Inverse Problems in Science and Engineering
Vol. 16, No. 6, September 2008, 677–692

Taylor & Francis
Taylor & Francis Group

Approximation of the likelihood function in the Bayesian technique for the solution of inverse problems

Helcio R.B. Orlande[a]*, Marcelo J. Colaço[b] and George S. Dulikravich[c]

[a]*Department of Mechanical Engineering, POLI/COPPE, Federal University of Rio de Janeiro, UFRJ, Rio de Janeiro, RJ, Brazil;* [b]*Department of Mechanical and Materials Engineering, Military Institute of Engineering, Rio de Janeiro, RJ, Brazil;* [c]*Department of Mechanical and Materials Engineering, Florida International University, Miami, Florida, USA*

(*Received 16 April 2007; final version received 1 May 2008*)

This work deals with the use of radial basis functions for the interpolation of the likelihood function in parameter estimation problems. The focus is on the use of Bayesian techniques based on Markov Chain Monte Carlo (MCMC) methods. The proposed interpolation of the likelihood function is applied to test cases of inverse problems in heat and mass transfer, solved with the Metropolis–Hastings algorithm. The use of the interpolated likelihood function reduces significantly the computational cost associated with the implementation of such Markov Chain Monte Carlo method without loss of accuracy in the estimated parameters.

Keywords: Bayesian technique; inverse problems; response surfaces

Nomenclature

a, b, c	Sides of the parallelepiped
C	Normalized concentration
c_j	Shape parameter for the radial basis functions
D	Dispersion coefficient
h_m	Mass transfer coefficient between the column and an outflow plenum
I	Number of measurements
k_1, k_2, k_3	Thermal conductivity components in the x, y and z directions, respectively
N	Number of unknown parameters
$\mathbf{P}$	Vector of unknown parameters
q_1, q_2, q_2	Uniform heat fluxes applied at the surfaces $x=a$, $y=b$ and $z=c$, respectively
R	Retardation factor
$\mathbf{T}$	Vector of estimated variables
V	Pore velocity
$\mathbf{W}$	Inverse of the covariance matrix of the measurements
$\mathbf{Y}$	Vector of measurements

*Corresponding author. Email: helcio@mecanica.coppe.ufrj.br

ISSN 1741–5977 print/ISSN 1741–5985 online

DOI: 10.1080/17415970802231677
http://www.informaworld.com

Greeks

π	Probability density
$\phi(\cdot)$	Radial basis functions, see Equations (8a–d)
σ	Standard deviation of the measurements

Superscripts

t	Current state in the Markov chain
n	Number of states in the Markov chain

1. Introduction

A variety of techniques are currently available for the solution of inverse problems. However, one common approach relies on the minimization of an objective function that generally involves the squared difference between measured and estimated variables, like the least-squares norm, as well as some kind of regularization term. Despite the fact that the minimization of the least-squares norm is indiscriminately used, it only yields *maximum likelihood* estimates if the following statistical hypotheses are valid [1]: the errors in the measured variables are additive, uncorrelated, normally distributed, with zero mean and known constant standard-deviation; only the measured variables appearing in the objective function contain errors; and there is no prior information regarding the values and uncertainties of the unknown parameters. Although very popular and useful in many situations, the minimization of the least-squares norm is a non-Bayesian estimator. A Bayesian estimator is concerned with the analysis of the *posterior probability density*, which is the conditional probability of the parameters given the measurements, while the likelihood is the conditional probability of the measurements given the parameters [1–4]. Recent examples of works dealing with Bayesian techniques for the solution of inverse heat transfer problems include [5–8].

If we assume the parameters and the measurements to be independent Gaussian random variables, with known mean and covariance matrices, and that the measurement errors are additive, a closed form expression can be derived for the posterior probability density. In this case, the estimator that maximizes the posterior probability density can be recast in the form of a minimization problem involving the *maximum a posteriori objective function* [1–4].

On the other hand, if different *prior* probability densities are assumed for the parameters, the posterior probability distribution does not allow an analytical treatment. In this case, Markov Chain Monte Carlo (MCMC) methods are used to draw samples of all possible parameters, so that inference on the posterior probability becomes inference on the samples [2–4]. As such, the number of samples required to accurately approximate the posterior distribution is generally large, resulting in prohibitive computational costs for many practical applications. Such is specially the case when the solution of the forward problem, which is needed for the computation of the likelihood function, requires large computational times.

In this work, we examine the use of radial basis functions (RBFs) to interpolate the likelihood function, in order to reduce the computational cost of MCMC methods in the Bayesian approach of solution of inverse problems. The likelihood function is

interpolated in the space of all possible parameters, by using a small number of solutions of the forward model as compared to that required for the implementation of the MCMC methods. Hence, the interpolated likelihood function, instead of the actual function, is used afterwards in the sampling procedure of the MCMC method, providing a substantial reduction on computational costs.

The use of RBFs followed by collocation, a technique first proposed by Kansa [9] after the work of Hardy [10] on multivariate approximation, is now becoming an established approach. Various applications to problems in mechanics have been made in recent years – see, for example, Leitão [11,12]. A systematic evaluation of the use of RBFs to approximate multivariable functions is presented in [13].

The proposed approach to reduce the computational cost of MCMC methods is examined below, as applied to the solution of inverse parameter estimation problems in heat and mass transfer.

2. Bayesian technique for the solution of inverse problems

In the Bayesian approach to statistics, an attempt is made to utilize all available information in order to reduce the amount of uncertainty present in an inferential or decision-making problem. As new information is obtained, it is combined with any previous information to form the basis for statistical procedures. The formal mechanism used to combine the new information with the previously available information is known as Bayes' theorem [14]. Therefore, the term *Bayesian* is often used to describe the so-called *statistical inversion approach*, which is based on the following principles [2]:

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

Consider the vector of parameters appearing in the physical model formulation as

$$\mathbf{P}^T \equiv [P_1, P_2, \ldots, P_N] \tag{1a}$$

and the vector of available measurements as

$$\mathbf{Y}^T \equiv [Y_1, Y_2, \ldots, Y_I] \tag{1b}$$

where N is the number of parameters and I is the number of measurements. Bayes' theorem can then be stated as [2]:

$$\pi_{\text{posterior}}(\mathbf{P}) = \pi(\mathbf{P}|\mathbf{Y}) = \frac{\pi_{\text{prior}}(\mathbf{P})\pi(\mathbf{Y}|\mathbf{P})}{\pi(\mathbf{Y})} \tag{2}$$

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

In practice, such normalizing constant is difficult to compute and numerical techniques, like MCMC methods, are required in order to obtain samples that accurately represent the posterior probability density. In order to implement the Markov chain, a density $q(\mathbf{P}^*, \mathbf{P}^{(t-1)})$ is required, which gives the probability of moving from the current state in the chain $\mathbf{P}^{(t-1)}$ to a new state $\mathbf{P}^*$.

The Metropolis–Hastings algorithm was used in this work to implement the MCMC method. It can be summarized in the following steps [2–4]:

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

$$\alpha = \min\left[1, \frac{\pi(\mathbf{P}^*|\mathbf{Y})q(\mathbf{P}^{(t-1)}, \mathbf{P}^*)}{\pi(\mathbf{P}^{(t-1)}|\mathbf{Y})q(\mathbf{P}^*, \mathbf{P}^{(t-1)})}\right]. \tag{3}$$

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

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

We assume in this work that the errors in the measured variables are additive, uncorrelated, normally distributed, with zero mean and known constant standard-deviation. Hence, the likelihood function is given by [1–4]:

$$\pi(\mathbf{Y}|\mathbf{P}) = (2\pi)^{-I/2}\left|\mathbf{W}^{-1}\right|^{-1/2}\exp\left\{-\frac{1}{2}[\mathbf{Y} - \mathbf{T}(\mathbf{P})]^T\mathbf{W}[\mathbf{Y} - \mathbf{T}(\mathbf{P})]\right\} \tag{4}$$

where $\mathbf{T}$ is the vector of estimated variables, obtained from the solution of the forward model with an estimate for the parameters $\mathbf{P}$, and $\mathbf{W}$ is the inverse of the covariance matrix of the measurements.

3. Interpolation of the likelihood function

We note that Equation (3) involves a ratio between the posterior distributions for $\mathbf{P}^*$ and $\mathbf{P}^{(t-1)}$. As a result, only the exponential term appearing in the likelihood function, Equation (4), actually needs to be computed in the implementation of the Metropolis–Hastings algorithm. In the approach proposed in this work, such exponential term is interpolated by using RBFs as described below [13].

Consider a function of L variables x_i, $i = 1, \ldots, L$. The RBF model used in this work has the following form:

$$f(\mathbf{x}) = \sum_{j=1}^{N}\alpha_j\phi(|\mathbf{x} - \mathbf{x}_j|) + \sum_{k=1}^{M}\sum_{i=1}^{L}\beta_{i,k}p_k(x_i) + \beta_0 \tag{5}$$

where $\mathbf{x} = \{x_1, \ldots, x_i, \ldots, x_L)$ and $f(\mathbf{x})$ is known for a series of N points $\mathbf{x}$. Here, $p_k(x_i)$ is one of the M terms of a given basis of polynomials. This approximation is solved for the α_j

and $\beta_{i,k}$ unknowns from the system of N linear equations, subject to the following conditions required for the sake of uniqueness:

$$\begin{aligned} \sum_{j=1}^{N} \alpha_j p_k(x_i) &= 0 \\ &\vdots \\ \sum_{j=1}^{N} \alpha_j p_k(x_L) &= 0 \end{aligned} \tag{6a}$$

$$\sum_{j=1}^{N} \alpha_j = 0. \tag{6b}$$

In this work, the polynomial term appearing in Equation (5) was taken as

$$p_k(x_i) = x_i^k \tag{7}$$

and the RBFs are selected among the following

$$\text{Multiquadrics}: \quad \phi\left(|\mathbf{x}_i - \mathbf{x}_j|\right) = \sqrt{(\mathbf{x}_i - \mathbf{x}_j)^2 + c_j^2} \tag{8a}$$

$$\text{Gaussian}: \quad \phi\left(|\mathbf{x}_i - \mathbf{x}_j|\right) = \exp\left[-c_j^2(\mathbf{x}_i - \mathbf{x}_j)^2\right] \tag{8b}$$

$$\text{Squared multiquadrics}: \quad \phi\left(|\mathbf{x}_i - \mathbf{x}_j|\right) = (\mathbf{x}_i - \mathbf{x}_j)^2 + c_j^2 \tag{8c}$$

$$\text{Cubical multiquadrics}: \quad \phi\left(|\mathbf{x}_i - \mathbf{x}_j|\right) = \left[\sqrt{(\mathbf{x}_i - \mathbf{x}_j)^2 + c_j^2}\right]^3. \tag{8d}$$

The choice of which polynomial order, which RBF and which shape parameter c_j are the best to a specific function, was made based on a cross-validation procedure. Let us suppose that we have P_{TR} training points, which are the locations on the multidimensional space where the values of the function are known. Such set of training points is equally subdivided into two subsets of points, named P_{TR1} and P_{TR2}. Equations (5) and (6a and b) are solved for a polynomial of order zero with a small value of the shape factor c_j and for each RBF given by Equations (8a–d) using the subset P_{TR1}. The initial value of the shape factor is taken as equal to the minimum distance among two points of the training data. Then, the value of the interpolated function is checked against the known value of the function for the subset P_{TR2} and the error is computed as

$$\text{RMS}_{P_{\text{TR1}}} = \sum_{i=1}^{P_{\text{TR2}}} [f_{\text{inter}}(\mathbf{x}_i) - f(\mathbf{x}_i)]^2 \tag{9a}$$

where $f_{\text{inter}}(\mathbf{x}_i)$ is the value of the interpolated function at $\mathbf{x}_i$.

Then, the same procedure is repeated by using the subset P_{TR2} to solve Equations (5) and (6a and b) and the subset P_{TR1} to calculate the error as

$$\text{RMS}_{P_{\text{TR2}}} = \sum_{i=1}^{P_{\text{TR1}}} [f_{\text{inter}}(\mathbf{x}_i) - f(\mathbf{x}_i)]^2. \tag{9b}$$

Finally, the total error is obtained as

$$\mathrm{RMS}_{\mathrm{RBF}} = \sqrt{\mathrm{RMS}_{P_{\mathrm{TR1}}} + \mathrm{RMS}_{P_{\mathrm{TR2}}}}. \tag{10}$$

This procedure is repeated for all polynomial orders, up to $M = 6$, for each one of the RBF expressions given by Equations (8a–d) and for increasing values of the shape factor c_j. The best combination that is used in the interpolation of the likelihood function is the one that returns the lowest value of the RMS error.

4. Results and discussions

We now examine the approach described above for the interpolation of the likelihood function, as applied to two different test cases of practical interest. These test cases include the transport of tracers in soil columns and the three-dimensional heat conduction in an orthotropic media. The results obtained with the interpolated approach described above are compared to those obtained without interpolation, as well as those interpolated with multiquadrics RBFs. For the sake of clarity, such tecnhiques are hereafter referred to as: (i) *Technique 1*: without interpolation; (ii) *Technique 2*: interpolation with multiquadrics RBFs with the shape factor given by the inverse of the standard-deviation of the measurement errors and (iii) *Technique 3*: interpolation with RBFs using the cross-validation procedure described above.

5. Transport of tracers in soil columns

Consider the identification of soil properties based on the dispersion of a tracer in a column. We assume that a column of length L is filled with a soil saturated with water. After establishing a constant flow of a solution with tracer concentration C_b, the inflow concentration is changed to C_0. Dispersion is assumed to be onedimensional along the longitudinal direction through the column. Also, we assume that the relation between adsorbed and solution concentrations is described by a linear isotherm, so that the diffusion–advection equation describing the salt dispersion through the column is given by:

$$R\frac{\partial c(z,t)}{\partial t} = D\frac{\partial^2 c(z,t)}{\partial z^2} - V\frac{\partial c(z,t)}{\partial z} \qquad \text{for } 0 < z < L \quad \text{and} \quad t > 0. \tag{11a}$$

The initial condition is given by:

$$c(z,0) = C_b \qquad \text{for } t = 0, \text{ in } 0 < z < L \tag{11b}$$

and the boundary conditions as:

$$c(0,t) = C_0 \qquad \text{at } z = 0 \text{ and for } t > 0 \tag{11c}$$

$$D\frac{\partial c(L,t)}{\partial z} + h_{\mathrm{m}} c(L,t) = h_{\mathrm{m}} C_{\mathrm{b}} \qquad \text{at } z = L \text{ and for } t > 0 \tag{11d}$$

where D is the dispersion coefficient, R is the retardation factor and V is the pore velocity.

Note in Equation (11c) that the boundary condition at $z = 0$ is taken as of the first kind, by assuming that the advective effects are locally dominant. Also, note in

Equation (11d) that the boundary condition at $z = L$ was taken as of the thirdkind, where h_m is the mass transfer coefficient between the column and an outflow plenum.

For the solution of the problem given by Equation (11) we use finite-differences with the McCormack predictor–corrector scheme [15].

The objective of the *inverse problem* is to estimate the vector of parameters

$$\mathbf{P}^{\mathrm{T}} = [D, R, h_{\mathrm{m}}, V] \tag{12}$$

from the measurements of the outflow concentration of the tracer at $z = L$. For the solution of the inverse problem, the concentration is considered to be normalized in the form:

$$C(z, t) = \frac{c(z, t)}{C_0}. \tag{13}$$

We utilize in this analysis simulated experimental data containing random errors with constant standard deviation of 0.05. The simulated measurements were generated in a column with length $L = 5.4\,\mathrm{cm}$, by considering the following values for the parameters: $R = 14.4$, $D = 11.08\,\mathrm{cm^2\,min^{-1}}$, $h_{\mathrm{m}} = 0.39\,\mathrm{cm\,min^{-1}}$ and $V = 0.59\,\mathrm{cm\,min^{-1}}$. It is assumed that 90 measurements of the outflow concentration are available for the inverse analysis, taken in intervals of 1 min.

For the solution of the inverse problem, the prior information for the parameters was considered in the form of uniform distributions, as follows:

$$\begin{aligned} 9 &\le R \le 20 \\ 9\,\mathrm{cm^2\,min^{-1}} &\le D \le 20\,\mathrm{cm^2\,min^{-1}} \\ 0.3\,\mathrm{cm\,min^{-1}} &\le h_{\mathrm{m}} \le 0.6\,\mathrm{cm\,min^{-1}} \\ 0.58\,\mathrm{cm\,min^{-1}} &\le V \le 0.60\,\mathrm{cm\,min^{-1}}. \end{aligned}$$

Due to the linear dependence between the pore velocity and the dispersion coefficient, a quite small interval was used as the prior for the pore velocity. Such a fact is not restrictive from the practical point of view, since this parameter can be directly obtained with relative high accuracy from the measurements of the volume flow at the column outlet. The number of samples used in the Metropolis–Hastings algorithm was 20,000 and, for the computation of the results, the first 500 samples were discarded.

Table 1 presents the results obtained with Technique 1, in terms of the mean and the standard-deviation for each parameter. This table shows that quite accurate estimates can be obtained for the mean values of the parameters, in comparison to the exact ones used to generate the simulated data. Such is the case despite the fact that the simulated errors are

Table 1. Results obtained with Technique 1.

Parameter	Mean	Standard deviation
R	16.7	1.9
$D\,(\mathrm{cm^2\,min^{-1}})$	12.7	1.5
$h_m\,(\mathrm{cm\,min^{-1}})$	0.50	0.06
$V\,(\mathrm{cm\,min^{-1}})$	0.59	0.01

actually quite large, as illustrated in Figure 1. Figure 1 presents the errorless simulated experimental data (solution of the forward problem with the exact parameters), the simulated data containing random errors that were used for the inverse analysis and the estimated concentration obtained with the mean values estimated for the parameters, by using Technique 1. The agreement between estimated and measured concentrations is quite good. The relatively large standard-deviations observed in Table 1 are due to the large errors assumed for the simulated measurements, as well as due to small and correlated sensitivity coefficients.

Tables 2 and 3 present the results obtained with Techniques 2 and 3, respectively, for two different numbers of interpolating points. A comparison of Tables 1 and 2 reveals that Technique 2 is not capable of reaching the same level of accuracy obtained with Technique 1. The mean values estimated for the parameters with Technique 2 are in fact

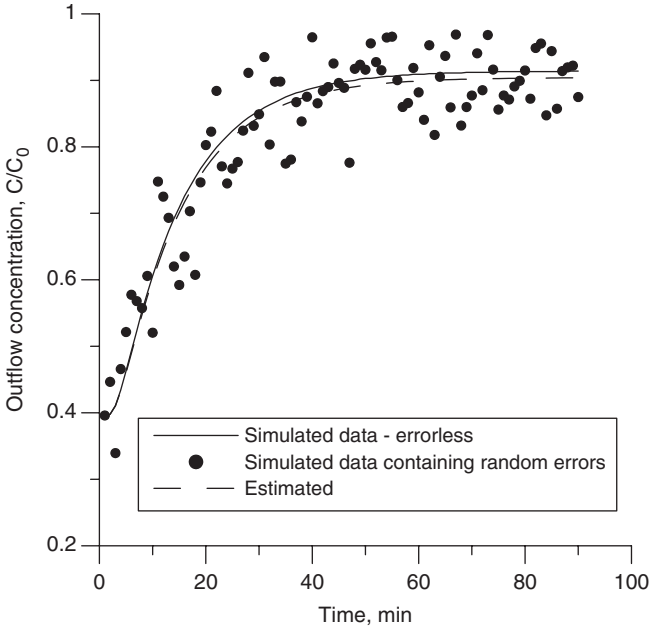

Figure 1. Comparison between measured and estimated concentrations: Technique 1.

Table 2. Results obtained with Technique 2.

Parameter	Mean	Standard deviation	Number of interpolating points
R	18.3	2.2	300
D ($cm^2\,min^{-1}$)	14.7	1.6	
h_m ($cm\,min^{-1}$)	0.57	0.03	
V ($cm\,min^{-1}$)	0.58	0.01	
R	18.9	0.7	500
D ($cm^2\,min^{-1}$)	15.2	0.7	
h_m ($cm\,min^{-1}$)	0.57	0.03	
V ($cm\,min^{-1}$)	0.59	0.01	

quite different from the exact ones, as well as from those obtained with Technique 1. On the other hand, results quite close to those obtained with Technique 1 were obtained with the use of Technique 3 (Tables 1 and 3). Tables 2 and 3 also show that the results obtained with the interpolated likelihood function were not significantly affected by the number of interpolating points used.

Figure 2(a–c) present the states of the Markov chain obtained for the retardation factor, dispersion coefficient and mass transfer coefficient, with Techniques 1–3, respectively. For Techniques 2 and 3; 300 points were used to interpolate the likelihood function with RBFs. A comparison of Figure 2(a)–(c) shows that Technique 2 results in a distribution for these parameters completely distinct from

Table 3. Results obtained with Technique 3.

Parameter	Mean	Standard deviation	Number of interpolating points
R	16.4	1.9	300
$D\,(\mathrm{cm^2\,min^{-1}})$	12.6	1.7	
$h_m\,(\mathrm{cm\,min^{-1}})$	0.41	0.06	
$V\,(\mathrm{cm\,min^{-1}})$	0.58	0.01	
R	17.4	2.0	500
$D\,(\mathrm{cm^2\,min^{-1}})$	13.4	1.7	
$h_m\,(\mathrm{cm\,min^{-1}})$	0.47	0.07	
$V\,(\mathrm{cm\,min^{-1}})$	0.59	0.01	

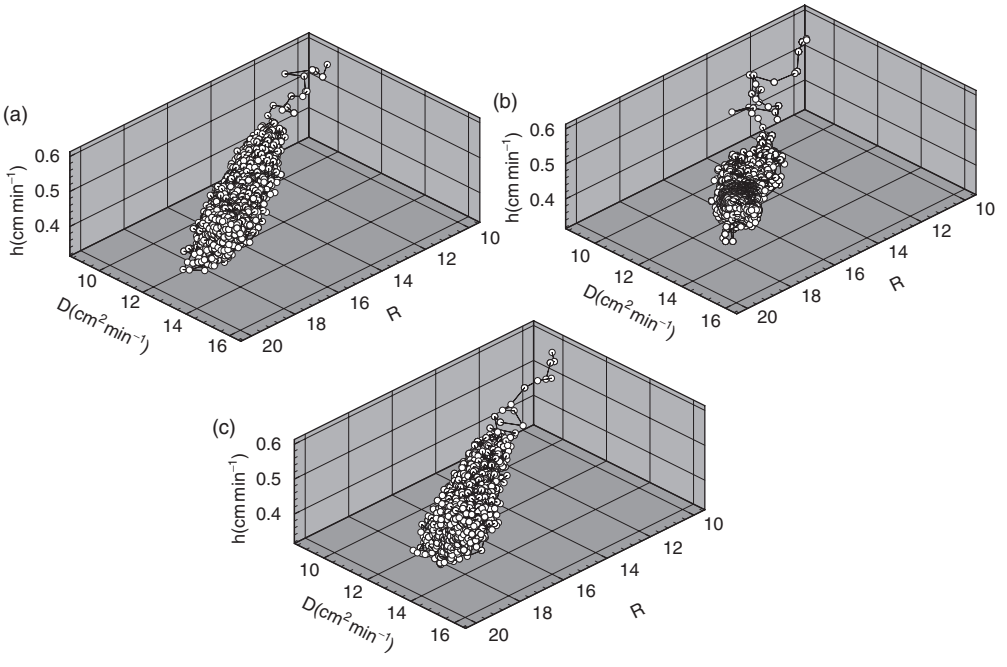

Figure 2. Samples obtained with (a) Technique 1; (b) Technique 2; (c) Technique 3.

Table 4. Computing times.

Technique	Number of interpolating points	Computing time (s)
1	–	1324
2	300	31
	500	52
3	300	38
	500	68

those obtained with Techniques 1 and 3. Such behaviour results from the poor interpolation of the likelihood function with Technique 2. On the other hand, the distributions obtained with Techniques 1 and 3 are quite similar, as a result of the accurate interpolation procedure developed in Technique 3, which automatically selects the most appropriate interpolation function for each specific case.

Table 4 presents the computing times required by the application of Techniques 1–3 to the solution of the present parameter estimation problem. Such computing times correspond to FORTRAN codes running under the Compaq Visual FORTRAN Professional Edition 6.6a platform, in an Intel Centrino Duo T2400 1.83 GHz processor, with 1 Gbyte of RAM memory. Table 4 shows a reduction of at least 20 times when the likelihood function interpolated with RBFs is used in the Metropolis–Hastings algorithm. If we compare Technique 1 with Technique 3, which accurately interpolate the likelihood function with 300 points, the reduction in computing time reaches 35 times. A comparison of the computing times for Techniques 2 and 3 shows a small increase in computational cost when the automatic selection of the interpolation function is applied, as described above, instead of using only multiquadric RBFs.

6. Heat conduction in orthotropic medium

As another example of application of the above interpolation procedure used to reduce the computational cost of the Metropolis–Hastings algorithm, we now consider the inverse problem of estimating the thermal conductivity components of an orthotropic solid. Such an inverse problem was addressed in [16] by using the Levenberg–Marquardt method, and is now revisited with the solution via the Bayesian MCMC approach.

The physical problem considered here involves the three-dimensional linear heat conduction in an orthotropic solid, with thermal conductivity components k_1^*, k_2^* and k_3^* in the x^*, y^* and z^* directions, respectively. The solid is a parallelepiped with sides a^*, b^* and c^*, initially at temperature T_0^*. For times $t^* > 0$, uniform heat fluxes $q_1^*(t)$, $q_2^*(t)$ and $q_3^*(t)$ are applied at the surfaces $x^* = a^*$, $y^* = b^*$ and $z^* = c^*$, respectively. The other three remaining surfaces at $x^* = 0$, $y^* = 0$ and $z^* = 0$ are maintained at a constant temperature (equal to the initial temperature). The mathematical formulation of such physical problem is given in *dimensionless form* by:

$$k_1 \frac{\partial^2 T}{\partial x^2} + k_2 \frac{\partial^2 T}{\partial y^2} + k_3 \frac{\partial^2 T}{\partial z^2} = \frac{\partial T}{\partial t} \quad \text{in } 0 < x < a,\ 0 < y < b,\ 0 < z < c;\ t > 0 \tag{14a}$$

$$T = 0 \text{ at } x = 0; \quad k_1 \frac{\partial T}{\partial x} = q_1(t) \quad \text{at } x = a, \text{ for } \quad t > 0 \tag{14b,c}$$

$$T = 0 \text{ at } y = 0; \quad k_2 \frac{\partial T}{\partial y} = q_2(t) \quad \text{at } y = b, \text{ for } t > 0 \tag{14d, e}$$

$$T = 0 \text{ at } z = 0; \quad k_3 \frac{\partial T}{\partial z} = q_3(t) \quad \text{at } z = c, \text{ for } t > 0 \tag{14f, g}$$

$$T = 0 \quad \text{for } t = 0; \quad \text{in } 0 < x < a,\ 0 < y < b,\ 0 < z < c. \tag{14h}$$

Here, the superscript * denotes dimensional variables and the following dimensionless groups were introduced:

$$t = \frac{k^*_{\text{ref}} t^*}{\rho^* C^* L^{*2}} \qquad T = \frac{T^* - T^*_0}{(q^*_{\text{ref}} L^*)/(k^*_{\text{ref}})} \quad x = \frac{x^*}{L^*} \qquad y = \frac{y^*}{L^*} \qquad z = \frac{z^*}{L^*} \tag{15a–e}$$

$$a = \frac{a^*}{L^*} \qquad b = \frac{b^*}{L^*} \qquad c = \frac{c^*}{L^*} \qquad k_1 = \frac{k^*_1}{k^*_{\text{ref}}} \qquad k_2 = \frac{k^*_2}{k^*_{\text{ref}}} \qquad k_3 = \frac{k^*_3}{k^*_{\text{ref}}} \tag{15f–k}$$

$$q_1 = \frac{q^*_1}{q^*_{\text{ref}}} \qquad q_2 = \frac{q^*_2}{q^*_{\text{ref}}} \qquad q_3 = \frac{q^*_3}{q^*_{\text{ref}}} \tag{15l–n}$$

where $\rho^* C^*$ is the volumetric heat capacity, L^* is a characteristic length of the solid, while q^*_{ref} and k^*_{ref} are reference values for heat flux and thermal conductivity, respectively.

The boundary heat fluxes are supposed to be pulses of finite duration t_h, written in dimensionless form as

$$q_j(t) = \begin{cases} \bar{q}_j & \text{for} \quad 0 < t \le t_h \\ 0 & \text{for} \quad t > t_h \end{cases} \quad \text{for } j = 1, 2, 3. \tag{16}$$

The inverse problem under examination is concerned with the estimation of the vector of unknown parameters $\mathbf{P}^{\mathbf{T}} = [k_1, k_2, k_3]$, by using temperature measurements of sensors located at the centres of the heated surfaces. In reference [16], the D-optimum approach was used in the experimental design in order to estimate these parameters with identical relative accuracies, by appropriately selecting the body dimensions and the magnitudes of the applied heat fluxes. The same optimally experimental variables selected in [16] were used here: for $k_1 = 1$, $k_2 = 15$ and $k_3 = 15$, we have $b/a = c/a = q_2/q_1 = q_3/q_1 = 3.87$ and the heating and final times were $t_{\text{h}} = t_{\text{f}} = 1$.

The solution of the direct problem was obtained analytically with the classical integral transform technique as [17]:

$$T(x, y, z, t) = \frac{8}{abc} \sum_{o=1}^{\infty} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \Theta(\lambda_m, x) \Omega(\beta_n, y) \Psi(\gamma_o, z) \tilde{\tilde{T}} \times (\lambda_m, \beta_n, \gamma_o)(1 - e^{-t(k_1 \lambda_m^2 + k_2 \beta_n^2 + k_3 \gamma_o^2)}) \tag{17a}$$

where

$$\tilde{\tilde{T}}(\lambda_m, \beta_n, \gamma_o) = \frac{\frac{(-1)^{m+1} \bar{q}_1}{\beta_n \gamma_o} + \frac{(-1)^{n+1} \bar{q}_2}{\lambda_m \gamma_o} + \frac{(-1)^{o+1} \bar{q}_3}{\lambda_m \beta_n}}{k_1 \lambda_m^2 + k_2 \beta_n^2 + k_3 \gamma_o^2} \tag{17b}$$

and the eigenfunctions are given by

$$\Theta(\lambda_m, x) = \sin(\lambda_m\, x)$$
$$\Omega(\beta_n, y) = \sin(\beta_n\, y) \tag{17c–e}$$
$$\Psi(\gamma_o, z) = \sin(\gamma_o\, z)$$

with eigenvalues

$$\lambda_m = \frac{(2m-1)}{2a}\pi; \quad \beta_n = \frac{(2n-1)}{2b}\pi; \quad \gamma_o = \frac{(2o-1)}{2c}\pi; \quad m, n, o = 1, 2, \ldots \tag{17f–h}$$

Simulated temperature measurements, obtained with the solution of the direct problem with $m = n = o = 50$, were used in the inverse analysis. Such number of points in the series (17a) was sufficient to achieve the desired convergence in the solution. In order to avoid the inverse crime of using the same solution of the direct problem in the generation of the simulated measurements and in the solution of the inverse problem [2], for the application of the Metropolis–Hastings algorithm we used $m = n = o = 20$ in the series-solution.

Results are presented below for the estimation of the thermal conductivity components $\mathbf{P}^{\mathrm{T}} = [k_1, k_2, k_3]$ with the Metropolis–Hastings algorithm, by using 20,000 samples and neglecting the first 2000. As for the test-case involving the dispersion in soils examined above, a uniform distribution is also used as prior information for the thermal conductivity components. The unknowns were assumed to be in the intervals given by:

$$0.5 \leq k_1 \leq 1.5$$
$$5 \leq k_2 \leq 25$$
$$5 \leq k_3 \leq 25$$

For this test-case, only the results obtained with Techniques 1 and 3 are presented below, since the use of Technique 2 did not result in an accurate interpolation of the likelihood function. We note, however, that for Technique 2 it was also examined the automatic selection of the shape factor for the multiquadrics RBFs that resulted in linear systems approaching ill-conditioning, as recommended in [18].

Figure 3(a) and (b) illustrate the exact and estimated temperatures, as well as the simulated measurements, at the centre of the heated surface at $x = a$, for standard-deviations of the errors of $\sigma = 0.01$ and $\sigma = 0.05$, respectively. The estimated temperatures were obtained with Technique 1.

Figure 4(a) and (b) present the states of the Markov chain resulting from the application of the Metropolis–Hastings algorithm without interpolation (Technique 1), for the cases with standard deviations of the measurement errors of $\sigma = 0.01$ and $\sigma = 0.05$, respectively. The mean parameters and their standard deviations obtained with Technique 1 are presented in Table 5. This table shows that the MCMC Bayesian approach with the Metropolis–Hastings algorithm provided very accurate estimates for the unknown thermal conductivity components, even for large magnitudes of the measurement errors, such as $\sigma = 0.05$. In fact, the agreement between exact and estimated temperatures is also excellent, as shown in Figure 3(a) and (b).

The results obtained with Technique 3 by using 500 interpolating points are presented in Table 6. A comparison of Tables 5 and 6 reveals that the interpolation utilized does not affect significantly the mean and standard deviations for the estimated parameters. In reality, such quantities are quite close to those obtained without

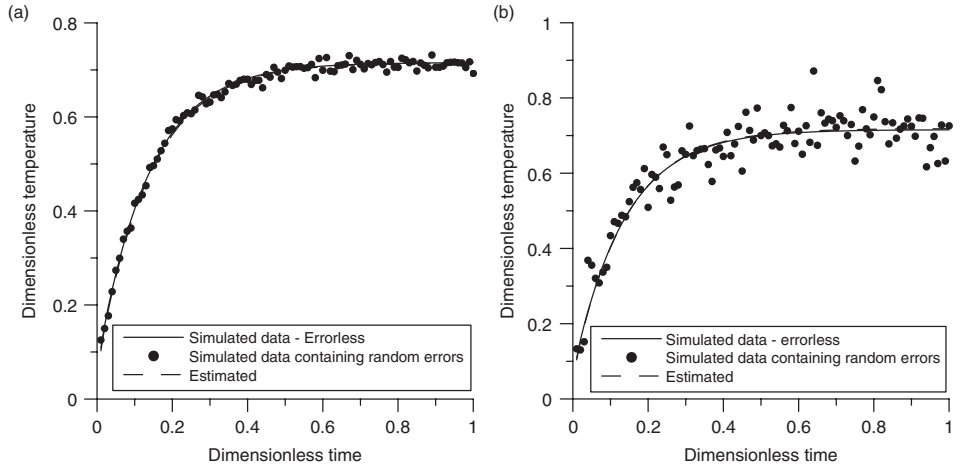

Figure 3. (a) Temperature at the centre of the surface at $x = a$. Simulated measurements with standard deviation $\sigma = 0.01$ (Technique 1). (b) Temperature at the centre of the surface at $x = a$. Simulated measurements with standard deviation $\sigma = 0.05$ (Technique 1).

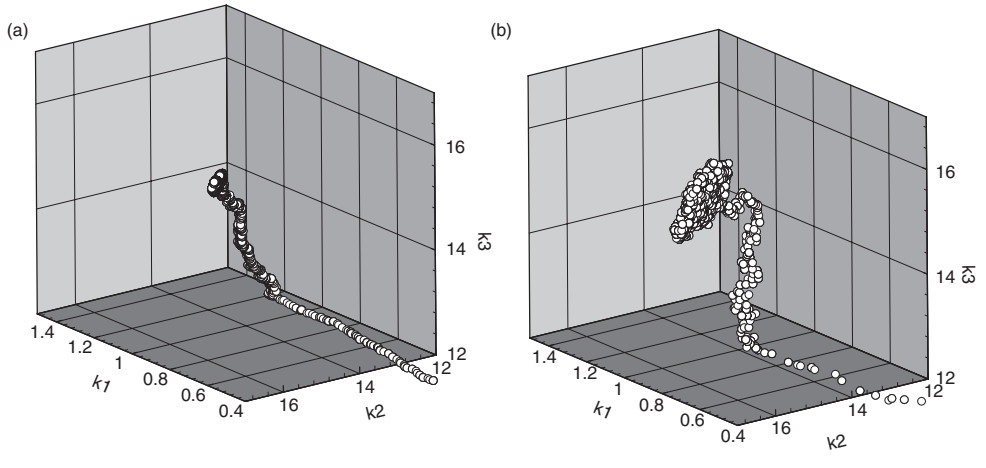

Figure 4. Samples obtained with (a) Technique 1 for standard deviation $\sigma = 0.01$; (b) Technique 1 for standard deviation $\sigma = 0.05$.

interpolation by using Technique 1. This is due to the accurate interpolation provided by Technique 3, which practically does not affect the states of the Markov chain. Such a fact is illustrated in Figure 5(a) and (b) which present the states of the Markov chain obtained with Technique 3, for $\sigma = 0.01$ and $\sigma = 0.05$, respectively.

A comparison of computing times for the present test-case is shown in Table 7. As for the test-case involving the dispersion in soil columns described above, the reduction in computational time with the use of our interpolation procedure was substantial. Even with a large number of interpolation points (500) the computational cost was reduced 25 times with the use of interpolated likelihood function; the computing time dropped from 13.3 h to 31 min.

Table 5. Results obtained with Technique 1.

Standard deviation for the measurements	Parameter	Mean	Standard deviation
$\sigma = 0.01$	k_1	1.000	0.004
	k_2	14.72	0.06
	k_3	14.84	0.07
$\sigma = 0.05$	k_1	0.97	0.02
	k_2	15.1	0.3
	k_3	15.1	0.2

Table 6. Results obtained with Technique 3.

Standard deviation for the measurements	Parameter	Mean	Standard deviation
$\sigma = 0.01$	k_1	0.989	0.004
	k_2	14.79	0.05
	k_3	14.94	0.05
$\sigma = 0.05$	k_1	0.98	0.02
	k_2	14.7	0.3
	k_3	14.9	0.2

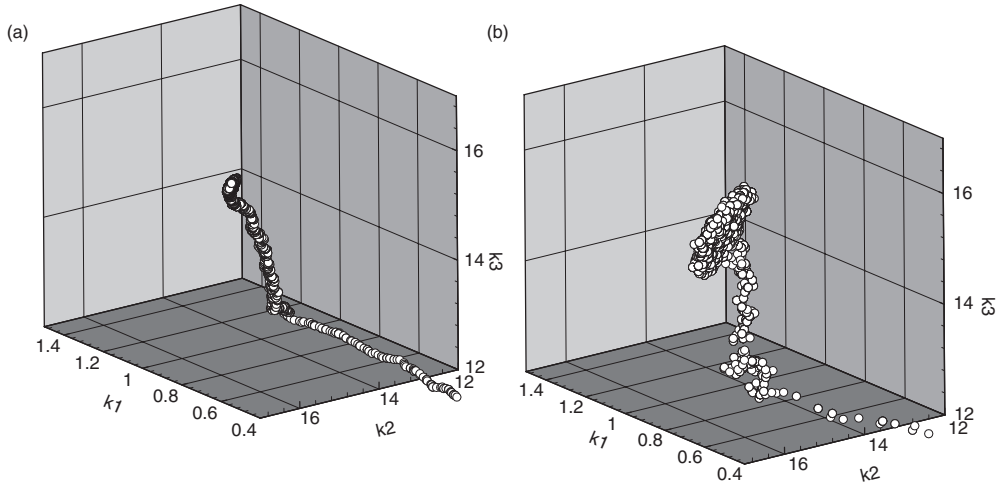

Figure 5. Samples obtained with Technique 3 for standard deviation (a) $\sigma = 0.01$; (b) $\sigma = 0.05$.

7. Conclusions

In this article we applied a Bayesian technique to the solution of inverse parameter estimation problems involving the dispersion of tracers in soil columns and the heat conduction in orthotropic media. The MCMC method, coded in the form of the

Table 7. Computing times.

Technique	Standard deviation for the measurements	Computing time (s)
1	$\sigma = 0.01$	48,008
	$\sigma = 0.05$	47,928
3	$\sigma = 0.01$	1873
	$\sigma = 0.05$	1875

Metropolis–Hastings algorithm, was used to obtain the posterior probability density for the parameters. Two different techniques were then proposed for the interpolation of the likelihood function with RBFs, in order to reduce the computational cost associated with the implementation of the Metropolis–Hastings algorithm. One of these techniques involved only multiquadric RBF for the interpolation, while the other one allowed for automatic selection of the RBF, shape factor and polynomial degree to be used for the interpolation. The automatic selection of such quantities was based on a cross-validation procedure described in the article.

For the two test-cases examined in this work, the use of only multiquadrics RBFs resulted on poor approximation of the likelihood function. As a consequence, the distributions for the parameters were distinct from those obtained with the actual likelihood function. Such was the case even if the automatic selection of the shape factor was used.

On the other hand, substantial reduction on the computational time could be obtained with the interpolation technique that automatically selected the interpolation function and its parameters by cross validation. Such interpolation technique did not cause loss of accuracy on the estimated parameters. In fact, for both test-cases the states of the Markov chain were not significantly affected if the interpolated likelihood function was used instead of the actual function. The use of the interpolated likelihood function shall allow the utilization of MCMC methods in problems requiring large computational times for the solution of the forward problem.

Acknowledgements

This work was mainly sponsored by CNPq, CAPES and FAPERJ, Brazilian agencies for the fostering of science. The authors are thankful to Prof. Vitor Leitão and Prof. Carlos Alves, from Instituto Superior Técnico de Lisboa in Portugal, for bringing to their attention the use of RBFs in different applications. A course on the Statistical Inversion Approach given in COPPE/UFRJ by Prof. Jari Kaipio and Prof. Ville Kolehmainen, from the University of Kuopio, Finland, is greatly appreciated. HRBO and MJC are thankful for the hospitality of Prof. George Dulikravich and his family during their technical visits to FIU. The authors are grateful for the partial financial support provided for this work by the US Air Force Office of Scientific Research under grant FA9550-06-1-0170 monitored by Dr Todd E. Combs, Dr Fariba Fahroo and Dr Donald Hearn and by the US Army Research Office/Materials Division under the contract number W911NF-06-1-0328 monitored by Dr William M. Mullins. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the US Air Force Office of Scientific Research, the US Army Research Office or the US Government. The US Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation thereon.

References

[1] J.V. Beck and K. Arnold, *Parameter Estimation in Engineering and Science*, Wiley Interscience, New York, 1977.
[2] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Applied Mathematical Sciences 160, Springer-Verlag, New York, USA, 2004.
[3] S. Tan, C. Fox, and G. Nicholls, *Inverse Problems*, Course Notes for Physics 707, University of Auckland, Auckland, NZ, 2006.
[4] P.M. Lee, *Bayesian Statistics*, Oxford University Press, London, 2004.
[5] N. Zabaras, *Inverse Problems in Heat Transfer, Chapter 17*, in *The Handbook of Numerical Heat Transfer*, 2nd ed., W.J. Minkowycz, E.M. Sparrow and J.Y. Murthy, eds., John Wiley & Sons, New Jersey, USA, 2004, pp. 525–558.
[6] J. Wang and N. Zabaras, *A Bayesian inference approach to the stochastic inverse heat conduction problem*, Int. J. Heat Mass Transfer 47 (2004), pp. 3927–3941.
[7] J. Wang and N. Zabaras, *Using Bayesian statistics in the estimation of heat source in radiation*, Int. J. Heat Mass Transfer 48 (2005), pp. 15–29.
[8] A. Emery, *The Effect of Correlations and Uncertain Parameters on the Efficiency of Estimating and the Precision of Estimated Parameters,* Chapter 7, in *Inverse Engineering Handbook*, K. Woodbury, ed., CRC, Boca Raton, Florida, USA, 2003.
[9] E.J. Kansa, *Multiquadrics – A scattered data approximation scheme with applications to computational fluid dynamics – II: Solutions to parabolic, hyperbolic and elliptic partial differential equations*, Comp. Math. Applic. 19 (1990), pp. 149–161.
[10] R.L. Hardy, *Multiquadric equations of topography and other irregular surfaces*, J. Geophys. Res. 176 (1971), pp. 1905–1915.
[11] V.M.A. Leitão, *A mesheless method for kirchhoff plate bending problems*, Int. J. Numer. Methods Eng. 52 (2001), pp. 1107–1130.
[12] V.M.A. Leitão, *RBF-Based Meshless Methods for 2D Elastostatic Problems, Eng*, Anal. Bound. Elem. 28 (2004), pp. 1271–1281.
[13] M.J. Colaço, G.S. Dulikravich, and D. Sahoo, *A Comparison of Two Methods for Fitting High Dimensional Response Surfaces*, in *Inverse Problems, Design and Optimization Symposium*, G.S. Dulikravich, H.R.B. Orlande, M.J. Colaco and M. Tanaka, eds., FIU, Miami, Florida, USA, 2007, pp. 749–756.
[14] R. Winkler, *An Introduction to Bayesian Inference and Decision*, Probabilistic Publishing, Gainsville, Florida, USA, 2003.
[15] D.A. Anderson, J.C. Tannehill, and R.H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, Hemisphere, New York, 1984.
[16] M.M. Mejias, H.R.B. Orlande, and M.N. Ozisik, *Effects of the heating process and body dimensions on the estimation of the thermal conductivity components of orthotropic solids*, Inverse Probl. Eng. 11 (2003), pp. 75–89.
[17] M.N. Ozisik, *Heat Conduction*, Wiley, New York, 1993.
[18] E. Divo and A.J. Kassab, *An efficient localized radial basis function meshless method for fluid flow and conjugate heat transfer*, J. Heat Transfer 129 (2007), pp. 124–136.