Inverse Problems, Design and Optimization Symposium
Miami, Florida,U.S.A., April 16-18, 2007

APPROXIMATION OF THE LIKELIHOOD FUNCTION IN THE BAYESIAN TECHNIQUE FOR THE SOLUTION OF INVERSE PROBLEMS

Helcio R. B. Orlande
Department of Mechanical Engineering, POLI/COPPE
Federal University of Rio de Janeiro, UFRJ
Rio de Janeiro, RJ, Brazil
helcio@mecanica.coppe.ufrj.br

Marcelo J. Colaço
Department of Mechanical and Materials Engineering
Military Institute of Engineering
Rio de Janeiro, RJ, Brazil
colaco@asme.org

George S. Dulikravich
Department of Mechanical and Materials Engineering
Florida International University
Miami, Florida, U.S.A.
dulikrav@fiu.edu

ABSTRACT

This work deals with the use of radial basis functions for the interpolation of the likelihood function, in a parameter estimation problem solved with the Bayesian technique. The proposed interpolation of the likelihood function is applied to a test-case involving the estimation of parameters in the dispersion of a tracer in saturated soils. The use of the interpolated likelihood function reduces significantly the computational cost associated with the implementation of Markov Chain Monte Carlo methods to the solution of the present inverse problem.

INTRODUCTION

A variety of techniques is nowadays 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 basically 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 means and covariance matrices, and that the measurement errors are additive, a closed form expression can be derived for the posterior probability density. In this case, the estimator that maximizes the posterior probability density can be recast in the form of a minimization problem involving the *maximum a posteriori objective function*.

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 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 the computational costs.

The use of radial basis functions (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 an inverse parameter estimation problem, which involves the dispersion of a tracer in saturated soils.

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 modeled as random variables.
2. The randomness describes the degree of information concerning their realizations.
3. The degree of information concerning these values is coded in probability distributions.
4. The solution of the inverse problem is the posterior probability distribution.Consider the vector of parameters appearing in the physical model formulation as

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

and the vector of available measurements as

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

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

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

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 Markov Chain Monte Carlo Methods, are required in order to obtain samples that represent accurately 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 [2-4] was used in this work to implement the MCMC method. It can be summarized in the following steps:

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

2. Calculate:

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

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]:

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

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.

INTERPOLATION OF THE LIKELIHOOD FUNCTION

We note that Eq. (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, Eq. (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 radial basis functions (RBFs) as described below [13].

Suppose a function of L variables x_i, i=1,…,L. The RBF model used in this work has the following form

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

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:

$$\sum_{j=1}^{N} \alpha_j p_k(x_i) = 0 \quad (6.a)$$

$$\vdots$$

$$\sum_{j=1}^{N} \alpha_j p_k(x_L) = 0$$

$$\sum_{j=1}^{N} \alpha_j = 0 \quad (6.b)$$

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

$$p_k(x_i) = x_i^k \quad (7)$$

and the radial basis functions are selected among the following

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

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

Squared multiquadrics:

$$\phi\left(\left|\mathbf{x}_i - \mathbf{x}_j\right|\right) = \left(\mathbf{x}_i - \mathbf{x}_j\right)^2 + c_j^2 \quad (8.c)$$

Cubical multiquadrics:

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

with the shape parameter c_j taken as equal to the inverse of the standard-deviation of the measurements, which is supposed constant. The shape parameter is used to control the smoothness of the RBF.

The choice of which polynomial order and which RBF are the best to a specific function, was made based on a cross-validation procedure. Let us supose 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 (6.a,b) are solved for a polynomial of order zero and for each RBF given by Eqs. (8.a-d) using the subset P_{TR1}. 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

$$RMS_{PTR1} = \sum_{i=1}^{P_{TR2}} \left[f_{inter}\left(\mathbf{x}_i\right) - f\left(\mathbf{x}_i\right) \right]^2 \qquad (9.a)$$

where $f_{inter}(\mathbf{x}_i)$ 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 Eqs. (5) and (6.a,b) and the subset P_{TR1} to calculate the error as

$$RMS_{PTR2} = \sum_{i=1}^{P_{TR1}} \left[f_{inter}\left(\mathbf{x}_i\right) - f\left(\mathbf{x}_i\right) \right]^2 \qquad (9.b)$$

Finally, the total error is obtained as

$$RMS_{RBF} = \sqrt{RMS_{PTR1} + RMS_{PTR2}} \qquad (10)$$

This procedure is repeated for all polynomial orders, up to M=6 and for each one of the RBF expressions given by Eqs. (8.a-d). The best combination is the one that returns the lowest value of the RMS error.

TRANSPORT OF TRACERS IN SOIL COLUMNS

In order to examine the approach proposed above to interpolate the likelihood function, we address the identification of soil properties based on the dispersion of a tracer in a column. We consider 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 one-dimensional 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:

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

The initial condition is given by:

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

and the boundary conditions as:

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

$$D\frac{\partial c(L,t)}{\partial z} + h_m\, c(L,t) = h_m C_b \quad \text{at } z = L \text{ and for } t > 0 \qquad (11.d)$$

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

Note in Eq. (11.c) 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 Eq. (11.d) that the boundary condition at z =L was taken as of the third-kind, where h_m is the mass transfer coefficient between the column and an outflow plenum.

For the solution of the problem given by Eqs. (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

$$\mathbf{P} = [\, D\,,\, R\,,\, h_m\,,\, V\,] \qquad (12)$$

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:

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

RESULTS AND DISCUSSIONS

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 cm, by considering the following values for the parameters: R = 14.4, D = 11.08 cm^2/min, h_m = 0.39 cm/min and V = 0.59 cm/min. It is assumed that 90 measurements of the outflow concentration are available for the inverse analysis, taken in intervals of 1 minute.

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

$9 \le R \le 20$

9 cm^2/min $\le D \le$ 20 cm^2/min

0.3 cm/min $\le h_m \le$ 0.6 cm/min

0.58 cm/min $\le V \le$ 0.60 cm/min

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 refered to as: (i) *Technique 1*: Without interpolation; (ii) *Technique 2*: Interpolation with Multiquadrics RBFs; and (iii) *Technique 3*: Iterpolation with RBFs using the cross-validation procedure described above.

The number of samples used in the Metropolis-Hastings algorithm was 20000 and, for the computation of the results, the first 500 samples were neglected.

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 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, which were used for the inverse analysis, and the estimated concentration obtained with the mean values estimated for the parameters, by using Technique 1. The agrement 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 measuremens, as well as due to small and correlated sensitivity coefficients.

Table 1. Results obtained with Technique 1

Parameter	Mean	Standard-Deviation
R	16.7	1.9
D (cm^2/min).	12.7	1.5
h_m (cm/min)	0.50	0.06
V (cm/min)	0.59	0.01

Tables 2 and 3 present the results obtained with Techniques 2 and 3, respectively, for two different number 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 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 (see Tables 1 and 3). Tables 2 and 3 also show that the results obtained with the interpolated likelihood function are not significantly affected by the number of interpolating points used.

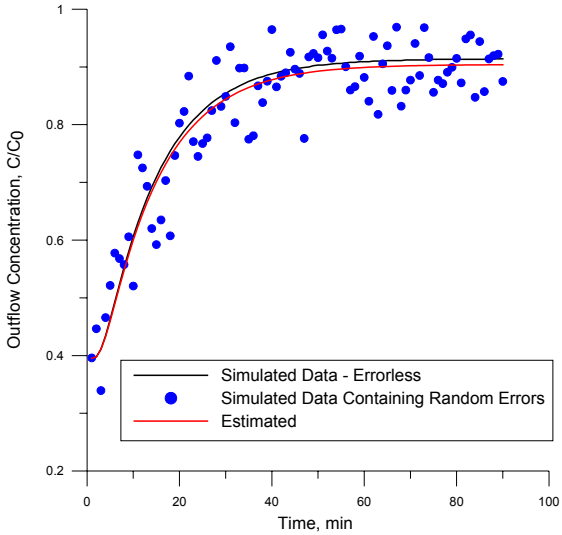

Figure 1. Comparison between measured and estimated concentrations – Technique 1.

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).	14.7	1.6	
h_m (cm/min)	0.57	0.03	
V (cm/min)	0.58	0.01	
R	18.9	0.7	500
D (cm^2/min).	15.2	0.7	
h_m (cm/min)	0.57	0.03	
V (cm/min)	0.59	0.01	

Table 3. Results obtained with Technique 3

Parameter	Mean	Standard-Deviation	Number of Interpolating Points
R	16.4	1.9	300
D (cm^2/min).	12.6	1.7	
h_m (cm/min)	0.41	0.06	
V (cm/min)	0.58	0.01	
R	17.4	2.0	500
D (cm^2/min).	13.4	1.7	
h_m (cm/min)	0.47	0.07	
V (cm/min)	0.59	0.01	

Figures 2-4 present the samples obtained for the retardation factor with Techniques 1-3, respectively. For techniques 2 and 3, 300 points

were used to interpolate the likelihood function with RBFs. A comparison of these figures shows that Technique 2 results on a distribution for this parameter completely distinct from those obtained with techniques 1 and 3. Such behavior 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. Results similar to those presented in figures 2-4 for the retardation factor were obtained for the other three parameters; they are omitted here for the sake of brevity.

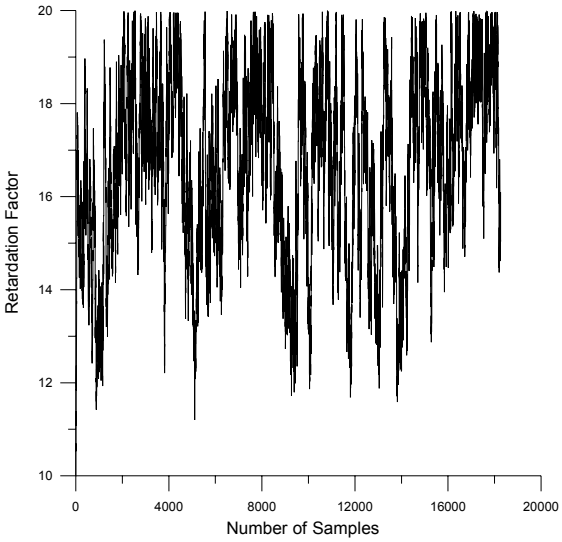

Figure 2. Samples for the Retardation Factor – Technique 1

Table 4 presents the CPU times resulting form the application of Techniques 1-3 to the solution of the present parameter estimation problem. Such CPU 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 Mbyte 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 CPU time reaches 35 times. A comparison of the CPU 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.

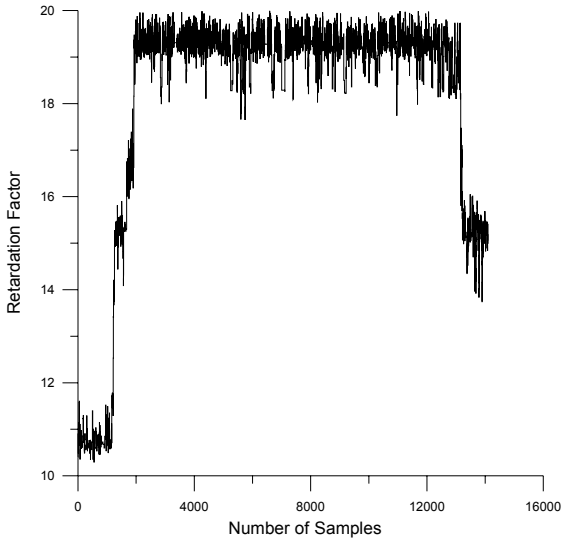

Figure 3. Samples for the Retardation Factor – Technique 2

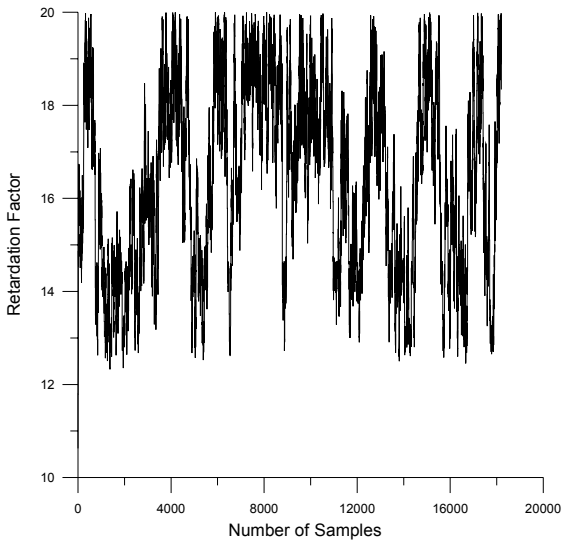

Figure 4. Samples for the Retardation Factor – Technique 3

Table 4 – CPU Times

Technique	Number of Interpolating Points	CPU Time (s)
1	-	1324
2	300	31
	500	52
3	300	38
	500	68

CONCLUSIONS

In this paper we apply the Bayesian approach to the solution of an inverse parameter estimation problem, involving the advection-diffusion of a tracer in soil columns. The interpolation of the likelihood function with radial basis functions is proposed in order to reduce the computational cost associated with the implementation of the Metropolis-Hastings algorithm.

Substantial reduction on the computational time can be obtained with the interpolation procedure developed, without loss of accuracy on the estimated parameters. On the other hand, the use of multiquadrics radial basis functions resulted on poor approximation of the likelihood function and, consequently, on distributions for the parameters distinct from those obtained with the actual likelihood function.

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 radial basis functions in different applications. A course on the Statistical Inversion Approach given in COPPE 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 during their technical visits to FIU.

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 (2004).

3. S. Tan, C. Fox, G. Nicholls, *Inverse Problems*, Course Notes for Physics 707, University of Auckland (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 Edt., John Wiley & Sons, (W.J. Minkowycz, E.M. Sparrow, J. Y. Murthy, editors.) (2004).

6. J. Wang and N. Zabaras, A Bayesian Inference Approach to The Stochastic Inverse Heat Conduction Problem, *International Journal of Heat and Mass Transfer*, 47: 3927–3941 (2004).

7. J. Wang and N. Zabaras, Using Bayesian statistics in the estimation of heat source in radiation, *International Journal of Heat and Mass Transfer*, 48: 15–29 (2005).

8. A. Emery, The effects of correlations and uncertain parameters on the efficiency of estimating and the precision of estimated parameters, in *Inverse Engineering Handbook*, edited by K. Woodbury, CRC Press, Boca Raton, (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*, Comput. Math. Applic., Vol. 19, pp. 149-161, (1990).

10. R.L. Hardy, Multiquadric Equations of Topography and Other Irregular Surfaces, *Journal of Geophysics Res*., Vol. 176, pp. 1905-1915, (1971).

11. V.M.A. Leitão, A Mesheless Method for Kirchhoff Plate Bending Problems, *International Journal of Numerical Methods in Engineering*, Vol. 52, pp. 1107-1130, (2001).

12. V.M.A. Leitão, RBF-Based Meshless Methods for 2D Elastostatic Problems, *Engineering Analysis with Boundary Elements*, Vol. 28, pp. 1271-1281, (2004).

13. M.J. Colaço, G. S. Dulikravich, D. Sahoo, A Comparison Of Two Methods For Fitting High Dimensional Response Surfaces, *Inverse Problems, Design and Optimization Symposium* Miami, Florida,U.S.A., April 16-18, (2007).

14. R. Winkler, *An Introduction to Bayesian Inference and Decision*, Probabilistic Publishing, Gainsville, (2003).

15. D.A. Anderson, J.C. Tannehill, R.H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, Hemisphere, New York, (1984).