SCIENTIFIC NOTEBOOK # 418

A Full-Bayesian Approach To Parameter Inference From Hydraulic Heads And Temperature Data

By:
Allan D. Woodbury
C. Lanczos & Associates Ltd.
Winnipeg, Manitoba, Canada

Contents

1 Initial Entries

Scientific Notebook: #418
Issued to: Allan D. Woodbury
Issue date: August 1, 2000
Printing Period: April 11, 2001 (final printout)
Project Title: A Full-Bayesian Approach To Parameter Inference From Hydraulic Heads And Temperature Data
Project Staff: David A. Farrell and (CNWRA).

By agreement with the CNWRA QA program, this notebook is to be printed at approximately quarterly intervals. This computerized Scientific Notebook is intended to address the criteria of CNWRA QAP-001. [Allan D. Woodbury, November 12, 2000]

2 Objectives and Summary

Within the Amargosa Desert and Fortymile Wash regions adjacent to Yucca Mountain, Nevada, vast areas exist along the projected radionuclide flow path for which little hydrogeologic and geologic data are available. As a result, groundwater flow and mass transport models are poorly constrained within this region. The conjunctive use of temperature and hydrogeological data for site characterization has the potential to reduce the concerns that have been raised. This potential is evident from published works relating thermal-energy transport to the aquifer's hydrogeological features, and from studies that attempted using temperature data for site characterization. Note that the thermal and hydraulic head fields are linked by the fluid specific discharge.

It is the hypothesis of this report that the full-Bayesian techniques mentioned in the above latter references can be successfully adapted to the inverse problem of coupled groundwater flow and heat transport. In the following sections the basic concepts of heat and fluid transfer in porous media are reviewed, followed by an adaptation of a "full-Bayesian" approach for coupled-nonlinear inversion of hydraulic head and thermal data.

A program was developed to illustrate the Bayesian solution to the regression of x against y for a linear relationship. The correspondence to least squares is near perfect and shows verification of the technique.

The procedure is also applied to a series of test cases in which the actual values of temperature and hydraulic head are generated, based on known aquifer characteristics. The test aquifer is identical to that described by Woodbury and Ulrych (2000). It is square, with 300 km sides and constant hydraulic head boundaries. The aquifer transmissivity is homogeneous and for this problem it is theoretically impossible to determine a aquifer value of transmissivity based on hydraulic head data alone. The use of the hydraulic head data is shown to resolve the log-transmissivity estimate, in comparison to hydraulic head data alone.

This notebook documents aspects of the work performed by C. Lanczos & Associates Limited (Dr. A. Woodbury) and CNWRA staff on this project.

2.1 Computers, Computer Codes, and Data Files

The computer codes used in this study are based on a suite of FORTRAN 77 codes developed or acquired by Dr. Allan Woodbury. The data analyses were carried out using computer versions of Windows NT 4.0. Processed data files, FORTRAN code and output will be included on floppy disk with the hard copy of this report.

3 Introduction

Difficulties associated with direct measurement of the hydrologic parameters needed for physically-based mathematical models are well known. Equally well known are the difficulties in the calibration procedure when trying to adjust parameters within preconceived limits until model output at selected points matches observed values. Quite often questions are raised as to the uniqueness and optimality of these models. A major focus of research over the last decade has been directed towards inversion techniques and parameter estimation as a way of both automatic calibration and as a statistical procedure to quantify the reliability of parameter estimates (see reviews by Ginn and Cushman, 1990; and McLaughlin and Townley, 1996, 1997; Kitanidis, 1997). These references

are by no means exhaustive and serve only to indicate the importance of the overall problem of site characterization. The understanding of the problem has improved, and while it is generally considered as yet unsolved (in the sense that no panacea has yet been developed) there are clear ideas of what the weak points are and what might be the remedies. In the inverse approach measurements of hydraulic head, hydraulic conductivity (or transmissivity), seepage flux, and the like are inputs to an inverse algorithm, and fitted hydraulic conductivity (or other parameters) become the output, along with the parameter covariance structure. Throughout this paper the term 'model' is defined as a configuration of material properties. Mathematically, the model takes the form of a vector $\mathbf{m}$ which consists of all unknown parameters (for instance, hydraulic conductivities, boundary conditions, etc.).

Traditionally, inverse techniques in hydrogeology rely on measurements of hydraulic conductivity and hydraulic heads, and they employ the groundwater flow equation for interpretation. Relatively few works have gone beyond this approach and introduced additional information such as tracer data (cf., Carrera et al., 1993), or geophysical measurements (Woodbury and Smith, 1988; Rubin et al., 1992; Hyndman et al., 1994; Copty and Rubin, 1995; Hubbard et al., 1997). The quest for diversifying the types of information stems from the recognition that sophistication of inverse algorithms cannot replace information and data. Along these lines, Abriola et al (1992) noted that "for most specialties, it was generally felt that the state-of-the-art has surpassed the ability to utilize the results in a practical scenario" and that "..applications of [mathematical models] is hindered by the lack of data required to implement or verify them". This recognition is well demonstrated in Carrera and Neuman (1986) where it is shown that the instability and non-uniqueness of solutions to the inverse problems can only be eliminated by introducing additional measurements and information. The challenge of course, is to find inexpensive and reliable sources of information.

The conjunctive use of temperature and hydrogeological data for site characterization has the potential to reduce the concerns that have been raised. This potential is evident from published works relating thermal-energy transport to the aquifer's hydrogeological features, and from studies that attempted using temperature data for site characterization (e.g. Woodbury et al., 1991). Note that the thermal and hydraulic head fields are linked by the fluid specific discharge. This linkage is discussed further below.

The idea of examining thermal fields to gain a better understanding of the hydrologic regime is best illustrated with a simple example. Figure 1 shows the thermal regime in a hypothetical basin for three different values of homogeneous, isotropic permeability: 1.0×10^{-18} m^2, 2.0×10^{-16} m^2, and 5.0×10^{-16} m^2. The basin is 40 km wide and 5 km deep, with a linear water table having a total relief of 500 m. Heat flow within the earth is predominantly vertical in conductive environments. However, because of a coupling between the groundwater and heat flow equations, the thermal field will show increased advective disturbance with increased fluid velocities. The upper plot in Figure 1 illustrates one end member in which the specific discharge is too low to have an effect on the conductive heat flow regime. The temperature within the basin in this simulation is governed entirely by the geometry of the basin, the thermal conductivity of the saturated porous medium, and the basal heat flux. As permeability is increased, fluid velocities become sufficient to redistribute heat in the system. Isotherms in the recharge area are depressed because of the downward flow of cooler water from the water table. In the middle of the basin, isotherms are tilted with respect to their conductive configuration, but for a broad region they remain subparallel to the surface. In this region, equipotentials are near vertical and have little sensitivity to variations in permeability. Isotherms in the discharge area are elevated because of the upward flow of warmer water at depth

Information potentially subject to copyright protection was redacted from this location. The redacted material, Figure 1, is from Smith and Chapman. The complete citation appears on page 21 of this scientific notebook.

Figure 1: From Smith and Chapman, J. Geop. Res., V. 88, p. 593-608, 1983. Copyright by Amer. Geophys. Union.

in the basin.

As mentioned, a number of studies have demonstrated the sensitivity of the thermal field to variations in the magnitude of hydraulic conductivity and anisotropic ratios, length/depth ratios of the flow system, thermal conductivity changes, varying basal heat flux, and the existence and location of aquifers (for example, Smith and Chapman, 1983). Three dimensional effects of fluid flow on the thermal regime have also been examined (Woodbury and Smith, 1985). Smith and Chapman (1983) describe an advective threshold, which is a relatively abrupt transition from a conductive to advectively-dominated system as fluid velocity increases with permeability. This transition occurs over approximately one order of magnitude in permeability for typical large-scale sedimentary basins. At high fluid velocities the thermal regime of a basin can be dominated by advective effects, with a near isothermal temperature field.

The basic idea behind a joint-hydrological/thermal inversion scheme is to exploit the sensitivity

of the thermal field to hydrogeologic parameters. Temperature measurements can be made in piezometers or boreholes along with hydraulic head measurements. Normally only a limited number of hydraulic head measurements can be taken in an individual piezometer. Temperatures, in contrast, can be profiled continuously along the length of a borehole or piezometer standpipe. This feature greatly increases the amount of temperature data compared to hydraulic heads that can be gathered at most sites. In addition, miniature digital temperature measuring equipment is now available and allows for easy use in the field (Woodbury et al., 1991; Woodbury, 1999).

The first works in joint thermal and groundwater inversion were carried out by Woodbury et al. (1987) and Woodbury and Smith (1988). Their joint inversion scheme found solutions by minimizing various multi-objective criteria. Additional unknown parameters such as thermal conductivity and thermal boundary conditions were also identified in the scheme. It is shown that for certain ranges of fluid flow, limitations in hydraulic head data sets, such as incompleteness or inaccuracy, can be overcome with the joint inverse scheme. Wang et al. (1987) advanced the joint inversion methodology by using Bayesian maximum aposteriori (MAP) techniques for a non-linear solution.

In recent publications (cf. Woodbury and Sudicky, 1992; Woodbury and Rubin, 2000) a full-Bayesian approach was used to obtain parameter estimates and variances. The "full-Bayesian" approach signifies that both parameter and hyperparameter determination is involved. It is the hypothesis of this report that the full-Bayesian techniques mentioned in the above latter references can be successfully adapted to the inverse problem of coupled groundwater and heat flow. In the following sections the basic concepts of heat and fluid transfer in porous media are reviewed, followed by an adaptation of a "full-Bayesian" approach for coupled-nonlinear inversion of hydraulic head and thermal data. Specifically, the following are dealt with in this report:

1. Solution to the non-linear thermal/groundwater inverse problem by a full-Bayesian approach.
2. The noise values in the data are assumed unknown a priori and their effects are removed from the problem by marginalization.
3. The prior pdfs for the model parameters, and noise values, are represented by priors developed from Minimum Relative Entropy considerations and are not assumed "ad-hoc".

4 Functional Relationships and Governing Equations

Stallman (1960) presented the governing equations for heat transfer in a saturated porous medium. Bear (1972) and Bear and Corapcioglu (1981) presented complete developments of fluid, momentum and energy transfer in a thermo-elastic medium. The following equations are valid for a single-phase Newtonian fluid in a saturated porous media. Solute concentrations are assumed to be negligible and cross-coupling phenomena such as the Soret and Dufour effect (Bear, 1972) are also assumed to be negligible. Normally, when considering deep regional flow systems, hydraulic conductivity shows significant departure from isothermal behavior because of the dependence of fluid density and viscosity on temperature. These dependencies need to be taken into account for deep groundwater flow systems; however, for the purposes of this work, (and without any loss of generality) these dependencies are neglected.

Only steady-state hydraulic head and temperature fields are considered in this study. This approximation is appropriate where yearly fluctuations in water table are small compared to the

depth of the flow system and where seasonal temperature variations effect only the near-surface temperature field.

4.1 Fluid Flow

This study will focus on the two-dimensional form of the steady-state groundwater flow equation:

$$\nabla \cdot \mathbf{K} \cdot \nabla \phi = \omega \delta(x - x', y - y') \tag{1}$$

subject to

$$-\mathbf{K} \cdot \nabla \phi \cdot \mathbf{n} = q_f \tag{2}$$

on Γ_1, and $\phi = g_1(x, y)$ on Γ_2, $\nabla = (\partial/\partial x, \partial/\partial y)$, $\mathbf{n}$ is the unit outward normal, and $\nabla \cdot = (i\partial/\partial x + j\partial/\partial y)$. Here, ϕ is the hydraulic head, $\omega\delta(x - x', y - y')$ is a fluid source/sink term of strength ω at location x', y', $\mathbf{K}$ is the hydraulic conductivity tensor, q_f is a specified fluid flux term on boundary Γ_1, and $g_1(x, y)$ is a function specifying Dirichlet boundary conditions on Γ_2. If a numerical scheme is used to solve (1), a matrix equation results:

$$\mathbf{A}\mathbf{h} = \mathbf{b}_1 \tag{3}$$

where $\mathbf{h}$ is the approximate value of ϕ due to the discretization, $\mathbf{A}$ is a global stiffness matrix which is a function of $\mathbf{K}$ and grid design, and $\mathbf{b}_1$ is a loading vector with the appropriate boundary conditions.

The fluid specific discharge is given by:

$$\mathbf{q} = -\mathbf{K} \cdot \nabla \phi \tag{4}$$

Heat Transfer

The appropriate heat transfer equation can be written as:

$$\nabla \cdot \Lambda \cdot \nabla T - \rho_f C_f \mathbf{q} \cdot \nabla T = 0 \tag{5}$$

subject to

$$-\Lambda \cdot \nabla T \cdot \mathbf{n} = q_T \tag{6}$$

on Γ_3, and

$$T = g_2(x, y) \tag{7}$$

on Γ_4. Here, T is the temperature, ρ_f is the fluid density, C_f is the fluid specific heat capacity, Λ is the thermal conductivity tensor of the fluid/porous medium composite, q_T is a specified heat flux on boundary Γ_3, and $g_2(x, y)$ is a function specifying Dirichlet boundary conditions on Γ_4. If the external domain to Γ is also a porous medium, and fluid flow occurs across the boundary, then a component of thermal energy is carried with the moving groundwater. In this case, the appropriate boundary condition is of the third type. If the temperature on both sides of the boundary is assumed to be equal, then the boundary condition degenerates to $\nabla T \cdot \mathbf{n} = 0$ (see Bear, 1972, p623).

For the cases under consideration in this study, thermal conductivity is considered to be homogeneous and isotropic. Therefore Λ is treated as a scalar, λ and (7) becomes:

$$\nabla^2 T - D\, \mathbf{q} \cdot \nabla T = 0 \tag{8}$$

where $D = \rho_f C_f / \lambda$ is a reciprocal-thermal diffusivity term. If a numerical scheme is used to solve (8), a matrix equation results:

$$\mathbf{A}_2 \mathbf{u} = \mathbf{b}_2 \tag{9}$$

where $\mathbf{u}$ is the approximate value of T due to the discretization, $\mathbf{A}_2$ is a global stiffness matrix which is a function of λ, grid design and $\mathbf{q}$, and $\mathbf{b}_2$ is a loading vector with the appropriate boundary conditions.

Within the upper crust, heat is transferred primarily in the vertical direction. The equation for the vertical component of heat flow in a conductive environment is:

$$J_z = -\lambda \frac{\partial T}{\partial z} \tag{10}$$

where J_z is the component of heat flow in the vertical direction. Variations in geothermal gradient $(\partial T / \partial z)$ with depth are useful in interpreting and separating advective effects in thermal profiles.

4.2 Aquifer Equations

The basic methodology that will be followed with respect to groundwater inversion is through relatively simple groundwater flow systems in which the hydraulic conductivity and transmissivity is considered statistically isotropic and homogeneous. The hydraulic head ϕ and the aquifer transmissivity $\mathbf{K}b$ satisfy the following partial differential equation in terms of the log-transmissivity, $Y = \ln(Kb)$.

$$\frac{\partial Y}{\partial x}\frac{\partial \phi}{\partial x} + \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial Y}{\partial y}\frac{\partial \phi}{\partial y} + \frac{\partial^2 \phi}{\partial y^2} = 0 \tag{11}$$

For the average specific discharge in the aquifer,

$$q_x = -K \frac{\partial \phi}{\partial x}$$

and

$$q_y = -K \frac{\partial \phi}{\partial y}$$

In terms of transmissivity,

$$q_x = -\frac{T}{b}\frac{\partial \phi}{\partial x} = \alpha \exp(Y) \frac{\partial \phi}{\partial x}$$

and

$$q_y = -\frac{T}{b}\frac{\partial \phi}{\partial y} = \alpha \exp(Y) \frac{\partial \phi}{\partial y}$$

For thermal transport in the aquifer, equation (8) becomes

$$\lambda \nabla^2 T - \rho_f C_f \alpha \exp(Y) \nabla \phi \cdot \nabla T = 0 \tag{12}$$

5 Bayesian Solution To Inverse Problems

As mentioned in the introduction, it is the goal of this work to reconstruct a vector of hydrogeologic model parameters from observations of hydraulic heads and temperatures. This is a non-linear inverse problem, and although for the cases presented in this report there are more data points than unknown parameters, the problem may be ill-posed and potentially lead to non-unique solutions. This is the case, say if one attempts an inversion to determine hydraulic gradients, hydraulic conductivity, porosity and retardation factors from measurements of a tracer cloud. The reader will note that the above parameters are all coupled and their unique determination is very difficult. To so called 'Bayesians', inverse problems are problems of inference and this is the philosophy adopted in this work to circumvent the aforementioned concern.

Much has been written on the subject of Bayesian inference and different points of view apply (for review see Ulrych et al., 2000). The reader will note that we refer to a "Full-Bayesian " approach and this is to signify that the inference problem will consist of both primary parameter and hyperparameter estimation (Mohammad-Djafari, 1996; Woodbury and Rubin, 2000; Woodbury and Ulrych, 2000).

Bayesian inference supposes that an observer can define a prior probability-density function (pdf) for some random variable $\mathbf{m}$. This pdf, $p(\mathbf{m})$, can in principle, be defined on the basis of personal experience or judgment. However, applications of Bayesian probability theory have been hampered by the precise meaning and interpretation of probabilities and controversy surrounding the appropriate choice of prior pdfs. An orthodox view of probabilities dictates that frequencies measured in an experiment are equated to probabilities and 'prior' information is not allowed. An alternative viewpoint of probability, denoted as the Jaynes-Cox viewpoint (Jowitt, 1979), is one in which probabilities are equated with the degree of plausibility of a proposition and may have no frequency interpretation whatsoever. This viewpoint is essentially Bayesian and is readily applicable to the questions that scientists and engineers typically ask. A necessary component of the Jaynes-Cox view is the 'principle of maximum entropy' (PME) which replaces the need for subjective prior information in the Bayesian approach and forces all observers who possess common information to produce consistent results (Woodbury and Ulrych, 1998).

Woodbury and Ulrych (1993), Woodbury et al. (1995) and Woodbury (1997) deal with the estimation of appropriate prior pdf's for hydrogeologic applications. As shown by Woodbury and Ulrych (1993), $p(\mathbf{m})$ may have the form of a multivariate-truncated exponential distribution. This pdf preserves the statistical independence of the parameters. That is, if no correlation is known beforehand the maximum entropy principle does not inject any correlation into the result. In this manner $p(\mathbf{m})$ has the most freedom in assigning realizations of the process. It is important to note that the above approach (PME) of determining $p(\mathbf{m})$ is the one which is the most uncommitted with respect to unknown information.

Bayes' rule (for example; Press, 1989) quantifies how the prior pdf can be changed on the basis of measurements. Simply stated, Bayes' rule is

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Consider a vector of observed data $\mathbf{d}^*$. If the conditional pdf of $\mathbf{d}^*$ given $\mathbf{m}$ and some prior information I, is given by $p(\mathbf{d}^* \mid \mathbf{m}, I)$, then Bayes' rule states that

$$p(\mathbf{m} \mid \mathbf{d}^*, I) = \frac{p(\mathbf{d}^* \mid \mathbf{m}, I)p(\mathbf{m} \mid I)}{\int p(\mathbf{d}^* \mid \mathbf{m}, I)p(\mathbf{m} \mid I)d\mathbf{m}} \tag{13}$$

In the above, $p(\mathbf{m} \mid I)$ is the prior probability density of the model parameters, given some form of prior information, I, and $p(\mathbf{d}^* \mid \mathbf{m}, I)$ is the likelihood of observing $\mathbf{d}^*$ given the model parameters and the prior information. This latter term is often referred to as a 'direct' as opposed to a subjective pdf. The term on the left hand side is called the posterior probability (after measurements are taken into account). Finally the term in the denominator is a constant that ensures the posterior is normalized, but is also the actual pdf of observing a set of data, with the uncertainty in the model parameters taken into account.

In the sections below we will outline how the various conditional pdfs and the prior information are defined and show how we can use Bayes' rule to reconstruct a vector of model parameters from heads and temperature data.

6 Inverse Problem Hydraulic Heads and Temperatures

Consider the finite element model for the hydraulic head predictions in an aquifer. Equation (3) is written in terms of a general non-linear model of the type

$$d_i = f_1(\mathbf{x}_i) \tag{14}$$

for $i = 1 \dots N$ where N is the number of predicted 'data' points and $\mathbf{x} = (x, y)$. Here, $f_1(x)$ depends upon a series of parameters $\mathbf{m}$ which could consist of log-transmissivities, flux conditions and the like.

In the case where head measurements are taken, the associated noise-corrupted case is

$$d_i^* = f_1(\mathbf{x}_i) + \epsilon_i \tag{15}$$

Where the data d_i^* consist of a collection of discrete values of hydraulic heads and ϵ_i is the noise.

The inverse problem consists of trying to reconstruct the parameter vector $\mathbf{m}$, based on the observed data. As mentioned, the inverse problem is viewed in a Bayesian context; that is the inversion is viewed as a problem of inference. In order to solve the inference problem, we will use a Bayesian framework to 'update' a prior probability based on consideration of measurements. To apply Bayes' Theorem we need to assign a noise probability density which is consistent with the available information about the noise. If one could predict the 'true' data, the difference between d_i and d_i^* is just ϵ_i, the noise. If it is assumed that the noise has a value ϵ given prior information I, and if the second moment of the noise is known, σ_1, then an application of the maximum entropy principle leads to a Gaussian distribution for ϵ (Bretthorst, 1988; Kapur, 1989):

$$p(\epsilon|\sigma_1, I) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{\epsilon^2}{2\sigma_1^2}\right) \tag{16}$$

Here σ_1 is taken as the root mean square (RMS) noise level and (16) is the least informative prior probability density for the noise that is consistent with the given second moment. Even if the second-moment of the noise is not known, the central limit theorem leads to the Gaussian form (Jaynes, 1983). In this work (shown later), we treat the noise explicitly as an unknown in Bayes' theorem and then proceed to integrate its effects out.

Having a pdf for the noise and adopting the notation that ϵ_i is the noise at distance x_i, one can apply the product rule of probability theory (assuming independence) to derive the pdf that one would obtain a set of noise values $(\epsilon_1, \epsilon_2, \ldots \epsilon_N)$:

$$p(\epsilon_1, \epsilon_2, \ldots \epsilon_N|\sigma_1, I) = \prod_{i=1}^{N} \left[\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp(-\frac{\epsilon_i^2}{2\sigma_1^2}) \right] \tag{17}$$

Kapur (1989) shows that (17) arises naturally in the multivariate case when entropy is maximized with correlations unknown.

Consider another non-linear model for the temperatures (9) of the type

$$d_j = f_2(\mathbf{x}_j)$$

for an additional M points and the associated noise-corrupted case is

$$d_j^* = f_2(\mathbf{x}_j) + \epsilon_j \tag{18}$$

Here, the model $f_2(\mathbf{x}_j)$ describes the physics of thermal transport and also depends on the same parameters $\mathbf{m}$, as in $f_1(x)$; namely the log-transmissivity, boundary conditions and the like. In this case the noise variance is different than the first N values and is equal to σ_2^2. The data in this case are M observed values of temperature.

In a similar line of reasoning with (17), the noise pdf now becomes

$$\begin{aligned} p(\epsilon_1, \epsilon_2, \ldots \epsilon_N, \epsilon_{N+1}, \epsilon_{N+2}, \ldots e_{N+M}|\sigma_1, \sigma_2, I) \;=\; & \prod_{i=1}^{N} \left[\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp(-\frac{\epsilon_i^2}{2\sigma_1^2}) \right] \\ & \prod_{j=1}^{M} \left[\frac{1}{\sqrt{2\pi\sigma_2^2}} \exp(-\frac{\epsilon_j^2}{2\sigma_2^2}) \right] \end{aligned} \tag{19}$$

Again, if the 'true' model is known, the difference between the data and the model is described by the noise. Taking into account (15) and (18) the pdf that one obtains a set of data $\mathbf{d}^* = (d_1^*, d_2^*, \ldots d_{N+M}^*)$, given a set of parameters and prior information, is proportional to the likelihood function, $\mathcal{L}$:

$$\begin{aligned} p(\mathbf{d}^*|\mathbf{m}, \sigma_1, \sigma_2, I) \;\propto\; & \mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) = \prod_{i=1}^{N} \sigma_1^{-1} \exp(-\frac{1}{2\sigma_1^2}[d_i^* - f_1(x_i)]^2) \\ & \prod_{j=1}^{M} \sigma_2^{-1} \exp(-\frac{1}{2\sigma_2^2}[d_j^* - f_2(x_j)]^2) \end{aligned} \tag{20}$$

or,

$$\mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) = \sigma_1^{-N} \sigma_2^{-M} \times \exp\left\{ -\frac{1}{2\sigma_1^2} \sum_{i=1}^{N} [d_i^* - f_1(x_i)]^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^{M} [d_j^* - f_2(x_j)]^2 \right\} \tag{21}$$

7 Full-Bayesian Solution

A non-linear least-squares approach would proceed by minimizing the combined sums in the argument in the exponential of (21), and the equivalent maximum likelihood procedure finds the parameter set that maximizes the logarithm of (21). Neither approach incorporates prior information about the model parameters. On the other hand, the Bayesian methodology readily lends itself to the problem of updating prior probabilities based on uncertain field measurements. For example, Kitanidis (1986) and Woodbury and Sudicky (1992) outlined the Bayesian approach in which relevant prior information about the model is incorporated. In the current work we adopt a similar approach but following the suggestions of Jaynes and others (for example Kitanidis, 1986; Loredo, 1990; Rubin and Dagan, 1992; Woodbury and Rubin, 2000) we treat the two noise variances σ_1^2, σ_2^2 as 'nuisance' parameters that are "removed" from further consideration by integration over these parameters (marginalization). This point is discussed further below.

7.1 Hyperparameters Unknown

Consider again our set of observed data $\mathbf{d}^*$ and Bayes' Theorem. If the conditional pdf of $\mathbf{d}^*$ given $\mathbf{m}, \sigma_1, \sigma_2$ is given by $p(\mathbf{d}^* \mid \mathbf{m}, \sigma_1, \sigma_2)$, then Bayes' rule states that

$$p(\mathbf{m}, \sigma_1, \sigma_2 \mid \mathbf{d}^*) = \frac{p(\mathbf{d}^* \mid \mathbf{m}, \sigma_1, \sigma_2) p(\mathbf{m}, \sigma_1, \sigma_2)}{\int p(\mathbf{d}^* \mid \mathbf{m}, \sigma_1, \sigma_2) p(\mathbf{m}, \sigma_1, \sigma_2) d\sigma_1 d\sigma_2 d\mathbf{m}} \tag{22}$$

In the above, the I dependence is dropped for convenience. $p(\mathbf{m}, \sigma_1, \sigma_2 \mid \mathbf{d}^*)$ is the conditional pdf of $(\mathbf{m}, \sigma_1, \sigma_2)$, given $\mathbf{d}^*$, and $p(\mathbf{d}^* \mid \mathbf{m}, \sigma_1, \sigma_2)$ represents the direct pdf from observations. $\mathbf{d}^*$ is a vector of $N + M$ data points.

The prior distribution of the model parameters is given by either a multivariate truncated exponential distribution (see Woodbury and Ulrych, 1993), or multivariate Gaussian. The prior pdf for the standard deviation of the noise (σ_1 or σ_2) is given by a Jeffrey's prior $p(\sigma) \propto 1/\sigma$ which has been used by other researchers in Bayesian methods (see Bretthorst, 1988; Woodbury and Ulrych, 2000) to represent complete ignorance of a parameter restricted to zero and infinity. The combined prior for the model and noise is given by

$$p(\mathbf{m}, \sigma_1, \sigma_2) = p(\mathbf{m}) p(\sigma_1) p(\sigma_2)$$

Note that σ_1 and σ_2 is independent of $\mathbf{m}$. Let the posterior pdf be written as:

$$p(\mathbf{m}, \sigma_1, \sigma_2 \mid \mathbf{d}^*) = \frac{1}{\nu} \times p(\mathbf{m}) p(\sigma_1) p(\sigma_2) \mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) \tag{23}$$

where

$$\mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) = \sigma_1^{-N} \sigma_2^{-M} \times \exp\left\{ -\frac{1}{2\sigma_1^2} \sum_{i=1}^{N} [d_i^* - f_1(x_i)]^2 - \frac{1}{2\sigma_2^2} \sum_{j=1}^{M} [d_j^* - f_2(x_j)]^2 \right\} \tag{24}$$

is the likelihood function. This can be rewritten as:

$$\mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) = \sigma_1^{-N} \sigma_2^{-M} \times \exp\left\{ -\frac{H(\mathbf{m})}{2\sigma_1^2} - \frac{T(\mathbf{m})}{2\sigma_2^2} \right\} \tag{25}$$

In order to remove the effects of uncertainty in the measurement noise, integration of (25) is necessary. The marginalized form is:

$$\mathcal{L}(\mathbf{m}) = \int_0^\infty \int_0^\infty \mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) \frac{1}{\sigma_1} \frac{1}{\sigma_2} d\sigma_1 d\sigma_2 \tag{26}$$

After some manipulation, integration yields

$$\mathcal{L}(\mathbf{m}) = 2^{-2+N/2+M/2} \Gamma(N/2) \Gamma(M/2) H(\mathbf{m})^{-N/2} T(\mathbf{m})^{-M/2} \tag{27}$$

which is a multivariate t-distribution. The normalizing constant ν in (23) then, is defined as

$$\nu = \int_{\mathbf{m}} \mathcal{L}(\mathbf{m}) p(\mathbf{m}) d\mathbf{m} \tag{28}$$

The first two moments of $p(\mathbf{m} \mid \mathbf{d}^*)$ (23) are the expected value

$$< \mathbf{m} >= \int_{\mathbf{m}} \mathbf{m} p(\mathbf{m} \mid \mathbf{d}^*) dd\mathbf{m} \tag{29}$$

and the posterior covariance

$$\mathbf{C}_q = \int_M \mathbf{m}\mathbf{m}^T p(\mathbf{m} \mid \mathbf{d}^*) d\mathbf{m} - < \mathbf{m} >< \mathbf{m} >^T \tag{30}$$

Unfortunately, when the forward model is nonlinear, then the integrations posed by (28, - 30) are intractable analytically and one cannot compute marginal densities and moments of the posterior distribution in terms of closed form expressions. However, the integrals in (28 - 30) can be evaluated using Monte-Carlo integration techniques, especially if the size of model space is small and the forward relationship can be computed relatively easily. Note from (28) ν is the expected value of $\mathcal{L}(\mathbf{m})$. Therefore an unbiased estimator of ν can be written as:

$$\nu = \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(\mathbf{m}_i) \tag{31}$$

where the $\mathbf{m}_i$ are generated out of a random population with pdf $p(\mathbf{m})$. Note that the likelihood function is evaluated for each model vector. The mean and covariance can then be calculated as

$$< m_k >= \frac{1}{N\nu} \sum_{i=1}^{N} m_{ik} \mathcal{L}(\mathbf{m}_i) \tag{32}$$

and

$$C_{kj} = \frac{1}{N\nu} \sum_{i=1}^{N} m_{ik} m_{ij} \mathcal{L}(\mathbf{m}_i) - < m_k >< m_j > \tag{33}$$

The above integrations can be accomplished using the method of Monte Carlo integration and the concept of importance sampling (for example; Shreider, 1966, p100-102, or Woodbury and Sudicky, 1992). The integrals posed by (28-30) are evaluated by generating a series of greater than 1,000 random model vectors using a multivariate random number generator with $p(\mathbf{m})$ or $p(\mathbf{m}, \sigma_1, \sigma_2)$ as the pdf. The generation of random points can be terminated when a criterion of relative precision of the integral (28) is satisfied (Tarantola, 1987).

The maximum likelihood point can be determined by examining the set of randomly generated models for that point which maximizes (23). The maximum likelihood point of a non-linear inverse problem may not be the same as the expected value. Because Gaussian distributions are symmetric, the maximum likelihood point coincides with the mean value. For an arbitrary non-linear surface (e.g., multi-modal, skewed, platykurtic) the maximum likelihood point can be far from the mean value.

The Bayesian technique used in this paper essentially samples the posterior pdf surface close to its maximum to find an average value of the model parameters. These values may be more appropriate than a single estimate. If the maximum likelihood point is reasonably close to $<\mathbf{m}>$ then one can place a certain amount of confidence that the posterior pdf is not significantly skewed or multimodal. Higher order statistics such as skew and kurtosis can also be generated by the approach and various tests can be made to determine if the resulting pdf is near Gaussian in form.

The covariance of the parameter estimates is a useful end-product of an inverse scheme. However, because the distribution for $\mathbf{m}$ may be non-Gaussian, the covariance of the estimated model parameters may be difficult to interpret, especially in terms of confidence intervals. In such cases the covariance of the parameters can be interpreted as the covariance from an equivalent linear problem, which may or may not be relevant. In these cases the best understanding may be achieved by determining the probability of a model parameter lying within a certain range. For example

$$P(\mathbf{a} \leq \mathbf{m} \leq \mathbf{b}) = \int_{\mathbf{a}}^{\mathbf{b}} \phi(\mathbf{m} \mid \mathbf{d}^*) d\mathbf{m} \tag{34}$$

which can be determined as part of the numerical integration of (28).

7.2 Hyperparameters Known

In the above section, the treatment of the Bayesian integrals was dealt with, considering that the hyperparameters, σ_1 and σ_2 were unknown, and the only information available was that each one has limits of 0 to infinity. In this section it is assumed that some additional information is available for these hyperparameters. This information could be accumulated by analysis of previous simulations from the section above, or some other approach.

The normalizing constant ν in (23) is defined as

$$\nu = \int_{\mathbf{m}} \int_{\sigma_1} \int_{\sigma_2} \mathcal{L}(\mathbf{m}, \sigma_1, \sigma_2) p(\mathbf{m}) p(\sigma_1) p(\sigma_2) d\sigma_1 d\sigma_2 d\mathbf{m} \tag{35}$$

Here the analysis of Woodbury and Rubin (2000) can be used to deal with the nuisance parameters by marginalization against all other parameters. Integrating (35), the first two moments of $p(\mathbf{m} \mid \mathbf{d}^*)$ (23) are the expected value

$$<\mathbf{m}>= \int_{\mathbf{m}} \mathbf{m} \int_{\sigma_1} \int_{\sigma_2} p(\mathbf{m}, \sigma_1, \sigma_2 \mid \mathbf{d}^*) d\sigma_1 d\sigma_2 d\mathbf{m} \tag{36}$$

and the posterior covariance

$$\mathbf{C}_q = \int_M \mathbf{m}\mathbf{m}^T \int \int p(\mathbf{m}, \sigma_1, \sigma_2 \mid \mathbf{d}^*)\ d\sigma_2 d\sigma_1 d\mathbf{m} - <\mathbf{m}><\mathbf{m}>^T \tag{37}$$

The numerical integrations proceed as before, with the exception that realizations out of $p(\mathbf{m})p(\sigma_1)p(\sigma_2)$ must be generated.

Consider (21) in which both σ_1 and σ_2 are also generated out of some random population. We can rewrite (21) as

$$\begin{aligned}\mathcal{L}(\mathbf{m},\sigma_1,\sigma_2) &= \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{N}[d_i^* - f_1(x_i)]^2\right.\\ &\left.-\frac{1}{2\sigma_2^2}\sum_{j=1}^{M}[d_j^* - f_2(x_j)]^2 - N\times\ln(\sigma_1) - M\times\ln(\sigma_2)\right\}\\ &= \exp(\zeta_i) \qquad (38)\end{aligned}$$

Then ν becomes

$$\nu = \frac{1}{N}\sum_{i=1}^{N}\exp(\zeta_i) \qquad (39)$$

8 Computational Considerations

The recommended procedure is as follows:

1. If the analysis is in the initial phases, no information may be available on the hyperparameters. In this case use $\mathcal{L}(\mathbf{m})$. Define which parameters are important and choose an appropriate prior pdf, $p(\mathbf{m})$. If information is available on the hyperparameters, choose $\mathcal{L}(\mathbf{m},\sigma_1,\sigma_2)$.
2. Generate realizations out of $p(\mathbf{m})$, or $p(\mathbf{m},\sigma_1,\sigma_2)$, increase counter *ic* by one.
3. For each realization, store the values of $\mathbf{m}_i$ and compute and store $\mathcal{L}(\mathbf{m})_i$ or $\mathcal{L}(\mathbf{m},\sigma_1,\sigma_2)_i$. For any realization, if the likelihood is below some threshold value, say 1×10^{-100}, store either of the above. Increase counter, *icount* by one.
4. Terminate realizations when *icount* reaches preset value.
5. Compute relative precision in integral ν.
6. Compute mean values, covariances, credibility intervals.
7. Using the final mean value for the parameters, compute a final run in which data is predicted, compare to observed values at the data points.

9 Verification Example I

A program was developed to illustrate the Bayesian solution to the regression of x against y for a linear relationship. This is a very simple but easy example to understand and allows for the comparison of the Bayesian solution against classical linear regression.

The first example consists of 50 values of x and y. The actual relationship is $y = ax + b$ where $a = 1.0$ and $b = 0$. Fifty values are sampled randomly from $x = 300$ to 500 and the corresponding y values were corrupted with Gaussian, additive noise $N[0,1]$. The fifty values were input to the program INVY to check if the Bayesian code could produce the same, or very similar, regression coefficients for a, b. The priors for the first run are given in Table 1.

The least squares estimates for the parameters can be computed with standard techniques. After 5,000 counting realizations, the Bayesian code with the Gaussian likelihood, produced the estimates listed in Table 2.

This first example illustrates that the Bayes code faithfully reproduced the expected results when compared to regression. Note that regression automatically assumes that there is no prior information, so one has to set very non-informative priors. The numerical accuracy of the integrations was a relative precision of 0.14455.

The second example consists of (again) the same 50 values of x and y, chosen before. The fifty values were input to the program INVY to check if the Bayesian code, with a t-likelihood could produce the same results as the Gaussian likelihood. Recall, the "t" distribution assumes no knowledge of the noise in the data. The priors for the second run are shown in Table 3:

The least squares estimates for the parameters can be computed with standard estimates. After 5,000 counting realizations, the Bayesian code, with the "t" likelihood produced the following estimates listed in Table 4.

This second example illustrates that the Bayes (t) code faithfully reproduced the expected results compared to regression. The numerical accuracy of the integrations was a relative precision of 0.15965. This result in interesting in that the correct solution was found without knowledge of the noise level in the data.

10 Verification Example II

The two-dimensional numerical application of the full-Bayesian approach to the joint inverse problem is tested in this section through simulation. A homogeneous log-transmissivity field is first generated along with its associated hydraulic head field. The hydraulic head values are solved by using the finite element procedure, utilizing linear basis functions and triangular elements. The groundwater specific discharges are first calculated and then used in a second phase (also finite elements) in which the temperature field is calculated. Samples of the computed hydraulic head and temperature fields are taken from the fictitious aquifer and corrupted with Gaussian additive noise (N[0,0.1] in each case). The number of samples are nine for each head and temperature. There is one unknown value of $\ln(T)$ to be solved, along with the noise in each set of measurements. The aquifer postulated as a test case is square, 33 rows by 33 columns, with each side being 300 km long. This test case is identical to that described by Woodbury and Ulrych (2000). The heads are prescribed on every boundary and there are 1089 nodal values of hydraulic head and temperature.

The first example used the Bayes (t) methodology and the results are shown in Tables 5 and 6. There were 1000 realizations and the relative precision of the integration was 0.4419. The priors for log-transmissivity in this case were a lower bound of -14.01, and expected value of -11.006 and an upper bound of -9.00 (truncated exponential type). The results are very good and virtually match the correct log-transmissivity value.

In a second run, the values determined above were used in a second computation (although perhaps not necessary) to try the Bayes solution with Gaussian likelihood. The priors for log-transmissivity in this case were an expected value of -10.9984, standard deviation 0.0816 (Gauss). For σ_1, the prior expected value is 0.0878, standard deviation of 0.0237 (Gauss) and for σ_2, the prior expected value is 0.0805, standard deviation of 0.0288 (Gauss). The is actually not much improvement in the parameter estimates; just a corresponding reduction in the numerical precision

Parameter	Lower	Expected	Upper	PDF Type
a (slope)	0.	0.9965	2.0	TE
b (intercept)	-300.	0.3939	300.	TE
σ (noise)	-	1.0	-	none

Table 1: Parameters and ranges adopted for first verification example. TE refers to truncated exponential distribution.

Parameter	Bayes	Variance	Least Squares	Variance
a (slope)	0.99694	0.356×10^{-5}	0.99662	0.368×10^{-5}
b (intercept)	1.38147	0.777	1.52466	0.79226
σ (noise)	1.02832	-	1.01807	-

Table 2: Output from first verification example and comparison to linear regression

Parameter	Lower	Expected	Upper	PDF Type
a (slope)	0.	0.9964	10.	TE
b (intercept)	-300.	1.53510	300.	TE

Table 3: Parameters and ranges adopted for second verification example

Parameter	Bayes (t)	Variance	Least Squares	Variance
a (slope)	0.99678	0.312×10^{-5}	0.99662	0.368×10^{-5}
b (intercept)	1.26778	0.845	1.52466	0.79226
σ (noise)	1.01785	-	1.01807	-

Table 4: Output from second verification example and comparison to linear regression

Parameter	Bayes (t)	Variance	True value
$\ln(T)$	-10.9984	0.33×10^{-3}	-11.000
σ_1 (head noise)	0.0878	-	0.1
σ_2 (temp noise)	0.0802	-	0.1

Table 5: Output from third verification example

of the integrations.

The numerical accuracy of the integrations was a relative precision of 0.1289. These results confirm that the method can be used to greater improve the resolution of the log-transmissivity of the aquifer over that obtained by hydraulic head information alone.

11 Discussion

Some of the strengths and weaknesses of the proposed approach are highlighted here. The strengths of the technique lie in its stability, being able to handle either over or underdetermined problems, arbitrarily non-linear natures of the forward model(s), and a variety of prior pdfs for the parameters. This leads to tremendous flexibility and being able to incorporate 'soft' geologic information. However, the immediate drawback of the technique is in the computational effort required. If, for example, we needed 10,000 forward model calculations and each one required 0.5 hr., this would indicate that 208 days of simulation time would be required. Therefore advances in the basic approach must be sought to make the full-Bayes technique a practical technique.

If one were to carry out the numerical integration of the Bayes integrals using a uniform prior pdf for each of the parameters, this would lead to an error term in the Monte Carlo integration that would decrease as $1/\sqrt{N}$. However, we already have improved on this basic approach by using the concept of importance sampling. Nevertheless, if more flexible-higher entropy prior pdfs are used, it is possible the maximum likelihood point for any of the parameters my lie out in the tail region of a prior distribution; an area that will only be sampled for a large number of total realizations.

One possibly that is certainly worth exploring is the "Latin Hypercube" sampling (Press et al., 1982). It has been shown that it is possible to greatly reduce the total number of Monte Carlo realizations in integration (McKay et al., 1979), and stochastic simulations (e.g. Lahkim et al., 1999). The idea is that each design parameter is actually tested in every one of its subranges (areas of equal probability). If the response of the system is dominated by one of the design parameters that parameter will be found with this sampling technique. It is recommended that this technique be investigated with respect to the integrals presented herein.

Other possibilities exist to speed up the process. The current SWRI code MULTIFLOW, solves a non-linear forward problem for fluid flow and thermal transport. It is suggested that during each nonlinear forward iteration, the likelihood be calculated. If during these initial phases the likelihood is extremely small, then this realization be immediately skipped over. This passing over is permitted as a result of the numerical integration. Those parameter realizations that produced low likelihoods do not contribute to the sums involved. This suggestion above could also greatly reduce the overall number of computations by discarding runs prior to their overall forward convergence.

12 Summary and Conclusions

This paper presents a methodology for the spatial inversion of transmissivity and other parameters from hydraulic head and thermal data for the two-dimensional steady state groundwater flow case. The methodology used is based on a full-Bayesian approach (Woodbury and Rubin, 2000). The approach deals with estimation of primary parameters and the hyperparameters governing noise in the observations. The pdfs of these hyperparameters are in turn determined from maximum entropy and consistency considerations. In other words, pdfs are chosen for each of the hyperparameters

that are maximally uncommitted with respect to unknown information. In the algorithm provided, the user selects a set of parameters and hyperparameters. Prior pdfs for these are chosen. Next, a Monte-Carlo procedure is followed in which a series of parameter (and possibly) hyperparameter values are generated. Each of the realizations and likelihoods of observing the 'data' are kept. When a criterion of relative accuracy is achieved the iterations stop and then expected values of the parameters and covariances are computed.

A program was developed to illustrate the Bayesian solution to the regression of x against y for a linear relationship. This is a very simple but easy example to understand and allows for the comparison of the Bayesian solution against classical linear regression. The first example consists of 50 values of x and y. The actual relationship is $y = ax + b$ where $a = 1.0$ and $b = 0$. Fifty values are sampled randomly from $x = 300$ to 500 and the corresponding y values were corrupted with Gaussian, additive noise $N(0,1)$. The fifty values were input to the program INVY to check if the Bayesian code could produce the same, or very similar, regression coefficients for a, b. The correspondence to least squares is near perfect and shows verification of the technique.

The procedure is also applied to a series of test cases in which the actual values of temperature and hydraulic head are generated, based on known aquifer characteristics. The test aquifer is identical to that described by Woodbury and Ulrych (2000). It is square, with 300 km sides and constant hydraulic head boundaries. The aquifer transmissivity is homogeneous and for this problem it is theoretically impossible to determine a aquifer value of transmissivity based on hydraulic head data alone. In a first series of tests nine temperature and hydraulic head values are used and are corrupted with N(0, 0.1) additive noise. In this series the hyperparameters are assumed unknown and the full-Bayesian produced is used. The use of the hydraulic head data is shown to resolve the log-transmissivity estimates, in comparison to hydraulic head data alone.

13 REFERENCES

Abriola, L.M., A.H. Demond, K.E. Hayes, 1992. *Summary Report: Second forum on NSF Research activities in Subsurface Systems*, The Univ. of Michigan at Ann Arbor, MI..

Bear, J., 1972. *Dynamics Of Fluids In Porous Media*, Elsevier, New York.

Bear, J. and M.Y. Corapcioglu, 1981. A mathematical model for consolidation in a thermoelastic aquifer due to hot water injection or pumping, *Water Resour. Res.*, **17(3)**, 723-736.

Bretthorst, G. L., 1988. Bayesian Spectrum Analysis and Parameter Estimation. In *Lecture Notes in Statistics No. 48*, (J. Berger, S. Fienberg, J. Gani, K. Krickeberg, and B. Singer (eds)), 209p.

Carrera, J., and S.P. Neuman, 1986. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information, *Water Resour. Res.*, **22(2)**, 199-210.

Carrera, J., A. Medina, X. Sanchez-Vila, 1993. *In Quantitative Geology and Geostatistics, 5, 779-792*, Kluwer Publications.

Copty, N. and Y. Rubin, 1995. A stochastic approach to the characterization of lithofacies from surface seismic and well data, *Water Resour. Res.*, **31(7)**, 1673-1686.

Ginn, T. R. and J. H. Cushman, 1990. Inverse methods for subsurface flow - A critical review of stochastic techniques, *Stochastic Hydrology and Hydraulics*, **4(1)**, 1-26.

Hubbard, S.S., Rubin, Y; Majer, E., 1997. Ground-penetrating-radar-assisted saturation and permeability estimation in bimodal systems, *Water Resour. Res.*, **33(5)**, 971-990.

Hyndman, D. W., Harris, J. M. and S. M. Gorelick, 1994. Coupled seismic and tracer test inversion for aquifer property characterization, *Water Resour. Res.*, **30(7)**, 1965-1977.

Jaynes, E. T., 1983. *Papers on Probability, Statistics and Statistical Physics*, D. Redel, Dordrecht-Holland.

Jowitt, P.W., 1979. Bayesian estimates of material properties from limited test data, *Eng. Struct.*, **Vol. 1**, 170-178.

Kapur, J.N., 1989. *Maximum Entropy Models In Science and Engineering*, Wiley, N.Y., 635p.

Kitanidis, P.K., 1986. Parameter uncertainty in estimation of spatial functions: Bayesian analysis, *Water Resour. Res.*, **22(4)**, 499-507.

Kitanidis, P.K., 1997. A reassessment of the groundwater inverse problem - Comment., *Water Resour. Res.*, **33(9)**, 2199-2202.

Lahkim, M.B., Garcia, L.A. and J.R. Nuckols, 1999. Stochastic modeling of exposure and risk in a contaminated heterogeneous aquifer. 2: Application of Latin Hypercube sampling, *Env. Eng. Sc.*, **16(5)**, 329-343.

Loredo, T.J., 1990. From Laplace to supernova SN 1987A: Bayesian Inference. In *Maximum Entropy And Bayesian Methods*, (P. F. Fougere), Kluwer Academic Publishers.

McKay, M.D., Beckman, R.J., and W.J. Conover, 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, **21(2)**, 239-245.

McLaughlin, D. and L.R. Townley, 1996. A reassessment of the groundwater inverse problem, *Water Resour. Res.*, **32(5)**, 1131-1161.

McLaughlin, D and L.R. Townley, 1997. A reassessment of the groundwater inverse problem - Reply., *Water Resour. Res.*, **33(9)**, 2203-2204.

Mohammad-Djafari, A. , 1996. A full Bayesian approach for inverse problems. In *Maximum Entropy And Bayesian Methods*, (K. M. Hanson and R. N. Silver), Kluwer Academic.135-144

Press, W.H., Flannery, B.P., Teukolsky, S.A., and W.T. Vetterling, 1992. *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press.

Press, S. J., 1989. *Bayesian Statistics: Principles, Models and Applications*, Wiley, New York, 237p..

Rubin, Y. and G. Dagan, 1992. Conditional estimation of solute travel time in heterogeneous formations - Impact of transmissivity measurements, *Water Resour. Res.*, **28(4)**, 1033-1040.

Rubin, Y., Mavko, G. and J. Harris, 1992. Mapping permeability in heterogeneous aquifers using hydrologic and seismic data, *Water Resour. Res.*, **28(7)**, 1809-1816.

Schreider, Yu. A., 1966. *The Monte Carlo Method*, Pergamon Press, 381pp.

Smith, L. and D.S. Chapman, 1983. On the thermal effects of groundwater flow, 1 Regional scale systems, *J. Geophys. Res.*, **88(B1)**, 593-608.

Stallman, R.W., 1960. Notes on the use of temperature data for computing groundwater velocity. In *6th Assembly on Hydraulics*, (Soc. Hydrotech. France), Nancy, June 28-30.

Tarantola, A., 1987. *Inverse Problem Theory*, Elsevier, 613p.

Ulrych, T.J., Velis, D.R., Sacchi, M.D. and A.D. Woodbury, 2000. *A Bayes tour of inversion: a tutorial*, Accepted, Geop. J..

Wang, K., Shen, P-Y., and A. Beck, 1989. A solution to the inverse problem of coupled hydrological and thermal regimes. In *Hydrogeologic Regimes And Their Subsurface Thermal Effects*, (Editors A. E. Beck, G. Garven and L. Stegena), AGU Monograph Series Vol 47/IUGG 7.p 7-21

Woodbury, A.D. and L. Smith, 1985. On the thermal effects of three dimensional groundwater flow, *J. Geophys. Res.*, **90(B1)**, 759-767.

Woodbury, A.D., J.L. Smith, and W.S. Dunbar , 1987. Simultaneous inversion of temperature and hydraulic head data, 1. Theory and application using hydraulic head data , *Water Resour. Res.*, **23(8)**, 1586-1606.

Woodbury, A.D. and J.L. Smith, 1988. Simultaneous inversion of temperature and hydraulic head data, 2. Application with thermal data , *Water Resources Research* , **24(3)**, 356-372.

Woodbury, A. D., Narod, B., Chandra, B., and B. Bennist , 1991. Temperature measurements in geotechnical studies using low-noise high-resolution digital techniques, *Canadian Geotechnical Journal*, **28(5)**, 639-649.

Woodbury, A.D., and E. A. Sudicky, 1992. Inversion of the Borden tracer experiment data: Investigation of stochastic moment models, *Water Resour. Res.*, **28(9)**, 2387-2398.

Woodbury, A.D., and T. J. Ulrych, 1993. Minimum relative entropy: Forward probabilistic modeling , *Water Resour. Res.*, **29(8)**, 2847-2860.

Woodbury, A. D., 1997. A probabilistic fracture transport model: Application to contaminant transport in a fractured clay deposit, *Can. Geotech. J.*, **34**, 784-798.

Woodbury, A.D., 1999. *Simultaneous inversion of thermal and hydrologic data: Where do we go from here*, Eos Trans. AGU 80(46), F337 .

Woodbury, A.D., F.W. Render and T.J. Ulrych, 1995. Practical probabilistic groundwater modeling , *Groundwater*, **33(4)**, 532-538.

Woodbury, A.D., and T.J. Ulrych , 1998. Minimum relative entropy and probabilistic inversion in groundwater hydrology, *Stoch. Hydrol & Hydr.*, **12**, 317-358.

Woodbury, A.D., and Y. Rubin, 2000. A Full-Bayesian approach to parameter inference from tracer travel time moments, *Water Resources Research*, **36(1)**, 159-171.

Woodbury, A.D. and T.J. Ulrych, 2000. A Full-Bayesian approach to inference problem for steady-state flow, *Water Resour. Res.*, **36(8)**, 2081-2093.

14 Figure Captions

Figure 1 Thermal effects of groundwater flow. Adapted from Smith and Chapman, J. Geop. Res., V. 88, p. 593-608.

Parameter	Bayes (t)	Variance	True value
$\ln(T)$	-11.0050	0.36×10^{-3}	-11.000
σ_1 (head noise)	0.0904	0.262×10^{-3}	0.1
σ_2 (temp noise)	0.0914	0.254×10^{-3}	0.1

Table 6: Output from fourth verification example

I have reviewed this scientific notebook and find it in agreement with QAP-001. There is sufficient information regarding methods used for conducting tests, acquiring and analyzing data so that another qualified individual could repeat the activity.

E. C. P

Element Manager

7/26/2001

Date