Available online at www.sciencedirect.com
ScienceDirect
International Journal of Heat and Mass Transfer 50 (2007) 5150–5160
International Journal of HEAT and MASS TRANSFER
www.elsevier.com/locate/ijhmt

Reconstruction of thermal conductivity and heat capacity using a tomographic approach

V. Kolehmainen [b,*], J.P. Kaipio [b], H.R.B. Orlande [a]

[a] Department of Mechanical Engineering, Universidade Federal de Rio de Janeiro, Rio de Janeiro, Brazil
[b] Department of Physics, University of Kuopio, P.O. Box 1627, FIN-70211 Kuopio, Finland

Received 2 March 2007
Available online 5 September 2007

Abstract

We consider the estimation of the volumetric heat capacity and the thermal conductivity as distributed parameters. The measurement scheme consists of sequentially heating the boundary of the object in different source locations and measuring the induced temperature evolutions in different measurement locations on the boundary. The estimation of the distributions of volumetric heat capacity and thermal conductivity based on these boundary data is an ill-posed inverse boundary value problem. We propose an approach which is based on transient data on the boundary and the modelling of the unknown coefficients as Markov random fields. The intended applications are non-destructive retrieval of defects as well as the estimation of macroscopic characteristics of novel materials. We evaluate the proposed approach by a numerical simulation.

Keywords: Inverse problem; Thermal tomography; Thermal conductivity; Volumetric heat capacity

1. Introduction

In different physical sciences the recovery of the internal properties of an object from non-intrusive boundary measurements is of extreme importance. Several examples very common in our daily life include imaging modalities such as ultrasound imaging and X-ray tomography, as well as the more recently studied diffuse tomographic techniques that make use of optical or electrical measurements [11,26,3]. From the mathematical point of view, the related parameter estimation problems are often unstable and belong to the class of ill-posed inverse problems [34,2,29,22,37,24]. This means that the questions related to the existence, uniqueness and stability with respect to the input data are not trivial [20]. There are two main frameworks for the solution of ill-posed inverse problems, regularization methods [34,15] and statistical inversion framework [24].

The regularization framework targets at the minimization of the least squares norm, containing the discrepancies between measured and computed data. From the statistical point of view, the minimization of the least squares norm is related to maximum likelihood estimation if the following statistical hypotheses are valid: the errors in the measured variables are additive, uncorrelated, normally distributed and have zero mean and identical variance [6]. If the inverse problem involves the estimation of only few unknown parameters, such as the uniform thermal conductivity of an object, the minimization of the least squares norm is usually a stable problem. However, if the task is to estimate a large number of parameters, such as the distribution of thermal conductivity inside the object, the problem is an ill-posed one and least squares approaches are not feasible. In such cases, regularization (stabilization) techniques are required, such as truncated singular value decomposition, projection methods, Tikhonov regularization or iterative methods, such as Landweber or Alifanov iterations [34,2,7,13,32,27,36,30,35]. A common feature of the regularization methods is that they force stability by

* Corresponding author. Tel.: +358 505407012; fax: +358 17162585.
E-mail address: Ville.Kolehmainen@uku.fi (V. Kolehmainen).

0017-9310/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijheatmasstransfer.2007.07.015

Nomenclature

$c(x)$	volumetric heat capacity distribution (J/m^3 K)
f	vector of unknown distributed parameters
f^*	vector of prior means for the parameters
f_{MAP}	maximum a posteriori estimate for parameters f
m_{d}	number of temperature sensors
m_{h}	number of heater elements
m_{t}	number of transient measurements per sensor
n_{p}	number of pixels in the image
$q(x,t)$	surface heat flux (W/m^2)
T	temperature (K)
t	time variable (s)
x	position vector in $\mathbb{R}^2$
y	vector of temperature measurements
α_ℓ	vector of the coefficients of the FEM approximation
Γ_f	prior covariance matrix for the parameters f
Ξ_k	patch of the object boundary modelling the kth heater element
$\kappa(x)$	thermal conductivity distribution (W/m K)
$\varphi(x)$	basis functions of the FEM approximation
Ω	object domain in $\mathbb{R}^2$
$\partial\Omega$	exterior boundary of the object domain
ξ_j	location of jth temperature sensor at the object boundary
$\pi(\cdot)$	probability density function
ζ_c	parameter of prior density (scalar)
ζ_κ	parameter of prior density (scalar)
$\mathcal{N}(f_*, \Gamma_f)$	Gaussian probability density with mean f^* and covariance Γ_f

modifying the associated least squares problem and thus they are not based on explicit analysis and modelling of the unknowns.

Statistical inversion is based on the construction of separate models for the measurement process and the prior information on the unknowns. The inverse problem is reformulated as a problem of Bayesian inference. Both the measurements and the unknown parameters are modelled as random variables. The randomness, which reflects the observer's uncertainty concerning their values, is coded in the probability distribution models of the parameters. From the perspective of the statistical inversion theory, the solution to an inverse problem is the posterior probability distribution of the parameters of interest, when all information available has been incorporated into the model [24]. While regularization techniques are typically aimed at producing a reasonable estimate of the unknown parameters based on the data available, in the statistical inversion theory the solution to an inverse problem is not considered as a single (point) estimate, but the posterior distribution is typically used to produce both point estimates and reliability measures for the uncertainty related to the unknowns given all underlying models and the measurements [24]. Most often, this necessitates the use of sampling which is most often carried out by the Metropolis–Hasting algorithm or the Gibbs sampler. For examples on sample based solutions for inverse problems, see [16,31,26]. However, in very large scale problems sampling can turn out to be infeasible and one often has to satisfy with the maximum a posteriori (MAP) estimates and some rough error estimates. However, the design and the feasibility of the likelihood and prior models is of great importance and sampling and otherwise accurate inference is pointless unless the models are good enough.

In this paper we apply the statistical inversion approach to the solution of a tomographic inverse thermal problem. The physical problem of interest consists of heating part of the boundary of an object and measuring the temperature evolution at specific points of its surface. Such temperature measurements are used in the inverse analysis for the reconstruction of the distributions of thermal conductivity and volumetric heat capacity inside the object.

Thermal tomography is a non-intrusive non-destructive technique for the detection of inclusions and non-homogeneities. Possible applications include detection of erosion inside steel furnaces or the existence of hydrates formations inside deep-sea oil pipelines, see e.g. [32]. The steady-state thermal tomographic problem has been addressed in reference [23], while in the present paper heat transfer within the object is formulated in terms of transient, two-dimensional linear conduction. However, in reality for many applications the steady-state is reached for excessively long times and the inverse analysis need to be undertaken during transients. Most importantly, the transient measurements contain much richer information on the unknown distributions than the steady state measurements. A linearized version of thermal tomographic problem with transient data has been previously discussed in [14]. In that paper, the volumetric heat capacity was assumed a known constant and the linearized inverse problem was to reconstruct a small unknown deviation in the thermal conductivity from a known constant background value. Also, related shape estimation type approaches for the detection of structural flaws or corrosion defects from materials with known thermal properties have been proposed in [5,8,10]. In these approaches, the goal is to reconstruct the shape of a part of the exterior boundary of the object based on the temperature data.

In this paper, the unknown distributions of volumetric heat capacity and thermal conductivity are modelled as mutually independent Gaussian Markov random fields. More specifically, both parameters are modelled as spatially smooth distributions, whose smoothness and uniformity with respect to the variance and other spatial properties can be controlled. Moreover, the natural constraint of

positivity of the thermal conductivity and volumetric heat capacity, is taken into account by using a logarithmic parametrization of both coefficients. With the logarithmic parameterization, the MAP estimation problem becomes a unconstrained optimization problem instead of a otherwise resulting constrained one. Another advantage of the logarithmic parameterization is that it reduces the magnitude of the difference in the values of the thermal conductivity and volumetric heat capacity parameters, and this scaling may improve the numerical stability of the optimization algorithm and is helpful in estimating target distributions with large contrast. We carry out a simulation study and show that the transient data and the prior model facilitate the stable estimation of spatially varying thermal parameters. Furthermore, the prior model is shown to be a robust model in the sense that even discontinuous targets can be relatively well located.

2. Methods

2.1. Forward problem

Let $\Omega \subset \mathbb{R}^2$ be the object domain with boundary $\partial\Omega$. In the measurement setup, m_h heater elements and m_d temperature sensors are attached to the boundary $\partial\Omega$ of the object. The heater elements are modelled by the disjoint surface patches $\Xi_k \subset \partial\Omega$ and the temperature measurement locations are denoted by $\xi_j \in \partial\Omega$. In the measurement process, a known heat flux is applied to one of the heater elements, say Ξ_k, and the evolution of the temperature is measured at time instants $\{t_1, t_2, \ldots, t_{m_\mathrm{t}}\}$ at each of the measurement locations $\xi_j, j = 1, 2, \ldots, m_\mathrm{d}$. This heating and measurement process is then repeated for all the m_h heater elements.

The forward model we use for these measurements is

$$c(x)\frac{\partial T(x,t)}{\partial t} = \nabla\cdot(\kappa(x)\nabla T(x,t)) \quad \text{in } \Omega \tag{1}$$

$$\kappa(x)\frac{\partial T(x,t)}{\partial n} = q(x,t) \quad \text{on } \partial\Omega \tag{2}$$

$$T|_{t=0} = T_0, \tag{3}$$

where $c(x)$ is the volumetric heat capacity distribution and $\kappa(x)$ is the thermal conductivity distribution and T_0 is the initial temperature.

In this paper, the numerical solution of the above initial value problem is based on the finite element method (FEM). The variational form of the model (1)–(3) is

$$\int_\Omega c\frac{\partial T}{\partial t}v\,\mathrm{d}x = \int_{\partial\Omega} qv\,\mathrm{d}S - \int_\Omega \kappa\nabla T\cdot\nabla v\,\mathrm{d}x, \tag{4}$$

$$T|_{t=0} = T_0, \tag{5}$$

where $v \in H_1(\Omega)$, $H_1(\Omega)$ being the associated Sobolev space.

In the FEM discretization, the domain Ω is divided into N_e disjoint triangular elements, joined at N_n vertex nodes. For the solution of (4) we write the approximation

$$T(x,t) = \sum_{k=1}^{N_n} a_k(t)\varphi_k(x) \in Q_h, \tag{6}$$

where φ_k are the nodal basis functions of the FEM mesh and $Q_h = \mathrm{span}\{\varphi_k\} \subset H_1(\Omega)$. Using the approximation (6), the semidiscrete Galerkin scheme gives

$$M(c)\frac{\partial\alpha}{\partial t} = -G(\kappa)\alpha + V, \tag{7}$$

where $\alpha = (a_1, a_2, \ldots, a_{N_n})^\mathrm{T} \in \mathbb{R}^{N_n}$. The elements of the mass and stiffness matrices and the boundary term are

$$M_{jk} = \int_\Omega c\varphi_j\varphi_k\,\mathrm{d}x, \quad G_{jk} = \int_\Omega \kappa\nabla\varphi_j\cdot\nabla\varphi_k\,\mathrm{d}x,$$

$$V_j = \int_{\partial\Omega} q\varphi_j\,\mathrm{d}S.$$

With fixed c and κ, the semidiscrete form (7) can be solved numerically for α for example with the implicit Euler iteration with a constant time step Δt. This leads to the form

$$(I + M^{-1}G\Delta t)\alpha_{\ell+1} = \alpha_\ell + M^{-1}V\Delta t. \tag{8}$$

where we denote $\alpha_\ell = (a_1(\ell\Delta t), \ldots, a_{N_n}(\ell\Delta t))^\mathrm{T}$. The computed temperature data $T(\xi_j, t_\mathrm{p})$ at the measurement location $\xi_j \in \partial\Omega$ at the specified measurement time t_p is obtained by interpolation from the values $\{T(\xi_j, \ell\Delta t), \ell = 0, 1, 2, \ldots, N_t\}$ obtained by Eq. (6) at the time steps used in the implicit Euler iteration.

To complete the specification of the forward model, we introduce the notation that will be used for data and forward problem in the subsequent sections. The temperature evolution is measured at time instants $\{t_1, t_2, \ldots, t_{m_\mathrm{t}}\}$ at each of the locations $\{\xi_j, j = 1, \ldots, m_\mathrm{d}\}$ as response to the excitation of the heat source at Ξ_k. The corresponding computed data is concatenated into a vector

$$T^{(k)} = (T(\xi_1, t_1), T(\xi_1, t_2), \ldots, T(\xi_{m_\mathrm{d}}, t_{m_\mathrm{t}-1}), T(\xi_{m_\mathrm{d}}, t_{m_\mathrm{t}}))^\mathrm{T} \in \mathbb{R}^{m_\mathrm{t}m_\mathrm{d}} \tag{9}$$

giving the forward solution for the source at Ξ_k. By concatenating the vectors $T^{(k)}$ for all the m_h heat source locations into a single column vector, we obtain the model for all computed (errorless) data

$$\bar{y} = (T^{(1)}, \ldots, T^{(m_\mathrm{h})})^\mathrm{T} \in \mathbb{R}^m, \quad m = m_\mathrm{t}m_\mathrm{h}m_\mathrm{d} \tag{10}$$

for the thermal tomography experiment.

At this stage, we specify the discretization for the unknowns (κ, c) in the inverse problem. We divide the domain Ω to a set of n_p disjoint image pixels (i.e., volume elements)

$$\Omega = \bigcup_{j=1}^{n_\mathrm{p}} \Omega_j \tag{11}$$

and use the piecewise constant approximations

$$\kappa(x) = \sum_{j=1}^{n_\mathrm{p}} \kappa_j\chi_j(x), \tag{12}$$

$$c(x) = \sum_{j=1}^{n_p} c_j \chi_j(x), \tag{13}$$

where χ_j denote the characteristic functions of the image pixels Ω_j. We identify the coefficients (distributed parameters) $(\kappa(x), c(x))$ with the vectors $\kappa \in \mathbb{R}^{n_p}$ and $c \in \mathbb{R}^{n_p}$ containing the pixel values of the thermal conductivity and volumetric heat capacity. With these choices, the parameter vector for the inverse problem becomes

$$f = \begin{pmatrix} \kappa \\ c \end{pmatrix} \in \mathbb{R}^n, \quad n = 2n_p. \tag{14}$$

For the FEM-based forward solver we use the notation

$$\bar{y} = T(f), \tag{15}$$

that is, the vector of computed (errorless) data for given distributions c and κ.

2.2. Likelihood, prior and posterior models

In the Bayesian framework for inverse problems, the first task is to construct the likelihood and prior models in terms of probability densities [24]. The likelihood is the stochastic model for the observations given that all the unknowns were known. The likelihood density is denoted by $\pi(y|f)$, where y denote all observations and f all unknowns. The prior model is the stochastic model for all unknowns and is denoted by $\pi_{\text{pri}}(f)$. The prior model should reflect all the prior information and especially all uncertainty in the unknowns. Thus, for example, if the unknown is a positive variable, we should have $\pi_{\text{pri}}(f) = 0$ for all $f < 0$.

Given the likelihood and prior models, the information and uncertainty in the unknowns once the measurements have been obtained, is contained in the posterior density $\pi(f|y)$. The posterior density is obtained by the Bayes' theorem and we have

$$\pi(f|y) \propto \pi(y|f)\pi_{\text{pri}}(f),$$

where the proportionality factor is the marginal density of the measurements and is not needed in most calculations. From the posterior density point and spread estimates can be calculated. The most straightforward estimate is the maximum a posteriori (MAP) estimate, the computation of which is a minimization problem. In contrast, the computation of the posterior mean estimate is formally an integration problem which in most high dimensional cases turns out to be a computationally massive task since it requires the use of Markov chain Monte Carlo (MCMC) methods. This applies also to the computation of exact spread estimates such as confidence interval type estimates. For general references on MCMC, see e.g. [21,17,19,18].

The standard choice is to model the measurement errors as zero mean additive Gaussian noise. This is the case for example when the noise is due to thermal noise from the electronics of the measurement equipment only, and the computational model for the forward problem is an adequately accurate one for the required accuracy, see [24,25,4] for the problem related to approximation errors.

In this paper we use relatively accurate forward models and shall also adopt the conventional measurement error model. Thus, we write for the observation model

$$y = T(f) + e$$

where y is vector of measured temperatures, $e \sim \pi_e = \mathcal{N}(0, \Gamma_e)$ and Γ_e is the covariance matrix of the noise e with $\Gamma_e^{-1} = L_e^{\mathrm{T}} L_e$. Assuming that the measurement errors and the unknown parameters are statistically independent, we obtain the likelihood model

$$\pi(y|f) = \pi_e(y - T(f)) \propto \exp\left\{-\frac{1}{2}\|L_e(y - T(f))\|^2\right\}.$$

The standard approach to model smooth coefficient functions is as follows. The thermal conductivity and volumetric heat capacity distributions are modelled as spatially smooth Gaussian Markov random fields with means κ_*, c_* and covariances Γ_κ, Γ_c, respectively. Furthermore, if they are modelled as mutually statistically independent, we have the prior model $\pi_{\text{G}}(f)$

$$f \sim \pi_{\text{G}}(f) = \mathcal{N}(f_*, \Gamma_f), \tag{16}$$

where

$$f_* = \begin{pmatrix} \kappa_* \\ c_* \end{pmatrix}, \quad \Gamma_f = \begin{pmatrix} \Gamma_\kappa & 0 \\ 0 & \Gamma_c \end{pmatrix}, \quad \Gamma_f^{-1} = L_f^{\mathrm{T}} L_f. \tag{17}$$

Details of the construction of the terms f^* and Γ_f are explained in Section 2.3.

The thermal conductivity and volumetric heat capacity are positive valued functions. To take the positivity of the parameter distributions into account, the standard approach is to include the positivity prior

$$\pi_+(f) = \prod_{i=1}^{n} \pi_+(f_i), \quad \pi_+(f_i) = \begin{cases} 1, & f_i > 0 \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

With these models, the overall prior $\pi_{\text{pri}}(f)$ becomes

$$\pi_{\text{pri}}(f) = \pi_+(f)\pi_{\text{G}}(f). \tag{19}$$

Thus we have for the posterior density

$$\begin{aligned} \pi(f|y) &\propto \pi(y|f)\pi_{\text{pri}}(f) = \pi_e(y - T(f))\pi_{\text{pri}}(f) \\ &\propto \pi_+(f) \exp\left\{-\frac{1}{2}(\|L_e(y - T(f))\|^2 + \|L_f(f - f_*)\|^2)\right\}. \end{aligned} \tag{20}$$

2.3. Construction of a proper smoothness prior

The conventional smoothness prior model for the mutually independent pair (κ, c) is defined as

$$\pi_{\text{G}}(f) \propto \exp\left\{-\frac{1}{2}\|Lf\|^2\right\} = \exp\left\{-\frac{1}{2} f^{\mathrm{T}} B f\right\}, \tag{21}$$

where $B = L^{\mathrm{T}}L$, L is a block matrix

$$L = \begin{pmatrix} \zeta_\kappa D & 0 \\ 0 & \zeta_c D \end{pmatrix}, \tag{22}$$

where D is a first order difference matrix between adjacent pixels in the image basis (see Eqs. (11)–(13)) and $\zeta_\kappa > 0$, $\zeta_c > 0$ are scalar parameters. The prior density (21) is improper (i.e., B is not invertible) because the matrix D has null space $\mathrm{Ker}(D) = \mathrm{span}\{u\}$, where $u = -\sqrt{2/n}(1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^{n/2}$. Mathematically, this type of a model has infinite (prior) variance for each variable.

In diffuse tomographic problems in which the measurements are obtained only on the boundary $\partial\Omega$ of the domain, the measurements reduce the uncertainty significantly more near the boundary than in the center of the domain. When conventional improper smoothness prior models are used the homogeneity of the resulting posterior estimates tend to be poor.

To avoid this problem, we construct a proper prior model which exhibits the desired property that the marginal distributions of all variables can be controlled. The prior model (21) is used as a starting point in the construction of proper smoothness prior model.

For this end, we follow the procedure in [24,9]. Assume that we have a prior model for the mean and the variability of the actual values of the thermal conductivity and volumetric heat capacity (κ, c) at some points or pixels in the domain Ω. Let k be the number of these pixels. Next, we want to write a joint marginal prior density for these pixels. By possibly reordering the elements of f, we may write the partition $f = (f_1, f_2)^{\mathrm{T}} \in \mathbb{R}^n$, where $f_2 \in \mathbb{R}^{2k}$ contain the thermal conductivity and heat capacity parameters of the "specified" pixels where we can specify prior model for the actual values, and $f_1 \in \mathbb{R}^{n-2k}$ contain the thermal conductivity and heat capacity parameters of the remaining,"unspecified" pixels. By partitioning accordingly the matrix

$$L^{\mathrm{T}}L = B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

in Eq. (21), we can write a proper conditional smoothness prior for f_1 conditioned on f_2 as [24,9]:

$$\pi_{\mathrm{G}}(f_1|f_2) \propto \exp\left\{-\frac{1}{2}(f_1 + B_{11}^{-1}B_{12}f_2)^{\mathrm{T}}B_{11}^{-1}(f_1 + B_{11}^{-1}B_{12}f_2)\right\}. \tag{23}$$

Next, assume that the thermal conductivity and heat capacity in the specified pixels is modelled by a Gaussian prior

$$\pi_{\mathrm{G}}(f_2) \propto \exp\left\{-\frac{1}{2}(f_2 - f_{*,2})^{\mathrm{T}}\Gamma_0^{-1}(f_2 - f_{*,2})\right\}, \tag{24}$$

where vector $f^*{}_{,2}$ is the mean and Γ_0 is the covariance matrix, respectively.

Using Eqs. (23) and (24), we can now obtain a new, proper smoothness prior for f as [24,9]:

$$\begin{aligned}\pi_{\mathrm{G}}(f_1, f_2) &= \pi_{\mathrm{G}}(f_1|f_2)\pi_{\mathrm{G}}(f_2) \\ &\propto \exp\left\{-\frac{1}{2}(f - f_*)^{\mathrm{T}}\Gamma_f^{-1}(f - f_*)\right\},\end{aligned} \tag{25}$$

where

$$f_* = \begin{pmatrix} -B_{11}^{-1}B_{12}f_2 \\ f_{*,2} \end{pmatrix}, \quad \Gamma_f^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{21}B_{11}^{-1}B_{12} + \Gamma_0^{-1} \end{pmatrix}. \tag{26}$$

Further, let us denote the Cholesky decompositions of B_{11} and Γ_0^{-1} by $B_{11} = L_1^{\mathrm{T}}L_1$ and $\Gamma_0^{-1} = L_0^{\mathrm{T}}L_0$. Then we can write the prior model (25) and (26) as

$$\pi_{\mathrm{G}}(f) \propto \exp\left\{-\frac{1}{2}\|L_f(f - Qf_{*,2})\|^2\right\}, \tag{27}$$

where

$$L_f = \begin{pmatrix} L_1 & L_1B_{11}^{-1}B_{12} \\ 0 & L_0 \end{pmatrix}, \quad Q = \begin{pmatrix} -B_{11}^{-1}B_{12} \\ I \end{pmatrix}. \tag{28}$$

For example, if we set the variables f_2 to be mutually independent and spatially homogeneously distributed in Ω, the distance between the fixed pixels controls the correlation distance of the coefficient estimates. Moreover, if the marginal variances of the fixed pixels are equal, the rest of the pixels can be adjusted to have approximately the same variances too, thus resulting in spatial homogeneity of the prior model, see [24,9] for details. For typical draws for the prior, see Section 3.

2.4. Change of variables and computation of the MAP estimate

The computation of the MAP estimate

$$f_{\mathrm{MAP}} = \arg\max \pi(f|y)$$

for the model (20) amounts to the constrained minimization problem

$$f_{\mathrm{MAP}} = \arg\min_{f>0}\{\|L_e(y - T(f))\|^2 + \|L_f(f - f_*)\|^2\}. \tag{29}$$

The use of constrained optimization techniques for minimizing (29) for high-dimensional nonlinear distributed parameter problems such as the thermal tomography problem can be computationally cumbersome. Moreover, if the expected ranges of the thermal conductivity and heat capacity values are several orders of magnitude apart, numerical instability problems may occur in the computation of the MAP estimate.

In order to avoid the constrained optimization problem and the range problem, we introduce the change of variables

$$\theta = \log(f) \in \mathbb{R}^n,$$

which allows us to drop the positivity constraints since $f = \exp(\theta) > 0$. Furthermore, if we have for example

$\theta_{\max} - \theta_{\min} = 7$, the range of the respective coefficient f covers three orders of magnitude.

In terms of the parameter θ the computation of the MAP estimates now takes the form

$$\begin{aligned}\theta_{\mathrm{MAP}} &= \arg\min_{\theta}\{\mathscr{L}(\theta)+\mathscr{P}(\theta)\} \\ &= \arg\min_{\theta}\{\|L_e(y-T(\exp\{\theta\}))\|^2+\|L_f(\exp\{\theta\}-f_*)\|^2\},\end{aligned} \tag{30}$$

where $\mathscr{L}$ and $\mathscr{P}$ are called the likelihood and prior potentials, respectively.

The mapping $f \mapsto T$ is nonlinear, and thus the introduction of a further exponential nonlinearity does not result in any increase in computational or implementational tasks.

Since all associated mappings are smooth, we compute the MAP estimate (30) with a (simplified) Newton method equipped with line search. The Newton type iteration can be written as

$$\theta^{(i+1)} = \theta^{(i)} + l^{(i)}\delta\theta^{(i)} \tag{31}$$

where the search direction $\delta\theta^{(i)}$ is obtained as the solution of

$$(J_\theta^{\mathrm{T}} L_e^{\mathrm{T}} L_e J_\theta + H_\theta)\delta\theta^{(i)} = J_\theta^{\mathrm{T}} L_e^{\mathrm{T}} L_e(y - T(\exp\{\theta^{(i)}\})) - g_{\theta^{(i)}},$$

where $J_\theta = J_f(\exp\{\theta^{(i)}\})$ is the Jacobian matrix of the map $T(\exp\{\theta\})$, the scalar $l^{(i)}$ is a step length parameter determined by a line search algorithm and

$$\begin{aligned} g_\theta &= \Lambda_\theta L_f^{\mathrm{T}} L_f(\exp\{\theta\} - f_*) \\ H_\theta &= \Lambda_\theta(L_f^{\mathrm{T}} L_f)\Lambda_\theta + \mathrm{diag}(g_\theta)\end{aligned}$$

are the gradient and Hessian of the prior potential $\mathscr{P}(\theta)$ in Eq. (30), respectively, and

$$\Lambda_\theta = \mathrm{diag}(\exp\{\theta\}).$$

A simplification compared to Newton's method in iteration (31) is that the Hessian of the likelihood term $\mathscr{L}(\theta)$ is approximated with the standard Gauss–Newton approximation [28,1].

Consider next the computation of the Jacobian matrix J_θ. For notational simplicity, we will derive the computation of the Jacobian

$$J_f = \begin{pmatrix} \frac{\partial T^{(1)}}{\partial \kappa_1}, & \cdots, & \frac{\partial T^{(1)}}{\partial \kappa_{n_\mathrm{p}}}, & \frac{\partial T^{(1)}}{\partial c_1}, & \cdots, & \frac{\partial T^{(1)}}{\partial c_{n_\mathrm{p}}} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial T^{(m_\mathrm{h})}}{\partial \kappa_1}, & \cdots, & \frac{\partial T^{(m_\mathrm{h})}}{\partial \kappa_{n_\mathrm{p}}}, & \frac{\partial T^{(m_\mathrm{h})}}{\partial c_1}, & \cdots, & \frac{\partial T^{(m_\mathrm{h})}}{\partial c_{n_\mathrm{p}}} \end{pmatrix},$$

where the vectors $\frac{\partial T^{(s)}}{\partial \kappa_k} \in \mathbb{R}^{m_\mathrm{t} m_\mathrm{d}}, \frac{\partial T^{(s)}}{\partial c_k} \in \mathbb{R}^{m_\mathrm{t} m_\mathrm{d}}$ contain the derivatives of the forward solution corresponding to the s:th heater element, see Eq. (9). The corresponding blocks of the Jacobian J_θ can be then obtained by the chain rule of differentiation as

$$\begin{aligned} \frac{\partial T^{(s)}}{\partial \theta_k} &= \frac{\partial T^{(s)}}{\partial \kappa_k}\frac{\partial \kappa_k}{\partial \theta_k} = \frac{\partial T^{(s)}}{\partial \kappa_k}\exp\{\theta_k\}, \quad k = 1,\ldots,n_\mathrm{p} \\ \frac{\partial T^{(s)}}{\partial \theta_{k+n_\mathrm{p}}} &= \frac{\partial T^{(s)}}{\partial c_k}\frac{\partial c_k}{\partial \theta_{k+n_\mathrm{p}}} = \frac{\partial T^{(s)}}{\partial c_k}\exp\{\theta_{k+n_\mathrm{p}}\}, \quad k = 1,\ldots,n_\mathrm{p}. \end{aligned}$$

Denote $\beta_\ell^k = \partial\alpha_\ell/\partial c_k \in \mathbb{R}^{N_n}$. We differentiate Eq. (8) with respect to coefficient c_k to get

$$\underbrace{\Delta t\frac{\partial M^{-1}}{\partial c_k}G}_{A_k}\alpha_{\ell+1} + \underbrace{(I + M^{-1}G\Delta t)}_{B}\beta_{\ell+1}^k = \beta_\ell^k + \underbrace{\Delta t\frac{\partial M^{-1}}{\partial c_k}V}_{C_k} \tag{32}$$

$$\beta_{\ell+1}^k = B^{-1}(\beta_\ell^k - A_k\alpha_{\ell+1} + C_k) \tag{33}$$

where $\alpha_\ell \in \mathbb{R}^{N_n}$ is the current iterate of the forward solution (see Eq. (8)) and we have

$$\frac{\partial M^{-1}}{\partial c_k} = -M^{-1}\frac{\partial M}{\partial c_k}M^{-1}.$$

Correspondingly, we have for $\gamma_\ell^k = \partial\alpha_\ell/\partial\kappa_k \in \mathbb{R}^{N_n}$

$$\underbrace{\Delta t M^{-1}\frac{\partial G}{\partial \kappa_k}}_{D_k}\alpha_{\ell+1} + \underbrace{(I + M^{-1}G\Delta t)}_{B}\gamma_{\ell+1}^k = \gamma_\ell^k \tag{34}$$

$$\gamma_{\ell+1}^k = B^{-1}(\gamma_\ell^k - D_k\alpha_{\ell+1}). \tag{35}$$

Note that both $\partial M/\partial c_k$ and $\partial G/\partial\kappa_k$ are very sparse low rank matrices and do not depend on either c or κ. The elements of the vectors

$$\frac{\partial T^{(s)}}{\partial c_k} \in \mathbb{R}^{m_\mathrm{t} m_\mathrm{d}}, \quad \frac{\partial T^{(s)}}{\partial \kappa_k} \in \mathbb{R}^{m_\mathrm{t} m_\mathrm{d}}$$

at the measurement locations ξ_j at the specified measurement times t_p can be obtained from solutions of Eqs. (33) and (35) by utilizing Eq. (6) and the temporal interpolation that was explained for the computation of the temperatures $T(\xi_j, t_p)$ in Section 2.1.

3. Numerical results

We evaluate the proposed estimation approach of the thermal conductivity and volumetric heat capacity by simulations in the case $\Omega \subset \mathbb{R}^2$. The object domain is a circle with radius of 25 mm. The measurement system in the simulations consist of $m_\mathrm{h} = 8$ heater elements Ξ_k, each covering 1/8 of the length of the boundary $\partial\Omega$, and $m_\mathrm{d} = 8$ point-like temperature detectors located between the heater elements on $\partial\Omega$ at $x = \xi_j$. In the simulated measurement, the heat flux $q_\mathrm{in} = 300\ \mathrm{W/m^2}$ is applied to each of the heater elements Ξ_k at a time for period of 120 s, and the temperatures at all eight detectors are measured for 600 s with measurement interval of 2 s. This measurement process is then repeated for all the eight sources, leading to a measurement vector $y \in \mathbb{R}^{19200}$.

In this study, we assume that the object is insulated from the surrounding space during the heating and temperature measurements, and thus the heat flux through $\partial\Omega$ is non-

zero only on the part of boundary Ξ_k modelling the active heating element. Thus, we write the model

$$q(x,t) = \begin{cases} q_{\text{in}}, & x \in \Xi_k, \quad t \in [0\ 120]\ \text{s} \\ 0, & \text{otherwise.} \end{cases} \tag{36}$$

for the Neumann data in the forward model.

The simulated target distributions (κ, c) for the simulation are shown in top left and right in Fig. 4, respectively. The simulated object Ω consist of two circular inclusions on homogeneous background. The values of the thermal conductivity and volumetric heat capacity in the background are $\kappa_0 = 1\ \text{W/m K}$ and $c_0 = 10^6\ \text{J/m}^3\ \text{K}$ corresponding to typical values of clay. The values of κ and c in the circular inclusions are $\kappa_1 = 0.03\ \text{W/m K}$ and $c_1 = 1000\ \text{J/m}^3\ \text{J K}$ corresponding to the thermal conductivity and heat capacity values of air.

For the simulation of the temperature data, the domain Ω is divided into a finite element mesh of $N_e = 3856$ triangular elements with $N_n = 2057$ nodes. The simulated measurement data is computed using time step parameter $\Delta t = 1$ s in the implicit Euler method (8). Gaussian random noise with standard deviation $\sigma = 0.5$ K is added to the simulated temperature measurements. Examples of temperature measurements for two measurement sensors are given in Fig. 1. The top figure shows the temporal evolution of the temperature for a detector next to the active heating element and the bottom figure for a detector opposite to the active heating element, respectively.

For the estimation of the thermal conductivity and heat capacity parameters, the domain Ω is divided into a mesh of $N_e = 2768$ triangular elements with $N_n = 1513$ nodes, and the images (κ, c) are represented in a lattice of *524* image elements, leading to unknown $f \in \mathbb{R}^{1028}$ in (29). The division of the domain Ω into the 524 image elements is displayed in Fig. 2. When solving (29) by the Newton type method, a time step of $\Delta t = 2$ s is used in the implicit Euler method for forward solutions and Jacobian. Thus in this case the time steps for the solution (8) coincide with the measurement times. If they do not, the measurements can be interpolated as explained in Section 2.1. Note that, in order to avoid the so-called "inverse crime",[1] the simulation of the measurements has been carried out in a computationally more accurate model than the one used in the estimation of the coefficients.

For the unknown f we construct a proper first order smoothness prior distribution as explained in Section 2.3. We aim to control the approximative correlation length of the coefficient distributions as follows. We choose seven pixels as "specified" pixels such that their distances from each other is approximately one fourth of the diameter of the domain Ω. The location of the seven specified pixels are shown with asterisks in Fig. 2. The thermal conductivity and volumetric heat capacity parameters of the specified

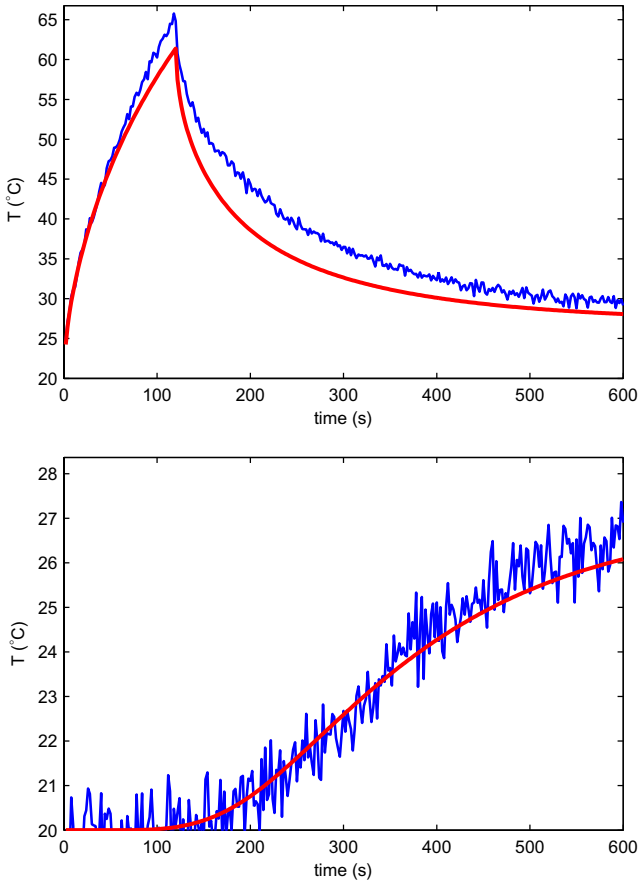

Fig. 1. An example of the temperature measurements. Top: Detector adjacent to the active heating element which is turned off at $t = 120$ s. Bottom: Detector opposite to the active heating element. Thin solid line: Noisy measurement for the object Ω shown in the top row of Fig. 4. Thick solid line: Forward solution corresponding to the object without the air-inclusions (i.e., noiseless computed data corresponding to constant values (κ_0, c_0) over the entire domain Ω).

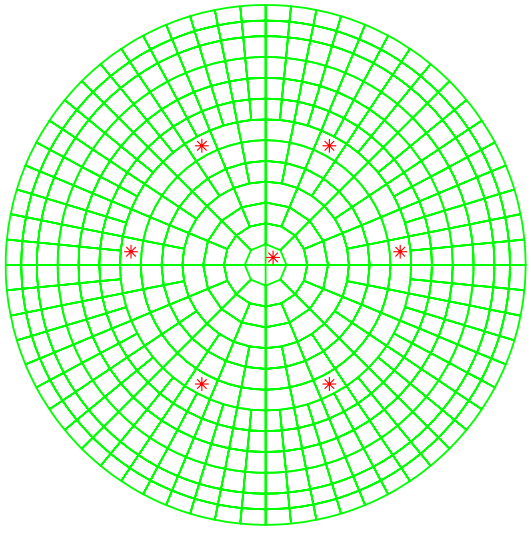

Fig. 2. The division of the domain Ω into a lattice of 524 image elements. The location of the seven specified pixels for the construction of the proper smoothness prior density are denoted by asterisks.

[1] An inverse crime refers to an overly optimistic test setting where an identical numerical model is used for both the simulation of the measurement data and solution of the inverse problem [24].

pixels define elements $f_2 \in \mathbb{R}^{14}$ of the unknown f. For the marginal distribution of the parameters f_2 we specify a Gaussian prior model of the form (24) with diagonal covariance matrix Γ_0 (i.e., elements of f_2 are assumed mutually independent). For the prior mean of f_2 we assume values $\kappa_* = 1\ \mathrm{W/m\,K}$ and $c_* = 10^6\ \mathrm{J/m^3\,K}$, respectively. The standard deviations $(\sigma_\kappa, \sigma_c)$ of f_2 are chosen such that $[\kappa_* - 3\sigma_\kappa,\ \kappa_* + 3\sigma_\kappa] = [0.5,\ 1.5]\ \mathrm{W/m\,K}$ and $[c_* - 3\sigma_c, c_* + 3\sigma_c] = [0.5, 1.5] \times 10^6\ \mathrm{J/m^3\,K}$, respectively. These choices correspond to the assumption that the values of (κ, c) in the specified pixels lie with prior probability 0.99 within these intervals.

Once the marginal distribution for f_2 has been specified, the parameters ζ_κ and ζ_c for the difference operator, see Eq. (22), are tuned such that the pixelwise prior variances in the smoothness prior model $\pi_G(f)$, Eq. (27), for the thermal conductivity and volumetric heat capacity images become approximately equalized over the entire image. That is, we choose ζ_κ and ζ_c such that we have for the marginal variances $\mathrm{var}(\kappa_k) \approx \sigma_\kappa^2$ and $\mathrm{var}(c_k) \approx \sigma_c^2$ for all k in the prior model. For a more systematic approach for the equalization of the prior variances over subdomains, see [9]. Two random samples from the prior density $\pi_G(f)$, Eq. (27), are shown in Fig. 3.

The reconstructed thermal conductivity and volumetric heat capacity images are shown on the left and right in the bottom row of Fig. 4, respectively. The reconstructions were obtained by solving the MAP estimation problem (30) with the parameterization θ, and the displayed images show the reconstruction $\exp\{\theta_{\mathrm{MAP}}\}$. The minimization problem (30) was solved with the Newton type algorithm (31) as described above. As can be seen, the air inclusions are reconstructed well in both, thermal conductivity and volumetric heat capacity, parameters. It is to be noted that we did not expect to find any inclusion type targets as such, which is reflected in the choice of the smooth prior model, see especially the typical draws in Fig. 3. However, due to the otherwise feasible construction of the prior with respect to the approximate range of expected values, the reconstructions can be argued to be at least indicative. Since the prior model is constructed for smooth targets, the boundaries of the inclusions are naturally smooth. If sharp boundaries are expected, total variation priors can be employed, see for example [12,26]. On the other hand, if

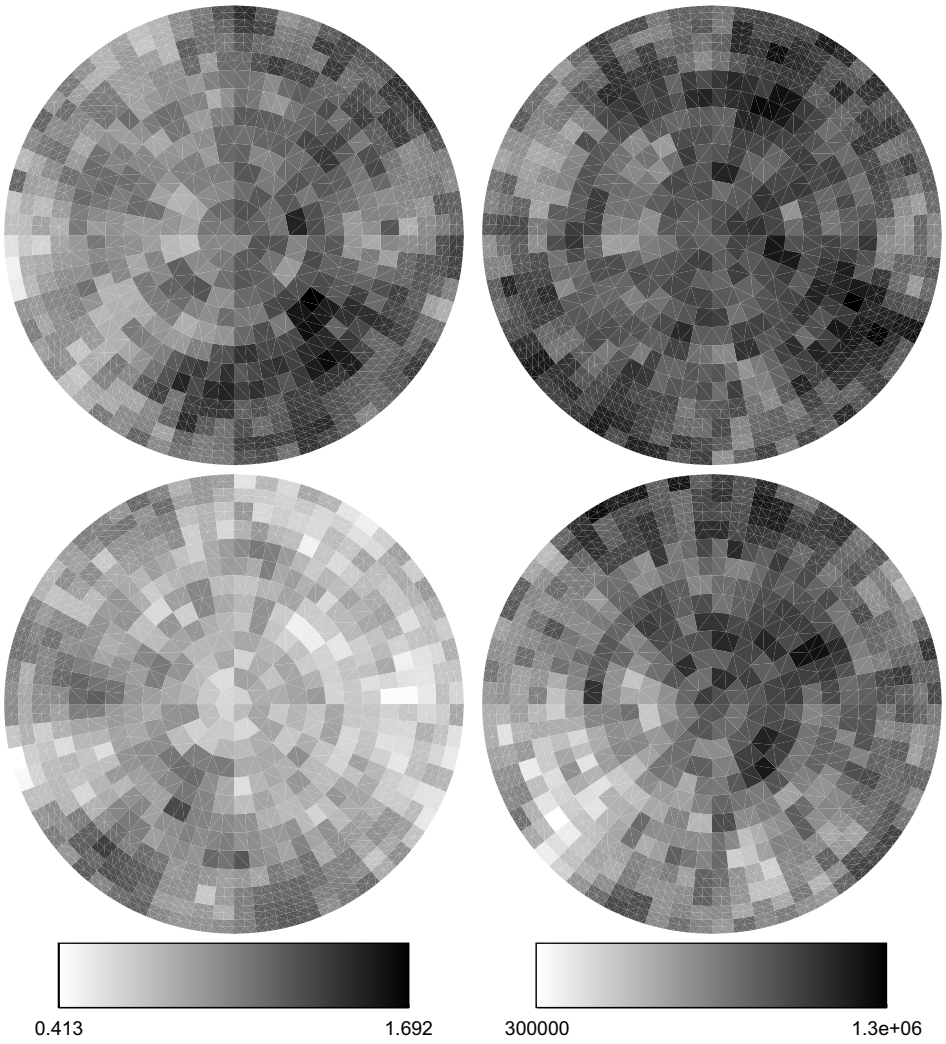

Fig. 3. Two random samples from the prior density $\pi_G(f)$, Eq. (27). Left column: Thermal conductivity κ. Right column: Volumetric heat capacity c.

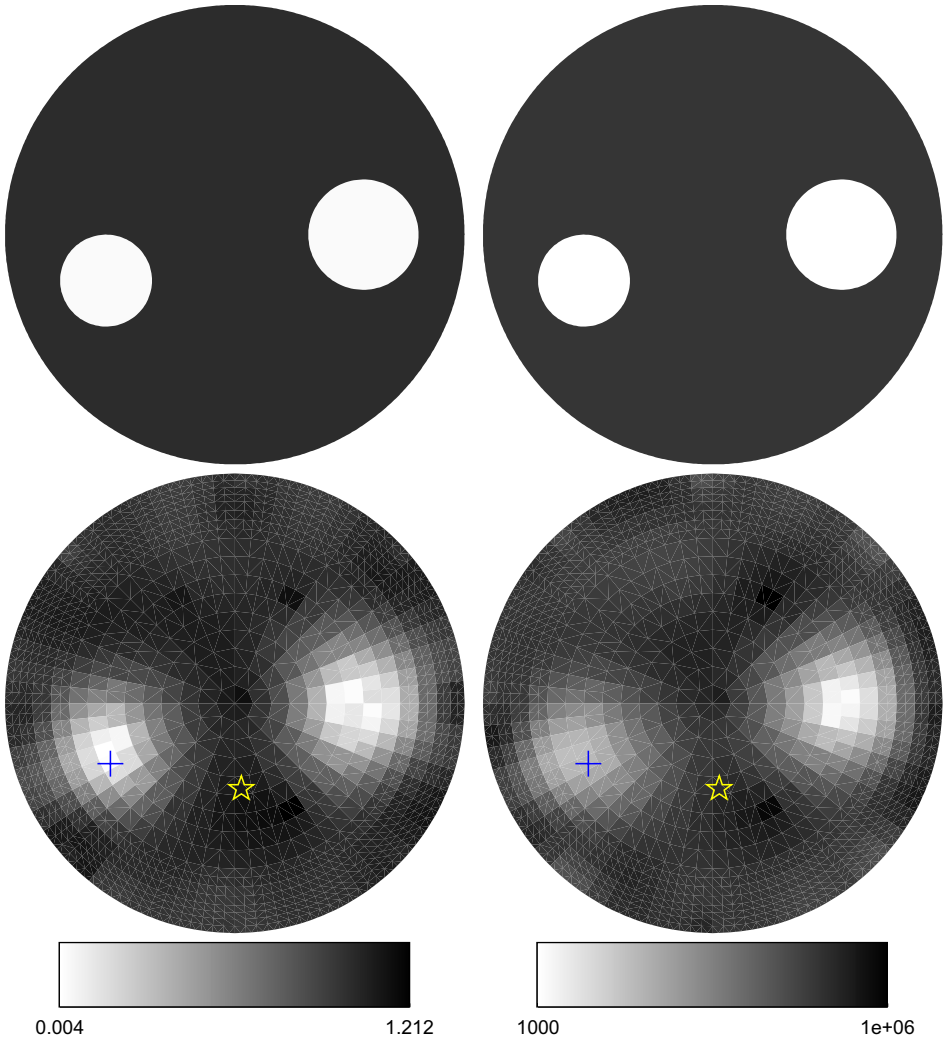

Fig. 4. Top left: Thermal conductivity κ. Top right: Volumetric heat capacity c. Bottom left: Reconstruction of the thermal conductivity. Bottom right: Reconstruction of the volumetric heat capacity.

smooth target distributions with very high contrast are expected, one could write the smoothness prior in terms of the parameterization θ instead of the parameterization f. In general, a smoothness prior model such as the one used here, can be a good robust first choice for the prior.

Fig. 5 shows approximate posterior marginal densities for the two pixels marked in the reconstructions in Fig. 4. The left column shows the approximate marginal densities of the thermal conductivities in the marked pixels, and the right column shows the approximate marginal densities of the volumetric heat capacity, respectively. The dashed vertical lines denote the values of the MAP estimate in the chosen pixels and the solid vertical lines denote the true values of the unknowns, respectively. The marginal posterior densities were approximated as follows. First, a Gaussian approximation for the posterior density of the parameters θ was obtained by linearizing the map $T(\exp\{\theta\})$ around the MAP estimate θ_{MAP}, and computing the covariance matrix by

$$\Gamma_{\theta_{\mathrm{MAP}}} = (J_{\theta_{\mathrm{MAP}}}^{\mathrm{T}} L_e^{\mathrm{T}} L_e J_{\theta_{\mathrm{MAP}}} + H_{\theta_{\mathrm{MAP}}})^{-1},$$

where $H_{\theta_{\mathrm{MAP}}}$ is the Hessian of the prior functional $\mathscr{P}(\theta)$ in Eq. (30). The approximate marginal density for the parameter θ_l can be then obtained as

$$\theta_l \sim \mathcal{N}((\theta_{\mathrm{MAP}})_l, (\Gamma_{\theta_{\mathrm{MAP}}})_{l,l}).$$

Next, the approximate marginal posterior density of the parameter $f_l = \exp\{\theta_l\}$ is obtained by standard techniques for transformations of functions of random variables, see for example [33]. Let

$$\Phi_{\theta_l}(t) = \mathbb{P}\{\theta_l \leqslant t\}$$

denote the (cumulative) distribution function of θ_l. Using the fact that $\exp\{\theta_l\}$ is monotonic function, the distribution function of f_l can be obtained by

$$\Phi_{f_l}(t) = \mathbb{P}\{\exp\{\theta_l\} \leqslant t\} = \mathbb{P}\{\theta_l \leqslant \log(t)\} = \Phi_{\theta_l}(\log(t)),$$

and the density function of f_l is then obtained as

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi_{\theta_l}(\log(t)).$$

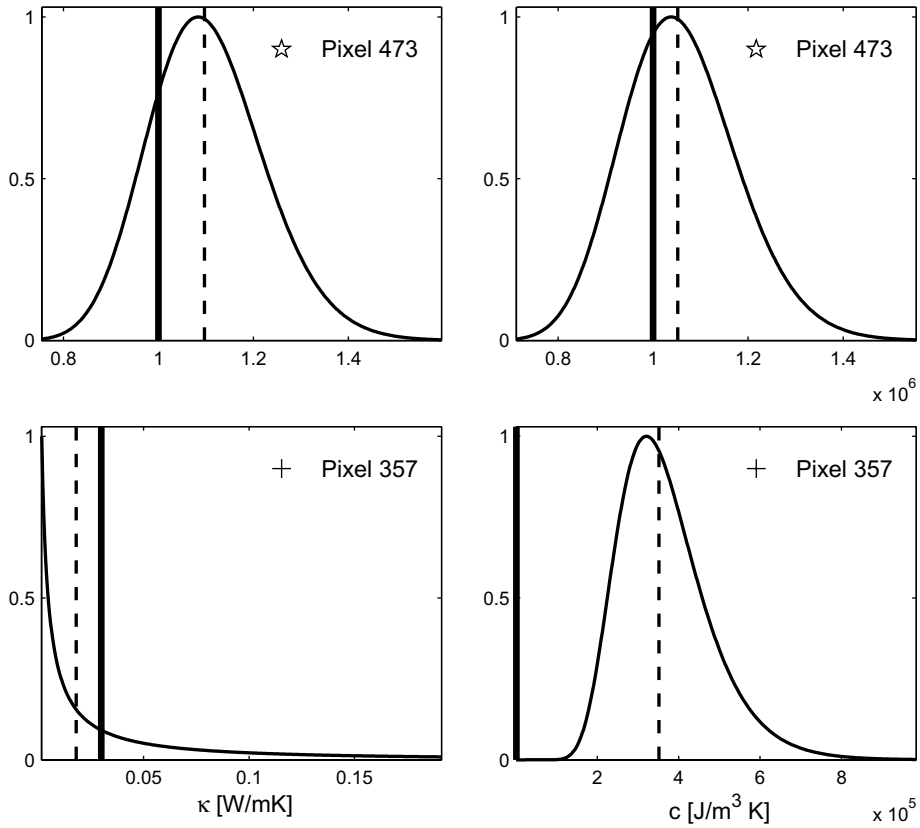

Fig. 5. Approximate marginal posterior densities for the two pixels marked in the bottom row of Fig. 4. Left column: Thermal conductivity κ. Right column: Volumetric heat capacity c. For illustration purposes, the values of the marginal densities have been scaled to the range $[0, 1]$. The dashed vertical line indicates the values of the thermal conductivity and volumetric heat capacity from the MAP estimate and the solid vertical line shows the respective true values. The line corresponding to the true value in the lower right hand figure cannot be distinguished from the vertical axis line.

The quality of the approximate error estimates depends on both the quality of the models and the nonlinearity of the mappings. In our case, using the smoothness prior model allows us to obtain feasible images of the spatial coefficient distributions but does not allow for very reliable error estimates, which is reflected by the approximate posterior marginal distribution of the heat capacity of pixel 357 (bottom right in Fig. 5). In the other three approximate marginal distributions the MAP estimate and the distribution coincide well with the actual coefficient values. If more reliable error estimates are to be obtained, the qualitative nature of the spatial target distribution has to be known better than was expected here. If small or blocky targets are to be recovered, ℓ_1 and total variation priors can be used, respectively, see for example [26]. Furthermore, the respective exact marginal distributions can in practice be obtained only via MCMC methods.

4. Conclusions

This paper deals with a tomographic inverse thermal problem, where the transient temperature measurements taken at the surface of a heated object are used for the identification of the spatial distributions of volumetric heat capacity and thermal conductivity. The Bayesian approach was used for the solution of the inverse problem, where an explicit prior model was constructed for the unknown parameters inside the domain. The computational model for the forward problem was based on a semidiscrete finite element approximation. A test case dealing with the identification of two inclusions in a conducting medium reveals that the proposed approach is stable and feasible. Moreover, the use of a proper smoothness prior model is a relatively robust one even in the case of spatially discontinuous parameter distributions.

Acknowledgements

The work of authors VK and JPK was supported by the Academy of Finland (projects 108299, 119270 and 213476, Finnish Programme for Centres of Excellence in Research 2006-2011). The visits of authors VK and JPK to COPPE/UFRJ were partially supported by CNPq, and agency for the fostering of science of the Brazilian government.

References

[1] Å. Björk, Numerical Methods for Least Squares Problems, SIAM, 1996.
[2] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, New York, 1994.

[3] S.R. Arridge, Optical tomography in medical imaging, Inverse Problems 15 (1999) R41–R93.
[4] S.R. Arridge, J.P. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, M. Vauhkonen, Approximation errors and model reduction with an application in optical diffusion tomography, Inverse Problems 22 (2006) 175–195.
[5] H.T. Banks, Boundary estimation problems arising in thermal tomography, Inverse Problems 6 (1990) 897–921.
[6] J.V. Beck, K.J. Arnold, Parameter Estimation in Engineering and Science, Wiley-Interscience, 1977.
[7] J.V. Beck, B. Blackwell, C.R. St. Clair, Inverse Heat Conduction: III – Posed Problems, Wiley-Interscience, New York, 1985.
[8] K. Bryan, L.F. Caudill Jr., Stability and reconstruction for an inverse problem for the heat equation, Inverse Problems 14 (1998) 1429–1453.
[9] D. Calvetti, J.P. Kaipio, E. Somersalo, Aristotelian prior boundary conditions, Int. J. Math. 1 (2006) 63–81.
[10] R. Chapko, R. Kress, J.-R. Yoon, On the numerical solution of an inverse boundary value problem for the heat equation, Inverse Problems 14 (1998) 853–867.
[11] M. Cheney, D. Isaacson, J.C. Newell, Electrical impedance tomography, SIAM Rev. 41 (1999) 85–101.
[12] D.C. Dobson, F. Santosa, An image enhancement technique for electrical impedance tomography, Inverse Problems 10 (1994) 317–334.
[13] G.S. Dulikravich, T.J. Martin, Inverse Shape and Boundary Condition Problems and Optimization in Heat Conduction, Number 1 in Advances in Numerical Heat Transfer, Taylor and Francis, 1996, pp. 381–426 (Chapter 10).
[14] A. Elayyan, V. Isakov, On an inverse diffusion problem, SIAM J. Appl. Math. 57 (1997) 1737–1748.
[15] H. Engl, Inverse Problems, Johannes–Kepler–Universität, Linz, Austria, 1992, Lecture notes.
[16] C. Fox, G. Nicholls, Sampling conductivity images via MCMC, in: K.V. Mardia, C.A. Gill, R.G. Aykroyd (Eds.), The Art and Science of Bayesian Image Analysis. Proceedings of the Leeds Annual Statistics Research workshop, Leeds, UK, 1–4 July, Leeds University Press, 1997, pp. 1–100.
[17] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. 6 (1984) 721–741.
[18] W.R. Gilks, S. Richardson, D.J. Spiegelhalter, Markov Chain Monte Carlo in Practice, Chapman & Hall, 1996.
[19] P.J. Green, Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, Biometrika (1995) 711–732.
[20] J. Hadamard, Lectures on Cauchy's Problem in Linear Differential Equations, Yale University Press, New Haven, CT, 1923.
[21] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1970) 97–109.
[22] V. Isakov, Inverse Problems in Partial Differential Equations, Springer, New York, 1998.
[23] M. Jones, A. Tezzuka, Y. Yamada, Thermal tomographic detection of inhomogeneities, J. Heat Transfer 117 (1995) 969–975.
[24] J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Springer, New York, 2005.
[25] J. Kaipio, E. Somersalo, Statistical inverse problems: discretization, model reduction and inverse crimes, J. Comput. Appl. Math. 198 (2007) 493–504.
[26] J.P. Kaipio, V. Kolehmainen, E. Somersalo, M. Vauhkonen, Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography, Inverse Problems 16 (2000) 1487–1522.
[27] K. Kurpisz, A.J. Nowak, Inverse Thermal Problems, WIT Press, Southampton, UK, 1995.
[28] J.L. Melsa, D.L. Cohn, Decision and Estimation Theory, McGraw-Hill, 1978.
[29] V.A. Morozov, Methods for Solving Incorrectly Posed Problems, Springer-Verlag, New York, 1984.
[30] D.A. Murio, The Mollification Method and the Numerical Solution of Ill-Posed Problems, Wiley-Interscience, New York, 1993.
[31] G.K. Nicholls, C. Fox, Prior modelling and posterior sampling in impedance imaging, in: A. Mohammad-Djafari (Ed.), Proceedings of SPIE, Bayesian Inference for Inverse Problems, vol. 3459, SPIE, Bellinghan, WA, USA, 1998, pp. 116–127.
[32] M.N. Ozisik, H.R.B. Orlande, Inverse Heat Transfer: Fundamentals and Applications, Taylor and Francis, New York, 2000.
[33] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, 1984.
[34] A.N. Tikhonov, V.Y. Arsenin, Solution of ill-posed problems, Winston & Sons, Washington, DC, 1977.
[35] D.M. Trujillo, H.R. Busby, Practical Inverse Analysis in Engineering, CRC Press, Boca Raton, 1997.
[36] K. Woodbury (Ed.), Inverse Engineering Handbook, CRC Press, Boca Raton, 2002.
[37] J.P. Zubelli, An Introduction to Inverse Problems: Examples, Methods and Questions, Institute of Pure and Applied Mathematics, Rio de Janeiro, Brazil, 1999.