Title
Stochastic Modeling of Spatial Heterogeneities Conditioned to Hydraulic Tracer Tests



Authors
Gupta, A. Datta
Vasco, D.W.
Long, J.C.S.
et al.

Publication Date
1994-07-01

LBL-35909
UC-400

# Lawrence Berkeley Laboratory

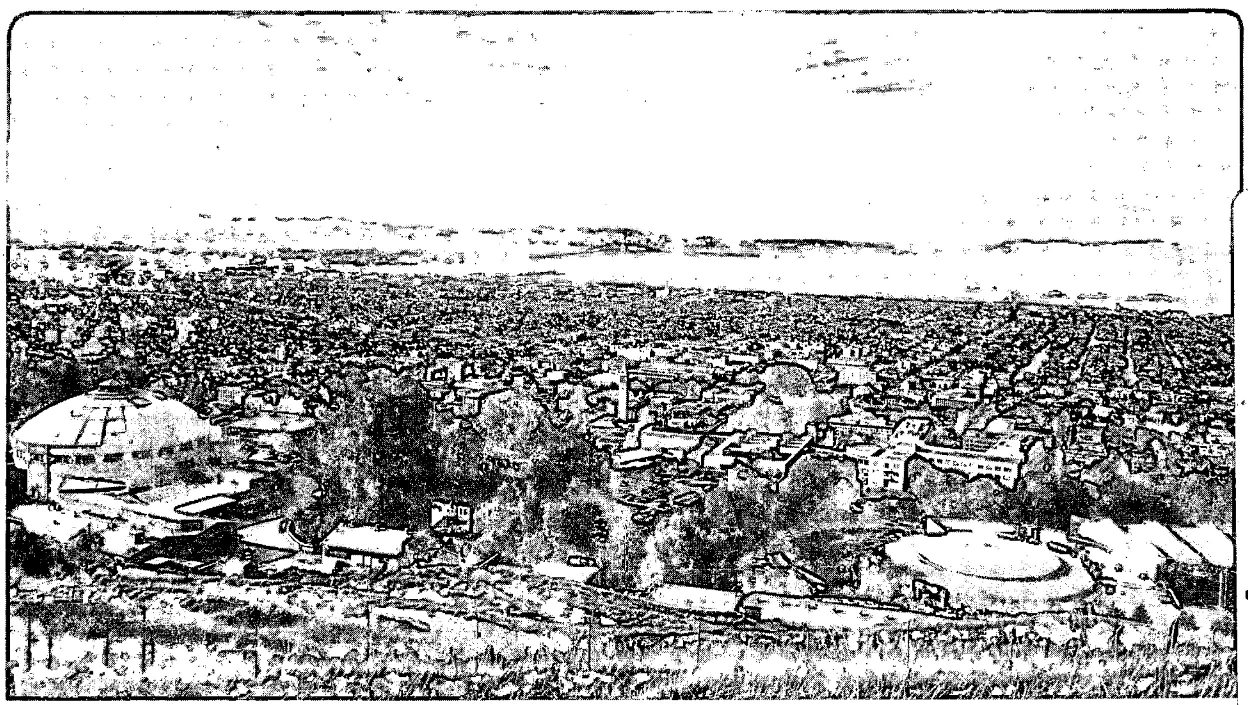
UNIVERSITY OF CALIFORNIA

## EARTH SCIENCES DIVISION

## Stochastic Modeling of Spatial Heterogeneities Conditioned to Hydraulic and Tracer Tests

A. Datta Gupta, D.W. Vasco, J.C.S. Long, and S. Vomvoris

July 1994





Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

LBL-35909
UC-400

# STOCHASTIC MODELING OF SPATIAL HETEROGENEITIES CONDITIONED TO HYDRAULIC AND TRACER TESTS

AKHIL DATTA GUPTA, D. W. VASCO, J. C. S. LONG

EARTH SCIENCES DIVISION
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

and

STRATIS VOMVORIS

NAGRA
Swiss National Cooperative for the Storage of Nuclear Waste
Hardstrasse 73, CH-543o
Wettingen, Switzerland

JULY 1994

This work was funded by the Director, Office of Civilian Radioactive Waste Management, Office of External Relations, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

This paper appeared in Proceedings of the Fifth Annual High Level Radioactive Waste Management Conference, Las Vegas, NV, May 22-26, 1994.

# STOCHASTIC MODELING OF SPATIAL HETEROGENEITIES CONDITIONED TO HYDRAULIC AND TRACER TESTS

Akhil Datta Gupta, D. W. Vasco and J. C. S. Long
Earth Sciences Division
Lawrence Berkeley Laboratory
Berkeley, CA 94720
(510) 486-6115

Stratis Vomvoris
Swiss National Coorporative
for the Storage of Nuclear Waste
Baden, Switzerland
41-56-371-324

## ABSTRACT

This paper discusses an approach to generation of stochastic permeability fields through simultaneous inversion of flow and transport data. For tracer transport calculations, we have used a semianalytic transit time algorithm which is fast, accurate and free from numerical dispersion. The inversion of data has been accomplished through the use of simulated annealing. We have addressed the non-uniqueness associated with our results by shifting the focus from the search for a single model that fits the data best to inferences about the properties that are shared by an ensemble of acceptable models. We then determine a most likely model for heterogeneity. The approach has been illustrated through application to tracer migration in a synthetic fracture plane.

## I. INTRODUCTION

Uncertainty concerning the physical and chemical nature of subsurface heterogeneities constitutes a severe technical barrier to assessing long term performance of nuclear waste repositories. Traditional stochastic imaging techniques to describe these heterogeneities (e. g. spectral methods, turning bands method etc.) are often poorly suited to reproduce the complex geological/morphological patterns and do not have the ability to incorporate fluid flow and transport data directly. Characterizing heterogeneous permeable media using flow and transport data typically requires solution of an inverse problem. Such inverse problems are computationally intensive and often involve iterative procedures requiring many forward simulations of the flow and transport problem. Previous attempts have been mostly limited to flow data such as pressure transient (interference) tests using multiple observation wells.[1,2] In this paper we examine the possibility of obtaining stochastic permeability fields conditioned to flow as well as transport data such as cross-hole tracer tests.

We assume permeability to be a spatially correlated random variable with a known distribution specified by a mean and a standard deviation. The flow and transport data are used to infer the underlying correlation structure. A critical aspect of our approach is an accurate and efficient solution of the transport problem. Modeling of tracer motion is complicated by the fact that any grid-based numerical method would be limited in resolution by the size of the grid block. If the characteristic length scale of heterogeneity is much smaller than the feasible grid block size, then dispersion would be impossible to model accurately, irrespective of the sophistication of the discretization scheme (higher order finite difference, curvilinear finite element, particle-in-cell, etc.). We circumvent the problem through the use of a novel semianalytic transit time algorithm which is fast, accurate and free of numerical dispersion.[3]

Another important issue that is common to all inverse problems is the uniqueness of the solution. Although incorporation of transport data in addition to flow data helps better constrain the problem, the permeability fields obtained by inversion are still non-unique. We address such non-uniqueness through the use of an ensemble approach, which is an attempt to go beyond previous work by developing new methods for understanding characteristics shared by all solutions of an inverse problem. The general approach is to generate and describe a collection of models that fit the data in some sense. A basic feature of this approach is that the focus of the problem shifts from the search for a single model that fits the data best to inferences about the properties that are shared by the ensemble of acceptable models.

We can then determine a most likely model and quantify the associated uncertainties in estimations.[4]

## II. THE APPROACH

Our approach involves a two-step procedure. First the problem is posed as an optimization question and then, the optimization problem is solved using simulated annealing which is a flexible and heuristic optimization technique with its foundation in statistical mechanics. An objective function is defined as the weighted sum of difference between the properties of any simulated image and reference values. The optimization problem consists of lowering the objective function enough so that the image has desired properties. Such an approach allows us to impose constraints that cannot be easily incorporated within the traditional framework of geostatistics.[5]

### A. Construction of an Objective Function

The objective function during inverse modeling quantifies the mismatch between the observed and computed data. It also provides a mechanism to incorporate data from various sources by imposing constraints during the minimization. We have used a two part objective function defined as follows,

$$E(\mathbf{m})= \zeta_1 \sum_{j=1}^{nw} \sum_{i=1}^{np_j} \omega_{pi}[P_{j,obs}(t^i) - P_{j,cal}(t^i)]^2 + \zeta_2 \sum_{j=1}^{nw} \sum_{i=1}^{nt_j} \omega_{ci}[C_{j,obs}(t^i) - C_{j,cal}(t^i)]^2 \ . \quad (1)$$

In the above expression, $\mathbf{m}$ is the parameter vector comprised of permeability values at grid blocks, nw is the total number of wells, $np_j$ and $nt_j$ are the number of pressure and tracer data recorded at the j-th well. The pressure or the tracer data can be weighted as desired, the weights being $\zeta_1$ and $\zeta_2$. In our application, we use $\zeta_1 = \zeta_2$.

The pressure transients are computed through a finite difference solution of the flow field. For tracer transport calculations, the flow field is assumed to have reached steady state conditions. Then the tracer history is obtained using the semianalytic transit time algorithm described in the next section. In Eq.(1), the weights $\omega_{pi}$ and $\omega_{ci}$ for the pressure and the concentration histories are chosen to emphasize the early times in pressure transients and tracer response.

### B. Modeling of Tracer Motion

For tracer transport calculations, we have used a semianalytic transit time approach which is computationally efficient and free from numerical dispersion. The method is *semianalytic* because we obtain the velocity field numerically. This generalizes the approach to any arbitrary configuration of wells and also to areally heterogeneous permeability fields. The semianalytic approach is based on the observation that in a velocity field derived by finite difference, streamlines can be approximated as piecewise hyperbolic intervals. Along each interval, we can solve the evolution equation exactly. Once the transit times to a producing well are determined, the tracer response can be obtained from simple integral expressions.

For steady flow in a non-deformable permeable medium, the stream function, $\psi$, must satisfy the differential equation

$$\nabla . (\lambda^{-1} \nabla \psi) = 0 \quad (2)$$

where, $\lambda = k/\mu$, is the fluid mobility. Equation (2) follows from the requirement that the curl of head gradient must vanish.[6]

The simplest representation of the stream function $\psi$ on a rectangular lattice grid block is a bilinear function,

$$\psi(x,y) = \psi_0 + ax + by + cxy \ , \quad (3)$$

which satisfies Eq (2) since $\lambda$ is constant on a grid block. The four constants are fixed by $\psi$ on the four corners of the grid block.

From Eq (3) one obtains,

$$v_x = \frac{\partial \psi}{\partial y} = b + cx \quad (4a)$$

$$v_y = - \frac{\partial \psi}{\partial x} = - (a + cy). \quad (4b)$$

Thus, the component velocities vary linearly in the respective directions only.

Consider a tracer particle initially ($t = 0$) at ($x_o,y_o$), which convects according to Eq (4). The particle trajectory can be calculated by direct integration, for example in the x direction,

$$\int \frac{dx}{b + cx} = \Delta t_x$$

$$\ln \frac{b + cx}{b + cx_0} = c\, \Delta t_x \qquad (5a)$$

$$x = x_0 + v_x(x_0) \,.[(e^{c\,\Delta t_x} - 1)/c]. \qquad (5b)$$

Analogously, in the orthogonal y direction,

$$\ln \frac{a + cy}{a + cy_0} = - c\, \Delta t_y \qquad (6a)$$

$$y = y_0 + v_y(y_0) \,.[(1 - e^{-c\,\Delta t_y})/c]. \qquad (6b)$$

Since the streamlines must enter and exit through grid block faces, the actual transit time across a grid block will be given by the minimum over the allowed edges.[7] Thus,

$$\Delta t = \text{Minimum}\,(\Delta t_x\,,\, \Delta t_y\,) \qquad (7)$$

where the minimum is examined only for positive values. One may verify that along these trajectories, $\psi$ is constant. From Eqs (5a, 6a)

$$(b + cx)(a + cy) = (b + cx_0)e^{c\,\Delta t}\,(a + cy_0)e^{-c\,\Delta t} = \text{constant} \qquad (8)$$

while $(b + cx)\,(a + cy) = c.\,\psi(x,y) + (ab - c\,\psi_0)$ .

Hence the streamlines are hyperbola with asymptotes at $x = -b/c$, $y = -a/c$.

In a block centered finite difference model, linear interpolation of the velocity components is consistent with the assumptions of the flow model.[8] Hence, in a finite difference scheme, we may approximate $\psi$ by piecewise hyperbolic intervals. On each of these streamlines, the tracer particle convects according to Eqs. (5b), 6b).

The transit time to any location in the domain is obtained by following the streamline backwards in time to the injector and summing up the travel times through successive grid blocks in the finite difference model. Given a configuration of permeabilities, first a transit time function, $\tau(\psi)$ is calculated at each producing well. For calculating the transit time function $\tau(\psi)$, we follow the procedure outlined above; however, we originate the trajectory at the producing well and follow it backwards in time until an injector is reached. Finally, we label the transit times as a function of the streamline. Once $\tau(\psi)$ for a producing well is obtained, the tracer history can be derived simply by integrating the response of individual streamlines, allowing for appropriate delay time. Thus,

$$C(t) = \int_{\text{all}\ \psi} C_i(t - \tau(\psi))\, d\psi \qquad (9)$$

where $C_i$ (t) is the injection history of the tracer slug. In Eq.(9) we have neglected physical dispersion and assumed that dispersion of the tracer occurs because of permeability heterogeneities only. However, this approach can be easily extended to incorporate longitudinal diffusion along streamlines.

### C. Ensemble Analysis

The inverse approaches to hydrologic characterization have the advantage that they can incorporate flow as well as transport data directly in deriving the spatial heterogeneity patterns. Thus the approach naturally emphasizes the underlying features that impact the fluid flow and transport. Most often one seeks the model which maximizes the likelihood of the data, i. e. minimizes the misfit to the data.[1] However, given the uncertainties associated with a set of measurements, it is probable that many sets of models may fit the data within some specified tolerance. In particular, a network of flow channels, which may be rather elaborate, is likely to be poorly constrained by limited data. Thus, hydrologic models derived by inversion are non-unique in general and it may not be informative to seek a single model. Instead, it may be best to generate a large collection or ensemble of models approximately satisfying the data. That is, models for which some measure of misfit, say $\chi^p$ misfit, is small. When a sufficient number of models have been accumulated various statistical quantities may be extracted from the ensemble.

## III. APPLICATION

We now illustrate the approach discussed above by application to tracer migration in a single fracture plane. The flow geometry consists of a single injection well located in the middle of the fracture plane and a single producing well located in the upper right hand

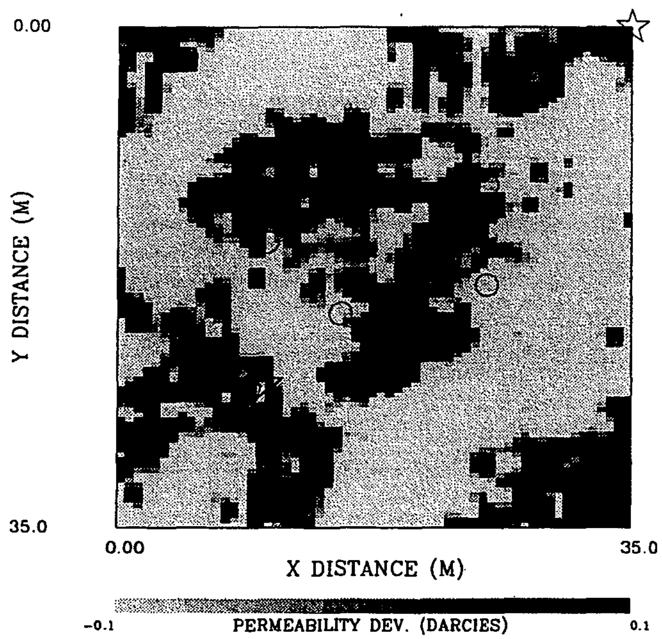



Fig. 1. Synthetic Fracture Plane

corner as shown in Fig. 1 using star symbols. The model can be thought of as an approximation to a tracer test in a fracture plane where the tracer is collected in a drift represented by the producing well.[9]

For generation of the synthetic fracture plane, we have assumed permeability to be a spatially correlated random variable specified by a probability distribution and a correlation length. A stochastic moving average method was then used in order to generate spatial patterns of permeability within the fracture plane as shown in Fig. 1. In Fig. 1, we have removed the mean and compressed the scale in order to highlight the extreme value features of the permeability field. Clearly, there is a central high permeability region surrounded by two low permeability regions. For this example, we have used a log-normal distribution of permeabilities with a mean of 0.5 Darcy, log standard deviation of 0.6 and an isotropic correlation length of 40% of the fracture plane dimensions.

A. The Forward Problem

A block centered finite-difference model was used for solving the forward problem. We have used a grid size of 35x35 with $\Delta x = \Delta y = 5$ ft. Pressures at the nodes were obtained using a five-point central difference scheme. The velocities at the block faces were then computed using Darcy's law. The tracer motion was calculated using the semianalytic transit time algorithm discussed before. First, the transit time function was computed at the producing well (Fig. 2a) and then, the tracer response was computed using Eq. (9) (Fig. 2b).

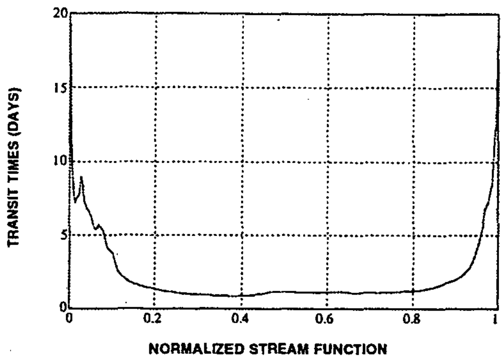



Fig. 2a. Transit Time Function for the Synthetic Fracture Plane.

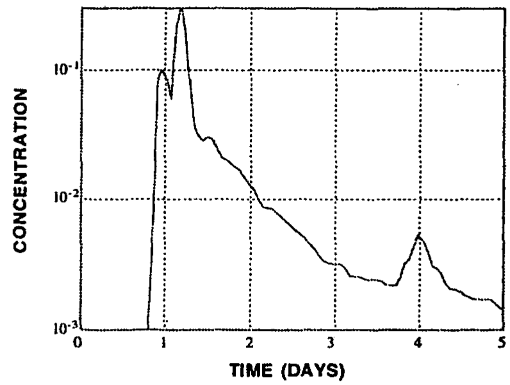



Fig. 2b. Tracer Response From the Synthetic Fracture Plane.

Figure 3a shows the isochrones for the synthetic fracture plane. These isochrones are contours of equal arrival times and thus, represent the tracer front at different time intervals. As expected, most of the tracer flows through the central high permeability region. This is evidenced by the early breakthrough of the tracer due to flow channelization. The same behavior is also reflected by the streamlines (Fig. 3b) which are grouped in the central region, indicating preferential fluid movement through these areas. While evaluating the conductivity patterns obtained by inversion of transport data, we will pay particular attention to this central high permeability region.

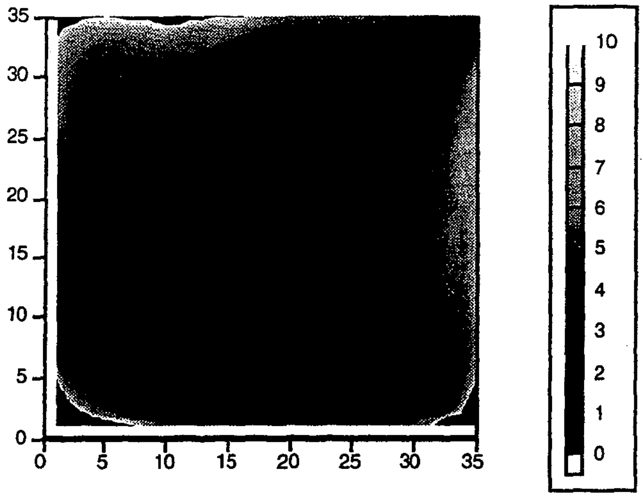


Fig. 3a. Isochrones (in Days) for the Synthetic Fracture Plane.

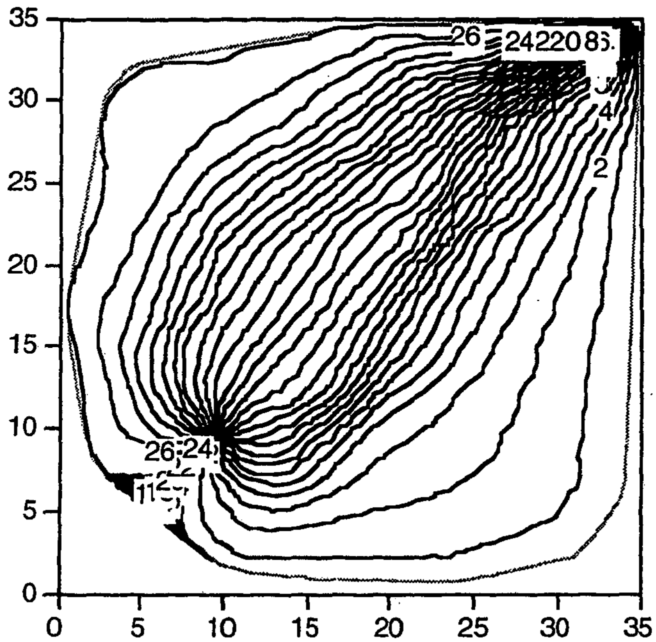


Fig. 3b. Streamlines for the Synthetic Fracture Plane.

B. The Inverse Problem

For inversion of flow and transport data, we chose steady head at the six locations indicated by circles in Fig. 1 along with the tracer response at the producing well as shown in Fig. 2b. As mentioned before, simulated annealing was used for inversion. We have attempted alternative approaches such as conjugate gradient for inversion of flow and transport data. However, for large number of parameters, simulated annealing was found to be more robust and computationally efficient for the example problem studied here. Starting with an initial distribution of permeabilities, we choose a location at random and replace the permeability of that location with a value sampled from a specified log-normal distribution. A change in misfit, also known as energy, is computed due to the perturbation. If the energy decreases, then the perturbation is accepted; otherwise the perturbation is accepted with a probability $P(\Delta E) = \exp(-\Delta E/T)$ where T is analogous to temperature in Gibbs distribution. By allowing to accept changes which result in an increase in energy, the simulated annealing approach to optimization provides a mechanism of *probabilistic hill climbing* which allows the method to escape from local minima.[10,11]

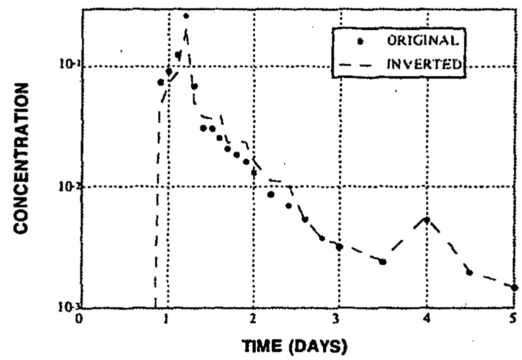



Fig. 4a. Inversion of Tracer Data.

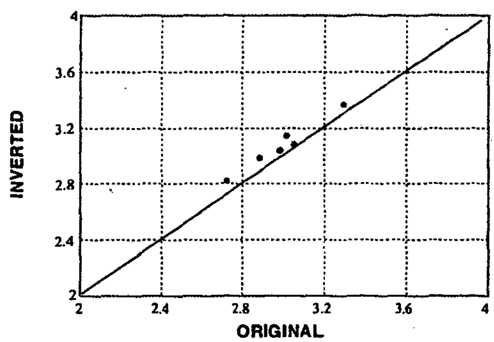



Fig. 4b. Inversion of Steady Head Data (Meters)

Figs. 4a and 4b compare the tracer history and the steady heads obtained from one such inversion with the synthetic data. Overall, the agreement with the data is reasonably good. However, since the tracer history is an integrated response, a more rigorous test of the inversion will be to compare the flow fields. Figs. 5a and 5b

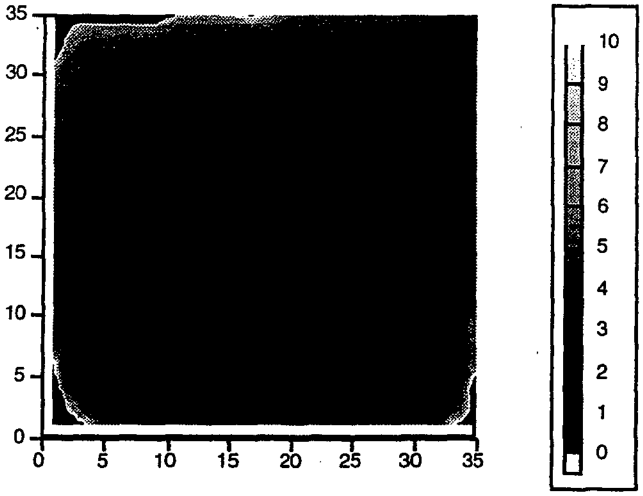



Fig. 5a. Isochrones (in Days) for the Inverted Permeability Field.

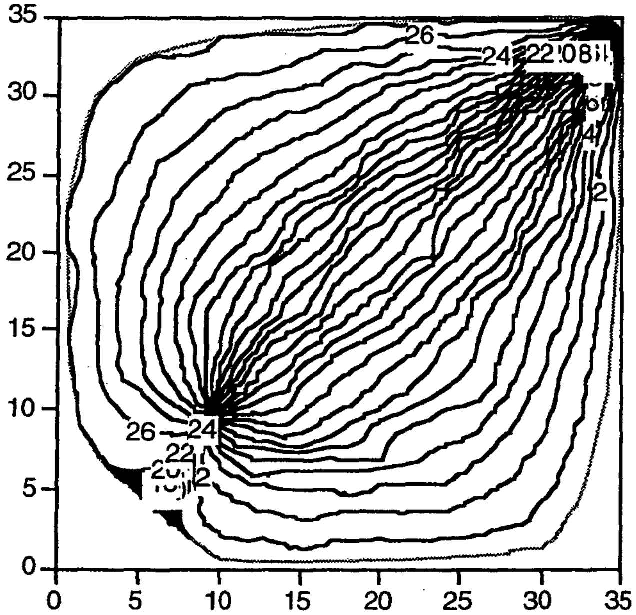



Fig. 5b. Streamlines for the Inverted Permeability Field.

show the isochrones and the streamlines from the inverted permeability field. On comparing with Fig. 3, we now see that many of the important features of the original flow field have not been reproduced by the inversion. For example, the streamlines are more uniformly spaced in Fig. 5b compared to Fig. 3b indicating lack of preferential flow patterns as present in the synthetic data. The permeability field obtained by the inversion is shown in Fig. 6. The permeability field appears to be noisy with scattered patches of high permeability regions. Clearly, the central high permeability streak present in the synthetic data has not been adequately reproduced by inversion. This example also emphasizes the inherent non-uniqueness underlying such inversions and the danger associated with basing performance predictions on a single model derived through inversion of limited data.

C. Ensemble Inversion

We now shift our focus from the search for a single model that fits the data best to inferences about the properties that are shared by an ensemble of acceptable models. Several inversions of the steady head and tracer data were performed and statistical analysis was carried out to discern the underlying features shared by the individual models.

Fig. 7 shows the energy vs. iterations from about 90 inversions of the data. The wide band of energy observed here indicates that many of the models actually converged to local minima. The ensemble mean and median models derived from these inversions are shown in Figs. 8a and 8b. The central high permeability region present in the data is now apparent in these models. On comparing with Fig. 6, which is a member of the ensemble, the mean or the median model appears much less noisy and the preferential flow paths are more clearly discernible here.

It is interesting to examine the characteristics of the individual models obtained by inversion in relation to the ensemble statistics. For this, we compared moving window semivariance estimates of individual models with the ensemble mean and median. The moving window semivariance describes the variance of a moving window of size h as a function of h (scale) and is defined as follows:[12]

$$\gamma_N(h) = \frac{1}{\int_0^h \xi^{d-1}\, d\xi} \int_0^h \xi^{d-1}\gamma(\xi)\, d\xi \qquad (10)$$

where d is the dimension in Euclidean space and $\gamma(h)$ is the classical semivariance estimator defined as follows:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2. \qquad (11)$$

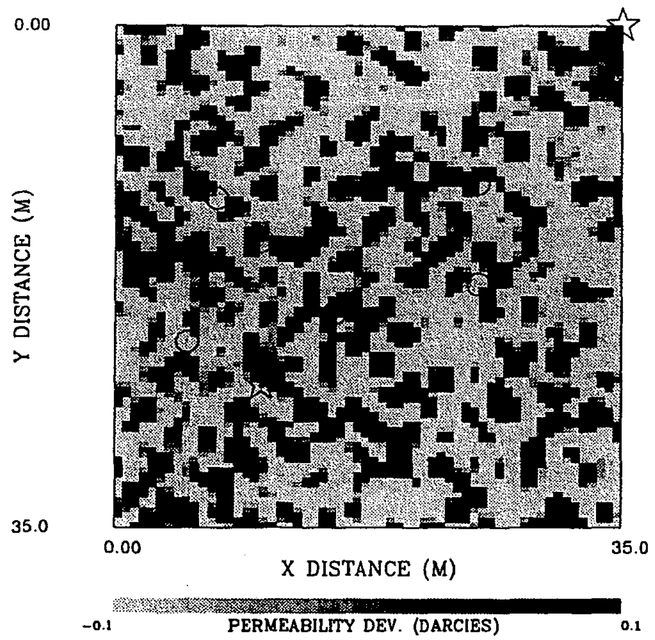



Fig. 6. A Single Realization of Permeability Field Obtained by Inversion.

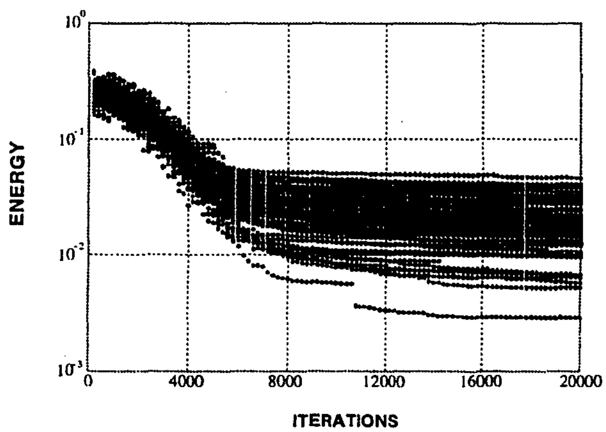



Fig. 7. Energy vs. Iterations for an Ensemble of Inversions (86 Models).

In Eq. (11), $z(x_i)$ is the data value at $x_i$ and N(h) is the number of data pairs for lag distance h. The moving window semivariance estimator has been shown to be robust and superior to other such existing estimators and can be used to examine the 'scale effects' (variance vs. scale) of heterogeneity.[12]

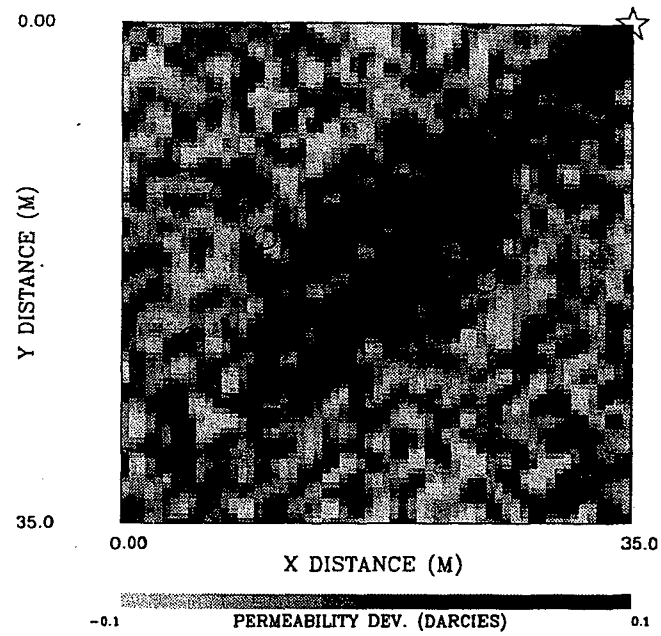



Fig. 8a. The Ensemble Mean Model.

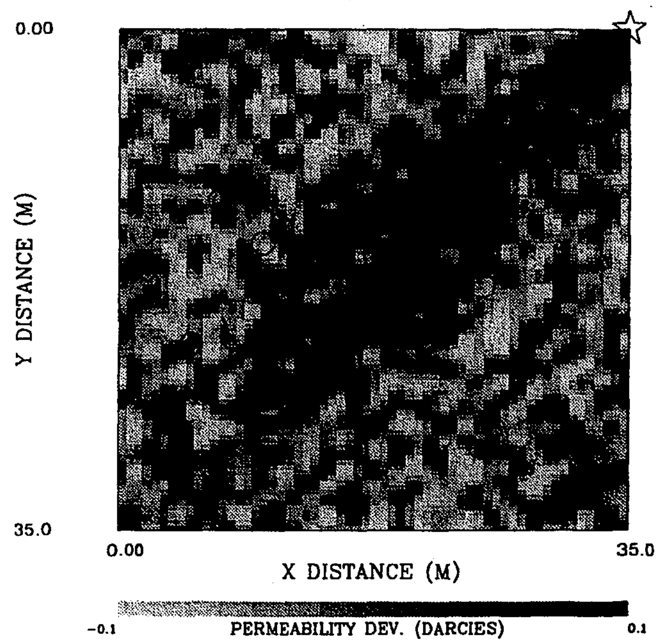



Fig. 8b. The Ensemble Median Model.

Fig. 9 shows the histograms of moving window semivariance estimates of individual models obtained by inversions. The semivariance for the ensemble median is also superimposed in the same figure (shown using asterisk symbols) along with the data from the synthetic fracture plane (shown as dashed lines). The re-

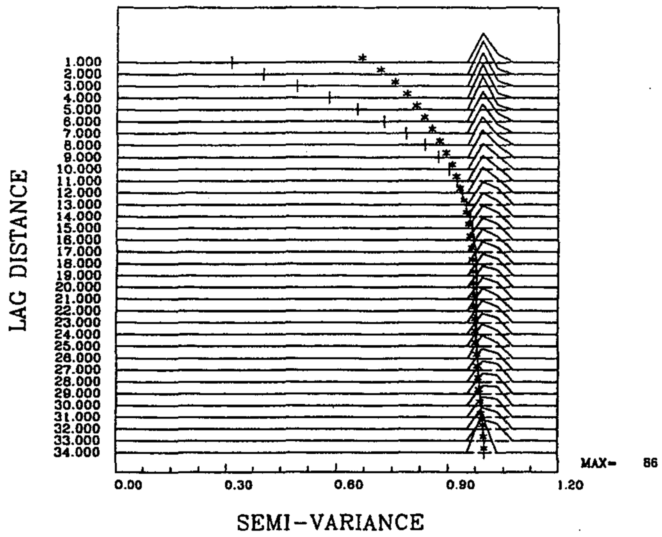



Fig. 9. Moving Window Semivariance Estimates for the Ensemble of Models and the Median Model.

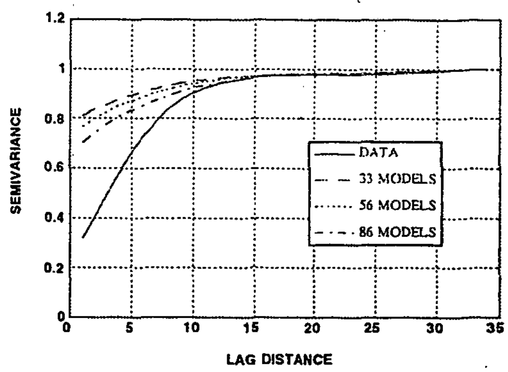



Fig. 10. The Ensemble Median as a Function of Number of Realizations.

sults clearly indicate that whereas the individual models are quite noisy, such small scale fluctuations disappear in the ensemble median. Thus the underlying structure can be inferred more clearly through the ensemble analysis. The behavior of the ensemble median as a function of number of realizations is shown in Fig. 10 which again underscores the need for a large number of models that fit the data.

## IV. SUMMARY AND CONCLUSIONS

1. An approach for generation of stochastic permeability fields through simultaneous inversion of flow and transport data has been presented. For data inversion, we have used simulated annealing which was found to be very robust and computationally effective, particularly for large number of parameters. The tracer transport was computed using a semianalytic transit time algorithm which is fast, accurate and free from numerical dispersion.

2. Whereas individual models obtained by inversion were found to be noisy and highly non-unique, we have addressed the non-uniqueness issue by focusing on the determination of underlying features shared by an ensemble of acceptable models rather than trying to obtain a model that fits the data the best.

3. The ensemble approach was particularly successful in reproducing the major features of heterogeneity and the preferential flowpaths. An examination of moving window semivariance estimates of the individual models indicated that the ensemble analysis resulted in natural cancellation of small scale fluctuations, thus facilitating the determination of the underlying correlation structure.

## ACKNOWLEDGMENTS

This work was carried out under U. S. Department of Energy Contract No. DE-AC03-76SF00098, for the Director, Office of Civilian Radioactive Waste Management, Office of External Relations, administered by the Nevada Operations Office in cooperation with the Swiss National Cooperative for Disposal of Radioactive Waste (NAGRA). The authors thank Yvonne Tsang and Janet Jacobsen for their careful review of the manuscript.

## NOMENCLATURE

| | |
|---|---|
| a,b,c | Parameters for bilinear streamfunction |
| $C_{j,obs}$ | Observed tracer response |
| $C_{j,cal}$ | Calculated tracer response |
| k | Permeability, $L^2$ |
| $P_{j,obs}$ | Observed pressure response, $ML^{-1}T^2$ |
| $P_{j,cal}$ | Calculated pressure response, $ML^{-1}T^2$ |
| $v_x$, $v_y$ | x and y velocities, L/T |
| $\Delta tx$, $\Delta ty$ | Transit times in the x and y directions, T |
| $\gamma(h)$ | Semivariance |
| $\lambda$ | Mobility ( = $k/\mu$) |
| $\mu$ | Viscosity, $ML^{-1}T^{-1}$ |
| $\psi$ | Streamfunction, $L^2/T$ |
| $\tau(\psi)$ | Transit time function, T |

REFERENCES


1. Carrera, J. and Neuman, S. P.,"Estimation of Aquifer Parameters Under Transient and Steady State Conditions: 3. Application to Synthetic and Field data," *Water Res. Res.*, 22(2), 228-242, 1986.

2. Long, J. C. S., Doughty, C., Hestir, K. and Martel, S., "Modeling Heterogeneous and Fractured Reservoirs with Inverse methods Based on Iterated Function Systems," in Reservoir Characterization III, Bill Linville, Ed., pp 471-503, PennWell Books, Tulsa, OK, 1993.

3. Datta Gupta, A. and King, M. J., "A Semianalytical Approach to Tracer Flow Modeling in Heterogeneous Permeable Media," Submitted to *Advances in Water Resources*, August, 1993.

4. Vasco, D. W., Johnson, L. R., Majer, E. L.,"Ensemble Inference in Geophysical Inverse Problems," *Geophys. J. Int.* (in press), 1993.

5. Deutsch, C. V. and Journel, A. G., *Geostatistical Software Library and User's Guide*, Oxford University Press, New York (1992).

6. Bear, J., *Dynamics of Fluids in Porous Media*, Dover Publications, Inc., New York, p 224.

7. Pollock, D. W., 'Semianalytical Computation of Pathlines for Finite-Difference Models," *Ground Water* (1988) 26, 6, 743-750.

8. Goode, D. J. "Particle Velocity Interpolation in Block-Centered Finite Difference Groundwater Flow Models," *Water Resources Research*, 26 (5), 925-940, 1990.

9. Moreno, L. and Tsang, Chin Fu, "Multiple-Peak Response to Tracer Injection Tests in Single Fractures: A Numerical Study," *Water Resources Research*, 27(8), 2143-2150, 1991.

10. Sen, M. K., Datta Gupta, A., Stoffa, P. L., Lake, L. W. and Pope, G. A., "Stochastic Reservoir Modeling Using Simulated Annealing and Genetic Algorithm," SPE 24754 (to be Published in *SPE Formation Evaluation*).

11. Kirkpatrick, S., Gelatt, C. D., Jr. and Vecchi, M. P., 1983, "Optimization by Simulated Annealing," *Science* (1983), 220, 671-680.

12. Li, Dachang and Lake, L. W., "A Moving Window Semivariance Estimator", submitted to *Water Resources Research*, 1993.

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720