.2. The performance of the DAMH with these two approximations was then compared. Figure 4 shows the empirical acceptance rate at the second step of DAMH for the two approximations of forward map. As seen, the second step of DAMH is close to one as expected in case of approximation with log-transformation; while it is well below one in case of approximation with ordinary local linearisation. Therefore, approximated forward map with log-transformation is a better approximation and enables the DAMH algorithm to perform better.

Second, DAMH with single-site update and single-site Metropolis (single site update in M-H) are compared on sampling the posterior distribution (2) of interest. Figure 5 shows the empirical acceptance rates at every sites for these two algorithms: As seen, the acceptance rates at the second step of DAMH are all close to 100%; while the ones in the first step of DAMH are all close to the acceptance rate of single-site Metropolis (single-site update in M-H). This finding indicates that the DAMH with single-site update has achieved the expected performance. In other words, single site update is about two times more efficient in DAMH than in M-H.

5 Conclusion

In section 3, we speed up the exact computation of forward map of EIT, by updating the Cholesky factor of the stiffness matrix [5]. As regard to the toy EIT model, our updating Choleksy factor method can be

as many as *five* times more efficient than the standard computation. We also make the approximation of forward map with local linearisation more accurate, by introducing a log-transformation into this process.

In section 4, we present two novel and efficient MCMC algorithms that perform much better than the single-site Metropolis for sampling a posterior distribution of our toy EIT model. One algorithm is made within the multiple prior update scheme. The surrogate chain in this scheme is updated by the constrained block update that leaves the mean value of state variable invariant. This multiple block update was found to be *six* times more efficient than the standard single-site Metropolis. The other MCMC algorithm is DAMH with single-site update. This algorithm was found *two* times more efficient than the standard single-site Metropolis.

As in [3], this study discusses MCMC sampling for a posterior distribution in the Bayesian approach to EIT. This study equipped single-site Metropolis with efficient forward map solver, on condition that the sparse system involved is solved by the direct method. We also present two novel and efficient MCMC algorithms that are more efficient than single-site Metropolis, which is considered hard to surpass [3].

On the other hand, the sparse system in the forward map computation of EIT is solved by the direct method in this paper. As EIT model gets larger and more practical, the resulting sparse system will get larger, and the iterative method will be more suitable. Therefore, it is important to study efficient computation forward map of EIT with iterative method for the sparse system. What's more, the proposal move in our efficient MCMC algorithms are still limited to a few components update. Future work should also focus on devising efficient MCMC algorithms with multi-variant update.

Acknowledgements

The author would like to thank Associate Professor Colin Fox for his support on this study. This work is also funded by NZIMA.

References

[1] C. Fox, G. Nicholls, *Sampling conductivity images via MCMC*, in *Proc. Leeds Annual Stat. Research Workshop* (1997), pp. 91–100.

[2] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer (2001).

[3] D. Higdon, C. S. Reese, J. D. Moulton, J. A. Vrugt, C. Fox, *Posterior exploration for computationally intensive forward models*, in *Handbook of Markov Chain Monte Carlo*, CHAPMAN & HALL/CRC (2011).

[4] A. Ern, J. Guermond, *Theory and Practice of Finite Elements*, Springer (2004).

[5] T. A. Davis, W. W. Hager, *Multiple-rank modifications of a sparse Cholesky factorization*, SIAM J. on Matrix Analysis and Applications, 22(4), (2001), 997–1013.

[6] G. K. Nicholls, C. Fox, *Prior modeling and posterior sampling in impedance imaging*, Proc. SPIE, 3459, (1998), 116–127.

[7] J. A. Christen, C. Fox, *MCMC using an approximation*, Journal of Computational and Graphical Statistics, 14(4), (2005), 795–810.

[8] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, *Equations of state calculations by fast computing machines*, Journal of Chemical Physics, 21, (1953), 1087–1091.

[9] M. Cheney, D. Isaacson, J. C. Newell, *Electrical impedance tomography*, SIAM Review, 41(1), (1999), 85–101.

[10] D. S. Holder, *Electrical Impedance Tomography:Methods, History and Applications*, Taylor & Francis (2004).

[11] W. K. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 57(1), (1970), 97–109.

[12] W. S. Cleveland, *Robust locally weighted regression and smoothing scatterplots.*, Journal of Computational Statistical Association, 74, (1979), 829–836.

[13] T. A. Davis, *CHOLMOD Version 1.0 User Guide*, Dept. of Computer and Information Science and Engineering, Univ. of Florida (2005).

[14] W. R. Gilks, S. Richardson, D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, CHAPMAN & HALL (1996).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

An Alternative Approach to Thermal Analysis Using Inverse Problems in Aluminum Alloy Welding

E. S. Magalhães, C. P. Silva, A. L. F. Lima e Silva, S. M. M. Lima e Silva*

Instituto de Engenharia Mecânica, Universidade Federal de Itajubá – UNIFEI, Itajubá, Minas Gerais, Brasil
e-mail: elisanmagalhaes@hotmail.com, gr9942@gmail.com, alfsilva@unifei.edu.br, metrevel@unifei.edu.br

Keywords: inverse problems, heat transfer, heat flux estimation, numerical and experimental methods.

Abstract
The determination of the temperature fields in a welded region has always been an obstacle to the improvement of welding processes. As an alternative, the use of inverse problems to determine the heat flux during the welding process allows a new analysis of the case. This article studies an alternative for the thermal analysis of the TIG (Tungsten Inert Gas) welding process on a 6065 T5 aluminum alloy. For this purpose, a C++ code was developed, based on a transient three-dimensional heat transfer model. To estimate the amount of heat delivered to the plate, the Specification Function technique was used. Moreover, an analysis of the duration time of the electrode in positive (t+) and negative (t-) polarities was carried out. The methodology was validated by accomplishing lab controlled experiments. The aluminum samples lay on four conical head screws and submitted to a heat flux on a surface by the TIG process torch. The torch displaces in the length of the sample thus simulating a real process. All the temperature sensors were fixed by capacitive discharge on the same surface and on the opposite heated surface. Type K thermocouples were used to obtain the temperatures by using HP 75000 B data acquisition system (DAS) connected to a computer. The numerical results presented low deviation when compared to the experimental results, which in turn confirms the validation of the methodology for the study of the welding process presented.

1 Introduction

In metal welding, it is very hard to directly measure the heat flux applied to the welded joint during the process as the heat flux supplied by the voltaic arch displays a very high thermal gradient. However, the calculation of the transferred heat is fundamental to determine the thermal efficiency of the process. Thus, the development of a methodology is necessary to determine the parameters that represent the optimal welding process, for example, what current and what speed the torch moves are necessary to control the temperature of the material. The analysis of the welded joints may be carried out for example, by studying the influence of the mechanical properties. In [1] the authors utilized an aeronautic aluminum alloy AA7075-T6 submitted to welding by friction. After the load tests, the authors observed the presence of small deflection in the micro structure and a uniform distribution of

MgZn2 precipitate particles which are highly important for the uniformity of the tension on the welded joint. Other welding processes have been studied concerning the structural analysis of the welding. Aluminum fusion was analysed by [2] in a laser welding process. For this purpose, the authors used several laser welding processes. In this analysis, the hybrid laser MAG (Metal Active Gas) was observed to have presented better results in relation to other processes. In [3] two aluminum plates AA6061-T6 were welded with the aim of analysing the effect of the GMAW - Gas Metal Arc Welding process on the mechanical properties of the aluminum. In this analysis, the differences between the theoretical and the experimental model were 8%. It was also observed that the welding behaviour is directly influenced by the amount of heat supplied to the base material in the diverse regions.

An alternative methodology for the analysis of the welding process is the resolution through inverse problem. One of the first articles using inverse problems in heat transfer was presented by [4], where a method, known today as Stolz method, was presented to estimate the heat flux applied to the surface of spheres during the tempera process having as main data, the temperature inside the samples. This method can also be extended to cylinders and plates. Stolz considered that the thermophysical properties were constant and that there was no internal heat generation, therefore, the problem could be treated as linear; thus Duhamel equations may be used for the elaboration of the method. This method presents good results for a great number of easily implemented. Based on the minimum square method and also on the Duhamel's model, [5] developed the sequential Specification Function method. This method presents stable results, not very influenced by experimental noises. In this method, the present heat flux is calculated by using previous fluxes. The posterior fluxes are considered to be null and a fictitious flux is assumed which may be constant, linear or quadratic under a certain number of future times. Another technique that may be used is the conjugated gradient method with adjoint equation [6]. This method is based on an optimization process with iterative regularization. Hence, the results of the minimization of the objective function tend to stabilize according to the number of iterations. This technique may also be used for the solution of linear and non-linear inverse problems as well as in parameter estimation problems [7]. The optimization method also called Broyden–Fletcher–Goldfarb–Shanno or BFGS is also used [8]. This is a non-linear optimization method obtained from the variation of the Newton method. From this method, [8] estimated the parameters of a continuous emission laser which aims to correct sight disorders.

[9] applied the optimization techniques and the enthalpy method to solve three-dimensional inverse problem in a TIG welding process. The authors estimated the heat flux based on the solution of a three-dimensional transient heat transfer with mobile source. The thermal field in the region of the plate or at any instant was determined from the estimation of the heat rate delivered to the piece. The direct problem was solved by the Finite Difference method with implicit formulation. For the solution of the algebraic equation system, the authors used the Successive Over Relaxation method (SOR) and the inverse problem resolution thorough the use of the Golden Section technique. The experimental temperatures were obtained from accessible points on the workpiece and the theoretical temperatures were calculated from three-dimensional thermal model. A non-linear inverse problem to estimate the heat in a welding process by friction [10]. The authors developed an algorithm, based on the conjugated gradient method and on the discrepancy principle. This methodology was used to estimate an unknown time dependent variable for the heat generation on the interface of the cylindrical bars during a welding process by friction. The temperature date were obtained from the direct problem and used to simulate the measurement of the temperature.

The inverse technique proposed in this work is based on a three-dimensional transient heat conduction model with moving heat sources. The Specification Function, which is an inverse problem technique, was implemented in a program to estimate the heat flux applied on the workpiece, from the experimental temperature records. The thermal field in any region of the plate or any time are determined from the estimation of the heat rate that is delivered to the workpiece. In addition, a new procedure is proposed to solve any experimental problem that appears in the aluminium welding application. In this kind of welding, when the voltaic arc is on, a very high level noise appears in the

experimental data signals if thermocouples are used to measure temperatures. In this case, practical procedures like electrical protection with Faraday Gail, use of electrical filters or grounding did not overcome the problem. In order to avoid this problem, this work presents a new experimental procedure: the temperature measurements just after the voltaic arc is cut off. The use of the measured temperatures together with the inverse technique allows the determination of the rate of heating necessary for welding as well as identification of the temperature field on the plate. The theoretical temperatures are calculated from the heat diffusion equation, which is solved numerically in three-dimensional Cartesian coordinates by using the Finite Volume Method for space and the implicit Euler method for time discretization. The temperature field can then be obtained through the solution of the heat diffusion equation. The thermal excitation was a moving heat source in directions *x* and *y*. The remaining surfaces are subject to convective and radiation heat losses.

2 Theoretical Development

2.1 Direct Problem

The thermal model presented in Fig. 1 may be described by the transient three-dimensional diffusion equation:

$$\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}=\frac{1}{\alpha}\frac{\partial T}{\partial t} \tag{1}$$

subject to the boundary conditions of convection and radiation

$$-\lambda\frac{\partial T}{\partial \eta_i}=h\left(T-T_\infty\right)+\sigma\cdot\varepsilon\left(T^4-T_\infty^4\right) \tag{2}$$

in the area defined by A_{xy}, is the following boundary condition

$$-\lambda\frac{\partial T}{\partial z}=q''(x,y) \tag{3}$$

and the initial condition

$$T(x,y,z,0)=T_0 \tag{4}$$

where T is the calculated temperature, i represents the number of flat surfaces (1 to 6), η_i is the normal surfaces, h is heat transfer coefficient by convection, λ the thermal conductivity, α the thermal diffusivity, σ the Stefan-Boltzmann constant, ε the emissivity, T_∞ the room temperature and q"(x,y) the unknown heat flux. The solution of the temperature field is obtained through the numerical approximation of Eq. (1) by using the Finite Volume method. The linear system of algebraic equations is solved by using SIP - Strongly Implicit Procedure.

Figure 1 presents a scheme model for the TIG welding process on an aluminum 6065 T5 sample. The heat flux q"(x,y), applied on the circular section area A_{xy} ,is obtained through the use of the Specification Function method [5]. The heat flux moves in constant speed, u (x,y), in direction s. The movement of the heat source may be written as:

$$q''_0(x,y)=q''(x,y)\times\left(\left(x-u(x)\cdot t\right)^2-\left(y-u(y)\cdot t\right)^2\right) \tag{5}$$

where u(x) and u(y) are the speed components, u (x,y) in the respective directions x and y and t is the time.

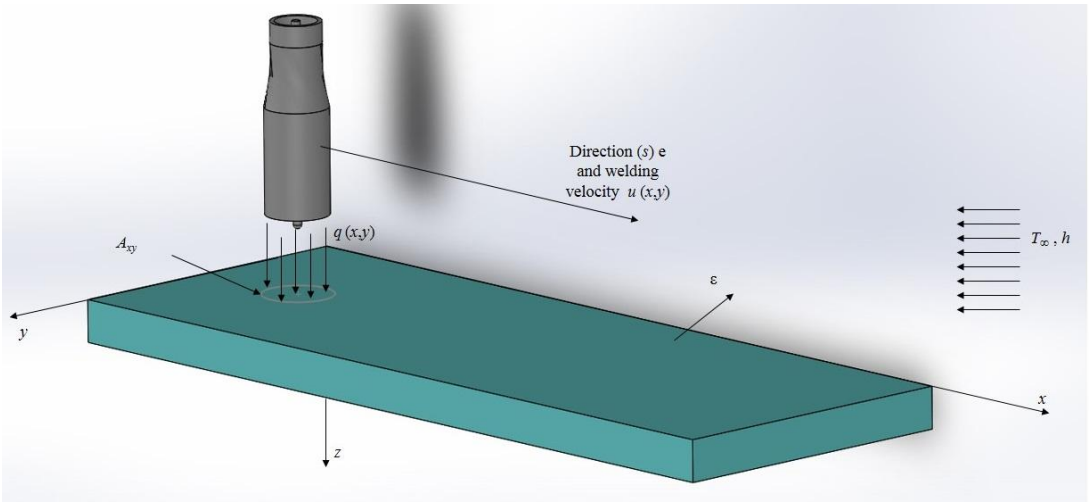


Figure 1. Three-dimensional representation of the welding process.

2.2 Objective Function

The application of the Specification Function requires an Objective Function to be minimized, defined by the square difference between the temperatures measured on the sample *Y*, and the temperatures calculated numerically *T*. Hence, it may be written as:

$$F = \sum_{p=1}^{r} \sum_{j=1}^{ns} \left(Y_{j,M+p-1} - T_{j,M+p-1} \right)^2 \tag{6}$$

where F is the Objective Function, j is the counter for the number of sensors, *nsens* represents the number of temperature sensors, p is the counter for future times steps and r is the number of future times used.

3 Experimental Procedure

3.1 Emissivity Determination

For the determination of the emissivity ε of aluminum 6065 T5, four controlled experiments were carried out in the Heat Transfer Laboratory (LabTC) at the Federal University of Itajubá (UNIFEI). The data acquisition Agilent 34980A, a digital power supply MCE 1051, an infrared thermometer Fluke 574, a 50x100x0.25 mm resistive Kapton heater, 2 aluminum 6065 T5 samples and thermocouples type K were used for these experiments. Figure 2 presents a picture of the experimental setting used for this experiment. The length of the full paper in its final layout should follow the rules given in Tab. 1.

To accomplish the experiment, the samples are heated on one surface by the resistive heater connected to the power supply MCE 1051. The surface is subjected to loss due to convection and radiation until they reach the permanent regime. The temperatures of the samples are measured through the use of the data acquisition controlled by a PC. When the steady state is reached, the emissivity is checked by comparing the temperature measured by the infrared sensor FLUKE 574 and the temperatures indicated by the acquisition. The temperature of the sensor is adjusted to the temperature of the sample

through the variation of the emissivity of the material. Table 1 presents the mean value for each experiment.

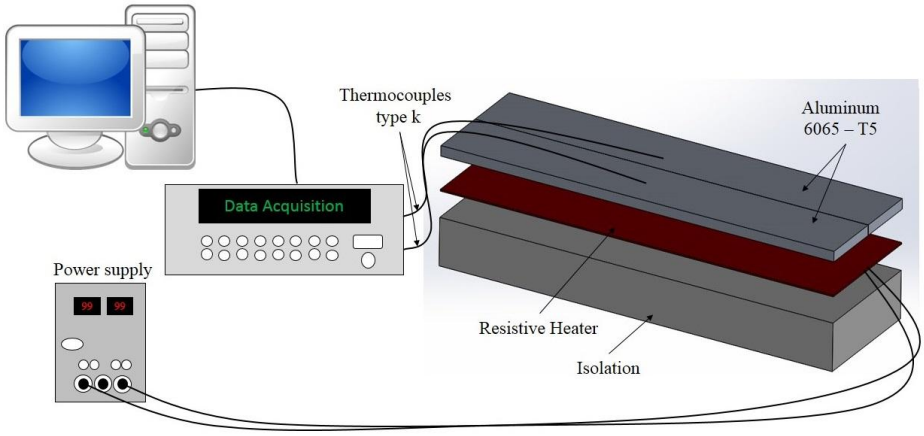


Figure 2: Experimental apparatus for the calculation of the emissivity

Table 1: Experimental results for the thermal emissivity of the aluminum 6065 T5.

Number of Experiments	**001**	**002**	**003**	**004**
ε	0.21 ± 0.02	0.22 ± 0.02	0.22 ± 0.02	0.20 ± 0.02

In all the calculations for the temperatures carried out afterwards, the average of all the values obtained experimentally for the emissivity was adopted, therefore, the mean calculated value is $\varepsilon = 0.21 \pm 0.02$.

3.2 Experimental Setting for the Real Welding Process

To accomplish this experiment, the following equipment were used: a data acquistion system HP 75000 Series B controlled by a PC, 4 thermocouples type K for the measurement of the experimental temperatures, twelve 0.25 x 0.038 x 0.0065 (m) aluminum 6065 T5 plates, a *xy* coordinate table, a counter to attach the aluminum plate to the coordinate table, an automated system to move the welding torch, an AC power supply and an acquisition system to measure the voltage and the electric current supplied for the welding. The thermocouples type K (Chromel-Alumel) are attached to the metal plate with the use of the capacitive discharge [11]. They must be positioned in the specific regions of the sample so that the voltaic arch does not interfere in the measuring of the temperature. The aluminum plate, the thermocouples, the counter and the TIG welding torch are presented in Fig. 3a. It should be highlighted that the aluminum plate must be attached to the coordinate table through the possible smallest area, so that the boundary conditions may be considered as free convection and radiation on all the surfaces. Figure 3b shows the details of the plate attached by four conical head screws, covered with a layer of thermal isolator which minimizes the contact area between the screw and the plate and reduces fin effect during the heat transfer process. It is worth mentioning that the layer of thermal isolator is applied on only three screws. The fourth screws is responsible for the grounding of the plate submitted to the welding process. Limiting lines may be verified on the sample with the purpose of determining the beginning and the end of the welding process.
Four *t*+ experimental conditions were tested and for each experiment 482 temperature points were collected. The velocity of the torch was of 62.5 mm/min. In the TIG welding process at AC, the time

the electrode remains in polarity *t+* is responsible for the cathodic cleaning of the weld bead due to the electric field emission from the test-plate to the tungsten electrode [12]. It can be observed that the oxide layer reminiscent is isolate and refractory what turns difficult the welding in the negative polarity (t-). Despite the cleaner qualities, the positive polarity has lesser efficiency and useful life than the negative polarity. The other aim of this work is to contribute to study the influence of the polarity in the temperature field. In order to research this influence four different welding conditions were analyzed (Tab. 2). It shows the power generation as the voltage increases. Table 3 presents the positions for each thermocouple, the coordinate system is the same as in Fig. 1.

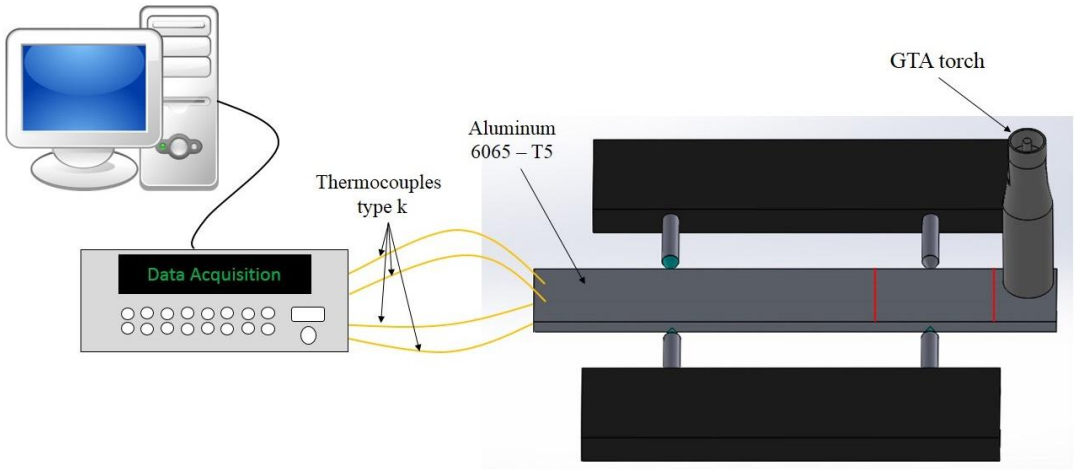


Figure 3: Experimental assembly with the position of the thermocouples and the marks indicating the beginning and end of the welding process on the aluminum plate.

Table 2: Welding Conditions.

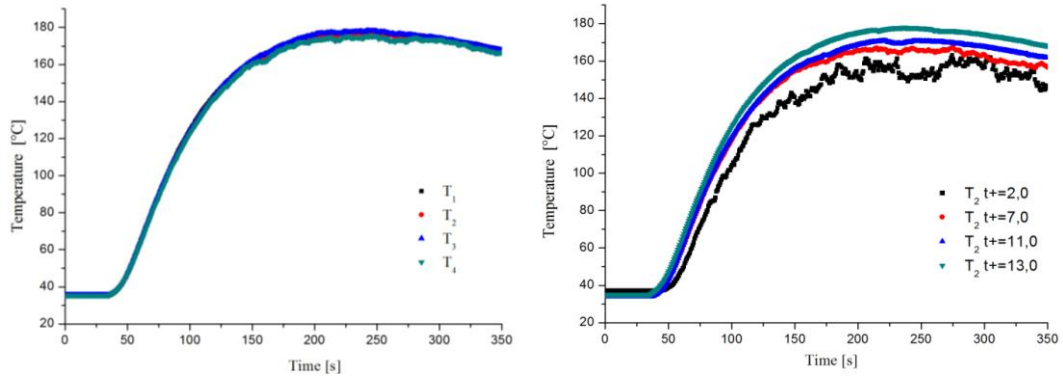
***(t+) adjusted* (*ms*)**	***(t-) adjusted* (*ms*)**	***Welding speed* (*u*) (*mm/min*)**	***Current* (*A*)**	***Voltage* (*V*)**	***Power Generated* (*W*)**
2	20	62.5	194	10.5	2037
7	20	62.5	193	11.3	2181
11	20	62.5	193	12.1	2335
13	20	62.5	193	14.0	2702

Table 3: Position of the thermocouple on the aluminum sample.

Thermocouple	**1**	**2**	**3**	**4**
x [mm]	238.0	245.0	237.0	244.0
y [mm]	33.00	11.00	25.00	11.00
z [mm]	6.500	6.500	0.000	0.000

4 Result Analysis

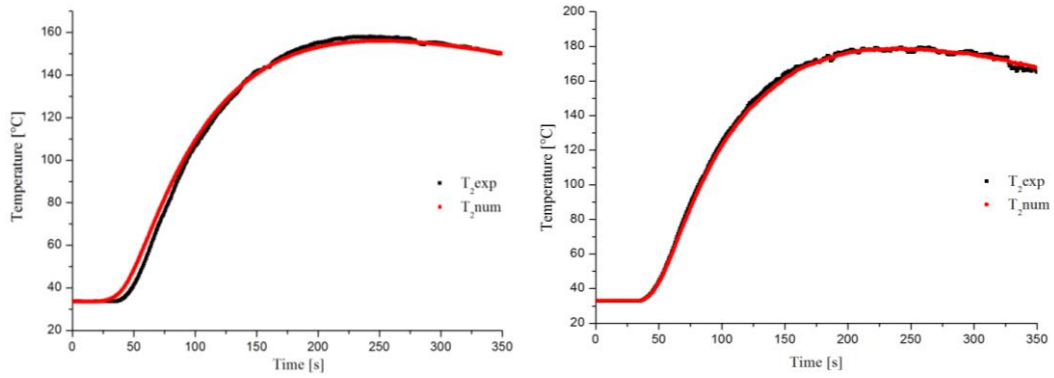
For each condition of $t+$, three experiments were carried with the aim of assessing the repeatability of the estimated heat flux results. As mentioned before, for each experiment 482 temperature points were observed at a time interval, Δt, of 0.66 s. Figure 4a illustrates the temperature signals measured by the thermocouples T_1, T_2, T_3 and T_4 for the welding the forth condition presented on Tab. 2 ($t+$ = 13ms). High temperature measured by the thermocouple after the voltaic arch was turned off may be observed. Figure 4b presents a comparison of the temperature measured in the four conditions for positive polarity ($t+$) with the thermocouple in position 2 (Tab. 3). An increase in the temperature may be observed with the increase of positive polarity ($t+$). It may also be seen that for values of $t+$ greater than 11, there is no increase in the temperature signal. This may be explained by the fact that for high values of $t+$, the greatest part of the heat generated remains on the electrode, which is undesirable for both the efficiency of the process and the electrode lifespan [12].



a) b)

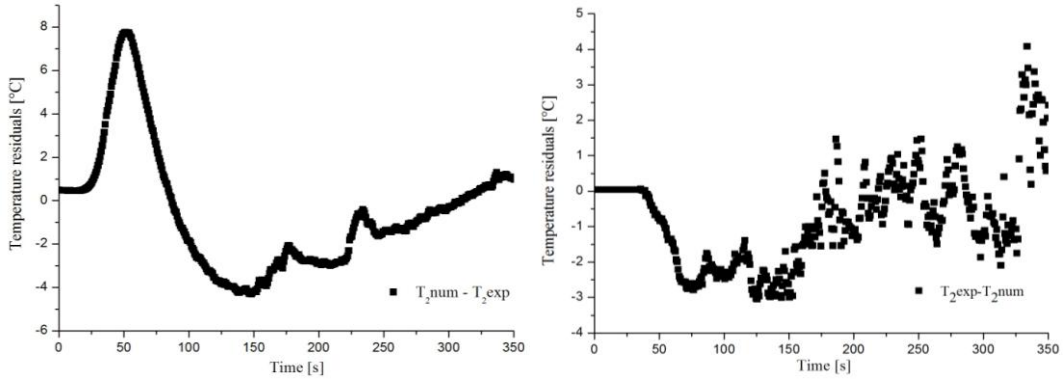
Figure 4: a) Evolution of the experimental temperature for condition of $t+$ = 13 ms and b) Comparison of the experimental temperature in position 2 for the four conditions of positive polarity (t+).

Thermal properties values of AA 6065-T5 were considered constants for the numerical solution: thermal conductivity 209 W/mK and thermal diffusivity 86.2 x 10^{-6} m^2/s [13]. Both the heat transfer coefficient by convection was also considered constant, h=20 W/m^2K for all the surfaces and the emissivity of the material obtained experimentally, ε = 0.21 were considered constant. The inverse problem technique utilized was Specification Function with multiple sensors. The adjustment of the multiple sensors to the technique allows a better determination of the heat flux supplied to the surface. This technique aims to minimize the Objective Function (Eq. 6). Another factor to consider in the simulation process is the total welding time which includes the beginning (static source and arch on), the middle (mobile source and arch on) and the end (static source e arch off). A comparison between the calculated and experimental temperatures for position 2 (Tab. 3) and the welding conditions $t+$ = 11 ms and $t+$ =13 ms are presented in Figs. 5a and 5b, respectively. The respective temperature residuals of Figs. 5a and 5b are shown in Figs. 6a and 6b. The heat fluxes estimated by the Specification Function technique for these two welding conditions are represented in Figs. 7.a and 7.b. By analyzing the temperature residuals of Figs. 6a e 6b, it may be observed that the software developed for this work presented satisfactory results. For these two cases, the maximum residual was approximately -8°C, which represents an error lesser than 5.0%. Several factors may be accounted for these differences. For example, the hypothesis of constant thermal properties and convection heat transfer coefficient. It may also be observed that the use of the discrete model, based on space and time for the movement of the heat source is appropriate for the analysis of the thermal phenomenon.



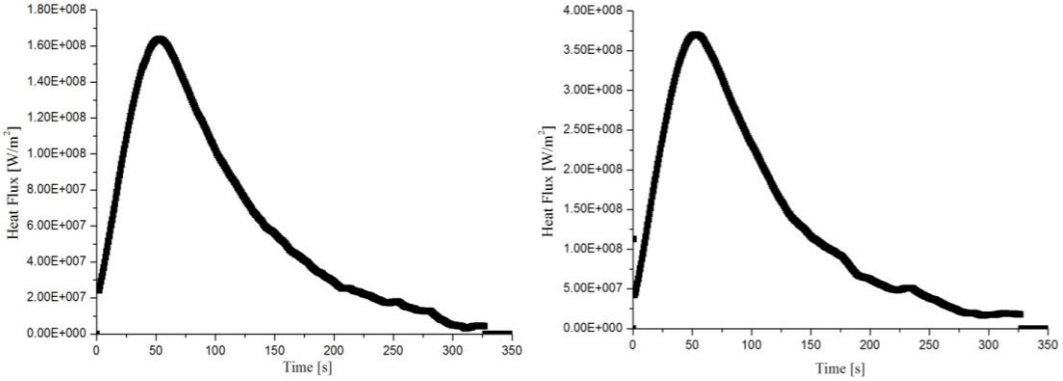
a) b)

Figure 5: Comparison of measured and calculated temperatures for a) $t+$ = 11 ms and b) $t+$ = 13 ms.



a) b)

Figure 6: Analysis of the residuals for a) $t+$ = 11 ms and b) $t+$ = 13 ms.



a) b)

Figure 7. Estimated heat flux for a) $t+$ = 11 ms and b) $t+$ = 13 ms

Figure 8 presents the temperature fields for $t+$ = 13 ms. It may be seen that the temperature gradient reaches its top at the point of the torch displacement. By analyzing Tab. 4, it may be observed that there is a power increase with the increase of $t+$. This occurs basically due to the voltage increase which goes from 10.5 V ($t+$ = 2.2 ms) to 14.0 V ($t+$ = 13.3 ms). This represents an increase of 33.0% in the voltage and almost the same percentage, 32.6%, occurs in the total power.

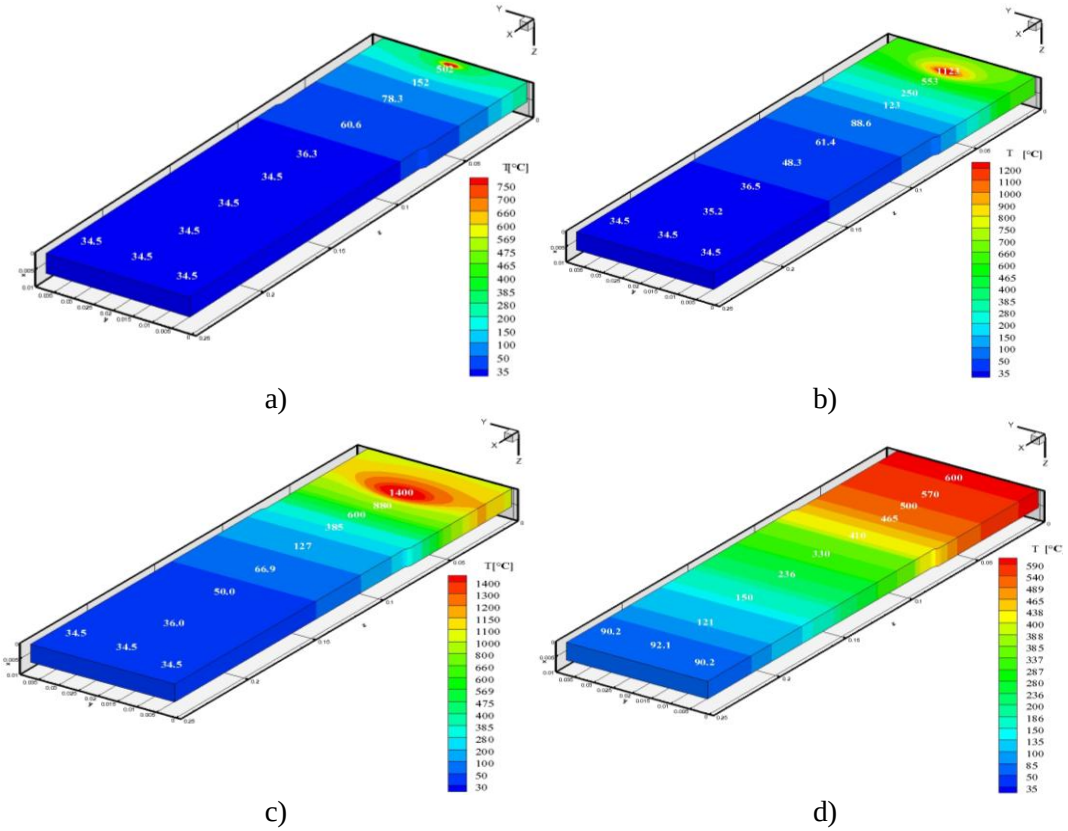


Figure 8: Evolution of the temperature field at instants: a) 6.6 s, b) 16.5 s, c) 26.4 s and d) 66 s.

Table 4: Average heat rate for each test.

Adjusted (t+) (ms)	Obtained (t+) (ms)	Total power (W)	Average useful power (W)	Thermal efficiency (%)
2	2.2	2037	914.3	44.9
7	6.9	2181	983.5	43.0
11	11.3	2335	1073.4	46.0
13	13.3	2702	1105.3	40.9

5 Conclusions

This paper presented a new methodology to determine the heat transfer rate on a aluminum 6065-T5 plate, under a TIG welding process, from the observation of the thermal fields developed after the welding torch is turned off. This alternative procedure has enabled the measurement of the temperature, used in the inverse problem solution, just after the voltaic arc is cut off. This has eliminated one of the greatest obstacles to measure the experimental temperature during the welding of aluminum plates: electrical noises coming from welding in AC. The applied heat flux on the surface was estimated by the Specification Function. The similarity between the experimental and calculated parameters validated and proved the

efficiency of the software developed when applied to the solution of welding thermal problems.

Acknowledgements

The authors would like to thank CNPq, FAPEMIG and CAPES for their financial support, in special Process APQ-1940-11 sponsored by FAPEMIG and Process 470971/2011-8 sponsored by CNPq.

References

[1] S. Rajakumar, C. Muralidaharan, V. Balasubramanian, *Influence of friction stir welding process and tool parameters on strength properties of AA7075-T6 aluminum alloy joints*, Materials and Design, 32, (2010), pp. 535-549.

[2] R. Quintino, R. Miranda, U. Dilthey, D. Iordachescu, M. Banasik, S. Stano, S.; *Laser welding of structural aluminium*, Adv. Struct. Mater., 8, (2011), pp. 33-57.

[3] J. A. Vargas, J. E. Torres, J. A. Pacheco, R. J. Hernandez, *Analysis of heat input effect on the mechanical properties of Al-6061-T6 alloy weld joints*, Materials and Design, 52, (2013), pp. 556-564.

[4] G. Stolz Jr., Numerical solution to an inverse problem of heat conduction for simple shapes, J. Heat Transfer, 82, (1960), pp. 20-26.

[5] J. V. Beck, B. Blackwell, C. St. Clair, *Inverse Heat Conduction: Ill-posed Problems*, Wiley-Interscience Publication, New York, (1985).

[6] O. M. Alifanov, *Solution of an inverse problem of heat conduction by iterations methods*, Journal of Engineering Physics, 26, (1974), pp. 471-476.

[7] K. J. Dowding, B. E. Blackwell, *Sensitivity analysis for nonlinear heat conduction*," J. Heat Transfer, 123, (2001), pp.1-10.

[8] M. V. C. Solza, *Otimização em Termos Fontes em Problemas de Biotransferência de Calor*, (in portuguese) Dissertação de Mestrado, Instituto Militar de Engenharia, Rio de Janeiro, Brasil, 2009).

[9] C. V. Gonçalves, S. R. Carvalho, G. Guimarães, *Application of optimization techniques and the enthalpy method to solve a 3D-inverse problem during a TIG welding process*, Applied Thermal Engineering, 30, (2010), pp 2396-2402.

[10] Y.-C. Yang, W.-L. Chen, H.-L. Lee, *A nonlinear inverse problem in estimating the heat generation in rotary friction welding*, Numerical Heat Transfer, Part A, 59, (2011), pp. 130-149.

[11] S. M. M. Lima e Silva, V. L. Borges, L. O. Vilarinho, A. Scotti, G. Guimarães, *Desenvolvimento de uma técnica experimental para a determinação do fluxo de calor em um processo de soldagem TIG*, 9th Brazilian Congress of Thermal Sciences and Engineering-ENCIT 2002, CDROM, Caxambu, Minas Gerais, Brasil, (2002).

[12] A. B. Vieira Jr., A. Scotti, *Influência das condições superficiais na estabilidade do arco em soldagem TIG de alumínio*, XXV Encontro Nacional de Tecnologia da Soldagem, Belo Horizonte, MG, Brasil, (1999).

[13] MatWeb, Aluminium 6065-T5, Available at: www.matweb.com, Access date: January 06, (2014).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

ICIPE 2014 Regularization using Truncated Singular Value Decomposition for estimating the Fourier spectrum of a noised space distribution over an extended support

D. Maillet*, S. André, B. Rémy, W. Al Hadad

LEMTA, University of Lorraine and CNRS, - ENSEM, 2 avenue de la Forêt de Haye - TSA60604 - 54518 Vandœuvre-lès-Nancy cedex - France
e-mail: denis.maillet@univ-lorraine.fr

Key words: integral transforms, thermal quadrupoles, heat transfer in minichannel, inverse heat conduction and convection

Abstract

This paper is devoted to a theoretical and numerical study of the different ways of calculating the Fourier transform of a noisy signal where the boundary conditions at the lateral boundaries of the measurement interval are not precisely known. This happens in different characterization problems where an infrared camera is used for temperature measurements. In order to overcome this difficulty, the interval where the Fourier transform is defined (its support) is supposed to be larger than the measurement domain. So this virtual interval larger than the measurement interval is thus used. We show regularization by Truncated Singular Value Decomposition is able to yield good estimates of this very ill-posed inverse problem.

Introduction

We consider here the problem of reconstructing the Fourier spectrum of N_x noisy discrete temperature measurements θ_i made on a solid surface, that is for discrete abscissa $x_i \in \left]-\ell\,;\,\ell\right]$, for $i = 1$ to N_x, in a 2D heat transfer case. The $N_h = 2\,n_h$ harmonics of this spectrum are :

$$\tilde{\theta}_n = \int_{-L}^{L} \theta(x)\,\exp\left(-\mathrm{i}\,\alpha_n\,x\right)\,\mathrm{d}x \quad \text{with } \mathrm{i}^2 = -1 \;,\; L \geq \ell$$
$$\text{and } \alpha_n = n\pi / L \text{ (for } n = -n_h + 1, -n_h, \cdots, -0, 1, 2, \cdots, n_h - 1,\ n_h) \tag{1}$$

Estimation of this spectrum is required in several thermal characterization experiments where infrared thermography (IR) is used, such as thermal diffusivity measurement of a composite flat plate made of an anisotropic material with a front face flash excitation and rear face IR temperature measurement [1] or for estimating the heat fluxes at the different interfaces of a minichannel heated locally over its front face (either steady state or transient heating), using front or rear face IR temperature measurement [2].

This stems from the fact that analytical solution of this type of heat transfer problem, can be obtained very easily using the Fourier integral transform over a finite space domain, for example through the

Thermal Quadrupoles technique [3]. Inversion of the corresponding model and experimental temperature distribution, written both in the Fourier domain, can be applied either to estimate thermophysical parameters of a sample (a parameter estimation problem) [1] or to recover temperature and fluxes at the different interfaces as well as the bulk temperature distribution of the fluid flow (inverse function estimation problem based on inverse heat conduction/convection) [2].

The main difficulty in this type of configuration is that the experimental boundary conditions at $x = \pm \ell$ are generally not precisely known: heat flux in the *x* direction by both natural convection and radiation can occur in the anisotropic diffusivity characterization problem [1] and the channel wall length is larger than the measurement length 2ℓ, which means that neither temperature nor heat flux are equal to zero at the two boundaries, especially at the downstream one if it is too close to the heated region ($x = \ell$ if the fluid flows in the x direction) in the channel thermal characterization problem [2].

So, if eigenvalues $\alpha'_n = n\pi / \ell$, that correspond to the zero temperature or flux boundary conditions are chosen, the solution of the direct problem may be biased.

This is why we have decided to define the eigenvalues $\alpha_n = n\pi / L$ over a larger interval $]-L\,;\,L]$, called here a 'virtual' interval, wider than the measurement interval $]-\ell\,;\,\ell]$, see equation (1), where $k_v = L/\ell \geq 1$ is a natural integer. If k_v becomes large enough, the virtual boundaries $x = \pm L$ are far enough from the heat source, which lays inside the $]-\ell\,;\,\ell]$ measurement interval, and the zero temperature or flux boundary conditions become valid and the α_n eigenvalues become exact.

The following part of this article is divided in two parts: in section 2, the spectrum estimation problem is studied without the notion of virtual length, that is for $k_v = 1$ and the virtual length case is studied in section 3, for $k_v = 2$. In both sections a simple theoretical temperature distribution $\theta(x)$ for the x_i points, that is a door function, with zero temperature at $x = \pm \ell$, which is later corrupted by a synthetic noise, is considered. In another work [4], the same technique has been applied on experimental IR temperature measurements for a flat minichannel in a transient heating case [4].

1 Estimation of the Fourier spectrum without any virtual length ($k_v = 1$)

We consider here the following symmetrical door function plotted in figure 1:

$$\theta(x) = \theta_{max}\left(H\,(x+\ell_h/2) - H\,(x-\ell_h/2)\right) \text{ for } x \in]-\ell\,;\,\ell] \tag{2}$$

where H is the Heaviside function with $\ell = 32.5$ mm, $\ell_h = 12$ mm and $\theta_{max} = 2$ °C. The exact analytical spectrum of this function is:

$$\tilde{\theta}_n^{exact} = \frac{2}{\alpha_n}\,\theta_{max}\,\sin(\alpha_n \ell_h/2) \quad \text{if} \quad n \neq 0 \quad \text{and} \quad \tilde{\theta}_0^{exact} = \ell_h\,\theta_{max} \tag{3}$$

An identically independently distributed noise ε_i of standard deviation $\sigma = 0.08$°C is added at each location $x_i = -\ell + i\,\Delta x$, for $i = 1$ to $N_x = 200$, where $\Delta x = 2\ell / N_x$:

$$\theta_i^{noised} = \theta(x_i) + \varepsilon_i \tag{4a}$$

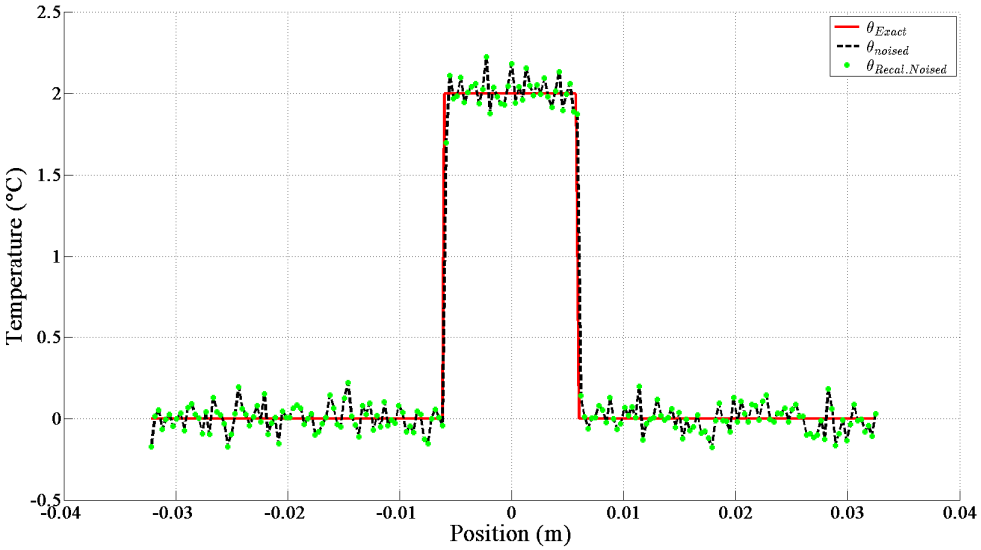


Fig. 1 – Exact, noised and reconstructed temperature profile using ordinary least squares (k_v = 1)

This generates a synthetic temperature measurement vector, of size N_x x 1:

$$\boldsymbol{\theta}^{noised} = \boldsymbol{\theta} + \boldsymbol{\varepsilon} \text{ where } \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0} \text{ and } \mathrm{cov}(\boldsymbol{\varepsilon}) = \sigma^2 \boldsymbol{I}_{Nx} \tag{4b}$$

where E () is the expectancy of a random column vector, cov() its variance-covariance matrix and and $\boldsymbol{I}_{Nx}$ the identity matrix of size N_x.

The N_h harmonics of the temperature profile (3) can be gathered in a $\tilde{\boldsymbol{\theta}}^{exact}$ column vector of size N_h x 1.

Once the synthetic (pseudo-experimental) temperature profile $\boldsymbol{\theta}^{noised}$ known, the unknown spectrum $\tilde{\boldsymbol{\theta}}$ can be calculated in three different ways:

1.1 Numerical quadrature

This technique is the simplest one, it consists in calculating an approximation of the definition (1) of each harmonics $\tilde{\theta}_n$ through a numerical integration of the noised signal [5] :

$$\tilde{\theta}_n \approx \Delta x \sum_{i=1}^{N_x} \theta_i^{noised} \exp(-\mathrm{i}\alpha_n x_i) = \sum_{i=1}^{N_x} G_{ni}\, \theta_i^{noised} \text{ with } G_{ni} = \exp(-\mathrm{i}\alpha_n x_i)\Delta x \tag{5}$$

Since only N_x data θ_i^{noised} are available, a number $N_h = 2\, n_h \le N_x$ of harmonics can be calculated.

This technique can provide good approximations of the harmonics of low order (low values of $|n|$) but its precision decreases for high space frequencies, because of the presence of noise in the signal.

1.2 Rectangular estimation

Instead of using the approximation of an integral of a noised signal (5), it is better to consider estimation of the spectrum as an inverse problem. So, we can start from an exact model, that is the definition of the inverse Fourier transform, for an exact output signal $\boldsymbol{\theta}$ depending on a limited number of harmonics $N_h = 2\, n_h = N_x$:

$$\theta(x_i) = \theta_i = \frac{1}{\ell} \sum_{n=-n_h+1}^{n_h} \tilde{\theta}_n \exp(+\mathrm{i}\alpha_n x_i) = \sum_{n=-n_h+1}^{n_h} S_{in}\, \tilde{\theta}_n \quad \text{with } S_{in} = \frac{1}{\ell} \exp(+\mathrm{i}\alpha_n x_i) \qquad (6)$$

This equation can be put under a matrix/column vector form, if the subscripts of the $\tilde{\theta}_n$ harmonics are increased by a simple translation equal to n_h, in order not to have any negative index in the components of the spectrum vector $\tilde{\boldsymbol{\theta}}$, that is $[\tilde{\boldsymbol{\theta}}]_k = \tilde{\theta}_{k-n_h}$. This spectrum vector, a parameter vector to be estimated, as well as the corresponding model are:

$$\tilde{\boldsymbol{\theta}} = [\tilde{\theta}_{-n_h+1} \quad \tilde{\theta}_{-n_h} \quad \cdots \quad 0 \quad \cdots \quad \tilde{\theta}_{n_h-1} \quad \tilde{\theta}_{n_h}]^T \text{ and } \quad \boldsymbol{\theta} = \boldsymbol{S}\, \tilde{\boldsymbol{\theta}} \qquad (7)$$

where the coefficients of the square matrix $\boldsymbol{S}$ of size N_x are $[\boldsymbol{S}]_{ik} = S_{i,k-n_h}$.
So, the ordinary least square solution is:

$$\hat{\tilde{\boldsymbol{\theta}}}_{OLS}^{square} = \arg(\min(J(\tilde{\boldsymbol{\theta}})) = \boldsymbol{S}^{-1}\boldsymbol{\theta}^{noised} \quad \text{where } J(\tilde{\boldsymbol{\theta}}) = \left\| \boldsymbol{r}(\tilde{\boldsymbol{\theta}}) \right\|^2 \quad \text{with } \boldsymbol{r}(\tilde{\boldsymbol{\theta}}) = \boldsymbol{\theta}^{noised} - \boldsymbol{S}\tilde{\boldsymbol{\theta}} \qquad (8)$$

Here the problem is square, which means that the residual vector $\boldsymbol{r}(\hat{\tilde{\boldsymbol{\theta}}}_{OLS-square})$ is equal to zero.

The inverse problem is here very well posed, since the condition number of matrix $\boldsymbol{S}$ is very close to unity. The recalculated signal $\boldsymbol{\theta}^{recalc} = \boldsymbol{S}\, \tilde{\boldsymbol{\theta}}_{OLS}^{square}$ is plotted together with the exact signal $\boldsymbol{\theta}$ and with the noised signal $\boldsymbol{\theta}^{noised}$ in figure 1: one can see that the fit is perfect.

The spectral energy density of the exact temperature profile, as well as its estimated value that stems from the ordinary least square estimation (8) deduced from the noised temperature profile, are plotted in figure 2: the estimation is very good, in spite of the noise in the data.

However, for different locations of the measurement points it may change and a natural solution would be to estimate an even number of harmonics α lower than the number of measurements N_x. In this case matrix $\boldsymbol{S}$ would be replaced by a rectangular matrix $\boldsymbol{S}_\alpha$, with a size N_x x α (the $\frac{N_x-\alpha}{2}$ first column as well as the $\frac{N_x-\alpha}{2}$ last columns of $\boldsymbol{S}$ have been removed) and the ordinary least square solution would be:

$$\hat{\tilde{\boldsymbol{\theta}}}_{OLS,\alpha}^{rect} = \arg\left(\min\left(J(\tilde{\boldsymbol{\theta}}_\alpha)\right)\right) = (\boldsymbol{S}_\alpha^* \boldsymbol{S}_\alpha)^T \boldsymbol{S}_\alpha^* \boldsymbol{\theta}^{noised} \quad \text{where} \quad J(\tilde{\boldsymbol{\theta}}_\alpha) = \left\| \boldsymbol{r}(\tilde{\boldsymbol{\theta}}_\alpha) \right\|^2 \quad \text{with} \quad \boldsymbol{r}(\tilde{\boldsymbol{\theta}}) = \boldsymbol{\theta}^{noised} - \boldsymbol{S}_\alpha^* \tilde{\boldsymbol{\theta}}_\alpha \tag{9}$$

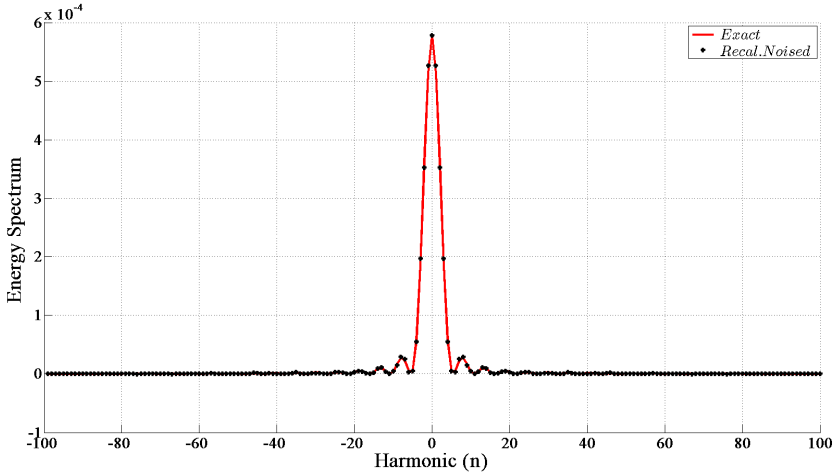


Fig. 2 – Spectral energy densities of temperature profile: exact ($\tilde{\boldsymbol{\theta}}_n^* \tilde{\boldsymbol{\theta}}_n$) and estimated ($\hat{\tilde{\boldsymbol{\theta}}}_{OLS}^{*square} \hat{\tilde{\boldsymbol{\theta}}}_{OLS}^{square}$)

Matrix $\boldsymbol{S}_\alpha^*$ is the adjoint of complex matrix $\boldsymbol{S}_\alpha$, that is the transpose of its conjugate ($\boldsymbol{S}_\alpha^* = \overline{\boldsymbol{S}}_\alpha^T$).

1.3 Truncated Singular Value estimation

Instead of trying to reduce the number of unknowns, the $\alpha \le N_x$ harmonics corresponding to the rectangular sensitivity matrix $\boldsymbol{S}_\alpha$ used in equation (9), it is possible to keep the total number of harmonics to be estimated equal to N_x, using regularization by Truncated Singular Value Decomposition (TSVD) [6]. This regularization technique is based on the square matrix $\boldsymbol{S}$ whose Singular Value Decomposition (SVD) is:

$$\boldsymbol{S} = \boldsymbol{U}\,\boldsymbol{W}\,\boldsymbol{V}^T \text{ with } \boldsymbol{U}^*\boldsymbol{U} = \boldsymbol{U}\,\boldsymbol{U}^* = \boldsymbol{V}^*\boldsymbol{V} = \boldsymbol{V}^*\boldsymbol{V} = \boldsymbol{I}_{N_x} \;;\; \text{with } \boldsymbol{W} = \operatorname{diag}(w_1, w_2, \cdots, w_{N_x})$$
$$\text{where } w_1 \ge w_2 \ge \cdots w_{Nx-1}, \ge w_{Nx} \ge 0 \;;\; \boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_1 & \boldsymbol{U}_2 & \cdots & \boldsymbol{U}_{Nx} \end{bmatrix} \;;\; \boldsymbol{V} = \begin{bmatrix} \boldsymbol{V}_1 & \boldsymbol{V}_2 & \cdots & \boldsymbol{V}_{Nx} \end{bmatrix} \tag{10}$$

and where the w_k, $\boldsymbol{U}_k$ and $\boldsymbol{V}_k$ are the k^{th} singular value, the k^{th} left singular vector and the k^{th} right singular vector respectively. The ordinary least estimator (8) can also be written:

$$\hat{\tilde{\boldsymbol{\theta}}}_{OLS}^{square} = \arg\left(\min\left(J(\tilde{\boldsymbol{\theta}})\right)\right) = \boldsymbol{V}\,\boldsymbol{W}^{-1}\,\boldsymbol{U}^*\,\boldsymbol{\theta}^{noised} \tag{11}$$

The truncated version of this estimator is:

$$\hat{\tilde{\boldsymbol{\theta}}}_\alpha^{TSVD} = \boldsymbol{V}\,\boldsymbol{W}_\alpha^{-1}\,\boldsymbol{U}^*\,\boldsymbol{\theta}_h^{exp}(t_k) \qquad \text{with } \boldsymbol{W}_\alpha^{-1} = \operatorname{diag}(w_1^{-1}, w_2^{-1}, \cdots, w_\alpha^{-1} \;\; 0 \;\; \cdots \;\; 0) \tag{12}$$

The interest of this estimator is to keep the number N_x of estimated unchanged, while using a number of internal degrees of freedom α (the number of the inverse of the singular values w_k^{-1} different from zero in (12)) lower than N_x, with a decrease of the dispersion of the estimates since:

$$\text{Trace}\,(\text{cov}(\hat{\tilde{\theta}}_{\alpha}^{TSVD})) = \sum_{n=1}^{N_x} \text{var}(\hat{\tilde{\theta}}_{n,\alpha}^{TSVD}) = \sigma^2 \sum_{k=1}^{N_x} \frac{1}{w_k^2} \leq \text{Trace}\,(\text{cov}(\hat{\tilde{\theta}}_{n}^{TSVD} = \hat{\tilde{\theta}}_{OLS}^{square})) \qquad (13)$$

where cov () is the varianc-covariance matrix of a column vector.

The mean square residuals $r_{rms} = \frac{1}{\sqrt{N_x}} \left\| \boldsymbol{r}\,(\tilde{\theta}) \right\|$ has been plotted as a function of α, for the rectangular estimation ($\tilde{\theta} = \hat{\tilde{\theta}}_{OLS,\alpha}^{rect}$, for an even value of α) and for the TSVD estimation ($\tilde{\theta} = \hat{\tilde{\theta}}_{\alpha}^{TSVD}$) in figure 3.

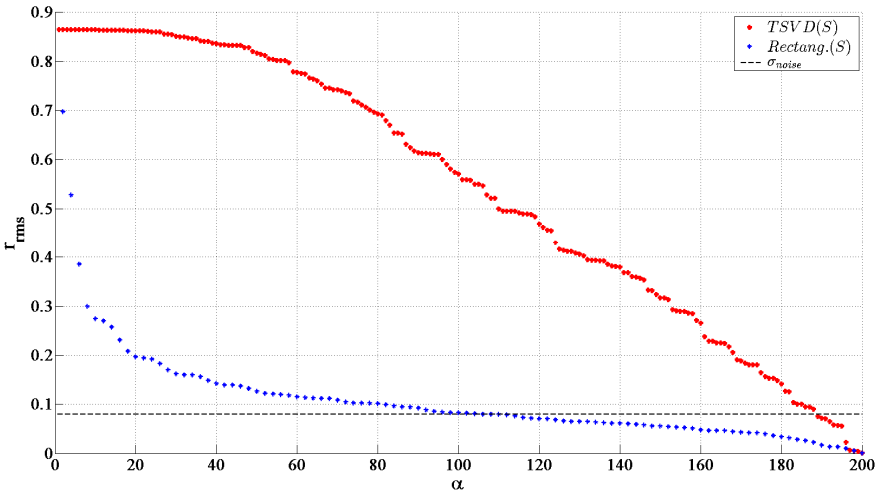


Fig. 3 – Root mean square residuals for rectangular and TSVD estimation (k_v = 1)

One can notice in this figure that the residuals corresponding to the rectangular estimation are lower than the TSVD ones.

The root mean square of the errors of the estimates $r_{rms} = \frac{1}{\sqrt{N_x}} \left\| \hat{\tilde{\theta}} \right\|$ are plotted as a function of α, for the rectangular estimation ($\tilde{\theta} = \hat{\tilde{\theta}}_{OLS,\alpha}^{rect}$, for an even value of α, where the $N_x - \alpha$ lacking harmonics have been given a zero value) and for the TSVD estimation ($\tilde{\theta} = \hat{\tilde{\theta}}_{\alpha}^{TSVD}$) in figure 4.

The error is lower for rectangular estimation and it is clear that regularization is not needed here, since the optimum value of α is $\alpha = N_x$.

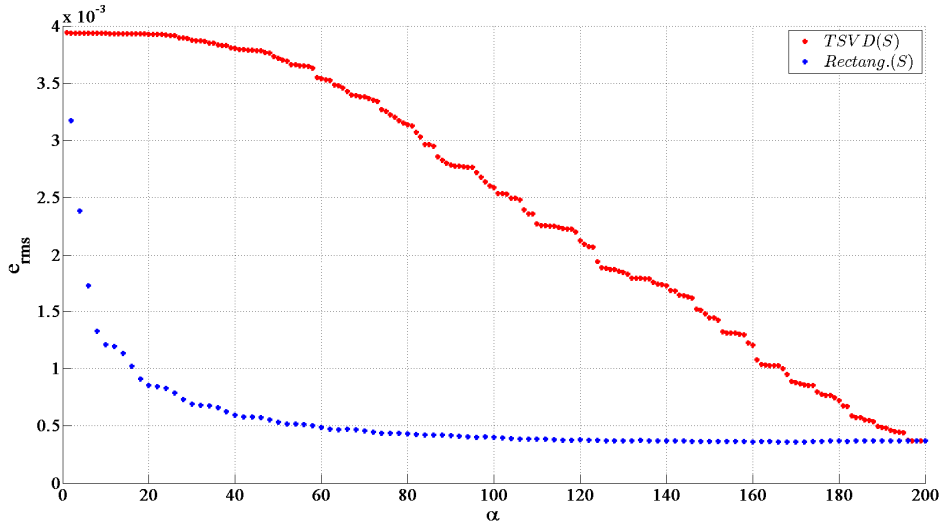


Fig. 4 – Root mean square error of the estimates of the harmonics for rectangular and TSVD estimation ($k_v = 1$)

2 Estimation of the Fourier spectrum with a virtual length ($k_v = 2$)

Exactly the same approach as in section 2 is followed here: the same N_x noised temperature are constructed on the $]-\ell\,;\,\ell]$ interval, but the integral Fourier transform (1) is defined on the $]-L\,;\,L]$ interval, with $L = 2\ell$ and, as a consequence N_x new eigenvalues $\alpha_n = \dfrac{n\pi}{L} = \dfrac{n\pi}{2\ell}$ are used. They that are the halves of the preceding ones (case $L = \ell$). This means that the space frequencies used for parameterizing the $\theta(x)$ profile are lower.

The exact analytical signal as well as its noised version shown in figure 8 are exactly the same as the ones plotted in figure 1.

The singular values of the sensitivity matrix $\boldsymbol{S}$ defined in equation (7) are plotted in figure 5.

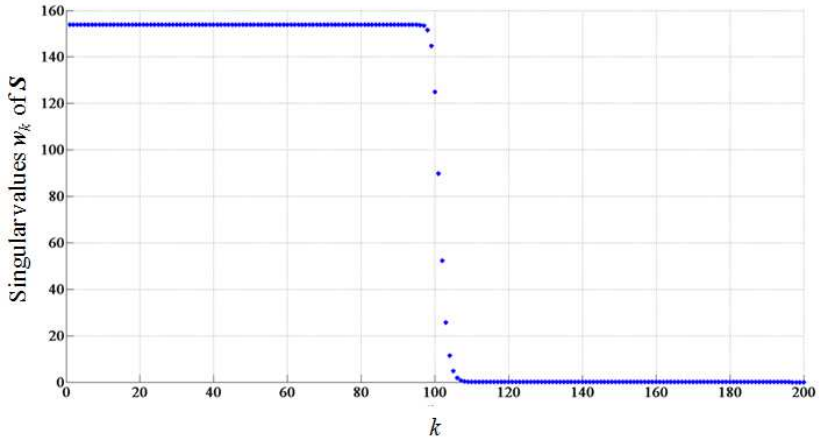


Fig. 5 – Singular values of the square sensitivity matrix $\boldsymbol{S}$ ($k_v = 2$)

It is very clear here that the inverse problem met here is severely ill-posed since the condition number of matrix $\boldsymbol{S}$ is $\mathrm{cond}(\boldsymbol{S}) = w_1 / w_{N_x=200} = 153/1.45\,10^{-15} = 1.55\,10^{17} \approx \infty$. This matrix is clearly singular and regularization is compulsory. If TSVD is used, it is obvious that the optimum value for the truncation parameter α will be in the region around $\alpha = N_x/2 = 100$ where the singular values show a sharp change of level.

Comparison of the variations of the root mean square residual with the α regularization hyperparameter for rectangular and TSVD estimation is shown in figure 6. Rectangular ordinary least square estimation cannot follow the simulated measurements. The residuals of TSVD estimation show an absolute minimum, slightly above the level of the standard deviation of the noise, for α between 120 and 140.

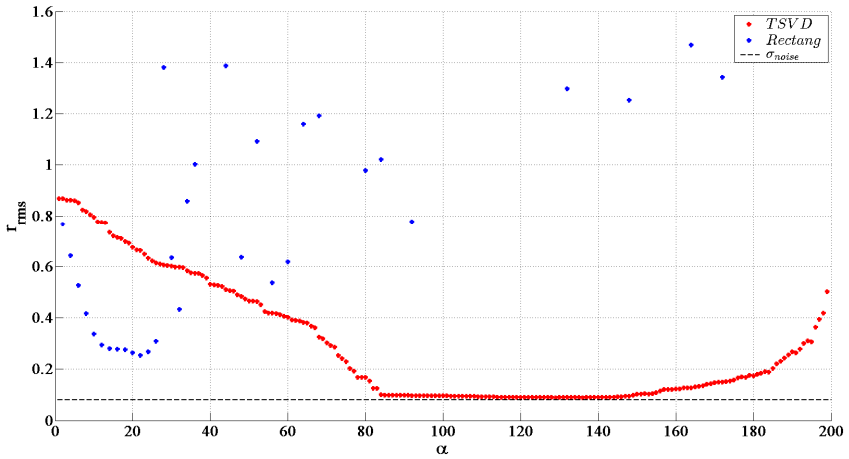


Fig. 6 – Root mean square residuals for rectangular and TSVD estimation (k_v = 2)

The root mean square of the errors of the estimates are plotted as a function of α, for the rectangular estimation and for the TSVD estimation in figure 7, where the vertical scale is logarithmic. For values of α lower than 120, the error is lower for the TSVD estimate, with a minimum slightly above α = 100. This corresponds roughly to the value where the rms residual is slightly above the noise level (flat region between 80 and 145) in figure 6, which corresponds to the discrepancy principle [6].

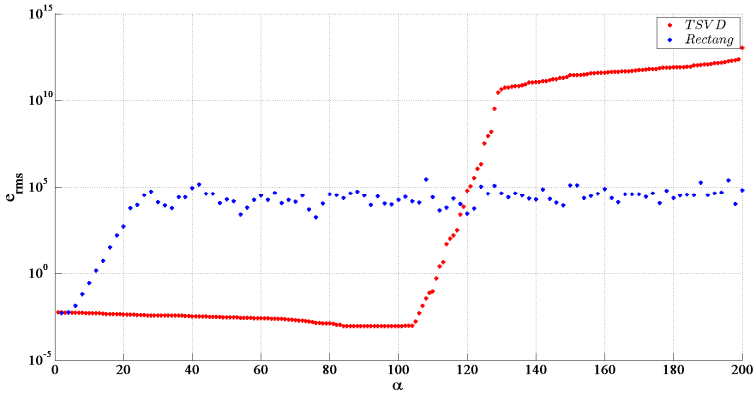


Fig. 7 – Root mean square error of the estimates of the harmonics for rectangular and TSVD estimation (k_v = 2)

The recalculated signal $\boldsymbol{\theta}^{recalc} = \boldsymbol{S}\,\tilde{\boldsymbol{\theta}}_{OLS}^{square}$ is plotted together with the exact signal $\boldsymbol{\theta}$ and with the noised signal $\boldsymbol{\theta}^{noised}$ in figure 8 for α=100 : the fit is good.

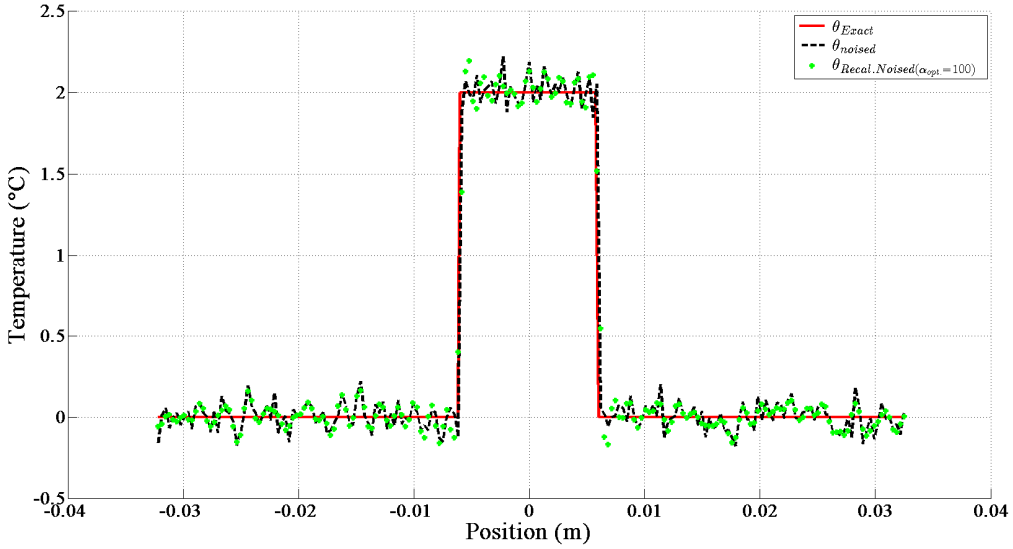


Fig. 8 – Exact, noised and reconstructed temperature profile using ordinary least squares (k_v = 2)

The corresponding spectral energy density is presented in figure 9: it shows that the estimation on a smaller x interval than the whole interval of width 2 L , even not perfect, is possible using TSVD regularization.

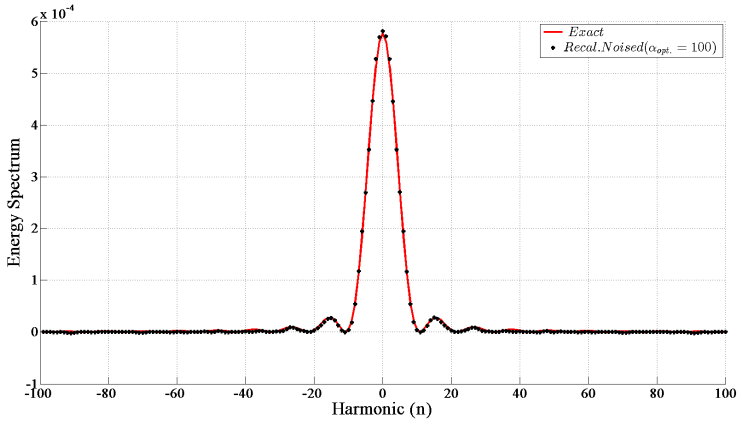


Fig. 9 – Spectral energy densities of temperature profile: exact ($\tilde{\boldsymbol{\theta}}_n^*\,\tilde{\boldsymbol{\theta}}_n$) and estimated ($\hat{\tilde{\boldsymbol{\theta}}}_{OLS}^{*square}\hat{\tilde{\boldsymbol{\theta}}}_{OLS}^{square}$)

Conclusions

We have shown in section 3 of this paper that estimation of the Fourier spectrum of a temperature profile was possible using measurements on an interval smaller than the space interval where its

Fourier transforms are defined. This can allow to take into account ill-defined lateral boundary conditions in problems involving inversion of a temperature profile measured by infrared thermography. These problems are met for example in thermal charactetization of heat transfer in a flat minichannel with outside temperature measurements, where the model can be written analytically in a simple way using Fourier transforms of temperature and flux (Thermal quadrupoles method) . Future efforts will be devoted to test the same type of simulated TSVD inversion for cases where neither temperature nor heat flux at one of the limits of the measurement interval is equal to zero: this happens for convection cases in the downstream area of the minichannel.

References

[1] Y. Souhar, B. Rémy, A. Degiovanni, *Thermal characterization of anisotropic materials at high temperature through integral methods and localized pulsed technique***,** International Journal of Thermophysics, Volume 34, Issue 2, (February 2013), pp. 322-340.

[2] Y. Rouizi, D. Maillet, Y. Jannot , *Fluid temperature distribution inside a flat mini-channel: Semi-analytical wall transfer functions and estimation from temperatures of external faces*, International Journal of Heat and Mass Transfer 64 (2013) 331–342.

[3] D. Maillet, S. André, J.C. Batsale, A. Degiovanni, C. Moyne, *Thermal Quadrupoles – Solving the Heat Equation through Integral transforms*, Wiley ed., Chichester, England, 370 pages, ISBN 0 471 98320 9, (October 2000).

[4] W. Al Hadad, Y. Rouizi, Y. Jannot, B. Rémy, D. Maillet, *Estimation of the heat transferred to a fluid minichannel by an inverse techniqu*e, 15 pages, paper 9190, submitted to 15th International Heat Transfer Conference, IHTC-15, Kyoto, August 10-15, (2014).

[5] A. Bendada, F. Erchiqui, M. Lamontagne, *Pulsed thermography in the evaluation of an aircraft composite using 3D thermal quadrupoles and mathematical perturbations*, Inverse Problems, 21 857. doi:10.1088/0266-5611/21/3/005 (March 2005).

[6] R.C. Aster, B. Borchers, C.H. Thurber, *Parameter Estimation and Inverse Problems*, Second edition, Elsevier ed., academic Press, Amsterdam, The Netherlands, 360 pages, (2012).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Analytical Observability of Symmetric Star Shaped Sources for Modified Helmholtz Model

Roberto Mamud G. S.[a], **Nilson C. Roberty**[*,b], **Carlos J. S. Alves**[c]

[a,b] Federal University of Rio de Janeiro
Av. Horácio Macedo, 2030, Bloco G, Sala 206, Cidade Universitria, Rio de Janeiro, Brazil
e-mail: nilson@con.ufrj.br
[c] Av. Rovisco Pais, 1 Department of Mathematics/IST
Technical University of Lisbon, Portugal

Key words: Observability, Analytical, Inverse Problem, Symmetric Star shaped, Helmholtz

Abstract

In this work we investigate the problem of reconstruction of symmetric star shaped characteristic sources in the modified Helmholtz inverse source problem. Uniqueness for the source centroid and effective radius is proved. It is not difficult to proof that only one Cauchy data at the boundary contain all information to be used in the reconstruction, see [1],and the lack of uniqueness is consequence of this. To resolve it, we must also prescribe field values for points in the interior of the domain, which signifies complete observability. The models for the inverse problem are typically systems of integral equations. Their derivation are based on the first and the second Green's formula. The first type are integral equations with the Dirichlet Green's function as kernel and the second are formulated with a Reciprocity gap functional based on moments with functions in the null space of the adjoint operator.

1 Introduction

In this work we investigate the reconstruction of characteristic sources in the modified Helmholtz model. As typical models for applications are those models based on second order equations, that is, the Laplace, Helmholtz and modified Helmholtz, wave and damped wave, heat and thermal wave inverse source problems. The central question on the inverse source problem for models with strongly elliptic operators and Cauchy data at the boundary is uniqueness. It is not difficult to proof that only one Cauchy data at the boundary contain all information to be used in the reconstruction, see [1],and the lack of uniqueness is consequence of this. To resolve it, we must also prescribe field values for points in the interior of the domain, which signifies complete observability. So, for boundary observability, these problems are classified according to the class of source we expected reconstruct. The main types are:

- Points sources given by Dirac delta like distributions;

- Regular sources f as harmonics functions or functions in a manifold defined by $\{A^*f = F\}$, where F is some given function and A is the model operator;
- Characteristic sources of convex or star shaped type.

The paper is structured as follows: The Section 2 presents the mathematical formulation and integral representation of the solution of the direct problem. The reciprocity gap equation is also introduced. The symmetric star shaped support, in which the formula for centroid determination is obtained, is presented in Section 3. We also show its uniqueness with respect to Cauchy data. An effective radius can also be uniquely defined. Spherical harmonics series expansion for Reciprocity gap functional are determined with the application of the Funk-Hecke Lemma in section 4. In Section 5 a numerical results for the two dimensional case is presented. Finally, we conclude in Section 6 by pointing out the advances introduced in the present work.

2 The Inverse Source Problem

2.1 Characterization of the problem

Let $\Omega = B_1(0) \subset \mathbb{R}^d$, $d \geq 2$, be the open unitary ball centred at the origin and $\Gamma := \partial\Omega$. The inverse source problem for the modified Helmholtz operator is, given Cauchy data $(g, g_\nu) \in H^{\frac{1}{2}}(\Omega) \times H^{-\frac{1}{2}}(\Omega)$, to find $(u, f) \in H^1(\Omega) \times L^2(\Omega)$ such that

$$\begin{cases} -\Delta u + \kappa^2 u = & f, \quad \text{in } \Omega \\ u = & g, \quad \text{on } \Gamma \\ \partial_\nu u = & g_\nu, \quad \text{on } \Gamma. \end{cases} \tag{1}$$

In this work, we are interested in reconstruct a source in the form $f(x)\chi_{\Omega_0}(x)$, where $\Omega_0 \subset \Omega$ is the support of the intensity of the source, f.

2.2 Uniqueness results for special cases

It is known that is useless to change the input Dirichlet data, g, and the unique information available is given by only one measurement, say, (g, g_ν), with $g = 0$, see [1].
The following lemma gives us a decomposition of the space $L^2(\Omega)$ where we can take the test functions.

Lemma 1 *Let $H_{-\Delta+\kappa^2}(\Omega) = \{v \in L^2(\Omega);\ (-\Delta + \kappa^2)v = 0\}$. If κ^2 and κ^4 are not eigenvalues of the Laplace and Bilaplace operators, respectively, then,*

$$L^2(\Omega) = H_{-\Delta+\kappa^2}(\Omega) \oplus (-\Delta - \kappa^2)(H_0^2(\Omega)).$$

Introducing a regular fundamental solution, v, to the modified Helmholtz operator and again using the Cauchy data, we obtain the *reciprocity gap equation*, or orthogonality condition,

$$\int_\Omega v(x)f(x)\chi_{\Omega_0}(x)dx + \int_\Gamma v(x)g_\nu(x)d\sigma_x - \int_\Gamma \partial_{\nu_x} v(x)g(x)d\sigma_x = 0, \tag{2}$$

where $v \in H_{-\Delta+\kappa^2}(\Omega)$ can be taken as

$$v(x) = e^{\kappa\omega\cdot(x-x_0)}, \tag{3}$$

where $\omega \in \mathbb{S}^{d-1}$ and $x_0 \in \mathbb{R}^d$ are arbitraries.
Depending on shaped of the set $\Omega_0 \subset \Omega$, we have uniqueness in the reconstruction of the source, that is, if two sources generates the same Cauchy data in the boundary, then these sources must be the same.

Theorem 2 *Let Ω_0^1 , $\Omega_0^2 \subset \Omega$ be star shaped domains with Lipschitz boundary. Consider the inverse problem* (1) *for $\kappa = 0$ with two sources $f\chi_{\Omega_0^1}$ and $f\chi_{\Omega_0^2}$. Suppose that $f(x) > 0$ for $x \in \Omega$. If the Cauchy data for the two problems are the same, then $\Omega_0^1 = \Omega_0^2$.*

Proof: See [3]. □

Theorem 3 *Let Ω_0^1 , $\Omega_0^2 \subset \Omega$ be convex domains with Lipschitz boundary. Consider the inverse problem* (1) *with two sources $f\chi_{\Omega_0^1}$ and $f\chi_{\Omega_0^2}$, where $f(x) > 0$ for $x \in \Omega$. If the Cauchy data for the two problems are the same and the potentials are equal on some ball $B \subset \Omega_0^1 \cap \Omega_0^2$, then $\Omega_0^1 = \Omega_0^2$.*

Proof: The proof is made by contradiction following the ideas of [3]. □

3 Symmetric Star Shaped Support

In this section we will be interested in define some integral operators related to the Reciprocity Gap Equation in order to study some properties of them, beside this we will investigate the series representation of these operators.

Observe that if we choose the polar coordinates system to coincide with the centroid of the star shaped source support, then the non linear integral equation, in polar coordinates, for the star shaped support reconstruction can be obtained by substituting the regular solution (3) in the reciprocity gap integral equation (2), for continuous $f > 0$,

$$\int_{\mathbb{S}^{d-1}} \int_0^{R(\theta)} \exp(\kappa\rho\omega\cdot\theta) f(\rho,\theta)\rho^{d-1} d\rho d\theta = h[g, g_\nu, x_c](\omega), \text{ for } \omega \in \mathbb{S}^{d-1}, \tag{4}$$

where $R(\theta)$ for $\theta \in \mathbb{S}^{d-1}$ is the star shaped boundary parametrization and

$$h[g, g_\nu, x_c](\omega) := -\int_\Gamma \exp(\kappa\omega\cdot(x-x_c)) g_\nu(x) d\sigma_x + \int_\Gamma \partial_{\nu_x} \exp(\kappa\omega\cdot(x-x_c)) g(x) d\sigma_x. \tag{5}$$

is the boundary function with exponential kernel.

Note that, in the analysis functional approach, (4) defines the following operator

$$F[\cdot] : L^d(\mathbb{S}^{d-1}) \to L^1(\mathbb{S}^{d-1})$$

by

$$F[R](\omega) := \int_{\mathbb{S}^{d-1}} \int_0^{R(\theta)} \exp(\kappa\rho\omega\cdot\theta) f(\rho,\theta)\rho^{d-1} d\rho d\theta \tag{6}$$

and (5) defines the operator

$$h[\cdot,\cdot,\cdot] : H^{\frac{1}{2}}(\Gamma) \times H^{-\frac{1}{2}}(\Gamma) \times \mathbb{R}^d \to L^1(\mathbb{S}^{d-1}). \tag{7}$$

Theorem 4 *There exist constants c_κ and C_κ such that*

$$c_\kappa \|R\|^d_{L^d(\mathbb{S}^{d-1})} \leq \|F[R]\|_{L^1(\mathbb{S}^{d-1})} \leq C_\kappa \|R\|^d_{L^d(\mathbb{S}^{d-1})}, \tag{8}$$

that is, the operator $F[\cdot] : L^d(\mathbb{S}^{d-1}) \to L^1(\mathbb{S}^{d-1})$ is bounded and coercive.

Proof: Firstly, note that for all bounded continuous function g with connected compact support $\Omega \subset \mathbb{R}^d$,

$$g(x_{min})\mu(\Omega) = \int_\Omega \min_\Omega\{g(x)\}dx \leq \int_\Omega g(x)dx \leq \int_\omega \max_\Omega\{g(x)\}dx = g(x_{max})\mu(\Omega),$$

where $\mu(\Omega)$ is the measure of Ω. So, considering some continuous path connecting x_{min} and x_{max}, for example $\gamma(t)$, $t \in [0,1]$, there exists some mean value $\overline{x} \in \gamma([0,1])$, such that

$$\int_\Omega g(x)dx = g(\overline{x})\mu(\Omega).$$

In this way, defining $g(x) = \exp(\kappa\omega \cdot x)f(x)\chi_{\Omega_0}(x)$, where χ_{Ω_0} is the characteristic function of the star shaped support of the source in (1), we have

$$F[R](\omega) = \int_\Omega g(x)dx = \int_\Omega \exp(\kappa\omega \cdot x)f(x)\chi_{\Omega_0}(x)dx = \exp(\kappa\omega \cdot \overline{x})f(\overline{x})\int_{\Omega_0} dx,$$

in which in polar coordinates gives us,

$$F[R](\omega) = \int_{\mathbb{S}^{d-1}}\int_0^{R(\theta)} \exp(\kappa\rho\omega \cdot \theta)f(\rho,\theta)\rho^{d-1}d\rho d\theta = \frac{\exp(\kappa\overline{\rho}\omega \cdot \overline{\theta})f(\overline{\rho},\overline{\theta})}{d}\int_{\mathbb{S}^{d-1}} R^d(\theta)d\theta.$$

Note that

$$\|F[R]\|_{L^1(\mathbb{S}^{d-1})} = \frac{f(\overline{\rho},\overline{\theta})}{d}\|R\|^d_{L^d(\mathbb{S}^{d-1})}\int_{\mathbb{S}^{d-1}} \exp(\kappa\overline{\rho}\omega \cdot \overline{\theta})d\omega.$$

In this way, since the exponential function is a continuous function, we have that there exists constants $c_\kappa = c_\kappa(\overline{\rho},\overline{\theta})$ and $C_\kappa = C_\kappa(\overline{\rho},\overline{\theta})$, such that

$$c_\kappa \|R\|^d_{L^d(\mathbb{S}^{d-1})} \leq \|F[R]\|_{L^1(\mathbb{S}^{d-1})} \leq C_\kappa \|R\|^d_{L^d(\mathbb{S}^{d-1})},$$

where

$$C_\kappa := \frac{f(\overline{\rho},\overline{\theta})}{d}\mu(\mathbb{S}^{d-1})\max_{\omega\in\mathbb{S}^{d-1}}\{\exp(\kappa\overline{\rho}\omega \cdot \overline{\theta})\}$$

and

$$c_\kappa := \frac{f(\overline{\rho},\overline{\theta})}{d}\mu(\mathbb{S}^{d-1})\min_{\omega\in\mathbb{S}^{d-1}}\{\exp(\kappa\overline{\rho}\omega \cdot \overline{\theta})\},$$

where $\mu(\mathbb{S}^{d-1}) = 2\frac{\pi^{d/2}}{\Gamma(d/2)}$ is the d-dimensional unitary sphere surface area. □

3.1 The symmetric star shaped case

Firstly, note that, since $f > 0$ is a known function, we can take, without loss of generality, its value 1. So, by using Taylor series expansion and boundedness, it is not difficult to show that

$$F[R](\omega) = \sum_{j=0}^{\infty}\frac{1}{j!}\int_{\mathbb{S}^{d-1}}\int_0^{R(\theta)} \kappa^j(\omega \cdot \theta)^j\rho^{j+d-1}d\theta,$$

that is,

$$F[R](\omega) = \sum_{j=0}^{\infty} \frac{\kappa^j}{(j+d)j!} \int_{\mathbb{S}^{d-1}} R^{j+d}(\theta)(\omega \cdot \theta)^j d\theta. \tag{9}$$

We can associate to F the already defined two others operators related to the *cosh* and *sinh* solutions of the modified Helmholtz model, respectively,

$$\begin{cases} F_c[R](\omega) = \frac{1}{2}[F[R](\omega) + F[R](-\omega)] = \int_{\mathbb{S}^{d-1}} \int_0^{R(\theta)} \cosh(\kappa\rho\omega \cdot \theta) f(\rho, \theta)\rho^{d-1} d\rho d\theta \\ F_s[R](\omega) = \frac{1}{2}[F[R](\omega) - F[R](-\omega)] = \int_{\mathbb{S}^{d-1}} \int_0^{R(\theta)} \sinh(\kappa\rho\omega \cdot \theta) f(\rho, \theta)\rho^{d-1} d\rho d\theta, \end{cases} \tag{10}$$

that is,

$$F_c[R](\omega) = \sum_{k=0}^{\infty} \frac{\kappa^{2k}}{(2k+d)(2k)!} \int_{\mathbb{S}^{d-1}} R^{2k+d}(\theta)(\omega \cdot \theta)^{2k} d\theta \tag{11}$$

and

$$F_s[R](\omega) = \sum_{k=0}^{\infty} \frac{\kappa^{2k+1}}{(2k+1+d)(2k+1)!} \int_{\mathbb{S}^{d-1}} R^{2k+1+d}(\theta)(\omega \cdot \theta)^{2k+1} d\theta. \tag{12}$$

Definition 1 *Let $R : \mathbb{S}^{d-1} \to (0,1)$, $R \in L^d(\mathbb{S}^{d-1})$, be some parametric representation of the star shaped boundary support of the source. We say that this source has a symmetric support when $R(\theta) = R(-\theta)$, for all $\theta \in \mathbb{S}^{d-1}$.*

Lemma 5 *Let χ_{Ω_0} be a star shaped source with symmetric support and $f \equiv 1$. Then, for all $\omega \in \mathbb{S}^{d-1}$, $F_s[R](\omega) = 0$.*

Proof: Suppose that ω' is some fixed direction in $\mathbb{S}^{d-1}$. Let $\mathbb{S}^{d-1}_{\pm\omega'} = \{\theta \in \mathbb{S}^{d-1}; \pm\omega' \cdot \theta > 0\}$ and $\mathbb{S}^{d-1}_0 = \{\theta \in \mathbb{S}^{d-1}; \omega' \cdot \theta = 0\}$. Then $\mathbb{S}^{d-1} = \mathbb{S}^{d-1}_{+\omega'} \cup \mathbb{S}^{d-1}_0 \cup \mathbb{S}^{d-1}_{-\omega'}$ and for $k \in \mathbb{N}$ and , by (10), we have

$$\begin{aligned} F_s[R](\omega') &= \int_{\mathbb{S}^{d-1}} \int_0^{R(\theta)} \sinh(\kappa\rho\omega' \cdot \theta)\rho^{d-1} d\rho d\theta \\ &= \int_{\mathbb{S}^{d-1}_{+\omega'}} \sinh(\kappa\rho\omega' \cdot \theta)\rho^{d-1} d\rho d\theta + \int_{\mathbb{S}^{d-1}_{-\omega'}} \int_0^{R(\theta)} \sinh(\kappa\rho\omega' \cdot \theta)\rho^{d-1} d\rho d\theta, \end{aligned}$$

since $\mu(\mathbb{S}^{d-1}_0) = 0$. Note that, by support symmetry and by definition of the sets $\mathbb{S}^{d-1}_{\pm\omega'}$,

$$\begin{aligned} \int_{\mathbb{S}^{d-1}_{-\omega'}} \int_0^{R(\theta)} \sinh(\kappa\rho\omega' \cdot \theta)\rho^{d-1} d\rho d\theta &= \int_{\mathbb{S}^{d-1}_{+\omega'}} \int_0^{R(\theta)} \sinh(-\kappa\rho\omega' \cdot \theta)\rho^{d-1} d\rho d\theta \\ &= -\int_{\mathbb{S}^{d-1}_{+\omega'}} \int_0^{R(\theta)} \sinh(\kappa\rho\omega' \cdot \theta)\rho^{d-1} d\rho d\theta. \end{aligned}$$

Therefore, $F_s[R](\omega') = 0$. Since $\omega' \in \mathbb{S}^{d-1}$ is arbitrary, we have that $F_s[R](\omega) = 0, \forall \omega \in \mathbb{S}^{d-1}$. □

Theorem 6 ***(Centroid Determination)*** *Suppose that the star shaped source has a symmetric support. Then, for a given Cauchy datum (g, g_ν), the centroid, x_c, can be calculated as solution of*

$$\omega \cdot x_c = \frac{1}{2\kappa} \ln \left(\frac{\int_\Gamma e^{\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{\kappa\omega\cdot x} g(x) d\sigma_x}{\int_\Gamma e^{-\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{-\kappa\omega\cdot x} g(x) d\sigma_x} \right), \tag{13}$$

where ω can be taken equal to $e_1, ..., e_d$, the canonical basis in $\mathbb{R}^d$.

Proof: Note that, by Lemma 5 and (10), we obtain the functional equation for centroid calculation

$$\int_\Gamma \sinh(\kappa\omega \cdot (x - x_c)) g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\omega_x} \sinh(\kappa\omega \cdot (x - x_c)) g(x) d\sigma_x = 0, \tag{14}$$

that is,

$$\int_\Gamma (\exp(\kappa\omega \cdot (x - x_c)) - \exp(-\kappa\omega \cdot (x - x_c))) g_\nu(x) d\sigma_x$$

$$- \int_\Gamma \partial_{\omega_x} (\exp(\kappa\omega \cdot (x - x_c)) - \exp(-\kappa\omega \cdot (x - x_c))) g(x) d\sigma_x = 0. \tag{15}$$

Observing that (15) is the same as

$$\begin{aligned} 0 &= e^{\kappa\omega\cdot(-x_c)} \left(\int_\Gamma e^{\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{\kappa\omega\cdot x} g(x) d\sigma_x \right) \\ &- e^{-\kappa\omega\cdot(-x_c)} \left(\int_\Gamma e^{-\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{-\kappa\omega\cdot x} g(x) d\sigma_x \right), \end{aligned}$$

we have that

$$e^{2\kappa\omega\cdot x_c} = \frac{\int_\Gamma e^{\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{\kappa\omega\cdot x} g(x) d\sigma_x}{\int_\Gamma e^{-\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{-\kappa\omega\cdot x} g(x) d\sigma_x},$$

for all $\omega \in \mathbb{S}^{d-1}$. This equation is invariant with respect to the euclidean group of orthogonal transformation in $\mathbb{R}^d$, SO^d. So

$$\omega \cdot x_c = \frac{1}{2\kappa} \ln \left(\frac{\int_\Gamma e^{\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{\kappa\omega\cdot x} g(x) d\sigma_x}{\int_\Gamma e^{-\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{-\kappa\omega\cdot x} g(x) d\sigma_x} \right),$$

for all $\omega \in \mathbb{S}^{d-1}$. This expression to calculated the centroid is well defined, since, by (4), $\int_\Gamma e^{\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{\kappa\omega\cdot x} g(x) d\sigma_x$ and $\int_\Gamma e^{-\kappa\omega\cdot x} g_\nu(x) d\sigma_x - \int_\Gamma \partial_{\nu_x} e^{-\kappa\omega\cdot x} g(x) d\sigma_x$ are null if and only if the source is null. □

Corollary 1 ***(Centroid Uniqueness)*** *Let $\chi_{\Omega_0^1}$ and $\chi_{\Omega_0^2}$ be star shaped sources with symmetric support in the modified Helmholtz model* (1)*. If they produces the same Cauchy data at the boundary, then they have the same centroid.*

4 Special Function Approach

In this section, we are interested in study the boundary functional (5) based on some properties of the modified Bessel function and the spherical harmonic space.

Definition 2 *Define the following integral representation of the modified Bessel function, see [4],*

$$I_\nu(z) = \frac{(z/2)^\nu}{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})}\int_{-1}^{1}(1-t^2)^{\nu-\frac{1}{2}}\cosh(zt)dt\,, \tag{16}$$

for $|argz| < \pi$ and $Re\nu > -\frac{1}{2}$. The d-dimensional regular spherical Bessel function is defined as

$$I_m^d(z) = (-1)^{d+m}\frac{I_{m+d/2-1}(z)}{z^{d/2-1}}. \tag{17}$$

The following lemma is very useful in the next proofs.

Lemma 7 ***(The Funk-Hecke Lemma)*** *Let $f : [-1,1] \to \mathbb{C}$ be a continuous function. If $m \geq 0$, then*

$$\int_{\mathbb{S}^{d-1}} f(\theta\cdot\omega)\psi_m(\omega)d\omega = \lambda\psi_m(\theta), \tag{18}$$

for all $\psi_m \in \mathcal{H}_m(\mathbb{S}^{d-1})$, where $\lambda_m = \mu(\mathbb{S}^{d-1})\int_{-1}^{1} f(t)P_m(d,t)(1-t^2)^{(d-3)/2}dt$, $P_m(d,t)$ is the $d-$dimensional Legendre polynomial and $\mathcal{H}_m(\mathbb{S}^{d-1})$ is the spherical harmonic space of dimension m of $\mathbb{S}^{d-1}$.

Proof: See [2]. □

Theorem 8 *Let $R \in L^d(\mathbb{S}^{d-1})$. Suppose that the source centroid coincide with the origin of the coordinate system and the Cauchy datum is referred to a system of coordinates in which the centroid coordinates is x_c. Then,*

$$\frac{(2\pi)^{\frac{d}{2}}}{(-1)^{m+d}}\int_{\mathbb{S}^{d-1}}\int_0^{R(\theta)} I_m^d(\kappa\rho)\psi_{m,p}(\theta)\rho^{d-1}d\rho d\theta = \int_{\mathbb{S}^{d-1}} h[g,g_\nu,x_c](\omega)\psi_{m,p}(\omega)d\omega, \tag{19}$$

for all $\psi_{m,p} \in \mathcal{H}_m(\mathbb{S}^{d-1})$, $p = 1,...,N_d(m)$, where $N_d(m) = \dim(\mathcal{H}_m(\mathbb{S}^{d-1}))$.

Proof: The proof is based on the Funk-Hecke lemma and on Proposition 2.26 in [2]. □

Theorem 9 *The boundary projection functional is given by*

$$\begin{aligned}\int_{\mathbb{S}^{d-1}} h[g,g_\nu,x_c](\omega)\psi_{m,p}(\omega)d\omega &= \frac{(2\pi)^{\frac{d}{2}}}{(-1)^{m+d}}\left(-\int_\Gamma I_m^d(\kappa\|x-x_c\|)\psi_{m,p}(\theta_c)g_\nu(x)d\sigma_x\right.\\ &+ \left.\int_\Gamma \partial_{\nu_x} I_m^d(\kappa\|x-x_c\|)\psi_{m,p}(\theta_c)g(x)d\sigma_x\right),\end{aligned} \tag{20}$$

for all $\psi_{m,p} \in \mathcal{H}_m(\mathbb{S}^{d-1})$, $p = 1,...,N_d(m)$, where $N_d(m) = \dim(\mathcal{H}_m(\mathbb{S}^{d-1}))$.

Proof: The proof is based on the Funk-Hecke lemma. □

Remark 1 *We must point out that a regular solution for the modified Helmholtz equation is given by*

$$\{I_m^d(\kappa\rho)\psi_{m,p}(\theta)\ ;\ m \geq 0 \text{ and } p = 1,...,N_d(m)\}.$$

So, if we introduce one of these test functions in the Reciprocity Gap Functional, we directly obtain a set of equations based on moments with the d-dimensional regular modified Bessel function. These solutions, as we have shown, express the projection of the functional resulting by substituting the exponential functions of plane waves.

Remark 2 *Combining Theorems 8 and 9, we can derive a new system associated to the reciprocity gap functional, that is,*

$$\int_{\mathbb{S}^{d-1}} \int_0^{R(\theta)} I_m^d(\kappa\rho)\psi_{m,p}(\theta)\rho^{d-1}d\rho d\theta = h[g, g_\nu, x_c]_{m,p}, \tag{21}$$

for all $\psi_{m,p} \in \mathcal{H}_m(\mathbb{S}^{d-1})$, $p = 1, ..., N_d(m)$, *where*

$$h[g, g_\nu, x_c]_{m,p} = -\int_\Gamma I_m^d(\kappa\|x - x_c\|)\psi_{m,p}(\theta_c)g_\nu(x)d\sigma_x + \int_\Gamma \partial_{\omega_x} I_m^d(\kappa\|x - x_c\|)\psi_{m,p}(\theta_c)g(x)d\sigma_x. \tag{22}$$

Once we have used the Cauchy data to reconstructed the source centroid, we can extract more information about the source by calculating an effective radius. For this, we treat the Cauchy data supposing that it is associated to a source with constant radius.

Definition 3 *The effective radius associated to the symmetric star shaped support of source is given by the solution,* $\overline{R}$*, of the integral equation*

$$\int_0^{\overline{R}} I_0^d(\kappa\rho)\rho^{d-1}d\rho = \frac{h[g, g_\nu, x_c]_0}{\mu(\mathbb{S}^{d-1})^{\frac{1}{2}}}. \tag{23}$$

Lemma 10 ***(Effective Radius Uniqueness)*** *If two symmetric star shaped sources generate the same Cauchy data at the boundary, then they have the same effective radius.*

We equivalently can use Taylor series to deduce the power series version of (4) for positive exponent.

Lemma 11 *Let* $R \in L^d(\mathbb{S}^{d-1})$, $0 < R(\theta) < 1$, *for all* $\theta \in \mathbb{S}^{d-1}$. *Then, for* $m > 0$,

$$\sum_{k=0}^{\infty} \frac{\pi^{\frac{d-1}{2}}}{2^m} \frac{\kappa^{m+2k}}{(m+2k+d)(2k)!} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(k+m+\frac{d}{2})} \int_{\mathbb{S}^{d-1}} R^{m+2k+d}(\theta)\psi_{m,p}(\theta)d\theta$$

$$= \int_{\mathbb{S}^{d-1}} h[g, g_\nu, x_c](\omega)\psi_{m,p}(\omega)d\omega, \tag{24}$$

for all $\psi_{m,p} \in \mathcal{H}_m(\mathbb{S}^{d-1})$.

The parametric representation of a star shaped boundary is giving by $R(\theta) = R_0 + \sum_{m=1}^{\infty}\sum_{p=1}^{N_d(m)} R_{m,p}\psi_{m,p}(\theta)$, for $\theta \in \mathbb{S}^{d-1}$. Note that since $L^p(\mathbb{S}^{d-1}) = \overline{\bigoplus_{n=0}^{m} \mathcal{H}_n(\mathbb{S}^{d-1}) \oplus \bigoplus_{n=m+1}^{\infty} \mathcal{H}_n(\mathbb{S}^{d-1})}|_{L^p(\mathbb{S}^{d-1})}$, for $p \geq 1$, we can decomposes $R(\theta) = R_{0:m}(\theta) + R_{>m}(\theta)$. Combining this decomposition with the symmetric support in (24), we have

$$\int_{\mathbb{S}^{d-1}} R^{m+2k+d}(\theta)\psi_m(\theta)d\theta = \int_{\mathbb{S}^{d-1}} (R_{0:m}(\theta) + R_{>m}(\theta))^{m+2k+d}\psi_m(\theta)d\theta$$

$$= \int_{\mathbb{S}^{d-1}} \sum_{l=0}^{m+2k+d} \frac{(m+2k+d)!}{(m+2k+d-l)!l!}(R_{0:m}(\theta))^{m+2k+d-l}(R_{>m}(\theta))^l\psi_m(\theta)d\theta.$$

Since $0 < R(\theta) < 1$, we are primarily interested in the first two terms of the expansion, that is, the separation of $m = 0$ and $m = 1$ from the $R(\theta)$ series.

5 Numerical Experiment

In this section we will study a two-dimensional numerical example involving a direct problem for two symmetric star shaped sources inside a circle. In this way, a quadratic Lagrangian finite element implementation is used to produces a Cauchy data on the boundary with a different model. The reconstruction operator, (6), and the reciprocity gap functional in the boundary, (5), are numerically evaluated taking $g \equiv 0$ and test functions $v_n = \frac{|n|2^{|n|}I_n(\kappa R)}{\kappa^{|n|}}$, $|n| = 0, 1, ..., N$, that is,

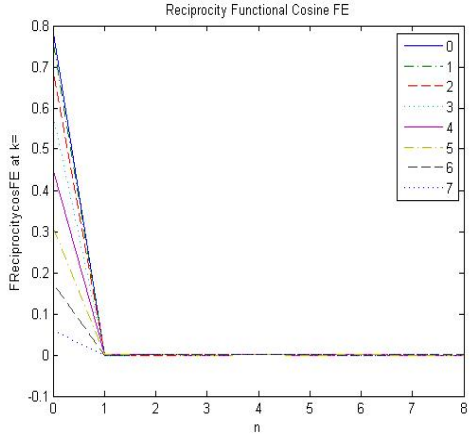
$$F_n^{\kappa}(r) = \int_0^{2\pi} \left[\int_0^{r(\theta)} \rho \frac{|n|2^{|n|}I_{|n|}(\kappa\rho)}{\kappa^{|n|}} d\rho \right] \exp(in\theta) d\theta \tag{25}$$

and

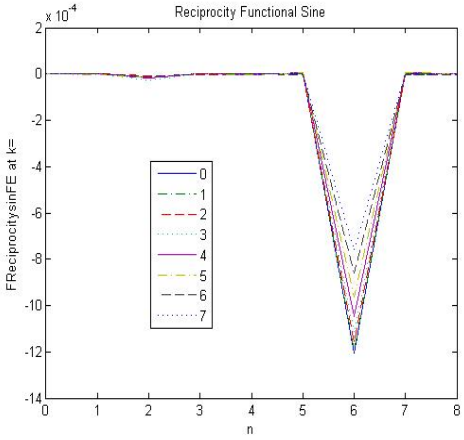
$$h_n^{\kappa}(g^{\nu}) = \int_0^{2\pi} g_{\nu} \frac{|n|2^{|n|}I_{|n|}(\kappa R(\theta))}{\kappa^{|n|}} \exp(in\theta) d\sigma. \tag{26}$$

Figures 1(a) and 1(b) shows typical values for this kind of data, for different values of κ.

The first characteristic source in the model is a symmetric star shaped source with boundary parametrized with the following finite Fourier series $r(t) = 0.5 + 0.05\ \sin(6t), t \in [0, 2\pi)$. This kind of source shape is shown with blue contour in Figure 1(c) and is almost exactly reconstructed in red for the modified Helmholtz parameters up to $\kappa = 5$. The values of the reciprocity gap functional calculated with the synthetic Cauchy data are showed in Figure 1(a) and 1(b). This data put in evidence the $|n| = 6$ component in the Fourier series of the shape. Data with 10% noise level are showed in Figure 1(d).



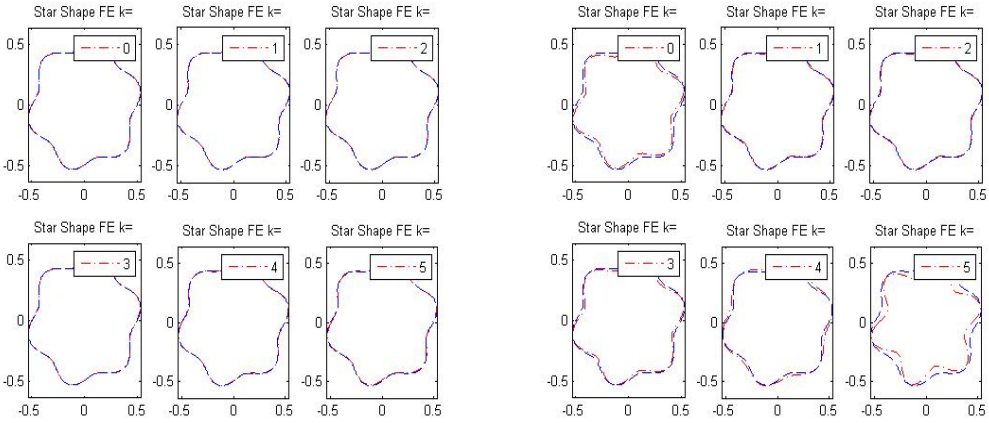
(a) Real part of the Reciprocity Gap Functional



(b) Imaginary part of the Reciprocity Gap Functional

6 Conclusions

An integral formulation for the inverse source problem with the modified Helmholtz equation is developed. The Reciprocity Gap equation, which solves the inverse problem independently of the direct problem, is formulated in polar coordinates for arbitrary d-dimensional problems. We proved that symmetric star shaped sources have centroid and effective radius uniqueness from the same Cauchy data. Some mathematical properties of functionals associated with the problem is established. Furthermore, the Funk-Hecke Theorem is used to established series representation of functionals and to approximate



(c) Reconstruction of the Star Shape source in the first model - no noise

(d) Reconstruction of the Star Shape source in the first model - 10% noise

it. A numerical experiment for reconstruction in the two-dimensional case shows a example of applications of the methodology developed. In this way, further works can be done in the three dimensional case, where it can be investigated with analogous tools.

Acknowledgements

The authors would like to thanks CNPq, Capes and FCT for the support obtained.

References

[1] C. J. S. Alves, N. F. M. Martins, N. C. Roberty, *Full identification of acoustic sources with multiple frequencies and boundary measurements*, Inverse Problems and Imaging, 3, (2009), pp. 275–294.

[2] K. Atkinson, W. Han, *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*, Springer, Lectures Notes in Mathematics, 2044, (2012).

[3] V. Isakov, *Inverse Source Problems*, Mathematical Surveys and Monographs 34, AMS, Providence, Rhode Island, 1st edition, (1990).

[4] N. N. Lebedev, *Special functions and their applications*, Selected Russian Publications in mathematical sciences Prentice Hall, New Jersey, 1st edition, (1965).

[5] N. C. Roberty, C. J. S. Alves, *Star shape transient source reconstruction from boundary data*, in Proceedings of ICCES'10, Las Vegas, Nevada, USA, (2010), pp.301-320.

[6] N. C. Roberty, D. M. S. de Sousa, *Source reconstruction for the Helmholtz equation*, in 68^o Seminário Brasileiro de Análise, USP, São Paulo, Brazil, (2008).

[7] D. M. S. de Sousa, N. C. Roberty, *An inverse source problem for the stationary diffusion-advection-decay equation*, Inverse Problems in Science and Engineering, (2012), pp. 891-915, DOI 10.1080/17415977.2011.609466 .

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

An Efficient Technique for Computing Sensitivity Coefficients of Solid Materials for Thermal Property Measurements

Antonio Marzola[1], Filippo de Monte[1,*], Dharmendra Mishra[2]

[1]Department of Industrial and Information Engineering and Economics
University of L'Aquila, Via G. Gronchi n.18, 67100 L'Aquila, Italy
e-mail: marzola.antonio@gmail.com, filippo.demonte@univaq.it
[2]Nestlé Nutrition, Fremont, Michigan 49412, USA
e-mail: mishradh@gmail.com

Key words: transient heat conduction, thermal properties, sensitivity coefficients, parameter estimation, two-layer slabs

Abstract
An analytical approach for computing the transient temperature and related sensitivity coefficients in a one dimensional, two-layer Cartesian body is described. In particular, we refer to an experimental apparatus for thermal property measurements of solid materials where one layer is the thin heater while the other is the sample of interest that can exchange heat on the backside. Once the temperature solution in both layers is obtained, the so-called 'scaled' sensitivity coefficients are calculated by using a semi-analytical technique that is computationally faster than fully numerical techniques available in the literature. A comparison with a simplified single-layer problem (where the heater is neglected) allows us to estimate the thermal disturbance of the heater on the scaled sensitivity coefficients of the specimen. Results are given in graphical form and shown a smaller or larger disturbance when changing the variables involved in the experiment.

1 Introduction

In the inverse problems such as parameter estimation, the sensitivity analysis of the temperature to the unknown parameters, such as thermal conductivity and volumetric heat capacity or thermal diffusivity, plays a fundamental role. It in fact allows a deep understanding of the phenomenon and provides insight into the estimation problem.
The parameter estimation technique (as described in [1]) requires measured temperature values. It is essential that the measured temperature be sensitive to the parameters of interest: the more sensitive the temperature (or large the sensitivity coefficient) is, the more valuable the temperature measurements are [2]. For this reason, the experimental apparatus have to be properly designed.
The sensitivity coefficients give a significant contribution to the optimal design of the experimental apparatus: on the one hand make possible a preliminary evaluation of the goodness of the experimental results (at least

from a qualitative point of view); on the other hand are directly involved in the estimation of the parameters when minimizing the ordinary least square norm [3-5]. A preliminary estimate can be made by observing the coefficients: to gain as much insight and information from the results as possible, it must be that the sensitivity coefficients with respect to the parameters of interest are uncorrelated and large in magnitudes.
In an apparatus that is not well designed, additional materials in the experimental configuration (for instance, the thin heater giving up heat to the sample) can have a larger impact on the temperature and, hence, on the sensitivity coefficients than the material of interest (sample). For this reason, in the current paper the focus will mainly be on investigating how the additional materials can modify and, hence, disturb the temperature and sensitivity coefficients of the specimen, with direct consequences on the quality of the results.
The heat conduction problem addressed is defined starting from the experimental design presented in [6], with the exception that the sample can exchange heat by convection/radiation with the surrounding ambient (air) through a heat transfer coefficient h. Two special cases can be considered: 1) $h = 0$ simulates a thermally insulated backside (boundary condition of Neumann type); and 2) $h \to \infty$ simulates a specimen whose backside is in contact with a cooling plate, such as aluminum (boundary condition of Dirichlet type). To estimate the effect of the thin heater on the sample sensitivity coefficients, a heat conduction problem in a one-dimensional two-layered slab is modeled and solved analytically by the well-established orthogonal expansion technique [7, chap. 8] based on the well-known separation-of-variable (SOV) method [7, chap. 2]. (For sake of completeness, multi-dimensional two-region solutions are discussed in [8, 9]). Once the temperature solution is obtained, the sensitivity coefficients are computed as partial derivative of the same temperature with respect to the parameters of interest and presented in a graphical form. Their computation is carried out in part analytically and in part numerically by using a first-order backwards difference; where the analytical part makes the computation very fast compared to fully-numerical techniques available in the parameter estimation literature.
Then a comparison between the sensitivity coefficients of two- and one-layered problems is carried out. It indicates that the measured temperature is very sensitive to the properties of the heater when the thermal conductivity ratio is large, but is not very sensitive to the Biot number on the sample backside and to the contact resistance at the interface.

2 Mathematical Formulation

Consider an experimental apparatus for the thermal property measurements of an orthotropic solid specimen, as proposed in Ref. [6]. A thin layer of mica heater is located at the interface of two samples of the same material and thickness and gives heat up at surface of both samples. For sake of thermal symmetry, the three-layer configuration (sample-heater-sample) reduces to the two-layer slab depicted in Fig. 1, where we assume that a time-independent heat flux is applied to the boundary $x = -L_1$ in place of a distributed heat source. The first layer represents ½ mica heater assembly while the second one is the specimen that can dissipate heat by convection/long-wavelengths radiation at $x = L_2$ with the surrounding air at temperature T_∞. Both layers are initially at a uniform temperature $T_{in} = T_\infty$ and there is no heat generation within them. Also, a thermal contact resistance is present at the interface $x = 0$ between the two adjacent layers whose thermal properties are considered temperature-independent.
The mathematical formulation of this transient, two-layer, linear, Cartesian heat conduction problem (denoted by X2C33B10T0; see the numbering system devised in [10]) is given in a dimensionless form by

$$\frac{\partial^2 \tilde{T}_1}{\partial \tilde{x}^2} = \frac{1}{\tilde{\alpha}} \frac{\partial \tilde{T}_1}{\partial \tilde{t}} \qquad (-L < \tilde{x} < 0;\ \tilde{t} > 0) \tag{1a}$$

$$\frac{\partial^2 \tilde{T}_2}{\partial \tilde{x}^2} = \frac{\partial \tilde{T}_2}{\partial \tilde{t}} \qquad (0 < \tilde{x} < 1;\ \tilde{t} > 0) \tag{1b}$$

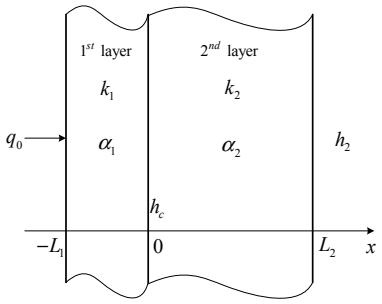


Figure 1: Schematic representation of the experimental apparatus. The first layer represents the heater, while the second one represents the specimen. Between them, there exists a contact resistance $R_c = 1/h_c$.

$$\tilde{k}\left(\frac{\partial \tilde{T}_1}{\partial \tilde{x}}\right)_{\tilde{x}=-L} = -1 \qquad (\tilde{t} > 0) \tag{1c}$$

$$\tilde{k}\left(\frac{\partial \tilde{T}_1}{\partial \tilde{x}}\right)_{\tilde{x}=0} + \frac{1}{\tilde{R}_c}\left[\tilde{T}_1(0,\tilde{t}) - \tilde{T}_2(0,\tilde{t})\right] = 0 \qquad (\tilde{t} > 0) \tag{1d}$$

$$\tilde{k}\left(\frac{\partial \tilde{T}_1}{\partial \tilde{x}}\right)_{\tilde{x}=0} = \left(\frac{\partial \tilde{T}_2}{\partial \tilde{x}}\right)_{\tilde{x}=0} \qquad (\tilde{t} > 0) \tag{1e}$$

$$\left(\frac{\partial \tilde{T}_2}{\partial \tilde{x}}\right)_{\tilde{x}=1} + Bi_2 \tilde{T}_2(1,\tilde{t}) = 0 \qquad (\tilde{t} > 0) \tag{1f}$$

$$\tilde{T}_1(\tilde{x},0) = 0 \qquad (-L < \tilde{x} < 0) \tag{1g}$$

$$\tilde{T}_2(\tilde{x},0) = 0 \qquad (0 < \tilde{x} < 1) \tag{1h}$$

The dimensionless variables appearing in the equations listed before are defined as:

$$\tilde{T} = \frac{T - T_\infty}{q_0 L_2 / k_2}, \quad \tilde{x} = \frac{x}{L_2}, \quad L = \frac{L_1}{L_2}, \quad \tilde{t} = \frac{\alpha_2 t}{L_2^2}, \quad \tilde{\alpha} = \frac{\alpha_1}{\alpha_2}, \quad \tilde{k} = \frac{k_1}{k_2}, \quad Bi_2 = \frac{h_2 L_2}{k_2}, \quad \tilde{R}_c = \frac{k_2 R_c}{L_2} \tag{2}$$

3 Solution

The problem defined by equations (1) has a nonhomogeneous boundary condition of the 2nd kind at $\tilde{x} = -L$ (driving term), as shown by Eq. (1c). Consequently, the SOV method cannot be applied directly. However, by using the principle of superposition for linear problems, the problem can be split up into a set of two simpler problems [7, Sect. 2-12]:

- a nonhomogeneous steady-state problem defined by the temperature $\tilde{T}_{i,ss}(\tilde{x})$ (with i = 1 or 2);
- and a homogeneous time-dependent problem (called complementary transient) defined by the temperature $\tilde{T}_{i,ct}(\tilde{x},\tilde{t})$ (with i = 1 or 2) and having $-\tilde{T}_{i,ss}(\tilde{x})$ as an initial temperature.

The two simpler problems stated above can be obtained by setting $\tilde{T}_i(\tilde{x},\tilde{t}) = \tilde{T}_{i,ss}(\tilde{x}) + \tilde{T}_{i,ct}(\tilde{x},\tilde{t})$ into Eqs. (1a)-(1h). Their solution is given in the following paragraphs.

3.1 Nonhomogeneous Steady-State Problem

The temperature solution $\tilde{T}_{i,ss}(\tilde{x})$ is:

$$\tilde{T}_{1,ss}(\tilde{x}) = 1 + \frac{1}{Bi_2} + \tilde{R}_c - \frac{1}{\tilde{k}}\tilde{x} \qquad (-L \le \tilde{x} \le 0) \tag{3a}$$

$$\tilde{T}_{2,ss}(\tilde{x}) = 1 + \frac{1}{Bi_2} - \tilde{x} \qquad (0 \le \tilde{x} \le 1) \tag{3b}$$

3.2 Homogeneous Complementary-Transient Problem

By applying the SOV-based orthogonal expansion technique [7, Chap. 8], the solution of the above homogeneous problem is:

$$\tilde{T}_{i,ct}(\tilde{x},\tilde{t}) = \sum_{n=1}^{\infty} c_n e^{-\beta_n^2 \tilde{t}} \psi_{i,n}(\tilde{x}) \qquad (i = 1, 2) \tag{4}$$

where β_n is the n-th eigenvalue whose computation is discussed ahead; and $\psi_{i,n}(\tilde{x})$ is the corresponding eigenfunction of the i-th layer which may be taken as

$$\psi_{1,n}(\tilde{x}) = A_{1,n} \sin\left(\frac{\beta_n}{\sqrt{\tilde{\alpha}}}\tilde{x}\right) + \cos\left(\frac{\beta_n}{\sqrt{\tilde{\alpha}}}\tilde{x}\right) \qquad (-L \le \tilde{x} \le 0) \tag{5a}$$

$$\psi_{2,n}(\tilde{x}) = A_{2,n} \sin\left(\beta_n \tilde{x}\right) + B_{2,n} \cos\left(\beta_n \tilde{x}\right) \qquad (0 \le \tilde{x} \le 1) \tag{5b}$$

The above coefficients may be derived by the inner and outer boundary conditions. After some algebra, we have

$$A_{1,n} = -\tan\left(\frac{\beta_n}{\sqrt{\tilde{\alpha}}}L\right), \qquad A_{2,n} = -\frac{\tilde{k}}{\sqrt{\tilde{\alpha}}}\tan\left(\frac{\beta_n}{\sqrt{\tilde{\alpha}}}L\right), \qquad B_{2,n} = 1 - \frac{\beta_n}{\sqrt{\tilde{\alpha}}}\tilde{R}_c\tilde{k}\tan\left(\frac{\beta_n}{\sqrt{\tilde{\alpha}}}L\right) \tag{6}$$

Also, the constant c_n appearing in the general solution Eq. (4) may be derived by applying the initial condition. Bearing in mind the orthogonality property of the eigenfunctions [7, Eq. (8.23-a)], we have:

$$c_n = -\frac{1}{N_n}\left[\frac{\tilde{k}}{\tilde{\alpha}}\int_{\tilde{x}'=-L}^{0} \tilde{T}_{1,ss}(\tilde{x}')\psi_{1,n}(\tilde{x}')\,d\tilde{x}' + \int_{\tilde{x}'=0}^{1} \tilde{T}_{2,ss}(\tilde{x}')\psi_{2,n}(\tilde{x}')\,d\tilde{x}'\right] \tag{7a}$$

where the norm N_n is defined as

$$N_n = \frac{\tilde{k}}{\tilde{\alpha}} \int_{\tilde{x}'=-L}^{0} \psi_{1,n}^2(\tilde{x}')\, d\tilde{x}' + \int_{\tilde{x}'=0}^{1} \psi_{2,n}^2(\tilde{x}')\, d\tilde{x}' \tag{7b}$$

Computation of the eigenvalues

The eigenvalues appearing in the complementary transient solution, Eqs. (4)-(7), may be computed as roots of the so-called eigencondition that is a transcendental equation coming from the application of the outer and inner boundary conditions. For the problem addressed, it is given by:

$$f(\beta) = \underbrace{\frac{\beta + Bi_2 \tan(\beta)}{Bi_2 - \beta \tan(\beta)}}_{f_1(\beta)} - \underbrace{\frac{\sqrt{\tilde{\alpha}} - \tilde{R}_c \beta \tilde{k} \tan\left(\frac{\beta}{\sqrt{\tilde{\alpha}}} L\right)}{\tilde{k} \tan\left(\frac{\beta}{\sqrt{\tilde{\alpha}}} L\right)}}_{f_2(\beta)} = 0 \tag{8}$$

The eigenvalues are computed by using the hybrid root-finder technique as described in [11, 12]. The starting point is that an eigenvalue is certainly located in the range between two successive asymptotes of the functions $f_1(\beta)$ and $f_2(\beta)$. These asymptotes are located where the denominator of f_1 and f_2 becomes zero. In particular, we note that the denominator of the $f_1(\beta)$ function is the eigencondition for the single-layer problem X23, whose eigenvalues can be obtained straight in an algebraic explicit form [13]. In other words, the locations of the asymptotes for $f_1(\beta)$ are the same as the eigenvalues for individual layers. For f_2, however, it is immediate to find the asymptotes.

Once the lower and upper bounds of each eigenvalue are established, the above hybrid computing technique provides the numerical value of that eigenvalue.

3.3 Temperature Solution

Applying the principle of superposition, the temperature of each layer is given by $\tilde{T}_i(\tilde{x},\tilde{t}) = \tilde{T}_{i,ss}(\tilde{x}) + \tilde{T}_{i,ct}(\tilde{x},\tilde{t})$. Therefore, bearing in mind Eqs. (4) and (6), we can write:

$$\tilde{T}_1(\tilde{x},\tilde{t}) = \left(1 + \frac{1}{Bi_2} + \tilde{R}_c - \frac{1}{\tilde{k}}\tilde{x}\right) + \sum_{n=1}^{\infty} c_n e^{-\beta_n^2 \tilde{t}} \psi_{1,n}(\tilde{x}) \qquad (-L \le \tilde{x} \le 0;\ \tilde{t} \ge 0) \tag{9a}$$

$$\tilde{T}_2(\tilde{x},\tilde{t}) = \left(1 + \frac{1}{Bi_2} - \tilde{x}\right) + \sum_{n=1}^{\infty} c_n e^{-\beta_n^2 \tilde{t}} \psi_{2,n}(\tilde{x}) \qquad (0 \le \tilde{x} \le 1;\ \tilde{t} \ge 0) \tag{9b}$$

A computer code in Matlab ambient was written for computing the above temperatures. A related code in Mathematica ambient is available at the internet site Exact Analytical Conduction Toolbox, or ExACT [14]. However, it refers to a two-layer slab with perfect contact between layers (X2C13B10T00 [15]). EXACT should be useful to engineers and scientists engaged in code verification, inverse problems, indirect

measurements, and anyone with a need for precise numerical values obtained from verified algorithms in heat conduction/diffusion. National Science Foundation funds this project.
However, for early times (less than the so-called deviation time t_d [10, 16, 17]), the temperature distribution of the first layer is not disturbed by the presence of the second layer. Hence, it behaves as a semi-infinite solid subject to a surface heat flux (X20B10T0) at its boundary with errors less than 10^{-A} (A = 1, 2, …, 10). The X20B1T0 solution is well-known and is given by [18]

$$\tilde{T}_1(\tilde{x},\tilde{t}) \approx \frac{2}{\tilde{k}}\sqrt{\tilde{\alpha}\tilde{t}}\,\mathrm{ierfc}\left(\frac{\tilde{x}+L}{2\sqrt{\tilde{\alpha}\tilde{t}}}\right) \qquad \left(-L \le \tilde{x} \le 0;\ 0 \le \tilde{t} \le \tilde{t}_d = \frac{1}{10A\tilde{\alpha}}(L-\tilde{x})^2\right) \tag{9c}$$

where $\mathrm{ierfc}(z)$ is the complementary error function integral.

4 Sensitivity Coefficients

Once the analytical solution of the thermal field is known, the scaled sensitivity coefficients for each layer can be obtained as partial derivatives of the temperature of the corresponding layer with respect to the model parameter η of interest (e.g., q_0, k_1, k_2, C_1, C_2, h_2, R_c) multiplied by the parameter itself:

$$S_{\eta,i} = \eta\left(\frac{\partial T_i}{\partial \eta}\right) \tag{10}$$

The scaled sensitivity coefficient has hence units of °C and their calculation requires the use of the chain rule [19]. In a dimensionless form,

$$X_{\eta,i} = \frac{S_{\eta,i}}{q_0 L_2 / k_2} = \eta\left(\frac{\partial \tilde{T}_i}{\partial \eta}\right) \tag{11}$$

4.1 Heat flux as a parameter

If the parameter of interest is the heat flux applied to the boundary surface $x = -L_1$, the corresponding dimensional scaled sensitivity coefficient is

$$S_{q_0,i} = q_0\frac{\partial T_i}{\partial q_0} = q_0\frac{\partial}{\partial q_0}\left(\tilde{T}_i\frac{q_0 L_2}{k_2}\right) = q_0\left[\frac{q_0 L_2}{k_2}\underbrace{\left(\frac{\partial \tilde{T}_i}{\partial q_0}\right)}_{=0} + \tilde{T}_i\frac{\partial}{\partial q_0}\left(\frac{q_0 L_2}{k_2}\right)\right] = \frac{q_0 L_2}{k_2}\tilde{T}_i \tag{12}$$

In dimensionless form, they are: $X_{q_0,1} = \tilde{T}_1$ and $X_{q_0,2} = \tilde{T}_2$.

4.2 Thermal conductivity and volumetric heat capacity as parameters

As $\tilde{T}_1$ and $\tilde{T}_2$ depend on the thermal properties of both heater and specimen, we will have eight coefficients, four for each layer, as shown in the following.

- Thermal conductivity:

$$S_{k_1,i} = k_1\left(\frac{\partial T_i}{\partial k_1}\right) = k_1\frac{\partial}{\partial k_1}\left(\tilde{T}_i\frac{q_0 L_2}{k_2}\right) = k_1\frac{q_0 L_2}{k_2}\left(\frac{\partial \tilde{T}_i}{\partial k_1}\right) \qquad (i = 1 \text{ or } 2) \tag{13a}$$

$$S_{k_2,i} = k_2\left(\frac{\partial T_i}{\partial k_2}\right) = k_2\frac{\partial}{\partial k_2}\left(\tilde{T}_i\frac{q_0 L_2}{k_2}\right) = k_2\left[\frac{q_0 L_2}{k_2}\left(\frac{\partial \tilde{T}_i}{\partial k_2}\right) + \tilde{T}_i\frac{\partial}{\partial k_2}\left(\frac{q_0 L_2}{k_2}\right)\right] \qquad (i = 1 \text{ or } 2) \tag{13b}$$

where $\tilde{T}_i = \tilde{T}_i\left[\tilde{k},\tilde{\alpha},\tilde{t},Bi_2,\tilde{R}_c,\beta_n(\tilde{k},\tilde{\alpha},Bi_2,\tilde{R}_c)\right]$, with $\tilde{k}(k_1,k_2)$, $\tilde{\alpha}[\alpha_1(k_1),\alpha_2(k_2)]$, $\tilde{t}[\alpha_2(k_2)]$, $Bi_2(k_2)$ and $\tilde{R}_c(k_2)$. By using the chain rule:

$$\frac{\partial \tilde{T}_i}{\partial k_1} = \frac{\partial \tilde{T}_i}{\partial \tilde{k}}\frac{\partial \tilde{k}}{\partial k_1} + \frac{\partial \tilde{T}_i}{\partial \tilde{\alpha}}\frac{\partial \tilde{\alpha}}{\partial \alpha_1}\frac{d\alpha_1}{dk_1} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{k}}\frac{\partial \tilde{k}}{\partial k_1} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{\alpha}}\frac{\partial \tilde{\alpha}}{\partial \alpha_1}\frac{d\alpha_1}{dk_1} \tag{14a}$$

$$\begin{aligned}\frac{\partial \tilde{T}_i}{\partial k_2} = {} & \frac{\partial \tilde{T}_i}{\partial \tilde{k}}\frac{\partial \tilde{k}}{\partial k_2} + \frac{\partial \tilde{T}_i}{\partial \tilde{\alpha}}\frac{\partial \tilde{\alpha}}{\partial \alpha_2}\frac{d\alpha_2}{dk_2} + \frac{\partial \tilde{T}_i}{\partial \tilde{t}}\frac{d\tilde{t}}{d\alpha_2}\frac{d\alpha_2}{dk_2} + \frac{\partial \tilde{T}_i}{\partial Bi_2}\frac{dBi_2}{dk_2} + \frac{\partial \tilde{T}_i}{\partial \tilde{R}_c}\frac{d\tilde{R}_c}{dk_2} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{\alpha}}\frac{\partial \tilde{\alpha}}{\partial \alpha_2}\frac{d\alpha_2}{dk_2} \\ & + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{k}}\frac{\partial \tilde{k}}{\partial k_2} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial Bi_2}\frac{dBi_2}{dk_2} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{R}_c}\frac{d\tilde{R}_c}{dk_2}\end{aligned} \tag{14b}$$

Substituting Eqs. (14) in Eqs. (13) gives the scaled sensitivity coefficients with respect to the thermal conductivities in both layers. In dimensionless form, we have:

$$X_{k_1,i} = \tilde{k}\frac{\partial \tilde{T}_i}{\partial \tilde{k}} + \tilde{\alpha}\frac{\partial \tilde{T}_i}{\partial \tilde{\alpha}} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\left(\tilde{k}\frac{\partial \beta_n}{\partial \tilde{k}} + \tilde{\alpha}\frac{\partial \beta_n}{\partial \tilde{\alpha}}\right) \qquad (i = 1,2) \tag{15a}$$

$$\begin{aligned}X_{k_2,i} = {} & \tilde{t}\frac{\partial \tilde{T}_i}{\partial \tilde{t}} - \tilde{T}_i - \left(\tilde{k}\frac{\partial \tilde{T}_i}{\partial \tilde{k}} + \tilde{\alpha}\frac{\partial \tilde{T}_i}{\partial \tilde{\alpha}} + Bi_2\frac{\partial \tilde{T}_i}{\partial Bi_2} - \tilde{R}_c\frac{\partial \tilde{T}_i}{\partial \tilde{R}_c}\right) \\ & - \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\left(\tilde{k}\frac{\partial \beta_n}{\partial \tilde{k}} + \tilde{\alpha}\frac{\partial \beta_n}{\partial \tilde{\alpha}} + Bi_2\frac{\partial \beta_n}{\partial Bi_2} - \tilde{R}_c\frac{\partial \beta_n}{\partial \tilde{R}_c}\right)\end{aligned} \qquad (i = 1,2) \tag{15b}$$

- Volumetric heat capacity. In a similar manner, we obtain

$$X_{C_1,i} = C_1\left(\frac{\partial \tilde{T}_i}{\partial C_1}\right) = -\tilde{\alpha}\frac{\partial \tilde{T}_i}{\partial \tilde{\alpha}} - \tilde{\alpha}\sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{\alpha}} \qquad (i = 1,2) \tag{16a}$$

$$X_{C_2,i} = C_2\left(\frac{\partial \tilde{T}_i}{\partial C_2}\right) = \tilde{\alpha}\frac{\partial \tilde{T}_i}{\partial \tilde{\alpha}} - \tilde{t}\frac{\partial \tilde{T}_i}{\partial \tilde{t}} + \tilde{\alpha}\sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{\alpha}} \qquad (i = 1,2) \tag{16b}$$

4.3 Heat transfer coefficient and contact resistance as parameters

Similarly to what was done in the previous subsection, we have:

- Heat transfer coefficient

$$X_{h_2,i} = h_2\left(\frac{\partial \tilde{T}_i}{\partial h_2}\right) = Bi_2\frac{\partial \tilde{T}_i}{\partial Bi_2} + Bi_2\sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial Bi_2} \qquad (i = 1,2) \tag{17a}$$

- Contact resistance

$$X_{R_c,i} = R_c\left(\frac{\partial \tilde{T}_i}{\partial R_c}\right) = \tilde{R}_c \frac{\partial \tilde{T}_i}{\partial \tilde{R}_c} + \tilde{R}_c \sum_{n=1}^{\infty} \frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \tilde{R}_c} \qquad (i=1,2) \tag{17b}$$

4.4 Thermal diffusivity as a parameter

Similarly, we obtain:

$$X_{\alpha_1,i} = \alpha_1\left(\frac{\partial \tilde{T}_i}{\partial \alpha_1}\right) = \alpha_1\left(\frac{\partial \tilde{T}_i}{\partial \widetilde{\alpha}}\frac{\partial \widetilde{\alpha}}{\partial \alpha_1} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \widetilde{\alpha}}\frac{\partial \widetilde{\alpha}}{\partial \alpha_1}\right) \tag{18a}$$

$$X_{\alpha_2,i} = \alpha_2\left(\frac{\partial \tilde{T}_i}{\partial \alpha_2}\right) = \alpha_2\left(\frac{\partial \tilde{T}_i}{\partial \tilde{t}}\frac{d\tilde{t}}{d\alpha_2} + \frac{\partial \tilde{T}_i}{\partial \widetilde{\alpha}}\frac{\partial \widetilde{\alpha}}{\partial \alpha_2} + \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \widetilde{\alpha}}\frac{\partial \widetilde{\alpha}}{\partial \alpha_2}\right) \tag{18b}$$

The sensitivity coefficients with respect to the thermal diffusivity in both layers are:

$$X_{\alpha_1,i} = \widetilde{\alpha}\frac{\partial \tilde{T}_i}{\partial \widetilde{\alpha}} + \widetilde{\alpha}\sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \widetilde{\alpha}} \tag{19a}$$

$$X_{\alpha_2,i} = \tilde{t}\frac{\partial \tilde{T}_i}{\partial \tilde{t}} - \widetilde{\alpha}\frac{\partial \tilde{T}_i}{\partial \widetilde{\alpha}} - \widetilde{\alpha}\sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \widetilde{\alpha}} \tag{19b}$$

By means of Eqs. (19), the sensitivity coefficients of Eqs. (15) and (16) can be rewritten in a more compact way as

$$X_{k_1,i} = X_{\alpha_1,i} + \widetilde{k}\frac{\partial \tilde{T}_i}{\partial \widetilde{k}} + \widetilde{k}\sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\frac{\partial \beta_n}{\partial \widetilde{k}} \tag{20a}$$

$$X_{k_2,i} = X_{\alpha_2,i} - \widetilde{T}_i - \left(\widetilde{k}\frac{\partial \tilde{T}_i}{\partial \widetilde{k}} + Bi_2\frac{\partial \tilde{T}_i}{\partial Bi_2} - \tilde{R}_c\frac{\partial \tilde{T}_i}{\partial \tilde{R}_c}\right) - \sum_{n=1}^{\infty}\frac{\partial \tilde{T}_i}{\partial \beta_n}\left(\widetilde{k}\frac{\partial \beta_n}{\partial \widetilde{k}} + Bi_2\frac{\partial \beta_n}{\partial Bi_2} - \tilde{R}_c\frac{\partial \beta_n}{\partial \tilde{R}_c}\right) \tag{20b}$$

$$X_{C_i,i} = -X_{\alpha_i,i} \tag{20c}$$

4.5 Computation of sensitivity coefficients

The partial derivatives appearing in Eqs. (15), (16), (17) and (19) are calculated analytically with the exception of $\partial\beta_n/\partial\tilde{k}$, $\partial\beta_n/\partial\tilde{\alpha}$, $\partial\beta_n/\partial Bi_2$ and $\partial\beta_n/\partial\tilde{R}_c$. These are computed numerically by using a first order backwards difference as, for multi-layer slabs, there are no explicit algebraic expressions of the eigenvalues such as for single-layer bodies [13].

It is relevant to note that all the scaled coefficients of sensitivity discussed above sum to zero for all values of time and position:

$$\begin{aligned}&(X_{q_0,1}+X_{q_0,2})+(X_{k_1,1}+X_{k_2,1}+X_{k_1,2}+X_{k_2,2})+(X_{C_1,1}+X_{C_2,1}+X_{C_1,2}+X_{C_2,2})\\&+(X_{h_2,1}+X_{h_2,2})-(X_{R_c,1}+X_{R_c,2})=0\end{aligned} \tag{21}$$

The same is true for each layer considered individually:

$$X_{q_0,1}+X_{k_1,1}+X_{k_2,1}+X_{C_1,1}+X_{C_2,1}+X_{h_2,1}-X_{R_c,1}=0 \tag{22a}$$

$$X_{q_0,2} + X_{k_1,2} + X_{k_2,2} + X_{C_1,2} + X_{C_2,2} + X_{h_2,2} - X_{R_c,2} = 0 \tag{22b}$$

Furthermore, it can be demonstrated that, for each layer, the following equations are valid:

$$X_{q_0,i} = \widetilde{T}_i \tag{23a}$$

$$X_{k_1,i} + X_{k_2,i} = \tilde{t}\frac{\partial \widetilde{T}_i}{\partial \tilde{t}} - \widetilde{T}_i \tag{23b}$$

$$X_{C_1,i} + X_{C_2,i} = -\tilde{t}\frac{\partial \widetilde{T}_i}{\partial \tilde{t}} \tag{23c}$$

They are exactly the same as the ones found for one-layer slabs [19].

5 Results

By using the numerical values of the experimental apparatus proposed in Ref. [6] for the thermal property measurements of an orthotropic carbon specimen, as given in the table below,

Table 1 – Thermal properties

layer	k [W m^{-1} °C^{-1}]	C [J m^{-3} °C^{-1}]	L [mm]
1st layer (mica heater)	0.14	2.03 x 10^6	0.44
2nd layer (carbon sample)	3.4	1.42 x 10^6	9.14

the dimensionless variables defined by Eq. (2) are:

$$L = 0.048\,, \quad \tilde{k} = 0.041\,, \quad \tilde{\alpha} = 0.0288\,, \quad \widetilde{R}_c = 0.05 \tag{24}$$

In particular, the numerical value of the dimensionless contact resistance comes out from R_c = 1.34 x 10^{-4} °C m^2 W^{-1} [20]. As the most critical parameter is the thermal conductivity, only its related sensitivity coefficient will here be taken into account.

Figure 2a shows a comparison between sensitivity coefficients of the sample at x = 0 for $Bi_2 \to \infty$, namely $X_{k_1,2}$ and $X_{k_2,2}$, with (two-layer) and without (one-layer) heater. It can be noted that the greater the thermal conductivity of the heater ($\tilde{k} = 10$ vs. $\tilde{k} = 0.1$), the lower the temperature sensitivity of the sample to its conductivity. In the one-layer transient problem (denoted by X21B10T0) the heater is completely neglected and, hence, the only layer is the specimen of interest. For sake of brevity, the solution of the X21B10T0 problem is not given here but it is available along with a related computer code in Matlab ambient for its computation at the internet site ExACT [21].

Figure 2b analyzes the effect of the contact resistance on both $X_{k_1,2}$ and $X_{k_2,2}$ calculated at x = 0 for $Bi_2 = 0$. The former is nearly constant when the contact resistance varies, while the latter is more influenced by $\tilde{R}_c$. The sensitivity coefficients for $\widetilde{R}_c = 0.05$ are not plotted in Fig. 2b as they are practically the same as the ones given for $\widetilde{R}_c = 0$. The temperature sensitivity of the sample to its conductivity reduces for large values of the contact resistance. A curve for the one-layer configuration (denoted by X22B10T0) is plotted on the same figure as a reference case. A computer code of the X22B10T0 problem is available in Ref. [22].

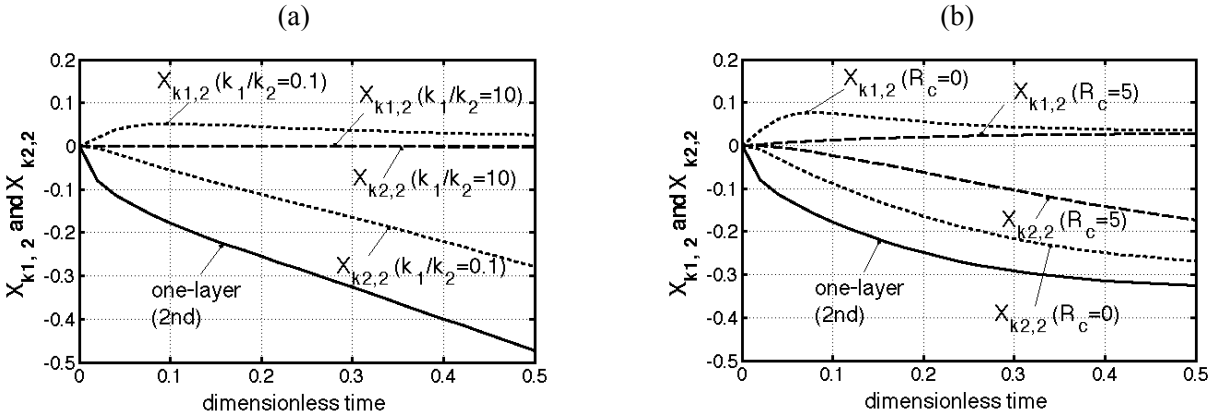


Figure 2: Dimensionless thermal conductivity scaled sensitivity coefficients of the sample at x = 0 vs. time for different values of a) thermal conductivity ratio $\tilde{k}$ (for $Bi_2 \to \infty$) and b) contact resistance $\tilde{R}_c$ (for $Bi_2 = 0$).

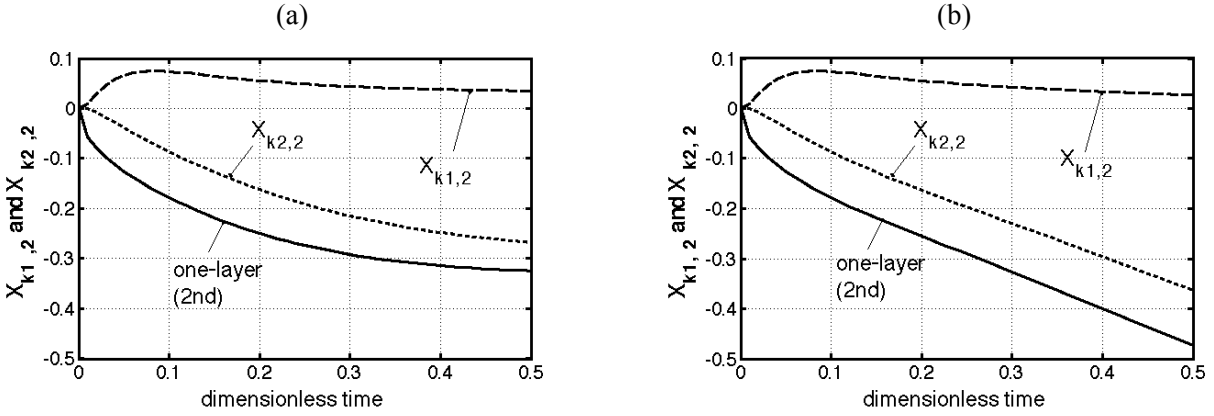


Figure 3: Dimensionless thermal conductivity scaled sensitivity coefficients of the sample versus time at x = 0 for (a) $Bi_2 = 0$ and (b) $Bi_2 \to \infty$.

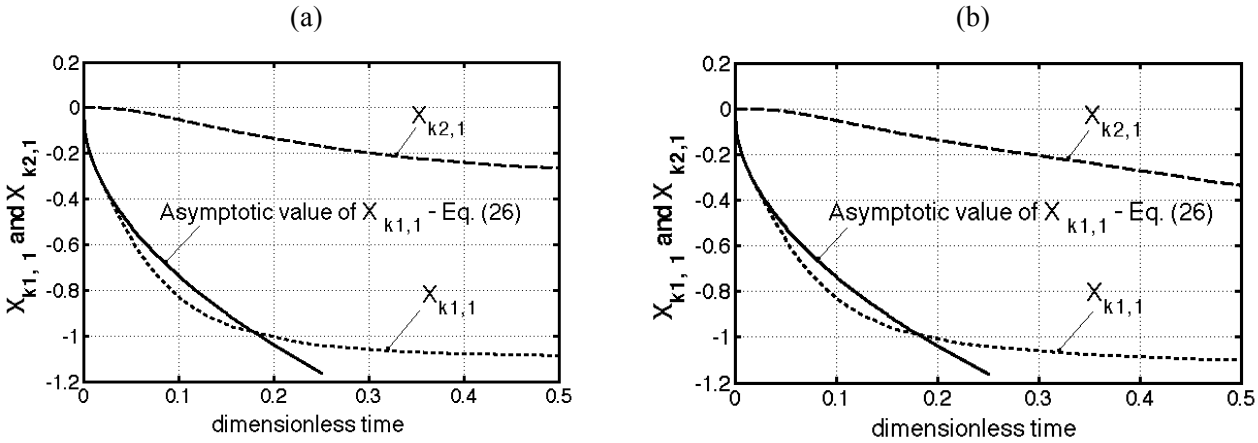


Figure 4: Dimensionless scaled sensitivity coefficients of the heater with respect to the thermal conductivity at $x = -L_1$ as a function of time for (a) $Bi_2 = 0$ and (b) $Bi_2 \to \infty$. The asymptotic behavior of $X_{k_1,1}$ is plotted too.

Figure 3 shows the dimensionless scaled sensitivity coefficients of the sample at $x = 0$ with respect to k_1 and k_2, namely $X_{k_1,2}$ and $X_{k_2,2}$, for two different values of the Biot number, $Bi_2 = 0$ (Fig. 3a - X2C32B10T0) and $Bi_2 \to \infty$ (Fig. 3b - X2C31B1T0; backside of the sample perfectly cooled). It is relevant to note that, for $Bi_2 \to \infty$, the $X_{k_2,2}$ coefficient is larger in magnitude. Therefore, except for early times, the sample temperature sensitivity to the thermal conductivity of the mica heater is small in comparison to the thermal conductivity of the same sample. A comparison with the one-layer configuration (only sample) is given too.
Finally, Figure 4 shows the sensitivity of the heater temperature calculated at $x = -L$ (mica heater mid-plane) for two special cases: $Bi_2 = 0$ (Figure 4a) and $Bi_2 \to \infty$ (Figure 4b). For both cases, the temperature sensitivity of the heater $X_{k_1,1}$ to its thermal conductivity is of the same order of magnitude as the sensitivity to the thermal conductivity of the specimen, $X_{k_2,1}$. Therefore, the thermal conductivity of the mica heater is as important as the thermal conductivity of the carbon sample and this aspect is not ideal for estimating the thermal conductivity of the specimen.
As for early times (less than the deviation time) the temperature distribution of the heater may be evaluated accurately by using Eq. (9c) in place of Eq. (9a) (that requires many eigenvalues), the sensitivity coefficients of Fig. 4 have been evaluated (with errors less than 10^{-A}) through the following algebraic expressions (instead of Eqs. (15a) and (15b) for $i = 1$, respectively)

$$X_{k_1,1} \approx -\frac{\tilde{T}_1}{2} + \left(\frac{\tilde{x}+L}{2\tilde{k}}\right) \underbrace{\operatorname{erfc}\left(\frac{\tilde{x}+L}{2\sqrt{\tilde{\alpha}\tilde{t}}}\right)}_{\tilde{q}_1} = -\frac{\tilde{T}_1}{2} + \left(\frac{\tilde{x}+L}{2\tilde{k}}\right)\tilde{q}_1, \qquad X_{k_2,1} \approx 0 \quad \text{for} \quad 0 \le \tilde{t} \le \tilde{t}_d = \frac{(L-\tilde{x})^2}{10A\tilde{\alpha}} \tag{25}$$

where $\tilde{q}_1 = q_1 / q_0$ is the conductive heat flux within the heater when it behaves as a semi-infinite solid. Also, as Fig. 4 refers to the location $\tilde{x} = -L$, the first of the two equations listed before simplifies to

$$X_{k_1,1} \approx -\frac{1}{\tilde{k}}\sqrt{\frac{\tilde{\alpha}\tilde{t}}{\pi}} \qquad \text{for} \quad 0 \le \tilde{t} \le \tilde{t}_d = \frac{4L^2}{10A\tilde{\alpha}} = 0.016 \tag{26}$$

where an accuracy of $A = 2$ (1 %) was considered. Hence, equation (26) represents the asymptotic behavior of Eq. (15a) when $i = 1$ for small values of the time.

6 Conclusions

The sensitivity analysis treated in this paper suggests that it is preferable to have the sensor for the temperature measurement embedded in the specimen, that is, at the interface between heater and specimen, rather than embedded in the heater. The experiment will in fact not be affected by the presence of the heater, regardless of its properties. On the contrary, if the temperature is measured in the heater, additional materials in the experimental configuration can have a larger impact on the temperature than the material of interest.

Acknowledgement

Support from the National Science Foundation is gratefully acknowledged from grant CBET-1250626 under program manager S. Acharya. ExACT Project PI: Prof. Kevin Cole; co-PI: Prof. Keith Woodbury.

References

[1] J.V. Beck, K.J. Arnold, *Parameter estimation in engineering and science*, John Wiley & Sons, New York, (1977).

[2] K.D. Dolan, D.K. Mishra, *Parameter Estimation in Food Science*, Annual review of food science and technology, 4 (2013), pp. 401-422.

[3] B. Garnier, D. Delauney, J.V. Beck, *Estimation Of Thermal Properties of Composite Materials without Instrumentation inside the Samples*, International Journal of Thermophysics, 13 (1992), pp. 1097-1111.

[4] A. Sakly, A. Jemni, P. Lagonotte, D. Petit, *Estimation of the Effective Thermal Properties in a Metallic Medium by an Inverse Technique using Temperature Measurements*, Journal of Applied Fluid Mechanics, 4 (2011), pp. 23-29.

[5] C. Aviles-Ramos, A. Haji-Sheikh, *Estimation of thermophysical properties of composites using multi-parameter estimation and zeroth-order regularization*, Inverse Problems in Engineering, 9 (2001), pp. 507-536.

[6] K.J. Dowding, B.F. Blackwell, J.V. Beck, A. Ulbrich, J. Hayes, *Estimation of Thermal Properties and Surface Heat Flux in Carbon-Carbon Composite*, Journal of Thermophysics and Heat Transfer, 9 (1995), pp. 345-351.

[7] N.M. Ozisik, *Heat Conduction*, John Wiley & Sons, Second Edition, (1993).

[8] C. Aviles-Ramos, A. Haji-Sheikh, J.V. Beck, *Exact solution of heat conduction in composite materials and application to inverse problems*, Journal of Heat Transfer, 120 (1998), pp. 592-599.

[9] C. Aviles-Ramos, A. Haji-Sheikh, J.V. Beck, K. *Dowding, Estimation of thermophysical properties by the spectral method—development and evaluation*, Journal of Heat Transfer, 123(1) (2001) 24-30.

[10] K. D. Cole, J. V. Beck, K. A. Woodbury, F. de Monte, *Intrinsic Verification and a Heat Conduction Database*, International Journal of Thermal Sciences, 78 (2014), pp. 36-47.

[11] C. Aviles-Ramos, K.T. Harris, A. Haji-Sheikh, *A hybrid root finder*, in: B. Bertram, C. Constanda, A. Struthers (Eds.), Integral Methods in Science and Engineering, Chapman & Hall/CRC Press, (2000).

[12] A. Haji-Sheikh, J.V. Beck, *Temperature solution in multi-dimensional multi-layer bodies*, International Journal of Heat and Mass Transfer, 45 (2002), pp. 1865-1877.

[13] A. Haji-Sheikh, J.V. Beck, *An efficient method of computing eigenvalues in heat conduction*, Numerical Heat Transfer, Part B: Fundamentals, 38 (2000), pp. 133-156.

[14] K. D. Cole, K. Woodbury, J. Beck, F. de Monte, D. Amos, A. Haji-Sheikh, et al., *Exact Analytical Conduction Toolbox (ExACT)*, exact.unl.edu, accessed 2 April 2013.

[15] A. Haji-Sheikh, *X2C13B10T00 Two-layer slab with perfect contact between layers; with jump in heat flux at one boundary, dissipating heat by convection into a zero temperature environment at x=L*, exact.unl.edu, August 21, 2013.

[16] F. de Monte, J. V. Beck, D. E. Amos, *Diffusion of thermal disturbances in two-dimensional Cartesian transient heat conduction*, International Journal of Heat and Mass Transfer, 51 (2008), pp. 5931-5941.

[17] F. de Monte, J. V. Beck, D. E. Amos, *Solving two-dimensional Cartesian unsteady heat conduction problems for small values of the time*, International Journal of Thermal Sciences, 60 (2012), pp. 106-113.

[18] F. de Monte, J. V. Beck, *X20B1T0 Semi-infinite Cartesian body with jump in boundary heat flux*, exact.unl.edu, February 1, 2013.

[19] B.F. Blackwell, R.J. Cochran, K.J. Dowding, *Development and Implementation of Sensitivity Coefficient Equations for Heat Conduction Problem*, Numerical Heat Transfer Part B: Fundamentals, 36 (1999), pp. 33-55.

[20] L.F.S. Carollo, A.L.F. Lima e Silva, S.M.M. Lima e Silva, *Applying different heat flux intensities to simultaneously estimate the thermal properties of metallic materials*, Measurement Science and Technology, 23 (2012), No. 6, 065601.

[21] F. de Monte, J. V. Beck, *X21B10T0 Slab with jump in heat flux at one boundary, zero temperature at other boundary*, exact.unl.edu, February 1, 2013.

[22] F. de Monte, J. V. Beck, *X22B10T0 Slab with jump in heat flux at one boundary, zero heat flux at other boundary*, exact.unl.edu, February 1, 2013.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Estimation of effective thermal properties within heterogeneous periodic media according to a multi-scale approach. Edge effect analysis

A. Matine, N. Boyard, Y. Jarny(*)
Université de Nantes - Laboratoire de Thermocinétique de Nantes ,UMR CNRS 6607,CNRS-, France
E-mail: yvon.jarny@univ-nantes.fr

Key words: heterogeneous media, effective properties, edge effects

Abstract

The homogenization theory is a powerful approach to determine the effective thermal conductivity tensor of heterogeneous materials such as composites, including thermoset matrix and fibers. Once the effective properties are calculated, they can be used to solve a heat conduction problem on the composite structure at the macroscopic scale. This approach leads to good approximations of both the heat flux and temperature fields in the interior zone of the structure. However, this approach provides a rather bad description of the heterogeneous fields close to the boundary. Consequently, the accuracy of the homogenized solutions in the vicinity of the boundary requires some additional correcting terms in order to account for edge effects. In the field of heat conduction, the solutions of classical inverse problems aim to determine unknown thermal properties and/or unknown heat flux boundary conditions from measurements which are most often located on the boundary of the sample. When these problems are formulated for heterogeneous media, the solution of the inverse problems is computed for estimating or by using the effective thermal properties. So the question of modeling the edge effects becomes important in order to perform reliable error analysis in such problems. In this paper, we will investigate the influence of edge effects on the solution of inverse heat conduction problems within heterogeneous periodic media, for which boundary measurements are used as additional data. Basic results concerning the homogenization approximation and the determination of correcting terms based on multi-scale asymptotic expansions will be first briefly presented, then the analysis of edge effects are performed and illustrated in the case of the classical "laser flash" experiment.

1.-Introduction

A usual approach to predict the thermal behaviour within complex heterogeneous media, consists in determining the effective thermal properties and/or temperature fields through homogenized heat transfer modelshomogenization theories. From an experimental point of view, specific devices such as classical transient laser flash (LF) [1], hot wire [2] methods (among others), or a specific hot disc method [3] can be used to estimate these effective properties.
The heat transfer modelling according to a multi-scale analysis [4]-[6] is complementary to the experimental approach and quite powerful. It aims on one hand to determine the effective thermal

properties from data known at the scale of the components, and on the other hand to have a better understanding of the "edge effects" [7]-[9] which may perturb the temperature field of the homogenized heat conduction model close to the boundaries of the spatial domain. Consequently in the estimation process of effective properties from experiments, which would result on the comparison of experimental surface temperatures with the solutions of homogenized models, more insight have to be done in the analysis of these "edge effects", to know when they can be neglected or not. More generally in the inverse analysis of experimental data, it is well-known that modelling errors should imply biased estimation of the unknown model parameters [10].
In a previous work [11], devoted to the heat conduction steady state analysis within heterogeneous periodic structures, it was shown how correcting terms can be introduced in the multi-scale asymptotic method to take into account these "edge effects", for homogenized 3-D heat conduction model. These results were extended to transient heat conduction [12] by using a space-time homogenization approach. Such approach was also studied in [13-14]. However, the introduction of additional terms to correct the edge effects in transient state was not considered.
In this paper, Wwe recall herefirst how these transient correcting terms can be introduced and calculated, depending on the classical boundary conditions in the heat transfer problem considered, and how correcting terms have also to be added to take into account "short time" effects. New numerical results are presented in the case of a periodic structure which models a simple UD (uni-directional) composite material, but the method is quite general and could be used for more complex periodic heterogeneous structures, like in plain weave fabric composites [15].
The last section is devoted to the discussion of numerical results of the heat transfer modelling in a LF experiment. The "heterogeneous solution" is compared to the homogenized one, computed with and without correcting terms, and to the analytical homogeneous solution. The spatial average temperature rise on the back face is measured as a function of time, and compared to the response of the homogenized problem in order to determine the effective thermal diffusivity of the sample. The bias, which results on the estimated value of the thermal diffusivity when edge effects are neglected in the homogenized model, is evaluated. The influence of the ratio between the size of the periodic cell and the size of the sample, so called the scale factor, is studied numerically for a UD structure.

2 Problem statement -Heat conduction in the heterogeneous medium

Let us consider a piece of heterogeneous periodic material, figure 1, defined in a bounded domain $\Omega \subset R^3$. The macroscopic coordinates of a point of Ω are denoted $\boldsymbol{x} = [x_1, x_2, x_3,]$ in a Cartesian coordinate system $\{0, e_1, e_2, e_3\}$. The boundary $\partial\Omega$ is subdivided in four distinct parts $\partial\Omega = \cup_{i=1}^{4} \Gamma_i$, in order to consider the different usual kinds of boundary conditions associated to the heat conduction problem:

- A Fourier condition on Γ_1 : the normal outward component $\phi^{\varepsilon}.n$ of the heat flux is fixed by an external temperature T_{ext} and a heat transfer coefficient h.
- A Neumann condition on Γ_2 : the normal outward component $\phi^{\varepsilon}.n$ of the heat flux is fixed
- A Dirichlet condition on Γ_3 : the temperature is fixed
- A periodic condition on Γ_4

The initial condition is fixed by the field $T_{ini}(x), x \in \Omega$, which is uniform or not.

The heterogeneous fields in the spatial domain Ω , are denoted respectively T^{ϵ} (temperature) and ϕ^{ε} (heat flux density). These fields over the time interval $(0, t_f)$ satisfy the following set of transient heat conduction equations together with the different kinds of boundary and initial conditions:

$$\begin{cases} \rho C_p \dfrac{\partial T^{\varepsilon}(x,t)}{\partial \tau} - div_x[\phi^{\varepsilon}(x,t)] = 0 \quad in \quad \Omega \times (0,t_f) \\ \phi^{\varepsilon}(x,t) = \boldsymbol{K}\nabla_x T^{\varepsilon}(x,t) \quad in \quad \Omega \times (0,t_f) \\ T^{\varepsilon}(x,t=0) = T_{ini}(x) \quad in \quad \Omega \\ \phi^{\varepsilon}(x,t).n = h(T^{\varepsilon} - T_{ext}) \quad on \quad \Gamma_1 \times (0,t_f) \\ \phi^{\varepsilon}(x,t).n = F \quad on \quad \Gamma_2 \times (0,t_f) \\ T^{\varepsilon}(x,t) = T_{ext} \quad on\ \Gamma_3 \times (0,t_f) \\ Periodic\ conditions\ on\ \Gamma_4 \times (0,t_f) \end{cases} \tag{1}$$

where n is the outward normal unit; ρ, C_p and $\boldsymbol{K}$ are respectively the density, the heat capacity and the thermal conductivity tensor of the heterogeneous medium which is assumed to have a periodic structure.

The periodic cell (see the figure 1), is denoted $\boldsymbol{Y} = \prod_{i=1}^{3}[0, l_i]$, and $\boldsymbol{y} = (y_1, y_2, y_3)\epsilon\boldsymbol{Y}$ are the coordinates of a cell point. The scale factor ε is the ratio between the size of $\boldsymbol{Y}$ and the size of Ω. Tthe microscopic coordinates are thus defined from $\boldsymbol{y} = \varepsilon^{-1}\boldsymbol{x}$.

Each component $k_{ij}(y); i,j = 1,2,3$ of the thermal tensor and the parameter $\rho C_p(\boldsymbol{y})$ are cell-periodic and depends on the local variable $\boldsymbol{y}$ (microscopic scale) in the cell domain $\boldsymbol{Y}$.

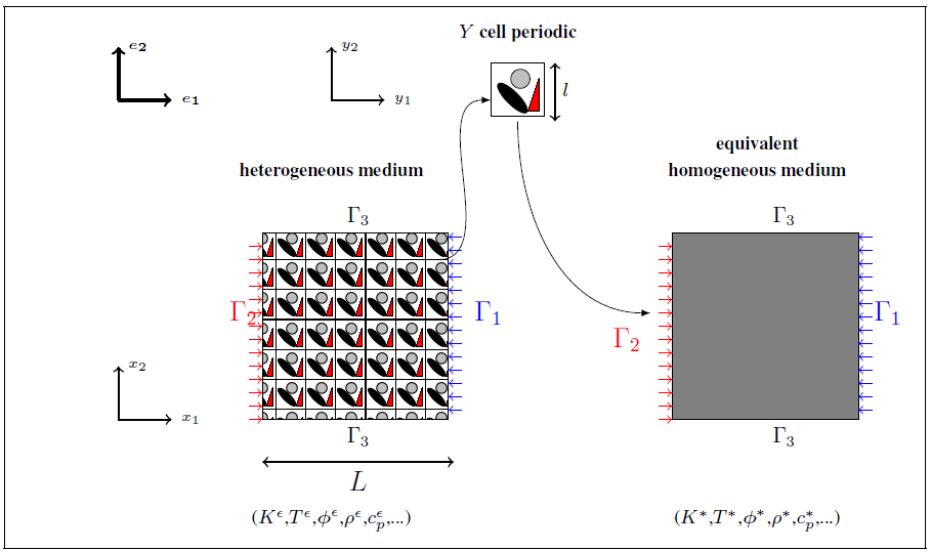


Figure 1: The spatial domain Ω of the heterogeneous periodic medium and the associated periodic cell $\boldsymbol{Y}$

3 Multi-scale asymptotic expansion method

It is assumed that the thermal conductivity of each components of the heterogeneous structure have the same order of magnitude, which means that the thermal contrast is not too large. The same assumption is done for the heat capacities. The influence of large contrast is not considered here and , it should lead to more developments, as described for example in [6]. Assuming that the scale factor ε is small enough, the asymptotic expansion method leads to expand the temperature T^{ϵ}, under the classical following form:

$$T^{\epsilon}(\boldsymbol{x},\boldsymbol{y},t) = T^0(\boldsymbol{x},\boldsymbol{y},t) + T^1(\boldsymbol{x},\boldsymbol{y},t).\epsilon + T^2(\boldsymbol{x},\boldsymbol{y},t).\epsilon^2 + \cdots; \quad \boldsymbol{x}\,\epsilon\,\Omega, y\,\epsilon\,\boldsymbol{Y} \tag{2}$$

To compute the effective thermal conductivity tensor $\boldsymbol{K}^*$ and $T^1(\boldsymbol{x},\boldsymbol{y},t)$, the new variables components $\omega_i(\boldsymbol{y}), i = 1,2,3$ of the variable $\boldsymbol{y}$ may then be introduced at the microscopic scale, as the solution of the cell problem (i=1,2,3):

$$\begin{cases} div_y\left[\boldsymbol{K}(\boldsymbol{y})(\boldsymbol{e_i} + \nabla_y \omega_i(\boldsymbol{y}))\right] = 0 \ \ in\ \boldsymbol{Y} \\ \omega_i \ \ periodic\ on\ \partial Y \end{cases} \tag{3}$$

therefore it can also be shown that the function T^0 does not depend on the microscopic variable, and satisfies the "homogenized" macroscopic scale heat conduction equation:

$$\begin{cases} \left(\rho C_p\right)^* \dfrac{\partial T^0(\boldsymbol{x},t)}{\partial \tau} - div_x[\boldsymbol{K}^* \nabla_x T^0(\boldsymbol{x},t)] = 0 \quad in \quad \Omega \times (0, t_f) \\ T^0(\boldsymbol{x}, t=0) = T_{ini}(\boldsymbol{x}) \quad in \quad \Omega \end{cases} \tag{4}$$

while the function $T^1(\boldsymbol{x},\boldsymbol{y},t)$ can be written as:

$$T^1(\boldsymbol{x},\boldsymbol{y},t) = \sum_{i=1}^{3} \frac{\partial T^0(\boldsymbol{x},t)}{\partial x_i} \omega_i(\boldsymbol{y}) \tag{5}$$

And the heat flux, at the order $k = 0$, is given by:

$$\boldsymbol{\varphi^0}(\boldsymbol{x},\boldsymbol{y},t) = (\boldsymbol{K}(\boldsymbol{y}).\left[\boldsymbol{e_i} + \nabla_y \omega w_i(\boldsymbol{y})\right] \boldsymbol{e_j})\, \nabla_x T^0(\boldsymbol{x},t) \tag{6}$$

where the effective thermal properties of the homogenized medium satisfy:

$$\left(\rho C_p\right)^* = \frac{1}{|\boldsymbol{Y}|} \int_Y \rho C_p dy \tag{7a}$$

$$\text{and } K^*_{i,j} = \frac{1}{|\boldsymbol{Y}|} \int_{\boldsymbol{Y}} \left[\boldsymbol{K}(\boldsymbol{y})(e_i + \nabla_y \omega_i(\boldsymbol{y}))e_j\right] dy,\ i,j = 1,2,3 \tag{7b}$$

It was discussed in a previous work [11] how periodic homogenization $(T^0 + T^1.\epsilon)$, at the order $k = 1$, provides good approximations of the heterogeneous solutions T^ϵ far enough from the boundary $\partial\Omega$ of the spatial domain. However, this approximation is not satisfactory anymore close to the boundary. This is first due to the loss of spatial periodicity. The second reason is that Neumann or Fourier conditions are only satisfied in a weak sense.
Consequently, in the vicinity of the boundary, correcting terms of "edge effects" have to be added. More generally in the transient state, the asymptotic expansion method developed above, provides a solution which satisfies the initial condition at the macroscopic scale, but does not guaranty that the initial condition of the functions $T^k(\boldsymbol{x},\boldsymbol{y},t=0), k \geq 1$, are equal to zero. Then correcting terms of "short time effects" have also to be introduced.

Finally, the temperature T^ϵ may be expanded (at the order $k = 1$), under the new following form

$$T^\epsilon = T^0(\boldsymbol{x},t) + \varepsilon.\left[T^1(\boldsymbol{x},\boldsymbol{y},t) + T^1_{st}(\boldsymbol{x},\boldsymbol{y},t,\tau) + \sum_{m=1}^{3} T^{1,m}_{bl}(\boldsymbol{x},\boldsymbol{y},t,\tau)\right] \ldots; \tag{8}$$

which involves the correcting terms T^1_{st} and $T^{1,m}_{bl}$ for short time and spatial boundary layer effects, respectively.

The spatial variations of $T_{bl}^{1,m}$ are exponentially decreasing [16] and vanish far enough from the boundary Γ_m : the superscript values *m=1,2,3* corresponds to one of the three different kinds of boundary condition usually associated to the heat conduction problem, see equations (1). The vicinity of the boundaries $\Gamma_{m=1,2,3}$ is defined as a strip G_m into the spatial domain Ω, in the normal direction to Γ_m [11]. By substituting this new asymptotic expansion in the heat conduction equation, and by identifying the powers of ε, we get a series of "spatial boundary layer" problems. At the order $k = 1$, it comes:

$$T_{blBL}^{1,m}(\boldsymbol{x},\boldsymbol{y},t,\tau) = \sum_{i=1}^{3} \frac{\partial T^0}{\partial x_i}(\boldsymbol{x},t)\, \chi_i^m(\boldsymbol{y},\tau) = \left(\boldsymbol{\chi}^m(\boldsymbol{y},\tau)\right)^t \nabla_{\boldsymbol{x}} T^0(\boldsymbol{x},t) \tag{9}$$

Consequently, the "edge effect" correcting term at the order $k = 0$, for associated to the approximation of the heat flux density can be written:

$$\boldsymbol{\phi_{bl}^{0,m}}(\boldsymbol{x},\boldsymbol{y},t,\tau) = \left[\boldsymbol{K}(\boldsymbol{y})(\nabla_y \chi^m(\boldsymbol{y},\tau))\right]\nabla_x T^0(\boldsymbol{x},t) \tag{10}$$

The functions $\chi_i^m(\boldsymbol{y},\tau)$, i=1,..,3; m=1,..,3, are the solutions of the following set of equations:

$$\begin{cases} \rho C_p(\boldsymbol{y})\dfrac{\chi_i^m(\boldsymbol{y},\tau)}{\partial t} - div_y\left[\boldsymbol{K}(\boldsymbol{y})(\nabla_y \chi_i^m(\boldsymbol{y},\tau))\right] = 0 \;\; in\; G_m \times \left]0, t_f\right[\\ \chi_i^m(\boldsymbol{y},\tau=0) = -\omega_i(\boldsymbol{y}) - z_i(\boldsymbol{y},\tau=0) \;\; in\; G_m \\ +Boundary\ Condition\ on\ \Gamma_m' = \Gamma_m \cap G_m \\ \chi_i^m\ periodic\ in\ the\ e_2, e_3\ direction \end{cases} \tag{11}$$

Like in the steady state analysis [11], depending on the condition chosen for determining the temperature field on the boundary Γ_m (Fourier, Neumann or Dirichlet conditions), the boundary condition for computing χ_i^m on Γ_m' takes different forms on $\Gamma_m' \times \left]0, t_f\right[$:

- Neumann or Fourier conditions (m=1 or 2):

$$\begin{aligned} \boldsymbol{K}(\boldsymbol{y})\left(\boldsymbol{\nabla_y}\chi_i^m(\boldsymbol{y},\boldsymbol{\tau})\right).\boldsymbol{n} \\ = -\boldsymbol{K}(\boldsymbol{y})\left((\boldsymbol{e_i} + \boldsymbol{\nabla_y} w_i(\boldsymbol{y}) + \boldsymbol{\nabla_y} \boldsymbol{z_i}(\boldsymbol{y},\boldsymbol{\tau}))\right).\boldsymbol{n} \\ + \frac{1}{|Y|}\int_{\boldsymbol{Y}} \boldsymbol{K}(\boldsymbol{y})\left(\boldsymbol{e_i} + \boldsymbol{\nabla_y} w_i(\boldsymbol{y})\right)\boldsymbol{dY}.\boldsymbol{n} \end{aligned} \tag{12a}$$

- Dirichlet condition (m = 3):

$$\chi_i^m(\boldsymbol{y},\tau) = -\omega_i(\boldsymbol{y}) - z_i(\boldsymbol{y},\tau) \tag{12b}$$

Furthermore, it was shown [11] how the depth of the "heat conduction boundary layer" (in the steady state), may be determined by solving numerically a specific eigenvalues problem set on G_m, depending on the kind of the boundary condition on Γ_m. For the above example, in the vicinity of $\Gamma_{m=1}$ the exponential decreasing in space of $T_{bl}^{1,m=1}$ in the e_1 direction, can be written [16]:

$$\chi_i^m(\boldsymbol{y}) = \Psi_i^m(\boldsymbol{y}).\, e^{-\delta_m\, y_1} \tag{13}$$

And where the parameter δ_m is given by the lowest solution of an eigenvalues problem, $\Psi_i^m(\boldsymbol{y})$ is the associated eigenvector on the sub-domain G_m. Hence the depth d_m of the spatial boundary layer close

to Γ_m where the correcting term $T_{bl}^{1,m}$ does not spatially vanish can be estimated by $d_m \approx \frac{3}{\delta_m}$. More mathematical developments and numerical results can be found in [11].

The short time correcting terms T_{st}^1 are determined, according to a double asymptotic scale in the time range (as in the spatial domain) [17]-[20], by introducing a "fast time" variable $\tau = t.\varepsilon^{-2}$. At the order $k = 1$, according to developments [12] similar to the above approach for $T_{bl}^{1,m}$, they may be computed as follows:

$$T_{st}^1(\boldsymbol{x}, \boldsymbol{y}, t, \tau) = \sum_{i=1}^{3} \frac{\partial T^0(\boldsymbol{x}, t)}{\partial x_i} z_i(\boldsymbol{y}, \tau) \tag{14}$$

and the "short time" correcting term at the order k = 0, for associated to the approximation of the heat flux density can be written:

$$\boldsymbol{\phi_{st}^0}(\boldsymbol{x}, \boldsymbol{y}, t, \tau) = \left[\boldsymbol{K}(\boldsymbol{y})(\nabla_y z_i(\boldsymbol{y}, \tau))\right] \nabla_x T^0(\boldsymbol{x}, t) \tag{15}$$

where the new variables $z_i(y, \tau), i = 1,2,3$ are solutions of the transient periodic cell problems

$$\begin{cases} \rho C_p \dfrac{z_i(\boldsymbol{y}, \tau)}{\partial t} - div_y\left[\boldsymbol{K}(\boldsymbol{y})(\nabla_y z_i(\boldsymbol{y}, \tau))\right] = 0 \;\; in\, \boldsymbol{Y} \times]0, t_f[\\ z_i(\boldsymbol{y}, \tau = 0) = -\omega_i(\boldsymbol{y}) \;\; in\, \boldsymbol{Y} \\ z_i \;\; periodic\ on\ \partial Y \end{cases} \tag{16}$$

The correcting terms T_{st}^1 are exponentially decreasing in time over the entire spatial domain, and vanishes after "short times".

4 Application to the modelling of a laser flash experiment

The laser flash [1] device is used to perform classical thermal experiments aiming to determine experimentally the effective thermal diffusivity of materials, heterogeneous or not. The data processing is based on the inverse analysis of the temperature rise measured on the back face of the sample which results of a heat pulse on the front face. To characterize heterogeneous structures, some "homogenized solution" of the transient heat conduction problem has to be compared to the measurements. Then it is important to take into account these « short time » and « spatial edge » effects, in the heat transfer modelling of the experiment (or at least to have them in mind) in order to avoid biased estimation.

4.1- The numerical measurements from the "heterogeneous solution"

The thermal characterization of a UD structure is considered (see figure 2). The initial state is supposed to be uniform, $T_{ini}(x) = 0$; the heat losses are modelled by a Fourier condition on both faces, with a heat transfer coefficient $h = 100W.m^{-2}.K^{-1}$ and the external temperature $T_{ext} = 0$ and. In practice, this h value should be not so high, but it has been chosen here to emphasize the "edge effects". Adiabatic condition is assumed on the others boundaries.

Three different values of the thickness sample are considered ($L = 3, 5, 10mm$), to study the influence of the scale factor (ε=0.33; 0.2; 0.1, respectively): $L = 3, 5, 10mm$. The heat pulse (duration = 0.001s, $\varphi = 2.10^5 W.m^{-2}$) is imposed at the front face of the rectangular sample. The resulting measurements $\widehat{T^{\varepsilon}}(t)$ are simulated on the back face. They are obtained by a spatial averaging of the

solution $T^{\varepsilon}(x,t)$ and plotted on the figure 3 (black lines). It is observed how the maximal temperature rises (time, magnitude) depends on the thickness of the sample.

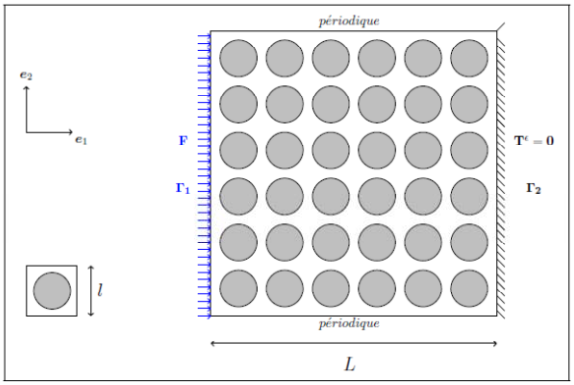


Figure 2: The UD medium and the associated periodic cell

k_m	$0.1\,W.m^{-1}.K^{-1}$
$\rho_m C_{pm}$	$10^5 J.m^{-3}.K^{-1}$
k_f	$1\ W.m^{-1}.K^{-1}$
$\rho_f C_{pf}$	$5*10^5 J.m^{-3}.K^{-1}$
Fiber volume ratio	$\frlvert\Omega_f\frrvert / \lvert\Omega\rvert = 0.64$
Cell size (l)	$l = 1mm$

Table 1: Thermal and geometrical data for the UD periodic structure-

4.2- The temperature response with the space-time expansion method.

Due to the symmetry of the periodic cell, the homogenized medium is isotropic and the effective thermal properties computed with the data given in table 1, are found to be:

- conductivity tensor : $\boldsymbol{K}^* = \begin{bmatrix} 0.333 & 0 \\ 0 & 0.333 \end{bmatrix}$ $(W.m^{-1}.K^{-1})$; and $\kappa = k_f/k_m = 10$
- heat capacity: $(\rho C_p)^* = \frac{1}{|\boldsymbol{Y}|}\int_Y \rho C_p dy = 3.54*10^5 J.m^{-3}.K^{-1}$; and $\rho_f C_{pf}\ /\rho_m C_{pm} = 5$
- diffusivity components: $a^*_{//} = a^*_{\perp} = 9.41.10^{-6} m^2.s^{-1}$

The homogenized solutions $(T^0 + T^1.\epsilon)$ are computed according to the multi-scale space-time expansion method and the first order correcting terms $T_{bl}^{1,m=1}$ is added or not. In this example, the correcting term T_{st}^1 for "short-time" effects can be neglected. The spatial average of these solutions on the back face are shown on figure 3 and compared to the heterogeneous solutions. The deviations which are observed without $T_{bl}^{1,m=1}$ (red lines) are well corrected when this term is considered (blue lines).

4.3- The estimation of the thermal diffusivity

In the LF experiment, the value of the effective thermal diffusivity component $a_{//}$ of the heterogeneous medium is estimated by matching the back face temperature measurements to the response of a homogenized heat conduction model. If no correcting terms (especially due to edge effects) are considered in the homogenized solution, a modelling error is done as illustrated on figure 3, and a biased estimation value $a_{//}^{est}$ is expected.

The matching process which aims to get $a_{//}^{est}$ from the measurement data $Y(t)$ is developed for three cases, in order to observe the influence on the thickness sample. Details of the method can be found in [1]. It allows to determine the parameter $a_{//}^{est}$ for which the exact 1-D analytical homogeneous solution gives the best fit to $Y(t)$. To verify the method, it is first applied to the data (red lines) without the correcting terms $Y(t) = \widehat{(T^0 + T^1.\epsilon)}$. For the three cases, the effective value $a^*_{//} = 9.41 \pm 0.05 e^{-6} e - 6 m^2.s^{-1}$ is well estimated. Then the method is applied to the data of the

heterogeneous model $Y(t) = \widehat{T^{\varepsilon}}(t)$. Three different estimated values $a_{//}^{est}$ of the thermal diffusivity component $a_{//}$ are found, and the resulting bias $\Delta a = \left|a_{//}^{*} - a_{//}^{est}\right|$ are summarized in table 2.

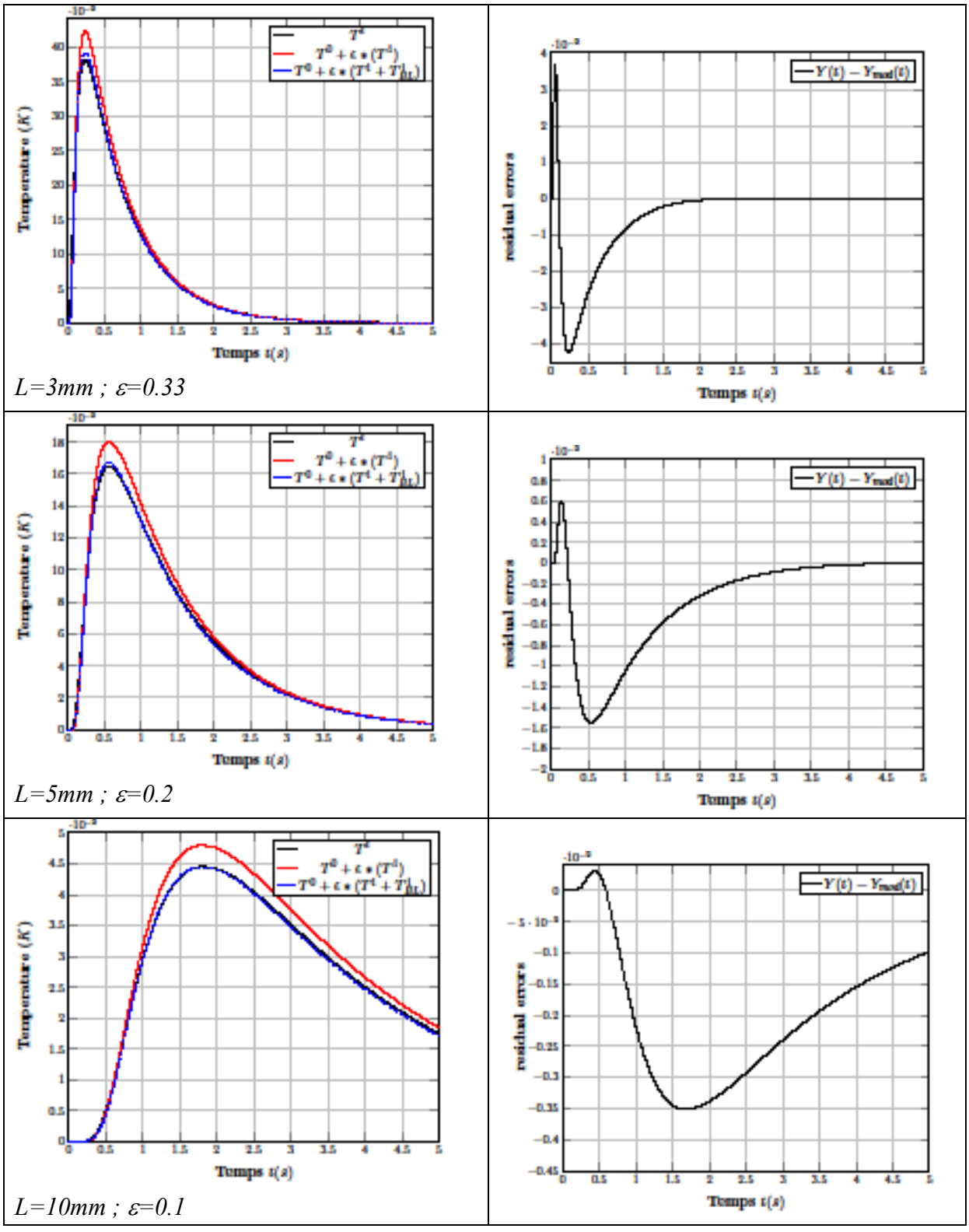


Figure 3: Left: The homogenized temperatures on the back face, computed with and without the correcting terms, and compared to the heterogeneous solution (influence of the scale factor). Right:- Deviations due to the correcting terms- Influence of the scale factor.

Thickness (mm)	Scale factor ε	Estimated value $a_{//}^{est}$ $(m^2.s^{-1})$	Bias $\Delta a = \left\|a_{//}^{*} - a_{//}^{est}\right\|$	$\frac{\Delta a}{a_{//}^{*}}$
0.01	0.1	9.70e-6	0.28e-6	2.97%
0.005	0.2	10.54e-6	1.13e-6	12%
0.003	0.33	12.51e-6	4.1e-6	43%

Table 2: Biased estimations of the thermal diffusivity component $a_{//}$ - Influence of the thicknessscale factor.

From these numerical experiments, three main results have to be underlined:

- the homogenized solution $(T^0 + T^1.\epsilon)$ without correcting terms is uniform on the back face, and its spatial average is identical to the 1-D analytical homogeneous solution computed with $a_{//}^{*}$, the "true" value determined from the effective properties
- the homogenized responses $(T^0 + (T^1 + T_{bl}^{1,m}).\epsilon)$ computed with the correcting terms $T_{bl}^{1,m}$ (and the "true" value $a_{//}^{*}$), are quite identical for the three cases to the solutions T^{ε} given by the heterogeneous model,
- the 1-D analytical homogeneous solution fits rather well the heterogeneous ones $\widehat{T^{\varepsilon}}(t)$, but the thermal diffusivity $a_{//}^{est}$ has to be chosen depending on thickness, which has no physical meaning.

It is observed in table 2, that the values of $a_{//}^{est}$ are always overestimated, compared to the effective value $a_{//}^{*}$. These results illustrate how the approximation of the heterogeneous model by a simple homogeneous one, that is a homogenized model without correcting terms, is not correct. Obviously the approximation becomes acceptable, for large enough thickness of the sample, that is for low value of the scale factor, here for $\varepsilon \leq 0.1$.

7 Conclusions

A space-time homogenization approach of a transient heat conduction problem in a periodic composite material has been developed according to a homogenization approach based on asymptotic expansion method. It leads to solve two kinds of problems, i.e. at the macroscopic and the microscopic scales. These results are available for any 3-D periodic heterogeneous structure, when the thermal contrast remains relatively low.

Following this approach, it was shown how the effective thermal diffusivity of such material may be biased, when it is estimated by fitting the experimental temperature rise from a laser flash experiment with the homogenized temperature computed without correcting terms. The influence of the scale factor has been illustrated. A good separation of scales is also required to ensure a good diffusivity estimation.

Acknowledgements

This work has been initialized in from the fruitful collaboration with G. Legrain and P. Cartraud from Ecole Centrale de Nantes [11]- [12].

References

[1] Degiovanni A., Laurent M. (1986)- *New Thermal-Diffusivity Identification applied to Flash Method*- Revue de Physique Appliquée, 21 (3), pp. 229-237

[2] Anderson A. (1976)- *Thermal conductivity of some rubbers under pressure by the transient hot-wire method* J. Appl. Phys 47, pp. 2424-2426

[3] Thomas M, Boyard N, Lefèvre N, Jarny Y and Delaunay D, (2010)- *An experimental device for the simultaneous estimation of the thermal conductivity 3-D tensor and the specific heat of orthotropic composite materials* Compos.Sci.Technol 53, pp. 5487-98.

[4] A Bensoussan, JL Lions, G Papanicolaou (1978)– *Asymptotic Analysis for periodic structures* – North Holland Ed., Amsterdam

[5] –E Sanchez-Palenzia, A Zaoui editors, *Homogenization techniques for composites media* (1987) Lecture Notes in Physics ,Volume 272 ,Springer Berlin Heidelberg

[6] J Auriault, C Boutin, C Geindreau (2010)- *Homogenization of Coupled Phenomena in Heterogeneous Media* – J Wiley & Sons

[7] H Dumontet (1986)–*Study of boundary layer problem in elastic composite materials- RAIRO-Modélisation mathématique et analyse numérique,* tome 20, n°2, p.265-286

[8] O Fudym, JC Batsale, D Lecomte (2004) – *Heat diffusion at the boundary of stratified media Homogenized temperature field and thermal constriction* – Int. J. of Heat and Mass transfer 47, pp. 2437-2447

[9] N Buannic, P Cartraud (2001) – *Higher-order effective modeling of periodic heterogeneous beams – Part 2 : derivation of the proper boundary conditions for the interior asymptotic solution* – Int. J. of Solids and Structures, 38, pp. 163-7180

[10] Beck J.V., B. Blackwell and C. St. Clair. (1985) *Inverse Heat Conduction, Ill-posed problems*. Wiley-Interscience, New York.

[11] A. Matine, N. Boyard, P. Cartraud, G.Legrain, Y. Jarny (2013) - *Modeling of thermo-physical properties in heterogeneous periodic media according to a multi-scale approach: effective conductivity tensor and edge effects* –Int. J. of Heat and Mass transfer 62, pp. 586-603

[12] A. Matine, N. Boyard, P. Cartraud, G.Legrain, Y. Jarny– *Transient heat conduction within periodic heterogeneous media- A space time homogenization approach* –submitted to Int. J. of Thermal Sciences (2014)

[13] M. Kaminski (2003)- *Homogenization of transient heat transfer problems for some composite materials* – International Journal of Engineering Science 41, pp. 1-29,

[14] F Larsson, K Runesson, Fsu (2010) – *Variationally consistent computational homogenization of transient heat flow*- Int. J. Numer. Meth. Engng., 81 , pp. 1659-1686

[15] A Dasgupta, R.K. Agarwal (1992) – *Orthotropic thermal conductivity of plain-weave fabric composites using a homogenization technique* – J. of Comp. Mat. 26, (18) pp. 2736-2759

[16] M Amar, M Tarallo, S Terracini (1999) – *On the exponential decay for boundary layer problems* – C.R. Acad. Sci., Paris, 328, série 1, pp. 1139-1144

[17] A. Piatnitski (1980). *A parabolic equation with rapidly oscillating coefficients*. Mosc. Univ. Math. Bull, 35(3):pp. 35-42.

[18] Denis Aubry and Guillaume Puel (2010). *Ccf modelling with use of a two-timescale homogenization model*. Procedia Engineering, 2(1): pp.787- 796.

[19] Qing Yu and Jacob Fish (2002). *Multiscale asymptotic homogenization for multiphysics problems with multiple spatial and temporal scales: a coupled thermo-viscoelastic example problem*. International Journal of Solids and Structures, 39(26) : pp. 6429-6452.

[20] J. Fish and Wen Chen (2004). *Space-time multiscale model for wave propagation in heterogeneous media.* Computer Methods in Applied Mechanics and Engineering, 193 (45-47): pp. 4837-4856.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

A Filter Based Solution for Non-Linear Inverse Heat Conduction Problems in Multi-Layer Medium

Hamidreza Najafi[a], Keith A. Woodbury*[a], James V. Beck[b], Ned R. Keltner[c]

[a]Mechanical Engineering Department, The University of Alabama, Tuscaloosa AL
[b]Department of Mechanical Engineering, Michigan State University, East Lansing, MI
[c]F.I.R.E.S, Inc, Albuquerque, NM, USA
*e-mail: keith.woodbury@ua.edu

Key words: Inverse heat conduction, real-time heat flux measurement, temperature dependent material properties multi-layer medium, filter based solution

Abstract
A solution for the non-linear inverse heat conduction problem (IHCP) in a two-layer medium is proposed and tested through numerical experiment. The temperature histories are considered to be known at two points on one layer and the heat transfer rate at the end of the layer exposed to a thermal environment is to be determined. A step-by-step solution is proposed for solving this problem based on the minimization of the sum of the squared errors between the computed and known values and by using of Tikhonov Regularization for stabilizing the solution. The solution is cast in digital filter form which allows a near real-time heat flux estimation in the multi-layer problem. The filter coefficients are determined for different temperatures. These data are used to train an artificial neural network (ANN) which calculates the filter coefficients based on the temperature at each time step. The ANN here serves to interpolate the filter coefficients to account for the temperature variation of the material properties. The proposed method is tested via numerical model developed in ANSYS and also the results are validated with the exact solution for constant properties. The filter algorithm can be used easily for near real-time heat flux estimation in industrial applications.

1 Introduction

The problem of estimating unknown surface conditions (temperature or heat flux) using internal temperature measurements is known as inverse heat conduction problem (IHCP. The IHCP is an ill-posed problem due to the lack of continuous dependence of the solution on the data. A small error in input data can significantly affect the outputs. Therefore, an appropriate regularization method needs to be applied to convert the ill-posed problem to a nearby well-posed problem which can be solved. Several techniques have been proposed and applied for solving IHCPs which can be found in references namely Beck [1], Alifanov [2], Ozisik and Orlande [3] and Murio [4]. Some of these

methods include the least-square method with regularization, the sequential function specification, conjugate gradient method and numerical approaches [1].
Heat conduction in multi-layer mediums has been discussed in several papers. Ozisik [5] studied conduction in one dimensional composite media using orthogonal expansions, Green's functions and Laplace transform. De Monte [6] studied the transient response of one-dimensional multilayered composite conducting slabs to sudden variations of the temperature of the surrounding fluid. An analytical method for solving multi-layer heat conduction problems using Laplace transform and separation of variables is developed by Lu et al. [7]. They show that the result from their proposed closed form solution is in good agreement with numerical techniques. Haji-Sheikh and Beck [8] studied the temperature field in multi-dimensional, multi-layer bodies for the boundary conditions of the first, second and third kind. A solution for transient heat conduction through a one-dimensional three-layer composite slab is proposed by Sun and Wichman [9].
A few studies discussed the solution of IHCPs in multi-layer medium. Al Najem and Ozisik [10] conducted an inverse heat conduction analysis using a splitting-up procedure and nonlinear least-squares technique for the whole time domain and estimated the surface condition in composite layers. Ruan et al. [11] used least square method and Beck's future time method for 1-D and 2-D geometries and calculated the unknown boundary cooling condition and contact heat transfer coefficient for solidification of alloys. The design of optimal transient heat conduction experiments on composite orthotropic materials is studied by Taktak et al. [12]. Al-Najem [13] developed a method of analysis for determining surface conditions from the knowledge of the time variations of the temperature at the insulated boundary. He used two segmented polynomial in time for the unknown surface temperature. An inverse solution is then developed over the whole time domain using the splitting-up procedure.
The necessity for real time heat flux measurement in variety of industrial applications attracted a lot of attentions to develop real-time solutions for IHCPs. A filter solution based on the idea of training neural networks is studied by Kowsari et al. [14]. Ijaz, et al. [15], used a Kalman filter to solve a two-dimensional transient IHCP. Feng et al. [16] used Laplace transforms to relate the measured conditions at one end of a domain to the unknown conditions at the remote surface. Woodbury and Beck [17] studied the structure of the Tikhonov regularization problem and concluded that the method can be interpreted as a sequential filter formulation for continuous processing of data. They show that the computed heat fluxes using the whole domain solution and the filter coefficient solution are virtually the same for the constant-property solutions.

While in most of the IHCP studies the remote boundaries is assumed as an insulated surface or cooled with a known heat transfer coefficient, e.g. [15, 18, 19], this is not always the case in practice. Woodbury et al. [20] developed a filter based solution to incorporate the temperature measurement history from a second subsurface sensor as a remote boundary condition in an IHCP solution. In real world problems, the material thermal properties can greatly vary during the heating/cooling process due to significant temperature changes. This paper presents a filter based solution for two-layer mediums when the material thermal properties are temperature dependent. For this purpose, two IHCPs are solved (one for each layer) and a coupled solution is determined and tested to estimate the unknown heat flux at the surface of the front layer. The solution is then written in a digital filter form and filter coefficients are calculated for different temperatures and corresponding material thermal properties. An artificial neural network is then developed and trained to interpolate the filter coefficients at every time step. The proposed solution is then verified through several numerical experiments using exact solutions and ANSYS simulation. The filter solution of the IHCP has several advantages including simplicity, continuous operation and application to moderate nonlinearity [21] which makes it an appropriate approach for real time heat flux estimation in industrial applications.

2 Problem Description

A two-layer slab is considered to demonstrate the application of the proposed approach. A schematic of a two-layer slab is shown in Fig. 1. The temperature measurement histories are available for $x=x_1$ and $x= x_2$ on layer 2 while no specific temperature/heat flux measurement is available on layer 1. The heat flux at the remote surface of the first layer (x=0) is unknown and to be determined by the proposed solution. An IHCP is solved for each layer, starting from the one with known temperature measurements (layer #2 in here), and the heat flux is estimated at the interface with the next layer.

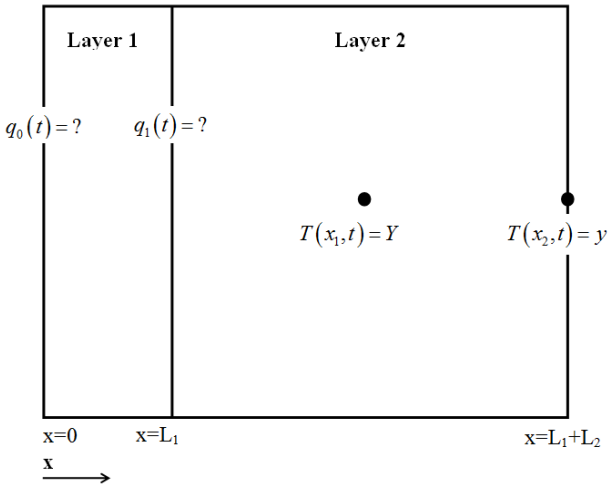


Figure 1: Schematic of the two-layer problem

3 Solution Scheme

A schematic of the problem is given in Fig. 1. The solution is started in the second layer, where two temperature measurements are available at x_1 and $x_2=L_1+L_2$ and q1 is the unknown heat flux. After solving the first IHCP, the heat flux and temperature are both known at the interface ($\mathbf{q_1}$ and $\mathbf{T_1}$). These values will be used as boundary conditions to solve the second IHCP associated with the first layer, where $\mathbf{q}$ at x=0 ($\mathbf{q_0}$) is unknown. The analysis of each layer is explained in detail as below.

3.1 Second Layer

The analysis starts with the second layer. The solution for the IHCP when the temperature measurement is given at two sub-surface locations is given by Woodbury et al. [20]. A similar approach is utilized in here to analyze the second layer. It is considered that the two temperature measurements histories are available at $x=x_1$ and $x= L_1+L_2$ ($L_1 \le x_1 \le L_1+L_2$) on the layer 2:

$$T(x_1,t) = Y(t) \tag{1}$$

$$T(L_1+L_2,t) = y(t) \tag{2}$$

The initial temperature is considered to be zero,

$$T(x,0) = 0 \tag{3}$$

The objective of the first IHCP is to estimate the heat flux ($\mathbf{q_1}$) at the interface between the two layers:

$$-k\frac{\partial T}{\partial x}(0,t) = q_1(t) = ? \tag{4}$$

Solving the IHCP for the second layer results in determining the heat flux at the interface with the first layer which then will be used as a boundary condition for solving the second IHCP.

The temperature at x_1 ($L_1 \le x_1 \le L_1+L_2$) is a function of the surface heat flux $\mathbf{q_{1(t)}}$ and the temperature at $x = L_1+L_2$. This problem is known as X21B10T0, which is herein shortened to X21 for reference [22]. The notation denotes a Cartesian geometry subjected to a type 2 condition at the first boundary and a type 1 condition at the second boundary, and that the first boundary has a step change in value while the second boundary is homogeneous, and that the initial condition is also homogeneous.
The connecting curves between the heat flux components, q_i and q_{i+1}, and also between the adjacent components, y_i and y_{i+1} are considered as constant between points (step function):

$$q(t) = q_i \ , \quad t_i < t < t_{i+1} \tag{5}$$

$$t_i = i\Delta t \tag{6}$$

A solution for the X21 case with a constant heat flux, q_c, at $x = L_1$ is

$$\frac{T_{X21}(x,t)}{\frac{q_c L_2}{k}\frac{L_2^2}{\alpha \Delta t}} = \left(1 - \frac{x}{L_2}\right) - \sum_{m=1}^{\infty} \frac{\cos\left(\beta_m \frac{x}{L_2}\right)}{\beta_m^2} \exp\left(-\beta_m^2 \frac{\alpha t}{L_2^2}\right) \tag{7}$$

where $\beta_m = (m - 1/2)\pi, \ \ m = 1, 2, \ldots$.

Analogous to the above equations for a constant heat flux, the equation for a constant temperature, T_C, at $x = L_1+L_2$ is denoted X12B10T0 and shortened herein to X12:

$$\frac{T_{X12}(x,t)}{T_c \frac{L_2^2}{\alpha \Delta t}} = 1 - 2\sum_{m=1}^{\infty} \frac{\sin\left(\beta_m \left(\frac{x}{L_2}\right)\right)}{\beta_m} \exp\left(-\beta_m^2 \frac{\alpha t}{L_2^2}\right) \tag{8}$$

The temperature at any location x_1 ($L_1 \le x_1 \le L_1+L_2$) caused by the heat flux $\mathbf{q_1}$ at $x = L_1$ and the temperature y at $x = L_1+L_2$ is

$$T_M = T_{q,M} + T_{y,M} = \sum_{i=1}^{M} q_{1,i} \Delta\phi_{M-i} + \sum_{i=1}^{M} y_i \Delta\eta_{M-i} \tag{9}$$

where ϕ and η are the response basis functions for the two cases. That is,

$$\phi(x,t) = \frac{\partial T_{X21}}{\partial q_c}; \qquad \eta(x,t) = \frac{\partial T_{X12}}{\partial T_c} \tag{10}$$

The $\Delta\phi$'s can be found as [1]:

$$\Delta\phi_0 = \phi_1, \ \ \Delta\phi_1 = \phi_2 - \phi_1, \ \ldots \ , \ \ \Delta\phi_i = \phi_{i+1} - \phi_i \tag{11}$$

Equation 9 can be described by the matrix equation of

$$\mathbf{T} = \mathbf{Xq} + \mathbf{Zy} \tag{12}$$

where

$$\mathbf{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix}, \ \mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix}, \ \mathbf{X} = \begin{bmatrix} X_1 & 0 & \cdots & 0 & 0 \\ X_2 & X_1 & \cdots & 0 & 0 \\ X_3 & X_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ X_n & X_{n-1} & \cdots & X_2 & X_1 \end{bmatrix}, \ \mathrm{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \ \mathrm{Z} = \begin{bmatrix} Z_1 & 0 & \cdots & 0 & 0 \\ Z_2 & Z_1 & \cdots & 0 & 0 \\ Z_3 & Z_2 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ Z_n & Z_{n-1} & \cdots & Z_2 & Z_1 \end{bmatrix} \tag{13 a,b,c,d,e}$$

The components of the $\mathbf{X}$ and $\mathbf{Z}$ matrices are related to the response basis functions by

$$X_i = \frac{L}{k}\Delta\phi_{i-1}, \quad Z_i = \Delta\eta_{i-1} \tag{14}$$

where $\Delta\phi_i$ is defined in Eq. 11, and $\Delta\eta_i$ is defined analogously.
The columns in the **X** matrix of Eq. 12 are the same as the sensitivity vectors in parameter estimation. The first column is for q at t1, the second for **q** at t_2 and so on. Using whole domain Tikhonov regularization, the estimated value heat flux vector at the interface, $\hat{\mathbf{q}}_1$ can be given as:

$$\hat{\mathbf{q}}_1 = [\mathbf{X}_2{}^T\mathbf{X}_2 + \alpha_T\mathbf{H}^T\mathbf{H}]^{-1}\mathbf{X}_2{}^T(\mathbf{Y}-\mathbf{Zy}) = \mathbf{F}_2(\mathbf{Y}-\mathbf{Zy}) \tag{15}$$

where the α_T is the Tikhonov regularization parameter [1]. Here the subscript 2 for **X** and **F** refers to Layer 2. Note that this equation is dimensional. $\mathbf{F_2}$ has the same definition as the filter matrix of Ref [17]. The value of $\hat{\mathbf{q}}_1$ from Eq. 15 will be used as a known boundary condition for the IHCP associated with the first layer. The heat flux at the remote surface (x=0) is the unknown parameters in the second IHCP.

3.2 First Layer

The temperature at any location x_1 ($0 \le x_1 \le$ L1) is a function of the surface heat fluxes $q_{0(t)}$ and $q_{1(t)}$. The results from the middle layer analysis determine q_1(t). A solution for the **X22** case with a constant heat flux at x = 0 and zero heat flux at $x=L_1$ (this is denoted the X22B10T0 case) is [22]:

$$22\,\frac{T_{X\,22}(x,t)}{q_c L_1/k} = \frac{\alpha t}{L_1^{\,2}} + \frac{1}{3} - \frac{x}{L_1} + \frac{x^2}{2L_1^{\,2}} - 2\sum_{m=1}^{m_{\max}} \frac{\cos(\beta_m x/L_1)}{\beta_m^2}\exp\left(-\beta_m^2\alpha t/L_1^{\,2}\right) \tag{16}$$

Note the same solution applies for a zero heat flux at x=0 and a constant heat flux at $x=L_1$ through a simple change of variables such as $\xi = L_1 - x$. Analogous to Eq. (9), the temperature response at x_1 ($0 \le x_1 \le L_1$) can be found due to the heat flux histories $\mathbf{q_{0(t)}}$ and $\mathbf{q_{1(t)}}$. For assumed piecewise constant variation in these heat fluxes, the temperature can be computed from

$$T_M = T_{q_0,M} + T_{q_1,M} = \sum_{i=1}^{M} q_{0i}\Delta\varphi_{M-i} + \sum_{i=1}^{M} q_{1i}\Delta\theta_{M-i} \tag{17}$$

where

$$\varphi(x_1,t) = \frac{\partial T_{X\,22}(x_1,t)}{\partial q_c}; \quad \theta(x_1,t) = -\frac{\partial T_{X\,22}(L_1-x_1,t)}{\partial q_c} \tag{18a,b}$$

The step basis function representation used here (and also others) for temperature given in Eq. (17) can be described by the matrix equation of

$$\mathbf{T} = \mathbf{X}_0\mathbf{q}_0 + \mathbf{X}_{L_1}\mathbf{q}_{L_1} \tag{19}$$

where

$$\mathbf{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix}, \quad \mathbf{q}_0 = \begin{bmatrix} q_{01} \\ q_{02} \\ \vdots \\ q_{0n} \end{bmatrix}, \quad \mathbf{q}_L = \begin{bmatrix} q_{L1} \\ q_{L2} \\ \vdots \\ q_{Ln} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} X_1 & 0 & 0 & \cdots & 0 & 0 \\ X_2 & X_1 & 0 & \cdots & 0 & 0 \\ X_3 & X_2 & X_1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ X_n & X_{n-1} & X_{n-2} & \cdots & X_2 & X_1 \end{bmatrix} \tag{20a,b,c,d}$$

The components of the $\mathbf{X_0}$ and $\mathbf{X}_{L_1}$ matrices are related to the response basis functions in Eq. (18) similar to the relation of **X** and **Z** in Eq. (14) to Eq. (10).
The whole domain Tikhonov regularization method is used to solve the IHCP. Later the filter coefficients are found from the solution. The IHCP solution for Layer 1 starts with a matrix form for the sum of squares with an added regularization term given by

$$S = (\mathbf{T_1} - \mathbf{X}_0\mathbf{q}_0 - \mathbf{X}_{L_1}\mathbf{q}_1)^T(\mathbf{T_1} - \mathbf{X}_0\mathbf{q}_0 - \mathbf{X}_{L_1}\mathbf{q}_1) + \alpha_T\mathbf{q}_0^T\mathbf{H}^T\mathbf{H}\mathbf{q}_0 \quad (21)$$

This is minimized with respect to the parameter vector $\mathbf{q_0}$. The symbol $\mathbf{T_1}$ is the temperature vector at L_1 and $\mathbf{q_1}$ is the known heat flux (from the solution of the IHCP in Layer 2) at $x = L_1$. The initial temperature is zero. The α_T symbol is the Tikhonov regularization parameter.

The estimated value of the heat flux vector, denoted $\hat{\mathbf{q}}_0$, is then given by

$$\hat{\mathbf{q}}_0 = [\mathbf{X}_0^T\mathbf{X}_0 + \alpha_T\mathbf{H}^T\mathbf{H}]^{-1}\mathbf{X}_0^T\left(\mathbf{T_1} - \mathbf{X}_{L_1}\mathbf{q}_1\right) = \mathbf{F}_1\left(\mathbf{T_1} - \mathbf{X}_{L_1}\mathbf{q}_1\right) \quad (22)$$

3.3 Coupling of the Solutions

To achieve a single expression for a two-layer IHCP, Eq (15) is substituted in Eq (22). An expression is:

$$\hat{\mathbf{q}}_0 = \mathbf{F}_1\mathbf{T}_1 - \mathbf{F}_1\mathbf{X}_{L_1}\left(\mathbf{F}_2\mathbf{T_1} - \mathbf{F}_2\mathbf{Z}\mathbf{y}\right) = \mathbf{F}_1\left(\mathbf{I} - \mathbf{X}_{L_1}\mathbf{F}_2\right)\mathbf{T}_1 + \mathbf{F}_1\mathbf{X}_{L_1}\mathbf{F}_2\mathbf{Z}\mathbf{y} \quad (23)$$

where $\mathbf{T}_1$ is the temperature at the interface and can be found from Eq. 12:

$$\mathbf{T}_1 = \mathbf{X}_2\mathbf{q}_1 + \mathbf{Z}\mathbf{y}$$

By substituting $\mathbf{q}_1$ in the above equation $\mathbf{T}_1$ can be found as:

$$\mathbf{T}_1 = \mathbf{X}_2\mathbf{F}_2\mathbf{Y} + \left(\mathbf{Z} - \mathbf{X}_2\mathbf{F}_2\mathbf{Z}\right)\mathbf{y} \quad (24)$$

Substituting Eq. 25 in Eq. 23, an expression is:

$$\hat{\mathbf{q}}_0 = \left[\mathbf{F}_1\left(\mathbf{X}_2\mathbf{F}_2 - \mathbf{X}_{L_1}\mathbf{F}_2\mathbf{X}_2\mathbf{F}_2\right)\right]\mathbf{Y} + \left[\mathbf{F}_1\left(\mathbf{Z} - \mathbf{X}_2\mathbf{F}_2\mathbf{Z} + \mathbf{X}_{L_1}\mathbf{F}_2\mathbf{X}_2\mathbf{F}_2\mathbf{Z}\right)\right]\mathbf{y} \quad (25)$$

Equation 25 can be used directly to calculate the heat flux at the remote surface on the first layer by using two sets of temperature measurements from the second layer.

3.4 Filter Form of the Solution

The concept of the filter algorithm is that the solution for the heat flux at any time is only affected by the recent temperature history and a few future time steps. Equation (25) can be written in filter form as:

$$\hat{\mathbf{q}}_0 = f\mathbf{Y} + g\mathbf{y} \quad (26)$$

$$f = row\left[\mathbf{F}_1\left(\mathbf{X}_2\mathbf{F}_2 - \mathbf{X}_{L_1}\mathbf{F}_2\mathbf{X}_2\mathbf{F}_2\right)\right] \;,\; g = row\left[\mathbf{F}_1\left(\mathbf{Z} - \mathbf{X}_2\mathbf{F}_2\mathbf{Z} + \mathbf{X}_{L_1}\mathbf{F}_2\mathbf{X}_2\mathbf{F}_2\mathbf{Z}\right)\right] \quad (27)$$

where f and g have the same characteristics as filter coefficients, which arediscussed in detail in Ref [17]. The meaning of "row()" in Eq. (27) designates a row in the middle of the indicated matrix. Note that these filter coefficients are not the same as filter factors suggested by Hansen [23].

All the f-filter coefficients can be found at one time by setting all the **Y** and **y** components equal to zero except the m_f component of **Y** is set equal to one (Y_{mf}=1). The solution of the IHCP with this data gives the f coefficients. To get the g-filter coefficients (those for the y vector), the same procedure is followed with now all the components of **Y** and y equal to zero except y_{mf}=1.

Assuming x_1=L_1, the dimensionless time step of 0.0052 (0.1 s) and for first order Tikhonov regularization with parameter α_T = 0.0001, and material properties given in Table. 1, the 150th rows of the matrices in Eq. (27) (the *f* and *g* filter coefficients) are plotted in figures 2 and 3.

Table 1: material properties

	Layer 1 (Inconel)	**Layer 2 (Ceramic Fiber)**
Thickness, m	0.01	0.02
T= 300 K	K= 14.9 W/m-K, α=3.9e-6 m^2/s	K= 0.046 W/m-K, α= 3.2e-7m^2/s
T= 450 K	K= 17.4 W/m-K, α= 4.2e-6m^2/s	K= 0.065W/m-K, α=3.9e-7 m^2/s
T= 800 K	K=22.6 W/m-K, α=4.9e-6 m^2/s	K=0.13 W/m-K, α=6.9e-7 m^2/s
T= 1000 K	K=25.4 W/m-K, α=5.3e-6 m^2/s	K=0.19 W/m-K, α=7.9e-7 m^2/s
T= 1200 K	K= 28.2W/m-K, α= 5.57e-6m^2/s	K=0.27 W/m-K, α= 1.2e-6.m^2/s

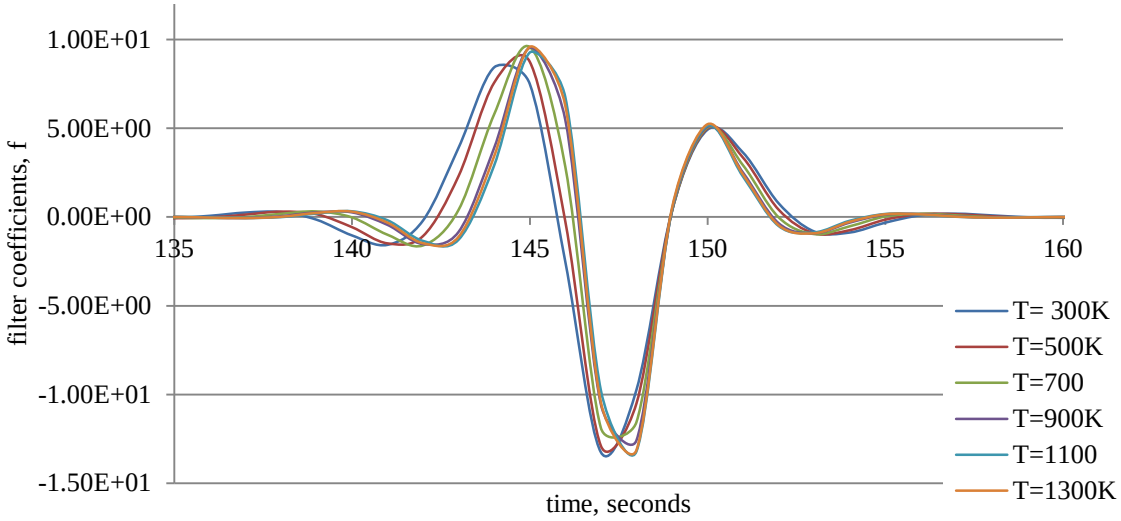


Figure 2: filter coefficients (*f*) for different temperatures

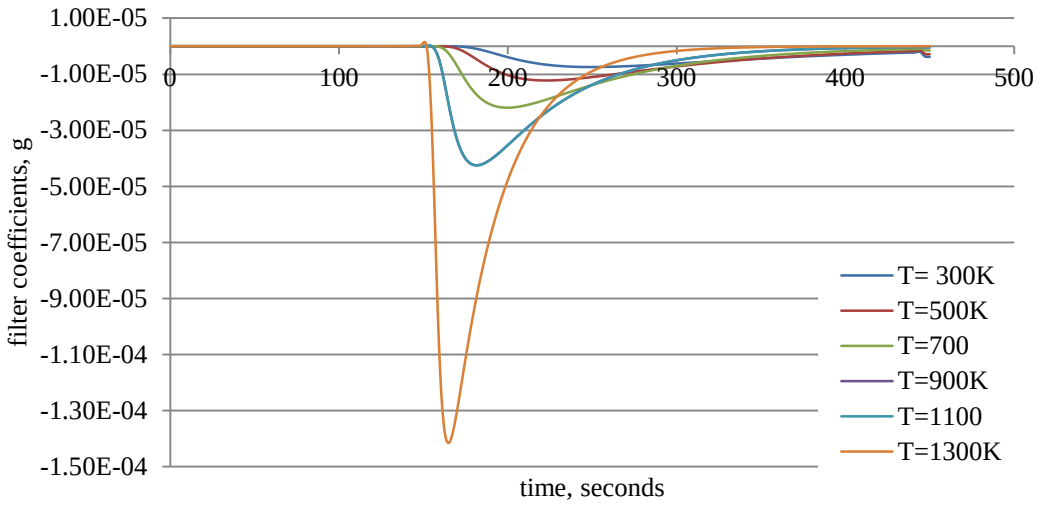


Figure 3: filter coefficients (*g*) for different temperatures

As can be seen, the filter coefficients can greatly vary as the temperature changes.

3.5 Filter Coefficients for the Case of Temperature Dependent Materials

If the material properties are temperature dependent, the filter coefficients are changing through the heating/cooling process as the temperature varies (as shown in Fig. 2 and 3). Therefore, it is necessary to find the filter coefficients at each time step with a particular temperature and use them accordingly. This can be done by finding the filter coefficients for a set of temperatures and linearly interpolate between those at each time step. However, this technique turned out to be time consuming. Therefore, a fast and accurate method is developed to find filter coefficients at each temperature by using Artificial Neural Networks (ANNs). ANN's are computational models inspired from human brain system which has been successfully used in several engineering applications namely time series prediction, pattern recognition, function approximation, classification and more. They have also been used to solve direct and inverse heat conduction problems [24, 25, 26, 27, 28].
ANN consisted of a set of interconnected neurons that can evaluate outputs from inputs by feeding information through the network and adjusting the weights. In the present work (test case 2), a feed forward multi-layered network is used which consists of a layer of input neurons (including temperature data and time step), a layer of output neurons (filter coefficients) and two hidden layers. Data enter the network through the input nodes and going through a non-linear transformation. The output data are subsequently generated by the output nodes. The inputs of the network are the temperature and the time step. Since the filter coefficients beyond m_p+m_f are all zero, the interpolations performed only for the time steps 1 through m_p+m_f. A schematic of the architecture of a three-layered neural network used in this work to calculate f filter coefficients is shown in Figure 4. A similar structure is used to calculate g filter coefficients.

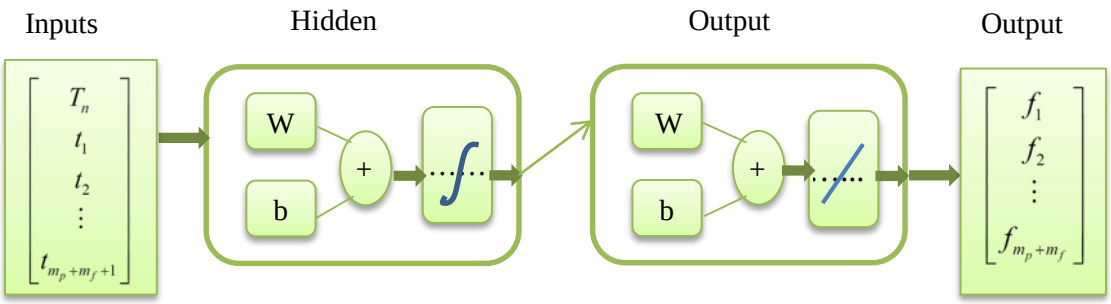


Figure 4: Neural network structure

4 Results and discussion

The developed method is verified via numerical experiments and ANSYS simulation. The verification of the method is described in this section. To verify the developed filter solution, a two layer slab is modeled in ANSYS. The material properties of the layers are given in Table 2.

Table 2: Material properties for test case 1

Parameter	Layer 1 (Steel)	Layer 2 (Aluminum)
Density, kg/m^3	7833	2702
Specific Heat, J/kg.K	465	903
Conductivity (k), W/m.K	54	237
Thickness (L), m	0.01	0.02

To validate the ANSYS model, the results from ANSYS simulation is compared with exact solution [29] for constant material properties. The created model is meshed and the loads are applied on the

geometry similar to the X1C11B10T0 case. A step change in temperature is applied at the front surface and the back surface temperature is kept at zero.

The sensor locations are at $x=L_1$ and $x=L_1+L_2$. The temperature data is also generated using the solution for the direct problem for a step change in temperature at the surface (X1C11B10T0) and used as inputs for the filter solution (Eq. 26) to determine the heat flux at the surface. The calculated heat flux by the filter solution, ANSYS simulation and exact solution are compared in Fig. 5 and a good agreement can be observed between all three set of results.

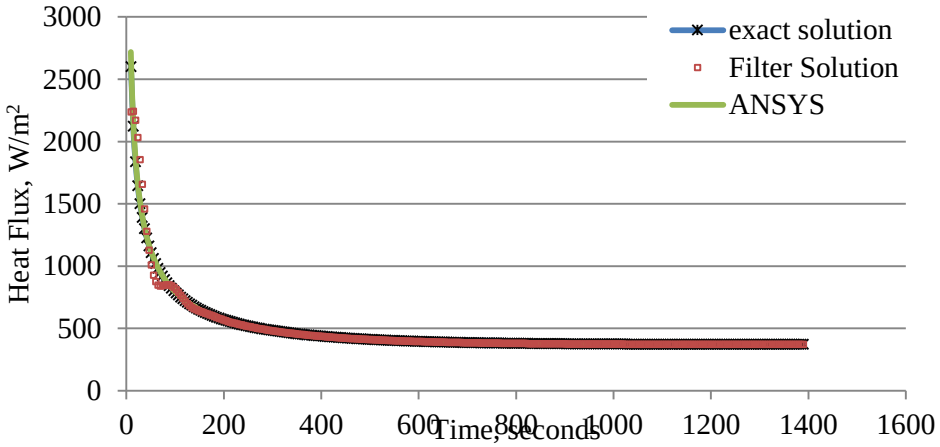


Figure 5: heat flux estimation for test case 1

It should be noted that the values of m_p and m_f are considered as 60 for this test case. Afterwards, a random error (0.5% of the temperature) is added to data and the heat flux is calculated. The calculated heat flux when error presents in the data is plotted in Fig. 6. It can be observed that the results are still in a good agreement with the exact solution. It should be noted that the first test case only examined the developed approach for constant material properties. The next test case has taken into account the temperature dependent material thermal properties.

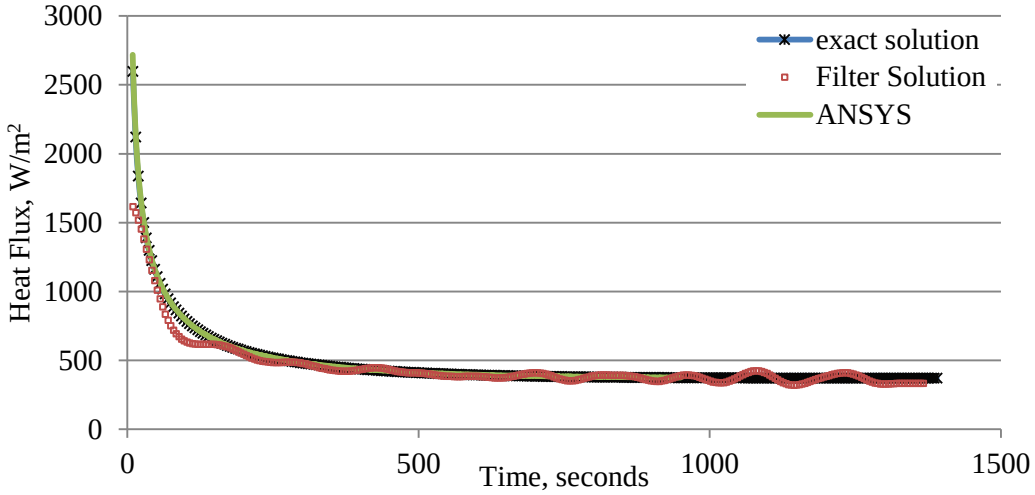


Figure 6: heat flux estimation for test case 1 with 0.5% error added to temperature data

In a second test case, a triangular heat flux profile is applied to the surface of the two-layer slab. It is assumed that geometry of the two-layer medium is similar to the previous test case. The material

properties, however, are changing with temperature as it is shown in Table 1. The temperature histories are then obtained at the desired surfaces using ANSYS and used as inputs for the filter solution. The calculated temperature profiles are shown in Fig. 7.

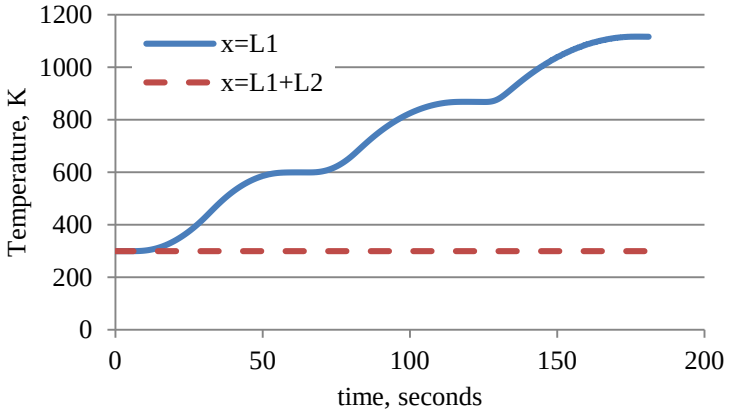


Figure 7: Temperature profile for the second test case

Using filter coefficients for constant material properties, the heat flux profile is estimated and shown in Fig. 8. The Tikhonov parameter is considered as 0.0001and the time step is 1 second.

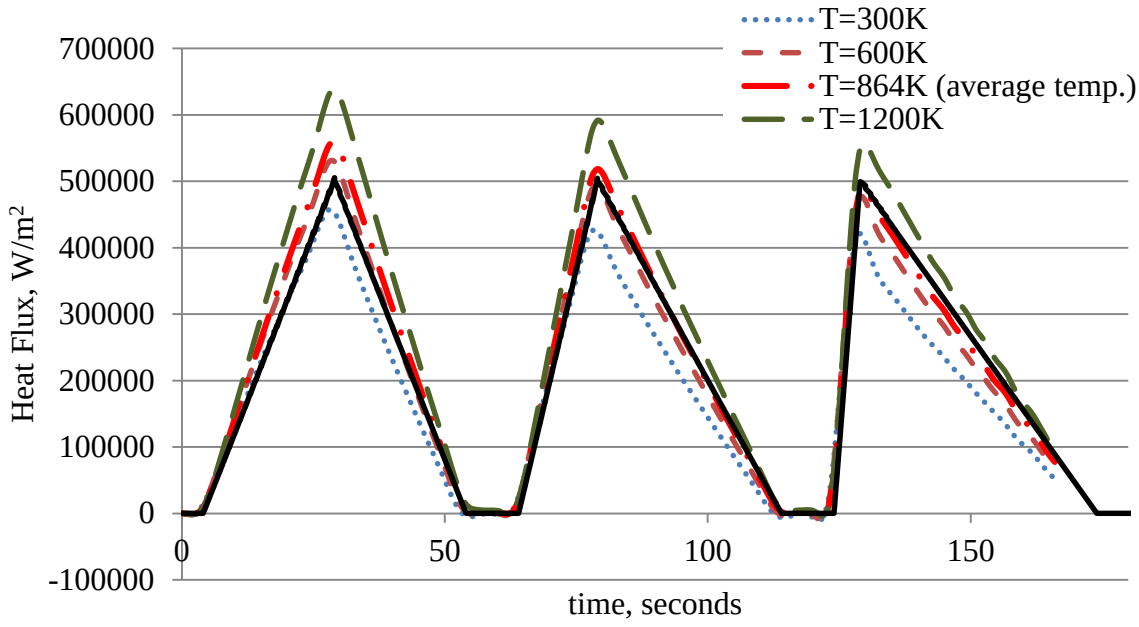


Figure 8: heat flux estimation using filter method-assuming constant thermal properties

As can be seen, the actual heat flux is significantly different from the estimated heat fluxes by the proposed method. The closest estimation is the one which uses average temperature of 864 K which still significantly off during the first 50 seconds, when the temperature is below the average temperature. This can clearly show the importance of accounting for variation of thermal properties due to temperature changes.

Next, using the developed neural networks, the filter coefficients are found for each time step based on the temperature. Each of the developed ANN's for f and g includes 2 hidden layers, 31 inputs ($1+m_p+m_f$) including one temperature and 30 time steps and 30 outputs (m_p+m_f) including the non-zero filter

coefficients. The network trained using 10 set of samples. (T=300,400,…,T=1200). The heat flux is then calculated using Eq. 26 and compared with the given heat flux at the surface in Fig. 9. As can be seen, the heat flux profile is very close to the heat flux from the ANSYS model and can do significantly better than the calculated heat flux by the filter coefficients from average temperature.

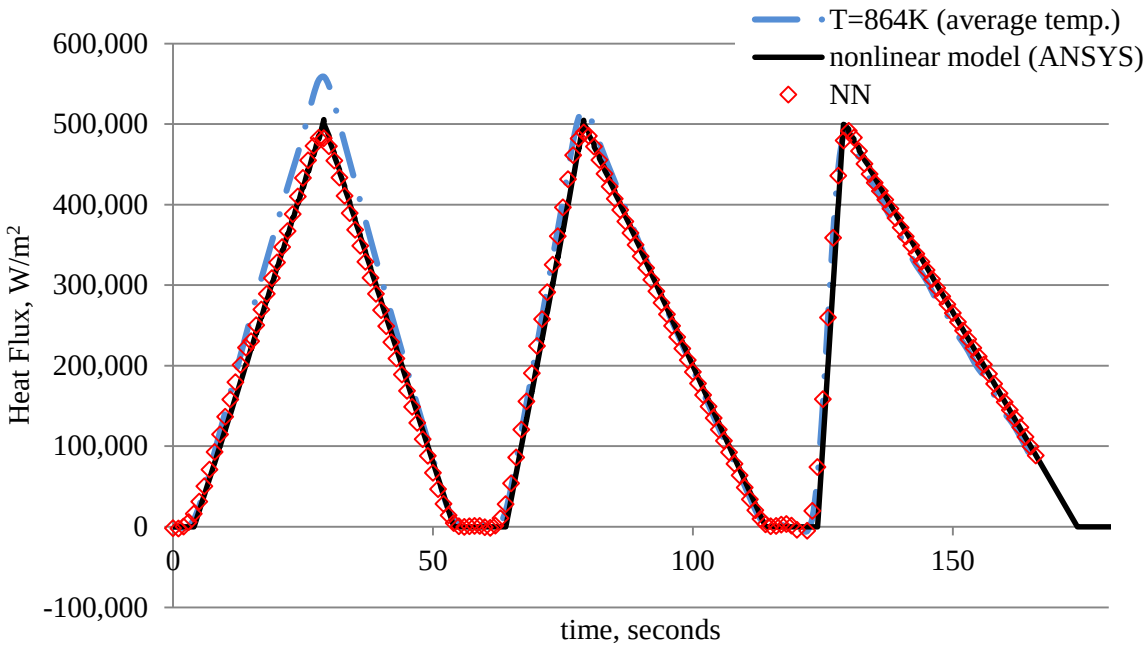


Figure 9: Heat flux estimation using filter solution and neural networks

5 Conclusion

A method for solving one dimensional non-linear IHCP in two-layer mediums with temperature dependent material thermal properties is developed and successfully tested. The method is discussed for a two-layer slab when the temperature measurement history is given in two interior locations of one layer. An IHCP is solved for each layer based on the minimization of the sum of the squared errors between the computed and known values and by using of Tikhonov Regularization (TR) for stabilizing the solution. The developed algorithm is then written in filter form. The filter form solution allows near real time heat flux estimation which can be used in several industrial applications. The proposed solutions are then validated by numerical experiments. Two numerical test cases are developed in ANSYS, the first one with constant thermal properties and the second one with temperature dependent thermal properties. The heat flux at the surface is obtained using the proposed solution and compared with the ANSYS simulation and also exact solution. For the second test case, two neural networks are developed to estimate the filter coefficients f and g at each time step and its corresponding temperature. The filter coefficients are then used to calculate the heat flux at the surface by using the proposed solution. The results from the proposed solution is in a very good agreement with the ANSYS simulation. The results showed that when the temperature varies significantly, accounting for variation in thermal properties of the materials can greatly improve the accuracy of the estimated heat flux.

References

[1] J.V. Beck, B. Blackwell, C.R. St Clair Jr., Inverse Heat Conduction: Ill-PosedProblems, John Wiley and Sons, New York, 1985.
[2] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Heidelberg,1994.
[3] M.N. Ozisik, H.R.B. Orlande, Inverse Heat Transfer, Taylor & Francis, New York,2000.
[4] D.A. Murio, The Mollification Method and the Numerical Solution of Ill-Posed Problems, John Wiley & Sons, New York, 1993.
[5] M.N. Ozisik, Heat Conduction, second ed., John Wiley & Sons, Inc., New York, 1993.
[6] F. de Monte, Transient heat conduction in one-dimensional composite slab.A `natural' analytic approach, International Journal of Heat and Mass Transfer 43 (2000) 3607-3619.
[7] X. Lu, P. Tervola, M. Viljanen , Transient analytical solution to heat conduction in multi-dimensional composite cylinder slab, International Journal of Heat and Mass Transfer 49 (2006) 1107–1114.
[8] A. Haji-Sheikh, J.V. Beck, Temperature solution in multi-dimensional multi-layer bodies, International Journal of Heat and Mass Transfer 45 (2002) 1865–1877.
[9] Y. Sun, I.S. Wichman, On transient heat conduction in a one-dimensional composite slab, International Journal of Heat and Mass Transfer 47 (2004) 1555–1559.
[10] N. M Al-Najem, M. N Ozisik, Estimating unknown surface condition in composite media. Int. Comm. Heat Mass Transfer 19 (1991) 69±77
[11] Y. Ruan, J. C.Liu, O. Richmond, Determining the unknown cooling condition and contact heat transfer coefficient during solidifcation of Alloys. Inverse Problems in Engineering 1 (1994) 45-69
[12] Taktak, R.; Beck, J. V.; Scott, E. P.: Optimal experimental design for estimating thermal properties of composite materials. Int. J. Heat Mass Transfer 36 (1993) 2977-2986.
[13] N. M. AL-Najem, Whole time domain solution of inverse heat conduction problem in multi-layer media, Heat and Mass Transfer 33 (1997) 233±240 Ó Springer-Verlag 1997.
[14] F. Kowsary, M. Mohammadzaheri and S. Irano, Training based, moving digital filter method for real time heat flux function estimation, International Communications in Heat and Mass Transfer, vol. 33, pp. 1291-1296, 2006.
[15] U. Ijaz, A. Khambampati, M.-C.Kim, S. Kim and K.-Y.Kim, Estimation of time-dependent heat flux and measurement bias in two-dimensional inverse heat conduction problems, International Journal of Heat and Mass Transfer, vol. 50, pp. 4117-4130, 2007.
[16] Z. C. Feng, J. K. Chen and Y. Zhang, Real-time solution of heat conduction in a finite slab for inverse analysis, International Journal of Thermal Sciences, vol. 49, pp. 762-768, 2012.
[17] K. A. Woodbury, J.V. Beck, *Estimation metrics and optimal regularization in a Tikhonov digital filter for the inverse heat conduction problem*, International Journal of Heat and Mass Transfer 62 (2013) 31–39
[18] H.-T. Chen, S.-Y.Lin, L.-C. Fang, *Estimation of surface temperature in two-dimensional inverse heat conduction problems*, Int. J. Heat Mass Transfer 44 (2001) 1455-1463
[19] J. M. Mulcahy, D. J. Browne, K. T. Stanton, F. R. C. Diaz , L. D. Cassady, D. F. Berisford, R. D. Bengtson, *Heat flux estimation of a plasma rocket helicon source by solution of the inverse heat conduction problem*, Int. Journal of Heat Mass Transfer 52 (2009) 2343-2357.
[20] K.A. Woodbury, J.V. Beck, H. Najafi, *Filter Solution of Inverse Heat Conduction Problem using Measured Temperature History as Remote Boundary Condition*, Int. Journal of Heat and Mass Transfer, Vol. 72, May 2014, Pages 139–147
[21] J.V. Beck, *Filter solutions for the nonlinear inverse heat conduction problem*, Inverse Problems in Science and Engineering, Vol. 16, No. 1, January 2008, 3–20.
[22] K.D. Cole, J.V. Beck, A. Haji-Sheikh, , B. Litkouhi, *Heat Conduction Using Green's Functions*, Taylor & Francis, 2nd Edition, 2010.
[23] P. C. Hansen, *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, SIAM, 1997

[24] M. Raudensky, J. Horsky, J. Krejsa, L. Slama, *Usage of artificial intelligence methods in inverse problems for estimation of material parameters*, Int. J. Numer. Methods Heat Fluid Flow 6 (8) (1996) 19–29.
[25] S. Deng, Y. Hwang, *Applying neural networks to the solution of forward and inverse heat conduction problems*, Int. J. Heat Mass Transfer 49 (25–26) (2006) 4732–4750.
[26] S.S. Sablani, A. Kacimov, J. Perret, A.S. Mujumdar, *A. Campo, Non-iterative estimation of heat transfer coefficients using artificial neural network models*, Int. J. Heat Mass Transfer 48 (3–4) (2005) 665–679.
[27] C. Balaji, T. Padhi, *A new ANN driven MCMC method for multi-parameter estimation in two-dimensional conduction with heat generation*, Int. J. Heat Mass Transfer 53 (23–24) (2010) 5440–5455.
[28] B. Czél, K.A. Woodbury, G. Gróf, *Simultaneous estimation of temperature-dependent volumetric heat capacity and thermal conductivity functions via neural networks*
[29] ExACT Analytical Conduction Toolbox, http://exact.unl.edu/exact/home/home.php

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Identification of radiative heat transfer parameters in multilayer thermal insulation of a spacecraft

Aleksey V. Nenarokomov[1], Leonid A. Dombrovsky[2], Irina V. Krainova[1], Oleg M. Alifanov[1], Sergey A. Budnik[1]

[1] Moscow Aviation Institute, Moscow, Russia
[2] Joint Institute for High Temperatures, NCHMT, Moscow, Russia

KEY WORDS: Radiation, Thermal insulation, Identification in spectral domain, Fibrous spacer, Vacuum experiment.

Abstract
The study is motivated by necessity to optimize multilayer vacuum thermal insulation (MLI) of modern high-weight spacecraft. The modern approaches to the design of space structures assume broad application of physical models and computational methods. The latter is impossible if information on physical properties is insufficient. In many cases, both theoretical prediction and direct measurements of radiative properties of composite materials is very problematic. There is only one way to overcome these complexities. It is the use of some indirect measurements. Mathematically, such an approach is usually formulated as a solution of the inverse problem: to retrieve the properties from the measurements of thermal characteristics of the insulation. The experiments in thermo-vacuum facilities are used to re-estimate some radiative properties of metallic foil / metallized polymer foil and spacer. The specimens of a real MLI of the BP-Colombo satellite (ESA) were examined in the experiments. The retrieved values of the effective emissivity can be used directly for heat transfer calculations in the case of similar thermal conditions. Therefore, it is important to use the identified parameters to validate and improve theoretical models. The latter is considered as an objective of our further work.

1. INTRODUCTION

The present study is a part of a work on optimization of passive thermal control systems (TCS) for space vehicles. The specific of the external heating of space vehicles during the flight enables us to consider some variants of passive TCS based on screening the vehicle surface from the direct solar radiation, the solar radiation reflected by the planets, and also from thermal radiation of closely spaced planets. Various multilayer insulations (MLI) are widely used to solve this engineering problem [1]. This type of vacuum insulations has obvious advantages such as high thermal resistance at a relatively low density and convenience of their use for the surfaces of complex shape (Fig.1).

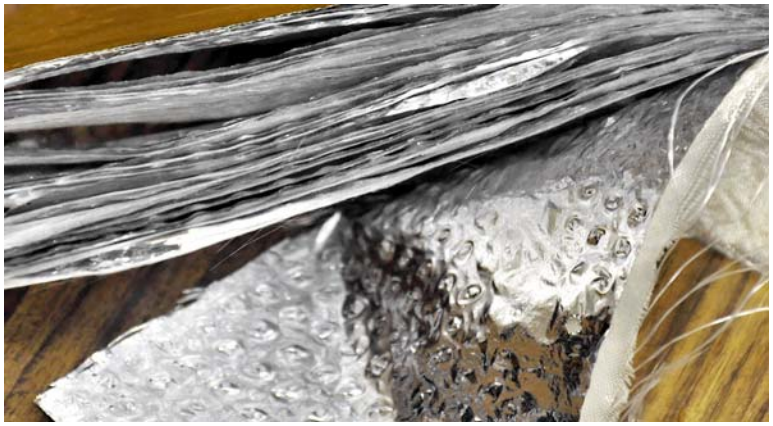
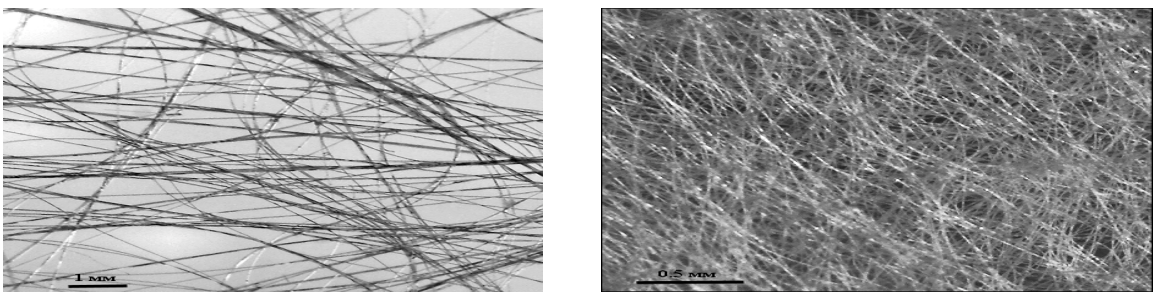


Fig. 1 A photo of the MLI sample.

A typical MLI looks as a set of thin metal screens of thickness about 5–9 μm with spacers between them. The spacer prevents a contact between the screens. Simple estimates show that the use of several layers (screens with spacers) may lead to a significant reduction of heat flux to the protected surface. MLI's with 10–30 shielding layers are usually used. The material choice for MLI depends on the expected level of temperature. The PET (polyethylene terephthalate) film coated with aluminum, silver, or gold can be used for screens at working temperatures of MLI up to 423 K. The aluminum foil with spacers of fiberglass is used at higher temperatures up to 723 K. At temperatures greater than 723 K, the foil is made of copper, nickel or steel and the spacer is made of quartz fibers. The surface density of ten screens of PET film is in the range of 0.2–0.3 kg/m^2, whereas the use of metal foils increases this value up to 1 kg/m^2 [1]. However, regardless of used materials, the principle of the MLI work is the same.

The traditional MLI thermal model [2] is very simplified. This model is based on the gray approximation and does not take into account the effects of both semi-transparent fibrous spacers between the foils and possible oxidation of aluminum foil. However the present-day computational models for radiative heat transfer in a single layer of the vacuum insulation are insufficient to solve this problem because the experimental data for the wide-range infrared radiative properties of materials are not quite reliable. Moreover, our preliminary estimates showed that the known physical effect of near-field radiative transfer between two closely spaced layers of aluminum foil may lead to a significant increase in the radiative flux. A theoretical prediction of this effect is very sensitive to the distance between the foils, and it is problematic to obtain accurate quantitative results for various flight conditions. In addition, the role of a highly-porous fibrous spacer between the foils (see Fig. 2) has been estimated in our recent papers by neglecting the near-field effect.



a b

Figure 2: Photographs of typical materials used for spacers between two foil layers:
a – material of a very low density, b – relatively dense material.

Obviously, the estimate of the spacer effect should be revised in a general theory taking into account a small thickness of the gap between the foils. This theoretical study is in the very beginning now. Therefore, the role of the inverse problem solution in obtaining the radiative flux at realistic conditions is considered as an important stage of the work. For brevity, we do not give here the complete mathematical problem statement, which can be found in our previous papers. At the same time, it should be noted that we are focused on identification of the conventional effective emissivity, which is the key integral parameter (over the spectrum) used to determine radiative heat transfer in the vacuum insulation of a spacecraft.

2. PHYSICAL MODEL AND EXPERIMENTAL FACILITY

The objective of the experimental study is to estimate total thermal resistance of the MLI blanket, which is a screen-vacuum thermal insulation elaborated for the temperature range from 300 to 900 K. Obviously, the accurate measurements for vacuum insulations are not simple. The space vacuum in the laboratory installation is not the only condition of successful measurements. The other problem is a very high thermal resistance of the MLI. We had no possibility to develop a specific method and use an additional facility to minimize an error of these measurements. Therefore, a previously developed vacuum installation TVS-1 is employed. Let us consider the experimental facility and the scheme of the temperature measurements. The MLI sample of size about 150×150×5 mm was placed in the experiment module EM-2 (see photo in Fig. 1 and scheme of the MLI in Fig. 3).

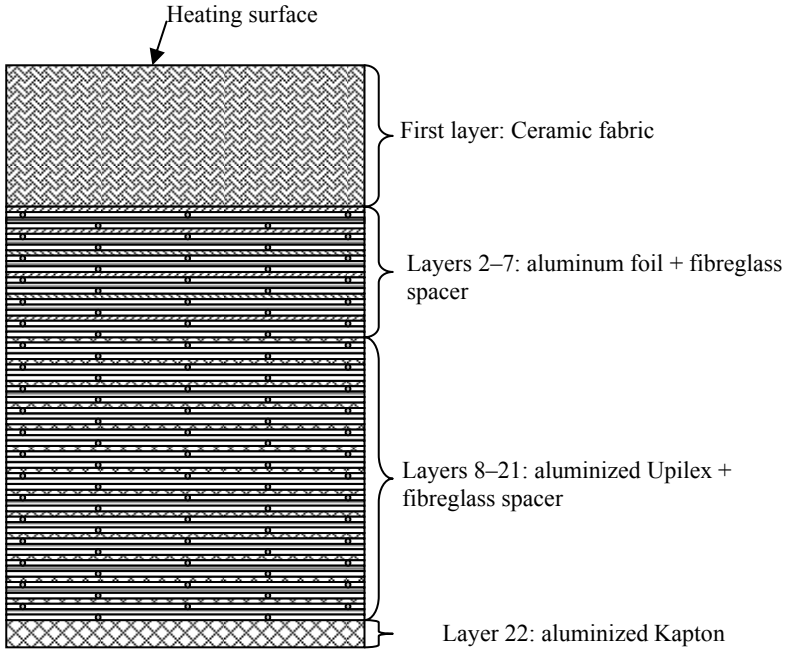


Figure 3: A scheme of the tested multilayered thermal-insulating blanket.

A scheme of temperature measurements with a set of thermocouples is presented in Fig. 4. The thermocouple T_1 is installed in the heated ceramic fabric. Two thermocouples T_6 and T_7 are installed

on the copper plate at the cold surface of sample A (T_7 is the standby thermocouple) and similar thermocouples T_8 and T_9 – in the copper plate at the cold surface of sample B (T_9 is the standby thermocouple). An electric heater made of a refractory stainless steel foil of 0.1 mm thickness is used. Elements of the insulating holder are arranged around the sample. The thermocouple conductors, covered by the glass-sleeve, are lead out through special ducts at the elements of holder. The setting frames press densely the sample in the insulating holder to the heater, and 5 mm in depth holders provide non-stress merit of the samples (preserve the initial thickness). The experimental module EM-2 is placed into a vacuum chamber of installation TVS-1. Also thermocouple T_2 is installed on the steel wall of installation cooled by water.

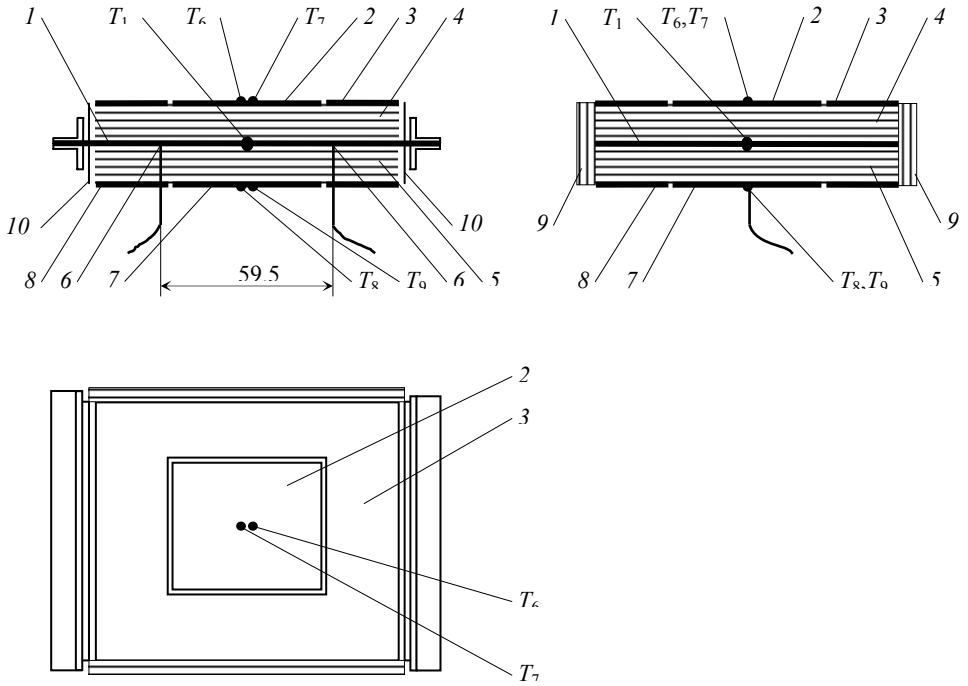


Figure 4: A scheme of temperature measurements:
1 – heating element of module EM-2 (steel foil); *2* – copper slab on specimen A; *3* – protection on specimen A; *4* – specimen A; *5* – specimen B; *6* – voltage measuring points on the heating element; *7* – copper slab on specimen B; *8* – protection of specimen B; *9* – elements of insulating holder of MLI; *10* – elements of insulating slab.

3. MATHEMATICAL MODEL

As was mentioned above, an improved theoretical model has been recently developed by the authors [3, 4]. A MLI is considered as the set of couples of screens and a spacer made of semi-transparent highly porous fibrous material between them. The resulting mathematical problem statement can be found in [4] and it is not reproduced below. At the same time, the main assumptions and special features of this approach should be briefly discussed. It was assumed that: (1) There is no any thermal contact between the screens and spacer, and thermal radiation is the only heat transfer mode to be considered; (2) The radiative flux in normal direction can be determined as a solution of a 1-D

problem neglecting small 2-D effects; (3) The isothermal metal screens are totally opaque for thermal radiation; (4) Both transmittance and reflectance of radiation by highly porous fibrous spacers can be determined on the basis of independent scattering hypothesis [5]; (5) The Mie theory for randomly oriented cylinders can be employed to calculate spectral radiative properties of a spacer [6, 7]; (6) One can use transport approximation without taking into account the details of angular dependence of radiation scattered by fibers [7, 8]; (7) The monodisperse approximation for the spacer's fibers can be used in engineering estimates [7]; (8) The polarization effects can be neglected; (9) The thicknesses of oxide films on both surfaces of the aluminum foil are the same as those in the manufacturing, and no subsequent changes of these oxide layers in space is expected.

The spectral dependences of radiative properties of all substances are taken into account. Nevertheless, the resulting mathematical formulation is rather simple because it is based on balance equations for the spectral radiation fluxes [4]. The model suggested in [3, 4] includes an approximate description of the effect of a thin but dense oxide film formed at the surface of an aluminum foil at normal atmospheric conditions. Infrared optical properties of fused silica have been studied during many years, and the optical constants we need for calculations are well known [9]. The temperature dependences of the optical constants of fused silica and are rather weak and can be neglected in the calculations. The model presented in this section is based on the method of papers [3, 4] generalized to the case of a transient problem for numerous layers of MLI. The resulting equations take into account a set of $L = 21$ isothermal opaque screens with one additional layer corresponding to copper plate placed at the outer side of the MLI. The fibrous spacers located between the first and the last layers are not considered as separate elements of MLI, but their effect on radiative heat transfer is taken into account as described in paper [4]. Thermal radiation from the layer of ceramic fabric is treated as an external radiative flux. There is also a radiative heat transfer between the copper plate and the steel wall of the chamber. This wall was cooled by water to avoid its heating during the experiment. As earlier, we assumed that heat transfer between the layers is a result of the far-field thermal radiation and there is no direct thermal contact between the layers. The resulting mathematical formulation of the problem is as follows:

$$c_l \rho_l \delta_l \frac{dT_l}{dt} = \int_{\lambda_{min}}^{\lambda_{max}} \left(q_{a,l,\lambda} - q_{l,2,\lambda} \right) d\lambda , \quad T_1(0) = T_0 , \quad T_a = T_a(t) \tag{1a}$$

$$c_l \rho_l \delta_l \frac{dT_l}{dt} = \int_{\lambda_{min}}^{\lambda_{max}} \left(q_{l-1,l,\lambda} - q_{l,l+1,\lambda} \right) d\lambda , \quad T_l(0) = T_0 , \quad l = 2,...,20 \tag{1b}$$

$$c_L \rho_L \delta_L \frac{dT_L}{dt} = \int_{\lambda_{min}}^{\lambda_{ma[}} q_{L-l,L,\lambda} d\lambda + \varepsilon_{L,c}^{eff} \sigma \left(T_c^4 - T_L^4 \right), \quad T_L(0) = T_0 \tag{1c}$$

$$c_c \rho_c \delta_c \frac{dT_c}{dt} = \varepsilon_{L,c}^{eff} \sigma \left(T_L^4 - T_c^4 \right) + \varepsilon_{c,s}^{eff} \sigma \left(T_s^4 - T_c^4 \right), \quad T_c(0) = T_0 \quad T_s = T_s(t) \tag{1d}$$

$$0 < t < t_{max}$$

where $\varepsilon_{L,c}^{eff} = \dfrac{\varepsilon_L \varepsilon_c}{\varepsilon_L + \varepsilon_c - \varepsilon_L \varepsilon_c}$ $\varepsilon_{c,s}^{eff} = \dfrac{\varepsilon_c \varepsilon_s}{\varepsilon_c + \varepsilon_s - \varepsilon_c \varepsilon_s}$ are the effective emissivities of the neighboring elements. Index "a*"* refers to the heated fabric, indices "*l"* and "L" refer to the screens of MLI (l = 1,..,6 – for aluminum foil, l = 7,..,20 – for aluminized polymer Upilex, and L – for aluminized Kapton), indices "c" and "s" refer to the copper plate and steel wall of experimental facility, respectively.

Spectral radiative flux through a layer of MLI, which consists of two screens and a spacer made of semi-transparent highly porous fibrous material, is expresses as follows:

$$q_{l,l+1,\lambda} = \frac{f_{l,\lambda} - f_{l+1,\lambda}}{\dfrac{1}{\varepsilon_{l,\lambda}} + \dfrac{1}{\varepsilon_{l+1,\lambda}} + \dfrac{2}{1 + T_\lambda - R_\lambda} - 2} \tag{2}$$

where $\varepsilon_{l,\lambda}, \varepsilon_{l+1,\lambda}$ are the spectral emittances of the screens at temperatures T_l and T_{l+1}, $f_\lambda = \pi B_\lambda(T)$ is the blackbody spectral radiative flux, $B_\lambda(T)$ is the Planck function, R_λ and T_λ are the spectral hemispherical reflectance and transmittance of the spacer.

Both transmittance and reflectance of radiation by highly porous fibrous spacers can be determined on the basis of independent scattering hypothesis and Mie theory for infinite homogeneous cylinders. Characteristic of highly porous fibrous spacer can be determined by the relation:

$$U_\lambda = T_\lambda - R_\lambda = 1 - \frac{4}{\pi}(1-p)\widetilde{Q}_{\text{tr}} \tag{3}$$

where p is the surface porosity of a spacer, $\widetilde{Q}_{\text{tr}}$ is transport efficiency factor of extinction. The values of efficiency factors for cylindrical particles at arbitrary illumination of fibers are calculated using known relations of Mie theory. It should be noted that model (1)–(3) includes some uncertainties in material properties, And the spectral emittance of heated fabric $\varepsilon_{a,\lambda}$ is the main of them. Therefore, the latter value should be identified to complete the computational model.

3. INVERSE PROBLEMS ALGORITHM

In a general case, the heat transfer is determined by (1) the parameters of interaction with an environment (experimental facility and heater), (2) the radiative heat transfer in MLI, (3) thermal properties, densities and thickness of the MLI layers, as well as by the system's initial temperature. As was noted, the material properties in the model (1)–(3) have some uncertainties in, and the spectral emittance of heated fabric $\varepsilon_{a,\lambda}$ is the main of them. The results of temperature measurements in the cooper slab (1d) can be considered as additional information, which is not necessary to solve direct heat transfer problem.

$$T_c^{\exp}(\tau_m) = f_m, \ m = \overline{1, M} \tag{4}$$

Strictly speaking, the retrieval of functions $\varepsilon_{a,\lambda}$ is related with to a minimization of the residual functional characterizing the deviation of temperature $T_c(\tau_m)$ calculated for certain estimates of $\varepsilon_{a,\lambda}$ from the measured temperature f_m in the corresponded metric. The following functional can be considered to characterize the least-square deviation of experimental and calculated temperatures:

$$J(\varepsilon_{a,\lambda}) = \int_0^{\tau_{max}} (T_c(\tau) - f(\tau))^2 d\tau \tag{5}$$

To solve the inverse problem, the conjugate gradient method of minimization as the base of iterative regularization method [1] can be used. According to the approach suggested in [2], the unknown coefficient can be approximated by a set of basic functions (in particular, peas-waste functions or B-splines):

$$\varepsilon_{a,\lambda}(\lambda)=\sum_{k=1}^{N} e_k \varphi_k(\lambda) \tag{6}$$

The gradient of the minimized functional is computed using the solution of an adjoint problem:

$$J'_{e_k}=\int_{\lambda_{min}}^{\lambda_{max}} \Psi_l \left(\frac{\varphi_k(\lambda)\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda)}{(\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{a,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{a,\lambda}\varepsilon_{1,\lambda}-2\varepsilon_{a,\lambda}\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda))} - \right.$$
$$\left. \frac{\varepsilon_{a,\lambda}\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda)\varphi_k(\lambda)((1+T_\lambda-R_\lambda)+2\varepsilon_{1,\lambda}-2\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda))}{(\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{a,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{a,\lambda}\varepsilon_{1,\lambda}-2\varepsilon_{a,\lambda}\varepsilon_{11,\lambda}(1+T_\lambda-R_\lambda))^2} \right)(f_{a,\lambda}-f_{1,\lambda})d\lambda \tag{7}$$

where ψ_l is the solution of the following adjoint problem:

$$-c_1\rho_1\delta_1\frac{d\Psi_1}{dt}=\int_{\lambda_{min}}^{\lambda_{max}}\left(\frac{\varepsilon_{a,\lambda}\varepsilon_{+1,\lambda}(1+T_\lambda-R_\lambda)(-)}{(\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{a,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{a,\lambda}\varepsilon_{l+1,\lambda}-2\varepsilon_{a,\lambda}\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda))}-\right.$$
$$\left.\frac{\varepsilon_{1,\lambda}\varepsilon_{12,\lambda}(1+T_\lambda-R_\lambda)(\Psi_1(\tau)df_{1,\lambda}/dT-\Psi_2(\tau)df_{2,\lambda}/dT)}{(\varepsilon_{2,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{1,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{1,\lambda}\varepsilon_{2,\lambda}-2\varepsilon_{1,\lambda}\varepsilon_{2,\lambda}(1+T_\lambda-R_\lambda))}\right)d\lambda,$$
$$\Psi_1(\tau_{\max})=0, \tag{8a}$$

$$-c_l\rho_l\delta_l\frac{d\Psi_l}{dt}=\int_{\lambda_{min}}^{\lambda_{max}}\left(\frac{\varepsilon_{l-1,\lambda}\varepsilon_{l,\lambda}(1+T_\lambda-R_\lambda)(\Psi_{l-1}(\tau)df_{l-1,\lambda}/dT-\Psi_l(\tau)df_{l,\lambda}/dT)}{(\varepsilon_{l,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{l-1,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{l,\lambda}\varepsilon_{l-1,\lambda}-2\varepsilon_{l,\lambda}\varepsilon_{l-1,\lambda}(1+T_\lambda-R_\lambda))}-\right.$$
$$\left.\frac{\varepsilon_{l,\lambda}\varepsilon_{l+1,\lambda}(1+T_\lambda-R_\lambda)(\Psi_l(\tau)df_{l,\lambda}/dT-\Psi_{l+1}(\tau)df_{l+1,\lambda}/dT)}{(\varepsilon_{l+1,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{l,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{l,\lambda}\varepsilon_{l+1,\lambda}-2\varepsilon_{l,\lambda}\varepsilon_{l+1,\lambda}(1+T_\lambda-R_\lambda))}\right)d\lambda,$$
$$\Psi_l(\tau_{max})=0,\ l=2,...,20 \tag{8b}$$

$$-c_L\rho_L\delta_L\frac{d\Psi_L}{dt}=\int_{\lambda_{max}}^{\lambda_{min}}\left(\frac{\varepsilon_{L-1,\lambda}\varepsilon_{L,\lambda}(1+T_\lambda-R_\lambda)(\Psi_{L-1}(\tau)df_{L-1,\lambda}/dT-\Psi_L(\tau)df_{L,\lambda}/dT)}{(\varepsilon_{L,\lambda}(1+T_\lambda-R_\lambda)+\varepsilon_{L-1,\lambda}(1+T_\lambda-R_\lambda)+2\varepsilon_{L,\lambda}\varepsilon_{L-1,\lambda}-2\varepsilon_{L,\lambda}\varepsilon_{L-1,\lambda}(1+T_\lambda-R_\lambda))}\right)d\lambda$$
$$+4\varepsilon_{L,c}^{eff}\sigma(T_c^3-T_L^3),$$
$$\Psi_L(\tau_{max})=0 \tag{8c}$$

$$-c_c\rho_c d_c\frac{d\psi_c}{d\tau}=4\varepsilon_{L,c}^{eff}\sigma(T_L^3-T_c^3)+4\varepsilon_{c,s}^{eff}\sigma(T_s^3-T_c^3)+2(T_c(\tau)-f(\tau)),\qquad \Psi_c(\tau_{\max})=0 \tag{8d}$$

4. COMPUTATIONAL RESULTS

An application of inverse problem methods for heat transfer problems, is related with a specific testing of the developed algorithms. The most universal approach hereby is a computational experiment, which is made in the following way: first a direct heat transfer problem in the specimen should be solved, on the assumption that all characteristics of material are known. Using the obtained values of temperature in the supposed points of thermosensors installation, then "experimental" data necessary for solving an inverse problem are formed, and after that an inverse problem on determining of $\varepsilon_{a,\lambda}$ is solved. Random errors in the "experimental" data are formed as follows

$$f(\tau) = \tilde{f}(\tau)\left(1 + \delta_f \omega(\tau)\right) \tag{9}$$

where $\tilde{f}(\tau)$ is the exact value, ω is a random value of normal distribution with a mean value equal 0.0 and a dispersion equal 1.0, $\delta_f = 0.05$ is a relative error. The computational results at conditions of experimental heating are presented in Fig. 5. The emissivities of oxidized copper plate surface and steel surface were taken equal to 0.6 and 0.45, respectively. As to parameters of fibrous spacers, the fiber radius was taken equal to $a = 3$ μm and the surface porosity – $p = 0.8$. The wavelength range from $\lambda_{min} = 2$ μm to $\lambda_{\max} = 20$ μm was used while integrating over the spectrum since this range makes the main contribution to the integral radiative flux.

A comparison of the calculated and "measured" temperatures of copper slab as a result of inverse problem solving for different approximations (6) is presented in Fig. 6 and in Table 1. The estimated values of $\varepsilon_{a,\lambda}$ for different approximation are presented in Fig.7. As one can expect, the minimal set of the approximation parameters provides the best accuracy and stability of inverse problems solving. Though in the case of approximation by peace-waste functions with N=2 we used the a-priori information about the spectral boundary of two-band model of spectral properties of quartz fibers [10] (with the "boundary" wavelength about 6.5 μm), but the estimate of this value can be obtained from other approximations (Fig. 7, curves 2-4).

Table 1: The deviation of the calculated temperatures and measured temperatures

Approximation	Least-squares temperature deviation (K)	Temperature deviation (%)	Maximum deviation (K)	Maximum deviation (%)
1	1.24	4.1	2.5	7.1
2	1.16	4.3	2.3	7.5
3	1.21	4.7	2.4	6.2
4	0.37	2.4	1.2	4.3

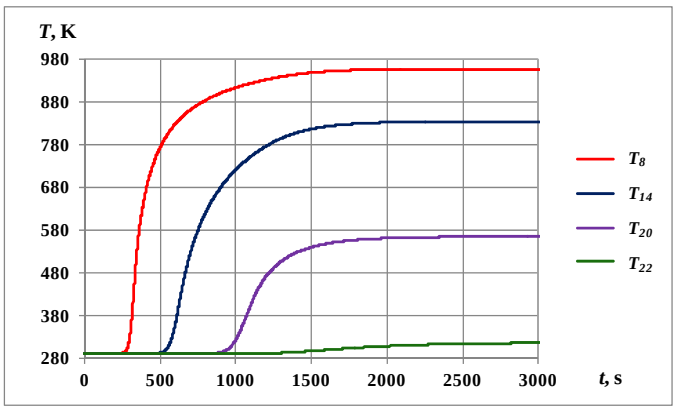


Figure 5: Temperature of MLI layers calculated for the stages of initial heating: the numbers of MLI layers are specified in panels of the figure.

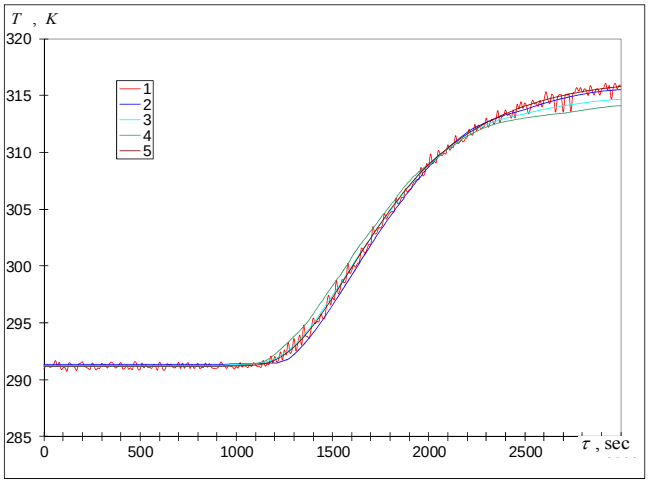


Figure 6: Comparison of theoretical predictions and experimental data for temperature of copper plate: 1 – "experimental", 2–4 – aproximations (2) by cubic B-splines at N=5, (3) by cubic B-splines at N=3, (4) – by peace-waste functions at N=36, (5) by peace-waste functions at N=2.

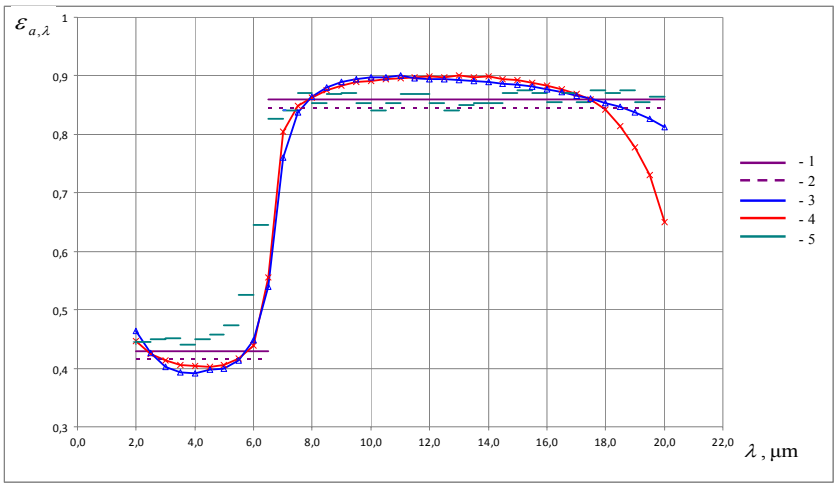


Figure 7: Reconstruction of $\varepsilon_{a,\lambda}$: 1 – "exact" value, 2–5 – estimations using (2) cubic B-splines at N=5, (3) cubic B-splines at N=3, (4) peace-waste functions at N=36, (5) peace-waste functions at N=2.

5. CONCLUSIONS

An identification procedure for mathematical model of the multilayer thermal insulation (MLI) of space vehicles with the use of the MLI sample of the BP-Colombo satellite (ESA) showed that a theoretical model developed recently by the authors can be used to estimate thermal properties of the insulation at conditions of space vacuum. The above comparison enables us to use the developed method to analyze wide-range spectral properties of the MLI. A further study of the external pressure effect on radiative transfer in single layers of MLI is expected to be interesting both theoretically (to

observe possible near-field effects) and practically (to predict variation of the heat shielding properties of the MLI at conditions of a planetary atmosphere).

ACKNOWLEDGMENTS

The authors are grateful to the Russian Foundation for Basic Research for the financial support of this study (grant 13-08-00022-a). This work was also supported by the Russian president grant for Scientific Team NSh- 6343.2014.8.

REFERENCES

[1] Alifanov, O.M., *Inverse Heat Transfer Problems*, Springer-Verlag, Berlin, 1994.

[2] Alifanov, O.M., Nenarokomov, A.V., and Gonzalez, V.M., *Study of multilayer thermal insulation by inverse problems method*, Acta Astronautica, 65, (2009).

[3] Gritsevich, I.V., Dombrovsky, L.A., and Nenarokomov, A.V., *Heat transfer by radiation in a vacuum thermal insulation of space vehicles*, Therm. Proc. Eng., 5(1), (2013), pp. 12-21. (in Russian)

[4] Gritsevich, I.V., Dombrovsky, L.A., and Nenarokomov, A.V., *Radiative transfer in vacuum thermal insulation of space vehicles*, Comput. Therm. Sci., 2014, in press.

[5] Tien, C.L. and Drolen, B.L., *Thermal radiation in particulate media with dependent and independent scattering*, in "Annual Review of Numerical Fluid Mechanics and Heat Transfer", New York: Hemisphere, (1987), v. 1, pp. 1-32.

[6] Bohren, C.F. and Huffman, D.R., *Absorption and Scattering of Light by Small Particles*, New York: Wiley, (1983).

[7] Dombrovsky, L.A. and Baillis, D., *Thermal Radiation in Disperse Systems: An Engineering Approach*, New York: Begell House, (2010).

[8] Dombrovsky, L.A., *The use of transport approximation and diffusion-based models in radiative transfer calculations*, Comput. Therm. Sci., (2012), 4(4), pp. 297-315.

[9] Kitamura, R., Pilon, L., and Jonasz, M., *Optical constants of silica glass from extreme ultraviolet to far infrared at near room temperatures*, Appl. Opt., (2007), 46(33), pp. 8118-8133.

[10] Dombrovsky, L.A., *Quartz-fiber thermal insulation: Infrared radiative properties and calculation of radiative-conductive heat transfer*, ASME J. Heat Transfer, (1996) 118(2), pp. 408-414.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Automated Hybrid Singularity Superposition and Anchored Grid Pattern BEM Algorithm for the Solution of the Inverse Geometric Problem

Marcus W. Ni, Alain J. Kassab*, Eduardo Divo

Department Mechanical and Aerospace Engineering
University of Central Florida, Orlando, Florida, USA
e-mail: alain.kassab@ucf.edu

Key words: Inverse geometric problems, Simplex optimization, Anchored grid pattern, Boundary element methods

Abstract
A method for solving the inverse geometrical problem is presented by reconstructing the unknown subsurface cavity geometry using boundary element methods, a genetic algorithm, and Nelder-Mead non-linear simplex optimization. The heat conduction problem is solved utilizing the boundary element method, which calculates the difference between the measured temperature at the exposed surface and the computed temperature under the current update of the unknown subsurface flaws and cavities. In a first step, clusters of singularities are utilized to solve the inverse problem and to identify the location of the centroid(s) of the subsurface cavity(ies)/flaw(s). In a second step, the reconstruction of the estimated cavity(ies)/flaw(s) geometry(ies) is accomplished by utilizing an anchored grid pattern upon which cubic spline knots are restricted to move in the search for unknown geometry. Solution of the inverse problem is achieved using a genetic algorithm accelerated with the Nelder-Mead non-linear simplex. To optimize the cubic spline interpolated geometry, the flux (Neumann) boundary conditions are minimized using a least squares functional. The automated algorithm successfully reconstructs single and multiple subsurface cavities within two dimensional mediums. The solver is also shown to accurately predict cavity geometries with random noise in the boundary condition measurements. Subsurface cavities can be difficult to detect based on their location. By applying different boundary conditions to the same geometry, more information is supplied at the boundary, and the subsurface cavity is easily detected despite its low heat signature effect at the boundaries. Extensions to three-dimensional applications are mentioned.

1 Introduction

Forward problems are defined as well posed, and require five specifications which are as follows: governing equation for field variable, physical properties, boundary conditions, initial condition(s) and system geometry [2]. The inverse geometrical problem is one of many inverse problems that arise in engineering today. This problem is broadly classified into the inverse problem category. The inverse problem differs from the forward problem in that one of the forward problem specifications is unknown, and there is at least one over specified condition. In most cases, including the inverse geometric problem, this over specified condition arises at the boundary, which lends itself to the boundary element method (BEM).

Subsurface cavity detection and geometry reconstruction methods using BEM have been well documented within the past decade [1], [3]. These methods are non-intrusive, and have been successfully proven to detect sub surface cavities, as well as predict its shape. In these problems the governing equation, physical properties, boundary conditions, initial conditions and external geometry are known, leaving the internal cavity wall geometry unknown. The boundary conditions at the cavity walls are also considered to be known as adiabatic, or very close to adiabatic. This can be attributed to the extremely low thermal conductivity of the material or void within the cavity. Radiation would have an effect on these subsurface cavity boundary conditions at high temperatures; however, radiation has been omitted because the temperature ranges should be kept relatively low.

An efficient way to detect cavity shape and location was proposed by E. Divo in 2004 [2]. In his article, the inverse geometric problem is stated much like the one in this article. The use of point source clusters is employed to search for subsurface cavities in 2-D and 3-D geometries by minimizing the first order boundary condition at the exposed surface. The efficiency of his model comes from lack of grid reconstruction. Rather than actually reconstructing the cavity walls during the optimization, a search for the adiabatic condition is run with the optimized cluster location and shape in place to act as the void within the medium. The shape is simplified to be elliptical to lower optimization parameters, and the technique has proven to be successful. Application of these techniques requires thermal imaging of the exposed surface with the use of infrared scanners as shown in Figure 1. At this point the boundary conditions at the exposed surface are over specified, and the internal cavity location and shape is unknown.

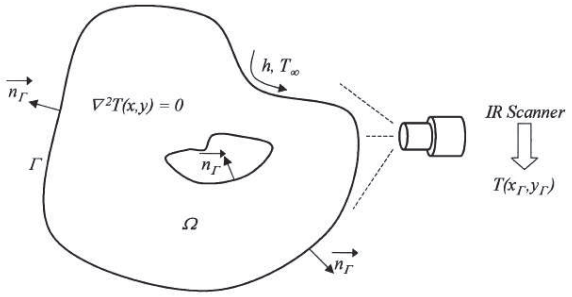


Figure 1: Problem setup using IR scanner to measure thermal footprint at the exposed boundary [2].

It is also possible to utilize elastics rather than heat conduction to search for these cavities. Kassab et al [3] presents a solution using elastostatics and the boundary element method. Rather than minimizing the Neumann or Cauchy boundary conditions of the heat transfer problem, the strains and deformations at the exposed surface boundaries are minimized. This paper also proposes a variety of differently shaped anchored grid patterns (AGP) that adapt to the shape of the internal cavity [3].
Another solution to this inverse problem is to use the method of fundamental solutions (MFS) rather than the boundary element method (BEM), proposed by A. Karageorghis [5]. In one of these methods,

Karageorghis utilizes a moving pseudo-boundary technique to detect void location and boundary location. The heat conduction equation and Neumann boundary condition minimization are also utilized similar to this article's approach. High resolution of these cavity boundaries are obtained by using radial polar parameterization, which simply searches for points radially around the cavity center [5], [8]. This technique requires a high number of parameters to be optimized which can lead to longer calculation times and instability. Karageorghis also proposed a solution to cavity detection within plane linear elastic bodies. The same MFS and radial polar parameterization technique is used; however, the boundary conditions to be minimized are described by the elastic properties of the medium. Boundary deformity, and strain levels are minimized to optimize the cavity location and shape [7].

In this article, a higher resolution algorithm for cavity shape is proposed by using an anchored grid pattern to map a cubic spline that wraps itself around the cavity. The boundary element method is used to solve the forward heat transfer problem at each step of the overall inverse problem that in turn employs simplex optimization techniques. To run these experiments numerically, the cavity is imposed and the boundary conditions are calculated based on this implication. In essence, the boundary element calculations are being compared to the analytical Laplace equation solution for the test problems (rather than actually using a thermal measuring device).

The automated algorithm for reconstruction of cavity geometry starts by searching for its general location using clusters of sources/sinks that satisfy the heat flux (Neumann) boundary conditions. These cluster(s) must locate themselves within the cavity or outside of the medium to satisfy the Laplace equation. The boundary element method is used to solve the forward problem, while the genetic algorithm optimizes the location of these source/sink clusters. Once the cavity location is determined, the same steps are used to predict the geometry's shape using cubic spline interpolation. By using an anchored grid pattern, the cavity is shaped using eight splines, to which the surface is attached [1]. This pattern is placed at the center of the detected cavity, adiabatic boundary conditions are applied to this cubic spline surface, and the spline lengths are optimized to satisfy the heat flux (Neumann) boundary conditions. This application has also been extended to include multiple cavities, and multiple boundary condition sets (MBCS). MBCS uses two or more boundary conditions sets that are applied to the same geometry set. This method increases the number of boundary conditions without increasing the number of boundary elements along the surface. In a lab setting, measuring points are limited, so MBCS becomes useful in increasing the resolution without lab expense. The concept of multiple boundary condition sets (MBCS) optimization was used to effectively enhance sensitivity and detect complicated geometrical shapes and locations. This study has been successful in adding shape resolution to the Efficient Singularity Super Position Technique [2], and should be extended to 3-D geometries using these same techniques

2 Solution Procedure

To begin to understand the solution to the geometric problem, a layout for the automated algorithm is presented. First, the forward problem is solved using boundary elements with the source/sink clusters to imitate subsurface cavity behavior. Second, the clusters are moved to the cavity location with the help of the genetic algorithm method. Once the cavity is located, an anchored grid pattern replaces the clusters, and a new surface is created to guess at the cavity geometry. These anchored grids are then optimized with the Nelder-Mead simplex method until the Neumann boundary conditions, that have been over specified, agree with the boundary element solution.

1) Set up problem parameters and initial guesses for the cavity location
2) Solve forward problem using BEM
3) Optimize the cavity location using the genetic algorithm

4) Use optimized location for central spline knot of cubic interpolation and set up initial guesses for the cavity geometry
5) Solve forward problem using BEM
6) Optimize the cavity geometry using the Nelder-Mead simplex.

2.1 The forward problem solver: Boundary Element Method

As stated previously, the general equation to be solved is the homogeneous Laplace equation, equation 1, where T is the temperature at the point (x, y) with in the medium Ω.

$$\nabla^2 T(x,y) = 0 \in \Omega \tag{1}$$

The boundary element method solves for T(x, y) numerically by discretizing the boundary into several elements. Green's second identity is applied and a free space test function, T*, is introduced to the Laplace equation. It is then integrated over the surface elements, Γ, as shown in equation 2 [8].

$$C(i)T(i) + \sum_{j=1}^{N} \oint_{\Gamma_j} Tq^* d\Gamma = \sum_{j=1}^{N} \oint_{\Gamma_j} T^* q d\Gamma \tag{2}$$

The temperature for the given point "i" has been solved by integrating over the surface elements labeled "j", where N is the number of surface elements. "C" is termed the jump coefficient, which is determined as 0.5 for completely smooth elements. Since T and q are constant over the element, they are taken out of the integral, and the test functions are labeled accordingly as shown in equations 3 and 4.

$$H_{ij} = \oint_{\Gamma_j} q^* d\Gamma \tag{3}$$

$$G_{ij} = \oint_{\Gamma_j} q^* d\Gamma \tag{4}$$

Once the elements have been discretized, equation 2 can be written in the standard form as equation 5 [8].

$$\sum_{j=1}^{N} G_{ij} q_j = \sum_{j=1}^{N} H_{ij} T_j \tag{5}$$

This formulation does not take into consideration point sources; however, the point sources can be simply added to the right hand side of this equation. Equation 6 shows the added point sources to the general equation, where {B} is the source vector shown in equation 7 [12].

$$[G]\{q\} = [H]\{T\} + \{B\} \tag{6}$$

$$B_i = \sum_{j=1}^{Ns} \frac{Q_j}{4\pi} \left[\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 \right] \tag{7}$$

The standard procedure to solve equation 6 is to move all the unknown boundary conditions to the same vector and use matrix multiplication to simplify. This simplification leads to equation 8, and is easily solved using pivoting methods for multiple sets of equations.

$$[A]\{x\} = \{b\} \tag{8}$$

The solution for "{x}" represents the unknown boundary conditions that are being solved for. Using this solution and equation 2 the entire solution is formed.

2.2 Geometry reconstruction: Cubic spline interpolation

To reconstruct the subsurface cavity geometry, a cubic spline interpolator generates a surface around a central point provided by the efficient hole finding optimization [2]. The first step is to define the anchored pattern as to which the cubic splines will attach themselves. Eight splines are defined to be positive, be equally separated and extend from the central knot as shown in Figure 2. The ends of these splines define the cubic spline endpoints, and continuity conditions define the shape. Since the angles between each spline are fixed and the shape is periodic, polar coordinates are used to define the locations along the cubic splines. According to Pollard and Kassab [1], the location along the cubic spline surface can be defined by equation 9,

$$r(\theta) = \mathrm{M}_{i-1}\frac{(\theta_i-\theta)^3}{6\Delta\theta_i} + M_i\frac{(\theta-\theta_{i-1})^3}{6\Delta\theta_i} + \left(r_{i-1} - \frac{M_{i-1}\Delta\theta_i^2}{6}\right)\frac{\theta_i-\theta}{\Delta\theta_i} + \left(r_i - \frac{M_i\Delta\theta_i^2}{6}\right)\frac{\theta-\theta_{i-1}}{\Delta\theta_i} \quad (9)$$

where "r" is the radial position along the polar coordinate system which depends on "θ", the angle along the polar coordinate system. The "i" corresponds to the intervals between each spline, adding up to eight in this case. The spacing between each spline is defined as $\Delta\theta i = \theta i - \theta(i-1)$, which is considered to be fixed in these problems. The continuity conditions state that the first and second derivatives at the spline endpoints are continuous from interval to interval. The periodic condition states that the first location is equal to the last location. These conditions lead to a set of tri-diagonal equations shown in equation 10, 11 and 12. This simple set of equations is solved during the optimization using the Thomas algorithm.

$$2M_1 + \lambda M_2 = d_1 \quad (10)$$

$$\mu_i M_{i-1} + 2M_i + \lambda_i M_{i+1} = d_i \in i = 2,3,...,7 \quad (11)$$

$$\mu_8 2M_8 = d_8 \quad (12)$$

The coefficients are defined by Pollard and Kassab's cubic spline anchored grid pattern [1].

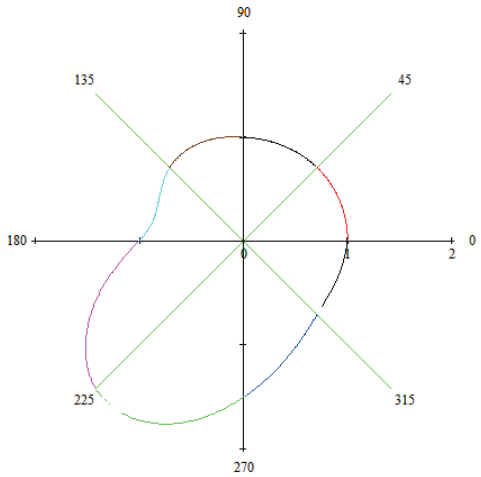


Figure 2: Cubic spline interpolation representation.

2.3 Geometry optimization technique: Nelder-Mead simplex

The inverse problem requires an iterative method that aids in the search for the subsurface cavity geometry by minimizing a least squares functional (LSF) comparing the second order (Neumann) boundary conditions. Since the forward problem is computationally expensive for complete grid reconstruction, the Nelder-Mead simplex method is ideal due to minimal forward problem calculations per iteration. The first step in any optimization technique is to define how the unknown parameters are altered to minimize the function/problem being analyzed. The simplex method is defined by the number of unknown parameters that are of interest. The "simplex" is a geometrical figure with N + 1 points and N dimensions, where 'N' is the number of parameters. For example, in a two dimensional problem the simplex is a triangle, and in a three dimensional problem the simplex is a tetrahedron [10]. As for the solution marching logic, the simplex can be manipulated by expansion, contraction and reflection. These three tools are utilized to move the solution with the highest residual error around the solutions of low residual error to find a new minimum. In essence, the simplex is moved in a downhill manner until a tolerance between solutions is reached.

The single cavity problem with known internal boundary conditions requires nine initial guesses of eight parameters. These eight parameters define the shape of the internal cavity. The cubic spline creates high resolution without the drawback of unknown parameter increases. Calculation time is thereby reduced, and the simplex optimization becomes more stable. To minimize these unknown parameters a least squares functional is defined by comparing the Neumann boundary conditions at the surface. Equation 13 defines this functional, where {r} is the array containing all of the spline lengths of the AGP, qc and qm are the BEM calculated and IR measured flux of each boundary element respectively. Nsc is the scale factor and Ne is the number of boundary elements at the surface.

$$LSF(\{r\}) = \sum_{i=1}^{Ne} \frac{1}{Nsc} \left(q_c - q_m\right)_i^2 + \alpha(r) \tag{13}$$

This 'α' function requires the simplex method to confine its search within the limits that have been setup using educated guesses. These limits may include, but not limited to, geometrical boundaries and anchored grid pattern sizing.

Logic behind choosing the initial N + 1 guesses is also addressed during the optimization technique. The user provides one educated initial guess for the geometric solution to the problem. The optimizer then chooses the remaining initial guesses by scaling the geometry, and randomly adding mutations to the scale. The cavity geometry must fit within the medium, so a method for checking geometry stability is proposed. Points along the cavity wall must be contained within the medium, so it can be said that the angles made between each boundary element along the exposed surface will add to 2π. If the point lies outside the medium, this summation becomes zero. Figure 3 shows how these angles add up to 2π when the point is located within the medium and zero when the point is located beyond the medium.

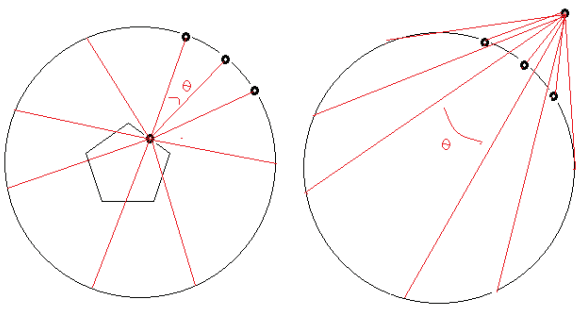


Figure 3: Left - Inside. Right - Outside.

To calculate these angles appropriately, the connections between the points of the exposed surface and cavity walls are treated as vectors. Simple vector algebra states that the angle between any two rays can be expressed by equation 14.

$$\cos(\theta_i) = \frac{r_{i+1} r_i}{|r_{i+1}||r_i|} \tag{14}$$

The convergence criteria of the simplex method consists of parameter comparison, and LSF residuals. As the simplex marches downhill, the N + 1 parameter sets begin to converge to a single set. Once the changes between each set reaches a tolerance set by the user, the optimization stops, and the solution is analyzed. If the solution did not yield a low residual between the Neumann conditions, the weighting parameters for the contraction, expansion, and reflection tools are tweaked. Once the residual condition is satisfied, the solution to the inverse geometry problem is found.

Multiple boundary condition sets (MBCS) can also be used to increase the resolution of the exposed surface without having to add more boundary element locations to the geometry. In a simple example, two boundary condition sets are used as shown in Figure 4. Two identical geometries are imposed with separate boundary conditions to increase the number of boundary conditions being minimized. This is a much more efficient method as compared to just increasing the number of boundary elements. Essentially two sets of boundary conditions are used to solve one inverse geometric problem. The new LSF looks quite similar to equation 13 with the exception of a second set of boundary conditions. Equation 15 defines the MBCS minimization functional.

$$LSF(\{r\}) = \left(\sum_{i=1}^{Ne} \frac{1}{Nsc}\left(q_c - q_m\right)_i^2\right)_{SET1} + \left(\sum_{i=1}^{Ne} \frac{1}{Nsc}\left(q_c - q_m\right)_i^2\right)_{SET2} \tag{15}$$

This technique is not restricted to two sets, and can be extended further with the implication of speed loss due to increased calculation times. Each set requires an additional forward problem solution per iteration.

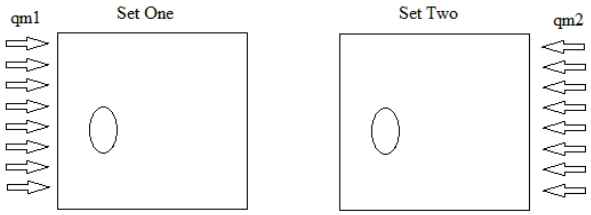


Figure 4: Multiple boundary condition sets experiment.

3 Results

3.1 Single cavity reconstruction

The single cavity experiment involved realistic boundary conditions of both temperature and heat flux values. An irregular shaped cavity was imposed with a very low constant heat flux at the surface. The outer boundary is insulated on the top and bottom walls, whereas the temperature is imposed at the left and right surfaces of the outer boundary. The cavity location is found using point source techniques. Once the location is found, the surface generator takes over and optimizes the cavity

geometry. The optimizer is initialized with two initial guesses as shown in Figure 5. The AGP splines are then used to generate the cavity surface and section into boundary elements.

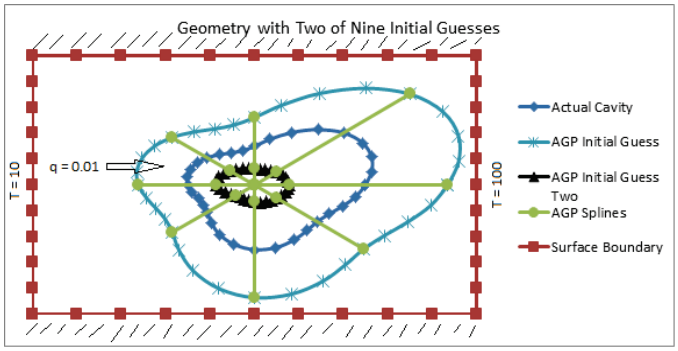


Figure 5: Geometry and initial guess locations

After several iterations, the unknown cavity surface was found. The final solution is shown in Figure 6 (Left), which is observed to have high accuracy. In practical applications where experiments provide the over-specified conditions required to solve the inverse problem, the exact boundary conditions cannot be measured perfectly. To simulate these conditions noise was added to the exact boundary conditions on the exterior surfaces (the boundary conditions that would be measured with instrumentation). Figure 6 (Right) depicts the solution with 1% noise added to both the flux and temperature boundary conditions. The noisy data causes some deviation from the exact geometry, but the solution still successfully captures the correct shape and location.

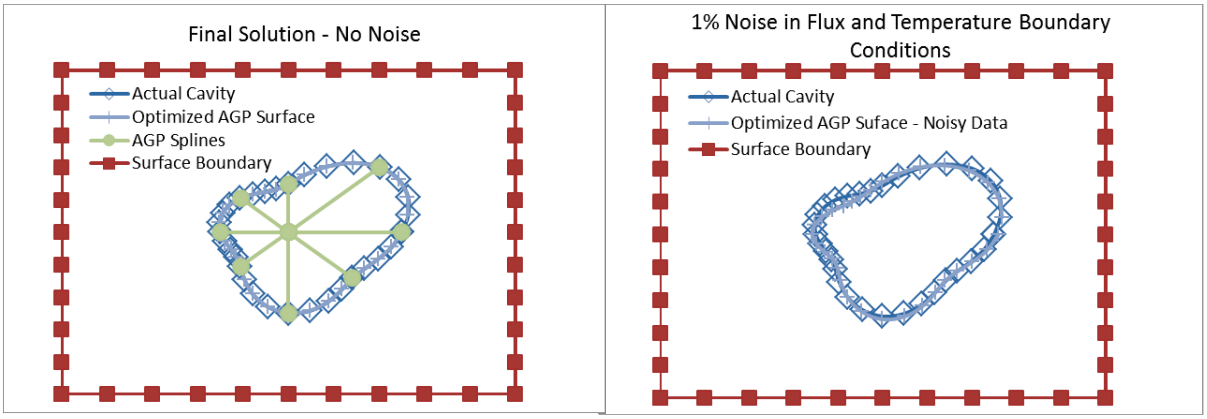


Figure 6: Final solution: No noise (Left), Noise (Right)

3.2 Multiple cavity reconstruction experiment

In this two cavity experiment the 8 parameter problem has now been increased to 16 geometrical parameters. The problem being solved is shown in Figure 7 (Left). These particular boundary conditions are chosen to create a temperature field that will show large differentiations on the boundaries due to interior cavities. The internal cavities are found using the simplex method and are shown in Figure 7 (Right). The calculated cavities are a close match to the test geometry that was set up as an experiment, but there are some deviations due to the approximations incurred from constant element BEM. This error can be reduced by creating higher resolution geometries, or applying MBCS.

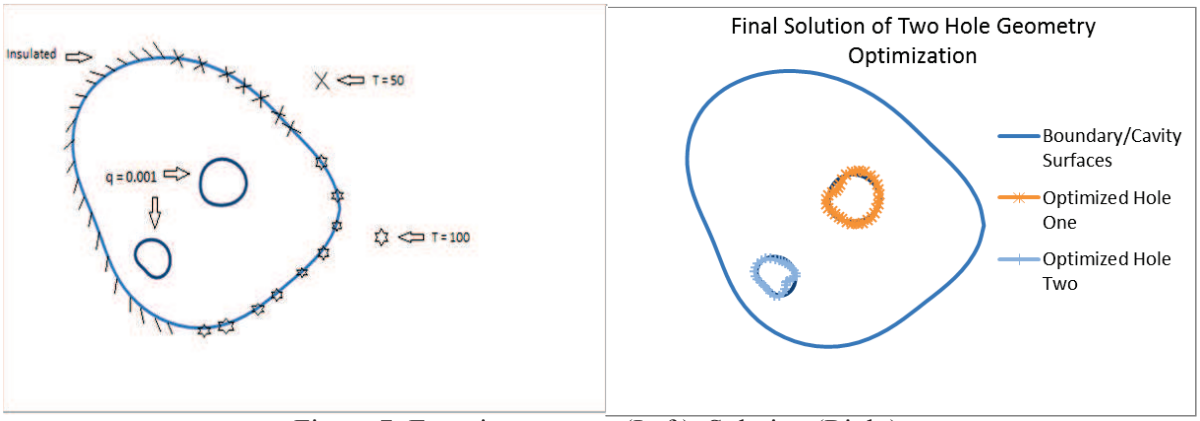


Figure 7: Experiment setup (Left), Solution (Right)

3.3 MBCS using environmentally realistic boundary conditions

Multiple boundary condition setups help to alleviate the error incurred by using star shaped anchored grid patterns. As discussed in the 'solution procedure' section, multiple boundary condition setups can be used to increase the resolution of the exposed boundary without adding more boundary elements. In this case a more realistic boundary condition setup is used. Up to this point the boundary conditions that have been applied to the test geometries have been chosen to create distinct heat signatures which are easily detected at the surface. If the cavity location or shape does not cause significant heat signature differences it may be difficult to locate. Hence, MBCS is proposed to solve this problem. These techniques can be extended to actual experiments, which would involve convective boundary conditions rather than imposed Cauchy conditions. The experimental setup involves the same 10x10 unit square; however, the cavity has been moved to the corner to limit the sensitivity of the cavity's effect along some of the square's boundaries. Figure 8 (Left) displays the four different boundary condition setups that are used in this experiment. In all four setups a heat flux is applied to one side, while the other three sides are subjected to convection conditions at room temperature.

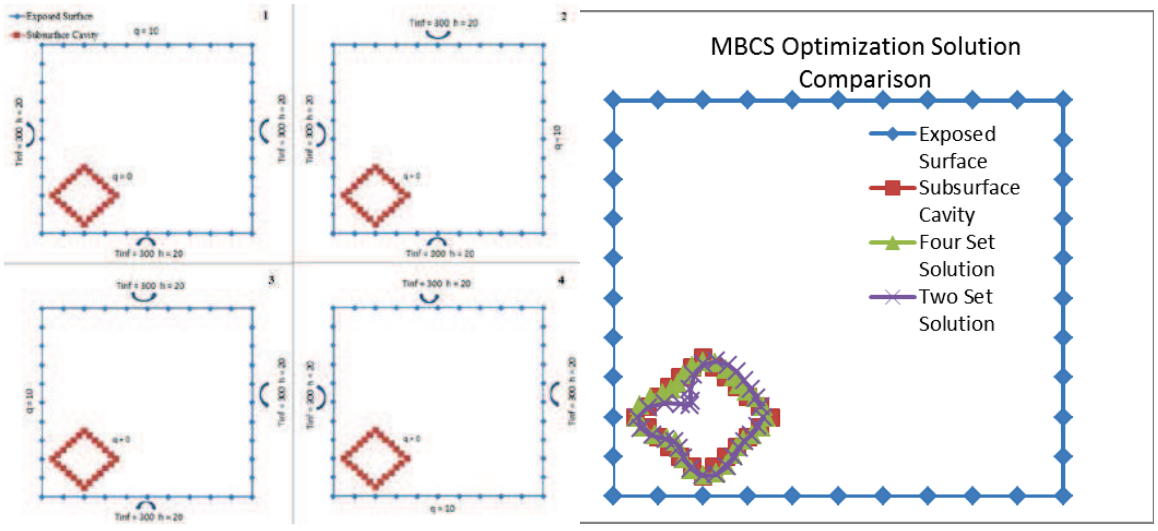


Figure 8: Setup (Left), Two set solution (Top Right), Four set solution (Bottom Right)

Simply put, a heater is placed on one side of the square while the rest of the sides are free to interact convectively with its surroundings. The heater is then moved to the next side and the rest of the sides interact convectively. This procedure increases the sensitivity of the cavities effect on the boundaries, even with difficult cavity locations or poor boundary conditions. First, setup 1 and 2 are used to compute the cavity geometry using MBCS. Figure 8 (Top Right) provides the results from this two boundary set problem. As the results indicate, the cavity geometry is successfully predicted at all edges except for top left side of the diamond. This error is a result of poor boundary condition setup with respect to the location and shape of the cavity. To prove the effectiveness of MBCS beyond two boundary condition sets, setups 1 to 4 are used to calculate the cavity geometry. Figure 8 (Bottom Left) depicts the four boundary sets solution and the result speaks for itself. The cavity geometry is predicted with high accuracy due to increased resolution provided by the four sets of boundary conditions. Separately, each boundary condition set is inefficient in its use for optimization; however, when all four sets are combined the needed information is exceeded.

3.4 Notch Cavity Reconstruction MBCS Trial

Cavity geometries with non-circular shapes are also of interest when detecting unknown cavities. The star shaped anchored grid pattern (AGP) comes into question when the shape of the cavity has sharp edges, turns or corners. Figure 9 outlines the geometrical setup used to test the star shaped AGP against notch shaped cavities. The block is a 10x10 unit structure with 40 boundary elements along the exposed surface.

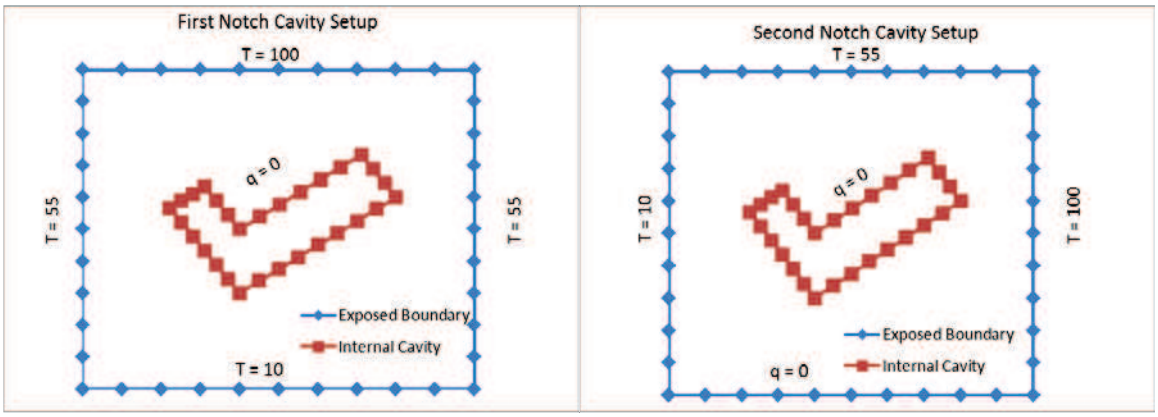


Figure 9: First boundary condition setup (Left), Second boundary condition setup (Right)

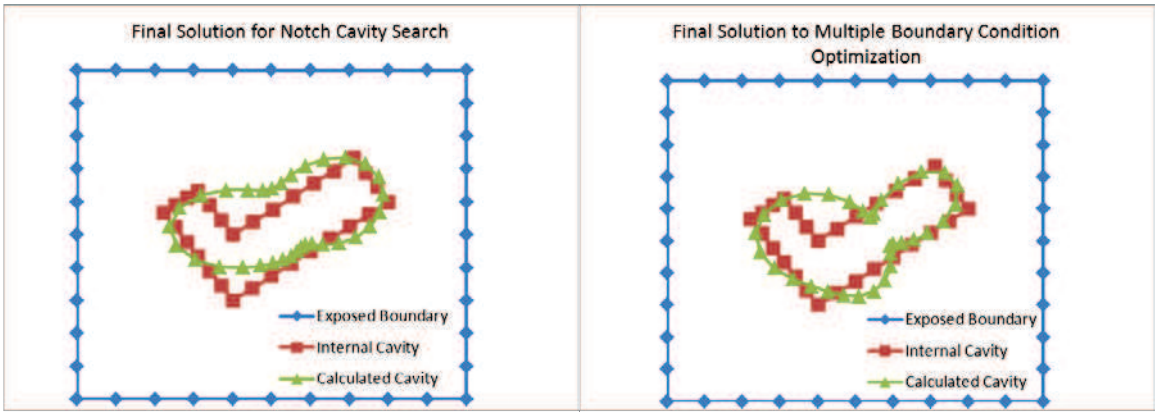


Figure 10: Solution for first setup (Left), MBCS solution for both setups (Right)

The shape optimization in this first case did not accurately predict the shape of the notch. This error is due to the AGP shape itself. As the splines are trying to find their way to the correct solution, they end up crossing the actual cavity walls twice due to the sharp corner of the notch. The AGP was successful in following the general shape of the notch, especially on the right side of the notch where there is no sharp corner. Figure 10 (Left) depicts the calculated cavity over the actual cavity. Error incurred by using star shaped anchored grid patterns is reduced by using multiple boundary condition setups. Figure 10(Right) suggests the MBCS final solution to the cavity geometry is more accurate. The star AGP correctly identifies the right portion of the notch, but has severe issues with the corner section of the notch. The interpolator compensates by cutting in and out of the correct geometry, but does not predict the sharp corner of the notch.

4 Conclusion

An automated algorithm for detecting cavities and reconstructing their shape has been successfully tested and used in hypothetical experiment setups. Constant elements were also proven to be accurate enough to detect the shape of the cavity through optimization by increasing the resolution of the boundary elements. The system performed to expectations by successfully increasing the resolution of subsurface cavity geometry prediction. Multiple boundary condition set optimization also succeeded in increasing the resolution of the problem without adding boundary elements to the surface. Complicated cavity geometry and location difficulties have proven to be avoidable with MBCS. These techniques can be extended to 3-D geometries, and shows promising calculation speeds due to the low parameter counts being optimized.

References

[1] A. J. Kassab and J. Pollard, Cubic Spline Anchored Grid Pattern Algorithm for High Resolution Detection of Subsurface Cavities by the IR-CAT Method, *Numerical Heat Transfer B*, vol. 26, no. 1, pp. 63-78, 1994.

[2] E. Divo, A. J. Kassab and F. Rodriguez, An Efficient Singular Superposition Technique for Cavity Detection and Shape Optimization, *Numerical Heat Transfer B*, vol. 46, pp. 1-30, 2004.

[3] A. J. Kassab, F. A. Moslehy and A. B. Daryapurkar, Nondestructive Detection of Cavities by an Inverse Elastostatics Boundary Element Method, *Engineering Analysis with Boundary Elements*, vol. 14, pp. 45-55, 1994.

[4] D. Ojeda, E. Divo and A. J. Kassab, Cavity Detection in Biomechanics by an Inverse Evolutionary Point Load BEM Technique, *Inverse Problems in Science and Engineering*, vol. 16, no. 8, pp. 981-993, 2008.

[5] A. Karageorghis, D. Lesnic and L Marin, A Moving Pseudo-boundary Method of Fundamental Solutions for Void Detection, *Wiley Online Library, 10.1002/num.21739, 2012*

[6] A. Karageorghis, D. Lesnic and L Marin, The Method of Fundamental Solutions for the Detection of Rigid Inclusions and Cavities in Plane Linear Elastic Bodies, *Computers and Structures,* vol. 106-107, pp. 176-188, 2012.

[7] A. Karageorghis, D. Lesnic and L Marin, A Survey of the MFS to Inverse Problems, *Inverse Problems in Science and Engineering,* vol. 19, no. 3, pp. 309-336, 2011

[8] C. A. Brebbia and J. Dominguez, *Boundary Elements an Introductory Course*, CMP, Great Britain, 1992.

[9] E. Divo and A. Kassab, *Boundary Element Method for Heat Conduction: with applications to non-homogeneous media, Anisotropic Heterogeneous Heat Conduction,* Wessex Institute of Technology Press, Southampton, 2003.

[10] W. H. Press, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes in FORTRAN 77*, CUB, Cambridge, 2001

[11] D. Pepper, A. Kassab and E. Divo, *An Introduction to the Finite Element, Boundary Element, and Meshless Methods – with applications to heat transfer and fluid flow,* ASME Press*,* New York (due 2014 – in press).

[12] G. S. Gipson, *Boundary Element Fundamentals: Basic Concepts and Recent Developments in the Poisson Equation (Topics in Engineering),* Wessex Institute of Technology Press, Southampton, 1987.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

A Meshless Method for Creating 3-D Wind Fields using Sparse Meteorological Data

D. W. Pepper[1], C. Rasmussen[2] and D. Fyda[3]

Abstract

We consider the problem of developing an efficient numerical method for creating 3-D wind fields utilizing sparse meteorological tower data. Node points are placed within a region of interest based on topological features. Meshless methods do not require connective mesh generation. The method was implemented using MATLAB. Meteorological data consisting of wind speeds and directions are obtained from instrumented towers located within a domain and used to create a temporary, interpolated data set to all nodal points. The meshless method, based on the Kansa technique, creates a mass-consistent, 3-D wind field. The meshless method yields close approximations to results obtained with other numerical techniques.

Keywords: Mesh-free method, 3-D wind field

1. Introduction

Constructing 3-D wind fields is important in weather forecasting, wind energy assessment, and wind turbine siting. However, generating accurate (or realistic) wind fields can be difficult. The primary reason is that measurements of atmospheric flow are generally sparse and insufficient to resolve important flow phenomena. Linking meteorological data with numerical approaches is fairly routine, and has been employed for many years. The majority of numerical models used in meteorological simulations are based on finite difference and finite volume techniques, and more recently finite element methods with adaptation (Pepper and Wang, 2009).

Meshless methods are now becoming popular, due principally to their abilities to deal with complex geometrical problems with inhomogeneous or variable properties, the use of general-purpose algorithms, and no mesh requirement. There has been a growth of research interest for applying meshless (or mesh-free) methods for obtaining approximate solutions of differential equations (see Atluri and Zhu, 1998; Balachandran et al, 2009). In meshless methods there is no need of computing mesh elements induced by the selected nodes in the domain. The meshless method uses a group of nodes to derive solution without the need of connectivity relationship among them. In many situations where explicit mesh generation is difficult, the meshless method is emerging as an alternative to finite element analysis. The node placement problem for mesh-free applications is addressed in Atluri and Zhu (1998); Choi and Kim (1999); Gewali and Pepper (2010).

[1] UNLV, NCACM, Dept. of Mech. Engr., Las Vegas, NV, U.S.A

[2] Major, USAF, Dept. of Engineering Mechanics, USAF Academy, USAF, Colorado, U.S.A.

[3] UNLV, NCACM, Dept. of Mech. Engr., Las Vegas, NV, U.S.A

Mass consistent models for creating 3-D winds have been used for many years, and have been found to be effective for modeling atmospheric dispersion as well as for wind energy assessment studies. Such modeling techniques are discussed in Lange (1978), Sherman (1978), Goodin et al. (1980), Pepper (1991), Ratto et al. (1994), Finardi et al. (1998), Montero and Sanin (2001), and Pepper and Wang (2007). The problems stemming from most of the earlier applications deal with the coarse meshes and inabilities to deal effectively with refinement where terrain and/or velocities vary dramatically. The application of a mass consistent approach essentially poses a least-squares problem in the computational domain, i.e., minimizing the differences between observed and adjusted values.

In this study, a meshless approach is used to create a 3-D wind field over the Nevada Test Site, which is located northwest of Las Vegas, NV. Initial topological data is obtained from Digital Elevation Map (DEM) data, developed by the USGS. Numerical results are validated using meteorological data recorded in 1993 from 15 towers, and the wind fields are compared with model output from an hp-adaptive finite element model (see Pepper and Wang (2009)). When compared to the use of mesh-dependent schemes, the meshless method is more computationally effective and efficient.

1. Mass Consistent Wind Field

A diagnostic mass consistent model is derived from the continuity equation and incorporation of actual field data. The term "diagnostic" was used by Pielke (1984) to discuss different mesoscale meteorological models. Early research work on mass consistent models was undertaken by Sherman (1978) and later applied by Pepper (1991). The main idea of the technique is to match simulation values with measured meteorological data, using weighted averaging around the (usually) sparse data points to fill in values to all the nodes within the computational domain.

A surface wind field is constructed from measured data by interpolation over the initial mesh using inverse squared weighting ($1/r^2$, where r is the radial distance between the grid points and the tower locations). A fixed radius, R, is specified which indicates the distance beyond which the influence of a station's value is no longer felt (Goodin, et al., 1979; Kitada et al., 1983, Pepper, 1991).

The velocity at a grid point within an upper layer is calculated using the velocity at the grid point which has the same horizontal level as the tower layer. The top layer velocity is obtained using log-linear interpolation.

Vertical wind velocity is not a commonly observed variable in meteorology and its estimation appears as one of the more difficult problems for modeling studies. The vertical velocity is an integral component in many diagnostic and prognostic problems. In this study, vertical velocities are calculated at all node locations from the equation of continuity using the horizontal wind observations and accounting for the divergence correction, i.e.

$$w = -\int_0^z \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) dz \tag{1}$$

where u, v, w are velocities in x, y, z direction

A 3-D mass-consistent model was first used to generate wind fields for the ADPIC

pollutant transport model developed at Lawrence Livermore National Laboratory by Lange (1978). Early research work was conducted by Sherman (1978), who developed a mass-consistent model for wind fields over complex terrain and Dickerson (1978) who developed the mass-consistent atmospheric model for regions with complex terrain. The basis of their technique stems back to an objective analysis approach using a Sasaki variational technique (Sasaki, 1958). This technique was later applied by Mathur and Peters (1990) for application in air pollution modeling and by Pepper (1991) using a finite element approach to predict mesoscale wind fields over Vandenberg Air Force Base. Montero et al. (2005) developed genetic algorithms for improved parameter estimation with local tetrahedral mesh refinement in a wind model. The selection of parameters and computational methods implemented by different mass consistent models are discussed by Ratto et al. (1994). We elected to use this approach in lieu of other techniques because of its simplicity and ease of implementation with adaptivity. Warner et al (1983) describes the use of observed winds versus predicted winds employing a 3-D dynamic model to predict medium range atmospheric transport, along with short comings in accuracies attributed to each technique. Schaefer and Doswell (1979) discuss issues attributed to the interpolation of a vector field for atmospheric simulations; their analysis was based on finite differences and regional scales.

In this model, an Euler-Lagrange method is used in an integral function that minimizes the variance of the difference between the observed and analyzed variables (Sasaki, 1958). The function can be written as

$$E\left(u,v,w,\lambda\right)=\int_{\Omega}\left[\alpha_1^2\left(u-u_0\right)^2+\alpha_1^2\left(v-v_0\right)^2+\alpha_2^2\left(w-w_0\right)^2+\lambda\left(\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}\right)\right]d\Omega \quad (2)$$

where u_0, v_0, w_0 are observed velocity value in x, y, z direction, Ω is physical domain ($d\Omega \equiv dxdydz$), α_i are Gauss precision moduli where $\alpha_i^2 \equiv 1/(2\sigma_i^2)$ (σ_i are observation tower errors and /or deviations of the observed field from the desired adjusted field). For horizontal directions, the Gauss precision moduli are assumed identical, since apparent distinctions exist between horizontal and vertical directions, but not between x and y coordinates (see Sherman, 1978). These moduli are important in determining nondivergence wind fields over irregular terrain. Sherman (1978) suggested that $(\alpha_1/\alpha_2)^2$ should be proportional to the magnitude of the expected $(w/u)^2$. Using this relation and studies from Kitada et al. (1983), the 3-D flow fields tested for the minimum residual divergence occurred at about $(\alpha_1/\alpha_2)^2 = 0.01$. In this study, the values of α_1 and α_2 and were taken to be 0.01 and 0.1 respectively.

The corresponding Euler-Lagrange equations whose solutions minimize equation (2) are given as (see Sherman, 1978; Kitada et al., 1983; Pepper, 1991 and Ratto et al., 1994)

$$u = u_0 + \frac{1}{2\alpha_1^2}\frac{\partial \lambda}{\partial x} \quad (3)$$

$$v = v_0 + \frac{1}{2\alpha_1^2}\frac{\partial \lambda}{\partial y} \quad (4)$$

$$w = w_0 + \frac{1}{2\alpha_2^2}\frac{\partial \lambda}{\partial z} \tag{5}$$

where λ is the Lagrange multiplier. Substituting equations (3-5) into the continuity equation (assuming air density is constant in the computational domain[4]),

$$\frac{\partial \mathrm{u}}{\partial \mathrm{x}} + \frac{\partial \mathrm{v}}{\partial \mathrm{y}} + \frac{\partial \mathrm{w}}{\partial \mathrm{z}} = 0 \tag{6}$$

a Poisson equation for $\lambda(x, y, z)$ can then be obtained of the form

$$\frac{\partial^2 \lambda}{\partial x^2} + \frac{\partial^2 \lambda}{\partial y^2} + \left(\frac{\alpha_1}{\alpha_2}\right)^2 \frac{\partial^2 \lambda}{\partial \mathrm{z}^2} = -2\alpha_1^2 \left(\frac{\partial u_0}{\partial x} + \frac{\partial v_0}{\partial y} + \frac{\partial w_0}{\partial z}\right) \tag{7}$$

The ratio of α_1/α_2 allows one to adjust between horizontal or vertical influential preference.

The gradient $\partial\lambda/\partial$n in general is not zero and adjustment of the observed velocities from equation (3)-(5) can be obtained. A non-zero adjustment of the velocity normal to the boundary implies mass entering or leaving the volume. The boundary condition $\lambda = 0$ is appropriate for open or "flow-through" boundaries. When $\partial\lambda/\partial$n = 0 on the boundary, the adjusted values of the normal velocity are the same as the observed value. Setting $\partial\lambda/\partial$n = 0 on the boundary doesn't affect the normal velocity on the boundary. If the observed normal velocity is zero, there will be no transport of mass through the boundary. Therefore, $\partial\lambda/\partial$n = 0 is used for closed or "no-flow-through" boundaries.

The diagnostic procedure produces a mass consistent wind field which is realistic and fairly accurate that can then be refined to account for microscale topographic flow features. Sherman (1978) and Dickerson (1978) found that the use of the Sasaki (1958) technique was within a factor of 2 around 50% of the time and within an order of magnitude about 90% of the time. They used a fairly coarse mesh density at the time of their results. We wish to emphasize that the application of adaptation, as a minimum, can provide more detailed visualization and help in assessing the dynamics of the flow.

2. Meshless Approach

PDEs are traditionally solved using such methods as finite difference (FDM), finite volume (FVM), finite element (FEM), boundary element (BEM), and spectral (SM) techniques. Complex geometries are generally found with varying boundary conditions. This makes using traditional methods difficult as they require detailed mesh generation. For such situations, a meshless method may be better suited as it does not require this initial step. Figure 1 shows the difference in nodal spacing for a meshed vs. meshless approach (using random nodal positioning). Optimal placement of node points is addressed in more detail in Gewali and Pepper (2010).

[4] Since the velocities are low in this case, we assumed a constant density; however, variable density can be added to account for vertical temperature variation if warranted (see Sherman, 1978).

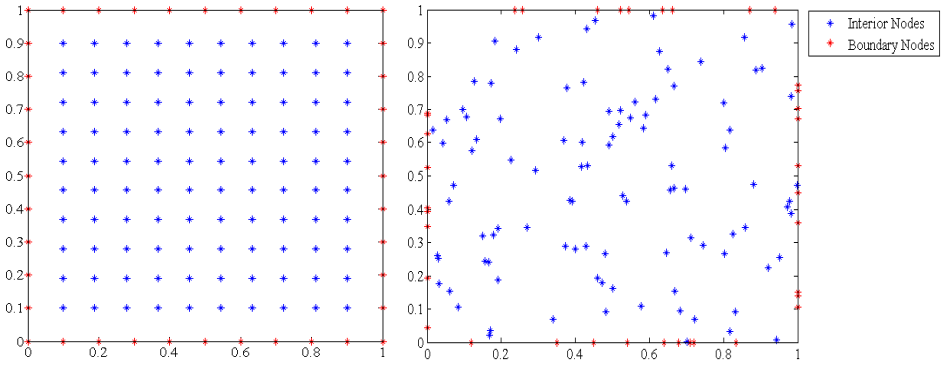


Figure 1. Structured meshed domain versus meshless domain with random spacing.

Although the roots of the development of meshless methods began in the 1970's, it is only in the last 15+ years where they have begun to see wider use. They have attracted notice for their ease in implementation, compared to the more traditional techniques, such as finite difference, finite volume, and finite element, that rely on a mesh consisting of interconnected node points to calculate values of interest. Although there are a variety of meshless method implementations, most have one property in common: they need not rely on a structured or unstructured mesh (there are some that claim to be meshless, but still rely on an mesh foundation). Strengths of a non-structured mesh become apparent when the problem of complex geometries or time-varying boundaries is considered.

An extensive explanation of the procedure involved in implementing a meshless method is discussed in Liu (2003). To begin, the problem domain must be conceptualized and represented with nodes scattered across the domain and boundary. The nodes are generally not distributed in a uniform manner and are denser in regions of greater change (Gewali and Pepper, 2010). Next, shape functions are used to relate the influence of each node to the other nodes in the domain. This is particularly necessary when the nodes are especially unevenly distributed. Shape functions are generally referred to as the support domain for the node of interest and there can be weighted influence. The support domain can take a circular or rectangular shape. The field variable, e.g. U, is then interpolated using the displacements at its nodes within the support domain. The shape functions can be used to write a PDE in nodal matrix form, and assemble global matrices for the entire problem domain. All that remains is to solve the matrices to obtain PDE solutions.

There are various categories of meshless methods. These include smoothed particle hydrodynamics method, reproducing kernel particle method, meshless Petrov-Galerkin method, local radial point interpolation method, finite point method, and the finite difference method with arbitrary irregular grids. Each of the methods has benefits and drawbacks. In this study, radial basis functions (RBF) were used as this did not require special consideration for nodal placement.

2.1 The RBF Method

One of the simplest implementations of a meshless method is the RBF method. Specifically, Kansa's approach (2005) is followed in this study. In this method, a basis function relates the influence of surrounding nodes on the node of interest. The nodes closest to the node of interest have the greatest influence. Nodes increasingly farther away have decreasing influence.

First, for a one-dimensional (1-D) problem, the distance d between radial position r is defined as:

$$d_i = [(r - r_i)^2]^{1/2} \tag{8}$$

Different basis functions ϕ have been proposed and can be viewed in Table 2.1:

Table 2.1 – Typical Radial Basis Functions

Item	Name	Expression	Shape Parameters
1	Multiquadrics (MQ)	$\phi_j(d) = (d_i^2 + c_j^2)^{\beta}$	c, β
2	Gaussian (EXP)	$\phi_j(d) = \exp(-C_j d_j^2)$	C
3	Thin Plate Spline (TPS)	$\phi_j(d) = d_j^{\eta}$	η
4	Logarithmic RBF	$\phi_j(d_i) = d_i^{\eta}\log(d_i)$	η

The most commonly used basis function is the MQ as proposed by Hardy (1971) with an exponent of β = +0.5. The MQ form shown in Table 2.1 is a general form with the exponent β as another parameter to be optimized. If the exponent is set at -0.5, this gives the inverse MQ form. As seen from the table, there are two shape parameters to be tuned: c and β. The shape parameter c varies with position and is defined in section 2.4.

The Gaussian RBF gives an exponential function of the distance. The shape parameter in this RBF controls the decay rate. The thin plate spline (TPS) is a special case of the MQ as it only has one shape parameter to optimize rather than two. Finally, the logarithmic RBF can also be used with one shape parameter to optimize.

The MQ form has been widely used in constructing approximate solutions to PDEs. Therefore, the following analysis will be conducted using the MQ approach. Upon substitution of the distance formula, the basis function ϕ is described as:

$$\phi_j = [(r - r_j)^2 + c_j{}^2]^{\beta} \quad ,j = 1,2 \dots N \tag{9}$$

where N is the total number of nodes.

Upon defining the domain Ω, the continuous function U(x) is:

$$U(x) = \sum_{j=1}^{N} \phi_j(x)\alpha_j \tag{10}$$

where α_j is a series of coefficients. These coefficients must be determined at each continuous function value.

In order to solve partial differential equations (PDEs), we apply a linear operator ($\mathcal{L}$) on the interior $\Omega \backslash \partial\Omega$ to the continuous function as defined above. This gives

$$\mathcal{L}U(x) = \sum_{j=1}^{N} \mathcal{L}\phi_j(x)\alpha_j \tag{11}$$

From this relation, the linear operator (the PDE) must be applied to the basis function. This requires calculation of the first and second derivatives of the basis functions. Applying the boundary operator ($\mathcal{B}$) on the boundary $\partial\Omega$ as:

$$\mathcal{B}U(x) = \sum_{j=1}^{N} \mathcal{B}\phi_j(x)\alpha_j \quad (12)$$

Dirichlet, Neumann, or Robin conditions can be used for the problem.

2.2 Radial-Polynomial Basis

The use of pure radial functions tends to be inconsistent and so for further accuracy, polynomials added to the basis functions can ensure consistency. However, addition of the polynomials complicates the solution matrices making them difficult to implement. They are, however, accounted for in the following manner:

$$U(x) = \sum_{j=1}^{N} \phi_j(x)\alpha_j + \sum_{k=1}^{M} p_k(x)\gamma_k \quad (13)$$

where $p_k(x)$ is a polynomial of degree k, and the expansion is subject to the constraint:

$$\sum_{j=1}^{N} p_k(x)\alpha_j = 0 \quad (14)$$

This constraint guarantees unique approximation (see Liu, 2003). The interior operator and boundary operator have similar expansions:

$$\mathcal{L}U(x) = \sum_{j=1}^{N} \mathcal{L}\phi_j(x)\alpha_j + \sum_{k=1}^{M} \mathcal{L}p_k(x)\gamma_k = f \quad (15)$$

$$\mathcal{B}U(x) = \sum_{j=1}^{N} \mathcal{B}\phi_j(x)\alpha_j + \sum_{k=1}^{M} \mathcal{B}p_k(x)\gamma_k = g \quad (16)$$

These are also subject to the constraints:

$$\sum_{j=1}^{N} \mathcal{L}p_k(x)\alpha_j = 0 \quad (17)$$

$$\sum_{j=1}^{N} \mathcal{B}p_k(x)\alpha_j = 0 \quad (18)$$

As stated previously, the addition of polynomials can be beneficial for increasing solution accuracy. It does, however, increase the computation time and complexity of implementation. Also, the solution still relies on well-chosen shape parameters.

2.3 Shape Parameters

In the previous discussion, the shape parameter c was introduced. The shape parameter is critical for determining the accuracy of PDE solutions when applying the radial basis functions. Because of this, considerable attention has been given for determining values of the shape parameter and some formulations can be seen in Table 3.2.

Table 2.2 –Shape Parameter Formulations

Number	Shape Parameter, c
1 [40]	$c = 0.815d$
2 [41]	$c = 1.25d/\sqrt{N}$
3 [42]	$c^2 = c^2 \min(\frac{c_{max}^2}{c_{min}^2})^{(j-1)/(N-1)}$
4 [43]	$c = 2/\sqrt{N}$
5 [39]	$c_j^2 = C_1[1 + C_2(-1)^j]$

Most of the shape parameter formulations (1,2,4) rely on the distance d and number of nodes, N. Other formulations show a changing shape parameter based on the node

position. In summary, the shape parameter can depend on many factors including (see Roque and Ferreira, 2009): number and distribution of nodes, the basis function, and computer precision.

Kansa (2005) made use of the fifth formulation and that was the formulation used for this solution procedure. Different shape parameter constants were used for the interior $\Omega \backslash \partial\Omega$ and the boundary $\partial\Omega$. For the interior, the parameter C_1 was defined as the ideal shape parameter for the domain and C_2 determined the amplitudes of C_1 depending on whether j is even or odd. The constant C_2 should be varied between 0.25 to 0.33. For the boundary, the constant C_2 should be varied between 0.49 and 0.55.

2.4 The Euler-Lagrange Equations

To illustrate the application of the meshless method with RBF, eqn (7) can be expressed as

$$\begin{aligned} \nabla^2\lambda &= f(\mathbf{x}), \quad \mathbf{x} \in \Omega \\ \lambda &= g(\mathbf{x}), \quad \mathbf{x} \in \Gamma \end{aligned} \tag{19}$$

where **x**=(x,y,z) and f(**x**) and g(**x**) denote the divergence of the observed velocity values. Now approximate λ assuming

$$\lambda(\mathbf{x}) = \sum_{j=1}^{N} \phi(r_j)\lambda_j \tag{20}$$

where r is defined, assuming MQ as the basis function, as

$$\phi(r_j) = \sqrt{r_j^2 + c^2} = \sqrt{(x - x_j)^2 + (y - y_j)^2 + (z - z_j)^2 + c^2} \tag{21}$$

Likewise, the derivatives can be expressed as

$$\frac{\partial\phi}{\partial x} = \frac{x - x_j}{\sqrt{r_j^2 + c^2}}, \quad \frac{\partial\phi}{\partial y} = \frac{y - y_j}{\sqrt{r_j^2 + c^2}}, \quad \frac{\partial\phi}{\partial z} = \frac{z - z_j}{\sqrt{r_j^2 + c^2}}$$

$$\frac{\partial^2\phi}{\partial x^2} = \frac{(y - y_j)^2 + (z - z_j)^2 + c^2}{\sqrt[3]{r_j^2 + c^2}}, \frac{\partial^2\phi}{\partial y^2} = \frac{(x - x_j)^2 + (z - z_j)^2 + c^2}{\sqrt[3]{r_j^2 + c^2}}, \frac{\partial^2\phi}{\partial z^2} = \frac{(x - x_j)^2 + (y - y_j)^2 + c^2}{\sqrt[3]{r_j^2 + c^2}} \tag{22}$$

Substituting into the original equation set, one obtains

$$\begin{aligned} \sum_{j=1}^{N} \nabla^2\phi(r_j)\lambda_j &= f(\mathbf{x}), \quad i = 1, 2, \cdots, N_I \\ \sum_{j=1}^{N} \phi(r_j)\lambda_j &= g(\mathbf{x}), \quad i = N_I + 1, N_I + 2, \cdots, N \end{aligned} \tag{23}$$

which produces an NxN linear system of equations for the unknown λ_j. Figure 2 shows an

arbitrary domain with randomly placed node points and the circle of influence surrounding each node.

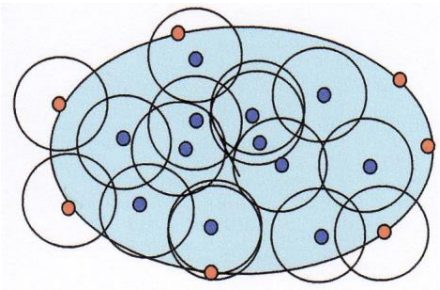


Figure 2. Node placement and circle of influence

3. Simulation Results

The Nevada Test Site (NTS) is a massive outdoor laboratory established in 1950 as the United States' continental nuclear weapons testing area. The location of the NTS is shown in Fig. 3 (a), near the southern part of Nevada and beginning about 65 miles northwest of Las Vegas. The NTS is defined by West longitude 115.5^0 to 116.5^0, and North latitude 36.5^0 to 37.5^0. A total of 26 (currently 15 operational) towers were placed within the NTS to support nuclear testing in the early 1950's. An unstructured surface mesh with the locations of the towers is shown in Fig. 3 (b). A set of four terrain following layers (10m, 50m, 300m and 1000m) was used to construct the initial node positions based on data from the USGS DEM (digital elevation map). This domain consists of 3750 nodes - a 3-D view of an equivalent mesh generated by an hp-adaptive finite element model is shown in Fig. 4 (a). In the Fig. 4(b) 2-D view, the wind velocities obtained at the tower locations are incorporated as nodal points within the domain. The hp-adapted finite element model subsequently resulted in a mesh consisting of 10,392 elements and 12,508 nodes with a horizontal fine mesh resolution is 950m.

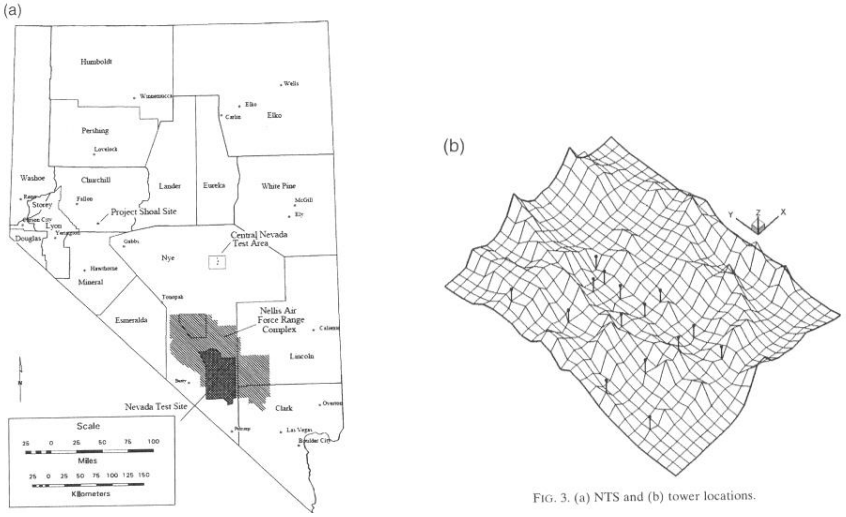


Figure 3. (a) Nevada Test Site and (b) location of meteorological towers

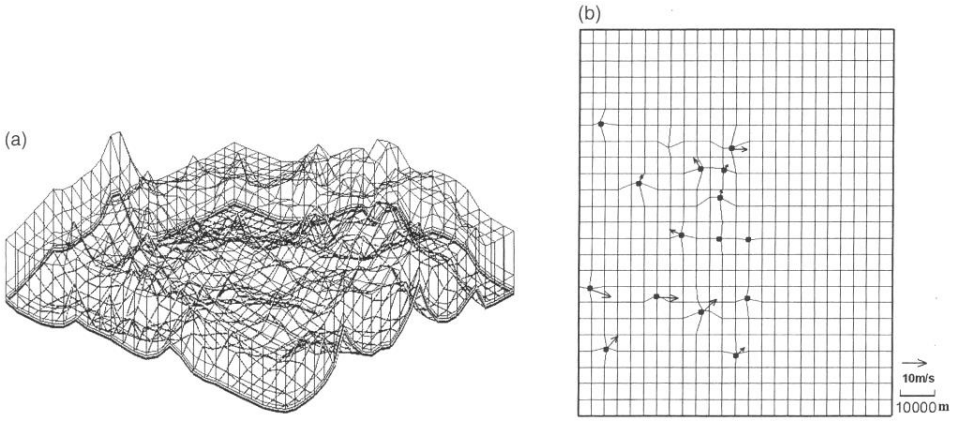


Figure 4. (a) 3-D equivalent mesh and (b) measured wind velocities

Three-dimensional, mass consistent wind fields were constructed utilizing the data obtained from the tower network. The recorded averaged measured data for each tower during the time interval 1:00am to 1:45am on Jan. 1, 1993 was used with 15 minute intervals. Wind speeds and directions associated with the fifteen towers during this period of time are shown in Table 3.1, where r refer to the wind speed magnitude (m/s) and θ refer to the wind direction (the angle between the wind vector and the horizontal direction in counter clock direction). The averaged wind velocity vectors at each tower location are shown in Fig. 4 (b).

TABLE 2. Wind speed records of NTS towers on 1 Jan 1993. The units are meters per second for r, U, and V, and degrees for θ.

Time (LT)	0100		0115		0130		0145		Avg	
Tower no.	r	θ	r	θ	r	θ	r	θ	U	V
2	6	313	7	311	6	319	6	13	−3.07	4.76
5	1	336	1	273	3	156	1	134	0.13	−0.62
6	1	359	1	193	1	348	3	348	−0.27	−0.98
9	3	31	4	9	4	15	3	102	1.53	2.44
11	1	103	2	151	1	96	2	23	0.93	−0.059
14	6	295	7	296	6	295	5	296	−5.42	2.57
15	3	43	7	95	7	98	9	103	6.18	−0.35
17	4	356	3	17	5	18	3	355	0.46	3.65
18	5	12	4	18	4	24	4	14	1.22	4.06
20	1	161	1	152	1	147	1	158	0.43	−0.90
23	4	65	6	37	5	49	4	28	3.22	3.32
24	10	107	10	108	9	115	9	11	8.91	−3.25
25	7	33	8	38	8	50	7	44	4.93	5.59
26	8	92	10	97	9	85	10	98	9.20	−0.52
27	9	38	9	37	8	56	8	65	6.21	5.54

Mass consistent wind fields generated by the meshless method are compared to the hp-adapted finite element results for the 10m and 50m levels in Figs. 5(a,b) and 6(a,b). As can be seen, the meshless results appear close in values and pattern compared with the high resolution finite element model. While both methods produce realistic wind fields, the computational work utilizing the meshless approach was significantly less, and was run using Matlab on a PC, versus the finite element model running on a supercomputer and employing FORTRAN.

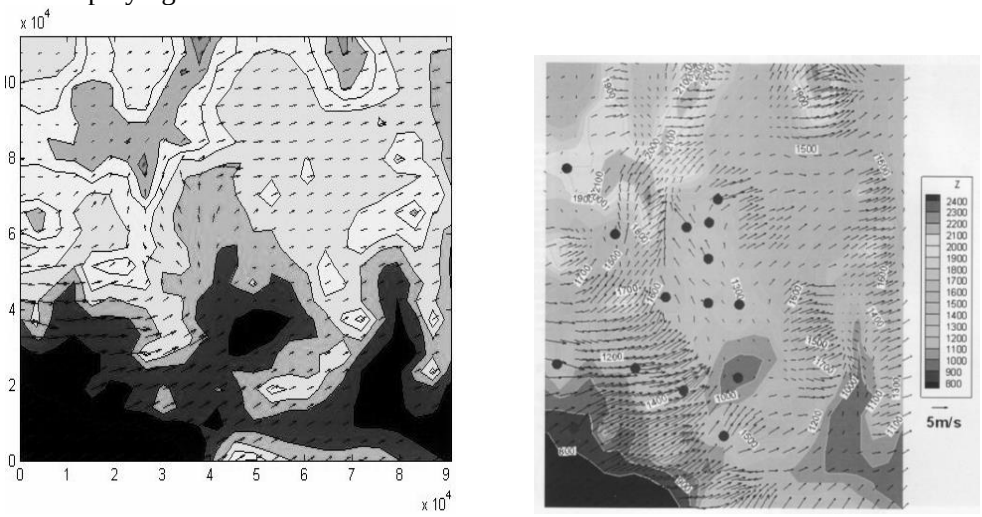


Figure 5. (a) Meshless results for 10 m level and (b) hp-fem results

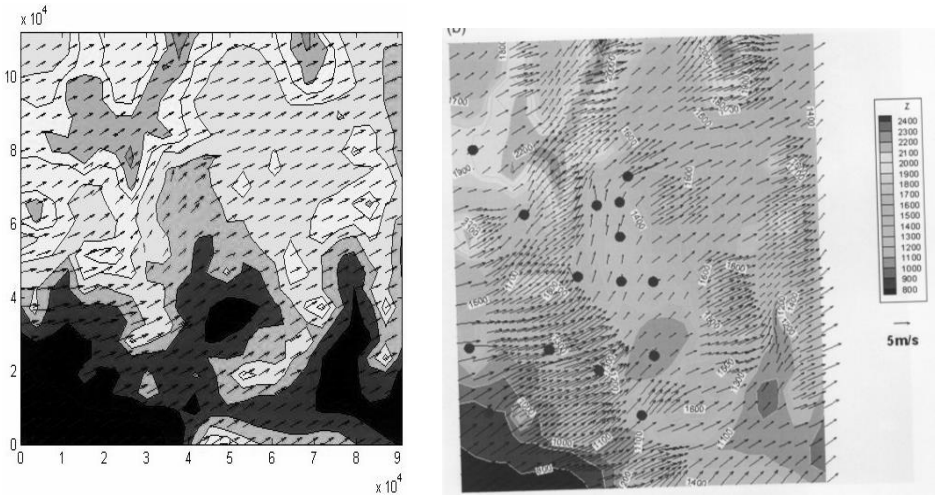


Figure 6. (a) Meshless results for 50 m level and (b) hp-fem results

3. Conclusion

A meshless method has been used to create a mass consistent, 3-D wind fields utilizing a Sasaki variational approach and sparse meteorological tower data. Wind fields are presented for the Nevada Test Site. Model results obtained with the meshless method are in good agreement with results obtained from an hp-adaptive finite element model. The application of the meshless method permits the rapid construction of localized high resolution wind field estimates in regions where conventional numerical methods would require large storage and computational times. The method appears useful in refining meteorological tower locations and quickly determining potential wind turbine sites in wind energy assessment studies.

Meshless methods are cost efficient – this is well documented in the literature. The application of a meshless approach for meteorological simulations appears promising when compared to large-scale calculations over regions where more conventional numerical approaches would require localized fine meshing to accurately capture rapidly changing flow features.

4. References

Acres, J. M., 2010: Modeling Arterial Blood Flow Using Radial Basis Functions, MS Thesis, University of Nevada Las Vegas, LV, NV, 55p.

Atluri, S. N. and T. Zhu, 1998: A New Mesh-less Local Petrov-Galerkin Approach in Computational Mechanics, *Computational Mechanics*, **22**, 117-127.

Balachandran, G. R., A. Rajagopal and S. M. Sivakumar, 2009: Mesh free Galerkin Method Based on Natural Neighbors and Conformal Mapping, *Computational Mechanics*, **42(6)**, 885-905.

Choi, Y. and S. J. Kim, 1999: Node Generation Scheme for the Mesh-less Method by Voronoi Diagram and Weighted Bubble Packing, *Fifth US National Congress on Computational Mechanics*, Boulder, CO.

Dickerson, M. H., 1978: MASCON-A mass consistent atmospheric flux model for regions with complex terrain. *J. Appl. Meteor.*, **17**, 241-253.

Fasshauer, G. E., 2007: *Meshfree Approximation Methods with MATLAB,* World Scientific Pub. Co., Singapore, 500p.

Finardi, S., G. Tinarelli, P. Faggian, and G. Brusasca, 1998: Evaluation of different wind field modeling techniques for wind energy applications over complex topography. *J. Wind Eng. Ind. Aero.*, **74-76**, 283-294.

Gewali, L. and D. W. Pepper, 2010: Adaptive Node Placement for Mesh-Free Methods, *ICCES'10*, Las Vegas, NV.

Goodin, W. R., G. J. Mcrae, and J. H. Seinfeld, 1980: An objective analysis technique for constructing three-dimensional urban-scale wind fields. *J. Appl. Meteor.*, **19**, 98-108.

Goodin, W. R., G. J. Mcrae, and J. H. Seinfeld, 1979: A comparison of interpolation methods for sparse data: application to wind and concentration fields. *J. Appl. Meteor.*, **18**, 761-771.

Hardy, R.L., 1971: Multiquadric equations of topography and other irregular surfaces, *J. of Geophysical Res*., **176**, 1905-1915.

Kansa, E. J., 2005: Highly accurate methods for solving elliptic and parabolic partial differential equations, WIT Transactions on Modelling and Simulation, **39**, 5-15.

Kitada, T., Kaki, A., Ueda, H. and L.K. Peters, 1983: Estimation of vertical air motion from limited horizontal wind data-a numerical experiment. *Atmos. Environ.*, **17**, 2181-2192.

Lange, R, 1978: A three-dimensional transport-diffusion model for the dispersal of atmospheric pollutants and its validation against regional tracer studies. *J. Appl. Meteor.*, **17**, 241-256.

Li, X. Y.; S. H. Teng; A. Ungor, 2000: Generating a Good Quality Point Set for the Mesh-less Methods, *Computer Modeling in Engineering Sciences (CMES), Vol. 1, No. 1, pp. 10-17.*

Liu, G.R., 2003: *Mesh Free Methods: Moving beyond the Finite Element Method*, CRC Press, Boca Raton.

Mathur, R. and L. K. Peters, 1990: Adjustment of wind fields for application in air pollution modeling. *Atmos. Environ.*, **24**, 1095-1106.

Montero, G. and N. Sanin, 2001: 3-D modeling of wind field adjustment using finite differences in a terrain conformal coordinate system. *J. Wind Eng. Ind. Aero.*, **89**, 471-488.

Pepper, D. W, 1991: A Finite Element Model for Calculating 3-D Wind Fields over Vandenberg Air Force Base. *29th AIAA Aerospace Sciences Meeting*, Jan. 7-10, Reno, NV, Amer. Institute Aeronaut. Astronaut., AIAA paper 91-0451.

Pepper, D. W. and X. Wang, 2009: An h-Adaptive Finite-Element Technique for Constructing 3D Wind Fields, *J. Appl. Meteor. and Climatology*, **48**, 580-599.

Pepper, D. W. and X. Wang, 2007: Application of an *h*-adaptive FEM for Wind Energy Assessment in Nevada, *Renew. Energy*, **32**, 1705-1722.

Pielke, R, 1984: *Mesoscale Meteorological Modeling*, Academic Press, New York, N. Y., 612pp.

Ratto, C. F., R. Festa, C. Romeo, O. A. Frumento, and M. Galluzzi, 1994: Mass-consistent models for wind fields over complex terrain: the state of the art, *Environ. Software*, **9**, 247-268.

Roque, C.M.C., and Ferreira, A.J.M., 2009: Numerical Experiments on Optimal Shape Parameters for Radial Basis Functions, *Num. Meth. Part. Diff. Eqns.*, **26(3)**. 675-689.

Sasaki, Y., 1958: An objective analysis based on the variational method, *J. Meteor. Soc. Japan*, **36**, 77-88.

Schaefer, J. T. and C. A. Doswell III, 1979: On the interpolation of a vector field. *Monthly Weather Review*, **107**, 458-476.

Sherman, C. A, 1978: A Mass-consistent model for wind field over complex terrain. *J. Appl. Meteor.*, **17**, 312-319.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Analysis of modelling capabilities of ablative fire-proof properties in polymer fibre composites

P. Przybyłek[1]*, R. Szczepaniak[1,2],W. Kucharczyk[3] ,R. Rudzki[2]

[1)] Faculty of Airframe and Engine
Air Force Academy, 35 Dywizjonu 303 Street, 08-521 Dęblin , Poland
e-mail: sqcdr@wp.pl , robert.szczepaniak@o2.pl , wojciech.kucharczyk@uthradom.pl

[2)]Faculty of Mechatronics and Aeronauitics
Military University of Technology, gen. S. Kaliskiego 2, 00-908 Warsaw, Poland

[3)]Faculty of Mechanical Engineering
University of Technology and Humanities, Krasickiego 54B, 26-600 Radom, Poland

Key words: ablative, passive fire-proof protections, epoxy composites, kevlar, carbon fibers

Abstract

The inverse heat conduction and radiation problems are advanced when they apply to many different layers of the composite extreme temperatures on the outer surface. The approximate formulation and solution to ablative problem were conducted by applying FEM Comsol software. A rectangular sample was replaced with a cylindrical sample of an equivalent volume. The tested model does not take into account the changes caused by the impact of geometrical dimensions of the flame heat. The working conditions for the test sample were non-standard to such an extent that we were forced to estimate the thermophysical properties as referred to the applied temperature range. In addition, the model does not take into account the combustion heat; the given forced heat transfer was modelled by adding, to the tested model, a plate with a given diameter of the burner with the condition of constant temperature, taking into account the radiation and convection. We assume that the considered simplifications are non-significant for the model. Typical results of numerical tests compared with our own experimental results have been listed in the paper. It can be concluded that the presented approximate model correctly represents the actual process of ablation.

1 Introduction

The use of modified plastics as ablative materials protecting against an excessive temperature increase was connected with the middle of 20th century, directly with arms industry as well as aeronautical, rocket and space techniques. These materials can also be used in the design of passive fire-proof protections for large cubature supporting elements in building structures[2], communication tunnels and for the protection of data stored in electronic, optical and magnetic carriers. This paper reports results of studies on modelling of ablative and thermal properties of epoxy composites with hybrid fabrics

reinforcement (kevlar and carbon fibers) filled with a mixture of epoxy resin and mineral nanoclays (layered silicates). The composites were treated with hot combustion gases to detect the temperature profiles across the studied samples, the back side temperature of specimen t_s, the average linear rate of ablation v_a, and their mass waste U_a during ablation processes. The paper briefs assumptions and requirements on research how to create ablative and thermal properties of epoxy composites with hybrid fabrics reinforcement.

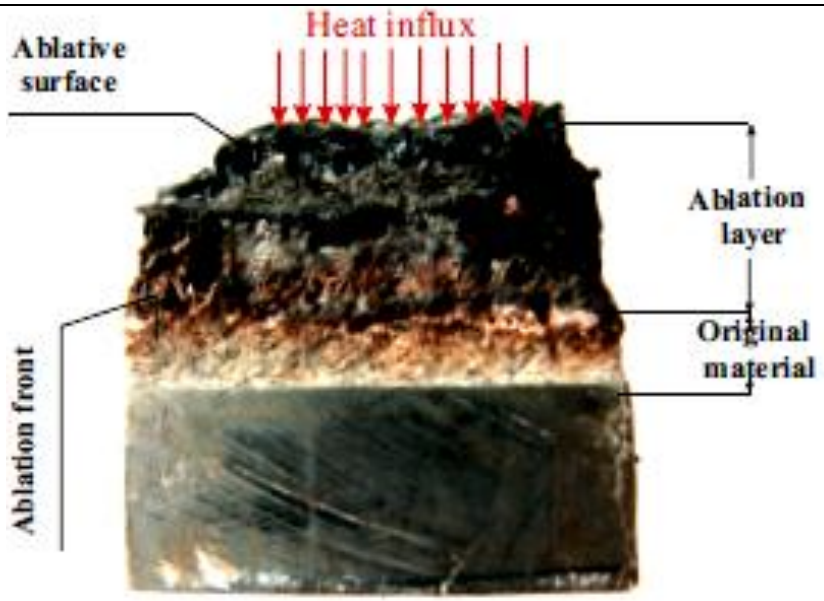


Fig. 1 Physical model of ablation [1]

2 Description of the method

Ablation is a self-regulating heat and mass transfer process, which, due to physical and chemical reactions, leads to irreversible structural and chemical changes of a material combined with heat absorption at the same time. The process is initiated and sustained by external sources of thermal energy. (fig. 1) [1]

Once an ablation surface interacts with a high-temperature heat stream, an ablation process is initiated. This is when the ablation surface, under the influence of temperature, undergoes internal structural changes, which protect lower layers and affect thermo protective properties of a material. If an ablation shield has a multi-surface structure, the process takes place in cycles: the external layer is sometimes burned, falling off the larger part; subsequently another layer undergoes ablation changes. Among the factors which are not directly linked with the ablation layer, the cohesion of native material plays a tremendous role in its exploitation reliability, particularly in the case of thermal interactions. The need to assess mechanical properties of a native ablative material is connected with the fact that high mechanical stresses arising in a non-degraded layer of polymer composites, during a process of high-temperature heat transfer and heat conduction exert a much stronger influence on composite cohesion rather than heat stresses in a degraded ablation layer. It is thermomechanical stresses of the native material that are responsible for the damage of a composite. Therefore it is vital to also evaluate the basic mechanical properties of the composites in question. [2, 3, 4]. Despite numerous years of applying ablation materials, there is still room to fully determine the quality and quantity relationships between the type-phase composition and ablation properties, in the context of

other exploitation features of composites used for thermo protective shields [5, 6].Polymer composites with fibre reinforcement in the form of glass fibres and carbon fibres have been quite thoroughly investigated, also in connection with various powder fillers [8, 9, 10, 11, 12].It is also true of composites with reinforcement in the form of carbon fibre as well as multilayer composites with hybrid reinforcement [13, 14].
The phase composition of samples and the number of conducted experiments *(N = 8)* were determined on the basis of a plan of experimental research, i.e. 2-level full factorial orthogonal design matrix for replicated 2^3 factorial experiments.[15].
On the basis of the analysis of thematic publications [16-÷19], we selected phase components of the composites: epoxide resine - Epidian 52 – base, cured in room temperature by the following curing agents: TFF (27 ns) produced by Z. Ch. Organika-Sarzyna S.A. in Nowa Sarzyna. The ablative properties of the composition of resin were modified with a layered aluminosilicate, *Bentonite Special Extra*, which contained 75% calcium montmorillonite – MMT – (*Zębiec* Mining-Metal Plant located in Zębiec). The hybrid reinforcements of the composite were balanced fabrics: aramid of 230g/m^2 basis weight and carbon of 160g/m^2 basis weight, layered in a alternating and even way in the composite layer (fig.2).

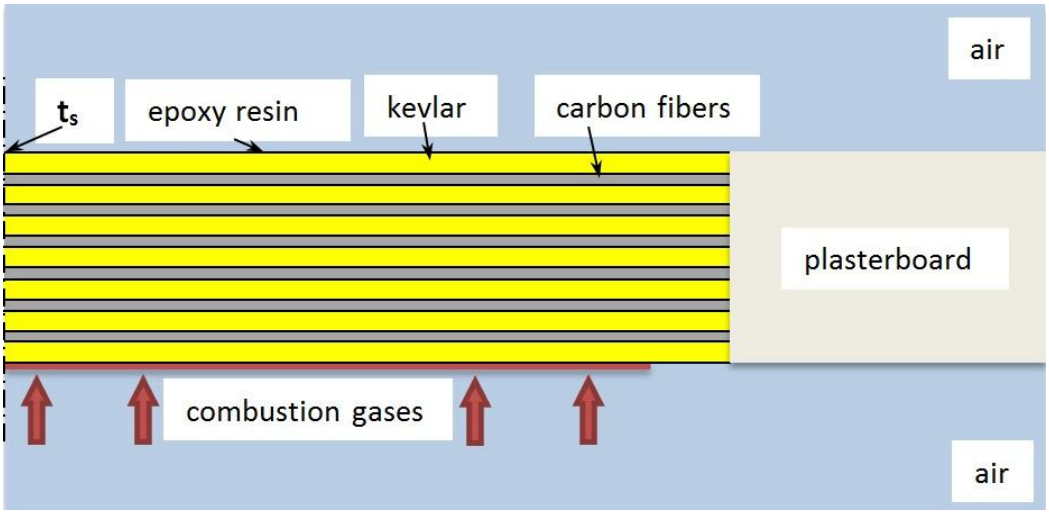
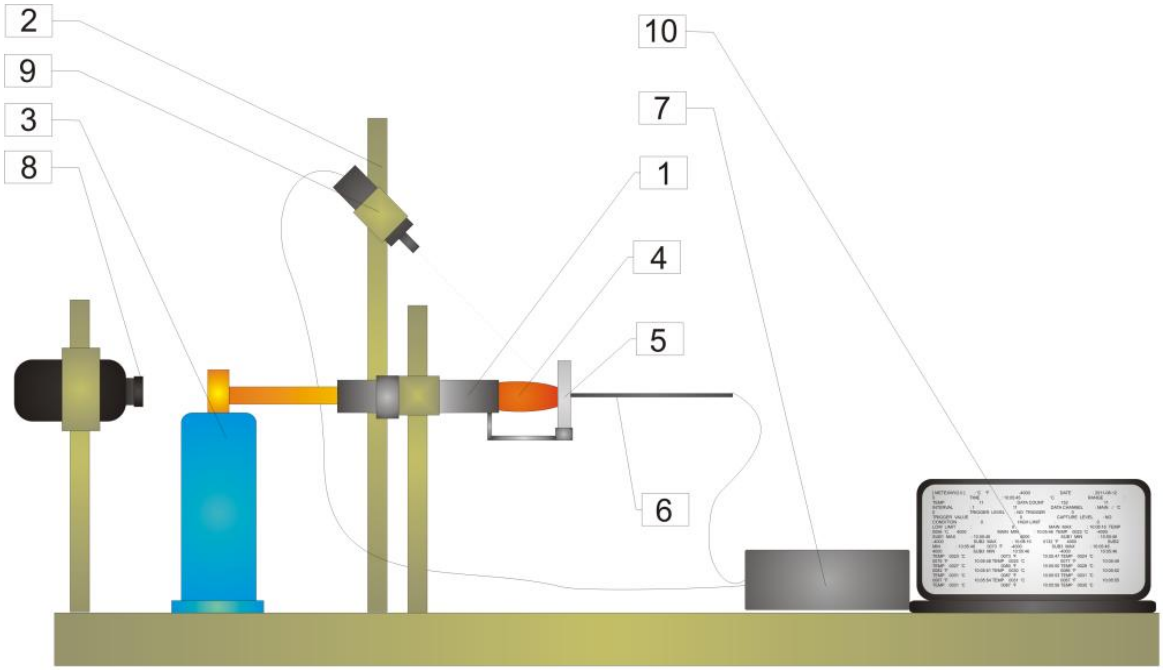


Fig. 2: Sample structure

3 Experimental verification

3.1 Description of the model and the real object of study

The thermo protective ablative research was realized on the author's own construction stand(fig. 2), in accordance with the assumptions and methodology described in the available studies [19,20]. The ablative specimens (composite blocks of 11 x 25 x 35) were placed in a fire resistant insulating board made with plasterboard, where they were treated with a stabilized stream of combustible gases for a time period of τ = 150 s. The source of heat was burning a mixture of liquid gases – propane butane, which gave off a flame temperature of 1100 ^{o}C.



Research stand: 1 - “ablative gun”,2 - stand, 3 - the gas cylinder with burner nozzle,4 -flame, 5 - pspeciment, 6 - thermoelement, 7 - measure, 8 - thermal imaging camera, 9 - pyrometer, 10 - computer.

Fig. 3 Laboratory research stand

3.2 Description of numerical model

We prepared a numerical model as well as performing numerical simulations in the COMSOL 3.5a environment, by COMSOL Multiphysics. The analysis was conducted by means of the heat transfer module, where we simulated the physical processes of heat conduction, heat transfer through radiation and convection.

During the numerical analysis we introduced simplifications of the following:

- the process of gas combustion was simulated through an introduction of a uniform distribution of the temperature field on the surface of the examined sample, with a geometry close to the radius of flame geometry, as well as introducing a replacement value for air due to a turbulent flow of heated air;
- the 3D model was replaced with a two-dimension axially symmetrical model, with the volume of the tested sample similar to the volume of the actual sample (Fig. 2);
- the introduced thermophysical values of the materials are replacement and approximate values (some values were used as values dependent upon temperatures – Table 1).

3.3 Modelling construction elements of the system

The thermophysical properties of the specimen and the nearest environment of the simulation examination of the numerical model (fig. 2) were adopted on the basis of available scientific journals, and are listed in the table below (see Table 1):

Table 1: Materials properties.

	Replacement parametres for numerical calculations		
Materials	Heat capacity [J/kg K]	Thermal conductivity [W/m K]	Density [kg/m^3]
Air	1	2+0.005(T-273)*	1.2-0.001(T-273)
Epoxy resin[22]	2030 if T=473 K	1	1200
Carbon fibers [28,29]	$1921/(1+2.189 \cdot e^{-0.0064\,(T-273)})$	10	1760
Kevlar [21,24,25]	$1380+7.5 \cdot (T-273)$	$0.05+(T-273) \cdot 10^{-6}$	1440
Plasterboard [27]	0.8	0.17-0.00025 (T-273)	600
*replacement value resulting from an introduction of air flow through a heat stream from the burner			

4 Analysis results

4.1 Temperature field and the ablation layer of the examined material

Fig. 4 shows a graphic distribution of the temperature field for the constant replacement value on the surface of the tested sample, at t=160 s. The measurements of the temperature field for the introduced numerical model prove that the temperature on the opposing surface to the surface of the thermal activation surface, after a time span of 160 sec, does not exceed 100° C. Moreover, the ablation surface of the composite material was marked, equalling approximately 1.8 mm, with a radius of r=0, where the temperature exceeded the critical value for the examined material.

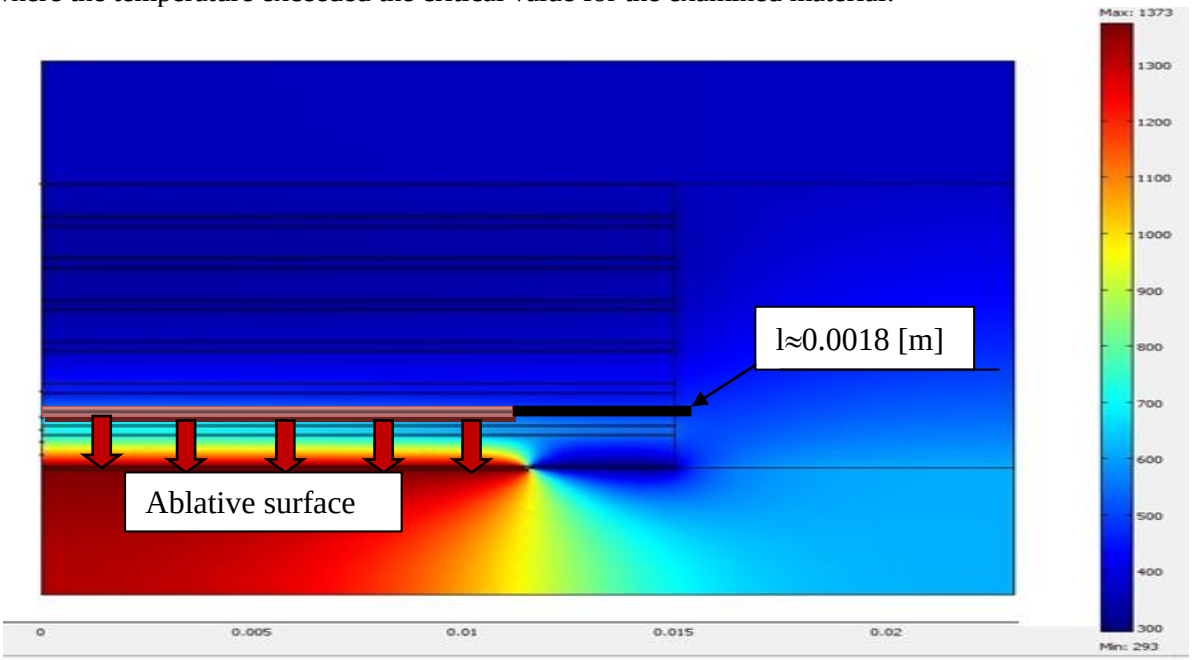


Fig. 4: Illustration of the ablation layer, after a simulation of the examined sample

In addition, we also presented temperature distribution in the function of time for different values of the distance to thermal activation (Fig. 5). It depicts rough values of the change of the ablation front across particular material layers in the time function.

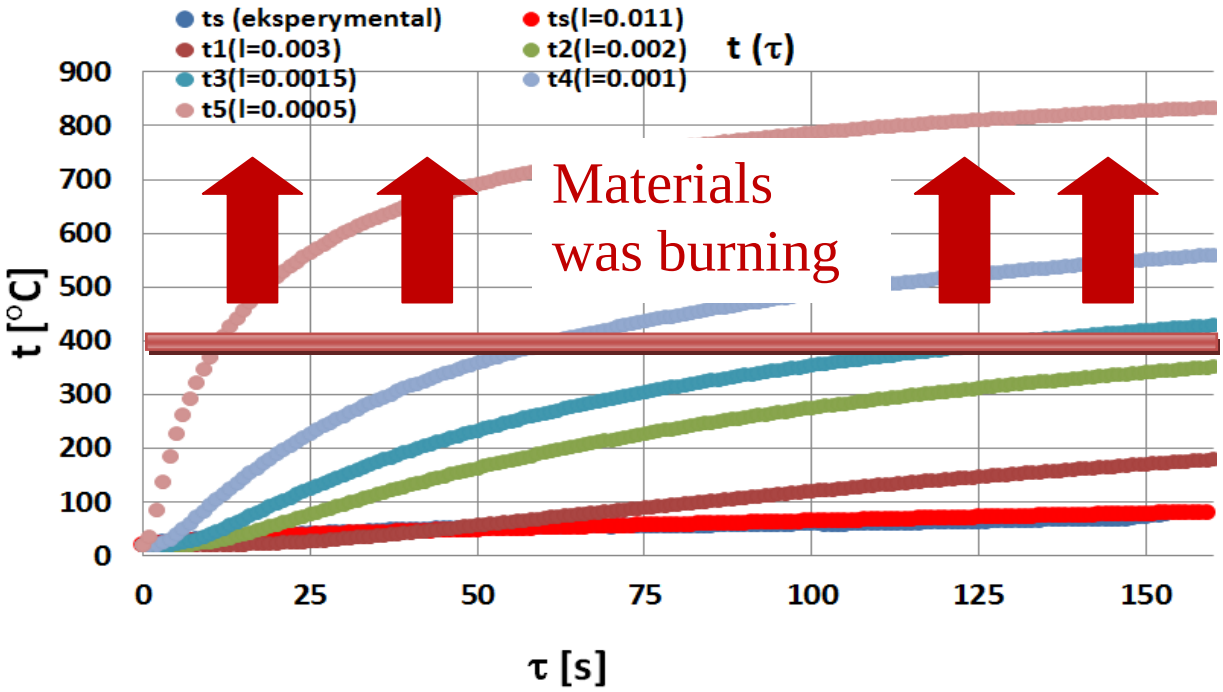


Fig. 5: Simulation result of the examined sample

4.2 Heat resistance properties of a material

The fundamental parametre which specifies the heat resistance properties of a material is the temperature on the rear surface of the examined sample (t_s). The temperature distribution on the rear surface is presented on the comparative diagram below. We can observe only a slight discrepancy between these findings and the experimental findings, which may indicate a proper approximation of the introduced numerical model simplifications as referred to the actual model.

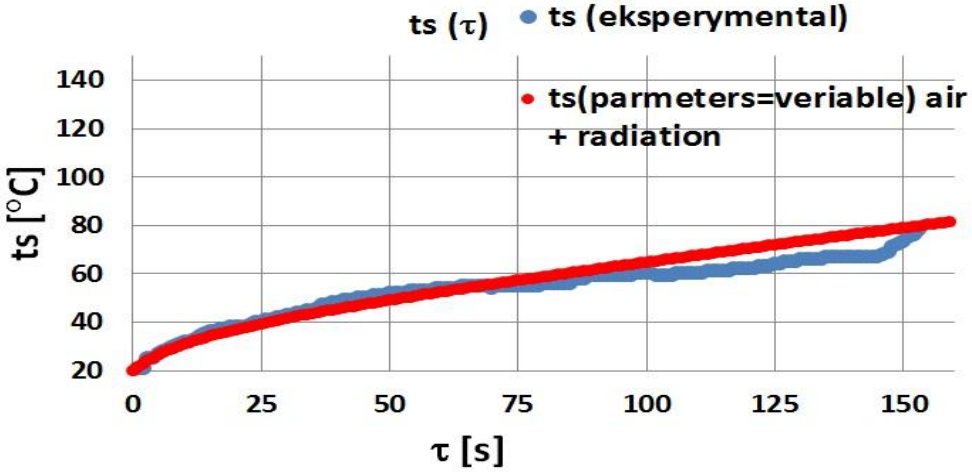


Fig. 6: Numerical and experimental results t_s.

The best ablative thermo protective properties were noticed in those composites which achieved, at the same time: the lowest final temperature of the rear surface area of the specimen t_s (ensuring the highest temperature reduction on the casing's width), the lowest rate of ablation v_a (the smallest degrading speed of the native material structure) and the smallest ablative mass loss U_a (the thickest ablative exploitation layer with low thermal diffusivity).
In the investigated range of temperature of the ablative area (900 ÷ 1300) ^{o}C, the lowest values of all the components of the response function are absorbed by the composite with phase composition no 3 , which contains 60% base (cured by *TFF* hardener), 14 layer of aramid fabric and 6 layer of carbon fabric as well as 15% aluminosilicate, *Bentonite Special Extra(fig.8).*

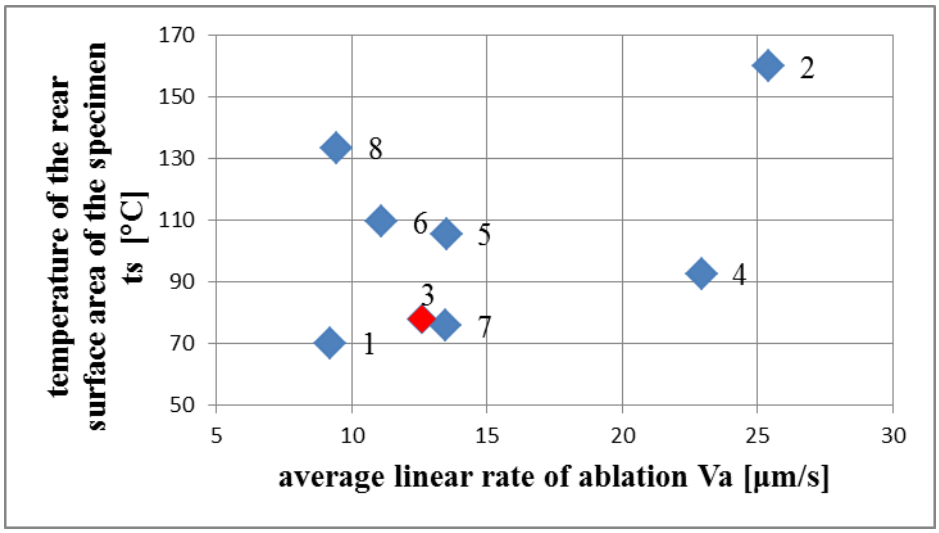


Fig. 7 Ablative thermo protective properties

5 Summary and conclusions

This paper presents the findings of a numerical analysis of a complex heat transfer for composite materials (polymer fibre composite of geometry illustrated in Fig 2). During the assessment of the analysis, it is essential to emphasise the introduction of several simplification assumptions:

- partial implementation of variables of thermophysical parametres, which has got an error due to an incomplete temperature range of the examination of parameters of particular composite materials,
- the pattern has got an error in the measurement of temperature, resulting from very high temperatures (1100° C) on the front surface, and the rear surface, after measuring the temperature by touching the surface with a thermoelement,
- lack of proper insulation of the pattern zone, where there are extremely high temperatures and the air flow is distant from the examined sample,
- assumption about a uniform temperature field, through gas combustion, on the surface of the heated sample,
- the composite material underwent an analysis as layered composition of several components, disregarding the permeation of its components.

The distance from the heated surface, where the material is burnt and the ablation surface is formed equals 1.8 mm, on the basis of a simulation, in comparison with the experimental value of 2.02 mm in the middle layer of the sample, and is a very good approximation. The discrepancy may also be linked to the actual change in the sample geometry while burning part of the material. In addition, the ablation volume is connected with the heating radius of the front surface (the radius of flame is lower than the radius of the sample), which can be observed in Fig. 6.
The presented findings of the simulation examination prove that it is essential to expand the range of the research by temperature measurements of different widths of the examined material, which will considerably allow to observe the thermal field distribution in a material. Thus, it will be easier to introduce more accurate values of thermophysical properties in a wider range of temperatures. In this way, it will be possible to analyse a numerical model, closer to reality, and more precisely determine the ablation surface of the examined materials.

Acknowledgements

"This publication has been co-financed with the European Union funds by the European Social Fund" Prioryty VIII ,Operation 8.2

References

[1] Kucharczyk W, Mazurkiewicz A, Żurowski W.: *Nowoczesne materiały konstrukcyjne wybrane zagadnienia*, wydanie II, Politechnika Radomska, Radom 2010.
[2] Dimitrienki Yu.I.:*Thermomechanical behaviour of composite materials and structures under high temperature: 1. Materials, Composites Part A 28a* (1997), pp453-461.
[3] Dimitrienki Yu.I.:*Thermomechanical behaviour of composite materials and structures under high temperature: 2. Structures, Composites Part A 28a* (1997), pp463-471.
[4] Dimitrienki Yu.I.: *Thermomechanical behaviour of composites under local intense heating by irradiation, Composites: Part A 31* (2000) pp 591–598.
[5] Przybyłek P., Opara T., Kucharczyk W.: *Możliwości zwiększenia odporności cieplnej rejestratorów lotniczych poprzez zastosowanie osłon z polimerowych kompozytów ablacyjnych*, Journal of Aeronautica Integra, 2/2011 (9), pp 50-56.
[6] Kucharczyk W.: *Kształtowanie ablacyjnych właściwości termoochronnych polimerowych kompozytów proszkowych*, Kompozyty 8: 3 (2008) 274-279.
[7] Ochelski S.: *Metody doświadczalne mechaniki kompozytów konstrukcyjnych*, Wydawnictwo Naukowo-Techniczne, Warszawa 2004.
[8] Zhang J, Chaisombat K, He S, Wang C.H.: *Hybrid composite laminates reinforced with glass/carbon woven fabrics for lightweight load bearing structures*, Materials and Design 36 (2012) 75–80
[9] Biswas S, Kindo S, Patnaik A.: *Effect of Fiber Length on Mechanical Behavior of Coir Fiber Reinforced Epoxy composites*, Fibers and Polymers 2011, Vol.12, No.1, pp 73-78.

[10] Kim S.Y., Baek S.J., Youn J.R.: *New hybrid method for simultaneous improvement of tensile and impact properties of carbon fiber reinforced composites*, Carbon 49 (2011) pp 5329 –5338

[11] Basavarajappa S., Ellangovan S.: *Dry sliding wear characteristics of glass–epoxy composite filled with silicon carbide and graphite particles*, Wear 296 (2012) pp 491–496. Garoushi S, Vallittu P.K.,

[12] Watts D.C., Lassila L.V.J.: *Effect of nanofiller fractions and temperature on polymerization shrinkage on glass fiber reinforced filling material*, dental materials 24 (2008) pp 606–610.

[13] Zhou S, Zhang Q, Wua C., Huang J.: *Effect of carbon fiber reinforcement on the mechanical and tribological properties of polyamide6/polyphenylene sulfide composites*, Materials and Design 44 (2013) pp 493–499

[14] Suresha B., Ramesh B.N., Subbaya K.M., Kumar B.N.R., Chandramohan G.: *Influence of graphite filler on two-body abrasive wear behaviour of carbon fabric reinforced epoxy composites*, Materials and Design 31 (2010) pp 1833–1841.

[15] Polański Z., *Planowanie doświadczeń w technice*, PWN, Warszawa, 1984.

[16] Bahramian A.R., Kokabi M.: Journal of Hazardous Materials 2009, 166, 445-454.

[17] Reis P.N.B., Ferreira J.A.M., Zhang Z.Y., Benameur T., Richardson M.O.W.: *Composites Part B* 2013, 46, 7-14.

[18] Srikanth I., Daniel A., Kumar S., Padmavathi N., Singh V.: *Scripta Materialia* 2010, 63, 200-203.

[19] Kucharczyk W.: *Przemysł Chemiczny* 2010, 89(12), 1673-1676.

[20] Kucharczyk W.: *Eksploatacja i Niezawodnosc–Maintenance and Reliability* 2012, 1, 12-18

[21] Shalin R.E. *Polymer matrix composites*, Chapman and Hall, 1995.

[22] Johnston A.A. *An integrated model of the development of process-induced deformation in autoclave processing of composite structures*, UBC, 1997

[23] Anderson A. *What is the specific heat capacity of carbon fibres* MadSci Network: Chemistry, 2000

[24] Polyaramid Polyparaphenylene Terephthalamide (Kevlar / Twaron) – Properties and Applications – Supplier Data by Goodfellow, 2003

[25] Chang K.K. *Aramid fibers* E.I. Du Pont de Nemours&Company,Inc

[26] Tables of *Thermal Properties of Gases*, NBS Circular 564, 1955

[27] *KNAUF Płyta zwykła K711B.pl/pol./PL/07.13*

[28] Schulz B., Huber N., Zasluft C.:High Temperatures-High Pressures, v34, 2002, p 431.

[29] Ohlhorst C. W., Vaughin W. L., Ronsone P. O., H-T. Tsou: *NASA Technical Meno* 4787 1997

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Parameter identification of transient models with multiple time scales

Guillaume Puel*, **Denis Aubry**

Laboratoire MSSMat
Ecole Centrale Paris / CNRS UMR8579, Grande Voie des Vignes, 92290 Chatenay-Malabry, France
e-mail: guillaume.puel@ecp.fr

Key words: parameter identification, adjoint state, time multiscale, homogenization

Abstract

The aim here is to study time-multiscale models and their associated parameter identification. When it is possible to consider several well-separated time scales, and when the applied loading consists of periodic functions with respect to some of these time scales, a periodic time homogenization scheme, similar to what exists in space homogenization, can be used to derive a homogenized model. The parameter identification of this latter then requires some specific formulations, which are described here in an academic example.

1 Introduction

As far as the numerical calculation of a time-dependent model is concerned, the question of the computational cost can be of utmost relevance, especially when the considered model deals with fast phenomena, which require the use of very small time steps, when compared with the length of the time interval. In order to drastically reduce the computational cost, a periodic time homogenization method, such as [1], can be used when well-separated time scales can be defined, and when the components of the applied loading are periodic with respect to some of these time scales. The resulting homogenized model is then cheaper computationally, for it can be solved using time steps related to the slower time scale only, whereas the other, faster time scales are taken into account in an average way in the homogenization scheme.

In order them to give accurate predictions, such time-homogenized models have to be compared with experimental data. The key point is to define an identification strategy able to deal with such models in such a way that the process remains cheap and efficient. The aim of this paper is to analyze on a specific academic example how the different choices can be made throughout the whole identification process, and what is their impact on the identification results.

2 Periodic time homogenization method

Periodic time homogenization, as it was initially proposed in [1], can be seen as a transposition to time of the classical periodic space homogenization methods, such as the techniques described in [2]. It consists in separating two time scales, a slow one t and a fast one $\tau = t/\xi$, by introducing, in the equations of the reference problem, an asymptotic expansion, in terms of the ratio ξ, for each time-dependent variable:

$$\alpha(t,\tau) = \alpha_0(t,\tau) + \xi\alpha_1(t,\tau) + \xi^2\alpha_2(t,\tau) + O(\xi^3) \tag{1}$$

When this ratio ξ is very small, it is possible to consider that the two time scales are independent, and that any derivative with respect to time has to use the partial derivatives with respect to the two time scales:

$$\mathrm{d}_t\alpha = \partial_t\alpha + \frac{1}{\xi}\partial_\tau\alpha \tag{2}$$

where d_t, ∂_t, et ∂_τ stand for the total time derivative, the partial derivative with respect to the slow time and the partial derivative with respect to the fast time respectively. Moreover, if the applied loading has a component, which is periodic with respect to the fast time τ, it is possible to assume that each variable is quasiperiodic, meaning that it is periodic with respect to τ.

Using the asymptotic expansion (1) for each quantity in the reference equations, and balancing the different orders of ξ, the time homogenized equations are determined by averaging over a fast period the different quantities:

$$< \alpha > = \frac{\xi}{T_f}\int_0^{T_f/\xi} \alpha(t,\tau)\,\mathrm{d}\tau \tag{3}$$

where T_f is the period associated with the fast periodic component of the loading.The residual associated with this average is then denoted as $\alpha^* = \alpha - < \alpha >$, and depends on both time scales t and τ a priori. Eventually, these homogenized equations are solved relatively to the slow time scale only, by introducing the averaged influence of the fast cycles, allowing to solve for all the zeroth-order time-homogenized variables.

Recently, this method has been used for different models in several frameworks, namely for simulations of structures withstanding fatigue loads with two periodic components:

- material fatigue with a viscoplastic law defining two hardening variables in [3] ;
- material fatigue with an isotropic damage law in [4] ;
- extension of the method to the dynamic framework in [5] ;
- extension of the method to three different time scales in [6].

Previously, investigations from other authors on this time homogenization method had also been proposed, such as the behavior of viscoelastic materials in [7], or the vibration of preloaded beams in [8].

3 Parametric identification

In this section, a general identification strategy is proposed: from this scheme, a specific strategy, relevant for the case of time-homogenized models, will be described in Section 4. The forward problem is considered as an implicit formulation with a functional $\mathcal{F}$ over a time interval $[0, T]$:

$$\begin{aligned} &\mathcal{F}\left(\mathbf{u}(t), \mathrm{d}_t\mathbf{u}(t), \mathbf{p}, t\right) = \mathbf{0} \quad \forall t \in [0; T] \\ &\mathbf{u}(0) = \mathbf{U}_0 \end{aligned} \tag{4}$$

where d_t is the time derivative. Whereas $\mathbf{u}$ is the state vector of size N, composed of all the time-dependent degrees of freedom (DOFs) describing the studied problem, $\mathbf{p}$ stands for the vector containing the P scalar parameters associated with the differential equation (4).

The identification problem consists in finding the parameter vector $\mathbf{p}_{opt}$ such that the solution $\mathbf{u}(t; \mathbf{p}_{opt})$ of (4) obtained with the parameters $\mathbf{p}_{opt}$ is as close to the available experimental data as possible. These latter are indeed compared with the corresponding quantities $\mathbf{A}\mathbf{u}(t; \mathbf{p})$, where $\mathbf{A}$ is a projection operator allowing to select, for each quantity, the closest DOF to the experimental measurement point. In order to use consistent notations, the corresponding experimental quantity is denoted $\mathbf{A}\mathbf{u}_{exp}(t)$; however, it does not mean that such a vector $\mathbf{u}_{exp}(t)$ actually exists.

The following misfit function is then introduced: it consists of a norm measuring the discrepancy between the quantities predicted with the forward model (4) and experimental data:

$$\mathcal{J}(\mathbf{p}) \quad = \quad \frac{1}{2}\int_0^T |\mathbf{A}(\mathbf{u}(t; \mathbf{p}) - \mathbf{u}_{exp}(t))|^2 \, \mathrm{d}t + \frac{1}{2}|\mathbf{R}(\mathbf{p} - \mathbf{p}_0)|^2 \tag{5}$$

where $\mathbf{u}(t; \mathbf{p})$ satisfies Equation (4). The L^2-norm proposed here is completed with a Tikhonov regularization term, allowing to deal with the ill-posedness of the identification problem, by bounding the magnitude of the parameter vector $\mathbf{p}$ to be identified: this regularization term uses a vector $\mathbf{p}_0$ containing nominal values corresponding to a priori experience, and a diagonal weighing matrix $\mathbf{R}$. Eventually, the solution of the identification problem can be sought as the parameter vector minimizing the misfit function $\mathcal{J}(\mathbf{p})$:

$$\mathbf{p}_{opt} = \arg\min_{\mathbf{p}} \mathcal{J}(\mathbf{p}) \tag{6}$$

The determination of this minimum is achieved using gradient-based minimization methods, therefore the question of avoiding local minima by means of an appropriate regularization process should be carefully addressed. In some cases, rather than using the classical Tikhonov regularization term, the a priori experience can be introduced in some specific ways, as in [9]. Similarly, the fact of using a homogenized model in the parameter identification process can introduce a regularizing effect, just as explained in [10].

To estimate the gradient of the misfit function, we solve here an adjoint state problem. A typical example in mechanical engineering is given in [11], where the parameters of an elastoplastic material law are identified with indentation tests. In the strategy proposed here, the generic form of the adjoint state problem is as follows:

$$\begin{aligned} &\nabla_{\mathbf{u}}\mathcal{F}^{\mathrm{T}}\mathbf{z} - \mathrm{d}_t\left(\nabla_{\mathrm{d}_t\mathbf{u}}\mathcal{F}^{\mathrm{T}}\mathbf{z}\right) = \mathbf{A}^{\mathrm{T}}\mathbf{A}(\mathbf{u} - \mathbf{u}_{exp}) \quad \forall t \in [0, T] \\ &\left(\nabla_{\mathrm{d}_t\mathbf{u}}\mathcal{F}^{\mathrm{T}}\mathbf{z}\right)_{|t=T} = \mathbf{0} \end{aligned} \tag{7}$$

where $\nabla_{\mathbf{u}}\mathcal{F}$ and $\nabla_{\mathrm{d}_t\mathbf{u}}\mathcal{F}$ stand for the directional derivatives of $\mathcal{F}$ with respect to $\mathbf{u}$ and $\mathrm{d}_t\mathbf{u}$ respectively. The adjoint state problem is then a time-backward differential equation with a final condition, and where the first-order sensitivities of the forward problem are concerned.

Once the adjoint state problem (7) is solved, it can be shown that the misfit function gradient with respect to the parameter vector $\mathbf{p}$ can be expressed as:

$$\nabla_{\mathbf{p}}\mathcal{J}(\mathbf{p}) = \mathbf{R}^{\mathrm{T}}\mathbf{R}(\mathbf{p}-\mathbf{p}_0) - \int_0^T \nabla_{\mathbf{p}}\mathcal{F}^{\mathrm{T}}\mathbf{z}(t)\,\mathrm{d}t \tag{8}$$

This specific way of estimating the misfit function gradient can be compared with a classical finite difference formula, such as the central finite difference scheme: in this latter case, when the parameter vector is of size P, the gradient calculation is obtained by evaluating the misfit function in $2P$ additional 'points', each couple of points corresponding to two symmetrical perturbations of the misfit function associated with each parameter in the vector $\mathbf{p}$. The resulting computational cost for each gradient evaluation consists of $2P$ solutions of the forward problem (4) and $2P$ time integral evaluations. By contrast, when the adjoint state solution is used, only two differential equation solutions are required: for the forward problem (4) and for the adjoint problem (7). The associated computational cost for each gradient evaluation is then 2 differential equation solutions, and P time integral evaluations. The resulting gain is as high as the number of parameters to be identified. Moreover, it is easier to control the accuracy of the gradient estimate with the adjoint state method than with finite difference formulas, for which the choice of the discretization steps has a strong influence on the final estimate.

4 Parameter identification of a model with two time scales

Here an academic example is proposed to discuss the different steps of the parameter identification problem associated with a time-homogenized model. This latter consists of a straight bar of length L withstanding at one end a normal force with two periodic components. The measured displacement at this end is then used to determine the parameter values of the material elastic viscoplastic law.

4.1 Reference problem

The quasistatic equilibrium of the bar is a scalar equation, with the normal stress $\sigma(x,t)$ defined at each point $x \in [0, L]$:

$$\partial_x \sigma = 0 \tag{9}$$

where ∂_x is the partial space derivative. It is assumed that no load is applied along the bar. Whereas this latter is clamped at $x = 0$, a surface force is applied at $x = L$. It follows that:

$$\sigma(x,t) = f_{\mathrm{s}}(t) \quad \forall x \in [0;L] \quad \forall t \in [0,T] \tag{10}$$

The constitutive relation links the normal stress to the longitudinal displacement $u(x,t)$ along the bar:

$$\sigma = E\left(\partial_x u - \varepsilon^{\mathrm{p}}\right) \tag{11}$$

where E is the Young's modulus, and $\varepsilon^{\mathrm{p}}(x,t)$ stands for the longitudinal plastic strain. This latter verifies a Norton's viscoplastic evolution law:

$$\mathrm{d}_t\varepsilon^{\mathrm{p}} = \left(\frac{|\sigma|}{K}\right)^n \tag{12}$$

with $\varepsilon^{\mathrm{p}}(x,0) = 0 \; \forall x \in [0,L]$. K and n are two constant parameters to be identified.

Since the stress σ is homogeneous, the plastic strain ε^{p}, as well as the total strain $\partial_x u$, are constant along the bar as well. The forward problem then consists in solving the following differential equation:

$$\mathrm{d}_t \partial_x u = \frac{\mathrm{d}_t \sigma}{E} + \mathrm{d}_t \varepsilon^{\mathrm{p}} \tag{13}$$

Eventually, the obtained displacement at $x = L$, which is compared with experimental data, has to verify:

$$\begin{aligned} &\frac{\mathrm{d}_t u(L,t)}{L} = \frac{\mathrm{d}_t f_{\mathrm{s}}(t)}{E} + \left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)^n \quad \forall t \in [0;T] \\ &u(L,0) = 0 \end{aligned} \tag{14}$$

4.2 Time-homogenized problem

The loading applied to the bar is typical to what occurs in Combined Cycle Fatigue (CCF), where the two considered frequencies are very different one from the other ($F = 0.05\mathrm{Hz}$ and $F/\xi = 500\mathrm{Hz}$ here):

$$f_{\mathrm{s}}(t,\tau) = < f_{\mathrm{s}} > (t) + f_{\mathrm{s}}^*(\tau) \tag{15}$$

where $< f_{\mathrm{s}} > (t) = f_0 + f_1 \cos 2\pi F t$ and $f_{\mathrm{s}}^*(\tau) = f_2 \cos 2\pi F \tau$, with f_0, f_1 and f_2 some fixed values such that $f_{\mathrm{s}}(0,0) = 0$.

The time-homogenized equations come from the zeroth-order expression of the forward problem (14), where the fast-time average (3) has been previously applied:

$$\begin{aligned} &\frac{\partial_t < u_0 > (L,t)}{L} = \frac{\partial_t < f_{\mathrm{s}} > (t)}{E} + \left\langle \left(\frac{|f_{\mathrm{s}}(t,\tau)|}{K}\right)^n \right\rangle \quad \forall t \in [0,T] \\ &< u_0 > (L,0) = \frac{< f_{\mathrm{s}} > (0) L}{E} \end{aligned} \tag{16}$$

The term associated with the evolution equation is evaluated by means of a numerical integration formula, such as the trapezoidal rule with $M+1$ points:

$$\left\langle \left(\frac{|f_{\mathrm{s}}(t,\tau)|}{K}\right)^n \right\rangle = \sum_{k=0}^{M} w_k \left(\frac{| < f_{\mathrm{s}} > (t) + f_{\mathrm{s}}^*(\frac{k}{MF})|}{K}\right)^n \tag{17}$$

where $w_0 = 1/2M, w_{0<k<M} = 1/M, w_M = 1/2M$.

Eventually, the displacement at $x = L$ is derived by adding to the zeroth-order time-homogenized displacement the associated residual $u_0^*(L,\tau)$, which is the quasi-static solution of an elastic tension problem, when the fast time variable is replaced by its actual expression $\tau = t/\xi$:

$$u(L,t) = < u_0 > (L,t) + \frac{f_{\mathrm{s}}^*(t/\xi) L}{E} \quad \forall t \in [0,T] \tag{18}$$

When compared with the reference problem, whose solution requires time steps of 10^{-4}s (if one considers that 20 times steps are needed to discretize well the fast period component), time steps of 0.01s allow to get the same estimations with the homogenized problem. It then becomes possible to deal with such time-dependent problems, for very long time intervals, whereas the solution could not be calculated using the reference problem. More details are given in [6] for this academic example, for a more complex evolution law (viscoplastic behavior with two hardenings after [12]).

4.3 Formulation of the identification problem

The associated identification problem consists in determining the material parameters E, K and n using measurements obtained during very long time intervals: in the CCF case indeed, a significant evolution of the structure's response is visible only if a very high number of loading cycles is applied to the structure. Here is studied how the choice of the misfit function can influence the identification results.

4.3.1 Misfit function using time-homogenized quantities

A first possible choice consists in modifying the misfit function (5) by introducing time-homogenized quantities. It is considered here that the displacement $u_{exp}(t)$ at the end at $x = L$ can be measured:

$$\begin{aligned}\mathcal{J}^0(E,K,n) = \frac{1}{2}\int_0^T | < u_0 > (L,t;E,K,n) - < u_{exp} > (t)|^2 \,\mathrm{d}t \\ + \frac{\alpha_E}{2}|E-E_0|^2 + \frac{\alpha_K}{2}|K-K_0|^2 + \frac{\alpha_n}{2}|n-n_0|^2\end{aligned} \tag{19}$$

Whereas $< u_0 > (L,t;E,K,n)$ is the solution of the zeroth-order time-homogenized forward problem (16), $< u_{exp} > (t)$ stands for the corresponding experimental quantity, which is obtained by fast-averaging the experimental data for each time step t_k of the time-homogenized displacement:

$$< u_{exp} > (t_k) \;= \frac{F}{\xi}\int_{t_k}^{t_k+\frac{\xi}{F}} u_{exp}(t)\,\mathrm{d}t \tag{20}$$

The gradient of this misfit function is evaluated as in Section 3, using the solution of an adjoint state problem. This latter is as follows:

$$\begin{aligned}&\mathrm{d}_t z^0(t) = L\left(< u_0 > (L,t;E,K,n) - < u_{exp} > (t)\right) \quad \forall t \in [0,T] \\ &z^0(T) = 0\end{aligned} \tag{21}$$

which is, as previously, a time-backward differential equation with a final condition equal to zero. This equation can be solved using the slow time steps t_k only, which allows to derive the solution in a way as efficient as for the time-homogenized forward solution (16). The misfit function gradient then consists of the three following partial derivatives:

$$\begin{aligned}\nabla_E\mathcal{J}^0(E,K,n) &= \int_0^T \frac{L}{E^2} < |\mathrm{d}_t f_{\mathrm{s}}(t)| > z^0(t)\,\mathrm{d}t + \frac{L}{E^2} < f_{\mathrm{s}} > (0) z^0(0) + \alpha_E(E-E_0) \\ \nabla_K\mathcal{J}^0(E,K,n) &= \int_0^T \frac{nL}{K}\left\langle\left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)^n\right\rangle z^0(t)\,\mathrm{d}t + \alpha_K(K-K_0) \\ \nabla_n\mathcal{J}^0(E,K,n) &= -\int_0^T L\left\langle\ln\left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)\left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)^n\right\rangle z^0(t)\,\mathrm{d}t + \alpha_n(n-n_0)\end{aligned} \tag{22}$$

These preceding relations (21) and (22) are actually the zeroth-order time-homogenized estimates of the adjoint state solution and the misfit function gradient of the identification problem associated with the reference problem (14). This is a classical result in periodic space homogenization, as described for example in [13].

4.3.2 Misfit function using 'exact' quantities

An alternative consists in choosing a misfit function where the 'exact' experimental displacement $u_{exp}(t)$ is compared with the corresponding displacement (18) :

$$\mathcal{J}^{0*}(E,K,n) = \frac{1}{2}\int_0^T | < u_0 > (L,t;E,K,n) + \frac{f_{\mathrm{s}}^*(t/\xi)L}{E} - u_{exp}(t)|^2 \,\mathrm{d}t \\ + \frac{\alpha_E}{2}|E-E_0|^2 + \frac{\alpha_K}{2}|K-K_0|^2 + \frac{\alpha_n}{2}|n-n_0|^2 \tag{23}$$

The misfit function's gradient is evaluated by means of the adjoint state z^{0*} verifying the following time-backward equation:

$$\begin{aligned} &\mathrm{d}_t z^{0*}(t) = L\left(< u_0 > (L,t;E,K,n) + \frac{f_{\mathrm{s}}^*(t/\xi)L}{E} - u_{exp}(t)\right) \quad \forall t \in [0;T] \\ &z^{0*}(T) = 0 \end{aligned} \tag{24}$$

Because of the fast term involving T/ξ, it is not relevant to directly solve this differential equation, unless one loses the drastic reduction in the computational cost offered by the time homogenization method. On the contrary, it is possible to homogenize Equation (24) with respect to time and solving for a zeroth-order adjoint state z_0^{0*} verifying:

$$\begin{aligned} &\mathrm{d}_t z_0^{0*}(t) = L\left(< u_0 > (L,t;E,K,n) - < u_{exp} > (t)\right) \quad \forall t \in [0;T] \\ &z_0^{0*}(T) = 0 \end{aligned} \tag{25}$$

which is solved with respect to the slow time scale only. This differential equation is actually identical to Equation (21).

The zeroth-order time-homogenized adjoint state z_0^{0*} is then used to evaluate the misfit function's gradient as follows:

$$\begin{aligned} \nabla_E \mathcal{J}^{0*}(E,K,n) &= \int_0^T \frac{L}{E^2} < |\mathrm{d}_t f_{\mathrm{s}}(t)| > z_0^{0*}(t)\,\mathrm{d}t + \frac{L}{E^2} < f_{\mathrm{s}} > (0) z_0^{0*}(0) + \alpha_E(E-E_0) \\ &\quad - \int_0^T \frac{L}{E^2} f_{\mathrm{s}}^*(t/\xi)\left(< u_0 > (L,t;E,K,n) + \frac{f_{\mathrm{s}}^*(t/\xi)L}{E} - u_{exp}(t)\right)\mathrm{d}t \\ \nabla_K \mathcal{J}^{0*}(E,K,n) &= \int_0^T \frac{nL}{K}\left\langle \left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)^n \right\rangle z_0^{0*}(t)\,\mathrm{d}t + \alpha_K(K-K_0) \\ \nabla_n \mathcal{J}^{0*}(E,K,n) &= -\int_0^T L\left\langle \ln\left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)\left(\frac{|f_{\mathrm{s}}(t)|}{K}\right)^n \right\rangle z_0^{0*}(t)\,\mathrm{d}t + \alpha_n(n-n_0) \end{aligned} \tag{26}$$

It can be seen that these partial derivatives are identical to those estimated in Equation (22), except for an additional term, in the expression of $\nabla_E \mathcal{J}^{0*}$, which is related to the fast loading component $f_{\mathrm{s}}^*(t/\xi)$. As a result, the main difference between the two misfit functions is whether one takes into account this fast component or not.

4.3.3 Identification results

In order to evaluate how the two misfit functions perform, synthetic data $u_{exp}(t)$ are created by solving the reference forward problem (14), using $E_{exp} = 200$GPa, $K_{exp} = 100$MPa, $n_{exp} = 10$ as parameter values. As previously, the bar withstands a CCF loading with $F = 0.05$Hz and $F/\xi = 500$Hz.

$E_0 = 220$GPa, $K_0 = 110$MPa, $n_0 = 11$ are chosen as initial parameter values for the identification process, which is based on an interior-reflective Newton method [14] in order to minimize the two misfit functions $\mathcal{J}^0$ and $\mathcal{J}^{0*}$, for which no regularization term is added. Subsequent results are listed in Table 1: the parameter values identified for the two misfit functions are very similar, even if the misfit function $\mathcal{J}^{0*}$ using 'exact' quantities seems to provide slightly better results. Figure 1 shows the comparison between the two identified models and the synthetic reference, more precisely the variations of the longitudinal plastic strain, which is not directly observable. In both cases, the computational cost associated with the identification process is significantly reduced when compared with what is obtained when the inverse problem related to the reference problem (14) is considered.

Table 1: Identified parameter values.

Misfit function	Number of iterations	E_{id}	K_{id}	n_{id}
$\mathcal{J}^0$	18	200.0GPa	97.6MPa	10.2
$\mathcal{J}^{0*}$	21	200.0GPa	97.9MPa	10.1

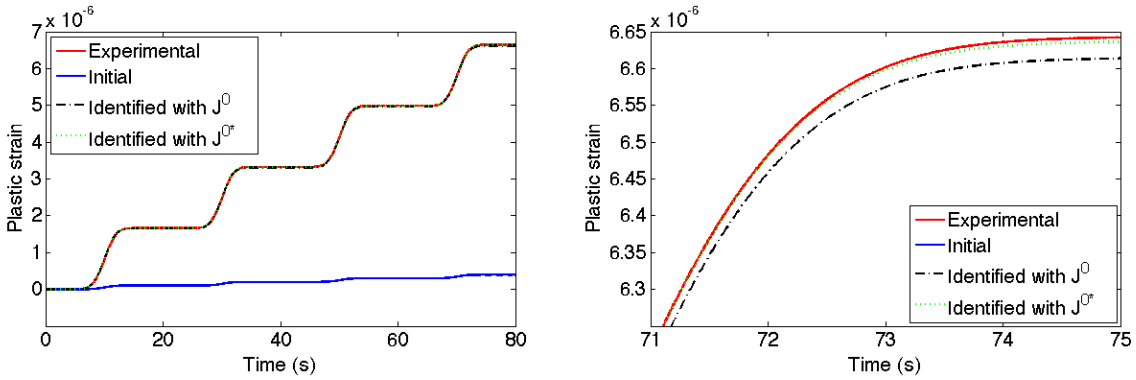


Figure 1: Longitudinal plastic strain: synthetic data (red), initial parameter values (blue), identified parameter values with $\mathcal{J}^0$ (black), identified parameter values with $\mathcal{J}^{0*}$ (green).

5 Summary

Here we have proposed a first study of a parameter identification process for time-homogenized models. An adaptation of the approach with an adjoint state formulation can be used: this leads to the determination of the time-homogenized counterpart of the adjoint solution associated with the reference identification problem. The academical example studied here showed the relevance of the strategy and its reduced computational cost. These results can be viewed as a first step before dealing with more complex cases of study. The extension to models with three different time scales, such as those described in [15], should be straightforward as well.

References

[1] T. Guennouni, D. Aubry, *Réponse homogénéisée en temps de structures sous chargements cycliques*, Comptes rendus de l'Académie des sciences. Série II, 303(20), (1986), 1765–1768, (in French).

[2] E. Sanchez-Palencia, *Non-homogeneous media and vibration theory*, Springer (1980).

[3] D. Aubry, G. Puel, *Ccf modelling with use of a two-timescale homogenization model*, in *Proceedings of the International Fatigue Conference - Fatigue 2010*, vol. Procedia Engineering 2 (2010), pp. 787–796.

[4] A. Devulder, D. Aubry, G. Puel, *Two-time scale fatigue modelling: application to damage*, Computational Mechanics, 45(6), (2010), 637–646.

[5] D. Aubry, G. Puel, *Two-timescale homogenization method for the modeling of material fatigue*, in *Proceedings of the 9th World Congress on Computational Mechanics - WCCM9*, vol. IOP Conference Series: Material Science and Engineering 10 (2010), p. 012113.

[6] G. Puel, D. Aubry, *Material fatigue simulation using a periodic time homogenization method*, European Journal of Computational Mechanics, 21(3–6), (2012), 312–324.

[7] Q. Yu, J. Fish, *Temporal homogenization of viscoelastic and viscoplastic solids subjected to locally periodic loading*, Computational Mechanics, 29(3), (2002), 199–211.

[8] E. V. Shishkina, I. I. Blekhman, M. P. Cartmell, S. N. Gavrilov, *Application of the method of direct separation of motions to the parametric stabilization of an elastic wire*, Nonlinear Dynamics, 54(4), (2008), 313–331.

[9] G. Puel, B. Bourgeteau, D. Aubry, *Parameter identification of transient nonlinear models for the multibody simulation of a vehicle chassis*, Mechanical Systems and Signal Processing, 36(2), (2013), 354–369.

[10] C.-B. Chung, C. Kravaris, *Identification of spatially discontinuous parameters in second-order parabolic systems by piecewise regularisation*, Inverse Problems, 4(4), (1988), 973–994.

[11] A. Constantinescu, N. Tardieu, *On the identification of elastoviscoplastic constitutive laws from indentation tests*, Inverse Problems in Engineering, 9(1), (2001), 19–44.

[12] J. Lemaitre, J.-L. Chaboche, *Mechanics of solid materials*, Cambridge University Press (1990).

[13] S. Kesavan, J. Saint Jean Paulin, *Homogenization of an optimal control problem*, SIAM Journal on Control and Optimization, 35(5), (1997), 1557–1573.

[14] T. F. Coleman, Y. Li, *On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds*, Mathematical Programming, 67(1-3), (1994), 189–224.

[15] G. Puel, D. Aubry, *Efficient fatigue simulation using periodic homogenization with multiple time scales*, International Journal for Multiscale Computational Engineering, (to appear).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Identification of multi-inclusion Statistically Similar Representative Volume Element

L. Rauch*, D. Szeliga, K. Bzowski, D. Bachniak

Department of Applied Computer Science and Modelling
AGH University of Science and Technology, Mickiewicza 30, 30-059 Krakow, Poland
e-mail: lrauch@agh.edu.pl

Key words: inverse analysis, statistically similar representative volume element, DP steels

Abstract
The paper presents

1 Introduction

Numerical simulations of metallurgical industrial processes require computational costly multiscale procedures [1]. Usage of such methods in the case of computer aided technology design is very time consuming due to applied multi iterative optimization approaches. Exertions, leading to improvement of numerical efficiency, are focused mainly on two areas of development i.e. parallelization of numerical procedures or simplification of virtual material representation in micro scale [2]. The approach presented in this paper covers both these subjects, because it applies idea of Statistically Similar Representative Volume Element (SSRVE) to simplify the computational domain, as well as it uses High Performance Computing hardware architectures to increase the computing performance.

SSRVE is an extension of RVE idea, which is very well known and widely applied in multiscale simulations [3]. In the micro-macro modelling approach a RVE representing the underlying microstructure is attached at each Gauss point of the macroscopic solution. The constitutive law describing material behavior in the macro scale is obtained by averaging the first Piola-Kirchoff stresses with respect to the RVE. The theoretical basis of the micro-macro modelling is well described in the scientific literature [4]. Considering micro-heterogeneous materials, the continuum mechanical properties at the macro scale are characterized by the morphology and by the properties of the particular constituents in the micro scale. In this work two phase microstructures composed of soft ferrite and hard martensite are considered. In the analysis we concentrate on the measures characterizing the hard martensite islands only. The material models of the individual constituents are assumed to be known. The description of the microstructure is the based on statistical consideration [5]. An usual RVE is determined by the smallest possible sub domain, which is still able to represent the macroscopic behavior of the material. Although these RVEs are the smallest possible by definition, they still can be too complex for the efficient calculations. Therefore, the construction of statistically similar RVEs, which

are characterized by a lower complexity than the smallest possible substructure, was proposed in [4]. The basic idea is to replace a RVE with an arbitrary complex inclusion morphology by a periodic one composed of optimal unit inclusions (Fig. 1), which is characterized by similar material properties as RVE.

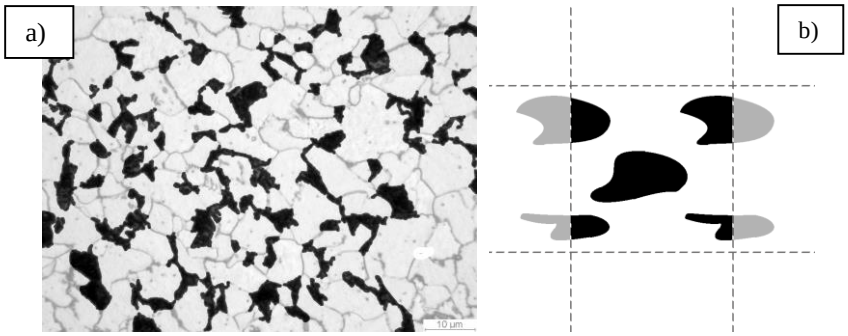


Figure 1: Original micrograph (can be binarized to obtain RVE) of DP steel (a) and periodical three element SSRVE (b).

The number of inclusions in SSRVE is usually not higher than three. The most of the papers, related to determination of SSRVE, are focused on theoretical considerations or creation of single element SSRVE. In this work, identification procedure, focused on creation of multi-inclusion SSRVE, is proposed. The second section contains description of SSRVE methodology including all the steps required to obtain final multi-inclusion SSRVE. The subsequent sections (three and four) present sensitivity analysis and optimization procedures. Afterwards, the results in 2D and 3D are shown and the paper is concluded with perspectives of the future work.

2 Procedure of SSRVE creation

The scheme of the procedure dedicated to SSRVE creation consists of four main steps dedicated to preparation of microstructure images characterizing analyzed steel, processing of images to determine features describing material, and performance of sensitivity analysis and optimization procedure. These steps are presented in Fig. 2.

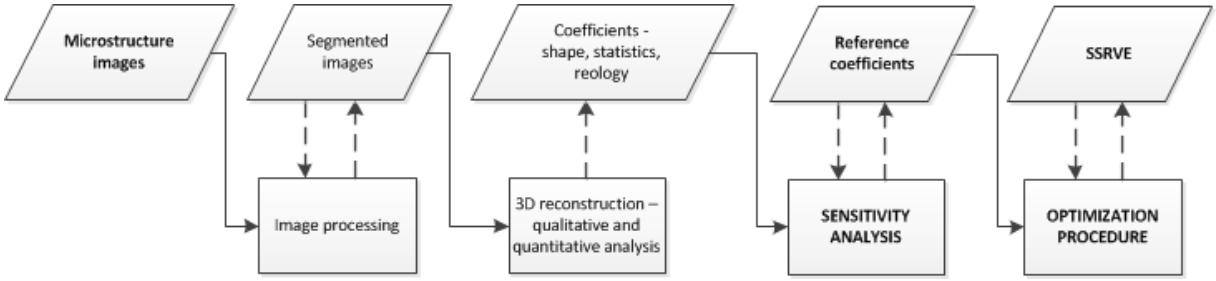


Figure 2: Procedure of SSRVE creation.

The procedure of SSRVE creation starts with analysis of original micrographs, which aims at creation of binary (segmented) images with separated martensitic inclusions. The algorithms for processing of various micrographs are presented in details in [6]. Then, in case of 2D SSRVE, the shape coefficients of inclusions in original images are estimated directly from the segmented pictures. In case of 3D procedure, the reconstruction of three dimensional microstructure on the basis of 2D images has

to be performed [7]. Afterwards, the shape coefficients of 3D inclusions are estimated. Ohser and Muecklich [8] proposed four basic parameters for statistical shape description. Few more parameters, which are used in the image analysis, were added and described by Authors in [9]. In consequence, the following parameters describing shape of inclusions were considered: volume fraction, area/volume, roundness, ellipsoid fit, contour to center ratio, border index, mean curvature, total curvature, Malinowska coefficient, Blair-Bliss coefficient, Danielsson coefficient and Haralick coefficient. Not all of these parameters can be adapted from 2D to 3D. Thus, some of the coefficients are used only for 2D SSRVE.

Additionally, statistical and rheological coefficients are calculated to obtain full set of reference coefficients, which describe all aspects of material properties and which are used further in optimization procedure. The statistical measures, used for microstructure characterization in the SSRVE for micro-macro modelling of DP steels, are described in details in [9]. Brandts et al. [10] introduced the higher order statistical measures for microstructures: *n*-point probability functions, spectral density and lineal-path function. The latter parameter is crucial in description of anisotropic microstructures, while it describes the probability that a complete line segment $\mathbf{a} = \overrightarrow{\mathbf{a}_1\mathbf{a}_2}$ is located in in specific direction in the same phase, where $a_1 = \{x_1,y_1\}$ and $a_2 = \{x_2,y_2\}$ are coordinates of the ends of the line segment. Lu and Torquato [11] gave general mathematical description of this measure for multi-phase anisotropic materials. Simplified approach, which is applicable to DP microstructures, can be defined by the following equation:

$$\zeta_{LP}(m,k) = \frac{1}{N_x N_y} \sum_{p=1}^{N_x} \sum_{q=1}^{N_y} \chi\left(\overrightarrow{\mathbf{a}_1\mathbf{a}_2}\right) \tag{1}$$

where N_x, N_y are the dimensions of micrograph in pixels, and $\chi^{(i)}\left(\overrightarrow{\mathbf{a}_1\mathbf{a}_2}\right)$ is modified indicator function defined for phase $D^{(i)}$ as:

$$\chi^{(i)}\left(\overrightarrow{\mathbf{a}_1\mathbf{a}_2}\right) = \begin{cases} 1 \text{ if } \overrightarrow{\mathbf{a}_1\mathbf{a}_2} \in D^{(i)} \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

The estimated shape coefficients as well as statistical measured are the main elements of optimization function aiming at SSRVE creation. Optimization procedure is based on approach proposed in [9]. Originally, a method for the construction of simple periodic structures for the special case of randomly distributed circular inclusions with constant equal diameters was proposed by Povirk [12]. In that work the positions of circular inclusions with given diameter were found by minimizing the objective function, which was defined as a square root error between spectral density of the periodic RVE and non-periodic real microstructure. In our work this function was adapted to the following form:

$$\Phi = \sqrt{\sum_{i=1}^{n} \left[w_i \left(\frac{\zeta_i - \zeta_{iSSRVE}}{\zeta_i} \right) \right]} \tag{3}$$

where: w_i – normalized weights, n – number of coefficients, ζ_i – i^{th} reference coefficient obtained from original microstructure, ζ_{iSSRVE} – i^{th} coefficient obtained from SSRVE. The coefficients includes all mentioned parameters describing shape, rheology and statistics.

The current implementation of optimization procedure is based on genetic algorithm (GA), where chromosome is composed of *m* elements representing coordinates of control points determining SSRVE shape. These points are connected with spline functions forming smooth shape of SSRVE inclusion. Calculations of the objective function are performed iteratively for each proposition of new SSRVE shape. The optimization loop is preceded by sensitivity analysis, which allows to determine the most influential parameters of the optimization or to determine the weights used in the objective function (3) . The sensitivity analysis is described in the next section, which is followed by detailed presentation of optimization procedure.

3 Sensitivity analysis

Sensitivity Analysis (SA) is the field of knowledge investigating the model behavior for various input data and model parameters [13]. It determines how the variations of input data and parameters are distributed on the variations of model outputs and influence them. In this work global sensitivity methods were applied, which calculate one (global) value expressing the sensitivity of a parameter for the whole parameter domain. These methods are derived from statistics and the probability theory. Global SA requires a definition of the following terms:

- expression which characterizes the measure of model output/outputs (it should be a scalar value),
- definition of the variation interval for each input parameter,
- selection of the points in the parameters domain (design of experiment techniques are applied),
- sensitivity measure – the sensitivities are estimated based on the model outputs measure variations, caused by changes in the model parameters.

Two algorithms were selected for the purposes of this work: Morris Design which is a screening method and Sobol' algorithm ,based on analysis of variance.

3.1 Morris Design algorithm

The term screening design characterizes the method of the input parameters domain processing. The methods of this group calculate the parameter sensitivities as the global indices and they search systematically the whole parameters domain, thus, they are called screening methods. The main idea of these algorithms is to investigate the model parameters, which have the biggest influence on variability of the model output, and to keep computing costs as low as possible. The methods deal with the question, which model parameters are really important. Due to the assumption of low computing costs, these procedures estimate the importance of the input parameters qualitatively, not quantitatively. The One-At-a-Time (OAT) approach, originally developed by Morris [14], was selected. This technique investigates the impact of the variation of each parameter in turn. The OAT design is called the global sensitivity analysis, because the algorithm explores the entire space over which the parameters vary. In the algorithm, the term of parameter main effect is introduced and it is determined by computing a number of local measures at different points in the input space and next estimated by mean value and standard deviation. The key definitions and steps of Morris design are presented below.

Assumptions and definitions. Let $\mathbf{x}$ be an *n*-dimensional vector of model parameters x_i. The primary assumption of the algorithm is that all x_i components are defined on $[0,1]$ interval. In most practical problems x_i components are of various physical units and the parameters have to be rescaled to $[0,1]$. Linear or logarithmic transformation can be applied. The conversion is necessary to compare the results obtained for various parameters. It is feasible only if estimated elementary effects are expressed with the same units for all parameters. Let the components x_i, $i=1,\ldots,n$, accept k values in the set $\{0,1/(k-1),2/(k-1),\ldots,1\}$. Then the parameters domain $\Omega \subseteq \mathbb{R}^n$ forms an *n*-dimensional *k*-level grid. The elementary effect x_i of the i^{th} parameter at a given point $\mathbf{x}$ calculated for *y* model output is defined as:

$$\xi_i(\mathbf{x}) := \frac{\tilde{y}(x_1,\ldots,x_{i-1},x_i+\tilde{\Delta}_i,x_{i+1},\ldots,x_n)-\tilde{y}(\mathbf{x})}{\Delta} \tag{4}$$

where $\mathbf{x}$ is any value in the Ω domain such that the perturbed point $\mathbf{x}+\Delta$ is also in Ω, $\tilde{\Delta}_i = \Delta\left(x_{i,s}-x_{i,e}\right)$ and $x_{i,s}$, $x_{i,e}$ are start and end points of parameter x_i variation interval, respectively, Δ depend on *k* and describe the side length of the grid element. A finite distribution F_i for each parameter x_i is obtained

by sampling **x** in Ω. The number of elements of F_i is equal to $(k-1)k^{n-1}$. Distribution F_i of elementary effects is described by mean μ and standard deviation σ. A mean characterizes the sensitivity of the model output with respect to i^{th} parameter. A high mean indicates that the parameter is important and it substantially influences the output. A high standard deviation implies that the parameter interacts with other parameters or its effect to the model is nonlinear.

The results of the sensitivity of the model output with respect to the input model parameters expressed as the estimated mean of elementary effects are dependent on Δ value, which is selected arbitrarily. This dependence rises with the nonlinearity of the model and the sensitivity calculations may be not reliable. Thus, the calculations with the Morris Design algorithm are performed for various Δ and next the results are compared. The comparison is feasible for normalized quantities:

$$\tilde{\mu}_i = \frac{\mu_i}{\|\boldsymbol{\mu}\|} \qquad \tilde{\sigma}_i = \frac{\sigma_i}{\|\boldsymbol{\sigma}\|} \tag{5}$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$, $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$ are vectors of means and standard deviations calculated for all the input parameters x_i, $i = 1, \ldots, n$. If means and standard deviations computed for various Δ are close to each other, the sensitivities are properly estimated. If not, the value of Δ should be narrowed down and the procedure is repeated.

3.2 Sobol' method

Sobol in [4] developed the method of the global SA based on the variance analysis and the Monte Carlo algorithm. Let us assume that the domain of the input model parameters x_i, $i = 1, \ldots, n$, is defined as an n-dimensional cube Ω:

$$\Omega = \{\mathbf{x} : 0 \le x_i \le 1 \quad \forall i = 1, \ldots, n\} \tag{6}$$

Let the function $y = y(x)$ represents a model. Sobol defined the decomposition of $y(x)$ as the sum of the increasing dimensionality addends:

$$y(x_1, \ldots, x_n) = y_0 + \sum_{i=1}^{n} y_i(x_i) + \sum_{1 \le i < j \le n} y_{ij}(x_i, x_j) + \ldots + y_{1,2,\ldots,n}(x_1, \ldots, x_n) \tag{7}$$

The decomposition (7) is held if y_0 is constant and the integrals of every addend over its own variables is zero:

$$\int_0^1 y_{i_1,\ldots i_s}(x_{i_1}, \ldots, x_{i_s})\, dx_{i_k} = 0 \qquad \forall k : 1 \le k \le s \tag{8}$$

From (7) and (8) it is concluded that all the addends in (7) are orthogonal:

$$\int_\Omega y_{i_1,\ldots i_s}(x_{i_1}, \ldots, x_{i_s})\, y_{j_1,\ldots j_k}(x_{j_1}, \ldots, x_{j_k})\, d\mathbf{x} = 0 \qquad \forall (i_1, \ldots, i_s) \ne (j_1, \ldots, j_k) \tag{9}$$

and

$$y_0 = \int_\Omega y(\mathbf{x})\, d\mathbf{x} \tag{10}$$

Sobol' in [15] proved that the decomposition (7) is unique and all the decomposition addends can be evaluated as multidimensional integrals:

$$y_i(x_i) = -y_0 + \int_0^1 \ldots \int_0^1 y(\mathbf{x})\, d\mathbf{x}_{\sim i} \tag{11}$$

$$y_{ij}(x_i, x_j) = -y_0 - y_i(x_i) - y_j(x_j) + \int_0^1 \ldots \int_0^1 y(\mathbf{x})\, d\mathbf{x}_{\sim (ij)}$$

where $d\mathbf{x}_{\sim i}$ and $d\mathbf{x}_{\sim(ij)}$ denote integration over all the variables except x_i and x_i, x_j, respectively. Following that, the total variance Var is of the form:

$$\mathrm{Var} = \int_{\Omega} y^2(\mathbf{x})\, d\mathbf{x} - y_0^2 \tag{12}$$

and partial variances are estimated based on the terms in equation (7):

$$\mathrm{Var}_{i_1 \ldots i_s} = \int_0^1 \ldots \int_0^1 y^2_{i_1 \ldots i_s}\left(x_{i_1} \ldots x_{i_s}\right) dx_{i_1} \ldots x_{i_s} \tag{13}$$

where $1 \le i_1 < \ldots < i_s \le n$, $s = 1, \ldots, n$. Squared and integrated over Ω equation (7) gives:

$$\mathrm{Var} = \sum_{i=1}^{n} \mathrm{Var}_i + \sum_{1 \le i < j \le n} \mathrm{Var}_{ij} + \ldots + \mathrm{Var}_{1,2,\ldots,n} \tag{14}$$

Thus, the sensitivity measures $S_{i_1 \ldots i_s}$ are defined by:

$$S_{i_1 \ldots i_s} = \frac{\mathrm{Var}_{i_1 \ldots i_s}}{\mathrm{Var}} \tag{15}$$

S_i is called the first order sensitivity index for the parameter x_i and it measures the main effect of x_i on the model output. S_{ij}, for $i \neq j$, is the second order sensitivity index and it measures the interacted effect of the two parameters x_i and x_j on the model output. The higher order sensitivity indices can be defined in the same way. The multidimensional integration in calculations of variances (13) and (14) is performed with the Monte Carlo method, hence the efficiency of Sobol's algorithm depends mostly of efficiency of the Monte Carlo procedure.

4 Optimization procedure

On the basis of results obtained from sensitivity analysis the reference coefficients, their values and weights are established. A set of these parameters is used further in optimization procedure as ς_i in (3) . In this work the procedure is implemented in two versions i.e. for 2D and 3D, however the core part of the applied optimization algorithm is the same for both solutions. Implemented multi-iterative genetic algorithm [16] is composed if the following steps:

a) Generation of initial population – random generation of *n* specimens containing information about coordinates of control points and their weights. Size of the specimen depends strongly on number of control points describing martensitic inclusions in SSRVE. At the beginning of calculations the control points form *m* circles, where *m* is a number of inclusions. The main assumption is that those circles cannot touch each other and the summary phase fraction of these inclusions has to be equal to the phase fraction determined for original microstructure.

b) Estimation of objective function value – the procedure calculates values of ς_{iSSRVE} on the basis of shapes of martensitic inclusions in succeeding SSRVEs. Shapes of these inclusions are described by Non-uniform Relational B-Spline (NURBS). These parametric curves are controlled by basic interpolation functions and a set of mentioned control points, which are described by coordinates and weights. The weight of each control point influences the position of curve near the particular control point. The shapes of martensitic inclusions influence also rheological properties and statistical description of microstructure. Rheological model of SSRVE is determined by processing of this element in virtual uniaxial compression, tensile and shear deformation tests. Obtained stress-strain relations allow to calculate equivalent tensile stress, which describes material rheology. The main statistical measure is lineal-path function, calculated directly from pixels or voxels values of SSRVE.

c) Stop conditions – two fundamental conditions are implemented i.e. number of iterations and mean square error between expected and actual objective function.
d) Application of genetic operators – the following operators were implemented in the presented approach: 1x crossing and 2x mutation operators. The former operator is responsible for exchange of random number of genes between two specimens (Fig. 3).

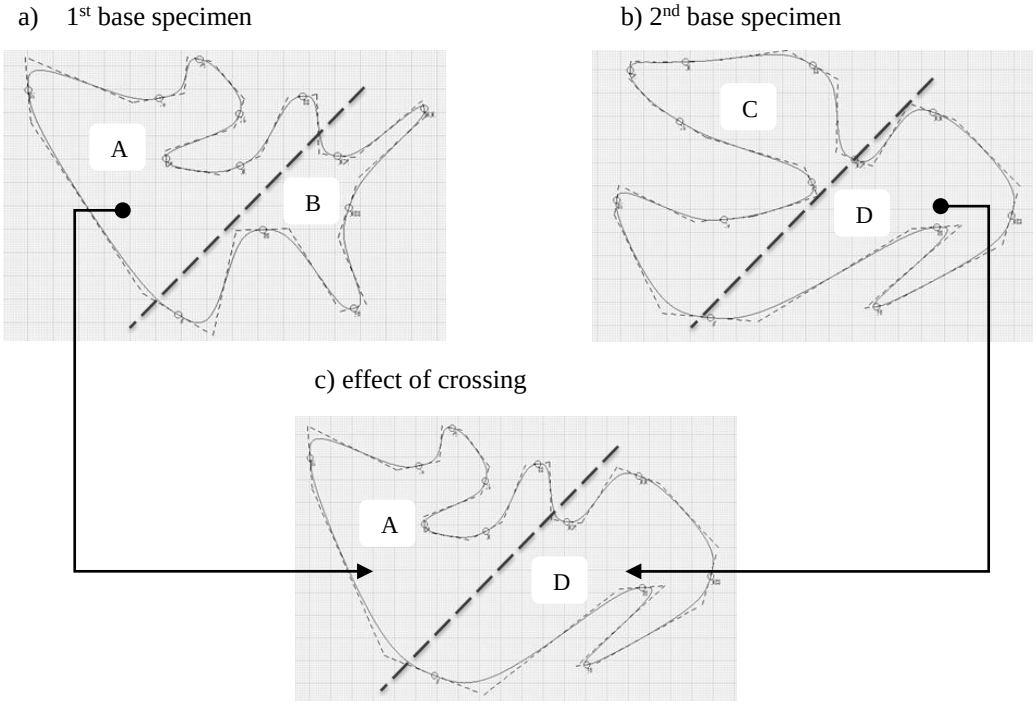


Figure 3: Illustration of crossing operator for 2D specimens.

The first of mutation operators (Fig. 4) changes positions of control points regarding center of gravity of the shape. The operator determines random set of control points, which will be mutated, and then the new position is calculated also on the basis of randomized operations. The second mutation operator changes the positions of control points in X or Y axis direction by using random values of coordinates in the vector of translation.

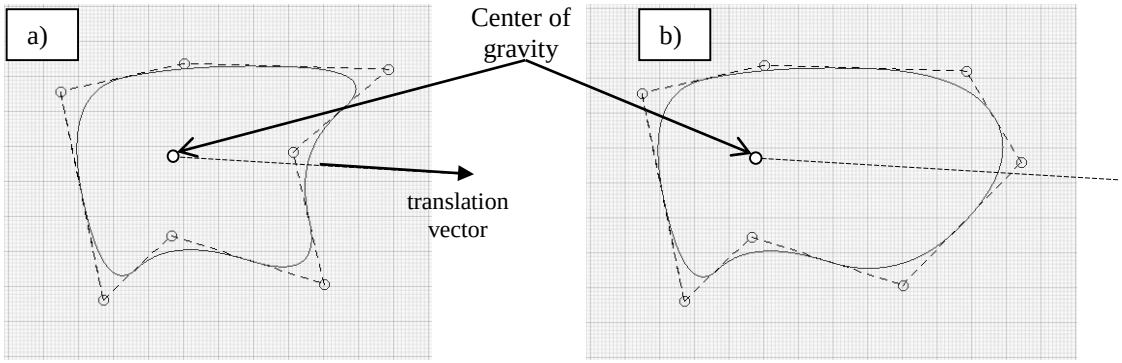


Figure 4: Illustration of mutation operator for 2D specimens.

e) Generation of new population – the specimens obtained after operations of crossing and mutation are included in the new population. In each subsequent iteration, each specimen is validated. On the basis of validation the worst specimens are removed from population, some of the best

specimens go further without changes and the rest is passed again to crossing and mutation operators.

The procedure of NURBS generation, crossing and mutation operators, representation of specimen are the main results between 2D and 3D approach. Shape of inclusion in 3D is composed of 2D layers, which were used in reconstruction algorithm mentioned at the beginning of this section. Thus, the specimen in 3D contains additional Z coordinates, which influences all the algorithms inside optimization procedure. Therefore, the crossing operator is based on exchange of a set of whole layers between two specimens instead of a set of single control points. The first mutation operator behaves similarly by translating of whole layers in Z axis direction, while the second mutation operator chooses two layers randomly and replace them inside one SSRVE.

5 Obtained results

The calculations in this work were performed for dual phase steel, which microstructure is presented in Fig. 5. The original micrograph was analysed by using filtering methods based on convolution algorithms and the method proposed by Authors, based on algorithm of dynamic particles [6]. The final segmentation is a result of watershed algorithm implemented by using cellular automata [17].

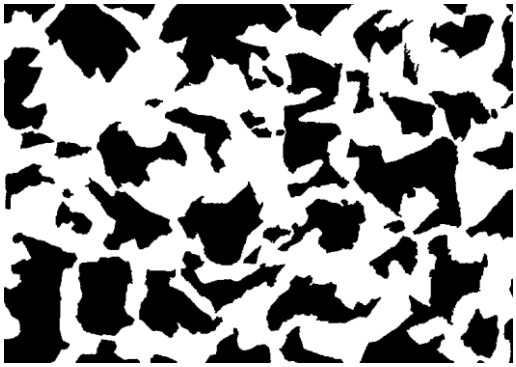


Figure 5: Results of sensitivity analysis.

Grains on the micrograph were analysed according to procedure described in [17], and the reference values were established. Afterwards sensitivity analysis by using Morris Design and Sobol method was applied. The results are presented in Fig. 6.

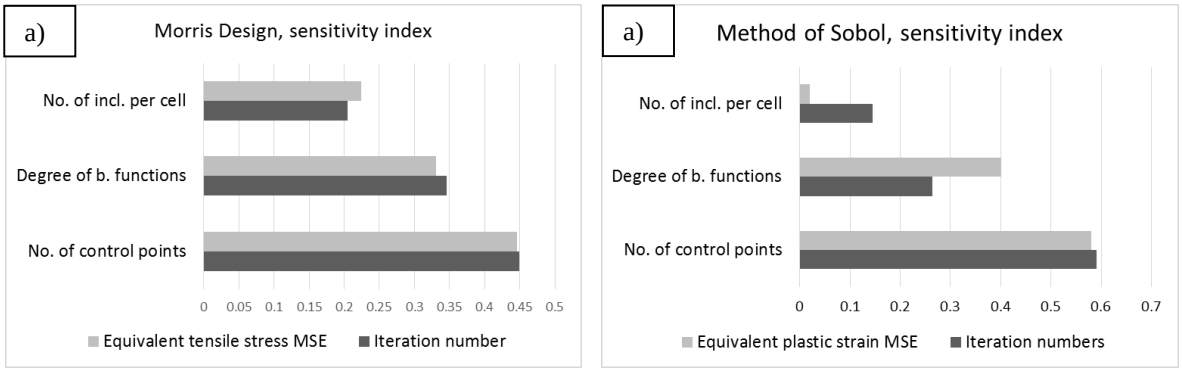


Figure 6: Results of sensitivity analysis for Morris Design (a) and Sobol method (b).

It can be seen that the parameter, which is the most influential on time of calculations and accuracy of the model, is the number of control points. The less important coefficient is number of inclusions per SSRVE, which does not increase qualitative nor quantitative results. However in the case of strong diversification of grains in a microstructure, it is recommended to design SSRVE based on two or three inclusions. The procedure applied in this work took into account four reference shape coefficients (maximum curvature = 4.964, border index = 0.831 and Malinowska = 0.341), rheological model, lineal-path function and phase fraction = 0.45. The results obtained for different number of control points are presented in Fig. 7a-d for 2D. Three dimensional result is presented in Fig. 7e. The values of objective function were in all cases below 0.01, while the best numerical performance was obtained in case of 8 and 12 control points. However, the best accuracy of rheological model was obtained for 20 control points. The solution based on 8 control points was selected and applied in 3D on 8 layers in Z axis direction. Therefore we obtained 64 control points, which offered value of objective function similar to 2D cases.

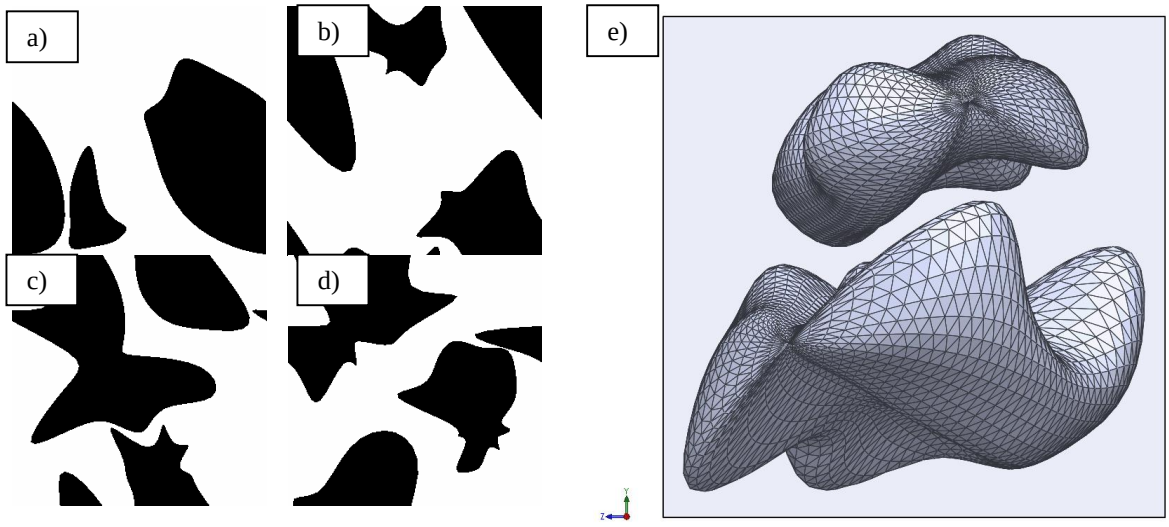


Figure 7: Results of optimization for 2D SSRVE in case of (a) 8, (b) 12, (c) 16, (d) 20 control points for each inclusion in SSRVE.

6 Conclusions

The paper presents procedure of identification of SSRVE, preceded by sensitivity analysis. The results obtained in this work proves high reliability of material representation by using SSRVE as well as satisfactory numerical performance. The optimization procedure based on the results of sensitivity analysis allows to setup better weights of parameters in objective function and to reduce its complexity. Therefore, 2D and 3D elements are created very efficiently. Further objectives of results will be focused on integration of the approach presented in this work with software dedicated to management of multiscale method in modern e-infrastructures (e.g. Scalarm software).

Acknowledgements

Financial assistance of the NCN, project 2011/03/B/ST8/06100, is acknowledged.

References

[1] L. Madej, A. Mrozek, W. Kuś, T. Burczyński, M. Pietrzyk, *Concurrent and upscaling methods in multi scale modelling -- case studies*, Computer Methods in Materials Science, 8, (2008), pp. 1-15.

[2] L. Rauch, R. Kuziak, M. Pietrzyk *From High Accuracy to High Efficiency in Simulations of Processing of Dual-Phase Steels*, Metallurgical and Material Transactions B, published online DOI: 10.1007/s11663-013-9926-5, (2014).

[3] R. Hill, *Elastic properties of reinforced solids: some theoretical principles*, Journal of the Mechanics and Physics of Solids, 11(5), (1963), pp. 357–372.

[4] J. Schroeder, D. Balzani, S. Brands, *Approximation of random microstructures by periodic statistically similar representative volume elements based on lineal-path functions*, Archives of Applied Mechanics, 81, (2011), pp. 975-997.

[5] M. Beran, Statistical continuum theories, Wiley, New York, (1968).

[6] L. Rauch, L. Madej, *Application of the automatic image processing in modelling of the deformation mechanisms based on the digital representation of microstructure*, International Journal for Multiscale Computational Engineering, 8, (2010), pp. 1-14.

[7] L. Rauch, K. Imiołek, Reconstruction of 3D material microstructure of one phase steels, Rudy i Metale Nieżelazne, 58(11), (2013), pp. 726–730.

[8] J. Ohser, F. Muecklich, *Statistical analysis of microstructure in materials science*, Wiley, New York, (2000).

[9] L. Rauch, M. Pernach, K. Bzowski, M. Pietrzyk, *On application of shape coefficients to creation of the statistically similar representative element of DP steels*, Computer Methods in Materials Science, 11, (2011), 531-541.

[10] S. Brands, J. Schroder, D. Balzani, *Statistically similar reconstruction of dual-phase steel microstructures for engineering applications*, in proc. CMM-2011 Computer Methods in Mechanics, Warsaw, (2011), CD ROM.

[11] B.L. Lu, S. Torquato, *Lineal-path function for random heterogeneous materials*, Journal of Applied Mechanics, 36, (1992), pp. 922-929.

[12] Povirk, G.L., Incorporation of microstructural information into models of two-phase materials, *Acta Metallurgica*, 43, (1995), 3199-3206.

[13] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis*, The Primer. John Wiley & Sons Ltd, (2008).

[14] M.D. Morris, Factorial sampling plans for preliminary computational experiments, Technometrics, 33, (1991), pp. 161–174.

[15] I. M. Sobol, *Sensitivity analysis for non-linear mathematical models*, Mathematical Modelling and Computational Experiment, 1, (1993), pp. 407–414.

[16] W. Kus, T. Burczynski, *Parallel evolutionary optimization in multiscale problems*, 9, (2009), pp.347-351.

[17] L. Rauch., L. Madej, J. Kusiak, *Modelling of microstructure deformation based on digital material representation integrated with watershed image segmentation algorithm*, Steel Research International, 81(9), 1446–1449, 2010.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Retrieval error due to model discordance in multiparameter nonlinear inversion of transient acoustic scattering response

Thierry Scotti, Armand Wirgin*

LMA, CNRS, UPR 7051
Aix-Marseille Univ, Centrale Marseille, F-13402 Marseille Cedex 20, France
e-mail: wirgin@lma.cnrs-mrs.fr

Key words: parameter retrieval, nonlinear inversion, model discordance, retrieval error

Abstract

The aim of this study is to simultaneously retrieve three parameters, of constitutive and geometrical nature, that define the state of a fluid-like object, by nonlinear inversion of synthetic transient acoustic response data. Contrary to most publications treating this subject, it is assumed herein that the retrieval model is not identical to the data model, this being so because some of the parameters (priors) in the retrieval model are different (by the simple fact of being more or less unknown) from their true counterparts in the data model. This so-called model discordance, which is the usual situation in real-world inverse problems, is a source of errors for the retrieval of the other parameters. Retrieval errors as a function of prior uncertainty are computed for a a lossy fluid-like cylinder object, and found to be large, for various model discordance scenarios involving one uncertain prior, and noiseless data.

1 Introduction

The ultimate goal of the majority of the applications of inverse scattering is to obtain a quantitative image [9] of a corrupted (such as by a crack: [1], [2]), diseased (such as affected by a tumor: [8], [17], or otherwise inhomogeneous [7], [3] body. To do this in an oft-employed manner (e.g., via a domain integral approach: [7], [16]) requires the previous knowledge of the (termed 'specific' in [16], [6]) Green's function of the corresponding uncorrupted, healthy [7]), or homogeneous body. The determination of this Green's function requires quantitative information about the size, shape and constitutive parameters of the healthy body, which must be determined by solving another inverse (scattering or vibration) problem. If the diseased body is a cylinder containing a tumor, then the healthy body is the same cylinder without the tumor; this cylinder is the object we wish to characterize.

We revisit an oft-studied acoustically-solicited (analogous to the electromagnetic wave: [5]) 'healthy' configuration involving a canonical object (simple circular shape, homogeneous viscous fluid [11], [12], [13] (analogous to lossy dielectric [5]) composition), thereby:

- enabling the direct-scattering problem to be solved in a quasi-analytic manner (e.g., [16]), and
- enabling nonlinear full (transient)- wave inversions (e.g., [15], [6], [14], [12], [10], [13]) to be carried out in a reasonable amount of time.

Numerous numerical tests of the multiparameter retrieval (such as in [10],[13]) are made herein in the hope of shedding light on the crucial issue of retrieval error of more-complicated (i.e., involving objects with non-canonical shapes and/or that are inhomogeneous) transient-wave inverse problems.

2 Ingredients of multiparameter inversion of time- domain data

The model of the physical process which accurately mimics the data during the inversion is termed *retrieval model*, or $\mathfrak{M}$, characterized by its mathematical ingredients $\mathbf{M}$ and numerical ingredients $\mathbf{N}$. We term the model that mimicks measured data, the *data simulation model*, or $\mathfrak{m}$, characterized by its mathematical ingredients $\mathbf{m}$ and numerical ingredient(s) $\mathbf{n}$. The physical/geometric parameters of $\mathfrak{m}$ and $\mathfrak{M}$ are $\mathbf{p}$ and $\mathbf{P}$ respectively, i.e., lower-case letters apply to the (true) data (simulated or experimental) and upper-case letters to the retrieval model.

Let $\mathbf{r} := \{r_j \ ; \ j = 1, 2, ..., N_r\}$ be the set of to-be-retrieved parameters, $\mathbf{q} := \{q_j \ ; \ j = 1, 2, ...N_q\}$ the set of priors (whose nature will be made more precise further on), $\mathbf{X} := \{\mathbf{x}_j \ ; \ j = 1, 2, .., N_x\}$ the set of sensor locations and $\mathbf{t} := \{t_l \ ; \ l = 1, 2, .., N_t\}$ the set of instants of the sampled signal. At present, $N_r = 3$ and $N_q = 11$.

The basic scheme is to replace the true (fixed) parameter set $\mathbf{r}$ by the trial (variable) parameter set $\mathbf{R} := \{R_1, R_2, ..., R_{N_r}\}$ and the true (fixed) set of priors $\mathbf{q}$ by another fixed set $\mathbf{Q} := \{Q_1, Q_2, ..., Q_{N_q}\}$, and search, in iterative manner, for the set $\mathbf{R} = \tilde{\mathbf{r}}$ that minimizes the discrepancy (the measure of which is a cost or objective functional) between trial signals (resulting from trial parameters) $\mathfrak{S}(\mathbf{X}, \mathbf{t}, \mathbf{Q}, \mathbf{R})$ and the true signal $\mathfrak{s}(\mathbf{X}, \mathbf{t}, \mathbf{q}, \mathbf{r})$ (resulting from the true parameters).

Note that if, as will be assumed in some of the examples presented hereafter, there exists some uncertainty in the priors, $\mathbf{Q}$ will not be identical to $\mathbf{q}$. In fact, this possible (model) *discordance* will affect the accuracy of the retrieval and it is of considerable interest to quantify this effect.

The discrepancy between $\mathfrak{S}$ and $\mathfrak{s}$ is the *cost functional* $\mathcal{K}$. We choose this as:

$$\mathcal{K}(\mathbf{r}, \mathbf{R}, \mathbf{q}, \mathbf{Q}, \mathbf{X}, \mathbf{t}) = \frac{\sum_{j=1}^{N_x} \sum_{k=1}^{N_t} [\mathfrak{s}(\mathbf{x}_j, t_k, \mathbf{q}, \mathbf{r}) - \mathfrak{S}(\mathbf{x}_j, t_k, \mathbf{Q}, \mathbf{R})]^2}{\sum_{j=1}^{N_x} \sum_{k=1}^{N_t} [\mathfrak{s}(\mathbf{x}_j, t_k, \mathbf{q}, \mathbf{r})]^2} \ . \tag{1}$$

The inverse problem is solved, at each stage, by minimization of the cost functional via the Nelder-Mead Simplex scheme [4] implemented in MATLAB by the function *fminsearch*. The minimum of the cost functional (in the parameter space explored by the Simplex scheme, anchored at the user-defined starting values of the to-be-retrieved parameters, is found for $\mathbf{R} = \tilde{\mathbf{r}}$, i.e.,

$$\tilde{\mathbf{r}} = \arg \min_{\mathbf{R} \subset \mathcal{S}} \ \mathcal{K}(\mathbf{r}, \mathbf{R}, \mathbf{q}, \mathbf{Q}, \mathbf{X}, \mathbf{t}) \ , \tag{2}$$

wherein $\mathcal{S}$ designates the parameter search space during a given stage of the search of a minimum, a stage being defined as a specific choice of the initial value set (since $fminsearch$ requests such a set to get started). We call the initial value set $\mathbf{R} := \{R_j \ ; \ j = 1, 2, 3, N_r = 4\}$. Note that $\mathbf{R}$ takes on

different values (from this initial value) as the minimization proceeds. The schemes considered in this investigation actually solve for all members of the set $\mathbf{r}$ at each stage.

Often, a priori information, concerning the solutions of the inverse problem, is of the form: 'the sought-for parameter r_j lies between R_{bj} and R_{ej}', which fact suggests that one should try to carry out the retrieval for a variety of starting values lying between R_{bj} and R_{ej}.

More precisely, we shall carry out the inversion for the *set of starting values* $\mathbf{R} := \{R_{jk} \; ; \; j = 1, 2, 3, N_r = 4 \; ; \; k = 1, 2, ..., N_{R_j}\}$, with $R_{jk} := R_{bj} + (k-1)d_{R_j} \; ; \; j = 1, 2, 3 \; ; \; k = 1, 2, ..., N_{R_j}$ and $d_{R_j} := \frac{R_{ej} - R_{bj}}{N_{R_j} - 1}$. Thus, supposing $N_{R_1} = N_{R_2} = N_{R_3} := N_R$, the number of stages, and the number of retrieved solutions, will be N_R^3, a number that increases rapidly with N_R since $2^3 = 8$, $3^3 = 27$, $4^3 = 48$,... The question (alluded-to previously) is then how to select the most appropriate (i.e., 'best') solution among this multitude of retrievals.

The result of an inversion scheme can lead to (one or more) r_j that are not within the initial search interval $[R_{bj}, R_{ej}]$, but a logical option (i.e., equivalent to a regularization scheme) is to reject such retrievals (otherwise, why impose initial search intervals other than to constrain the retrievals?) and then apply the min min cost rule (select the solution, amongst those obtained by the Simplex minimization scheme, corresponding to the smallest minimum of the cost functions) to the remaining solutions. In the sequel, we shall apply this regularization scheme [13] to our canonical inverse problem.

2.1 The retrieval error generated by prior uncertainty

Nonlinear inversion usually leads to error-ridden retrievals of the $j-$th parameter (recall that the true value of this parameter is r_j). Our aim will be to evaluate the relative error of the final retrieved parameter $\tilde{r}_j$ ($j = 1, 2, ..., N_r = 4$)

$$\varepsilon_j := \frac{\tilde{r}_j - r_j}{r_j} \; , \tag{3}$$

as a function of the prior uncertainties, which, for the k-th prior, is

$$\delta_k := \frac{Q_k - q_k}{q_k} \; . \tag{4}$$

Note that it is the theoretical (rather than experimental) nature of our study which makes it possible to evaluate the retrieval errors since in real life the true values of the various parameters are not known.

3 The models for computing the signals $\mathfrak{s}(\mathbf{x}_j, t_k, \mathbf{q}, \mathbf{r})$ and $\mathfrak{S}(\mathbf{x}_j, t_k, \mathbf{Q}, \mathbf{R})$

3.1 Model description

An infinitely-long (in the z-direction of an $Oxyz$ cartesian reference system) lossy fluid-like cylinder, (known to be) centered at the origin O and of radius a, is immersed in another (host, non-lossy) fluid and is submitted to an acoustic plane wave whose wavevector $\mathbf{k}^i$ lies in the $x - y$ (sagittal) plane. Consequently, the incident and diffracted fields are independent of z, i.e., the problem is 2D, with z the ignorable coordinate.

The plane wave carries a pulse, and the purpose of this excitation field is to produce a pulse-like diffracted field (i.e., signal) constituting the data, which by means of an inversion scheme, is analyzed to enable the retrieval of the radius of the cylinder and constitutive parameters of the fluid enclosed therein.

The constitutive parameters of the host medium, as well as the excitation parameters, may, or may not, be precisely known beforehand.

The analysis takes place in the sagittal plane wherein Ω_0 denotes the domain exterior to the cylinder, Γ the circular bounding curve of the cylinder, and Ω_1 the domain within Γ.

The homogeneous fluids in Ω_l are $M^{[l]}$, $l = 0, 1$. The acoustic bulk phase velocities in these media are $c^{[l]} = \Re c^{[l]} + i\Im c^{[l]}$, with $\Im c^{[1]} \leq 0$ and $\Im c^{[0]} = 0$ (note that $\Im c^{[l]}$ is a measure of attenuation which is equal to zero in a non-attenuating medium). The densities in $M^{[l]}$, $l = 0, 1$ are $\rho^{[l]}$. The corresponding wavenumbers are $k^{[l]} := \omega/c^{[l]}$, with $\omega = 2\pi f$ the angular frequency and f the frequency.

The bandwidth of the incident pulse is, for all practical purposes, finite, and within this band the phase velocities of the two fluids are assumed to be constant with respect to the frequency f.

The (incident) angle between $\mathbf{k}^i$ and the x axis is θ^i.

The diffracted field is sensed at various (N_x) points $\mathbf{X}$ on a circle of radius $b > a$ enclosing the cylinder and having the common origin O. The polar angle at which a generic point-like sensor is located is θ_j (with respect to the positive x axis). $\mathbf{X}$ is assumed to be precisely known.

The objective is to retrieve the parameters $c^{[1]r} := \Re c^{[1]}$, $c^{[1]i} := \Im c^{[1]}$, $\rho^{[1]}$ and radius a from data relating to the diffracted (scattered) field signals recorded at one or more sensors.

Recall that in this study, the experiment for obtaining the diffracted field data signal(s), rather than being carried out physically in the laboratory or field, is simulated by computational means via the $data\ simulation\ model$, whereas the inversion requires another model, termed the $retrieval\ model$. We assume the mathematical and numerical ingredients of these two models to be identical; the two models differ only by the values of various physical/geometrical parameters entering therein and eventually by noise added to the data signal (we treat the noiseless case here; otherwise see [13]).

Each model is employed to solve a $direct\ scattering\ problem$, once as concerns the generation of the synthetic data, and as many times as necessary during the inversion process.

3.2 Acoustic field solutions

The incident plane-wave pressure field is:

$$u^{[0]i}(\mathbf{x}, \omega) = S(\omega) e^{-ikr\cos(\theta - \theta^i)} = S(\omega) \sum_{m \in \mathbb{Z}} \gamma_m J_m(k^{[0]} r) e^{im\theta} \,, \tag{5}$$

wherein $\mathbf{x} = (r, \theta)$, $S(\omega)$ is the spectrum of the pulse associated with the incident wave, and $\gamma_m = e^{-im\left(\theta^i + \frac{\pi}{2}\right)}$. We choose a pseudo-Ricker type of pulse excitation whose spectrum $S(\omega) = \frac{\omega^2}{4\alpha^3\sqrt{\pi}} \exp\left(i\beta\omega - \frac{\omega^2}{4\alpha^2}\right)$. The diffracted pressure field outside the cylinder is [16]:

$$u^{[0]d}(\mathbf{x}, \omega) = S(\omega) \sum_{m \in \mathbb{Z}} \mathfrak{A}_m H_m^{(1)}(k^{[0]} r) e^{im\theta} \,, \tag{6}$$

wherein $\mathfrak{A}_m = \gamma_m A_m$, $\eta := \frac{k^{[0]}\rho^{[1]}}{k^{[1]}\rho^{[0]}}$, $\dot{Z}_m(z) := \frac{d}{dz} Z_m(z)$, with $Z_m(z) = J_m(z)$ (Bessel function) or $Z_m(z) = H_m^{(1)}(z)$ (Hankel function). The time-domain diffracted field outside the cylinder (termed 'signal' when $r = b$) is

$$U^{[0]d}(\mathbf{x}, t) = 2\Re \int_0^\infty u^{[0]d}(\mathbf{x}, \omega) e^{-i\omega t} d\omega \,. \tag{7}$$

Recall that that the field is sensed on a circle of radius $r = b > a$. We take into account only the first $N + 1$ terms of the series in $u^{[0]d}(b, \theta, \omega)$, with $N_f = \mathfrak{f}k^{[0]}b + 3$ and $\mathfrak{f}$ a user-chosen factor that is usually equal to 2 or 3. We carry out the Fourier transform as follows

$$U^{[0]d}(b, \theta, t) \approx 2\Re \int_{2\pi f_d}^{2\pi f_f} u^{[0]d\ N_f}(b, \theta, \omega) e^{-i\omega t} d\omega \ , \tag{8}$$

with f_d a very small number, f_f chosen large enough to prevent the generation of numerical artifacts, and the finite integral evaluated by the Simpson scheme.

4 The parameters involved in the simulated data and retrieval data

4.1 The parameters involved in the (simulated) data

Our simulated data involves fourteen physical/geometric parameters (PGP) and two numerical parameter (NP) sets (see Tab. 1), both designated by lower-case characters. The PGP consist of: the four to-be-

Table 1: Correspondences between the different (lower case) symbols relating to the simulated/experimental (fixed during the retrieval) parameters. The entries of a given column represent the same parameter.

p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}	p_{14}
r_1	r_2	r_3	r_4	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}
$c^{[1]r}$	$c^{[1]i}$	$\rho^{[1]}$	a	$c^{[0]}$	$\rho^{[0]}$	θ^i	α	β	b	θ_b	θ_e	N_θ	$\mathcal{P}$

retrieved parameters of the cylinder $p_1,\ p_2,\ p_3,\ p_4$, the two background medium priors $p_5,\ p_6$, the three excitation priors $p_7,\ p_8,\ p_9$ and the five receiver priors $p_{10},\ p_{11},\ p_{12},\ p_{13},\ p_{14}$.

The polar angle of the j-th receiver (the total number of receivers being N_θ) is

$$\theta_j = \theta_b + (j - 1)d_\theta \ , \tag{9}$$

and

$$d_\theta = \frac{\theta_e - \theta_b}{N_\theta - 1} \ , \tag{10}$$

it being understood that

$$\mathfrak{s}(\mathbf{x}_j, t_k, \mathbf{q}, \mathbf{r}) = U^{[0]d}(b, \theta_j, t_k, \mathbf{q}, \mathbf{r}) \ , \tag{11}$$

is the (simulated) data signal at the j-th receiver and instant t_k, and that

$$\mathfrak{S}(\mathbf{x}_j, t_k, \mathbf{Q}, \mathbf{R}) = U^{[0]d}(b, \theta_j, t_k, \mathbf{Q}, \mathbf{R}) \ , \tag{12}$$

is the retrieval signal at the same receiver and instant.

In practice, the retrieved signal is affected by noise whose amplitude is accounted-for by the parameter $\mathcal{P}$, but in the sequel we consider only noiseless data corresponding to $\mathcal{P} = 0$.

The two NP sets are: $\mathbf{n}_1 = (f_d, f_f, N_f, \mathfrak{f})$, $\mathbf{n}_2 = (t_d, t_f, N_t)$. We stress the fact that the first two NP employed in the retrieval model are identical to those in the data simulation model, i.e., $\mathbf{N}_l = \mathbf{n}_l \ ; \ l = 1, 2$.

4.2 The parameters involved in the retrieval data

Our retrieval data involves fourteen physical/geometric (PGP) parameter sets and two numerical parameter (NP) sets (see Tab. 2), both designated by upper-case characters.

Table 2: Correspondences between the different (upper case) symbols relating to the retrieval parameters. The entries of a given column represent the same parameter or vector.

$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$	$\mathbf{P}_4$	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}	P_{14}
$\mathbf{R}_1$	$\mathbf{R}_2$	$\mathbf{R}_3$	$\mathbf{R}_4$	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}
$(C_b^{[1]r}$, $C_e^{[1]r}$, $N_{C^{[1]r}})$	$(C_b^{[1]i}$, $C_e^{[1]i}$, $N_{C^{[1]i}})$	$(\mathcal{R}_b^{[1]}$, $\mathcal{R}_e^{[1]}$, $N_{\mathcal{R}^{[1]}})$	$(A_b$, A_e, $N_A)$	$C^{[0]}$	$\mathcal{R}^{[0]}$	Θ^i	$\mathcal{A}$	$\mathcal{B}$	B	Θ_b	Θ_e	N_Θ	$\mathcal{P}$

The PGP consist of: the three parameter sets relative to the cylinder $\mathbf{P}_1,\ \mathbf{P}_2,\ \mathbf{P}_3,\ \mathbf{P}_4$, the two background priors $P_5,\ P_6$, the three excitation priors $P_7,\ P_8\ P_9$, and the five receiver priors $P_{10},\ P_{11}\ P_{12},\ P_{13},\ P_{14}$, it being understood that

$$\mathbf{P}_j = \mathbf{R}_j = (R_{bj}, R_{ej}, N_{R_j})\ ;\ j = 1, 2, 3\ , \tag{13}$$

and that we choose: $P_j = p_j\ ;\ j = 10, 12, 13, 14$.

Furthermore,

$$\mathfrak{S}(\mathbf{x}_j, t_k, \mathbf{Q}, \mathbf{R}) = U^{[0]d}(b, \theta_j, t_k, \mathbf{Q}, \mathbf{R})\ , \tag{14}$$

is the retrieval signal at the j-th receiver and instant t_k.

The three numerical parameter sets are: $\mathbf{N}_1 = (F_d, F_f, N_F, \mathfrak{F})$, $\mathbf{n}_2 = (t_d, t_f, N_t)$, $\mathbf{n}_3 = (MFE)$, with MFE the maximum number of function evaluations per stage in each call to the Simplex function.

We always chose $\mathbf{N}_l = \mathbf{n}_l\ ;\ l = 1, 2$ and $MFE = 5000$.

5 Parameters and conventions common to the figures

The true (target) parameters are: $c^{[1]r} = 1700$ m/s, $c^{[1]i} = -210$ m/s, $\rho^{[1]} = 1300$ kg/m^3 , $a = .1$ m. The true priors are: $c^{[0]} = 1500$ m/s, $\rho^{[0]} = 1000$ kg/m^3, $\theta^i = 0°$, $\alpha = 2000$ Hz, $\beta = .1$ s, $b = 1$ m. The numerical parameters are: $\mathbf{f}$ =(.003 Hz,2000 Hz,201), $\mathbf{t}$ = (.098 s, .103 s,181). We take $\boldsymbol{\theta} = (0°, 180°, 5)$.

As stated earlier, we *do* possess some knowledge at the outset of the values of the parameters we wish to retrieve. This knowledge takes the form of intervals which are used to initialize the Simplex optimization algorithm. To parametrize this degree of knowledge, we employ the variable $d_{R_j} = d = .9\ ;\ j = 1, 2, 3$ for the initial parameter intervals $[C_b^{[1]r} = c^{[1]r}(1-d),\ C_e^{[1]r} = c^{[1]r}(1+d)]$, $[C_b^{[1]i} = c^{[1]i}(1-d),\ C_e^{[1]i} = c^{[1]i}(1+d)]$, $[\mathcal{R}_b^{[1]} = \rho^{[1]}(1-d),\ \mathcal{R}_e^{[1]} = \rho^{[1]}(1+d)]$, $[A_b = a(1-d),\ A_e = a(1+d)]$.

6 Tables expressing the retrieval errors of three parameters

The following tables indicate the relative retrieval errors (ε) incurred by a fixed relative uncertainty (δ) of a selected prior (one at a time). We take $\delta = \pm 0.1$ for all but the prior $\mathcal{B}$ and $\delta = \pm 0.0004$ for prior $\mathcal{B}$. This is done for the various programs we employed to simultaneously retrieve three parameters amongst $c^{[1]r}$, $c^{[1]i}$, $\rho^{[1]}$, a.

Table 3: The entries of this table are the retrieval error $\varepsilon_{c^{[1]r}}$, given in this, and the other following tables, as a function of the various prior uncertainties (one at a time, the other priors taking their true values) and the programs by which the retrievals were carried out. For example, program $r1r2r3$ retrieves $r_1 = c^{[1]r}$, $r_2 = c^{[1]i}$, and $r_3 = \rho^{[1]}$ simultaneously. The prior uncertainty is: $\delta = -0.1$ for the priors $C^{[1]i}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = -0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r3				.040	-.191	.000	.014	.039
r1r2r4			-.064		-.201	.071	.008	.045
r1r3r4		-.009			-.189	.000	.004	.228

Table 4: The entries of this table are the retrieval error $\varepsilon_{c^{[1]r}}$. The prior uncertainty is: $\delta = +0.1$ for the priors $C^{[1]i}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = +0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r3				-.028	.203	.000	-.017	-.057
r1r2r4			.065		.171	-.058	-.008	-.056
r1r3r4		.009			.429	.000	-.005	-.067

Table 5: The entries of this table are the retrieval error $\varepsilon_{c^{[1]i}}$. The prior uncertainty is: $\delta = -0.1$ for the priors $C^{[1]r}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = -0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r3				.522	.147	.000	.134	-.644
r1r2r4			-.593		.612	1.145	.034	-.627
r2r3r4	-.829				1.908	.000	.046	-.742

Table 6: The entries of this table are the retrieval error $\varepsilon_{c^{[1]i}}$. The prior uncertainty is: $\delta = +0.1$ for the priors $C^{[1]r}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = +0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r3				-.281	-1.045	.000	-.154	.402
r1r2r4			1.000		-1.004	-.551	-.033	.486
r2r3r4	-2.120				-1.000	.000	-.040	1.362

Table 7: The entries of this table are the retrieval error $\varepsilon_{\rho^{[1]}}$. The prior uncertainty is: $\delta = -0.1$ for the priors $C^{[1]r}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = -0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r3				.058	-.052	-.100	.018	-.006
r1r3r4		-.013			-.069	-.100	.005	.113
r2r3r4	-.154				-.012	-.100	.012	-.042

Table 8: The entries of this table are the retrieval error $\varepsilon_{\rho^{[1]}}$. The prior uncertainty is: $\delta = +0.1$ for the priors $C^{[1]r}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = +0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r3				-.041	.019	.100	-.023	-.007
r1r3r4		.013			.112	.100	-.006	-.053
r2r3r4	.171				-.158	.100	-.012	.049

Table 9: The entries of this table are the retrieval error ε_a. The prior uncertainty is: $\delta = -0.1$ for the priors $C^{[1]r}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = -0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r4			.333		-.124	-.165	.025	-.012
r1r3r4		.030			.052	.000	.026	-.179
r2r3r4	.898				-.275	.000	.021	.090

Table 10: The entries of this table are the retrieval error ε_a. The prior uncertainty is: $\delta = +0.1$ for the priors $C^{[1]r}$, $\mathcal{R}^{[1]}$, A, $C^{[0]}$, $\mathcal{R}^{[0]}$, $\mathcal{A}$ and $\delta = +0.0004$ for the prior $\mathcal{B}$.

	$C^{[1]r}$	$C^{[1]i}$	$\mathcal{R}^{[1]}$	A	$C^{[0]}$	$\mathcal{R}^{[0]}$	$\mathcal{A}$	$\mathcal{B}$
r1r2r4			-.153		.040	.286	-.038	-.018
r1r3r4		-.025			-1.860	.000	-.039	.119
r2r3r4	-.224				.988	.000	-.035	-.145

7 Discussion of results and conclusion

We consider a retrieval to be:
- accurate for a given parameter r_j with respect to a prior uncertainty δ_k of prior R_k if $|\varepsilon_j| \leq |\delta_k|$,
- fairly accurate if $|\delta_k| < |\varepsilon_j| \leq 3|\delta_k|$,
- inaccurate if $3|\delta_k| < |\varepsilon_j| \leq 10|\delta_k|$.
- without signification (or helplessly inaccurate) if $|\varepsilon_j| > 10|\delta_k|$.

A first comment applicable to all the tables Tab. 3-Tab. 10: the retrievals of all four parameters (three at a time) are hopelessly inaccurate with respect to the excitation prior $\mathcal{B}$ (which is related to the instant of onset of the pulse), even for the choice made herein of the extremely-small prior uncertainties $\delta_\beta = \pm.0004$, this being true whatever program is employed.

A second comment again applicable to all the tables Tab. 3-Tab. 10: the choice of $\pm 10\%$ uncertainties of the other priors is relatively modest, in that in many real-life situations, either an excitation prior is not properly measured, and/or the value of a constitutive prior is simply guessed or taken from the literature and not necessarily-transposable to the sample under investigation.

Tab. 3 shows that the retrieval of $c^{[1]r}$ is accurate with respect to all priors except $C^{[0]}$ and $\mathcal{B}$. The retrieval of $c^{[1]r}$ is fairly inaccurate with respect to $C^{[0]}$.

Tab. 4 shows that the retrieval of $c^{[1]r}$ is accurate with respect to all priors except $C^{[0]}$ and $\mathcal{B}$, this being true by all the programs. The retrieval of $C^{[0]}$ is fairly accurate by $r1r2r3$, $r1r2r4$, and inaccurate by $r1r3r4$.

Tab. 5 shows that the retrieval of $c^{[1]i}$ is fairly accurate with respect to priors $C^{[0]}$ and $\mathcal{A}$ by program $r1r2r3$. The retrieval is accurate with respect to prior $\mathcal{A}^{[0]}$ by $r1r2r4$ and $r2r3r4$ and with respect to $\mathcal{R}^{[0]}$ by $r1r2r3$ and $r2r3r4$. The retrievals of $c^{[1]i}$ are inaccurate with respect to priors $C^{[1]r}$ and $\mathcal{R}^{[1]}$ and A . They are also inaccurate with respect to $C^{[0]}$ by $r1r2r4$. Helplessly inaccurate retrievals are obtained with respect to prior $\mathcal{R}^{[0]}$ by $r1r2r4$, and with respect to $C^{[0]}$, by $r2r3r4$.

Tab. 6 shows that the retrieval of $c^{[1]i}$ is accurate or fairly accurate with respect to prior $\mathcal{A}$. The retrievals are accurate with respect to $\mathcal{R}^{[0]}$ by $r1r2r3$ and $r2r3r4$, while being inaccurate by $r1r2r4$. The retrievals are hopelessly inaccurate with respect to priors $C^{[1]r}$ and $\mathcal{R}^{[1]}$, and they are fairly inaccurate with respect to A. Helplessly inaccurate retrievals are obtained with respect to prior $C^{[0]}$ for all programs.

Tab. 7 shows that the retrievals of $\rho^{[1]}$ with respect to prior $\mathcal{R}^{[0]}$, are such that $\varepsilon_{\rho^{[1]}} = \delta_{\mathcal{R}^{[0]}}$. The retrievals of $\rho^{[1]}$, with respect to the other priors except $\mathcal{B}$, are either accurate or fairly accurate.

Tab. 8 shows that he retrievals of $\rho^{[1]}$, with respect to prior $\mathcal{R}^{[0]}$, are such that $\varepsilon_{\rho^{[1]}} = \delta_{\mathcal{R}^{[0]}}$. The retrievals of $\rho^{[1]}$, with respect to the other priors except $\mathcal{B}$, are either accurate or fairly accurate.

Tab. 9 shows that the retrievals of a are either accurate or fairly accurate with respect to all the priors except $C^{[1]r}$, $\mathcal{R}^{[1]}$ and $\mathcal{B}$. Inaccurate retrievals are obtained with respect to $C^{[1]r}$ and $\mathcal{R}^{[1]}$.

Tab. 10 shows that the retrievals of a are either accurate or fairly accurate with respect to all the priors except $C^{[0]}$ and $\mathcal{B}$. A hopelessly inaccurate retrieval is obtained with respect to $C^{[0]}$ by program $r1r3r4$.

In conclusion: this investigation shows that the accuracy of a retrieved parameter (especially as it relates to attenuation) is largely conditioned by prior uncertainty and can be low (i.e., $|\varepsilon|$ large compared to $|\delta|$) even for small prior uncertainty (especially as it relates to the excitation prior $\mathcal{B}$).

References

[1] J.D. Achenbach , *Quantitative nondestructive evaluation*, Int.J.Solids Struct., 37, (1980), pp. 13-27.

[2] G. Alessandrini , A. Bilotta, G. Formica, A. Morassi, E. Rosset, E. Turco, *Evaluating the volume of a hidden inclusion in an elastic body*, J.Comput.Appl.Math., 198, (2007), pp. 288-306.

[3] X. Chen, T.M. Grzegorczyk, B.-I. Wu , J. Pacheco Jr., J.A. Kong, *Robust method to retrieve the constitutive effective parameters of metamaterials*, Phys.Rev. E, 70, (2004), 016608.

[4] J. F. Devlin, *A simple and powerful method of parameter estimation using simplex optimization*, Ground Water, 32, (1994), pp. 323-327 .

[5] A. Dubois A., J.-M. Geffrin, K. Belkebir, *Imaging of dielectric cylinders from experimental stepped-frequency data*, Appl.Phys.Lett., 88, (2006), 164104 .

[6] C. Gelis, J. Virieux, G. Grandjean, *Two-dimensional elastic full waveform inversion using Born and Rytov forumulations in the frequency domain*, Geophys.J.Intl., 168, (2007), 605-633.

[7] R. Guillermin, *Caractérisation d'objets enfouis dans des sédiments marins par imagerie acoustique*, Phd thesis, Université de la Méditerranée Aix-Marseille II, Marseille, France, (2000).

[8] P. Huthwaite, F. Simonetti, *High-resolution imaging without iteration: A fast and robust method for breast ultrasound tomography*, J.Acoust.Soc.Am., 130, (2011), pp. 1721-1734.

[9] R.J. Lavarello Montero, *New developments on quantitative imaging using ultrasonic waves*, Phd Thesis, University Illinois, Urbana-Champaign, USA (2009).

[10] G. Lefeuve-Mesgouez, A. Mesgouez, E. Ogam, T. Scotti, A. Wirgin, *Retrieval of the physical properties of an anelastic solid half space from seismic data*, J.Appl.Geophys., 88, (2013), pp. 70-82.

[11] E. Ogam, C. Depollier, Z.E.A. Fellah, *The direct problem of acoustic diffraction of an audible probe radiation by an air-saturated prous cylinder*, J.Appl.Phys.,108, (2010), 113519.

[12] E. Ogam, Z.E.A. Fellah, *The direct and inverse problems of an air-saturated poroelastic cylinder submitted to acoustic radiation*, AIP Advances, 1, (2011), 032174.

[13] T. Scotti, A. Wirgin, *Multiparameter identification of a lossy fluid-like object from its transient acoustic response*, Inverse Prob.Sci.Engrg., (2013), doi: 10.1080/17415977.2013.867485.

[14] C.-H. Sun, C.-C. Chiu, C.-L. Li, C.H. Huang, *Time domain image reconstruction for homogenous dielectric objects by dynamic differential evolution*, Electromagnetics, 30, (2010), pp. 309-323 .

[15] A. Tarantola, *A strategy for nonlinear elastic inversion of seismic reflection data*, Geophys., 51, (1986), 1893-1903.

[16] A. Wirgin, *Some quasi-analytic and numerical methods for acoustical imaging of complex media*, in Wirgin A., ed., *Wavefield Inversion*, Springer, Wien, Austria, (1999), pp. 241-304.

[17] J. Wiskin, D.T. Borup, S.A. Johnson, M. Berggren, *Non-linear inverse scattering: High resolution quantitative breast tissue tomography*, J.Acoust.Soc.Am., 131, (2012), pp. 3802-3813.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Damping Estimation from Measured Random Vibration Response in Operation

Ibrahim A. Sever*

Vibration Engineering-ERTS
Rolls-Royce plc. Derby, DE24 BBJ, England
ibrahim.sever@rolls-royce.com

Key words: Damping estimation, Rice distribution, random excitation, non-stationary signals

Abstract

Reliable damping estimates are crucial for accurate forced response predictions when assessing vibration response of aero-engine components. This requires the solution of inverse problem for determination of dissipation parameters from measured data. In this paper a number of damping estimation methods applicable to random vibration data are presented. Methods applicable to stationary, as well as non-stationary, signals are studied. The latter is very important in real engine environment where natural frequencies often change as a function of time. An empirically derived new method, based on Rice distribution operating on probability distribution of zero-crossing intervals is also introduced. The applicability of the new method is demonstrated for various ranges of prevailing damping and modal densities.

1 Introduction

Harsh operating environment of aero engines often necessitates introduction of dissipation mechanisms to keep vibration response below harmful levels. These dissipation mechanisms may be introduced deliberately, such as in the case of under-platform dampers used in high pressure turbine stages, or may occur naturally, such as in the case of blade-disc contact. Whatever the mechanism, the analysts rely on accurate estimates of damping available to perform representative forced response calculations, which in turn are used to assess suitability of design from vibration view point. Therefore methods for solution of inverse problem to determine dissipation parameters from measured data are required.

The present work gives an account of various methods that were used to produce damping estimates from measured engine data. Given the complex nature of interactions, producing such estimates is often a challenge. The task is further complicated by the fact that in real engine measurements, upon which the said methods are operating, the exact excitation forces are unknown. This means that certain assumptions will have to be made about the character of excitation force before any estimates for damping can be produced. Adequacy of these assumptions is discussed.

2 Methods for Identification of Instantaneous Frequency and Amplitude

Some of the methods used in estimation of damping rely heavily on variation of natural frequency and amplitudes as a function of time. In cases where these variations are not taken into consideration it is important to justify the underlying assumption. And for methods that can account for these variations, they need to be identified with enough fidelity such that the techniques can be applied reliably. In this paper zero-crossing, nonlinear curve fitting and Hilbert transform methods are used for identification of instantaneous properties.

Of these, zero-crossing or level-crossing method is applicable to mono-harmonic signals. Therefore its use on real engine data is almost always preceded by band-pass filtering to isolate the mode of interest. This method is used for construction of Rice distributions used in Section 5.

In cases when isolation is not possible, nonlinear curve fitting can be applied, provided that a suitable model is selected. The method employed here for curve fitting uses Levenberg–Marquardt (L-M) algorithm [3]**.** Use of the L-M algorithm in this paper is based on the one given in [4] but modified for increased stability and speed. For signals that present local variations in time, nonlinear curve fitting can provide unique advantages in terms of resolution of both amplitude and frequency. However, curve fitting methods do not quite reveal instantaneous properties in the signal as these properties are estimated from a portion of the signal over a period of time. For such cases Hilbert transform methods are used. Use of Hilbert transform in structural dynamics application has gained prominence in recent years [6]. Here Hilbert transform has been used to identify signal instantaneous amplitude (i.e. envelope) and frequency as a function of time. Nonlinear curve fitting and Hilbert transform based methods are predominantly used in stationarisation of signals in Section 4.

3 Standard Damping Estimation

Power Spectral Density (PSD), $R_x(\omega)$, of a Single Degree of Freedom (SDOF) system to broadband random excitation can be expressed as the multiplication of its power transfer function, $|H(\omega)|^2$, and input PSD, $R_y(\omega)$, [1]. For a lumped-mass parameter SDOF system with mass m and stiffness k, $R_x(\omega)$ can be expressed as follows:

$$R_x(\omega) = \left[\frac{{}^{1}/_{k} \cdot {\omega_n}^4}{({\omega_n}^2 - \omega^2)^2 + \left(\frac{\omega_n \cdot \omega}{Q}\right)^2} \right] \cdot R_y(\omega) \qquad (1)$$

Here the first term in square brackets is $|H(\omega)|^2$ and since it is the power transfer function between output and input, it is in u^2/o^2 units, where u denotes output units (e.g. m, MPa, etc.) and o denotes input units (e.g. N). ω_n is the natural frequency and Q is the measure of damping, defined as the ratio of natural frequency to modal bandwidth and is therefore non-dimensional. Output PSD, $R_x(\omega)$, is given in units of u^2/Hz and likewise input PSD, $R_y(\omega)$, in o^2/Hz units. Since $R_y(\omega)$ can be considered constant (i.e. broadband random input), and, assuming time-invariant system properties, the whole of numerator in equation (1) can be expressed in terms of a constant; $C = 1/k \cdot {\omega_n}^4 \cdot R_y$. Given an PSD curve for a mode of interest, a nonlinear curve fitting algorithm such as L-M [3] can be used to fit equation (1) for identification of unknown quantities ω_n , Q, and, C (i.e. the numerator). The output PSD is obtained by successively segmenting the measurements and averaging PSD curves obtained in each segment. The theory of PSD calculations can be found in various good text books such as in [1].

The relation given in equation (1) relies on underlying excitation force to be of random broadband in nature. It is important that this is verified before any damping estimation is attempted. When excited by a random input, the mode of interest will act as a band-pass filter resulting in a normal distribution of instantaneous values. In this case the peak response amplitudes will follow the Rayleigh distribution

[8] given in equation (2). Therefore random input requirement of the excitation can be assessed by fitting equation (2) to probability distribution of the response peaks.

$$P(\delta;\sigma) = \frac{\delta}{\sigma^2} \cdot e^{-\frac{1}{2}\cdot\left(\frac{\delta}{\sigma}\right)^2} \qquad (2)$$

Figure 1 (a) shows two modes measured from a real engine test. The measurements are acquired on an external component of a civil engine using strain gauges under gradually increased engine speed conditions. Rayleigh distributions for both modes were constructed after narrow-band filtering of the individual modal responses. Figure 1 (b) shows an adequate degree of correlation whereby justifying use of standard damping estimation given in equation (1). However, the same cannot be said for Figure 1 (c). Upon close inspection it can be seen that a nearby strong engine order excitation introduces a bias in peak amplitude distribution thereby invalidating the underlying assumption for damping estimation. Therefore use of the method for this mode is not justified.

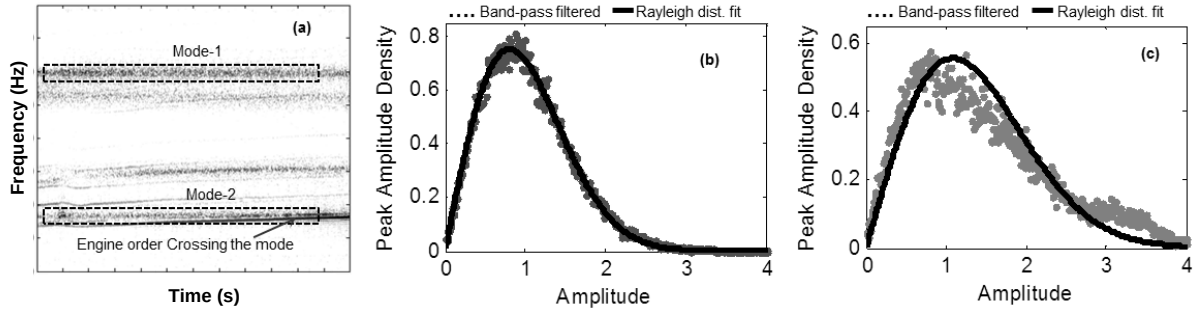


Figure 1. (a) Zmod measured from a real engine component , peak amplitude density and Rayleigh distribution fits for (b) Mode-1, and, (c) Mode-2 shown on the Zmod.

4 Damping Estimation from Non-Stationary Data

Standard damping estimation method given in Section 3 assumes that natural frequency of the mode for which the damping is being estimated is time invariant. However there are circumstances in which natural frequencies of engine components change sufficiently enough that this assumption is no longer valid. These changes may be due to temperature effects, centrifugal stiffening etc. When standard estimation method is applied to these cases, the effect will be spurious stretching of modal bandwidth which in turn will lead to overestimation of the available damping. The effect is schematically demonstrated in Figure 2.

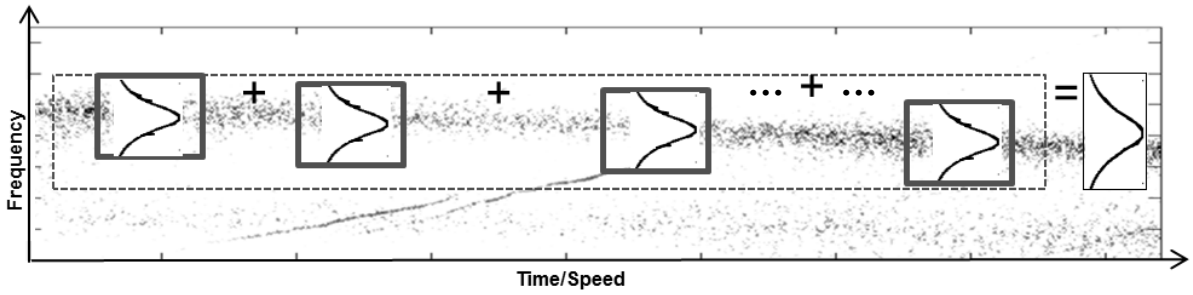


Figure 2. Time varying natural frequency of an engine component and illustration of error introduced in damping estimation via standard analysis technique.

Assuming that the effective bandwidth of the mode does not change with changing natural frequency, variation in the natural frequency can be compensated for to obtain reliable damping estimations. An effective method for achieving this is given in [2]. In essence the method constructs a carrier signal from the identified natural frequency variation and convolves this with the original signal to create a

component of vibration that is stationarised at an arbitrarily selected mean frequency. Assuming this frequency to be ω_s and identified natural frequency variation, $\omega_{nat}(t)$, a carrier signal can be created as follows:

$$S_C(t) = \sin\left(\int (\omega_s - \omega_{nat}(t)\,)dt\right) \qquad (3)$$

Here the integration in the argument of sine function computes the carrier phase. Given an original time varying signal of $S_O(t)$, a modulated signal $S(t)$ can be ccomputed as:

$$S(t) = S_O(t) \cdot S_C(t) \qquad (4)$$

The modulation process creates sum and difference frequencies, so-called upper and lower sidebands respectively. Stationarised response is contained in the upper sideband.

Since resultant stationarised signal has a resonant frequency higher than that of the original variation, higher Q values are estimated. This is because the resonant frequency is shifted up by modulation process but the mechanical bandwidth is not changed. When new resonant frequency is close to the mean of the original variation then the error introduced may be negligible. However as a good practice, the calculated value can be corrected by mean value of the original frequency variation.

Application of this method to a response measured on a real engine component due to random excitation in it is normal operation is given in Figure 3. Here the variation in natural frequency is evident in Zmod given in Figure 3 (a). Averaged PSD calculated for this case together with corresponding SDOF fit are given in Figure 3 (b) where a damping Q value of 27 is calculated. The variation in natural frequency is identified through methods summarised in Section 2 and it is given by the dashed line in Figure 3 (a). For stationarisation purposes mean frequency of the mode in question is centred at 850 Hz as showing Figure 3 (c) where time dependency of variation is evidently removed. Corresponding averaged PSD given in Figure 3 (d) now shows a much cleaner curve with the SDOF fit revealing a damping Q value of 80. In other terms estimate from non-stationary case suggests 3 times as much damping than there really is present, leading to false sense of security. Although variation in natural frequency is relatively modest with respect to mean value of natural frequency, it is significant compared with the bandwidth of the mode, which is what matters here.

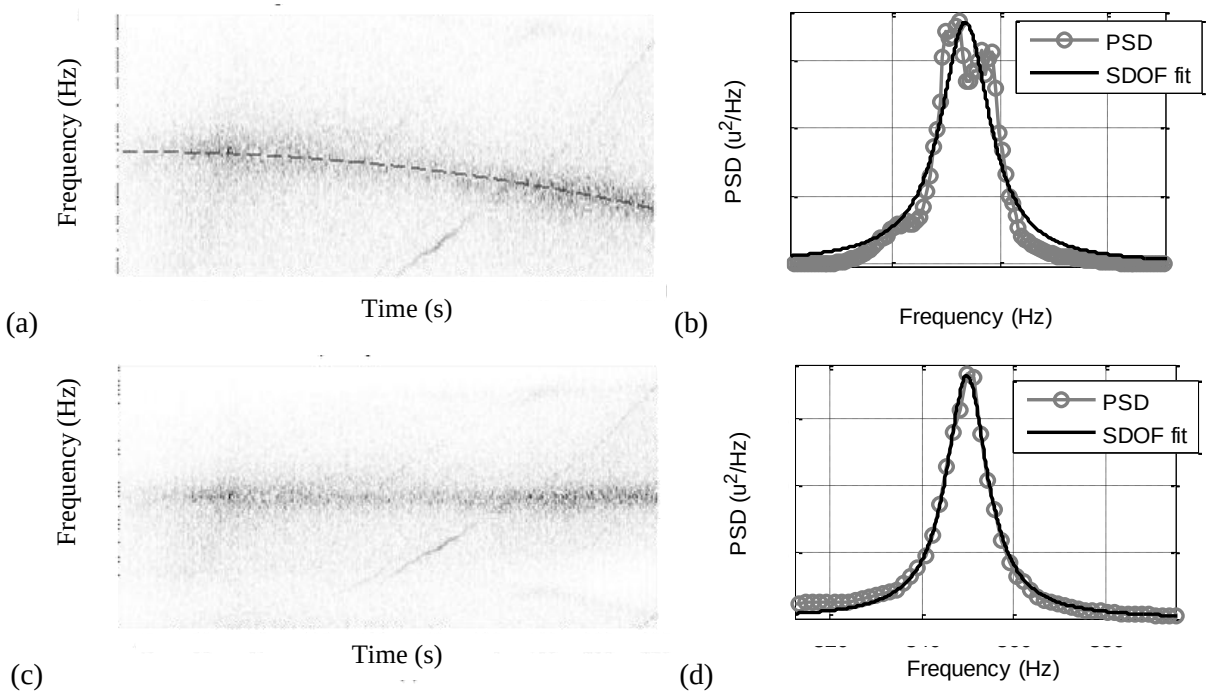


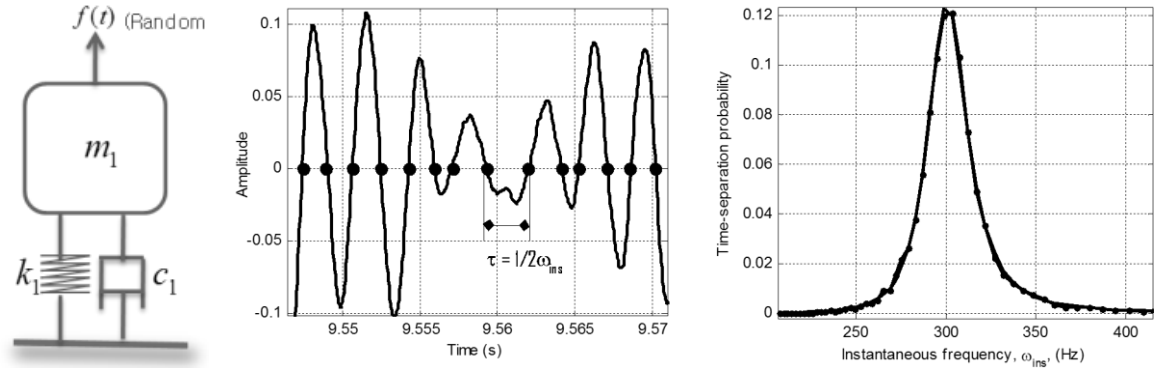
Figure 3. (a) Zmod showing a mode with varying natural frequency (dashed line), (b) best fit to its PSD, (c) the same mode after stationarisation, (d) best fit to its PSD after stationarisation.

5 Damping Estimation from Time-Separation Distribution

The connection between the bandwidth of an 'ideal' band pass (electrical) filter to the characteristics of the time separation distribution from the resulting narrowband signal has been made in the classic paper by Rice, [7]. Utilising Rice's original notation, for an ideal narrow band-pass filter the probability, p, that the distance between two successive zeros lies between τ and $\tau + \Delta\tau$ is approximately;

$$p = \frac{\Delta\tau}{2} \cdot \mathrm{a} \sqrt{\frac{1}{(1 + \mathrm{a}^2(\tau - \tau_{\mathrm{nat}})^2)^3}} \qquad (5)$$

where $\tau_{\mathrm{nat}} = \frac{1}{2 \cdot \omega_{\mathrm{nat}}}$ is half the period of natural frequency. For electrical systems, the bandwidth is approximated by the parameter value "a". Figure 4 (a) shows an SDOF mechanical system. The response of this system to random forcing is computed and locations of zero-crossings in the response are identified as shown in Figure 4 (b). Time-separation probability distribution of these crossing for natural frequency of 300 Hz, mass of 0.1 kg and a Q value of 100 is given by the dotted line in Figure 4 (c). Best fit to the given probability distribution using equation (5) and employing nonlinear curve fitting algorithm given in Section 2, is also plotted in Figure 4 (c), as a solid line. It is evident from the plot that the probability distribution is accurately described by equation (5).



(a) (b) (c)

Figure 4. (a) SDOF system with random excitation, (b) its response showing zero-crossings, and (c), distribution of instantaneous frequencies computed from these crossings [**...**Computed distribution, — Fitted distribution].

The contribution of this paper is its attempt to appropriate abovementioned concept for mechanical systems as a means of damping estimation. The main concept and empirical identification of bandwidth and damping from a given zero-crossing distribution are presented below. The adaptation of method and its application to simulated and real data, as well as its limitations and scope are also presented.

In an attempt to find a relationship for bandwidth and damping in case of mechanical systems from distribution given above, the SDOF model shown in Figure 4 (a) is used. The idea is to construct an association involving frequency of the mode in the narrowband response and parameter "a" empirically. For this purpose the SDOF model given above is simulated for a range (20-300) of damping Q values. The SDOF system is designed to have a natural frequency of 200 Hz and a lumped mass of 1 kg. The simulation range for Q values is kept wide enough to cover possible observations in practice. For each case the SDOF model is simulated to a random input for 20 seconds. This input is kept the same for all cases. Zero-crossing time distribution for each case is constructed. Once again, equation (5) is fitted to data. Variation of input Q value as a function of fit parameter "a" is plotted in

Figure 5 (a). The plot reveals a smooth nonlinear relationship which is quadratic in nature. In an attempt to describe this relationship better, input Q is plotted against a^2 in Figure 5 (b). The plot now demonstrates a near-perfect linear variation. Given this relationship, it becomes possible to "scale" the "a^2" term with a constant featuring natural frequency, such that the slope of this linear relationship is as close to 1 as possible (i.e. an estimate for Q value). A suitable relationship is found when a^2 term is multiplied by $\frac{{\tau_{nat}}^2}{\pi}$, as shown in Figure 5 (c). The slope of linear variation show in Figure 5 (c) is nearly 1. As a result the empirical description for estimation of Q value using time separation distribution resulting from a narrowband signal is defined in the following form:

$$Q_e \cong \frac{1}{\pi} \cdot a^2 \cdot {\tau_{nat}}^2 \qquad (6)$$

Since parameters a and τ_{nat} are obtained directly from fitting equation (5) to time-separation probability distribution, estimate for damping in the form of Q_e can readily be obtained.

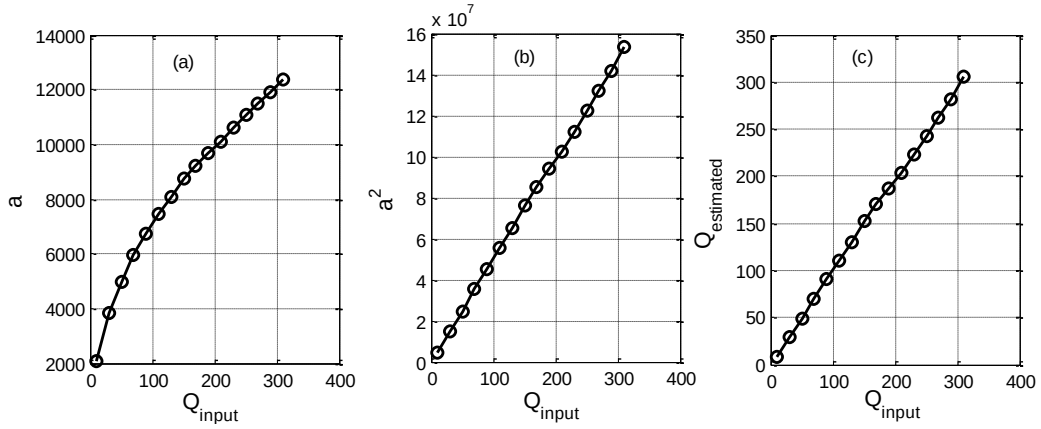


Figure 5. (a) Variation of parameter "a", (b) "a^2" with respect to input Q value, and, (c) Estimated Q value against input Q value.

5.1. Time-separation distribution of narrow-band filtered response

Although the applicability of the time-separation method is shown on an SDOF system, in practice acquired vibration responses will feature a multitude of harmonic content at different natural frequencies. For the method to work in these cases, total response will have to be band–pass filtered around the mode of interest. To study the effect of narrow-band filtering on time-separation method, the SDOF system introduced above is designed and used with following properties: $m_1 = 0.1\ \text{kg}, \omega = 2000\text{Hz}, Q = 100$. Its response to a random input is then simulated for a 20 seconds period. The response obtained is filtered using an equiripple FIR band-pass filter, [5], around the natural frequency. For both un-filtered and filtered responses time-separation probability distributions are obtained. Figure 6 shows the comparison of distributions before and after filtering. Also shown in the figure is the fitted distribution to the filtered data.

Figure 6 presents a surprising picture. The effect of filtering on distribution obtained is significant. This is unexpected as the bandwidth of the designed mechanical SDOF system is 3 Hz which is much smaller compared with filter bandwidth of 240 Hz used here. The reason for this is that when high-frequency component of response is filtered, the nature of zero-crossing intervals change significantly. Crossing time intervals which were much shorter when unfiltered (i.e. outliers in original distribution), are removed in case of filtered response and are moved towards the mean of the distribution.

An important observation from equation (5) is that, time-separation distribution is only a function of parameter "a" and system natural frequency. In other words, filtered response distribution is identical to the unfiltered distribution in characteristics but its parameter "a" is simply scaled by a constant. This is more clearly demonstrated in Figure 6 (b) where the response for a different SDOF system with a natural frequency of 600 Hz is filtered with a series of band-pass filters which are progressively getting narrow in the pass band. It is evident that kurtosis of the distribution increases as the harshness of the filters increase. Therefore, any scaling constant must be a function of filter characteristics and a correction factor will have to be identified in relation to unfiltered case for each filter individually. The question of whether such a correction factor can be derived in a robust manner such that it can be used in real cases with confidence is explored in the following sections.

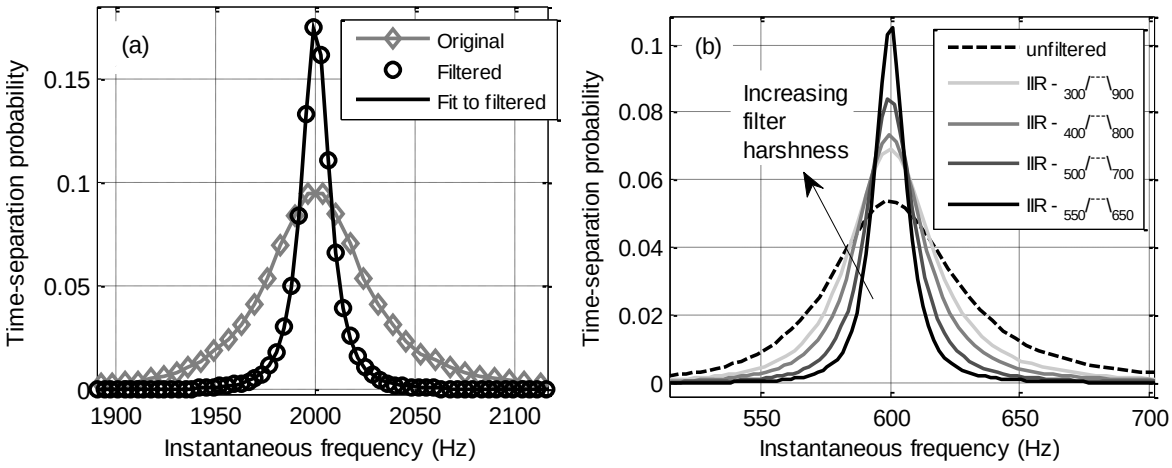


Figure 6. (a) Unfiltered and filtered time-separation probability distribution, (b) Effect of increased filter harshness on time-separation probability distribution.

5.2. Identification of filter correction factor

The process followed in correction factor identification for a given filter is as follows: First an SDOF system with desired natural frequency is designed where a rough estimate for expected damping value is assumed. Using time-separation damping estimation method damping from simulation of this system is estimated. Subsequently the simulated response is passed through the filter for which the correction factor is to be identified. Time-separation method is then applied to the filtered response and a new damping estimation is obtained. The correction factor for the filter at hand is then simply a ratio of these two damping estimates.

The process is applied to an SDOF system with ω_n=300 Hz, m=0.1 kg, and Q=100, simulated for 20 seconds at a sampling rate of 25 kHz. Correction factor is identified for band-pass FIR filter with following properties: First transition band of 250-270 Hz and second transition band of 330-350 Hz (i.e. pass-band=270-330 Hz), stop-band attenuation of 20dB, and, pass-band ripple of 0.1dB. For this case the response from simulation is first filtered with the given band-pass filter and then time-separation damping estimation is applied to the filtered time domain data. Damping value with applied correction factor is identified as 92.3. The deviation of this value from the original Q value of 100 is 7.7% and the deviation from Q value identified using time-separation for un-filtered response is 6.7%.

5.3. Dependency of correction factor on damping value

Table 1 provides a number of correction scenarios for the SDOF model given above with different Q values. Here the correction factors, C_f (the ratio of Q values after and before filtering), are identified for a range of Q values starting from 20 to 300 using the band-pass filter given above. For very low Q values deviations up to 20% are observed. As the Q values increase stability of correction factor also increases.

Given the evident stability of correction factor with varying Q values, the SDOF system is simulated for the Q of 100 and filtered and then corrected in turn by different C_f factors in Table 1 to quantify the dependency of Q value to choice of reference Q value. Estimated Q values (Q_C) and their deviation from original value (Q_O) of 100, and, estimate from unfiltered response (Q_{UE}) of 98.9 are given in Table 1. Estimates calculated using correction factors for Q values ranging from 50-300 are all within 10% of original design value. Therefore it is concluded that correction factor is reasonably insensitive to choice of Q value within this range.

$Q_{Reference}$	20	30	40	50	70	100	130	150	170	200	300
C_f	10.5	9.8	9.4	9.2	8.9	8.6	8.4	8.4	8.4	8.4	8.7
Q_C ($Q_{Filtered}/C_f$)	79.9	85.6	89.2	91.1	94.2	97.5	99.8	99.8	99.8	99.8	96.4
$(Q_C/Q_{UE})*100$	19.3	13.5	9.8	7.9	4.8	1.5	0.9	0.9	0.9	0.9	2.6
$(Q_C/Q_O)*100$	20.1	14.4	10.8	8.9	5.8	2.5	0.2	0.2	0.2	0.2	3.6

Table 1. Q values estimated by application of correction factors identified from simulations of an SDOF system with different Q values.

5.4. Effect of choice of SDOF system on filter correction factor

Effect of filtering on zero-crossing time-distribution was studied earlier with an SDOF system. A detailed examination was carried out to see if correction factors obtained for different SDOF systems with the same natural frequency would yield the same damping value. To this end, mass value was gradually varied from 0.1 kg to 5 kg and stiffness value adjusted accordingly for a natural frequency of 300 Hz. Correction factors obtained from different SDOF systems showed variations smaller than 1%. Therefore it is concluded that the choice of SDOF system is not a critical parameter in determination of the correction factor.

5.5. Effect of modal density on filter correction factor

Effect of proximity of neighbouring/bracketing modes on the filtering correction factors was explored. To this end, a number of 3-DOF systems were designed. Natural frequency separation for these systems were gradually decreased (i.e. modal density increased) to simulate close modes scenarios. Each system was then simulated and their responses at 1st DOF were filtered and passed through time separation damping estimation algorithm. Properties of these systems, together with identified correction coefficients and filter properties are given in Table 2. Also given in this table are an equivalent SDOF system and its correction coefficient as reference. It is evident from the table that when the modes of the system are well separated and filter cut-off frequencies are sufficiently far away from neighbouring modes, correction coefficients for 3-DOF system are very close to that of reference SDOF system. However, as the cut off frequencies get closer to bracketing mode frequencies, as is the case in System-4, significant deviations are observed.

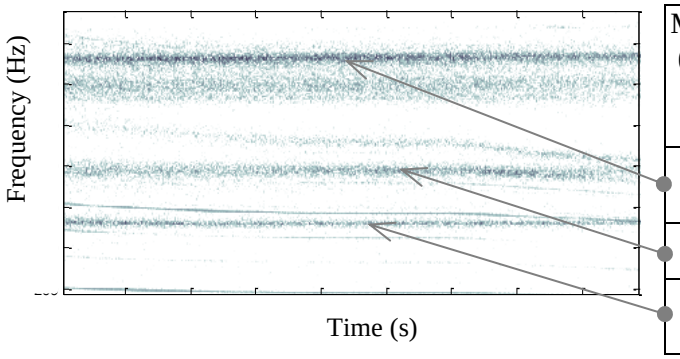
The same process is repeated with a much harsher band-pass filter to see if the C_f from all 3-DOF systems can be aligned with that from the SDOF system. The results for this case are given in "Filter-2" column of Table 2. As the new filter is now much harsher, the resultant filtering coefficients are much larger. Nevertheless, all systems now result in similar correction coefficients with maximum deviation from reference SDOF system being less than 9%.

	m1,m2,m3 (kg)	k1,k2,k3 (N/m)	c1,c2,c3 (kg/s)	ω1,ω2,ω3 (Hz)	Q1,Q2,Q3	Filter 1		Filter 2	
						$Q2_C$	C_f	$Q2_C$	C_f
Sys-1	0.788381 0.455614 1.19597	1935330 1599560 1921130	10 10 10	100 300 500	285.2 99.8 55.6	1772.3	17.8	3622.2	36.3
Sys-2	1.19312 0.578852 0.0174768	4020170 1759710 60468.6	0.34 0.34 0.34	200 300 400	2693.4 99.9 1345.6	1737.6	17.4	3583.5	35.9
Sys-3	2.51416 0.134519 0.0064339	10122900 461284 19756.6	0.202 0.202 0.202	250 300 350	146.0 100.6 435.2	1630.0	16.2	3541.9	35.2
Sys-4	4.23869 0.0877328 0.00167354	15444700 310571 5703.85	0.058 0.058 0.058	270 300 330	183.0 100.2 261.7	1181.7	11.8	3322.0	33.2
SDOF	1	3553060	18.8496	300	100	1763.3	17.6	3640.8	36.4

Table 2. Effect of modal density on identification of filter coefficient [IIR Filter-1: Fs=25000Hz, Fc1=280 Hz, Fc2=320 Hz, Order=4, IIR Filter-2: Fs=25000Hz, Fc1=290 Hz, Fc2=310 Hz, Order=3]

5.6. Application of time-separation damping estimation method to real data

The data set used in this section is acquired from a deceleration manoeuvre engine test. The raw time-domain data is processed at sampling rate of 25 kHz. A Zmod for one of the channels from this manoeuvre is given in Figure 7. The modes at 765 Hz, 500 Hz and 360 Hz, as indicated on the Zmod, are selected for analysis. First 2 modes are analysed for 30-80 seconds interval, and the mode at 360 Hz is analysed for 50-90 seconds interval as they show reasonably stationary behaviour in these intervals, respectively. Table shown in Figure 7 gives a summary of filter parameters used and correction factors obtained. Maximum difference in damping values estimated using time-separation and PSD based methods is 10%.



Mode (Hz)	Filter band (Hz)	Filter Order	C_f	Q_{TS}	Q_{PSD}	%Δ from Q_{PSD}
765	720-810	4	20.3	68	76	10
500	450-550	4	11.5	27	30	10
360	340-380	4	20.9	67	64	4.5

Figure 7. Zmod showing target modes, and, comparison of damping estimations from time-separation distribution and standard PSD SDOF curve fit methods.

5.7. Sensitivity of time-separation method to data length

Sensitivity of time separation damping estimation method to signal length used in construction of time-separation distribution is also investigated. For this purpose an SDOF system with a natural frequency of 300 Hz is excited by a random input and simulated for a 20 second period. Figure 8 shows damping estimates from 1, 2, 3 and 4 seconds-long segments taken from this simulation. Also shown on these plots is the original distribution for whole 20 seconds of simulation. It is evident from distribution and subsequent damping estimates that 1 second segment does not capture the process accurately. However, as the segment length is increased distribution converges very rapidly to original distribution with 4 seconds segment providing near full convergence.

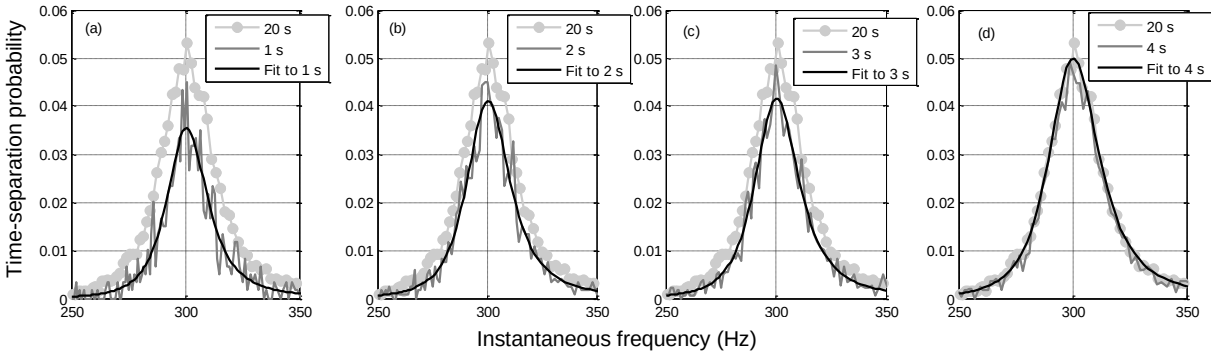


Figure 8. Sensitivity of Time Separation damping estimation methods to time-domain data length used.

5.8. Effect of signal length on damping estimation – real case

Dependency of damping estimations to length of signal used is investigated on the same engine data used in Section 5.6. For this purpose the mode at 765 Hz is used. Data samples of 10s, 5s, 2s, 1s and 0.5s are evaluated by time-separation distribution and the standard PSD curve fitting methods. Estimates for natural frequencies and damping values are given in Table 3. As the signal lengths get shorter, PSD based method deviates from expected value significantly whilst time-separation estimates are always within 10%.

Sample size	ω_{TS}(Hz)	Q_{TS}	% dev from Q_{Exp}	ω_{PSD}(Hz)	Q_{PSD}	% dev from Q_{Exp}
10s	763.1	85	-7.6	764.5	84	-6.3
5s	763.4	86	-8.9	764.8	94	-19.0
2s	762.6	75	5.1	763.8	85	-7.6
1s	762.5	73	7.6	762.7	49	38.0
0.5s	761.1	76	3.8	759.8	148	-87.3

Table 3. Effect of signal length on damping estimation by PSD and time-separation.

6 Discussion and Concluding Remarks

Obtaining reliable damping estimates from real engine running measurements is very difficult. Responses acquired are often non-stationary due time-dependent running conditions. In addition the forcing functions are complicated and not precisely known. In the absence of input force data, estimates obtained are a function of response amplitude at which they are extracted and are only valid

at those amplitudes. The argument here is that this may be acceptable from practical assessment point of view if these amplitudes can be assumed to be representative of the normal operating conditions. Nevertheless, the absence of precise forcing measurements requires that certain assumptions will have to be made about the character of the forcing to enable calculation of damping estimates. In this paper the focus is on the broadband random vibration measurements. As such necessary tools are introduced to ensure the random nature of excitation before the relevant methods are applied.

When the properties of the component being analysed are stationary, and the vibration modes are reasonably well isolated, reasonable damping estimates can be obtained by approximating the mode for which the damping estimate is sought by a SDOF model. Application of this method to simulated, as well as to real engine data has been successfully demonstrated. However, as referred to earlier, engine operating environment is often non-stationary with component natural frequencies changing in time due to various effects such as temperature and rotational speed. In such cases application of SDOF model without taking these variations into account yield optimistic values, suggesting more damping than is actually present. Obviously this is not a desirable situation. It is demonstrated on real engine measurements that using stationarisation methods introduced in earlier studies, it is possible to account for the variations, and that estimates can be out by a factor of 3 if stationarisation is not performed. It must be noted that the critical factor here is the total variation with respect to bandwidth of the mode in question, and not the natural frequency. As the success of stationarisation depends on accuracy with which natural frequency variations can be identified, various means of achieving this are also introduced.

Majority of this paper is dedicated to a new damping estimation method operating on broadband random vibration response and utilising Rice distribution of zero-crossing time intervals. This method is particularly important since it provides estimates that are independent of the way in which PSD based estimates given above are obtained. As such it is very useful in verifying estimates and gaining confidence in damping values produced. Nevertheless it is application is less straight forward. The method is derived empirically for a practical range of damping values. Care should be taken to ensure it is being applied within the boundaries of its scope. As the method is applicable to mono harmonic signals, it is almost always preceded by narrow-band filtering to isolate the mode of interest. Narrowband filtering is shown to have significant effects on zero-crossing distribution; however, a means of accounting for this effect is also given. The performance of the method on real data is compared with that of PSD-based methods and a very good correlation is obtained. It is shown to outperform PSD-based methods when suitable signal durations are short, providing reasonable estimates for signals that are up to an order of magnitude shorter. The work presented here amounts to first attempt made at making this method work. More work is needed to extend it applicability to a much wider damping range (i.e. very low and very high) as well as reducing its sensitivity to narrow-band filtering.

Acknowledgement

The author is grateful to Rolls-Royce plc. and TSB for providing the financial support under the SILOET project and for giving permission to publish this work. The author is particularly grateful to Dr. H. Kurt-Elli whose original works have provided the foundation for this publication.

References

[1]. **Newland, D. E.,** *An Introduction to Random Vibrations, Spectral & Wavelet Analysis*; Dover Publications Inc., 3rd edition, 1 Jan 2009.

[2]. **Kurt-Elli, H.,** *Damping estimation from engine data with varying natural frequencies*, IMAC XXIII, 2005.

[3]. **Marquardt , D.** , *An Algorithm for Least-Squares Estimation of Nonlinear Parameters.* SIAM Journal on Applied Mathematics 11 (2): 431–441, 1963

[4]. **Press W.H., Flannery, B.P., Teukolsky S.A., Vetterling, W. T.,** *Numerical recipes in C*, Cambridge University Press, 1989
[5]. **Rorabaugh, C.,** *Digital Filter Designer's Handbook*, McGraw Hill, 1993
[6]. **Feldman, M.,** *Hilbert Transform Application in Mechanical Vibration*, John Wiley and Sons Ltd., Chichester, United Kingdom, 2011.
[7]. **Rice, O.S.**, *Mathematical Analysis of Random Noise*, Bell System Technical Journal, 24 (1): p46-156, Jan. 1945
[8]. **Thomson, W.,** *Theory of Vibration with Applications*, Second Edition, Prentice-Hall, New Jersey, 1981.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Identification of a finite number of defects in an elastic body using boundary data

Efim I. Shifrin*, Pavel S. Shushpannikov

A.Yu. Ishlinsky Institute for Problems in Mechanics RAS
Prosp. Vernadskogo 101-1, 119526 Moscow, Russia
e-mail: shifrin@ipmnet.ru

Key words: identification, anisotropic elasticity, multiple defects, reciprocity gap functional, boundary data

Abstract
A problem of identification of a finite number of embedded defects (inclusions, cavities, cracks) in an anisotropic linear elastic body by means of boundary data obtained in a single static test is considered. It is assumed that the defects are small and well-separated. If the defect is a cavity or a crack we suppose that its boundary is unloaded. If the defect is an inclusion we suppose complete bonding between the matrix and inclusion. A method for determination of the number of defects and positions of their centres using the available boundary data is developed. If the defects have an ellipsoidal shape the developed method enables also to determine their sizes and orientations. Numerical example illustrating the efficiency of the developed method is considered.

1 Introduction

An analytical-numerical method for identification of a single ellipsoidal defect in a linear elastic body was developed by the authors in a series of publications. In particular, a problem of identification of a single ellipsoidal defect in an isotropic and anisotropic linear elastic solid was solved in [1,2] and [3,4], respectively. A method for identification of multiple defects in an isotropic linear elastic body was developed in [5]. The developed methods were based on the application of the reciprocity gap functional (RGF) method first applied to inverse elastostatic problems in [6]. In the present paper we combine the ideas used in the publications [1-5] to develop a method for identification multiple ellipsoidal defects in a linear elastic, anisotropic body.
The paper is organized as follows. The mathematical formulation of the problem is given in Section 2. To identify the multiple defects we determine first their number and projections of their centers on an arbitrary plane. The problem is solved by means of the RGF method with the use of the regular elastic fields of special type. Construction of the regular elastic fields, reduction of the initially considered 3d inverse problem to the 2d problem of identification of the projections of defects centers on an arbitrary plane and development of the algorithm for solving the 2d problem are presented in Section 3. For

identification of the sizes and orientations of ellipsoidal defects it is necessary to use some additional regular elastic fields. Construction of such fields and their application for identification of ellipsoidal defects is given in Section 4. In Section 5 a numerical example illustrating efficiency of the developed method is considered.

2 Statement of the problem

Let $V \subset R^3$ be a bounded domain with a boundary ∂V. $G_k \subset V$, $k = 1, 2, \cdots, n$ are small, simply connected subdomains. We suppose that $\overline{G}_i \cap \overline{G}_j = \varnothing$, $i \neq j$, $\overline{G} = \bigcup_{k=1}^{n} \overline{G}_k \subset V$, where $\overline{G}_k$ is a closure of the subdomain G_k. Let us suppose that an anisotropic linear elastic body with the elastic moduli C_{ijkl}^M occupies the domain $\Omega = V \setminus \overline{G}$. The defects G_k can be cavities or inclusions (rigid or linear elastic). If G_k is a cavity we suppose that its boundary ∂G_k is unloaded. If G_k is an inclusion, it is supposed complete bonding between the matrix and inclusion. We assume that typical sizes of the defects have the same order. Denote the typical size l. Assume also that the typical distances between the defects have the same order and denote the typical distance L. We assume that the defects are small in the following sense

$$l \ll L \tag{1}$$

Let us introduce Cartesian coordinates $Ox_1x_2x_3$. We suppose that the loads $\mathbf{t}^d = \left(t_1^d, t_2^d, t_3^d\right)$ and displacements $\mathbf{u}^d = \left(u_1^d, u_2^d, u_3^d\right)$ are measured on ∂V in a single static test. We will mark with the superscript d the stress-strain state in the body Ω: σ_{ij}^d is the stress tensor, e_{ij}^d is the strain tensor and $\mathbf{u}^d$ is the displacement vector, $t_i^d = \sigma_{ij}^d n_j$, where $\mathbf{n} = \left(n_1, n_2, n_3\right)$ is a unit outward normal to the boundary ∂V and convention of summation for repeated indices is used. Below we will suppose that the defects are linear elastic inclusions. The cases of cavities and rigid inclusions can be considered as limit cases when the elastic moduli tend to zero or infinity, respectively. The stress-strain state in the inclusion G_k we will mark with the superscript Ik (σ_{ij}^{Ik}, e_{ij}^{Ik}, $\mathbf{u}^{Ik} = \left(u_1^{Ik}, u_2^{Ik}, u_3^{Ik}\right)$ are the stress tensor, the strain tensor and the displacement vector, respectively). The elastic moduli of the inclusion G_k we denote by C_{ijpl}^{Ik}.

According to our suppositions the following equalities are valid for $x = \left(x_1, x_2, x_3\right) \in \Omega$:

$$C_{ijpl}^M u_{p,lj}^d = 0 \tag{2}$$

The elastic field with the superscript Ik satisfies in the domain G_k the equations analogical to Eq. (2) with the replacement of the values C_{ijpl}^M by the values C_{ijpl}^{Ik}. The conditions of complete bonding between the matrix and inclusion G_k have the following form:

$$\mathbf{u}^{Ik}\left(x\right) = \mathbf{u}^d\left(x\right), \quad \sigma_{ij}^{Ik}\left(x\right) N_j\left(x\right) = \sigma_{ij}^d\left(x\right) N_j\left(x\right), \quad x \in \partial G_k \tag{3}$$

Here $\mathbf{N}\left(x\right) = \left(N_1\left(x\right), N_2\left(x\right), N_3\left(x\right)\right)$ is a unit normal to the boundary ∂G_k at the point x.

We will call the elastic fields in the body V without defects as regular elastic fields and mark by a superscript r (σ_{ij}^r, e_{ij}^r, $\mathbf{u}^r = \left(u_1^r, u_2^r, u_3^r\right)$ are the stress tensor, the strain tensor and the displacement vector, respectively). The regular elastic fields satisfy the Eq. (2) in the domain V.

The RGF, depending on two stress states with superscripts d and r, is defined as follows:

$$RG\left(d,r\right) = \int_{\partial V} \left(t_i^d u_i^r - t_i^r u_i^d\right) dS\,, \qquad t_i^r = \sigma_{ij}^r n_j \tag{4}$$

The problem is to reconstruct the defects G_k using the known loads $\mathbf{t}^d$ and displacements $\mathbf{u}^d$ on the boundary ∂V. Because the vector-functions $\mathbf{t}^d$ and $\mathbf{u}^d$ are known, the values $RG\left(d,r\right)$ can be calculated for all regular elastic fields r. So, the problem will be solved if we express the parameters of the domains G_k by means of the values $RG\left(d,r\right)$. According to the results [1-5] the expression (4) can be written in the following form:

$$RG\left(d,r\right) = \sum_{k=1}^{n} \int_{G_k} \Delta\sigma_{ij}^{Ik} e_{ij}^r \, dx \tag{5}$$

where $\Delta\sigma_{ij}^{Ik} = \sigma_{ij}^{Ik} - \bar{\sigma}_{ij}^{Ik}$, $\bar{\sigma}_{ij}^{Ik}$ are the stresses corresponding to the strains e_{ij}^{Ik} in the material with the elastic moduli C_{ijkl}^M.

Let us denote the centers of the defects G_k by $x^k = \left(x_1^k, x_2^k, x_3^k\right)$ and the volumes of the domains G_k by $\left|G_k\right|$. Consider a regular elastic field in the body V without defects subjected to the loads $\mathbf{t}^d$ on the boundary ∂V. The elastic field we will mark by a superscript dr. Because the defects are small, we suppose also that the stress state in the defect G_k is close to the stress state in the inclusion G_k located in an infinite elastic solid and subjected to the constant stresses $\sigma_{ij}^{dr}\left(x^k\right)$ at the infinity. It follows from the supposition and the Eshelby results [7,8] that the stresses σ_{ij}^{Ik} are approximately constant in the ellipsoidal inclusion G_k. Finally, we assume that the values $\max\limits_{i,j}\left|\sigma_{ij}^{Ik}\right|$ have the same order for different G_k.

3 Reduction of the problem to 2d problem of the centers of defects projections identification

According to our suppositions, formulated in the preceding Section, we will approximate the values of the RGF by the principal term of the asymptotic expansion of the Eq. (5) provided that $l / L \to 0$

$$RG\left(d,r\right) \approx \sum_{k=1}^{n} \Delta\sigma_{ij}^{Ik}\left(x^k\right) e_{ij}^r\left(x^k\right) \left|G_k\right| \tag{6}$$

Consider, for example, projections of the defects on the plane x_1x_2. To determine projections of the defects centers on the plane x_1x_2 we will use the regular elastic fields depending only on the coordinates x_1 and x_2. Consider a regular elastic field $\mathbf{u}^r\left(x_1,x_2\right) = \left(u_1^r, u_2^r, u_3^r\right)$. Let us search for the functions $u_p^r\left(x_1,x_2\right)$ in the form $u_p^r\left(x_1,x_2\right) = f_p\left(x_1 + sx_2\right)$.

Because the functions u_p^r do not depend on the coordinate x_3, the elasticity equations have the form

$$C_{i\alpha p\beta}^M u_{p,\beta\alpha}^r = 0, \;\; \alpha = 1,2, \;\; \beta = 1,2 \tag{7}$$

It follows from the Eq. (7)

$$m_{ip}(s)\, f_p'' = 0\,, \quad m_{ip}(s) = C^M_{i1p1} + \left(C^M_{i1p2} + C^M_{i2p1}\right)s + C^M_{i2p2}s^2\,, \quad i = 1,2,3, \quad p = 1,2,3 \tag{8}$$

Consider the matrix $\mathbf{M}(s) = \left(m_{ip}(s)\right)$. Eq. (8) has a nonzero solution if and only if the following condition is valid:

$$\det\left(\mathbf{M}(s)\right) = 0 \tag{9}$$

This equality leads to an algebraic equation of sixth order relative to s. Let us suppose for simplicity that the roots of the equation are simple. According to [9] the imaginary parts of the roots are not zero. Because the coefficients of the algebraic equation are real, the roots have the following form:

$$s_1 = \alpha_1 + i\beta_1\,, \quad s_2 = \alpha_2 + i\beta_2\,, \quad s_3 = \alpha_3 + i\beta_3\,, \quad s_4 = \overline{s}_1\,, \quad s_5 = \overline{s}_2\,, \quad s_6 = \overline{s}_3\,, \quad \beta_j > 0, \quad j = 1,2,3 \tag{10}$$

The over-bar denotes complex conjugation.

Because the roots of the Eq. (9) are simple, the space of the solutions of the system of equations $m_{ip}\left(s_j\right) f_p'' = 0$, for $j = 1,2,3$, is one-dimensional. Let $\gamma_j = \left(\gamma_{j1}, \gamma_{j2}, \gamma_{j3}\right)^T$ be a normalized solution of the equations. The superscript T denotes transposition of a matrix. Let $g\left(x_1 + s_j x_2\right)$ be an arbitrary smooth function. It follows from the Eqs. (8)-(10) that vector-function $u^{r_j} = \mathrm{Re}\left(\gamma_j g\left(z_j\right)\right)$ is a regular elastic field. Here $z_j = x_1 + s_j x_2$. The strain tensor corresponding to the elastic field is as follows

$$e^{r_j}_{11}\left(z_j\right) = \mathrm{Re}\left(\gamma_{j1}\varphi\left(z_j\right)\right), \quad e^{r_j}_{22}\left(z_j\right) = \mathrm{Re}\left(\gamma_{j2}s_j\varphi\left(z_j\right)\right), \quad e^{r_j}_{12}\left(z_j\right) = \frac{1}{2}\mathrm{Re}\left[\left(\gamma_{j1}s_j + \gamma_{j2}\right)\varphi\left(z_j\right)\right]$$

$$e^{r_j}_{13}\left(z_j\right) = \frac{1}{2}\mathrm{Re}\left(\gamma_{j3}\varphi\left(z_j\right)\right), \quad e^{r_j}_{23}\left(z_j\right) = \frac{1}{2}\mathrm{Re}\left(\gamma_{j3}s_j\varphi\left(z_j\right)\right), \quad e^{r_j}_{33}\left(z_j\right) = 0 \tag{11}$$

where $\varphi\left(z_j\right) = g'\left(z_j\right)$.

Let us construct another regular field $u^{\rho_j} = \mathrm{Re}\left(i\gamma_j g\left(z_j\right)\right)$. The strain tensor corresponding to the elastic field is as follows

$$e^{\rho_j}_{11}\left(z_j\right) = -\mathrm{Im}\left(\gamma_{j1}\varphi\left(z_j\right)\right), \quad e^{\rho_j}_{22}\left(z_j\right) = -\mathrm{Im}\left(\gamma_{j2}s_j\varphi\left(z_j\right)\right), \quad e^{\rho_j}_{12}\left(z_j\right) = -\frac{1}{2}\mathrm{Im}\left[\left(\gamma_{j1}s_j + \gamma_{j2}\right)\varphi\left(z_j\right)\right]$$

$$e^{\rho_j}_{13}\left(z_j\right) = -\frac{1}{2}\mathrm{Im}\left(\gamma_{j3}\varphi\left(z_j\right)\right), \quad e^{\rho_j}_{23}\left(z_j\right) = -\frac{1}{2}\mathrm{Im}\left(\gamma_{j3}s_j\varphi\left(z_j\right)\right), \quad e^{\rho_j}_{33}\left(z_j\right) = 0 \tag{12}$$

It follows from the Eqs. (6), (11) and (12)

$$RG\left(d, r_j\right) - iRG\left(d, \rho_j\right) = \sum_{k=1}^{n} A_{kj}\varphi\left(z_{jk}\right), \quad z_{jk} = x_1^k + s_j x_2^k \tag{13}$$

$$A_{kj} = \left[\Delta\sigma^{Ik}_{11}\gamma_{j1} + \Delta\sigma^{Ik}_{22}s_j\gamma_{j2} + \Delta\sigma^{Ik}_{12}\left(s_j\gamma_{j1} + \gamma_{j2}\right) + \Delta\sigma^{Ik}_{13}\gamma_{j3} + \Delta\sigma^{Ik}_{23}s_j\gamma_{j3}\right]\left|G_k\right| \tag{14}$$

Let us take $\varphi\left(z_j\right) = \varphi_p\left(z_j\right) = \left(z_j / L\right)^p = w_j^p$, $p = 0,1,2,\cdots$, the value L was introduced in Eq. (1). The regular elastic fields corresponding to the function $\varphi_p\left(z_j\right)$ we denote by r_{jp} and ρ_{jp}. Eq. (13) for these regular elastic fields has the following form

$$\sum_{k=1}^{n} A_{kj} w_{jk}^p = b_{jp}\,, \quad w_{jk} = \frac{z_{jk}}{L}\,, \quad b_{jp} = RG\left(d, r_{jp}\right) - iRG\left(d, \rho_{jp}\right), \quad p = 0,1,2,\cdots \tag{15}$$

Eq. (15) coincide with the equations arising in the problem of simple poles of a meromorphic function identification, see for example [10,11]. In [5], where an isotropic case was considered, the method, proposed in [11], was used to determine the number of defects and projections of their centers. The same method we use here. We remind briefly the main ideas of the method below simply for the sake of making the paper self-contained. Let us first suppose that we know the number of the defects n. Consider a polynomial

$$P_{nj}(w)=\prod_{k=1}^{n}\left(w-w_{jk}\right)=w^{n}+\sum_{m=0}^{n-1}q_{jm}w^{m} \tag{16}$$

Here $\deg P_{nj}(w)=n$ and w_{jk}, $k=1,2,\cdots,n$ are the roots of the polynomial $P_{nj}(w)$, q_{jm} are unknown coefficients.

Using Eq. (16) it is possible to obtain a system of linear algebraic equations relative to the coefficients q_{jm}. It follows from the Eqs. (15) and (16)

$$b_{j\,p+n}+\sum_{m=0}^{n-1}b_{j\,p+m}q_{jm}=0,\quad p=0,1,2,\cdots,n-1 \tag{17}$$

Eq. (17) form a system of linear algebraic equations relative to unknowns q_{jm}, $m=0,1,2,\cdots,n-1$. After determination of the values q_{jm} it is possible to construct the polynomial $P_{nj}(w)$ and to find its roots. Thus, the projections of the defects on the plane x_1x_2 are found. After determination of the values w_{jk} we consider a system of Eq. (15) for $m=0,1,2,\cdots,n-1$. It is a system of linear algebraic equations relative to A_{kj}. So, the values w_{jk} and A_{kj} are determined. Here we suppose that $A_{kj}\neq 0$.

Usually we do not know the number of defects n, but we can suppose that we know an upper bound of the number. Let us suppose that we know that $n\le N$. In this case applying the considered procedure for a polynomial $P_{Nj}(w)$, constructed according to Eq. (16), we obtain the roots of the polynomial w_{jk} and coefficients A_{kj}, $k=1,2,\cdots,N$. Among the obtained values w_{jk} there are some roots corresponding to the defects projections and some spurious roots. In [5] there were proposed the following criteria for excluding the spurious roots: 1) some of the projections of defects corresponding to spurious roots are located outside of the body projection; 2) the values $\left|A_{kj}\right|$ corresponding to spurious roots w_{jk} are small relative to the values $\left|A_{kj}\right|$ corresponding to projections of real defects; 3) the projections of defects corresponding to spurious roots w_{jk} are not stable relative to the chosen value of the upper bound N and can change significantly when various values N are considered; 4) if projections on several planes are considered then it is possible to see that projections determined by spurious roots are not projections of some points in R^3.

In the anisotropic case, considered in the paper, it is possible to add one additional criterion. 5) The proposed procedure can be fulfilled for all values s_j, $j=1,2,3$. The coordinates $\left(x_{1j}^{k},x_{2j}^{k}\right)$, obtained by means of the values w_{jk}, correspond to the projections of real defects only in case when they are close to each other for different $j=1,2,3$.

The value of w_{jk} is excluded if it should be deleted according at least one of the criteria.

Let us suppose that the spurious values of w_{jk} are excluded and we have found points $\left(x_1^k,x_2^k\right)$ corresponding to the defects projections. The number of the points can exceed the number of projections of real defects because several points can correspond to one defect, see for example [12,5]. To determine the exact number of defects projections it is possible to use Eq. (1). So, if we obtain, for example, three points 1, 2 and 3 and the distance between the points 1 and 2 is much less than the

distances between the points 1, 3 and 2, 3 then we can suppose that the points 1 and 2 correspond to the projection of the same defect. After determination of the number of defects we repeat the described above procedure for the obtained number of defects. As a result, we obtain points located close enough to the projections of defects centers. Numerical example confirming that the use of the proposed algorithm enables to determine the number of defects and their centers is considered in the Section 5.

4 Identification of small ellipsoidal defects

Let us suppose now that the defects G_k are ellipsoidal inclusions. To determine the geometrical parameters of the inclusions (the magnitudes and directions of their axes) we will use an approach developed in [1] and applied to reconstruction of a single ellipsoidal defect in anisotropic body in [3,4]. According to our suppositions the stress state in the defect G_k is approximately constant and close to the stress state in the ellipsoidal inclusion G_k located in an infinite elastic solid and subjected to the stresses $\sigma_{ij}^{dr}\left(x^k\right)$ at the infinity. For definiteness, let us find the geometrical parameters of the defect G_1. To determine the geometrical parameters of the defect G_1, we construct a regular elastic field so that the contribution of the first term in the sum on the right side of the Eq. (5) was significantly greater than that of the remaining terms. First introduce Cartesian coordinates with the origin in the center of the defect G_1: $x_i = \xi_i + x_i^1$, $i = 1,2,3$. Denote coordinates of other defects centers in the coordinate system by $\left(\xi_1^k, \xi_2^k, \xi_3^k\right)$, $k = 2,3,\cdots,n$. Define holomorphic functions, $\chi_{mj}\left(\zeta_j\right) = \dfrac{1}{L^{m(n-1)}}\prod_{k=2}^{n}\left(\zeta_j - \zeta_{kj}\right)^m$, $m \geq 3$, $\zeta_j = \xi_1 + s_j\xi_2$, $\zeta_{kj} = \xi_1^k + s_j\xi_2^k$. Denote

$$\chi_{mj}(0) = \prod_{k=2}^{n}\left(-\frac{\zeta_{kj}}{L}\right)^m = P_{mj} \tag{18}$$

Consider a function

$$\varphi_{mj}^*\left(\zeta_j\right) = \left(\frac{\zeta_j}{L}\right)^2 \chi_{mj}\left(\zeta_j\right) \tag{19}$$

Let us consider regular elastic fields r_{mj}^* and ρ_{mj}^* constructed according to Eqs. (11) and (12) by means of the function $\varphi_{mj}^*\left(\zeta_j\right)$. It follows from the Eqs. (5), (11), (12), (14) and suppositions formulated in Section 2

$$RG\left(d, r_{mj}^*\right) - iRG\left(d, \rho_{mj}^*\right) \approx \sum_{k=1}^{n}\frac{A_{kj}}{\left|G_k\right|}\int_{G_k}\varphi_{mj}^*\left(\zeta_j\right)d\xi, \quad d\xi = d\xi_1 d\xi_2 d\xi_3 \tag{20}$$

Using arguments similar to those given in [5], one can show that for sufficiently large m, the contribution of the first term in Eq. (20) is much greater than the sum of contributions of the remaining terms. Thus we have the following equation

$$RG\left(d, r_{mj}^*\right) - iRG\left(d, \rho_{mj}^*\right) \approx \frac{A_{k1}}{\left|G_1\right|}\int_{G_1}\varphi_{mj}^*\left(\zeta_j\right)d\xi \tag{21}$$

It follows from the Eqs. (10), (18), (19) and (21)

$$RG\left(d, r_{mj}^*\right) - iRG\left(d, \rho_{mj}^*\right) \approx \frac{A_{k1}P_{mj}}{\left|G_1\right|L^2}\int_{G_1}\left[\xi_1^2 + 2\alpha_j\xi_1\xi_2 + \left(\alpha_j^2 - \beta_j^2\right)\xi_2^2 + 2i\beta_j\left(\xi_1\xi_2 + \alpha_j\xi_2^2\right)\right]d\xi \tag{22}$$

Let us remind that according to preceding results of the authors [1], to identify the ellipsoidal defect G_1, it is sufficiently to construct a matrix $\mathbf{Z}^1 = \left(Z_{ij}^1\right)$, where $Z_{ij}^1 = \dfrac{5}{|G_1|}\int\limits_{G_1} \xi_i \xi_j \, d\xi, \quad i = 1,2,3, \quad j = 1,2,3$.

As it was shown in [1], the eigenvalues of the matrix $\mathbf{Z}^1$ equal $\left(a_1^1\right)^2$, $\left(a_2^1\right)^2$ and $\left(a_3^1\right)^2$, where a_j^1, $j = 1,2,3$ are the semiaxes of the ellipsoid G_1. The corresponding eigenvectors are directed along the axes of the ellipsoid. Eq. (22) is a complex-valued linear equation with respect to three unknowns $Z_{pq}^1, \quad p = 1,2, \quad q = 1,2$. Using the described procedure for different s_j, $j = 1,2,3$, we obtain three linear complex-valued equations relative to the unknowns. It follows from the Eq. (22)

$$Z_{11}^1 + 2s_j Z_{12}^1 + s_j^2 Z_{22}^1 = \frac{5L^2\left[RG\left(d, r_{mj}^*\right) - iRG\left(r, \rho_{mj}^*\right)\right]}{A_{k1} P_{mj}}, \quad j = 1,2,3 \tag{23}$$

Because $s_i \neq s_j$ for $i \neq j$ the determinant of the system is not zero and the unknowns are determined uniquely. Considering projections on other coordinate planes we obtain all elements of the matrix $\mathbf{Z}^1$.

5 A numerical example

Consider a numerical example illustrating efficiency of the method of defects reconstruction presented in Sections 3 and 4. Let us assume that elastic body V is a cube $\left\{x : |x_i| \le 10, \, i = 1,2,3\right\}$, containing two ellipsoidal cavities. The loads applied to the boundary ∂V of elastic body V are chosen correspond to uniaxial tension in the direction of the axis x_3: $\mathbf{t}^d(x) = \left(0, 0, \sigma n_3(x)\right), \; x \in \partial V$, where $\sigma = 200$ MPa (see Fig. 1).

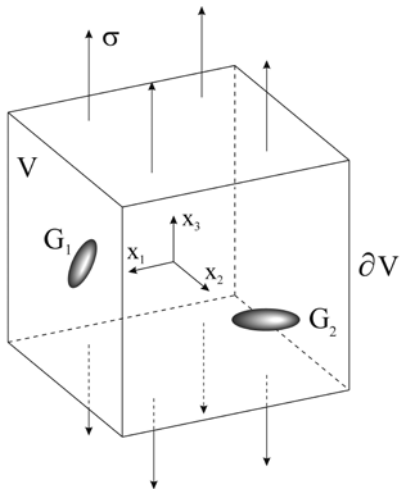


Figure 1: Uniaxial tension of a cube with embedded ellipsoidal defects.

We suppose that material of the cube V is orthotropic topaz. The elastic moduli of topaz in the crystallographic coordinate system can be found in [13] (see also [3]). The orientation of the crystallographic coordinate system relative to the given Cartesian coordinates we define by the Euler angles $\left(\beta_1, \beta_2, \beta_3\right)$ (see, [2]). In the considered example we took $\left(\beta_1, \beta_2, \beta_3\right) = \left(45^\circ, 45^\circ, 45^\circ\right)$. Two considered ellipsoidal cavities are defined by their centers $x^1 = (3, -4, -5)$, $x^2 = (-5, 4, 3)$, volumes ($|G_k| = 4\pi/3 \approx 4.189$, $k = 1,2$ corresponding to the volume of the unit ball), aspect ratios ($\rho_1^1 = 0.25$,

$\rho_2^1 = 0.50$, $\rho_1^2 = 0.50$, $\rho_2^2 = 0.75$, where $\rho_1^k = a_3^k / a_1^k$, $\rho_2^k = a_2^k / a_1^k$) and orientations. The orientation is defined by the Euler angles $(\varphi_k, \theta_k, \psi_k)$ (see, [2,3]). The Euler angles in the example are as follows: $(\varphi_1, \theta_1, \psi_1) = (30^\circ, 45^\circ, 60^\circ)$, $(\varphi_2, \theta_2, \psi_2) = (60^\circ, 45^\circ, 30^\circ)$.

The direct Neumann problem for the considered example is solved using FEM and displacements $\mathbf{u}^d$ are determined on the boundary ∂V. After that the corresponding values of the RGF are calculated. Using calculated values of the RGF the number of defects and their geometrical parameters are determined by means of the results presented in Sections 3, 4.

5.1 Determination of the number of ellipsoidal defects and coordinates of their centers

According to the approach proposed in Section 3, to determine the number of defects and their centers it is necessary to construct polynomials $P_{Nj}(w)$ for different values of the upper bound N and roots s_j, $j = 1,2,3$. The defects projections, corresponding to the roots w_{1k}, $k = 1,2,\cdots,10$ of the polynomial $P_{101}(w)$, constructed for the root s_1 in planes x_1x_2, x_2x_3, and x_1x_3, are presented on the Fig. 2. Here and below the projections of the given defects on the planes are grey dashed on the figures ((a) – plane x_1x_2, (b) – plane x_2x_3, (c) – plane x_1x_3). The projections of the body V are marked with the solid lines. The obtained defects projections, corresponding to the roots w_{1k} of the polynomials $P_{101}(w)$ are marked with the thick points ($\bullet$). The arrows ($\rightarrow$) denote the projections located outside the figures bounds. Among the obtained defects projections, presented on the Fig. 2, there are some spurious points which do not correspond to any defects. The spurious points are excluded using the criteria formulated in Section 3. To illustrate these criteria let us consider for example the coordinates of the projections of defects constructed in the plane x_1x_2 (see Fig. 2a). The points excluded by the criteria $1,2,\cdots,5$ are marked by the symbols ($\times$), ($\bigcirc$), ($\square$), ($\Diamond$) and ($\triangle$), respectively. To illustrate the application of the criterion 2, let us present the values of $|A_{k1}|$ corresponding to the Fig. 2a: $|A_{11}| = 0.105 \cdot 10^9$, $|A_{21}| = 0.123 \cdot 10^9$, $|A_{31}| = 0.573 \cdot 10^8$, $|A_{41}| = 0.403 \cdot 10^8$, $|A_{51}| = 0.152 \cdot 10^2$, $|A_{61}| = 0.748 \cdot 10^4$, $|A_{71}| = 0.231 \cdot 10^5$, $|A_{81}| = 0.139 \cdot 10^5$, $|A_{91}| = 0.450 \cdot 10^3$, $|A_{101}| = . = 0.408 \cdot 10^{-1}$. It is possible to see that the values $|A_{k1}|$, $k = 1,\cdots,4$ are much larger than the others.

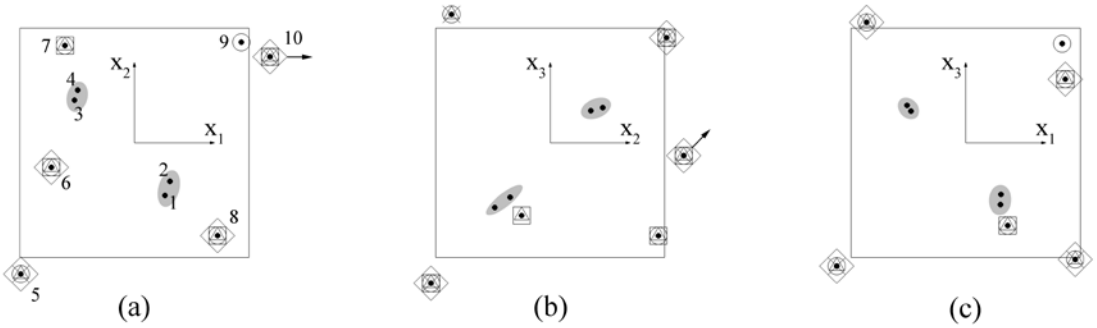
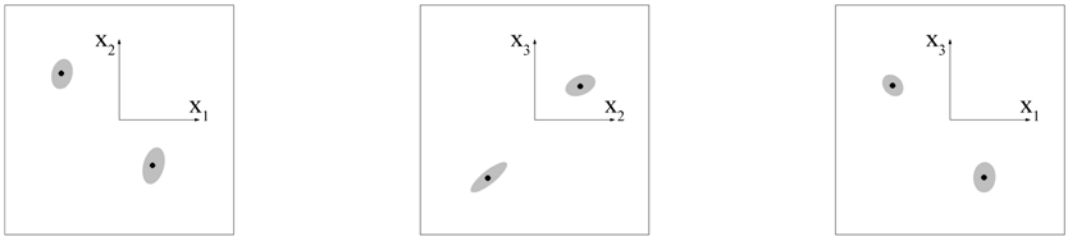


Figure 2: The defects projections determined by the polynomials $P_{101}(w)$.

As shown on Fig. 2a, the defects projections marked with the numbers 5-10 on the figure satisfy at least one of the formulated criteria and hence are spurious. Therefore, only the projections marked with the numbers 1-4 should be considered. It can be seen from Fig. 2a that the distance between the points 1, 2 is much less than the distances between the points 1, 3 and 2, 3 (or 1, 4 and 2, 4).

Accounting for supposition (1), it can be concluded that the points 1 and 2 correspond to the projection of the same defect. The same holds for the pair of points marked with the numbers 3, 4 on Fig. 2a. So, the number of defects $n = 2$. Applications of the considered procedure to the roots s_2, s_3 and other values of upper bounds N lead to the same number of defects.
Let us take the upper bound of the number of defects $N = n = 2$ and repeat the procedure. The obtained points corresponding to the defects projections are presented on Fig. 3. As can be seen the points are very close to the projections of the centers of the sought-for ellipsoidal cavities.

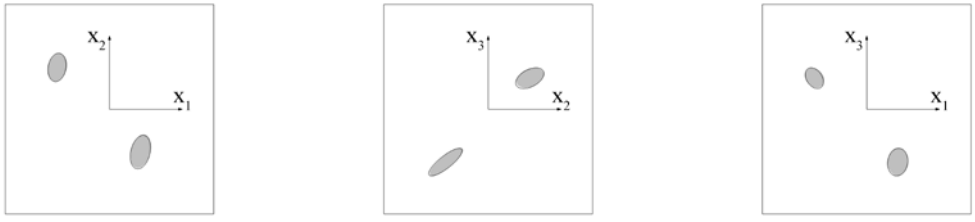


(a) (b) (c)

Figure 3: The defects projections determined by the polynomials $P_{21}(w)$.

5.2 Determination of the geometrical parameters of ellipsoidal defects

After determination of the number of ellipsoidal defects n and coordinates of their centers, the magnitudes and directions of the ellipsoids semiaxes can be calculated using the formulas obtained in Section 4. The geometrical parameters of the defects are determined by the use of regular elastic fields r_{mj}^* and ρ_{mj}^*, $j = 1,2,3$, depending on integer-valued parameter m ($m \geq 3$). The values of parameter m are chosen so big as to eliminate the contribution of all terms in the sum on the right side of Eq. (20) except one corresponding to the defect which parameters are determined. It follows from Eq. (20) that the bigger the number of defects n the bigger the value of parameter m should be used to obtain stable identification results.
Let us note that proposed identification method is based on the measurements which are usually subjected to some errors. Increasing of the parameter m leads to increasing of the degree of the polynomials corresponding to regular elastic fields r_{mj}^* and ρ_{mj}^*. As consequence, the errors in the calculated values of the RGF corresponding to the fields can also increase. Due to this reason using big values of the parameter m for identification of defects requires the measured data with higher accuracy. In the considered numerical examples the measured data were simulated by FEM. The level of accuracy of such data enabled us to consider only the values of parameter up to $m = 6$. The identification results obtained for $m = 4$ are presented on Fig. 4. Here the boundaries of the identified ellipsoidal defects projections are marked with solid lines and the projections of real defects are grey dashed.



(a) (b) (c)

Figure 4: The results of two ellipsoidal cavities identification.

The results presented on Fig. 4 show that for chosen value of the parameter m the identified defects projections are in exact agreement with the projections of the given ellipsoidal defects. Using $m = 5,6$ gives the results with similar accuracy. Using $m = 3$ in the considered case is not enough and leads to big errors in the results of identification.

Acknowledgements

The support of RFBR grants 13-01-00257, 13-01-00923 and grant MK-5261.2014.1 of President of Russian Federation for Government support of young Russian scientists is gratefully acknowledged.

References

[1] E.I. Shifrin, *Ellipsoidal defect identification in an elastic body from the results of a uniaxial tension (compression) test*, Mechanics of Solids, 45, (2010), pp. 417-426.

[2] E.I. Shifrin, P.S. Shushpannikov, *Identification of an ellipsoidal defect in an elastic solid using boundary measurements*, International Journal of Solids and Structures, 48, (2011), pp. 1154-1163.

[3] E.I. Shifrin, P.S. Shushpannikov, *Reconstruction of an ellipsoidal defect in anisotropic elastic solid, using results of one static test*, Inverse Problems in Science and Engineering, 21, (2013), pp. 781-800.

[4] E.I. Shifrin, P.S. Shushpannikov, *Identification of Defects in an Elastic Body by Means of the Boundary Measurements,* in Y.M. Hwang, Y.R. Jeng and C.P. Jiang, eds., Key Engineering Materials, Deformation and Fracture in Technological Processes, Trans Tech Publications, Switzerland, 528, (2013), pp. 101-110.

[5] E.I. Shifrin, P.S. Shushpannikov, *Identification of small well-separated defects in an isotropic elastic body using boundary measurements*, International Journal of Solids and Structures, 50, (2013), pp. 3707-3716.

[6] S. Andrieux, A. Ben Abda, H.D. Bui, *Reciprocity principle and crack identification*, Inverse Problems, 15, (1999), pp. 59-65.

[7] J.D. Eshelby, *The determination of the elastic field of an ellipsoidal inclusion and related problems*, Proc. Roy. Soc., A241, (1957), pp. 376 – 396.

[8] R.J. Asaro, *Somigliana dislocations and internal stresses: with application to second phase hardening*, Int. J. Engineering Science, 13, (1975), pp. 271 – 286.

[9] S.G. Lekhnitskii, *Theory of Elasticity of an Anisotropic Body*, Mir Publishers, Moscow, (1981).

[10] A. El Badia, T. Ha-Duong, *An inverse source problem in potential analysis*, Inverse Problems, 16, (2000), pp. 651-663.

[11] H. Kang, H. Lee, *Identification of simple poles via boundary measurements and an application of EIT*, Inverse Problems, 20, (2004), pp. 1853-1863.

[12] M. Hanke, W. Rundell, *On rational approximation methods for inverse source problems*, Inverse Problems and Imaging, 5, (2011), pp. 185-202.

[13] O.L. Anderson, *Determination and some uses of isotropic elastic constants of polycrystalline aggregates using simple crystal data*, in Physical Acoustic: Principles and Methods, W. Mason, ed., Academic Press, New York, (1965), pp. 43-95.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Inverse Problem in Anomalous Diffusion with Uncertainty Propagation

Luciano G. Silva[1,a], Diego C. Knupp[1,b], Luiz Bevilacqua[2], Augusto C. N. R. Galeão[3], Antônio J. Silva Neto[1,c]

[1]Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, Brazil,
e-mail: [a]luciano@iprj.uerj.br, [b]diegoknupp@gmail.com, [c]ajsneto@iprj.uerj.br

[2]Universidade Federal do Rio de Janeiro, UFRJ/COPPE, Rio de Janeiro, Brazil
e-mail: bevilacqua@coc.ufrj.br

[2]Laboratório Nacional de Computação Científica, LNCC, Petrópolis, Brazil
e-mail: acng@lncc.br

Key words: Anomalous diffusion, Inverse problems, Uncertainty propagation, Monte Carlo method

Abstract

Recently Bevilacqua, Galeão and co-workers developed a new analytical formulation for the simulation of the phenomena of diffusion with retention. This new formulation aims at the reduction of all diffusion processes with retention to a unifying phenomenon that can adequately simulate the retention effect. This model may have relevant applications in a number of different areas such as: population spreading with partial hold up of the population to guarantee territorial domain; chemical reactions inducing adsorption processes; and multiphase flow through porous media, just to mention a few. In the new formulation a discrete approach is firstly formulated taking into account a control parameter which represents the fraction of particles that are able to diffuse. The resulting governing equation for the modelling of diffusion with retention in a continuum medium requires a fourth order differential term. Specific experimental techniques, together with an appropriate inverse analysis, need to be settled to characterize the complementary parameters. The present work investigates an inverse problem which does not allow for simultaneous estimation of all model parameters and a characterization procedure in two steps is proposed: the first one in order to estimate the diffusion coefficient and the second one in order to estimate the complementary parameters. In this paper it is assumed that the first step is already done and the diffusion coefficient is known within certain degree of reliability. Therefore this work is aimed at investigating the confidence intervals of the complementary parameters estimates considering both the uncertainties due to measurement errors in the experimental data, and due to the uncertainty propagation of the estimated value of the diffusion coefficient. The inverse problem solution is carried out through the maximum likelihood approach, with the minimization problem solved with the Levenberg-Marquardt method, and the estimation of the confidence intervals is carried out through a Monte Carlo analysis.

1 Introduction

Spreading of particles or microorganisms immersed in a given medium or deployed on a given substratum is frequently modeled as a diffusion process, given by the well-known diffusion equation derived from the Fick's law. This model represents quite satisfactorily the behavior of several physical phenomena related to dispersion processes, but for some cases the approach fails to represent the real physical behavior. For instance, population spreading or dispersing particles may be partially and temporarily blocked when immersed in some particular media; an invading species may hold a fraction of the total population stationary on the conquered territory in order to guarantee territorial domain; and chemical reactions may induce absorption processes for the diffusion of solutes in liquid solvents in the presence of absorbent material [1]. Among other physicochemical phenomena that need improvement on the analytical formulation due to side effects not accounted for in the classical diffusion theory we may cite flows though porous media [2], and diffusion processes for some dispersing substances immersed in particular supporting media [3-7]. In most cases appearing in the literature addressing this issue, it is assumed the well-known second order parabolic equation as the basic governing equation of the dispersion process, but the anomalous diffusion effect is modeled with the introduction of fractional derivatives [8], or imposing an arbitrary variation of the diffusion coefficient with time or concentration [9, 10]. Nevertheless, trying to overcome the anomalous diffusion issue by imposing an artificial dependence of the diffusion coefficient on the particle concentration, or introducing extra differential terms while keeping the second order rank of the governing equation disguises the real physical phenomenon occurring in the process. In 2011, Bevilacqua, Galeão and co-workers derived a new analytical formulation for the simulation of the phenomena of anomalous diffusion [11], explicitly taking into account the retention effect in the dispersion process, aiming at the reduction of all diffusion processes with retention to a unifying phenomenon that can adequately simulate the retention effect. The new parameters introduced, besides the diffusion coefficient, characterize the blocking process and specific experimental techniques, together with an inverse analysis, need to be settled to determine these complementary parameters.

The present work investigates an anomalous diffusion inverse problem which does not allow for simultaneous estimation of all model parameters [12], and a characterization procedure in two steps is proposed. It is considered possible to impose a case in which the blocking process does not occur, for instance as it occurs in the flow of ferrofluid in microchannels. In this example, the problem may present an anomalous diffusion effect or not, depending on whether under the presence of a magnetic field or not [13]. Therefore, in the first step, assuming all particles are able to diffuse, the diffusion coefficient of the model may be estimated. Then the second step may be carried out in a situation in which anomalous diffusion occurs for the given problem, with the diffusion coefficient already characterized, in order to estimate the complementary anomalous diffusion parameters. In this paper it is assumed that the first step of the two-step procedure is already done and the diffusion coefficient is known within certain degree of reliability. Therefore this work is aimed at investigating the confidence intervals of the complementary parameters estimates, considering both the uncertainties due to measurement errors in the experimental data, and due to the uncertainty propagation of the considered value of the diffusion coefficient. The inverse problem solution is carried out through the maximum likelihood approach, with the minimization problem solved with the Levenberg-Marquardt method [14], and the estimation of the confidence intervals is carried out through a Monte Carlo analysis [15].

2 Problem formulation and solution methodology

Consider the process schematically represented in figure 1. The redistribution of the contents of each cell indicates that a fraction of the contents αp_n is retained in the n^{th} cell and the exceeding volume is evenly transferred to the neighboring cells, that is, $0.5\beta p_n$ to the left, to the $(n\text{-}1)^{th}$ cell and $0.5\beta p_n$ to the right, to the $(n+1)^{th}$ cell, at each time step, where $\beta = 1-\alpha$. This means that the dispersion runs slower than for the classical diffusion problem. Note that if $\beta = 1$, the problem is reduced to the classical Gaussian distribution.

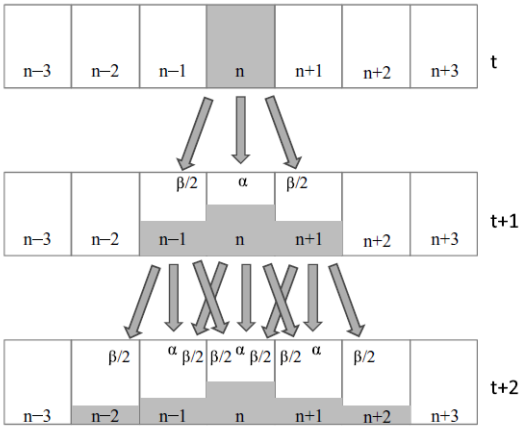


Figure 1: Schematic representation of the symmetric distribution with retention $\alpha = (\beta - 1)$.

This process can be written as the following algebraic expressions:

$$p_n^t = (1-\beta)\, p_n^{t-1} + \frac{1}{2}\beta\, p_{n-1}^{t-1} + \frac{1}{2}\beta\, p_{n+1}^{t-1} \tag{1a}$$

$$p_n^{t+1} = (1-\beta)\, p_n^{t} + \frac{1}{2}\beta\, p_{n-1}^{t} + \frac{1}{2}\beta\, p_{n+1}^{t} \tag{1b}$$

Manipulating eqs. (1a,b) in order to obtain finite difference terms yields:

$$\frac{\Delta p_n^{t+\Delta t}}{\Delta t} = \beta \left\{ \frac{1}{2}\frac{L_0^2}{T_0}\frac{\Delta^2 p_n}{\Delta x^2} + \frac{O(\Delta x^2)}{\Delta x^2} - (1-\beta)\frac{1}{4}\frac{L_1^4}{T_0}\frac{\Delta^4 p_n}{\Delta x^4} \right\}^{t-\Delta t} \tag{2}$$

where T_0, L_0 and L_1 are integration parameters. Calling $K_2 = L_0^2 / 2T_0$ and $K_4 = L_1^4 / 4T_0$, both considered constant in this work, and taking the limit as $\Delta x \to 0$ and $\Delta t \to 0$, we have:

$$\frac{\partial p(x,t)}{\partial t} = \beta K_2 \frac{\partial^2 p(x,t)}{\partial x^2} - \beta(1-\beta) K_4 \frac{\partial^4 p(x,t)}{\partial x^4} \tag{3a}$$

The fourth order term with negative sign introduces the anomalous diffusion effect, which turns up naturally, without any artificial assumption, as an immediate consequence of the temporary retention imposed by the redistribution law. Further discussion on the model derivation can be found in reference [11].

As the test case for the present work, consider the governing equation given by eq. (3a) valid for $0 < x < 1$ and $t > 0$, with the following boundary and initial conditions:

$$p(0,t)=p_0,\quad p(1,t)=p_1\ ,\ \left.\frac{\partial p(x,t)}{\partial x}\right|_{x=0}=0\ ,\ \left.\frac{\partial p(x,t)}{\partial x}\right|_{x=1}=0,\quad t>0 \tag{3b}$$

$$p(x,0)=f(x)=2\sin^{100}(\pi x)+1,\quad 0\le x\le 1 \tag{3c}$$

The problem given by eqs. (3a-c) is solved in this work with the *NDSolve* routine of the *Mathematica* system, under automatic absolute and relative error control. Concerning the inverse problem solution, observing the problem defined in eqs. (3), it is evident that the three parameters appearing in the model cannot be estimated simultaneously since there are three parameters defining two coefficients in eq. (3a), i.e. there are infinite sets of values for the parameters $\mathbf{Z}=\{\beta,K_2,K_4\}$ that lead to the exactly same mathematical formulation, yielding non-unicity of the inverse problem solution, which was also illustrated by means of a sensitivity analysis in ref. [12]. Since the most interesting aspect of this work is the identification of the three parameters appearing in the model, due to their direct physical interpretation [11], we choose not to rewrite the problem in terms of two coefficients (which would multiply the second and fourth order differential terms). Next we shall consider that the parameter K_2 can be obtained through an independent experiment, for example by means of an inverse problem in a physical situation where the blocking process that characterize the anomalous diffusion phenomenon does not occur, i.e. $\beta=1$. Then the main goal becomes to estimate β and K_4, and provide an accurate uncertainty analysis due to errors in the experimental data and in the considered value of K_2. The inverse problem formulation and solution is addressed at the following sections.

3 Inverse problem formulation and solution

In order to investigate the inverse problem solution concerning the estimation of the three model parameters, $\mathbf{Z}=\{\beta,K_2,K_4\}$, we consider a vector of experimental data $\mathbf{Y}$, simulated with the solution of eq. (3), and the addition of noise simulated from a normal distribution with known variance:

$$Y_i=p_i(\mathbf{Z}_{\text{exact}})+\varepsilon_i,\quad \varepsilon\sim N(0,\sigma_e^2) \tag{4}$$

In this case, the maximum likelihood approach leads to the ordinary least-squares norm as objective function, given by the sum of the squared residues between the experimental data and the predicted values from the solution of eq. (3):

$$S(\mathbf{Z})=\sum_{i=1}^{N_d}\left(p_i(\mathbf{Z})-Y_i\right)^2 \tag{5}$$

So the vector $\mathbf{Z}$ that minimizes S yields the maximum likelihood estimates for the model parameters under investigation. In order to minimize S, in this work, we use the Levenberg-Marquardt method [14]. Starting with an initial guess $\mathbf{Z}_0$ an iterative procedure is constructed, in which new estimates are obtained with

$$\mathbf{Z}_{n+1}=\mathbf{Z}_n+\Delta\mathbf{Z}_n,\quad n=0,1,2,\cdots \tag{6a}$$

being the correction $\Delta\mathbf{Z}_n$ calculated from

$$\Delta\mathbf{Z}_n=-\left[\mathbf{J}_n^T\mathbf{J}_n+\lambda_n\boldsymbol{\Gamma}\right]^{-1}\mathbf{J}_n^T\mathbf{R}_n \tag{6b}$$

where λ is a damping parameter, $\boldsymbol{\Gamma}$ is the identity matrix, the elements of the sensitivity matrix $\mathbf{J}$, known as the sensitivity coefficients, are

$$J_{ij} = \frac{\partial p_i}{\partial Z_j},\ i = 1,2,\cdots,N_d,\ j = 1,2,\cdots,N_p \tag{7}$$

where N_p is the number of parameters being estimated, i.e., the dimension of the vector $\mathbf{Z}$, and $\mathbf{R}$ is the vector of residues, whose elements are given by:

$$R_i = p_i(\mathbf{Z}) - Y_i \tag{8}$$

The iterative procedure of sequentially calculating $\Delta\mathbf{Z}_n$ and $\mathbf{Z}_{n+1}$ with eqs. (6a,b) is continued until the convergence criterion

$$\left|\Delta Z_{n,j}\right| < \varepsilon_{\text{tol}}, \quad \text{for } j = 1,2,\cdots,N_p \tag{9}$$

is satisfied, where ε_{tol} is a prescribed tolerance. The damping factor λ_n is varied during the iterative procedure, such that when convergence is achieved its value is close to zero.

The derivatives that must be calculated in order to obtain the sensitivity coefficients in eq. (7) can be computed with a finite difference scheme. Nevertheless, the finite difference approximations must be employed with care because of the choice of the increment. If a large value is used, it is possible that the approximations will not be sufficiently accurate. On the other hand, if very small values are used for the increment, large numerical errors can occur due to the difference of numbers very close to each other, motivating the use of more involved techniques for the computation of the sensitivity coefficients, such as the complex-step method [16], or the derivation and solution of the sensitivity coefficient equations [17]. In the present work, in order to have a safer computation of the sensitivity coefficients, the sensitivity equations have been derived and numerically solved, as presented in details in ref. [12], which will be omitted here due to space limitation.

It should also be highlighted that the sensitivity analysis plays a major role in several aspects related to the formulation and solution of inverse problems [17]. In order to obtain good estimates, within reasonable confidence intervals, it is required the sensitivity coefficients to be relatively high and, when two or more unknowns are simultaneously estimated, their sensitivity coefficients must be linearly independent, what graphically means that they should not present the same slope in absolute value. Otherwise $\left|\mathbf{J}^T\mathbf{J}\right| \approx 0$ and the problem is ill-conditioned. Since the problem investigated in this work involves parameters with different orders of magnitude, the scaled sensitivity coefficients are employed in order to allow for more evident comparisons between the sensitivity coefficients with respect to different parameters and identification of linear dependence. The scaled sensitivity coefficients are obtained by multiplying the sensitivity coefficient by the value of the concerned parameter, i.e. $\bar{X}_{Z_i} = Z_i \dfrac{\partial p}{\partial Z_i}$. Therefore we have:

$$\bar{X}_{\beta}(x,t) = \beta \frac{\partial p(x,t)}{\partial \beta};\ \bar{X}_{K_2}(x,t) = K_2 \frac{\partial p(x,t)}{\partial K_2};\qquad \bar{X}_{K_4}(x,t) = K_4 \frac{\partial p(x,t)}{\partial K_4} \tag{10}$$

In this work, we are interested in investigating the uncertainty propagation of K_2, assumed to be estimated independently, in the first step, into the estimates of β and K_4, and it is also assessed the additional uncertainties due to error in the measured experimental data employed in the inverse problem solution. In order to calculate the confidence intervals for the parameters β and K_4, it is solved the inverse problem of estimating β and K_4 through the minimization of the maximum likelihood objective function, eq. (5), assuming the parameter K_2 is known. In order to calculate the confidence associated to the estimated parameters, taking into account the uncertainty in the given

value of K_2, the Monte Carlo error propagation analysis is introduced [15]. The idea is to simulate M virtual noisy experiments, employing different values of K_2 (randomly simulated from the a priori known probability distribution), and different simulated experimental data, and then examine the statistics of the corresponding estimated parameters. This procedure can be seen as solving the inverse problem several times, and for each time it is solved all inputs (experimental data and K_2 in this case) are varied within their uncertainty limits randomly, obeying their known statistical distributions, independent of the others. After a sufficiently large number of independent calculations is performed, the distribution of the computed results (the estimated values of β and K_4 in this case) nearly describes the distribution of all possible results from the combination of the input data.

4 Results and discussion

In the following results, consider the case with $\beta = 0.2$, $K_2 = 10^{-3}$ and $K_4 = 10^{-5}$ in eq. (3a). In reference [12] this test case is investigated, and it is shown that for $t > 95$ the sensitivity coefficients become essentially constant, suggesting that measurements beyond that time may not aggregate useful information for the inverse problem solution. In the present work it also investigated the influence of the measurement position on the sensitivity coefficients. For instance, consider figure 2, which depicts the scaled sensitivity coefficients with respect to the three parameters, β, K_2, and K_4, at $t = 10$, for $0 \le x \le 1$. It should be observed in this figure that $x_m = 0.5$ may be the best position for performing transient measurements concerning the inverse problem solution. In the present we work, besides $x_m = 0.5$, it is also investigated transient measurements acquired with a single sensor at $x_m = 0.4$ and $x_m = 0.45$, for the inverse problem solution, in order to illustrate the influence of the measurement position choice on the confidence intervals of the estimates. In all results hereafter presented, it is considered 90 experimental data for the inverse problem solution, obtained from $t = 5$ up to $t = 95$, using a single sensor located at (i) case 1: $x_m = 0.5$, (ii) case 2: $x_m = 0.4$, and (iii) case 3: $x_m = 0.45$. For instance, figure 3 illustrates a set of experimental data at $x = 0.5$, simulated employing eq. (4) with $\sigma_e = 0.02$, yielding, in average, up to 4% noise in the data. In this figure, together with the experimental data, it is also plotted the curve obtained from the solution of problem (3) for the test case under consideration.

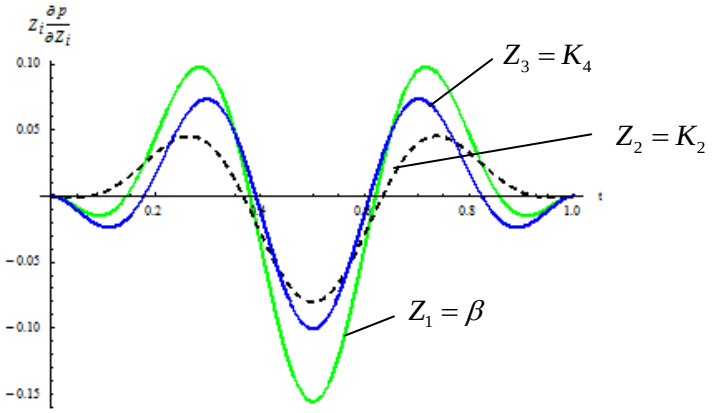


Figure 2: Scaled sensitivity coefficients along the spatial domain, x, at $t = 10$.

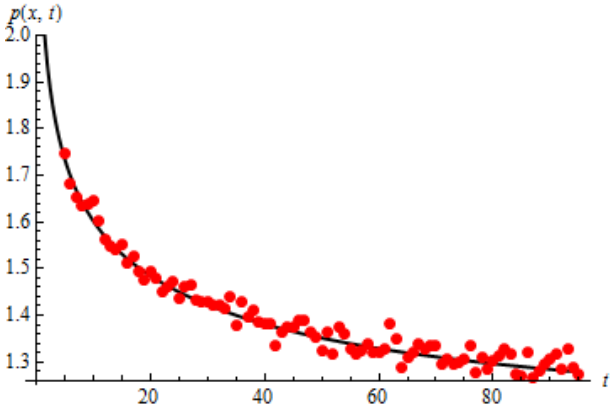
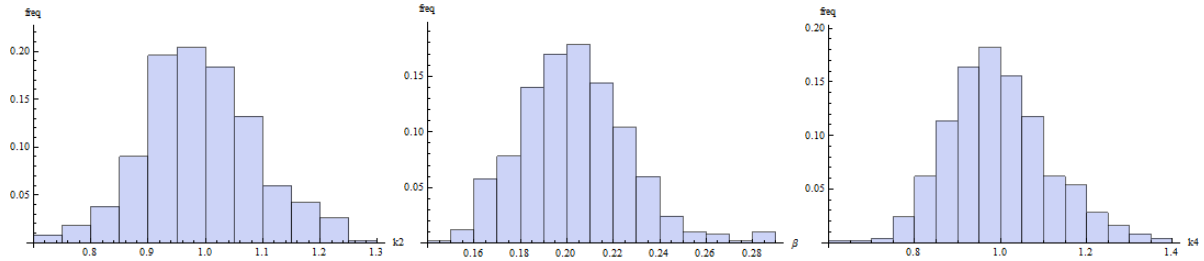


Figure 3: Simulated experimental data (red dots) for transient measurements of a sensor located at $x = 0.5$. The black curve shows the numerical solution employed to simulate the experimental data.

In the present work, for each one of the three measurement positions considered ($x_m = 0.5$, $x_m = 0.4$, and $x_m = 0.45$, named case 1, 2 and 3, respectively), $M = 500$ virtual noisy experiments have been simulated employing $\sigma_e = 0.02$ in eq. (4), and for each simulated independent experiment, different values of K_2 have been employed, randomly obtained from a normal distribution with 10^{-3} mean and 0.1×10^{-3} standard deviation (10% of the mean value), which means the 95% confidence interval for K_2 is $\left[0.8\times10^{-3},1.2\times10^{-3}\right]$. This information concerning K_2 is supposed to be obtained in the first step of the procedure herein proposed, with an experiment where the blocking process does not occur and the diffusion coefficient could be estimated. The goal here is to investigate how this uncertainty on the value of K_2 propagates into the estimates of β and K_4, in the presence of measurement errors, in the second step, now considering an experiment with anomalous diffusion and with this a priori information available for the parameter K_2.

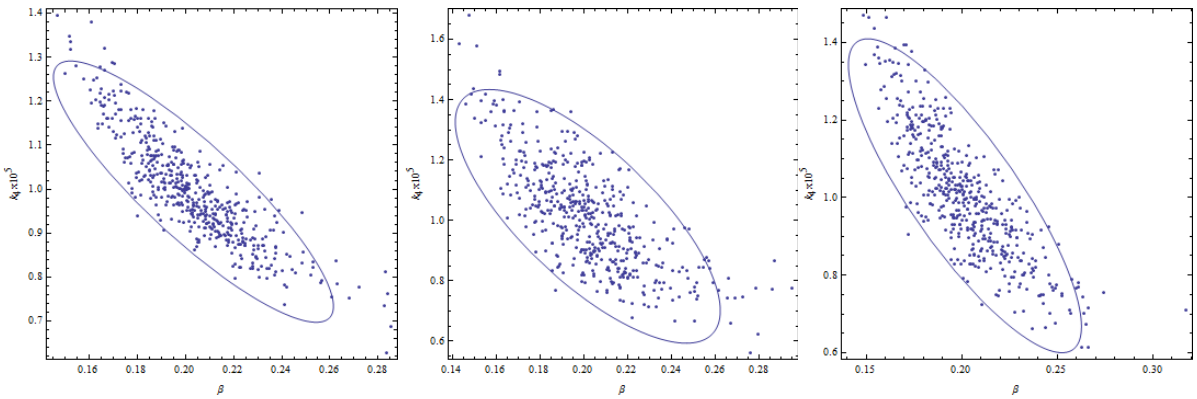
First, figure 4(a) illustrates the histogram plotted from the 500 values of K_2 employed in the simulations for case 1. For a sufficiently large number of simulations, this histogram approaches the exact normal distribution from which the values of K_2 have been sampled. In fact, figure 4(a) demonstrates that with 500 experiments it is already obtained a fairly good approximation of the distribution considered known for this distribution. In fact, the calculated 95% confidence interval from those 500 samples illustrated in figure 4(a) is $\left[0.804\times10^{-3},1.196\times10^{-3}\right]$, which approaches very well its exact interval, which is $\left[0.8\times10^{-3},1.2\times10^{-3}\right]$. Figures 4(b,c) depict the corresponding histograms for the estimates of the parameters β and K_4, respectively, obtained from the 500 simulated noisy experiments. It should be noted that both the histograms for β and K_4 seem to be slightly asymmetric with respect to the mean. Whilst this can be a collateral effect of the approximation obtained with a limited number of simulated experiments, it can be due to the nonlinearity of the problem, which means that even if both the experimental errors and the assumed values of K_2 are normally distributed, the statistical distributions of the estimated parameters β and K_4 are not necessarily normal.



(a) (b) (c)

Figure 4: histogram of (a) sampled values of K_2; (b) estimates of β; (c) estimates of K_4.

Proceeding to the investigation of the influence of the measurement position on the reliability of the estimates, figure 5 shows the elliptic joint confidence intervals for the estimated parameters β and K_4 for the three cases studied in this work: (a) case 1: $x_m = 0.5$, (b) case 2: $x_m = 0.4$, and (c) case 3: $x_m = 0.45$. A direct comparison of those three confidence regions is shown in figure 6, which, in agreement with the sensitivity analysis performed earlier, in fig. 2, show that the measurements performed at $x_m = 0.5$ produces the most reliable estimates, whereas the measurements performed at $x_m = 0.4$ produces the least reliable estimates. It is interesting to notice that this effect is much more evident in the estimation of K_4: in the direct comparison of the results obtained with the three different measurement positions (fig. 6), it should be observed that the elliptic confidence region narrows much more in the vertical direction of the graph (K_4) than in the horizontal direction of the graph (β).



(a) (b) (c)

Figure 5: Elliptic joint confidence region for β and K_4: (a) case 1; (b) case 2; (c) case 3.

The influence of the measurement position on the reliability of the estimates is better quantified in table 1, which besides the mean estimates and confidence intervals for each case, also shows the ratio between the standard deviation and the mean of each distribution. Remembering that this ratio was 10% in the assumed knowledge on K_2, it can be concluded that the uncertainties were not largely amplified into the estimates of β and K_4, noting that for case 1 their calculated ratios are 11.38% and 11.89%, which besides the uncertainty of K_2 also include the effect of the measurement errors. Nonetheless, if other less favorable measurement position is chosen (cases 2 or 3, for instance), these rations increase remarkably, especially for K_4, clearly illustrating the importance of the adequate choice of the experimental measurements employed for the inverse problem solution.

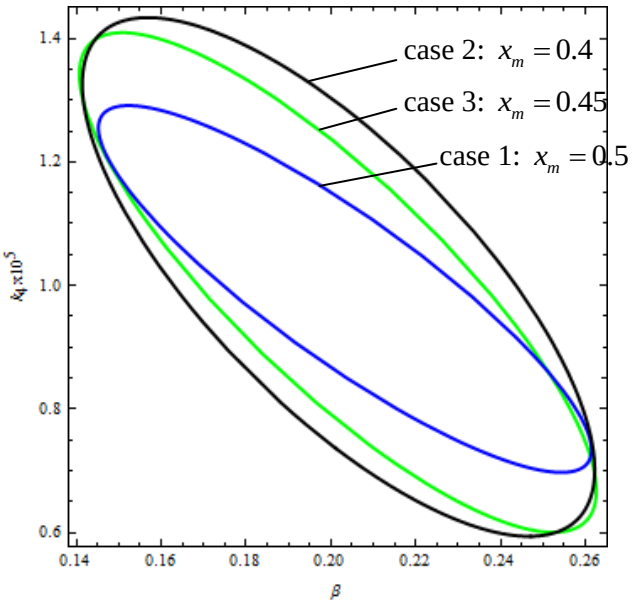


Figure 6: Direct comparison of the elliptic joint confidence regions for β and K_4 estimated in cases 1, 2 and 3.

Table 1: Estimated mean and 95% confidence intervals for β and K_4.

Meas. position	β	$\frac{\sigma_\beta}{\mu_\beta}\times 100\%$	$K_4 \times 10^5$	$\frac{\sigma_{K_4}}{\mu_{K_4}}\times 100\%$
Case 1, $x_m = 0.5$	0.203 [0.154, 0.246]	11.38%	0.994 [0.764, 1.24]	11.89%
Case 2, $x_m = 0.4$	0.201 [0.150, 0.250]	12.39%	1.01 [0.652, 1.35]	17.14%
Case 3, $x_m = 0.45$	0.202 [0.152, 0.248]	12.00%	1.00 [0.678, 1.32]	16.00%

5 Conclusions

It has been investigated in this work the inverse problem formulation and solution related with a new analytical formulation for the simulation of the phenomena of anomalous diffusion. The inverse problem investigated does not allow for the simultaneous estimation of all parameters and a characterization procedure in two steps is proposed. The reliability of the anomalous diffusion parameters estimates is studied concerning the uncertainty in the experimental data as well as the propagation of error concerning the value of the diffusion coefficient, estimated in the first step. The inverse analysis was carried out for transient measured experimental data obtained with a single sensor, whose position was investigated with respect to the corresponding estimates obtained. The results show that the errors present in the inputs do not amplify significantly into the estimates of the anomalous diffusion parameters. Nevertheless, the measured data employed in the inverse problem solution may be chosen with care, since it affects remarkably the reliability of the estimates.

Acknowledgements

The authors acknowledge the financial support provided by the Brazilian sponsoring agencies CNPq, CAPES and FAPERJ.

References

[1] Brandani, S., Jama, M., and Ruthven, D., *Diffusion, self-diffusion and counter-diffusion of benzene and p-xylene in silicalite*, Micropor Mesopor Mat, 35/36, (2000), pp. 283-300.

[2] Liu, H., and Thompson, K. E., *Numerical modeling of reactive polymer flow in porous media*, Comput Chem Eng, 26, (2002), pp. 1595-1610.

[3] Atsumi, H., *Hydrogen bulk retention in graphite and kinetics of diffusion*, J Nuc Mat, 307/311, (2002), pp. 1466-1470.

[4] D'Angelo, M. V., Fontana, E., Chertcoff, R., and Rosen, M., *Retention phenomena in non-Newtonian fluid flows*, Physics A, 327, (2003), pp. 44-48.

[5] Deleersnijder, E., Beckers, J.-M., and Delhez, E. M., *The residence time of setting in the surface mixed layer*, Environ Fluid Mech, 6, (2006), pp. 25-42.

[6] Green, P. F., *Translational dynamics of macromolecules in metals*, in: Neogi, P. (ed.), Diffusion in Polymers, Marcel Dekker Inc., (1996).

[7] Muhammad, N., Hydraulic, diffusion, and retention characteristics of inorganic chemicals in bentonite, Ph.D. Thesis, Department of Civil and Environmental Engineering, College of Engineering, University of South Florida, (2004), 251 p.

[8] Metzler, R., and Klafter, J., *The random walk's guide to anomalous diffusion: a fractional dynamics approach*, Phys. Rep. 339, (2000), pp. 1–77.

[9] Joannès, S., Mazé, L., and Bunsell, A. R., *A concentration-dependent diffusion coefficient model for water sorption in composite*, Composite Structures 108, (2014), pp. 111-118.

[10] Wu, J., and Berland, K. M., *Propagators and time-dependent diffusion coefficients for anomalous diffusion*, Biophysical Journal 95, (2008), pp. 2049-2052.

[11] Bevilacqua, L., Galeão, A. C. N. R. and Costa, F. P., A new analytical formulation of retention effects on particle diffusion process, *Annals of the Brazilian Academy of Sciences*, 83 (4) , (2011), pp. 1443-1464.

[12] Silva, L. G., Knupp, D. C., Bevilacqua, L., Galeão, A. C. N. R., Simas, J. G., Vasconcellos, J. F., and Silva Neto, A. J., *Investigation of a new model for anomalous diffusion phenomena by means of an inverse analysis*, 4th Inverse Problems, Design and Optimization Symposium, Albi, France, (2013).

[13] Derec, C., Smerlak, M., Servais, J., and Bacri, J.-C., *Anomalouos diffusion in microchannel under magnetic field*, Physics Procedia 9, (2010), pp. 109-112.

[14] Marquardt, D. W., *An algorithm for least-squares estimation of non-linear parameters*, J. Soc. Industr. Appl. Math., 11, (1963), pp. 431-441.

[15] Hahn, G. J., and Shapiro, S. S., *Statistical models in engineering*, Wiley, (1967).

[16] Thomsom, N. H., and Orlande, H. R. B., *Computation of sensitivity coefficients and estimation of thermophysical properties with the line heat source method*, Proceeding of the III European Conference on Computational Mechanics Solids, Structures and Coupled Problems in Engineering, 05-08 June 2006, Lisbon, Portugal, (2006).

[17] Blackwell, B. F., Dowding, K. J., and Cochran, R. J., *Development and implementation of sensitivity coefficient equations for heat conduction problems*, Numerical Heat Transfer – Part B, **36**, (1999), pp. 15-32.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Inverse Problem in Anomalous Diffusion with Uncertainty Propagation

Luciano G. Silva[1,a], Diego C. Knupp[1,b], Luiz Bevilacqua[2], Augusto C. N. R. Galeão[3], Antônio J. Silva Neto[1,c]

[1]Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, Brazil,
e-mail: [a]luciano@iprj.uerj.br, [b]diegoknupp@gmail.com, [c]ajsneto@iprj.uerj.br

[2]Universidade Federal do Rio de Janeiro, UFRJ/COPPE, Rio de Janeiro, Brazil
e-mail: bevilacqua@coc.ufrj.br

[2]Laboratório Nacional de Computação Científica, LNCC, Petrópolis, Brazil
e-mail: acng@lncc.br

Key words: Anomalous diffusion, Inverse problems, Uncertainty propagation, Monte Carlo method

Abstract

Recently Bevilacqua, Galeão and co-workers developed a new analytical formulation for the simulation of the phenomena of diffusion with retention. This new formulation aims at the reduction of all diffusion processes with retention to a unifying phenomenon that can adequately simulate the retention effect. This model may have relevant applications in a number of different areas such as: population spreading with partial hold up of the population to guarantee territorial domain; chemical reactions inducing adsorption processes; and multiphase flow through porous media, just to mention a few. In the new formulation a discrete approach is firstly formulated taking into account a control parameter which represents the fraction of particles that are able to diffuse. The resulting governing equation for the modelling of diffusion with retention in a continuum medium requires a fourth order differential term. Specific experimental techniques, together with an appropriate inverse analysis, need to be settled to characterize the complementary parameters. The present work investigates an inverse problem which does not allow for simultaneous estimation of all model parameters and a characterization procedure in two steps is proposed: the first one in order to estimate the diffusion coefficient and the second one in order to estimate the complementary parameters. In this paper it is assumed that the first step is already done and the diffusion coefficient is known within certain degree of reliability. Therefore this work is aimed at investigating the confidence intervals of the complementary parameters estimates considering both the uncertainties due to measurement errors in the experimental data, and due to the uncertainty propagation of the estimated value of the diffusion coefficient. The inverse problem solution is carried out through the maximum likelihood approach, with the minimization problem solved with the Levenberg-Marquardt method, and the estimation of the confidence intervals is carried out through a Monte Carlo analysis.

1 Introduction

Spreading of particles or microorganisms immersed in a given medium or deployed on a given substratum is frequently modeled as a diffusion process, given by the well-known diffusion equation derived from the Fick's law. This model represents quite satisfactorily the behavior of several physical phenomena related to dispersion processes, but for some cases the approach fails to represent the real physical behavior. For instance, population spreading or dispersing particles may be partially and temporarily blocked when immersed in some particular media; an invading species may hold a fraction of the total population stationary on the conquered territory in order to guarantee territorial domain; and chemical reactions may induce absorption processes for the diffusion of solutes in liquid solvents in the presence of absorbent material [1]. Among other physicochemical phenomena that need improvement on the analytical formulation due to side effects not accounted for in the classical diffusion theory we may cite flows though porous media [2], and diffusion processes for some dispersing substances immersed in particular supporting media [3-7]. In most cases appearing in the literature addressing this issue, it is assumed the well-known second order parabolic equation as the basic governing equation of the dispersion process, but the anomalous diffusion effect is modeled with the introduction of fractional derivatives [8], or imposing an arbitrary variation of the diffusion coefficient with time or concentration [9, 10]. Nevertheless, trying to overcome the anomalous diffusion issue by imposing an artificial dependence of the diffusion coefficient on the particle concentration, or introducing extra differential terms while keeping the second order rank of the governing equation disguises the real physical phenomenon occurring in the process. In 2011, Bevilacqua, Galeão and co-workers derived a new analytical formulation for the simulation of the phenomena of anomalous diffusion [11], explicitly taking into account the retention effect in the dispersion process, aiming at the reduction of all diffusion processes with retention to a unifying phenomenon that can adequately simulate the retention effect. The new parameters introduced, besides the diffusion coefficient, characterize the blocking process and specific experimental techniques, together with an inverse analysis, need to be settled to determine these complementary parameters.

The present work investigates an anomalous diffusion inverse problem which does not allow for simultaneous estimation of all model parameters [12], and a characterization procedure in two steps is proposed. It is considered possible to impose a case in which the blocking process does not occur, for instance as it occurs in the flow of ferrofluid in microchannels. In this example, the problem may present an anomalous diffusion effect or not, depending on whether under the presence of a magnetic field or not [13]. Therefore, in the first step, assuming all particles are able to diffuse, the diffusion coefficient of the model may be estimated. Then the second step may be carried out in a situation in which anomalous diffusion occurs for the given problem, with the diffusion coefficient already characterized, in order to estimate the complementary anomalous diffusion parameters. In this paper it is assumed that the first step of the two-step procedure is already done and the diffusion coefficient is known within certain degree of reliability. Therefore this work is aimed at investigating the confidence intervals of the complementary parameters estimates, considering both the uncertainties due to measurement errors in the experimental data, and due to the uncertainty propagation of the considered value of the diffusion coefficient. The inverse problem solution is carried out through the maximum likelihood approach, with the minimization problem solved with the Levenberg-Marquardt method [14], and the estimation of the confidence intervals is carried out through a Monte Carlo analysis [15].

2 Problem formulation and solution methodology

Consider the process schematically represented in figure 1. The redistribution of the contents of each cell indicates that a fraction of the contents αp_n is retained in the n^{th} cell and the exceeding volume is evenly transferred to the neighboring cells, that is, $0.5\beta p_n$ to the left, to the $(n\text{-}1)^{th}$ cell and $0.5\beta p_n$ to the right, to the $(n+1)^{th}$ cell, at each time step, where $\beta = 1-\alpha$. This means that the dispersion runs slower than for the classical diffusion problem. Note that if $\beta = 1$, the problem is reduced to the classical Gaussian distribution.

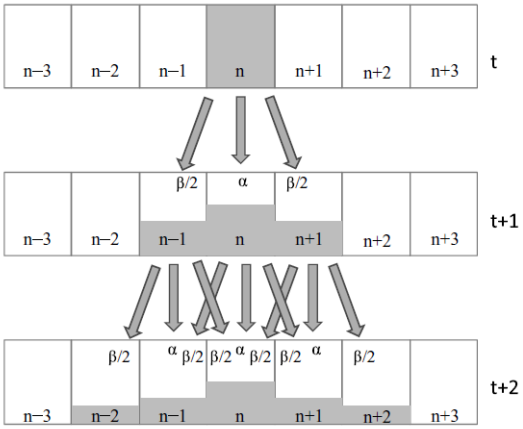


Figure 1: Schematic representation of the symmetric distribution with retention $\alpha = (\beta - 1)$.

This process can be written as the following algebraic expressions:

$$p_n^t = (1-\beta)\,p_n^{t-1} + \frac{1}{2}\beta\,p_{n-1}^{t-1} + \frac{1}{2}\beta\,p_{n+1}^{t-1} \tag{1a}$$

$$p_n^{t+1} = (1-\beta)\,p_n^{t} + \frac{1}{2}\beta\,p_{n-1}^{t} + \frac{1}{2}\beta\,p_{n+1}^{t} \tag{1b}$$

Manipulating eqs. (1a,b) in order to obtain finite difference terms yields:

$$\frac{\Delta p_n^{t+\Delta t}}{\Delta t} = \beta\left\{\frac{1}{2}\frac{L_0^2}{T_0}\frac{\Delta^2 p_n}{\Delta x^2} + \frac{O(\Delta x^2)}{\Delta x^2} - (1-\beta)\frac{1}{4}\frac{L_1^4}{T_0}\frac{\Delta^4 p_n}{\Delta x^4}\right\}^{t-\Delta t} \tag{2}$$

where T_0, L_0 and L_1 are integration parameters. Calling $K_2 = L_0^2 / 2T_0$ and $K_4 = L_1^4 / 4T_0$, both considered constant in this work, and taking the limit as $\Delta x \to 0$ and $\Delta t \to 0$, we have:

$$\frac{\partial p(x,t)}{\partial t} = \beta K_2 \frac{\partial^2 p(x,t)}{\partial x^2} - \beta(1-\beta)K_4 \frac{\partial^4 p(x,t)}{\partial x^4} \tag{3a}$$

The fourth order term with negative sign introduces the anomalous diffusion effect, which turns up naturally, without any artificial assumption, as an immediate consequence of the temporary retention imposed by the redistribution law. Further discussion on the model derivation can be found in reference [11].

As the test case for the present work, consider the governing equation given by eq. (3a) valid for $0 < x < 1$ and $t > 0$, with the following boundary and initial conditions:

$$p(0,t) = p_0, \quad p(1,t) = p_1, \quad \left.\frac{\partial p(x,t)}{\partial x}\right|_{x=0} = 0, \quad \left.\frac{\partial p(x,t)}{\partial x}\right|_{x=1} = 0, \quad t > 0 \tag{3b}$$

$$p(x,0) = f(x) = 2\sin^{100}(\pi x) + 1, \quad 0 \le x \le 1 \tag{3c}$$

The problem given by eqs. (3a-c) is solved in this work with the *NDSolve* routine of the *Mathematica* system, under automatic absolute and relative error control. Concerning the inverse problem solution, observing the problem defined in eqs. (3), it is evident that the three parameters appearing in the model cannot be estimated simultaneously since there are three parameters defining two coefficients in eq. (3a), i.e. there are infinite sets of values for the parameters $\mathbf{Z} = \{\beta, K_2, K_4\}$ that lead to the exactly same mathematical formulation, yielding non-unicity of the inverse problem solution, which was also illustrated by means of a sensitivity analysis in ref. [12]. Since the most interesting aspect of this work is the identification of the three parameters appearing in the model, due to their direct physical interpretation [11], we choose not to rewrite the problem in terms of two coefficients (which would multiply the second and fourth order differential terms). Next we shall consider that the parameter K_2 can be obtained through an independent experiment, for example by means of an inverse problem in a physical situation where the blocking process that characterize the anomalous diffusion phenomenon does not occur, i.e. $\beta = 1$. Then the main goal becomes to estimate β and K_4, and provide an accurate uncertainty analysis due to errors in the experimental data and in the considered value of K_2. The inverse problem formulation and solution is addressed at the following sections.

3 Inverse problem formulation and solution

In order to investigate the inverse problem solution concerning the estimation of the three model parameters, $\mathbf{Z} = \{\beta, K_2, K_4\}$, we consider a vector of experimental data $\mathbf{Y}$, simulated with the solution of eq. (3), and the addition of noise simulated from a normal distribution with known variance:

$$Y_i = p_i(\mathbf{Z}_{\text{exact}}) + \varepsilon_i, \quad \varepsilon \sim N(0, \sigma_e^2) \tag{4}$$

In this case, the maximum likelihood approach leads to the ordinary least-squares norm as objective function, given by the sum of the squared residues between the experimental data and the predicted values from the solution of eq. (3):

$$S(\mathbf{Z}) = \sum_{i=1}^{N_d} \left(p_i(\mathbf{Z}) - Y_i \right)^2 \tag{5}$$

So the vector $\mathbf{Z}$ that minimizes S yields the maximum likelihood estimates for the model parameters under investigation. In order to minimize S, in this work, we use the Levenberg-Marquardt method [14]. Starting with an initial guess $\mathbf{Z}_0$ an iterative procedure is constructed, in which new estimates are obtained with

$$\mathbf{Z}_{n+1} = \mathbf{Z}_n + \Delta\mathbf{Z}_n, \quad n = 0, 1, 2, \cdots \tag{6a}$$

being the correction $\Delta\mathbf{Z}_n$ calculated from

$$\Delta\mathbf{Z}_n = -\left[\mathbf{J}_n^T \mathbf{J}_n + \lambda_n \boldsymbol{\Gamma} \right]^{-1} \mathbf{J}_n^T \mathbf{R}_n \tag{6b}$$

where λ is a damping parameter, $\boldsymbol{\Gamma}$ is the identity matrix, the elements of the sensitivity matrix $\mathbf{J}$, known as the sensitivity coefficients, are

$$J_{ij} = \frac{\partial p_i}{\partial Z_j},\ i = 1,2,\cdots,N_d,\ j = 1,2,\cdots,N_p \tag{7}$$

where N_p is the number of parameters being estimated, i.e., the dimension of the vector $\mathbf{Z}$, and $\mathbf{R}$ is the vector of residues, whose elements are given by:

$$R_i = p_i(\mathbf{Z}) - Y_i \tag{8}$$

The iterative procedure of sequentially calculating $\Delta\mathbf{Z}_n$ and $\mathbf{Z}_{n+1}$ with eqs. (6a,b) is continued until the convergence criterion

$$\left|\Delta Z_{n,j}\right| < \varepsilon_{\text{tol}}, \quad \text{for } j = 1,2,\cdots,N_p \tag{9}$$

is satisfied, where ε_{tol} is a prescribed tolerance. The damping factor λ_n is varied during the iterative procedure, such that when convergence is achieved its value is close to zero.

The derivatives that must be calculated in order to obtain the sensitivity coefficients in eq. (7) can be computed with a finite difference scheme. Nevertheless, the finite difference approximations must be employed with care because of the choice of the increment. If a large value is used, it is possible that the approximations will not be sufficiently accurate. On the other hand, if very small values are used for the increment, large numerical errors can occur due to the difference of numbers very close to each other, motivating the use of more involved techniques for the computation of the sensitivity coefficients, such as the complex-step method [16], or the derivation and solution of the sensitivity coefficient equations [17]. In the present work, in order to have a safer computation of the sensitivity coefficients, the sensitivity equations have been derived and numerically solved, as presented in details in ref. [12], which will be omitted here due to space limitation.

It should also be highlighted that the sensitivity analysis plays a major role in several aspects related to the formulation and solution of inverse problems [17]. In order to obtain good estimates, within reasonable confidence intervals, it is required the sensitivity coefficients to be relatively high and, when two or more unknowns are simultaneously estimated, their sensitivity coefficients must be linearly independent, what graphically means that they should not present the same slope in absolute value. Otherwise $\left|\mathbf{J}^T\mathbf{J}\right| \approx 0$ and the problem is ill-conditioned. Since the problem investigated in this work involves parameters with different orders of magnitude, the scaled sensitivity coefficients are employed in order to allow for more evident comparisons between the sensitivity coefficients with respect to different parameters and identification of linear dependence. The scaled sensitivity coefficients are obtained by multiplying the sensitivity coefficient by the value of the concerned parameter, i.e. $\bar{X}_{Z_i} = Z_i \dfrac{\partial p}{\partial Z_i}$. Therefore we have:

$$\bar{X}_{\beta}(x,t) = \beta \frac{\partial p(x,t)}{\partial \beta};\ \bar{X}_{K_2}(x,t) = K_2 \frac{\partial p(x,t)}{\partial K_2};\qquad \bar{X}_{K_4}(x,t) = K_4 \frac{\partial p(x,t)}{\partial K_4} \tag{10}$$

In this work, we are interested in investigating the uncertainty propagation of K_2, assumed to be estimated independently, in the first step, into the estimates of β and K_4, and it is also assessed the additional uncertainties due to error in the measured experimental data employed in the inverse problem solution. In order to calculate the confidence intervals for the parameters β and K_4, it is solved the inverse problem of estimating β and K_4 through the minimization of the maximum likelihood objective function, eq. (5), assuming the parameter K_2 is known. In order to calculate the confidence associated to the estimated parameters, taking into account the uncertainty in the given

value of K_2, the Monte Carlo error propagation analysis is introduced [15]. The idea is to simulate M virtual noisy experiments, employing different values of K_2 (randomly simulated from the a priori known probability distribution), and different simulated experimental data, and then examine the statistics of the corresponding estimated parameters. This procedure can be seen as solving the inverse problem several times, and for each time it is solved all inputs (experimental data and K_2 in this case) are varied within their uncertainty limits randomly, obeying their known statistical distributions, independent of the others. After a sufficiently large number of independent calculations is performed, the distribution of the computed results (the estimated values of β and K_4 in this case) nearly describes the distribution of all possible results from the combination of the input data.

4 Results and discussion

In the following results, consider the case with $\beta = 0.2$, $K_2 = 10^{-3}$ and $K_4 = 10^{-5}$ in eq. (3a). In reference [12] this test case is investigated, and it is shown that for $t > 95$ the sensitivity coefficients become essentially constant, suggesting that measurements beyond that time may not aggregate useful information for the inverse problem solution. In the present work it also investigated the influence of the measurement position on the sensitivity coefficients. For instance, consider figure 2, which depicts the scaled sensitivity coefficients with respect to the three parameters, β, K_2, and K_4, at $t = 10$, for $0 \le x \le 1$. It should be observed in this figure that $x_m = 0.5$ may be the best position for performing transient measurements concerning the inverse problem solution. In the present we work, besides $x_m = 0.5$, it is also investigated transient measurements acquired with a single sensor at $x_m = 0.4$ and $x_m = 0.45$, for the inverse problem solution, in order to illustrate the influence of the measurement position choice on the confidence intervals of the estimates. In all results hereafter presented, it is considered 90 experimental data for the inverse problem solution, obtained from $t = 5$ up to $t = 95$, using a single sensor located at (i) case 1: $x_m = 0.5$, (ii) case 2: $x_m = 0.4$, and (iii) case 3: $x_m = 0.45$. For instance, figure 3 illustrates a set of experimental data at $x = 0.5$, simulated employing eq. (4) with $\sigma_e = 0.02$, yielding, in average, up to 4% noise in the data. In this figure, together with the experimental data, it is also plotted the curve obtained from the solution of problem (3) for the test case under consideration.

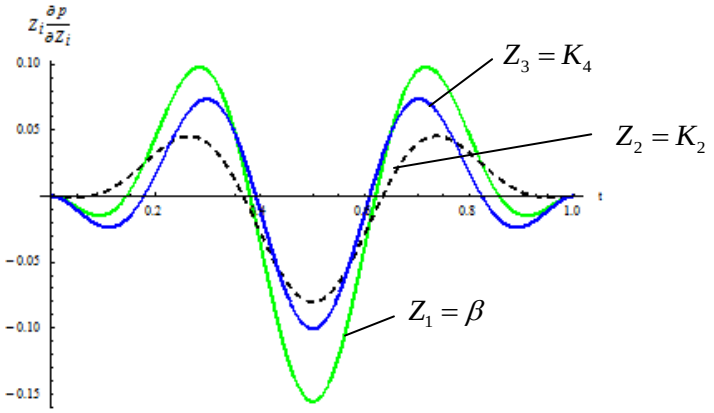


Figure 2: Scaled sensitivity coefficients along the spatial domain, x, at $t = 10$.

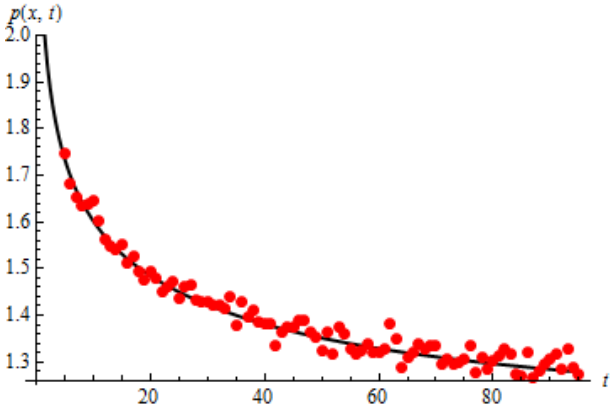
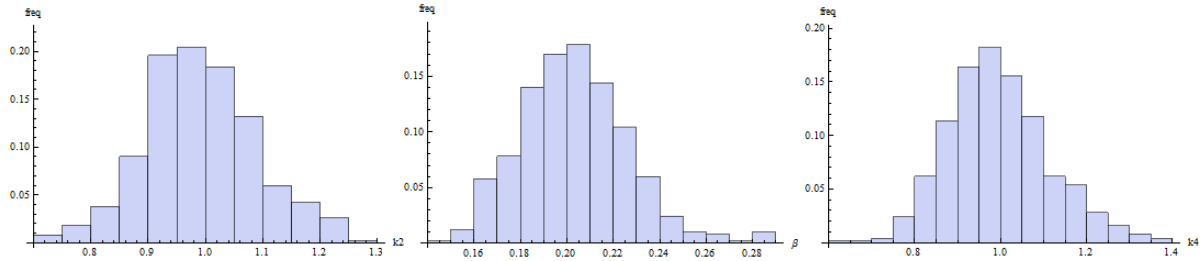


Figure 3: Simulated experimental data (red dots) for transient measurements of a sensor located at $x = 0.5$. The black curve shows the numerical solution employed to simulate the experimental data.

In the present work, for each one of the three measurement positions considered ($x_m = 0.5$, $x_m = 0.4$, and $x_m = 0.45$, named case 1, 2 and 3, respectively), $M = 500$ virtual noisy experiments have been simulated employing $\sigma_e = 0.02$ in eq. (4), and for each simulated independent experiment, different values of K_2 have been employed, randomly obtained from a normal distribution with 10^{-3} mean and 0.1×10^{-3} standard deviation (10% of the mean value), which means the 95% confidence interval for K_2 is $\left[0.8\times10^{-3},1.2\times10^{-3}\right]$. This information concerning K_2 is supposed to be obtained in the first step of the procedure herein proposed, with an experiment where the blocking process does not occur and the diffusion coefficient could be estimated. The goal here is to investigate how this uncertainty on the value of K_2 propagates into the estimates of β and K_4, in the presence of measurement errors, in the second step, now considering an experiment with anomalous diffusion and with this a priori information available for the parameter K_2.

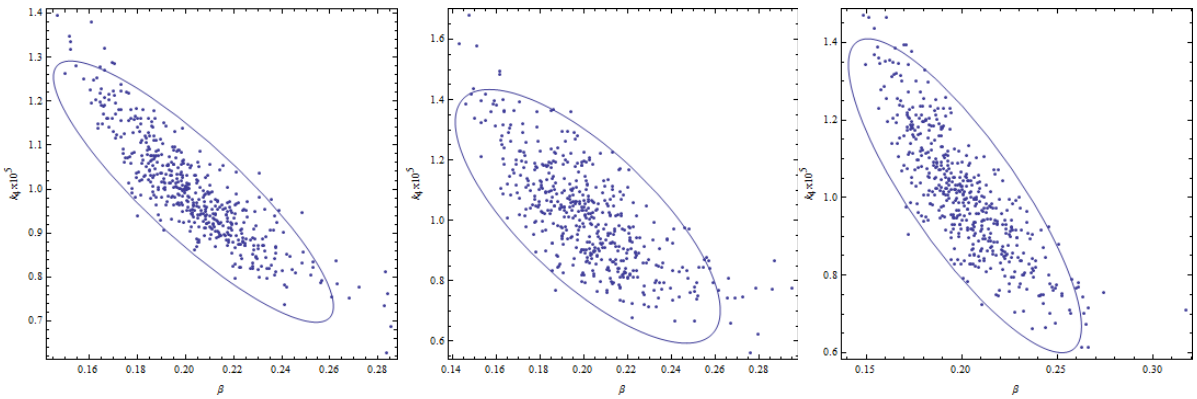
First, figure 4(a) illustrates the histogram plotted from the 500 values of K_2 employed in the simulations for case 1. For a sufficiently large number of simulations, this histogram approaches the exact normal distribution from which the values of K_2 have been sampled. In fact, figure 4(a) demonstrates that with 500 experiments it is already obtained a fairly good approximation of the distribution considered known for this distribution. In fact, the calculated 95% confidence interval from those 500 samples illustrated in figure 4(a) is $\left[0.804\times10^{-3},1.196\times10^{-3}\right]$, which approaches very well its exact interval, which is $\left[0.8\times10^{-3},1.2\times10^{-3}\right]$. Figures 4(b,c) depict the corresponding histograms for the estimates of the parameters β and K_4, respectively, obtained from the 500 simulated noisy experiments. It should be noted that both the histograms for β and K_4 seem to be slightly asymmetric with respect to the mean. Whilst this can be a collateral effect of the approximation obtained with a limited number of simulated experiments, it can be due to the nonlinearity of the problem, which means that even if both the experimental errors and the assumed values of K_2 are normally distributed, the statistical distributions of the estimated parameters β and K_4 are not necessarily normal.



(a) (b) (c)

Figure 4: histogram of (a) sampled values of K_2; (b) estimates of β; (c) estimates of K_4.

Proceeding to the investigation of the influence of the measurement position on the reliability of the estimates, figure 5 shows the elliptic joint confidence intervals for the estimated parameters β and K_4 for the three cases studied in this work: (a) case 1: $x_m = 0.5$, (b) case 2: $x_m = 0.4$, and (c) case 3: $x_m = 0.45$. A direct comparison of those three confidence regions is shown in figure 6, which, in agreement with the sensitivity analysis performed earlier, in fig. 2, show that the measurements performed at $x_m = 0.5$ produces the most reliable estimates, whereas the measurements performed at $x_m = 0.4$ produces the least reliable estimates. It is interesting to notice that this effect is much more evident in the estimation of K_4: in the direct comparison of the results obtained with the three different measurement positions (fig. 6), it should be observed that the elliptic confidence region narrows much more in the vertical direction of the graph (K_4) than in the horizontal direction of the graph (β).



(a) (b) (c)

Figure 5: Elliptic joint confidence region for β and K_4: (a) case 1; (b) case 2; (c) case 3.

The influence of the measurement position on the reliability of the estimates is better quantified in table 1, which besides the mean estimates and confidence intervals for each case, also shows the ratio between the standard deviation and the mean of each distribution. Remembering that this ratio was 10% in the assumed knowledge on K_2, it can be concluded that the uncertainties were not largely amplified into the estimates of β and K_4, noting that for case 1 their calculated ratios are 11.38% and 11.89%, which besides the uncertainty of K_2 also include the effect of the measurement errors. Nonetheless, if other less favorable measurement position is chosen (cases 2 or 3, for instance), these rations increase remarkably, especially for K_4, clearly illustrating the importance of the adequate choice of the experimental measurements employed for the inverse problem solution.

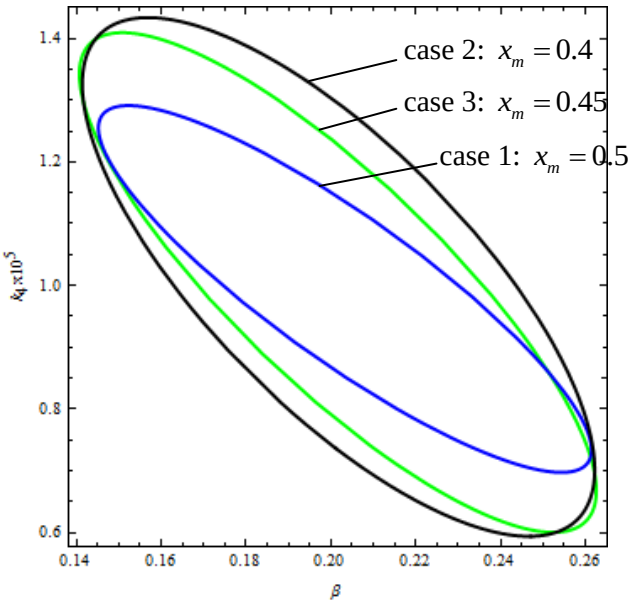


Figure 6: Direct comparison of the elliptic joint confidence regions for β and K_4 estimated in cases 1, 2 and 3.

Table 1: Estimated mean and 95% confidence intervals for β and K_4.

Meas. position	β	$\frac{\sigma_\beta}{\mu_\beta}\times 100\%$	$K_4 \times 10^5$	$\frac{\sigma_{K_4}}{\mu_{K_4}}\times 100\%$
Case 1, $x_m = 0.5$	0.203 [0.154, 0.246]	11.38%	0.994 [0.764, 1.24]	11.89%
Case 2, $x_m = 0.4$	0.201 [0.150, 0.250]	12.39%	1.01 [0.652, 1.35]	17.14%
Case 3, $x_m = 0.45$	0.202 [0.152, 0.248]	12.00%	1.00 [0.678, 1.32]	16.00%

5 Conclusions

It has been investigated in this work the inverse problem formulation and solution related with a new analytical formulation for the simulation of the phenomena of anomalous diffusion. The inverse problem investigated does not allow for the simultaneous estimation of all parameters and a characterization procedure in two steps is proposed. The reliability of the anomalous diffusion parameters estimates is studied concerning the uncertainty in the experimental data as well as the propagation of error concerning the value of the diffusion coefficient, estimated in the first step. The inverse analysis was carried out for transient measured experimental data obtained with a single sensor, whose position was investigated with respect to the corresponding estimates obtained. The results show that the errors present in the inputs do not amplify significantly into the estimates of the anomalous diffusion parameters. Nevertheless, the measured data employed in the inverse problem solution may be chosen with care, since it affects remarkably the reliability of the estimates.

Acknowledgements

The authors acknowledge the financial support provided by the Brazilian sponsoring agencies CNPq, CAPES and FAPERJ.

References

[1] Brandani, S., Jama, M., and Ruthven, D., *Diffusion, self-diffusion and counter-diffusion of benzene and p-xylene in silicalite*, Micropor Mesopor Mat, 35/36, (2000), pp. 283-300.

[2] Liu, H., and Thompson, K. E., *Numerical modeling of reactive polymer flow in porous media*, Comput Chem Eng, 26, (2002), pp. 1595-1610.

[3] Atsumi, H., *Hydrogen bulk retention in graphite and kinetics of diffusion*, J Nuc Mat, 307/311, (2002), pp. 1466-1470.

[4] D'Angelo, M. V., Fontana, E., Chertcoff, R., and Rosen, M., *Retention phenomena in non-Newtonian fluid flows*, Physics A, 327, (2003), pp. 44-48.

[5] Deleersnijder, E., Beckers, J.-M., and Delhez, E. M., *The residence time of setting in the surface mixed layer*, Environ Fluid Mech, 6, (2006), pp. 25-42.

[6] Green, P. F., *Translational dynamics of macromolecules in metals*, in: Neogi, P. (ed.), Diffusion in Polymers, Marcel Dekker Inc., (1996).

[7] Muhammad, N., Hydraulic, diffusion, and retention characteristics of inorganic chemicals in bentonite, Ph.D. Thesis, Department of Civil and Environmental Engineering, College of Engineering, University of South Florida, (2004), 251 p.

[8] Metzler, R., and Klafter, J., *The random walk's guide to anomalous diffusion: a fractional dynamics approach*, Phys. Rep. 339, (2000), pp. 1–77.

[9] Joannès, S., Mazé, L., and Bunsell, A. R., *A concentration-dependent diffusion coefficient model for water sorption in composite*, Composite Structures 108, (2014), pp. 111-118.

[10] Wu, J., and Berland, K. M., *Propagators and time-dependent diffusion coefficients for anomalous diffusion*, Biophysical Journal 95, (2008), pp. 2049-2052.

[11] Bevilacqua, L., Galeão, A. C. N. R. and Costa, F. P., A new analytical formulation of retention effects on particle diffusion process, *Annals of the Brazilian Academy of Sciences*, 83 (4) , (2011), pp. 1443-1464.

[12] Silva, L. G., Knupp, D. C., Bevilacqua, L., Galeão, A. C. N. R., Simas, J. G., Vasconcellos, J. F., and Silva Neto, A. J., *Investigation of a new model for anomalous diffusion phenomena by means of an inverse analysis*, 4th Inverse Problems, Design and Optimization Symposium, Albi, France, (2013).

[13] Derec, C., Smerlak, M., Servais, J., and Bacri, J.-C., *Anomalouos diffusion in microchannel under magnetic field*, Physics Procedia 9, (2010), pp. 109-112.

[14] Marquardt, D. W., *An algorithm for least-squares estimation of non-linear parameters*, J. Soc. Industr. Appl. Math., 11, (1963), pp. 431-441.

[15] Hahn, G. J., and Shapiro, S. S., *Statistical models in engineering*, Wiley, (1967).

[16] Thomsom, N. H., and Orlande, H. R. B., *Computation of sensitivity coefficients and estimation of thermophysical properties with the line heat source method*, Proceeding of the III European Conference on Computational Mechanics Solids, Structures and Coupled Problems in Engineering, 05-08 June 2006, Lisbon, Portugal, (2006).

[17] Blackwell, B. F., Dowding, K. J., and Cochran, R. J., *Development and implementation of sensitivity coefficient equations for heat conduction problems*, Numerical Heat Transfer – Part B, **36**, (1999), pp. 15-32.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Pressure/Rate Deconvolution Problem and Reconstruction of Solution Satisfying to All A Priori Constraints

Georgy G. Skorik, Vladimir V. Vasin*

Institute of Mathematics and Mechanics UB RAS, S. Kovalevskaya 16, 620990 Ekaterinburg, Russia,
Ural Federal University, Lenin avenue 51, 620000 Ekaterinburg, Russia
e-mail: skorik@imm.uran.ru, vasin@imm.uran.ru

Key words: deconvolution problem, Volterra equation, well test, regularization algorithm, a priori information

Abstract
The pressure/rate deconvolution problem is formulated in the form of a linear Volterra equation of the first kind. Besides its ill-posedness, the problem is characterized by multi-scale behaviour of solution and discontinuous input data with large measurement errors, which do not allow one to apply standard regularizing algorithms for solving the Volterra equation. Therefore, in the most cases, algorithms constructed for this problem contain ideas of regularization and take into account *a priori* information usually in the form of certain conditions on solution, f.e., positivity, monotony or convexity (see Introduction). However, as it is known, solution of the convolution problem satisfies the infinite system of inequalities. In this paper, first for this problem, we construct two effective algorithms, which generate approximate solutions satisfying all *a priori* constraints, prove the convergence of the methods, and discuss results of numerical experiments for noisy input data.

1 Introduction

The pressure/rate deconvolution problem is described by the firs kind Volterra equation [1]

$$Ag \equiv \int_0^t q(t-\tau)g(\tau)d\tau = \Delta p(t), \quad 0 \le t \le T \tag{1}$$

and arises in well test interpretation. Equation (1) follows from the Duhamel principle, which means that the measured pressure drop $\Delta p(t) = p_0 - p(t)$ is a convolution of the measured flow rate $q(t)$, and the reservoir response function $g(t)$, where $g(t) = dp_u(t)/dt$, $p_u(t)$, is the constant-unit-rate

pressure drop. Here, p_0 is the initial reservoir pressure (in undisturbed state of reservoir), $p(t)$ is a pressure measured somewhere in the well, for example, on the surface or at the bottom hole. The functions $g(t)$ and $p_u(t)$ are unknown and should be calculated from measured with noise $\Delta p(t)$ and $q(t)$. The pressure/rate deconvolution problem attracts the grate attention of researchers and reservoir engineers because the log-log plots of functions $p_u(t)$ and $v(t) = tg(t)$ contain very important information about the wellbore/reservoir system. Namely, the first parts of these curves characterize the wellbore storage, a type of the reservoir permeability can be extracted from the middle parts of the curves, and a type of the boundaries can be determined from their final parts (see [2,3]).

As it is known [4], under certain conditions onto the wellbore/reservoir system, the reservoir response function $g(t)$ satisfies the following *a priori* conditions:

$$(-1)^k d^k g(t)/dt^k \geq 0, \quad k = 0,1,2,\ldots . \tag{2}$$

It should be noted that in all numerous works (see a survey in [5]) devoted to solving this problem, one usually used not more than three relations from (2):

$$g(t) \geq 0, \quad g'(t) \leq 0, \quad g''(t) \geq 0 . \tag{3}$$

In particular, constraints (3) were taken into account in [5], where two new regular algorithms were constructed. In both algorithms, the desired solution was represented as a decomposition into a system of specific base functions, which satisfy constraints (3).

In the work [6], very interesting idea was suggested that is based on transition from the linear equation (1) to the nonlinear statement in the form

$$\overline{A}(z) = \int_{-\infty}^{\ln t} q(t - e^{\sigma}) e^{z(\sigma)} d\sigma = \Delta p(t) \tag{4}$$

by replacements of the variables and functions:

$$\sigma = \ln \tau, \quad z(\sigma) = \ln(\tau g(\tau)) , \tag{5}$$

and, by this, the rather stable algorithm was constructed. These replacements allow one to decrease influence of different scales in the solution and make the condition $g(t) \geq 0$ to be satisfied. The further development of this approach was done, f.e., in [3,7].

The specific features of the pressure/rate deconvolution problem and sensitivity of the solution to noise in the input data stipulate application of the regularization technique. For example, the methods presented in [5] for approximation of solution of equation (1) include the Ivanov quasi-solutions algorithm and the Tikhonov regularization procedure together with an algorithm of gradient type for solving the regularized minimization problem with constraints (3).

In this paper, the same regularization algorithms are applied to solving equation (1), but instead of (3), the infinite system of constraints (2) is used to construct approximate solutions satisfying all *a priori* constraints (2). It is attained by expansion of the sought for solution on a system of specific base functions, for which all *a priori* constraints are automatically satisfied. Besides, the regularized minimization problem, which should be solved, has the simplest constraints in the form of non-negativity of the coefficients.

The paper has the following structure . Investigation of the convergence of the algorithms based on the quasi-solutions method and the Tikhonov regularization scheme is given in Section 2. Section 3 is devoted to numerical experiments on solving equation (1) for noised flow rate and pressure drop.

2 Regular algorithms using all *a priori* information on solution

2.1 Quasi-solutions method

Assume that the operator A in (1) acts from $L_2[0,T]$ into $L_2[0,T]$. The operator A is compact on this pair of spaces. Hence, the inverse operator A^{-1} is discontinuous It means that problem (1) is ill-posed and its solution is very sensitive to errors in the input data $q(t)$, $\Delta p(t)$. Real input data have errors of measurements, *i.e.*, instead of the exact data $q(t)$ and $\Delta p(t)$, the approximate functions $q_\varepsilon(t)$, $\Delta p_\delta(t)$ are given, for which the following estimates hold:

$$\| q - q_\varepsilon \|_{L_2} \le \varepsilon, \ \ \| \Delta p - \Delta p_\delta \|_{L_2} \le \delta . \tag{6}$$

In the sequel, we denote A_ε as the operator defined by relation (1) where q is replaced by q_ε. Let us consider the following system of functions:

$$h_k(t) = e^{-\lambda_k t}, \ k = 1,2,..., \ h_0 = 1, \ \sum_{k=1}^{\infty} \lambda_k^{-1} = \infty, \ \lambda_k > 0 . \tag{7}$$

It is known [8] (see Theorem 3.3.5) that system (7) generates a complete system in the space $L_2[0,T]$. Hence, any function $g \in L_2[0,T]$ can be approximated by a linear combination of functions from (7), *i.e.*,

$$g(t) = \lim_{n\to\infty} \sum_{k=0}^{n} c_k e^{-\lambda_k t} . \tag{8}$$

Note that for every $k = 0,1,2,...$, $h_k(k) = e^{-\lambda_k t}$ satisfies all constraints (2).

Lemma 1. Let $g(t) = \sum_{k=0}^{n} c_k e^{-\lambda_k t}$, where $\lambda_i > 0, \ \lambda_i \ne \lambda_j$ as $i \ne j$, and for g the infinite system of inequalities (2) is satisfied, *i.e.*,

$$(-1)^m g^{(m)}(t) \ge 0, \quad m = 0,1,2,.... .$$

Then for all $i = 1,2,...,n$ the coefficients $c_i \ge 0$.

Proof. Assume the contrary that there exists a coefficient $c_k < 0$. Then for $t_m = \dfrac{m}{\lambda_k}$ and the sequence

$\left\{(-1)^m g^{(m)}(t_m)\frac{e^m}{\lambda_k^m}\right\}$, $m=0,1,2,...$ we have the following relations

$$(-1)^m g^{(m)}(t_m)\frac{e^m}{\lambda_k^m}=(-1)^m\frac{e^m}{\lambda_k^m}\sum_{i=0}^{n}c_i(-\lambda_i)^m e^{-\lambda_i t_m}=\sum_{i=0}^{n}c_i\left(\frac{\lambda_i}{\lambda_k}\right)^m e^{\left(1-\frac{\lambda_i}{\lambda_k}\right)m}=\sum_{i=0}^{n}c_i z_i^m e^{(1-z_i)m}=\sum_{i=0}^{n}c_i\left(z_i e^{1-z_i}\right)^m,$$

where $z_i=\frac{\lambda_i}{\lambda_k}$. The function ze^{1-z} has maximum at $z=1$. For the numbers $i\neq k$ and $z_i\neq 1$, the inequalities $0<z_i e^{1-z_i}<1$ are valid. If $i=k$ then $z_k=1$ and $z_k e^{1-z_k}=1$. It means that

$$(-1)^m g^{(m)}(t_m)\frac{e^m}{\lambda_k^m}=\sum_{i=1,i\neq k}^{n}c_i\left(z_i e^{1-z_i}\right)^m+c_k\overset{m\to\infty}{\to}c_k<0,$$

therefore, there exists $\overline{m}$ such that $(-1)^m g^{(m)}(t_m)<0$ for $\overline{m}\geq m$. It contradicts to the conditions of Lemma 1.

Now because of Lemma 1, we may assume that the exact solution $\hat{g}(t)$ of equation (1) can be represented as

$$\hat{g}(t)=\sum_{k=0}^{n}\hat{c}_k e^{-\lambda_k t},\quad \hat{c}_k\geq 0, \tag{9}$$

Then it allows us to find approximate solutions of equation (1) by the quasi-solutions method [9]

$$\min\{\|A_\varepsilon g-\Delta p_\delta\|^2:\ \ g=\sum_{k=0}^{n}c_k h_k,\ 0\leq c_k,\ \sum_{k=0}^{n}c_k\leq\overline{c}\ \}, \tag{10}$$

where $h_k=e^{-\lambda_k t}$, $\sum_{k=0}^{n}\hat{c}_k\leq\overline{c}$.

Theorem 1. Let equation (1) be uniquely solvable and $\hat{g}(t)$ be its solution represented by (9). Let approximation conditions be satisfied. Then there exists a solution $g_{\varepsilon,\delta}(t)$ of the minimization problem (10) and the uniform convergence of functions and its derivatives convergence holds:

$$\lim_{\varepsilon,\delta\to 0}\max_{t\in[0,T]}|\,g_{\varepsilon,\delta}^{(m)}(t)-\hat{g}^{(m)}(t)\,|=0,\quad m=0,1,2,..., \tag{11}$$

Proof. The set represented in the form

$$Q=\{g(t):\ \ g(t)=\sum_{k=0}^{n}c_k e^{-\lambda_k t},\ \ 0\leq c_k,\ \ \sum_{k=0}^{n}c_k\leq c\}$$

is the uniformly bounded and closed set of monotone functions g. Therefore, Q is a compact set. From this fact and continuity of the operator A_ε, it follows that in problem (9) there exists a

minimizer $g_{\varepsilon,\delta}$. Completion of the proof for the convergence of $g_{\varepsilon,\delta}(t)=\sum_{k=0}^{n} c_{\varepsilon\delta,k}e^{-\lambda_k t}$ to the exact solution $\hat{g}(t)=\sum_{k=0}^{n}\hat{c}_k e^{-\lambda_k t}$ in the space $L_2[0,T]$ can be performed similarly as in [5]. It implies the convergence of the coefficients for any k: $c_{\varepsilon\delta,k}\to\hat{c}_k$ as $\varepsilon,\delta\to 0$. From this fact it follows that relations (11) are valid.

2.2 Tikhonov regularization technique

Numerical experiments with using *a priori* constraints (3) had shown [5] that for large errors in the flow rate $q(t)$ and pressure $p(t)$, the Tikhonov regularization method with the non-quadratic stabilizing functional

$$\Omega[g]=\int_a^T \tau^3[g''(\tau)/g(\tau)]^2\,d\tau,\quad a>0 \tag{12}$$

takes better into account the multi-scale structure of solution and gives more appropriate results.
Now assuming that relation (9) is fulfilled, consider the regularization method in the form

$$\min\{\|A_\varepsilon g-\Delta p_\delta\|_{L_2}+\Omega[g]:\ 0<b\le g=\textstyle\sum_{k=0}^{n} c_k e^{-\lambda_k t},\ 0\le c_k,\ \sum_{k=0}^{n} c_k\le\overline{c}\} \tag{13}$$

where $\sum_{k=0}^{n}\hat{c}_k\le\overline{c}$ and, in contrast to (12), the stabilizing functional is expressed by the formula

$$\Omega[g]=\int_a^T[\tau^{3/2}g''(\tau)/g(\tau)-\tau^{1/2}g'(\tau)/g(t)+\tau^{3/2}(g'(\tau)/g(\tau))^2]^2\,d\tau,\quad a>0. \tag{14}$$

Let us explain why such a type of the stabilizer is taken. As it was noted in Introduction about the transition from linear equation (1) to nonlinear problem (4), the replacements (5) are used. Therefore, we need to pass from the interval $[0,T]$ to one $[a,T]$, where $a>0$. Note that in real computations, usually $\ln a>0.001$.
Now, if for the nonlinear deconvolution problem the stabilizing functional is defined by the traditional way [10]

$$\Phi[z]=\int_{\ln a}^{\ln T}[z''(\sigma)]^2\,d\sigma \tag{15}$$

and in (15) the inverse replacement is carried out, then we arrive to functional (14).
Theorem 2. Let the premises of Theorem 1 be fulfilled. Then for $\alpha>0$ problem (13) is solvable, possibly in a non-unique way. If $\alpha(\varepsilon,\delta)\to 0$, $(\delta+\varepsilon)^2/\alpha(\varepsilon,\delta)\to 0$ as $\varepsilon,\delta\to 0$, then the extreme functions $g^{\alpha(\varepsilon,\delta)}$ converge to the exact solution $\hat{g}$ in the space $C[a,T]$ together with their derivatives, *i.e.*,

$$\lim_{\varepsilon,\delta\to 0}\max_{t\in[a,T]}\left|\frac{d^m}{dt^m}g^{\alpha(\varepsilon,\delta)}(t)-\frac{d^m}{dt^m}\hat{g}(t)\right|=0,\quad m=0,1,2,\ldots. \tag{15}$$

Proof. Modifying a little the technique of the proof from [5] (see Theorem 5 there), one can prove solvability of problem (13) and the convergence of the minimisers $g^{\alpha(\varepsilon,\delta)}$

$$\lim_{\varepsilon,\delta\to 0} \max_{t\in[a,T]}\{|\, g^{\alpha(\varepsilon,\delta)}(t)-\hat{g}(t)\,|, |\, dg^{\alpha(\varepsilon,\delta)}(t)/\,dt-d\hat{g}(t)/\,dt\,|\}=0\,. \qquad (16)$$

Taking into account that the functions $g^{\alpha(\varepsilon,\delta)}$, $\hat{g}$ have expansions of form (9), we obtain that (16) implies

$$\lim_{\varepsilon,\delta\to 0} |\, c_k^{\alpha(\varepsilon,\delta)}-\hat{c}_k\,|=0 \quad \text{for any } k=0,1,2,\ldots,$$

hence, also the relations (15) hold.

3 Numerical solution of the deconvolution equation

3.1 Discrete approximation and solving algorithms

To solve equation (1) numerically, we need to use discrete approximation of the integral operator A. For this let us take a collocation method with a grid $\{t_i\}_{i=1}^m$, where t_i are time points of pressure measurements in the well. A grid $\{\overline{\tau}_j\}_{j=0}^{n-1}$ for the variable τ is constructed in the form:

$$\ln\overline{\tau}_j=\ln\overline{\tau}_0+\frac{j}{n-1}(\ln\overline{\tau}_{n-1}-\ln\overline{\tau}_0),\quad j=0,1,\ldots,n-1\,,$$

i.e., the grid is uniform in the logarithmic scale. As above, the desired solution is presented as $g_k(\tau)=\sum_{j=0}^k c_j e^{-\lambda_j\tau}$, where the parameters λ_j are chosen from the relations $e^{-\lambda_j\overline{\tau}_j}=\gamma$ for a certain $\gamma>0$ (in numerical experiments γ=0.01).

The integral operator A of problem (1) is approximated by the following relations:

$$(A_{mk}c)_i=\int_0^{t_i} q_\varepsilon^n(t_i-\tau)g^k(\tau)d\tau=\sum_{j=0}^k c_j\int_0^{t_i} q_\varepsilon^n(t_i-\tau)e^{-\lambda_j\tau}d\tau,\ \ i=1,2,\ldots,n\,,$$

where $c=(c_0,c_1,\ldots,c_k)^T$, q_ε^n is the piecewise linear function constructed on the grid $\{\tau_j\}_{j=1}^n$ with values $q_\varepsilon^n(\tau_i)$. Now problem (10) is reduced to the minimization problem

$$\min\{\|\, A_{mk}\,c-p\,\|_{l_2^m}^2:\ \ c_j\ge 0,\ \sum_{j=0}^k c_j\le\overline{c}\}\,, \qquad (17)$$

where $\|\,p\,\|_{l_2^m}^2=\sum_{i=1}^m \mu_i\,p_i^2$, $\mu_i=t_i-t_{i-1}$, $t_0=0$, $p_i=\Delta p_\varepsilon(t_i)$. Problem (17) is resolved by the method of conditional gradients ([11], Chapter 5, Section 4).

Remark 1. Under large errors in input data ($q_\varepsilon(t)$ up to 15 % and $\Delta p_\delta(t)$ up to 5 %) sometimes for the quasi-solutions method necessity arises in an additional correction of the function $q_\varepsilon(t)$. It can be done by a special procedure described in [5].

Remark 2. Discrete approximation in the Tikhonov method is carried out like the quasi-solution method. After discretization of (13), the minimization problem can be solved by the conjugate gradient method [11].

3.2 Numerical experiments

The quasi-solutions algorithm was tested for synthetic data (Figure 1) taken from the paper [5]. Functions $p(t)$ and $q(t)$ were given on a grid with 5% levels of noise. In this case the exact solution, *i.e.,* the pair $(p_u(t), dp_u(t)/d\ln t)$ is known. Therefore, it allows us to estimate quality of numerical solution obtained by the constructed algorithm. In Figures 2 and 3 the results of numerical solution of equation (1) are presented on log-log scale for exact data and pressure $p(t)$ with error 5% respectively. In Figure 4 the numerical solution is shown for input data $p_\delta(t), q_\varepsilon(t)$ perturbed by 5% and 1% noise correspondingly. In contrast to [5] we have obtained more smooth solution that is important for well test interpretation. Accuracy of numerically retrieved solution for exact and noisy data is quite satisfactory. For input data $p_\delta(t), q_\varepsilon(t)$ up to 5% level of noise and noiseless p_0 the quasi-solutions method gives sufficiently smooth approximate solution but we can see that the solution is rather sensitive to errors in input data. To obtain a reasonable numerical solution for 5% noise in p_0, $p_\delta(t)$ and 15% nose in $q_\varepsilon(t)$ it is required to use the Tikhonov method, which provides the greater regularizing effect. Besides, in this situation we need in additional correction of the function $q_\varepsilon(t)$ (see [5]). Thus, the numerical experiments showed the efficiency of the proposed methods. It is achieved by using all *a priori* information on the desired solution and choosing a special basis of smooth functions $h_i(t)$, which satisfy the infinite system of inequalities (2).

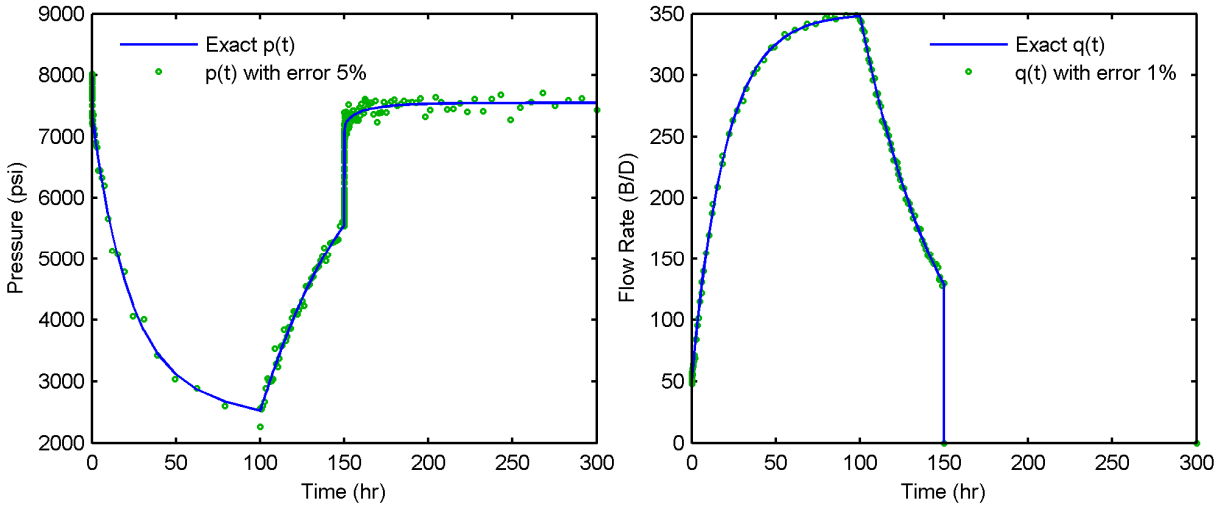


Figure 1: Synthetic data for pressure and flow rate.

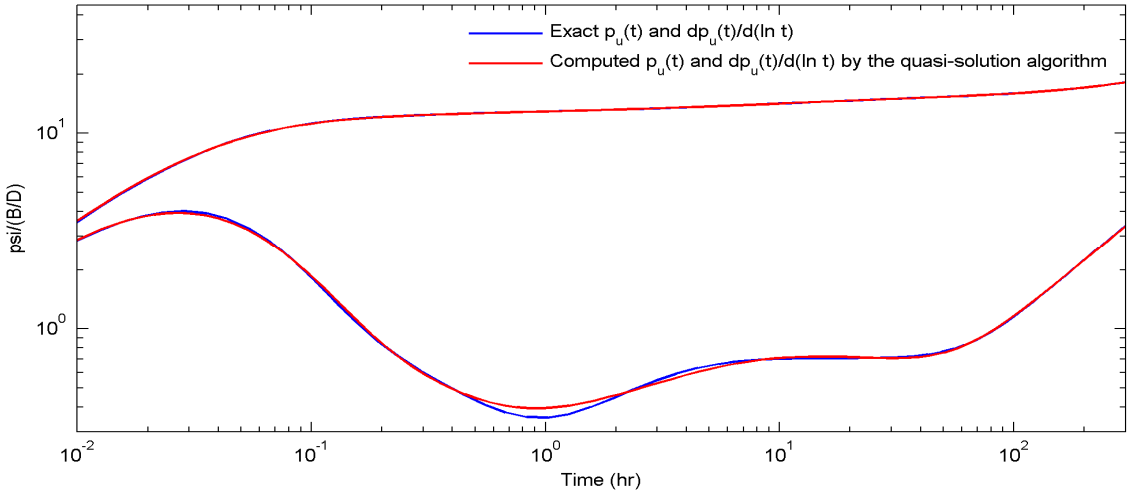


Figure 2: Solution by the quasi-solutions method with exact $p(t)$ and $q(t)$.

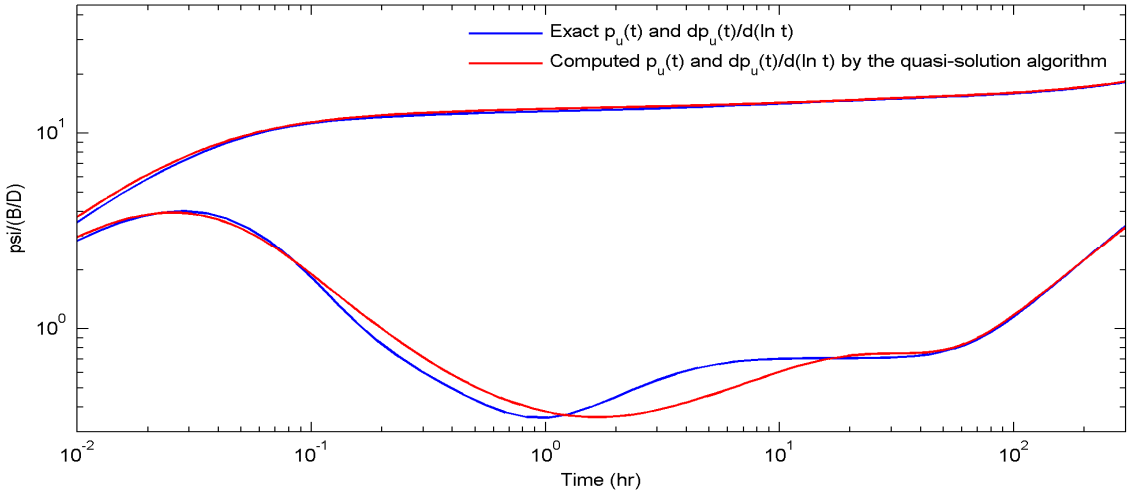


Figure 3: Solution by the quasi-solutions method with 5% error in $p(t)$ and exact $q(t)$.

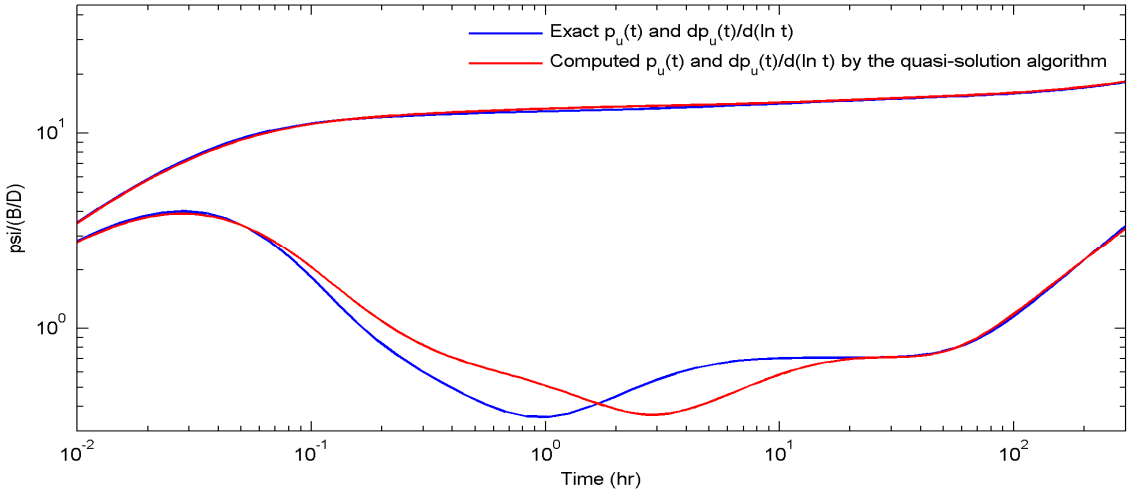


Figure 4: Solution by the quasi-solutions method with 5% error in $p(t)$ and 1% in $q(t)$.

Acknowledgements
The work was supported by the RFBR grant, project no.12-01-00106 and the Ural Branch of RAS, project no.12-П-15-2019.

References

[1] A. F. von Everdinger and W. Hurst, *The application of the Laplace transformation to flow problem in reservoir,* Transaction of American Institute of Mining, Metallurgical and Petrolleum Engineers, 186, (1949), pp. 305-324.

[2] D. Bourdet, J. A. Ayoub, and Y.M.Pirard, *Use of pressure derivative in well-test interpretation,* SPE Formation Evaluation, SPE Paper 12777, (1989), pp. 293-302.

[3] E. A. Pimonov, M. Onur, and F. J. Kuchuk, *A new robust algorithm for solution of pressure/rate deconvolution problem*, Journal of Inverse and Ill-Posed Problems, 17, (2009), pp.611-627.

[4] K. H. Coats, L. A. Rapport, J. R. McCord, W.P. Drews, Determination of aquifer influence function from field data, Transactions of SPE (AIME), 231, (1964), pp. 1417-1424.

[5] V. Vasin, G. Skorik, E. Pimonov, and F. Kuchuk, *New regular algorithms for solution of the convolution problem in well test interpretation*, Journal of Inverse and Ill-Posed Problems, Applied Mathematics, 1, (2010), pp. 387-399.

[6] T. Von Schroeter, F. Hollaender, A. Gringarten, *Deconvolution of well test data analysis as a nonlinear total least squares problem,* SPE Annual Technical Conference and Exibition, New Orlean, Louisiana, 30 September – 3 October, (2001), SPE 71574.

[7] M. M. Levitan, *Practical application of pressure-rate deconvolution to analysis of real well test,* SPE Reservoir Evaluation & Engineering, 8 (2005), pp. 113-135.

[8] A. F. Leontiev, *Sequences of polynomials from exponents*, Nauka, Mocow, (1980). (in Russian)

[9] V. K. Ivanov, V. V. Vasin and V. P. Tanana*, Theory of Linear Ill-Posed Problems and its Applications,* VSP, Utrecht etc, 2002.

[10] A. N. Tikhonov and V.Y. Arsenin, *Methods of Solving Ill-Posed Problems*, "Nauka English Translation, Moscow; Wiley, New York, (1977).

[11] F. P. Vasilyev, Numerical Methods for Solving Extremal Problems, Nauka, Moscow, (1988). (in Russian)

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

The inverse heat conduction problem solution for a laser flash studies of a thin layer coatings

W. Stryczniewicz[1,2*], **J. Zmywaczyk**[2], **A. J. Panas**[2]

[1)] Applied Aerodynamics Laboratory
Institute of Aviation, Krakowska 110/114, 02-256 Warsaw, Poland

[2)]Faculty of Mechatronics and Aeronautics
Military University of Technology, gen. S. Kaliskiego 2, 00-908 Warsaw, Poland
e-mail: Wit.Stryczniewicz@ilot.edu.pl, Andrzej.Panas@wat.edu.pl

Key words: inverse heat convection problem, laser flash experiment, multi layer structure, thin layer thermal diffusivity, flake graphite coating

Abstract
An issue of the inverse heat conduction methodology in solution of a certain parameter identification problem has been discussed. The problem itself concerns determination of the thermophysical properties of a thin layer coatings by the laser flash experimental investigation of the thermal diffusivity. The analysed model coating was a flake graphite thin layer conventionally applied in the laser flash experiments for the specimen surface emissivity enhancement. The paper presents a methodology developed for a three layer specimen studies. The procedure assumes identification of the out of plane layer diffusivity from the inverse problem solution. The flake graphite heat capacity and density are supposed to be known from other experiments. The procedure utilises a finite element method model for the direct problem analyses and multi-parametrical identification using Levenberg-Marquardt algorithm. The presentation is restricted to the numerical modelling only. Typical results of the performed numerical studies illustrate performance of the developed procedure.

1 Introduction

Advancements of last two decades in laser excitation techniques and contactless measurements of the temperature resulted in development of robust and accurate thermal properties measurement devices [1]. Particularly the development of the high power impulse lasers and sensitive infrared signal detectors helped to take full advantage of the simplicity of procedure introduced by Parker [2], [3]. Application of a modern flash apparatus involves radiative excitation and contactless measurements of the specimen temperature changes in course of the experimental study [4]. In order to improve both the thermal excitation and the thermal response infrared signal recording conditions special procedures of the specimen surface emissivity enhancement are usually applied. Typically a flake graphite coatings are utilised. Although the standard absorbing graphite layer is extremely thin, its presence still influences the thermal diffusivity investigation result. Because of that several studies have been performed aiming to determine the final effect of the layer existence [5, 6]. However, the

issue of the thermal properties of the layer material itself is still being open. In the course of a present study a methodology of the thin coating layer thermal properties (TP) determination has been developed. The flake graphite layer density and the heat capacity are determined in complementary measurements. The thermal diffusivity is identified applying results of a laser flash experiments performed on a three layer specimen. The identification method involves the inverse heat conduction problem solution. The present paper covers the issue of the inverse problem formulation [7], solution of the associated direct heat conduction problem applying FEM Comsol software and multi-parametrical identification using Levenberg-Marquardt algorithm. The identified parameters - properties are: the planar heat source density, the layer material out of plane thermal diffusivity, and, finally, the heat transfer coefficient for the convection heat losses from the investigated specimen. Typical results of the performed numerical studies are shown in the paper. These results illustrate performance of the developed algorithm. The algorithm itself will be applied in evaluation of the directional, out of plane thermal diffusivity of the Graphite 33 coating from the laser flash experimental data.

2 Problem definition

Transient heat transfer in solids is govern by a Fourier equation in a form

$$\rho c_p \frac{\partial T}{\partial \tau} = \nabla \cdot (k \nabla T) + q_V \tag{1}$$

Where ρ - material density, c_p- heat capacity at constant pressure, T - temperature, k - thermal conductivity of a material, τ - time and q_V stands for external heat sources. Assuming lack of internal heat sources ($q_V = 0$) above equation can by transformed into following relation:

$$\frac{\partial T}{\partial \tau} = \frac{k}{\rho c_p} \Delta T = a \Delta T \tag{2}$$

where a - is a thermal diffusivity of a material. This relation assumes temperature independent values of thermophysical properties. Accounting for it most of the thermophysical property investigation procedures were developed, in that range several methods for the thermal diffusivity studies. However, it is a common practice to apply the extended relation

$$a(T) = \frac{k(T)}{\rho(T) c_p(T)} \tag{3}$$

for the experimental data processing [1].
The thermal diffusivity of the material in the flash method proposed by Parker is determined from the thermal response measured at back surface of the sample exposed to impulse excitation at the front surface (Fig.1.a). The analysis of the temperature response of the sample leads to simple relation between the thickness of the sample l, the time at the value of the temperature reach half of its maximum value $t_{0,5}$ and the thermal diffusivity of the material:

$$a = \frac{1{,}38 l^2}{\pi^2 t_{0,5}}, \tag{4}$$

Nowadays, however more sophisticated models are utilised for analysis of the sample thermal response. In order to increase the accuracy of the measurements the effect of finite laser pulse time [3], radiation heat losses and the sample internal structure needs to be included [8]. Therefore a variety of models have been developed in a course of thermal diffusivity investigation by the flash method. Moreover, when processing the data many parametric estimation methods are usually used.

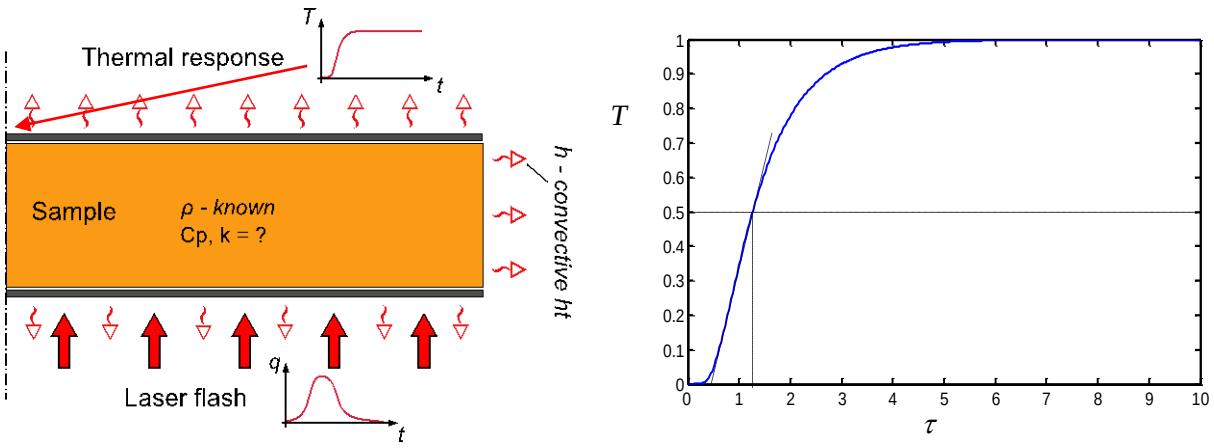


Fig.1. Scheme of the analysed laser flash experiment performed on the graphite coated specimen (a) and typical system response (b; abscissa and ordinate data in arbitrary units)

Although modern laser flash apparatus provide robust results in investigations of thermal diffusivity of most materials, special sample treatment procedures are necessary for samples with highly reflective surfaces. Therefore, it is a common practice to improve the thermal excitation and infrared signal recording conditions of the thermal response by specimen surface emissivity enhancement. Typically front and rear surface of the sample is covered with flake graphite coating [4]. The thickness of typical sample ranges from 5 to 20 mm, whereas the thickness of typical absorbing layer does not exceed a two dozens of micrometers. Therefore the influence of thermal resistance of absorbing layer are usually assumed to be negligible. Nevertheless for highly conductive materials and for very thin samples the presence of additional layers influences the thermal diffusivity investigation result [5]. In scope of determination of this effect several researches have been performed. The interest of researchers has focused on determination of the thermal diffusivity measurement error, that the layer introduces, in correlation to sample thickness and thermal properties of investigated material. In order to minimise the error caused by the graphite layer, Lim et al. proposed the correction procedure based on three-layer model [6]. Although efforts to improve the measurement procedure of covered samples has been made, the issue of determination of thermal diffusivity of the coating itself is still being open.

Multi-layered samples are commonly used in determination of thermal properties of transparent or highly conductive materials by flash method. Typically two-layered sample are utilised. The layer of investigated material is deposed on a layer of material of know thermal diffusivity. The unknown thermal diffusivity is determined form the experimental thermal response of the sample with use of mathematical model of heat transfer in two-layered samples exposed to surface impulse excitation [8]. In standard procedure the thickness of both layers is similar [4], [6].

Unfortunately, due to a a huge disproportion between the thin layer thickness and the thickness of the base specimen the standard procedures of the two-layered samples data processing revealed to be ineffective for determination of thermal properties of a thin layer covering the sample. Therefore, an alternative methodology based on solution of inverse heat transfer problem was proposed.

3 Method

3.1 Inverse problem

In the proposed methodology the thermal diffusivity of a thin layer is estimated by fitting of a numerical model response T (Fig. 1. b) to the experimental data Y (Fig. 1. a). The solution of this inverse heat problem is based on the minimization of the ordinary least squares norm in a form:

$$S(P) = \sum_{i=1}^{I} \left[Y_i - T_i(P) \right]^2 , \tag{2}$$

where: S - is a sum of squares error, P - vector of unknown parameters, $T_i(P)$ - estimated temperature at time t_i, Y_i – "measured" temperature at time t_i. In order to simulate the laser flash excitation and convective heat transfer between the ambient and the sample it is necessary to know the planar heat source density q and the convective heat transfer coefficient k. Unfortunately, it is impossible to measure this parameters during actual experiment. Therefore, the multi-parametrical identification needs to be performed. The vector of unknown parameters of the model P=[k,h,q] is estimated by the Levenberg-Marquardt algorithm by utilizing the iterative procedure in the form:

$$P^{n+1} = P^n + [(J^n)^T J^n + \mu^n \Omega^n]^{-1} (J^n)^T [Y - T(P^n)], \tag{3}$$

where: n - is a iteration number, μ^n - is a positive scalar named damping parameter, Ω^n - is a diagonal matrix and J- is a matrix of the sensitivity coefficients in the form:

$$J(P) = \left[\frac{\partial T^T(P)}{\partial P} \right]^T , \tag{4}$$

Generally the inverse heat transfer problems are very ill-conditioned around the initial guesses used in the iterative procedure. This causes the oscillations and instabilities in the iterative estimation algorithms. Therefore, in order to damp this perturbations of the solution, in the Levenberg-Marquardt procedure the damping parameter μ^n is made large in the beginning of iterations and it is gradually reduced while the identification procedure advances. In the developed algorithm the stopping criteria of the estimation procedure is following:

$$S(P^k) = \sum_{i=1}^{I} \left[Y_i - T_i(P^k) \right]^2 < \varepsilon , \tag{5}$$

where ε - is a user prescribed parameter.

3.2 Sensitivity coefficient analysis

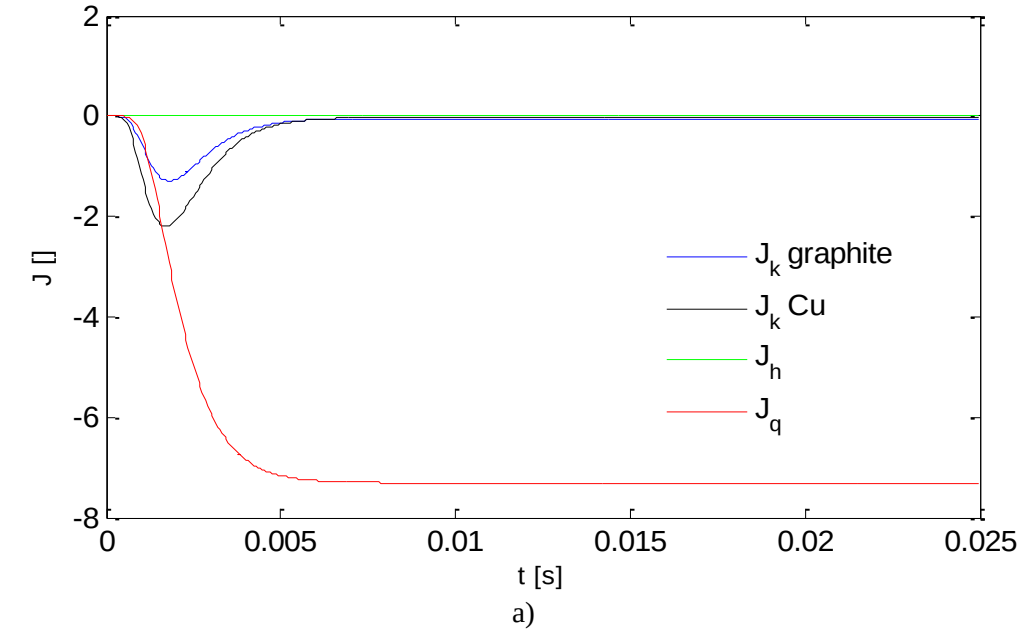
In order design optimum experiments and to estimate the influence of measurement errors it is important to analyse the sensitivity coefficients for each estimated parameter. The sensitivity coefficients are defined as a first derivative of the estimated temperature T_i with respect to the unknown parameter P_j [7]

$$J_{ij} = \frac{\partial T_i}{\partial P_j} , \tag{6}$$

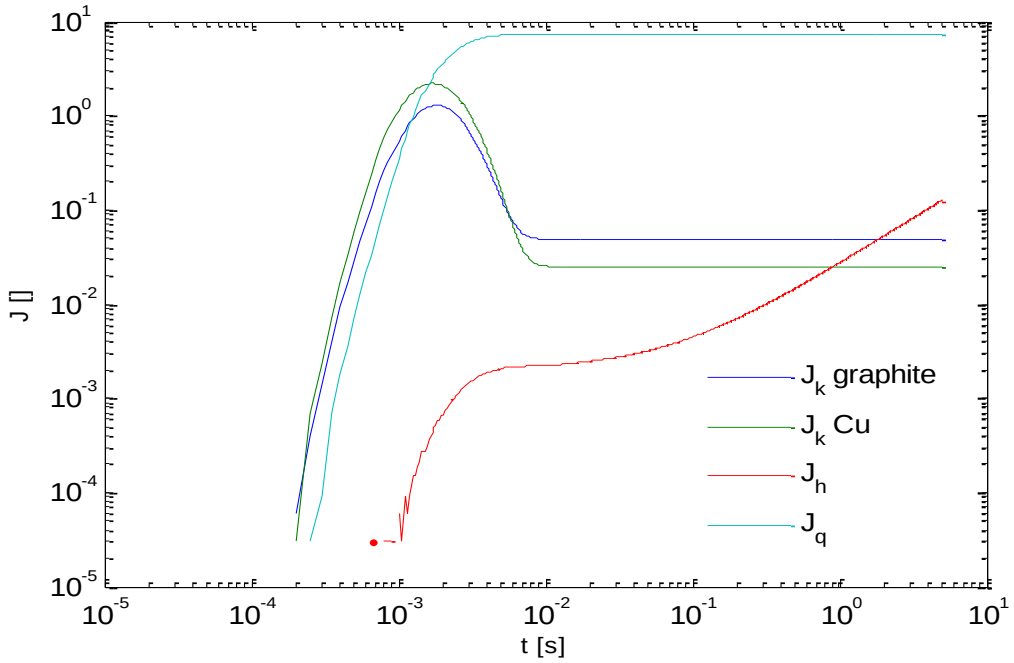
For small values of the sensitivity coefficients and for their linear dependence the inverse problem are ill-conditioned. Therefore it is desirable to design the experiments in such a way that absolute value of

the sensitivity coefficients are as high as possible and they are linearly independent. This can be determined from the timewise variations of the sensitivity coefficients.
The sensitivity coefficients for the analysed inverse heat transfer problem were determined as follows:

$$J_k = \frac{T_{k+\Delta k} - T_k}{\Delta k}, \qquad J_h = \frac{T_{h+\Delta h} - T_h}{\Delta h}, \qquad J_q = \frac{T_{q+\Delta q} - T_q}{\Delta q}, \tag{7}$$



a)



b)

Fig. 2. Results of the sensitivity coefficient calculation in the time scale of the real experiment (a) and in the extended time scale (b)

Typical results of calculations are depicted in Fig. 2. Analysing the displayed data one can easily notice that the problem is well conditioned in view of the heat source planar density *q* estimation within almost the whole time interval. The conditioning of the layer thermal conductivity *k* identification is relatively good, as was expected, for the initial part of the "experimental". However, proper estimation of the heat transfer coefficient *h* for this particular (*q*, *k*, *h*) parameter vector is possible only on the basis of the time extended thermal response analysis. Separation of domains of the nonzero values of the sensitivity coefficients foretells optimistically for the analyses of real experimental results that are planned to be performed.

3.3 Numerical model for direct problem solution

The numerical model was developed applying Finite Element Method (FEM) Comsol/Multiphysics software. The sample in the flash method has a shape of a disc. Therefore it was possible to take advantage form the axial symmetry and simplify the geometry of the model. The axisymmetric 2D model of a three layer specimen was developed accounting for real specimen dimensions in the Heat Transfer module. Since in the experiment there is strong temperature gradient in the sample between the top and the bottom surface and one direction of the heat transfer dominate, the structural mesh in all domains was chosen (Fig. 3.). The thermophysical properties of the graphite and copper layers the subdomains were set as shown in Table 1.

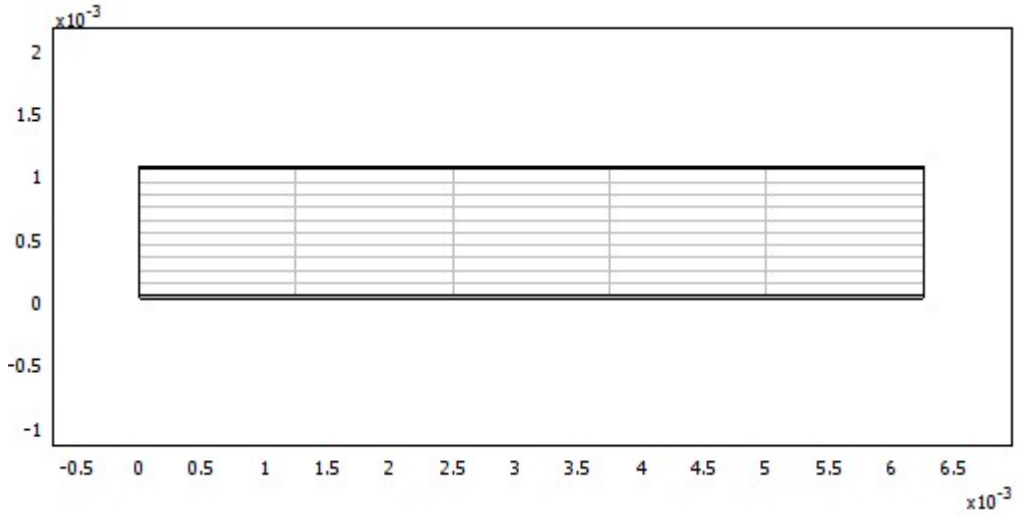


Fig. 3. Axisymmetric geometry of the analysed multilayer specimen - dimensions are indicated in [mm]. The density of the horizontal meshing is equal for all three domains. In the vertical direction all there domains are equally subdivided for 10 elements.

Table 1. The assumed values of the thermophysical properties of the analyzed multilayer specimen

Subdomain	Density $\rho[\mathrm{kg}\cdot\mathrm{m}^{-3}]$	Specific heat $c_p[\mathrm{J}\cdot\mathrm{kg}^{-1}\cdot\mathrm{K}^{-1}]$	Thermal conductivity $k[\mathrm{W}\cdot\mathrm{m}^{-1}\cdot\mathrm{K}^{-1}]$
Cooper – basic middle domain	8900	385	400
Graphite – coating	800	700	0,1 ; 1; 10

The following boundary conditions was set: axial symmetry on boundary simulating the symmetry axis of the disc sample and heat flux condition on three other boundaries in order to simulate convective heat transfer between the sample and the ambient fluid. Additionally, the laser flash excitation was simulate the on the bottom surface of the sample with the Inward heat flux condition. The time of the surface excitation was set to be equal to the time of laser pulse in real experiment. Thin thermally resistive layers boundary condition was set on the internal boundaries between the graphite and the cooper subdomains in order to simulate the thermal contact resistance. The numerical solver was set to follow strictly the assumed time steps.

3.4 Procedure of the inverse procedure testing

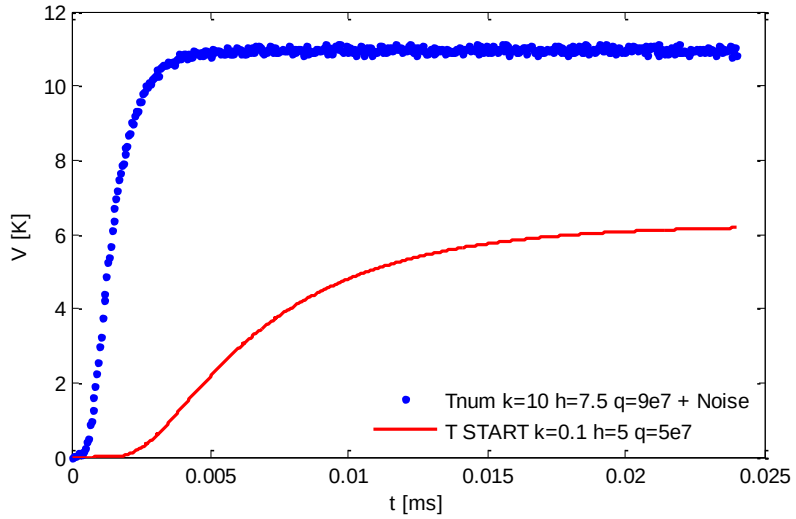
In order to validate the performance of the algorithm the experimental data was simulated by disturbing the numerical model data. The white Gaussian noise was added the temperature response of the FEM model. The disturbed data simulated the input experimental data Y of the algorithm. The performance of the algorithm was tested for different set of parameters of the simulated experimental data to be estimated and for different initial guesses. Different levels of the Signal to Noise Ratio (SNR) ranging from 5 to 60 dB was tested.

4 Results and discussion

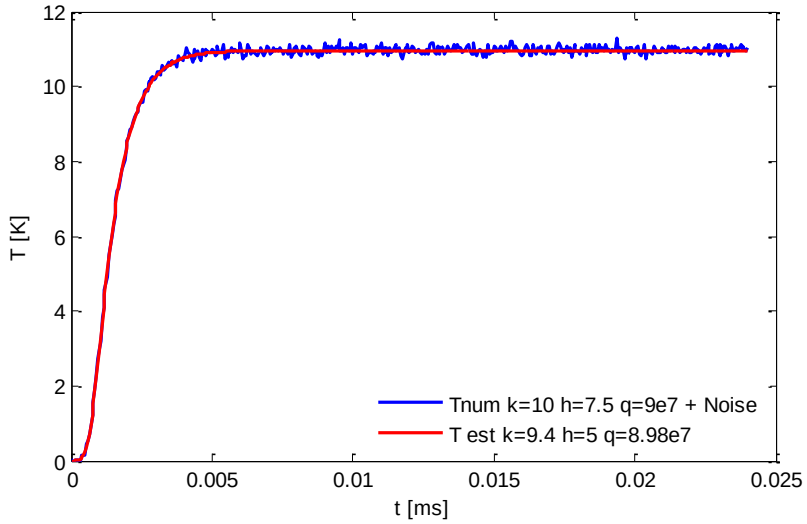
The developed procedure of parametric estimation was tested for several sets of the assumed thermophysical parameters of the analyzed system and for several different amplitudes of the imposed noise. Typical results in the form of three selected tests outcome are shown in Table 2. These tests were conducted applying a 30 dB SNR noise imposed onto simulated thermal response signal (comp. Fig. 1.b). A supplementary illustration of the data fitting result is provided in Fig. 4 while the evolution of the Levenberg-Marquard iteration procedure result in Fig. 5.
As was expected, in the analyzed time scale of the input data the procedure proved to be effective in estimation of the heat flux density and the thermal conductivity of the model graphite layer. This does not concern the heat transfer coefficient. However, this parameter is not of a primary importance in view of the planned experimental investigations.
Concluding the discussion it is worth to mention that the algorithm has proved to be effective in estimation of the layer thermal conductivity ranging for two orders of magnitude. The expected value of the transversal component of the flake graphite thermal conductivity seems to be much lower than that reported for pyrolitic graphite (comp. eg. [9] and [10]). According to the available data (comp. e.g. [6]) it should be comprised within the interval of the analyzed thermal conductivity values.

Table 2. Comparison between the assumed parameter values and those obtained from the inverse problem solution when analyzing the Gaussian white noise affected "laser flash" response data

	Direct problem - assumed	Inverse - initial guess	Inverse - results
	k h q	k0, h0, q0	k_est h_est q_est
Test 1	k=1, h=5, $q=9\cdot10^7$	k0=10, h0=5, $q=4\cdot10^7$	k_est=0.997, h=4.67, $q=8.90\cdot10^7$
Test 2	k=0.1, h=5, $q=9\cdot10^7$	k0=10, h0=5, $q=4\cdot10^7$	k_est=0.1, h=5, $q=8.29\cdot10^7$
Test 3	k=10, h=7.5, $q=9\cdot10^7$	k0=0.1, h0=5, $q=5\cdot10^7$	k_est=9.4, h=5, $q=8.98\cdot10^7$

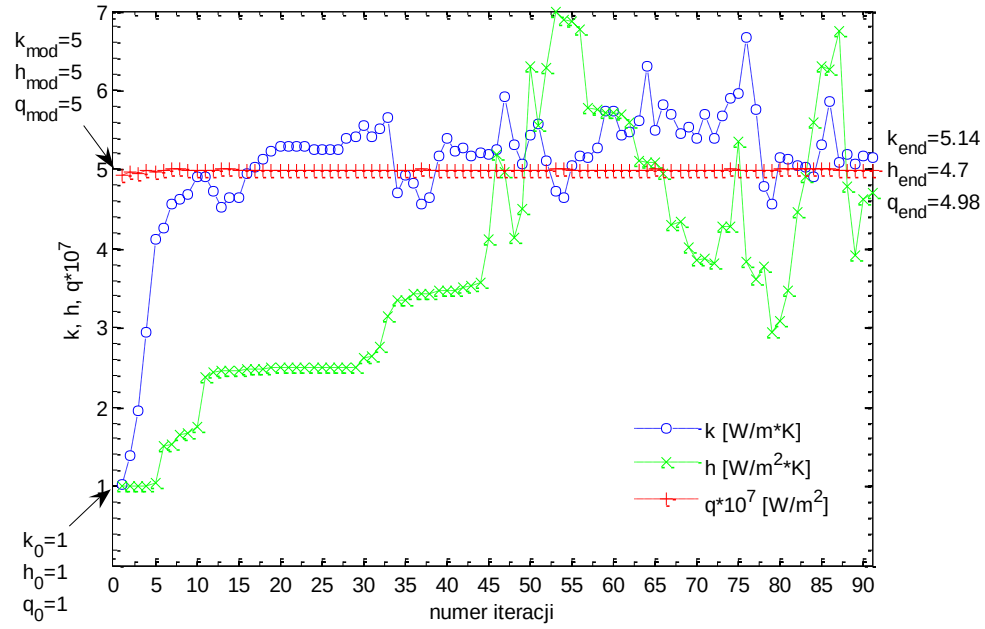


a)

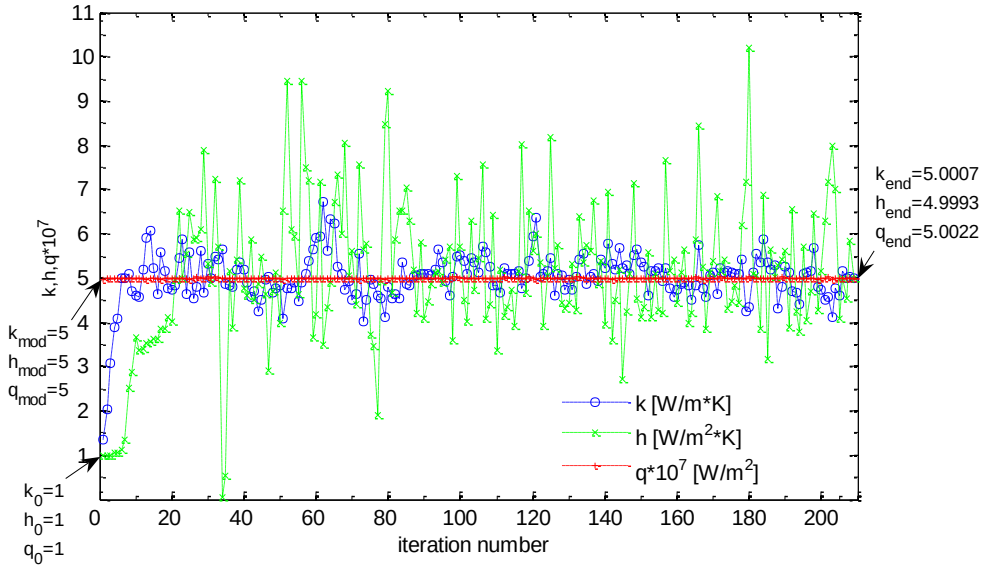


b)

Fig. 4. Typical data fitting result of the Test 3 (see also Table 2): a -comparison between the model noisy signal data (indicated by blue coloured points; the signal to noise ratio was equal to 30 dB) and the initial guess (red line), b - iillustration of the estimation result



a)



b)

Fig. 5. Evolution of the estimated parameters at the beginning of the iteration procedure (a) and illustration of the thin layer thermal conductivity *k* and the heat transfer coefficient *h* oscillation for the increased iteration number (b)

5 Summary

In the course of a present study a methodology of the thin coating layer thermal properties (TP) determination has been developed. The procedure has been designed for the flake graphite studies. The layer material density and the heat capacity should be determined in complementary measurements. The thermal diffusivity is identified applying results of a laser flash experiments performed on a three layer specimen. The identification method involves the inverse heat conduction problem solution. The present paper has covered the issue of the inverse problem formulation, solution of the associated direct heat conduction problem applying FEM Comsol software and multi-parametrical identification using Levenberg-Marquardt algorithm. The identified parameters have been: the planar heat source density, the layer material out of plane thermal diffusivity and the heat transfer coefficient for the convection heat losses from the investigated specimen. Typical results of the performed numerical studies are shown in the paper. This results illustrates performance of the developed algorithm that will be applied in ongoing investigations.
The algorithm has proved to estimate the selected parameters well. It is also worth to notice that the combination of the FEM model and multi-parametrical identification in one algorithm opens up possibilities to estimate large variety of parameters.
The developed procedure will be used to determine the thermal diffusivity of thin graphite coatings deposed on a sample form a laser flash experimental data.

References

[1] S. Min, J. Blumm, A. Lindemann, *A new laser flash system for measurement of the thermophysical properties,* Thermochimica Acta, 455, (2007), pp. 46-49.

[2] J. W. Parker, R. J. Jenkins, C. P. Butler, G. L. Abbott, *Flash Method of Determining Thermal Diffusivity, Heat Capacity, and Thermal Conductivity*, Journal of Applied Physics, 9, (1961), pp. 1679-1684.

[3] J. A. Cape, G. W. Lehman, *Temperature and Finite Pulse-Time Effects in the Flash Method for Measuring Thermal Diffusivity*, Journal of Applied Physics, 7, (1963), pp. 1909-1913.

[4] LFA apparatus manual, Netzsch, Germany, (2013).

[5] F. Cernuschi, L. Lorenzoni, P. Bianchi, A. Figari, *The effects of sample surface treatments on laser flash thermal diffusivity measurements*, Infrared Physics & Technology, 43, (2002), pp. 133-138.

[6] S. Kim, Y. Kim, *Determination of apparent thickness of graphite coating in flash method,* Thermochimica Acta, 468, (2008), pp. 6-9.

[7] M. N. Ozisik, H. R. B. Orlande, *Inverse Heat Transfer*, Taylor&Francis, New York, United States of America, (2000).

[8] K. B. Larson, K. Koyama, *Measurement by the Flash Method of Thermal Diffusivity, Heat Capacity, and Thermal Conductivity in Two-Layer Composite Samples*, Journal of Applied Physics, vol. 39 No. 9., (1986), pp. 4408-4416

[9] A. J. Panas, *IR Support of Thermophysical Property Investigation. Medical and Advanced Technology Materials Study*, Chapter 4 in *Infrared Thermography*, Raghu V. Prakash ed., Intech, (2011), pp. 65-90

[10] Material Property Database MPDG v.7.08, 2009, JAHM Software, Inc., USA

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

MPCA for Flight Dynamics Parameters Determination

Ivana Y. Sumida[1], Haroldo F. de Campos Velho[1], Eduardo F. P. Luz[1], Ronaldo V. Cruz[2], Luiz Carlos S. Góes[2]

[1]National Institute of Space Research
São José dos Campos, Brazil
e-mail: ivanayoshie@yahoo.com.br, haroldo@lac.inpe.br, eduardofpl@gmail.com

[2]Technological Institute of Aeronautics
São José dos Campos, Brazil
e-mail: ronaldocruz@hotmail.com, goes@ita.br

Key words: dynamic flight, parameter identification, multiple particle collision algorithm

Abstract

Aircraft have become increasingly costly and complex. Military and civil pilots and engineers have used flight simulators in order to increase the security through the training of crew. It is necessary to calibrate the simulation for simulators to have good adherence to reality, that is, to identify the parameters that make the simulation as close as possible to the real dynamic. After determining these parameters, the simulator will be ready to be used in human resources training or assessing aircraft. Parameter identification characterizes the aerodynamic performance of the aircraft and can be formulated as a problem optimization. The calibration of a dynamic flight simulator is achieved by a new meta-heuristic: Multiple Particle Collision Algorithm (MPCA). Preliminary results show a good performance of the approach employed.

1 Introduction

The dynamics of flight are critical for aircraft design, looking at the adaptation of automatic control systems and flight simulators. Identification techniques applied to dynamics flight has had an improvement in the increasing availability of faster computers.

The identification of parameters methodology has various applications areas (eg astronomy, aerospace, economics, biology, electrical, geological, etc.) [1, 2, 3]. The strategy to adjust the unknown parameters in order to have the best fit of a mathematical model of the phenomenon with the observations. Tools and techniques of identification have evolved to match the complexity and the increasing need for correction and precision in the results. This methodology is more accurate than the corresponding values

predicted by other methods such as analytical and numerical differentiation [4, 7]. Thus, identification of parameters has become a significant tool for applications such as model validation, handling qualities evaluation, control law design, and flight-vehicle design and certification [6].

Specifically for helicopter parameter estimation, Hamel and Kaletka [8] presented a general vision of the progress in this field up to 1997 and Padfield [5] described a comprehensive flight dynamic theoretical model, flight qualities criteria development, flight test techniques, and several results of this research in the United Kingdom.

It is important to notice that most works use local optimization algorithms based on gradient search methods such as Gauss-Newton and Levenberg-Marquadt methods for finding a local minimum of the prediction error function at the system output. Concerning the use of global optimization algorithms and, more specifically, stochastic algorithms (e.g. genetic algorithm (GA)) have been used by Hajela and Lee [9] in rotor blade design, by Wells et al. [10] in the acoustic level reduction rotor design, and by Zaal et al. [11] in the estimation of parameters multichannel pilot models, among others.

Regarding helicopter system identification techniques, very few articles have used GA for global optimization of a cost function based on the prediction error. In this framework, one can cite Cruz et al. [14, 15] in the longitudinal mode system identification of the Twin Squirrel helicopter and Cerro [17] in the identification of a small unmanned helicopter model.

Thus, parameter identification associated with the aerodynamic performance of the aircraft can be formulated as an optimization problem. In this paper, a new meta-heuristics, named multiple particle collision algorithm (MPCA) [12], was applied to the calibration of a dynamic flight simulator [13]. The MPCA optimization algorithm was inspired from some typical phenomena inside of nuclear reactors during the neutron travel: absorption and scattering of multiple particles. The results obtained with MPCA are compared to the ones obtained by Cruz [13], where a GA was used to find the parameters.

2 Parameter Identification Methodology

The methodology for the identification of parameters used in this work is the well-known Quad-M, proposed by Jategaonkar [19]. This methodology takes into account the main elements of rotorcraft system identification, including the rotorcraft excitation maneuvers, the aerodynamic data measurements, the mathematical model of the helicopter equations of motion, and the parameter estimation methods used to minimize the predicted output-error between the model and the real data, as shown in Figure 1. Here it is important to notice that the method used was MPCA. Each one of these elements will be discussed below.

2.1 Maneuver

The dynamic response and results from the application of inputs such as pulse, step, doublet, multistep, sinusoidal, 3-2-1-1, among others. Thus, a wide variety of maneuvers can be specified. We considered the following basic principle in identification systems: the data records of the flight test must contain the information of the dynamic characteristics to be obtained in the model [13]. In this work, only the sinusoidal frequency input will be used.

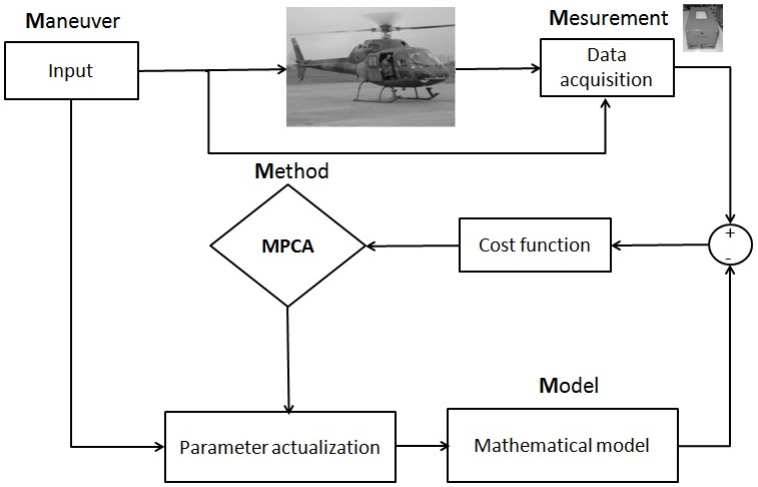


Figure 1: Adaptation Quad-M Methodology.

2.2 Measurements

The tested helicopter was equipped with the Aydin Vector Data Acquisition System (AVDAS) PCU-816-I, ATD-800 digital recorder, this system measures a total of thirty-five different parameters. Some of the measured data channels include fuel quantity in each tank, nose boom static and dynamic pressures, external stagnation temperature, aerodynamic angle of attack (α) and sideslip (β), roll, pitch, and yaw rates (p, q and r, resp.), load factors, longitudinal (θ) and lateral (ϕ) body attitudes, heading, collective, longitudinal and lateral cyclic, and pedal command deflections (δ_c, δ_b, δ_a and δ_p, resp.) .

The Earth axis speeds (u, v, w) are obtained with the aid of a Z12 Differential Global Positioning System DGPS, from Astech, whose antenna is fixed in the top of the vertical fin. The DGPS and AVDAS data synchronization is made by inserting a simultaneous event in both systems. The DGPS data is represented with the same AVDAS data sampling rate by means of linear interpolation procedure.

The wind direction and intensity are obtained comparing the body axis speeds with the aerodynamic speed from the flight-test air data system, mounted on a nose boom, at trim conditions. Consequently, the body axis speeds (u, v, w) are easily calculated adding wind vector to the Earth axis speeds [16].

2.3 Linearized Model Flight Dynamic

The helicopter equations of motion, deduced from the Newton second law for translational and rotational movements are given by [20] and [21] as:

$$X = m(\dot{u} - rv + qw) + mg\sin\theta \quad (1)$$

$$Y = m(\dot{v} - pw + ru) - mg\cos\theta\sin\phi \quad (2)$$

$$Z = m(\dot{w} - qu + pv) - mg\cos\theta\cos\phi \quad (3)$$

$$L = I_{xx}\dot{p} - I_{zx}(\dot{r} + pq) - (I_{yy} - I_{zz})qr \quad (4)$$

$$M = I_{yy}\dot{q} - I_{zx}(r^2 - p^2) - (I_{zz} - I_{xx})rp \quad (5)$$

$$N = I_{zz}\dot{r} - I_{zx}(\dot{p} - qr) - (I_{xx} - I_{yy})pq \tag{6}$$

where X, Y and Z represents the external force components (longitudinal, lateral and vertical); L, M and N are respectively, the roll, pitch and yaw moments; and $I_{()}$ corresponds to the moments and product of inertia of a rotating body. The kinematic relation for the pitch rate and roll rate about Y and X-axis are written as:

$$\dot{\theta} = q\cos\phi - r\sin\phi \tag{7}$$

$$\dot{\phi} = p + q\sin\phi\tan\theta + r\cos\phi\tan\theta \tag{8}$$

The helicopter equations of motion are nonlinear, but a meaningful analysis can be employed by converting them into linear differential equations, by considering only small perturbations on a trimmed equilibrium point (represented by subscript 0) in the rotorcraft flight envelope. In matrix notation, a linearized dynamical model is given by [13]:

$$\begin{bmatrix} \Delta\dot{u} \\ \Delta\dot{w} \\ \Delta\dot{q} \\ \Delta\dot{\theta} \\ \Delta\dot{v} \\ \Delta\dot{p} \\ \Delta\dot{\phi} \\ \Delta\dot{r} \end{bmatrix} = \begin{bmatrix} \frac{X_u}{m} & \frac{X_w}{m} & \frac{X_q}{m} - w_0 & -g\cos\theta_0 & \frac{X_v}{m} & \frac{X_p}{m} & 0 & \frac{X_r}{m} + v_0 \\ \frac{Z_u}{m} & \frac{Z_w}{m} & \frac{Z_q}{m} + u_0 & -g\cos\phi_0\sin\theta_0 & \frac{Z_v}{m} & \frac{Z_p}{m} - v_0 & -g\sin\phi_0\cos\theta_0 & \frac{Z_r}{m} \\ \frac{M_u}{I_{yy}} & \frac{M_w}{I_{yy}} & \frac{M_q}{I_{yy}} & 0 & \frac{M_v}{I_{yy}} & \frac{M_p}{I_{yy}} & 0 & \frac{M_r}{I_{yy}} \\ 0 & 0 & \cos\phi_0 & 0 & 0 & 0 & 0 & -\sin\phi_0 \\ \frac{Y_u}{m} & \frac{Y_w}{m} & \frac{Y_q}{m} & -g\sin\phi_0\sin\theta_0 & \frac{Y_v}{m} & \frac{Y_p}{m} + w_0 & g\cos\phi_0\cos\theta_0 & \frac{Y_r}{m} - u_0 \\ L'_u & L'_w & L'_q & 0 & L'_v & L'_p & 0 & L'_r \\ 0 & 0 & \sin\phi_0\tan\theta_0 & 0 & 0 & 1 & 0 & \cos\phi_0\tan\theta_0 \\ N'_u & N'_w & N'_q & 0 & N'_v & N'_p & 0 & N'_r \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta w \\ \Delta q \\ \Delta\theta \\ \Delta v \\ \Delta p \\ \Delta\phi \\ \Delta r \end{bmatrix}$$

$$+ \begin{bmatrix} \frac{X_{\delta_B}}{m} & \frac{X_{\delta_C}}{m} & \frac{X_{\delta_A}}{m} & \frac{X_{\delta_P}}{m} \\ \frac{Z_{\delta_B}}{m} & \frac{Z_{\delta_C}}{m} & \frac{Z_{\delta_A}}{m} & \frac{Z_{\delta_P}}{m} \\ \frac{M_{\delta_B}}{I_{yy}} & \frac{M_{\delta_C}}{I_{yy}} & M_{\delta_A} I_{yy} & \frac{M_{\delta_P}}{I_{yy}} \\ 0 & 0 & 0 & 0 \\ \frac{Y_{\delta_B}}{m} & \frac{Y_{\delta_C}}{m} & \frac{Y_{\delta_A}}{m} & \frac{Y_{\delta_P}}{m} \\ L'_{\delta_B} & L'_{\delta_C} & L'_{\delta_A} & L'_{\delta_P} \\ 0 & 0 & 0 & 0 \\ N'_{\delta_B} & N'_{\delta_C} & N'_{\delta_A} & N'_{\delta_P} \end{bmatrix} \begin{bmatrix} \delta_b \\ \delta_c \\ \delta_a \\ \delta_p \end{bmatrix} \tag{9}$$

Therefore, Equation 9 may also be written as:

$$\frac{d}{dt}\begin{bmatrix} X_l \\ X_d \end{bmatrix} = \begin{bmatrix} A_l & C_1 \\ C_2 & A_d \end{bmatrix}\begin{bmatrix} X_l \\ X_d \end{bmatrix} + \begin{bmatrix} B_l & D_1 \\ D_2 & B_d \end{bmatrix}\begin{bmatrix} \Delta\delta_l(t-\tau) \\ \Delta\delta_d(t-\tau) \end{bmatrix} + \dot{x}_{bias} \tag{10}$$

where X_l and X_d represent the longitudinal and lateral movements. Therefore, the longitudinal movement is expressed by:

$$\frac{dX_l}{dt} = A_l X_l + B_l \Delta\delta_l(t-\tau) + \dot{x}_{bias} \tag{11}$$

$$X_l = [\Delta u \quad \Delta w \quad \Delta q \quad \Delta\theta]^T \tag{12}$$

$$\Delta\delta_l = [\Delta\delta_b \quad \Delta\delta_c]^T \tag{13}$$

Values of interest for system identification are the elements of matrix A (stability derivatives), matrix B (control derivatives), and τ the delays associated with the aircraft response. Furthermore, the addition of tendency vector, x_{bias}, is constant and unknown. This vector is introduced in the mathematical model to

represent measurement errors and noise produced by transducers and instrumentation [18]. Let $J(\Omega)$ be the cost function, given by:

$$J(\Omega) = \sum_{i=1}^{n} \|X_i^{obs} - X_i^{mod}(\Omega)\|_2^2 \tag{14}$$

$$\begin{aligned} \Omega \quad = \quad & \left(\frac{X_u}{m}, \frac{X_w}{m}, \frac{X_q}{m}, \frac{Z_u}{m}, \frac{Z_w}{m}, \frac{Z_q}{m}, \frac{M_u}{I_{yy}}, \frac{M_w}{I_{yy}}, \frac{M_q}{I_{yy}}, \frac{X_{\delta_B}}{m}, \frac{X_{\delta_C}}{m}, \frac{Z_{\delta_B}}{m}, \frac{Z_{\delta_C}}{m}, \frac{M_{\delta_B}}{I_{yy}}, \frac{M_{\delta_C}}{I_{yy}}, \right. \\ & \left. \Delta\dot{u}_{bias}, \Delta\dot{w}_{bias}, \Delta\dot{q}_{bias}, \Delta\dot{\theta}_{bias}, \Delta u_{ref}, \Delta w_{ref}, \Delta q_{ref}, \Delta\theta_{ref}, \tau_c, \tau_b \right) \end{aligned} \tag{15}$$

where n is the number of measurements.

2.4 Method: Multiple Particle Collision Algorithm

The cost function to be minimized is a function of the parameters of the dynamic model, such as the helicopter aerodynamic stability and control derivatives, sensor bias, and sensitivities. Therefore, the determination of a parameter vector Ω that minimizes the cost function given by Equation 14 can be seen as an optimization problem and will be solved by a new meta-heuristics, named the MPCA.

MPCA is a meta-heuristics based on the canonical PCA [22]. This version uses multiple particles in a collaborative way, organizing a population of candidate solutions. The PCA was inspired by the traveling process (with absorption and scattering) of a particle (neutron) in a nuclear reactor. The use of the PCA was effective for several test functions and real applications [23].

The PCA starts with a selection of an initial solution (Old-Config), it is modified by a stochastic perturbation (*Perturbation*$\{.\}$), leading to the construction of a new solution (New-Config). The new solution is compared (function *Fitness*$\{.\}$), and a decision is made on whether it can ou cannot be accepted. If the new solution is not accepted, the scheme of scaterring (*Scaterring*$\{.\}$) is used. The exploration around closer positions is guaranteed by using the functions *Perturbation*$\{.\}$ and *Small-Perturbation*$\{.\}$. If the new solution is better than the previous one, this new solution is absorbed. If a worse solution is found, the particle can be sent to a different location of the search space, such that it enables the algorithm escape from a local minimum [12].

The implementation of the MPCA algorithm is similar to PCA, but it uses a set with N particles, where a mechanism to share the particle information is necessary. A blackboard strategy is adopted, where the best-fitness information is shared among all particles in the process. This process was implemented in Message Passing Interface (MPI), looking for application into a distributed memory machine [12]. The pseudo-code for the MPCA is presented by Table 1.

3 Results

The computational results obtained with MPCA and GA are show in Figures 2-5. The GA used Matlab Toolbox, and the MPCA was also implemented in Matlab R2001b. Computer tests were conducted under Linux operating system, in an Intel Core I5 2.27 GHz. The sinusoidal maneuver is represented by δ and the results presented take into consideration the average of 4 experiments with seeds generate different random numbers and experimental data generating artificially. The parameters used are: 2 particles; 10 iterations (exploration). The stopping criterion used was the total number of iterations (30).

Table 1: MPCA: psedo-code for the algorithm.

```
Generate an initial solution: Old-Config
Best-Fitness = Fitness{Old-Config}
Update Blackboard
For n = 0 to # of particles
    For n = 0 to # iterations
    Update Blackboard
    Perturbation{.}
        If Fitness{New-Config} > Fitness{Old-Config}
            If Fitness{New-Config} > Best-Fitness
                Best-Fitness = Fitness{New-Config}
            End If
            Old-Config = New-Config
            Exploration{.}
        Else
            Scattering{.}
        End If
    End For
End For
```

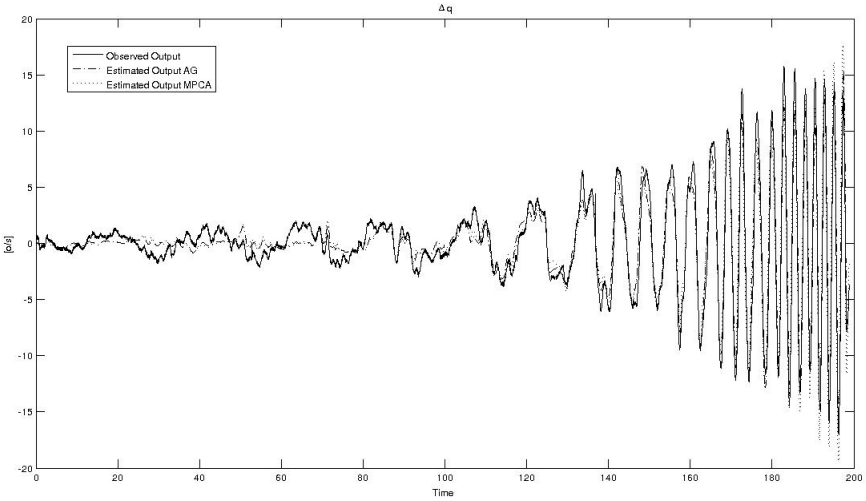


Figure 2: Pitch Rate.

The solid curve corresponds to the real data obtained during the test, the dashdotted is the result of identification produced by the GA and the results achieved by the MPCA are represented by the dotted curve. The results show that especially for u and w there is a slight discrepancy between the measured data and the data obtained by both algorithms. However, for high frequencies there is an improvement in the identification.

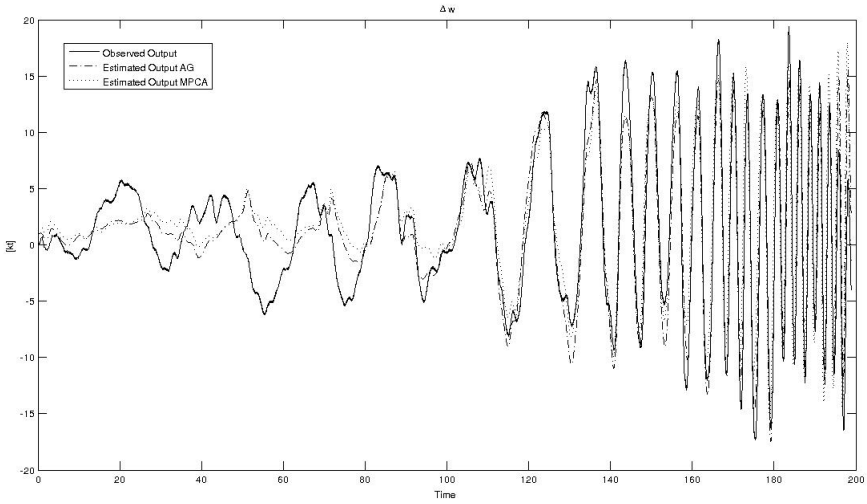


Figure 3: Vertical Velocity in Body Axis.

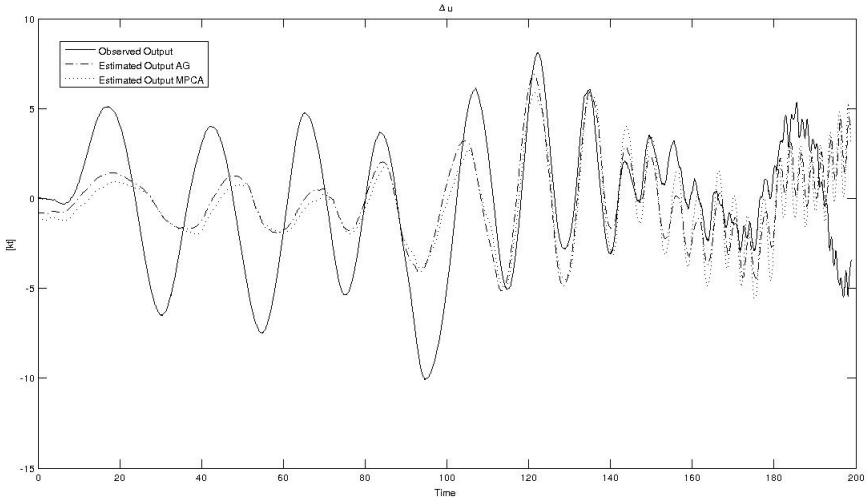


Figure 4: Longitudinal Velocity in Body Axis.

4 Conclusions

In this work, we compared two stochastic algorithms, GA and MPCA, for a helicopter parameter identification. The techniques were applied only in the estimation of the aerodynamic parameters of the longitudinal motion. The problem is formulated as an optimization process. The algorithms (GA and MPCA) were employed to address the solution of the optimization problem.

The results indicate that GA and MPCA present a good agreement, but it is a little bit better for MPCA implementation. Further work is suggested to apply MPCA in lateral-directional dynamic mode and in a more complex model which includes both longitudinal and lateral-directional dynamic modes.

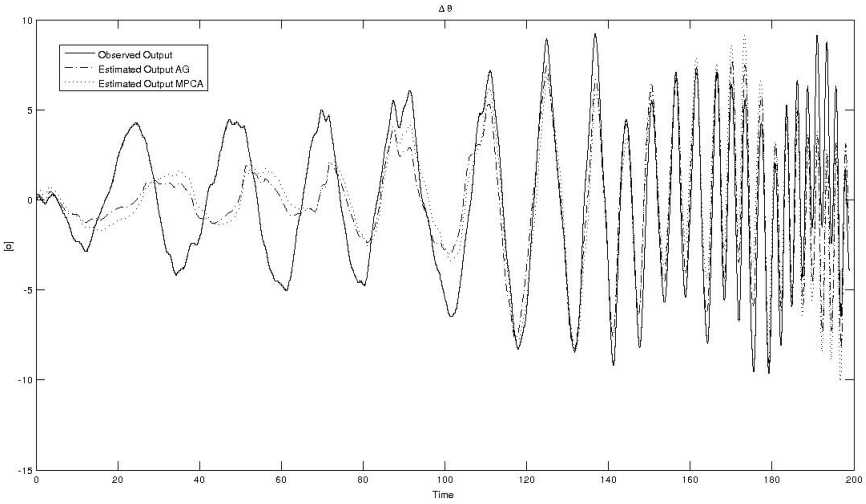


Figure 5: Pitch Attitude.

Acknowledgements

We wish to thank the partial support given by ITA and IPEV, specially the 2013 class Flight Teste Course (EFEV) Students, for providing the experimental data. Also we like to thank FINEP under agreement 01.13.0518.00 that funded the development of this tool and the presentation trip.

References

[1] H.F. Campos Velho, F.M. Ramos *Numerical Inversion of Two-Dimensional Geoelectric Conductivity Distributions from Electromagnetic Ground Data*, Journal of Geophysics, 15(2), 133-143, Brazilian, (1997).

[2] E.S. Chalhoub, H.F. Campos Velho *Estimation of the Optical Properties of Sea Water from Measurements of Exit Radiance*, Journal of Quantitative Spectroscopy Radiative Transfer, 72(5), 551-565, (2002).

[3] S. Stephany, H.F. Campos Velho, F.M. Ramos, C.D. Mobley *Identification of Inherent Optical Properties and Bioluminescence Source Term in a Hydrologic Optics Problem*, Journal of Quantitative Spectroscopy Radiative Transfer, 67(2), 113-123, (2000).

[4] G.D. Padfield *Helicopter Flight Dynamics: The Theory and Application of Flying Qualities and Simulation Modeling*, AIAA Education Series, AIAA, Reston, (1999).

[5] G.D. Padfield *Applications of System Identification in Rotorcraft Flight Dynamics*, Vertica, 13(3), 207-412, (1989).

[6] R. Kumar, R. Ganguli, S.N. Omkar *Rotorcraft Parameter Identification from Real Time Flight Data*, Journal of Aircraft, vol. 45, n. 1, 333-341, (2008).

[7] Agard *Rotorcraft System Identification*, AGARD AR 280, LS 178, (1991).

[8] P.G. Hamel, J. Kaletka *Advances in Rotorcraft System Identification*, Progress in Aerospace Sciences, 33(3-4), 259-284, (1997).

[9] P. Hajela, J. Lee *Genetic Algorithms in Multidisciplinary Rotor Blade Design*, Proceedings of the 36th AIAA Structures, Structural Dynamics, and Materials Conference, New Orleans, USA, (1995).

[10] V.L. Wells, A.Y. Han, W.A. Crossley *Acoustic Design of Rotor Blades Using a Genetic Algorithm*, AGARD CP 552, (1995).

[11] P.M.T. Zaal, D.M. Pool, Q.P. Chu, M.M.V. Paassen, M. Mulder, J.A. Mulder *Modeling Human Multimodal Perception and Control Using Genetic Maximum Likelihood Estimation*, Journal of Guidance, Control, and Dynamics, 32(4), 1089-1099, (2009).

[12] E.F.P. Luz, *Meta-heurísticas paralelas na solução de problemas inversos*, Tese de Doutorado, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, (2012).

[13] R.V. Cruz, *Desenvolvimento de um Modelo Dinâmico para Simuladores de Helicóptero*, Tese de Doutorado, Instituto Tenológico de Aeronáutica (ITA), São José dos Campos, (2009).

[14] R.V. Cruz, N.S.B. Neto, L.C.S. Góes *Preliminary Results of Longitudinal Helicopter System Identification Using Genetic Optimization Algorithm*, Proceedings of the 5th Brazilian Conference on Dynamics, Control and their Apllications, Guaratinguetá, Brazil, (2006).

[15] R.V. Cruz, L.C.S. Góes *Results of Longitudinal Helicopter System Identification Using Output-Error and Both Genetic and Levenberg-Marquardt Optimization Algorithm*, Proceedings of the 19th Brazilian Congress of Mechanical Engineering, Brasília, Brazil, (2007).

[16] R.V. Cruz, L.C.S. Góes *Results of Short-Period Helicopter System Identification Using Output-Error and Hybrid Search-Gradient Optimization Algorithm*, Mathematical Problems in Engineering, 2010, (2010).

[17] J. del Cerro, J. Valero, A. Barrientos *Identification of Small Unmanned Helicopter Model Using Genetic Algorithms*, Proceedings of the International Conference on Intelligent Robots and Systems, Edmonton, Canada, (2005).

[18] M.B. Tischler, R.K. Remple *Aircraft and rotorcraft system identification*, AIAA Education Series, (2006).

[19] R.V. Jategaonkar *Flight Vehicle System Identification: A Time Domain Methodology*, Progress in Aerospace Sciences, Reston, USA, (2006).

[20] R.W. Prouty *Helicopter Performance, Stability and Control*, Krieger Publishing Company, Malabar, (2005).

[21] A.K. Cooke, E.W.H Fitzpatrick *Helicopter Test and Evaluation*, AIAA Education Series, AIAA, Reston, (2002).

[22] W.F. Sacco, C.R.E.A. Oliveira, *A new stochastic optimization algorithm based on a particle collision metaheuristic*, Proceedings of the 6th World Congress of Structural and Multidisciplinary Optimization, (2005).

[23] W.F. Sacco, D.C. Knupp, E.F.P. Luz, A.J.S. Neto *Algoritmo de Colisão de Partículas*, Técnicas de Inteligência Computacional Inspiradas na Natureza Aplicação em Problemas Inversos em Transferência Radiativa, (2009).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Application of a Continuum Damage Identification Approach via Markov Chain Monte Carlo Method using the Impulsive Response Function

Josiele S. Teixeira[a], Leonardo T. Stutz[b], Diego C. Knupp[c], Antônio J. Silva Neto[d]

Departamento de Engenharia Mecânica e Energia
Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, Brazil, 28630-050
e-mail: [a]josyelly1@gmail.com, [b]ltstutz@iprj.uerj.br, [c]diegoknupp@gmail.com, [d]ajsneto@iprj.uerj.br

Key words: Damage identification, Impulsive response function, Markov Chain Monte Carlo method

Abstract
The present work is concerned with the problem of structural damage identification built on the statistical inversion approach. Here, the damage state of the structure is continuously described by a cohesion parameter, which, in turn, is spatially discretized by the finite element method. The inverse problem of damage identification is then posed as the determination of the posterior probability densities of the nodal cohesion parameters. The Markov Chain Monte Carlo (MCMC) method, implemented with the Metropolis-Hastings algorithm, is considered in order to approximate the posterior probabilities by drawing samples from the posterior probability density function. With this approach, prior information on the sought parameters can be used and the uncertainty concerning the known values of the material properties can be quantified in the estimation of the cohesion parameters. The assessment of the proposed approach has been performed by means of numerical simulations on a simply supported Euler-Bernoulli beam. Different damage scenarios and noise levels were addressed and three different strategies were considered in the damage identification process.

1 Introduction

Structural damage identification and health monitoring are essential issues for determining safety reliability and remaining lifetime of aerospace, civil and mechanical structures. The technological and scientific challenges posed by damage identification problems yielded a great research activity on this subject within the scientific community [1]. Although different damage identification approaches have been proposed in the specialized literature, one may observe a special attention to the non-destructive ones built on the dynamic behavior of the structures. These ones, encompassing deterministic or statistical perspective, consider different types of data (modal parameters, time series, frequency responses) and distinct mathematical formulations and numerical algorithms for solving the corresponding inverse problem.

Most prior works are built on the vibrational behavior of the structure, more specifically on the traditional modal analysis, being conceived within the general framework of Finite Element Model (FEM) updating methods [2,3]. Those methods are aimed at identifying structural damage through determination of changes in some parameters of a FEM of the structure in comparison to the known values in the undamaged scenario. Therefore, the basic idea of the vibration-based damage

identification approach is that the modal properties of the structure (frequencies, mode-shapes and modal damping) are functions of the physical properties (mass, stiffness and damping) and changes in these last ones, due to the presence of a damage, will be reflected in the modal characteristics, which can be measured and used to infer about damage location and severity.

More recently, damage identification techniques based on time-domain responses have been proposed, allowing for the damage identification problem formulation in non-linear structures, highly damped, and systems with high modal density, situations in which the modal approach may be extremely cumbersome. In order to identify the FEM parameters, an objective function is formulated, given by the squared residues between the vibration data identified by modal testing and those computed from the mathematical model, and then minimized [4]. In most cases the ordinary least-squares norm is employed as objective function, and although very useful in many situations, it yields maximum likelihood estimates (with statistical significance) only if some hypotheses are valid [5,6]: (i) the measured data contain errors which are additive, uncorrelated, and follow a normal distribution with zero mean and constant standard deviation; (ii) only the measured variables appearing in the objective function contain errors (that is, in the model both the independent variables and the considered known parameters are exactly known); and (iii) there is no prior information regarding the values and uncertainties of the unknown (sought) parameters. For example, typical sources of uncertainties which are commonly neglected are the values of the material properties.

In the present work, the damage state of the structure is continuously described by a cohesion parameter. The cohesion field is discretized by the finite element method and this spatial discretization is not necessarily coincident with that used for the displacement field. Therefore, the damage model adopted in the present work is different from that commonly adopted in the specialized literature, where the damage is supposed to be constant within a finite element of the structure. The damage identification problem is, then, formulated as an inverse problem where we seek to estimate the cohesion parameters. In this work, we use the statistical inverse approach [7,8]. The inverse problem of damage identification is then posed as the determination of the posterior probability densities of the nodal cohesion parameters. The Markov Chain Monte Carlo (MCMC) method, implemented with the Metropolis-Hastings algorithm, is considered in order to approximate the posterior probabilities by drawing samples from the posterior probability density function.

The remainder of the paper is organized as follows. Section 2 presents all the mathematical formulation required for the definition of the problem of continuum damage identification in the time domain. Hence, the continuum damage model and the vector of generalized response are presented in this section. Section 3 presents the formulation of the inverse problem of damage identification built on the statistical inversion approach. Section 4 presents the numerical assessment of the potentiality of the proposed damage identification approach applied on a simply supported Euler-Bernoulli beam for different damage scenarios and noise levels. Finally, Section 5 presents the concluding remarks.

2 Direct problem formulation

In the present damage identification approach, the damage state of the structure is continuously described by a structural parameter $\beta \in [0,1]$, named cohesion parameter [9]. This parameter is related with the connections among material points and can be interpreted as a measure of the local cohesion state of the material. If $\beta = 1$ it is assumed that all connections between the material points

are preserved and, therefore, there is no damage in the structure. If $\beta = 0$, a local rupture is considered, since all connections between the material points are broken.

In the present work, it is assumed that, during the vibration test, the internal forces within the structure do not suffice to cause the continuation of the damage process and, besides, that the damage affects only the elastic properties of the structure. Hence, for the special case of an Euler-Bernoulli beam, the stiffness matrix of a finite element is given by

$$\mathbf{K^e} = \int_0^{l_e} \beta^e(x) E_0 I_0 \frac{\partial^2 \mathbf{N^e}(x)}{\partial x^2} \frac{\partial^2 \mathbf{N^e}(x)}{\partial x^2}^T dx \tag{1}$$

where $\beta^e(x)$ stands for the cohesion field within the element, E_0 and I_0 are, respectively, the nominal young modulus and the nominal area moment of inertia, $\mathbf{N}^e(x)$ is the vector composed of the appropriate interpolation functions for the displacement filed, l_e is the element length and T is the matrix transpose operator.

The present work considers a beam element with two degrees of freedom per node, so that the cohesion field, within an element, is interpolated by linear Lagrangian functions from its nodal parameters β_1^e and β_2^e. It must be emphasized here that, the approach considered in the present work to continuously describe the structural damage differs from the ones presented in the specialized literature. The damage is generally supposed to be uniform within the damaged region, since it is associated with a classical parameter of the structure, generally the young modulus, which is uniform within a finite element.

Therefore, after a spatial discretization by the finite element method and considering eq. (1), the equation of motion of a system with n degrees of freedom is given by

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{D}\dot{\mathbf{q}}(t) + \mathbf{K}(\boldsymbol{\beta})\mathbf{q}(t) = \mathbf{f}(t) \tag{2}$$

where $\mathbf{M}$, $\mathbf{D}$ and $\mathbf{K}(\boldsymbol{\beta})$ are, respectively, the mass, damping and stiffness matrices of the structure, $\boldsymbol{\beta}$ is the vector comprised of the nodal cohesion parameters, $\mathbf{q}$ is the vector of nodal displacements, $(\cdot)$ represents the differentiation with respect to time and $\mathbf{f}$ is the external excitation vector.

In order to formulate the damage identification problem in the time domain, the vector of generalized displacement is defined as follows. Be $\mathbf{y}_i^m$ the vector comprised with the response, predicted by the model given in eq. (2), at the position $\mathbf{p}_i$:

$$\mathbf{y}_i^m = \begin{bmatrix} y_i^m(t_1) & y_i^m(t_2) & \cdots & y_i^m(t_N) \end{bmatrix}^T \tag{3}$$

where $y_i^m(t_j)$ is the response, at the position $\mathbf{p}_i$, acquired at the time instant t_j and N is the number of transient measurements that will be considered in the experimental procedure of the damage identification process. The vector of generalized displacement is, then, defined as:

$$\mathbf{Y}^m = \begin{bmatrix} \mathbf{y}_1^m \\ \mathbf{y}_2^m \\ \vdots \\ \mathbf{y}_{m_s}^m \end{bmatrix} \tag{4}$$

where m_s is the number of positions where the response of the structure is measured.

3 Inverse problem formulation and solution

The associated inverse problem consists of estimating the cohesion parameters, β_i, $i = 2,3,...,24$, using the impulse response function of the mechanical system. It is considered to be available transient deflection measurements at $x = 0.233m$, acquired with a 1 kHz sample frequency, after an impulse excitation at the same position. Synthetic experimental data $\mathbf{Y}^m$ have been simulated by adding random noise, from a Normal distribution with zero mean and prescribed standard deviation, to the model predicted impulsive response of the structure.

It has been considered three levels of experimental noise, which were computed considering the impulse response of the structure in the time interval of 10s and the following signal to noise ratios (SNR): 80dB, which represents a very low noise level, being primarily used to test the inverse problem implementation; 30dB which is a realistic level of noise for the experiment herein considered, and 20dB, which is related to very noisy measurements, and has been used to challenge the methodology. In all cases, it has been considered the presence of two local damages in the beam, one located at $x = 0.5475m$ and the other at $x = 1.2775m$, corresponding to $i = 10$, and $i = 22$, respectively. Both damages have been simulated with $h / h_o = 0.9$, yielding $\beta_{10} = \beta_{22} = 0.729$. All the other nodes are considered undamaged and therefore $\beta_i = 1.0$. Since, due to damping effects, most of the information carried by the impulse response of the structure is retained in its beginning, only the first 125 measurements were considered in the damage identification process, i.e. $N = 125$.

Assuming that the prior information can be modeled as a probability density $\pi_{pr}(\boldsymbol{\beta})$, the Bayes' theorem of inverse problems can be expressed as [7]:

$$\pi_{post}(\boldsymbol{\beta}) = \pi(\boldsymbol{\beta} \mid \mathbf{Y}^{\exp}) = \frac{\pi_{pr}(\boldsymbol{\beta})\pi(\mathbf{Y}^{\exp} \mid \boldsymbol{\beta})}{\pi(\mathbf{Y}^{\exp})} \tag{5}$$

where $\pi_{post}(\boldsymbol{\beta})$ is the posterior probability density, $\pi_{pr}(\boldsymbol{\beta})$ is the prior information on the unknowns, modeled as a probability distribution, $\pi(\mathbf{Y}^{\exp} \mid \boldsymbol{\beta})$ is the likelihood function and $\pi(\mathbf{Y}^{\exp})$ is the marginal density and plays the role of a normalizing constant. Considering that the measurement errors related to the data $\mathbf{Y}^{\exp}$ are additive, uncorrelated, and have normal distribution, the likelihood function $\pi(\mathbf{Y}^{\exp} \mid \boldsymbol{\beta})$, i.e. the probability density for the occurrence of the measurements $\mathbf{Y}^{\exp}$ given the model parameters $\boldsymbol{\beta}$ can be expressed as [6]:

$$\pi(\mathbf{Y}^{\exp} \mid \boldsymbol{\beta}) = \frac{1}{\sqrt{(2\pi)^{N_d}}} \frac{1}{\sqrt{\det(\mathbf{W})}} \exp\left[-\frac{1}{2} \left(\mathbf{Y}^{\exp} - \mathbf{Y}^m \right)^T \mathbf{W}^{-1} \left(\mathbf{Y}^{\exp} - \mathbf{Y}^m \right) \right] \tag{6}$$

where $\mathbf{W}$ is the covariance matrix of the errors related to the data $\mathbf{Y}^m$, and $\mathbf{Y}^m$ are the predicted deflections, at the same location and time instants as $\mathbf{Y}^{\exp}$, calculated by the model with the given values of the cohesion parameters, $\boldsymbol{\beta}$.

In this work we use Markov Chain Monte Carlo (MCMC) methods [7] in order to approximate the posterior probabilities by drawing samples from the posterior probability density function. In order to implement the Markov Chain we need a candidate-generating density, $q(\boldsymbol{\beta}^t, \boldsymbol{\beta}^*)$, which denote a source density for a candidate draw $\boldsymbol{\beta}^*$ given the current state $\boldsymbol{\beta}^t$. Then the Metropolis-Hastings algorithm [7,8], which is used in this work to implement the MCMC method, is defined by the following steps:

Step 1: Sample a candidate $\boldsymbol{\beta}^*$ from the candidate-generating density $g(\boldsymbol{\beta}^t, \boldsymbol{\beta}^*)$
Step 2: Calculate

$$\alpha = \min\left[1, \frac{\pi(\boldsymbol{\beta}^* \mid \mathbf{Y}) g(\boldsymbol{\beta}^*, \boldsymbol{\beta}^t)}{\pi(\boldsymbol{\beta}^t \mid \mathbf{Y}) g(\boldsymbol{\beta}^t, \boldsymbol{\beta}^*)}\right] \tag{7a}$$

Step 3: If $U(0,1) < \alpha$, then

$$\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^* \tag{7b}$$

else,

$$\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^t \tag{7c}$$

where $U(0,1)$ is a random number from an uniform distribution between 0 and 1.
Step 4: Return to Step 1 in order to generate the chain $\left\{\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, ..., \boldsymbol{\beta}^{N_{MCMC}}\right\}$. We should stress that the first states of this chain must be discarded until the convergence of the chain is reached. These ignored samples are called the burn-in period, whose length will be denoted by $N_{burn-in}$.
In the present work we have used a random walk process in order to generate the candidates, so that $\boldsymbol{\beta}^* = \boldsymbol{\beta}^t + \boldsymbol{\eta}$, where $\boldsymbol{\eta}$ follows the distribution q, which was defined as a normal density. In this case g is symmetric and $g(\boldsymbol{\beta}^*, \boldsymbol{\beta}^t) = g(\boldsymbol{\beta}^t, \boldsymbol{\beta}^*)$, so **Step 2** is simplified and eq. (7a) may be rewritten as:

$$\alpha = \min\left[1, \frac{\pi(\boldsymbol{\beta}^* \mid \mathbf{Y})}{\pi(\boldsymbol{\beta}^t \mid \mathbf{Y})}\right] \tag{7d}$$

which can be directly calculated form eqs. (5) and (6).

The main difficulty associated with this inverse problem solution is due to the large number of parameters being estimated (β_i, $i = 2,3,...,24$) and the choice of the step-size in the random walk

process (characterized by the standard deviation, $\sigma_{g,i}$, of the normal distribution g associated with each parameter β_i). If the step-size is too small the posterior density is explored very slowly, and the Markov chain takes long time to visit all parts of the density. When the step-size is increased, the coverage also increases, but if the step-size is too large the acceptance ratio of the candidates becomes too small. This acceptance ratio feature is associated with the sensitivity of the model predictions employed in eq. (6) to each one of the parameters, which can be rather different one from another, and is also associated with the covariance matrix of the experimental data: it's intuitive from eq. (6) that for experimental data with higher noise the likelihood function is less sensitive to variations in the parameters. In this work we have considered three different strategies for the estimation of the cohesion parameters, β_i, which are described below:

Strategy 1: the candidate generating normal density, g, in the random walk process of the Metropolis-Hastings algorithm, has the same standard deviation value, σ_g, for all parameters.

Strategy 2: after a prescribed number of states in the Markov Chain employing **Strategy 1**, large enough to achieve convergence of the chain, it is identified the most probable damage locations: if the mean of the last 1,000 states of the chain is less than a threshold value ($\beta_{\text{threshold}} = 0.9$), the point is considered as probably damaged. The Markov Chain is then reinitialized, being assigned two different standard deviations for the candidate-generating density, g, in the random walk process of the Metropolis-Hastings algorithm: σ_d for the cohesion parameters β_i identified as probably damaged and σ_{nd} for the cohesion parameters β_i identified as probably not damaged.

Strategy 3: it is similar to **Strategy 2**, but instead of simply assigning different step-sizes for the cohesion parameters associated with identified probably damaged and undamaged locations, it is considered, for the identified undamaged locations, the use of a strong a priori information for the corresponding cohesion parameters: the parameters β_i identified as probably undamaged have their value set as known a priori with $\beta_i = 1.0$, and the Markov Chain is reinitialized for exploring the posterior distributions of only the parameters associated with the damaged locations and their neighbor nodes.

4 Results and discussion

Considering the low noise level case (80dB), all three strategies described in Section 3 performed well, being able to accurately identify the damage locations and intensities. These results are not presented here for brevity.

Considering the mild noise level case (30dB), employing the following step-sizes for the candidate-generating density in Strategies 1, 2, and 3: $\sigma = 5.0\times10^{-4}$, $\sigma_d = 5.0\times10^{-3}$, $\sigma_{nd} = 1.0\times10^{-6}$, the estimated $h(x)/h_0$ profile (obtained from the estimated parameters β_i) are presented in figures 1(a-c) for strategies 1, 2, and 3, respectively. It should be observed in these figures that all strategies were able to identify the damaged regions, being the most accurate results obtained by Strategies 2 and 3. The main error associated with the inferred posterior distribution associated with Strategy 1 is due to not converged and/or poor mixed chains. We highlight that due to the low chain convergence rate of Strategy 1, 100,000 states were used in the MCMC chain (being the first 20,000 sates just to achieve

convergence, and discarded as the burn-in period), whereas for Strategies 2 and 3 only 20,000 states were needed (the first 4,000 states discarded as the burn-in period), reducing significantly the computational cost. Table 1 summarizes the mean values and estimated 99% confidence intervals obtained for the damaged cohesion parameters, β_{10} and β_{22}, with the three strategies for this case. It should be observed that the Strategy 3 was the only one to accurately identify the 99% confidence region encompassing the exact value. The poorest estimates for the damaged cohesion parameters were obtained by Strategy 1. The Markov chains are illustrated in figures 2 (a-c), which depict the MCMC evolution for β_{10} for the three strategies, respectively.

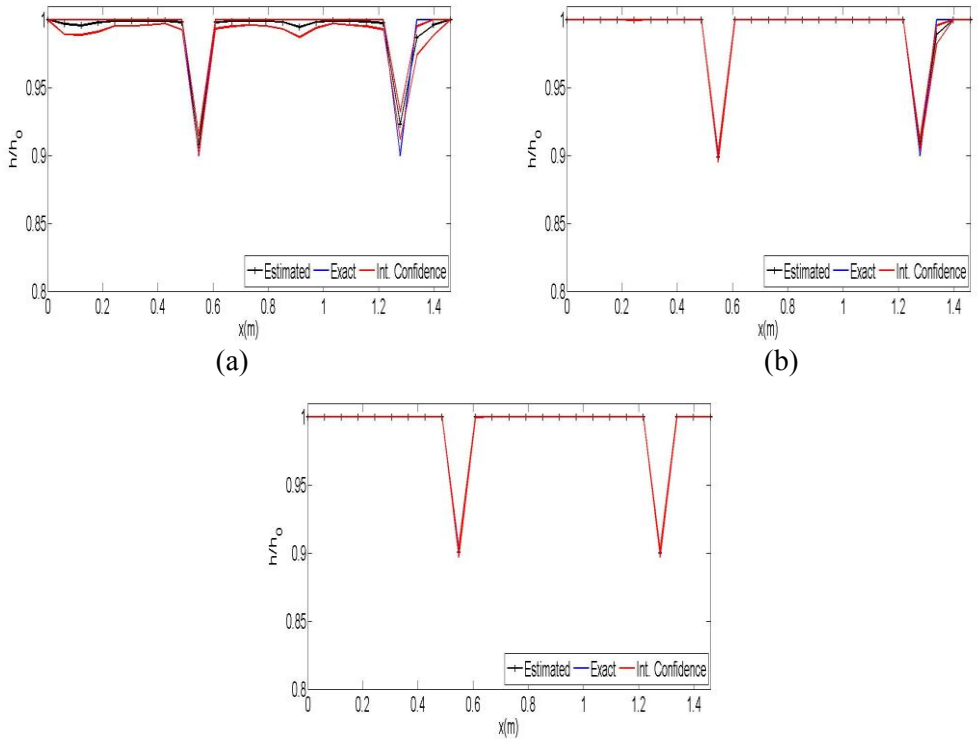


Figure 1: Estimated $h(x)/h_0$ profile for the 30dB SNR case with (a) Strategy 1; (b) Strategy 2; and (c) Strategy 3.

Table 1: Summary of the results obtained for β_{10} and β_{22} for the 30dB SNR case.

	β_{10} [99% confidence interval]	β_{22} [99% confidence interval]
Strategy 1	0.749 [0.731, 0.764]	0.786 [0.758, 0.810]
Strategy 2	0.726 [0.717, 0.733]	0.747 [0.738, 0.757]
Strategy 3	0.729 [0.720, 0.738]	0.729 [0.722, 0.733]
Exact	**0.729**	**0.729**

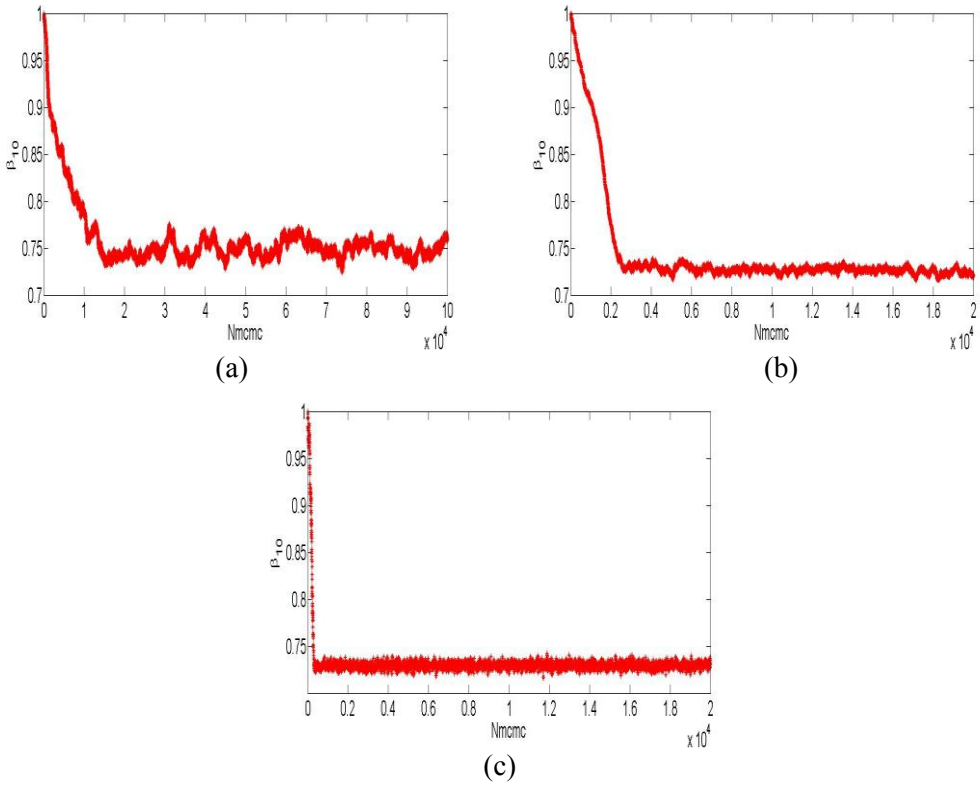
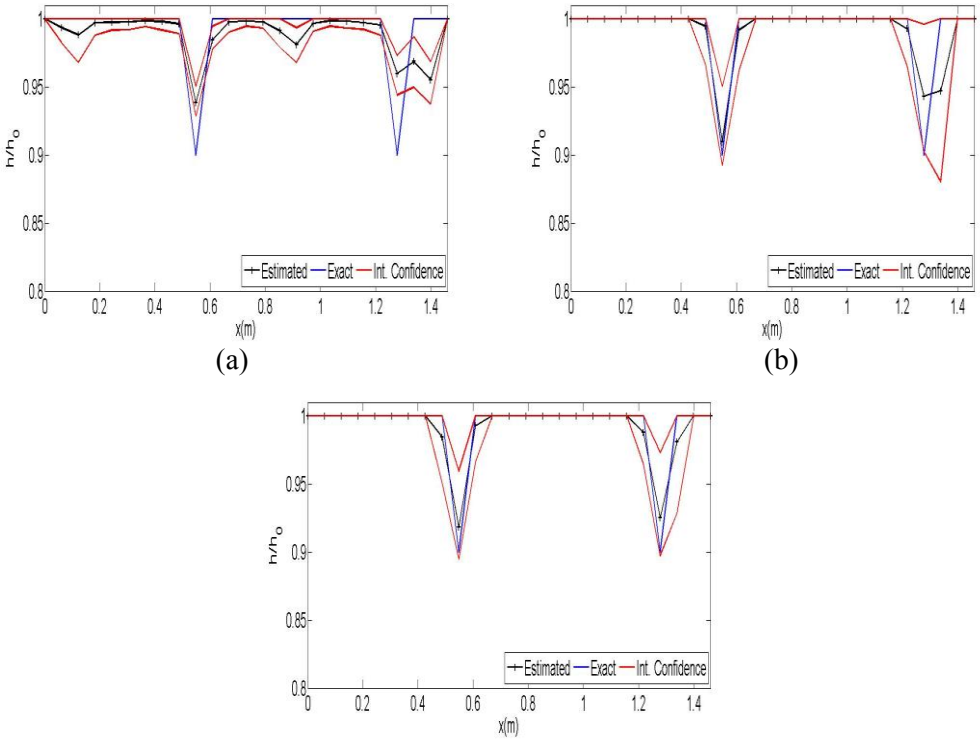


Figure 2: MCMC evolution for β_{10} for the 30dB SNR case with strategies 1, 2 and 3 (a-c).

Considering the high noise level case (20dB), employing the following step-sizes for the candidate-generating density in Strategies 1, 2, and 3: $\sigma = 5.0\times10^{-4}$, $\sigma_d = 5.0\times10^{-2}$, $\sigma_{nd} = 1.0\times10^{-6}$, the estimated $h(x)/h_0$ profile (obtained from the estimated parameters β_i) are presented in figures 3(a-c) for strategies 1, 2, and 3, respectively. These results show that the estimates obtained by Strategy 1 under the presence of high noise level are quite poor, being the damaged regions roughly estimated, with the estimated damage intensity rather far from the exact value. In fact, it has been observed that even after 100,000 states, that Markov Chain was not fully converged to the equilibrium distribution, and the inferred distribution, in this case, represents very poorly the real posterior distribution. On the other hand, strategies 2 and 3 identified the location and intensity of the damage related to β_{10} ($x = 0.5475m$) quite accurately, whereas only Strategy 3 was able to identify accurately the damaged region related to β_{22}, $x = 1.2775m$ - Strategy 2 identified β_{23} as the most damaged location, which in reality is an undamaged cohesion parameter. This result is probably due to the sensor location ($x = 0.233m$) which most probably presents higher sensitivity to the cohesion parameter β_{10} rather than β_{22}, yielding better estimates for the first parameter rather than the latter. Table 2 summarizes the mean values and estimated 99% confidence intervals obtained for the damaged cohesion parameters, β_{10} and β_{22}, with the three strategies for this case. These results confirm that Strategy 1 failed to identify the damage intensity related to both β_{10} and β_{22}, whereas only Strategy 3 was able to estimate 99% confidence intervals which encompass the exact values of the cohesion parameters. It is also clear that the confidence interval for β_{10} is narrower than for β_{22}, suggesting once again a higher sensitivity of the measurement location to β_{10}. Finally, figures 4(a,b) illustrate, respectively, the Markov Chain with 100,000 states for β_{10} and the corresponding histogram after neglecting the

first 10,000 states, which as shown in figure 4(a) is more than enough to achieve convergence of the chain. It should be recalled that this case employs a very high level of noise and is here used to challenge the methodology and evaluate the behavior difference of the three strategies.



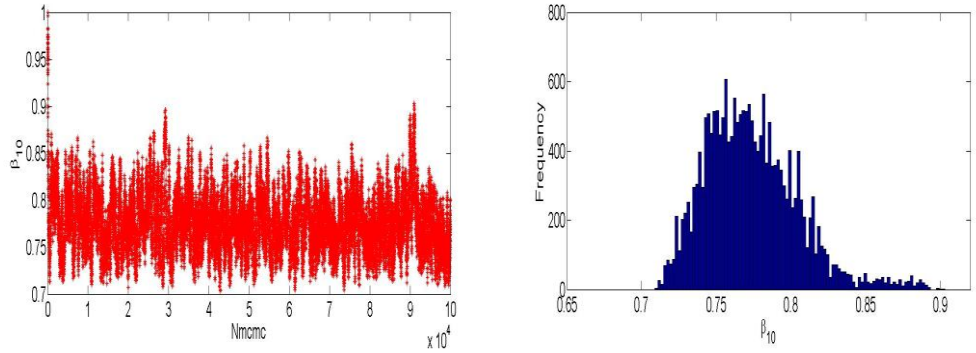
(a) (b)

(x)

Figure 3: Estimated $h(x)/h_0$ profile for the 20dB SNR case with strategies 1, 2 and 3 (a-c).

Table 3: Summary of the results obtained for β_{10} and β_{22} for the 20dB SNR case.

	β_{10} [99% confidence interval]	β_{22} [99% confidence interval]
Strategy 1	0.826 [0.800, 0.859]	0.884 [0.841, 0.921]
Strategy 2	0.753 [0.710, 0.858]	0.838 [0.734, 0.988]
Strategy 3	0.773 [0.716, 0.882]	0.791 [0.721. 0.921]
Exact	**0.729**	**0.729**



(a) (b)

Figure 4: (a) MCMC evolution in Strategy 3 for β_{10} in the 20dB SNR case; (b) Corresponding historgram.

5 Conclusions

In the present work the problem of structural damage identification was investigated by considering that the damage state of the structure is continuously described by a cohesion parameter. The inverse problem was then posed within a Bayesian framework, as the determination of the posterior probability densities of the nodal cohesion parameters. The Markov Chain Monte Carlo (MCMC) method, implemented with the Metropolis-Hastings algorithm, was employed in order to approximate the posterior probabilities by drawing samples from the posterior probability density function. The methodology herein developed, of first identifying the potentially damaged regions, and then assigning different step-sizes to the corresponding cohesion parameters, has been demonstrated to remarkably enhance the process of exploration of the posterior distribution. This work has also shown that assigning an informative a priori to the cohesion parameters related to probably undamaged regions brings more accuracy to the damage intensity identification of the damaged regions.

Acknowledgements

The authors acknowledge the financial support provided by the Brazilian sponsoring agencies CNPq, CAPES and FAPERJ.

References

[1] Santos, J. V. A, Maia, N. M. M., and Soares, C. M. M., *Structural damage identification: A survey*, In: Topping, B. H. V., Papadrakakis, M., Editors, Trends in computational structures technology, Stirlingshire, Scotland: Saxe-Coburg Publications, (2008).

[2] Doebling, S. W., Farrar, C. R., Prime, M. B., and Sheritz, D. W., *Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review*, Los Alamos National Lab., Rept. LA-13070-MS, Los Alamos, NM, (1996).

[3] Stutz, L. T., Castello, D. A., and Rochinha, F. A., *Structural damage identification based on a finite element updating*, Proceedings of the sixth US national congress on computational mechanics, Dearborn, Michigan, (2001).

[4] Hao, H., and Xia, Y., *Vibration-Based Damage Detection of Structures by Genetic Algorithm*, J. Comput. Civ. Eng., 16, (2002), pp. 222-229.

[5] Beck, J. V., and Arnold, K. J., *Parameter Estimation in Engineering and Science*, Wiley Interscience, New York, (1977).

[6] Schwaab, M., and Pinto, J. C., *Experimental Data Analysis I: Statistics Fundamentals and Parameter Estimation* (in portuguese), E-papers, Rio de Janeiro, (2007).

[7] Kaipio, J., and Somersalo, E., *Statistical and Computational Inverse Problems*, Applied Mathematical Sciences 160, Springer-Verlag, (2004).

[8] Paez, M., and Gamerman, D., *Bayesian Approaches for the Solution of Inverse Problems*, In: Orlande, H. R. B., Fudym, O., Maillet, D., and Cotta, R. M., Editors, Thermal Measurements and Inverse Techniques, CRC Press, Boca Raton, (2011).

[9] Stutz, L. T., Castello, D. A., and Rochinha, F. A, *A flexibility-based continuum damage identification approach*, Jounal of Sound and Vibration, 279, (2005), pp. 641-667.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Identification of rough impedance profile using an improved acoustic wave propagation algorithm

Roberto A. Tenenbaum, Filipe O. Taminato, Kennedy M. Fernandes
Computer Modeling Program
Rio de Janeiro State University, Rua Bonfim, 25, Nova Friburgo, RJ, Brazil
e-mail: ratenenbaum@gmail.com

Key words: Acoustic wave propagation, Impedance identification, Damage assessment, Luus-Jaakola optimization.

Abstract

As it is well known, damage identification is an essential issue for evaluating safety reliability and remaining lifetime of aerospace, civil and mechanical structures. The scientific and technological challenges posed by damage identification problems yielded a great research activity in this subject. The damage identification can be successful addressed from the point of view of the wave propagation approach. The main idea is to solve the direct problem using the Sequential Algebraic Algorithm (SAA), which provides a straightforward solution for the acoustical wave propagation in inhomogeneous media. Then, the inverse problem is solved by an optimization procedure. Several optimization methods were tested in this sense. The main conclusion was that for this particular problem, the Luus-Jaakola optimization method provided the best accuracy and computational cost compromise. The identification of a smooth profile by using the SAA for running the direct problem and the Luus-Jaakola (LJ) procedure for optimization purposes to identify the impedance profile successful, even with an average additive noise in the pulse and echo signals. One of the main advantages of using the SAA is that it furnishes a closed form solution for the direct problem, that means, the running time is negligible. Another important feature is the possibility to compute the parameters in a sequential way, reducing the computing time. However, when trying to identify a rough variation impedance profile, the identification does not succeed, showing a significative error propagation. This happens because the transmission coefficients, which are in the form $1 \pm R_i$, where R_i is the i^{th} reflection coefficient and the signal depends on the propagation direction, are approximated in the algorithm. When these reflection coefficients are not small, the product of several transmission coefficients deviates from unity and the identification result diverges, as will be shown in the numerical examples. Here, it is presented the Improved Sequential Algebraic Algorithm (ISAA) that provides also a closed form for the direct acoustic wave propagation phenomenon but without the approximation found in the SAA. The main features are preserved, i.e., it runs in a negligible amount of time and the identification procedure can be done with one parameter per turn. Identification results with an average noise in the signals for rough impedance profiles, showing a 80 parameters recovery with low error are presented in this paper.

1 Introduction

Structural health monitoring (SHM) and damage identification (DI) are prime concerns in the realm of civil, mechanical and aerospace engineering. They represent essential issues to determine the safety and reliability of their systems and structures.

Different nondestructive SHM and DI approaches are proposed in the literature. Most of them, however, are built on changes in the vibration characteristics of the structures under concern [1]. The basic idea of these approaches is that the modal properties are functions of the physical properties of the structure and, therefore, changes due to damage in the physical properties will be reflected in the modal ones, which, can be measured and used to infer about the damage state.

Although the vibration-based damage identification approaches have been successfully applied to practical problems [5], it is well known that small defects may yield to excessively small or even no effects on the modal properties of the structure, making the damage identification a more difficult task. Damage identification methods built on the acoustic wave propagation approach, on the other hand, are highly sensitive to changes in local mechanic impedance such as those caused by small defects [2]. Besides the higher sensitivity to small defects, the wave propagation based approaches are directly defined in the time domain and, therefore, differently from most of the vibration based methods, they do not require any signal processing for compressing the acquired data to the modal space, which inherently results in some loss of information.

Applications of the elastic wave propagation approach in the framework of structural damage identification are recently reported in the literature. For instance, damage assessment in laminated beams, fatigue damage identification and applications in damage assessment and structural health monitoring in aircraft fuselages are reported. The modeling of the acoustic wave propagation phenomenon in the frequency domain by the Espectral Element Method for damage identification purposes has been extensively considered. Applications are reported for beam and plate type structures [3].

For each damage scenario, a pulse-echo synthetic experiment is performed and the excitation and corresponding response are considered in the damage identification procedure. In this work, the damage state is described by the generalized acoustical impedance of the bar, —see Section 2—, $Z(x)$ were x is the position variable along the bar. Therefore, the damage identification is performed by minimizing, with respect to $Z(x)$, the squared norm between the experimental (synthetic) echo and the predicted one.

The main goal of this research is to study the inverse problem of damage identification in bars within the framework of acoustic wave propagation approach. In a previous work [10], the direct problem of one-dimensional acoustic wave propagation was addressed by considering the Sequential Algebraic Algorithm (SAA). The inverse problem of damage identification was then casted as a minimization one, in the time domain, and several optimization schemes were considered for minimizing the squared difference between the experimental (or synthetic) echo and the one predicted by the SAA. As shown in [11], the Luus-Jaakola optimization method has shown to present the better compromise between accuracy and computational cost to solve this problem.

In the following, the difference between the SAA, used previously, and the ISAA will be explained and it will be showed, in the numerical examples, the actual improvement obtained in the damage assessment.

2 Mathematical Foundations

2.1 The algebraic sequential algorithm

The one-dimensional longitudinal acoustic wave propagation in a non-homogeneous slender bar can be described by the hyperbolic second-order differential equation [7]

$$\sigma_{tt} - c^2 \left[\sigma_{xx} + \left(\frac{A'}{A} - \frac{\rho'}{\rho} \right) \sigma_x + \rho \left(\frac{A'}{\rho A} \right)' \sigma \right] = 0, \tag{1}$$

where $\sigma(x,t)$ is the longitudinal stress field, depending on the position x and time t, $A(x)$ is the bar cross section area, $\rho(x)$ is the bar density, both depending on the position x, c is the longitudinal acoustic wave speed, the prime stands for total derivative, and the subscripts, as usual, represent partial derivatives.

The general D'Alembert solution for Eq. (1) cannot be obtained in a closed form. However, it can be shown, [8], that Eq. (1) can be written in an alternative form, in the characteristic plane (r,s), as the following system of first-order equations:

$$\begin{aligned} U_r + \frac{\dot{Z}}{4Z} U &= 0; \\ V_s - \frac{\dot{Z}}{4Z} V &= 0, \end{aligned} \tag{2}$$

where $Z = \rho c A$ is called the *generalized acoustical impedance*, $U(r,s)$ and $V(r,s)$ are, respectively, the *progressive* and *regressive* stress wave components traveling along the characteristic plane, defined as

$$r = t + \tau; \qquad s = t - \tau, \tag{3}$$

and the dot stands as derivative with respect to the independent variable τ, the *travel time*, defined as

$$\tau(x) = \int_0^x \frac{d\xi}{c(\xi)}. \tag{4}$$

Equations (2) are a *compact and uncoupled* pair of first order differential equations that describes the longitudinal acoustic wave propagation phenomenon in a more convenient way. To integrate it, boundary conditions in the (r,s) plane must be provided, corresponding to the physical situation under concern. Let us, for instance, consider the probing of a medium, $x \geq 0$, by a pulse excitation at $x = 0$. Assuming also the Sommerfeld radiation hypothesis [6], the boundary conditions can be stated as:

$$U(s,s) = F(s) = f(t); \quad V(r,0) = 0, \tag{5}$$

where $f(t)$ is the incident longitudinal stress being applied at the boundary $r = s$ ($x = 0$) and the second equation ensures that there is no disturbance at $s \leq 0$ ($t \leq x/c$). Note that $f(t)$, being the longitudinal stress at the physical boundary $x = 0$, corresponds to $U(s,s)$, a *progressive* wave component. Analogously, the echo observed at $x = 0$, due to the inhomogeneity, will be the output signal $g(t) = V(s,s)$, a *regressive* wave component.

Assuming now that the bar under study has length l and is discretized in n sections with equal length $\Delta x = c\Delta t$, so that $l = n\Delta x$. The (known) discrete incoming pulse is written then as

$$F_j = f\big(2(j-1)\Delta t\big), \quad j = 1, 2, \ldots, N, \tag{6}$$

and the discrete outgoing echo is

$$G_j = g(2j\Delta t), \quad j = 1, 2, \ldots, N, \tag{7}$$

where $N\Delta t \leq \Delta T$ is the total time interval under consideration.

It can be shown that Eqs. (2), with the boundary conditions given in Eqs. (5), have, after the discretization given in Eqs. (6) and (7), the following algebraic solution for the echo [8]

$$G_j = \sum_{k=1}^{j} \left(R_k + \sum_{p=1}^{k-2} Q_k^p \right) F_{j-k+1}, \quad j = 1, 2, \ldots, N, \tag{8}$$

where the polynomials Q_k^p have the general recursive formula

$$Q_k^p = R_{k-p} \left[\frac{Q_{k-1}^p}{R_{k-p-1}} - R_{k-p-1} \left(R_{k-1} + \sum_{l=1}^{p-1} Q_{k-1}^l \right) \right], \quad k = 1, 2, \ldots, \quad p = 1, 2, \ldots, k-2. \tag{9}$$

In Eqs. (8, 9), R_i stands for the *reflection coefficient* at the i-th layer of the medium, defined as

$$R_i = \frac{Z_i - Z_{i-1}}{Z_i + Z_{i-1}}, \quad i = 1, 2, \ldots, n. \tag{10}$$

The mathematical procedure, in the direct wave propagation approach, consists then in the following steps. The medium, with a nominal cross-section area A_0 and nominal generalized acoustical impedance Z_0, is discretized into n elements. Then, the reflection coefficients are computed by Eq. (10). In the sequel, the polynomials Q_k^p are calculated from Eq. (9). Finally, the output echo is computed from Eq. (8). The described technique is called *Sequential Algebraic Algorithm* (SAA).

It is worth stressing that the mathematical model above provides an original algebraic formula to solve the direct acoustic wave propagation problem. It also permits, in the identification procedure, to identify one parameter per step. As it will be seen in the damage assessment results, the number of parameters that may be updated in the identification is significantly larger than what is usually found in an optimization processes.

It is worth noting that it is not necessary to consider an infinite or even a semi-infinite medium. The echoes will be observed in the interval ΔT, that means, $t \in (0, 2l/c)$. This means that the echo originated by the right end of the bar — whatever is its boundary condition — is irrelevant for the analysis.

2.2 The improved sequential algebraic algorithm

In this section, an improvement will be introduced in the SAA, resulting in the *improved sequential algebraic algorithm* (ISAA).

Let us consider the progressive and regressive terms in the D'Alembert solution for one dimensional wave propagation as $\Phi(x,t)$ and $\Psi(x,t)$. The expansions in Taylor series for the estimate of Φ_i^j and Ψ_i^j, respectively, were approximated, in the SAA, with a truncation error of order $\Delta\tau$. Having in mind to increase the algorithm accuracy, we propose estimations of second order for these functions.

Expanding then $\Phi_{i+1/2}^{j+1/2}$, $\Phi_{i-1/2}^{j-1/2}$, $\Psi_{i-1/2}^{j+1/2}$ and $\Psi_{i+1/2}^{j-1/2}$ in Taylor series around the point (i,j) one get,

$$\Phi_{i+1/2}^{j+1/2} = \Phi_i^j + \Delta\tau \left(\frac{\partial\Phi}{\partial r}\right)_i^j + \mathcal{O}(\Delta\tau^2); \tag{11}$$

$$\Phi_{i-1/2}^{j-1/2} = \Phi_i^j - \Delta\tau \left(\frac{\partial\Phi}{\partial r}\right)_i^j + \mathcal{O}(\Delta\tau^2); \tag{12}$$

$$\Psi_{i-1/2}^{j+1/2} = \Psi_i^j + \Delta\tau \left(\frac{\partial\Psi}{\partial s}\right)_i^j + \mathcal{O}(\Delta\tau^2); \tag{13}$$

$$\Psi_{i+1/2}^{j-1/2} = \Psi_i^j - \Delta\tau \left(\frac{\partial\Psi}{\partial s}\right)_i^j + \mathcal{O}(\Delta\tau^2). \tag{14}$$

Adding now Eq. (11) with Eq. (12) and Eq. (13) with Eq. (14), one arrives to:

$$\Phi_i^j = \frac{\Phi_{i+1/2}^{j+1/2} + \Phi_{i-1/2}^{j-1/2}}{2} + \mathcal{O}(\Delta\tau^2); \tag{15}$$

$$\Psi_i^j = \frac{\Psi_{i-1/2}^{j+1/2} + \Psi_{i+1/2}^{j-1/2}}{2} + \mathcal{O}(\Delta\tau^2). \tag{16}$$

Denoting $\overline{U}_i^j$ and $\overline{V}_i^j$, respectively, the approximation of $\Phi(\tau_i, t_j)$ and $\Psi(\tau_i, t_j)$, where $\tau_i = i\Delta\tau$ and $t_j = j\Delta t$, one obtain the discretized form for the progressive and regressive waves of order $\Delta\tau^2$:

$$\frac{\overline{U}_{i+1/2}^{j+1/2} - \overline{U}_{i-1/2}^{j-1/2}}{2\Delta\tau} + \frac{1}{4}\frac{2}{\Delta\tau} R_i \left(\frac{\overline{V}_{i-1/2}^{j+1/2} + \overline{V}_{i+1/2}^{j-1/2}}{2}\right) = 0, \quad \text{for } s \text{ constant;} \tag{17}$$

$$\frac{\overline{V}_{i-1/2}^{j+1/2} - \overline{V}_{i+1/2}^{j-1/2}}{2\Delta\tau} - \frac{1}{4}\frac{2}{\Delta\tau} R_i \left(\frac{\overline{U}_{i+1/2}^{j+1/2} + \overline{U}_{i-1/2}^{j-1/2}}{2}\right) = 0, \quad \text{for } r \text{ constant.} \tag{18}$$

Solving Eq. (17) and Eq. (18) for $\overline{U}_{i+1/2}^{j+1/2}$ and $\overline{V}_{i-1/2}^{j+1/2}$, one get:

$$\overline{U}_{i+1/2}^{j+1/2} = \overline{U}_{i-1/2}^{j-1/2} - \frac{R_i}{2}\left(\overline{V}_{i-1/2}^{j+1/2} + \overline{V}_{i+1/2}^{j-1/2}\right), \quad \text{for } s \text{ constant;} \tag{19}$$

$$\overline{V}_{i-1/2}^{j+1/2} = \overline{V}_{i+1/2}^{j-1/2} + \frac{R_i}{2}\left(\overline{U}_{i+1/2}^{j+1/2} + \overline{U}_{i-1/2}^{j-1/2}\right), \quad \text{for } r \text{ constant.} \tag{20}$$

Changing now the indexes i, j to n, p one arrives to:

$$\overline{U}_{n+1}^{p} = \overline{U}_n^p - \frac{R_{n-p}}{2}\left(\overline{V}_n^{p+1} + \overline{V}_n^p\right); \tag{21}$$

$$\overline{V}_n^{p+1} = \overline{V}_n^p + \frac{R_{n-p}}{2}\left(\overline{U}_{n+1}^p + \overline{U}_n^p\right). \tag{22}$$

One cannot obtain a sequential form for the components $\overline{U}_{n+1}^p$ e $\overline{V}_n^{p+1}$, since these components are unknown in Eqs. (21–22). However, one can modify these relationships in order to eliminate the term

$\overline{V}_n^{p+1}$ from Eq. (21) and $\overline{U}_{n+1}^{p}$ from Eq. (22). Substituting then Eq. (22) in Eq. (21) and Eq. (21) in Eq. (22), one get:

$$\overline{U}_{n+1}^{p} = \left(\frac{1-\frac{R_{n-p}^2}{4}}{1+\frac{R_{n-p}^2}{4}}\right)\overline{U}_n^p - \left(\frac{R_{n-p}}{1+\frac{R_{n-p}^2}{4}}\right)\overline{V}_n^p; \tag{23}$$

$$\overline{V}_n^{p+1} = \left(\frac{1-\frac{R_{n-p}^2}{4}}{1+\frac{R_{n-p}^2}{4}}\right)\overline{V}_n^p + \left(\frac{R_{n-p}}{1+\frac{R_{n-p}^2}{4}}\right)\overline{U}_n^p. \tag{24}$$

Defining now $\overline{R}_i$ and $\overline{T}_i$ as the *apparent reflection coefficients* and the *apparent transmission coefficients*, respectively, a new system of recurrence relationships is obtained:

$$\overline{R}_i = \frac{R_i}{1+\frac{R_i^2}{4}}; \tag{25}$$

$$\overline{T}_i = \frac{1-\frac{R_i^2}{4}}{1+\frac{R_i^2}{4}} \tag{26}$$

obeying to the condition,

$$\overline{R}_i^2 + \overline{T}_i^2 = 1. \tag{27}$$

Rewriting the system of equations (23–24) in terms of $\overline{R}$ and $\overline{T}$ one get:

$$\overline{U}_{n+1}^{p} = \overline{T}_{n-p}\overline{U}_n^p - \overline{R}_{n-p}\overline{V}_n^p; \tag{28}$$

$$\overline{V}_n^{p+1} = \overline{T}_{n-p}\overline{V}_n^p + \overline{R}_{n-p}\overline{U}_n^p. \tag{29}$$

This system, to be numerically integrated, needs two boundary conditions, in similar form of Eqs. 5, such as:

$$\overline{U}_{n+1}^{n} = F_n; \quad \overline{V}_n^0 = 0. \tag{30}$$

It is worth noting that when the generalized acoustical impedance variations are smooth, the apparent transmission coefficients $\overline{T}_i \approx 1$ and the apparent reflection coefficients $\overline{R}_i \approx R_i$, resulting that the SAA obtain results practically identical to the ISAA. However, when the generalized acoustical impedance variations are abrupt, the apparent transmission coefficients $\overline{T}_i$ become relevant and, as it will be shown in the numerical results section, this is an important feature in the damage identification accuracy.

3 Optimization Technique

The inverse problem of damage assessment can be stated as: by using the experimental (synthetic) echo $\overline{G}_i$, $i = 1, 2, \ldots, N$, at the time instant t_i, where N is the total number of data considered in the identification process, determine the elements of the unknown vector $\mathbf{A}$, defined as

$$\mathbf{A} = \big(A_1, A_2, \ldots, A_N\big), \tag{31}$$

where A_i is the i^{th} point of the discretization of the damaged region.

The damage identification problem is solved as an optimization problem in which the following objective function is to be minimized,

$$\mathfrak{F}(\mathbf{A}) = \sum_{i=1}^{N} \left[G_i(\mathbf{A}) - \overline{G}_i \right]^2, \tag{32}$$

where $G_i(\mathbf{A})$ is the computed echo by the direct problem model.

To consider the damage assessment with noisy data, the generated synthetic data is considered as

$$\overline{G}_i = G_i^{\text{exact}}(\mathbf{A}_{\text{exact}}) + \epsilon \Gamma_i, \tag{33}$$

where the exact echo $G_i^{\text{exact}}(\mathbf{A}_{\text{exact}})$ are computed from the solution of the direct problem, using exact parameters $\mathbf{A}_{\text{exact}}$; ϵ is the standard deviation of a measurement error and Γ is a computer generated pseudo-random number in the range $[-1, 1]$.

To minimize the objective function (32) the *Luus-Jaakola* (LJ) optimization procedure is considered. This is an well known method and will not be described here. The interested reader will find a detailed description of the method in [4].

4 Numerical Results

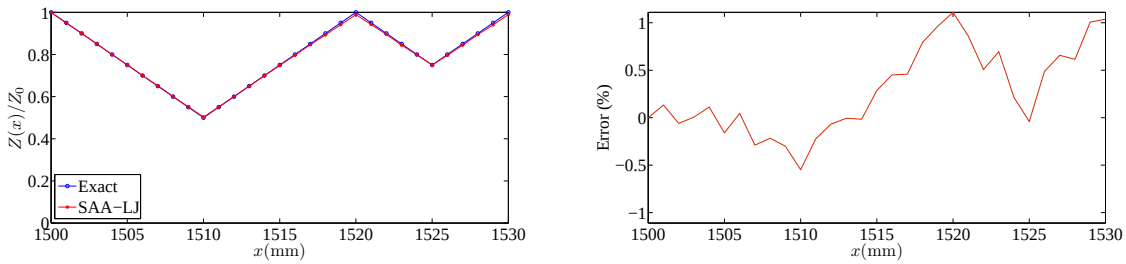
In order to verify the differences in the two presented formulations a bar of length $L = 3,000$ mm was chosen. A level of random null average noise was added to the pulse and echo signals. In all the identifications were used $\epsilon = 0.01$. This corresponds to a signal to noise ratio of about 15 dB. Four different damage scenarios were imposed to the bar. Two of them are smooth (Case 1 and 2) impedance profiles and the other two are rough (Case 3 and 4). In Case 1, two superposed triangular variations were considered, as shown in Fig. 1a. In Case 2, the damage is modeled as a circular hole with diameter of one half of the bar height inserted at the middle of the beam and crossing its entire width. In Cases 3 and 4, the damages are computer-generated random with maximum reduction of 25% and 50% of the cross section area, respectively. The assumed material properties are: density $\rho = 7,894$ kg/m^3 and Young's modulus $E = 167.1$ GPa, so that the sound speed is $c = \sqrt{E/\rho} \approx 4,600$ m/s.

4.1 Smooth damage recovery

For Cases 1 and 2, the Figs. 1–4 present the damage recovery (left) and the remaining error in the identification (right). The identification is almost perfect both using the SAA or the ISAA; however, the errors with the ISAA become lower.

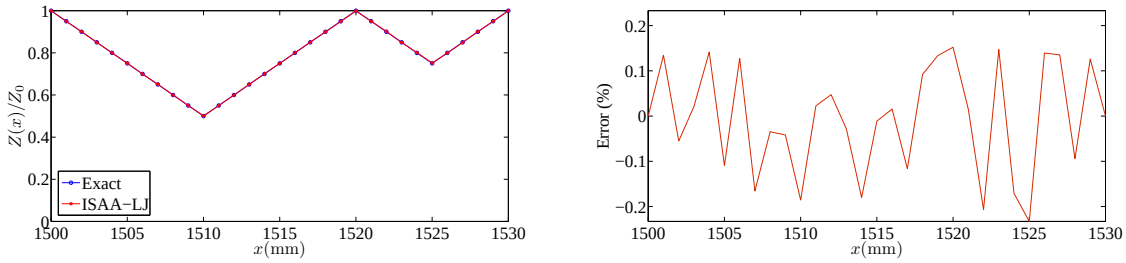
4.2 Rough damage recovery

For Cases 3 and 4 the Figs. 5–8 present the damage recovery (left) and the remaining error in the identification (right). It is worth noting that there is an error propagation in Fig. 5b (SAA) that is almost not noticeable in Fig. 6b (ISAA). For Case 4, with stronger impedance variations, only the ISAA shows to identify almost well, as can be seen in Fig. 8a. However, with this noise level and strong impedance variation still remains some error propagation, as seen in Fig. 8b.



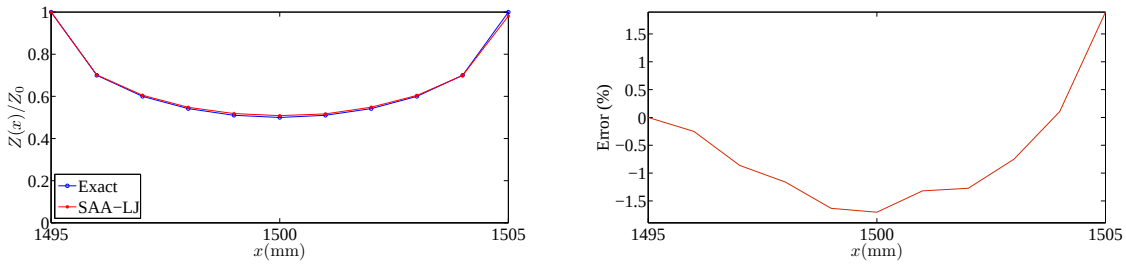
(a) Identification result (b) Relative Error

Figure 1: Damage identification with SAA-LJ for Case 1.



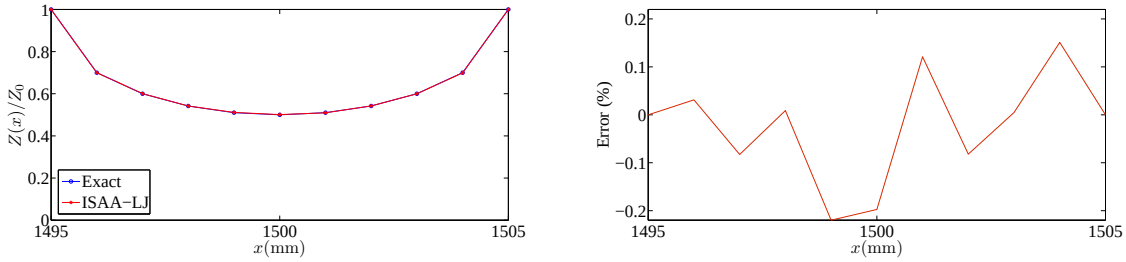
(a) Identification result (b) Relative Error

Figure 2: Damage identification with ISAA-LJ for Case 1.



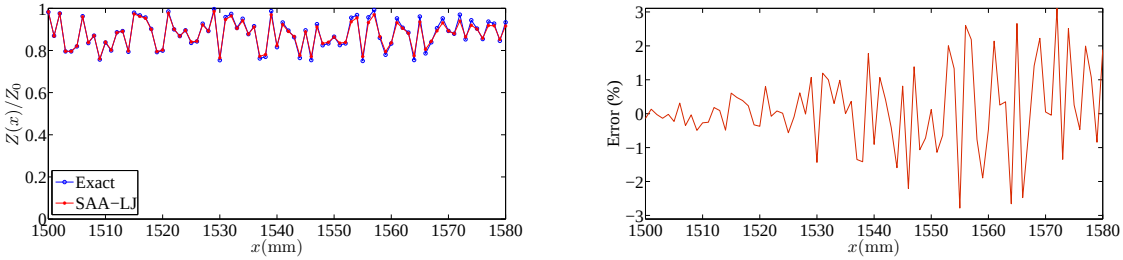
(a) Identification result (b) Relative Error

Figure 3: Damage identification with SAA-LJ for Case 2.



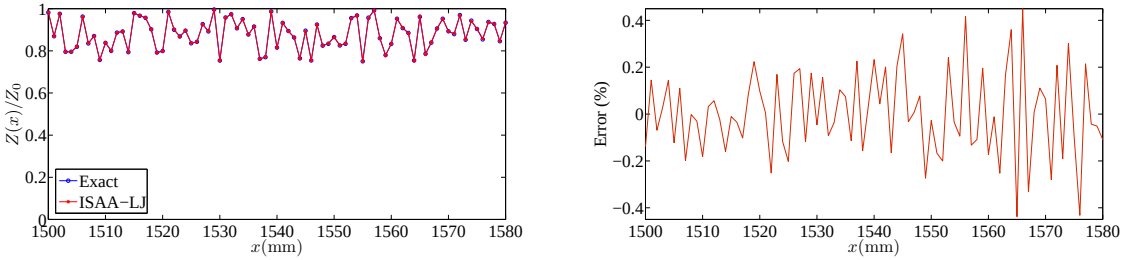
(a) Identification result (b) Relative Error

Figure 4: Damage identification with ISAA-LJ for Case 2.



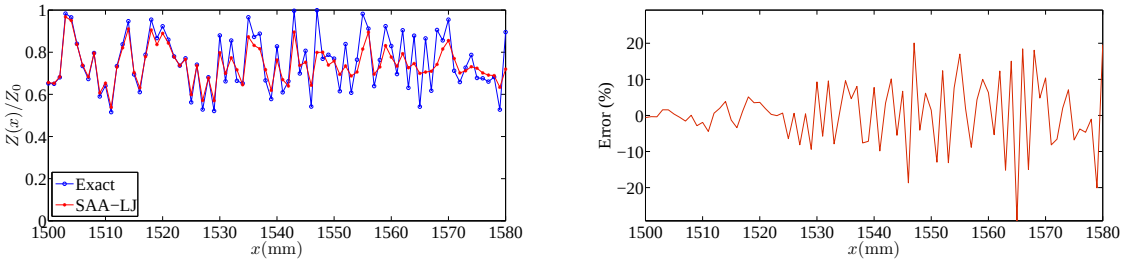
(a) Identification result (b) Relative Error

Figure 5: Damage identification with SAA-LJ for Case 3.



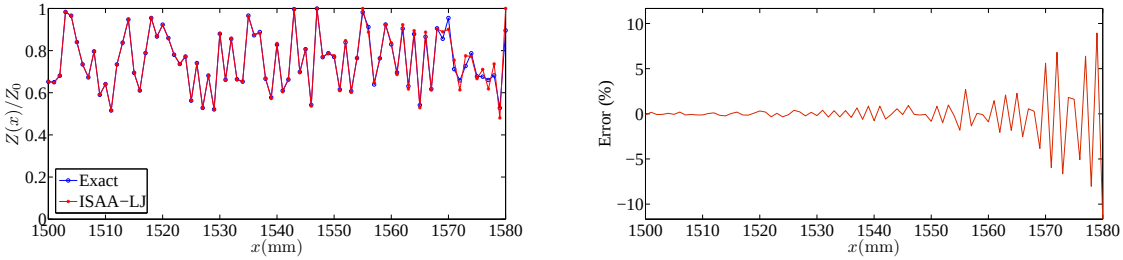
(a) Identification result (b) Relative Error

Figure 6: Damage identification with ISAA-LJ for Case 3.



(a) Identification result (b) Relative Error

Figure 7: Damage identification with SAA-LJ for Case 4.



(a) Identification result (b) Relative Error

Figure 8: Damage identification with ISAA-LJ for Case 4.

5 Conclusions

In this paper, an improvement in the sequential algebraic algorithm (SAA), called improved sequential algebraic algorithm (ISAA) to solve the direct wave propagation in inhomogeneous media is presented. The inverse problem of damage assessment is cast as an optimization procedure. It was shown that the ISAA minimizes the identification error in any situation, being actually important for rough impedance variation cases. It remains some error propagation. This phenomenon, due to the sequential identification procedure, can probably be reduced by using a regularization term in the functional and by applying an hybridization technique, with a deterministic optimization method. It is worth noticing that the most rough impedance variation was assessed with 80 parameters and an average absolute error lower than 4%.

References

[1] J.V. Araújo dos Santos, N.M.M. Maia, C.M. Mota Soares, C.A. Mota Soares, "Structural Damage Identification: A Survey", in B.H.V. Topping, M. Papadrakakis, (Editors), "Trends in Computational Structures Technology", Saxe-Coburg Publications, Stirlingshire, UK, Chapter 1, pp 1-24, 2008. doi:10.4203/csets.19.1

[2] R. Gangadharan, D. Roy Mahapatra, S. Gopalakrishnan, C.R.L. Murthy, and M.R. Bhat. On the sensitivity of elastic waves due to structural damages: Timefrequency based indexing method. *Journal of Sound and Vibration*, 320, 915–941, 2009.

[3] U. Lee. A reduced-domain method of structural damage identification: Application to a spectral element beam model. *Shock and Vibration*, 10, pp. 313–324, 2003.

[4] R. Luus, T.H.L. Jaakola. Optimization by direct search and systematic reduction of the size of search region. *AIChE Journal*, V. 19, 760–788, 1973.

[5] G. Manson and R.J. Barthorpe. Advanced feature selection for simplified pattern recognition within the damage identification framework. *Shock and Vibration*, 17, 589–599, 2010.

[6] A.D. Pierce. Acoustics: An Introduction to its Physical Principles and Applications. Acoustical Society of America, Woodbury, 1989.

[7] E.A. Robinson and S. Treitel. Geophysical Signal Analysis. Englewood Cliffs, Prentice Hall, 1980.

[8] R.A. Tenenbaum and M. Zindeluk. An exact solution for the one-dimensional elastic wave equation in layered media. *J. Acoust. Soc. Am.*, v. 92, n.6, 3364–3370, 1992.

[9] R.A. Tenenbaum, L.T. Stutz, and K.M. Fernandes. Comparison of vibration and wave propagation approaches applied to assess damage influence on the behavior of Euler-Bernoulli beams. *Computers and Structures*, 89, 1820–1828, 2011.

[10] R.A. Tenenbaum, K.M. Fernandes, L.T. Stutz, and A.J. Silva Neto. Damage identification in bars with a wave propagation approach and a hybrid optimization method. *Shock and Vibration*, Vol. 19, pp. 301–321, 2012.

[11] R.A. Tenenbaum, L.T. Stutz, K.M. Fernandes. Damage identification in bars with a wave propagation approach: Performance comparison of five hybrid optimization methods. *Shock and Vibration*, Vol. 20, pp. 863–878, 2013.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Freeze profile identification in an aluminium reduction cell using thermal measurements

Ádám Ugron*, László I. Kiss, Jean-François Bilodeau, Sébastien Guérard

Research Group for Process and Systems Engineering
Université du Québec à Chicoutimi, 555 boul. de l'Université; Chicoutimi, QC, Canada
e-mail: adam.ugron1@uqac.ca

Key words: aluminium reduction, sidewall freeze, thermal measurement

Abstract
The frozen bath on the sidewalls of aluminium electrolysis cells plays a crucial role in their stable operation. The efforts today for increasing the production affect the stability of the freeze layer. We aim to adapt the detection of the phase change front by means of in-situ temperature measurements under industrial conditions, in aluminium electrolysis cell sidewalls. This way we can identify the thickness and shape of the protective freeze layer. We carried out continuous temperature measurements over several months at two locations in the sidewalls of a cell. We built a simplified, two-dimensional model of the sidewall and the side-ledge and used it to identify the isothermal surface of the freeze. With data generated by numerical simulations the inverse estimation worked well, while in case of measured temperatures, the inverse algorithm successfully converged to a smooth solution providing an estimate for the freeze shape. The identified profiles showed reasonable agreement with profiles obtained by mechanical probing and captured tendencies over time correctly. With the further development of the method we aim to develop a monitoring method suitable for industrial application.

1 Introduction

Nowadays aluminium in industrial scale is exclusively produced using the Hall-Héroult process patented in 1886. During this process reduction cells are used to produce aluminium from alumina by electrolysis. The alumina is dissolved in molten cryolite which, with a small amount of additives like alkali fluorides, has a melting point of about 960 °C [1]. The high temperature is maintained by the Joule heat produced by the more than 100 kiloamperes of current passing through the resistances of a cell. The cryolite bath has an excellent solvent capacity, rendering it chemically aggressive, attacking almost all known materials, including the cell sidewalls. As a result of the electrolysis, molten aluminium descends onto the surface of the cathode and carbon dioxide is produced as the alumina reacts with the carbon on the surface of the anodes. The gas travels along the anode surface in the form of bubbles resulting in a fluctuating resistance between the anode and the cathode. During

normal operation, the bath in which alumina is dissolved solidifies along the sidewalls of the cells as those are intentionally less insulated.

The thickness and form of the so-called freeze layer or side-ledge is a result of the thermal equilibrium between the bath and sidewall. The movement of the molten cryolite driven by the drag force exerted by the gas bubbles and the magneto-hydrodynamic forces in the molten aluminum also affect the freeze shape. The resulting freeze is bound by the sidewall and the isothermal liquid-solid interface. The typical steady-state temperature distribution resulting across the sidewalls is depicted in Figure 1.

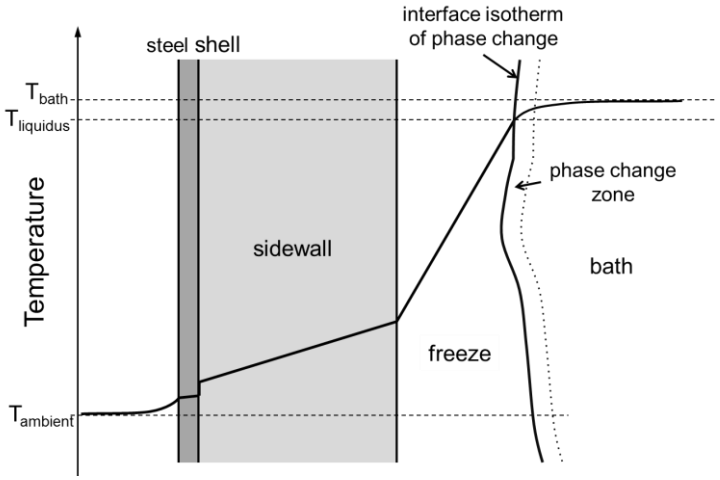


Figure 1: Steady-state temperature distribution across the sidewall of aluminium electrolysis cells

The freeze has a threefold role in the aluminium reduction process. First, the freeze layer on the sidewalls provides an effective protection against the corrosive bath. Without this protective layer the bath may come in direct contact with the sidewall destroying the cell. Second, the dynamic freezing and melting of the side-ledge helps to maintain the thermal stability of the cell. Finally, the freeze layer helps maintaining the optimal direction of the vertical electrical current between the anodes and the cathode by acting as an insulator on the sidewalls.

During steady operation, the shape of the side-ledge in the aluminium electrolysis cells remains stable, corresponding roughly to the isotherm of the liquidus temperature of the bath (see Figure 1), but transient effects occur during normal operation. These effects can strongly modify the side-ledge profile as a result of the system trying to regain its thermal stability. Knowing the importance of the freeze, the thickness is of special interest in aluminium production. Commonly a manual measurement method is used to probe the shape and thickness of the side-ledge. Due to its intrusive and labour-intensive nature this method of freeze detection is only utilised 2-3 times a year [2] and the results are uncertain. There are huge efforts today to increase the production, which is mainly achieved by increasing the useful area of the cells using larger anodes. The higher production area demands higher current, increasing the internal heat generation in the cells. These developments affect the stability of the freeze layer; therefore, there is a need for methods, which provide information on the side-ledge in a continuous way at strategically selected points. For this purpose the detection of the phase change front by means of temperature measurements can be a useful method.

The detection of the phase change front of a solidifying molten liquid or the melting of a solid medium is of special interest in various industries like casting, heat storage systems [3], welding [4, 5] and furnaces [6]. It was demonstrated that the phase change front can be identified with the help of temperature measurements [3- 7] but the studies are limited by their generated test cases or reference

values originating from well controlled experiments. Sorli and Skaar [8] presented a methodology to detect the wear-line in a melting furnace by associating it with an isothermal line. They carried out measurements in the sidewall of a furnace involving nine thermocouples distributed along a line of the wall contour and compared these with their identified profiles. We developed and tested a similar methodology to theirs that can detect the phase change front on the sidewall of industrial aluminium electrolysis cells based on continuously monitored temperatures inside the cell walls. Our approach is adapted to electrolysis cells and able to give a first rough estimate of the freeze front based on the measured data on which the inverse identification can be based.

2 Methodology

We carried out continuous temperature measurements over several months at two locations in the sidewalls of an industrial point-fed prebake electrolysis cell operated by Rio Tinto Alcan. Two arrays of appropriately protected thermocouples were embedded in two sidewall blocks along the longitudinal sides of the electrolysis cell to measure the temperatures. The thermocouples in the arrays were arranged so that the local temperature gradients and second order derivatives (excluding the mixed derivatives) could be estimated. All thermocouples were calibrated and measures were taken to minimize the thermal contact resistance between the tip of the thermocouples and the material of the sidewall block. We registered one thermocouple signal in the array as reference temperature; all other thermocouple signals were recorded as differences to the reference. The temperature signals were recorded with a sampling interval of one minute using the data acquisition system of the plant. For comparison, during the measurement period the freeze shape was measured several times by means of mechanical probing.

We used the ANSYS parametric design language (APDL) to implement the inverse algorithm of the freeze shape identification. The algorithm was based on a simplified, two-dimensional model of the sidewall and the side-ledge (Figure 2).

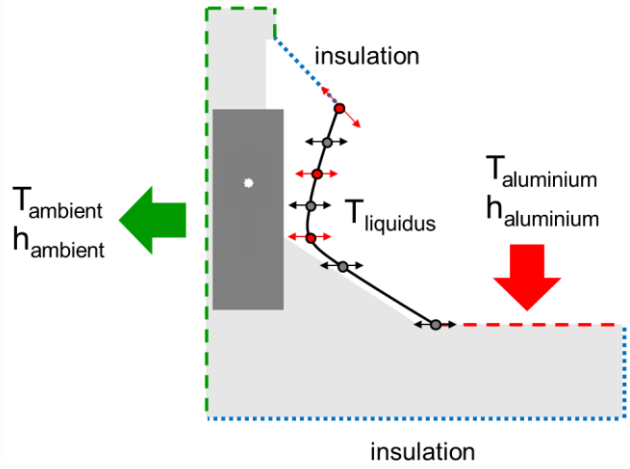


Figure 2: Numerical model of the sidewall and the freeze

We modeled the influence of the liquid aluminium with a heat transfer coefficient on the cathode surface. Similarly, the heat loss on the external surface of the sidewall was modeled with the ambient temperature and a heat transfer coefficient. All other boundaries were considered adiabatic. The positions of the bounding surfaces on the bottom were chosen so that they do not influence the

temperature at the temperature measurement points. The truncated surface of the freeze profile at the top of the model is set as a 45° line through the corner point of the side block edge and a point on the ledge profile at the electrolytic bath upper level, an approach adopted after Dupuis [9]. Finally, the isothermal phase change surface is defined by seven control points along its contour, which are moving according to the progress of the optimisation cycle of the inverse algorithm.

The inverse identification was driven by the built-in APDL gradient based optimisation; a least square cost function was used with uneven weights of the temperature and temperature difference agreement, the latter having a stronger influence on the results. Additionally, a second order regularisation [10] was used to obtain smooth freeze profiles by prohibiting instabilities arising from the ill-posed nature of similar problems [8]. The cost function used in our algorithm is given in Equation (1).

$$J = \beta \left(Y_{ref} - T_{ref}\right)^2 + \sum_{i=1}^{n} \left(\Delta Y_i - \Delta T_i\right)^2 + \alpha \sum_{j=1}^{r-2} \left[c_j \cdot s_{j+2} - s_{j+1} + \left(1 - c_j\right) s_j\right]^2 \tag{1}$$

$$c_j = \frac{y_{j+1} - y_j}{y_{j+2} - y_j} \tag{2}$$

Here Y (ΔY) and T (ΔT) denote the measured and the calculated temperatures (temperature differences), respectively. While n denotes the number of temperature difference sampling points, r refers to the number of active control points in the optimisation loop. α is the regularisation constant the value of which is chosen so that we obtain a smooth approximation of the isothermal profile. β is a weight factor to reduce the importance of absolute temperature match and emphasize the match of temperature gradients. The basic idea behind this lies in the nature of thermocouple accuracy, namely, the temperature readings are significantly less accurate than the differential signals between the thermocouples. Also, due the high temperature level, the difference between the calculated and measured temperatures gives a much higher numerical value than the difference between the calculated and measured temperature differences. The value of β is chosen so that the two different sources in the cost function are compensated. The c_j weight factors of the s_j freeze thickness control parameters are responsible for the compensation of the non-uniform spacing of the control points along the vertical axis (y) in the second order regularisation.

The direct problem is assumed to be steady, since the temperature changes during normal operation are slow. This assumption was verified with the measured gradients and second order derivatives, which showed well-established temperature distributions in the planar sidewall.

The developed inverse algorithm consists of three main steps, each of which gives an estimate of the freeze thickness and profile. The subsequent steps are created so that they increase the accuracy of the result of the previous step. The first step directly gives a rough guess based on the measured data. We apply a second order extrapolation using the Taylor series to directly estimate the freeze layer thickness and shape locally. This direct estimation provides an initial guess for the identification routine. In the second stage we select three control points (solid red points in Figure 2) on the top half of the freeze that have higher influence on the temperature field in the region of the thermocouple array. These points are optimized while we apply a linear interpolation between them to calculate the values of the unused, passive control points. The bottom part of the freeze is set to a constant thickness equal to the thickness at the lowest active control point in this step. In the last step we include all control points in the optimization loop to refine the converged solution of the optimisation in the second step.

3 Results and discussion

3.1 Generated data

Using typical test profiles, we generated artificial temperature distributions in the solid domain that were fed as inputs of the inverse algorithm. In Figure 3 the estimation of a profile at the end of each of the three algorithmic steps is compared with the reference profile used to generate the artificial temperature fields.

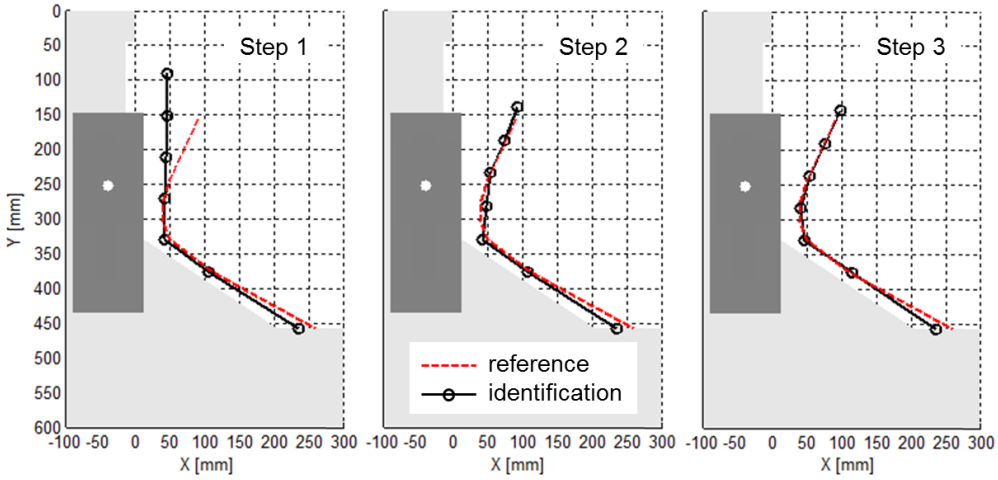


Figure 3: Comparison of the three steps of the identification with a prescribed freeze profile (red)

Figure 3 shows that the sample test profile was successfully reconstructed from the simulated temperature values. The first approximation already gives a safe estimation of the freeze thickness, the cost function at the end of this stage has a value of J = 270, with a negligible computation time. The second step gives an excellent match of the reference and identified profiles with only three control points. The cost function is J = 0.9 after approximately 15 minutes of computation time on an average desktop PC. The third step in the present case takes approximately 7 additional minutes resulting in a perfect match on most of the reference profile (J = 0.4). The largest difference of the resulting profile appears at the bottom. The reduced sensitivity of the measurement array at this distance is the result of the limited "field of view" of the array.

To identify the uncertainties related to our simplified model we carried out a sensitivity study on the parameters affecting the resulting side-ledge profiles. These parameters included the freeze thermal conductivity, the liquid aluminium and the ambient temperatures, the heat transfer coefficients and the liquidus temperature. The model was found to be most responsive to the changes of the freeze thermal conductivity and the ambient heat transfer coefficient (Figure 4). We found that with these parameters the positions of the freeze control points have a near to unity sensitivity coefficient. The only exception is the thinnest part of the freeze considered as the critical section, which shows a significantly lower sensitivity in a wide range of parameter variation. These results of the sensitivity study point to the inaccuracy of model parameters, but at the same time they also show that the inverse estimation is safely approximating the minimal thickness of the freeze.

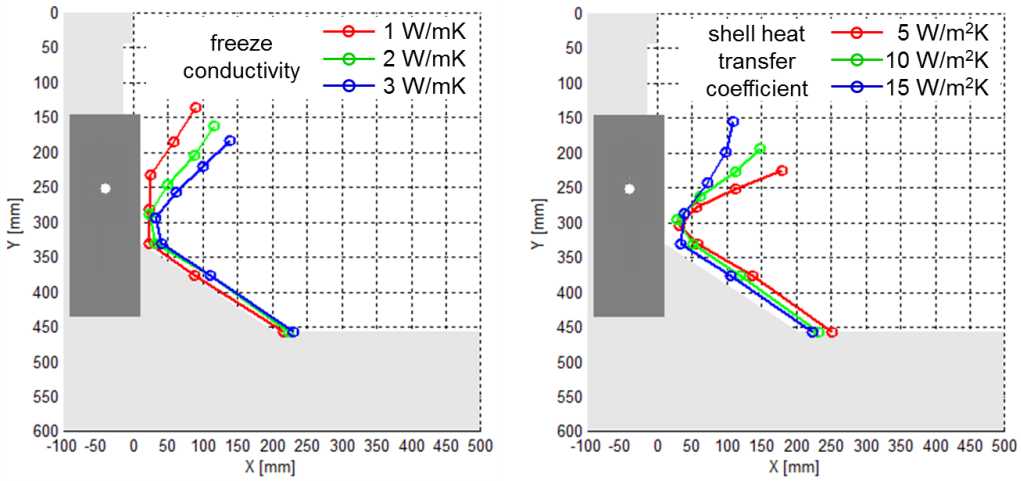


Figure 4: Main findings of the profile sensitivity study: the two most significantly influencing parameters are the freeze conductivity (left) and the heat transfer coefficient on the shell (right)

3.2 In-situ data

In case of the measured temperatures in the electrolysis cell wall, the presented algorithm reliably converged to a smooth solution providing an estimate for the freeze shape. Two resulting freeze profiles at one of the measured sections of the cell sidewall with a one-month follow-up interval are given in Figure 5. The identified profiles are compared with the corresponding mechanical probing data.

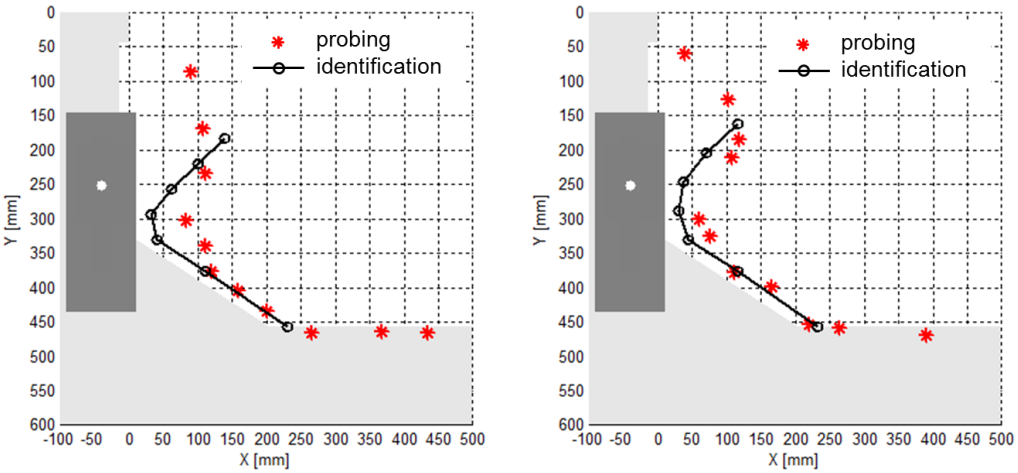


Figure 5: Identifications based on measured temperatures compared with mechanical probing data

Although the comparison of the mechanical probing and the inverse identification contains significant uncertainties, the profiles in general show reasonable agreement. Additionally, the identification successfully captured the tendency of evolution between different states of the freeze probed with a one- month difference. On the left-hand side of Figure 5 the top part of the freeze is apparently thicker than on the right-hand side image as reflected by the identified profiles. While capturing the tendency

of the freeze thickness change it can be observed that the identified profiles underestimate the thickness. Despite of the deviations in the estimated thickness, in a real monitoring situation this identification improves the security of cell operation during control.

3.3 Application of the method

We applied the presented algorithm to a typical event in the operation of an electrolysis cell. To see how an anode change near the detection point affects the freeze thickness and shape, we carried out a series of profile identifications. In Figure 6 the profiles are plotted along the evolution of the reference temperature signal in time.

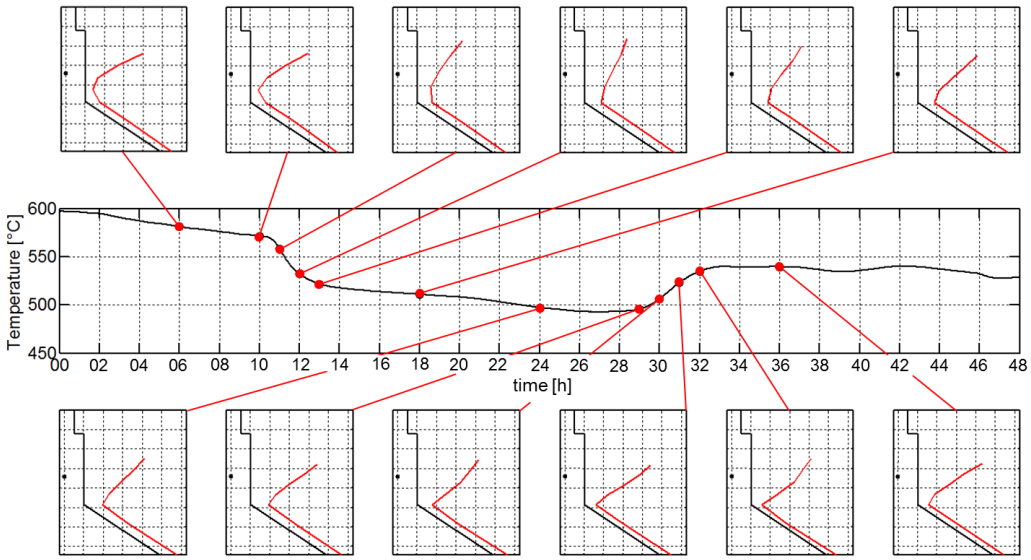


Figure 6: Development of the freeze profile in case of an anode change near the thermocouple array

The temperature signal allows us to estimate the anode removal around the 10th hour of the presented period. Before this event we see the reference temperature decreasing slightly but no change is detected in the freeze parameters. As a result of the removed anode the identification predicts a significant increase of thickness at the weakest point of the freeze. Simultaneously, the top part of the freeze shrinks, which is the direct consequence of the process of an anode removal, during which the crust together with the top of the freeze is crushed around the anode to facilitate its removal. In the following hours the identification captures simultaneous thickening of the thinnest and top parts of the freeze. As the reference temperature starts to increase in the 29th hour, as the new anode becomes working, the freeze starts to thin in the critical section. The bottom part of the freeze is also thinning slightly.

These identified profiles suggest that the melting and freezing process happens rapidly compared to the overall timescale of the processes in the electrolysis cell. Such changes that may play an important role not only in the control of the cell, but in its overall lifetime, could not be followed with the intrusive mechanical probing method.

4 Conclusions

The general tendency of increasing the thermal loads in aluminium electrolysis cells demands regular follow-up of the sidewall thermal state during its operation. We are developing a method to actively follow the protective freeze layer on the sidewalls. The developed inverse model, based on continuous temperature measurements in the cell walls, is able to identify the position of the varying isothermal boundary associated with the frozen bath layer. The identification of the profiles are more accurate in case of simulated data, which is understandable knowing that the modelling assumptions cannot reflect perfectly the real conditions of the thermal measurements. The analysis of the influence of routine operations, such as an anode change, on the freeze shape gives a better understanding on the thermal load of the pot structure and it gives support for improvements.

With further development, we aim to create a monitoring method suitable for helping the efficient and long-lasting operation of industrial cells.

Acknowledgements

The authors would like to thank for the support of the Natural Sciences and Engineering Research Council of Canada and that of Rio Tinto Alcan of the project RDCell; a part of which is related to the present work.

References

[1] J. Thonstad, P. Fellner, G.M. Haarberg, J. Hives, H. Kvande, A. Sterten, *Aluminium electrolysis: Fundamentals of the Hall-Héroult process,* Aluminium-Verlag, Düsseldorf, Germany, (2001).

[2] K. Grjotheim, H. Kvande, *Introduction to aluminium electrolysis*, Aluminium-Verlag, Düsseldorf, Germany, (1993).

[3] M.H. Chun, H.O. Choi, H.G. Jun, Y.S. Kim, *Phase-change front prediction by measuring the wall temperature on which solidification occurs,* International Journal of Heat and Mass Transfer, 30, (1987), pp. 2641-2650.

[4] Y. Ruan, N. Zabaras, *An inverse finite-element technique to determine the change of phase interface location in two-dimensional melting problems*, Communications in Applied Numerical Methods, 7, (1991), pp. 325-338.

[5] R.G. Keanini, N.N. Desai, *Inverse finite element reduced mesh method for predicting multi-dimensional phase change boundaries and nonlinear solid phase heat transfer*, International Journal of Heat and Mass Transfer, 39, (1996), pp. 1039-1049.

[6] M.A. Marois, M. Désilets, M. Lacroix, *Prediction of a 2-D solidification front in high temperature furnaces by an inverse analysis*, Numerical Heat Transfer, Part A, 59, (2011), pp. 151-161.

[7] P. Boily, L.I. Kiss, R.T. Bui, P. Desclaux, *Sensitivity analysis of the thermal detection of the freeze profile in an aluminium reduction cell*, TMS Light Metals 2001, (2001), pp. 1209-1215.

[8] K. Sorli, I.M. Skaar, *Monitoring the wear-line of a melting furnace*, in Proceedings of 3ICIPE Port Ludlow, WA, (1999).

[9] M. Dupuis, *Computation of Aluminum Reduction Cell Energy Balance using ANSYS Finite Element Models*, TMS Light Metals 1998, (1998), pp. 409-417.

[10] A.N. Tikhonov, V.Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston & Sons, Washington D.C., (1977).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Computational algorithms for solving the coefficient inverse problem for parabolic equations

Petr N. Vabishchevich*, Vasily I. Vasil'ev

Nuclear Safety Institute, Russian Academy of Sciences
B. Tulskaya 52, Moscow, Russia
e-mail: vabishchevich@gmail.com

North-Eastern Federal University
Belinskogo 58, Yakutsk, Russia
e-mail: vasvasil@mail.ru

Key words: inverse problem, identification of the coefficient, parabolic partial differential equation, finite difference scheme

Abstract

Among inverse problems for PDEs we distinguish coefficient inverse problems, which are associated with the identification of coefficients and/or the right-hand side of an equation using some additional information. When considering time-dependent problems, the identification of the coefficient dependences on space and on time is usually separated into individual problems. In some cases, we have linear inverse problems (e.g., identification problems for the right-hand side of an equation); this situation essentially simplify their study. This work deals with the problem of determining in a multidimensional parabolic equation the lower coefficient that depends on time only. To solve numerically a nonlinear inverse problem, linearized approximations in time are constructed using standard finite difference approximations in space. The computational algorithm is based on a special decomposition, where the transition to a new time level is implemented via solving two standard elliptic problems.

1 Introduction

In the theory and practice of inverse problems for partial differential equations (PDEs), much attention is paid to the problem of the identification of coefficients from some additional information [1, 2]. Particular attention should be given to inverse problems for PDEs [3, 4]. In this case, a theoretical study includes the fundamental questions of uniqueness of the solution and its stability both from the viewpoint of the theory of differential equations [4, 5] and from the viewpoint of the theory of optimal control for distributed systems [6]. Many inverse problems are formulated as non-classical problems for PDEs. To

solve approximately these problems, emphasis is on the development of stable computational algorithms that take into account peculiarities of inverse problems [7, 8].

Much attention is paid to the problem of determining the right-hand side, lower and leading coefficients of a parabolic equation of second order, where, in particular, the right-hand side and the coefficients depends on time only. An additional condition is most often formulated as a specification of the solution at an interior point or as the average value that results from integration over the whole domain. The existence and uniqueness of the solution to such an inverse problem and well-posedness of this problem in various functional classes are examined, for example, in the works [9, 10, 11, 12].

Numerical methods for solving problems of the identification of the right-hand side, lower and leading coefficients for parabolic equations are considered in many works [13, 14, 15, 16, 17]. In view of the practical use, we highlight separately the studies dealing with numerical solving inverse problems for multidimensional parabolic equations [18, 19, 8]. To construct computational algorithms for the identification of the lower coefficient of a parabolic equation, there is widely used the idea of transformation of the equation by introducing new unknowns that results in a linear inverse problem.

Approximation in space is performed using the standard finite differences [20, 21] or finite elements [22, 23]. The main features of the nonlinear inverse problem are taken into account via a proper choice of the linearized approximation in time. In this paper, for a multidimensional parabolic equation, we consider the problem of determining the lower coefficient that depends on time only. Linear problems at every time level are solved on the basis of a special decomposition into two standard elliptic problems.

2 Problem formulation

For simplicity, we restrict ourselves to a 2D problem in a rectangle. Let $\boldsymbol{x} = (x_1, x_2)$ and

$$\Omega = \{\boldsymbol{x} \mid \boldsymbol{x} = (x_1, x_2)\,, \quad 0 < x_\alpha < l_\alpha, \quad \alpha = 1, 2\}.$$

The direct problem is formulated as follows. We search $u(\boldsymbol{x}, t)$, $0 \le t \le T$, $T > 0$ such that it is the solution of the parabolic equation of second order:

$$\frac{\partial u}{\partial t} - \mathrm{div}(k(\boldsymbol{x})\mathrm{grad}u) + p(t)u = f(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \Omega, \quad 0 < t \le T. \tag{1}$$

The boundary and initial conditions are also specified:

$$k(\boldsymbol{x})\frac{\partial u}{\partial n} = 0, \quad \boldsymbol{x} \in \partial\Omega, \quad 0 < t \le T, \tag{2}$$

$$u(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \tag{3}$$

where n is the normal to Ω. The formulation (1)–(3) presents the direct problem, where the right-hand side, coefficients of the equation as well as the boundary and initial conditions are given.

Let us consider the inverse problem, where in equation (1), the coefficient $p(t)$ is unknown. An additional condition is often formulated as

$$\int_\Omega u(\boldsymbol{x}, t)\omega(\boldsymbol{x})d\boldsymbol{x} = \varphi(t), \quad 0 < t \le T, \tag{4}$$

where $\omega(\boldsymbol{x})$ is a weight function. In particular, choosing $\omega(\boldsymbol{x}) = \delta(\boldsymbol{x} - \boldsymbol{x}^*)$ ($\boldsymbol{x}^* \in \Omega$), where $\delta(\boldsymbol{x})$ is the Dirac δ-function, from (4), we get

$$u(\boldsymbol{x}^*, t) = \varphi(t), \quad 0 < t \leq T. \tag{5}$$

We assume that the above inverse problem of finding a pair of $u(\boldsymbol{x}, t),\ p(t)$ from equations (1)–(3) and additional conditions (4) or (5) is well-posed. The corresponding conditions for existence and uniqueness of the solution are available in the above-mentioned works. In the present work, we consider only numerical techniques for solving these inverse problems omitting theoretical issues of the convergence of an approximate solution to the exact one.

From the nonlinear inverse problem, we can proceed to the linear one. Suppose

$$v(\boldsymbol{x}, t) = \chi(t) u(\boldsymbol{x}, t), \quad \chi(t) = \exp\left(\int_0^t p(\theta) d\theta\right).$$

Then from (1)–(3), we get

$$\frac{\partial v}{\partial t} - \operatorname{div}(k(\boldsymbol{x}) \operatorname{grad} v) = \chi(t) f(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \Omega, \quad 0 < t \leq T,$$

$$k(\boldsymbol{x}) \frac{\partial v}{\partial n} + g(\boldsymbol{x}) v = 0, \quad \boldsymbol{x} \in \partial\Omega, \quad 0 < t \leq T,$$

$$v(\boldsymbol{x}, 0) = u_0(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega.$$

The additional conditions (4) and (5) to identify uniquely $v(\boldsymbol{x}, t),\ \chi(t)$ take the form

$$\int_\Omega v(\boldsymbol{x}, t) \omega(\boldsymbol{x}) d\boldsymbol{x} = \chi(t) \varphi(t), \quad 0 < t \leq T,$$

$$u(\boldsymbol{x}^*, t) = \chi(t) \varphi(t), \quad 0 < t \leq T.$$

The above transition from the nonlinear inverse problem to the linear one is in common use for numerically solving problems of identification. In our work, we focus on the original formulation of the inverse problem (1)–(4) (or (1)–(3), (5)) without going to the linear problem.

3 Semi-discrete problem

To solve numerically the time-dependent convection-diffusion problem, we introduce the uniform grid in the domain Ω:

$$\omega = \{\boldsymbol{x} \mid \boldsymbol{x} = (x_1, x_2), \quad x_\alpha = \left(i_\alpha + \frac{1}{2}\right) h_\alpha, \quad i_\alpha = 0, 1, ..., N_\alpha, \quad (N_\alpha + 1) h_\alpha = l_\alpha, \quad \alpha = 1, 2\}.$$

For grid functions, we define the Hilbert space $H = L_2(\omega)$, where the scalar product and norm are given as follows:

$$(y, w) \equiv \sum_{\boldsymbol{x} \in \omega} y(\boldsymbol{x}) w(\boldsymbol{x}) h_1 h_2, \quad \|y\| \equiv (y, y)^{1/2}.$$

The difference operator for the diffusion transport D has the following additive representation:

$$D = \sum_{\alpha=1}^{2} D_\alpha, \quad \alpha = 1, 2, \quad \boldsymbol{x} \in \omega, \tag{6}$$

where $D_\alpha,\ \alpha = 1, 2$ are associated with the corresponding differential operator in one spatial direction.

For all nodes except ajoining the boundary, and for sufficiently smooth diffusion coefficients $k(\boldsymbol{x})$, the grid operator D_1 can be written as:

$$D_1 y = -\frac{1}{h_1^2} k(x_1 + 0.5h_1, h_2)(y(x_1 + h_1, h_2) - y(\boldsymbol{x})) + \frac{1}{h_1^2} k(x_1 - 0.5h_1, h_2)(y(\boldsymbol{x}) - y(x_1 - h_1, h_2)),$$
$$\boldsymbol{x} \in \omega, \quad x_1 \neq 0.5h_1, \quad x_1 \neq l_1 - 0.5h_1.$$

At the nodes ajoining the boundary, approximation should take into account the boundary condition (2):

$$D_1 y = -\frac{1}{h_1^2} k(x_1 + 0.5h_1, h_2)(y(x_1 + h_1, h_2) - y(\boldsymbol{x})),$$
$$\boldsymbol{x} \in \omega, \quad x_1 = 0.5h_1,$$
$$D_1 y = \frac{1}{h_1^2} k(x_1 - 0.5h_1, h_2)(y(\boldsymbol{x}) - y(x_1 - h_1, h_2)),$$
$$\boldsymbol{x} \in \omega, \quad x_1 = l_1 - 0.5h_1.$$

The grid operator D_2 is constructed in a similarly way. Direct calculations yield (see, e.g., [21, 24]):

$$D_\alpha = D_\alpha^* \geq 0, \quad \alpha = 1, 2.$$

This grid operator of diffusion approximates the corresponding differential operator with an accuracy of $O\left(|h|^2\right)$. As in the differential case, the difference operator of diffusive transport (6) is self-adjoint and positive definite in H:

$$D = D^* \geq 0. \tag{7}$$

In view of (7), we can obtain the corresponding a priori estimates for the solution of the boundary value problem (1)–(3) in H that ensure the stability of the solution with respect to the initial data and the right-hand side.

After discretization in space, from the problem (1)–(3), we arrive at the Cauchy problem for the semi-discrete equation:

$$\frac{dy}{dt} + Dy + p(t)y = f(t), \quad 0 < t \leq T, \tag{8}$$
$$y(0) = u_0. \tag{9}$$

We consider the case, where the lower coefficient in the parabolic equation (1) is negative:

$$p(t) < 0, \quad p_m = \max_{0 \leq t \leq T} |p(t)|. \tag{10}$$

Then we have

$$\|y(t)\| \leq \exp(p_m t)\|u_0\| + \int_0^t \exp(p_m(t - \theta))\|f(\theta)\| d\theta, \tag{11}$$

i.e., the norm of the solution of the homogeneous equation ($f(\boldsymbol{x}, t) = 0$ in (1)) may grow exponentially with time. The a priori estimate (11) holds in the Banach space of grid functions $L_\infty(\omega)$, where

$$\|\cdot\| = \|\cdot\|_\infty, \quad \|y\|_\infty \equiv \max_{\boldsymbol{x} \in \omega} |y|.$$

This fact can be established on the basis of the maximum principle for grid functions and the relevant comparison theorems [21] taking into account the diagonal dominance of the matrix (operator) D.

4 Time-stepping techniques

Let us define a uniform grid in time $t^n = n\tau,\ n = 0, 1, ..., N,\ \tau N = T$ and denote $y^n = y(t^n),\ t^n = n\tau$. We start with discretization in time for the numerically solving direct problem (8), (9). To solve numerically boundary value problems for transient diffusion-reaction equation (1), approximation in time is carried out depending on the sign of the coefficient $p(t)$ [25, 26]. In particular, for $p(t) \geq 0$, unconditionally stable schemes are constructed using the implicit approximation for the lower coefficient, where

$$\frac{y^{n+1} - y^n}{\tau} + Dy^{n+1} + p^{n+1}y^{n+1} = f^{n+1}.$$

If $p(t) \leq 0$, then we need to focus on the explicit approximation:

$$\frac{y^{n+1} - y^n}{\tau} + Dy^{n+1} + p^{n+1}y^n = f^{n+1}, \quad n = 0, 1, ..., N-1. \tag{12}$$

The initial condition (9) yields

$$y^0 = u_0. \tag{13}$$

For the coefficient $p(t)$ with alternating signs, unconditionally stable schemes are based on the explicit-implicit approximation for the lower term. If we apply the explicit approximation (12) for the problem (8), (9), then, for $p(t) \geq 0$, the stability of the scheme (12), (13) holds with rather weak restrictions on the time step [26].

Under the assumptions (10), the difference solution of the problem (12), (13) satisfies the following level-wise estimate in $L_\infty(\omega)$:

$$\|y^{n+1}\| \leq (1 + \tau p_m)\|y^n\| + \tau\|f^{n+1}\|, \quad n = 0, 1, ..., N-1. \tag{14}$$

The estimate (14) is a discrete analog of the estimate (11) for the solution of the problem (8)–(10). To prove (14), we cam apply the maximum principle for grid functions [21, 27]. The second possibility to check the a priori estimate (14) is associated with the use of the concept of the logarithmic norm [28, 29].

5 Algorithm for solving the inverse problem

For the fully discretized (both in space and in time) direct problem (1)–(3), we can solve the inverse problem of the identification of the lower coefficient $p(t)$. We restrict ourselves to the case, where an additional information on the solution is defined (see (5)) at some interior node $\boldsymbol{x}^* \in \omega$ of the grid:

$$y^{n+1}(\boldsymbol{x}^*) = \varphi^{n+1}, \quad n = 0, 1, ..., N-1. \tag{15}$$

This inverse problem is nonlinear, but for discretization in time, we can use the explicit-implicit scheme (12). This approximation in time leads to linear problems at each time level.

For the approximate solution of the problem (12), (13), (15) at the new time level y^{n+1}, we introduce the following decomposition (see, e.,g., [8, 30]):

$$y^{n+1}(\boldsymbol{x}) = v^{n+1}(\boldsymbol{x}) + p^{n+1}w^{n+1}(\boldsymbol{x}), \quad \boldsymbol{x} \in \omega. \tag{16}$$

To find $v^{n+1}(\boldsymbol{x})$, we employ the equation

$$\frac{v^{n+1} - y^n}{\tau} + Dv^{n+1} = f^{n+1}, \quad n = 0, 1, ..., N-1. \tag{17}$$

The function $w^{n+1}(\boldsymbol{x})$ is determined from

$$\frac{1}{\tau}w^{n+1} + Dw^{n+1} = -y^n, \quad n = 0, 1, ..., N-1. \tag{18}$$

Using the decomposition (16)–(18), equation (12) holds automatically for any p^{n+1}.

To evaluate p^{n+1}, we apply the condition (15). The substitution of (16) into (15) yields

$$p^{n+1} = \frac{1}{w^{n+1}(\boldsymbol{x}^*)}(\varphi^{n+1} - y^{n+1}(\boldsymbol{x}^*)). \tag{19}$$

The fundamental point of applicability of this algorithm is associated with the condition $w^{n+1}(\boldsymbol{x}^*) \neq 0$. The auxiliary function $w^{n+1}(\boldsymbol{x})$ is determined from the grid elliptic equation (18). The property of having fixed sign for $w^{n+1}(\boldsymbol{x})$ is followed, in particular, from the same property of the solution at the previous time level $u^n(\boldsymbol{x})$. Such constraints on the solution can be provided by the corresponding restrictions on the input data of the inverse problem.

It is relatively easy to establish that the solution of the inverse problem for $u_0 > 0$, $f^{n+1} > 0$, $n = 0, 1, ..., N-1$ in the class $p^{n+1} < 0$, $n = 0, 1, ..., N-1$ may be represented in the form (16)–(19). Correctness of the computational algorithm follows from fulfilling the condition $w^{n+1}(\boldsymbol{x}^*) \neq 0$.

On the basis of the maximum principle, for the solution of the grid problem (18) with $y^n > 0$, we have $w^{n+1} < 0$. Thus, it is sufficient to show that $y^{n+1} < 0$. The proof is performed by induction. We have $y^0 > 0$ and let $y^n > 0$. From (18), we obtain $w^{n+1} < 0$. Similarly, from (17), we have $v^{n+1} > 0$.

For $y^{n+1} - v^{n+1}$, from (12) and (17), we get

$$\frac{y^{n+1} - v^{n+1}}{\tau} + D(y^{n+1} - v^{n+1}) = -p^{n+1}y^n.$$

Because of this, we have $y^{n+1} - v^{n+1} > 0$ and therefore $y^{n+1} > 0$. It should be noted that $\varphi^{n+1} - v^{n+1}(\boldsymbol{x}^*) > 0$ and for the definition of the coefficient according to (18), we have $p^{n+1} < 0$. Thus, we remain in the class of negative coefficients.

6 Numerical examples

To demonstrate possibilities of the above linearized schemes for solving the coefficient identification problem for the parabolic equation, we consider a model problem. In the examples below, we consider the problem in the unit square ($l_1 = l_2 = 1$). Suppose

$$k(\boldsymbol{x}) = 1, \quad f(\boldsymbol{x}, t) = x_1 x_2, \quad u_0(\boldsymbol{x}) = 1, \quad \boldsymbol{x} \in \Omega.$$

The problem is considered on the grid $N_1 = N_2 = 51$, the observation point is located at the square centre ($\boldsymbol{x}^* = (0.5, 0.5)$). The coefficient $p(t)$ is taken in the form

$$p(t) = \begin{cases} -1000t, & 0 < t \leq 0.5T, \\ 0, & 0.5T < t \leq T. \end{cases}$$

The solution of the direct problem (1)–(3) at the observation point is depicted in Fig.1. The solution at the final time moment is presented in Fig.2.

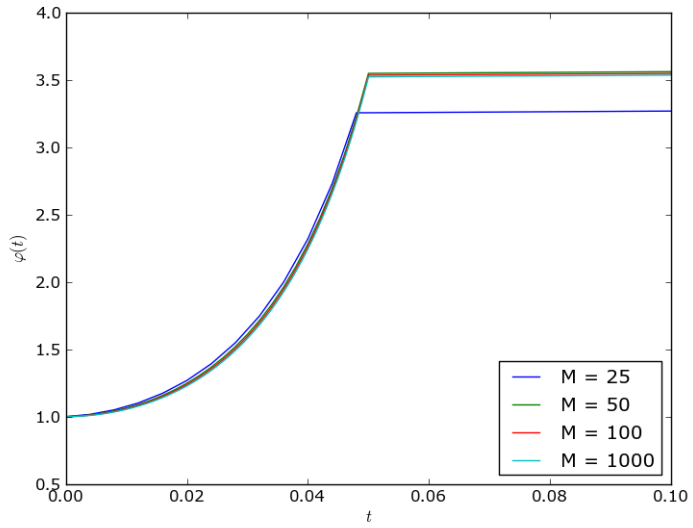


Figure 1: The solution of the direct problem at the point of observation

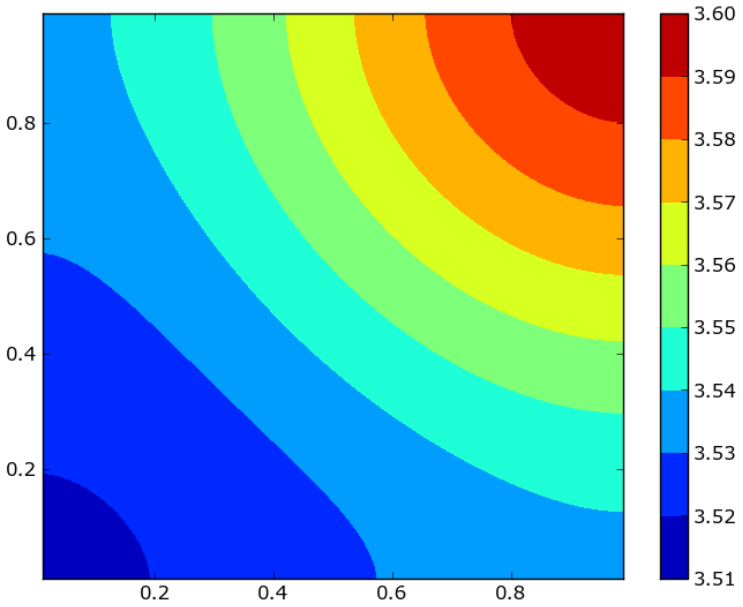


Figure 2: The solution of the direct problem at $t = T$

The results of solving the inverse problem with various grids in time are shown in Fig.3. To study the influence of parameters of the computational algorithm, we need to use the same input data. In our case, as the input data we use the numerical solution of the direct problem obtained using a very fine grid in time. The solution of the direct problem obtained with $N = 1000$ is employed as the input data (the function $\varphi(t)$ in the condition (5)). It is easy to see that the approximate solution of the inverse problem converges with decreasing the time step.

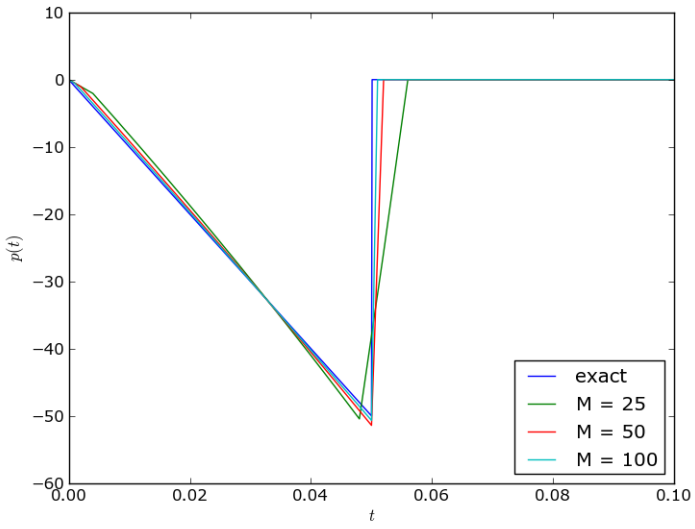


Figure 3: The solution of the inverse problem

Acknowledgements

This research was supported by RFBR (project 14-01-00785).

References

[1] O. M. Alifanov, *Inverse Heat Transfer Problems*, Springer (2011).

[2] R. C. Aster, B. Borchers, C. H. Thurber, *Parameter Estimation and Inverse Problems*, Elsevier Science (2011).

[3] M. M. Lavrent'ev, V. G. Romanov, S. P. Shishatskii, *Ill-posed Problems of Mathematical Physics and Analysis*, American Mathematical Society (1986).

[4] V. Isakov, *Inverse problems for partial differential equations*, Springer (1998).

[5] A. I. Prilepko, D. G. Orlovsky, I. A. Vasin, *Methods for Solving Inverse Problems in Mathematical Physics*, Marcel Dekker, Inc (2000).

[6] V. I. Maksimov, *Dynamical inverse problems of distributed systems*, VSP (2002).

[7] C. R. Vogel, *Computational Methods for Inverse Problems*, Society for Industrial and Applied Mathematics (2002).

[8] A. A. Samarskii, P. N. Vabishchevich, *Numerical Methods for Solving Inverse Problems of Mathematical Physics*, De Gruyter (2007).

[9] A. I. Prilepko, D. G. Orlovskii, *Determination of evolution parameter of an equation, and inverse problems in mathematical physics, part I and II*, Differ. Equat., 21(1, 4), (1985), 119–129, 694–701, in Russian.

[10] A. I. Prilepko, V. V. Solo'ev, *Solvability of the inverse boundary value problem of finding a coefficient of a lower order term in a parabolic equation*, Differ. Equat., 23(1), (1987), 136–143, in Russian.

[11] J. A. MacBain, J. B. Bednar, *Existence and uniqueness properties for the one-dimensional magnetotellurics inversion problem*, Journal of mathematical physics, 27, (1986), 645–649.

[12] J. R. Cannon, Y. Lin, *An inverse problem of finding a parameter in a semi-linear heat equation*, Journal of Mathematical Analysis and Applications, 145(2), (1990), 470–484.

[13] S. Wang, Y. Lin, *A finite-difference solution to an inverse problem for determining a control function in a parabolic partial differential equation*, Inverse problems, 5(4), (1989), 631–640.

[14] J. R. Cannon, Y. Lin, S. Xu, *Numerical procedures for the determination of an unknown coefficient in semi-linear parabolic differential equations*, Inverse Problems, 10(2), (1994), 227–243.

[15] M. Dehghan, *Identification of a time-dependent coefficient in a partial differential equation subject to an extra measurement*, Numerical Methods for Partial Differential Equations, 21(3), (2005), 611–622.

[16] C. Ye, Z. Sun, *On the stability and convergence of a difference scheme for an one-dimensional parabolic inverse problem*, Applied mathematics and computation, 188(1), (2007), 214–225.

[17] W. Wang, B. Han, M. Yamamoto, *Inverse heat problem of determining time-dependent source parameter in reproducing kernel space*, Nonlinear Analysis: Real World Applications, 14(1), (2013), 875–887.

[18] M. Dehghan, *Numerical methods for two-dimensional parabolic inverse problem with energy overspecification*, International journal of computer mathematics, 77(3), (2001), 441–455.

[19] D. S. Daoud, D. Subasi, *A splitting up algorithm for the determination of the control parameter in multi dimensional parabolic problem*, Applied mathematics and computation, 166(3), (2005), 584–595.

[20] R. J. LeVeque, *Finite difference methods for ordinary and partial differential equations. Steady-state and time-dependent problems*, SIAM (2007).

[21] A. A. Samarskii, *The theory of difference schemes*, Marcel Dekker, New York (2001).

[22] A. Quarteroni, A. Valli, *Numerical approximation of partial differential equations*, Springer (2008).

[23] S. C. Brenner, L. R. Scott, *The mathematical theory of finite element methods*, Springer (2008).

[24] A. A. Samarskii, E. S. Nikolaev, *Numerical methods for grid equations. Vol. I, II*, Birkhauser Verlag, Basel (1989).

[25] U. M. Ascher, S. J. Ruuth, B. T. R. Wetton, *Implicit-explicit methods for time-dependent partial differential equations*, SIAM Journal on Numerical Analysis, 32(3), (1995), 797–823.

[26] P. N. Vabishchevich, M. V. Vasileva, *Explicit-implicit schemes for convection-diffusion-reaction problems*, Siberian J. Num. Math., 15(4), (2012), 359–369.

[27] A. A. Samarskii, A. V. Gulin, *Stability of Difference Schemes*, Nauka, Moscow (1973), in Russian.

[28] W. H. Hundsdorfer, J. G. Verwer, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Verlag (2003).

[29] N. M. Afanas'eva, A. G. Churbanov, P. N. Vabishchevich, *Unconditionally monotone schemes for unsteady convection-diffusion problems*, Computational Methods in Applied Mathematics, 13(2), (2013), 185–205.

[30] V. T. Borukhov, P. N. Vabishchevich, *Numerical solution of the inverse problem of reconstructing a distributed right-hand side of a parabolic equation*, Computer physics communications, 126(1), (2000), 32–36.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

On-line monitoring of surfacic mobile heating sources

A. Vergnaud[1], L. Perez[2], L. Autrique[1]

[1] LARIS-ISTIA, University of Angers
62 avenue notre dame du lac, 49000 Angers, France
e-mail: alban.vergnaud@univ-angers.fr ; laurent.autrique@univ-angers.fr

[2] LTN-IUT, University of Nantes
La Chantrerie, Rue Christian Pauc, BP50609, 44306 Nantes cedex 03, France
e-mail: laetitia.perez@univ-nantes.fr

Key words: Parametric identification, sliding time window, partial differential equations

Abstract
In this communication, an adaptation of the conjugate gradient method is presented for an identification of flux from two mobile heating sources on a 2D geometry. This adaptation is based on sliding-time window to achieve quasi-online identification of unknown parameters. Different strategies are presented and discussed (adaptive integration interval, sliding time window).

Nomenclature

General symbols

t	Time, s
x	Space variable, *m*
y	Space variable, *m*
e	Thickness, *m*
h	Heat transfer coefficient, $W.m^{-2}.K^{-1}$
t_f	Final time, *s*
$r_{j=1,2}$	Heat flux radius, *m*
l	Plate width, *m*
L	Plate length, *m*
$\vec{n}$	Unit external outward-pointing vector
N_s	Total number of sensors
N_t	Number of time steps

Greek symbols

θ	Temperature, *K*
λ	Thermal conductivity, $W.m^{-1}.K^{-1}$
ρ	Density, $kg.m^{-3}$
c	Specific heat capacity, $J.kg^{-1}.K^{-1}$
$\phi(t)$	Heat flux density, $W.m^{-2}$
θ_0	Initial temperature, *K*

Indices

0	Initial
j	For heating sources
i	For sensors
I	Centre of the disk

1 Introduction

This study is based on a resolution of an inverse thermal problem in 2D geometry. In this context it is usual to wait the acquisition of all measurements before starting the identification procedure. However, when the inversion is realized for control or diagnosis purposes, an online identification can be an interesting mean to reduce the calculation time.

Thus, in this paper a presentation of a quasi online identification of two heating flux of two mobile heating sources is presented, based on an adaptation of a regularization method (conjugate gradient method named CGM). This inverse heat conduction problem is an ill-posed problem in Hadamard sense [1] and numerous methods [2-6] were developed to obtain a reliable digital resolution.

First, the identification of two time depend parameters is presented considering all measurements obtained and will be used as reference. Then the iterative regularization method will be adapted in order to identify online these two heat flux density. Several strategies will be implemented and compared (adaptive integration interval, sliding time window). The robustness of these methods will be demonstrated for each proposed strategies.

2 Direct problem

Let us consider two mobile heating sources S_1 and S_2 evolving on a thin metallic plate $\Omega \subset \mathbb{R}^2, \Omega = (0 \times L) \times (0 \times l)$, with boundaries $\partial\Omega \in \mathbb{R}$ and thickness e, see Figure 1. Considering that the plate is thin, thermal transfers within the plate can be neglected and a two dimensional mathematical model is validated. Both heat losses (convective exchange) and heating fluxes are formulated in the heat equation (issued from energy balance considering Fourier's law). For each source, the heat density $\phi_j \ (j = 1,2)$ was considered as uniform on a mobile disk $D_j\left(I_j, r_j\right)$ with center $I_j\left(x_j, y_j\right)$ and radius r_j .The total heat flux applied on the plate can be written:

$$\Phi(x,y;t) = \begin{cases} \phi_1(t) & \text{if} \quad (x,y) \in D_1\left(I_1(t), r_1\right) \\ \phi_2(t) & \text{if} \quad (x,y) \in D_2\left(I_2(t), r_2\right) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

To describe this heat flux in continuous and differentiable way, we express it as follows:

$$\Phi(x,y;t) = \sum_{j=1}^{2} \frac{\phi_j(t)}{\pi} \left(-\arctan\left(\mu \left(\sqrt{\left(x - x_j\right)^2 + \left(y - y_j\right)^2} - r_j \right) \right) + \frac{\pi}{2} \right), \tag{2}$$

Where $\mu > 0$ is a parameter to be prescribed. The time interval $\left[0, t_f\right]$ is divided into N segments, namely, $\left[0, t_f\right] = \bigcup_{i=0}^{N} \left[t_i, t_{i+1}\right]$ with $t_i = \tau i$ and a discretization step defined by $\tau = \frac{t_f}{N+1}$. We also define the piecewise continuous linear functions as:

$$s_i(t)=\begin{cases}\dfrac{t}{\tau}-i+1 & \text{if} \quad t\in[t_{i-1},t_i]\\ -\dfrac{t}{\tau}+i+1 & \text{if} \quad t\in[t_i,t_{i+1}]\\ 0 & \text{otherwise}\end{cases}$$

The two heat flux densities are expressed as follows: $\phi_j(t)=\sum_{i=1}^{N}\phi_i^j s_i(t),\ j=1,2..$ Thus, the time dependent and spatial distribution of temperature $\theta(x,y;t)$ within the domain is solution of the following partial derivative equation system (3) with parameters listed in Table 1.

$$\begin{cases}\rho c\dfrac{\partial\theta(x,y;t)}{\partial t}-\lambda\nabla^2\theta(x,y;t)=\dfrac{\Phi(x,y;t)-2h\left(\theta(x,y;t)-\theta_0\right)}{e} & \forall(x,y;t)\in\Omega\times[0,t_f]\\ \theta(x,y;0)=\theta_0 & \forall(x,y)\in\Omega\\ -\lambda\dfrac{\partial\theta(x,y;t)}{\partial\vec{n}}=0 & \forall(x,y;t)\in\partial\Omega\times[0,t_f]\end{cases} \quad (3)$$

The two flux densities $\phi_1(t)$ and $\phi_2(t)$ do not follow the same temporal evolution and are represented in Figure 2. The thermal properties of the system are considered constant during the experiment. Numerical resolutions presented in this communication use the finite element method of Comsol® software interfaced with Matlab©.

$\rho c\left(\text{J.m}^{-3}.K^{-1}\right)$	$h\left(\text{W.m}^{-2}.\text{K}^{-1}\right)$	$t_f\,(s)$	$\lambda\left(\text{W.m}^{-1}.\text{K}^{-1}\right)$	$\theta_0\,(K)$
$2.43.10^6$	10	600	160	291
$r_{j=1,2}\,(m)$	$e(m)$	$L(m)$	$l(m)$	
6.10^{-2}	2.10^{-3}	1	1	

Table 1: Model parameters.

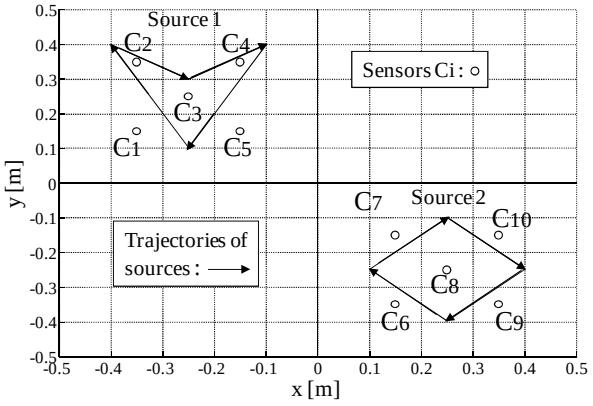


Figure 1: Direct problem representation.

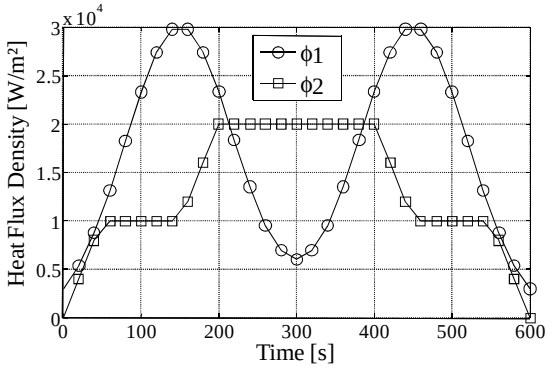


Figure 2: Exact fluxes.

Exact fluxes is given by :

$$\phi_1(t) = 3.10^4\left(\exp\left(-\frac{(t-150)^2}{2\times 70^2}\right)+\exp\left(-\frac{(t-450)^2}{2\times 70^2}\right)\right) \text{ and } \phi_2(t) = \begin{cases} 200\,t & \text{if} \quad t\in[0,50] \\ 1.10^4 & \text{if} \quad t\in[50,150] \\ 200t - 2.10^4 & \text{if} \quad t\in[150,200] \\ 2.10^4 & \text{if} \quad t\in[200,400] \\ -200t + 10e4 & \text{if} \quad t\in[400,450] \\ 1.10^4 & \text{if} \quad t\in[450,550] \\ -200t + 12e4 & \text{if} \quad t\in[550,600] \end{cases}$$

3 Inverse problem

3.1 Formulation

To identify the heat flux densities $\overline{\phi}_1(t)$ and $\overline{\phi}_2(t)$ from temperature evolutions $\left(\hat{\theta}_i(t)\right)_{i=\overline{1,N_c}}$ provided by $N_c = 10$ sensors placed on the plate at locations $(C_i)_{i=\overline{1,N_c}}$, an inverse problem can be formulated and solved by minimizing the following quadratic criterion:

$$J(\overline{\Phi}) = \frac{1}{2}\sum_{i=1}^{N_c}\int_0^{t_f}\left(\theta(C_i,t;\overline{\Phi}) - \hat{\theta}_i(t)\right)^2 dt. \tag{4}$$

The conjugate gradient method is implemented to identify the unknown parameters [2,3]. This algorithm requires iterative resolution of three well-posed problems:

- The direct problem (3) to calculate the cost-function (4) and estimate the quality of the estimate $\overline{\Phi}^k$ at iteration k.
- The adjoint problem to determine the gradient of the cost-function (4) and thus to define the next descent direction [4-7].
- The sensitivity problem to calculate the descent depth (in the descent direction).

3.2 Sensitivity problem

Let us consider $\delta\theta(x,y;t)$ the temperature variation induced by heat flux variation $\phi_j\ (j=1,2)$ which is solution of the following system:

$$\begin{cases} \rho c\dfrac{\partial\delta\theta(x,y;t)}{\partial t} - \lambda\nabla^2\delta\theta(x,y;t) = \dfrac{\delta\Phi(x,y;t) - 2h\delta\theta(x,y;t)}{e} & \forall(x,y;t)\in\Omega\times\left[0,t_f\right] \\ \delta\theta(x,y;0) = 0 & \forall(x,y)\in\Omega \\ -\lambda\dfrac{\partial\delta\theta(x,y;t)}{\partial\vec{n}} = 0 & \forall(x,y;t)\in\partial\Omega\times\left[0,t_f\right] \end{cases} \tag{5}$$

Where

$$\delta\Phi(x,y;t) = \sum_{j=1}^{2}\left(\sum_{i=1}^{N}\frac{\partial\Phi}{\partial\phi_j}\delta\phi_j\right) = \sum_{j=1}^{2}\frac{\delta\phi_j}{\pi}\left(-\arctan\left(\mu\left(\sqrt{(x-x_j)^2+(y-y_j)^2}-r\right)\right)+\frac{\pi}{2}\right). \tag{6}$$

The solution $\delta\theta(x,y;t)$ of the sensitivity problem is used to calculate the descent depth at each iteration k [2]:

$$\gamma^k = \underset{\gamma\in\mathbb{R}^*}{\text{Arg min}}\ J\left(\Phi-\gamma d^k\right) = \frac{\int_0^{t_f}\sum_{i=1}^{N_c}\left(\theta\left(C_i,t;\Phi^k\right)-\hat{\theta}_i(t)\right)\delta\theta\left(C_i,t;\Phi^k\right)dt}{\int_0^{t_f}\sum_{i=1}^{N_c}\left[\delta\theta\left(C_i,t;\Phi^k\right)\right]^2 dt}. \tag{7}$$

3.3 Adjoint problem

In order to determine the gradient of the cost-function $\nabla J^n = \left(\frac{\partial J}{\partial \Phi^i}\right)_{i=(1,\ldots,N)}$ at each iteration of the minimization algorithm, a Lagrangian formulation $\ell(\Phi,\psi)$ is introduced:

$$\ell(\Phi,\psi) = J(\Phi) + \int_0^{t_f}\int_\Omega\left[\rho c\frac{\partial\theta(x,y;t)}{\partial t} - \lambda\Delta\theta(x,y;t) - \frac{\Phi(x,y;t)-2h\left(\theta(x,y;t)-\theta_0\right)}{e}\right]\psi\, dt d\Omega \tag{8}$$

The Lagrangian variation is given by:

$$\begin{aligned}\delta\ell(\Phi,\psi) = &\int_0^{t_f}\sum_{i=1}^{N_c}\left(\theta\left(x_i,y_i;t;\bar{\Phi}\right)-\hat{\theta}_i(x_i,y_i;t)\right)\delta\theta(x_i,y_i;t)\,dt\\ &+\int_0^{t_f}\int_\Omega\left[\rho c\frac{\partial\delta\theta(x,y;t)}{\partial t} - \lambda\nabla^2\delta\theta(x,y;t) - \frac{\delta\Phi-2h\left(\delta\theta(x,y;t)\right)}{e}\right]\psi(x,y;t)\,dt d\Omega\end{aligned}$$

Considering equations of sensitivity problem, $\psi(x,y;t)$ is solution the following problem:

$$\begin{cases}\rho c\dfrac{\partial\psi(x,y;t)}{\partial t} + \lambda\nabla^2\psi(x,y;t) = E(x,y;t) + \dfrac{2h\psi(x,y;t)}{e} & \forall(x,y;t)\in\Omega\times\left[0,t_f\right]\\ \psi(x,y;0)=0 & \forall(x,y)\in\Omega\\ -\lambda\dfrac{\partial\psi(x,y;t)}{\partial\vec{n}}=0 & \forall(x,y;t)\in\partial\Omega\times\left[0,t_f\right]\end{cases} \tag{9}$$

Where $E(x,y;t) = \sum_{i=1}^{N_c}\left(\theta\left(x_i,y_i;t;\bar{\Phi}\right)-\hat{\theta}_i(x_i,y_i;t)\right)\delta_D(x-x_i)\,\delta_D(y-y_i)$ and $\delta_D(x-x_i)\,\delta_D(y-y_i)$ is the Dirac distribution for sensor $C_i(x_i,y_i)$.

When $\psi(x,y;t)$ is solution of the adjoint problem (9) then, $\delta J(\Phi) = -\int_0^{t_f}\int_\Omega\left[\frac{\delta\Phi\,\psi}{e}\right]d\Omega dt..$

The time discretization of the heat flux is written such as:

$$\begin{aligned}\Phi(t) = \sum_{i=1}^{N} s_i(t)\Phi_i = &\sum_{i=1}^{N}\delta\phi_{1_i}\frac{s_i(t)}{\pi}\left(-\arctan\left(\mu\left(\sqrt{(x-x_1)^2+(y-y_1)^2}-r\right)\right)+\frac{\pi}{2}\right)+\\ &\sum_{i=1}^{N}\delta\phi_{2_i}\frac{s_i(t)}{\pi}\left(-\arctan\left(\mu\left(\sqrt{(x-x_2)^2+(y-y_2)^2}-r\right)\right)+\frac{\pi}{2}\right)\end{aligned} \tag{10}$$

Thus, the gradient is expressed by:

$$\frac{\partial J}{\partial \phi_{ji}} = -\int_0^{t_f}\int_\Omega \frac{s_i(t)}{\pi}\left(-\arctan\left(\mu\left(\sqrt{\left(x-x_j\right)^2+\left(y-y_j\right)^2}-r\right)\right)+\frac{\pi}{2}\right)\left[\frac{\psi}{e}\right]d\Omega dt,\ j=1,2,i=1,...,N. \quad (11)$$

4 Numerical results

4.1 Conjugate gradient method (CGM)

Considering for example the initials heat flux $\phi_1(t) = 3020\ \mathrm{W}.m^{-2}$ and $\phi_2(t) = 0$, the CGM is implemented in order to identify the heat flux density of these two mobile sources with noisy numerical data perturbed by $\mathcal{N}(0, 0.5)$ on 10 sensors. The cost-function is presented in Figure 3. The identification has converged in 70 minutes with a stop-criterion $J_{Stop} = 0.5 N_s N_t \sigma^2$, [2, 3]. The two identified densities are represented in Figure 4. To validate these two identified heat fluxes, temperature residual between exact and numerical solutions are presented in Table 2.

	Average temperature residual (K)	*Standard deviation of residual (K)*
For sensors $C_{j=(1,..,5)}$	0.02	0.48
For sensors $C_{j=(6,..,10)}$	0.04	0.47

Table 2: Temperature residual.

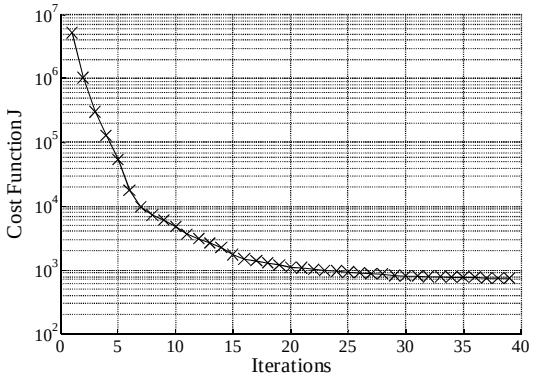


Figure 3: Criterion evolution.

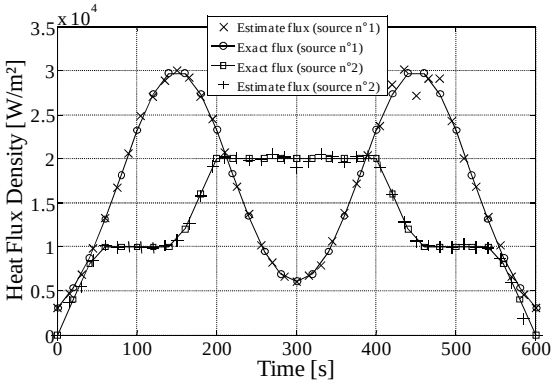


Figure 4: Estimate heat flux (Sources 1 & 2).

4.2 Quasi-online identification

Previous results highlight the efficiency of the CGM. However, the calculation time was 70 minutes in addition to the 10 minutes of the experimentation. The main inconvenient of this method is the convergence time which can be very important according to the complexity problem and the identification time interval. To reduce this computation time, the CGM is implemented considering a time interval $\mathcal{T}_i = \left[\tau_i^-, \tau_i^+\right] \subset \mathcal{T}$ which will slide on the total time horizon with a step $\Delta t_i > 0$ to identify the values of unknown parameters $\phi_{1,\mathcal{T}_i}(t), \phi_{2,\mathcal{T}_i}(t)$. When the values of the parameters are accurately

estimated on the interval $\mathcal{T}_i$, the identification window $\mathcal{T}_{i+1} = \left[\tau_i^- + \Delta t_i^-, \tau_i^+ + \Delta t_i^+\right] = \left[\tau_{i+1}^-, \tau_{i+1}^+\right]$ moves on the horizon time considering initialization $\phi_{j,\mathcal{T}_{i+1}}^{k=0}(t) = \phi_{j,\mathcal{T}}\left(\tau_i^+\right)$ and the initial temperature of the direct problem corresponding to $\theta(x, y; \tau_{i+1}^-)$. Several strategies based on this approach are proposed and analyzed below.

4.2.1 Strategy 1: offset with constant time step

For this first strategy, the time interval of the window $\mathcal{T}_i = \left[\tau_i^-, \tau_i^+\right] \subset \mathcal{T}$ used to identify $\phi_{1,\mathcal{T}_i}(t), \phi_{2,\mathcal{T}_i}(t)$ is fixed at a constant interval $\tau_i^+ - \tau_i^- = 60\,s$ (10% of the overall time of identification). This first strategy is based on a constant offset of $\mathcal{T}_i$ with $\Delta t < \tau_i^+ - \tau_i^-$ to ensure identification ranges overlap. Initially, an offset value $\Delta t = 15\,\mathrm{s}$ (25% of the overall time of identification on $\mathcal{T}_i$) is studied, the results of the identification of the two heating fluxes are presented in Figure 5. The convergence of this identification was obtained after 12.58 minutes. To compare with the CGM resolved on the entire entire time horizon, several time-steps have been tested, see Table 3.

Offset (s)	*Average temperature residual (K)*	*Standard deviation of residual (K)*	*Convergence time (min)*	*Average delay on the identification (s)*
$\Delta t = 15$	0.012	0.546	12.58	115
$\Delta t = 30$	0.027	0.549	9.36	35
$\Delta t = 45$	0.055	0.585	8.66	46

Table 3 : Temperature residual results

Considering results presented in Table 3, this method reveals its effectiveness for a quasi-online identification of the two vectors of unknown parameters. Averages and standard deviations of temperature residual are acceptable given the noise Gaussian considered ($\mathcal{N}(0, 0.5)$) on sensor data. The identification results for this method are presented in Figure 5. Considering that measurements are taken every second, delays between measurements and the results of the identification on the interval are presented in Figure 6.

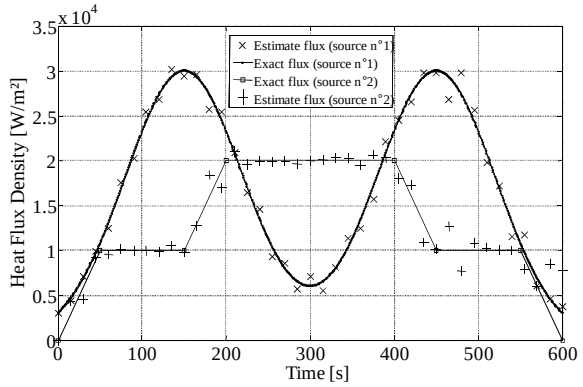


Figure 5: Results sliding mode with (Δt = 15s).

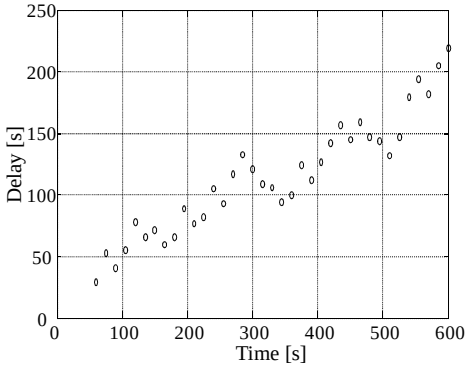


Figure 6: Delay between the identification results and measurement for (Δt = 15s).

This first CGM adaption allows the reduction of the computation time of the identification procedure. Considering results presented in Table 3 this method can be considering as effective. However, as shown in Figure 6 the delay between identifications results and measurement grows up over time. For Δt = 15s, unknown parameters are identified 4 minutes after the experimentation (duration 10 minutes)

4.2.2 Strategy 2: offset with adaptive time step

For this second method, consider the time interval $\mathcal{T}_i = \left[\tau_i^-, \tau_i^+\right] \subset \mathcal{T}$ such as $\tau_i^+ - \tau_i^- = 60\,s$. Identification on this interval is performed during a CPU time equivalent to Δt_i. When the identification is satisfactory then immediately a new interval is considered $\tau_{i+1}^+ = \tau_i^+ + \Delta t_i$ or, if $\tau_{i+1}^+ - \tau_i^+ > 60$ then $\tau_{i+1}^+ = \tau_i^+ + 30$ and $\tau_{i+1}^- = \tau_{i+1}^+ - 60$. So the identification process is launched only if the new measurements are not in adequacy with the temperatures predicted using the previous identification of $\phi_{1\,\mathcal{T}_i}(t), \phi_{2\,\mathcal{T}_i}(t)$. The results of the identification of $\phi_{1\,\mathcal{T}_i}(t), \phi_{2\,\mathcal{T}_i}(t)$ for this method are presented in Figure 7. Considering that measurements are taken every second, delays between measurements and the results of the identification on the intervals are presented in Figure 8. This method presents the advantage of a decrease in the average delay between the identification results and measurements.

Average temperature residual (K)	*Standard deviation of residual (K)*	*Convergence time (min)*	*Average delay on the identification (s)*
0.060	0.548	10.07	28.9

Table 4: Results of temperature residuals.

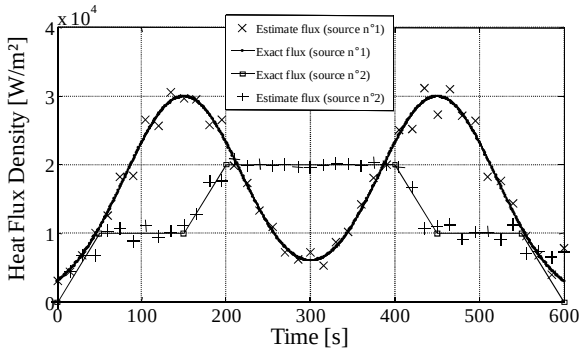


Figure 7: Results adaptive mode.

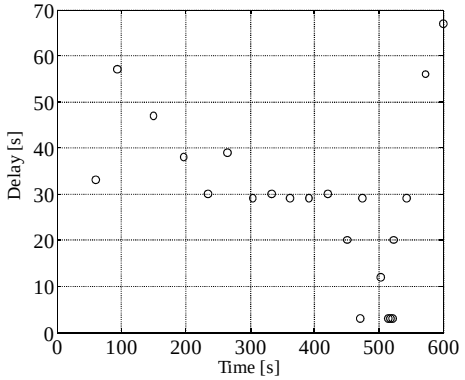


Figure 8: Delay between the identification results and measurement with "adaptive mode".

Considering results presented in Table 4 and in Figure 7, this method can be considering as effective. The main advantage compared to the first strategy is that the delay between the estimated results and measurement can be caught up (see figure 8) and thus reduce the total time of the identification procedure. Unknown parameters are identified 67 seconds after the experimentation (duration 10 minutes)

5 Concluding remarks

In this paper, several approaches for quasi online identification of the heat flux density of heat flow of two mobile sources have been presented. These two strategies are based on the Conjugate Gradient Method recognized for its effectiveness on inverse problems described by partial differential equations. These two strategies show their effectiveness taking into account the time of convergence of the algorithm and the results obtained. To test this method of identifying a concrete case and open new prospects of studies (choice of relevant sensors [8] and adaptive deployment of a network of sensors) experimentation is currently studied.

Acknowledgements

This work was partially supported by the ANR-12-BS03-008-03.

References

[1] J. Hadamard, *Le Problème de Cauchy et les Équations aux Dérivées Partielles Linéaires Hyperboliques*, ED Hermann et C^{ie}, Paris, pp. 560, (1932).

[2] S. Beddiaf, L. Autrique, L. Perez, J.C Jolly, *Time-dependent heat flux identification: Application to a three-dimensional inverse heat conduction problem*, 4th International Conference on Modeling, Identification and Control, Wuhan- China, (2012), pp.1242- 1248.

[3] L. Perez, L. Autrique, M. Gillet, *Implementation of a conjugate gradient algorithm for thermal diffusivity identification in a moving boundaries system*, Journal of Physics, Conference Series, Vol. 135, (2008), doi:10.1088/1742-6596/135/1/012082.

[4] S. Rouquette, L. Autrique, C. Chaussavoine, L.Thomas, *Identification of influence factors in a thermal model of a plasma-assisted chemical vapor deposition process*, Inverse Problems in Science and Engineering, Vol. 15, n° 5, (2007), pp. 489-515.

[5] S. Beddiaf, L. Perez, L. Autrique, J.C. Jolly, *Simultaneous determination of time-varying strength and location of fixed heat sources in 3D*, Inverse Problems in Science and Engineering, Vol. 22, Issue 1, (2013), pp. 166-183.

[6] L. Ling, M. Yamamoto, Y.C. Hon, T. Takeuchi, *Identification of source locations in two-dimensional heat equations*, Inverse Problems, 22, (2006), pp. 1289 – 1305.

[7] C.H. Huang, W.C. Chen, *A three-dimensional inverse forced convection problem in estimating surface heat flux by conjugate gradient method*, International Journal of Heat and Mass Transfer, 43, (2000), pp. 317 – 3181.

[8] L. Perez, *Observation strategies for mobile heating source tracking*, 4th Inverse Problems, Design and Optimization Symposium, Albi, France, (2013), (ISBN 979-10-91526-01-2).

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

An inverse problem for a weakly degenerate parabolic equation in a rectangular domain

Vitaliy Vlasov*

Department of differential equations
Ivan Franko Lviv National University, 1 Universytetska str., 79000 Lviv, Ukraine
e-mail: siphiuel@gmail.com

Key words: parabolic equation, degeneration, Schauder fixed point theorem, two-dimensional, Green's function

Abstract

This paper deals with an inverse problem of identification of a leading time-dependent coefficient in a degenerate two-dimensional parabolic equation. We consider a case of weak degeneration by the time variable.

First treatises of this type of inverse problems for parabolic equations were given in [1], [2]. A variety of identification problems for one-dimensional domains were studied in [3]. Inverse problems for degenerate one-dimensional equations have been considered in [4], [5]. Non-degenerate inverse problems in two-dimensional domains have been considered in e.g. [6], for multidimensional domains see e.g. [7].

In the paper both existence and uniqueness of the solution are proven in corresponding theorems. Existence is established via transformation of the problem to a system of operator equations and subsequent application of Schauder fixed point theorem. Properties of second order Volterra integro-differential equations are used for the proof of solution's uniqueness.

1 Problem statement

In the domain $Q_T = (0,h) \times (0,l) \times (0,T)$ we consider an inverse problem of finding an unknown coefficient $a(t) > 0$, $t \in [0,T]$ in a parabolic equation

$$u_t = t^\beta a(t)\Delta u + b_1(x,y,t)u_x + b_2(x,y,t)u_y + c(x,y,t)u + f(x,y,t) \tag{1}$$

initial condition

$$u(x,y,0) = \varphi(x,y), \quad (x,y) \in \overline{D} := [0,h] \times [0,l], \tag{2}$$

boundary conditions

$$u(0,y,t)=\mu_1(y,t),\quad u(h,y,t)=\mu_2(y,t),\quad (y,t)\in[0,l]\times[0,T], \tag{3}$$

$$u_y(x,0,t)=\nu_1(x,t),\quad u_y(x,l,t)=\nu_2(x,t),\quad (x,t)\in[0,h]\times[0,T] \tag{4}$$

and overdetermination condition

$$a(t)u_x(0,y_0,t)=\varkappa(t),\quad t\in[0,T], \tag{5}$$

where $0<\beta<1$, $0<y_0<l$ are some fixed constants.

2 Existence

Theorem 1 (existence) Suppose that the following conditions hold:

(A1) $\varphi\in C^1(\overline{D});\mu_i\in C^{2,1}([0,l]\times[0,T]),\ \nu_i\in C^{1,0}([0,h]\times[0,T]),\ b_i,c,f\in C^{1,0}(\overline{Q}_T),\ i=1,2,$
$\varkappa\in C[0,T];$

(A2) $\varphi_x(x,y)>0,\quad (x,y)\in\overline{D},\ \varkappa(t)>0,\quad t\in[0,T];$

(A3) $\mu_1(y,0)=\varphi(0,y),\ \mu_2(y,0)=\varphi(h,y),\quad y\in[0,l],$
$\nu_1(x,0)=\varphi_y(x,0),\ \nu_2(x,0)=\varphi_y(x,l),\quad x\in[0,h],$
$\mu_{1y}(0,t)=\nu_1(0,t),\ \mu_{1y}(l,t)=\nu_2(0,t),\ \mu_{2y}(0,t)=\nu_1(h,t),\ \mu_{2y}(l,t)=\nu_2(h,t),\quad t\in[0,T].$

Then for $t\in[0,T_0]$, where $0<T_0\le T$, and T_0 is determined by given data, there exists a solution (a,u) to the problem (1)-(5), that belongs to $C[0,T_0]\times(C^{2,1}(Q_{T_0})\cap C^{1,0}(\overline{Q_{T_0}})$, where $a(t)>0$, $t\in[0,T_0]$.

Proof. Suppose that $a=a(t)$ is known. Let us write an integro-differential equation, that is equivalent to direct problem (1)-(4):

$$u(x,y,t)=u_0(x,y,t)+\int_0^t\int_0^l\int_0^h G_{12}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)v_1(\xi,\eta,\tau)+b_2(\xi,\eta,\tau)v_2(\xi,\eta,\tau)$$
$$+c(\xi,\eta,\tau)u(\xi,\eta,\tau))d\xi d\eta d\tau, \tag{6}$$

where the derivatives $u_x(x,y,t)$ and $u_y(x,y,t)$ are denoted as v_1 and v_2 respectively, and

$$u_0(x,y,t)=\int_0^l\int_0^h G_{12}(x,y,t,\xi,\eta,0)\varphi(\xi,\eta)d\xi d\eta+\int_0^t\int_0^l G_{12\xi}(x,y,t,0,\eta,\tau)\tau^\beta a(\tau)\mu_1(\eta,\tau)d\eta d\tau$$
$$-\int_0^t\int_0^l G_{12\xi}(x,y,t,h,\eta,\tau)\tau^\beta a(\tau)\mu_2(\eta,\tau)d\eta d\tau-\int_0^t\int_0^h G_{12}(x,y,t,\xi,0,\tau)\tau^\beta a(\tau)\nu_1(\xi,\tau)d\xi d\tau$$
$$+\int_0^t\int_0^h G_{12}(x,y,t,\xi,l,\tau)\tau^\beta a(\tau)\nu_2(\xi,\tau)d\xi d\tau+\int_0^t\int_0^l\int_0^h G_{12}(x,y,t,\xi,\eta,\tau)f(\xi,\eta,\tau)d\xi d\eta d\tau.$$

By $G_{ij}(x,y,t,\xi,\eta,\tau)$, $i=1,2$ we denote Green's functions for the heat equation $u_t = t^\beta a(t)\Delta u + f(x,y,t)$ with boundary conditions of the i-th kind by x and j-th kind by y. They are determined by the equality:

$$G_{ij}(x,y,t,\xi,\eta,\tau) = \frac{1}{4\pi(\theta(t)-\theta(\tau))}\sum_{m,n=-\infty}^{+\infty}\Bigg(\Big(\exp\big(-\frac{(x-\xi+2nh)^2}{4(\theta(t)-\theta(\tau))}\big) + (-1)^i\exp\big(-\frac{(x+\xi+2nh)^2}{4(\theta(t)-\theta(\tau))}\big)\Big)\Big(\exp\big(-\frac{(y-\eta+2ml)^2}{4(\theta(t)-\theta(\tau))}\big) + (-1)^j\exp\big(-\frac{(y+\eta+2ml)^2}{4(\theta(t)-\theta(\tau))}\big)\Big)\Bigg), \text{ where } \theta(t)=\int_0^t \tau^\beta a(\tau)d\tau,\ i,j=1,2. \tag{7}$$

Note that the Green's function G_{ij} can be presented as a product of two one-dimensional Green's functions: $G_{ij}(x,y,t,\xi,\eta,\tau) = G_i(x,t,\xi,\tau)G_j(y,t,\eta,\tau)$. Using (6), integration by parts and the Green's function properties

$$G_{1x}(x,t,\xi,\tau) = -G_{2\xi}(x,t,\xi,\tau),\ G_{2\xi\xi}(x,t,\xi,\tau) = -\frac{G_{2\tau}(x,t,\xi,\tau)}{a(\tau)},$$

we calculate v_1 and v_2:

$$v_1(x,y,t) = u_{0x}(x,y,t) + \int_0^t\int_0^l\int_0^h G_{12x}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)v_1(\xi,\eta,\tau) + b_2(\xi,\eta,\tau)v_2(\xi,\eta,\tau) + c(\xi,\eta,\tau)u(\xi,\eta,\tau))d\xi d\eta d\tau, \tag{8}$$

$$v_2(x,y,t) = u_{0y}(x,y,t) + \int_0^t\int_0^l\int_0^h G_{12y}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)v_1(\xi,\eta,\tau) + b_2(\xi,\eta,\tau)v_2(\xi,\eta,\tau) + c(\xi,\eta,\tau)u(\xi,\eta,\tau))d\xi d\eta d\tau, \tag{9}$$

where

$$u_{0x}(x,y,t) = \int_0^l\int_0^h G_{22}(x,y,t,\xi,\eta,0)\varphi_\xi(\xi,\eta)d\xi d\eta + \int_0^t\int_0^l G_{22}(x,y,t,h,\eta,\tau)(\mu_{2_\tau}(\eta,\tau) - \tau^\beta a(\tau)\mu_{2_{\eta\eta}}(\eta,\tau) - f(h,\eta,\tau))d\eta d\tau - \int_0^t\int_0^l G_{22}(x,y,t,0,\eta,\tau)(\mu_{1_\tau}(\eta,\tau) - \tau^\beta a(\tau)\mu_{1_{\eta\eta}}(\eta,\tau) - f(0,\eta,\tau))d\eta d\tau - \int_0^t\int_0^h G_{22}(x,y,t,\xi,0,\tau)\tau^\beta a(\tau)\nu_{1_\xi}(\xi,\tau)d\xi d\tau + \int_0^t\int_0^h G_{22}(x,y,t,\xi,l,\tau)\times\tau^\beta a(\tau)\nu_{2_\xi}(\xi,\tau)d\xi d\tau + \int_0^t\int_0^l\int_0^h G_{22}(x,y,t,\xi,\eta,\tau)f_\xi(\xi,\eta,\tau)d\xi d\eta d\tau, \tag{10}$$

$$u_{0y}(x,y,t)=\int\limits_0^l\int\limits_0^h G_{11}(x,y,t,\xi,\eta,0)\varphi_\eta(\xi,\eta)d\xi d\eta+\int\limits_0^t\int\limits_0^l G_{11\xi}(x,y,t,0,\eta,\tau)\tau^\beta a(\tau)\mu_{1\eta}(\eta,\tau)d\eta d\tau$$
$$-\int\limits_0^t\int\limits_0^l G_{11\xi}(x,y,t,h,\eta,\tau)\tau^\beta a(\tau)\mu_{2\eta}(\eta,\tau)d\eta d\tau+\int\limits_0^t\int\limits_0^h G_{11\eta}(x,y,t,\xi,0,\tau)\tau^\beta a(\tau)\nu_1(\xi,\tau)d\xi d\tau$$
$$-\int\limits_0^t\int\limits_0^h G_{11\eta}(x,y,t,\xi,l,\tau)\tau^\beta a(\tau)\nu_2(\xi,\tau)d\xi d\tau+\int\limits_0^t\int\limits_0^l\int\limits_0^h G_{11}(x,y,t,\xi,\eta,\tau)f_\eta(\xi,\eta,\tau)d\xi d\eta d\tau. \tag{11}$$

We obtained a system (6), (8), (9) equivalent to direct problem (1)-(4). In order to obtain the equation equivalent to overdetermination condition (5), we must first show that $v_1(0,y_0,t)>0$. First, we have to estimate $v_1(0,y_0,t)$ from below. Denote first integral in the equality for $v_1(0,y_0,t)$ as I_1. Note that the following equality holds:

$$\int\limits_0^h G_2(x,t,\xi,\tau)d\xi=1. \tag{12}$$

Using this equality and conditions **(A2)**, we obtain:

$$I_1\geq\min_{\overline{D}}\varphi_x(x,y)\coloneqq M_1>0. \tag{13}$$

It follows from the theorem conditions and the fact that other integrals in $v_1(0,y_0,t)$ converge to zero when $t\to+0$ that one can choose such $T_1:0<T_1\leq T$, that the following inequality holds:

$$v_1(0,y_0,t)\geq\frac{M_1}{2},\quad t\in[0,T_1]. \tag{14}$$

Hence we get from (5)

$$a(t)=\frac{\varkappa(t)}{v_1(0,y_0,t)},\quad t\in[0,T_1]. \tag{15}$$

Thus we have obtained a system of equations (6), (8), (9), (15). It easy to show that the problem (1)-(5) is equivalent to the system of equations (6), (8), (9), (15).

To prove the existence of the solution, we use Schauder fixed point theorem. We will find apriori estimates of the solutions of the system (6), (8), (9), (15).

Estimate $a(t)$ from above. Applying (14) to (15), we obtain:

$$a(t)\leq A_1,\quad t\in[0,T_1], \tag{16}$$

where A_1 is determined by given data.

Now we find the estimate for $a(t)$ from below. For that we need estimates for $u(x,y,t)$, $v_1(x,y,t)$, $v_2(x,y,t)$ from above. Denote $U(t)\coloneqq\max\limits_{(x,y)\in\overline{D}}|u(x,y,t)|$, $V_k(t)\coloneqq\max\limits_{(x,y)\in\overline{D}}|v_k(x,y,t)|$, $k=1,2$. Estimate u_0, u_{0x}, u_{0y} first. Using (12), (16) and known estimate [3]:

$$G_2(x,t,\xi,\tau)\leq C_1+\frac{C_2}{\sqrt{\theta(t)-\theta(\tau)}}, \tag{17}$$

we get

$$|u_0(x,y,t)| \le C_3,\ |u_{0x}(x,y,t)| \le C_4 + C_5 \int\limits_0^t \frac{d\tau}{\sqrt{\theta(t)-\theta(\tau)}},\ |u_{0y}(x,y,t)| \le C_6. \tag{18}$$

Then, using estimates for u_0, u_{0x}, u_{0y}, and equations (6), (8), (9), we obtain:

$$U(t) \le C_7 + C_8 \int\limits_0^t (V_1(\tau) + V_2(\tau) + U(\tau))\, d\tau, \quad t \in [0,T], \tag{19}$$

$$V_1(t) \le C_9 + C_{10} \int\limits_0^t \left(\frac{V_1(\tau) + V_2(\tau) + U(\tau) + 1}{\sqrt{\theta(t)-\theta(\tau)}} \right) d\tau, \quad t \in [0,T], \tag{20}$$

$$V_2(t) \le C_{11} + C_{12} \int\limits_0^t \left(\frac{V_1(\tau) + V_2(\tau) + U(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} \right) d\tau, \quad t \in [0,T]. \tag{21}$$

In order to estimate U let us multiply and divide subintegral expression in (19) by $\sqrt{\theta(t)-\theta(\tau)}$. By using (16), we obtain:

$$U(t) \le C_7 + C_{13} \int\limits_0^t \frac{V_1(\tau) + V_2(\tau) + U(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau, \quad t \in [0,T]. \tag{22}$$

Denote $W(t) \coloneqq U(t) + V_1(t) + V_2(t) + 1$. From (19)-(21) we have:

$$W(t) \le C_{14} + C_{15} \int\limits_0^t \frac{W(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau. \tag{23}$$

From the overdetermination condition (5) it follows that $a(t)W(t) \ge \varkappa(t)$, and thus $1 \le \frac{a(t)W(t)}{\varkappa(t)}$. Let us substitute the obtained inequality into (23):

$$W(t) \le C_{14} + C_{16} \int\limits_0^t \frac{a(\tau)W^2(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau. \tag{24}$$

Squaring both parts in (24) and using Cauchy-Bunyakovsky inequality we get:

$$W^2(t) \le 2C_{14}^2 + 2C_{16}^2 \int\limits_0^t \frac{A_1 W^4(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau \int\limits_0^t \frac{a(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau. \tag{25}$$

Estimate the last integral in this inequality. From the definition of $\theta(t)$ it follows that $\theta(t) \le \frac{t^{\beta+1}}{\beta+1} A_1$, that is, $\left(\frac{\theta(t)(\beta+1)}{A_1}\right)^{\frac{\beta}{\beta+1}} \le t^\beta$. We have:

$$I \coloneqq \int\limits_0^t \frac{a(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau \le C_{17} \int\limits_0^t \frac{a(\tau)\tau^\beta d\tau}{(\theta(\tau))^{\frac{\beta}{\beta+1}} \sqrt{\theta(t)-\theta(\tau)}}.$$

By performing a substitution $z = \dfrac{\theta(\tau)}{\theta(t)}$ the last inequality is transformed to:

$$I \le C_{18}(\theta(t))^{\frac{1-\beta}{2(\beta+1)}} \int\limits_0^1 \frac{dz}{z^{\frac{\beta}{\beta+1}}\sqrt{1-z}}.$$

Taking into account that $\dfrac{\beta}{\beta+1} < \dfrac{1}{2}$ for $0 < \beta < 1$, the last inequality assumes the form

$$I \le C_{18}(\theta(t))^{\frac{1-\beta}{2(\beta+1)}} \int\limits_0^1 \frac{dz}{\sqrt{z(1-z)}} \le C_{19}.$$

Applying this result to inequality (25) yields:

$$W^2(t) \le C_{20} + C_{21} \int\limits_0^t \frac{W^4(\tau)}{\sqrt{\theta(t)-\theta(\tau)}} d\tau.$$

Let us substitute t by σ, multiply both parts by $\dfrac{a(\sigma)}{\sqrt{\theta(t)-\theta(\sigma)}}$ and integrate from 0 to t. We get:

$$\int\limits_0^t \frac{a(\sigma)W^2(\sigma)}{\sqrt{\theta(t)-\theta(\sigma)}} d\sigma \le C_{22} \int\limits_0^t \frac{a(\sigma)}{\sqrt{\theta(t)-\theta(\sigma)}} d\sigma + C_{23} \int\limits_0^t \frac{a(\sigma)}{\sqrt{\theta(t)-\theta(\sigma)}} d\sigma \int\limits_0^\sigma \frac{W^4(\tau)}{\sqrt{\theta(\sigma)-\theta(\tau)}} d\tau.$$

Transform the last summand in this inequality. By changing the order of integration and making a substitution $z = \dfrac{\theta(\sigma)-\theta(\tau)}{\theta(t)-\theta(\tau)}$, we obtain:

$$\int\limits_\tau^t \frac{a(\sigma)\sigma^\beta d\sigma}{\sqrt{(\theta(t)-\theta(\sigma))(\theta(\sigma)-\theta(\tau))}} = \pi.$$

Then

$$\int\limits_0^t \frac{a(\sigma)W^2(\sigma)}{\sqrt{\theta(t)-\theta(\sigma)}} d\sigma \le C_{24} + C_{25} \int\limits_0^t \frac{W^4(\tau)d\tau}{\tau^\beta}.$$

Substituting this estimate into (24) yields:

$$W(t) \le C_{26} + C_{27} \int\limits_0^t \frac{W^4(\tau)d\tau}{\tau^\beta}.$$

Let us denote right part of the inequality by $\widetilde{W}(t)$. Let's find the derivative

$$\widetilde{W}'(t) = C_{27}\frac{W^4(t)}{t^\beta} \le C_{27}\frac{\widetilde{W}^4(t)}{t^\beta}.$$

By solving this inequality, we get

$$W(t) \le \frac{C_{28}(1-\beta)^{\frac{1}{3}}}{(1-\beta-3C_{28}^3C_{27}t^{1-\beta})^{\frac{1}{3}}}, \quad t \in [0,T_2], \tag{26}$$

where $T_2 : 0 < T_2 \le T_1$ satisfies the inequality $1-\beta-3C_{28}^3C_{27}T_2^{1-\beta} > 0$. Hence we get the following estimate: $W(t) \le M_2,\ t \in [0,T_2]$. Therefore we have:

$$|v_1(x,y,t)| \le M_2,\ |v_2(x,y,t)| \le M_2, |u(x,y,t)| \le M_2, \quad t \in [0,T_2]. \tag{27}$$

Using (15) we get an estimate from below for $a(t)$:

$$a(t) \ge A_0 > 0, \quad t \in [0,T_2], \tag{28}$$

where A_0 is determined by given data. System (6), (8), (9), (15) can be presented as:

$$(a,u,v_1,v_2) = \mathbf{P}(a,u,v_1,v_2),$$

where $\mathbf{P}$ is determined by right parts of (6), (8), (9), (15). Operator $\mathbf{P}$ maps

$$\mathcal{N} = \Big\{(a,u,v_1,v_2) \in C[0,T_0] \times \left(C(\overline{Q}_{T_0})\right)^3 :\ A_0 \le a(t) \le A_1,\ |u| \le M_2, \frac{M_1}{2} \le v_1 \le M_2,$$
$$|v_2| \le M_2\Big\}, \quad T_0 = \min\{T_1,T_2\}$$

into itself. This operator is completely continuous [4]. Hence $\mathbf{P}$ has a fixed point, therefore there exists a solution to (1)-(5). □

3 Uniqueness

Theorem 2 (uniqueness) Suppose that the following conditions hold:

(A4) $\varphi \in C^2(\overline{D}),\ \mu_i \in C^{3,1}([0,l] \times (0,T]) \cap C^{2,1}([0,l] \times [0,T]),\ \nu_i \in C^{2,1}([0,h] \times (0,T]) \cap$
$\cap\, C^{1,0}([0,h] \times [0,T]);$ there exist finite limits $\lim\limits_{t \to +0} t^\beta \mu_{iyyy},\ \lim\limits_{t \to +0} t^\beta \nu_{ixx};\ b_i, c, f \in C^{1,0}(\overline{Q}_T);$

(A5) $\varkappa(t) \ne 0, \quad t \in [0,T].$

Then the solution of the problem (1)-(5) is unique in $C[0,T] \times (C^{2,1}(Q_T) \cap C^{1,0}(\overline{Q_T}))$, $a(t) > 0$, $t \in [0,T]$.

Proof. Suppose that (a_1,u_1) and (a_2,u_2) are two solutions to the problem (1)-(5). For their difference $(a,u) = (a_1-a_2, u_1-u_2)$ we have the following problem:

$$u_t = t^\beta a_1(t)\Delta u + t^\beta a(t)\Delta u_2 + b_1(x,y,t)u_x + b_2(x,y,t)u_y + c(x,y,t)u, \quad (x,y,t) \in Q_T, \tag{29}$$
$$u(x,y,0) = 0, \quad (x,y) \in \overline{D}, \tag{30}$$
$$u(0,y,t) = u(h,y,t) = 0, \quad (y,t) \in [0,l] \times [0,T], \tag{31}$$
$$u_y(x,0,t) = u_y(x,l,t) = 0, \quad (x,t) \in [0,h] \times [0,T], \tag{32}$$
$$a_1(t)u_x(0,y_0,t) = -a(t)u_{2x}(0,y_0,t), \quad t \in [0,T]. \tag{33}$$

By using Green's function $\tilde{G}_{12}(x,y,t,\xi,\eta,\tau)$ for the parabolic equation (29), the problem (29)-(33) is reduced to the following integral equation:

$$a(t) = -\frac{a_1(t)}{u_{2x}(0,y_0,t)} \int\limits_0^t \int\limits_0^l \int\limits_0^h \tilde{G}_{12_x}(0,y_0,t,\xi,\eta,\tau)\tau^\beta a(\tau)\Delta u_2(\xi,\eta,\tau)d\xi d\eta d\tau, \tag{34}$$

Let us estimate the kernel of this equation. Denote by $G^{(2)}_{12}(x,y,t,\xi,\eta,\tau)$ the Green's function for the heat equation $u_{2t} = t^\beta a_2(t)\Delta u_2$ with boundary conditions (3), (4). By differentiating (6), (8), (9) twice by x and y, we get:

$$u_{2xx}(x,y,t) = u_{0xx}(x,y,t) + \int\limits_0^t \int\limits_0^l G^{(2)}_{12\xi}(x,y,t,h,\eta,\tau)\Big(b_1(h,\eta,\tau)u_{2\xi}(h,\eta,\tau) + b_2(h,\eta,\tau)\mu_{2\eta}(\eta,\tau)$$

$$+ c(h,\eta,\tau)\mu_2(\eta,\tau)\Big)d\eta d\tau - \int\limits_0^t \int\limits_0^l G^{(2)}_{12\xi}(x,y,t,0,\eta,\tau)\Big(b_1(0,\eta,\tau)u_{2\xi}(0,\eta,\tau) + b_2(0,\eta,\tau)\mu_{1\eta}(\eta,\tau)$$

$$+ c(0,\eta,\tau)\mu_1(\eta,\tau)\Big)d\eta d\tau - \int\limits_0^t \int\limits_0^l \int\limits_0^h G^{(2)}_{12\xi}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)u_{2\xi\xi} + b_2(\xi,\eta,\tau)u_{2\eta\xi}$$

$$+ (b_{1_\xi}(\xi,\eta,\tau) + c(\xi,\eta,\tau))u_{2\xi} + b_{2_\xi}(\xi,\eta,\tau)u_{2\eta} + c_\xi(\xi,\eta,\tau)u_2)d\xi d\eta d\tau, \tag{35}$$

$$u_{2yy}(x,y,t) = u_{0yy}(x,y,t) - \int\limits_0^t \int\limits_0^l \int\limits_0^h G^{(2)}_{12\eta}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)u_{2\xi\eta}$$

$$+ b_2(\xi,\eta,\tau)u_{2\eta\eta} + (b_{2_\eta}(\xi,\eta,\tau) + c(\xi,\eta,\tau))u_{2\eta} + b_{1_\eta}(\xi,\eta,\tau)u_{2\xi} + c_\eta(\xi,\eta,\tau)u_2)d\xi d\eta d\tau, \tag{36}$$

$$u_{2xy}(x,y,t) = u_{0xy}(x,y,t) - \int\limits_0^t \int\limits_0^l \int\limits_0^h G^{(2)}_{21\xi}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)u_{2\xi\eta} + b_2(\xi,\eta,\tau)u_{2\eta\eta}$$

$$+ (b_{2_\eta}(\xi,\eta,\tau) + c(\xi,\eta,\tau))u_{2\eta} + b_{1_\eta}(\xi,\eta,\tau)u_{2\xi} + c_\eta(\xi,\eta,\tau)u_2)d\xi d\eta d\tau, \tag{37}$$

$$u_{2yx}(x,y,t) = u_{0yx}(x,y,t) - \int\limits_0^t \int\limits_0^l \int\limits_0^h G^{(2)}_{21\eta}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)u_{2\xi\xi} + b_2(\xi,\eta,\tau)u_{2\eta\xi}$$

$$+ (b_{1_\xi}(\xi,\eta,\tau) + c(\xi,\eta,\tau))u_{2\xi} + b_{2_\xi}(\xi,\eta,\tau)u_{2\eta} + c_\xi(\xi,\eta,\tau)u_2)d\xi d\eta d\tau. \tag{38}$$

It is clear that we can split ((35), (36), (37), (38)) in two independent systems ((35), (38)) and ((36), (37)). Then for ((35), (38)) we obtain:

$$u_{2xx}(x,y,t) = \hat{u}_{2xx}(x,y,t) - \int\limits_0^t \int\limits_0^l \int\limits_0^h G^{(2)}_{12\xi}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)u_{2\xi\xi} + b_2(\xi,\eta,\tau)u_{2\eta\xi}$$

$$+ (b_{1_\xi}(\xi,\eta,\tau) + c(\xi,\eta,\tau))u_{2\xi} + b_{2_\xi}(\xi,\eta,\tau)u_{2\eta} + c_\xi(\xi,\eta,\tau)u_2)d\xi d\eta d\tau,$$

$$u_{2yx}(x,y,t) = \hat{u}_{2yx}(x,y,t) - \int\limits_0^t \int\limits_0^l \int\limits_0^h G^{(2)}_{21\eta}(x,y,t,\xi,\eta,\tau)(b_1(\xi,\eta,\tau)u_{2\xi\xi} + b_2(\xi,\eta,\tau)u_{2\eta\xi}$$

$$+ (b_{2_\eta}(\xi,\eta,\tau) + c(\xi,\eta,\tau))u_{2\eta} + b_{1_\eta}(\xi,\eta,\tau)u_{2\xi} + c_\eta(\xi,\eta,\tau)u_2)d\xi d\eta d\tau, \tag{39}$$

where $\hat{u}_{2xx}(x,y,t)$ and $\hat{u}_{2yx}(x,y,t)$ denote corresponding free members in equations for $u_{2xx}(x,y,t)$ and $u_{2yx}(x,y,t)$. Their behaviour will be established first. We will need the equations for u_{0xx} and u_{0yx}:

$$u_{0xx}(x,y,t) = \int_0^l \int_0^h G_{12}^{(2)}(x,y,t,\xi,\eta,0)\varphi_{\xi\xi}(\xi,\eta)d\xi d\eta - \int_0^t \int_0^l G_{12\xi}^{(2)}(x,y,t,h,\eta,\tau)(\mu_{2\tau}(\eta,\tau)$$

$$- \tau^\beta a_2(\tau)\mu_{2\eta\eta}(\eta,\tau) - f(h,\eta,\tau))d\eta d\tau + \int_0^t \int_0^l G_{12\xi}^{(2)}(x,y,t,0,\eta,\tau)(\mu_{1\tau}(\eta,\tau)$$

$$- \tau^\beta a_2(\tau)\mu_{1\eta\eta}(\eta,\tau) - f(0,\eta,\tau))d\eta d\tau + \int_0^t \int_0^h G_{12}^{(2)}(x,y,t,\xi,l,\tau)\tau^\beta a_2(\tau)\nu_{2\xi\xi}(\xi,\tau)d\xi d\tau$$

$$- \int_0^t \int_0^h G_{12}^{(2)}(x,y,t,\xi,0,\tau)\tau^\beta a_2(\tau)\nu_{1\xi\xi}(\xi,\tau)d\xi d\tau - \int_0^t \int_0^l \int_0^h G_{12\xi}^{(2)}(x,y,t,\xi,\eta,\tau)f_\xi(\xi,\eta,\tau)d\xi d\eta d\tau,$$

$$u_{0yx}(x,y,t) = \int_0^l \int_0^h G_{21}^{(2)}(x,y,t,\xi,\eta,0)\varphi_{\xi\eta}(\xi,\eta)d\xi d\eta + \int_0^t \int_0^l G_{21}^{(2)}(x,y,t,h,\eta,\tau)(\mu_{2\tau\eta}(\eta,\tau)$$

$$- \tau^\beta a_2(\tau)\mu_{2\eta\eta\eta}(\eta,\tau) - f_\eta(h,\eta,\tau))d\eta d\tau - \int_0^t \int_0^l G_{21}^{(2)}(x,y,t,0,\eta,\tau)(\mu_{1\tau\eta}(\eta,\tau)$$

$$- \tau^\beta a_2(\tau)\mu_{1\eta\eta\eta}(\eta,\tau) - f_\eta(0,\eta,\tau))d\eta d\tau - \int_0^t \int_0^h G_{21\eta}^{(2)}(x,y,t,\xi,l,\tau)\tau^\beta a_2(\tau)\nu_{2\xi}(\xi,\tau)d\xi d\tau$$

$$+ \int_0^t \int_0^h G_{21\eta}^{(2)}(x,y,t,\xi,0,\tau)\tau^\beta a_2(\tau)\nu_{1\xi}(\xi,\tau)d\xi d\tau - \int_0^t \int_0^l \int_0^h G_{21\eta}^{(2)}(x,y,t,\xi,\eta,\tau)f_\xi(\xi,\eta,\tau)d\xi d\eta d\tau.$$

Using known estimates [4]:

$$\int_0^t G_{1\xi}(x,t,0,\tau)d\tau \le \frac{C_{29}}{t^\beta}, \ -\int_0^t G_{1\xi}(x,t,h,\tau)d\tau \le \frac{C_{30}}{t^\beta}, \quad (x,t) \in [0,h]\times[0,T], \tag{40}$$

from (35), (38) and corresponding equations for u_{0xx} and u_{0yx} it is easy to see that:

$$\hat{u}_{2xx}(x,y,t) \le C_{31}t^{-\beta}, \hat{u}_{2yx} \le C_{32}, \quad t \in [0,T]. \tag{41}$$

We now estimate the kernel in first equation in (39), denoting it by I. Using (7), we get:

$$I \le \int_0^t \int_0^h \frac{C_{33} + C_{34}\tau^{-\beta}}{\big(\theta_2(t) - \theta_2(\tau)\big)^{3/2}} \sum_{n=-\infty}^{+\infty} \Bigg(|x - \xi + 2nh| \exp\Big(- \frac{(x-\xi+2nh)^2}{4(\theta_2(t)-\theta_2(\tau))}\Big)$$

$$+ |x + \xi + 2nh| \exp\Big(- \frac{(x+\xi+2nh)^2}{4(\theta_2(t)-\theta_2(\tau))}\Big)\Bigg) d\xi d\tau := I_1 + I_2,$$

where $\theta_2(t) = \int\limits_0^t a_2(\tau)d\tau$. In the integral I_1 substitute $z = \dfrac{x-\xi+2nh}{2\sqrt{\theta_2(t)-\theta_2(\tau)}}$:

$$I_1 \le \sum_{n=-\infty}^{+\infty} \int\limits_0^t \frac{C_{35}+C_{36}\tau^{-\beta}}{\sqrt{\theta_2(t)-\theta_2(\tau)}} \int\limits_{\frac{x-\xi+(2n-1)h}{2\sqrt{\theta_2(t)-\theta_2(\tau)}}}^{\frac{x-\xi+2nh}{2\sqrt{\theta_2(t)-\theta_2(\tau)}}} |z|\exp(-z^2)dz \le \int\limits_0^t \frac{C_{37}\tau^{-\beta}d\tau}{\sqrt{\theta_2(t)-\theta_2(\tau)}}$$

$$\times \int\limits_{-\infty}^{+\infty} |z|\exp(-z^2)dz \le \int\limits_0^t \frac{C_{38}+C_{39}\tau^{-\beta}}{\sqrt{\theta_2(t)-\theta_2(\tau)}}d\tau \le \int\limits_0^t \frac{C_{40}+C_{41}\tau^{-\beta}}{\sqrt{t^{\beta+1}-\tau^{\beta+1}}}d\tau.$$

By substituting $z = \tau/t$ in the last integral, we get the next estimate:

$$I_1 \le t^{\frac{1-\beta}{2}} \left(C_{42} \int\limits_0^1 \frac{dz}{\sqrt{1-z}} + C_{43}t^{-\beta} \int\limits_0^1 \frac{dz}{z^\beta\sqrt{1-z}} \right) \le t^{\frac{1-\beta}{2}} \left(C_{44} + C_{45}t^{-\beta} \right).$$

I_2 is estimated similarly. Then, applying similar procedure to u_{2yx}, from (39) we obtain

$$u_{2xx}(x,y,t) \le C_{46} + C_{47}t^{-\beta},\ u_{2yx}(x,y,t) \le C_{48}, \quad t \in [0,T]. \tag{42}$$

For u_{2xy} and u_{2yy}, using similar calculations, we have:

$$u_{2xy}(x,y,t) \le C_{49},\ u_{2yy}(x,y,t) \le C_{50}, \quad t \in [0,T]. \tag{43}$$

From the given estimates it follows, that the kernel of the equation (34) is integrable, therefore the solution to (29)-(33) is trivial. This condludes the proof of the theorem. □

References

[1] B. Jones, *The determination of a coefficient in a parabolic differential equation. I. Existence and uniqueness*, J. Math. Mech., Vol. 11, No. 5, (1962), pp. 907-918.

[2] J. Cannon, *Determination of an unknown coeffient in a parabolic differential equation*, Duke Math. J., Vol. 30, No. 2, (1963), pp. 313-323.

[3] M. Ivanchov, *Inverse problems for equations of parabolic type*, VNTL Publishers, Lviv, (2003).

[4] N. Saldina, *An inverse problem for a weakly degenerate parabolic equation*, Math. Methods and Phys.-Mech. Fields, Vol. 49, No. 3, (2006), pp.7- 17.

[5] M. Ivanchov, N. Saldina, *An inverse problem for strongly degenerate heat equation*, J. Inv. Ill-Posed Problems, Vol. 14, No. 5, (2006), pp. 465-480.

[6] I. Barans'ka, M. Ivanchov, *Inverse problem with free boundary for the two-dimensional heat equation*, Ukrainian Mathematical Bulletin, Vol. 4, No. 4, (2007), pp. 457-484.

[7] M. Ivanchov, A. Lorenzi, N. Saldina, *Solving a scalar degenerate multidimensional identification problem in Banach space*, J. Inv. Ill-posed Problems, Vol. 16, No. 4, (2008), pp. 397-415.

ICIPE2014 8th International Conference on Inverse Problems in Engineering, May 12–15, 2014, Poland
Eds: I. Szczygieł, A.J. Nowak, M. Rojczyk

Numerical modelling of gas-particle countercurrent flow in the installation for clinker production

J. Wydrych, G. Borsuk, B. Dobrowolski, G. Nowosielski

Department of Thermal Engineering and Industrial Facilities
Opole University of Technology, Opole, Poland
e-mail: j.wydrych@po.opole.pl, g.borsuk@po.opole.pl, b.dobrowolski@po.opole.pl, gn.opole.pl@gmail.com

Key words: two-phase flow, precalcination, particle distribution, CFD

Abstract
One of elements cement kiln installation is a precalciner, which improves calcination process efficiency and product quality. During operation of presented type of precalciner particle distribution between upper and lower outlets is separated in a proportion 30 and 70%, respectively. This paper presents results of numerical calculations of the gas-particle countercurrent flow in precalciner. Gas-particle flow was calculated using Euler method for air motion and Lagrange method for the particles motion. By iterative scaling the particle diameters is possible achieve proper particle distribution between outlets from precalciner chamber.

1 Introduction

Industrial installations for clinker production have a very complex structure and they are difficult to analyses by theoretical methods [1]. One of the most important elements of this installation is precalciner, where is the initial stage of the raw material calcination [2,3]. The main part of the work is modelling the phenomena flow in the chamber with atypical construction. Gas-particle multiphase flow in some kind of precalcination chamber is considered. Fig. 1 presents calculation domain with the information about inlets, outlets and geometrical shape of the chamber. The distribution of particle size in inlets to precalciner chamber has influence on heat and mass transfer between particles and gas. As a result of sieve analysis of raw material and Rosin-Rammler-Sperling calculation method, the authors obtained a particle size distribution as shown in Fig. 3. This distribution was assumed in the early stages of calculations.
Due to Euler-Lagrange method used for the calculation, it was necessary to set the correct fractional distribution of particles. For assumed particles distribution, authors noted that all particles were carried away by upper outlet to preheater tower (Fig. 2). Sieve analysis, shown in Fig. 3, was made for particles after cooling to room temperature. This situation creates a risk that during the sieving particle were crushed and sieve residues do not represent the fractional composition of raw material in real flow. The possibility of particles coagulation increases with the temperature. Because of that, particles

from cyclones IV may have larger diameters than those from cyclones III. The authors proposed a method based on the displacement of the upper range of diameters in the direction of greater values. Because of the lack of possibility to determine the true particles size distribution by measurements, therefore authors used the method of gradually increasing particle diameters. Due to the frequent change of the upper limit, this method can be classified as iterative. The method used for the calculation and numerical results are presented in the following parts of the work.

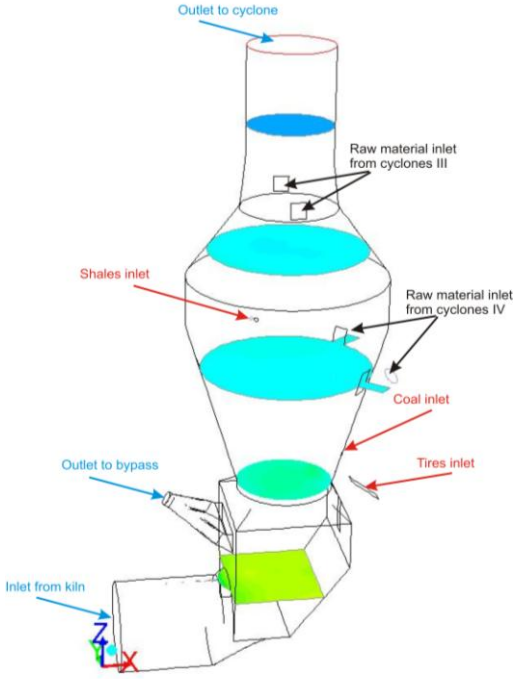


Figure 1. Calculation domain with control cross-sections.

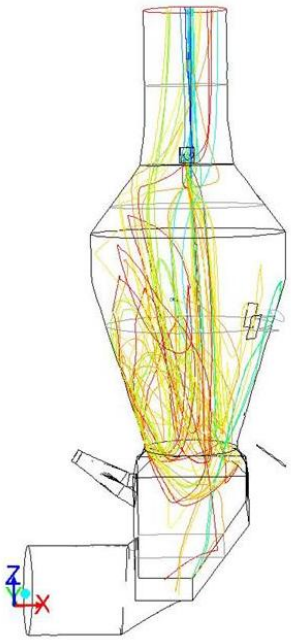


Figure 2. Particles trajectory from upper inlets for all diameters in range15-500µm.

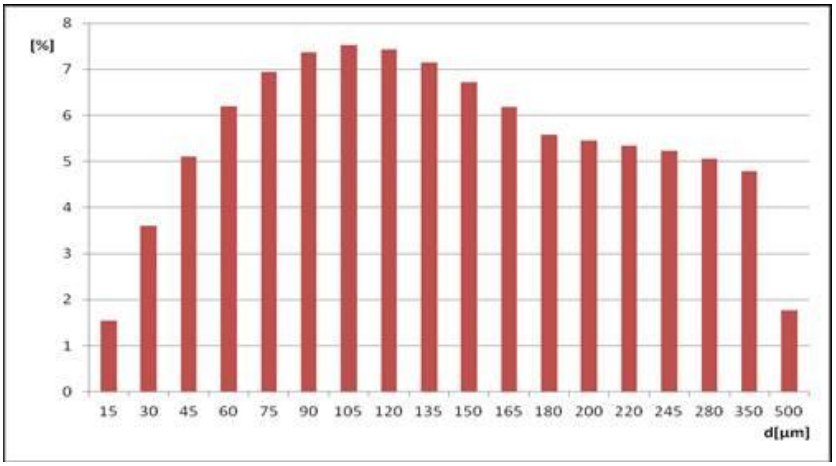


Figure 3. Particle size distributions of raw material.

2 Method of calculations

In order to realize numerical calculations, the mathematical model containing equations of motion for the gaseous phase and particles was formulated. Gas motion was described with the Euler method, and the particle motion – by the Lagrange method. To analyze motion of the gas-particle polydispersive mixture, in this paper the PSI-Cell method was applied [4,5,6].
Numerical calculations are based on the following assumptions: considered flow is stationary, without phase changes, and both phases are incompressible. Gas motion is described in the uniform, generalized conservative form, isolating convection, diffusion and source components. In a consequence we obtain

$$\frac{\partial(\rho\phi)}{\partial t}+\frac{\partial(\rho U_i\phi)}{\partial x_i}=\frac{\partial}{\partial x_i}\left(\Gamma_\phi\frac{\partial\phi}{\partial x_i}\right)+S_\phi+S_{\phi p}, \tag{1}$$

where ϕ is a generalized dependent variable, Γ_ϕ is the coefficient of diffusion transport, and the source term S_ϕ contains all the remaining components of the differential equations [4]. The coefficients Γ_ϕ and S_ϕ are dependent on the variable ϕ. In the PSI-Cell method it is assumed that particles of the disintegrated phase are the sources of mass, momentum and energy occurring as additional components $S_{\phi p}$ in equations of the continuous (gaseous) phase.
The system of equations is accompanied by suitable boundary and initial conditions. The above system of partial differential equations is non-linear. Particular equations are coupled, so they have to be solved with special numerical techniques.
In order to calculate turbulence, model k-ε was used. The standard k-ε model is a semi-empirical model based on model transport equations for the turbulence kinetic energy k and its dissipation rate ε. The model transport equation for k is derived from the exact equation, while the model transport equation for ε was obtained using physical reasoning and bears little resemblance to its mathematically exact counterpart.
In the derivation of the k-ε model, it was assumed that the flow was fully turbulent, and the effects of molecular viscosity were negligible. The standard k-ε model is therefore valid only for fully turbulent flows [7]. The turbulence kinetic energy, k, and its rate of dissipation, ε, are calculated from the following transport equations [8]:

$$\frac{\partial}{\partial t}(\rho k)+\frac{\partial}{\partial x_i}(\rho k u_i)=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right]+G_k+G_b-\rho\varepsilon-Y_M+S_k, \tag{2}$$

$$\frac{\partial}{\partial t}(\rho\varepsilon)+\frac{\partial}{\partial x_i}(\rho\varepsilon u_i)=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial x_j}\right]+C_{1\varepsilon}\frac{\varepsilon}{k}(G_k+C_{3\varepsilon}G_b)-C_{2\varepsilon}\rho\frac{\varepsilon^2}{k}+S_\varepsilon, \tag{3}$$

In these equations, G_k represents the generation of turbulence kinetic energy due to the mean velocity gradients. G_b is the generation of turbulence kinetic energy due to buoyancy. Y_M represents the contribution of the fluctuating dilatation in compressible turbulence to the overall dissipation rate. $C_{1\varepsilon}$, $C_{2\varepsilon}$, and $C_{3\varepsilon}$ are constants. σ_k and σ_ε are the turbulent Prandtl numbers for k and ε, respectively. S_k and S_ε are user-defined source terms. The turbulent (or eddy) viscosity μ_t is computed by combining k and ε as follows:

$$\mu_t=\rho C_\mu\frac{k^2}{\varepsilon}, \tag{4}$$

The model constants $C_{1\varepsilon}$, $C_{2\varepsilon}$, C_μ, σ_k and σ_ε have the following default values $C_{1\varepsilon}=1.44$, $C_{2\varepsilon}=1.92$, $C_\mu=0.09$, $\sigma_k=1.0$ and $\sigma_\varepsilon=1.3$.

The particle trajectory should be known during calculation of the mentioned above source components of differential equations. The particle trajectory is calculated according to its equation of motion [8]. If the phase density difference is big, the equation of particle motion can be written as:

$$\frac{d\vec{u}_p}{dt} = \frac{1}{\tau_p}\left(\vec{u} - \vec{u}_p\right) + g + 3{,}084\left(\frac{\mu}{\rho}\right)^{1/2}\left(\frac{1}{|\vec{\omega}|}\right)^{1/2}\left\{\vec{\omega}\times\left(\vec{u} - \vec{u}_p\right)\right\} f\left(\mathrm{Re}_p, \mathrm{Re}_s\right), \tag{5}$$

Equations presented above were solved with the using ANSYS Fluent Package [9]

3 Results

In order to perform the calculations authors adopted the boundary conditions, which are shown in Tab. 1. Inlet from kiln is the source of gas produced during the combustion of fuel and chemical transformation of raw material in the high temperature. The same cross-section is treated as a outlet for calcined material. Raw material is supplied from two inlets. In the following descriptions raw material inlets from cyclones III and VI are called upper and lower inlets, respectively, regarding outlets inlet from kiln and outlet to cyclones are called lower and upper outlets.

Table 1. Boundary conditions for inlets and outlets.

Boundary	v [m/s]	T [K]	p_n [Pa]
Inlet from kiln	15	1370	-400
Outlet to bypass	-30	1100	-400
Outlet to cyclones	From the balance of the inflow mass	1112	-800
Coal inlet	100	300	0
Shale's inlet	45	300	0
Tires inlet	Closed – treated as a wall		
Raw material inlet from cyclones III	From flow rate of particles	1000	0
Raw material inlet from cyclones IV		1100	0

To reach from the initial ranges of diameters, as a result the sieve analysis to the final range, corresponding to the mass distribution between the upper and lower outlets, authors made about 20 steps in the search for the appropriate range. In the first of the result series, maximum limits of the particle diameters from the upper and lower inlets were set to 1500 and 2000μm, respectively.

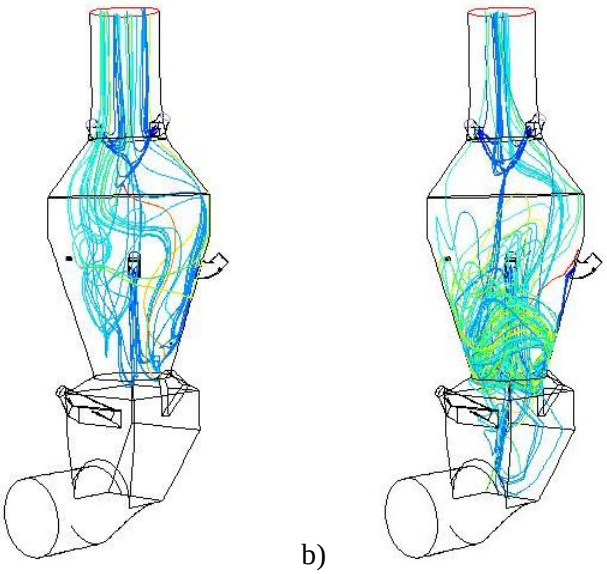
The results that were presented in Table 2 allow to conclude that all particles with diameter less than 200μm from the upper inlet and all particles with diameter less than 105μm from the lower inlet flow to the upper outlet to cyclones. For all particles from the two inlets mass flow rate is separated in a proportion 76.22% to 23.78% for the upper and lower outlets, respectively. Ranges used in this iteration are not those that cause the correct distribution of the mass flow rate between the upper and lower outlets.

Since the upper diameter ranges have been changed to higher values, the numbering does not indicate the particle diameters but rather their names, and for each fraction, the actual particle diameters are set, as in example shown in Tab. 3.

Table 2. Percentage distribution of particles from all, upper and lower inlets to the upper and lower outlets considered size of particles.

	All inlets		Upper inlet		Lower inlet	
d	Upper	Lower	Upper	Lower	Upper	Lower
15	100.00%	0.00%	100.00%	0.00%	0.00%	0.00%
30	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
45	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
60	100.00%	0.00%	100.00%	0.00%	100.00%	0.00%
75	99.40%	0.60%	100.00%	0.00%	98.59%	1.41%
90	99.83%	0.17%	100.00%	0.00%	99.52%	0.48%
105	93.60%	6.40%	100.00%	0.00%	79.79%	20.21%
120	88.26%	11.74%	100.00%	0.00%	65.37%	34.63%
135	82.82%	17.18%	100.00%	0.00%	50.50%	49.50%
150	79.92%	20.08%	100.00%	0.00%	43.18%	56.82%
165	70.79%	29.21%	99.59%	0.41%	26.42%	73.58%
180	67.19%	32.81%	100.00%	0.00%	28.16%	71.84%
200	56.72%	43.28%	97.99%	2.01%	11.45%	88.55%
220	46.10%	53.90%	82.86%	17.14%	1.96%	98.04%
245	39.77%	60.23%	94.51%	5.49%	0.40%	99.60%
280	34.00%	66.00%	91.28%	8.72%	0.00%	100.00%
350	27.15%	72.85%	76.52%	23.48%	0.00%	100.00%
500	2.75%	97.25%	6.30%	93.70%	0.00%	100.00%

For next series of the results, upper limits of the particle diameters from the upper and lower inlets were set to 1500 and 2500μm, respectively. Results for this step were presented in form of particle trajectories.



a) b)

Figure 4. Trajectory particles of size a) 45, b) 105.

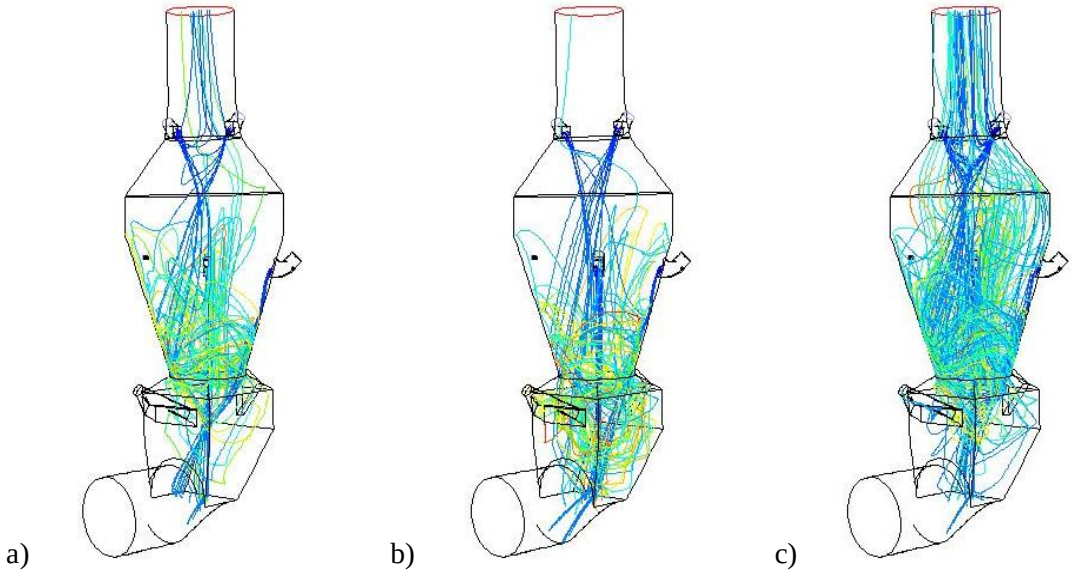


Figure 5. Trajectory particles of size a) 245, b) 500, c) all diameters.

The results presented in this form allow for the identification of areas not reached by particles. Detection of these areas can be helpful in modifying the geometry and efficiency increasing of the flow system. Particles calcination process is dependent on the surrounding gas temperature and the particles residence time in the system. This time is indicated by trajectory colors. Analysis of Fig. 4 and 5 lead to the conclusion that the largest particles end up in the lower outlet, the smallest in the upper outlet, while the rest are distributed between the upper and lower in different proportion.
For all particles from the two inlets mass flow rate is distributed in the ratio 54.02% to 45.98% for the upper and lower outlets, respectively. But also, this ranges are not those that cause the correct distribution of the mass flow rate between the upper and lower outlets.
In last series of the results, upper limits of the particle diameters from the upper and lower inlets were set to 1500μm and 2900μm, respectively.

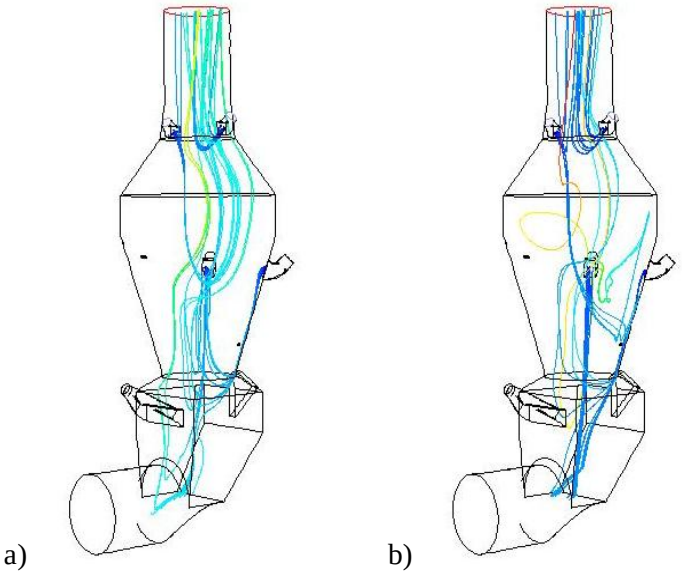
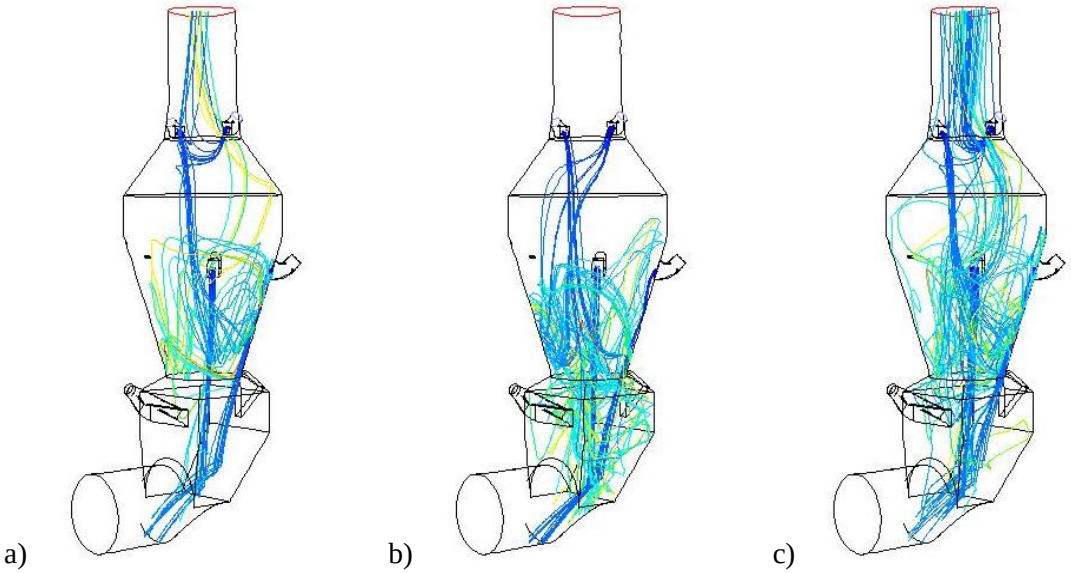


Figure 6. Trajectory particles of size a) 45, b) 105.



a) b) c)

Figure 7. Trajectory particles of size a) 245, b) 500, c) all diameters.

Analysis of trajectory presented in Fig.4÷7 let conclude, that change of diameter range for the lower inlet causes significant changes in the trajectory of the particles from the upper inlet. In the Fig. 6 and 7 can be seen that the particles are concentrated in a smaller areas than it was before. This is due to the fact that almost all of the particles from the lower inlet flow down along the chamber wall in form of "rope" and get into lower outlet, which results in more free motion of the gas phase [10,11]. For all particles from the two inlets mass flow rate is separated in a proportion 27.01% to 72.99% for the upper and lower outlets, respectively. This ranges was used as a correct for final calculations.

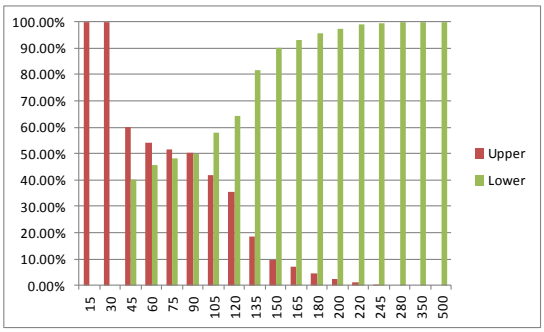


Figure 8. Percentage distribution of particle from all inlets to the upper and lower outlets.

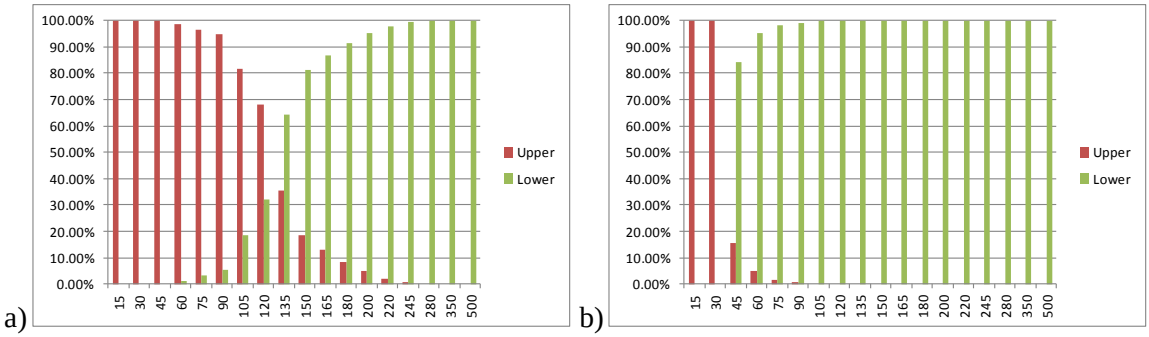


Figure 9. Percentage distribution of particle to the upper and lower outlets from a) upper, b) lower inlets.

Charts presented in the Fig. 8 and 9 show percentage particle distributions to the upper and lower outlets: supplied from all (Fig. 8), upper and lower inlets (Fig. 9). Fig. 9 b shows, that from lower inlet all the particles flow into the lower outlet, except the particle sizes of 15 and 30.

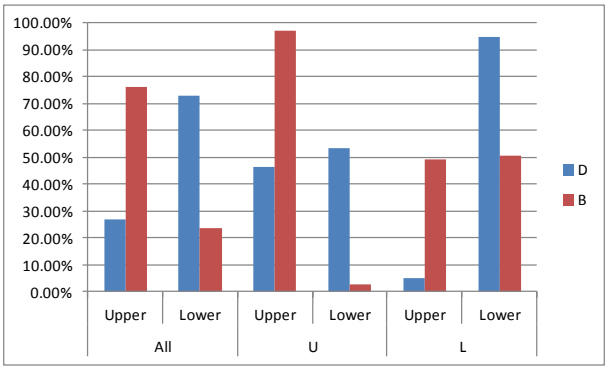


Figure 10. Percentage distribution of particle from all (All), upper (U) and lower (L) inlets to the upper and lower outlets for samples B and D.

Results presented in this work refer to change the upper diameter ranges for examples: A – upper inlet 500 µm, lower inlet 500 µm, B – upper inlet 1500 µm, lower inlet 2000 µm, C – upper inlet 1500 µm, lower inlet 2500 µm and D – upper inlet 1500 µm, lower inlet 2900 µm. In the Fig. 10 was shown percentage distribution of particles from upper, lower and both of inlets to the upper and lower outlets for cases B and D. This comparison confirms influence of particle diameter range flowing from lower inlets on particles flow from upper inlets. Particles flowing from lower inlet along the chamber wall reduces the disturbances in the throat of the chamber between the kiln and the precalciner. This gas flow reduces the amount of particles flowing into the upper outlet from the upper inlets.

4 Conclusions

The number of steps of numerical calculations used at this stage gave the effect of the mass distribution between top and bottom of the chamber similar to the obtained from the experimental observations. In order to achieve the separation of the particles mass flow in correct proportions, authors made series of numerical calculations, for which the fractional distribution of the particles was varied. In the Fig. 11 was shown percentage distribution of particle from all inlets to the upper and

lower outlets for all considered examples. Percentage distribution for case D has reached the required level of 30% to the upper and 70% to the lower outlet.

Table 3. Corrected values of particle diameters of raw material and flow rates for upper and lower inlets

d	d_U[μm]	M_U[kg/s]	d_L[μm]	M_L[kg/s]
15	15	0.428	15	0.385
30	60	1	105	0.9
45	105	1.417	195	1.275
60	155	1.722	285	1.55
75	200	1.928	370	1.735
90	245	2.047	460	1.843
105	290	2.089	550	1.88
120	335	2.064	640	1.858
135	380	1.986	730	1.788
150	430	1.867	820	1.68
165	475	1.717	905	1.545
180	520	1.55	995	1.395
200	580	1.794	1115	1.615
220	645	1.486	1235	1.338
245	720	1.456	1385	1.31
280	825	1.406	1590	1.265
350	1040	1.331	2010	1.198
500	1500	0.492	2900	0.443

By scaling the particle diameters we can observe, that mass flow rate reaching directly to the kiln and returning to the heat exchangers tower were significantly changed. After changing the range of diameters to 15-2900 μm, large diameter particles fall into the rotary kiln, while the remaining fractions are carried away to the outlet from the chamber. In Tab. 3 was presented real particle diameters and mass flow rates for correct distribution, which was used on the final stage of calculations. First column numbering does not mean the particle diameters but rather their names, and for each fraction, the actual particle diameters are set for upper and lower inlets.

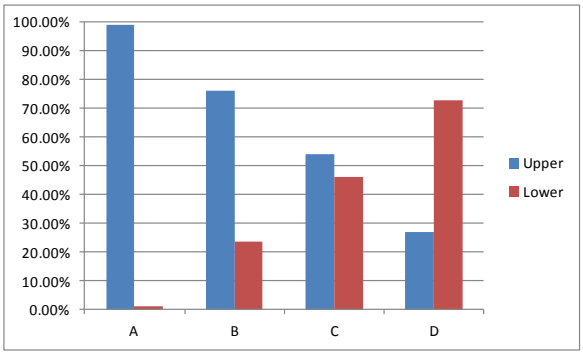


Figure 11. Percentage distribution of particle from all inlets to the upper and lower outlets for all considered samples.

References

[1] R. Saidura, M.S. Hossaina, M.R. Islama, H. Fayazb, H.A. Mohammed, *A review on kiln system modeling*, Renewable and Sustainable Energy Reviews, 15, (2011), pp. 2487-2500.

[2] D.K. Fidaros, C.A. Baxevanou, C.D. Dritselis, N.S. Vlachos, *Numerical modelling of flow and transport processes in a calciner for cement production*, Powder Technology, 171, (2007), pp. 81-95.

[3] Z. Hu, J. Lu, L. Huang, S. Wang, *Numerical simulation study on gas-solid two-phase flow in pre-calciner*, Communications in Nonlinear Science and Numerical Simulation, 11, (2006), pp.440-451.

[4] G. Borsuk, B. Dobrowolski, J. Wydrych, *Gas - solids mixture flow through a two-bend system*, Chemical and Process Engineering, vol. 27, no. 3/1, (2006), pp. 645-656.

[5] J. Wydrych, *Comparative analysis of the methods of simulation of flow in boiler dust systems*, Chemical and Process Engineering, vol. 31, no. 4, (2010), pp. 603-623.

[6] D.K. Fidaros, C.A. Baxevanou, C.D. Dritselis, N.S. Vlachos, *Numerical modelling of flow and transport processes in a calciner for cement production*, Powder Technolog,y 171, (2007), pp. 81–95.

[7] Z. Hu, J. Lu, L. Huang, S. Wang, *Numerical simulation study on gas–solid two-phase flow in pre-calciner*, Communications in Nonlinear Science and Numerical Simulation, 11, (2006), pp. 440–451.

[8] K.S. Mujumdara, K.V. Ganesha, S.B. Kulkarnia, V.V. Ranade, *Rotary Cement Kiln Simulator (RoCKS): Integrated modeling of pre-heater, calciner, kiln and clinker cooler*, Chemical Engineering Science, 62, (2007), pp. 2590–2607.

[9] ANSYS Fluent, Release 14.0, ANSYS Inc.

[10] H. Bilirgen, E.K. Levy, *Mixing and dispersion of particle ropes in lean phase pneumatic conveying*, Powder Technology, 119, (2001), pp. 134–152.

[11] A. Yilmaz, E.K. Levy: *Formation and dispersion of ropes in pneumatic conveying*, Powder Technology 114, (2001), pp. 168–185.

Authors Index

Keywords Index

Institute of Thermal Technology
Silesian University of Technology

Konarskiego 22
44–100 Gliwice, Poland

www.itc.polsl.pl

ISBN 978-83-61506-23-2