Available online at www.sciencedirect.com

ScienceDirect

International Journal of Heat and Mass Transfer 51 (2008) 2313–2334

International Journal of HEAT and MASS TRANSFER

www.elsevier.com/locate/ijhmt

Estimation of parameters in multi-mode heat transfer problems using Bayesian inference – Effect of noise and a priori

Siddharth Parthasarathy, C. Balaji *

Heat Transfer and Thermal Power Laboratory, Department of Mechanical Engineering, Indian Institute of Technology, Madras, Chennai 600 036, TamilNadu, India

Received 15 May 2007; received in revised form 27 July 2007
Available online 5 November 2007

Abstract

Parameter estimation problems and heat source/flux reconstruction problems are some of the most frequently encountered inverse heat transfer problems. These problems find their application in many areas of science and engineering. The primary focus of this paper is on the heat transfer parameter estimation for a two-dimensional unsteady heat conduction problem with (a) convection boundary condition and (b) convection and radiation boundary condition. The paper demonstrates the effect of a priori model on the performance of the algorithm at different noise levels in the measured data. The inverse problem is solved using three different a priori models namely normal, log normal and uniform. The posterior PDF is sampled using the Metropolis–Hastings sampling algorithm. Both single-parameter estimation and multi-parameter estimation problems are addressed and the effects of corresponding a priori models are studied. It was found that the mean and maximum a posteriori estimates for thermal conductivity and the convection heat transfer coefficient were insensitive to the a priori model at all the considered noise levels for the single-parameter estimation problem. At high noise levels in the two-parameter estimation problem, the estimates for thermal conductivity and convection coefficient were sensitive to the a priori model. It was also found that the standard deviation of the samples was correlated to the error in estimation in the single-parameter estimation case. In three parameter estimation case, alternate solutions to the same problem were retrieved due to a strong correlation between the convection coefficient and the emissivity. However, a more informative a priori model could address this issue.

Keywords: Bayesian inference; A priori model; Parameter estimation; Inverse problems; Noise; Unsteady heat conduction

1. Introduction

Parameter estimation problems and heat source/flux reconstruction problems are some of the most frequently encountered inverse heat transfer problems. These problems find their application in many areas of science [1] and engineering [2]. Generally parameter estimation problems are ill-posed by their very nature and lead to solutions that may not unique and are sensitive to the input data. Inverse problems often demand regularization like the one due to Tikhonov [3] to address the ill-posed nature of the problem.

Estimation methods can broadly be classified as (a) deterministic methods and (b) stochastic methods. Due to the ill-posed nature of these problems, stochastic methods fare better when compared to deterministic methods. Stochastic methods are data driven and the data is collected by solving many cases of the direct problem. In the jargon of inverse problems, the direct problem is referred to as the forward model. The quality of the data collected affects the solution to the inverse problem and the task of data collection that involves repeated forward calculations makes the solution computationally expensive.

Stochastic methods are becoming increasingly relevant in the context of estimation or retrievals with "polluted"

* Corresponding author.
E-mail address: balaji@iitm.ac.in (C. Balaji).

0017-9310/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijheatmasstransfer.2007.08.031

Nomenclature

A	acceptance ratio
B	breadth of the domain, m
C_p	specific heat, J/kg K
E	expectation
F	forward problem
h	convective heat transfer coefficient, W/m^2 K
k	thermal conductivity W/m K
L	length of the domain, m
M	number of samples
N	normal distribution
n	dimension of the measurement vector
P/p	probability distribution function
q	proposal distribution
SD	standard deviation
t	time s
$U(a,b)$	uniform distribution between a and b
u	random sample
$\mathbf{x}$	state vector (parameters)
$\hat{x}$	estimate of x
$\mathbf{Y}$	measurement vector

Greek symbols

α	gamma distribution parameter
β	gamma distribution parameter
Γ	boundary of the domain
δ	dirac delta function
ε	emissivity
v	variance
π	numerical value of $\pi = 3.143$ kg/m^3
ρ	density
σ	standard deviation of the instrument, K
Ω	domain
ω	Gaussian noise

Superscripts

T	transpose
i	ith sample
$*$	sample from proposal distribution

Subscripts

i	ith sample
MAP	maximum a posteriori
MEAN	mean estimate

data. There are many stochastic methods available to solve parameter estimation problems. A review of these techniques is presented in Alifanov [4] and Beck et al. [5]. Genetic Algorithms (GA), Artificial Neural Networks (ANN) and Bayesian Inference are some of the commonly used methods. Most of the stochastic methods including Genetic Algorithms and ANN, model the problem as an optimization problem wherein some sort of least square minimization is done. Bayesian Inference, in philosophy, is different from the above methods.

Bayesian inference is one of the emerging techniques in solving the parameter estimation problems. It works on the Bayes' conditional probability concept. This method is based on using probability to represent all forms of uncertainty in the problem. Bayes' equation is then applied to relate the experimental data $\mathbf{Y}$ and the parameters $\mathbf{x}$ as follows:

$$P(x|Y) = \frac{P(Y|x) \times P(x)}{P(Y)} \tag{1}$$

Here

$P(x|Y)$ is the posterior probability density function (PPDF),
$P(Y|x)$ is the likelihood function,
$P(x)$ is the prior distribution function,
$P(Y)$ is a normalizing constant.

A Bayesian approach to a problem starts with the formulation of a model that best describes the situation of interest, i.e. the forward problem or direct problem. The next important step is to formulate a prior distribution over the unknown parameter to capture the beliefs about the situation before seeing the data. Lastly, Bayes' rule is applied to obtain the posterior distribution that relates the observed data $\mathbf{Y}$, the parameters $\mathbf{x}$ and the prior beliefs of the parameters. From the posterior distribution, the parameters of interest can be estimated in many ways and the most commonly used estimates are (a) maximum a posteriori (MAP) and (b) the mean estimate

$$\hat{x}_{\mathrm{MAP}} = \arg\max_x(P(x|Y)) \tag{2}$$

and

$$\hat{x}_{\mathrm{MEAN}} = E(P(x|Y)) \tag{3}$$

This theoretically simple process comes with two big challenges, (a) the prior specification and (b) computation. The computational complexity of the problem can be tackled in part by the Markov Chain Monte Carlo (MCMC) techniques but many are still uncomfortable in use of this method, often because they view the selection of prior as being arbitrary and subjective [6]. The difficulty that is discussed above is outside of the difficulty associated with solving the forward problem, that could in principle be from any branch of science and engineering.

This work applies the Bayesian paradigm for solving an inverse heat transfer problem to investigate the effect of a priori selection on the performance of the algorithm when applied to heat transfer problems, at different noise levels in the measured data (**Y**).

2. Bayesian inference and inverse heat transfer problems

Bayesian inference is one of the relatively new methods employed in solving inverse heat transfer problems and the applicability of the method to inverse heat transfer problems can be attributed to developments in computational techniques such as Markov Chain Monte Carlo (MCMC) and their related sampling algorithms. Modeling is one of the important steps in Bayesian inference. The modeling step starts with modeling the current situation (Likelihood function) and is followed by the a priori modeling.

2.1. Modeling the heat transfer problem

Data collection is the foremost step in Bayesian inference. The modeling step starts with the modeling of the data collected and the conditions in which it was collected. In most of the inverse heat transfer problems, the data collected is in the form of temperature. Modeling the temperature data is relatively an easy task, as the uncertainty in the measurement of temperature, i.e. noise, can be easily modeled as a Gaussian distribution. Consider **Y** to be a vector of measured temperatures at different spatial positions and at different times. The measured temperatures y_i are nothing but the calculated temperatures plus a Gaussian random noise ω, i.e.

$$Y = F(x) + \omega \tag{4}$$

$F(x)$ is the solution to the forward problem with the given state of parameters $\mathbf{x}$. Since $F(x)$ is not directly invertible, estimation of $\mathbf{x}$ given the set of temperatures $\mathbf{Y}$ is not an easy task. The very nature of the problem is ill-posed. As one could observe from Eq. (4), $(Y - F(x))$ has the same distribution of $\omega \cdot \omega$ is a random sample from $N(0, \boldsymbol{\sigma})$, where σ is the standard deviation of the measuring instrument and N is normal distribution function.

Hence

$$P(Y|x) = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^n} \exp\left(-\frac{(Y - F(x))^{\mathrm{T}}(Y - F(x))}{2\sigma^2}\right) \tag{5}$$

where n is the dimension of the vector **Y**. In the above equation, the term in the numerator of the exponential function is the sum of the squares of the differences between the measured temperatures and the calculated temperatures. It is easy to infer from the above expression that least square minimization is in-built in Bayesian inference.

2.2. Modeling the a priori

In a priori modeling the attempt is to capture the prior knowledge of the parameters. Prior information could be in the form of bounds for parameters, distribution and so on. In a few cases, the distribution of the parameter is known and the interest in the problem might be the precise estimation of the parameter. In such cases the a priori model is the known distribution and in the parlance of Bayesian statistics is known as an informative prior. However, in many problems this information is not available.

Inverse heat transfer problems mainly fall into two main categories: (a) Parameter estimation as for example, estimation of thermal conductivity of a material, emissivity of a surface and so on. (b) Heat source/flux estimation. In heat flux estimation problems, the heat flux is discretized both spatially and temporally and this forms a field. The heat source is approximated as a linear combination of weights at these nodes (both spatial and temporal) and the basis functions which are similar to shape functions in finite element analysis. This field is modeled using the concepts of Markov Random Field (MRF) and the weights are estimated. For a further insight, one may refer to the article by Wang and Zabaras [7]. In most cases of the parameter estimation problem, prior information is not readily available. In order to handle these problems, non-informative priors are employed. By employing these priors, new parameters related to the prior distribution come into the picture. These are termed hyper parameters. For example, if a Gaussian distribution with unknown mean and variance is employed, the mean and the variance of the distribution become the hyper parameters which also form a part of the retrieval using Bayesian inference.

2.3. Sampling

In most of the cases, the PPDF is of a non-standard form, is non-linear or has an implicit likelihood [7]. Due to the above reasons, numerical sampling becomes necessary. Markov Chain Monte Carlo (MCMC) is one of the most powerful and popular sampling techniques. A good review of the available sampling techniques is presented in [8,9]. The idea here is to draw M identical independently distributed (i.i.d) samples $\{x^i\}$ $i = 1, \ldots, M$ and approximate the PPDF from these samples as follows:

$$P(x|Y) = \frac{1}{M}\sum_{i=1}^{M} \delta(x - x^i) \tag{6}$$

In order to draw these samples from a rather complex PPDF, various sampling algorithms are employed. Metropolis–Hastings (MH) sampling algorithm and Gibbs sampling algorithm are the widely used algorithms to solve the parameter estimation problem and the heat source/flux estimation problem respectively. Hybrid Monte Carlo sampler [10], a relatively less popular sampler, is also gaining momentum in the field of parameter estimation.

2.3.1. Metropolis–Hastings sampling algorithm

Both Metropolis–Hastings and Gibbs sampling algorithms are discussed in several references (see for example [9,11]). The algorithm can be summarized as follows: For a one-dimensional **x**, the algorithm is as follows:

1. Initialize x^1
2. for $i = 1, \dots, M$
   a. Draw a sample $u \sim U(0,1)$ i.e. from a uniform distribution between 0, 1.
   b. Draw a sample $x^* \sim q(x^*|x^i)$
   c. If $u < A(x^*, x^i) x^{i+1} = x^*$
   d. else $x^{i+1} = x^i$

In the above algorithm, q is called the proposal distribution (easy-to-sample) and A is defined as follows:

$$A(x^*, x^i) = \min\left\{1, \frac{p(x^*) \cdot q(x^i|x^*)}{p(x^i) \cdot q(x^*|x^i)}\right\} \tag{7}$$

A is called acceptance. M is the number of samples. One must make a note that in Eq. (7), $p(x)$ refers to the function to be explored i.e. target function and not the prior. In the cases considered, the target function is the PPDF. The above algorithm converges to the target function for any reasonable proposal distribution though the speed of convergence itself varies. The most commonly used proposal distribution is $q(x^*|x^i) \sim N(x^i, \sigma)$ where σ is 5% of x^i. One must also observe that the method is computationally expensive as in every iteration, the forward sample is solved to determine the next sample.

In case of problems where the parameter vector is multidimensional, a certain variant of the above scheme is used and is as follows

1. Initialize $x^1 = \{x_1^1, x_2^1, \dots, x_n^1\}$
2. for $i = 1, \dots, M$
   a. for $j = 1, \dots, n$
      i. Draw a sample $u \sim U(0,1)$
      ii. Draw a sample $x^* \sim q(x^*|x_{-j}^{i+1}, x_j^i)$
      iii. If $u < A(x^*, x_j^i)$ then $x_j^{i+1} = x^*$
      iv. Else $x_j^{i+1} = x_j^i$

where n is the dimension of the parameter vector and

$$x_{-j}^{i+1} = \{x_1^{i+1}, x_2^{i+1}, \dots, x_{j-1}^{i+1}, x_{j+1}^{i}, \dots, x_n^i\}^{\mathrm{T}}$$

and

$$A(x^*, x_j^i) = \min\left(1, \frac{p(x^*|x_{-j}^{i+1}) \cdot q(x_j^i|x_j^*, x_{-j}^{i+1})}{p(x_j^i|x_{-j}^{i+1}) \cdot q(x^*|x_j^i, x_{-j}^{i+1})}\right) \tag{8}$$

and $q(x^*|x_j^i, x_{-j}^{i+1}) \sim N(x_j^i, \sigma_{x_j}^2)$, $\sigma_{x_j}^2$ is 5% of the proposal mean x_j^i.

2.3.2. Gibbs sampling algorithm

This algorithm is used when the dimension of the parameter vector is high. This is mainly used in heat source/flux estimation problems. Wang and Zabaras [7] demonstrate its application in a heat transfer problem. The algorithm can be summarized as follows:

1. Initialize $x^1 = \{x_1^1, x_2^1, \dots, x_n^1\}$
2. For $i = 1, \dots, M$
   a. $x_1^{i+1} \sim p(x_1|x_2^i, x_3^i, \dots, x_n^i)$
   b. $x_2^{i+1} \sim p(x_2|x_1^{i+1}, x_3^i, \dots, x_n^i)$
   ⋮
   n. $x_n^{i+1} \sim p(x_n|x_1^{i+1}, x_2^{i+1}, x_3^{i+1}, \dots, x_{n-1}^{i+1})$

In this algorithm the acceptance is always 1. This is an important feature, as this leads to quick convergence.

2.4. Bayesian inference using Metropolis–Hastings sampler

Bayesian inference and MCMC sampling supplement each other and the combination results in a powerful tool. The algorithm for a single parameter estimation (say x) problem using Bayesian inference and MH sampler is presented in the flow chart (Fig. 1).

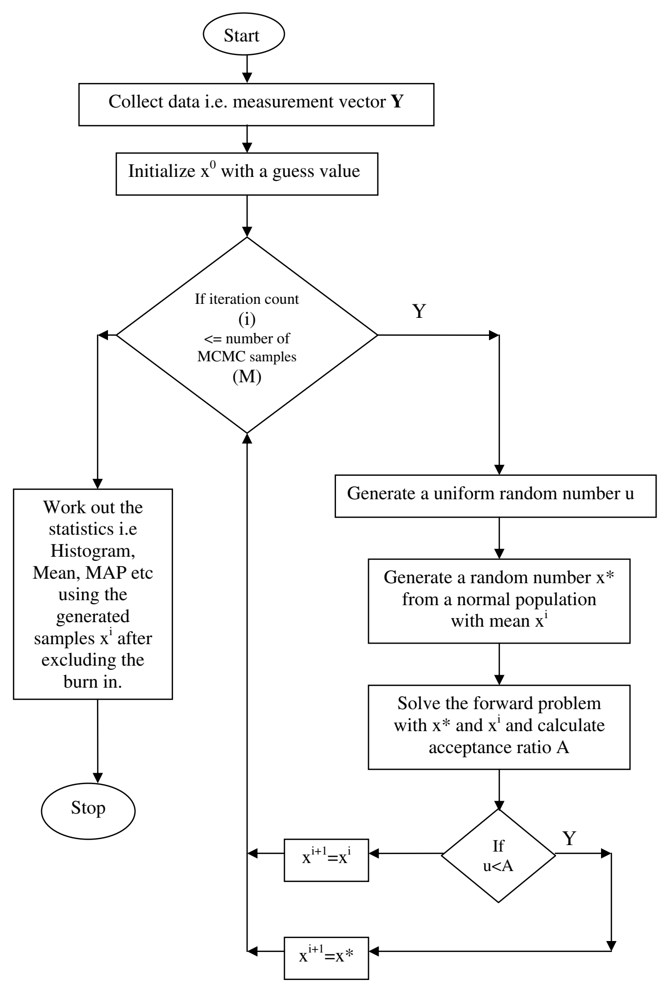


Fig. 1. Flow chart of single-parameter estimation problem using Bayesian inference and MH sampling scheme.

3. Parameter estimation

In the current study, a typical two-dimensional transient heat conduction problem (Fig. 2) is chosen and the performance of the algorithm is studied under different conditions. The left wall of the two-dimensional slab is hot and isothermal at T_w. The other three surfaces are bathed by a fluid at T_∞ that gives rise to convective heat transfer out of the domain. The governing equations of the forward problem are

$$\rho C_p \frac{\partial T}{\partial t} = \nabla(k \cdot \nabla T) \quad (9)$$

$$T(x,y,0) = T_I \quad (10)$$

$$-k \cdot \frac{\partial T}{\partial y} = h \cdot (T - T_\infty) \quad \forall (x,y) \in (\Gamma_1 \cup \Gamma_3) \quad (11)$$

$$-k \cdot \frac{\partial T}{\partial x} = h \cdot (T - T_\infty) \quad \forall (x,y) \in \Gamma_2 \quad (12)$$

$$T(x,y,t) = T_w \quad \forall (x,y) \in \Gamma_4 \quad (13)$$

$$T_I = 573\ K$$

$$T_w = 573\ K$$

$$T_\infty = 298\ K$$

The outline of the current study is as follows:

1. Single parameter estimation problem
   a. Thermal conductivity (k) ,
   b. Convective heat transfer coefficient (h).
2. Two-parameter estimation problem – (k, h),
3. Three-parameter estimation problem – (k, h, ε), where ε is the emissivity.

3.1. Single parameter estimation in 2D transient heat conduction

Single parameter estimation is carried out for two different sub-cases.

3.1.1. Estimation of thermal conductivity in 2D transient heat conduction

Thermal conductivity estimation is an important inverse heat transfer problem and is often encountered in material recognition experiments. These problems are becoming more common with the advancements in material science. In this study, the material properties chosen are similar to that of steel and the thermal conductivity of the same is retrieved using Bayesian inference. In this case, a value of $h = 10\ \text{W/m}^2\ \text{K}$, $k = 25\ \text{W/m K}$ was used to construct the data vector **Y** and the value of k was estimated using Bayesian inference. The problem specifications are as follows:

Length, L 0.1 m;
Breadth, B 0.1 m;
Observation time 50 s;
Thermal conductivity, k to be estimated;
Convection coefficient, h 10 W/m^2 K;
Density 7850 kg/m^3;
Specific heat 460 J/kg K.

Grid independence studies (Fig. 3) were carried out to arrive at the right grid size. A 31×31 grid was found to be the optimum. The time discretization was 0.1 s. An explicit finite difference scheme was used to solve the forward problem. Temperature data was collected with $k = 25\ \text{W/m K}$ at points shown in Fig. 2 and at times $t = 10$, 20, 30, 40, 50 s (the dimension of the **Y** matrix is 45) and this value of k was retrieved. The data collected (**Y**) was corrupted with Gaussian noise of $\sigma = \{0.0\ \text{K}, 0.1\ \text{K}, 0.5\ \text{K}, 1.0\ \text{K}\}$ and at each noise level, the inverse problem was solved using three different a priori distributions, namely normal, log normal and uniform. Normal and uniform priors are commonly used in Bayesian inference while the rationale behind the use of a log normal prior is to investigate the effect of such a non-obvious choice on the performance of the algorithm. The PPDFs

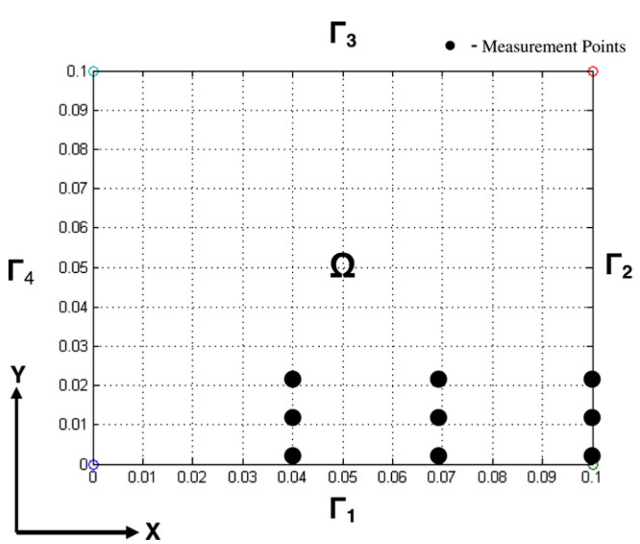


Fig. 2. Problem geometry, Cases 3.1 and 3.2.

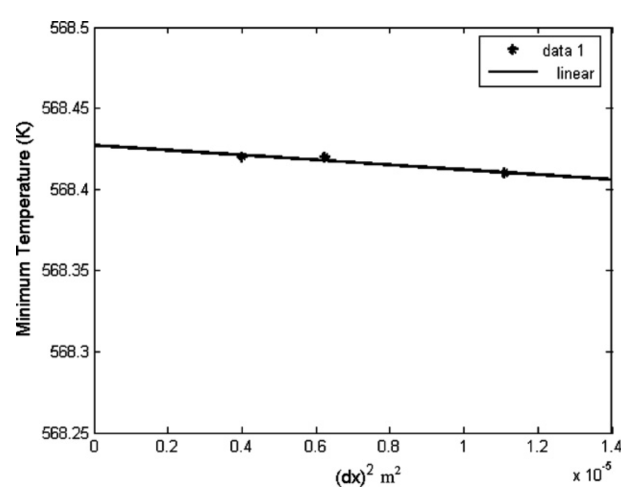


Fig. 3. Grid independence check.

for each of the above priors are constructed following the hierarchical Bayesian models suggested in [11] and are listed below.

a. Normal

$$P(k,\mu_k,\upsilon_k|Y) \propto \exp\left(-\frac{(Y-F(k))^{\mathrm{T}}(Y-F(k))}{2\sigma^2}\right) \cdot \upsilon_k^{-0.5}\exp\left(-\frac{(k-\mu_k)^2}{2\upsilon_k}\right)\cdot\upsilon_k^{-(1+\alpha)} \cdot \exp(-\beta\cdot\upsilon_k^{-1}) \quad (14)$$

b. Log normal

$$P(k,\mu_{lk},\upsilon_{lk}|Y) \propto \exp\left(-\frac{(Y-F(k))^{\mathrm{T}}(Y-F(k))}{2\sigma^2}\right) \cdot \upsilon_{lk}^{-0.5}\exp\left(-\frac{(\log(k)-\mu_{lk})^2}{2\upsilon_{lk}}\right) \cdot \upsilon_{lk}^{-(1+\alpha)}\cdot\exp(-\beta\cdot\upsilon_{lk}^{-1}) \quad (15)$$

c. Uniform

$$P(k|Y) \propto \exp\left(-\frac{(Y-F(k))^{\mathrm{T}}(Y-F(k))}{2\sigma^2}\right) \quad (16)$$

for $0 < k < k_{\max}$, where $k_{\max}$ can be an arbitrarily large value.

Here $\mu_k, \upsilon_k, \mu_{lk}, \upsilon_{lk}$ are hyper parameters. The values of α and β chosen are $\alpha = \beta = 0.001$ following [11].

Eqs. (14) and (15) are basically of the form $P(Y|k) \times (P(k|\mu, v) \times P(v))$ where the first term is the likelihood function and the second (in bracket) is the prior distribution. The distributions for $P(k|\mu, v)$ is the prior distribution of k given the hyper parameters, i.e. μ, v, while $P(v)$ is the distribution for the variance. The prior distributions are conjugate. In the case of uniform distribution since the distribution is a constant, no prior terms feature in Eq. (16). The constant is taken care of by the proportionality.

The results from this study are presented in Tables 1 and 3 and Fig. 4. The spread of the samples increased as noise in measurement increased. Point estimates (i.e. statistics) of the posterior probability distribution function did not depend on the a priori model at all considered noise levels.

It is evident that for noise levels around 0.5 K which translates to a measurement error of 1.5 K i.e. $\pm 3\sigma$, there is a 10% error in value of k estimated both by the mean and the MAP. Hence, when working with non-informative priors i.e. minimal prior information, it is imperative that the instrument error is reduced to the extent possible inorder to improve the accuracy of the estimate. While this may seem pretty obvious, it is heartening to note that for noise levels up to 0.3 K or a measurement error of 1 K, the retrieval is remarkably accurate regardless of the a priori used. The relatively weak dependence of the estimates on the priors also subdues substantially the general criticism associated with retrievals using hierarchical Bayesian models. Ghosh and Samanta [12,13] present an excellent discussion on this topic.

Table 1
Thermal conductivity estimates for different priors at various noise levels

A priori	Mean estimate (W/m K)	MAP estimate (W/m K)	SD (W/m K)
Noise level 0.0 K			
Normal	25.08	25.28	0.42
Log normal	25.00	24.98	0.44
Uniform	25.01	25.03	0.44
Noise level 0.1 K			
Normal	24.59	24.58	0.42
Log normal	24.58	24.54	0.43
Uniform	24.58	24.52	0.43
Noise level 0.5 K			
Normal	22.49	22.29	2.02
Log normal	22.01	21.39	1.80
Uniform	22.24	22.16	1.83
Noise level 1.0 K			
Normal	20.62	20.39	2.92
Log normal	20.74	19.19	3.46
Uniform	20.80	19.74	3.37

Table 2
Convection coefficient estimates for different priors at various noise levels

A priori	Mean estimate (W/m^2 K)	MAP estimate (W/m^2 K)	SD (W/m^2 K)
Noise level 0.0 K			
Normal	10.00	10.00	0.08
Log normal	10.00	9.99	0.08
Uniform	10.00	10.00	0.08
Noise level 0.1 K			
Normal	10.09	10.07	0.08
Log normal	10.09	10.08	0.07
Uniform	10.09	10.07	0.08
Noise level 0.5 K			
Normal	10.31	10.29	0.38
Log normal	10.32	10.21	0.38
Uniform	10.32	10.35	0.37
Noise level 1.0 K			
Normal	11.52	11.93	0.71
Log normal	11.48	11.44	0.76
Uniform	11.51	11.49	0.76

3.1.2. Estimation of convection heat transfer coefficient in 2D transient heat conduction

Even in this case, the properties of the material properties used were similar to that of steel. In this case, value of k was fixed and the convection coefficient was retrieved. A value of $h = 10\ \mathrm{W/m^2\,K}$, $k = 25\ \mathrm{W/m\,K}$ was used to construct data vector $\mathbf{Y}$ and the value of the convection

Table 3
Posterior probability density functions in the thermal conductivity estimation case (Case 3.1.1)

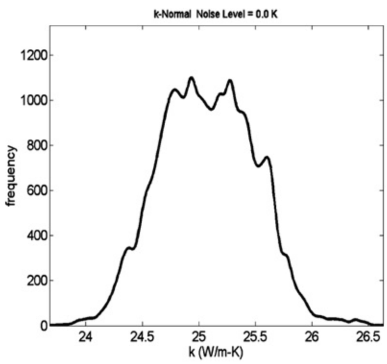
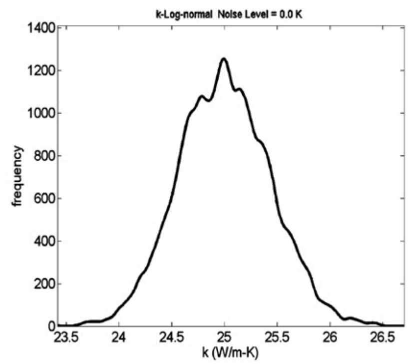
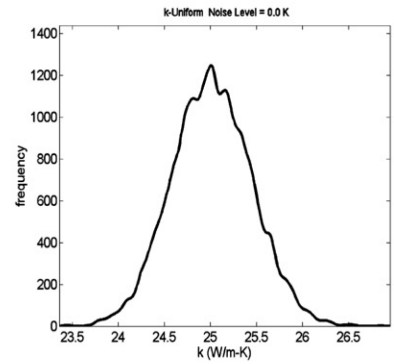
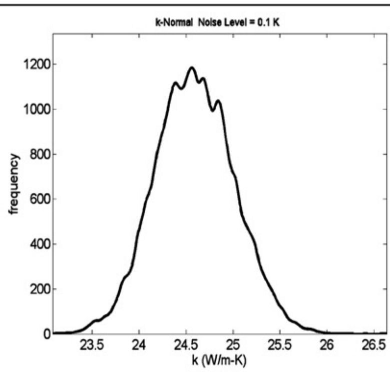
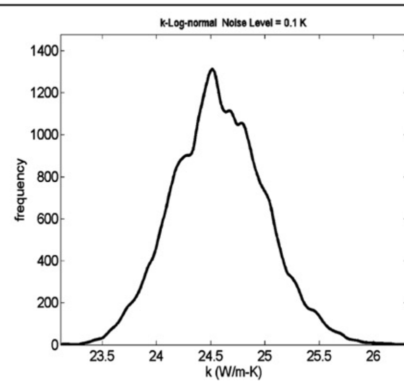
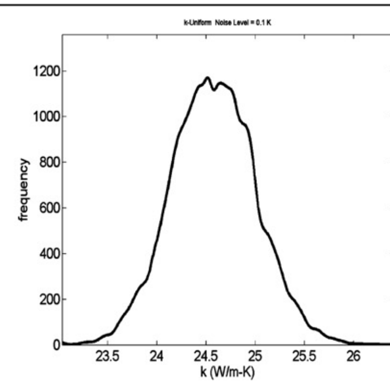
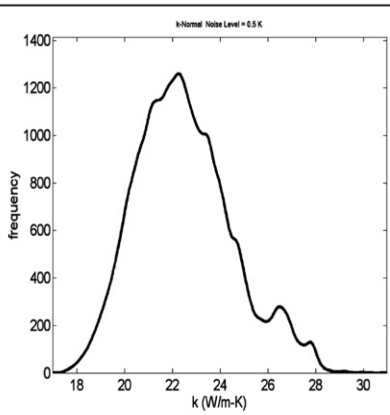
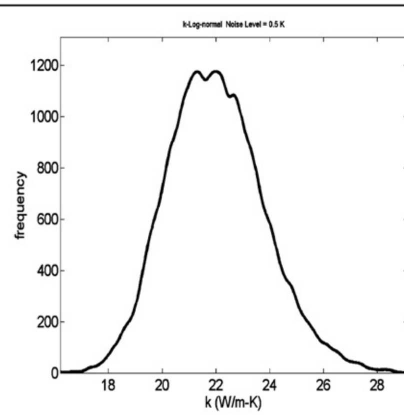
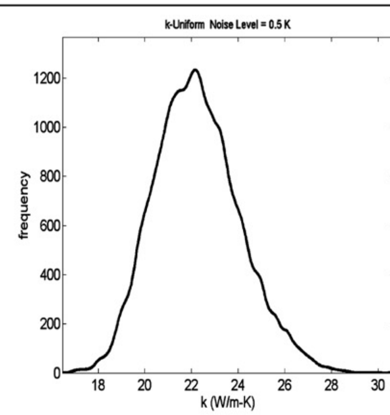
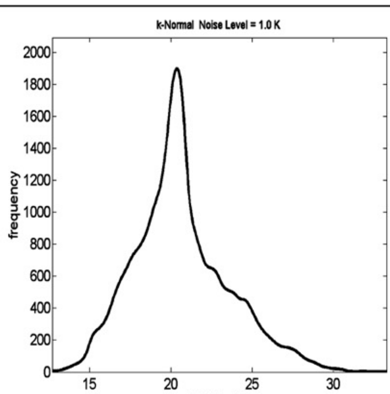
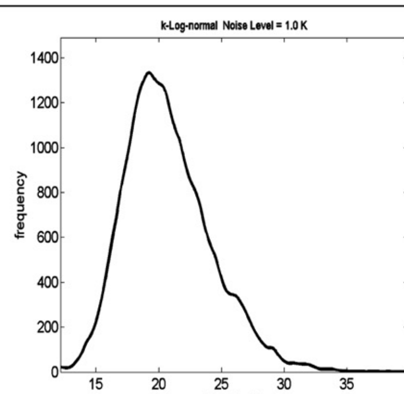
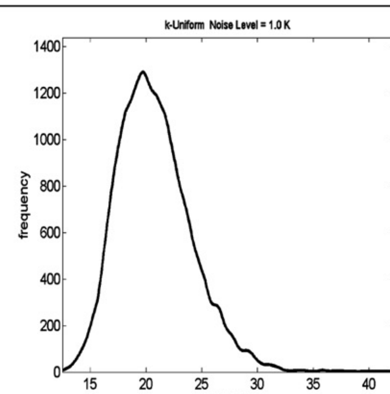
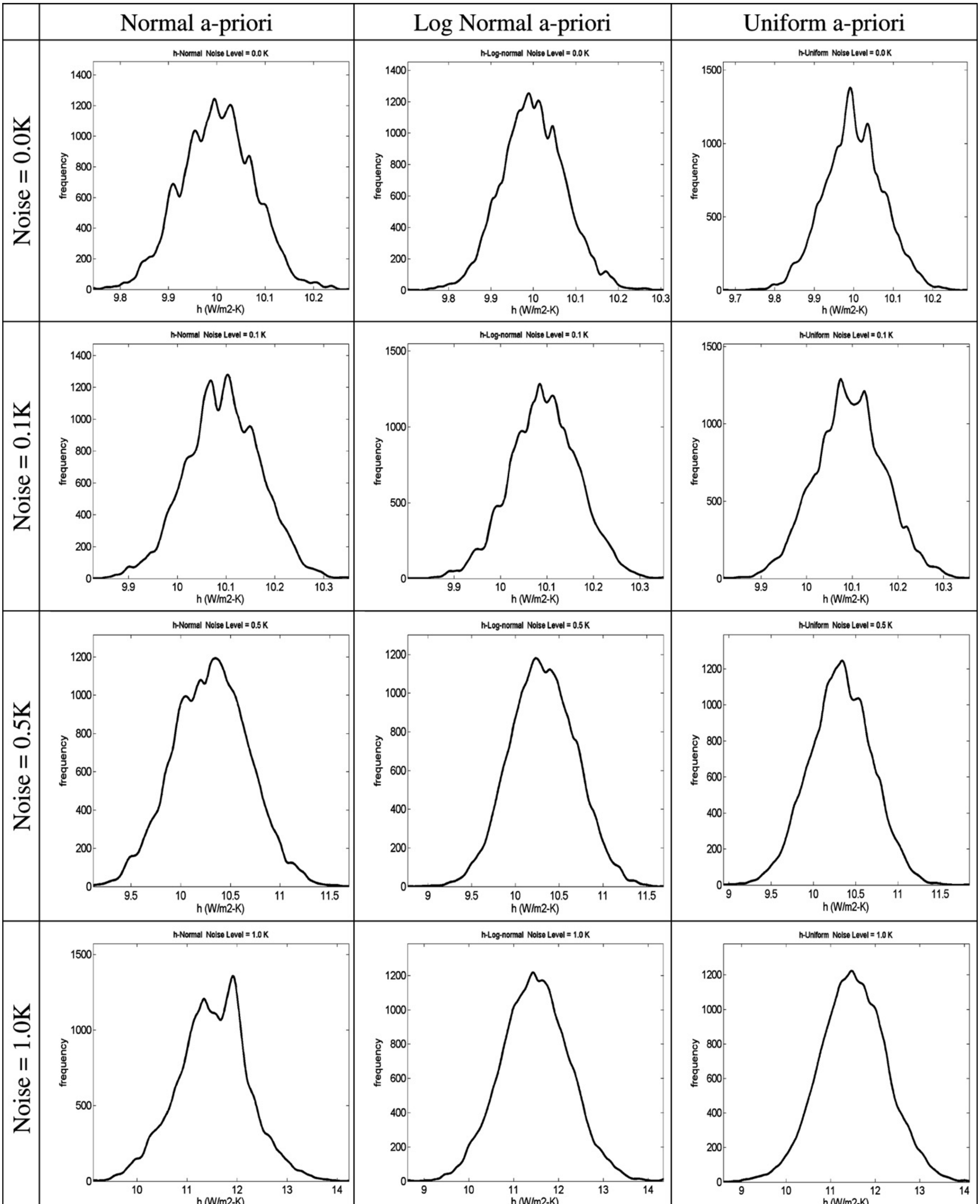
	Normal a-priori	Log Normal a-priori	Uniform a-priori
Noise = 0.0K			
Noise = 0.1K			
Noise = 0.5K			
Noise = 1.0K			

Table 4
Posterior probability density functions in convection coefficient estimation case (Case 3.1.2)



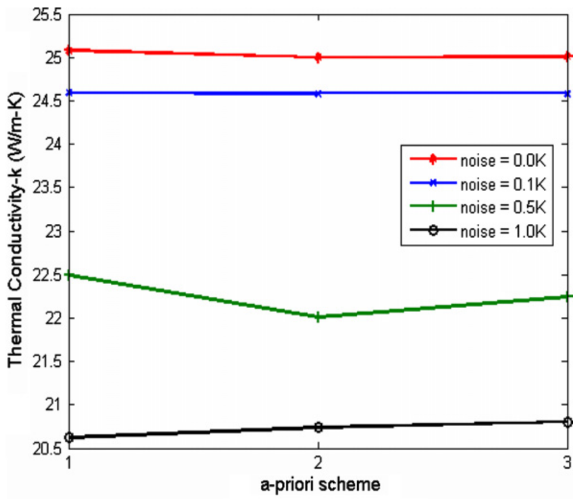


Fig. 4. Effect of a priori scheme on mean thermal conductivity estimate in Case 3.1.1. Scheme 1 – normal, Scheme 2 – log normal, Scheme 3 – uniform.

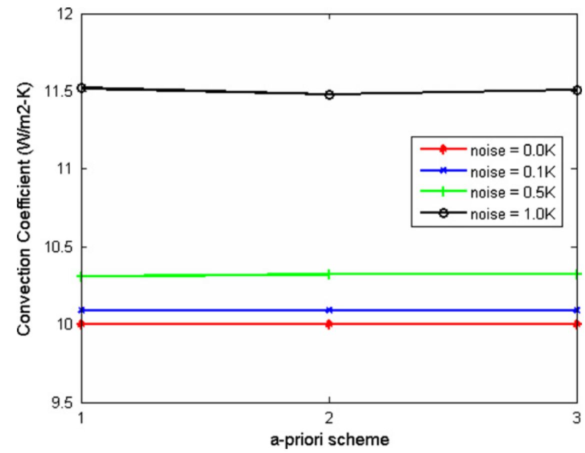


Fig. 5. Effect of a priori scheme on mean convection coefficient estimate, Case 3.1.2. Scheme 1 – normal, Scheme 2 – log normal, Scheme 3 – uniform.

coefficient was estimated using Bayesian inference keeping k at 25 W/m K. The problem specifications are as follows:

Length, L 0.1 m;
Breadth, B 0.1 m;
Observation time 50 s;
Thermal conductivity, k 25 W/m K;
Convection coefficient, h to be estimated;
Density 7850 kg/m^3;
Specific heat 460 J/kg K.

As in the previous case, a 31×31 uniform mesh was used and the time discretization was 0.1 s. An explicit finite difference scheme was used to solve the problem and temperature data was collected with $h = 10$ W/m^2 K. The data collected was corrupted with noise (Gaussian i.i.d) values $\sigma = \{0.0\text{ K},0.1\text{ K},0.5\text{ K},1.0\text{ K}\}$ and at each noise level the inverse problem was solved by Bayesian inference using three different a priori distributions namely normal, log normal and uniform. The PPDFs are similar to that in the previous case with k replaced by h.

The results from this study are presented in Tables 2 and 4 and Fig. 5. The estimates for the convection coefficient are more accurate when compared to thermal conductivity. Even at a noise level of 1 K, i.e. measurement error of 3 K the error in the estimate, both by the MAP and the mean scheme, is around 10%. This is possibly due to the more constrained nature of problem by the convention coefficient; the performance of the algorithm was independent of the a priori model. The discussion pertaining to the thermal conductivity estimation case, in respect of the retrievals, holds for this case as well.

Another interesting trend was observed between the SD/MEAN ratio of the samples and the accuracy of estimation in the thermal conductivity estimation and convective heat transfer coefficient estimation case. A second degree curve was fitted in both cases with an R^2 value of 0.98 (Fig. 6) and

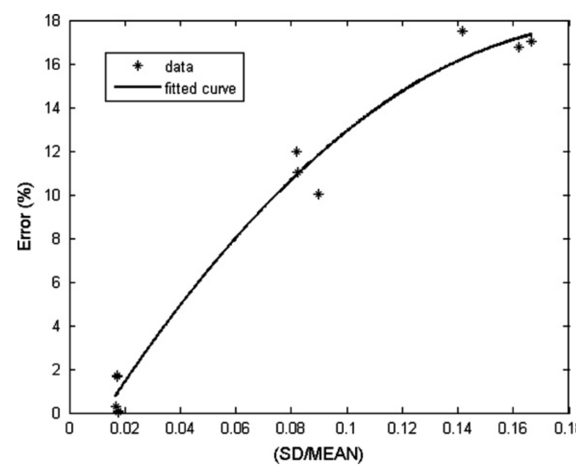


Fig. 6. Relation between SD and error, $R^2 = 0.98$ (Case 3.1.1).

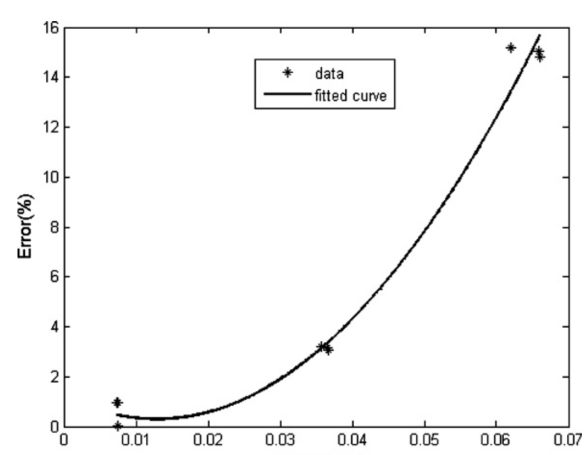


Fig. 7. Relation between SD and error, $R^2 = 0.99$ (Case 3.1.2).

Table 5
Thermal conductivity estimates and convection coefficient estimates in the two-parameter estimation problem for various prior models at different noise levels (Case 3.2)

A priori of k	A priori of h	Estimates of k (W/m K)			Estimates of h (W/m^2 K)		
		Mean	MAP	SD	Mean	MAP	SD
Noise level 0.0 K							
Normal	Normal	25.01	24.56	1.20	10.00	10.00	0.21
Normal	Log normal	24.99	25.04	1.20	9.99	10.03	0.21
Normal	Uniform	25.19	25.02	1.18	10.03	10.00	0.21
Log normal	Normal	25.09	25.04	1.24	10.01	10.04	0.22
Log normal	Log normal	25.06	24.83	1.21	10.01	10.02	0.21
Log normal	Uniform	25.22	25.71	1.24	10.03	9.96	0.21
Uniform	Normal	25.08	24.72	1.27	10.01	9.90	0.22
Uniform	Log normal	25.08	24.54	1.21	10.01	9.93	0.21
Uniform	Uniform	25.02	24.78	1.20	10.00	10.08	0.21
Noise level 0.1 K							
Normal	Normal	25.15	25.98	0.96	9.98	10.10	0.17
Normal	Log normal	25.68	25.93	1.26	10.07	10.09	0.21
Normal	Uniform	25.48	25.13	1.24	10.03	10.02	0.21
Log normal	Normal	25.42	25.83	1.28	10.02	10.09	0.22
Log normal	Log normal	25.54	25.23	1.27	10.04	10.15	0.22
Log normal	Uniform	25.57	25.11	1.25	10.05	10.02	0.21
Uniform	Normal	25.81	25.89	1.25	10.09	10.14	0.21
Uniform	Log normal	25.70	25.84	1.24	10.07	10.09	0.21
Uniform	Uniform	25.75	25.36	1.25	10.08	10.05	0.21
Noise level 0.5 K							
Normal	Normal	24.63	23.06	5.36	9.76	9.54	0.92
Normal	Log normal	20.61	19.65	2.33	9.11	8.78	0.54
Normal	Uniform	26.21	23.29	6.55	9.98	9.66	1.05
Log normal	Normal	26.70	27.55	5.69	10.13	10.02	0.91
Log normal	Log normal	24.45	21.49	5.60	9.72	9.59	0.96
Log normal	Uniform	25.07	22.81	6.00	9.82	9.96	1.01
Uniform	Normal	25.54	23.78	6.39	9.87	9.78	1.05
Uniform	Log normal	25.76	25.08	6.13	9.92	9.79	1.02
Uniform	Uniform	26.60	24.72	6.12	10.07	9.67	1.01
Noise level 1.0 K							
Normal	Normal	31.13	45.37	11.05	10.58	10.25	1.76
Normal	Log normal	42.84	40.52	16.97	11.89	11.99	2.06
Normal	Uniform	33.02	26.97	11.75	10.84	10.05	1.74
Log normal	Normal	36.36	21.74	18.33	11.19	10.43	2.37
Log normal	Log normal	27.21	20.19	13.38	9.84	9.62	2.11
Log normal	Uniform	30.00	13.35	16.53	10.22	10.38	2.30
Uniform	Normal	35.19	26.21	13.77	11.08	9.88	1.65
Uniform	Log normal	34.61	22.07	16.38	10.85	9.48	2.11
Uniform	Uniform	33.17	24.00	14.05	10.79	9.85	1.80

0.99 (Fig. 7), respectively. The standard deviation of the sample indicates the spread of the sample and hence in a way it confirms the accuracy of the estimate.

3.2. Two-parameter estimation problem in 2D transient heat conduction

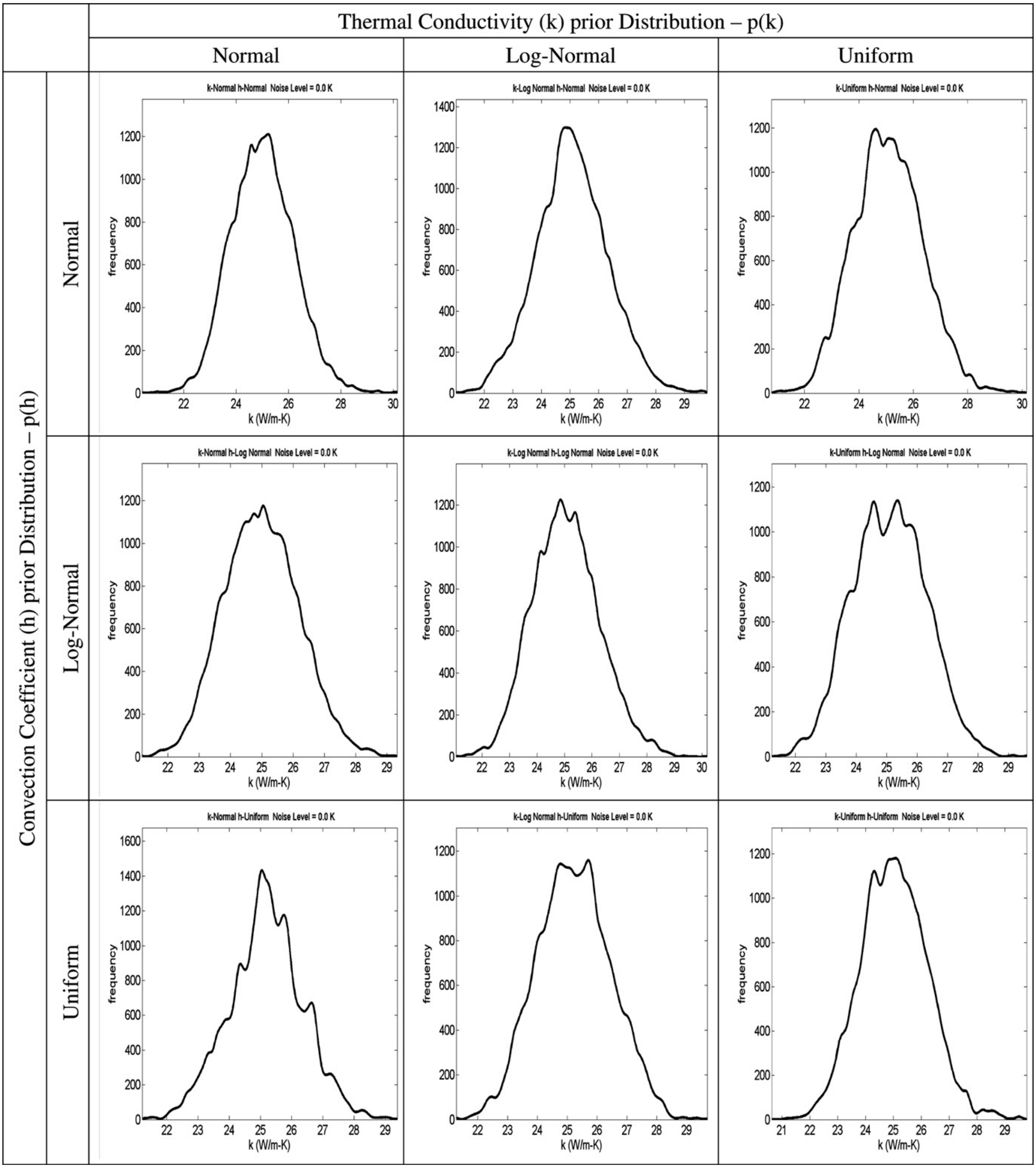
In this problem, the data vector **Y** is constructed by solving the forward problem with $k = 25$ W/m K and $h = 10$ W/m^2 K and corrupting them with Gaussian i.i.d noise. The noise levels chosen are same as in the previous case, $\sigma = \{0.0\text{ K}, 0.1\text{ K}, 0.5\text{ K}, 1.0\text{ K}\}$. The problem is solved using three different priors (normal, log normal and uniform) for k and h, respectively. In other words, nine different combinations of prior selection at each noise level. The specifications of the problem are as follows:

Length, L 0.1 m;
Breadth, B 0.1 m;
Observation time 50 s;
Thermal conductivity, k to be estimated;
Convection coefficient, h to be estimated;
Density 7850 kg/m^3;
Specific heat 460 J/kg K.

The PPDF is of the form

$$P(k, \mu_k, \upsilon_k, h, \mu_h, \upsilon_h|Y) \propto \exp\left(-\frac{(Y - F(k,h))^{\mathrm{T}}(Y - F(k,h))}{2\sigma^2}\right) \cdot f_1(k, \mu_k, \upsilon_k) \cdot f_2(h, \mu_h, \upsilon_h) \quad (17)$$

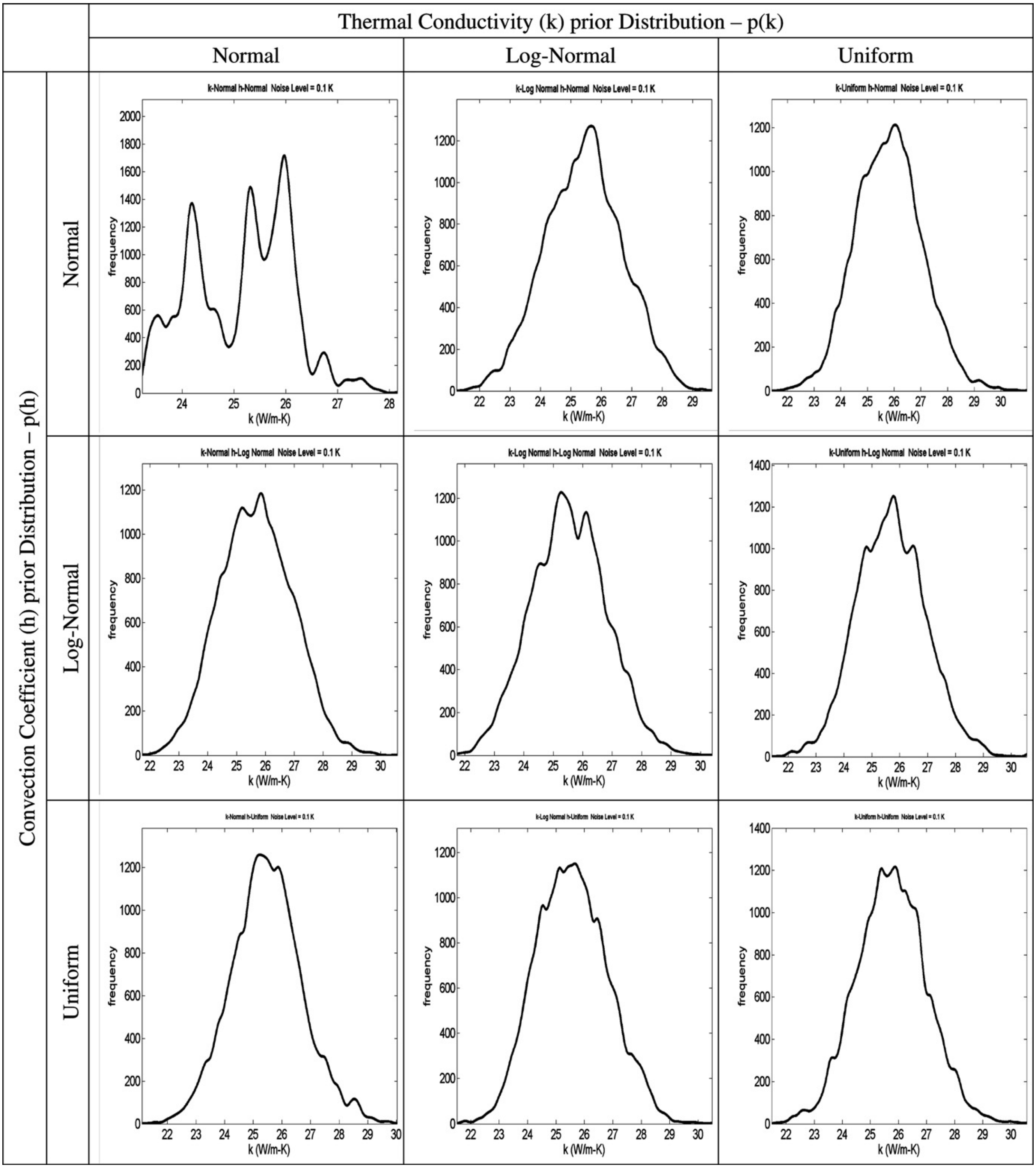
Table 6
Marginal PPDFs for thermal conductivity $P(k|Y)$ at noise level: 0.0 K, Case 3.2



where f_1 and f_2 assume different forms depending on the prior selection and are presented in Table 14. The grid and the scheme used for data collection are same as in the previous case.

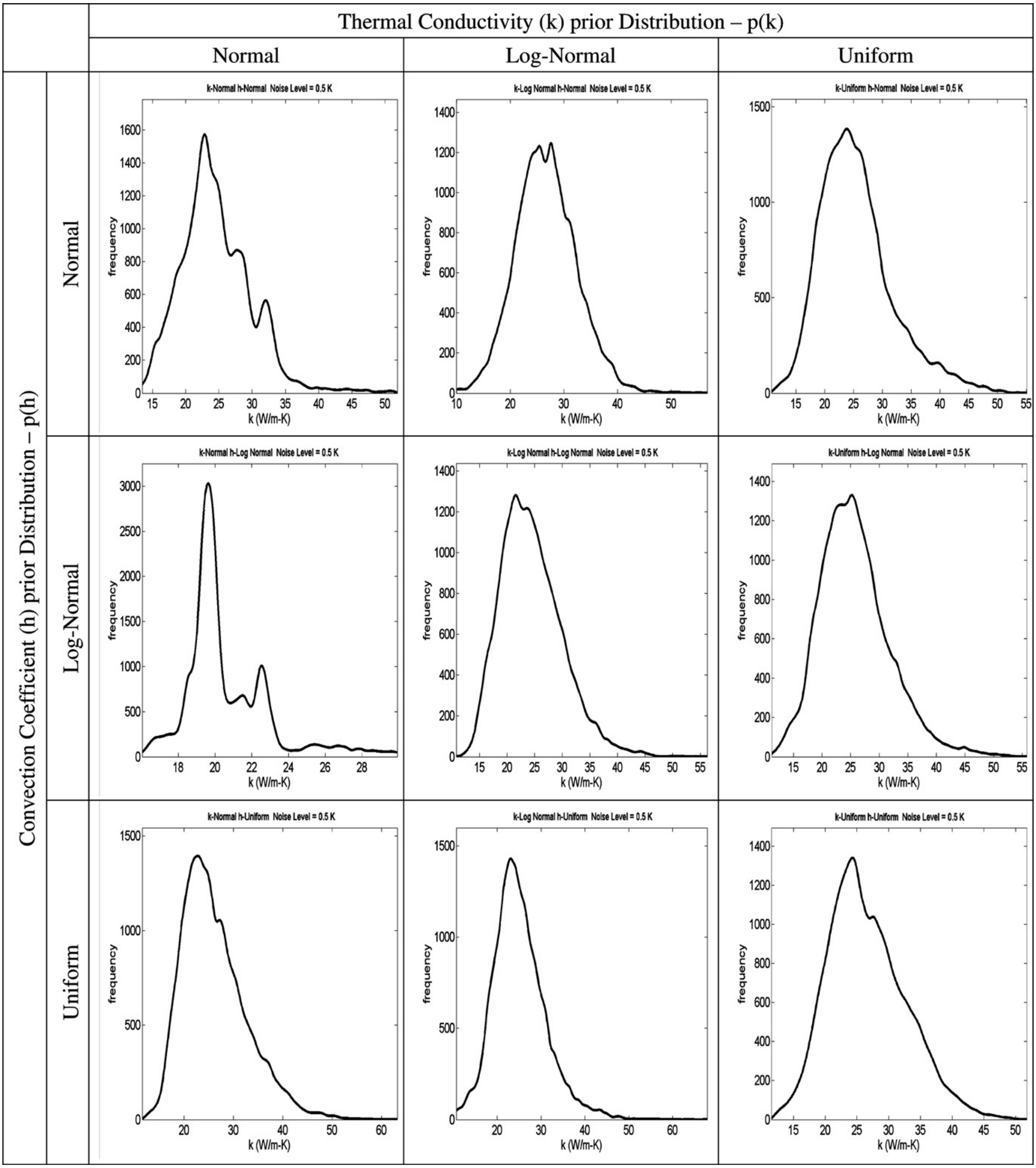
The results of the study are presented in Tables 5–13. While at low noise levels of 0.1 K the estimates were pretty accurate, as the noise levels increased, the estimates became increasingly inaccurate, as expected intuitively.

Table 7
Marginal PPDFs for thermal conductivity $P(k|Y)$ at noise level: 0.1 K, Case 3.2



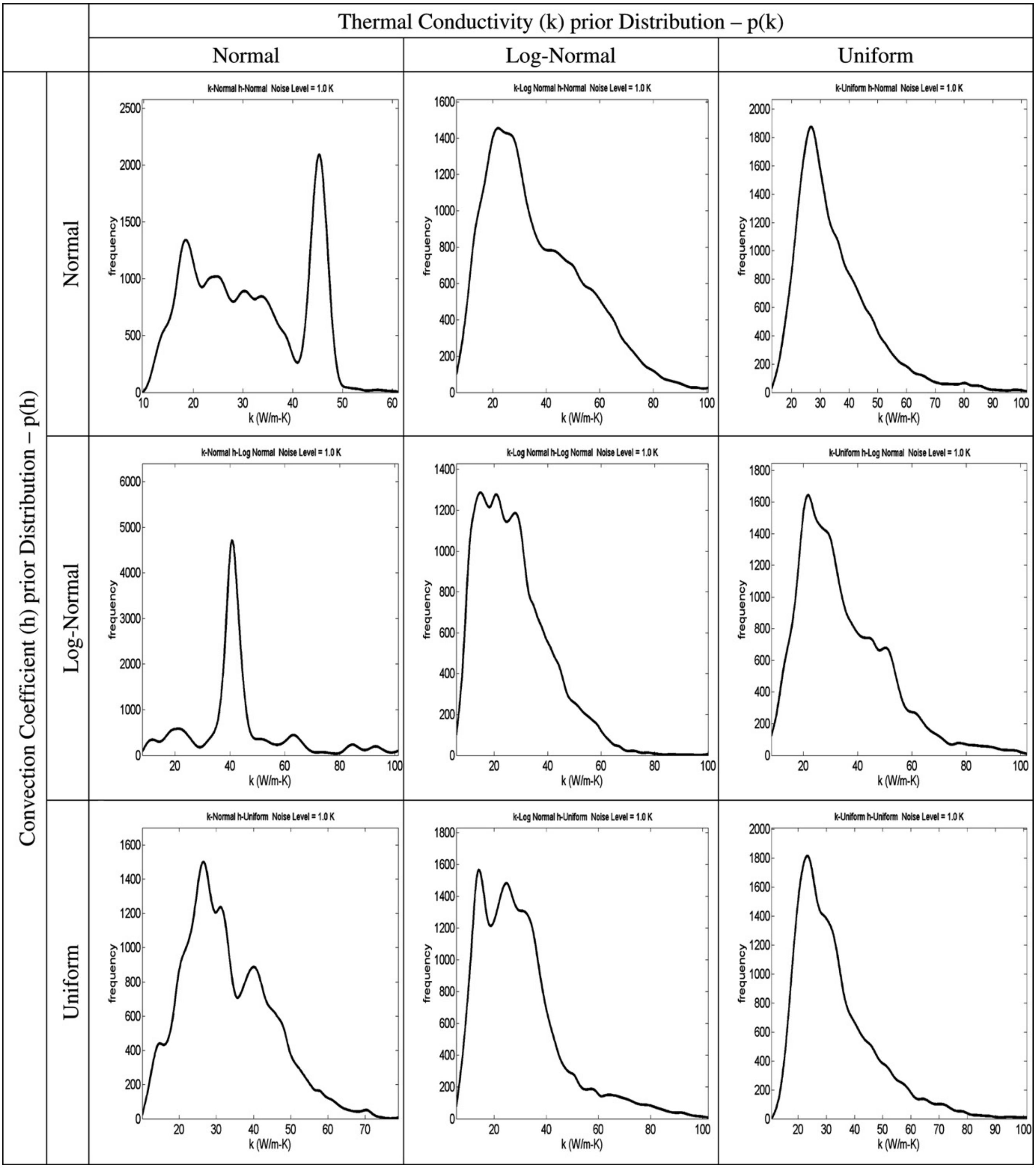
A priori had little effect on point estimates at low noise levels. At noise level 0.5 K and 1.0 K, the estimates for thermal conductivity and convection coefficient were sensitive to the a priori model. The trends observed in this case are plotted in Figs. 8 and 9. The low accuracy of the algorithm can be attributed to two reasons (a) increased noise

Table 8
Marginal PPDFs for thermal conductivity $P(k|Y)$ at noise level: 0.5 K, Case 3.2



level and (b) correlation between thermal conductivity and convection coefficient. A higher k increases rate of heat transfer in the domain. In-order to maintain the same temperatures (i.e. $\mathbf{Y}$) the convective heat transfer must

Table 9
Marginal PPDFs for thermal conductivity $P(k|Y)$ at noise level: 1.0 K, Case 3.2

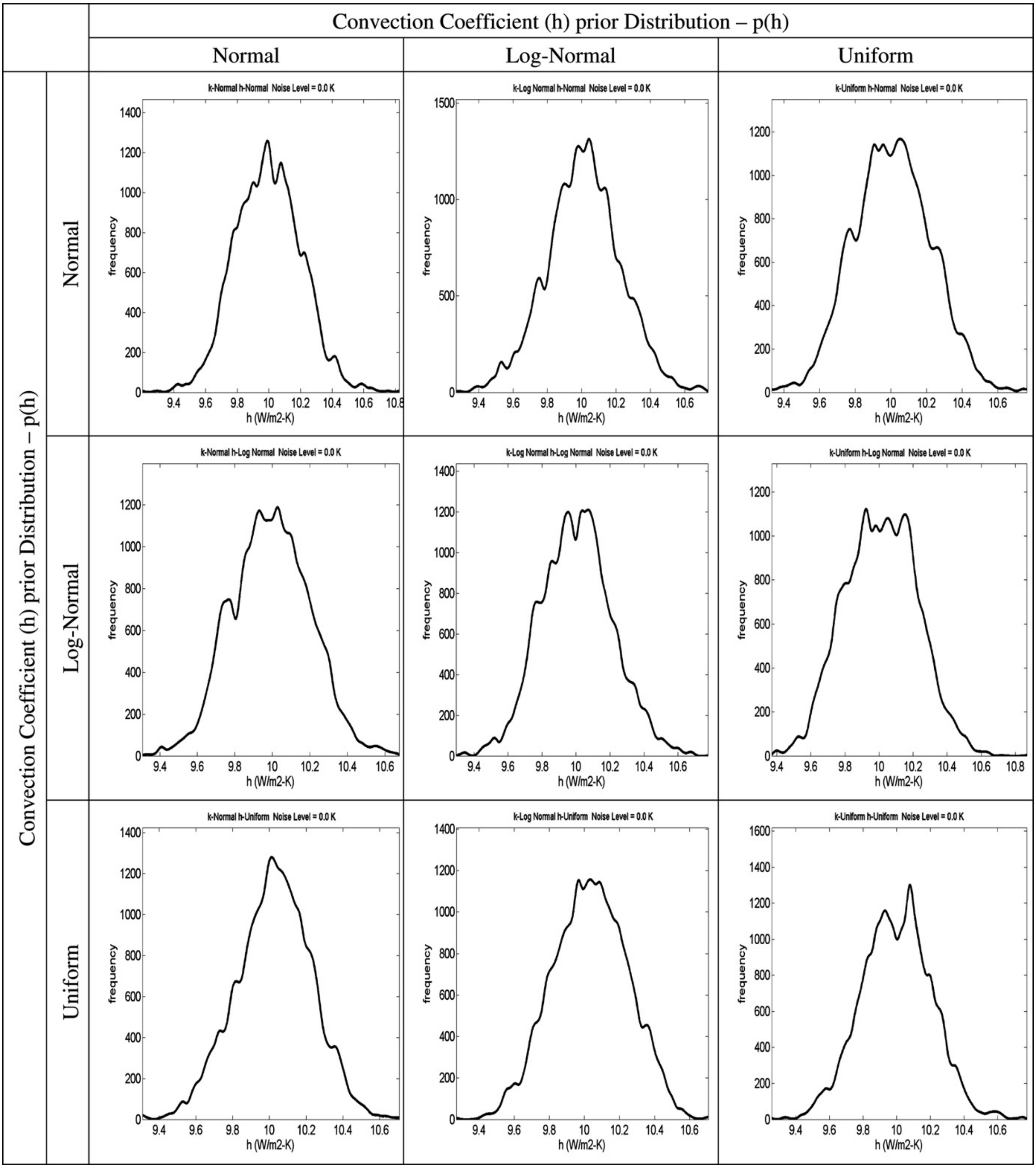


increase to increase heat transfer out of the domain. Hence a higher k estimate leads to a higher h estimate or vice-versa. This trend can also be observed from Figs. 8 and 9.

3.3. Three-parameter estimation in 2D transient conduction

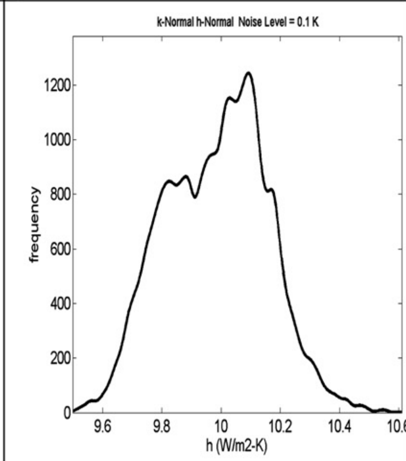
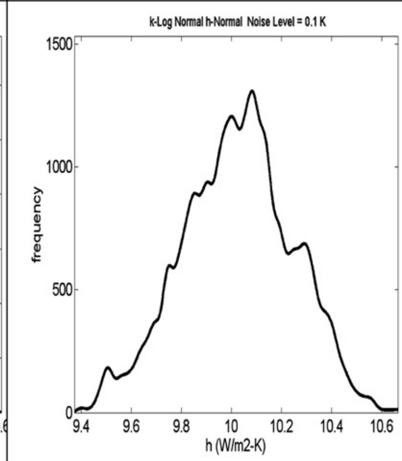
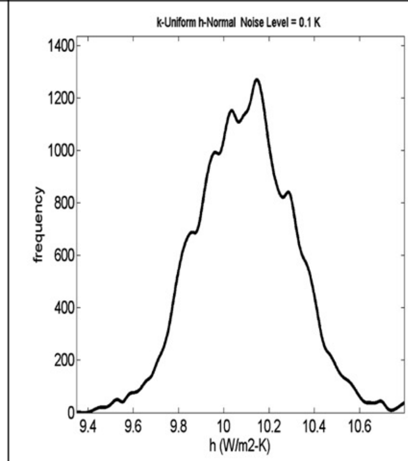
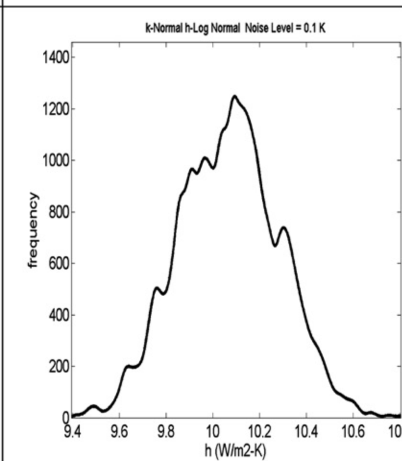
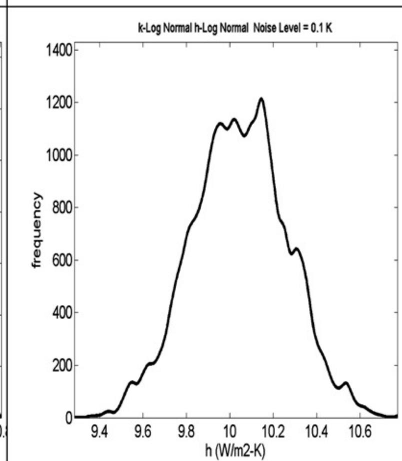
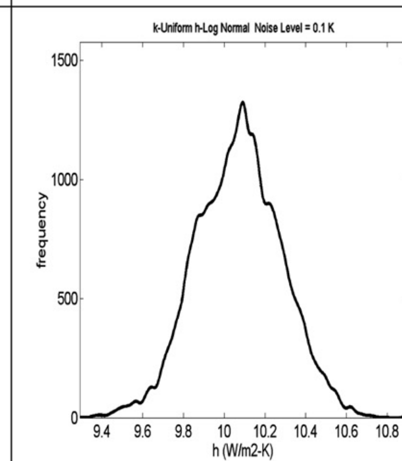
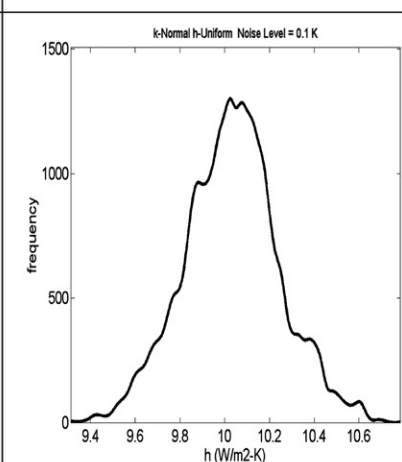
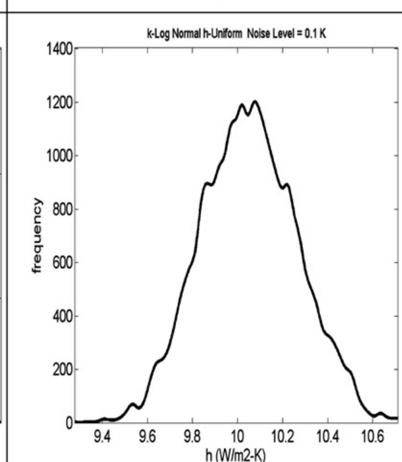
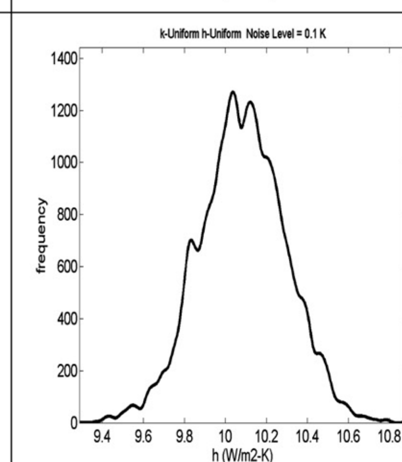
In this case, a heat transfer to the outside by radiation is introduced apart from the convection. Furthermore, the

Table 10
Marginal PPDFs for convection coefficient $P(h|Y)$ at noise level : 0.0 K, Case 3.2



aspect ratio is changed to 10:1 so as to simulate heat transfer in a fin. Array of such fins are widely used in cooling of electric motors, electronics and so on. This is less restrictive than the case considered in Section 3.2, as radiation is

Table 11
Marginal PPDFs for convection coefficient, $P(h|Y)$ at noise level : 0.1 K, Case 3.2

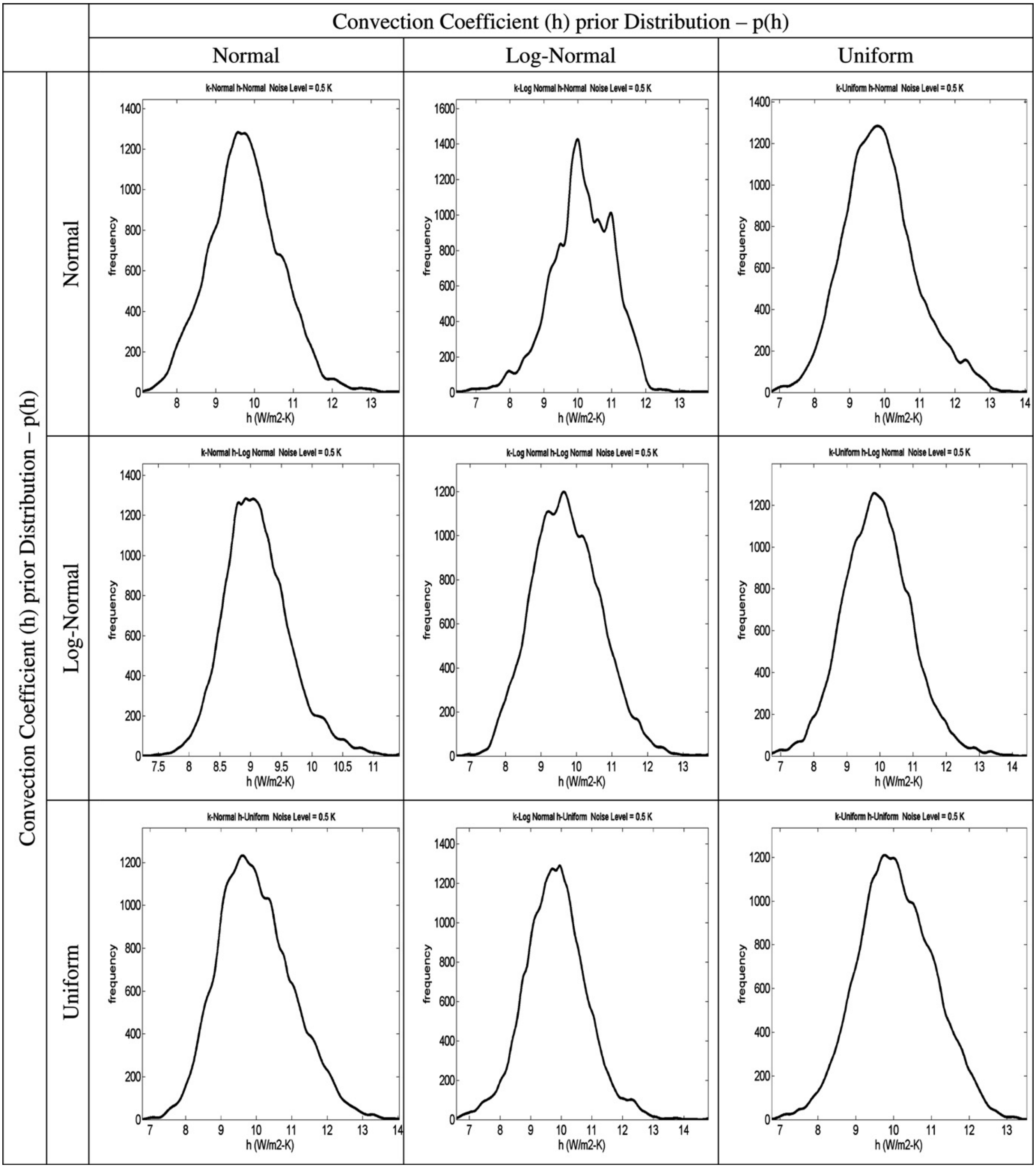
Convection Coefficient (h) prior Distribution – p(h)		Convection Coefficient (h) prior Distribution – p(h)		
		Normal	Log-Normal	Uniform
Convection Coefficient (h) prior Distribution – p(h)	Normal			
	Log-Normal			
	Uniform			

inevitably present in such applications. The geometry is presented in Fig. 10 and the governing equation and boundary conditions are listed below:

$$\rho C_p \frac{\partial T}{\partial t} = \nabla(k \cdot \nabla T) \tag{18}$$

$$T(x, y, 0) = T_I \tag{19}$$

Table 12
Marginal PPDFs for convection coefficient, $P(h|Y)$ at noise level : 0.5 K, Case 3.2

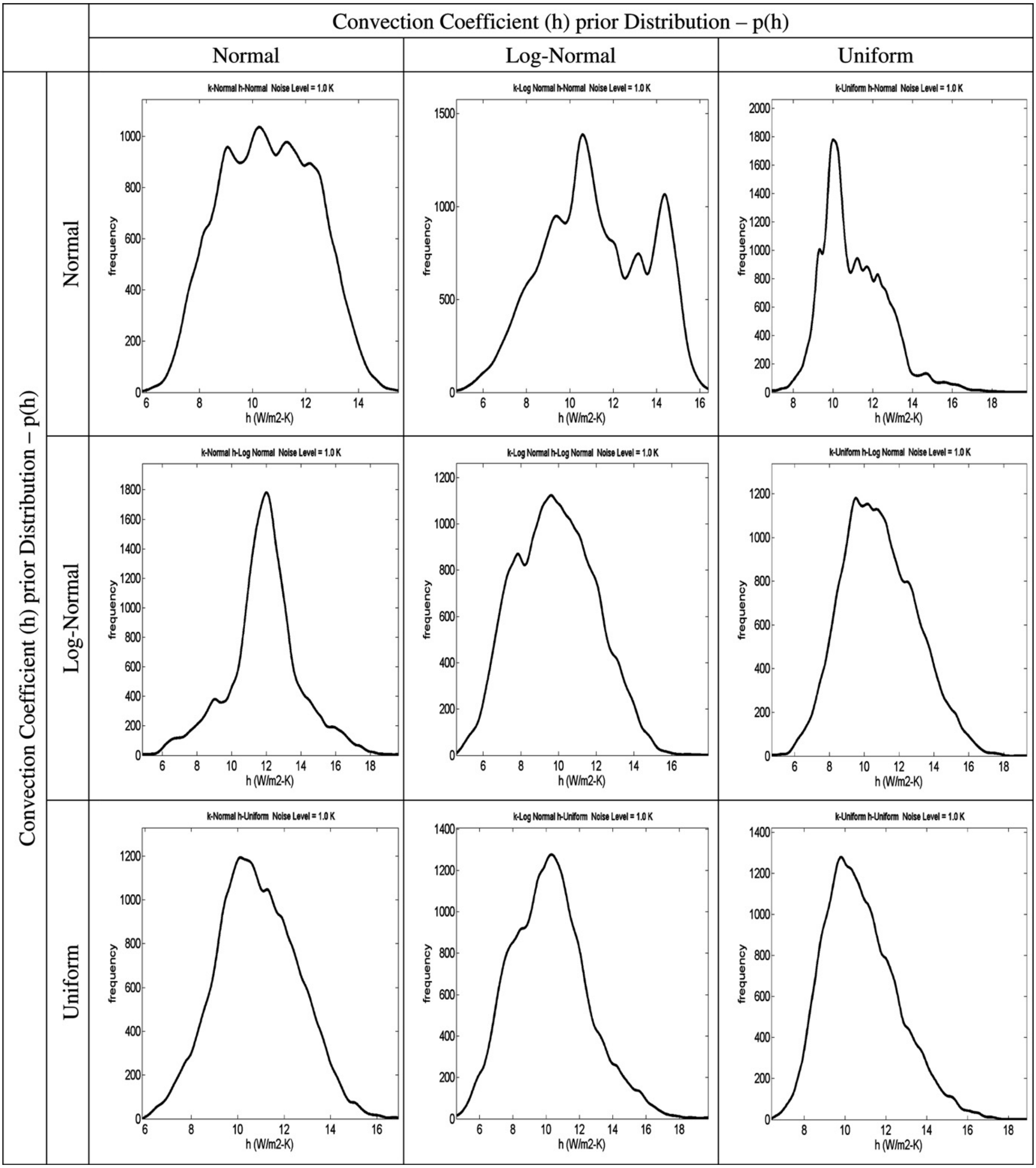


$$-k \cdot \frac{\partial T}{\partial y} = h \cdot (T - T_\infty) + \sigma_s \cdot \varepsilon \cdot (T^4 - T_\infty^4) \quad \forall (x,y) \in (\Gamma_1 U \Gamma_3) \qquad (20)$$

$$-k \cdot \frac{\partial T}{\partial x} = h \cdot (T - T_\infty) + \sigma_s \cdot \varepsilon \cdot (T^4 - T_\infty^4) \quad \forall (x,y) \in \Gamma_2 \qquad (21)$$

$$T(x,y,t) = T_w \quad \forall (x,y) \in \Gamma_4 \qquad (22)$$

Table 13
Marginal PPDFs for convection coefficient, $P(h|Y)$ at noise level : 1.0 K, Case 3.2



$T_I = 573$ K

$T_w = 573$ K

$T_\infty = 298$ K

The data vector Y is constructed by solving the forward problem with $k = 25$ W/m K, $h = 10$ W/m^2 K and $\varepsilon = 0.8$ in one sub-case and with $k = 25$ W/m K, $h = 10$ W/m^2 K and $\varepsilon = 0.83$ in another sub-case. The data is corrupted

Table 14
Prior distributions in two-parameter estimation problem

Prior	f_1	f_2
Normal	$\upsilon_k^{-0.5}\exp\left(-\frac{(k-\mu_k)^2}{2\upsilon_k}\right)\cdot\upsilon_k^{-(1+\alpha)}\cdot\exp(-\beta\cdot\upsilon_k^{-1})$	$\upsilon_h^{-0.5}\exp\left(-\frac{(h-\mu_h)^2}{2\upsilon_h}\right)\cdot\upsilon_k^{-(1+\alpha)}\cdot\exp(-\beta\cdot\upsilon_h^{-1})$
Log normal	$\upsilon_{lk}^{-0.5}\exp\left(-\frac{(\log(k)-\mu_{lk})^2}{2\upsilon_{lk}}\right)\cdot\upsilon_{lk}^{-(1+\alpha)}.\exp(-\beta\cdot\upsilon_{lk}^{-1})$	$\upsilon_{lh}^{-0.5}\exp\left(-\frac{(\log(k)-\mu_{lh})^2}{2\upsilon_{lh}}\right)\cdot\upsilon_{lh}^{-(1+\alpha)}.\exp\left(-\beta\cdot\upsilon_{lh}^{-1}\right)$
Uniform	Constant	Constant

Note: The above functions are defined for $0 < k < k_{\max}$ and $0 < h < h_{\max}$ where $k_{\max}, h_{\max}$ can be arbitrarily large values and $\mu_h, \upsilon_h, \mu_{lh}, \upsilon_{lh}, \mu_{lk}, \upsilon_{lk}$ are hyper parameters and these values are also retrieved in Bayesian inference.

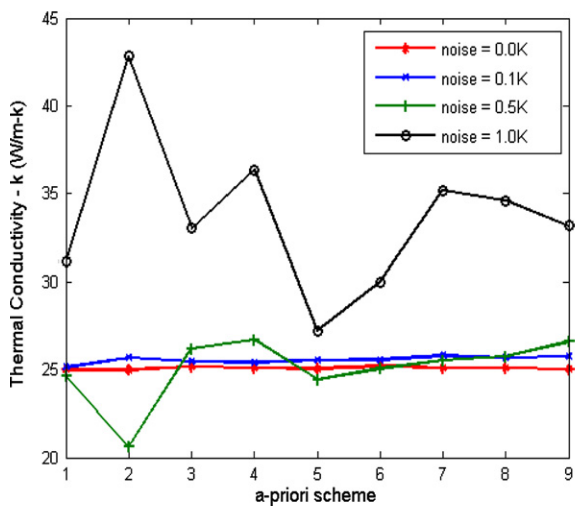


Fig. 8. Effect of a priori scheme on mean thermal conductivity estimate in Case 3.2. Schemes – (K,H): 1 – (N,N), 2 – (N,LN), 3 – (N,U), 4 – (LN,N), 5 – (LN,LN), 6 – (LN,U), 7 – (U,N), 8 – (U,LN), 9 – (U,U).

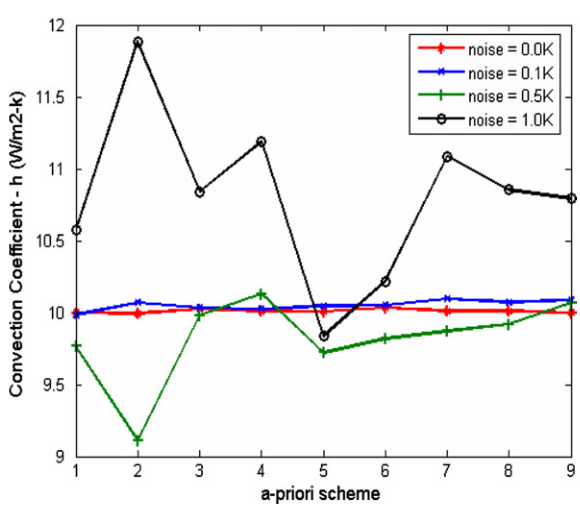


Fig. 9. Effect of a priori scheme on mean convection coefficient estimate in Case 3.2. Schemes – (K,H): 1 – (N,N), 2 – (N,LN), 3 – (N,U), 4 – (LN,N), 5 – (LN,LN), 6 – (LN,U), 7 – (U,N), 8 – (U,LN), 9 – (U,U).

with Gaussian i.i.d noise with $\sigma = 0.1$ K in the latter subcase.

A commercial FEM solver was used to solve the forward problem. An 11×21 grid and full order scheme was used. The time discretization was 1 s and this scheme was arrived at after the grid independence checks (Fig. 11). The measurement vector **Y** was constructed by the temperatures at (0.3, 0.001), (0.6, 0.001), (0.9, 0.001), (0.9, 0.004), (0.5, 0.003) and (0.7, 0.003) at time $t = 1,2,3,\ldots,10$ s. The dimension of **Y** vector is 60.

Coming to the inverse problem, the problem is a case of alternate solution [14]. The convection heat transfer coefficient and the emissivity of the surface are highly correlated. A fall in heat transfer by one mode can be compensated by increase in the heat transfer by the other mode. The motive for working on this case is to investigate if the algorithm showed any indications of alternate solution. Stated explicitly, there is a possibility of the algorithm converging to low emissivity and unusually high convection coefficient. Even though "domain knowledge" can be used to disregard this solution, from a Bayesian perspective we need to investigate whether constraining the prior leads to the true solution.

Since the noise level is low, i.e. 0.1 K, uniform priors are used for conductivity and convection coefficient. In the first case, an unconstrained uniform a priori between 0 and 1, i.e. $U(0, 1)$, is used for emissivity and a preliminary estimate is obtained from 1000 samples after convergence of the Markov chain. The mean estimates obtained for k, h, ε were 24.43 W/m K, 25.98 W/m^2 K, and 0.02, respectively. Fig. 12 is a good test for the retrieval, wherein a parity plot of the measured temperatures and calculated temperatures with estimated values of the parameters using forward model is given. This plot confirms the validity of the solution. The R^2 value for the line was 1.0.

From a heat transfer perspective, it is clear that the h value of 25 W/m^2 K for natural convection of air is highly unlikely. Therefore this solution can be considered "infeasible" not from the point of view of the Bayesian statistics, but from the physics of the problem.

Further constraining the problem, a $U(0.5, 1)$ i.e. uniform prior between 0.5 and 1.0 was used. Using 1000 samples after convergence the estimates were 26.08 W/m K, 15.53 W/m^2 K, 0.527. Fig. 13 validates the above solution. The R^2 value for the line is 1.00 too. In the above two sub cases the measurement vector was not corrupted by noise but the σ value used was 0.1 K.

In the final case, a normal a priori with mean 0.8 and SD 5% of mean was used. The data vector was constructed using $\varepsilon = 0.83$. In this case, the values estimated using 10000 converged samples were 27.15 W/m K, 10.39 W/m^2 K, 0.815. Fig. 14 validates the solution.

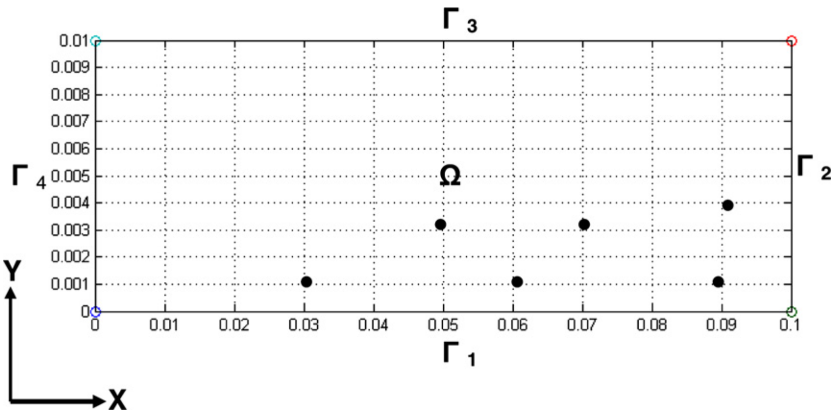


Fig. 10. Problem geometry, Case 3.3. $L = 0.1$ m, $B = 0.01$ m, observation time = 10 s.

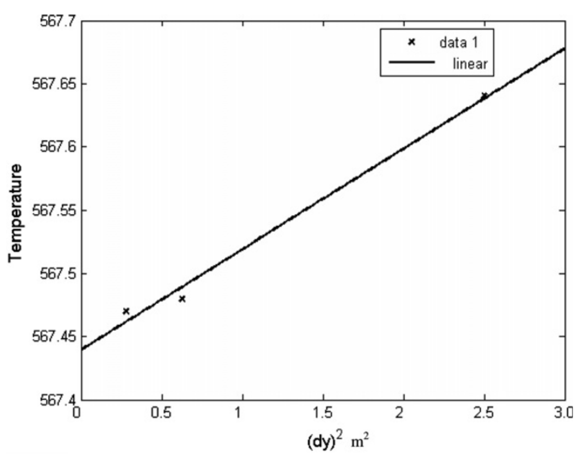


Fig. 11. Grid independence check – (Richardson's Extrapolation).

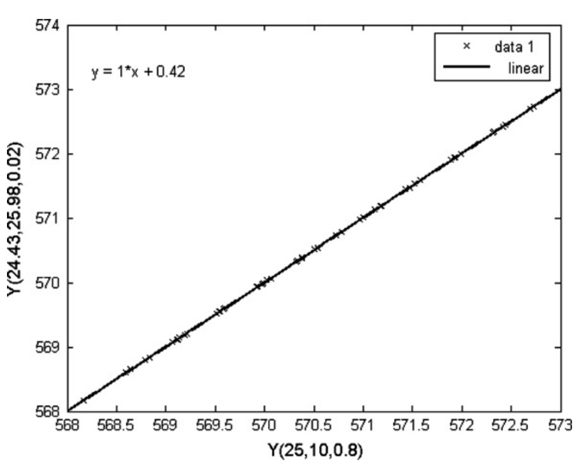


Fig. 12. Parity plot of Y (24.43, 25.98, 0.02) vs. Y(25,10,0.8), Case 3.3, $R^2 = 1$.

In all the above three cases, the results did not show any aberration so as to indicate an alternate optimum. It is also important to note that better priors, in terms of information, gave better estimates. The posterior probability density functions and the statistics of the estimates in the final case are presented in Figs. 15–17.

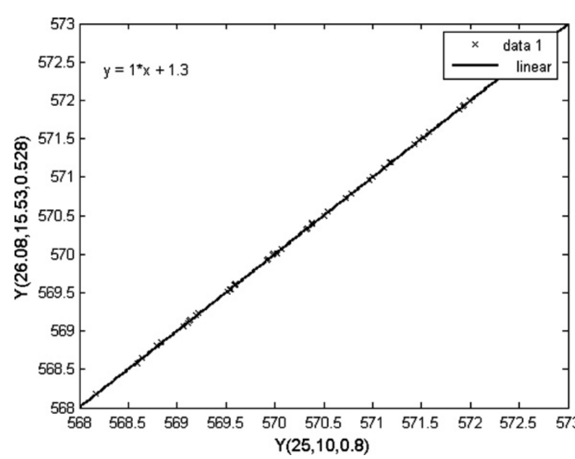


Fig. 13. Parity plot of Y (26.08, 15.53, 0528) vs. Y (25, 10, 0.8), Case 3.3, $R^2 = 1$.

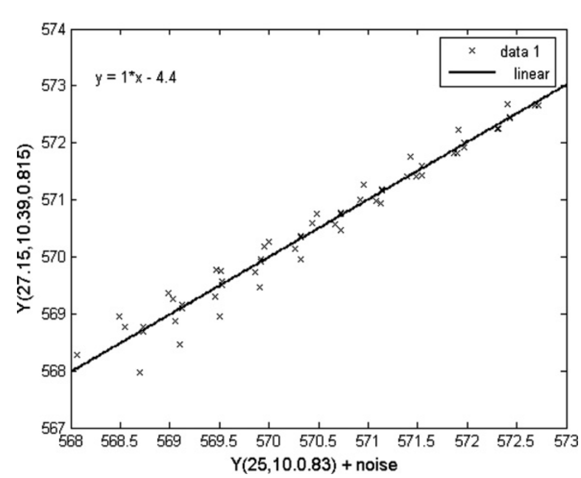


Fig. 14. Parity plot of Y (27.15, 10.39, 0.815) vs. Y (25, 10, 0.83), $R^2 = 0.97$.

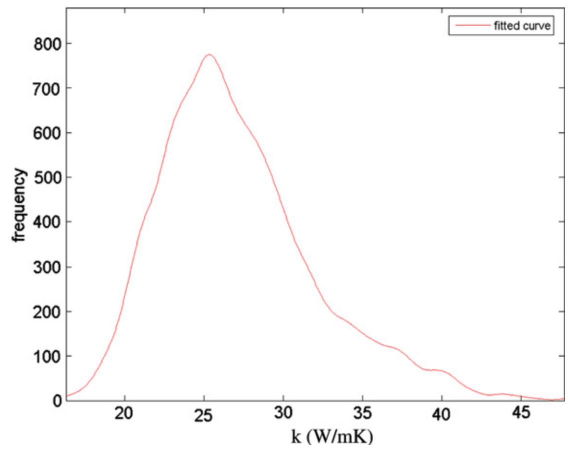


Fig. 15. Marginal posterior probability density function of thermal conductivity – $P(k|Y)$.

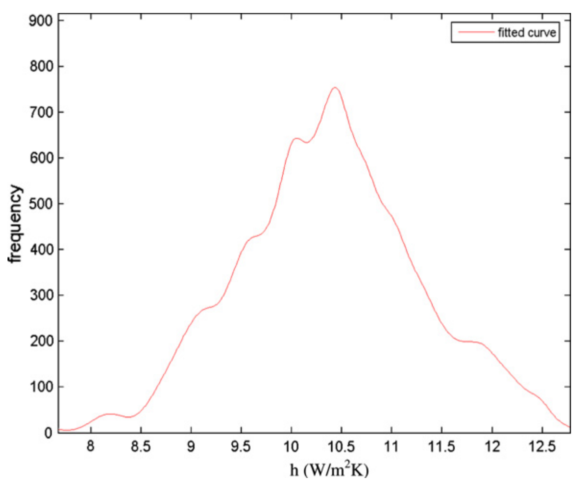


Fig. 16. Marginal posterior probability density function of convection coefficient – $P(h|Y)$.

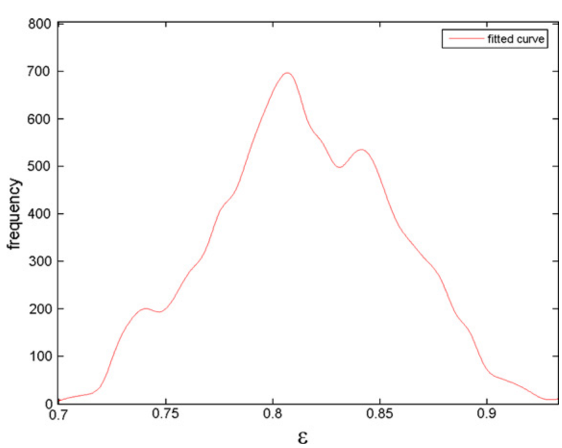


Fig. 17. Marginal posterior probability density function of emissivity – $P(\varepsilon|Y)$.

4. Conclusions

For a two-dimensional, unsteady inverse heat conduction problem, the sensitivity of the solution on the a priori model has been studied in this work at different levels of noise in measured data for both single-parameter and two-parameter estimation problems, when hierarchical Bayesian prior models are applied.

Estimates were found insensitive to the a priori model at all the considered noise levels in the single-parameter estimation problem. However, in the two-parameter estimation problem at noise levels of 0.5 K and 1.0 K, due to the increased ill-posed nature of the problem when compared to a single-parameter problem, the a priori model had a significant effect on the estimates. At 1 K noise level, mean estimates for k varied between 27 W/m K and 43 W/m K and mean estimates of h varied between 9.75 W/m^2 K and 12 W/m^2 K for different combinations of a priori models.

Bayesian inference tends to point to alternate solutions when highly correlated parameters are retrieved using non-informative prior models. This was demonstrated in the three parameter estimation problem. An improvement in the sampling technique or better priori knowledge of the concerned parameters, i.e. a priori, may address this issue.

Bayesian inference could be used as a powerful tool to design experiments by exploiting the dependence of variance and the accuracy. It is desired to have minimal variance for the estimated samples and this could be achieved by appropriate selection of the measurement points which were considered frozen in this study. The selection of these points for achieving minimum variance is currently under progress.

References

[1] M. Deiveegan, V. Swaminathan, C. Balaji, S.P. Venkateshan, A Bayesian algorithm for the retrieval of geophysical parameters in the Atmosphere, in: Proceedings from 13th International Heat Transfer Conference, Sydney, Australia, 2006.
[2] P. Lauret, F. Miranville, H. Boyer, F. Garde, L. Adelard, Bayesian parameter estimation of convective heat transfer coefficients of a roof-mounted radiant barrier system, J. Solar Energy 128 (2006) 213–225.
[3] A.N. Tikhonov, Solution of Ill-posed Problems, Halsted Press, Washington, 1977.
[4] O.M. Alifanov, Inverse Heat Transfer Problems, Springer-Verlag, Berlin, 1994.
[5] J.V. Beck, B. Blackwell, C.R. St-Clair Jr., Inverse Heat Conduction: Ill-posed Problems, Wiley-Interscience, New York, 1985.
[6] R.M. Neal, Philosophy of Bayesian Inference, http://www.cs.toronto.edu/~radford/res-bayes-ex.html, 1998.
[7] J. Wang, N. Zabaras, A Bayesian inference approach to the inverse heat conduction problem, Int. J. Heat Mass Transfer 47 (2004) 3927–3941.
[8] R.M. Neal, Probabilistic Inference Using Markov Chain Monte Carlo Methods, Technical Report, CRG-TR-93-1, Department of Computer Science, University of Toronto, September, 1993.
[9] J.K. Ghosh, M. Delampady, T. Samanta, An Introduction to Bayesian Analysis, Theory and Methods, Springer, New York, USA, 2006.
[10] L. Chen, Z. Qin, J.S. Liu, Exploring hybrid Monte Carlo in Bayesian computation, in: Proceedings of International Society for Bayesian Analysis, Sixth World meeting, 2000.

[11] J. Wang, N. Zabaras, Hierarchical Bayesian models for inverse problems in heat conduction, Inverse Problems 21 (2005) 183–206.
[12] J.K. Ghosh, T. Samanta, Non-subjective Bayes testing – an overview, J. Stat. Planning Inference 103 (2002) 205–223.
[13] J.K. Ghosh, T. Samanta, Towards a non-subjective Bayesian paradigm, Uncertainty and Optimality, World Scientific, 2002, pp. 1–69.
[14] A.F. Emery, E. Valenti, D. Bardot, Using Bayesian inference for parameter estimation when the system response and experimental conditions are measured with error and some variables are considered as nuisance variables, Measur. Sci. Technol. 18 (2007) 19–29.