Bayesian approach for thermal diffusivity mapping from infrared images with spatially random heat pulse heating

O Fudym\textsuperscript{1,4}, H R B Orlande\textsuperscript{2}, M Bamford\textsuperscript{3}, JC Batsale\textsuperscript{3}

\textsuperscript{1}RAPSODEE UMR 2392 CNRS, Mines Albi, Campus Jarlard, Albi, France
\textsuperscript{2}PEM/COPPE, Federal University of Rio de Janeiro, Brazil
\textsuperscript{3}TREFLE-ENSAM, UMR 8508 CNRS, Talence, France

e-mail: fudym@enstimac.fr

Abstract. This paper aims at the identification of spatially-dependent thermophysical properties and is focused on the mapping of thermal diffusivity. A one-dimensional heat conduction problem is used for the comparison of ordinary least-squares, maximum a posteriori and Markov Chain Monte Carlo methods. Simulated temperature measurements are used in the inverse analysis, which is based on a nodal strategy that results in a linear estimation problem. The nodal strategy relies on the availability of temperature measurements with fine spatial resolution and high frequency, typical of nowadays infrared cameras.

1. Introduction

Thermophysical properties mapping from thermal images provided by an infrared camera is a difficult specific inverse problem, due to both the high flux of data to be processed and parameters to be estimated, as well as the low signal-to-noise ratio. It is thus of interest to implement a linear estimation approach, in order to obtain low computational costs. That can be achieved by minimizing a prediction error model \cite{1}. Usual inversion methods implemented so far to retrieve the parameter fields are generally based on local estimation algorithms that study correlations between pixels. Moreover, in the case of the heat pulse response analysis, the signal-to-noise ratio and consequently the condition number of the so-called Fischer information matrix can be increased by applying a spatially random heat pulse heating \cite{2}. Unfortunately, this kind of approach makes necessary to fill in the sensitivity matrix directly with some linear operations on the measured data. Then the stochastic nature of the regression matrix may yield a bias in the estimates when an ordinary least squares approach is computed. In a previous work, this problem was addressed with the Total Least Squares (TLS) estimation method \cite{3}.

In this paper, we make use of a Bayesian approach \cite{4-12} for the solution of an inverse problem involving the identification of non-homogeneities or inclusions in a one-dimensional heat conduction problem. The physical problem is formulated in terms of the nodal strategy advanced in \cite{2,3}, so that the inverse problem is linear. Both the minimization of the Maximum a Posteriori Objective Function (MAP) and Monte Carlo Markov Chain (MCMC) methods are examined as applied to the inverse problem under picture, as described below.
2. Physical problem and mathematical formulation

The physical problem examined in this paper involves heat conduction in a one-dimensional medium, with spatially varying thermal conductivity and volumetric heat capacity. Boundaries at \( x=0 \) and \( x=L \) are supposed insulated and there is no internal heat generation. The initial temperature within the medium is non-uniform. The mathematical formulation for this problem is given by:

\[
C(x) \frac{\partial T(x,t)}{\partial t} = \frac{\partial}{\partial x} \left[ k(x) \frac{\partial T}{\partial x} \right] \quad \text{in } 0 < x < L, \text{ for } t > 0 \tag{1.a}
\]

\[
\frac{\partial T}{\partial x} = 0 \quad \text{at } x = 0, \text{ for } t > 0 \tag{1.b}
\]

\[
\frac{\partial T}{\partial x} = 0 \quad \text{at } x = L, \text{ for } t > 0 \tag{1.c}
\]

\[
T = F(x) \quad \text{for } t = 0, \text{ in } 0 < x < L \tag{1.b}
\]

In the direct problem associated with the mathematical formulation of such physical problem, the thickness of the medium and the spatial variations of the volumetric heat capacity, thermal conductivity and initial temperature are known. The objective of the direct problem is then to determine the transient temperature variation within the medium.

3. Nodal strategy for the solution of the inverse problem

In this paper we deal with the solution of an inverse problem involving the identification of inclusions or non-homogeneities within the medium. This is accomplished through the identification of the spatially-dependent physical properties, by making use of the nodal strategy advanced in [2,3].

For the solution of the inverse problem, we assume that transient temperature measurements are available at several positions within the medium. The temperature measurements are assumed to be taken with an infrared camera. Such measurement technique is quite powerful because it can provide accurate non-intrusive measurements, with fine spatial resolutions and at large frequencies.

For the application of the nodal strategy [2,3] we rewrite equation (1.a) in the following non-conservative form:

\[
\frac{\partial T(x,t)}{\partial t} = \alpha(x) \frac{\partial^2 T}{\partial x^2} + \delta(x) \frac{\partial T}{\partial x} \tag{2}
\]

where

\[
\alpha(x) = \frac{k(x)}{C(x)} \quad \text{and} \quad \delta(x) = \frac{1}{C(x)} \frac{dk}{dx} \tag{3.a,b}
\]

An explicit discretization of equation (2) using finite-differences result in:

\[
Y_i^{n+1} = L^i Y_i^n + D^i \delta_i \tag{4}
\]

where the subscript \( i \) denotes the finite-difference node at \( x_i=i\Delta x, \ i=1,\ldots,I \), and the superscript \( n \) denotes the time \( t_n=n\Delta t, \ n=1,\ldots,N \). The other quantities appearing in equation (4) are given by:

\[
Y_i^{n+1} = T_i^{n+1} - T_i^n \tag{5.a}
\]

\[
L_i^n = \frac{\Delta t}{(\Delta x)^2} (T_{i+1}^n - 2T_i^n + T_{i-1}^n) \tag{5.b}
\]

\[
D_i^n = \frac{\Delta t}{2\Delta x} (T_{i+1}^n - T_{i-1}^n) \tag{5.c}
\]

By writing equation (4) for all finite-difference nodes and all time-steps, we obtain:
\[ Y = JP \]  

Equation (6) shows a linear dependence of the system response \( Y \) with respect to the vector of parameters \( P \), based on the knowledge of the sensitivity matrix \( J \). With the spatial resolution and frequency of measurements made available by infrared cameras, the sensitivity matrix can be approximately computed with the measurements [2,3]. In this paper, we utilize a Bayesian approach [4-12] for the solution of the parameter estimation problem (6).

4. Bayesian approach for the solution of inverse problems

In the Bayesian approach to statistics, an attempt is made to utilize all available information in order to reduce the amount of uncertainty present in an inferential or decision-making problem. As new information is obtained, it is combined with any previous information to form the basis for statistical procedures. The formal mechanism used to combine the new information with the previously available information is known as Bayes’ theorem [4,5,9-11]. Therefore, the term Bayesian is often used to describe the so-called statistical inversion approach, which is based on the following principles [4]:

1. All variables included in the model are modeled as random variables.
2. The randomness describes the degree of information concerning their realizations.
3. The degree of information concerning these values is coded in probability distributions.
4. The solution of the inverse problem is the posterior probability distribution.

Bayes’ theorem can then be stated as [4,5,9-11]:

\[
\pi_{\text{posterior}}(P) = \pi(P|Y) = \frac{\pi_{\text{prior}}(P)\pi(Y|P)}{\pi(Y)}
\]  

where \( \pi_{\text{posterior}}(P) \) is the posterior probability density, that is, the conditional probability of the parameters \( P \) given the measurements \( Y \); \( \pi_{\text{prior}}(P) \) is the prior density, that is, the coded information about the parameters prior to the measurements; \( \pi(Y|P) \) is the likelihood function, which expresses the likelihood of different measurement outcomes \( Y \) with \( P \) given; and \( \pi(Y) \) is the marginal probability density of the measurements, which plays the role of a normalizing constant.

If we assume the parameters and the measurement errors to be independent Gaussian random variables, with known means and covariance matrices, and that the measurement errors are additive, a closed form expression can be derived for the posterior probability density. In this case, the likelihood function can be expressed as [4,5,9-11]:

\[
\pi(Y|P) = (2\pi)^{-M/2}[W^{-1}]^{1/2}\exp\left[-\frac{1}{2}(Y - Y_e)^T W (Y - Y_e)\right]
\]  

where \( Y_e \) is the vector of estimated variables, obtained from the solution of the forward model with an estimate for the parameters \( P \), \( M = IN \) is the number of measurements and \( W \) is the inverse of the covariance matrix of the errors in \( Y \).
Similarly, for the case involving a prior normal distribution for the parameters we can write:

\[
\pi(P) = \frac{1}{2\pi^{\frac{n}{2}}|V|^{\frac{1}{2}}} \exp \left[ -\frac{1}{2}(P - \mu)^T V^{-1}(P - \mu) \right]
\]  

(11)

where \(\mu\) and \(V\) are the known mean and covariance matrix for \(P\), respectively. By substituting equations (10) and (11) into Bayes' theorem, except for the normalizing constant in the denominator we obtain:

\[
\ln\pi(P | Y) \propto -\frac{1}{2} \left[ (M + N) \ln 2\pi + \ln |W^{-1}| + \ln |V| + S_{MAP}(P) \right]
\]  

(12)

where

\[
S_{MAP}(P) = [Y - T(P)]^T W [Y - T(P)] + (\mu - P)^T V^{-1}(\mu - P)
\]  

(13)

Equation (12) reveals that the maximization of the posterior distribution function can obtained with the minimization of the objective function given by equation (13), denoted as the \textit{maximum a posteriori objective function} [4,5,9-11]. Equation (13) clearly shows the contributions of the likelihood and of the prior distributions in the objective function, given by the first and second terms on the right-hand side, respectively.

For linear estimation problems such as the one under picture, the minimization of the maximum a posteriori objective function is obtained with (14):

\[
P = [J^T W J + V^{-1}]^{-1} [J^T W Y + V^{-1} \mu]
\]  

(14)

It is clearly apparent in equation (14) how the prior information is used as a regularization term. On the other hand, if different prior probability densities are assumed for the parameters, the Posterior Probability Distribution may not allow an analytical treatment. In this case, Markov Chain Monte Carlo (MCMC) methods are used to draw samples of all possible parameters, so that inference on the posterior probability becomes inference on the samples [4,5,9].

In order to implement the Markov Chain, a density \(q(P^*,P^{(t-1)})\) is required, which gives the probability of moving from the current state in the chain \(P^{(t-1)}\) to a new state \(P^*\).

The Metropolis-Hastings algorithm [4,5,9] was used in this work to implement the MCMC method. It can be summarized in the following steps:

1. Sample a Candidate Point \(P^*\) from a jumping distribution \(q(P^*,P^{(t-1)})\).
2. Calculate:

\[
\alpha = \min \left[ 1, \frac{\pi(P^* | Y) q(P^{(t-1)},P^*)}{\pi(P^{(t-1)} | Y) q(P^*,P^{(t-1)})} \right]
\]  

(15)

3. Generate a random value \(U\) which is uniformly distributed on (0,1).
4. If \(U \leq \alpha\), define \(P^{(t)} = P^*\); otherwise, define \(P^{(t)} = P^{(t-1)}\).
5. Return to step 1 in order to generate the sequence \(\{P^{(1)}, P^{(2)}, \ldots, P^{(n)}\}\).

In this way, we get a sequence that represents the posterior distribution and inference on this distribution is obtained from inference on the samples \(\{P^{(1)}, P^{(2)}, \ldots, P^{(n)}\}\). We note that values of \(P^{(i)}\) must be ignored until the chain has not converged to equilibrium. For more details on theoretical aspects of the Metropolis-Hastings algorithm and MCMC methods, the reader should consult references [4,5,9].

In the nodal strategy described in section 3, the sensitivity matrix is approximately computed with the measurements. Therefore, for the implementation of the Metropolis-Hastings algorithm the uncertainties in the computation of the sensitivity matrix need to be taken into account. By assuming that \(P\) and \(J\) are independent random variables, the sought \textit{posterior probability density} is then given by [4]:

\[
\pi (P, J | Y) \propto \pi (Y | P, J) \pi (P) \pi (J)
\]  

(16)

where \(\pi (J)\) is the \textit{a priori} distribution for the sensitivity matrix \(J\).
5. Results and Discussions

The results presented below are obtained from simulated measurements. The direct model is solved with a specific finite volume approach [12], where a simulated noise is added to the simulated temperatures. The errors are additive, uncorrelated, normally distributed, with zero mean and known standard constant-deviation \( \sigma \). Note that the measurements \( Y \) used in the linear model (6) are actually composed of the nodal temperature differences at subsequent times (see equation 5.a). Hence, the likelihood function \( \pi(Y|P,J) \) is given by equation (14) with covariance matrix \( W^{-1} = 2\sigma^2 I \). With such hypotheses regarding the temperature measurements, the prior distribution for the elements of the sensitivity matrix is also normal, with mean values given by equations (7.b,c) and standard-deviations given by:

\[
\sigma_L = \frac{\Delta t}{(\Delta x)^2} \sqrt{6\sigma} \quad \text{and} \quad \sigma_D = \frac{\Delta t}{2\Delta x} \sqrt{2\sigma}
\]

(17.a,b)

for \( L^i \) and \( D^i \), respectively, \( i=1,\ldots,I \), and \( n=1,\ldots,(N-1) \) (see equations 5.b,c, 7.b and 8.b).

For the results obtained with the maximum a posteriori objective function, the prior information for the elements of the vector of parameters \( P \) was taken in the form of normal distributions. The mean values for the thermal diffusivities \( \alpha_i \), \( i=1,\ldots,I \), were given by \( \alpha_i \), that is, the value of thermal diffusivity at the first finite-difference node, which was assumed as known. The standard-deviations for \( \alpha_i \) were taken as \( 0.5 \times 10^{-4} \) m\(^2\)/s, for \( i=1,\ldots,I \). With respect to the parameters \( \delta_i \) the mean values were taken as zero and the standard-deviations as \( 5 \times 10^{-6} \) m/s, for \( i=1,\ldots,I \). In fact, for cases involving sharp interfaces between different materials, such as those examined below, \( \delta_i \) is zero except at the interfaces.

On the other hand, a smoothness prior was used for the estimation of the thermal diffusivities \( \alpha_i \), \( i=1,\ldots,I \), for the cases where the MCMC method was used for the solution of the inverse problem. Such parameters were assumed to be within the interval \([10^{-7}, 10^{-4}] \) m\(^2\)/s. Note that such interval encompasses most substances and does not provide any practical restriction to the values to be estimated for the parameters. The smoothness prior was taken in the form:

\[
\pi(P) \propto \exp \left( -\frac{1}{2\gamma} ||BP||_L^2 \right)
\]

(18)

where \( B \) is a \((I-1)xI\) first-order difference matrix and \( || . ||_2 \) is the vector \( L_2 \) norm. For the cases examined below \( \gamma \) was of the order of \( 10^7 \). The Markov chain consisted of 50000 states, where the first 10000 were neglected for the computation of the statistics for the parameters. Generally, the acceptance ratio in the Metropolis-Hastings algorithm was of the order of 28%.

The cases examined below involved a slab with thickness \( L=0.4 \) m. The slab was basically composed of two materials with thermal properties \( k_A = 1\) Wm\(^{-1}\)K\(^{-1}\), \( C_A = 1x10^6\) Jm\(^{-1}\)K\(^{-1}\), \( k_B = 0.03\) Wm\(^{-1}\)K\(^{-1}\) and \( C_B = 1000\) Jm\(^{-1}\)K\(^{-1}\), which approximates a poor conductor solid material and a gas, respectively. Materials \( A \) and \( B \) were interposed in different forms in order to examine test-cases involving different functional forms for \( \alpha(x) \). In order to avoid the inverse crime of using the same direct solution for the generation of the simulated measurements and for the solution of the inverse problem, a finite-volume solution was developed in order to obtain the simulated measured data. For the solution of the inverse problem, the slab was discretized with \( I=80 \) internal nodes where the values of \( \alpha_i \) and \( \delta_i \) were estimated. The final time was taken as 50 s and the slab was initially randomly excited with a temperature between 0 and 10 °C [2].

Figures 1.a,b present a comparison of the estimations obtained with ordinary-least squares (OLS), maximum a posteriori (MAP) and the Metropolis-Hastings (MH) algorithm, for \( \alpha(x) \) and \( \delta(x) \), respectively. A step decreasing function for the thermal diffusivity was considered in this case, which involved simulated measurements with standard-deviation \( \sigma=0.1 \) °C, which corresponds to 1% of the maximum temperature in the medium. Figure 1.a shows that the thermal diffusivity variation estimated with OLS and MAP became very unstable. On the other hand, the discontinuous thermal
diffusivity could be accurately predicted with the Metropolis-Hastings algorithm. Figure 1.b shows that the estimations obtained for $\delta(x)$ with OLS were unstable. On the other hand, the values estimated with MAP and MH for such parameters were practically null, as a result of the prior information considered available. It is interesting to note that predicting a null value for $\delta(x)$ did not affect significantly the estimation of the thermal diffusivity when the Metropolis-Hastings algorithm was used. Similar conclusions can be taken for an increasing step function for the thermal diffusivity, as illustrated in figures 2.a,b.

The results also showed the main problem arising with the MAP approach when a higher regularization term is used, yielding an important bias in the estimates towards the prior information (these results are not included herein).

In fact, figures 1 and 2 show that, among the techniques examined, only the Markov Chain Monte Carlo method implemented with the Metropolis-Hastings algorithm was capable of accurately predicting the enclosed material, via the estimation of the spatially varying thermal diffusivity. Such a technique was applied for more strict cases, like the one presented in figure 3. Such a case involved nine discontinuities in the thermal diffusivity. Notice in figure 3 that the thermal diffusivity estimated with the Metropolis-Hastings algorithm followed very closely the exact one used to generate the simulated measurements.

---

Figure 1.a. Estimated thermal diffusivity – step decreasing function

Figure 1.b. Estimated $\delta(x)$

Figure 2.a. Estimated thermal diffusivity – step increasing function

Figure 2.b. Estimated $\delta(x)$

---

In fact, figures 1 and 2 show that, among the techniques examined, only the Markov Chain Monte Carlo method implemented with the Metropolis-Hastings algorithm was capable of accurately predicting the enclosed material, via the estimation of the spatially varying thermal diffusivity. Such a technique was applied for more strict cases, like the one presented in figure 3. Such a case involved nine discontinuities in the thermal diffusivity. Notice in figure 3 that the thermal diffusivity estimated with the Metropolis-Hastings algorithm followed very closely the exact one used to generate the simulated measurements.
The effect of the random measurement error is examined in figure 4, which presents results for the same functional form examined in figure 3, but for simulated data with standard-deviation $\sigma=0.2$ °C. This figure shows that the Metropolis-Hastings algorithm provides stable estimates for the present problem even for simulated measurements containing large errors.

6. Conclusions
In this paper we presented the solution of an inverse problem dealing with the identification of non-homogeneities or inclusions in a medium, through the estimation of spatially-dependent thermophysical properties. A one-dimensional heat conduction problem was used for the comparison of different techniques for the solution of the inverse problem. The heat conduction equation was written in a non-conservative form in order to apply a nodal strategy that results in a linear estimation problem. In such form of the heat conduction equation the spatial-dependent thermal diffusivity...
appears explicitly, in addition to another quantity that involves the gradient of the thermal conductivity and the volumetric heat capacity. The one-dimensional approach is realistic when the flash experiment is implemented with a convenient periodic or random grid used as a spatial mask.

The linear inverse problem was solved with the Bayesian approaches of the maximum a posteriori estimator and of the Markov Chain Monte Carlo method based on the Metropolis-Hastings algorithm. The results obtained with these approaches were compared to those obtained with the ordinary least-squares method. Among the techniques examined, only the Markov Chain Monte Carlo method was capable of accurately predicting the spatially varying thermal diffusivity. Such was the case even for functional forms with several discontinuities, as well as with measurements containing random errors with relatively large standard-deviations.

Future works include the implementation of this approach for 2D images sequences and the comparison of results with the Total Least Squares methods.

7. Acknowledgements
This work is co-sponsored by CNPq (Brazil) and CNRS (France). HRBO would like to acknowledge the hospitality of Ecole des Mines d’Albi during his visit in July 2007 for the preparation of this work. A course on the Statistical Inversion Approach given in COPPE/UFRJ by Prof. Jari Kaipio and Prof. Ville Kolehmainen, from the University of Kuopio – Finland, is greatly appreciated.

References
[1] L. Ljung, System identification: theory for the user, Prentice Hall, 1987.
[2] J.C. Batsale J.C., J.L. Battaglia, O. Fudym, Autoregressive algorithms and spatially random flash excitation for 3D non destructive evaluation with infrared cameras. QIRT Journal 1 5-20, 2004.
[3] M. Bamford, J.C. Batsale, D. Reungoat, and O. Fudym, Two dimensional velocity and diffusion mapping in the case of three dimensional transient diffusion: “Flash” method and infrared image sequence analysis, 8th Conf. on Quantitative Infrared Thermography, QIRT 2006, Padova, Italy, June 27-30, 2006.
[4] Kaipio, J. and Somersalo, E., Statistical and Computational Inverse Problems, Applied Mathematical Sciences 160, Springer-Verlag, 2004.
[5] Tan, S., Fox, C., and Nicholls, G., Inverse Problems, Course Notes for Physics 707, University of Auckland, 2006.
[6] N. Zabaras, Inverse problems in heat transfer, chapter 17, in the Handbook of Numerical Heat Transfer, 2nd Edt., John Wiley & Sons, (W.J. Minkowycz, E.M. Sparrow, J. Y. Murthy, editors.), 2004.
[7] J. Wang and N. Zabaras, A Bayesian Inference Approach to The Stochastic Inverse Heat Conduction Problem, International Journal of Heat and Mass Transfer, 47: 3927–3941, 2004.
[8] J. Wang and N. Zabaras, Using Bayesian statistics in the estimation of heat source in radiation, International Journal of Heat and Mass Transfer, 48: 15–29, 2005.
[9] P. M. Lee, Bayesian Statistics, Oxford University Press, London, 2004.
[10] Winkler, R., An Introduction to Bayesian Inference and Decision, Probabilistic Publishing, Gainsville, 2003.
[11] Beck, J. V. and Arnold, K. J., Parameter Estimation in Engineering and Science, Wiley Interscience, New York, 1977.
[12] Fudym O., B. Ladevie, J.C. Batsale, A semi-numerical approach for heat conduction in heterogeneous media. One extension of the analytical Quadrupole method. Numerical Heat Transfer, Part B : Fundamentals, 42 (4) 325-348, 2002.
[13] H.R.B. Orlande, V. Kolehmainen, J.P. Kaipio, Reconstruction of thermal conductivity and heat capacity using a tomographic approach, Int. Journal Heat Mass Transfer, 50, 5150-5160, 2007.