Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 281 (2003) 265–280
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Non-uniqueness of inverse transmissivity field calibration and predictive transport modeling

Sean A. McKenna[a,*], John Doherty[b], David B. Hart[c]

[a]*Department of Geohydrology, Sandia National Laboratories, P.O. Box 5800 MS 0735, Albuquerque, NM 87187-0735, USA*
[b]*Watermark Numerical Computing, 336 Cliveden Ave., Corinda 4075, Australia*
[c]*Department of Computer Engineering, Utah State University, Logan, UT, USA*

Received 17 August 2001; revised 15 May 2002; accepted 8 May 2003

Abstract

Recent work with stochastic inverse modeling techniques has led to the development of efficient algorithms for the construction of transmissivity (T) fields conditioned to measurements of T and head. Small numbers of calibration targets and correlation between model parameters in these inverse solutions can lead to a relatively large region in parameter space that will produce a near optimal calibration of the T field to measured heads. Most applications of these inverse techniques have not considered the effects of non-unique calibration on subsequent predictions made with the T fields. Use of these T fields in predictive contaminant transport modeling must take into account the non-uniqueness of the T field calibration. A recently developed 'predictive estimation' technique is presented and employed to create T fields that are conditioned to observed heads and measured T values while maximizing the conservatism of the associated predicted advective travel time. Predictive estimation employs confidence and prediction intervals calculated simultaneously on the flow and transport models, respectively. In an example problem, the distribution of advective transport results created with the predictive estimation technique is compared to the distribution of results created under traditional T field optimization where model non-uniqueness is not considered. The predictive estimation technique produces results with significantly shorter travel times relative to traditional techniques while maintaining near optimal calibration. Additionally, predictive estimation produces more accurate estimates of the fastest travel times.

Keywords: Groundwater; Geostatistics; Inverse modeling; Advective transport

1. Introduction

In recent years, the practice of inverse parameter estimation has become relatively commonplace in the groundwater modeling field. Inverse parameter estimation has been applied to a wide variety of problems including hydraulic test analysis, basin scale modeling, tracer test analysis and the estimation of spatially variable property fields. It is this latter area that is the focus of this paper. The essence of the problem is to estimate the value of a spatially variable property, usually transmissivity, T, at every grid cell in a model domain based on a limited set of spatially distributed T measurements and measurements of hydraulic head at the same or other locations.

* Corresponding author. Tel.: +1-505-844-2450; fax: +1-505-844-7354.
E-mail addresses: samcken@sandia.gov (S.A. McKenna), jdoherty@gil.com.au (J. Doherty), dbhart@cc.usu.edu (D.B. Hart).

0022-1694/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0022-1694(03)00194-X

A number of approaches to this inverse estimation problem have been developed. A critical review and assessment of various techniques from the perspective of functional analysis is given by McLaughlin and Townley (1996). A series of blind tests of seven geostatistically based inverse estimation approaches on exhaustively known, synthetic transmissivity fields is reported in Zimmerman et al. (1998). These two papers, as well as other review articles, provide general background and highlight the strengths and weaknesses of different inverse approaches to the problem of estimating spatially variable property fields.

One inverse technique, generally known as the 'pilot point' method, has been used by a number of authors (Certes and De Marsily, 1991; Lavenue and Pickens, 1992; RamaRao et al., 1995) to estimate heterogeneous transmissivity fields. The pilot point technique consists of selecting locations (the pilot points) within a domain at which the value of T is adjusted. The values of T at locations surrounding each pilot point are also adjusted through a kriging procedure using a spatial covariance model estimated from the original T measurements. The values of T at each pilot point are calculated so as to reduce the mismatch between the observed and modeled value of head. Early applications of the pilot point technique determined the locations of the pilot points subjectively (Certes and De Marsily, 1991) or added pilot points iteratively at locations where the modeled heads were most sensitive to the local T value (Lavenue and Pickens, 1992). Recent work has extended the pilot point method to three-dimensional applications (Hendricks Franssen et al., 1999; Lavenue and de Marsily, 2001). Advantages of the pilot point technique over other inverse parameter estimation techniques are that it does not require a small log T variance or zonation of the model domain and it can incorporate qualitative information through the initial estimate of T at the pilot point locations. The main disadvantage of the pilot point technique is that it can be computationally expensive.

An extension of the pilot point approach, known as the *self-calibrated algorithm*, has been developed by Gómez-Hernández et al. (1997) and demonstrated on synthetic fields (Capilla et al., 1997) and at the Waste Isolation Pilot Plant site in southern New Mexico, USA (Capilla et al., 1998). This extension allows for the simultaneous estimation of T at multiple 'master locations' (essentially pilot points). The work by Gómez-Hernández and others shows that a large number of master locations, even greater than the number of original T observations, located on a nearly uniform grid, can produce T fields that are well calibrated to the observed heads. As pointed out by McLaughlin and Townley (1996), large numbers of master locations/pilot points can overparameterize the model, making it necessary to somehow regularize the parameters. Regularization amounts to constraining the optimization of the T fields with bounds on the estimated T values (see examples in Capilla et al. (1997)) or minimizing the difference between estimated T values at adjacent pilot points.

Most documented applications of the pilot point technique and its extensions have been motivated by the need to produce T fields, calibrated to measured head data, that can be used as the basis for contaminant transport models. Uncertainty in the spatial distribution of T can be addressed through a stochastic approach where multiple, equiprobable realizations of the T field are created (RamaRao et al., 1995). Most authors acknowledge that the estimation of T fields from a limited number of T and head measurements is an ill-posed problem and leads to non-unique solutions (i.e. a variety of different solutions). All of these solutions satisfy a geostatistical characterization of the model domain, as encapsulated in a variogram, and all yield nearly identical model-generated heads at the locations of the observed heads. To date, a quantitative technique for addressing the effect of T field non-uniqueness on model-calculated transport results has not been presented.

The objective of this paper is to present and demonstrate an approach for addressing the relationship between non-uniqueness in estimated T fields and resulting groundwater transport calculations. This objective is accomplished by using an inverse parameter estimation approach to stochastic T field modeling that is similar to that developed by Gómez-Hernández et al. (1997) and can include a strong regularization constraint. Additionally, the approach of simultaneous determination of confidence and prediction intervals introduced by Vecchia and Cooley (1987) is applied to the evaluation of nonlinear confidence intervals of transport outcomes as calculated by a flow and transport model. In the case examined here, the major contributor to the uncertainty

in transport predictions is the transmissivity field heterogeneity. Combining inverse T field generation and simultaneous calculation of confidence and prediction intervals into a 'predictive estimation' approach allows for generation of a calibrated T field while maximizing the conservatism of transport predictions based upon that T field. This approach to addressing the non-uniqueness of T fields with respect to transport predictive error, as implemented in the PEST parameter estimation software (Doherty, 2000), is demonstrated on a hypothetical example.

2. Modeling approach

We employ a three-tiered approach to T field generation and advective transport modeling to examine the effects of different calibration data and calibration constraints on the overall nature of the resulting T fields and on the distributions of advective travel times. The three approaches are: (1) generation of *seed* transmissivity fields conditioned only to the available T measurements, (2) *calibration* to both T and head measurements, and (3) *predictive estimation* for calibration to T and head measurements under the constraint of minimizing an advective travel time prediction while quasi-maintaining the calibration.

2.1. Seed transmissivity fields

The parameter to be estimated is the $\log_{10}$ value of the transmissivity, denoted as Z. Multiple, equally probable, two-dimensional realizations of the spatial distribution of Z are created using a multiGaussian (MG) geostatistical simulation algorithm. For each (x, y) location within the model domain, a conditional Gaussian cumulative distribution function is created from the ordinary kriging (OK) estimate of the mean (Z^*) and the variance (σ^2) of Z based on n surrounding data points. The OK estimate of the mean is a weighted linear combination of surrounding original and previously simulated data

$$Z^*_{\mathrm{OK}}(x,y) = \sum_{i=1}^{n(x,y)} \lambda_i^{\mathrm{OK}}(x,y) Z(x,y)_i, \tag{1}$$

where λ_i^{OK} is the kriging weight applied to the ith surrounding Z datum. The kriging weights are determined from solution of the kriging system under a single unbiasedness constraint such that:

$$\sum_{i=1}^{n(x,y)} \lambda_i^{\mathrm{OK}}(x,y) = 1.0. \tag{2}$$

The corresponding OK estimate of the variance of the local cdf of Z^* is

$$\sigma^2_{\mathrm{OK}}(x,y) = C(0) - \sum_{i=1}^{n(x,y)} \lambda_i^{\mathrm{OK}}(x,y) C((x,y)_i - (x,y)) - \mu_{\mathrm{OK}}(x,y), \tag{3}$$

where $C(x, y)$ is the covariance function defined by the variogram and under the assumption of second-order stationarity and the argument to the covariance function is the separation distance between the location being simulated and a surrounding data location. The Lagrange parameter, μ_{OK} is included in the kriging system to account for the unbiasedness constraint. At each location, an independent random number in [0,1] is drawn and the corresponding Z value from the conditional cumulative distribution function is assigned. The assigned Z then becomes part of the dataset used for generation of Z at other locations within the model domain. The process is repeated until every location has been simulated and the realization is complete. More details on the formulation of the kriging equations and their application in stochastic simulation can be found in Goovaerts (1997) and Deutsch and Journel (1998).

It is noted that the assumption of an MG model for a contaminant transport problem has been shown to be a non-conservative choice of conceptual model (Gómez-Hernández and Wen, 1998). However, for the example application shown herein, the choice of an MG model allows for ease of computation and also allows us to focus on non-uniqueness in the estimated T fields as a previously unexplored issue in conservative predictions of transport performance. Extension of the approaches shown herein to discrete fields created through indicator geostatistical simulation is relatively straightforward (Doherty, 1998).

2.2. Transmissivity field calibration

In solving the inverse problem, the transmissivity fields created through geostatistical simulation are

updated through calibration to the available observations of head made in observation wells. For this work, we consider the case of the head observations being collocated with the T measurements used to condition the seed realizations. A number of pilot points are located throughout the model domain and the Z values at the pilot point locations are then adjusted such that an 'optimal Z field' is determined. The values of Z around the pilot points are also updated through the spatial covariance function defined by the variogram. The original measured values of Z are not modified during the calibration process. The optimal parameter set is defined as that for which the sum of squared errors (SSE) between model-generated observations and experimental observations is reduced to a minimum. The objective function, Φ, is defined as

$$\Phi = (\mathbf{c}' - \mathbf{X}\mathbf{b})^{\mathrm{t}}\mathbf{Q}(\mathbf{c} - \mathbf{X}\mathbf{b}), \tag{4}$$

where the vector $\mathbf{b}$ holds the model parameters (Z values at the pilot points), the matrix $\mathbf{X}$ is an m (number of observations) by n (number of parameters) operator that contains equations for the solution of the groundwater flow model as a function of the parameters describing the model domain as well as the discretization scheme and the boundary conditions, $\mathbf{c}'$ contains the set of field measurements and $\mathbf{Q}$ is an observation weight matrix. The observations, generally measurements of head, can be equally weighted, as is done in this work, or different weights can be assigned to capture varying levels of measurement error in each observation. The 't' superscript indicates the matrix transpose operation. If $\mathbf{X}$ is a linear model with respect to parameters, the vector $\mathbf{b}$ that minimizes Φ of Eq. (4) is given by (Bard, 1974):

$$\mathbf{b} = (\mathbf{X}^{\mathrm{t}}\mathbf{Q}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{t}}\mathbf{Q}\mathbf{c}'. \tag{5}$$

Solution of the inverse problem in the realm of nonlinear models can be done with a simple extension of Eq. (5) provided that the function, M, that maps n-dimensional parameter space into m-dimensional observation space is continuously differentiable with respect to all estimated model parameters. First the inverse problem must be linearized. To do this, suppose that a set of parameters comprising the vector $\mathbf{b}_0$ results in a set of 'model-calculated observations' $\mathbf{c}_0$ through the equation:

$$\mathbf{c}_0 = M(\mathbf{b}_0). \tag{6}$$

Using Taylor's theorem, a set of observations, $\mathbf{c}$, corresponding to a parameter vector, $\mathbf{b}$, that differs only slightly from $\mathbf{b}_0$ can be generated using the approximate relationship

$$\mathbf{c} = \mathbf{c}_0 + \mathbf{J}(\mathbf{b} - \mathbf{b}_0), \tag{7}$$

where $\mathbf{J}$ is the Jacobian matrix of M, i.e. the matrix composed of m rows (one for each observation), the n elements of each row being the derivatives of one particular observation with respect to each of the n parameters. Therefore, Eq. (7) is a linearization of Eq. (6).

In implementing this relationship for the solution of the inverse problem, the goal is to derive a set of model parameters, $\mathbf{b}$, for which the model-generated observations, $\mathbf{c}$, are as close as possible to the set of experimental observations, $\mathbf{c}'$, in the least squares sense given the initial set of parameter estimates encapsulated in the vector $\mathbf{b}_0$. The objective function to be minimized is now defined as:

$$\Phi = (\mathbf{c}' - \mathbf{c}_0 - \mathbf{J}(\mathbf{b} - \mathbf{b}_0))^{\mathrm{t}}\mathbf{Q}(\mathbf{c}' - \mathbf{c}_0 - \mathbf{J}(\mathbf{b} - \mathbf{b}_0)). \tag{8}$$

The above equations pertinent to linear parameter estimation can then be used to calculate the parameter upgrade vector $(\mathbf{b} - \mathbf{b}_0)$ on the basis of the vector $(\mathbf{c}' - \mathbf{c}_0)$ that defines the discrepancy between the model-calculated observations $\mathbf{c}_0$ and their field-measured counterparts observations, $\mathbf{c}'$.

Based on Eq. (5), the parameter upgrade vector, $\mathbf{u}$, is the difference between the initial, or current, parameter vector and the updated parameter vector and is calculated as:

$$\mathbf{u} = (\mathbf{J}^{\mathrm{t}}\mathbf{Q}\mathbf{J})^{-1}\mathbf{J}^{\mathrm{t}}\mathbf{Q}(\mathbf{c}' - \mathbf{c}_0). \tag{9}$$

For a nonlinear model, $\mathbf{b}$, is defined by adding the parameter upgrade vector $\mathbf{u}$ of Eq. (9) to the current parameter values $\mathbf{b}_0$, and is not guaranteed to produce the minimum value of the objective function. Hence the new set of parameters contained in $\mathbf{b}$ must be used as a starting point, $\mathbf{b}_0$, in determining a further parameter upgrade vector. This updating continues until either the global Φ minimum is reached or a maximum number of updates is exceeded.

The updating of **u** is also controlled by the use of a Marquardt parameter, to render the optimization process more stable, especially in cases where high parameter correlation and low parameter sensitivity make the optimization process more difficult to accomplish (Doherty, 1998).

The iterative solution process described above requires that an initial set of parameters $\mathbf{b}_0$ be supplied to begin the optimization process. In this work, the initial estimates of Z at the pilot points are taken from the seed realizations created through geostatistical simulation. The range of possible Z values at each pilot point is constrained to be within the maximum and minimum Z values as measured across all observed Z values.

The objective function defining the mismatch between model outputs and observed data, defined in Eq. (4), is denoted as Φ_{obs}. If regularization constraints are to be imposed, then in addition to minimizing Φ_{obs}, it is also necessary to impose the requirement that a 'regularization objective function', Φ_r be minimized by the estimated parameter set. The regularization objective function is the SSE of the differences of the Z values between pilot points as weighted by the spatial covariance function (complement of the variogram). These weighted differences can be considered as regularization measures, for which a regularization objective function is defined as

$$\Phi_r = (\mathbf{d}' - \mathbf{N}\mathbf{b})^t \mathbf{Q}_r (\mathbf{d}' - \mathbf{N}\mathbf{b}), \quad (10)$$

where $\mathbf{Q}_r$ is a diagonal matrix comprised of the squares of weights assigned to the various regularization measures that collectively comprise the vector $\mathbf{d}'$. The relationships by which the model-generated counterparts to these regularization measures are calculated from the parameter values (constituting the vector **b**) are encapsulated in the matrix **N**. In the present case, **N** is a simple linear matrix composed of respective parameter differences. The 'preferred' or 'observed' value of these differences, $\mathbf{d}'$, is set to zero for enforcement of the 'maximally homogenous' regularization constraints.

The goal of the regularization process is to minimize Φ_r while ensuring that Φ_{obs}, is below a threshold value. This threshold value is set slightly above the minimum value for Φ_{obs} that could have been achieved without regularization and is denoted as Φ^l_{obs} (i.e. the 'limiting measurement objective function'). This value of Φ^l_{obs} is chosen by the user and is generally 1.05–1.10 times the minimum achievable value of Φ_{obs} where this minimum achievable value can be estimated as the weighted SSE of the measurement errors. To decrease Φ_{obs} below the measurement errors would force the calibration to fit noise in the data. The ability to define Φ^l_{obs} in advance of the parameter estimation makes the regularization methodology an extremely powerful tool. The regularization process must minimize Φ_r while enforcing the condition that $\Phi_{obs} \leq \Phi^l_{obs}$ (or, in practice, that $\Phi_{obs} = \Phi^l_{obs}$ because a decrease in Φ_r will nearly always require an increase in Φ_{obs} when parameter values are close to optimum). Thus the enforcement of regularization constraints is not allowed to take place at the expense of obtaining a good fit between model outputs and field observations.

The constrained minimization problem that implements the regularization can be formulated as an unconstrained minimization problem through the use of a Lagrange multiplier, λ. With regularization constraints imposed, the parameterization problem consists of determining the parameter vector that minimizes the total objective function, Φ_t, defined by the equation

$$\Phi_t = \Phi_r + \lambda \Phi_{obs}, \quad (11)$$

while simultaneously finding an appropriate value for λ that allows Φ_{obs} to equal Φ^l_{obs}. The Lagrange multiplier can be thought of as a factor by which all of the observation weights are multiplied to ensure that $\Phi_{obs} = \Phi^l_{obs}$ during the minimization of Φ_t.

2.3. Predictive estimation

Predictive estimation is based on determination of the true nonlinear range of predictive uncertainty by calibrating a model so that a key model parameter (prediction) is either maximized or minimized while maintaining calibration constraints. The theory is based on the work of Vecchia and Cooley (1987).

The traditional parameter estimation problem consists of minimizing the objective function defined by Eq. (4). For predictive estimation, let **K** represent the same linear model as **X** in Eq. (4) when run under predictive conditions (in the present case the key

model prediction is the advective travel time to a regulatory boundary). The action of the model when used in predictive mode can be represented by the equation

$$d = \mathbf{Kb}, \tag{12}$$

where d is a scalar representing a single model outcome (i.e. prediction) and $\mathbf{K}$ is a $1 \times n$ matrix where, once again, n is the number of parameters requiring estimation. This is the same parameter vector as that requiring estimation through imposition of calibration constraints in the traditional parameter estimation problem. The aim of predictive estimation is to maximize, or minimize, d while maintaining a calibration. Maintaining calibration is defined as maintaining the value of the objective function, Φ of Eq. (4) no higher than a value of Φ_0^l (similar to the upper measurement objective function constraint discussed in relation to the regularization problem). As shown in Fig. 1, d will be maximized (minimized) when $\mathbf{b}$ lies on the Φ_0^l contour and the equality $\Phi = \Phi_0^l$ holds.

For a linear model, the predictive estimation problem can be formulated as follows:

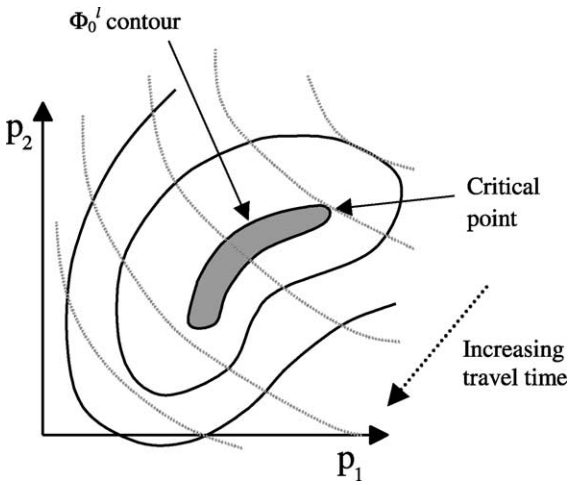

Fig. 1. Schematic diagram of minimizing a prediction (e.g. travel time) while maintaining a calibration where the calibration is defined as being within δ of Φ_{min} (shaded region). The black lines indicate contours of Φ values (SSE) and the contour surrounding the shaded region indicates where $\Phi = \Phi_0^l$. The gray dotted lines are contours of the prediction value (e.g. advective travel time). This example diagram is shown for the simple case of only two parameters, p_1 and p_2 (after Doherty, 2000).

Find $\mathbf{b}$ such as to maximize (minimize) $\mathbf{Kb}$ subject to

$$(\mathbf{c}' - \mathbf{Xb})^t \mathbf{Q} (\mathbf{c}' - \mathbf{Xb}) = \Phi_0^l,$$

where Φ_0^l is slightly above the minimum objective function. The value of Φ_0^l is set by the user with typical values being 1.05–1.10 times the value of the minimum value of Φ (after Vecchia and Cooley, 1987). The minimum value of Φ can be determined as the SSE of the weighted observation error as was discussed for the regularization approach.

The solution to this problem (Vecchia and Cooley, 1987) is

$$\mathbf{b} = (\mathbf{X}^t\mathbf{QX})^{-1}\left\{\mathbf{X}^t\mathbf{Qc} - \frac{\mathbf{K}}{2\lambda}\right\}, \tag{13}$$

where the Lagrange parameter, λ, is defined by the equation:

$$\left(\frac{1}{2\lambda}\right)^2 = \frac{\Phi_0 - \mathbf{c}^t\mathbf{Qc} + \mathbf{c}^t\mathbf{QX}(\mathbf{X}^t\mathbf{QX})^{-1}\mathbf{X}^t\mathbf{Qc}}{\mathbf{K}^t(\mathbf{X}^t\mathbf{QX})^{-1}\mathbf{K}}. \tag{14}$$

This problem is similar to the regularization problem discussed above, in that one function is minimized while the objective function is held at some upper limit. However, in the predictive analysis case, an analytical solution to the Lagrange multiplier is available through Eq. (14) (see details in Doherty (2000)).

3. Example problem

A groundwater flow and transport problem on a hypothetical aquifer demonstrates the changes in the T fields and advective travel times that occur using the three approaches to modeling described in the previous section. An exhaustively known T field is sampled at 89 locations (Fig. 2) to provide the T and head measurements. In order to compare results across the three different modeling approaches, the same set of 100 seed realizations is used as the basis for each approach. In the first approach, the seed realizations are used as is, with no calibration to the observed head values. In the second approach, the seed realizations are calibrated to the observed heads using 17 pilot points (Fig. 2). These pilot point locations were chosen with a focus on placing them in

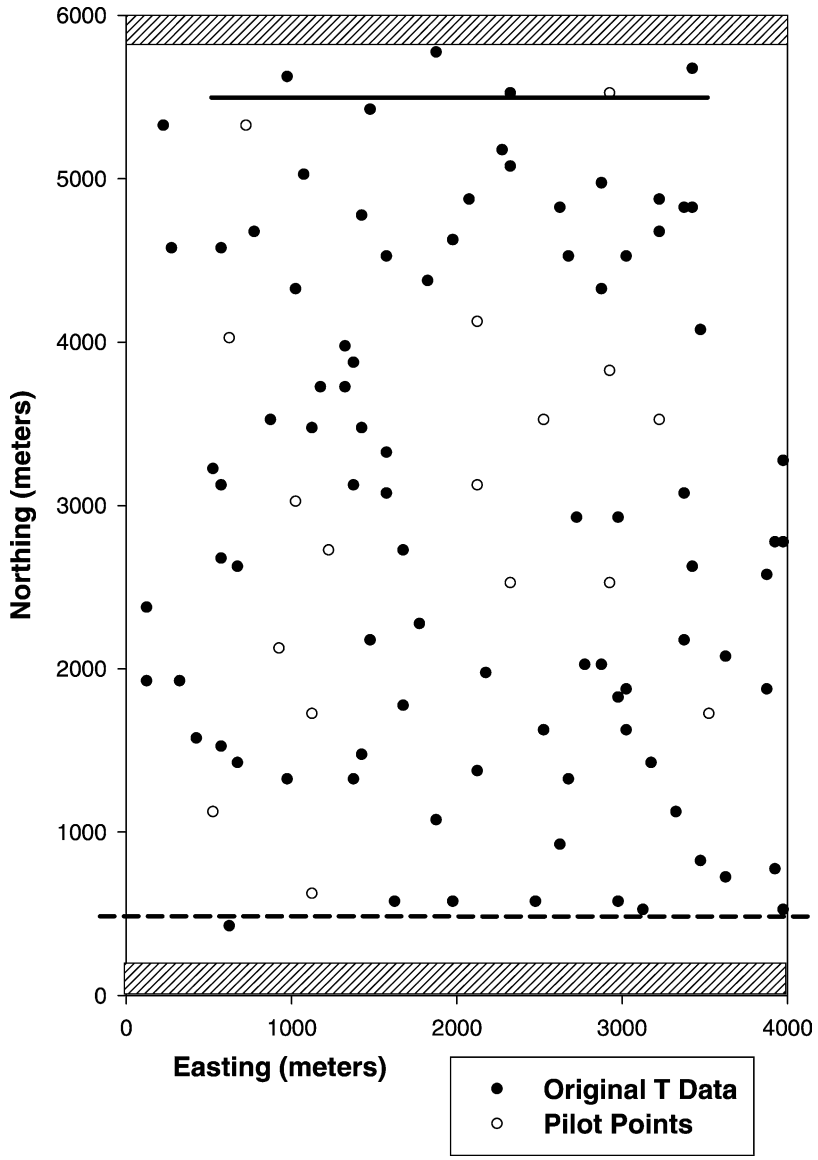

Fig. 2. Groundwater model domain, data locations and pilot points for the example problem. The high T buffer zones at the top and bottom of the model are cross-hatched. The source zone for the 400 particles is shown by the solid black line near the top of the model and the regulatory boundary is shown as a dashed black line near the bottom of the model.

areas distant from existing T measurements. For this example problem with 89 measurements and only 17 pilot points, it was not necessary to use the regularization constraints to achieve a stable solution. Therefore, the only constraint on the Z values at the pilot point locations is that they remain within the minimum and maximum of the 89 measured Z values.

In the third approach, predictive estimation is used to calibrate to the observed heads, as in the second approach, using the same 17 pilot points with the added constraint of minimizing the fifth percentile particle travel time. The predictive estimation is conducted such that for each realization the objective function is held to 1.10 times the minimum value of

the objective function determined in the second ('calibration') approach (Eq. (4) for this work). The PEST parameter estimation software (Doherty, 1998) was used to accomplish the calibration and predictive estimation modeling.

The mean and standard deviation of the 89 $Z = \log_{10} T$ data are -7.48 and 0.67, respectively. An anisotropic spherical variogram model with a nugget of 0.08 and a sill of 0.36 was used to model the spatial variation of Z. The range in the N–S direction is 5000 m and the range in the E–W direction is 2800 m. It is noted that the spatial structure of the exhaustive (true) T field is not obtained from a random field simulator, but is taken from a centimeter-scale digital X-ray transmission image of the porosity of a dolomite sample. Any uncertainty inherent in the estimation of the true variogram and its effect on the final predictive estimation is not considered in this example problem.

A total of 100 unique seed T fields created by a MG geostatistical simulation algorithm are used as input to a steady-state, 5-point, finite difference groundwater flow simulator. For this work, we use the MODFLOW (MacDonald and Harbaugh, 1988) groundwater flow model with the preconditioned conjugate gradient solver to solve the steady-state flow problem. The interblock transmissivities are calculated as harmonic averages of the T values in any two adjacent cells. The model has fixed heads on the upstream and downstream boundaries and zero-flux boundaries along the sides of the model. As the present study is focused specifically on T field generation, the boundary conditions are assumed to be known quantities.

For the geostatistical modeling, the grid blocks are 50×50 m and the domain is 4000 by 5600 m. The same gridblock discretization is used in the ground water flow model, but the size of the groundwater flow model domain is increased to 4000 by 6000 m with the addition of 4 rows of high T (1.0×10^{-5} m^2/s) buffer zones at both the upstream and downstream ends of the model (Fig. 2). These buffer zones allow the uniform fixed head applied across these boundaries to equilibrate with the heterogeneous T field.

A total of 400 particles were tracked from a 3000 m wide source region at a Northing coordinate of 5500 m to a regulatory boundary at a Northing coordinate of 500 m (Fig. 2). These particles are tracked along streamlines using the MODPATH software developed by the US Geological Survey (Pollock, 1989). For this hypothetical problem, the travel time of the fifth percentile arrival (the 20th fastest particle) is of particular regulatory interest.

To efficiently solve the groundwater flow and inverse problem for 100 realizations using each of three different modeling approaches, a parallel computing platform was employed. The heterogeneous platform, constructed for this project, comprises 17 PC's linked together and running the Linux operating system. Each PC contained at least one 200 MHz, or faster, Pentium Pro processor and 128Mb of RAM. The PEST software used to implement all three modeling approaches is capable of parallelizing the calculation for each T field by determining the Jacobian at every iteration in a parallel manner; however, we opted to parallelize the calibration process for all realizations by letting each processor solve all aspects of the estimation problem for a single realization. Custom shell scripts were written to maintain processor use across the PCs. The 100 runs took between 12 and 31 h depending on the modeling approach.

4. Results

The results across all three stages of modeling are compared to each other and to the true values obtained from the true T field from which the observed T and head values were obtained. The comparisons focus on the ability of each of the three ensembles of 100 T fields to match the observed heads while predicting advective travel times. Additionally, the reasons for the different calibration and travel time results are examined with respect to changes in the Z fields and changes in the arrival times and arrival locations of the particles.

4.1. Head calibration and travel times

The objective function defined in Eq. (4) is the SSE of the 89 head observations. The cumulative distribution, across 100 realizations, of the objective function is shown in Fig. 3A for all three modeling stages. Fig. 3A demonstrates the effect of conditioning to both T and heads relative to only conditioning to T values (seed T fields). The decrease in the median

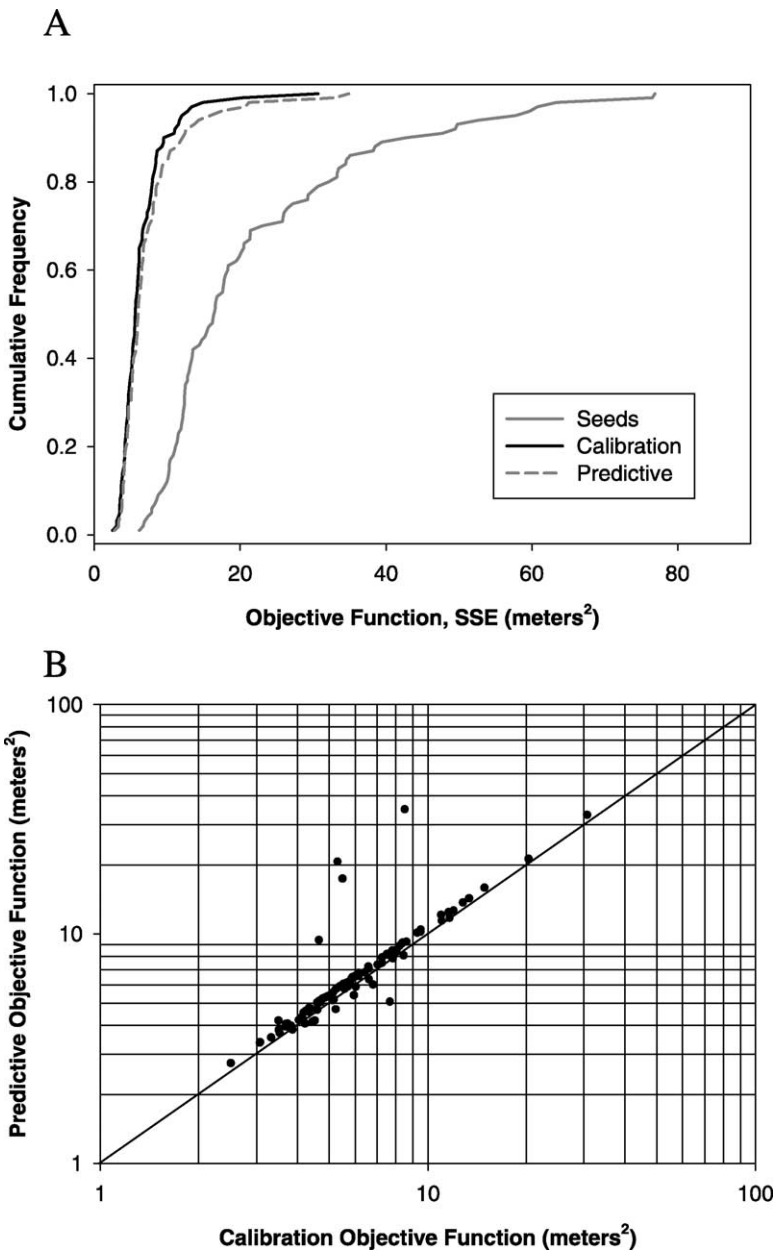

Fig. 3. Comparison of the cumulative distributions of objective function (SSE) values between the three different modeling stages (A) and a scatterplot comparison of the individual T field SSE values between the calibration and predictive estimation stages (B). The 1:1 line is also shown in the scatterplot comparison.

objective function value from the seed T fields to the calibrated fields is roughly 10 m^2. Fig. 3A also shows the similarity of the objective function distributions between the calibrated T fields and the T fields created with predictive estimation. The objective function distribution from the predictive estimation T fields is only slightly higher than the distribution from the calibrated T fields. A Kolmogorov–Smirnov (K–S) test shows that these two distributions are not different from one another at the 5% significance level.

A comparison of the change in the objective function, Eq. (4), from the calibration stage to the predictive estimation stage for each initial seed T field is shown in Fig. 3B. The comparison in Fig. 3B shows that for a given seed T field, the predictive estimation SSE is generally 10% higher than the SSE obtained under calibration mode. While most of the points in Fig. 3B lie above the 1:1 line, there are several points that fall below this line. For these T fields, it was possible to determine a better calibration to

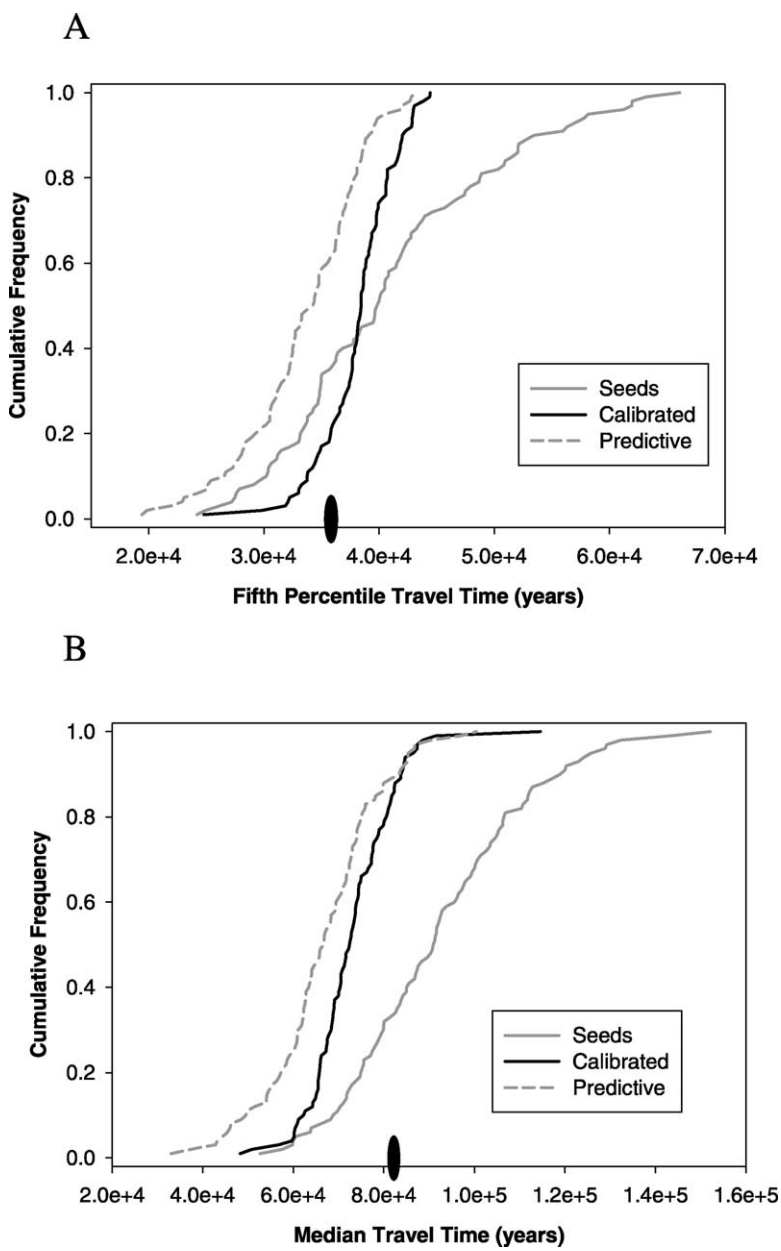

Fig. 4. Comparison of the fifth percentile arrival time, (A) and median arrival time (B) cumulative distributions for the three different modeling stages. The true values of the fifth percentile and median arrival times are shown by the black ellipses along the X-axis in each figure and are given in Table 1.

the observed heads under the minimum travel time constraint relative to the calibration mode. There are also four realizations that result in a very large value of the objective function under predictive estimation—those points lie well above the 1:1 line. For these predictive estimation runs, PEST was unable to find a solution with an SSE that was 10% above the minimum SSE value. This result may be due to these solutions getting trapped in local minima. Even though these four solutions were not better than the initial calibration, their results are included in the final travel time and exit location summaries.

Fig. 3A and the results of the K–S test show that across 100 T fields the match to the observed heads is not significantly different when the T fields are created under the calibration and the predictive estimation modes. The question remains as to whether or not it is possible to get significantly more conservative fifth percentile travel times out of the T fields created in the predictive estimation mode. Fig. 4A shows the distributions of fifth percentile travel times created under all three stages of the modeling. The parameters of these distributions and the value of the actual (true) fifth percentile travel time are shown in the top of Table 1.

Fig. 4A, shows that all three modeling stages produce an accurate distribution of fifth percentile travel times—all three distributions capture the true value. The distribution of fifth percentile travel times created with the seed T fields is the widest, least precise, distribution. Table 1 shows that the predictive estimation stage decreases the mean fifth percentile travel time by approximately 4700 years (12%) relative to the calibration stage. The minimum travel time determined by the predictive estimation is 19,420 years (22%) less than the minimum determined by the calibration stage of modeling. A K–S test on the distributions of fifth percentile travel times created by the calibration and predictive estimation modeling stages shows that the distributions are different at the 5% significance level.

These results show that using the predictive estimation model yields significantly shorter groundwater travel times without significantly changing the calibration as measured across an ensemble of 100 T fields. To check the effect of minimizing the fifth percentile travel time on the rest of the groundwater travel times, the results for the median travel time through each T field are also examined (Fig. 4B and the bottom of Table 1). Fig. 4B shows that all three distributions capture the true value of the median travel time and similar to the fifth percentile travel time results, the seed realizations produce the least precise estimate of the median travel time and the predictive estimation produces the fastest estimates of the median travel time. Also similar to the fifth percentile travel time results, the K–S test of the distributions of median travel time produced by the calibration and predictive estimation modes show that they are significantly different.

4.2. Changes in T fields, arrival locations and arrival times

The changes in the T fields and particle tracks that are necessary for the calibration and travel time results shown in the previous section are examined here. Fig. 5A shows the true Z field along with the expectation maps from the 100 realizations created at each stage of modeling: seed fields (B), calibration (C) and predictive estimation (D). The expectation

Table 1
Parameters of fifth percentile and median travel time distributions (all values given in years)

	Model stage	Mean	Median	Std. dev.	Min.	Max.
Fifth percentile travel time (years)	Seed	40,939	39,895	9401	24,200	66,100
	Calibrated	38,206	38,440	3274	24,790	44,430
	Predictive	33,527	34,315	5115	19,420	43,130
	True	35,740	N/A	N/A	N/A	N/A
Median travel time (years)	Seed	92,152	91,025	19,592	52,650	152,100
	Calibrated	72,944	72,460	9177	48,290	114,600
	Predictive	66,623	66,750	12,384	33,040	100,400
	True	81,950	N/A	N/A	N/A	N/A

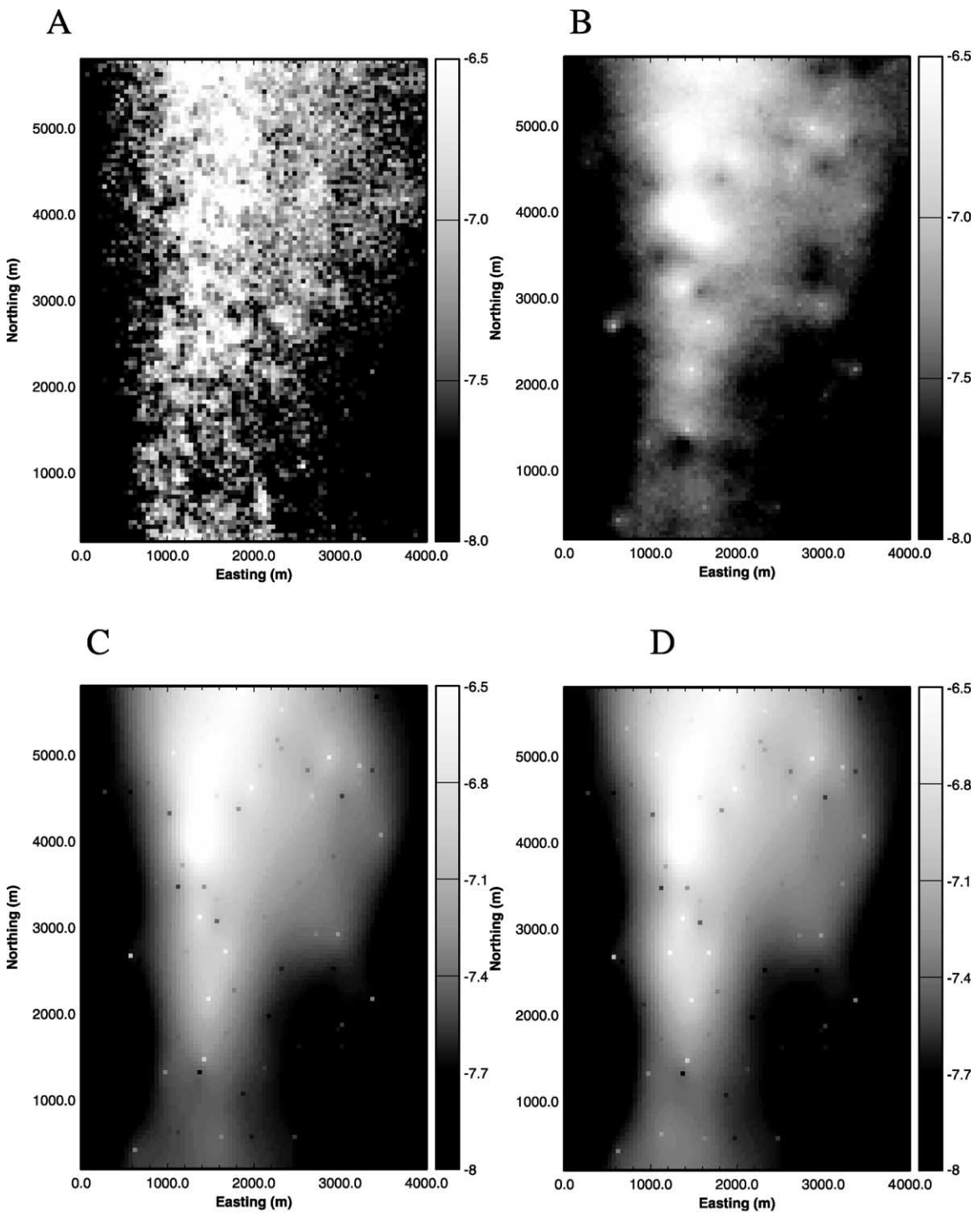

Fig. 5. Comparison of the true $\log_{10}$ transmissivity, Z, field (A) with the expectation maps created from the seed realizations (B), the calibrated realizations (C) and the realizations created through the predictive estimation approach (D). The color scale defines the Z value.

maps shown in Fig. 5B–D are all noticeably smoother than the true field, as would be expected from averaging Z over 100 realizations. However, the general shape of the high T zone down the left-center of the domain is reproduced by all three modeling stages including the higher T values at the top and the narrowing and decreasing values of the high T region towards the bottom of the domain.

There are several noticeable differences between the expectation maps in Fig. 5B–D. The expectation map from the seed realizations shows disconnected zones of high T, (white areas) from the top of the domain to near the bottom. The expectation maps from the calibration and predictive estimation stages (Fig. 5C and D) show a smoother and more continuous area of high T from the top to near the bottom of the domain relative to the expectation map from the seed realizations. This high T zone is also somewhat wider (east–west) in the expectation map created through predictive estimation relative to the calibration stage expectation map.

The differences between the T fields created in the calibration and the predictive estimation stages are further examined in Fig. 6. Fig. 6 shows the difference map created by subtracting Fig. 5C from Fig. 5D (predictive estimation minus calibration). The majority of the domain is unchanged between the two modeling approaches—difference values near zero. The major difference between the expectation maps is that the predictive estimation increases the transmissivity in the area surrounding the pilot points at (1225, 2725) and, to a lesser extent, at (1125, 625) relative to the calibration approach. These increases in T are in the path of the fastest particle tracks.

The effect of the different modeling stages on the location and travel time of individual particles (streamlines) is examined in Fig. 7. The particles are uniformly distributed along the source line and are numbered from 1 to 400 from west to east (left to right in Fig. 7). Fig. 7 shows the mean and two standard deviations of the arrival coordinate (Easting coordinate in meters where each particle crosses the regulatory boundary) compared to the true arrival coordinate (Fig. 7A, C and E). This figure also shows the mean and $+/-$ two standard deviations of the travel time compared to the true arrival time for every particle (Fig. 7B, D and F). Fig. 7 shows that the seed

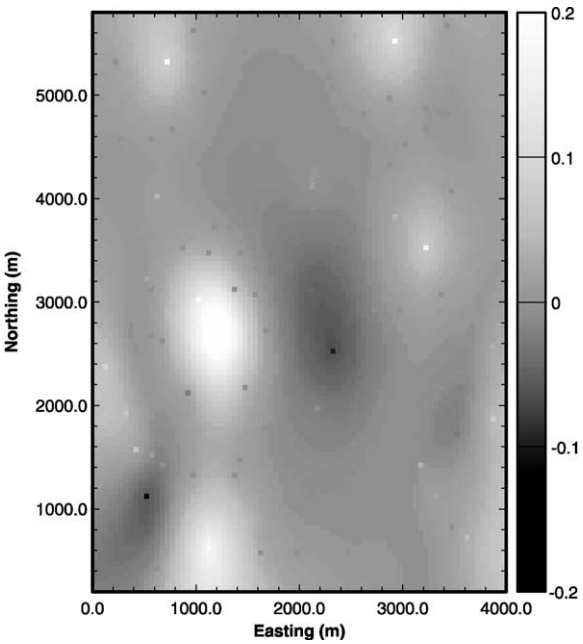

Fig. 6. The difference map showing the difference in $Z = \log_{10} T$ values as calculated by subtracting the calibration expectation map (Fig. 5C) from the predictive estimation expectation map (Fig. 5D). The color scale defines the difference in $Z = \log_{10} T$.

realizations are the least precise, largest standard deviations, for both the arrival coordinates and the travel times. The calibration stage of modeling provides the most precise estimates of both arrival location and travel times for the particles.

The graph of the true arrival coordinates shows very few particles passing the regulatory boundary between the easting coordinates of 1700 and 2050 m and a large number of particles exiting the boundary between 2050 and 2200 m. This result is not well captured by any of the modeling approaches and appears to be controlled by a local high T zone in the true field (Fig. 5A) between the Easting coordinates of 2050 and 2200 m at a Northing coordinate of 500 m. This local high T zone is not captured, on average, by the simulated T fields (Fig. 5B–D). Examination of Fig. 2 shows that there are no data or pilot points that sample this local high T zone.

The mean and standard deviation of the travel times change with both the particle number and the modeling approach (Fig. 7B, D and F). In general, the seed realizations overestimate the true travel time for all particles. This may be due to the less connected high

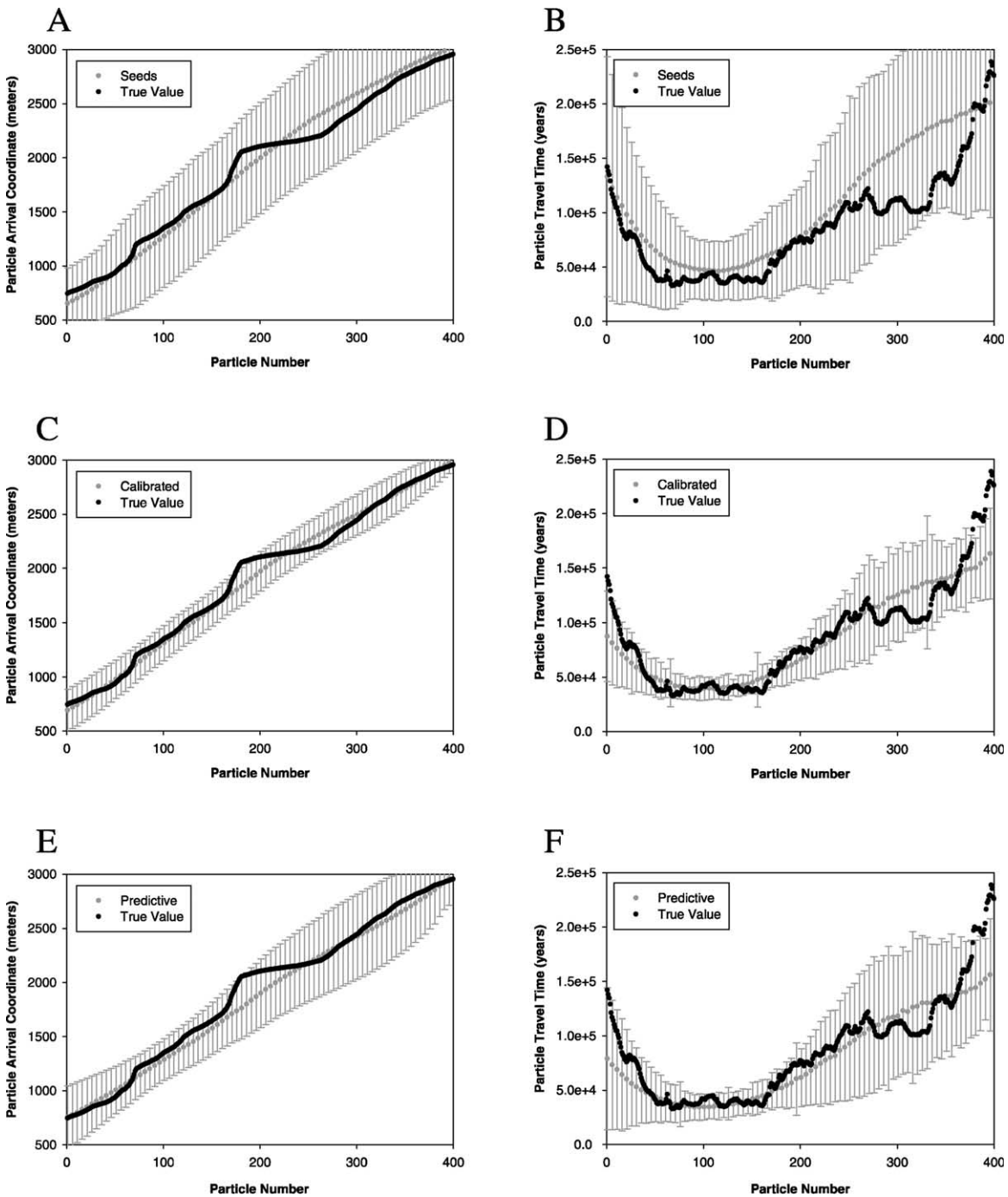

Fig. 7. Results of the particle arrival locations (images on left-hand side) and the particle travel times (images on right-hand side). The particles are numbered from 1 to 400 from left to right. The results of the seed realizations are shown in (A) and (B), the calibration results in (C) and (D) and the predictive estimation results in (E) and (F). The results are shown as the mean and $+/-$ two standard deviations as calculated across the 100 realizations for each particle (gray lines) and are compared to the true arrival location and travel time values (black circles). For readability, the mean and standard deviation of only every fifth particle are shown.

Table 2
Average absolute differences between the mean simulated and true arrival coordinates (upper) and the mean simulated and true travel times (lower)

		Seed realizations	Calibrated	Predictive estimation
Arrival coordinate (meters)	All particles	84	50	73
	Particles 50–160	56	32	81
Travel time (years)	All particles	22,713	12,540	13,206
	Particles 50–160	12,976	4897	3693

T zones seen in the seed realization expectation map (Fig. 5B) relative to the expectation maps produced by the calibration and predictive estimation approaches. The predictive estimation approach creates the fastest set of travel times across all particles.

Table 2 summarizes the ability of the different modeling approaches to accurately predict the true travel times. Table 2 shows the average absolute difference between the mean of the simulated arrival coordinates and the true arrival coordinate and the average absolute difference between the mean simulated travel time and the true travel time for each particle. The results in Table 2 show that across all particles the calibration stage of modeling produces the most accurate estimates (smallest absolute differences) of the arrival coordinate and the travel times. However, the focus of the predictive estimation is not the entire suite of travel times, but the fastest travel times. From the results on the true field, the fastest particles are number 50–160 (Fig. 7). Table 2 also shows the absolute average deviation between the mean simulated and the true travel time for particles 50–160. These results show that the predictive estimation modeling approach produces the most accurate travel times for the fastest particles. These results indicate that the predictive estimation approach not only maximizes conservatism while maintaining a calibration, but, for this example problem, it also produces the most accurate results in the area where the optimization was focused (fifth percentile arrival times).

5. Conclusions

Many authors have recognized the non-uniqueness of parameter estimates in stochastic inverse modeling of T fields. However, the impact of this non-uniqueness on the prediction of groundwater transport has not previously been addressed in a quantitative manner. This paper presents the predictive estimation technique to address non-uniqueness in T field generation by adding a constraint of maximizing conservatism in the predictive model while maintaining a near optimal calibration to the observed data.

In the advective transport example problem shown here, the results demonstrate that using the predictive estimation approach can produce significantly shorter groundwater travel times without significantly changing the calibration as measured across an ensemble of 100 T fields. Not only were the predictions of advective travel time more conservative, but they were also more accurate in the area where the optimization was focused (fifth percentile arrival times) relative to a more traditional T field generation technique. It is noted that results of any T field calibration, including the predictive estimation presented herein, are dependent on the number and distribution of pilot points. Future work will investigate this dependency and examine the robustness of the predictive estimation approach on additional test cases.

Predictive estimation is a powerful technique for addressing non-uniqueness in parameter estimates by focusing results on maximizing the conservatism of model predictions. This approach is not limited to ground water travel time problems and may provide solutions to problems where other risk or even financial-based prediction must be viewed from the perspective of conservatism.

Acknowledgements

The authors wish to acknowledge critical reviews of this article by Randy Roberts, Scott James, Erik Webb, Daryl Mercer, Manuel Maria Pacheco Figueiredo and an anonymous reviewer. We also acknowledge the help of Lane Yarrington in extending the parallel PC cluster for this work. The third author participated in this study while in the Sandia National

Laboratories Student Intern Program. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL-85000.

References

Bard, J., 1974. Nonlinear Parameter Estimation, Academic Press, New York, pp. 341.

Capilla, J.E., Gómez-Hernández, J.J., Sahuquillo, A., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 2. Demonstration on a synthetic aquifer. J. Hydrol. 203 (1–4), 175–188.

Capilla, J.E., Gómez-Hernández, J.J., Sahuquillo, A., 1998. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 3. Application to the Culebra Formation at the waste isolation pilot plant (WIPP). J. Hydrol. 207 (3/4), 254–269.

Certes, C., De Marsily, G., 1991. Application of the pilot point method to the identification of aquifer transmissivities. Adv. Water Resour. 14 (5), 284–300.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide, second ed, Oxford University Press, New York, pp. 369.

Doherty, J., 1998. PEST-Model Independent Parameter Estimation, second ed, Watermark Numerical Computing, Brisbane, Australia, pp. 191, http://members.ozemail.com.au/~wnc/wnc.htm.

Doherty, J., 2000. PEST2000 Upgrade Notes, Watermark Numerical Computing, Brisbane, Australia, pp. 34, http://members.ozemail.com.au/~wnc/wnc.htm.

Gómez-Hernández, J.J., Wen, X.-H., 1998. To be or not to be multiGaussian? A reflection on stochastic hydrology. Adv. Water Resour. 21 (1), 47–61.

Gómez-Hernández, J.J., Sahuquillo, A., Capilla, J.E., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. I. Theory. J. Hydrol. 203 (1–4), 162–174.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation, Oxford University Press, New York, pp. 483.

Hendricks Franssen, H.-J.W.M., Cassiraga, E.F., Gómez-Hernández, J.J., Sahuquillo, A., Capilla, J.E., 1999. Inverse modeling of groundwater flow in a 3D fractured media. In: J., Gómez-Hernández, A., Soares, R., Froidevaux (Eds.), GeoENV II: Geostatistics for Environmental Applications, Kluwer Academic Publishers, Dordrecht, pp. 283–294.

Lavenue, M., de Marsily, G., 2001. Three-dimensional interference test interpretation in a fractured aquifer using the pilot-point inverse method. Water Resour. Res. 37 (11), 2659–2675.

Lavenue, A.M., Pickens, J.F., 1992. Application of a coupled adjoint-sensitivity and kriging approach to calibrate a groundwater flow model. Water Resour. Res. 28 (6), 1543–1569.

MacDonald, M.G., Harbaugh, A.L., 1988. A modular three-dimensional finite-difference ground-water flow model, US Geological Survey Techniques of Water Resources Investigations, Book 6, US Geological Survey, Denver, Colorado, pp. 1-1–E-8 (Chapter A1).

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

Pollock, D.W., 1989. Documentation of computer programs to compute and display pathlines using results from the US geological survey modular three-dimensional finite-difference ground-water flow model, Open File Report 89-381, US Geological Survey, Reston, VA, pp. 188.

RamaRao, B.S., Lavenue, A.M., de Marsily, G., Marietta, M.G., 1995. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1. Theory and computational experiments. Water Resour. Res. 31 (3), 475–493.

Vecchia, A.V., Cooley, R.L., 1987. Simultaneous confidence and prediction intervals for nonlinear regression models with application to a groundwater flow model. Water Resour. Res. 23 (7), 1237–1250.

Zimmerman, D.A., de Marsily, G., Gotway, C.A., Marietta, M.G., Axness, C.L., Beauheim, R.L., Bras, R.L., Carrera, J., Dagan, G., Davies, P.B., Gallegos, D.P., Galli, A., Gomez-Hernandez, J., Grindrod, P., Gutjahr, A.L., Kitanidis, P.K., Lavenue, A.M., McLaughlin, D., Neuman, S.P., RamaRao, B.S., Ravenne, C., Rubin, Y., 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. Water Resour. Res. 34 (6), 1373–1413.