RAPID TOOLS FOR JOINT INVERSION AND IMAGING

Final Report

D. M. Rizzo and D. E. Dougherty

DOE Patent Clearance Granted

Date 2/2/01
Daniel D. Park
(630) 252-2308
E-mail: daniel.park@ch.doe.gov
Office of Intellectual Property Law
DOE Chicago Operations Office

Subterranean Research, Inc.
Burlington, Vermont 05401

August 2000

Prepared for

THE U. S. DEPARTMENT OF ENERGY
Under Award No. DE-FG02-97ER82485

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

TABLE OF CONTENTS

TABLE OF CONTENTS.....	2
PROJECT SUMMARY.....	2
1. INTRODUCTION.....	3
1.1 JEDII.....	3
1.2 RITA.....	4
2. RESCALING.....	6
2.1 UPSCALING.....	6
2.2 DOWNSCALING.....	8
3. SEGMENTATION AND DOMAIN DECOMPOSITION.....	11
4. KALMAN FILTER.....	14
5. EXTENDED KALMAN FILTER.....	17
6. GPR RITA.....	21
6.1 COMPONENT 1 – USING AN ANN TO PROVIDE INITIAL STRUCTURE	21
6.2 COMPONENT 2 – USING APPRIZE™-GPR FOR INVERSION OF GPR TRAVELTIMES	21
7. HYDRAULIC RITA.....	25
7.1 COMPONENT 1 – USING AN ANN TO PROVIDE INITIAL STRUCTURE	26
7.2 COMPONENT 2 – USING APPRIZE™-HYDRO FOR ESTIMATING HYDRAULIC CONDUCTIVITY	26
8. TECHNICAL AND ECONOMIC FEASIBILITY.....	28
8.1 TECHNICAL FEASIBILITY.....	28
8.2 ECONOMIC FEASIBILITY.....	28
9. REFERENCES.....	29
10. GLOSSARY.....	30
APPENDIX A -- CPT SOIL CLASSIFICATION	32

PROJECT SUMMARY

The objectives of this Small Business Innovative Research (SBIR) Phase II project were related to the joint use of sparse environmental data to map or characterize the distribution of one or more subsurface material properties (parameters) or conditions (states). Two distinct classes of data were considered: a point or grab sample (e.g., a soil core or aliquot of water), and a signal that propagates through a region filled with material to a receiver from a source (e.g., a pumping test or cross-hole electrical resistivity tomography). Artificial neural network and extended Kalman filter methods were researched and developed. Methods of upscaling, downscaling, and segmenting data for analysis were investigated. Testing was performed using ground penetrating radar and groundwater hydraulic applications. Segmentation of data for analysis was shown to offer advantages for large data sets. Upscaling has an important role in robust inversion, while downscaling was found to be unhelpful. The combination of different data types was shown to provide significant benefit, at least in some cases.

1. INTRODUCTION

This final report summarizes a Small Business Innovative Research (SBIR) Phase II project entitled "Rapid Tools for Joint Inversion and Imaging".

The objectives of this project were related to the joint use of sparse environmental data to map or characterize the distribution of one or more subsurface material properties (parameters) or conditions (states). Two distinct classes of data were considered: a point or grab sample (e.g., a soil core or aliquot of water), and a signal that propagates through a region filled with material to a receiver from a source (e.g., a pumping test or cross-hole electrical resistivity tomography).

The remainder of this section describes the key elements of this work, which are called JEDII and RITA. Sections 2 through 7 present summaries of technical tasks and results. Section 8 reviews technical and economic feasibility of the project. References and a glossary conclude the body of the report.

1.1 JEDII

The Joint Environmental Data Imaging and Inversion (JEDII) software technology is devised for joint, fully three-dimensional data analysis and interpretation. JEDII (see Figure 1) may be viewed as a controller that orchestrates the joint data inversion algorithms or modules known as RITAs (Rapid Image and Inverse Tomography Algorithms). JEDII contains a number of RITAs for inverting multiple data types, as well as the software to coordinate the RITAs and manage associated data. Each active RITA processes a user-designated set of data to estimate a "target" parameter (e.g., ground-penetrating radar (GPR) traveltimes data might be inverted to "target" propagation velocities).

Figure 1. Schematic of the Joint Environmental Data Imaging and Inversion (JEDII) software technology.

The inversion is augmented using one or more types of data that JEDII has determined are correlated to the “target” data type. (For example, if electrical conductivities and grain-size distributions are available at a specific site and if these data are determined to be statistically correlated to GPR “slowness”, then they would be used as “secondary” data in the traveltime-data inversion process.) JEDII uses an ANN in conjunction with traditional statistical regression methods to determine (1) correlation *between* data types and (2) horizontal and vertical correlations *within* data types.

1.2 RITA

A major component of the JEDII technology is the Rapid Image and Inverse Tomography Algorithm (RITA). RITAs for ground penetrating radar data and one for hydraulic data were developed during Phases I and II. These use proprietary approaches to imaging and inverting specific geophysical signals and are described in greater detail in Sections 2 through 5.

A key conclusion from our Phase I research was that an artificial neural network approach to inversion of geophysical signals was not the best way to attack the entire problem. Although technically feasible, the approach had limitations related to its technical *efficiency*—especially related to rates of convergence and sensitivity of inversion results to the way that data are segmented or partitioned. However, the ANN approach did provide exceptional value during the initial estimation phase of geophysical reconstruction. Whereas detailed inversion was more efficiently performed using other more traditional inversion methods, once a good initial estimate is in hand. These observations led us to modify our technical approach for Phase II.

As a result, data inversion methods were incorporated into the RITAs during Phase II that are not based on artificial neural networks (ANNs). This allowed us to capitalize on the strengths of the ANNs for rapid assimilation of large amounts of disparate data into preliminary estimates of the statistical structure of the parameter space and then, where appropriate, further condition the estimates using optimal estimation techniques. Additional innovations enhanced the computational efficiency. The advantages of this approach are shown the examples of Sections 6 and 7.

Each of the RITAs that make up JEDII are a system comprised of (1) an artificial neural network, SCANN (Site Characterization using Artificial Neural Networks), which is used to provide an initial estimate of the parameter structure using multiple types of data and (2) an approximated extended Kalman Filter approach called APPRIZE™ (Automatic Progressive Parameter-Reducing Inverse Zonation and Estimation) which is used to provide customized inversion (see Figure 2). A patent for the APPRIZE™ technologies was issued in May 2000 [Eppstein and Dougherty, 2000].

Finding the right ANN architecture and using customized inversion tools with reasonable model parameter limits, is simpler, less time consuming, and, in our experience, more robust than considering larger, more general ANNs (trained to incorporate many model parameters and data types) and inversion tools. As a result, we devised a RITA for each specific data

type, while JEDII contains and orchestrates the various RITAs working on a particular multi-signal data set.

Figure 2. Schematic of the RITA used for inversion of GPR traveltimes comprised of an artificial neural network, SCANN and a customized inversion tool, APPRIZE.

Section 6 describes the RITA developed for inverting GPR traveltime data to “target” propagation velocities. Two imaging and inversion methods are applied in a tandem strategy to 3-D cross-well GPR traveltime data. Section 7 describes the RITA developed for the inversion of hydraulic data.

2. RESCALING

Rescaling refers to either *upscaling* (taking a state or parameter from one length scale to a longer one), or *downscaling* (the converse).

2.1 UPSCALING

Upscaling is the process of converting information that is resolved on one scale (let's call the original scale of resolution δ) to a larger scale (that is, information at some lesser resolution). Let us denote the new, coarser representation by $a\delta$, where a is a constant greater than 1.

There are several reasons for upscaling distributions of geologic media properties. Among these are the following:

- An upscaled representation requires less computer memory.
- If an upscaled representation is used by a simulation model that has as characteristic length that scales with the media properties, then an upscaled representation will lead to reduced computer resources for simulation (e.g., time, memory, communications, class of computer).
- An upscaled representation may be sufficient for the purposes to which the media properties are being applied, i.e., a decision may be needed about regional issues rather than localized ones, so localized information can be upscaled.
- An upscaled representation of medium properties may be more compatible with the information used to estimate the medium property values. In the context of inverse methods, for example, the signals that are used to infer medium properties may not contain enough information to estimate at scale δ , but may be sufficient for estimating at the larger scale $a\delta$.

The basic idea is illustrated with an example. Say that the medium occupies a right parallelepiped with sides X , Y , and Z , and volume V . Assume that this volume is divided into small cubic volume elements ("voxels") with sides of length δ such that there are N cubes that do not overlap and that they totally fill the volume V . Let a property (or state) P be represented by a constant value within each voxel and that the value can change from voxel to voxel.

Now let us consider how to upscale. The basic idea is to establish a new set of larger volume elements and, using some procedure, assign the N values of P from the cubic voxels to the larger voxels. Let us assume for discussion that the new voxels are cubic and are formed from the union of $2 \times 2 \times 2$ of the original voxels. That is, the new voxels are twice as long per side and eight times the volume. Figure 3 illustrates this example.

Upscaling can be accomplished in several ways: *ad hoc* averaging, theory-based averaging, process-simulation based averaging.

Figure 3: Domain with 16 cubic voxels is upscaled by a factor of 2 in each of 3 dimensions to 2 cubic voxels.

Figure 4. Example numbering of a group voxels for use in upscaling calculations.

An *ad hoc* method is not driven by physics or mechanistic explanations of phenomena. Instead it may be developed from convenient expressions or intuitive rules. Consider the 8 voxels in Figure 3 above and let P_i for $i=1\dots 8$ be the values of P in each of these 8 voxels, which are numbered as in Figure 4. Let P^{eff} be the upscaled, or effective, value. Following are some examples:

- Convenient: Assume that P is log-normally distributed. Then the geometric mean might be used as the effective value. That is,

$$P^{eff} = \left(\prod_{i=1}^8 P_i \right)^{1/8}.$$

- Intuitive: Use the minimum value of P_i because an expert's opinion is that the low values of P dominate the behavior of the system. That is,

$$P^{eff} = \min_{i=1\dots 8} P_i.$$

There are a variety of theoretically based methods, which vary with the parameter being upscaled and the type of signal being used to infer the parameter values. We include both statistical and mechanistic theoretical methods under this classification. Following are a couple of examples:

- Statistical: Assume all 8 values are drawn from the same population; then an unbiased single representative value is the mean. If the mean is not known, then use the average of data values as the best estimate of the mean. That is,

$$P^{eff} = \frac{1}{8} \sum_{i=1}^8 P_i.$$

Similar analyses could be used to upscale the variance.

- Analytical: If the parameter P represents electrical resistivity and if the electrical current is left to right, then continuity and Ohm's law can be used to show that the effective resistivity is determined from the harmonic averages of resistivities in each left-right pair of voxels (resulting in 4 values) followed by taking the arithmetic average of these harmonic averages.
- Stochastic Analysis: If the parameter P represents hydraulic conductivity and if the mean head gradient J is left to right, and if P is log-normally distributed with small variance σ^2 , then the effective conductivity is

$$P^{eff} = P_G \left[1 + \frac{\sigma_Q^2}{2} (1 - \lambda) \right],$$

where P_G is the geometric mean of P , σ_Q^2 is the variance of the log of P , and λ is a structurally dependent coefficient. (For P isotropic in all three dimensions $\lambda = 2/3$.)

- Equivalent Medium (Renormalization): In this method, we assume that the 8 voxels exist within an infinite homogeneous medium and we assume that we specify appropriate boundary conditions. Then, introducing 1 heterogeneity at a time into the infinite homogeneous medium results in perturbations of the observable quantity (e.g., head, travel time, etc.). These perturbations are averaged over all the heterogeneous voxels. To obtain the effective parameter value for the specified boundary conditions. Since the value of the homogeneous parameter in which these are imbedded is not known ahead of time, we can set it equal to the determined effective parameter value. For isotropic media with small variance, the equivalent medium theory is exactly the same as the properly linearized stochastic analysis.

For this project, we have used clustering methods to identify the zones that will be combined to form an upscaled (or rezoned) variable. Statistical upscaling is used to find the mean of the upscaled zone. In addition, we use another statistical method to upscale the variances among zones, called random field union [Eppstein and Dougherty, 1996]. We have applied these methods for normally- and log-normally-distributed variables. This upscaling approach preserves mean values and variances.

2.2 DOWNSCALING

Downscaling is the converse of *upscaling*. It is the process of converting information that is resolved on one scale (let's call the original scale of resolution δ) to a smaller scale (that is, information at some greater resolution). Let us denote the new, finer representation by $b\delta$, where b is a constant less than 1.

In subsurface science downscaling has received much less attention than upscaling. However, downscaling has been around for many years in the literature of numerical methods for fluids, in particular for meteorology and turbulence modeling. When there are property variations or fluctuations at scale $b\delta$ and simulations occur on scale δ , "sub-grid parameterization" models are used to convert "subgrid" processes to "grid-scale" processes

using effective coefficients that are based on models of what happens at the sub-grid scale. For example, one may assume isotropic stationary turbulence to develop "eddy diffusion" coefficients for grid-scale models that diffuse momentum at approximately the same rates as would be observed if the sub-grid scale momentum transfer were fully modeled. After calculating the grid-scale fluxes and knowing the sub-grid parameterization model, the grid-scale information can be applied approximately on the smaller scale, i.e., sub-grid scale.

Downscaling is also related to multigrid methods. One takes an interim solution to a numerical linear algebra problem and projects it onto a coarser grid (upsampling the solution), where other iterations occur. Eventually, the coarse grid solution is extrapolated back onto the fine grid (downsampling the solution). Multigrid methods are very well developed for linear systems arising from isotropic media, and semi-coarsening methods are being used to solve strongly isotropic methods with increasing robustness.

There are several reasons for downsampling distributions of geologic media properties. Among these are the following:

- A downscaled representation provides a more highly resolved representation of the distribution.
- If a downscaled representation is used by a simulation model that has a characteristic length that scales with the media properties, then a downscaled representation will lead to reduced errors for simulation (e.g., discretization errors are reduced due to smaller grid size, model error may be reduced by more detailed and resolved parameters).
- A downscaled representation may be necessary for the purposes to which the media properties are being applied, i.e., a decision may be needed about localized issues rather than regional ones, so it is beneficial to downscale the regional-scale information.
- A downscaled representation of medium properties may be more compatible with the information used to estimate the medium property values. In the context of inverse methods, for example, the signals that are used to infer medium properties may contain more than enough information to estimate the parameter at scale δ , and may be sufficient for estimating at the smaller scale $b\delta$.

Now let us consider how to downscale. The basic idea is to start with one set of cubic N volume elements and establish a set of smaller volume elements, such that both sets occupy the same volume. We then devise a downsampling procedure to assign the N values of P from the larger cubic voxels to the smaller voxels. Let us assume for discussion that the new voxels are cubic and are formed from the division of each of the original voxels into $2 \times 2 \times 2$ smaller voxels. That is, the new voxels are half as long per side and one-eighth the volume of the original, larger voxels. This is illustrated in Figure 5.

Downsampling can be accomplished in exact or approximate ways. Exact downsampling can be accomplished if upsampling has occurred *and* if the information typically lost during upsampling is retained. This is similar to a hierarchical storage method that retains 100% of the content of an image and allows one to repaint the image on the screen at reduced scales of resolution. Approximate downsampling can be thought of as trying to recover an image after it has been passed through a lossy image compression algorithm, such as JPEG. There is no

Figure 5: Domain with 2 cubic voxels is downsampled by a factor of 2 in each of 3 dimensions to 16 cubic voxels.

way to actually recover the full quality of the original resolution. However, there are several ways to go from very blocky resolution to more representative, improved resolution.

When applying the APPRIZE™ methodology, one has the parameter value on the coarse block and the variance on the same coarse block. One approach is therefore to assign these quantities to the finer blocks. A second method is to assign the 8 finer blocks 8 differing values whose average is the coarse block parameter value and whose variances upscale to the variance of the coarse voxel. Then the APPRIZE™ estimation algorithm is applied to the new downsampled parameterization.

As in weather modeling, one can postulate a model of variability on the subgrid (small) scale. For example, one may infer that the parameter is “grading upward”, which means that the parameter value is essentially constant in the horizontal plane and changes linearly with the vertical location. An estimate of the strength of the gradation (i.e., the magnitude of the gradient) is all that is necessary to estimate the 8 parameter values. Additional information is needed to obtain the variances; it is typical to say either that variance is independent of solution variable or that it is proportional. In either case, this allows us to assign variances from the coarse voxel to the smaller (downsampled) voxels. Again, the APPRIZE™ method for data assimilation allows these downsampled estimates to be adjusted on the basis of observational data.

For this project, we have investigated downscaling methods for use with inversion methods. The approach for improving the performance of inverse methods is to start with coarse scale application of APPRIZE™ and then downscale, so that additional degrees of freedom are introduced and can be adjusted by APPRIZE™ in a second round of data assimilation.

Studies were conducted using the APPRIZE™-MODFLOW assimilator-model pair, i.e., assimilating data into a diffusion equation to estimate the diffusion coefficient (in groundwater, this coefficient is called the hydraulic conductivity). Both of the downscaling methods described in preceding paragraphs were tested in which the downscaling coefficient b was 2. Results were disappointing; essentially no improvements are recoverable by downscaling. We believe this is due to the fact that the number of observations is quite small relative to the number of hydraulic conductivity values being estimated. Inasmuch as the data density used in our experiments exceeds the typical field data density, we have concluded that this downscaling approach works, but is not currently useful.

3. SEGMENTATION AND DOMAIN DECOMPOSITION

Segmentation is the process of partitioning a set of information or data into segments or subsets, which can be analyzed segment-by-segment instead of all-at-once. This process is in some ways analogous to *domain decomposition* (which is used to partition sets of linear equations) and to *mesh generation* (which is used to partition a domain into cells or elements). Let us denote a set of information or data as D , with members d_n and $j=1 \dots N$. Note that $D = \cup d_n$ and $\cap d_n = 0$. In segmentation, we partition D into segments D_m where $m \in I$ and I is a numbering that refers uniquely to all the segments, $m=1 \dots M$. Each segment contains one or more members of the set D . We require that $D = \cup D_m$. If $\cap D_m = 0$, then the segments are *non-overlapping*; otherwise the segmentation is *overlapping*.

With these definitions in hand, it is a simple matter to devise a segmentation for any discrete set. Selecting a specific segmentation requires that we state an objective for segmenting; then we seek a reasonable and effective algorithm to perform the segmentation.

In the context of inverse problems, segmentation has several possible objectives.

- Reduce the number of parameters being estimated at one time. By loading only part of the segments at a time, we reduce the amount of computer memory resources needed to hold information or data while it is being processed. For example, we might partition the parameter space into as many segments as needed to match computer resources available. For example, if D requires twice as much computer resources as we have available, then we will segment D into at least 2 segments. For two segments we can send the first $N/2$ members of D to the first segment (D_1) and the remainder to the second segment (D_2).
- Reduce the number of signals being used at one time to perform parameter estimation. In time-recursive estimation methods this is common, we process data from time t_1 before incorporating data from time t_2 and so on. Although the procedure is less frequently encountered when discussing the spatial domain, it is follows directly. We partition the spatial data into segments and assign an ordering to the partitions. Then the spatial data segments are used in the sequence given by the ordering, which give a time-like nature to the segments.
- Reduce the number of receivers being used at one time to perform parameter estimation. Transmitter-receiver (i.e., source-receiver) pairs are used for GPR (antenna-antenna, seismic (hammer-geophone), hydraulic (well-piezometer), and other geophysical tests. Therefore, a splitting of the data can occur by receivers equally well as by transmitters.
- Reduce the number of parameters that are free to be modified on the basis of transmitter-receiver data by using mechanistic understanding to select only the important ones. For example, say there is a GPR source and a distant receiver. The travel time between the two is dependent on the near-source details, on the velocity of the wave along the path taken between the source and receiver, and to a lesser extent on material properties non-local to the raypath. This idea is used in the interpretation of CT scans and most commonly used GPR interpretation methods. Usually, a straightline approximation of the raypath from source to receiver is used

to determine the parameter values that are free and should be modified on the basis of that particular traveltime measurement. In some cases however, the raypaths are not straight and need to be more carefully evaluated. For example, a waveguide can elicit an anomalous response.

As part of this project, several segmentation methods have been devised, implemented, and tested. We focus here on non-overlapping methods:

Contiguous cluster approach. Given an ordering for the voxel-based discretization of parameter values G , an estimate $P(g)$ for $g \in G = 1 \dots n_G$, and a C clusters of distance parameter values. Ignoring the spatial locations of the voxels g , find a new ordering G' such that the estimated material properties. Here, and in APPRIZE™, we use a k -means clustering algorithm to partition the voxel data into different clusters. Once each voxel has been assigned to one of C clusters based purely on parameter value, then a spatial “disaggregation” step is applied to each cluster in turn. Call c the vector that assigns each voxel an index value into the C clusters. The members of c are reindexed (regrouped) according to spatial- as well as value-based similarities based on contiguity. The algorithm is simple:

0. Let $L=0$.
1. If c is empty, go to step 6. If not, continue.
2. Let $L=L+1$.
3. Choose a member of c , call it the “seed” s_L , remove it from the list c , and assign it to a set c_L .
4. Find all members of c that have the same value as s_L and that are spatially adjacent to the voxel corresponding to s_L , assign them to c_L , and delete them from c .
5. Go to Step 1.
6. For each of the voxels v , find the set c_l ($l=1 \dots L$) to which it belongs and assign to $c(v)$ the value of the set c_l .

Each element $c(v)$ of the resulting vector c contains an integer index that describes which parameter value (of the C clustered values) has been assigned to voxel v .

“Card dealer” approach. This method is simpler than the contiguous cluster approach. To create S segments, we simply assign voxel v to segment $v \bmod S$ (that is the remainder from the integer division of v by S). This is like “dealing” out voxels to segments.

“Group-deal” approach. This method is like the “card-dealer” approach except that more than one voxel (which are adjacent in terms of voxel number) at a time is assigned to the segments.

Slices approach. In many cases, the geometric centers of the voxels occur in regular lattices or grids. In many applications, there are planes (or planelike structures) of voxel centers and these planes are Cartesian (that is, perpendicular to x , y , or z axes). We choose a slice axis (e.g., x , y , or z) and a number of segments S . Then we can deal slices that are perpendicular to the given slice axis into the segments S using either the dealer or group-deal approaches.

Mode-based approach. The voxels that parameterize a domain can be described using voxel centers and by adjacency matrix. A common approach used in parallel computing on finite element meshes is to use the adjacency list to construct a connectivity matrix. The modes (eigenfunctions) of the mesh can be found with some linear algebra. The modal responses, and in particular the modal lines, can be used to assign voxels to different segments. One either looks for an eigenfunction with S regions of positive and negative sign, or one can recursively construct 2 segments, then 2 segments of each of these, and so on.

Experiments show that domain decomposition (i.e., windowing, or segmenting) is beneficial, and sometimes provides the only way to assimilate a full data set. Unfortunately, we have not been able to determine a “universal” domain decomposition method that works for all problems.

We note that segmentation or domain decomposition is a bookkeeping proposition. There is no up- or down-scaling, which require corresponding losses or gains in information. There is no need to modify any covariance matrices, as may be used with the APPRIZE™ methodology.

The sliding window approach described by Eppstein and Dougherty [1998a, 1998b] and Eppstein et al. [1999a, 1999b] has been quite useful. It is an overlapping segments approach. Our general recommendation is to choose the largest segments possible (i.e., the smallest number of segments) and to use an easy-to-visualize segmentation approach.

4. KALMAN FILTER

The *Kalman Filter* (KF) is a method for improving estimates of states (e.g., piezometric heads) using observational data with a simulation model in a stochastic setting. The KF is strictly applicable to linear systems only; the extended Kalman filter (described in the next section) addresses nonlinear systems. The idea in KF is to use a simulation model to predict states, actually observe states, and then correct the model outputs in a “rational” manner before proceeding to the next round of observations. See Gelb [1979 and Menke [1984] for background.

The starting point is the representation of a state variable, of the measurement process, and of its dynamics. Say the true state variable of a system is denoted by ζ . If we take measurements, we convolve the state with a measurement operator and errors are committed. In discrete form, let us therefore approximate the dynamics as

$$z = H\zeta(p) + w + v \quad (1)$$

where H is the measurement operator, p is a vector of parameter values, w is a random system error, v is a random measurement error, and z is a recorded measurement. The expected value of the vector v is zero, i.e., $E[v]=0$, and it has a variance matrix $\text{Var}[v]=R$. Similarly, $E[w]=0$ and $\text{Var}[w]=Q$, where Q is a matrix. The state ζ is itself never actually observed; only measurements z are recorded.

Consider, for example, Darcy's law for flow of a homogeneous fluid through a porous medium at low Reynolds number. When combined with the mass balance equation, the Darcy equation may be written as

$$S \frac{\partial h}{\partial t} - \text{div}(K \text{grad } h) = q \quad (2)$$

where K is the hydraulic conductivity (parameter), S is the storativity, h is the head (state), and q is a controllable quantity for experimentation. Given K , q , and auxiliary data, we can use a simulation model to predict h . In discretized form, we solve $Ah=f$ for h , where A is a matrix, f is a forcing vector, and h is a vector of heads. Measurements of h will surely indicate errors. An “innovation” is defined to be the difference between a predicted head h_{pred} and an observed head z for the same time and spatial location. If there are estimates of measurement errors (R) and systematic errors (Q) that result from a number of sources (e.g., spatially correlated errors in the estimate of K , which are expressed as a covariance matrix), then the KF theory can be applied. The result is an updated head h_{update} as well as an updated covariance matrix.

We use the discrete form of the groundwater flow equation, which is written in matrix form as

$$Ah^{n+1} = Bh^n + f^{n+1} \quad (3)$$

where A is the so-called stiffness matrix, B is the so-called mass matrix, f is the forcing vector, and the superscript n indicates time level t^n . This is rearranged into the state-space representation

$$h^{n+1} = \Phi h^n + w^{n+1}$$

which is used in the KF method. Here h and w are $N \times 1$ vectors, and Φ is an $N \times N$ matrix, where N is the number of discretized states.

The KF algorithm follows, and is inserted in a recursive loop that represents the advance of time:

FORECASTING PHASE:

$$\begin{aligned} h_{\text{pred}}^{n+1} &= \Phi h_{\text{update}}^n + w^{n+1} \\ P_{\text{pred}}^{n+1} &= \Phi P_{\text{update}}^n \Phi^T + Q^n \end{aligned}$$

DATA ASSIMILATING PHASE:

$$\begin{aligned} K_G^{n+1} &= P_{\text{pred}}^{n+1} H^{n+1T} [H^{n+1} P_{\text{pred}}^{n+1} H^{n+1T} + R^{n+1}]^{-1} \\ h_{\text{update}}^{n+1} &= h_{\text{pred}}^{n+1} + K_G^{n+1} [z^{n+1} - H^{n+1} h_{\text{pred}}^{n+1}] \\ P_{\text{update}}^{n+1} &= [I - K_G^{n+1} H^{n+1}] P_{\text{pred}}^{n+1} \end{aligned}$$

Here the superscript n indicates time level t^n , superscript T indicates the transpose operation, h is the vector of states (a $N \times 1$ vector), w is an $N \times 1$ vector of forcing, Φ is a transition matrix, P is the autocovariance matrix of states (a $N \times N$ matrix), Q is a systematic noise matrix (a $N \times N$ matrix), H is a measurement matrix (a $M \times N$ matrix where M is the number of measurements at time $n+1$), K_G is a so-called Kalman gain matrix (an $N \times M$ matrix), z is the vector of measurements (a $M \times 1$ vector), and the superscript -1 indicates the matrix inverse.

The forecasting phase makes predictions about the future value of the state, based on a current estimate of the state plus information about the physics and forces acting on the system. The first equation forecasts the state and the second predicts the autocovariance matrix relating states at different spatial locations and the same instant of time.

The data assimilating phase uses the differences between observations (measurements) and predictions, the innovations, to improve the estimates of state and autocovariance. The first step is to complete the forecasting phase. Then the Kalman gain matrix K_G is calculated. The forecasts of state and autocovariance are then corrected or improved using final two equations.

If $M = 1$, then the Kalman gain matrix is 1×1 and it is calculated using a trivial inversion of a 1×1 matrix. Generally, the measurement matrix H is very sparse and computer implementations should take advantage of both sparse storage and matrix-vector operations.

Note that if the number of measurements M is zero, then the updated state is the same as the predicted value. Therefore, if there are S timesteps between measurements, then the simulation model (transition matrix) can be applied for S steps to propagate h and P , and then the data assimilation phase can be performed.

The last paragraph leads to the following implication. If the state variable is aqueous concentration of some chemical of concern, then the transition matrix represents the propagation of the chemical (due to advection, dispersion, reaction, etc.). It follows that measurements are used to provide an updated state estimate that forms the initial condition for the subsequent S simulation timesteps—that is, the initial condition for the next S

timesteps until the next set of measurements. The KF procedure can be considered as a method of "objective analysis" to develop data-driven initial conditions based on monitoring data. KF therefore can play a substantive role in the design, analysis, and execution of long term monitoring and stewardship plans.

When using geophysical information to estimate parameters, which typically involves one or more field equations that couple the parameters and observable states, there exists a nonlinear combination of the parameter and the state in the governing equation. This nonlinearity is outside the range of applicability of the standard Kalman Filter, and the Extended Kalman Fiter is employed.

5. EXTENDED KALMAN FILTER

The *Extended Kalman Filter* (EKF) is a method for estimating parameters p in equation (1) (typically material properties or boundary conditions) using observational data with a simulation model in a stochastic setting.

The Extended Kalman Filter is the culminating method in a hierarchy of estimators beginning with the well-known linear, deterministic, least squares method. Figure 6 illustrates this hierarchy.

Figure 6. Hierarchy of estimators. The green box describes the basic assumptions we would like to include. The pale yellow boxes indicate how estimates of a parameter are accomplished, moving from simpler to more complicated.

The EKF extends the original Kalman Filter (KF) from linear systems to nonlinear ones. As described in the previous section, the idea in KF is to use a simulation model to predict states, actually observe states, and then correct the model outputs in a “rational” manner before proceeding to the next round of observations. The nonlinearity is generally resolved using a first order Taylor series expansion, which linearizes the problem. See Gelb [1974].

Consider Darcy’s law for flow of a homogeneous fluid through a porous medium at low Reynolds number as an example for parameter estimation. The Darcy equation may be written as

$$-K \text{ grad } h = q$$

where K is the hydraulic conductivity (parameter), h is the head (state), and q is a controllable quantity for experimentation. A typical inverse problem is to find K given h and q . In the Kalman scenario, we predict quantities that are then measured and corrected. In this case, we are predicting head and conductivity and updating these predicted values on the basis of observations. It is clear from the Darcy equation that there is a product between K and the gradient of h —this constitutes a nonlinearity that must be accommodated to estimate K . In addition, like in KF, the EKF allows us to make corrections to estimates of covariances among parameters being estimated at various spatial locations.

In this project, several approximations of the EKF have been used in order to drastically improve performance on real-world data sets. The standard EKF algorithm is given in Figure 7 below.

Loop for each timestep

$$\mathbf{h} = \Phi \mathbf{h} + \mathbf{w} \quad (\text{forward simulator})$$

$$\mathbf{J} = \frac{\partial \mathbf{h}}{\partial \mathbf{K}}$$

$$\mathbf{P}_{hK} = \mathbf{J} \mathbf{P}_{KK} \Phi^T + \Phi \mathbf{P}_{hK}$$

$$\mathbf{P}_{hh} = \mathbf{P}_{hK} \mathbf{J}^T + \mathbf{Q}_z + \Phi \mathbf{P}_{hh} \Phi^T + \mathbf{J} \mathbf{P}_{hK} \Phi^T + \Phi \mathbf{P}_{zK} \mathbf{J}^T$$

If measurements $\mathbf{z}$ are available this timestep

$$\mathbf{K}^0 = \mathbf{P}_{hK} \mathbf{H}_z^T (\mathbf{R}_z + \mathbf{H}_z \mathbf{P}_{hh} \mathbf{H}_z^T)^{-1}$$

$$\mathbf{K}^K = \mathbf{P}_{hK}^T \mathbf{H}_z^T (\mathbf{R}_z + \mathbf{H}_z \mathbf{P}_{hh} \mathbf{H}_z^T)^{-1}$$

$$\mathbf{h} = \mathbf{h} + \mathbf{K}^0 (\mathbf{z} - \mathbf{H}_z \mathbf{h})$$

$$\mathbf{K} = \mathbf{K} + \mathbf{K}^K (\mathbf{z} - \mathbf{H}_z \mathbf{h})$$

$$\mathbf{P}_{hh} = (\mathbf{I} - \mathbf{K}^0 \mathbf{H}_z) \mathbf{P}_{hh}$$

$$\mathbf{P}_{KK} = \mathbf{P}_{KK} - \mathbf{K}^K \mathbf{P}_{hK}$$

$$\mathbf{P}_{hK} = (\mathbf{I} - \mathbf{K}^K \mathbf{H}_z) \mathbf{P}_{hK}$$

End if

End loop

Figure 7. Extended Kalman Filter algorithm. Highlighted lines are dropped in the Approximate Extended Kalman Filter method.

The so-called Approximate Extended Kalman Filter (AEKF), shown in Figure 8 below, is developed by neglecting second order terms in the full EKF. Removal of these terms, which are indicated by cyan-colored highlighting in the EKF figure (Figure 7), dramatically reduces execution times (up to 2 orders of magnitude).

Loop for each subset of measurements z

$$\mathbf{h}_{\text{all}} = \text{forwardsimulator}(\mathbf{K}_{\text{all}})$$

$$\mathbf{J} = \frac{\partial \mathbf{h}}{\partial \mathbf{K}}$$

$$\mathbf{P}_{hK} = \mathbf{J} \mathbf{P}_{KK}$$

$$\mathbf{P}_{hh} = \mathbf{P}_{hK} \mathbf{J}^T + \mathbf{Q}_z$$

$$\mathbf{K}^K = \mathbf{P}_{hK}^T \mathbf{H}_z^T (\mathbf{R}_z + \mathbf{H}_z \mathbf{P}_{hh} \mathbf{H}_z^T)^{-1}$$

$$\mathbf{K} = \mathbf{K} + \mathbf{K}^K (\mathbf{z} - \mathbf{H}_z \mathbf{h})$$

$$\mathbf{P}_{KK} = \mathbf{P}_{KK} - \mathbf{K}^K \mathbf{P}_{hK}$$

End loop

Figure 8. Approximate Extended Kalman Filter algorithm.

A further approximation can sometimes be made for the propagation of the covariance matrix. If we can write that the state and parameter covariances are a percentage of the state and parameter. In this case, once the state has been advanced in time and updated on the basis of observations, the covariances are determined by scaling the updated states. This obviates the need to repeatedly performing the more involved propagation.

As the systems being investigated become more nonlinear, the first order expansion used to develop the EKF becomes inaccurate. This leads to updates based on measurements that are too large and can lead to nonconvergence. In addition, this condition can lead to the calculation of non-physical negative covariances. To overcome these problems, we have borrowed an idea from transport simulation and limited the amount that a parameter can be changed from one iteration to the next. This slows, yet stabilizes, convergence, as shown in Figure 9.

Figure 9. Effects of limited degrees of updating on the state error, parameter error, and iterations to convergence for varying degrees.

6. GPR RITA

6.1 COMPONENT 1 – USING AN ANN TO PROVIDE INITIAL STRUCTURE

The first component of each RITA is a method of characterizing porous media that incorporates both quantitative and qualitative data developed by Rizzo and Dougherty [1994a, 1994b], called SCANN (Site Characterization using Artificial Neural Networks). It is a pattern completion method that has been successfully applied and compared to a number of geostatistical methods (*e.g.*, co-kriging, indicator kriging, and probability kriging) on a variety of site characterization problems. Applications include developing maps of discrete spatially-distributed fields (*e.g.*, log-hydraulic conductivity fields given estimates of hydraulic conductivity from pumping tests), and classifying soil lithology given soil sample descriptions from drillers' well logs [Rizzo and Dougherty 1994a].

SCANN uses a feed-forward counterpropagation training approach to determine a "best estimate" or map of a discrete spatially-distributed field, given some fixed amount of information. The method can be readily applied in 3-D mappings, and can be used as part of conditional simulation. The combination of simplicity and computational speed make SCANN ideally suited for a variety of large site characterization problems. Details of this method have been presented in Rizzo et al. [1994a, b, 1996] and Rizzo and Dougherty [2000], and are omitted here.

6.2 COMPONENT 2 – USING APPRIZE™-GPR FOR INVERSION OF GPR TRAVELTIMES

APPRIZE™ (automatic progressive parameter-reducing inverse zonation and estimation) is a statistically powerful computational method under development for accurate and efficient data inversion in large domains [Eppstein & Dougherty 2000]. At the heart of the inversion algorithm is the approximate extended Kalman filter (AEKF). The AEKF is a recursive, Bayesian, minimum-variance estimator that optimally estimates parameter values and covariance. Measurement updates are weighted based on the relative (spatially and temporally variable) covariances of the measurement noise, the system noise, and the current parameter estimates. The recursive nature of the filter enables the user to employ data and/or domain decomposition, using subsets of data measurements to condition subsets of the domain during various recursive passes through the filter [*e.g.* Eppstein & Dougherty 1998a, b]. In APPRIZE™, data conditioning is alternated with data-driven zonation, in which stochastic variables modeling spatially adjacent portions of the domain are clustered and merged into a fewer number of stochastic variables representing larger zones as warranted [Eppstein & Dougherty 1996]. Data-driven zonation can drastically improve the computational efficiency, accuracy, and stability of the AEKF [Eppstein & Dougherty 1996, 1999a, b].

The set of (pre-injection) crosswell GPR measurements as used in Eppstein & Dougherty [1998b]. Sparse, nonuniform measurements were taken between 5 vertical wells located approximately at the corners and one in the center of a 7.01 m × 7.32 m rectangular test site

extending 18.01 m deep into unsaturated soils in South Royalton, Vermont. Further details of the site and the measurements used are described in Eppstein & Dougherty [1998b].

In this example problem, the forward model was discretized into $23 \times 24 \times 55$ nodes, and transformed velocities (see Section 6.2.1) were assumed to be negatively exponentially correlated with a correlation length of 5 nodes.

6.2.1 PARAMETER TRANSFORMATIONS

In previous work using APPRIZE™ for GPR inversion, distributed slowness (inverse velocity) variables have been modeled as log-normally distributed stochastic variables, through the familiar use of log transforms. In consequence, some of the low velocity estimates were outside the feasible range [Eppstein & Dougherty 1998b]. Here, we model slowness as double-bounded β -distributed stochastic variables by employing novel pseudo- β transforms [as in Eppstein *et al.* 1999a, b]. In this manner we implicitly constrain velocity estimates to lie between the designated feasible bounds of 0.056 m/ns (*i.e.* a slowness of 17.97 ns/m or a dielectric of 30, as in wet quartz sand) and 0.176 m/ns (*i.e.* a slowness of 5.68 ns/m or a dielectric of 3, as in dry quartz sand). The transformed vector y that is directly estimated in the approximate extended Kalman filter is as follows:

$$y = \ln \left[\frac{17.97 - 5.68}{\frac{1}{v} - 5.68} - 1 \right]$$

where v is a vector of uncertain velocities (m/ns). Note that the parameter covariance matrix thus refers to the transformed variables y . The inverse transform is used to convert estimates y back to velocities v for output.

6.2.2 PRECLUSTERING

In a previously reported study we preclustered and merged (using random field union) the 30360 nodes into 7260 zones, most of which (except at the edges) were $2 \times 2 \times 1$ nodes in the x , y , and z dimensions, respectively [Eppstein & Dougherty 1998b].

In this example, we compare results using the $2 \times 2 \times 1$ preclustering strategy with a new preclustering strategy based on prior information using SCANN. SCANN is first applied using 1) the raw cone penetrometer testing (CPT) data shown in Appendix A and 2) to the CPT soil resistivity data obtained at the 5 wells to classify each of the 30360 nodes into one of a prespecified number (here, we used 6) of spatially discontinuous soil classification categories. Although SCANN is capable of estimating the horizontal correlation length of each layer, that information was not used in this example.

A partitional cluster analysis is applied to each horizontal layer of the output of SCANN, in order to identify spatially contiguous single-layer zones to use as an initial zonation structure. In the results reported here, we obtained 134 initial zones. The fully distributed stochastic parameters for transformed velocity are then merged according to this initial identified zonation structure. During subsequent estimation with APPRIZE™, the number and

geometry of the 3-D zonation structure is further determined as the values and covariance of the stochastic zones are estimated.

6.2.3 RESULTS

Using APPRIZE™ alone with the $2 \times 2 \times 1$ preclustering strategy, we achieved the 455 zone final velocity estimate shown in Figure 10a) in 4.5 hrs of cpu time on a 350 MHz Pentium II.

SCANN classification of the Vermont test site is accomplished in a only few seconds using a previously trained neural network (an untrained network requires approximately 0.25 hrs to train). Using SCANN and APPRIZE™ in the tandem fashion described above, we achieved the 78 zone final estimate shown in Figure 10b) in only 0.23 hrs.

In both cases the average final velocity estimate was 0.07 m/ns. The overall structure of velocity distribution is fairly consistent between the two estimates and root mean square traveltime errors were similarly reduced by approximately 25% in both cases, as compared to traveltime errors generated by the initial homogeneous 0.08 m/ns velocity model.

Figure 10. a) GPR velocity estimates using APPRIZE with $2 \times 2 \times 1$ preclustering required 4.5 hrs cpu time; b) GPR velocity estimating using APPRIZE with preclustering based on SCANN output of CPT data required 0.23 hrs cpu time.

By using appropriate parameter transformations, estimates of those parameters can be implicitly kept within feasible bounds without ad hoc corrections. This is proving to be a simple and elegant way to estimate double-bounded β -distributed parameters using an estimator (such as AEKF) that assumes normal distributions.

Using SCANN to provide a rapid initial estimate of soil structure based on CPT data (rather than using a much larger number of small rectangular zones) speeds up GPR velocity estimation by the APPRIZETM inverse method nearly 20 times for the example 3-D test site, without significant loss of accuracy in the model predictions.

This approach has the potential to be applied to a wide variety of site characterization problems. A similar tandem strategy is implemented in the RITA for 3-D groundwater inverse problems.

7. HYDRAULIC RITA

This RITA is designed to use hydraulic data to perform parameter estimation for distributed hydraulic conductivities in either 2-D (vertical slice of 3-D domain) or 3-D. MODFLOW96 [Harbaugh and McDonald, 1996a, 1996b] is used for the forward simulator. Synthetic tests were used to evaluate this RITA for both 2- and 3-D examples.

The synthetic tests used a hydraulic conductivity field (see Figure 11) specially constructed for controlled inversion experiments for the hydraulic RITA. A brick mesh is used with cell-centered nodes. The mesh consists of 18 rows, 34 columns, and 13 layers. Grid spacing is 50 ft. in the horizontal directions and a 5 ft. spacing in the vertical. The x, y, and z axes run

Figure 11. Hydrostratigraphic units for test problem. This domain is discretized into 13 layers for simulations.

from 0 to 1700 ft., 0 to 900 ft. and 0 to 65 ft. respectively. Although the domain of interest is hydraulically modeled as with 13 horizontal layers, it contains only five hydrostratigraphic units (water-related geologic layers), as shown in Figure 2. The aquifer material shown in brick color has a homogeneous hydraulic conductivity of 0.002835 ft/day. The top hydrostratigraphic unit (HSU) is 10 feet thick and has a homogeneous hydraulic conductivity of 28.35 ft/day. The second HSU is also 10 ft thick and contains a conductivity field that is classified into two categories. Most of this HSU has a hydraulic conductivity of 28.35 ft/day and there is a block heterogeneity with conductivity of 0.002835 ft/day). In particular, this HSU contains a 300 by 500 ft. window of material with a conductivity value that is four orders of magnitude larger than the conductivity value in the rest of the HSU). The third

HSU is 20 ft. thick with a hydraulic conductivity value of 28.35 ft/day. The fourth HSU is classified into two conductivity values and is similar (symmetric) to the second HSU, except flipped left-for-right. The bottom HSU is 15 ft thick and has a conductivity of 28.35 ft/day.

7.1 COMPONENT 1 – USING AN ANN TO PROVIDE INITIAL STRUCTURE

SCANN was not used to provide the initial structure for these simple 2-D and 3-D test cases using the synthetic domain described above. The intent of these experiments was to demonstrate that the inversion program functioned correctly, rather than to assess which assumptions produced the best inversion results. As a result, many test cases were performed (493 in total) in which the selection of measurements and the sequence for conditioning were arbitrary. Only two the 493 test cases (one for 2-D and one for 3-D) are shown below for the purposes of demonstrating convergence criteria.

7.2 COMPONENT 2 – USING APPRIZE™-HYDRO FOR ESTIMATING HYDRAULIC CONDUCTIVITY

Determining convergence criteria for conditioning is not a trivial matter, since (a) in a real system the parameter error will be unknown and (b) if we are conditioning on a variety of batches of heads at various time periods, the head error will bounce around significantly, depending on which head measurements we have used. The convergence test is comprised of two parts. The first part is the moving average of change in head errors over the last 10 conditioning steps, and the second part is the moving average of change in parameters (not parameter errors) over the last 10 conditioning steps (if 10 steps are not available, then the average is over whatever is available). Based on empirical studies using the synthetic test case, these moving averages are compared against arbitrary low values to determine convergence. Alternatively, the user may specify premature convergence at any time by clicking on a radio button displayed beneath the dynamically updated error graphs.

The results are dynamically updated and displayed in a figure. See Figures 12 and 13 for 2- and 3-D examples, both of which used upscaling of 4x4 cells, assumed heterogenous structure, and initial guesses of 0.002835 ft/day and 28.35 ft/day. The user can select whether or not to show a graphical representation of the current domain estimate (by setting the flag variable visualize). If visualize is set to true (non-zero), then a 2-D or 3-D image of the domain is displayed in the left-hand portion of the window, depending on the nature of the domain. Note the limits on the color map are always set to the lower and upper bounds on the beta-distribution of conductivity, and that the axes are labeled with field coordinates (not node coordinates). Also, note that the measurement locations of the measurements used in the most recent estimation step are displayed.

Whether or not a graphical representation of the domain is visualized, the system computes and displays dynamically updated error plots. Either 3 or 4 error plots are shown. The first (leftmost) is the root mean square head error over all batches of head measurements. The second is the change in head error from each step to the next. The third is the maximum absolute change in any parameter from one step to the next, relative to the size of the feasible range. The fourth metric is optional and is only calculated and displayed for

synthetic sites, for which the true parameters are known. This is the root mean square parameter error at all nodes in the domain.

Figure 12. Display for 2-D domain process.

Figure 13. Display for 3-D domain process.

8. TECHNICAL AND ECONOMIC FEASIBILITY

8.1 TECHNICAL FEASIBILITY

The methods subjected to the RR&D in this project have generally been proven to be technically feasibility. Downscaling was not successful and has been set aside for future reconsideration.

8.2 ECONOMIC FEASIBILITY

The economic feasibility for environmental applications of this project has shifted significantly since its initiation in 1997. Environmental software as a product, rather than as a service, is not the economic foundation of a viable small business for several reasons.

- First, costs to provide telephone or on-line support are high relative to the number of copies that might be sold.
- Second, the economics of the environmental business have changed over the past few years in the United States. Characterization and construction are winding down, while operations, maintenance, and monitoring are winding up.
- Third, site restoration program managers (whether on the regulatory or responsible-party sides) are much more sensitive to the use of all the data that they must collect. The typical "modeling" effort leaves most lukewarm because of the disconnect between the data and the model. These same program managers are demanding that decision support be incorporated with information systems (most often with a geographical information system or GIS).
- Fourth, the environmental business has become quite concentrated in the hands of large businesses like URS (which now includes the Dames and Moore Group, Woodward-Clyde Consultants, and others), reducing the number of potential customers.
- Finally, some Federal actions are reshaping the marketplace. For example, the U.S. Geological Survey has undertaken a new version of MODFLOW that includes parameter estimation. Because MODFLOW is the *de facto* standard for groundwater modeling software, the built-in parameter estimation software will become the *de facto* standard for groundwater modelers, regardless of other options. Another example is the DoD Groundwater Modeling System (GMS), a Federally-funded RR&D project, which has depressed the market for software not in the public domain.

These challenges lead to reconsideration of the commercialization. One alternative is to sell services, which happen to include the software developed under this project. A second alternative is to sell the methods to a larger environmental contractors, where it would be developed (most likely) for internal use. Third, it is possible to apply the methods in different domains of scientific and engineering interest. Finally, it is possible to extract methods from the characterization milieu, and combine them in new ways to address and solve problems that exist in the current marketplace.

9. REFERENCES

- Eppstein, M.J., & Dougherty, D.E., 1996. Simultaneous estimation of transmissivity values and zonation. *Water Resources Research*, 32(11):3321-3336.
- Eppstein, M.J., & Dougherty, D.E., 1998a. Optimal three-dimensional travelttime tomography. *Geophysics*, 61(4):1053-1061.
- Eppstein, M.J., & Dougherty, D.E., 1998b. Efficient three-dimensional data inversion: Soil characterization and moisture monitoring from cross-well ground-penetrating radar at a Vermont test site. *Water Resources Research*, 34(8):1889-1900.
- Eppstein, M.J., & Dougherty, D.E., 2000. Three-dimensional stochastic tomography with upscaling. U.S. Patent No. 6,067,340.
- Eppstein, M.J., Dougherty, D.E., Hawrysz, D.J., & Sevic-Muraca, E.M., 1999a. Three-dimensional optical tomography. Optical Tomography and Spectroscopy of Tissue III, *SPIE Proceedings*, 3597:97-105.
- Eppstein, M.J., Dougherty, D.E., Troy, T.L. & Sevic-Muraca, E.M., 1999b. Biomedical optical tomography using dynamic parameterization and Bayesian conditioning on photon migration measurements. *Applied Optics*, 38(10):2138-2150.
- Gelb, A., 1979. *Applied Optimal Estimation*, MIT Press,.
- Harbaugh, A.W., and McDonald, M.G., 1996a. "User's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model", U.S. Geological Survey Open-File Report 96-485, Reston, VA, 56 p.
- Harbaugh, A.W. and McDonald, M.G., 1996b. "Programmer's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model," U.S. Geological Survey Open-File Report 96-486, Reston, VA, 220 p.
- Menke, W., 1984. *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press.
- Rizzo, D.M., & Dougherty, D.E., 1994a. Characterization of aquifer properties using artificial neural networks, Neural kriging. *Water Resources Research*, 30(2):483-497.
- Rizzo, D.M. & Dougherty, D.E., 1994b. Application of artificial neural networks for site characterization using 'hard' and 'soft' information, *X Intl. Conf. Computational Methods in Water Resources*, A. Peters et al. (eds.), Kluwer Academic, Vol. 1:793-799.
- Rizzo, D.M., Lillys, T.P. & Dougherty, D.E., 1996. Comparison of site characterization methods that use mixed data. *Proc. Uncertainty in the Geologic Environment: From Theory to Practice*, C.D. Shackelford, et al. (eds.), ASCE:167-179.
- Rizzo, D.M., & Dougherty, D.E., 2000. Artificial neural networks in subsurface characterization, Chapter 6 in *Artificial Neural Networks in Hydrology*, R. S. Govindaraju and A. R. Rao (eds.), Kluwer Academic.

10. GLOSSARY

<i>Area element</i>	Pixel
<i>Class</i>	Voxels having essentially identical parameter values that are grouped together
<i>Domain</i>	Volume of the ROI or its discretization; limiting case of subdomain
<i>Downscaling</i>	Converse of <i>upsampling</i> . It is the process of converting information that is resolved on one scale (let's call the original scale of resolution δ) to a smaller scale (that is, information at some greater resolution).
<i>Linesegment</i>	Intersection of a usual line segment (or arc) with the region of interest; also the discrete form, which is a set of adjacent "points" in the discretization
<i>Line</i>	Intersection of line with region of interest; limiting case of linesegment
<i>Material property</i>	Coefficient used relate states to fluxes (e.g., hydraulic conductivity)
<i>Parameter</i>	Either a state, material property, boundary condition, etc. that is of interest and that can be estimated using local data
<i>Pixel</i>	Area element
<i>Point</i>	Point (Usual definition)
<i>Point element</i>	Point
<i>Pointset</i>	A set of points
<i>Region of Interest</i>	Region occupying space and containing material(s) of interest
<i>ROI</i>	Region of interest
<i>Segment</i>	A portion of the domain used for decomposition of calculations, computer storage, etc.
<i>State</i>	Observable quantity, used to characterize the situation in the region of interest
<i>Subdomain</i>	A volume contained within the region of interest, can be a union of disjoint voxels (i.e., not necessarily adjacent voxels)

<i>Upscaling</i>	Converse of <i>downscaling</i> . It is the process of converting information that is resolved on one scale (let's call the original scale of resolution δ) to a larger scale (that is, information at some lesser resolution)
<i>Volume element</i>	Voxel, grid cell, finite element
<i>Voxel</i>	Volume element
<i>Zone</i>	Similar to class except that, in addition to considering parameter values, all members of a zone must have essentially identical value and must be geometrically adjacent

APPENDIX A -- CPT SOIL CLASSIFICATION

The depth from the ground surface, tip resistance, sleeve friction, and pore water pressure are electronically recorded. The function f_bincpt calculates the normalized tip resistance (above and below the water table) as $Q_{tnorm} = (q_t - \sigma) / \sigma'$, where q_t is the sleeve stress, σ is the soil overburden and σ' is the effective overburden. (Note: The effective overburden is a function of the depth, the depth to the groundwater table, and the overburden's density, porosity, and degree of saturation, and hence these data must be specified or introduced through assumption to software.) The normalized friction ratio is calculated as $F_{snorm} = F_s / (q_t - \sigma) \times 100$, where F_s is the sleeve friction.

The normalized tip resistance (or cone bearing) is then plotted against the normalized friction ratio on the Robertson soil classification chart [Robertson, 1990] of Figure A.1.

Figure A.1. Soil classification chart modified from the Robinson [1990] chart.

Douglas and Olsen [1981] used cones of various dimensions and pressed them into soils of unknown composition, structure, and moisture content in a controlled laboratory environment. The cone resistance or cone bearing was plotted against the friction ratio. The cone resistance and friction ratio for each type of soil under various conditions were within discrete regions of the domain. These discrete regions were compiled to form a soil classification chart, similar to the chart shown in Figure A.1.

Robertson [1990] expanded upon the work by accounting for the increased resistance and friction due to overburden and moisture content by normalizing the cone resistance and friction ratio.

The Robertson soil classification system embodied an algorithm that uses cone tip stress, sleeve friction, depth from surface, overburden, and effective overburden to provide a soil classification number ranging from 1 to 9, corresponding to the regions in Figure A.1.

(rawcpt.dat)
Electric Cone Penetrometer Data

Test Id:	L1008DT	Date: 03/08/98					
Site:	Site Name Here	Cone Id: 108.042					
Location:	Location Here	GWT (ft): 102					
Project:	8828	Soil Density (pcf): 115.					
Depth	Sleeve Stress	Tip Stress	Tip Stress	Ratio	Pore Press.	Resistivity	Excitation
		UNC	COR	COR			
(ft)	(psi)	(psi)	(psi)	(%)	(psi)	(ohm-m)	(Vdc)
.00	.00	.00	.00	.00	.00	.00	3.979
.15	.00	213.93	213.93	.00	-.01	684.77	3.979
.22	.00	255.98	255.98	.00	-.02	741.34	3.980
.31	.59	300.45	300.45	.20	-.02	804.57	3.980
.39	.61	301.68	301.68	.20	-.01	846.92	3.979
.47	.71	326.90	326.90	.22	-.01	922.61	3.979
.54	.91	360.56	360.55	.25	-.01	986.55	3.979
.62	1.12	424.26	424.25	.26	-.01	1201.61	3.979
.70	1.50	549.25	549.25	.27	-.01	1720.33	3.979
.78	1.87	688.67	688.65	.27	-.06	2753.05	3.979
.86	2.19	802.81	802.81	.27	-.02	3865.51	3.980
.94	2.56	919.42	919.42	.28	-.01	4580.28	3.979
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
68.49	57.52	2311.28	2311.53	2.49	1.22	2272.21	3.979
68.55	57.86	2276.42	2276.65	2.54	1.12	2310.62	3.979
68.62	57.63	2252.30	2252.53	2.56	1.12	1979.36	3.979
68.69	56.94	2217.44	2217.67	2.57	1.12	1926.54	3.979
69.14	27.52	2814.88	2815.15	.98	1.31	.00	3.979
69.20	27.61	2849.74	2850.00	.97	1.31	.00	3.979
69.26	31.89	2864.05	2864.30	1.11	1.26	.00	3.979
69.32	35.55	2878.58	2878.85	1.23	1.31	.00	3.979
69.50	44.04	2930.15	2930.58	1.50	2.11	.00	3.979
69.56	47.55	2861.64	2862.00	1.66	1.73	.00	3.979
69.64	48.74	2767.90	2768.24	1.76	1.68	.00	3.979
69.70	.00	2640.50	2640.80	.00	1.45	.00	3.979
69.78	.00	2473.44	2473.73	.00	1.40	.00	3.979
69.85	.00	2370.08	2370.36	.00	1.35	.00	3.979

Figure A.2. Example Cone Penetrometer Test Raw Data File

(D13D1.dat)

10.5	SILTY CLAY
28.2	SANDY GRAVEL
30.1	SILTY GRAVEL
55.8	SAND
77.0	SILTY SAND
78.0	CLAYEY GRAVEL

· ·

· ·

∨

∨

Figure A.3. Example Driller's Well Log Raw Data File