

INFORMED ELECTRICAL RESISTIVITY IMAGING FOR MONITORING
INFILTRATION DYNAMICS IN THE NEAR SURFACE

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Vanessa Renae Mitchell
July 2010

ABSTRACT

Determining the spatial distribution of subsurface properties is critical to developing efficient groundwater management strategies. Though hydrologic methods can accurately measure these properties, they often have sparse spatial coverage and cannot capture the spatial variation. Integrating geophysical and standard methods can provide more comprehensive measures of spatial and temporal variation in hydrologic properties. Electrical resistivity imaging (ERI) provides continuous maps of the subsurface electrical conductivity, which is related to water content in the vadose zone, making it particularly useful for groundwater applications. However, differences in measurement scales and the challenge of obtaining accurate images of electrical conductivity can limit the ability to integrate ERI with existing hydrologic data. To address these issues I present a framework for informed imaging, the incorporation of all available information into the acquisition, inversion and interpretation of electrical resistivity data. This framework includes methods for informed experimental design, time-lapse inversion, and integration of different data types into a consistent hydrogeologic model.

I present an application of informed imaging to monitoring flow beneath an artificial recharge pond, which is part of an aquifer storage and recovery project for managing coastal groundwater resources. As a first step in informed imaging, I create a hydrostratigraphic model, which incorporates hydrologic, geologic, and geophysical data into a systematic framework. I use this model with the proposed experimental design method to select acquisition arrays for ERI experiments. Monitoring data acquired at the pond are inverted using an implementation of the extended Kalman filter, which incorporates all previously acquired information into the inversion of data at each time-step. This is the first application of this method to surface-ERI data collected in the field. Furthermore, I present an analysis of the impact of measurement support on the estimated correlation structure of properties in anisotropic systems, which must be accounted for when integrating processed geophysical data into hydrologic models. The successful application of these methodologies to monitoring

flow at an artificial recharge pond leads me to conclude that using an informed imaging approach can enhance the use of geophysical methods to developing groundwater management strategies.

ACKNOWLEDGMENTS

I would like to begin by thanking my advisor Rosemary Knight. Rosemary has been a constant source of motivation throughout this research. She has given me the freedom to explore my own research paths, but wasn't afraid to pull me off tangents that I was too stubborn to abandon. Her excitement in this research was unflinching, even when my own got bogged down in the details, and she taught me to hold onto the thrill of doing science through all the ups and downs. Rosemary has been a wonderful mentor; she has helped me to grow as a scientist and person; I am thankful for the opportunity to have worked with her.

Adam Pidlisecky provided the forward modeling code for electrical resistivity measurements without which this research would not have been completed. He has been an excellent collaborator over the years, providing ideas and suggestions for improving and extending this research, and has been my electrical-resistivity-coding guru. Adam has taught me much about conducting resistivity imaging experiments in the field and the many uses of Cat-5 cable.

Many thanks to the Pajaro Valley Water Management Agency for graciously allowing us to do all the science we could dream at their recharge pond. Thanks especially to Brian Lockwood and Jon Lear for providing access to the hydrologic and geologic data on the Harkins Slough Recharge Pond.

I would like to acknowledge the efforts of Sandra Jenni, Bob Will, Angeline Kneppers, Jean Pierre Delhomme, and the many collaborators at Schlumberger Water Services for their help in developing the structural and flow model of the recharge pond and for their vast knowledge of operating and manipulating the PETREL and ECLIPSE software.

Jerry Harris, Greg Beroza, and Peter Kitanidis, the other members my reading committee, have offered excellent advice throughout this project. I thank them all for the perspective and insight they brought to this research.

Thanks to Nigel Crook for all his help in acquiring the field data. Nigel was a fellow combatant against miles of tangled Cat-5 cable that were the bane of this field research. I thank him for his patience and tireless unknotting.

I would like to extend my gratitude to all the members of the Environmental Geophysics Group. In particular my officemates: Kristina Keating, Katherine Dlubac, and especially Elliot Grunewald. Elliot has shared my entire Stanford experience and has patiently endured frustrated rants and too many conversations about shoes.

Many thanks are owed to the great friends who made this experience worthwhile and whose support got me through the completion of this thesis. In particular, Matt Coble and Emily Pope were the best of roommates and better happy-hour-instigators. I have also had the good fortune at Stanford to work and play with fantastic people, all of whom I hope will remain close friends: Justin Brown, Annemarie Baltay, Andrea and Greg Earhart, Katie Maier-Coble, Kristin Mayer, Phil Sandlin, Kristina Keating, Julie Fosdick, Elliot Grunewald, and the Maddux house.

I owe many thanks to my family for supporting me in 20+ years of education. They have always believed in me and shown (or at least feigned) interest in my studies and research. I would not have made it to this point without their love and encouragement. Special thanks to my father for his comments and corrections to Chapter 4.

Finally, I would like to thank Fil Nenna: lifting buddy, marathon-training team-member, travel companion, and best friend. Fil has challenged me to achieve goals I would have never considered on my own and stood by me throughout, though I still maintain running for more than 10 miles at a go is overkill. I am a better person and more capable scientist for having known him. My time at Stanford has been blessed by his love and companionship.

Funding for my studies came from the Stanford Graduate Fellowship and a Department of Geophysics graduate student fellowship. Support for fieldwork came from a grant to Rosemary Knight from Schlumberger Water Services.

Table of Contents

ABSTRACT	IV
ACKNOWLEDGMENTS	VI
TABLE OF CONTENTS	VIII
LIST OF TABLES	X
LIST OF FIGURES	XI
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. HYDROSTRATIGRAPHIC MODEL OF THE HARKINS SLOUGH RECHARGE POND FROM INTEGRATION OF HYDROLOGIC, GEOLOGIC, AND GEOPHYSICAL DATA	5
2.1 Introduction	5
2.2 Data Sets	9
2.2.1 Well Data	9
2.2.2 Core Data	12
2.2.3 CPT Logs	12
2.2.4 Infiltration Data	18
2.2.5 Geophysical Surveys	20
2.3 Structural Model from Lithologic Data	29
2.3.1 Lithology Transforms	29
2.3.2 Grid Definition	33
2.4 ECLIPSE Modeling	38
2.5 Structural Model including Geophysical Data	42
2.5.1 Structural Observations from Geophysical Data	42
2.5.2 Grid Zonation	45
2.6 Conclusions	47
CHAPTER 3. INFORMED EXPERIMENTAL DESIGN FOR ELECTRICAL RESISTIVITY IMAGING	49
3.1 Introduction	49
3.1.1 Optimized Design	51
3.1.2 Description of Field Problem	52
3.2 Methodology	56
3.3 Measures of Data Sensitivity	60
3.3.1 Singular Value Decomposition of \mathbf{J}	60
3.3.2 Region of Investigation	61
3.4 Informed Experimental Design: Application to a Recharge Pond	63
3.5 Assessment of II Arrays	66
3.5.1 Example 1	68
3.5.2 Example 2	73
3.6 Conclusions	79
CHAPTER 4. ASSESSMENT OF AN EXTENDED KALMAN FILTER APPROACH TO INVERSION OF TIME-LAPSE ELECTRICAL RESISTIVITY IMAGING DATA FOR MONITORING INFILTRATION	81
4.1 Introduction	81
4.1.1 Monitoring the Vadose Zone with Electrical Resistivity	81
4.1.2 Inversion of Time-Lapse ERI Data	82

4.1.3 Description of Field Problem	83
4.2. Inversion Methods	85
4.2.1 Regularized, Gauss-Newton Update	85
4.2.2 Extended Kalman Filter	86
4.3 Synthetic Example	91
4.3.1 Description of Synthetic Data Set	91
4.3.2 Regularized Gauss-Newton Inversion Results	95
4.3.3 EKF Inversion Results	101
4.4 Field Experiment	105
4.4.1 Acquisition of Field Data	105
4.4.2 Regularized Gauss-Newton Inversion Results	108
4.4.3 EKF Results	114
4.5 Summary and Conclusions	119
CHAPTER 5. THE IMPACT OF GEOPHYSICAL MEASUREMENT SUPPORT IN QUANTIFYING THE CORRELATION STRUCTURE OF ANISOTROPIC SYSTEMS	121
5.1 Introduction	121
5.2 Geostatistical Analysis.....	125
5.3 Numerical Models.....	127
5.4 Results and Discussion	132
5.5 Conclusions	141
CHAPTER 6. CONCLUSIONS	143
APPENDIX A.	149
A1. Data	150
4.1.1 PETREL Data Formats	150
4.1.2 Pre-processing Data	152
4.1.3 Importing Data into PETREL	158
A2. Generating Structural Models in PETREL	162
4.2.1 Defining Model Boundaries	162
4.2.2 Making Grids.....	163
4.2.3 Populating a Grid with Properties.....	165
4.2.4 Incorporating Time-Dependent Data	168
A3. Flow Simulations in ECLIPSE	170
4.3.1 Exporting Data to ECLIPSE using the PETREL Interface	170
4.3.2 Creating an ECLIPSE Driver File Outside PETREL	171
APPENDIX B.....	185
REFERENCES	187

List of Tables

Table 2.1 Texture, Modifier, and Qualifier descriptors and values applied to the percent coarse sediment classification.....	10
Table 2.2 Texture description conversions to soil behavior type.....	30
Table 2.3 Porosity of wet-packed sand sample as a function of median diameter and sorting, after Beard and Weyl [1973].	32
Table 2.4 Mean grain diameter, porosity and permeability estimates for each soil behavior type classification	32
Table 5.1 Defined correlation length and initial estimates of the correlation length for the seven synthetic parameter fields analyzed.....	127

List of Figures

Figure 2.1. Conceptual model of operations at the Harkins Slough Recharge Pond [after Haines et al., 2009]. Depths are approximate, and the image is not to scale.	8
Figure 2.2. Location of recovery wells and monitoring wells surrounding the HSRP. The dashed line indicates the boundary of the recharge pond; the solid black line indicates the edges of the model space. A deep monitoring well, labeled in gray as HSPDMW1, is known to exist northwest of mw7, though the exact location is unknown.	11
Figure 2.3. Core locations collected along the base of the HSRP in Summer 2007. Cone penetrometer test logs are indicated by black dots and are labeled cpt 1 to cpt 3; 9-m cores are indicated by gray dots and are labeled bh 1 to bh 5.	14
Figure 2.4. Median grain size distribution at 9-m cores bh 2 through bh 4.	15
Figure 2.5. Mode grain size distribution for 9-m cores bh 2 through bh 4.	16
Figure 2.6. CPT logs acquired at location cpt 2 [after Haines et al., 2009]. The first panel shows the tip pressure, the second panel records the induced pore pressure, and the third panel shows the interpreted soil behavior type.	17
Figure 2.7. Average infiltration rate recorded for the HSRP for diversion season 2007-2008. Time is given in days after 10 January 2008.	19
Figure 2.8. Transects of seismic, electrical resistivity, and ground-penetrating radar surveys along the base of the HSRP in Summer 2007.	23
Figure 2.9. Processed seismic section after Haines et al. [2009]. The large reflector at ~0.2 seconds is the suspected shallow clay layer. The deep reflector at ~0.4 s is the clay layer identified in the drillers' logs from the recovery wells.	24
Figure 2.10. Inverted model of the conductivity structure below the HSRP. Electrode locations have been depth-corrected to account for topography, and the white line indicates the base of the pond. Depths are reported in m msl. The region of higher conductivity is interpreted to be a combined effect of the water table and deep clay layer from the conceptual model.	25
Figure 2.11. Processed ground-penetrating radar depth section collected along the same transect as the seismic and ERT sections. High magnetite content in the sands creates the large reflectors identified as dune foresets in the image.	26
Figure 2.12. Schematic of electrical resistivity monitoring surveys. Black circles mark the locations of probes. The locations of 2D surface lines are indicated for years 2008-2009 and 2009-2010. Line 4 in 2008-2009 is an aggregation of lines 1 and 2; line 3 in 2009-2010 is also an aggregation of lines 1 and 2 from that year.	27
Figure 2.13. Picture of a second-generation 1D electrical resistivity sampling probe. Probes are inserted to a depth of 2 m, with ring electrodes left above the surface to sample bulk fluid conductivity. All probes are fitted with multi-level sampling ports, identified as the white tubes in the image.	28
Figure 2.14. PETREL grid definition with x-y cell dimensions.	36

Figure 2.15. PETREL hydrostratigraphic model using lithologic data. Solid blue lines indicate the lateral extents of the model space. The topography is overlaid over the top boundary of the model. The green cells on the top boundary indicate the constant gas aquifer; blue cells at the surface indicate connection of the constant flux water aquifer. The cells within the model space show upscaled porosity data. 37

Figure 2.16. Relative Permeability - Water Saturation curve applied in the ECLIPSE flow simulation. 40

Figure 2.17. Comparison of observed versus simulated water table elevations at monitoring wells 2, 3, 4, 6, and 8 from Nov 2007 to Oct 2008..... 41

Figure 2.18. Time-lapse plot of the 1D electrical resistivity profile at probe 2. The time is given in days after 01:00:00, 24 December 2008. Depth is given in meters below ground surface. The log conductivity profile shows repeated wetting and drying below the fine-sediment layer, which remains saturated throughout the imaging experiment. 44

Figure 2.19. Soil behavior type distribution through the Y=3250 m slice of the PETREL structural model. The distribution includes the fine-sediment layer at the surface and the suspected shallow clay layer at ~30 m depth. Cool colors are indicative of sand-rich SBTs; warm colors represent clay-rich SBTs. 46

Figure 3.1. Conceptual model of the Harkins Slough Recharge Pond [after Haines et al., 2009]. Water percolates through the thick sand package and is extracted from wells screened within the sand and gravel layer. A massive clay at ~40-50 m depth is the basal aquitard for the unconfined aquifer. 55

Figure 3.2. Pseudo-code describing the proposed methodology for informed experimental design..... 59

Figure 3.3. *A priori* model of conductivity beneath the recharge pond. The region of interest (ROI) is indicated by the dashed lines. 65

Figure 3.4. Schematic representation of Wenner and dipole-dipole arrays. A and B indicate source electrodes; M and N are potential electrodes. The spacing, a , is the offset of the potential electrodes; n is an integer multiplier of a 67

Figure 3.5. Conductivity model used to create synthetic data for Example 1. Conductivity values are given in S/m..... 70

Figure 3.6. Log singular values of $(J^T J)$ as a function of rank for the II, Wenner, and dipole-dipole arrays for Example 1. The greater area under the spectrum of the II arrays indicates greater row independence, and, hence, more information content for these arrays. 71

Figure 3.7. Estimated conductivity structures from Example 1 for the II arrays, Wenner array and dipole-dipole array from a homogeneous reference model. Conductivity values are in S/m. The contoured R-values from the DOI measure are overlaid on the estimated structures. 72

Figure 3.8. True conductivity structure used to create the synthetic data for Example 2. Conductivity values are reported in S/m. 75

Figure 3.9. Log singular values of the $(J^T J)$ decompositions as a function of rank for the II, Wenner, and dipole-dipole arrays in Example 2..... 76

Figure 3.10. Estimated conductivity structures from Example 2 for the II, Wenner, and dipole-dipole data. Estimated conductivity values are given in S/m. Contoured R-values from the DOI calculation are overlaid on the conductivity models.....	77
Figure 3.11. Estimated conductivity structures with overlaid R-values for the field imaging experiments using the II, Wenner and dipole-dipole data contaminated with 3% Gaussian noise. Estimated conductivity values are reported in S/m.....	78
Figure 4.1. Time-lapse conductivity models created based on the numerical solution to Richard's equation for effective saturation during an infiltration experiment. Conversion to conductivity is achieved through an Archie's Law relationship. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.	94
Figure 4.2. Time-lapse conductivity models inverted using regularized GN and a reference conductivity model equal to the estimated conductivity structure at the previous time-step. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.....	97
Figure 4.3. Time-lapse conductivity models inverted using regularized GN with 1.5% Gaussian noise added to the simulated data. The reference conductivity model is equal to the estimated conductivity structure at the previous time-step, but is reset to the initial reference model if the inversion fails to converge. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.	100
Figure 4.4. Time-lapse conductivity models inverted using EKF with 1.5% Gaussian noise added to the simulated data. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.	103
Figure 4.5. Time-lapse conductivity models inverted using EKF with 1.5% Gaussian noise added to the simulated data and model updates occurring only when the relative difference in observed data between time-steps is greater than the data variance. A: 2 hours; B: 4 hours; C: 6 hours; D: 10 hours; E: 16 hours; F: 20 hours; G: 26 hours; H: 30 hours.	104
Figure 4.6. Schematic of field layout for acquisition of time-lapse ERI data.	107
Figure 4.7. Six time-lapse relative difference conductivity structures with respect to the initial conductivity model inverted using GN-algorithm. A: 13.6 days (rain event); B: 22.8 days; C: 48.4 days (diversion begins); D: 50.6 days; E: 75.2 days; F: 112.6 days.	111
Figure 4.8. Bulk fluid conductivity in S/m as a function of time after monitoring begins. Times corresponding to the time-steps of inverted 2D conductivity models are labeled.	112
Figure 4.9. Time-lapse relative difference conductivity profile with respect to the initial conductivity profile at probe 2. Times corresponding to the time-steps of inverted 2D conductivity models are labeled.	113
Figure 4.10. Six time-lapse relative difference conductivity structures with respect to the initial conductivity model inverted using an EKF algorithm. A: 13.6 days (rain event); B: 22.8 days; C: 48.4 days (diversion begins); D: 50.6 days; E: 75.2 days; F: 112.6 days.	117

Figure 4.11. Model error variances associated with elements for the estimated conductivity structure at 112.6 days, shown in (F) of Figure 4.11..... 118

Figure 5.1. Examples of initial parameter fields simulated for three model conditions. Model 1 is an isotropic model, while Models 4 and 6 are anisotropic models showing increased anisotropy..... 130

Figure 5.2. Schematic of the numerical averaging experiment where n is the number of averaged cells. Only non-equidimensional increases in the x-direction are shown. White blocks indicate cell values not included in the variogram calculation for the estimated correlation length for a given averaging length. .. 131

Figure 5.3. Effect of block averaging of measurement support for isotropic parameter fields on the apparent correlation length from 1D x- and y-direction variograms. 133

Figure 5.4. Apparent correlation length from variograms calculated in the x- and y-directions for x-directional averaging of measurement support in isotropic parameter fields. 134

Figure 5.5. Apparent correlation length from variograms calculated in the x- and y-directions for block averaging of measurement support in anisotropic parameter fields. 139

Figure 5.6. Estimated correlation lengths for directional averaging of measurement support in anisotropic parameter fields. Results are plotted as a function of the normalized measurement support. Circles denote variogram calculation in the x-direction; triangles denote variogram calculation the y-direction. The support dimension in panes A and B increase in the x-direction, while the support dimension in panes C and D increase in the y-direction. 140

CHAPTER 1. INTRODUCTION

Developing efficient and cost-effective strategies for managing groundwater resources requires accurate models of subsurface hydrogeologic structure. In turn, these models need information about the hydrologic properties that govern the transport and long-term fate of water or contaminants as they percolate from the surface, through the vadose zone, to groundwater aquifers. In-situ measurements of hydrologic properties from boreholes or probes provide accurate values of measured properties but are sparsely sampled, limiting their ability to determine the variability of those properties in space. Geophysical measurements are often spatially exhaustive and can be acquired at relatively low cost, making them very useful for supplementing conventional measurements. Electrical resistivity imaging (ERI) is a geophysical method that is particularly useful to studies of hydrologic processes in the near surface because of the high correlation between electrical conductivity, the inverse of resistivity, and water saturation in the vadose zone. Repeatedly imaging the same subsurface region provides maps of changes in the electrical conductivity structure with time. These maps of variation in electrical conductivity can be used to track the infiltration of water from the surface [Daily et al., 1992; Michot et al., 2003; Robinson et al., 2008], or in conjunction with other methods to estimate hydrologic properties [Binley et al., 2002; Day-Lewis et al., 2007; Slater, 2007].

Although ERI has been used successfully to monitor hydrologic processes in the near surface, the technique still needs refinement to produce more accurate and reliable images of changing electrical conductivity structure. The focus of this thesis is to develop a framework for “informed imaging:” the incorporation of all existing information about a site into the acquisition, inversion, and interpretation of electrical resistivity data. This framework includes a methodology for informed experimental design, in which an *a priori* model of the subsurface structure and known physical and temporal constraints on the acquisition are used to select arrays that provide the greatest amount of information about an imaged region. The informed imaging framework also incorporates problem-specific inversion algorithms. An application of

the extended Kalman filter (EKF) is proposed for the inversion of time-lapse ERI data. The EKF is suggested for monitoring applications because it incorporates all prior information about both the evolution of the physical state and measurement process acquired from parameter estimation at previous time-steps into the current estimation. The concept of informed imaging requires validation of the theory as well as practical application. Consequently, this dissertation focuses on the application of the informed imaging framework to both synthetic and field data.

This thesis is separated into six chapters and two appendices. Chapters 3-5 are in preparation for submission to peer reviewed journals. For continuity the abstracts of individual chapters are removed.

Chapter 2 includes both a description of the field site at which the concept of informed imaging is tested and a definition of the *a priori* hydrostratigraphic model. Model development requires incorporating diverse hydrologic, geologic, and geophysical data sets into a single structural and flow model of infiltration at a recharge pond. The *a priori* model represents the current understanding of controls on hydrologic processes at the recharge pond and identifies regions where interest for monitoring exists.

In chapter 3, a method for informed experimental design is proposed and is applied using the hydrostratigraphic model developed in the previous chapter. Informed experimental design incorporates prior information about the site as well as spatial and temporal constraints on the acquisition process to select arrays, combinations of source and potential electrodes, which have high data sensitivity in the region to be imaged. Selection of arrays in this fashion is a computationally efficient, robust way to increase the information content of ERI data over what can be attained using standard 4-electrode arrays. The performance of arrays selected through informed experimental design is validated through comparison to the performance of standard arrays on two synthetic, example problems. Chapter 3 has been submitted for publication in *Near Surface Geophysics* (Mitchell, Pidlisecky, and Knight, 2010).

Chapter 4 details two methods for inverting time-lapse electrical resistivity data acquired in the field. The first method uses a Gauss-Newton update algorithm with a regularization term that incorporates information about the electrical conductivity model from the previous time-step. While this method produces good results, it has challenges in applications to monitoring experiments: introduction of noise produces artifacts in the estimated structure and can cause the inversion to fail. When this occurs a “restart” of the algorithm, which erases all previous information, is required. Furthermore, this approach also requires that imaging experiments be treated as ‘snapshots’ of the subsurface conductivity structure, limiting the number of measurements that can be acquired to avoid temporal smearing. Consequently, we propose a second method using the EKF, which uses information about the model and measurement structures attained from all previous time-steps to estimate model parameters at the current time-step. This method has great potential for continuous monitoring because it models both changes in the physical state and the measurement process. Hence, the same arrays need not be used in acquisition of sequential images, removing the need to limit the number of acquired measurements. This chapter is in preparation for submission to *Water Resources Research* (Mitchell, Pidlisecky, and Knight, 2010).

Chapter 5 examines one of the challenges when incorporating geophysical data into hydrologic models that results from the impact of measurement support, the area or volume of the subsurface sample in a single measurement, on the estimated correlation length of the measured property. Geostatistical analysis is used to examine how anisotropy in a 2D synthetic parameter field and how directionally dependent increases in measurement support affect the correlation length estimated from directional variograms relative to the correlation lengths specified in simulating the parameter field. Chapter 5 is in preparation for submission to *Vadose Zone Journal* (Mitchell, Gulati, and Knight, 2010).

Chapter 6 summarizes the contributions of this thesis and outlines directions of future work that build on this research.

Appendix A contains notes on integrating various types of data into a single near-surface hydrogeologic structural and flow model using the PETREL and ECLIPSE software developed by Schlumberger. This software is used to develop the *a priori* model of recharge as described in Chapter 2. This appendix is not an introduction to the software, but provides suggestions and pseudo-code for formatting different data types typical to near-surface applications. It also contains modifications to the model and flow simulations required to use the software, originally developed to model oil and gas reservoirs at depth, for near-surface hydrologic applications.

Appendix B is a reference of the arrays selected using the informed experimental design methodology.

CHAPTER 2. HYDROSTRATIGRAPHIC MODEL OF THE HARKINS SLOUGH RECHARGE POND FROM INTEGRATION OF HYDROLOGIC, GEOLOGIC, AND GEOPHYSICAL DATA

2.1 Introduction

One of the first steps in the informed imaging process for geophysical monitoring is to compile and synthesize all available information about a site. The purpose of this step is to guide the articulation of the scientific question and the definition of imaging goals, as well as to begin the process of creating an experimental design targeted at monitoring the process of interest. While the method chosen to synthesize prior information will depend on the goals of the geophysical experiment, we present below an example of this process for monitoring infiltration at the Harkins Slough Recharge Pond (HSRP). In this example, a model of the hydrogeologic structure is developed and used in a flow simulation to aid the design of an electrical resistivity imaging experiment for monitoring infiltration in the top 10 m of the subsurface. The process described here is iterative; the model and experimental design are updated yearly as more information about the site and monitored processes is acquired. The initial hydrostratigraphic model includes only the geologic and hydrologic data available prior to acquisition of geophysical data. Fluid flow simulations based on the structure defined in the initial model demonstrate that a model based solely on existing geologic data is insufficient for explaining flow behavior below the HSRP. The hydrostratigraphic model is then updated by integrating acquired geophysical data.

The HSRP is located above a coastal aquifer between Monterey and Santa Cruz, CA. The region surrounding the pond is agricultural with year-round crop rotation. Precipitation in the winter months meets much of the water demand during that period, but throughout the dry summer months the area is highly dependent on groundwater resources. Over the past several decades, groundwater pumping in the basin has exceeded natural recharge, causing drawdown in the aquifer and allowing

saltwater intrusion. As saltwater has intruded landward, there have been instances of increased salinity in groundwater wells that are used to supply irrigation water. Salination of the aquifer could be highly detrimental to the local agricultural economy.

Between November 1 and May 31 each year the water agency in charge of operating the pond has a right to divert up to 2000 acre-feet ($2.5E6 \text{ m}^3$) of excess surface-flow from Harkins Slough. Diverted water is filtered of sediments and pumped to the HSRP. We refer to this period as the “diversion season.” During the diversion season water percolates from the base of the pond at ~ 29 m msl into the underlying unsaturated sands, replenishing the unconfined aquifer tens of meters below the ground surface. A clay layer at approximately -5 m msl is assumed to be the base of the unconfined aquifer, preventing the continued downward migration of water. As a result, it is thought there is a mounded water table below the recharge pond. A conceptual model of operations at the HSRP is included as Figure 2.1 [after Haines et al., 2009]. Water that infiltrates the subsurface from the HSRP is intended to serve two purposes: the first is to store water during the winter and spring to increase water supplies to meet demand for irrigation in the summer; the second is to replenish the aquifer sufficiently to slow the landward progression of the freshwater-saltwater interface.

Over the past nine years of operation, two significant challenges to achieving the above goals have arisen. The first is a decrease in infiltration rates at the pond from ~ 1 m/day to roughly 0.1 m/day during the diversion season. The decreased infiltration rates mean operators cannot divert the total amount of water to which they have rights each year, decreasing the efficiency of the pond. The second challenge is a recovery rate of roughly 25% from wells surrounding the pond, indicating hydrologic controls not currently captured in the subsurface model greatly affect the transport of water beneath the HSRP. In response to these challenges, geophysical monitoring at the pond has focused on acquiring information at two different scales. At the shallow scale, imaging focuses on monitoring infiltration within the top 2 m of the surface as water enters the vadose zone. It is hypothesized the observed decrease in infiltration rate results from accumulation of fine sediments in the top meter below the surface

and/or bioactivity in the same zone. At a deeper scale, we focused on monitoring the movement of the water in the vadose zone as it drains to the water table.

To aid the development of an electrical resistivity imaging survey designed to monitor flow in the near subsurface to address these issues, as well as to examine the current understanding of flow dynamics beneath the HSRP, we used Schlumberger's PETREL and ECLIPSE software programs to create a structural model of the hydrostratigraphy at the pond and to simulate flow within the modeled region during the 2007-2008 water year. This chapter describes the development of that model and provides a preliminary discussion of the value of information provided by geophysical methods in developing an understanding of local hydrologic processes. The chapter is organized as follows. Section 2 describes all data sets available for use in the model development. Section 3 looks at the first iteration of the structural model, using only existing hydrologic and geologic information. Section 4 describes modifications to the original model in a second iteration incorporating geophysical data sets. In Section 5, we compare the flow simulation results with a brief discussion of the value of information from the additional geophysical information. Section 6 offers some conclusions and suggestions for further development and uses of the structural model and flow simulations.

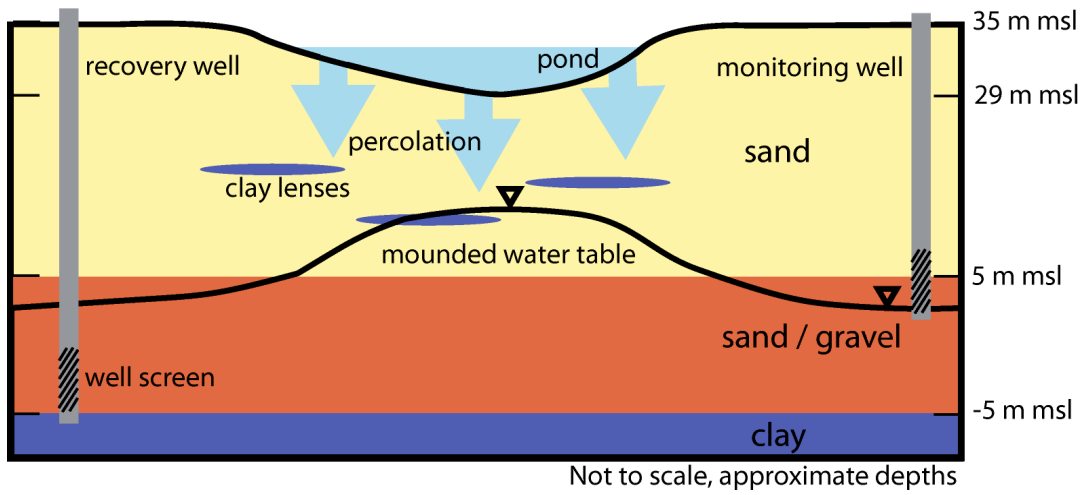


Figure 2.1. Conceptual model of operations at the Harkins Slough Recharge Pond [after Haines et al., 2009]. Depths are approximate, and the image is not to scale.

2.2 Data Sets

Because the HSRP is dry during the summer, it is possible to acquire several data sets from the base of the pond each year. A number of hydrologic, geologic, and geophysical data sets have been acquired at the HSRP and from a monitoring network surrounding the pond. In the initial structural model and flow simulation only the geologic and hydrologic data available prior to beginning geophysical investigations are included. This creates a baseline for the conceptualization and quantification of hydrologic processes at the pond. We update the structural model using geophysical data, which enhance our understanding of structure and hydrologic properties beneath the recharge pond as well as changes in physical properties as a result of hydrologic processes. Descriptions of the acquired geophysical, hydrologic, and geologic data are included below. Details on the integration of each type of data into the hydrostratigraphic model appear in subsequent sections.

2.2.1 Well Data

There are ten recovery wells, nine monitoring wells, and one deep monitoring well surrounding the HSRP, located as shown in Figure 2.2. Wells labeled 'rw' are the recovery wells, while those labeled 'mw' indicate monitoring wells. All recovery wells are drilled to a depth of approximately 55-70 m below ground surface. We have adopted mean sea level (msl) as the datum for reporting depths. The deepest point in the HSRP is approximately 29 m msl. Given topographic variations, most recovery wells are completed to a depth of -10 to -20 m msl and screened from approximately 5 m msl to -5 m msl. The screened section in each well is positioned to coincide with a gravel layer which overlays a roughly 5-m thick clay aquitard. Driller's logs are available for each of the recovery wells, and the local water management agency has defined a percent-coarse classification of the sediments on 1-ft intervals over the entire depth of each well. The percent coarse definition is derived from textural descriptions in the driller's logs. For each interval the adjectives are divided into one of three classes: texture, modifier or qualifier. A numeric value is assigned to each adjective

according to Table 2.1, and the percent coarse value is tabulated as a weighted combination of values within each class according to equation (1) below [Lear, 2009].

$$value = Text + \begin{cases} 0.2 * Mod, & Mod \geq Text \\ -0.2 * Mod, & Mod < Text \end{cases} + \begin{cases} 0.05 * Qual, & Qual \geq Text \\ -0.05 * Qual, & Qual < Text \end{cases} \quad (1)$$

In addition to the lithologic information, electrical conductivity logs were collected in rw1 through rw9 during drilling. Production information from the recovery wells is also collected throughout the water year, which runs from October 1 to September 30. The production data are reported in the form of extraction rates from each well, recorded at hour intervals.

The monitoring wells are drilled to a depth of 40-55 m below ground surface. Most wells are completed to ~3 m msl and are screened over the bottom 15 m of the well. The pressures in the monitoring wells are recorded every 15 minutes throughout the year. By correcting for barometric pressure, also recorded at 15-minute intervals from a weather station installed near the HSRP, the hydraulic heads in the wells, which correspond to the water table elevation, is determined. In model development, hydraulic head data in the monitoring wells are averaged on hour intervals. In addition to the nine monitoring wells there is one deep monitoring well northwest of monitoring well 7. The location of this well has not been surveyed, but it is completed to a depth of 76.2 m below ground surface. Given local topography, we estimate the bottom of this well is at approximately -30 m msl, and that it monitors the hydraulic head in the confined aquifer below the basal clay in the developed model.

Texture (Value)		Modifier (Value)		Qualifier (Value)	
Clay	(0.2)	Clayey / Sticky	(0.2)	Fine/ Cemented	(0.2)
Sand	(0.7)	Silty	(0.4)	Coarse-Cemented / Tight	(0.3)
Gravel	(1.0)	Sandy	(0.7)	Hard / Packed	(0.4)
		Gravelly	(0.9)	Soft / Medium	(0.6)
				Loose	(0.7)

Table 2.1 Texture, Modifier, and Qualifier descriptors and values applied to the percent coarse sediment classification.

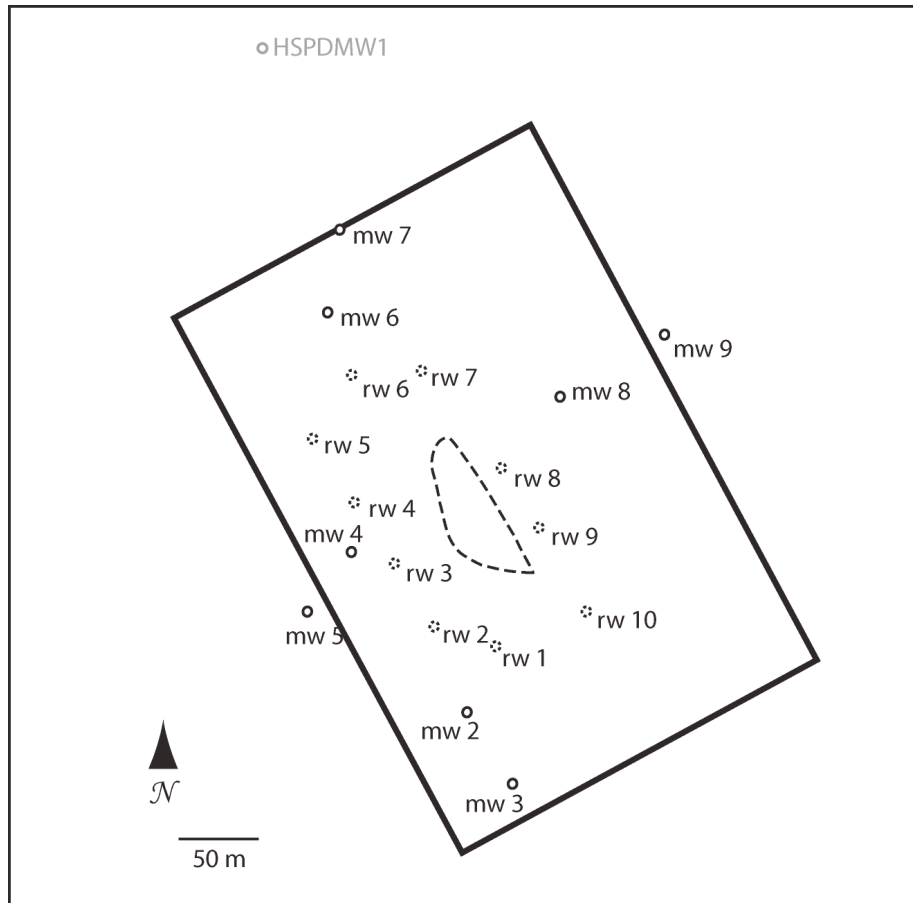


Figure 2.2. Location of recovery wells and monitoring wells surrounding the HSRP. The dashed line indicates the boundary of the recharge pond; the solid black line indicates the edges of the model space. A deep monitoring well, labeled in gray as HSPDMW1, is known to exist northwest of mw7, though the exact location is unknown.

2.2.2 Core Data

A total of five shallow cores, ~9 m in length, were collected by A. Pidlisecky from base the HSRP in the summer of 2007 using direct push methods. The locations of the 5 cores are shown in Figure 2.3 for reference; cores are labeled bh1 to bh5 in the figure. A. Racz from UC Santa Cruz analyzed these cores for mean, median, and mode grain size; magnetic susceptibility; and water content. However, at the time of model building, only data from bh 2 through bh 4 were available. Grain size and water content data were collected at discrete points on meter intervals, and magnetic susceptibility is reported at 0.5 m intervals. The median grain size distributions for cores 2, 3, and 4 are plotted in Figure 2.4. The mode grain size distributions for the same cores are included as Figure 2.5. These plots show that, while the mode grain size is largely consistent, there is an increase in the median grain size with depth up to approximately 0.6 m, especially for the core from bh 3.

2.2.3 CPT Logs

Cone tip pressure, pore pressure, sleeve friction, and shear wave seismic velocity data are available from three cone penetrometer test (CPT) logs collected from the base of the pond in August 2007. The locations of the three CPT logs are also plotted in Figure 2.3. Each log extends to a depth of ~1 m msl. Pore pressure, tip pressure, and sleeve friction are sampled every 0.05 m and are recorded with approximately 0.2 m resolution. The sleeve friction is used to derive a normalized friction ratio, which, with the cone tip pressure, defines the soil behavior type (SBT) every 0.2 m. The SBT is a soil texture definition based on mechanical properties. The data from the CPT logs are described in detail in Haines et al. [2009], and the reader is referred to their work for a complete description of the collected CPT data. We reproduce their results here for the logs collected at location cpt 2, Figure 2.6, to highlight some key observations. Primarily, we wish to emphasize the drop in tip pressure and spike in induced pore pressure at ~27 m depth in the plots. This is approximately 3-4 m msl with respect to our datum and indicates a thin clay layer that is not apparent in the lithology data observed at the recovery wells. Below this thin

clay layer, cone penetration was limited by a dense sediment layer; this behavior is consistent with what we would expect given the existence of a gravel layer at that depth. Lithology logs support the existence of such a layer at this depth. Similar behavior with respect to a spike in induced pore pressure followed by refusal beneath the thin clay is observed in the logs collected at locations cpt1 and cpt 3. In addition, shear wave seismic velocity measurements are available approximately every meter.

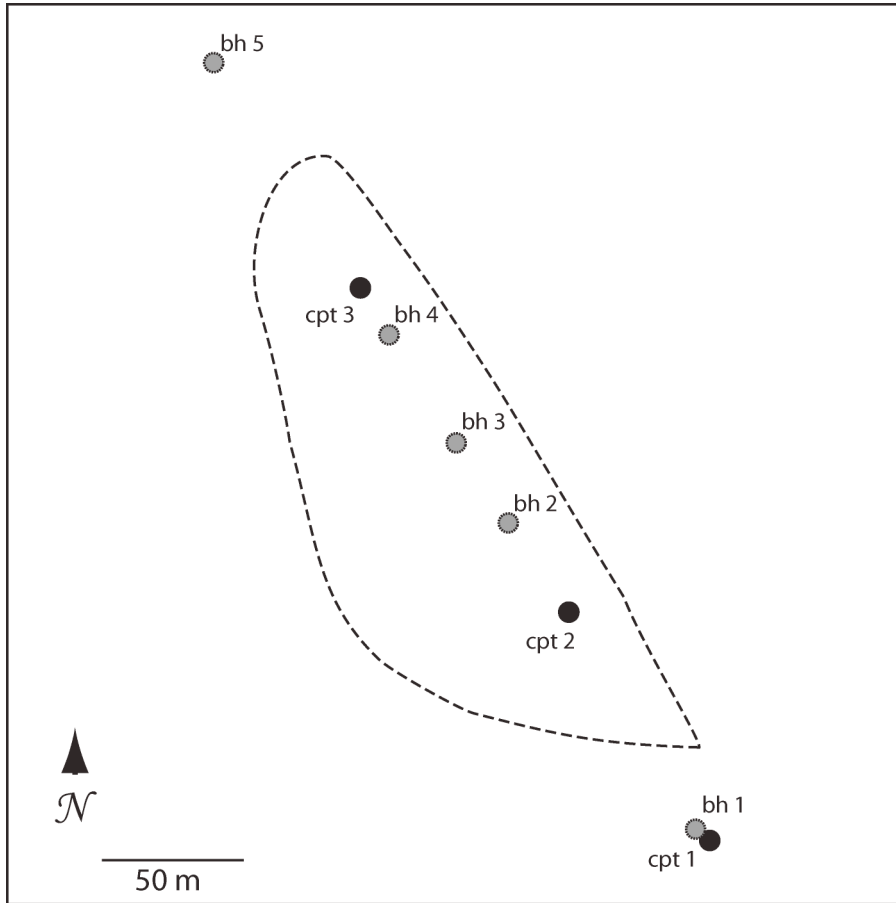


Figure 2.3. Core locations collected along the base of the HSRP in Summer 2007. Cone penetrometer test logs are indicated by black dots and are labeled cpt 1 to cpt 3; 9-m cores are indicated by gray dots and are labeled bh 1 to bh 5.

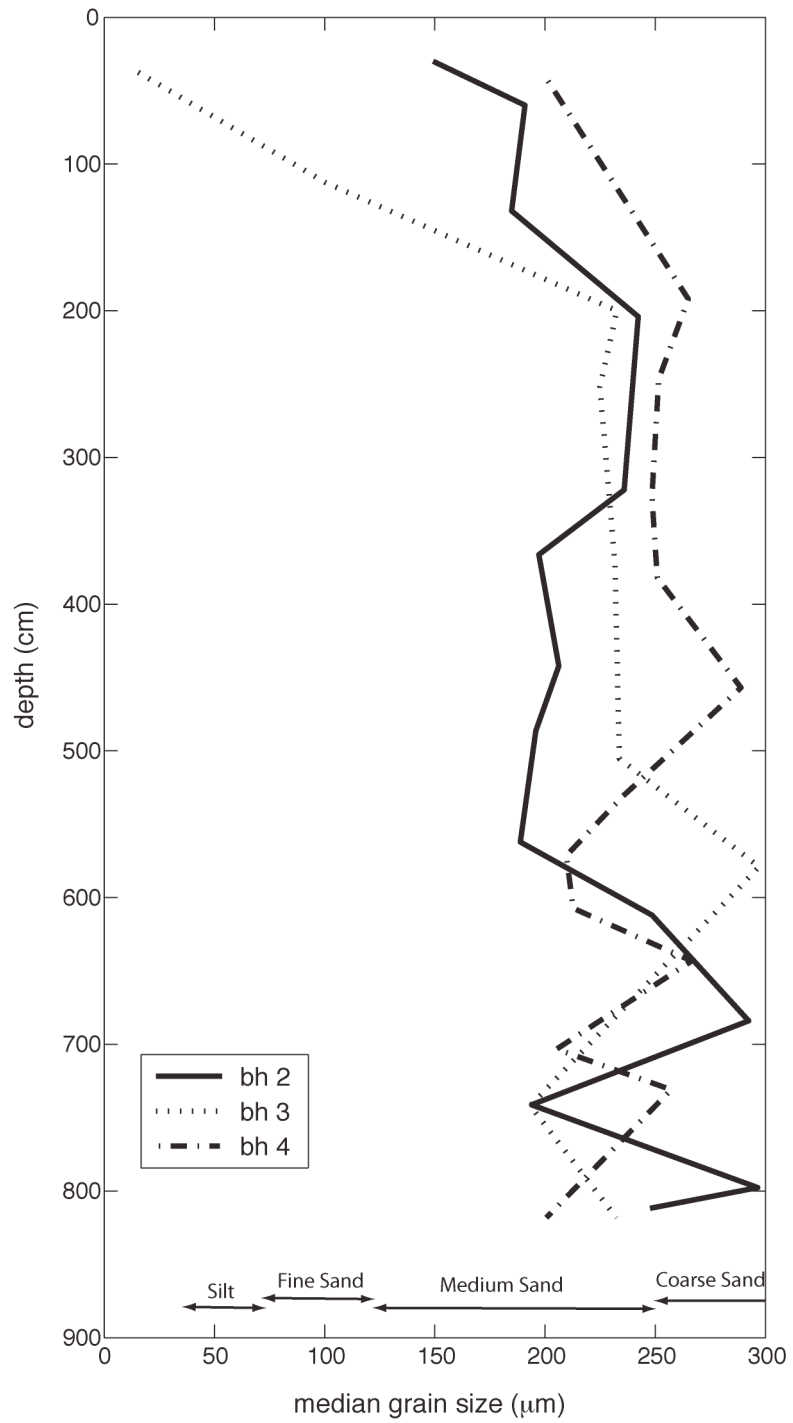


Figure 2.4. Median grain size distribution at 9-m cores bh 2 through bh 4.

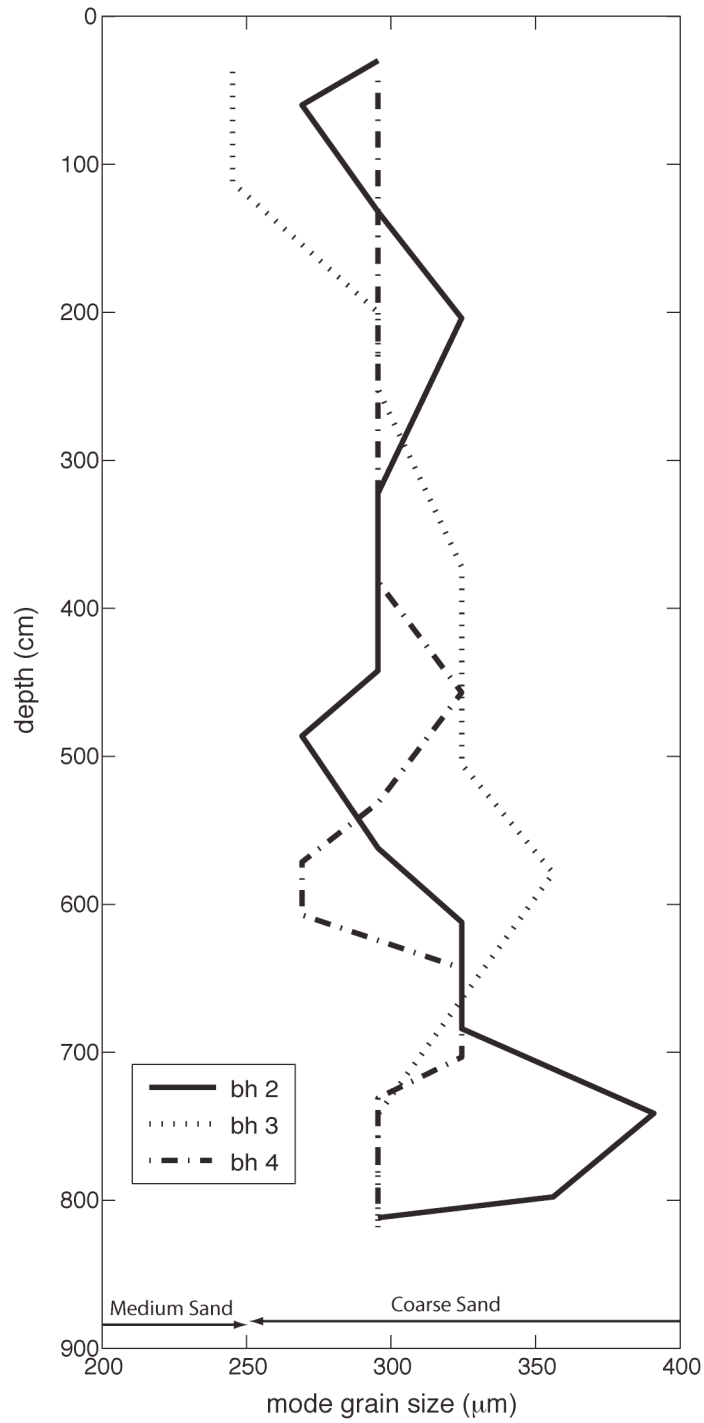


Figure 2.5. Mode grain size distribution for 9-m cores bh 2 through bh 4.

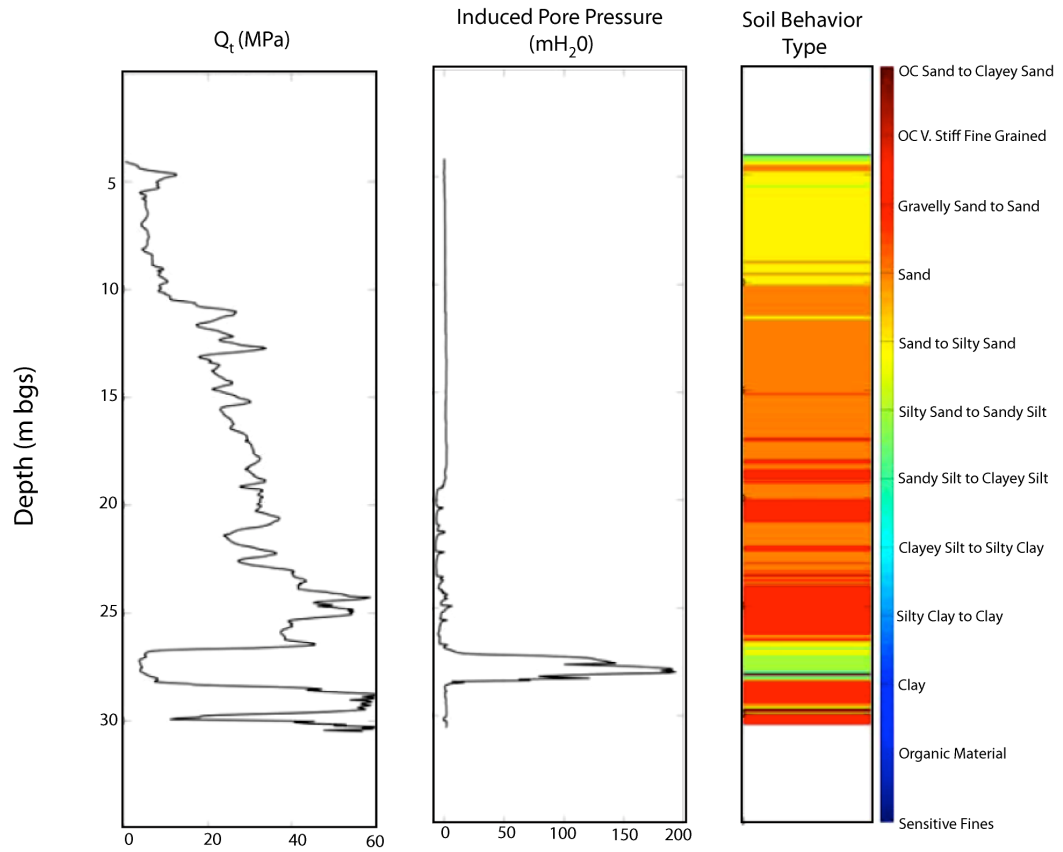


Figure 2.6. CPT logs acquired at location cpt 2 [after Haines et al., 2009]. The first panel shows the tip pressure, the second panel records the induced pore pressure, and the third panel shows the interpreted soil behavior type.

2.2.4 Infiltration Data

For three months from the time diversion begins, the water level in the pond is recorded on 15-minute intervals. Samples taken from a pressure gauge near the point of influx are corrected using barometric pressure from a weather station at the pond. The corrected pressure is converted to depth using a standard hydrostatic pressure-depth conversion. In addition to water depth at the gauge, researchers at UC Santa Cruz have acquired a LiDAR survey of the pond topography, and combined the surface topography with water depth information to create volume estimates of the pond as a function of time. The derived values of pond volume provide estimates of infiltration rates with time over the entire pond on six-hour intervals during the 2007-2008 diversion season. The average infiltration during the diversion season is plotted in Figure 2.7. In addition to pond water elevation data, point infiltration information is available from 1D profiles along electrical resistivity probes (described below) every 12 hours.

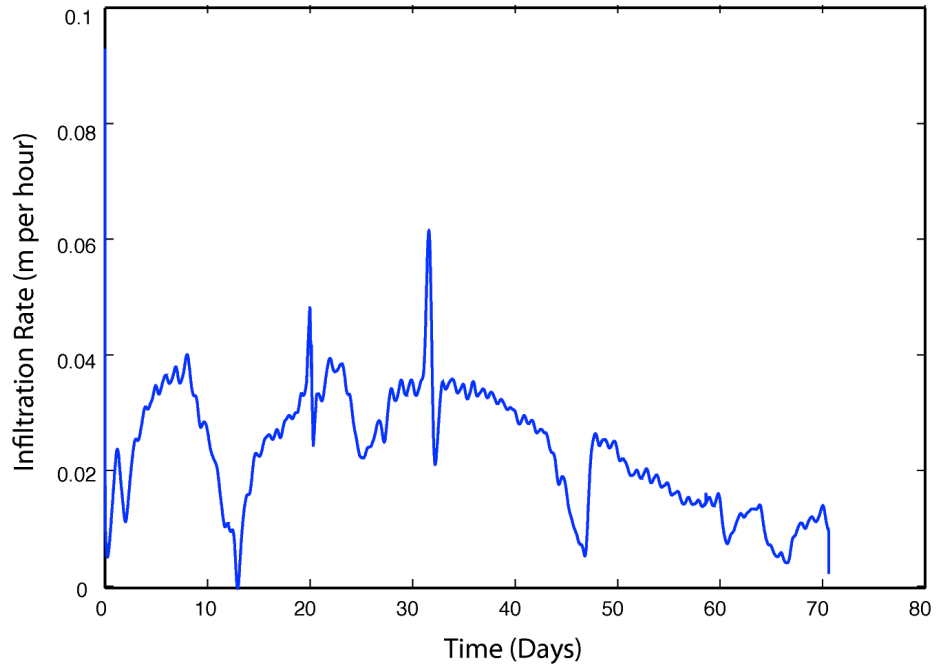


Figure 2.7. Average infiltration rate recorded for the HSRP for diversion season 2007-2008. Time is given in days after 10 January 2008.

2.2.5 Geophysical Surveys

Shear-wave seismic, ground-penetrating radar (GPR), and electrical resistivity data are available from surveys conducted on the base of the dry pond in the summer of 2007. The locations of survey transects appear in Figure 2.8. The shear-wave reflection seismic data are from two coincident, 2-dimensional lines with geophone spacings of 3 m and 1m. The 3-m seismic line is ~ 285 m long, while the 1-m seismic line covers ~80 m in the center of the longer line. The seismic data have been processed as travelttime amplitude data, extending to 0.4 s. The data acquisition and processing details, as well as subsequent shear seismic velocity inversions, are reported in Haines et al. [2009]. The processed seismic section is included as Figure 2.9. A deep reflection in the seismic section corresponds well to the deep clay layer identified in geologic driller's logs, considered to be the aquitard at the bottom of the unconfined aquifer. However, the seismic section has another, shallower reflection that indicates a possible second clay layer nearer the surface that could impact flow behavior. This clay layer occurs at approximately the same depth as the clay layer identified in the CPT logs.

Coincident with the seismic line, a 2-dimensional line of electrical resistivity imaging (ERI) data with 6-m electrode spacing was acquired in the summer of 2007. The data were collected using an IRIS Syscal unit, and have been inverted using the RESINV2D inversion algorithm of Pidlisecky et al. [2006]. The inverted conductivity model is shown in Figure 2.10. A standard L2 regularization was applied to the inversion. The electrode locations have been depth corrected for the inversion; the conductivity above the electrodes is 0 S/m, and cells above electrodes in the model space are not updated in the inversion. The region of higher conductivity at the base of the imaged region is interpreted as a combined increase in water saturation near the water table and higher conductivity resulting from a saturated clay layer. The clay layer identified by this survey is interpreted to be the deep clay layer that is the aquitard for the unconfined aquifer.

Thirteen lines of ground-penetrating radar (GPR) data collected using a PulseEcho system at 100 MHz have been processed to time and depth sections. Each

line is between 200 and 300 m in length with approximately 2 m between lines, and one line was located to coincide with the seismic and ERT lines. Transects of the GPR surveys are also plotted in Figure 2.8. Software developed by the Environmental Geophysics group at Stanford University has been used to process data and convert the time-sections to depth sections assuming a wave velocity of 0.12 ms/m. The GPR depth section coincident with the ERI and seismic sections is shown in Figure 2.11. Reflections are interpreted as migrating dunes faces, apparent because of increased magnetite-concentration from erosion of magnetite-rich rocks in hills to the east of the pond.

In addition to the surveys above, three electrical resistivity surveys have been acquired during the 2007-2008, 2008-2009, and 2009-2010 diversion seasons. The 2007-2008 resistivity survey consists of a single 1D probe and was a proof of concept experiment for monitoring the electrical conductivity profile in the near surface. The probe was 3 m long with 0.25 m electrode spacing. The approximate location of this probe is included in the upper panel of Figure 2.12. During installation a 2-m core was removed, and the hole was backfilled with a slurry of soil and water. The top meter of the probe sampled above ground to monitor the bulk water conductivity. Temperature sensors on the probe monitored ambient temperature in the pond and the ground to temperature-correct estimated conductivity values.

The 2008-2009 and 2009-2010 imaging experiments include probe data and 2D conductivity sections collected from surface surveys. A second generation of probes was constructed. The probes were again 3 m in length, but electrode spacing was 0.1 m, with 25 electrodes per probe. Eight thermistors attached to each probe monitored temperature, and multi-level sampling ports were included. In the second deployment, a direct push truck inserted the four probes to avoid the rapid-flow paths created by back filling. Once again, the top meter of the probe sampled above ground surface to monitor bulk-fluid conductivity in the pond. Figure 2.13 shows a second-generation probe after deployment in the base of the HSRP in December 2008. The four probes were placed in the base of the pond in a T-shape with 20 m between each probe, as shown in the second panel of Figure 2.12. Electrical resistivity profiles

along the probes were collected every 15 minutes during the diversion season using a dipole-dipole configuration. The same 1D probes were used in the 2009-2010 imaging experiments.

Surface lines were added to the imaging experiments in 2008-2009 and 2009-2010. In 2008-2009, three surface lines were installed with the probes acting as endpoints, as shown in the second panel of Figure 2.12. Two-dimensional electrical resistivity images were collected along each line. Lines 1 and 2 consisted of 40 electrodes, each with 0.5 m electrode spacing, and were laid between Probes 3 and 1, and Probes 1 and 2, respectively. Line 3 used only 30 electrodes, again with 0.5 m electrode spacing, and was centered between probes 1 and 4. Furthermore, lines 1 and 2 were combined to create a 40-m line, though only every other electrode was used to create 1 m electrode spacings. The lines were combined and the spacing opened to increase the depth of investigation with line 4. A standard Wenner array configuration and a non-standard array, selected using an informed experimental design algorithm, were used to collect 2D electrical profiles every 90 minutes. The non-standard array was deployed on all 4 lines, while the Wenner array was only used on lines 1, 2, and 3. Similar Wenner and non-standard acquisition arrays were used for all the surface lines in 2009-2010. However, as shown in the bottom panel of Figure 2.12, the location of lines 1, 2 and 3 was shifted to create one long line of electrodes that incorporated probes 1, 2, and 3. Line 1 contained 40 electrodes with 0.25 m spacing, extending from 5 m NW of probe 3 to 5 m SE of probe 3. Line 2 consisted of 131 electrodes with 0.5 m spacing extending SE from the last electrode of Line 1. Line 3 in 2009-2010 was a combination of lines 1 and 2, but the electrode spacing used for this line is 1 m. Hence only every fourth electrode of Line 1 and every other electrode of Line 2 were used in acquiring data on Line 3.

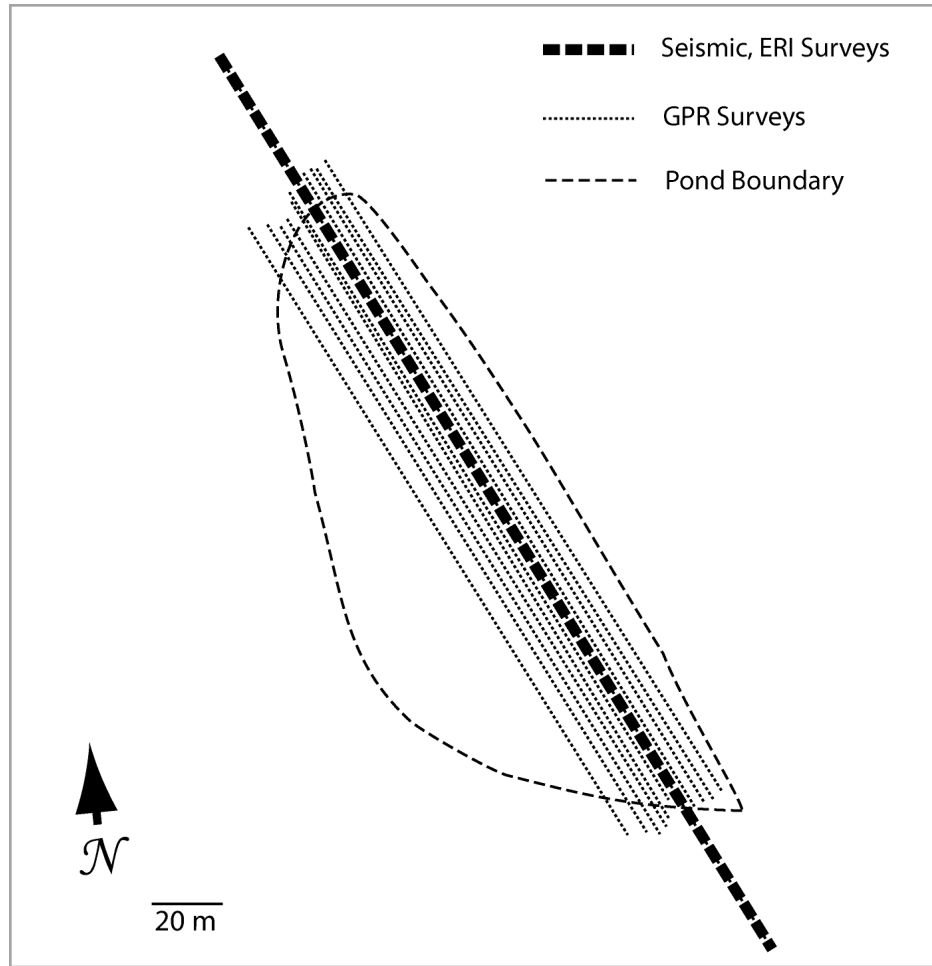


Figure 2.8. Transects of seismic, electrical resistivity, and ground-penetrating radar surveys along the base of the HSRP in Summer 2007.

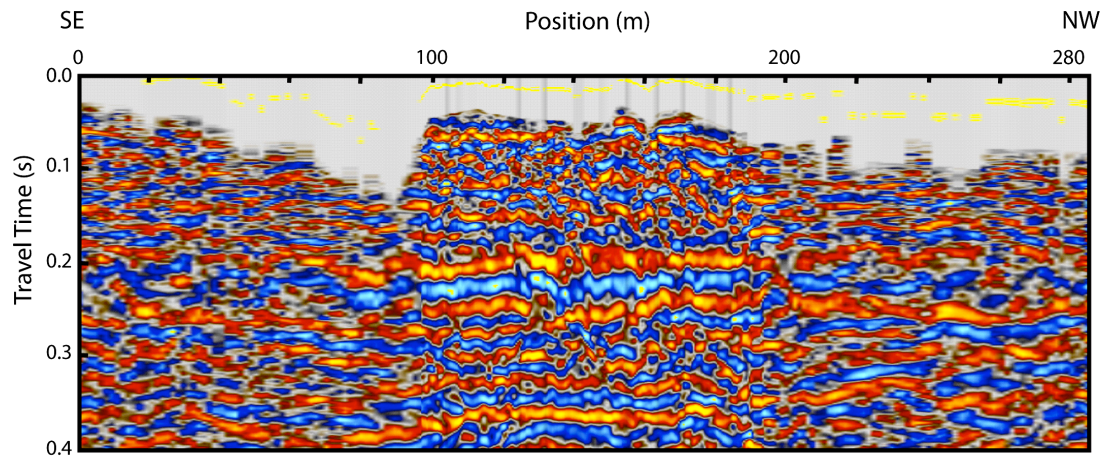


Figure 2.9. Processed seismic section after Haines et al. [2009]. The large reflector at ~0.2 seconds is the suspected shallow clay layer. The deep reflector at ~0.4 s is the clay layer identified in the drillers' logs from the recovery wells.

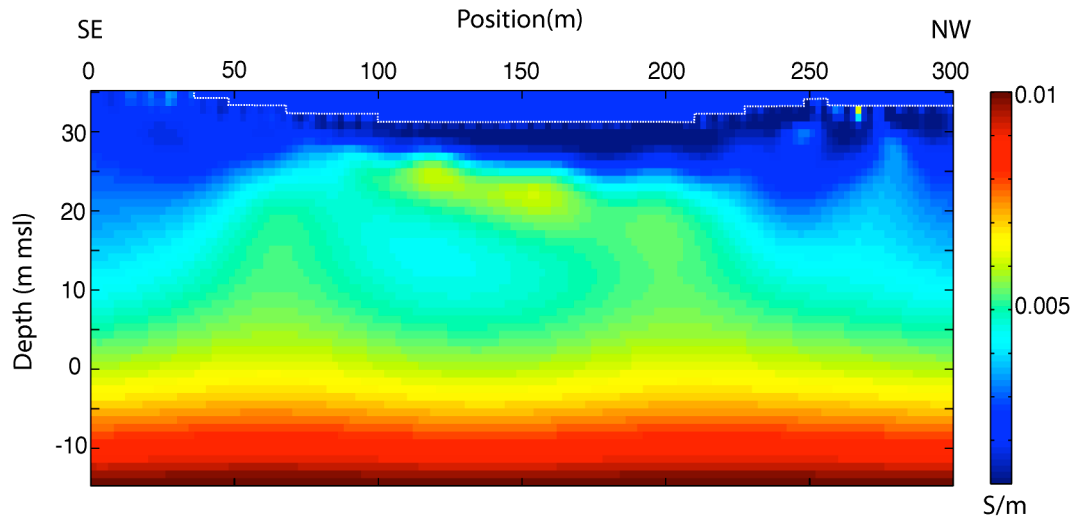


Figure 2.10. Inverted model of the conductivity structure below the HSRP. Electrode locations have been depth-corrected to account for topography, and the white line indicates the base of the pond. Depths are reported in m msl. The region of higher conductivity is interpreted to be a combined effect of the water table and deep clay layer from the conceptual model.

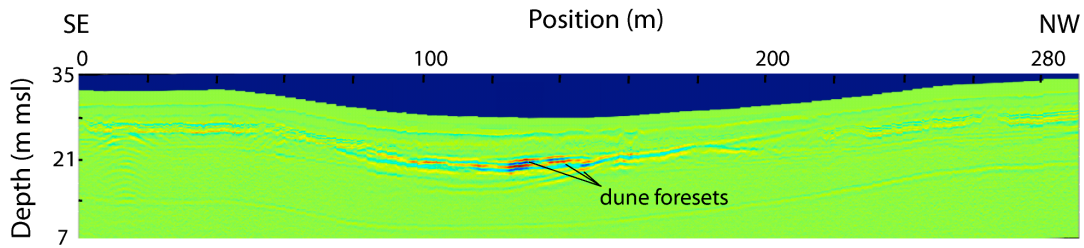


Figure 2.11. Processed ground-penetrating radar depth section collected along the same transect as the seismic and ERT sections. High magnetite content in the sands creates the large reflectors identified as dune foresets in the image.

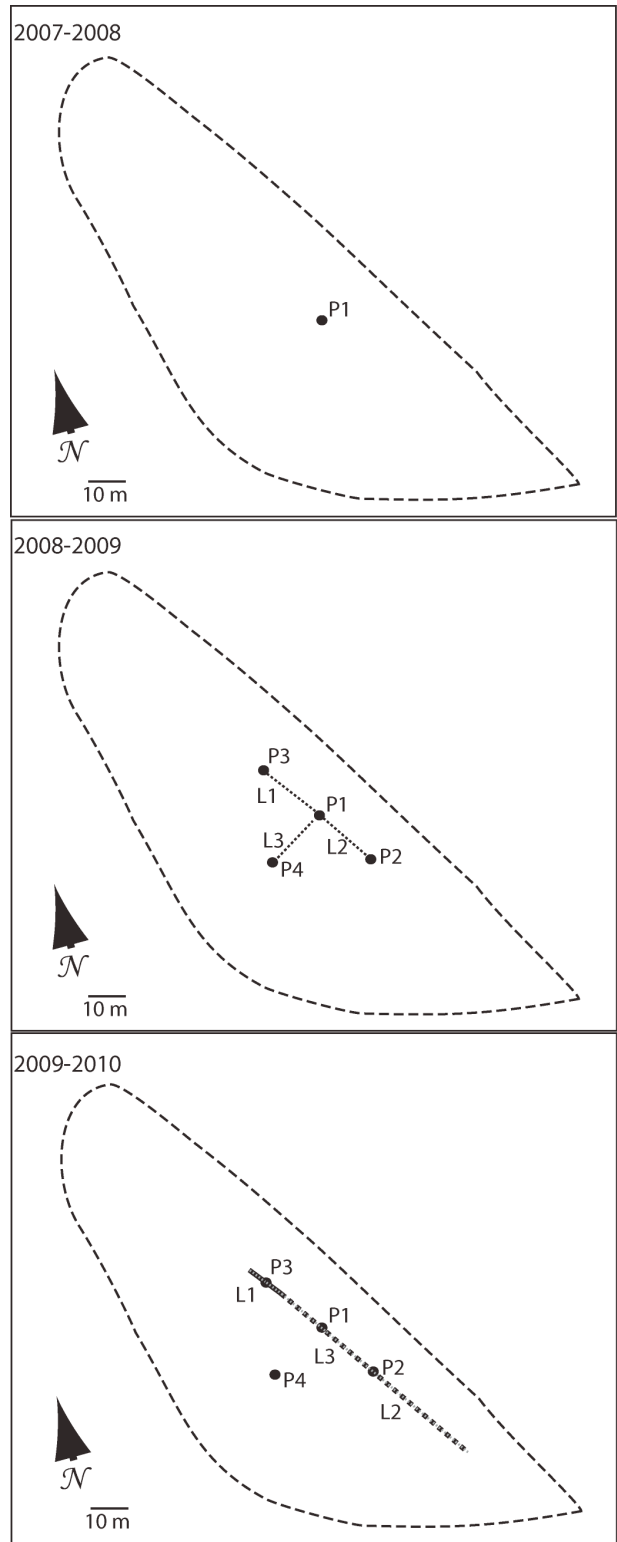


Figure 2.12. Schematic of electrical resistivity monitoring surveys. Black circles mark the locations of probes. The locations of 2D surface lines are indicated for years 2008-2009 and 2009-2010. Line 4 in 2008-2009 is an aggregation of lines 1 and 2; line 3 in 2009-2010 is also an aggregation of lines 1 and 2 from that year.



Figure 2.13. Picture of a second-generation 1D electrical resistivity sampling probe. Probes are inserted to a depth of 2 m, with ring electrodes left above the surface to sample bulk fluid conductivity. All probes are fitted with multi-level sampling ports, identified as the white tubes in the image.

2.3 Structural Model from Lithologic Data

PETREL and ECLIPSE are the software packages developed by Schlumberger for advanced visualization and multiphase fluid flow modeling. PETREL is visualization software used to create models of geologic structure. Associated with the structure are the hydrologic properties, such as porosity and permeability, that control flow behavior in the model space. These properties are exported from the structural model to ECLIPSE where, combined with fluid properties as well as boundary and initial conditions, they are used to model multiphase flow and transport. The output from ECLIPSE, time-dependent pressure and flux information, can then be reloaded into PETREL and visualized in conjunction with the underlying structure. Although originally developed for gas-oil-water reservoirs at depth, the software can be adapted to air-water systems in the near surface. As a first iteration of model building, we use only textural data from recovery wells, core data, soil behavior type, and infiltration rates to inform the structural and flow models.

2.3.1 Lithology Transforms

As a first step in building the structural model, a unified framework in which to define subsurface property information is required. The subsurface properties required for flow simulation are porosity, permeability, initial water saturation, and capillary pressure curves. The first two properties are derived from the existing geologic data. As will be discussed below, initial water saturation and capillary pressure curves are unknown for this site and are used as calibration parameters in the flow simulation.

Because the recovery well data are given in terms of percent coarse values, the cores in terms of grain size, and the CPT logs in terms of soil behavior type (SBT), a simple petrophysical transformation from soil description to porosity and permeability for all three data types cannot be obtained directly. This causes difficulties when attempting to combine the data into a single structural model. We begin by creating a transform to change the numeric values of the percent coarse description to SBT based on the texture, modifier and qualifier descriptions used to determine the percent coarse classification. Ultimately, we wish to have all data in

terms of grain size, from which it is possible to estimate porosity and permeability. However, we chose to transform all data sets to SBT given that the majority of data was already defined in a textural context. SBT was chosen because it allows for material types that are analogous to lithologies, for which the PETREL software is well suited. This classification also limits the number of material types to 12 discrete classification types. Furthermore, SBT classifications are easily correlated to average grain size. Table 2.2 includes the information used to make this transformation.

Classification Number	Soil Behavior Type	Texture	Modifier	Qualifier
1	Sensitive Fine Grained			
2	Organic Material			
3	Clay	Clay	Clayey / Sticky	Hard / Fine / Tight
4	Silty Clay - Clay	Clay	Silty	Tight
5	Clayey Silt - Silty Clay	Clay	Gravelly / Sandy	Soft / Medium
6	Sandy Silty - Clayey Silt	Sand	Clayey / Silty	Hard / Fine / Tight
7	Silty Sand - Sandy Silt	Sand	Silty	Tight
8	Sand - Silty Sand	Sand	Silty	Medium / Loose
9	Sand	Sand	Sandy	Soft / Medium
10	Gravelly Sand - Sand	Gravel / Sand	Sandy / Gravelly	Loose
11	Very Stiff Fine Grain (Cemented)	Clay		Fine / Cemented
12	Sand - Clayey Sand (Cemented)	Sand / Gravel		Fine / Packed / Cemented

Table 2.2 Texture description conversions to soil behavior type.

Once all the material classification information is in a common format, which will be referred to hereafter as the lithology, grain size data from the five cores and studies from the literature are used to define an average grain size and associated

porosity for each soil behavior type. The porosity transformation relies primarily on the correlations presented in Beard and Weyl [1973] for the porosity of artificially mixed, wet-packed sand as a function of grain size, reported in Table 2.3. Based on differences in the available values of mean, median and mode grain size at each 9-m core, it is assumed that the poorly sorted sediment samples are most representative of sediments at the recharge pond. Based on the porosity and grain size estimates for each SBT, a Kozeny – Carman equation [Boadu, 2000] is used to estimate the transverse permeability as a function of porosity according to:

$$k = \left(\frac{\phi^3}{(1 - \phi)^2} \right) \left(\frac{d_m^2}{180} \right), \quad (2)$$

where k is the permeability, ϕ the effective porosity, and d_m the median grain size. In the definition above effective porosity is the porosity that is connected and can contribute to flow.

While the above relation works well for samples that are primarily sands and gravels, the relationship does not hold well for clay-rich SBT classifications, and values for these regions are taken primarily from the soil science literature. [See e.g. Fetter, 2001; Rawls et al., 1982]. Rawls et al. [1982] compile several studies estimating soil water properties as a function of soil type. The range of effective porosities for clay rich sediments reported in their work is between ~30-50%, and permeabilities are close to those used here. When complete, SBT classifications are transformed to the needed porosity and permeability values given in Table 2.4. For this model the transverse permeability is isotropic and defined to be an order of magnitude greater than the vertical permeability based on previous studies in the area [Lear, 2009]. We note that the transforms applied at this stage are subject to potentially large errors as they are based on general studies for synthetic samples. Ideally, grain size, water content, porosity, and permeability analyses of core sections corresponding to each SBT classification at this site should be used to derive site-specific values for these hydrologic parameters. As these data are unavailable for this site, the transform described above was required to form some estimate of the hydrologic properties of the sediments.

Size →	Gravel		Sand					Silt
Sorting ↓			Coarse		Medium		Fine	
Well sorted	38.0	38.4	38.1	38.8	39.1	39.7	40.2	39.8
Moderately sorted	32.4	33.3	34.2	34.9	33.9	34.3	35.6	33.1
Poorly sorted	27.1	29.8	31.5	31.3	30.4	31.0	30.5	34.2

1.000 0.710 0.500 0.350 0.250 0.177 0.125 0.088 0.044
Median Diameter, mm

Table 2.3 Porosity of wet-packed sand sample as a function of median diameter and sorting, after Beard and Weyl [1973].

SBT Number	SBT	Mean Diameter (mm)	Effective Porosity	k_x (mDarcy)	k_y (mDarcy)
3	Clay	0.002	0.3	1	0.1
4	Silty Clay - Clay	0.002	0.3	1	0.1
5	Clayey Silt - Silty Clay	0.044	0.35	10	1
6	Sandy Silt - Clayey Silt	0.088	0.50	1000	100
7	Silty Sand - Sandy Silt	0.088	0.47	1600	160
8	Sand - Silty Sand	0.177	0.43	4.2E4	4.2E3
9	Sand	0.35	0.35	6.9E4	6.9E3
10	Gravelly Sand - Sand	0.7	0.32	2E5	2E4
11	V Stiff Fine Grain (Cemented)		0.15	1	0.1
12	Sand -Clayey Sand (Cemented)	0.25	0.15	100	10

Table 2.4 Mean grain diameter, porosity and permeability estimates for each soil behavior type classification

Having transformed all the geologic data into a single classification scheme and derived estimates of porosity and permeability consistent with the lithology, data are imported into PETREL. Because PETREL was originally developed for oil reservoir applications, the standard input for any one-dimensional data profile is as a ‘well.’ In our case, all well logs, cores, and CPT logs fit under this generic heading, and are referred to generally as wells throughout the rest of the paper. One advantage

of this approach is that all the data sets associated with a given 1D profile can be included as part of the same well regardless of measurement resolution. For example, the SBT information contained in the CPT logs with 0.2 m resolution are easily combined with the shear velocity profile data, which have 1 m resolution in the same well. In the first iteration of model building, however, we only used lithologic information (SBT) and the corresponding derived porosity and permeability values, all with the same resolution, for each well.

2.3.2 Grid Definition

Once all 1D well data have been imported, definition of the model grid is required. The model space is defined to have lateral extents of 500 m by 750, coinciding with six cells from a regional scale model developed for the water management agency by the USGS, San Diego. The vertical extents of the model space are chosen such that our entire model space occupies the top layer of the regional scale model. These extents are used so the fine-scale modeling of the pond here can be easily integrated into the larger scale model for context. The x-discretization is set equal to the y-discretization at 10 m. Despite differences in absolute length, the model is defined at an angle relative to Cartesian coordinates of PETREL, and, hence, has nearly the same number of total cells in the x-direction, 84, as in the y-direction, 86. The x-y grid is shown in Figure 2.14 for reference.

Variable spacing is used for the z-discretization. We set the top boundary of the model space according to a topographic model derived from surveyed points along the base and at well locations around the pond. The lower boundary is set to a constant depth of -37 m msl, the same constant depth used in the regional model. The base of the model is below the deep clay layer, assumed to be the aquitard bounding the unconfined aquifer, that is identified in the recovery wells and seismic data. The average total depth of the model space is then 79 m; the total maximum depth is 96 m. The number of layers defined for the model space then determines the z-discretization of each cell. For example, if we set the number of layers to 79, the average z-discretization of cells will be 1 m, and the maximum z-discretization will be 1.2 m. The z-discretization can be made finer or coarser by dividing the model space into

vertical subspaces, known as regions, and defining different numbers of layers in each region. For this application, we assume that the entire model space has 80 layers, resulting in an average z-discretization of roughly 1 m.

The well properties are upscaled onto the grid spacing using a linear interpolation. Linear interpolation is also used to populate the model space in the x-y directions based on values observed in the wells at each z-depth weighted by the inverse of distance to the well raised to the fourth power. Although PETREL offers several geostatistical methods for accomplishing interpolation along wells and throughout the model, we have no information at the pond to guide or justify the use of these more powerful methods at this stage.

As a final step in defining the structural grid, we set the boundary conditions of the model space. At the top boundary, we define two aquifers. No flow boundary conditions are applied at the base and on the sides of the model space. Although this is a poor assumption for the side boundaries of the model space, we have no information about the regional flow conditions to justify constant flux or constant head conditions along these model boundaries. At the top boundary, a constant pressure gas aquifer with the properties of air is defined to connect to the model space along the top face of all cells in the first layer. Defining this aquifer has the effect of making the top boundary of the model space the ground surface. Because ECLIPSE is designed for modeling flow in oil reservoirs at depth, it is not automatically equipped to allow free gas exchange along the top boundary as is expected at the ground surface. The second aquifer is a constant flux aquifer, representing the recharge pond. The constant flux aquifer connects to the top face of cells in the top layer that lie within the 35 m msl contour line of the topographic surface. Defining the gas and water aquifers in this way allows for gas exchange at the base of the pond when there is no defined inward flux of water, as well as specified infiltration rates that are consistent with observations during diversion. At this stage only the connection of the aquifers to the top face of cells in the first layer of the model are defined. The following section contains details on simulating changing infiltration rates at the HSRP throughout the

recharge season. The final hydrostratigraphic model including the aquifer connections is shown in Figure 2.15 for reference.

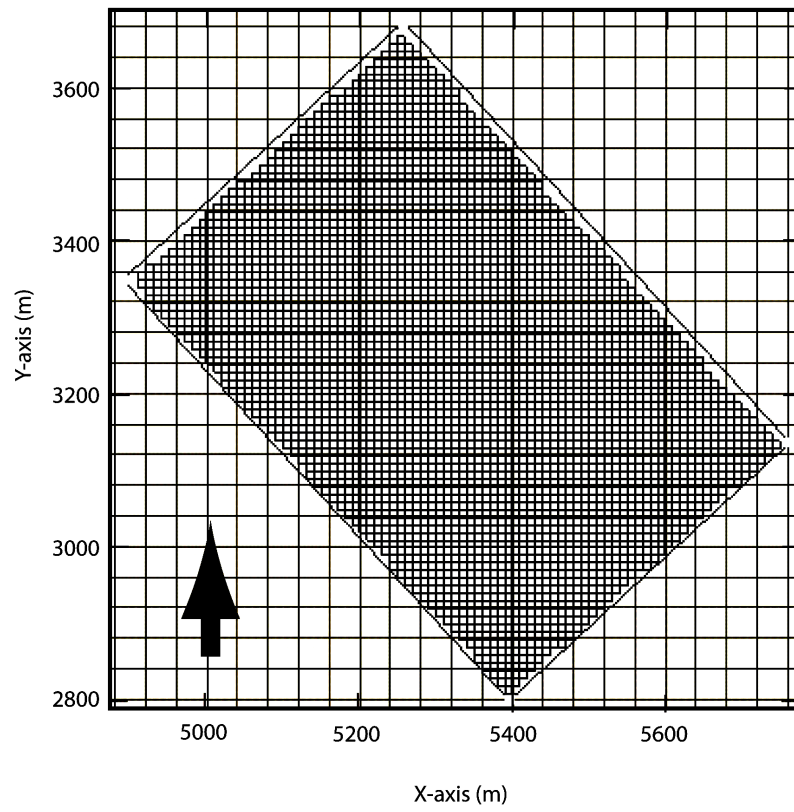


Figure 2.14. PETREL grid definition with x-y cell dimensions.

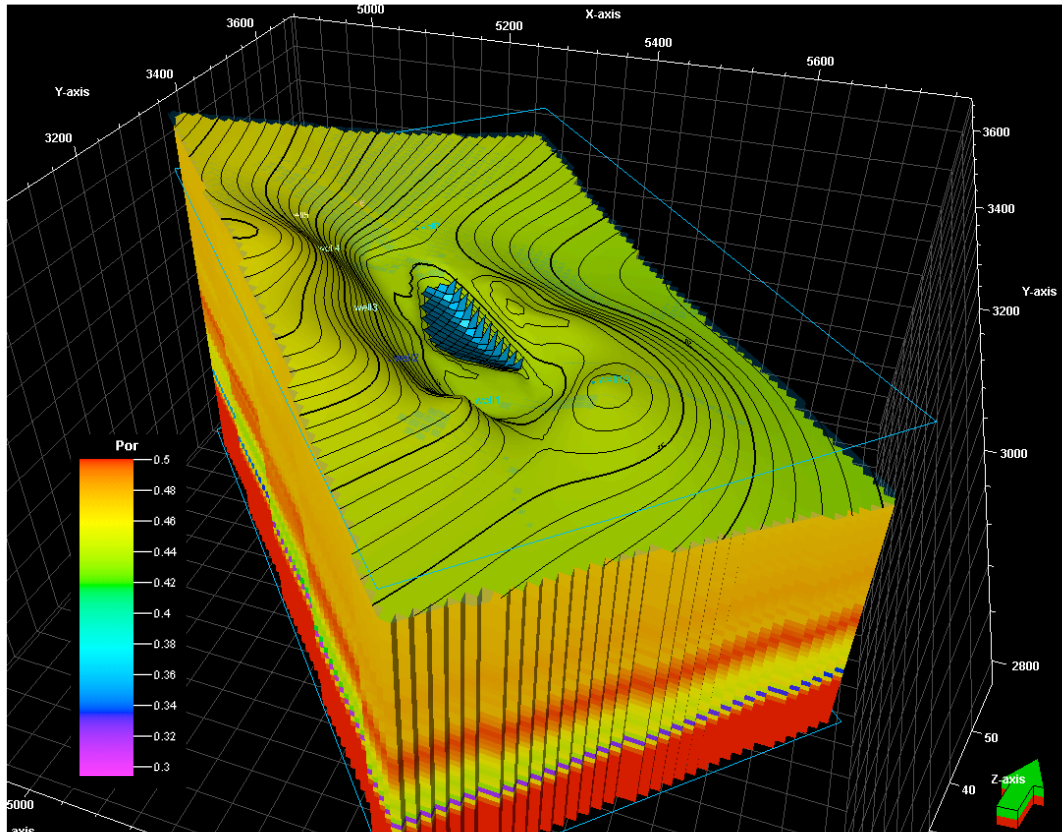


Figure 2.15. PETREL hydrostratigraphic model using lithologic data. Solid blue lines indicate the lateral extents of the model space. The topography is overlaid over the top boundary of the model. The green cells on the top boundary indicate the constant gas aquifer; blue cells at the surface indicate connection of the constant flux water aquifer. The cells within the model space show upscaled porosity data.

2.4 ECLIPSE Modeling

Once the structural information is in place, the hydrologic properties determined in the structural model are transferred to the ECLIPSE environment. A driver file controls the ECLIPSE flow simulation. Within the driver file, information passes into the program through keywords and arguments contained within one of seven sections. These sections define the constraints on the flow simulation including grid properties, fluid and rock properties, initial conditions, time-varying boundary conditions, and simulation outputs. Detailed information on the keywords within each section and how information is passed into the simulation through keyword arguments can be found in Appendix A to this thesis.

In the initial simulation, fluid flow is modeled from November 2007 to November 2008, using the structural and property model derived from the geologic data as described in the section above. Because no information exists on the relative permeability-saturation curve for lithologies within the model space, we use a general relationship derived from van Genuchten-type equations. The relative permeability-saturation curve used in this simulation is plotted in Figure 2.16; this curve is treated as a calibration parameter. Before infiltration begins, the water table is assumed to be flat at 16 m msl, the spatially weighted average of water table heights collected in the monitoring wells in November 2007. Ideally, the water table should appear as a surface representing the spatial variability of the heights observed in the monitoring wells. However, there are no data on the water table below the pond to validate the surface in this model. The flux information for the pond aquifer is taken from the pond-averaged infiltration rate as described in Section 2.2.4. These values are constant on two-week intervals with flux values equal to the average flux measurement over the same period. However, because measurements of the infiltration rate ended three months after diversion began, while we know there is still a large volume of water in the pond at that time, a constant, low infiltration rate equal to the last averaged value of infiltration rate is used until July. At this time local water users are permitted to drain the remaining water in the pond, rapidly emptying the remaining volume of water. Infiltration is, therefore, assumed to stop in July.

Given these boundary and temporal flow conditions, the monthly output for the estimated water table at the location of the monitoring wells is included as Figure 2.17. In this figure, we only plot the calculated versus true water table elevations for monitoring wells 2, 3, 4, 6, and 8, as these were the only wells that fell within the boundaries of the defined model space. The figure shows the poor fit between the simulated and measured water table elevations throughout the simulation. Although similar trends can be recognized, such as the increasing water table elevations after January 2008, the simulated heads underestimate the actual height of the water table, reach maximum values of water table elevations slower, and remain elevated for longer than the true head values as observed in the wells. We attribute this to several issues with the initial model. First, because we had no information on regional flow behavior near the pond, we assumed no flow boundary conditions, essentially turning the model volume into a bathtub with no drain. Furthermore, the simulation allows greater influx of water across the base of the pond than is observed in the field and does not account for extraction from the recovery wells, both of which serve to keep the simulated water table higher after infiltration has slowed and stopped. We seek to address some of the issues with the structural model that were highlighted in the flow simulation by incorporating information observed from geophysical data into the model.

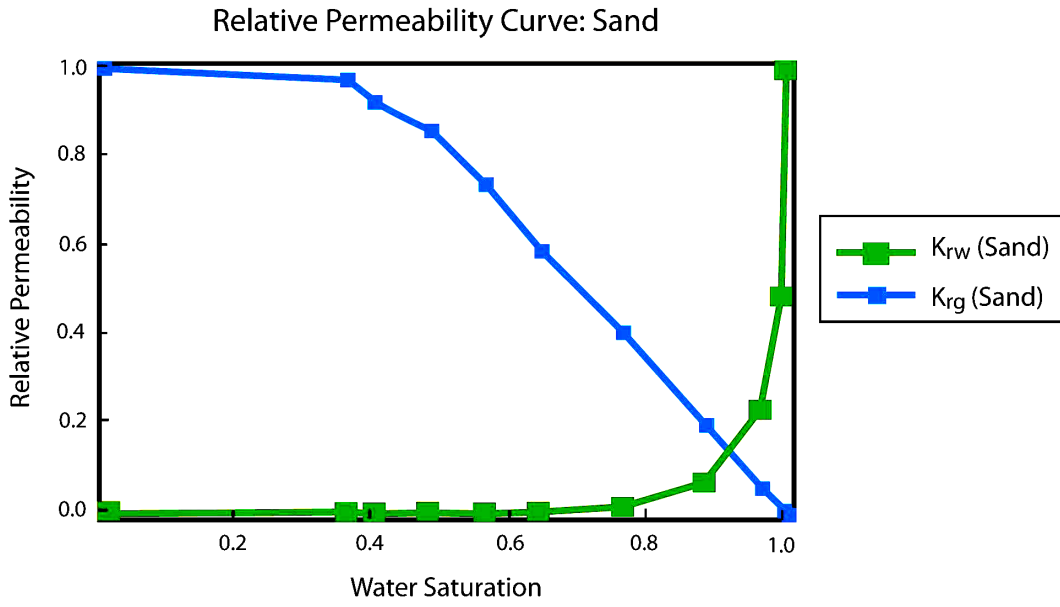


Figure 2.16. Relative Permeability - Water Saturation curve applied in the ECLIPSE flow simulation.

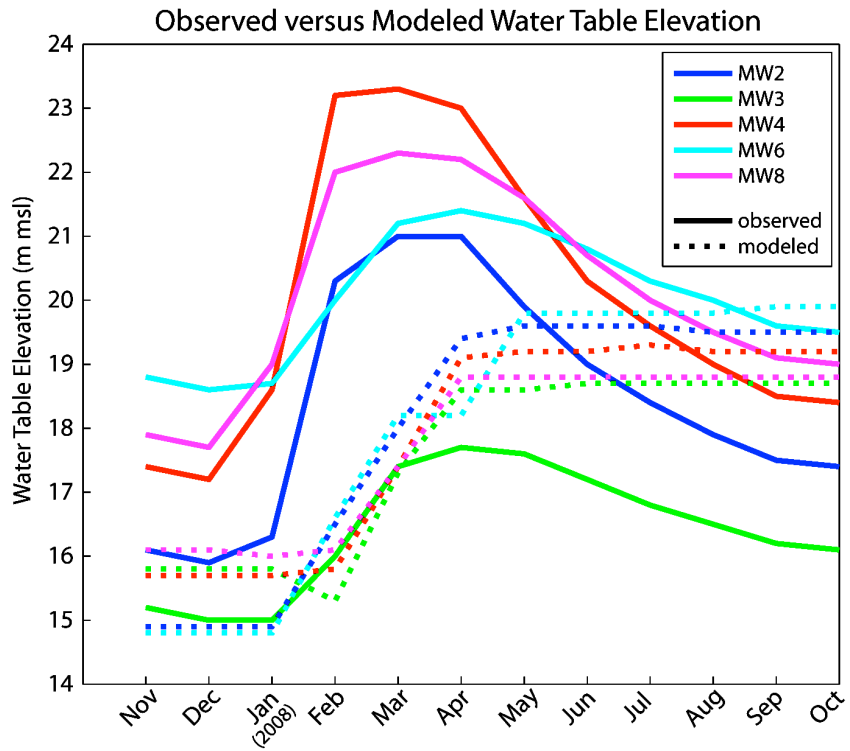


Figure 2.17. Comparison of observed versus simulated water table elevations at monitoring wells 2, 3, 4, 6, and 8 from Nov 2007 to Oct 2008.

2.5 Structural Model including Geophysical Data

2.5.1 Structural Observations from Geophysical Data

A modified version of the structural model is created based on interpretation of geophysical surveys resulting in three significant changes: zonation of the regional model, addition of a reduced permeability layer at the pond surface, and the addition of a possible, shallow clay layer. Results from the geophysical surveys and the initial flow simulations also highlight uncertainties in the geologic model, prompting a new approach to grid population and definition of the boundary conditions.

The first, and possibly most significant, addition to the lithologic model is a clay layer at approximately 3 m msl, as identified in the CPT logs shown in Figure 2.5. This clay layer is shallower than the basal clay identified as the aquitard of the unconfined aquifer in the model based on lithologic data. Although not obvious from the SBT data, which is derived from tip pressure and sleeve resistance of the cone from the CPT logs, the induced pore pressure shows a spike near 27 m below the surface on all three logs. As high induced pore pressure is indicative of clay, the peak in these logs implies there is a thin (~1 m thick) clay layer above the deep clay identified in the conceptual model. This layer does not appear in the recovery well lithology logs, indicating this shallower clay layer may be discontinuous or simply too thin to be seen at the observation scale of the drillers' logs. The presence of the shallow clay is also supported by the shear-wave seismic data. The processed seismic section (Figure 2.9) has a clear reflection at 0.2 s, which corresponds to the same depth as the peak in induced pore pressure from the CPT logs. Joint interpretation of these geophysical data provides strong evidence for the existence of a shallow, thin clay layer. However, they do not resolve whether the layer is continuous, which has significant implications for the flow behavior. A continuous clay layer will act as a flow barrier, while a discontinuous layer may have little impact on the flow behavior.

In addition to a shallow clay layer, which may impede flow, the ERT probes identify a low permeability zone in the top meter of the subsurface, which would tend to decrease infiltration rates. The 1D electrical profile collected at Probe 2 as a

function of time is shown in Figure 2.18. High values of electrical conductivity, red shades, indicate saturated soils while blue shades, low conductivity values, are associated with reduced water saturation. Figure 2.18 shows two infiltration events; the first is from a rainstorm, and the second is infiltration from ponding of diverted water. The figure also indicates a half-meter thick layer at the surface that remains saturated, regardless of the presence of water in the pond. Below this layer, however, soils never fully saturate and show evidence of drying out after an initial wetting phase when diversion begins, even when the water in the pond is at maximum levels. This behavior may be explained in part by grain size analysis of the 9-m cores. While the mode grain sizes are consistent along the length of the cores, the median grain sizes in the top half-meter are smaller than along the rest of the cores (see Figures 2.4 and 2.5 for reference). This implies more fine-grained sediment in the top half-meter, but a consistent dominant grain size, accounting for consistency in the measured soil behavior type. Because of increased fines content the top meter should have lower saturated hydraulic conductivity than underlying the sediments. As the fines-rich layer is always largely saturated, this would tend to reduce the amount of water delivered to the subsurface, with water held in narrow pore throats. By contrast, studies by Yeh et al. [2005] have shown that sands have a lower matric potential than silts, and, by consequence, a higher unsaturated hydraulic conductivity, even at low volumetric water contents. Hence, the deeper sands will tend to have a higher hydraulic conductivity, even at low water saturations, allowing water to drain more quickly than it is delivered through the top half-meter of sediments. This may account for the drying of sand below the top layer with time despite the continued ponding of water at the surface. Furthermore, the contrast in pore size between the fine and coarser regions of sand further limits the amount of water that can flow into the coarse sands because of the higher air entry pressure of the larger pores.

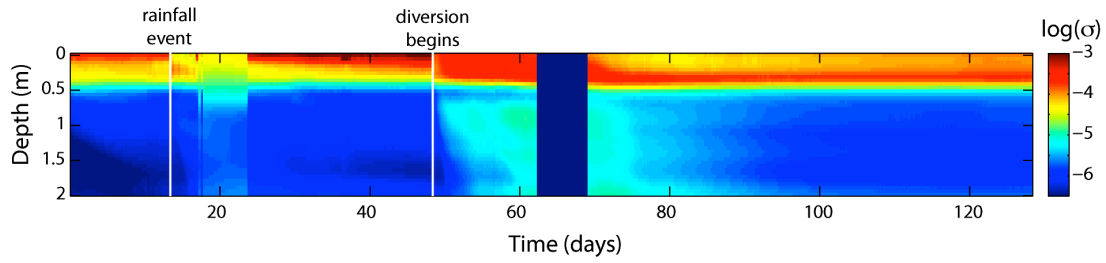


Figure 2.18. Time-lapse plot of the 1D electrical resistivity profile at probe 2. The time is given in days after 01:00:00, 24 December 2008. Depth is given in meters below ground surface. The log conductivity profile shows repeated wetting and drying below the fine-sediment layer, which remains saturated throughout the imaging experiment.

2.5.2 Grid Zonation

The presence of the fine sediment and clay layers identified by the geophysical surveys prompted model redefinition through zonation. Five model zones are defined in the redefinition of the structural model. The first zone represents the fine sediment layer at the surface. It extends from the model surface to ~0.5 m depth below the pond and is divided into 2 layers. The SBT in this zone is unchanged, but the porosity is decreased by 10% and the permeability by a factor of 10. The second zone extends from the base of the fines-rich layer to the top of the predicted shallow clay; the properties in this zone are unchanged, and it has 20 layers. The third zone is the hypothesized shallow clay. The SBT in this zone is defined as clay or silty clay with the corresponding porosity and permeability values. The fourth zone covers the region from the base of the shallow clay to the base of the deep clay, and the fifth zone to the base of the model. Properties in these zones are unchanged, and they are segmented into 30 and 3 layers, respectively.

In addition to zoning the model itself, in the redefined structural model a zoned approach to property population away from the wells is applied. The original averaging scheme tended to smooth property values across distinct layers, such as the deep clay. To avoid this, the model space is populated according to SBT away from the wells using sequential indicator simulation. Property values are then assigned to each SBT as described above. While this corrects the smoothing problem of the initial model, such an approach requires an input variogram. As this information is not available, an appropriate variogram is estimated through trial and error, resulting in an laterally isotropic model with correlation length equal to 500 m. The locations of the aquifers are unchanged from the initial model, completing redefinition of the structural model. The updated lithology distribution that is part of the new model definition is included as Figure 2.19. This updated structural model is the starting point for the design of subsequent geophysical monitoring surveys.

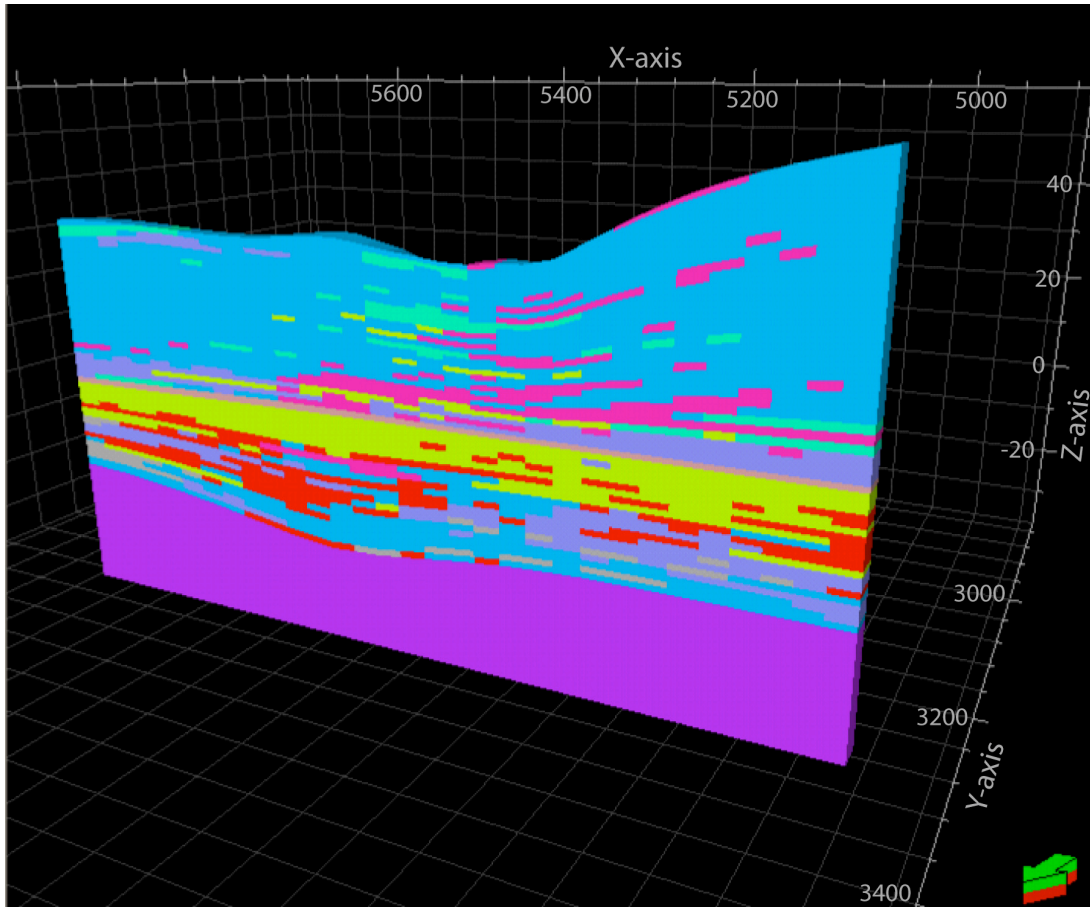


Figure 2.19. Soil behavior type distribution through the Y=3250 m slice of the PETREL structural model. The distribution includes the fine-sediment layer at the surface and the suspected shallow clay layer at ~30 m depth. Cool colors are indicative of sand-rich SBTs; warm colors represent clay-rich SBTs.

2.6 Conclusions

Because the concept of informed imaging centers on incorporating all available information into the design, acquisition, and interpretation of geophysical surveys, a method for synthesizing preliminary information about a site is required. We have demonstrated that one effective way to do this is to combine existing geologic, hydrologic, and geophysical data into an advanced visualization database. We do this in the creation of a structural and flow model of the Harkins Slough Recharge Pond located near Watsonville, CA. Operators require better understanding of flow processes at the pond to increase operational efficiency, bettering the management of groundwater resources in the coastal aquifer.

Before data can be compiled, however, it is necessary to create transforms that convert different data sets that directly or indirectly measure the same physical property into a consistent framework. In creating a model of flow beneath the HSRP, we give an example of this type of data synthesis for combining geophysical data (CPT logs) and geologic data (core grain-size analysis and soil texture logs) to derive measures of the porosity and permeability of sediments, the physical properties of interest. Once various types of data have been reconciled to a single format, it is possible to create an informed initial model of subsurface structure.

Evaluation of the structural model using a variably-saturated flow simulation identifies important gaps in understanding between what is predicted from the existing conceptual model and what is observed at the site. This evaluation highlights the need to enhance an understanding of hydrologic properties controlling flow direction and rates below the pond. Identification of poorly understood parameters serves as a basis for directing the design and implementation of geophysical surveys that will best supplement existing information. Through an iterative process, data can be continuously assimilated into the model, refining an understanding of the structural controls on flow behavior at a site. For geophysical surveys at the HSRP, a refined model incorporating acquired geophysical data has become the initial state for guiding the design of 1D and 2D electrical resistivity imaging experiments. These imaging experiments will monitor changes in subsurface conductivity structure as a result of

infiltration from the HSRP in an attempt to characterize the behavior that causes an observed fall in infiltration rates. Development of hydrologic models from integrated data sources drives the scientific questions that are the target of the geophysical surveys - what is the nature of infiltration in the very near surface? and how does the conductivity structure change in response to subsurface flow? In addition, the *a priori* model of hydrostratigraphic structure in the informed imaging framework is used to guide the design of arrays in terms spatial and temporal sampling that will be used to monitor these changes.

CHAPTER 3. INFORMED EXPERIMENTAL DESIGN FOR ELECTRICAL RESISTIVITY IMAGING

3.1 Introduction

Electrical resistivity imaging (ERI) is a non-invasive geophysical method that can be used to generate continuous maps of the electrical resistivity structure of the subsurface. Given the relationships between electrical resistivity and subsurface properties (reviewed in detail in Lesmes and Friedman [2005] and Knight and Endres [2005]) these maps can be related to properties such as clay content, porosity, water saturation, or fluid chemistry. The ability of ERI to identify spatial and temporal changes in such properties has led to its increased application in hydrological and environmental problems [see e.g. Daily, 1992; Halihan et al., 2005]. In this study we consider the problem of designing a time-lapse ERI experiment to monitor infiltration at a recharge pond above a coastal aquifer south of Santa Cruz, CA. Specifically, our interest is in developing an efficient method for selecting the electrode arrays that should be used for data acquisition.

Electrical resistivity imaging can be described by Poisson's equation, subject to appropriate boundary conditions, in the form:

$$\begin{aligned} \mathbf{j} &= -\sigma \cdot \nabla \phi \\ \nabla \cdot \mathbf{j} &= I \left[\delta(r - r_{s+}) - \delta(r - r_{s-}) \right] \end{aligned} \quad (1)$$

where σ is the electrical conductivity structure; ϕ is the electric potential; \mathbf{j} is the current density; I is the current; and $\delta(r - r_s)$ is a Dirac delta function centered at r_s , where r_s takes on the point current source or sink location. We refer to a unique combination of the locations of the source and sink electrodes as a “source pair” or simply as a “source.” The conductivity term in the above expression is simply the inverse of resistivity, and the terms will be used interchangeably throughout this paper. Equation (1) provides the physical relations between current sources, the electric potential field, and the conductivity structure that are used for ERI

experiments. In these experiments, the conductivity distribution is unknown, while the potential differences, or the difference in the electric potential field, between a series of “potential pairs” is sampled for several sources. The conductivity structure is then estimated from the potential measurements associated with different arrays, where each array is a unique combination of a source pair and a potential pair. Using the entire collection of arrays to acquire potential difference measurements is termed the acquisition of a single image. We refer to the acquisition of multiple images, in order to monitor changes in electrical conductivity, as a time-lapse imaging experiment.

In order to recover the conductivity structure from a collection of arrays (i.e. solve the inverse problem for conductivity), we require a numerical solution to the forward problem in equation (1). The numerical solution, subject to appropriate boundary conditions, of the forward problem in equation (1) can be written as:

$$\mathbf{D} \cdot \mathbf{S}(\mathbf{m}) \cdot \mathbf{G} \mathbf{u}(r_s) = \mathbf{q}(r_s), \quad (2)$$

where \mathbf{D} and \mathbf{G} are the 2D or 3D divergence and gradient operators; $\mathbf{S}(\mathbf{m})$ is the gridded conductivity structure of the model space; and $\mathbf{u}(r_s)$ and $\mathbf{q}(r_s)$ are the potential field and source vectors, respectively. We note that the measured potentials, which depend on the locations of the potential electrodes, are simply a subset of the potential field vector, $\mathbf{u}(r_s)$. Hence, the locations of the potential and source electrodes will affect the region of the subsurface sampled for each measurement. While the selection of electrode locations for the source pair and potential pair most often relies on one of a few standard, 4-electrode arrays (e.g. a Wenner array or dipole-dipole array), it has been shown that data sensitivity to the imaged region can be improved by using “non-standard” electrode arrays [Furman et al., 2002]. The objective of our research is to determine electrode arrays that will maximize the sensitivity of the acquired data in the subsurface region of interest, taking advantage of multichannel systems and non-standard array configurations. Central to our approach is the concept we refer to as informed imaging, where we incorporate all prior knowledge about a site into each stage of the geophysical experiment, from acquisition to processing to interpretation of the data.

A number of studies in the geophysical literature have investigated the sensitivity of data acquired using different arrays from the perspective of the forward problem [e.g. Furman et al., 2002; Oldenborger, 2005; Spitzer, 1998; Lehman, 1994; and Loke, 1999] as well as the inverse problem [e.g. Oldenburg and Li, 1999; Dahlin and Loke, 1998; Dahlin and Bing, 2004; Friedl, 2003; Miller and Routh, 2007]. Several of the studies demonstrate that a standard array might not produce data that have the highest sensitivity to the imaged region. Given the observed increase in sensitivity for some arrays, the question arises: how can data sensitivity be optimized through the selection of arrays used to acquire a single image of the subsurface conductivity structure? This is the question of optimized experimental design.

In this study, where we need to monitor infiltration beneath a recharge pond, we develop an approach that we refer to as informed experimental design. We term the method “informed” as opposed to “optimized” design because, unlike previous studies, we do not seek a globally optimal solution to the survey design problem. Rather, we seek to use the concept of informed imaging, the incorporation of prior knowledge into array design, with a simple, robust method that aims to improve data quality at our site over what would be obtained using a standard 4-electrode array. This efficient, practical approach can be used for a wide range of problems, and provides a computationally attractive alternative to methods of optimized design.

3.1.1 Optimized Design

Optimized experimental design has been treated in a number of disciplines in Earth sciences [Rabinowitz and Steinberg, 1990; Barth and Wunsch, 1990], as well as in applied mathematics and the physical and biological sciences. A detailed discussion of this topic can be found in the texts of Pazman [1986] and Pukelsheim [1993]. While a number of algorithms and approaches have been developed to address complexities unique to different measurement processes, the basic mathematical description is common to many applications and is concisely presented by Boyd and Vandenberghe [2004]. The work in this field has been extensive, so we limit this review to applications of relevance to this study, specifically applications to electrical resistivity and electromagnetics.

The approach to optimized design for the acquisition of electrical resistivity or electromagnetic data has tended to focus on optimization algorithms for under-determined, ill-posed problems, which require large computation times. Haber et al. [2008] propose a number of design criteria that can be used to regularize the underdetermined problem to obtain a globally optimal design. Most studies, however, have focused on a general linearization of the Jacobian, also referred to as the sensitivity matrix, or its inverse [e.g. Curtis, 1999; Bardow, 2006], where the Jacobian is composed of the elements:

$$\mathbf{J}_{ij} = \frac{\partial d_i}{\partial m_j}. \quad (3)$$

d_i represents the change to the potential field at discrete element i , for a change in the conductivity model component m_j . Experimental design approaches based on linearization of the Jacobian generally use global optimization algorithms, such as genetic algorithms [Maurer and Boerner, 1998; Furman et al., 2004], or statistical descriptions, such as Bayesian optimization approaches [Maurer et al., 2000; van den Berg, 2003; Curtis, 2004a/b]. While the above studies focus on four-electrode arrays, other studies have begun to consider experimental design for multi-channel systems or the use of novel search algorithms to speed the optimization [Stummer et al., 2002; Xu and Noel, 1993; Stummer et al., 2004; Coles and Morgan, 2009].

In addition to work focused on generally optimal solutions to the design problem, a number of studies have examined 4-electrode array selection applied to a specific hydrologic or geophysical target [Cherkeva and Tripp, 1996; Goes and Meekes, 2004; Wilkinson et al., 2006]. This problem-specific approach reduces the complexity of the design problem by disregarding general optimality and instead solving for an experimental design that is optimal for recovering an expected conductivity structure.

3.1.2 Description of Field Problem

The recharge pond of interest in this study is hydraulically connected to an unconfined coastal aquifer south of Santa Cruz, CA. The pond is located in an agricultural region where precipitation occurs primarily in the winter months while the

summer is very dry. The water management agency that operates the pond has the right to divert 2 000 acre-feet of water from November 1 to May 31 each year, referred to as the “diversion season.” When flow rates are sufficiently high, water is diverted from a near-by slough, filtered of sediments, and pumped into the recharge pond. The water percolates into the unconfined aquifer below the pond, where it is stored and recovered in the summer to meet demand for irrigation. The purpose of the pond is two-fold: to provide additional water resources during peak delivery demands and to replenish the aquifer to slow salt water intrusion into the aquifer.

In the current operational model, shown schematically in Figure 3.1, water percolates through a ~20 to 30 m thick unconsolidated sand package of relatively homogeneous medium to coarse-grained sand with possible isolated clay lenses. Below the sand package, a ~3 to 15 m thick gravel and sand layer overlies a thick clay which acts as an aquitard, separating the unconfined aquifer from the underlying saturated sandstone. The pond is designed with the intent that all diverted water percolates downward from the pond and mounds above the aquitard, remaining available for extraction during the summer. However, decreasing infiltration rates over the diversion season limit the amount of water that can be diverted to the pond each winter to less than the water agency’s entitled amount. It is presumed at present that fine sediments being deposited and/or microbial activity clog the pore space in the few meters immediately below the pond, causing the observed decrease in infiltration rate. Operators require a better understanding of the hydrologic processes controlling infiltration to improve operational efficiency of the recharge pond. As part of that effort, we propose to use electrical resistivity imaging (ERI) to monitor infiltration in the top 5 m of the subsurface.

Our objective is to define a time-lapse imaging experiment consistent with the concept of informed imaging where we use all prior knowledge about the site. In designing an imaging experiment for our particular site, we rely on an existing hydrostratigraphic model that synthesizes hydrologic, geologic and geophysical data. The developed *a priori* model identifies regions where there is a poor understanding of the subsurface properties and where there is an interest to monitor hydrologic

processes. Furthermore, the developed *a priori* model guides the experimental design process.

Information about the material properties of the subsurface is available from three 9-m cores, and electrical resistance measurements made using a 2-m probe. Both data sources indicate that the top 0.5-1 m below the pond is a fines-rich sandy layer, that grades into coarser sand. Information about the infiltration process is available as estimates of infiltration rates. Pond-scale estimates are obtained from a detailed LiDar model of the pond base and measured pond height [Schmidt et al., 2009]. Site-specific estimates are also obtained from 1D electrical resistance measurements on a 2-m probe [Pidlisecky and Knight, 2010]. The largest observed infiltration rates are on the order of 1 m/day, falling off by an order of magnitude over the diversion season.

Given this prior knowledge, we design a shallow time-lapse imaging experiment to monitor infiltration at the contact between the fines-rich layer and the underlying coarser sand layer. We constrain the acquisition of each image such that the sampling window is sufficiently small to avoid effects of temporal smearing at early times [Day-Lewis et al., 2005]. To meet these temporal constraints we are required to limit the number of measurements such that the time for acquisition of a single image is significantly less than the time it takes water to infiltrate 0.25 m, which we take as a conservative estimate of the rate of change in the imaged process.

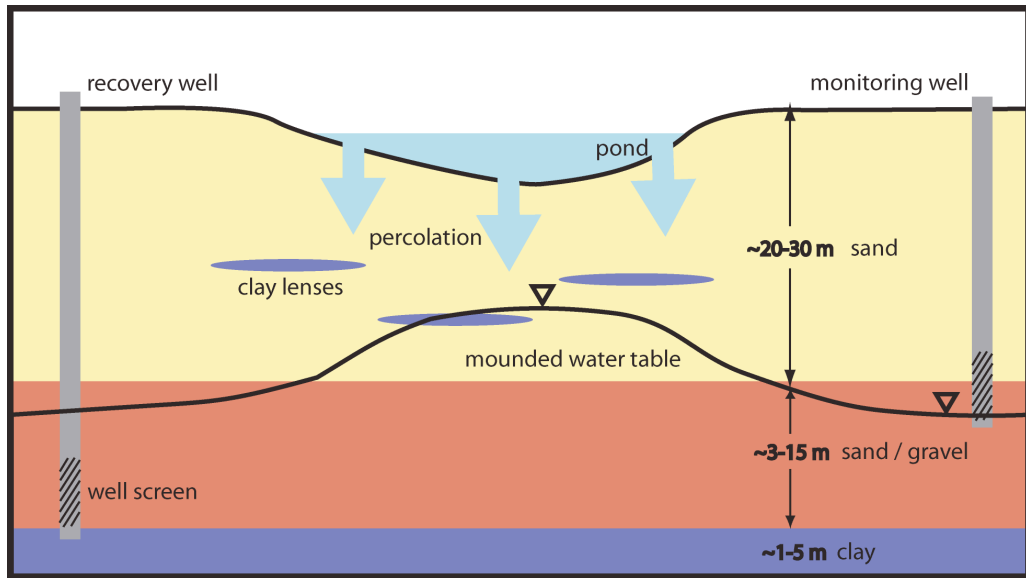


Figure 3.1. Conceptual model of the Harkins Slough Recharge Pond [after Haines et al., 2009]. Water percolates through the thick sand package and is extracted from wells screened within the sand and gravel layer. A massive clay at ~40-50 m depth is the basal aquitard for the unconfined aquifer.

3.2 Methodology

The pseudo-code describing our method for applying the concepts of informed imaging to experimental design is included as Figure 3.2. We refer to the arrays determined through use of this method as informed imaging (II) arrays.

The first step of informed experimental design, [1], is to develop an *a priori* estimate of the conductivity structure of the region to be imaged. This model should incorporate all available geologic, hydrologic, and geophysical data that give insight into the controls on the conductivity structure before the time-lapse imaging experiment begins.

In the next step, [2], we identify the spatial and temporal constraints imposed on the experiment. The spatial constraints include such factors as the total number of available electrodes, electrode-spacing, spatial extent of the measurements, and access to surface and subsurface electrode locations. Temporal constraints limit the total time for acquiring a single image given the rate at which the physical system is changing. Within the temporal constraints we incorporate the instrument configuration, whether it is a single or multi-channel acquisition system, the settling time for changing source pairs, and the time required for each measurement. For example, with a multi-channel system we can obtain multiple measurements per source pair, reducing the total time for the time-lapse imaging experiment.

In [3] we define the region of interest (ROI). In this context, the ROI is a subset of the imaged region within which we wish the data to be most sensitive, or where we expect greatest variations in the conductivity structure over the entire time-lapse imaging experiment. If desired, the ROI can be set to the full imaged region.

In step [4] we use the *a priori* conductivity model developed in [1] and the design constraints identified in [2] to generate all possible electrode pairs that can be sources for the imaging experiment, indexed from 1 to K . We then calculate the current density, \mathbf{j}_k , generated by each source pair over the imaged region. The magnitudes of the calculated current density vectors are integrated over the ROI in step [5] as an approximate measure of how well a source pair illuminates the ROI.

In [6] we filter source pairs based on the integrated magnitudes within the ROI calculated in [5]. This limits the total number of electrode pairs evaluated in subsequent steps, reducing computation time while preserving source pairs that provide a large amount of current to the ROI. We filter out source pairs producing integrated current density magnitudes that are less than 20% the maximum summed ROI value. This can be considered simply a soft constraint on the design process. However, we also note this filter can be used to limit the dynamic range of signals generated by source pairs to be consistent with the measurement capabilities of the instrument being used to acquire data. Limiting the dynamic range improves the signal to noise ratio of measurements and also provides a greater amount of stability to the inversion of collected data.

In [7] we begin to evaluate the amount of independent information contained in different source pairs to maximize the unique contribution of each source pair selected for the imaging experiment. For all source pairs not filtered out in [6], we calculate the dot product of the current density produced by each source i with itself and all other sources j . Each calculated dot product is summed within the ROI. The values of summed dot products, a_{ij} , are then ranked in ascending order in [8].

In [9] we select as sources the electrode pairs i and j that produce the $n/2$ smallest values of a_{ij} , where n is the total number of sources permitted in the final set of arrays. Small values of a_{ij} imply that the mutual sensitivity of two sources is minimized, or that the two sources sample the least amount of mutual information about the ROI, according to definitions of mutual sensitivity in Park and Van [1991]. By selecting the smallest values of a_{ij} , we maximize the unique information content for each source pair.

In [10] we turn our attention to the selection of potential pairs that will be used with each source pair to create the arrays used for the imaging experiment. Here we apply the constraints associated with the instrument configuration. For each selected source k , we accept as potential electrodes the electrode pairs, j , that produce the m largest a_{kj} –values from [9], where m is the number of unique potential pairs per source pair k . We require that selected potential pairs, j , share no common electrodes

with source pair k . In selecting the potential pairs in this manner, we maximize the sensitivity of our data over the region of interest according to the definition of sensitivity given by Spitzer [1998], which expresses sensitivity as:

$$s = \iint_{\mathbf{R}} \mathbf{j}_x \mathbf{j}_y dR . \quad (4)$$

In this measure, sensitivity is maximized when the integrated dot product of the two current density fields created by source and potential pairs x and y over some region is the greatest. In our case, a discretized version of the above expression is summed over the ROI, and we accept as potential pairs those electrode pairs that maximize the expression for a given source pair. The combinations of selected source and potential pairs identified through this process define the informed imaging (II) arrays.

1. Develop an *a priori* model of the conductivity structure of the imaged region.
2. Identify spatial and temporal constraints on the experimental design.
3. Identify a region of interest (ROI) within the imaged region.
4. For the given *a priori* conductivity model and spatial design constraints, evaluate the current density \mathbf{j}_k for all possible source pairs, $k = 1$ to K .
5. For all sources $k = 1$ to K , calculate $\sum_{ROI} |\mathbf{j}_k|$.
6. Set a threshold for the minimum value of $\sum_{ROI} |\mathbf{j}_k|$. Reject sources for which the value is below the threshold.
7. For the L remaining source pairs, calculate $a_{ij} = \sum_{ROI} (\mathbf{j}_i \cdot \mathbf{j}_j)$, where $i = 1$ to L , and $j = i$ to L .
8. Sort the values a_{ij} in ascending order.
9. Select the first $n/2$ values of a_{ij} , where n is the number of sources permitted given the identified design constraints. Use the unique electrode pairs generating the selected values of a_{ij} as the source pairs of the imaging experiment.
10. For each selected source pair k , select the last m values of a_{kj} , where j is an electrode pair that shares no electrode locations with source k . Here, m is the total number of potential pairs allowed per source pair.

Figure 3.2. Pseudo-code describing the proposed methodology for informed experimental design.

3.3 Measures of Data Sensitivity

In order to assess the performance of the II arrays selected for monitoring infiltration at the recharge pond we compare the data sensitivity of these arrays to the data sensitivity of two standard arrays. This analysis can be done as part of the informed experimental design process, but in most cases will not be needed. We complete this comparison here as a way of demonstrating the validity and value of our approach.

The data sensitivity from the acquisition of a single image is considered both from the forward and inverse perspectives. From the forward perspective we evaluate the sensitivity of data acquired with each array independent of any inversion or regularization. This gives a direct measure of the information content of the data. We also consider the extent to which data from different arrays inform estimated conductivity models in light of measurement noise and the inversion process. Although the selection of II arrays focuses on increasing sensitivity within the ROI, we evaluate the data sensitivity over the entire imaged region. We do this to ensure that the II arrays show improved performance over standard arrays even if the defined ROI (used to guide the experimental design) does not accurately capture the region where significant conductivity variations occur over the entire time-lapse imaging experiment.

3.3.1 Singular Value Decomposition of \mathbf{J}

As a first measure, we assess the data sensitivity produced by the arrays used to acquire a single image throughout the entire imaged region. Using the method first proposed by Tripp et al. [1984], and more recently implemented by Curtis [2004a,b], we examine the information content of an imaging experiment through the Jacobian, which gives the relationship between the model update, $\Delta\mathbf{m}$, and the difference between acquired and predicted data, $\Delta\mathbf{d}$, at each iteration in the solution of the non-linear problem, from the approximate relationship:

$$\Delta\mathbf{d} \approx \mathbf{J}\Delta\mathbf{m}. \quad (5)$$

The general least squares solution to (5) can be expressed

$$\Delta \mathbf{m} \approx (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T \Delta \mathbf{d}. \quad (6)$$

Referring to the definition of the Jacobian given in (3), it becomes evident that the rows of the matrix describe the amount of information the data contain on the model conductivities for a single array. Considering the singular value decomposition of \mathbf{J}^T , or the eigenvalue or singular value decomposition of $(\mathbf{J}^T \mathbf{J})$, we can immediately identify the independent information obtained in the acquisition of an image, as well as identify where measurements contain little or no independent information [Curtis, 2004a]. Comparing the array decompositions, we can see the relative amount of unique information obtained by the arrays for each acquired image. This measure is fundamental to evaluating data quality for a time-lapse imaging experiment; a set of arrays producing data with greater information content will always have more independent rows of the Jacobian.

3.3.2 Region of Investigation

While the independence of the rows of the sensitivity matrix is important in assessing the information content of acquired data, the estimated conductivity structure is also of interest. Furthermore, as the Jacobian only depends on the measurement geometry [Curtis, 2004a], it does not give a measure of how robust data will be in the presence of noise. Therefore, we also evaluate array performance from the inverse perspective. We do this using the “depth of investigation” method, first proposed by Oldenburg and Li [1999] and then expanded by Oldenborger et al. [2007]. While originally designed to identify the depth range investigated by the acquired data, this method can be used more generally to determine the region of investigation, i.e. the region in which the inverted model is sensitive to, or dependent on, measured data. Outside the region of investigation model values are more a reflection of the reference model applied to the inversion than the data.

To find the region of investigation, data are inverted using two half-space reference models; while the model structure is the same for both models, the first has a relatively higher conductivity. At each point in the inverted imaged region, we calculate the R-value as follows:

$$R(x,z) = \frac{m_h(x,z) - m_l(x,z)}{m_{refh} - m_{refl}} \quad (7)$$

where m_h and m_l are the inverted conductivity models estimated with reference models m_{refh} and m_{refl} , respectively. The R-value will be near zero when the inverted conductivity is the same regardless of the reference model, indicating high sensitivity to the measured data. When the value of R is equal to 1, the inverted conductivity model is entirely dependent on the reference model.

3.4 Informed Experimental Design: Application to a Recharge Pond

We follow the methodology outlined in Figure 3.2 to identify the II arrays to be used to monitor infiltration at the recharge pond.

In the *a priori* model of the conductivity structure at the pond, shown in Figure 3.3, the imaged region is 20 m in length, 5 m in depth. The model has three layers representing the fines-rich layer at the surface, a transition region, and the coarser underlying sediments. The conductivity value in each layer of the model is assigned based on conductivity values observed at 1D electrical resistivity probes when sediments are partially saturated. The model is centered along the upper boundary, extending from -10 m to 10 m along x, and from 0 m to 5 m in the z-direction. The model contains 80 cells in the x-direction and 20 cells in the z-direction, resulting in a cell dimension of 0.25 m, which is the minimum electrode spacing.

We limit possible electrode locations to the surface as there are no boreholes in the base of the pond. A line of 40 electrodes with uniform 0.25 m-electrode spacing is implanted in the base of the pond. This spacing is determined as a compromise between fine spatial resolution, which is related to the electrode offsets used in the arrays, and keeping electrodes adequately spaced to limit noise in the measurements from nearby, highly conductive electrodes. We are constrained temporally to less than 10 minutes for the acquisition of a single image as this study is part of a larger monitoring effort in which other experiments are being conducted simultaneously with the same acquisition system. Given the constraints of our acquisition system, as well as the timing of other ERI experiments taking place at the site, we are limited to acquiring a total of 200 measurements for a single image (the acquisition is multi-channel, so this amounts to 5 sources per potential pair). During the acquisition process, we allow a settling time of 15 ms when sources are switched, a measurement time of 3 ms for a single measurement, and stacks of 10 measurements to increase the signal to noise ratio.

The ROI is defined from -7.5 m to 7.5 m in x and 0.25 m to 1.5 m in z, maximizing sensitivity to the region in the model space that we associate with the contact between the fines-rich and coarser sediment layers while minimizing boundary

effects from numerical modeling during our evaluation of the various source electrode pairs. As described earlier, this contact region is identified as the region of most interest for time-lapse imaging because it is hypothesized that processes in the fines-rich layer at the interface control the infiltration rate.

We then evaluate the magnitude of the current density within the ROI for all possible electrode locations and rank the dot-products for electrode pairs producing at least 20% of the maximum value of summed magnitude of current density. Because the ROI is very shallow, we expect the dynamic range of signals from accepted source pairs will be sufficiently constrained by this filter. We accept 20 unique sources with 5 potential measurements per source, which provides the greatest number of unique sources while taking full advantage of the multichannel capabilities of the acquisition system. We also use as arrays the reflections of the selected arrays across the $x=0$ line, for a total of 200 measurements. Arrays that are reflected locations of source and potential pairs across the x -axis of the imaged region increase the symmetry of the measurements, thus minimizing the bias associated with the acquisition footprint. The 100 unique arrays (excluding reflections across the $x=0$ plane) are included as an appendix for reference. The selected II arrays require a total acquisition time of less than 7 minutes for each image.

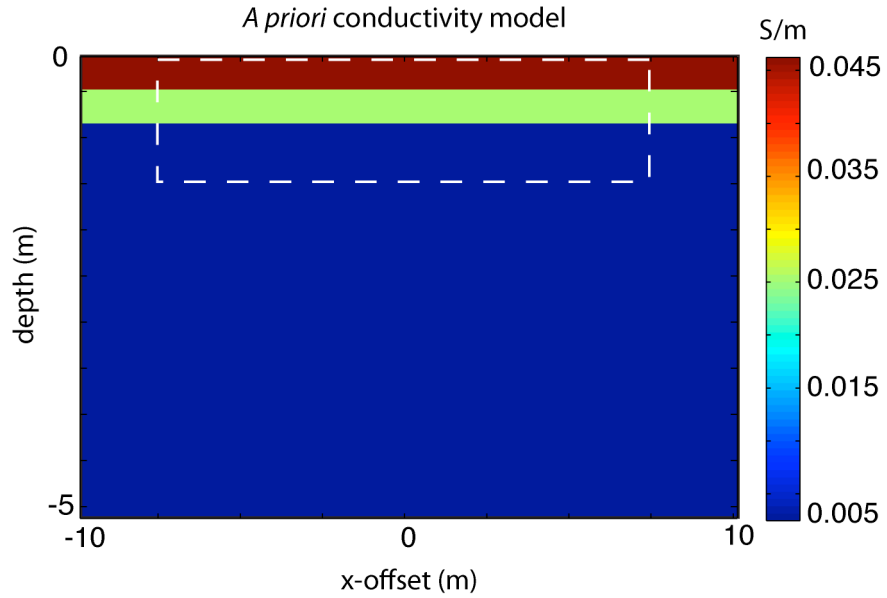


Figure 3.3. *A priori* model of conductivity beneath the recharge pond. The region of interest (ROI) is indicated by the dashed lines.

3.5 Assessment of II Arrays

We evaluate and demonstrate the performance of the II arrays, selected for monitoring infiltration at the recharge pond, using the two previously discussed measures of data sensitivity and two synthetic examples. In both synthetic examples we consider cases where the “true” subsurface conductivity model that arises due to infiltration differs from the presumed a priori model given in Figure 3.3. This acknowledges the fact that our prior information can represent imperfect knowledge about the subsurface at the pond. The first example uses a very simple buried conductive structure as a first order approximation of the conductivity variation we might expect to see at the pond, where relatively conductive water infiltrates relatively resistive sediments. This example uses an intentionally simple structure to make the comparison of the performance of different arrays straightforward. In the second example we consider a conductivity structure that is more representative of an infiltration front that we might expect to find when monitoring an infiltration pond. While still a simplified model, the second example has a diffusive, curved front between the relatively conductive saturated zone and the relatively resistive dry sediments.

The performance of the II arrays is compared to the performance of two commonly used standard arrays: a Wenner array and a dipole-dipole array. The general configurations for these two arrays are shown in Figure 3.4. In the following synthetic examples 125 Wenner arrays with a-spacing from 1 to 10 are used to acquire a single image. Dipole-dipole imaging acquires 141 measurements with a-spacings of 1 to 4, and 5 n-values. In the image acquisitions for both standard array configurations, measurements are taken evenly across the entire line of electrodes. Furthermore, for both standard arrays, and the acquisition conditions described above, the time required to acquire a single image using a standard array (~10-12 minutes) is greater than that required when using the II arrays.

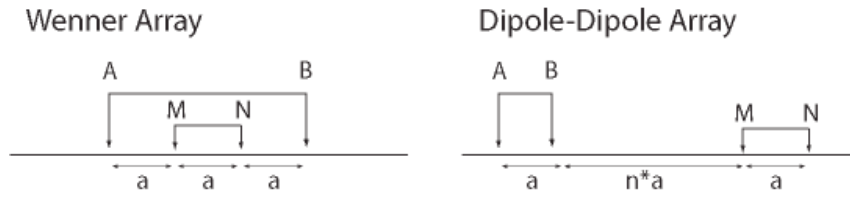


Figure 3.4. Schematic representation of Wenner and dipole-dipole arrays. A and B indicate source electrodes; M and N are potential electrodes. The spacing, a , is the offset of the potential electrodes; n is an integer multiplier of a .

3.5.1 Example 1

Data Simulation

As a first order approximation of conductivity variation at the recharge pond we assume a simple geometry for the true subsurface conductivity. We use a conductive structure, with conductivity similar to what we would expect for saturated sediments in the vadose zone, within a resistive background. The model conductivity structure used to simulate data is included as Figure 3.5. The spatial extents and grid spacing used to simulate data remain unchanged from the *a priori* model. Synthetic data for the II arrays and two standard array configurations are created using the forward modeling code described in Pidlisecky and Knight [2008].

SVD of $\mathbf{J}^T\mathbf{J}$

Based on the synthetic data, we calculate the singular value decomposition of the Jacobian for all three acquired images. The resulting singular values are plotted on a log-linear scale in Figure 3.6. This plot shows that the singular values for the standard arrays are consistently smaller than those of the II arrays, indicating they contain less unique information about the imaged region. The Wenner arrays do track the II arrays when the rank of the singular value is small. This is not surprising as we expect these arrays to locate the top and lateral edges of the conductive structure fairly accurately, indicating a high data sensitivity to those model cells. However, the trace of the singular value matrix, which approximates the area under the singular value spectrum, for the II arrays and this conductivity structure is noticeably larger than the traces for the standard arrays.

Region of Investigation

For the three acquired images, we invert the data using a 2D version of the RESINVM3D code developed by Pidlisecky et al. [2007]. This inversion uses a Gauss-Newton iterative method with standard L2 regularization and a preconditioned conjugate gradient (PCG) approach to finding the model update. This method regularizes the inversion both through the model smoothness constraint and by

limiting the number of PCG iterations for finding the model update. The reference models for the two inversions used to calculate R for this example are both homogeneous half-spaces, with log-normal conductivity values of -3 and -4. These values are close to the background log conductivity we would expect for partially saturated, sandy sediments. All data sets are inverted to approximately the same misfit tolerance of $1e-2$ with a maximum of 8 iterations. In this case the model is depth-weighted by $1/z$ [Li and Oldenburg, 1998], and the data weighting is inversely proportional to the absolute value of the observed data for all acquired images.

From the inverse perspective, the sensitivity of the inverted model to the data shows similar results to the Jacobian decomposition. The inverted conductivity models obtained with the high-conductivity reference model are shown in Figure 3.7. Calculated R-values are overlaid on the plots. All three arrays seem to locate the top and center of the conductive structure fairly accurately. The R-values suggest that the II arrays capture the conductivity contrast at the lateral edges of the conductive structure more accurately and consistently than the other two arrays. While it appears that the Wenner and dipole-dipole arrays more accurately capture the lower bound of the conductive structure, the R-values indicate that the model is poorly informed by the data below the conductive structure for both standard arrays. The conductivity estimates below the conductive structure simply equal the value of the reference model.

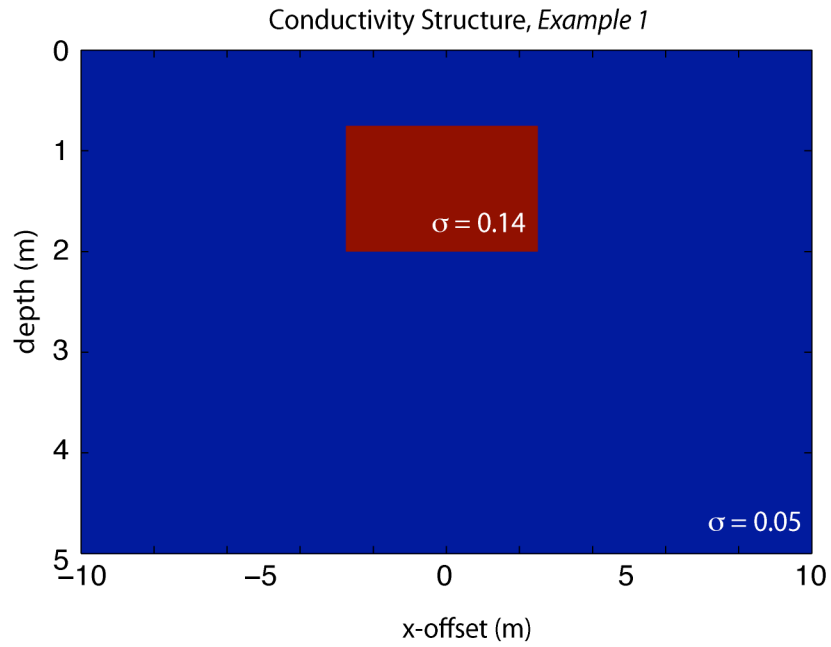


Figure 3.5. Conductivity model used to create synthetic data for Example 1. Conductivity values are given in S/m.

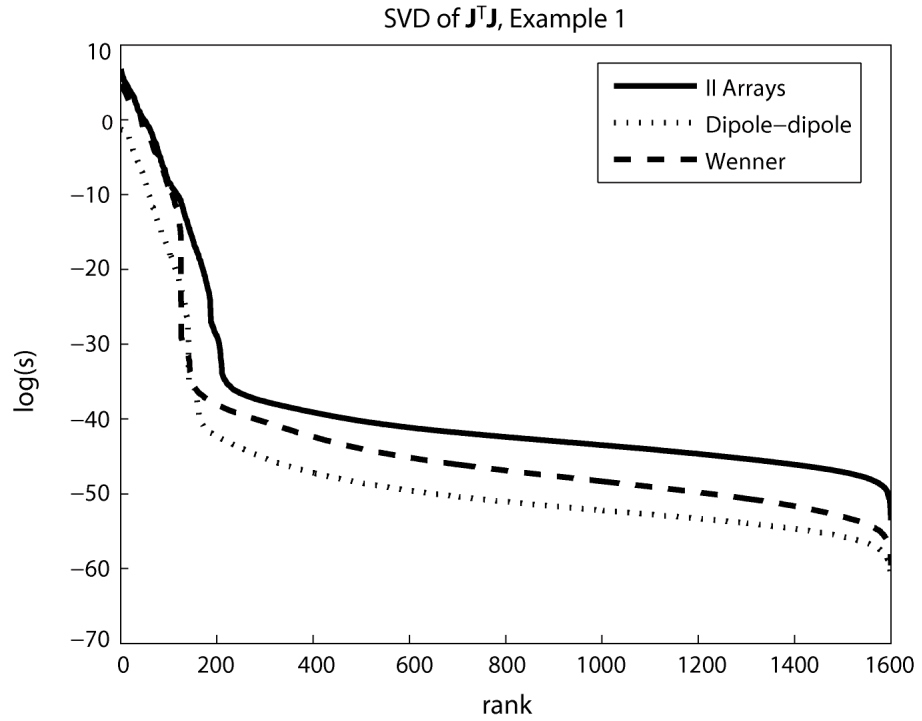


Figure 3.6. Log singular values of $(J^T J)$ as a function of rank for the II, Wenner, and dipole-dipole arrays for Example 1. The greater area under the spectrum of the II arrays indicates greater row independence, and, hence, more information content for these arrays.

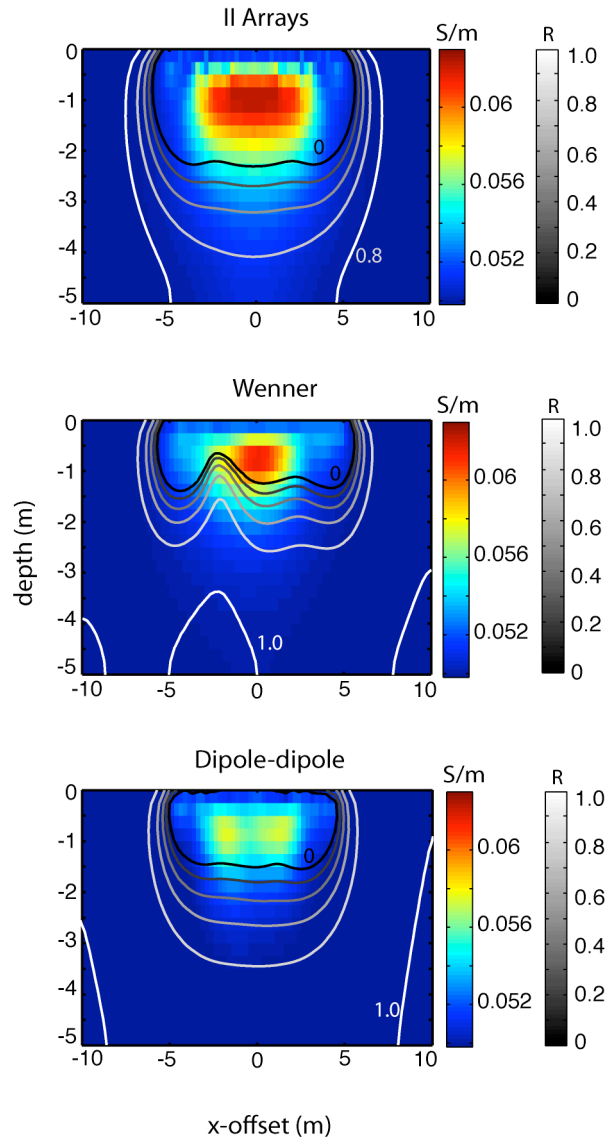


Figure 3.7. Estimated conductivity structures from Example 1 for the II arrays, Wenner array and dipole-dipole array from a homogeneous reference model. Conductivity values are in S/m. The contoured R-values from the DOI measure are overlaid on the estimated structures.

3.5.2 Example 2

Data Simulation

In the second example we compare the performance of selected II arrays to the same standard arrays, but use as our true conductivity structure one that is more representative of an infiltration front that we might expect at the recharge pond. The conductivity model for this example is included as Figure 3.8. Rather than representing the infiltrating pond water as a rectangular region with homogeneous conductivity, we represent the pond water as a diffusive front with a range of conductivity values that match observations at 1D electrical resistivity probes. The grid and cell sizes remain unchanged from the previous example, and data are generated using the same algorithm. We again consider three resistivity images acquired using Wenner, dipole-dipole and II arrays.

SVD of $\mathbf{J}^T\mathbf{J}$

The results of the singular value decomposition of $(\mathbf{J}^T\mathbf{J})$ for the three acquired images of the infiltration front are plotted in Figure 3.9 as the log singular values as a function of rank. Near the high end of the spectrum, the results for the II and Wenner arrays are clustered, but, as the rank increases, the II arrays show larger singular values, again implying greater independent information content for those arrays. Again evaluating the trace of the singular value matrix we find the area under the curve of the II spectrum is nearly an order of magnitude greater than that of the Wenner array and more than two orders of magnitude greater than that of the dipole-dipole array.

Region of Investigation

The region of investigation analysis of data from this second example uses the same parameters used for the first, simpler, structure. However, the conductivity values for the two homogeneous half-space models are -3 and -3.5, respectively. As before, the values chosen for the two models are consistent with conductivity values of partially saturated sediments observed at the 1D resistivity probes. Data are inverted

to the same misfit tolerance of $1e-2$, with a maximum of 8 iterations. We note that for this example, most inversions converge within 4 or 5 iterations.

The estimated conductivity values from data inverted using the higher-conductivity reference model are shown in Figure 3.10 with the R-values from the region of investigation measure overlaid. Again, we see the II arrays capture much more of the structure of the conductivity distribution than the standard arrays. In all three acquired images, however, the estimates over-predict the magnitude and under-predict the extent of the true conductivity structure. This over prediction is a function of reduced smoothing in our inversion models. We note that with true field data, where errors are unknown, we would be forced to increase smoothness and would likely underestimate the magnitudes of the conductivity, and over-estimate the extent of the conductivity. Nonetheless, the II arrays show a much deeper range on the region of inverted model space that is informed by the data. The image acquired with the dipole-dipole arrays shows a very restricted DOI region, indicating that these data only inform a very small portion of the entire imaged region.

For this example, we repeat the DOI analysis for the same image acquisitions after the data have been contaminated with 3% Gaussian noise to assess the robustness of the various imaging experiments in the presence of noise. The results are included as Figure 3.11. In all three acquired images we see minimal impacts on the recovered conductivity structures; values and extents increase only slightly but all three types of arrays capture attributes very similar to the noise-free data. However, examining the R-values of the region of investigation shows that the portion of the model space that is informed by the data decreases in all three images. The II arrays still inform the model space to a larger degree, indicated by higher R-values where variations in the conductivity structure are greatest. This implies that image acquisition with II arrays is robust in the presence of noise, even when the *a priori* estimated conductivity structure used to generate the II arrays differs significantly from the true conductivity structure of the subsurface.

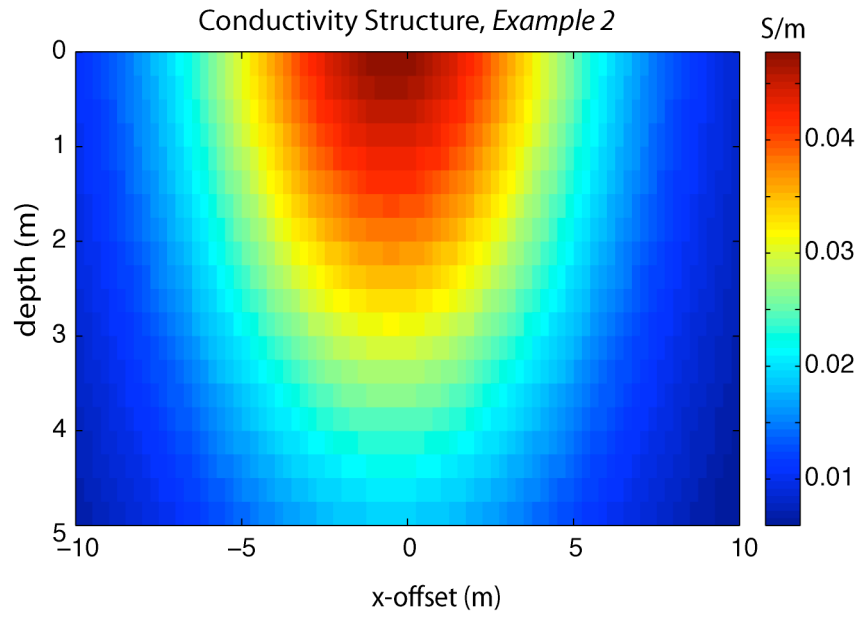


Figure 3.8. True conductivity structure used to create the synthetic data for Example 2. Conductivity values are reported in S/m.

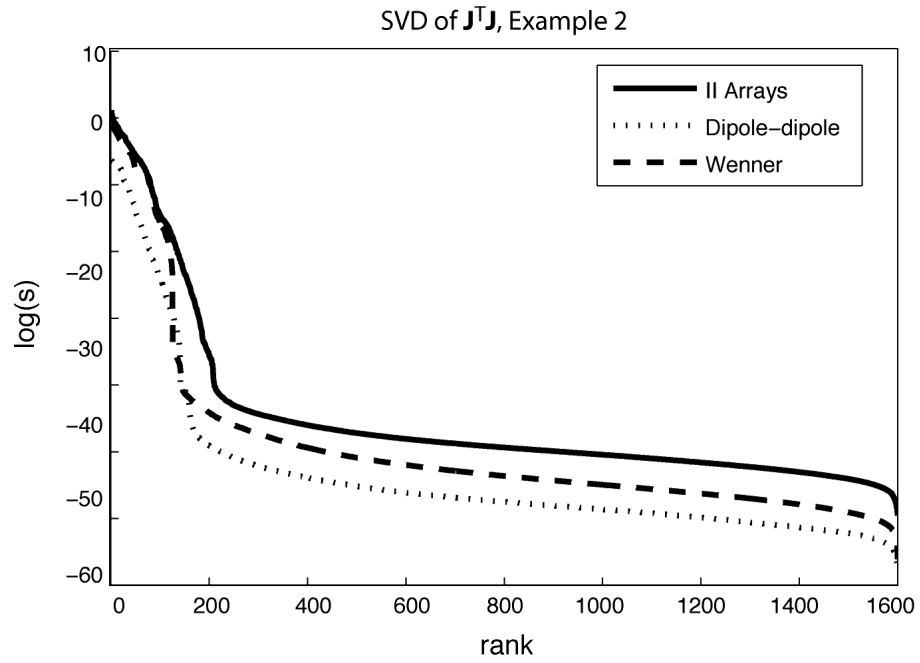


Figure 3.9. Log singular values of the $(\mathbf{J}^T\mathbf{J})$ decompositions as a function of rank for the II, Wenner, and dipole-dipole arrays in Example 2.

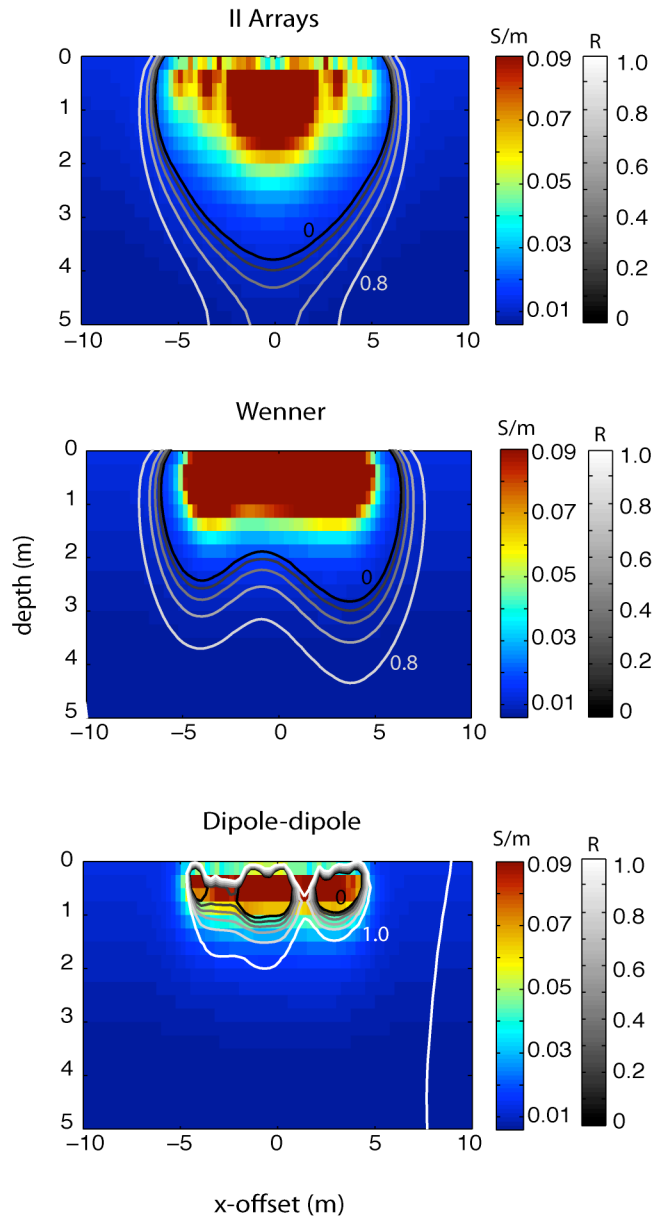


Figure 3.10. Estimated conductivity structures from Example 2 for the II, Wenner, and dipole-dipole data. Estimated conductivity values are given in S/m. Contoured R-values from the DOI calculation are overlaid on the conductivity models.

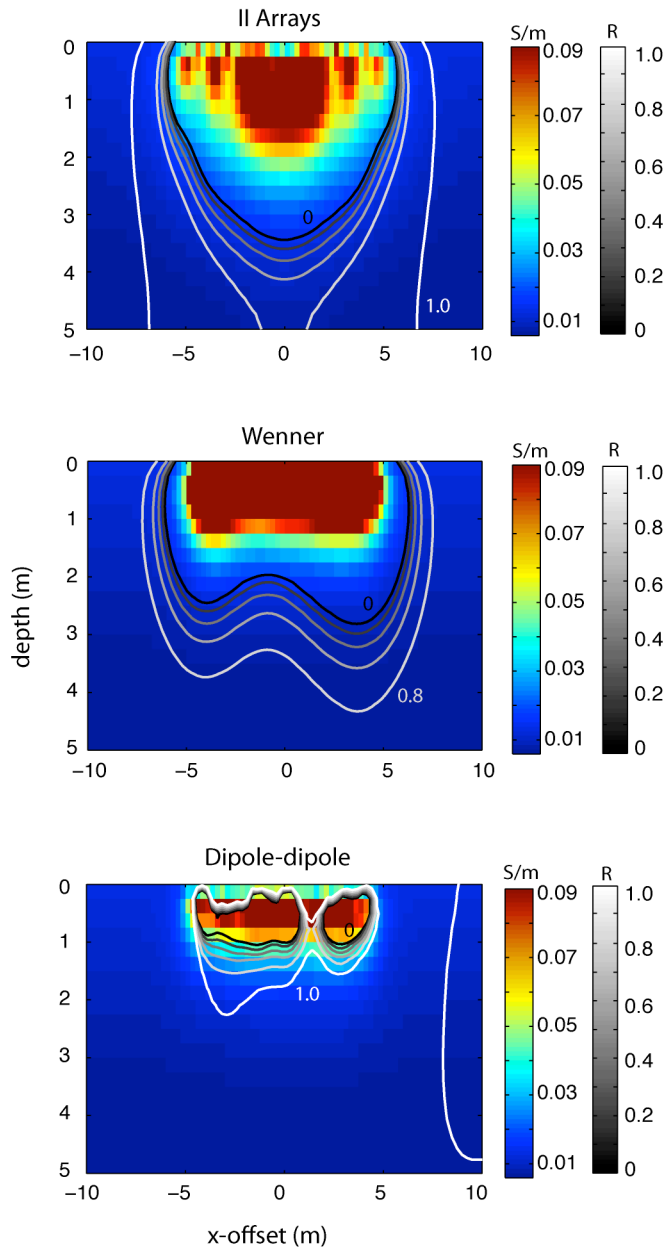


Figure 3.11. Estimated conductivity structures with overlaid R-values for the field imaging experiments using the II, Wenner and dipole-dipole data contaminated with 3% Gaussian noise. Estimated conductivity values are reported in S/m.

3.6 Conclusions

Our current research involves the use of ERI to improve the understanding of the processes responsible for the dramatic decrease in infiltration rates observed in a recharge pond. Faced with the challenge of designing an imaging experiment, with spatial and temporal constraints on electrode locations, we asked the question: How can we use all prior information available at the site to design the electrode arrays?

In this study we present a methodology for informed experimental design, the incorporation of prior information about a field site into the design process. We apply the methodology to the challenge of selecting arrays for monitoring infiltration in the top few meters beneath a recharge pond. When we compare the performance of the II arrays, selected with the informed design methodology, to two standard arrays we find that the II arrays that have greater data sensitivity to the imaged region and are more robust in the presence of noise. We conclude that this robust, computationally-inexpensive, and fast array-design method allows us to design arrays that are better than standard arrays for monitoring changes in conductivity structure during infiltration.

We note that in this study we selected a single set of arrays to be used for the entire imaging experiment. There is no reason, however, not to continuously update the array design as the experiment progresses. Further research is required to assess the benefits of using different imaging arrays that are designed to best capture the evolving conductivity structure over the course of a time-lapse imaging experiment. Future studies should also examine the increased data sensitivity that can be acquired by iteratively applying the method to update the experimental design as more information about the subsurface becomes available. In the spirit of informed imaging, it is clear that every acquired image potentially adds more information that could be used for “real-time” experimental design. These investigations could further clarify the sensitivity of the method to the *a priori* model, identifying the minimum amount of prior information required for the method to perform adequately as well as identifying scenarios in which the method does not greatly improve the quality of acquired data over that obtained with standard arrays.

CHAPTER 4. ASSESSMENT OF AN EXTENDED KALMAN FILTER APPROACH TO INVERSION OF TIME-LAPSE ELECTRICAL RESISTIVITY IMAGING DATA FOR MONITORING INFILTRATION

4.1 Introduction

Monitoring hydrologic processes in the vadose zone is of great importance in developing management strategies for groundwater resources. Flow and transport behavior in variably saturated porous media can greatly alter the rate at which, and path along which, fluids reach groundwater aquifers at different sites. While point measurements, such as Time-Domain Reflectometry probes, can give very accurate information about changes in water saturation with time [Robinson et al., 2008], these data are often insufficient for characterizing changes in saturation at the modeling scale. Electrical resistivity imaging (ERI) is a geophysical method that is sensitive to changes in the subsurface conductivity structure, which is highly dependent on the water saturation distribution in the vadose zone. Time-lapse ERI, therefore, has been widely used for environmental and hydrologic monitoring experiments. In this study we use ERI specifically to monitor infiltration beneath the Harkins Slough Recharge Pond (HSRP), part of a coastal aquifer management project in an agricultural region of Northern California.

4.1.1 Monitoring the Vadose Zone with Electrical Resistivity

In electrical resistivity imaging (ERI), currents are injected into the ground through electrodes on the surface or in boreholes, and the induced voltages or potentials are measured at a number of different locations. This is repeated for several current and potential electrode locations for the acquisition of a single image. The electrical conductivity structure of the imaged region is estimated based on the measured potential data and the applied current configurations. If we consider repeatedly acquiring images of the same region at different times, it is possible to

conduct a time-lapse imaging experiment from which changes in conductivity structure with time can be identified. Using appropriate petrophysical transforms [e.g. Archie, 1940], we can convert variations in the conductivity structure with time to temporal changes in hydrologic properties, such as water saturation.

ERI monitoring of vadose zone processes has included tracking infiltration and estimating transport properties. Daily et al. [1992] used cross-borehole ERI arrays to monitor infiltration from point and line (trench) sources. ERI has also been used in conjunction with ground penetrating radar measurements to estimate transport properties in the subsurface [Binley et al., 2002].

4.1.2 Inversion of Time-Lapse ERI Data

Reconstruction of time-lapse ERI data has largely relied on two approaches: differenced and absolute imaging. Differenced imaging has generally taken two forms when applied to monitoring vadose zone processes. The first form, proposed by Daily et al. [1992] inverts the ratio of a measured potential value at later times and the measured value in the initial image. A homogeneous reference model is used to invert each time-lapse image for the relative conductivity structure with respect to the initial conductivity structure. This approach has been applied to mapping fluid migration in 3D ERI inversions [Park, 1998] and to estimating hydrologic parameters such as hydraulic conductivity or break-through curves [Binley et al., 2002; Chambers et al., 2004]. The other form of differenced imaging, proposed by LaBrecque and Yang [2001], inverts a transformed data set, the difference between measured voltages at a given time-step and measured voltages at the initial time-step. The differenced data are then inverted with respect to a homogeneous reference model. The output of this approach to differenced imaging is an estimated change in conductivity structure with respect to the first time-step. Kemna et al. [2002] expanded this approach, using the initial inverted conductivity model as the reference model in later time-steps, removing the condition of a homogeneous reference model applied in the initial derivation.

As opposed to differenced imaging, absolute imaging inverts data sets for each time-step independently, sometimes using the inverted conductivity structure from the

previous time-step as the reference model for the current time-step. This approach has been applied to 3D data sets for studying vadose zone transport [Zhou et al., 2001; Slater et al., 2002; Labrecque et al., 2004; Looms et al., 2005; and Singha and Gorelick, 2005]. Oldenborger et al. [2007] assess the performance of differenced imaging and absolute imaging using results from the previous time-step for regularization on inversion of 3D time-series data from imaging of an injection/withdrawal experiment in a shallow aquifer. The authors found that both approaches result in greater model structure than is attained without using a time-lapse approach, but that neither approach can be used to predict adequately the full amount of injected/withdrawn fluid.

4.1.3 Description of Field Problem

Our primary interest at the HSRP is to image infiltration in the top few meters of the near surface. The pond is located in an agricultural region near the northern California coast, which has a Mediterranean climate, meaning rainfall occurs primarily in the winter while the summer is very dry. The purpose of the recharge pond is two-fold. First, it is used to store diverted storm-flow run-off during the winter to meet demands for groundwater deliveries; second, it is used to replenish the unconfined aquifer, which has been overdrawn for several decades, thus allowing saltwater intrusion. The water agency in charge of operating the pond has a right to divert approximately $2.5 \times 10^6 \text{ m}^3$ of surface water from the Harkins Slough to the pond from November to May each year. However, decreasing infiltration rates over the diversion period reduce the operational efficiency of the pond, infiltrating only $\sim 1 \times 10^6 \text{ m}^3$ in a year.

Prior investigations using 1D electrical resistivity probes recording to a depth of 2 m, and grain-size analysis of 5 9-m cores suggest a 0.5-m thick layer of fines-rich sediment at the surface, with medium to coarse-grained sand below this layer. It is hypothesized that clogging from very-fine particulates in the water column or bioactivity in the fines-rich layer and below causes the observed decrease in infiltration rate with time. However, it may be that infiltration rates are simply limited by the permeability within the top layer, and initial rates are temporarily increased as a

result of tilling of the pond base prior to infiltration. Operators require a better understanding of flow behavior below the pond to understand the parameters controlling infiltration rates.

In this study we use ERI to monitor infiltration beneath the HSRP to develop a more complete understanding of infiltration processes. We adopt an informed imaging framework [Mitchell et al., 2010], meaning all existing information about a site and experiment is incorporated into the acquisition, processing, and interpretation of ERI data. We consider two methods for the inversion of data within this framework. Both of the approaches incorporate information gained from the inversion of data acquired from imaging experiments at earlier time-steps into the inversion of data at the current time-step. In the first approach, we apply an absolute inversion scheme where the inverted conductivity model from the previous time-step is a reference model for the inversion at the current time-step, incorporated as a regularization term within a Gauss-Newton inversion algorithm [Pidlisecky et al., 2007]. We also propose an inversion method using the extended Kalman filter (EKF), which incorporates information about both the evolution of the physical system being monitored and the structure of the model error covariance to update estimates of model parameters. The EKF model updates are recursive; hence, information from all previous time-steps is included in the current update.

4.2. Inversion Methods

4.2.1 Regularized, Gauss-Newton Update

In the first inversion approach, we use a modified version of the Gauss-Newton (GN) inversion algorithm presented by Podlisecky et al. [2007]. This approach is representative of the current state of the science for inverting time-lapse ERI data. The image collected at each time-step is treated as an independent “snapshot” of the subsurface. Information from the previous time-step is incorporated through a regularization term as the reference model.

In the original formulation, the algorithm uses GN updates to minimize the cost functional

$$\Phi(\mathbf{m}) = \frac{1}{2} \|\mathbf{W}_d(\mathbf{d}(\mathbf{m}) - \mathbf{d}_{obs})\|^2 + \frac{\beta}{2} \|\mathbf{W}_m(\mathbf{m} - \mathbf{m}_{ref})\|^2 \quad (1)$$

The first term in (1) minimizes the weighted misfit between data predicted using the current conductivity model, \mathbf{m} , and the observed data, \mathbf{d}_{obs} , where \mathbf{W}_d is the data weighting. This term is known as the data misfit and is used as a convergence criterion for the inversion. The second term is a weighted Tikhonov-style regularization term that incorporates prior knowledge about the solution through the reference model, \mathbf{m}_{ref} [Tikhonov and Arsenin, 1977], where β is the weighting coefficient and \mathbf{W}_m the regularization term. A preconditioned gradient solver is used at each iteration to solve for a model update, $\Delta\mathbf{m}$, that moves \mathbf{m}_{ref} toward the solution \mathbf{m}^* that minimizes (1).

We incorporate information derived from the inversion of data from previous time-steps through the regularization term in (1). The regularization term penalizes large differences between the current model estimate and the reference model, taken here to be the inverted conductivity structure from the previous time-step. Inclusion of the regularization term incorporates our knowledge that the sampling interval is short relative to the time-scale at which the water moves through the subsurface. This means there are only small variations between adjacent time-steps, so we expect the

conductivity structure at the current time-step to be close to the conductivity structure at the preceding time-step.

A modified version of the implementation by Pidlisecky et al. [2007] employs two-grids in the model update. At each iteration, a “fine” grid of 318 by 124 cells is used to calculate the Jacobian, the sensitivity of measurements to changes in model parameters, and predicted data. A “coarse” grid of 147 by 51 cells is then used to estimate model parameters. Using a fine grid improves the approximation of both the data and the Jacobian, especially near source locations. Inverting on a coarse grid limits the number of model parameters and is more representative of the information content in the data. This combination of coarse and fine grids increases the stability of the inversion process and improves the accuracy of model updates.

4.2.2 Extended Kalman Filter

Kalman [1960] first proposed a prediction-filter method for evaluating time-series data of random variables that incorporates knowledge acquired from all previous time-steps to estimate parameters at the current time-step. This method is based on iterative updates to the expression describing the evolution of the measured physical system and the measurement process. The EKF is a modified version of the original Kalman filter that is adapted to predict and filter nonlinear processes.

Within the geosciences, the EKF has been applied to time-lapse seismic data [Bayless, 1970; Eppstein, 1998; and Quan and Harris, 2008], as well as in ocean sciences [Fukumori, 1994; Verron, 1999], and in soil sciences for predicting hydrologic parameters [Eppstein and Dougherty, 1996; Parada, 2004; Ines and Mohanty, 2008; Reichle, 2008; Vrugt, 2008; and Camporese, 2009]. However, much of the methodology for the application of EKF to electrical resistivity data has been developed within the medical literature for electrical impedance tomography (EIT). The medical EIT measurement is very similar to ERI measurements, except that EIT uses a high-frequency signal, incorporating real and imaginary components of the conductivity, whereas ERI applies a DC signal and only accounts for real components of the conductivity.

Vaukhonen et al. [1998] first proposed the use of extended Kalman filters for processing time-series electrical impedance data to address temporal smearing, which can result from rapid variations in the conductivity structure with respect to the time required to collect measurements for a single image. The theory has been expanded to incorporate spatial smoothing and known internal structure, which is often required for EIT and ERI because they are ill-posed inverse problems [Kaipio, 1999; Kim et al., 2001; Kim et al., 2002; and Trigo, 2002].

An EKF approach to inversion of time-lapse ERI data has only recently been proposed [Lehikoinen et al., 2009]. The authors of that study demonstrate that the EKF is a feasible method when applied to the inversion of synthetic cross-borehole data. Using the EKF approach, the authors apply a stochastic representation of Richard's equation to simulate changes to the water saturation distribution. They then apply a petrophysical transform to create a synthetic data set of the changing conductivity structure with time and invert the data with a general EKF methodology. We present here a very similar approach to inversion of ERI data using the EKF. Our work, however, addresses significant issues not encountered in the previous study as our application involves 1) acquisition using a surface-based array of electrodes and 2) field data. A surface-based acquisition geometry results in decreased measurement sensitivity to the imaged region relative to cross-borehole data. Working with actual field data, rather than synthetic data, involves the additional complexity of uncharacterized measurement noise, as well as greater heterogeneity in the imaged field than is often used in synthetic examples.

We present a derivation of the EKF below, which closely follows the derivation presented by Kim et al. [2002] with some modifications adopted from Lehikoinen et al. [2009]. We consider a system of two stochastic processes describing the evolution of the physical system (the state equation) and the measurement process (the observation equation). These are represented by (2) and (3), respectively.

$$m_{k+1} = \mathbf{F}_k m_k + \omega_k \quad (2)$$

$$U_k = \mathbf{V}_k(e^{m_k}) + v_k, \quad (3)$$

where m_k is a vector of model elements representing the log conductivity structure at time index k , where m_k is equal to $\log(\sigma_k)$. \mathbf{F}_k is the system evolution matrix describing how the conductivity structure changes physically with time; and ω_k is the system noise. In (3), U_k represents the predicted potential measurements acquired from using the measurement process described by \mathbf{V}_k , which is a function of the conductivity structure represented by e^{m_k} , with measurement noise, v_k . The matrix \mathbf{V} contains information on the location of source and potential electrodes as well as the physics of current flow. We consider that in both the state and observation systems the noise (ω or v) is Gaussian with 0-mean and covariance matrix:

$$\Gamma_k^i = E\left[i_k i_k^T\right] \approx s_i^2 \mathbf{I}. \quad (4)$$

Here, s_i is the standard deviation of the noise, and i is either ω or v .

Linearization of (3) about the latest estimate of the state or model structure, m_k , transforms (3) into:

$$U_k = V_k\left(e^{m_{k|k-1}}\right) + J_k\left(m_{k|k-1}\right)\left(m_k - m_{k-1}\right) + \mathbf{H.O.T} + v_k. \quad (5)$$

where $\mathbf{H.O.T}$ represents higher order terms, which are small in comparison to the first order term. In the above notation the subscript $k|k-1$ represents an estimate of the parameter at time-step k based on the parameter estimate at $k-1$. This notation for predicted parameters is adopted throughout the derivation. In (5), J_k is the Jacobian matrix calculated in terms of the log conductivity. We then define a pseudo-measurement to separate out terms dependent on m_k . The pseudo-measurement has the form

$$y_k \equiv U_k - V_k\left(e^{m_{k|k-1}}\right) + J_k\left(m_{k|k-1}\right)m_{k-1} + \bar{v}_k, \quad (6)$$

where \bar{v}_k is the modified error vector incorporating v_k and higher order terms from the linearization about the current estimate of log conductivity. The modified error covariance matrix can be written

$$\bar{\Gamma}_k^v = E\left[\bar{v}_k \bar{v}_k^T\right] \approx s_{\bar{v}}^2 \mathbf{I}. \quad (7)$$

The cost functional to be minimized in estimating the conductivity model is expressed:

$$\min \frac{1}{2} \left\{ \|m_k - m_{k-1}\|_{C_{k|k-1}^{-1}} + \left\| y_k - J_k(m_{k|k-1})m_k \right\|_{\Gamma_k^{-1}} + \alpha^2 \|R(m_k - m_*)\| \right\}, \quad (8)$$

where R is a spatial regularization matrix, and m_* is a reference model. Here we apply a standard smoothness regularization on the conductivity model. We note an important distinction between (8) and (1) in that an EKF approach does not have an explicit data misfit minimization term. Unlike the GN-approach, which uses the misfit between predicted and observed data as a stopping criterion, there is no such condition on a Kalman filter update. This increases the complexity to comparing how well the two methods fit the data. Combining the last two terms of the cost functional, we can build augmented measurement vectors and matrices

$$\bar{y}_k = \begin{pmatrix} y_k \\ \alpha R m_* \end{pmatrix}, \quad H_k = \begin{pmatrix} J_k \\ \alpha R \end{pmatrix}; \quad \Gamma_k = \text{block}(\bar{\Gamma}_k^v, \mathbf{I}) \quad (9)$$

reducing the cost functional (8) to

$$\min \frac{1}{2} \left\{ \|m_k - m_{k-1}\|_{C_{k|k-1}^{-1}} + \|\bar{y}_k - H_k m_k\|_{\bar{\Gamma}_k^{-1}} \right\} \quad (10)$$

The update prediction-filter iterations can then be expressed as:

Prediction:

$$\begin{aligned} C_{k+1|k} &= F_k C_k F_k^T + \Gamma_k^\omega \\ m_{k+1|k} &= F_k m_k \end{aligned} \quad (11)$$

Filter:

$$\begin{aligned} G_k &= C_{k|k-1} H_k \left[H_k C_{k|k-1} H_k^T + \Gamma_k \right]^{-1} \\ C_k &= [\mathbf{I} - G_k H_k] C_{k|k-1} \\ m_{k|k} &= m_{k|k-1} + G_k \left[y_k - H_k m_{k|k-1} \right] \end{aligned} \quad (12)$$

The prediction equations in (11) estimate the conductivity structure and model error covariance based on data acquired up to the $k-1$ time-step. The equations in (12) then filter the predictions based on new observations. In this application, where the sampling interval is short relative to the time-scale of the imaged process, we can approximate the state evolution matrix as a simple linear process. Our linear assumption is consistent with applications from impedance tomography in which there is assumed to be little evolution in the state model between time-steps [Vaukhonen et

al., 1998; and Kim et al., 2002]. An important distinction in our application is that we consider an EKF approach to inversion of surface-ERI data. Although the acquisition geometry does not change the form of the prediction-filter equations above, it has important implications for measurement resolution of our data relative to cross-borehole applications. Furthermore, we apply the method to field data, whereas previous studies have considered only synthetic infiltration data.

4.3 Synthetic Example

4.3.1 Description of Synthetic Data Set

Because the EKF has not previously been applied to surface ERI data, we first compare the two time-lapse inversion approaches for a synthetic, scaled model of the HSRP. The goal of the synthetic example is to validate the assumption that both inversion methods produce similar estimates for time-lapse ERI data and that we are justified comparing the performance of the different approaches on field data. We also use the synthetic examples to verify that the EKF remains a feasible approach in light of reduced measurement resolution (when compared with borehole or medical tomography), a result of the acquisition geometry.

The model space for simulating infiltration is 8 m across by 5 m deep, and is a scaled and simplified model of the field experiment. The model space is divided into two horizontal layers. The transport properties in each layer are assigned to be consistent with values in the hydrostratigraphic model of the pond developed in Chapter 2. The first layer in this scaled model is 0.1 m thick, corresponding to the 0.5-m layer of fines-rich sand observed at the pond. The saturated permeability is defined to be $5\text{e-}9\text{ m}^2$ with a residual saturation of 0.2 and 35% porosity. The second layer is 7.9 m thick and represents the coarse sand of the unconfined aquifer. The saturated permeability of this layer is $5\text{e-}8\text{ m}^2$; the porosity is 40%, and the residual water content is 0.05. We incorporate a thin, fines-rich sediment layer to simulate the flow behavior based on the current understanding of structure and flow parameters. Although the low permeability of the first layer does slow infiltration, we do not observe the same infiltration rate behavior observed at the pond. This suggests we do not capture the heterogeneity and boundary flow conditions at the pond in the numerical simulation.

Flow Model

Unsaturated flow in COMSOL is modeled using Richard's equation in the form [Dogan and Hotz, 2005]:

$$\left[C(h_p) + \frac{\theta}{\phi} S_s \right] \frac{\partial h_p}{\partial t} - \nabla \cdot \left[\frac{k(h_p) \rho_w g}{\mu} \nabla (h_p + z) \right] = 0 \quad (13)$$

where is h_p is the pressure head (m); C is the specific moisture capacity; θ is the moisture content; ϕ is the porosity, and S_s is the specific storage (m^{-1}). $k(h_p)$ is the pressure-head dependent relative permeability; ρ_w is the density of water (1000 kgm^{-3}); g is the gravitational constant (9.8 ms^{-2}); and μ is the fluid viscosity (0.01 Pa-s). Using a standard van Genuchten approach to estimate the water retention or relationship between water content and pressure head, the water content can be expressed as [Van Genuchten and Nielsen, 1985]:

$$\theta(h_p) = \theta_r + \frac{\theta_s - \theta_r}{\left[1 + |\alpha h_p|^n \right]^{1-\frac{1}{n}}} \quad (14)$$

where θ_s is the saturated liquid fraction, θ_r the residual water content, α and n are empirical parameters set here to .01 and 1.5, respectively.

We solve Richard's equation for variably saturated flow using COMSOL Multiphysics, applying the Richard's equation solver from the Earth Sciences Module with a hydraulic head interpretation. In determining the numerical solution to the water saturation with time, we use a finite element mesh with 5816 elements and 3069 nodes. No flow boundary conditions are set on the edges of the model space, which are sufficiently far from the area of infiltration to introduce minimal boundary effects. From 2 to 6 m along the top boundary, which represents the base of the pond, we apply a constant head boundary of 0.5 m, a scaled average of the pond depth during diversion. Internal boundaries between layers are set to continuity of flux boundary conditions. We simulate flow behavior over 50 hours and output the water saturation distribution every hour.

Electrical Conductivity Model

Saturation values from the COMSOL model are exported to MATLAB on a regular grid of 85 by 25 cells. The exported data include regular (x,y)-grid locations and saturation values for each of the 51 time-steps. Using an Archie's Law approximation for the conductivity of the sand layers of the form [Archie, 1940]:

$$\sigma = \sigma_w \phi^m \theta^n, \quad (15)$$

we convert saturation to values of electrical conductivity. In (15) σ_w is the bulk water conductivity (0.5 Sm^{-1}); m is the cementation exponent (set equal to 1.5); and n is the saturation index (set equal to 2). The estimated conductivity structures at a number of times during the simulation are included as Figure 4.1 for reference.

Potential measurement data from the conductivity structure are simulated using the electrical resistivity forward model developed by Pidlisecky and Knight [2008]. To simulate the data we assume 125 Wenner arrays with source-potential offsets ranging from 1 to 5 electrodes. Measurements sweep across a line of 40 electrodes from 2 to 6 m at the top boundary of the model space. Synthetic data are generated for each conductivity structure created from the output hydrologic model parameters from the 51 time-steps and 1.5% Gaussian noise is added. We note that, with the acquisition geometry used, the change in observed data between images after 30 hours is less than the measurement noise. We, therefore, do not expect that inversion of these data will result in meaningful updates to the estimated conductivity structure.

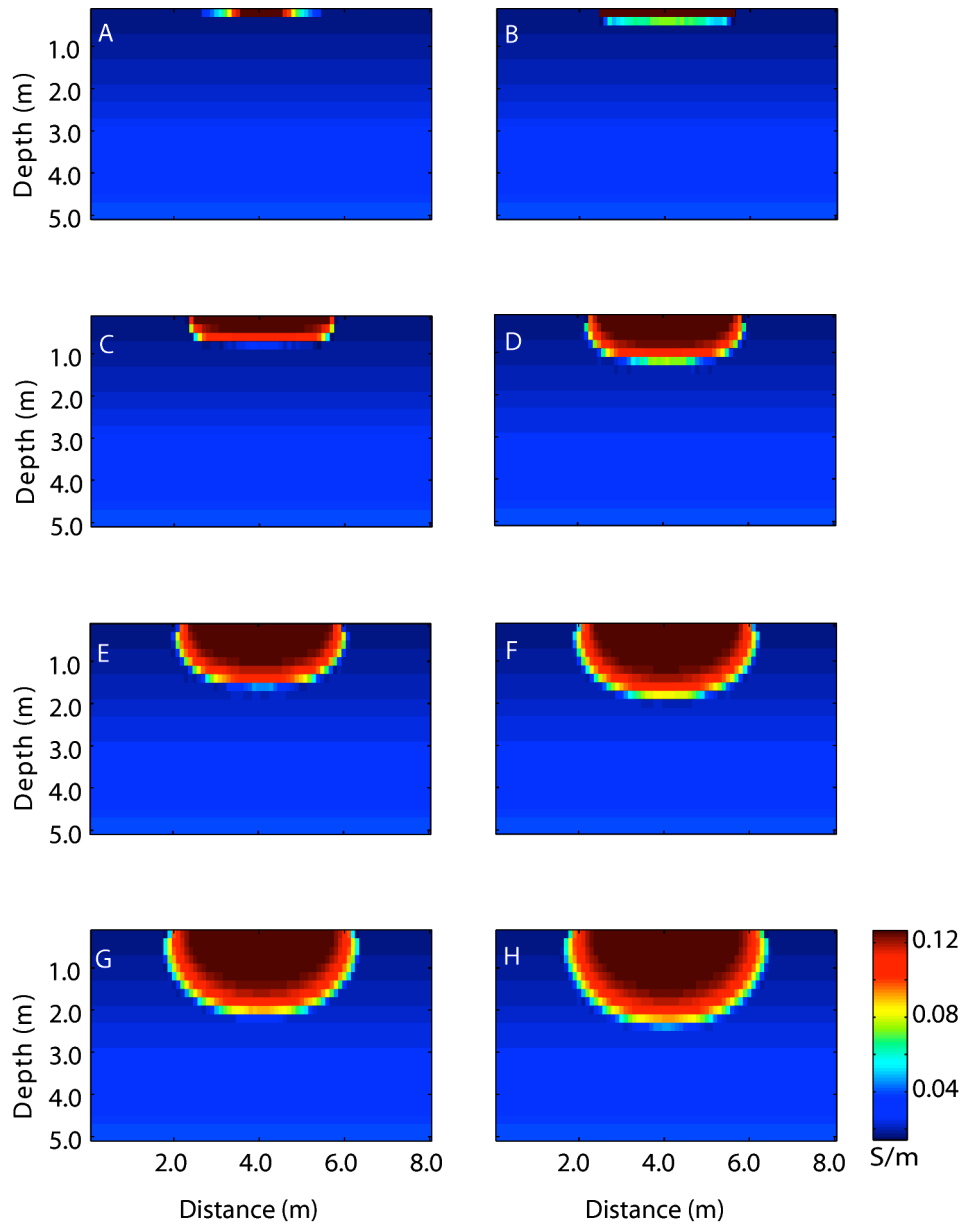


Figure 4.1. Time-lapse conductivity models created based on the numerical solution to Richard's equation for effective saturation during an infiltration experiment. Conversion to conductivity is achieved through an Archie's Law relationship. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.

4.3.2 Regularized Gauss-Newton Inversion Results

We first invert the synthetic data using the regularized Gauss-Newton (GN) algorithm. In this example data are weighted by the inverse of the standard deviation of the noise component times the inverse of observed data value plus a small value to penalize extremely small data values. Regularization terms are weighted by a smallness factor, $w_s = 1e-6$, and x- and z- smoothness weights both equal 0.01. The initial model regularization weighting term, β , is set equal to $1e-2$. This value of β is used throughout the inversion unless the decrease in the objective function for a model update is less than 10%, in which case it is decreased by one-third. Data are inverted to a data misfit tolerance of $1e-1$ with a maximum of 8 GN updates. The initial reference model is a simple layered model with 25 layers and log conductivity values from -3.6 to -4.6 spaced logarithmically from the top to the base of the model. The conductivity values are selected in this manner to incorporate prior knowledge. At the first time-step the model space is only partially saturated, reflected in the values chosen, and the imaged region is more likely to be saturated near the surface than at the base of the region, reflected in the spatial distribution of values. After the first time-step the inverted conductivity model from the previous time-step is used as the reference model at the current time-step. At the initial iteration of each time-step, if the misfit between the predicted data and the observed data is less than the data misfit tolerance, the conductivity model is not updated, and the process proceeds directly to the next time-step. This reduces the amount of noise introduced into the inverted conductivity model through the inversion process.

Results from the inversion of synthetic data using this process are shown in Figure 4.2 for the 6 times corresponding to the synthetic data shown in Figure 4.1. This plot clearly shows that after 5 hours the inverted results do a very poor job of capturing information about the magnitude and distribution of conductivities in the imaged region. Despite the use of regularization and a high misfit tolerance to help smooth the inversion results, a large amount of noise is introduced into the parameter estimates. This is clear from the plots and can also be seen in the progress of the inversion at each time-step. After the first time-step, all inversions fail to converge to

a solution below the misfit tolerance by the final permitted iteration and the line-search algorithm frequently fails to find an update capable of reducing the misfit within the first few iterations of the inversion at each time-step.

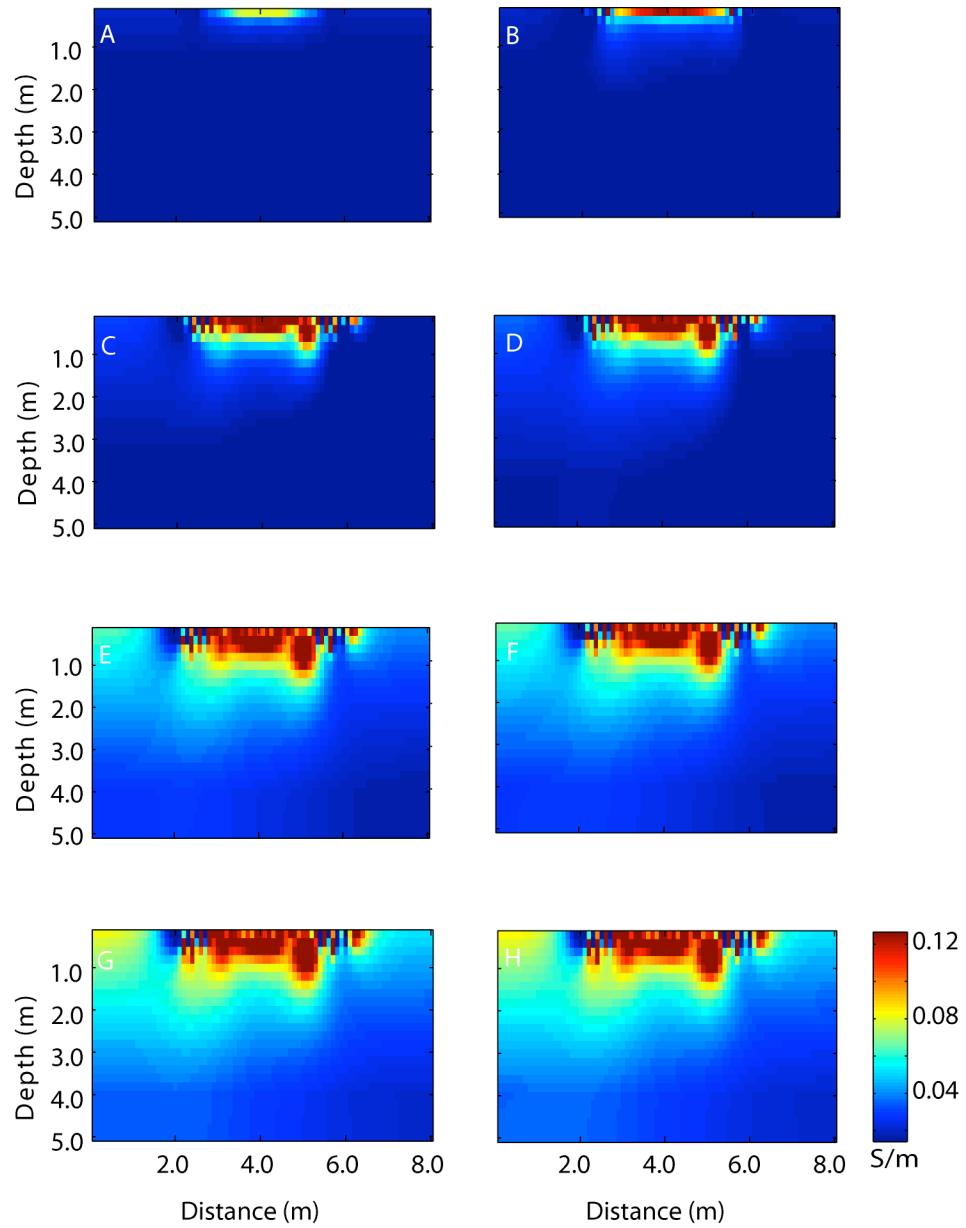


Figure 4.2. Time-lapse conductivity models inverted using regularized GN and a reference conductivity model equal to the estimated conductivity structure at the previous time-step. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.

We address this issue of failure to converge to a solution by re-inverting all data from the imaging experiment with the reference model reset to the initial model when the inversion fails to converge. If an inversion fails to converge for a given time-step, the data from that time-step are re-inverted with the initial reference model. In this case, the inversion “restart” is applied at hours 3, 5, 6, 9, and 12, while the model updates are skipped at hours 13, 14, 16-18, and 20-28 and 30-51 because the misfit between predicted and observed data at the first iteration is below the specified tolerance. After hour 12 the observed data change more slowly, evidenced by the number of time-steps in which the initial data prediction error is less than the misfit tolerance, and regularizing on the model from the previous time-step becomes more stable. These results suggest that this approach to temporal regularization of the inversion works best in situations in which the model does not change rapidly with time. Therefore, this approach will work well in situations where the process being monitored is very slow relative to the frequency with which images are acquired. When the process evolves on a time-scale similar to the data imaging, this approach is less stable.

The inverted conductivity models estimated using the modification above for the same time-steps plotted previously are shown in Figure 4.3. At early time-steps, panes A-E, these models capture much more of the structure seen in the “true” conductivity distributions. In fact, the first time-step nearly replicates the conductivity structure used to generate the data. The models also increase the size of the region of higher conductivity at approximately the same rate at which water infiltrates in the simulation. However, at later times the true size of the higher conductivity region is not fully recovered. Because of the acquisition geometry and the shape of the high conductivity region, electrodes are entirely within the high conductivity region at later time-steps, making it difficult to identify the boundary between the higher and lower conductivity regions. Consequently, the inversion attempts to account for increased conductivity from an enlarged region of higher conductivity by simply increasing conductivity values near electrodes. This causes the “noisy” results observed in panes F-H. Pane G is a replica of pane F because the misfit tolerance between predicted and

observed data is less than the specified tolerance for all images between those times. Hence, no new information about the system is obtained for those iterations. Despite the noisiness of estimated models at later time-steps, the inversion results do a much better job of capturing the conductivity structure throughout the imaging experiment. Nonetheless, each time the reference model is reset to the initial model, all information about the image region acquired up to that time-step is lost.

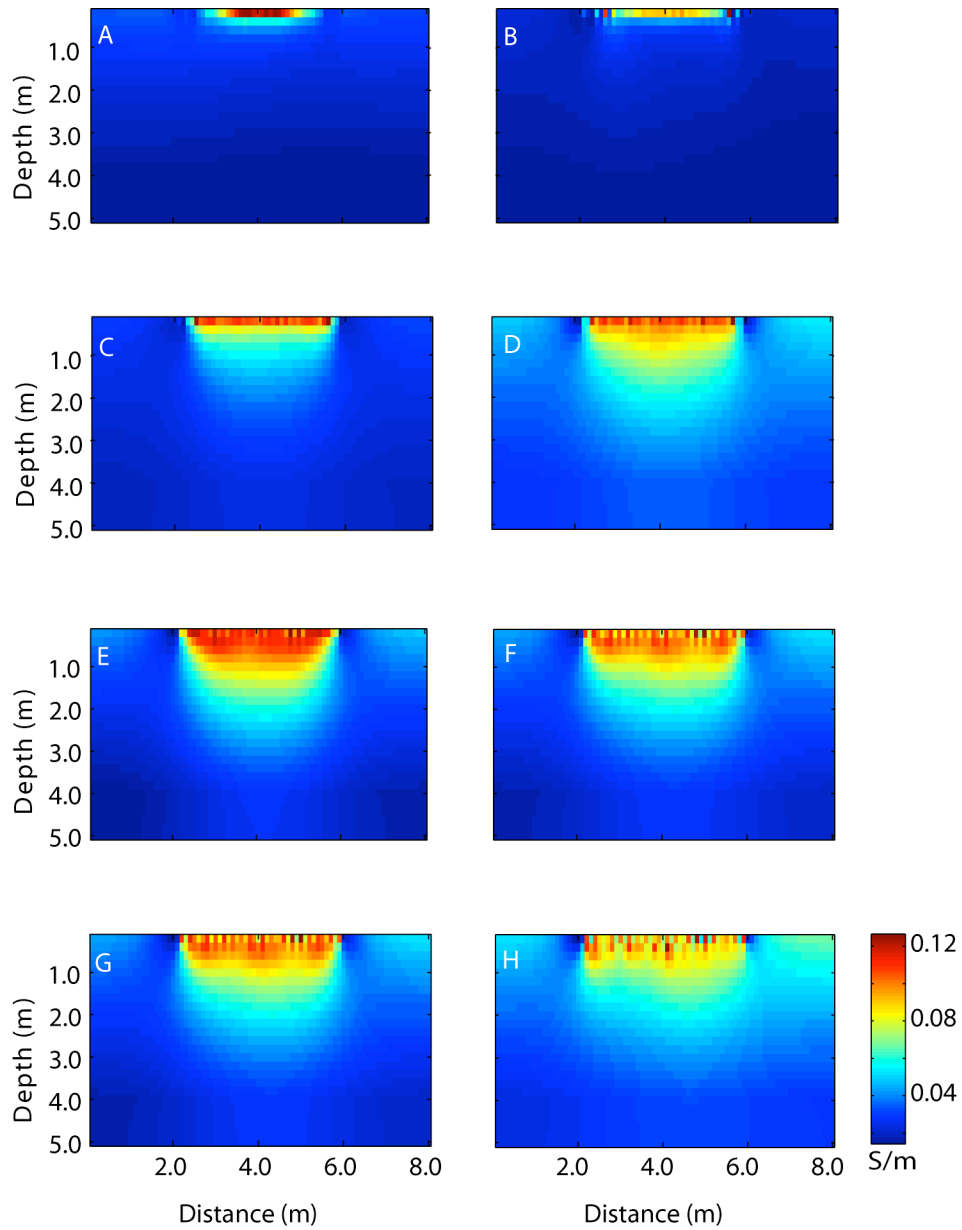


Figure 4.3. Time-lapse conductivity models inverted using regularized GN with 1.5% Gaussian noise added to the simulated data. The reference conductivity model is equal to the estimated conductivity structure at the previous time-step, but is reset to the initial reference model if the inversion fails to converge. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.

4.3.3 EKF Inversion Results

We estimate the conductivity structure at each time-step using the EKF prediction-filter recursions described above. We assume a model covariance of $1e0\mathbf{I}$; observation covariance of $3e-2\mathbf{I}$; and initial prediction error covariance matrix of the form \mathbf{I} . This choice for the form of the model error covariance matrix has little impact on the model estimates after the first few iterations [Vauhkonen et al., 1998]. We assume a standard smoothing operator for R , with weighting coefficient equal to $\sqrt{2}$. Because time-steps are too small to allow distinguishable changes in evolution of the saturated region between time-steps, we approximate \mathbf{F} as the identity matrix. This form of the state update incorporates our knowledge that the conductivity structure at any time-step should be close to the structure at the previous time-step, avoiding the computational expense of simulating the full solution of Richard's equation at each time-step and the assumption of several unknown values of hydrologic parameters. We use the same two-grid application described in the previous approach. A fine grid of 320 by 100 cells is used to calculate the predicted data and Jacobian, while a coarse grid of 80 by 25 cells is used to update model parameters.

In Figure 4.4, panes A-H, we plot the results of the EKF estimated models at the same times as the true conductivity models of Figure 4.1. As is evident from panes A-C of Figure 4.5, when filtering begins, it takes a few iterations for the estimated conductivity structure and magnitudes to become consistent with the true conductivity structure. By hour 10, pane D, however, both the shape and magnitudes of the estimated conductivity structure are similar to the true structure. The model estimates from the EKF continue to track the evolution of the saturated region quite accurately through hour 20 of the simulation, panes E-F. Past this time, the estimated conductivity models have what appears to be a reflection of the saturation plume at the base of the model region below 4.0 m depth, which eventually merges into the conductivity structure at the surface. By hour 30, pane H, the shape of the estimated structure relative has almost no resemblance to the true conductivity structure.

The delay in model response at early times has been noted in other applications of the EKF to electrical imaging; it is known as the filter lag. This lag is a function of

the variance of the data error; as the variance increases, the estimated model changes more slowly [Vaukhonen et al., 1998]. Kalman smoothers, which are applied recursively to each time-step once the entire time-series of data has been processed, have been developed to reduce this lag effect [Kaipio et al., 1998]. However, the application of these smoothing algorithms is beyond the scope of this work. At later times, panes F to H, we attribute the introduction of erroneous structure to the relative insensitivity of data to changes in the conductivity structure, a result of the acquisition geometry. Examining the relative norm of differences in data between time-steps shows that after 20 hours the change in data is less than the specified data variance. Therefore, beyond this iteration the model is largely fitting noise, which is augmented with each additional iteration. To examine the impact of repeatedly fitting data below the specified variance, we limit the EKF update to operate only at time-steps where the norm of the difference in data from the previous time-step regularized by the data norm of the previous time-step is greater than the variance of the observation error. We plot the results of implementation of the EKF with limited updates in Figure 4.5. The amount of erroneous structure in these models at all times is significantly less than when the full time series of data is considered. The results of the synthetic example clearly show that when monitoring systems that do not change on the time-scale of the sampling interval or when measurement resolution cannot distinguish changes between time-steps, attention should be paid to variations in the data measurements relative to the variance of the modeled measurement error.

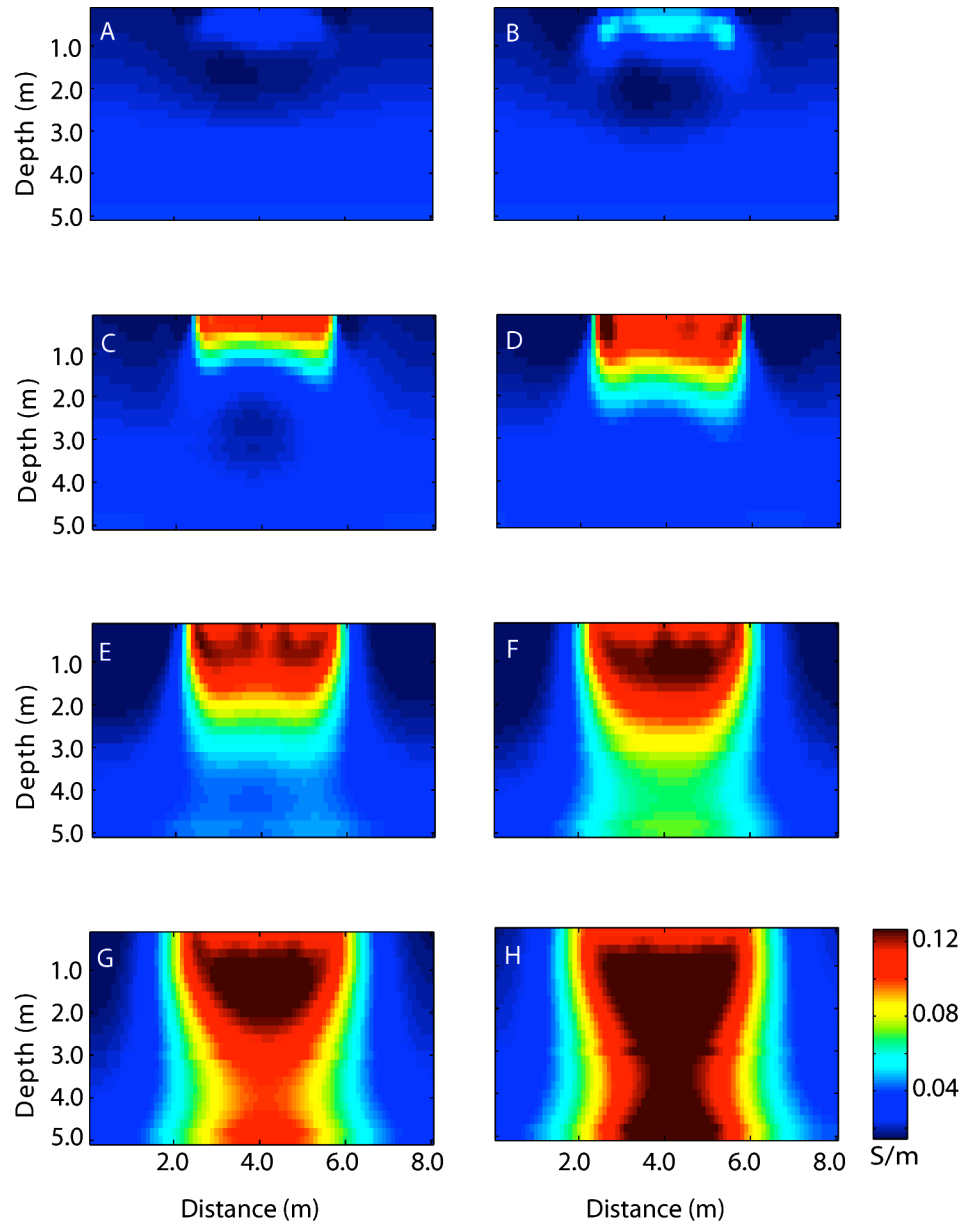


Figure 4.4. Time-lapse conductivity models inverted using EKF with 1.5% Gaussian noise added to the simulated data. A: 1 hour; B: 2 hours; C: 5 hours; D: 10 hours; E: 15 hours; F: 20 hours; G: 25 hours; H: 30 hours.

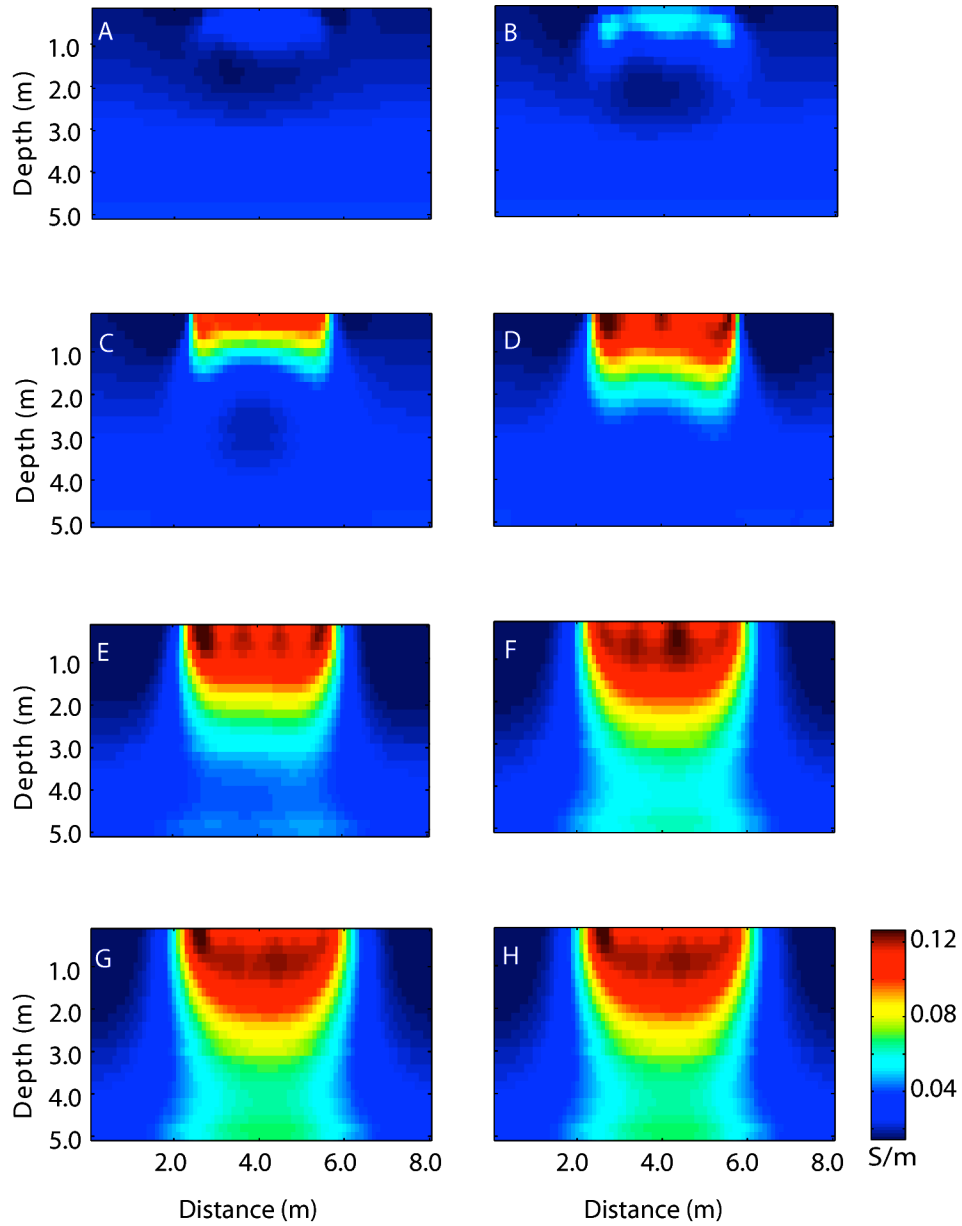


Figure 4.5. Time-lapse conductivity models inverted using EKF with 1.5% Gaussian noise added to the simulated data and model updates occurring only when the relative difference in observed data between time-steps is greater than the data variance. A: 2 hours; B: 4 hours; C: 6 hours; D: 10 hours; E: 16 hours; F: 20 hours; G: 26 hours; H: 30 hours.

4.4 Field Experiment

4.4.1 Acquisition of Field Data

Prior to the diversion period in 2008-2009, we implanted electrodes along two 20-m lines at a depth of 0.25 m below the base of the pond, which we refer to as line 1 and line 2; and along a third 15-m line also at a depth of 0.25 m. Lines 1 and 2 consisted of forty electrodes each with a regular separation of 0.5 m; line 3 had 30 electrodes, also with regular 0.5-m electrode spacing. A fourth surface line was created by joining lines 1 and 2; only every other electrode of these lines were used, resulting in 40-electrode line with meter electrode spacing. At either end of each of the first three lines, electrical resistivity probes were installed to provide detailed 1D measures of the electrical conductivity profile through time. The 1D probes had 25 electrodes with 0.1 m spacing between electrodes, except for a 0.5 m gap after the top 4 electrodes, which were left above ground to monitor bulk fluid conductivity in the pond. The remaining electrodes monitored to a depth of 2 m below ground surface. In addition to the bulk fluid conductivity, we also acquired the water depth throughout the diversion period. We use the water depth and fluid conductivity values to constrain the inversion of surface and probe electrical resistivity data. A schematic of the field lay-out is included as Figure 4.6 for reference.

The acquisition of a single image of the subsurface conductivity structure for lines 1 and 2 used 154 Wenner arrays with source-potential electrode pair offsets ranging from 0.5 m to 2.5 m. The same acquisition arrays were used for line 4 with source-potential pair offsets ranging from 1.0 to 5.0 m. Acquisition of a single image of the subsurface conductivity structure for line 3 used 108 Wenner arrays. Again source-potential electrode offsets ranged from 0.5 to 2.5 m. Measurements were collected across the lengths of the lines in all cases. For all four lines, collected images are available at 1.5-hour intervals for 3 months during the diversion period. For each 1.5 hour-interval, we have one set of data acquired from each of the surface lines, as well as 5 sets of data for each probe, collected between the acquisition of data along each surface line. A number of power surges causing incomplete or lost data

means that our time-lapse imaging experiment consists of 1936 complete data sets over approximately 18 weeks for each of the four lines.

A rain event at 13.6 days after monitoring began results in observable infiltration of water in images from all lines and probes, but did not cause accumulation of significant amounts of water in the base of the pond. Water diversion began approximately 48 days after the start of the monitoring experiment. Once the pumps were turned on there was a significant, a highly conductive water column was present throughout the remainder of the imaging experiment. Because electrodes were buried shallowly, this water column impacted the measured resistance values, and we must account for the presence of water when calculating the predicted data at each iteration for both inversion algorithms. For this study, we examine inversion results for the imaging experiment conducted on line 2. We also include the time-lapse inverted conductivity structure along probe 2, located at one end of line 2, as a means of validating inversion results. We expect to see similar results in the inversion of both surface and probe data. Unfortunately, probe 1, located on the other end of line 2 and near the pond influx, was miswired and the data could not be inverted to estimate the conductivity profile at that probe.

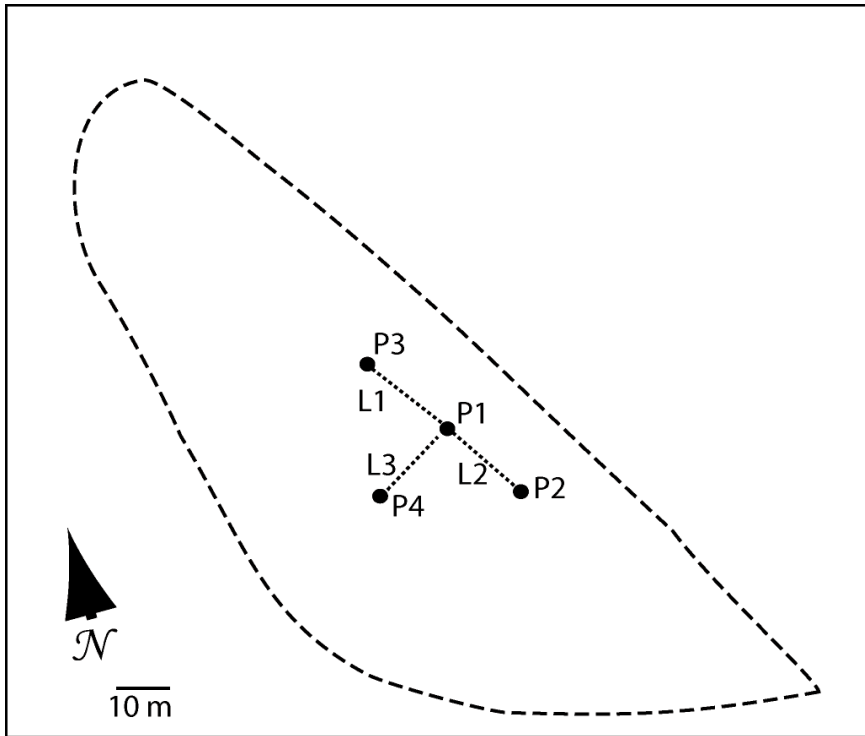


Figure 4.6. Schematic of field layout for acquisition of time-lapse ERI data.

4.4.2 Regularized Gauss-Newton Inversion Results

We first invert the field data using the modified version of the Gauss-Newton algorithm. In this implementation, the coarse grid is 147 by 51 cells with variable grid spacing from 0.25 m to 4 m in both the x- and z-directions, which incorporates padding at either lateral end and the base of the model space to minimize boundary effects. The fine grid is 318 by 124 cells, with variable spacing of 0.1 m to 4 m. The data from each time-step of the field imaging experiment are inverted to a misfit tolerance of 4.5%. The reference model, after the initial time-step, is again the estimated conductivity structure from the previous time-step. The reference model for the initial time-step is a simple, layered conductivity model. Unlike in the synthetic example, when inverting the field data we must augment the reference model at each time-step to account for the conductive water in the pond when forward modeling the potential electric field.

To calculate the predicted data at each time-step, we begin by using records of the water level in the pond, recorded every 15 minutes during the diversion period, to calculate the average water depth during a time-step. The water depth determines the thickness of a conductive layer representing the water column. The conductivity of the layer is set to the bulk conductivity measured at probe 2 during acquisition of an image. The reference conductivity model at each time-step is the concatenation of the water column and the subsurface conductivity model from the previous time-step. The augmented reference model and shifted electrode positions are used to calculate predicted data for each iteration in the inversion. Once the predicted data have been estimated, cells representing the water in the model are made inactive. Hence, we only update the estimated conductivity structure in the subsurface.

At the first iteration of each time-step, if the misfit between the predicted data and observed data is less than the prescribed misfit tolerance, the inversion process is skipped. The conductivity structure remains equal to the portion of the reference model in the subsurface. This reduces the amount of spurious noise and structure introduced into the estimated conductivity structure by reducing the likelihood that the

inversion is fitting noise in the data. If the misfit between predicted and observed data is greater than the misfit tolerance, data are inverted according to the criteria above.

A number of inverted models from different time-steps are included as Figure 4.7. The results are presented as relative differences with respect to the initial image to better show changes in the estimated structure. Probe 1 is located at 10 m offset in the x-direction, and represents the deepest part of the pond. Probe 2 is located at the -10 m offset. The inverted conductivity models clearly show wetting from a rain event at 13.6 days, (A), followed by drying back to initial subsurface conditions prior to the diversion of water to the pond, shown in (B). After diversion begins, shown in (C), the inverted models again show what appear to be alternating periods of wetting and drying in the deeper sediments, (D-F) even though the hydraulic head in the pond remains relatively constant until monitoring ends. However, examining the bulk fluid conductivity as a function of time, shown in Figure 4.8, clearly indicates the increase in subsurface conductivity at late times is a result of the increased bulk water conductivity. At late time during the imaging experiment, the temperature in the pond water increases with the ambient temperature. As temperature is positively correlated with conductivity, the bulk fluid conductivity increases. Hence, though the sediments do not fully saturate, as the conductivity in the deeper sediments is always less than that in the top half meter, water continuously infiltrates the subsurface. These inversion results indicate a complex infiltration process that cannot be explained by a simple model of clogging in the fines-rich layer as initially assumed.

Although the amount of spurious noise and structure introduced into the model is limited by only processing data with variation above the misfit tolerance, the inversion results still become noisy after a large number of time-steps using this method. Resetting the reference model to the initial reference model after ~200 inversions and proceeding with the method as described above does help to limit this effect. However, comparing the results of the inversion from the surface data to the time-lapse conductivity profile for probe 2, computed by A. Pidlisecky, shows that, as the inversion continues, some structure is still added into the estimated conductivity structure. The conductivity profile at probe 2 relative to the initial conductivity

profile, as a function of time, is included in Figure 4.9. Times coinciding with the time-steps shown in Figure 4.7 are labeled on the plot. At early times the conductivity structure in the 2D model in the region of probe 2 agrees well with the recorded conductivity profile, showing the same wetting and drying trends. At late times, however, captured in (F), the 2D relative conductivity map shows increasing conductivity near probe 2 that is not evident in the 1D profile, suggesting the 2D section shows more structure than may actually exist. This additional structure is initially interpreted as an artifact of the inversion. It may be possible to suppress these artifacts by restarting the inversion after fewer iterations.

The observed behavior of the time-lapse inverted conductivity models highlights some of the challenges of using this method. First, the weighting of the reference model relative to the data misfit must be set such that there is an appropriate balance of informing the inversion through previous information and allowing variations in data to drive the inversion process. Determining the correct weights to use can be difficult in a field setting where the true subsurface conductivity structure is unknown. The second challenge is limiting the introduction of artifacts into the model estimates from the inversion process. Even using a high misfit tolerance and smoothing the inversion through regularization, it is not straightforward to identify when the inversion has become “too noisy” and when the need to reset the reference conductivity model arises. Finally, in resetting the reference conductivity model back to the initial reference model all information gained about the subsurface conductivity structure is lost to the inversion. Resetting the reference conductivity model reduces noise and improves convergence of the algorithm, but is contrary to the concept of informed imaging.

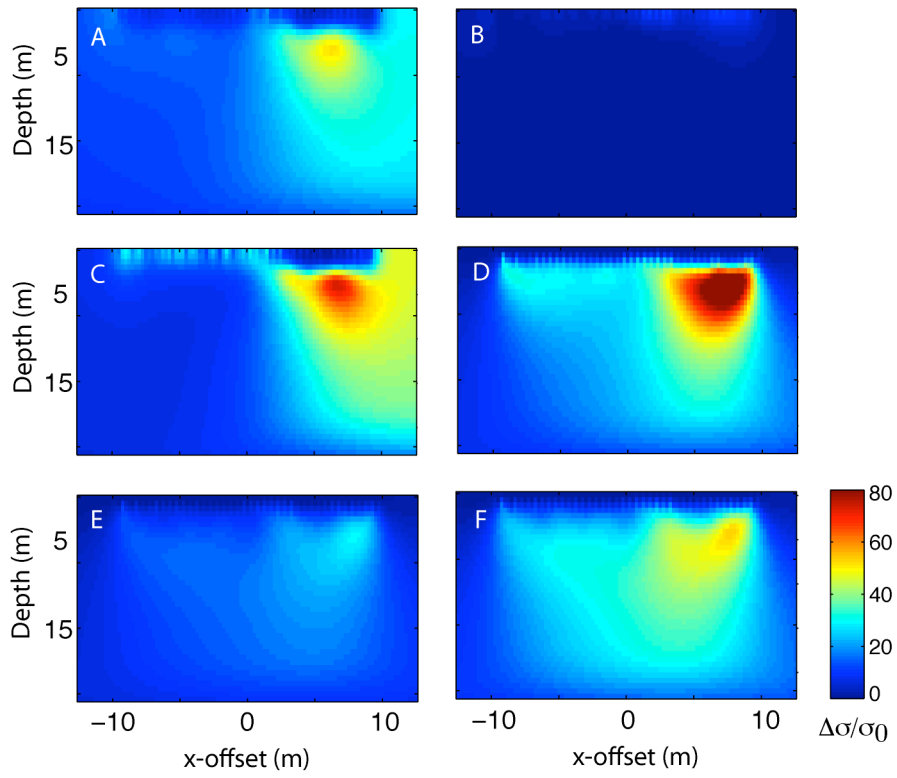


Figure 4.7. Six time-lapse relative difference conductivity structures with respect to the initial conductivity model inverted using GN-algorithm. A: 13.6 days (rain event); B: 22.8 days; C: 48.4 days (diversion begins); D: 50.6 days; E: 75.2 days; F: 112.6 days.

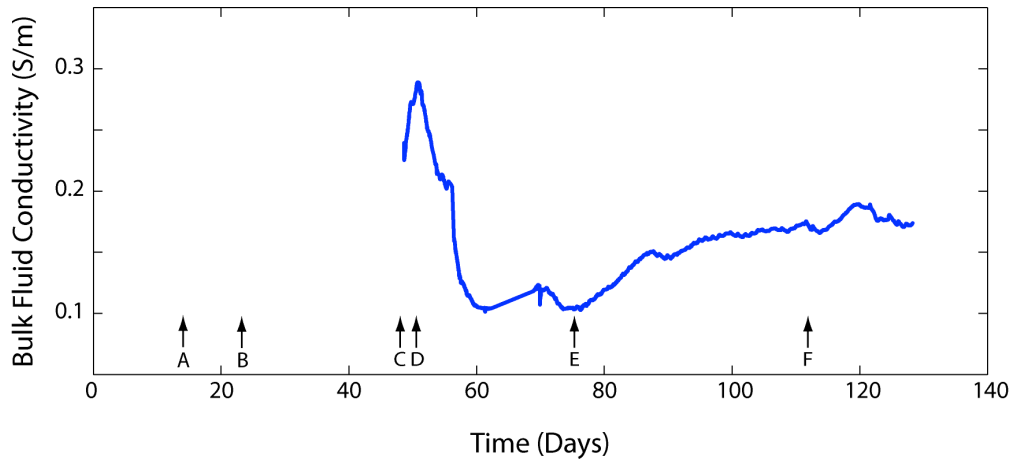


Figure 4.8. Bulk fluid conductivity in S/m as a function of time after monitoring begins. Times corresponding to the time-steps of inverted 2D conductivity models are labeled.

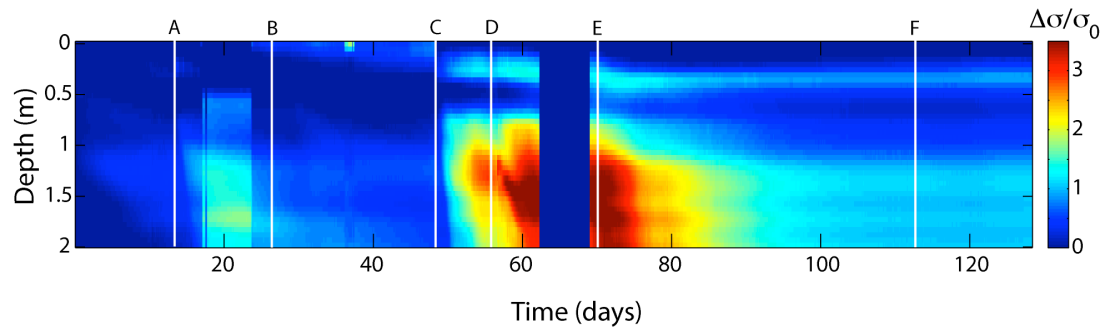


Figure 4.9. Time-lapse relative difference conductivity profile with respect to the initial conductivity profile at probe 2. Times corresponding to the time-steps of inverted 2D conductivity models are labeled.

4.4.3 EKF Results

We invert the same time-lapse field data using the Extended Kalman filter. As before, the model covariance matrix has the form $\beta\mathbf{I}$, where β takes a value of 2 in this case, and the matrix is again time-invariant. We also use the same expression for the initial prediction error covariance matrix. However, we do not assume the observation error covariance of $\gamma\mathbf{I}$ is static, but update the value of γ at every time-step based on the acquired data. The initial estimate of γ is equal to 1% of the difference between the maximum and minimum observed resistance values. This is the same determination of γ applied by Lehtikoinen et al. [2009] for a time-invariant estimate of the observation error covariance. At each time-step, γ is updated based on a weighted combination of the value from the previous time-step and the difference in data observed at the current time-step. Using a time-dependent update, γ ranges from approximately 0.1 to 0.4 throughout the imaging experiment, and tends to be higher at early time-steps. The value tends to stabilize at ~ 0.3 as the number of time-steps increases. The same smoothing operator R used in the synthetic example is applied here with a weighting of $\sqrt{2}$. Because the frequency of data acquisition is still relatively high relative to the rate at which the physical system evolves, we again approximate \mathbf{F} as the identity matrix. We again use two grids when calculating the model update. The fine grid is 320 by 120 cells with equal x- and z-dimensions equal to 0.125 m, and the coarse grid is 80 by 30 cells with 0.5 m x- and z-discretization.

Acquiring reasonable results for the predicted data requires that we account for the presence of the water column above the electrodes using the same process described in the previous section for the GN-algorithm. While the conductive water layer is incorporated into the calculation of the predicted data and the Jacobian, the model update is only applied to conductivity values in the subsurface; it does not update parameter estimates within the water.

We also incorporate a conditional model update based on the relative difference in observed data between time-steps to reduce the likelihood of updating the conductivity model to fit data noise. In cases where there is less than a 1% change

in observed data from the last model update to the current time-step, no model update is performed. However, we still incorporate information obtained at each time-step into updating the estimate of the observation error.

EKF inversion results for the same six time-steps plotted for the GN-algorithm conductivity model estimates are included in Figure 4.10, (A-F). As expected from the synthetic results, the inverted conductivity models using the EKF are very similar to those obtained using the GN-algorithm. In (A), the rain event creates a region of higher conductivity around 7 m offset in the x-direction, near the deepest part of the pond where rainwater is most likely to pool. However, the EKF inversion shows a more pronounced region of higher conductivity at roughly -5 m offset, as well. This region is only suggested in the corresponding GN-inverted image. Both inversion approaches show similar drying patterns in (B); the high conductivity region associated with infiltrated water quickly disappears. In (C) and (D), similar wetting behavior is observed when the pumps are turned on and the pond fills. However, the EKF model does not show the wetting immediately, a result of the lag time noted for this method that was also seen in the synthetic example. Furthermore, EKF models once again show a greater increase in conductivity at negative offsets than are observed in GN-results. The EKF conductivity models show the same drying behavior in (E), represented by a decrease in the relative conductivity contrast, followed by increased conductivity in (F), which corresponds to a higher conductivity contrast. Interestingly, the increased fluid conductivity is not evident in the 1D profile, though it clearly appears in the estimated 2D conductivity sections using both inversion approaches.

Comparison of the inverted models at each time-step reveals that models estimated with EKF are coarser than those from GN-updates, a result of coarser gridding required to maintain stability of the model update. One disadvantage to an EKF approach is that the update can become very unstable as the number of model elements increases. Studies within the medical literature have increased model resolution by creating regions of conductivity to decrease the number of elements updated [Trigo et al., 2004]. However, this requires assumptions about the shapes and

locations of structures within the imaged region, which are largely unknown in earth imaging. Although, the EKF results show slightly poorer resolution, they do have the advantage of being more consistent with the principles of informed imaging and better suited to automatic processing. Because the model update is recursive and based on both the system evolution and errors in the observation process, the estimated conductivity structure at each time-step is dependent on information acquired at all previous time-steps. Unlike in the GN-algorithm, in which all prior information is lost each time we reset the reference model, the EKF approach does not need to be restarted and all prior information is integral to the model update.

An EKF approach is also flexible in terms of incorporating a conductive water column into the data prediction step of the algorithm. An alternative approach for applying the EKF to conductivity data given known model structure is given in Kim et al. [2002]. Their approach incorporates structure into the full EKF update by defining regions of assigned conductivity values within the model space that are not updated at each time-step, but remain part of the model space. Though their approach was not applied here, further studies into comparisons of estimates obtained with both approaches for incorporating the presence of a highly conductive water layer is an extension of this work as it applies to applications of EKF to field data.

Furthermore, the EKF approach has distinct advantages in estimating model uncertainty. The estimate of the predicted model error covariance matrix is updated at each time-step, and shows the evolution of model uncertainty associated with each element in the model space. Looking at the diagonal elements of this matrix for the model estimated at 112.6 days, shown in Figure 4.11, gives an immediate approximation of the uncertainty associated with the estimated conductivity values of model elements, where high model error variances are associated with greater uncertainty. While the model error covariance matrix is dependent on our assumptions about model and observation errors, it is also driven by the error in predicted and observed data at each time-step. Hence, by tracking the variances of model elements through time, we can obtain a measure of model uncertainty as an integral part of the estimation process.

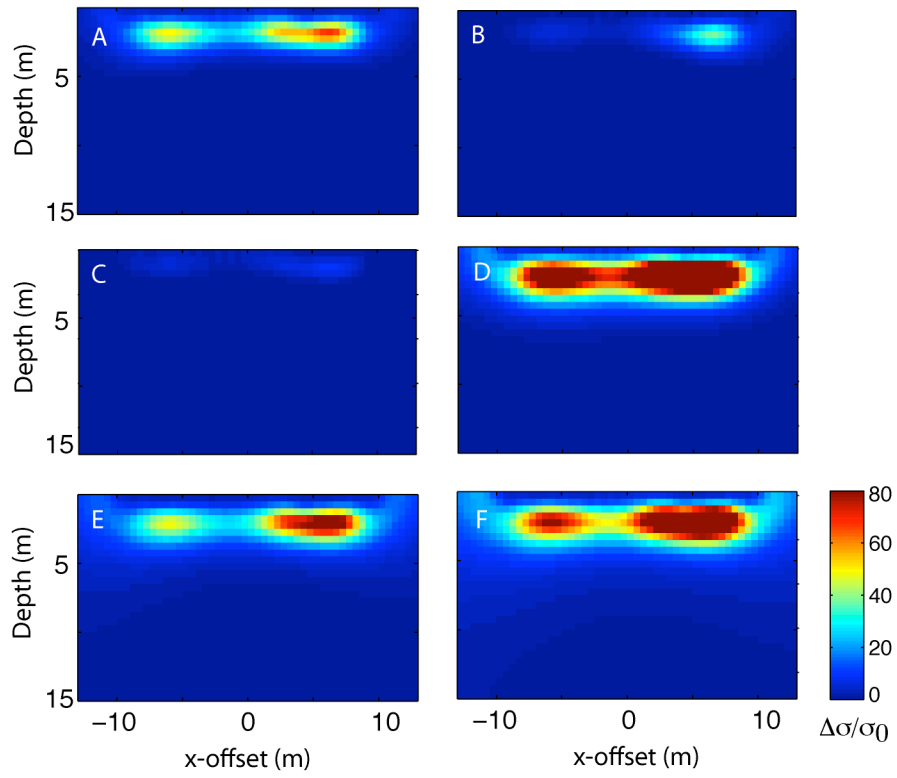


Figure 4.10. Six time-lapse relative difference conductivity structures with respect to the initial conductivity model inverted using an EKF algorithm. A: 13.6 days (rain event); B: 22.8 days; C: 48.4 days (diversion begins); D: 50.6 days; E: 75.2 days; F: 112.6 days.

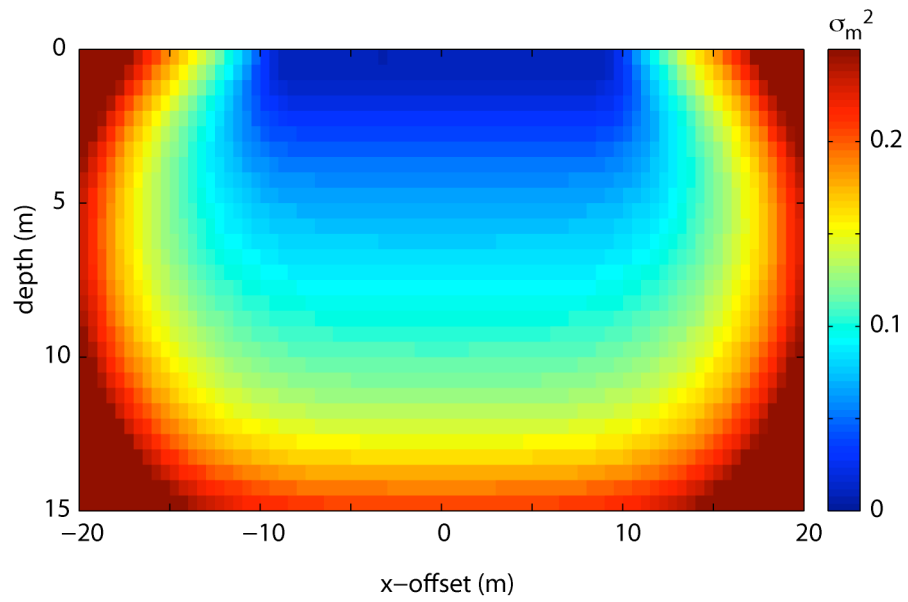


Figure 4.11. Model error variances associated with elements for the estimated conductivity structure at 112.6 days, shown in (F) of Figure 4.11.

4.5 Summary and Conclusions

We present two approaches for inversion of surface ERI monitoring data collected during infiltration events at a recharge pond in northern California. The first approach represents the current state of the science for inversion of time-lapse ERI data: each time-step is treated as an independent image and prior information is incorporated through regularization as a reference model. The second approach uses an extended Kalman filter, which concurrently models evolution of the monitored physical system and the observation process. Hence, an imaging experiment can be treated as a single time-series signal. Using a synthetic example, we validate that an EKF approach is a feasible method for inverting surface ERI monitoring data. We then apply both approaches to a field imaging experiment and demonstrate that similar results can be obtained with both methods. We also introduce update criteria for the EKF. This constrained implementation improves estimates for physical systems that evolve slowly with respect to the sampling interval or for experiments with lower data sensitivity, as is encountered with surface ERI measurements, that do not reflect small changes to the system. Our results demonstrate that EKF is a valid method for inverting surface ERI data collected in the field and that it provides advantages over current approaches. The most significant advantage from an informed imaging framework is that EKF integrates prior information from all previous time-steps into the model update. Furthermore, the EKF provides an estimate of model uncertainty at each time-step as an integral part of the inversion process.

Having demonstrated the feasibility of applying the EKF to surface ERI field data, we anticipate many opportunities for a new paradigm in ERI monitoring using this approach to data inversion. Because it is designed for time-series data with an evolving physical system, the EKF does not require current assumptions about systems remaining static during data acquisition. Consequently, monitoring can move to a continuous process in which the arrays used to acquire data can change at each time-step. We can, therefore, remove strict limitations on the number of measurements made during each time-step in an imaging experiment. Instead, we can monitor the system continuously, adjusting the number of arrays used at each time-

step to reflect the rate at which the physical system changes. The option of dividing an imaging experiment into time-steps of varying length improves our ability to automate the imaging process, diversifying the types of monitoring experiments in which we can use ERI to study subsurface processes.

Implementation of an EKF inversion of time-lapse ERI data provides several advantages over current approaches through estimates of the uncertainty at each time-step and the incorporation of evolution of the physical system into the estimation process. The method also presents several opportunities in creating a new paradigm for monitoring experiments using ERI methods. Further research is required to develop automated, continuous imaging methods that can be processed in an informed imaging framework using an EKF approach to incorporate all prior information into the inversion of ERI data.

CHAPTER 5. THE IMPACT OF GEOPHYSICAL MEASUREMENT SUPPORT IN QUANTIFYING THE CORRELATION STRUCTURE OF ANISOTROPIC SYSTEMS

5.1 Introduction

Fundamental to the concept of informed imaging is the integration of data collected with a number of methods from different disciplines that have a range of measurement scales. Combining hydrologic, geologic and geophysical data is critical to building a comprehensive *a priori* model of the subsurface to inform and guide the experimental design process. Furthermore, incorporating prior knowledge into data-processing leads to more meaningful and useful interpretations. The concept of data integration is by no means original and has already been widely used within near-surface geophysics to regularize inversions [Hyndman et al., 1994; Kowalsky et al., 2005], calibrate models [Singha and Gorelick, 2006], and improve interpretations [Day-Lewis, et al., 2005].

An example of a data integration problem from near-surface hydrology is the description of the spatial and temporal variations in soil water content to develop models of hydrologic processes. An accurate, spatially- and temporally-extensive characterization of this property and dependent processes has a large impact on the value of models used to describe surface-subsurface water interactions. Because it governs the flux of water across the land surface, soil moisture has a significant role in modeling weather, developing agricultural management practices, and describing the ecosystems and biodiversity in a region. Consequently, a recent area of focus in vadose zone hydrology research is developing methods and measurements to describe spatial and temporal variations in soil moisture at the field and catchment scales [Vereecken et al., 1998]. This has led to an increase in the types of data collected to measure or approximate the soil water content [Robinson et al., 2008].

In addition to the traditional types of soil moisture measurements which result in point measurements, such as TDR or capacitance probes, non-invasive methods such as ground-penetrating radar [Eppstein and Dougherty, 1998b; Huisman et al., 2001], electrical resistivity imaging [Michot et al., 2003; Robinson et al., 2008], and electromagnetic induction [Sheets and Hendricks, 1995; Corwin and Lesch, 2005] are being used to provide spatially continuous estimates of soil water content in transverse or depth planes. Furthermore, remote sensing and wireless sensing networks are being used to provide soil water information at the catchment scale with high temporal resolution [see e.g. Cardell-Oliver et al., 2005; Choi and Jacobs, 2007; Njoku and Entekhabi, 1996]. However, as Vereecken et al. [2008] describe, the measurements described above occur at a number of scales, and integrating the information collected at different spatial resolutions to describe the spatial variability of soil moisture across a field or catchment requires translating data into a common framework.

Several studies have sought to describe the spatial and temporal structure of soil water content [Grayson et al., 1997; Rodriguez-Iturbe et al., 1995; Western and Grayson, 1998]. However, there is a complex relation between measurement scale and the resulting estimated correlation structure of a property. This complex relationship has been widely documented for conventional hydrologic measurements. [e.g. Cushman, 1986; Gelhar, 1993; Beckie, 1996; Corwin et al., 2006; Blöschl and Sivapalan, 1995]. Gelhar [1986] developed analytical expressions for the covariance of a variable averaged over a given depth for a field that is isotropic in the orthogonal plane. Western et al. [1999] expanded on these expressions by decomposing the measurement scale into three components, termed the scale triplet. The components of the scale triplet are: spacing, the linear separation of measurements; support, the area or volume over which a measurement is averaged; and extent, the total areal or volumetric coverage of measurements. Western et al. [1999] used the notion of the scale triplet to derive analytical expressions for the relationship between increasing measurement support and the apparent correlation length. The derivation assumes that increasing measurement support can be treated as a simple linear filter (i.e. averaging several measurements over a square area) and solves for the variance within the

support by the method of Rodriguez-Iturbe and Mejia [1974], which assumes a stationary, isotropic process. Skøien and Blöschl [2006a,b] used the analytical relationships to investigate bias in the mean, variance, and integral length, a measure of the correlation length, of samples and applied the results to adjust monitoring networks to be consistent with the scale of properties of interest.

Similar approaches have been used when attempting to integrate geophysical and hydrologic data to understand the structure of soil water content [Daily et al., 1992; Rea and Knight, 1998]. However, as noted by Knight et al., [2004] decreasing the resolution of ground penetrating radar data in the vertical direction impacts estimates of apparent correlation length in the horizontal direction. This impact is independent of the horizontal range and the size of the Fresnel zone, indicating the estimated correlation structure in the horizontal direction depends on the vertical measurement resolution. Because measurement resolution is inversely proportional to measurement support, the results of this study imply that the introduction of anisotropy into the measurement scale, in the form of directionally-varying measurement support dimensions, or into the parameter field, in the form of anisotropic distributions of a property like water content, may lead to bias in the apparent correlation length that is not accounted for in existing analytical descriptions.

The focus of this study is to investigate numerically how the relationship between measurement support and the apparent correlation length deviates from analytical predictions when the assumptions of a stationary, isotropic process and square support are violated. We initially verify our model with equidimensional increases in measurement support on an isotropic model. We then consider sampling scenarios often encountered when using remote sensing or geophysical methods to acquire data over a 2D region of Earth's surface or a 2D cross-section of the subsurface. To do this we first violate the assumption of equidimensional measurement support when sampling a synthetic, isotropic parameter field, allowing the orthogonal dimensions defining support of the measurement to increase independently and be different in the two directions. We then consider an anisotropic parameter field (defined by a maximum and a minimum correlation length, λ_x and λ_y),

calculating the error in the apparent correlation length in both the x- and y-directions when measurement support increases by block-averaging. Finally, we examine non-equidimensional increases in measurement support in sampling an anisotropic parameter field. The support dimensions in the x- and y-directions are referred to as S_x and S_y , respectively. In this final scenario, we need to account for the independent affects of changes in S_x and S_y on the apparent correlation length in both the x- and y-directions. The results demonstrate that errors in the apparent correlation length change with increasing complexity in the measurement support and sampled parameter field.

5.2 Geostatistical Analysis

We adopt a geostatistical framework for estimating the correlation structure of a synthetic parameter field. It is assumed the hydrologic system of interest is a random field that is spatially correlated over some defined distance. Using the GSLIB software package [Deutsch and Journel, 1998], it is possible to calculate one-dimensional experimental semivariograms. The semivariogram is a measure of how similar the parameter field is at increasing spatial separations. The one-dimensional semivariogram can be described as [Journel and Huijbregts, 1978]:

$$\gamma(\mathbf{h}) = \frac{1}{2N} \sum_1^N [z(\mathbf{x}_i + \mathbf{h}) - z(\mathbf{x}_i)]^2. \quad (1)$$

In (1), \mathbf{h} is the lag, or the vector of separation between the points at which the parameter field, z , is evaluated. N is the number of data pairs separated by \mathbf{h} over which summation occurs. For a data set with a given correlation structure, γ increases with \mathbf{h} at small offsets and asymptotically approaches the sill as the lag becomes large. The sill is the value at which γ , the calculated variance, is equal to the variance of the samples, and the range is defined to be the value of \mathbf{h} at which the sill is reached. In this study directional semivariograms are calculated, meaning \mathbf{h} is specified in either the horizontal or the vertical direction. Following Journel and Huijbregts [1978], the lag is limited to one-half the extent of the data set, ensuring a sufficient number of data pairs are used to calculate a reliable γ -value.

Analytical models are fit to the experimental semivariograms, and the model coefficients are used to describe the correlation structure of the system. The exponential semivariogram is the model most commonly used to fit experimental semivariograms in stochastic hydrology because this model is consistent with a first-order Markov process, a simple assumption about the variability of a random field [Webster and Oliver, 2001]. The exponential model is expressed as follows:

$$\gamma(\mathbf{h}) = C \left(1 - e^{-\frac{3\mathbf{h}}{a}} \right), \quad (2)$$
$$\gamma(0) = 0;$$

where C is the sill, and a is the effective range. Because the exponential model approaches the sill asymptotically, the effective range is taken as the lag, \mathbf{h} , at which γ reaches 95% of the sill. The correlation length, λ , is equal to one-third the effective range and is the average separation over which parameter values may be considered spatially correlated. The expression in (2) describes a one-dimensional variogram model. Hence, in a two-dimensional parameter field, it is possible to calculate directional variograms by solving (2) in terms of $\gamma_x(\mathbf{h})$ and $\gamma_y(\mathbf{h})$, where the lag is specified in the x- or y-direction, which are related to λ_x and λ_y , the correlation lengths in the x- and y-directions, respectively.

5.3 Numerical Models

Seven numerical models are analyzed to determine the effect of increasing the measurement support dimension in the x- and y-directions on estimates of λ_x and λ_y from an exponential semivariogram. The models include 3 isotropic and 4 anisotropic synthetic parameter fields. In all anisotropic models, λ_x is in the direction of maximum correlation, and λ_y in the direction of minimum correlation. Each model is defined on a 100 m x 100 m grid with initial cell dimensions $dx = dy = 0.1$ m, resulting in 1000 cells in each direction. A summary of the key parameters from each of the 7 models is provided as Table 1, and examples of an isotropic parameter field and two anisotropic parameter fields showing increasing anisotropy are included as Figure 5.1 for reference.

Model	λ_x (m)	λ_y (m)	λ_{xapp_i} (m)	λ_{yapp_i} (m)
1	1.0	1.0	0.91	0.92
2	4.0	4.0	3.47	3.39
3	7.0	7.0	5.91	5.58
4	2.0	0.3	1.82	0.23
5	4.0	0.3	3.67	0.29
6	8.0	0.3	7.20	0.29
7	2.0	1.0	1.79	0.91

Table 5.1 Defined correlation length and initial estimates of the correlation length for the seven synthetic parameter fields analyzed.

For each modeled correlation structure 25 realizations of the parameter field are generated. In generating the 25 realizations of each model, the same assumption used in Skøien and Blöschl [2006a] are applied: 1) fields are stationary, or the correlation length is independent of location; 2) the univariate distribution of the field is normal with defined mean and variance; and 3) an exponential variogram adequately represents the spatial variation. Furthermore, it is assumed the spatial correlation structure is cyclic in nature, or that the underlying geologic process is periodic, and there are at least three equally-spaced measurements. This final

assumption follows from the idea of sampling at least at the Nyquist frequency used in signal processing literature and permits the use of twice the lag distance as a starting value of the range in calculating the exponential semivariogram model.

We then simulate the measurement process for a number of support sizes, both square and non-equidimensional, to construct experimental semivariograms in the x- and y-directions. We consider the measurement process as a filter that produces an average of parameter values over the measurement support [Cushman, 1986; Beekie, 1996]. Starting with the initial set of realizations for each model, our measurement simply involves retrieving the parameter value in a single cell. For the initial case $S_x = S_y = dx = dy = 0.1$ m. For block-averaging, we treat our measurement process as a very simple filter and increase the support by averaging values of the initial parameter field in both the x- and y-directions. This averaging creates blocks of 2^2 cells, 3^2 cells, etc., up to 15^2 cells. For non-equidimensional supports we first increase the support in the x-direction by averaging values only in the x-direction, creating averaged-cells out of 2 cells, 3 cells, etc., up to 40 cells. The parameter value in each block or averaged-cell is the average of the values in all the included cells. Averaging of the initial parameter field is then repeated in the same manner in the y-direction. A schematic of the averaging process for block-averaging and non-equidimensional supports is included as Figure 5.2.

The non-equidimensional measurement support is varied from 0.1m x 0.1m to 4.0m x 0.1m (or 0.1m x 4.0m) in 0.1m increments along the averaging direction. As cells are averaged, the measurement location is shifted to the center of the averaged cell block. This has the additional effect of incrementally increasing the spacing of the measurement from 0.1m to 1.5m for block averaging or 0.1m to 4.0m in the direction of averaging for non-equidimensional increases in measurement support. When the increase in measurement support is non-equidimensional, the spacing does not change in the direction perpendicular to averaging. We allow the support and spacing to increase concurrently because previous studies have shown that allowing samples to overlap (i.e. increasing the support without increasing the spacing as described above)

introduces exceptionally large errors into the apparent correlation length [Skøien and Blöschl, 2006a].

For each realization of a given model, experimental semivariograms are calculated in the x- and y- directions at all measurement supports as a function of the lag, \mathbf{h} . This results in 30 semivariograms for each realization under block-averaging (15 variograms in the x-direction and 15 variograms in the y-direction) and 80 semivariograms for each realization with non-equidimensional supports. An exponential semivariogram model is then fit to each of the experimental variograms. The model is calculated numerically using the *lsqnonlin* function in MATLAB to estimate the value of the sill, C , and the range, a , in (2); and the apparent correlation length is calculated from the range. The apparent correlation lengths in the x- and y- directions calculated for the initial parameter fields using this method are included in the last two columns of Table 5.1. The *lsqnonlin* optimization function uses an interior reflective Newton method to determine the parameter values that minimize the misfit between the model and the experimental data points assuming bounds on the possible parameter values are given [Coleman and Li, 1994]. The determined correlation lengths for each support are averaged over the 25 realizations. Thus, we obtain representative estimates of the apparent correlation length in the x- and y- direction for each increase in the x-dimension of support as well as each increase in the y-dimension.

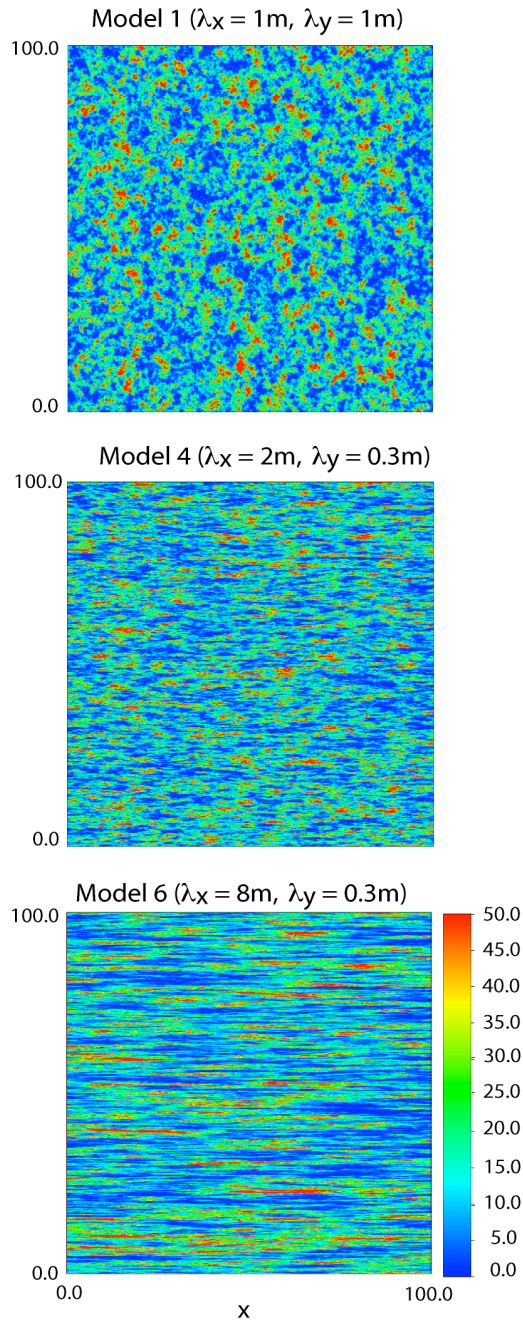


Figure 5.1. Examples of initial parameter fields simulated for three model conditions. Model 1 is an isotropic model, while Models 4 and 6 are anisotropic models showing increased anisotropy.

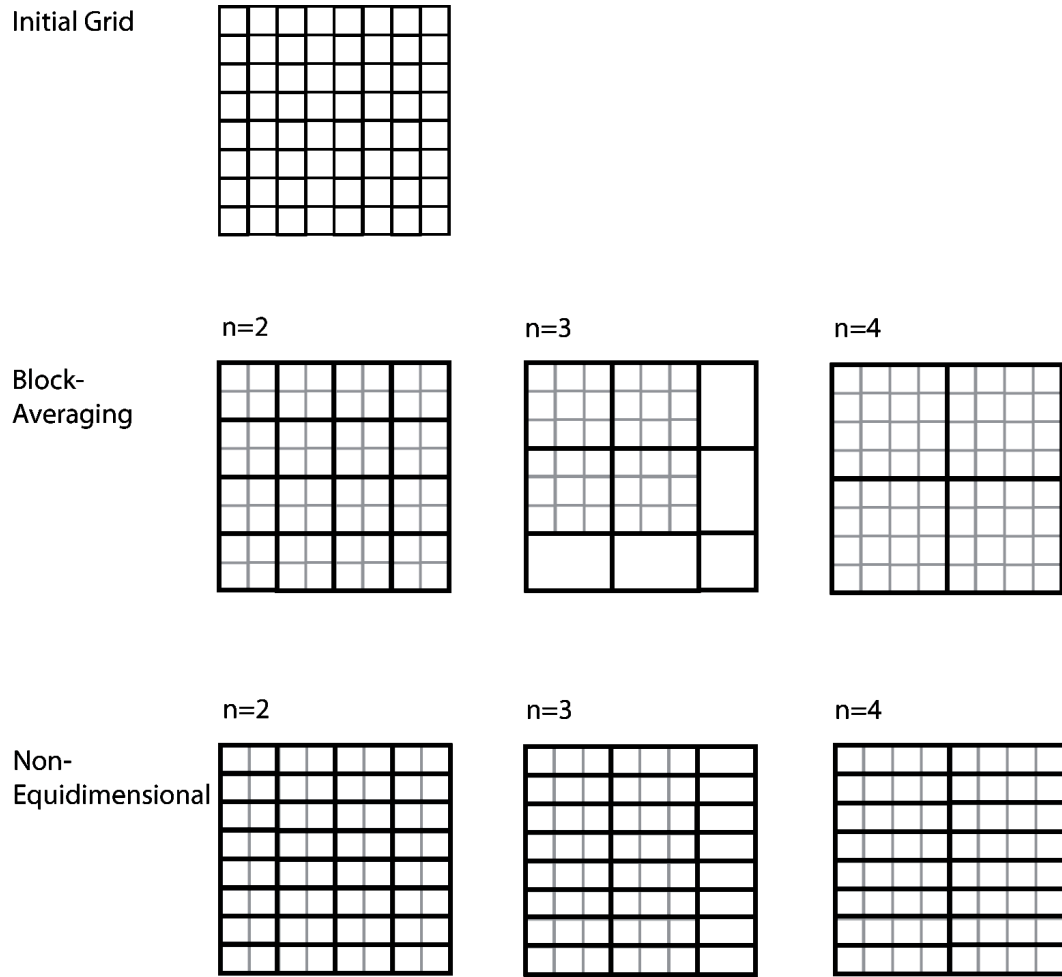


Figure 5.2. Schematic of the numerical averaging experiment where n is the number of averaged cells. Only non-equidimensional increases in the x-direction are shown. White blocks indicate cell values not included in the variogram calculation for the estimated correlation length for a given averaging length.

5.4 Results and Discussion

The geostatistical analyses of the 25 realizations for the 7 initial models ($S_x = S_y = 0.1\text{m}$) yield apparent correlation lengths that are less than the “true” correlation lengths used in generating the parameter fields in all cases. These initial estimates are included in Table 1 as $\lambda_{x\text{app } i}$ and $\lambda_{y\text{app } i}$. The underestimation of the correlation length of the initial model is also observed in a study by Skøien and Blöschl [2006b] when estimating the correlation length as a function of support for a large number of samples. We use the values of apparent correlation length estimated for the initial model to normalize the apparent correlation lengths with increasing measurement support. Normalizing the apparent correlation lengths in this manner and plotting as a function of the total support, defined as $S = (S_x S_y)^{1/2}$, normalized by the initial apparent correlation length is consistent with the definitions use by Western and Blöschl [1999], allowing us to compare our results to their analytical model. We refer to the increase in the apparent correlation lengths with increasing measurement support as the error response curve.

In Figure 5.3 we plot the results of block-averaging for the isotropic fields, Model 1 ($\lambda_x = \lambda_y = 1\text{m}$), Model 2 ($\lambda_x = \lambda_y = 4\text{m}$) and Model 3 ($\lambda_x = \lambda_y = 7\text{m}$). Error response curves for the apparent correlation lengths calculated from both the x- and y-direction variograms are shown, and they track well with each other and the analytical curve. The initial ratio of apparent correlation length to true correlation length is always lower for the numerical results than predicted, a result of normalizing by the initial value, which makes this value always equal to 1. At all but the initial support the numerical results fall very close to the predicted error for all models and supports, verifying that the numerical model is consistent with the analytical model.

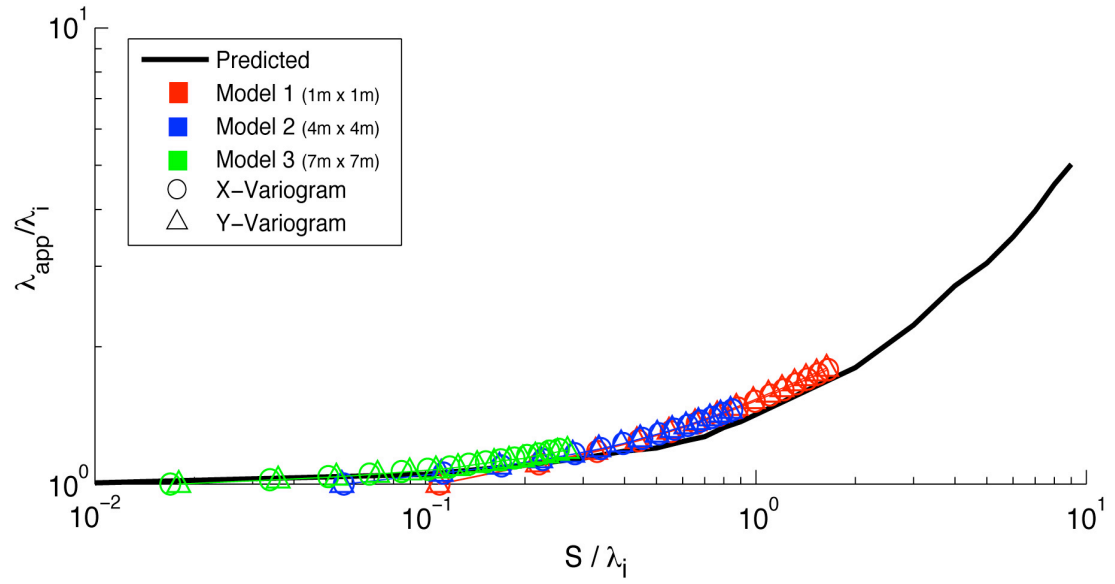


Figure 5.3. Effect of block averaging of measurement support for isotropic parameter fields on the apparent correlation length from 1D x- and y-direction variograms.

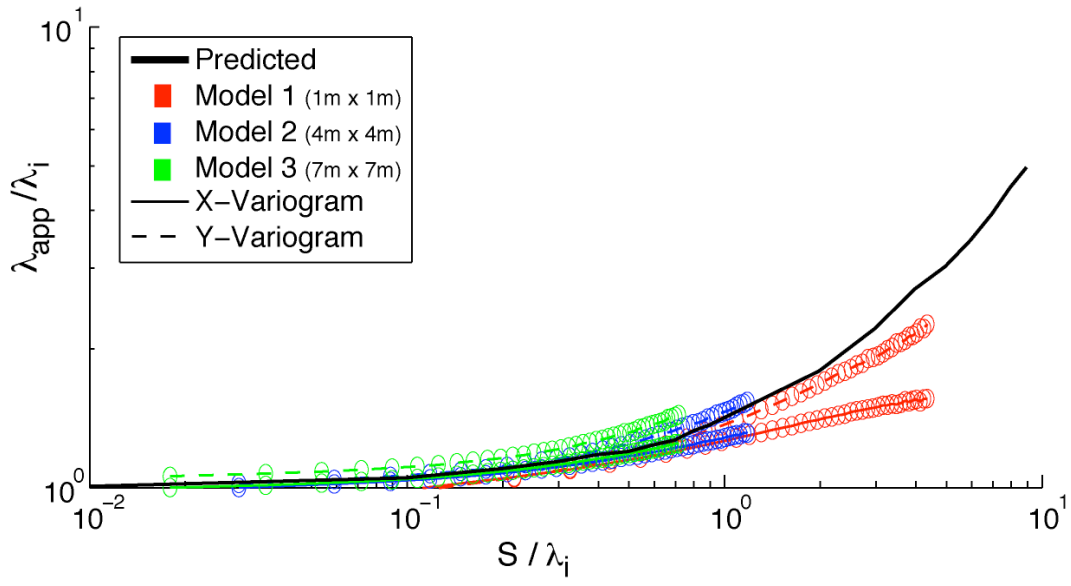


Figure 5.4. Apparent correlation length from variograms calculated in the x- and y-directions for x-directional averaging of measurement support in isotropic parameter fields.

In Figure 5.4 we compare the analytical result to those obtained with non-equidimensional increases in measurement support for isotropic parameter fields. Again, results from variograms calculated in both the x- and y-directions are shown, though we only show results for increasing support dimension in the x-direction. Increasing support dimension in the y-direction produced similar results. The errors in apparent correlation lengths demonstrate that when the support increases only along one dimension, the slope of the error response curve depends on the direction in which the variogram is calculated. The error response curves for variograms calculated in the y-direction are concave up while the curve for those calculated in the x-direction are nearly linear. Moreover, this increase in the error is greater when the variogram is calculated in the y-direction for all cases. As the values in the x-direction are averaged over a greater number of cells, the measurement becomes so biased in that direction there is little information about variation in the y-direction. Even for isotropic parameter fields the differences are noticeable between the x- and y-directions. Furthermore, the slope of the error response curve decreases as the true correlation length increases, evidenced by the greater amount of error in apparent correlation length for Model 1 than for Models 2 and 3.

We also consider the alternate situation: the increases in measurement support are equidimensional, but the parameter fields are anisotropic. These results are shown in Figure 5.5 and include Models 4 through 7. In this case the shape of the error response curves is different for the variograms in the x- and y-directions, though both deviate greatly from the analytical prediction. Model 7, the closest to an isotropic field, shows results that are much closer to the analytical model than the other anisotropic models, but still has similar trends to those observed for all anisotropic models. The shape of the error-response curves for apparent correlation lengths calculated from both the x- and y-direction variograms for Model 7 are concave up, but the slope of the curve is steeper when apparent correlation length is calculated in the y-direction than when calculated in the x-direction. Furthermore, both curves are less concave than predicted by the analytical model. In contrast with the results for Model 7, the slope of the error response curve for apparent correlation lengths

calculated in the x-direction for Models 4 through 6 decreases with increasing support; it is concave down rather than concave up. This flattening of the error response curve for the x-direction is more pronounced for Model 6 than for Models 4 or 5, and the magnitude of error is also smaller for Model 6. These results indicate that the error in the estimated correlation length for anisotropic fields is sensitive to the degree of anisotropy in the parameter field. This is also evidenced by the similarity of Model 7 to the analytical response, and the greater deviation from the analytical curve of Models 4 through 6, which are increasingly anisotropic. By contrast, the error response curves for apparent correlation lengths in the y-direction for these same models show slightly increasing slopes and a greater total magnitude in error. Furthermore, the error response curves for estimates in the y-direction of Models 4 through 6, which have the same underlying correlation length in the y-direction, fall nearly on top of one another. These observations indicate the increase of error in the apparent correlation length is also sensitive to the underlying correlation length in the direction of increasing support dimension. As the underlying correlation length increases, the error in the apparent correlation length grows at a slower rate with increasing support.

The final scenario considered is of anisotropic parameter fields with non-equidimensional changes in measurement support. The results for this scenario are given in Figure 5.6, panes A-B, plotting the error response curves for all anisotropic models, as well as the error response curves of Model 3 for comparison to the response in an isotropic parameter field.

When both increases in measurement support and calculation of the apparent correlation length are in the x-direction, shown in Figure 5.6(A), the shape of the error response curves is similar to the shape for the isotropic model. However, the increase in the error is less for Model 6 than the isotropic model, whereas it is greater Models 4, 5, and 7. Furthermore, the results for Models 4 and 7 track very closely, suggesting the slopes are inversely related to the correlation length in the direction of increasing support dimension. This is because, as the support increases, shorter layers are averaged into single units more rapidly than layers with longer correlation lengths.

Similar patterns are observed in Figure 5.6(D), which shows the results for increasing the measurement support dimension and calculation of the apparent correlation length in the y-direction. However, the error in the apparent correlation length is much greater relative to the underlying correlation length in this case, further supporting the interpretation of an inverse relationship between the error in the apparent correlation length and the correlation length in the direction of increasing measurement support.

For increasing measurement support in the x-direction, the direction of maximum correlation, and calculation of the apparent correlation length in the y-direction, shown in Figure 5.6(B), the slope is flatter than when both increases in measurement support and calculation of the variogram occur in the x-direction. The greater error in the latter case is attributed to changing the underlying physical state in the direction in which the variogram is calculated through the averaging that occurs in the measurement process. When the support dimension increases in a direction that is perpendicular to the direction of variogram calculation, the underlying physical state being measured is better preserved, and the error in the apparent correlation length is smaller at each support, though still greater than for block-averaging. The shape of the error-response curve is similar for all models, including the isotropic Model 3. However, the total magnitude of the relative error is smallest for Model 6 and increases for Models 5 and 4, respectively, which have decreasing anisotropy. Furthermore, Models 4 and 7 have nearly identical error response curves despite having different true correlation lengths in the y-direction. These results suggest the relative error is highly dependent on the correlation length in the direction of increasing measurement support. Moreover, the relative error decreases as the correlation length in the direction of increasing measurement support increases relative to the correlation length in the direction in which we calculate the apparent correlation length.

The above interpretation is supported when measurement support increases in the y-direction and the apparent correlation length is calculated in the x-direction. However, in this case the shape of the error response curves is very different than for the isotropic model. The curves for Models 4 through 7 initially have very steep

slopes, but show a significant decrease in slope as the support becomes large relative to the underlying correlation length in the direction of averaging. This is attributed to merging of layers as the support increases in the y-direction. There is a diminishing effect on the estimated correlation length as the support continues to increase because averaging is perpendicular in direction to the calculation of the apparent correlation length. This limits the impact of the measurement process on the underlying property structure being approximated. Because the correlation length in the x-direction is so much greater than in the y-direction, the limited effect with increasing support causes the slope to plateau. This flattening of the error response curve occurs first for Model 6, which has the greatest anisotropy, while there is only a slight flattening for Model 7, the least anisotropic model. Also, in contrast to the previous scenario, Model 7 has a smaller total magnitude of error than the other anisotropic models, and the response curve for Models 4 and 7 are very different even though both models have the same underlying correlation length in the direction in which the apparent correlation length is calculated. This further supports the interpretation that the relative error depends on the ratio of the underlying correlation length in the direction of increasing support dimension to the underlying correlation length in the direction of calculation of the apparent correlation length. As the ratio decreases, the amount of error in the apparent correlation length increases. This plot also clearly shows a decrease in the slope of the error response curve with increasing support, a phenomenon not seen in the isotropic models under any averaging scheme, but apparent in block-averaging increases in measurement support when sampling anisotropic fields.

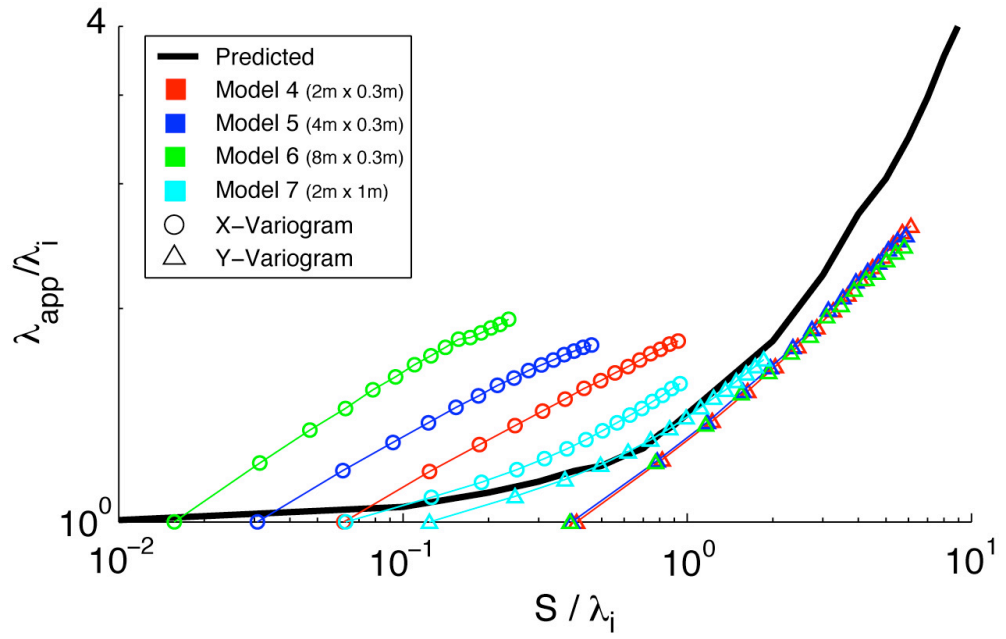


Figure 5.5. Apparent correlation length from variograms calculated in the x- and y-directions for block averaging of measurement support in anisotropic parameter fields.

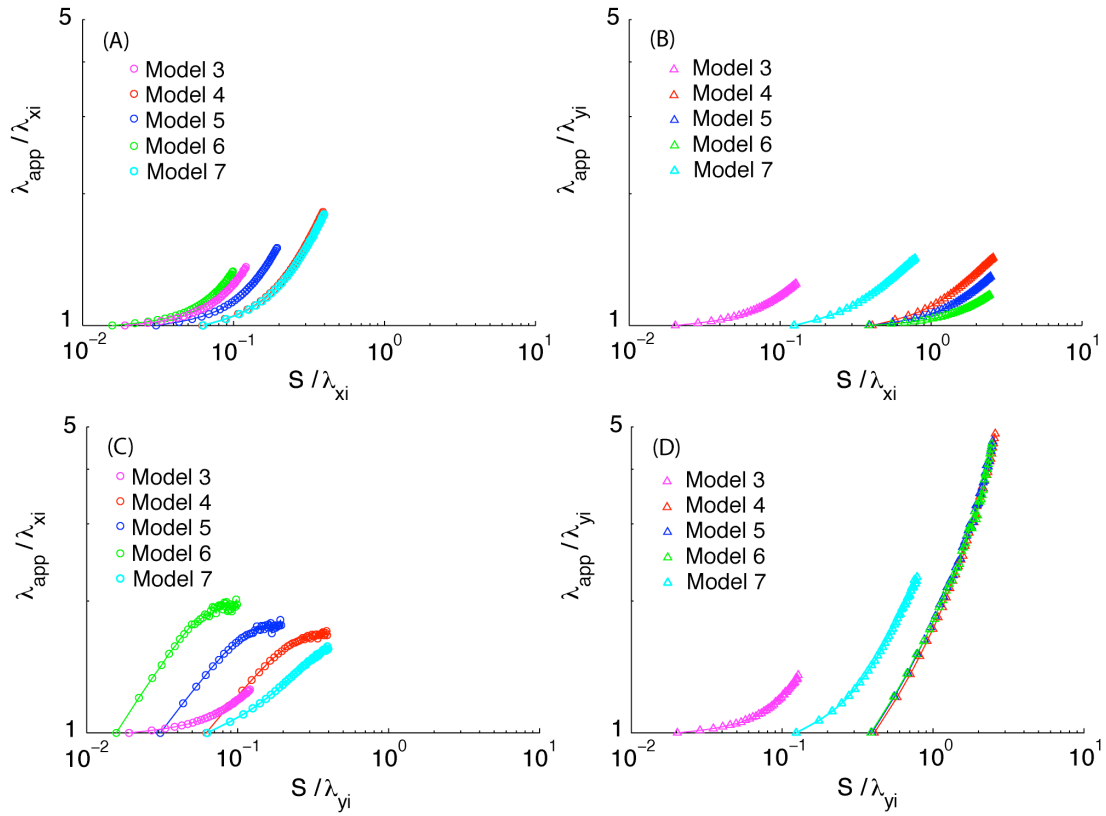


Figure 5.6. Estimated correlation lengths for directional averaging of measurement support in anisotropic parameter fields. Results are plotted as a function of the normalized measurement support. Circles denote variogram calculation in the x-direction; triangles denote variogram calculation the y-direction. The support dimension in panes A and B increase in the x-direction, while the support dimension in panes C and D increase in the y-direction.

5.5 Conclusions

Combining measurements from different methods that have a range of measurement scales to quantify the correlation structure of hydrologic properties is a challenge in developing surface or subsurface models to study Earth processes. While a number of studies have examined the complex relationship between measurement scale and the apparent correlation to develop analytical models, we have demonstrated these results are only valid for an isotropic parameter field where increasing measurement support can be treated by block averaging. However, most earth processes are anisotropic and several remote sensing techniques used to characterize relevant properties have non-equidimensional measurement supports

When anisotropy is introduced into the parameter field of interest or when measurement support changes with direction, the error in the apparent correlation length becomes sensitive to the degree of anisotropy in the sampled property as well as the correlation length in the direction of increasing measurement support. When the increasing measurement support dimension is in the same direction as calculated apparent correlation length the errors are consistently higher than when the two processes are perpendicular to one another or when the measurement support is equidimensional. We attribute this to the continuous alteration of the field in the direction in which we calculate the apparent correlation length by the measurement process. When the two processes are perpendicular, the amount of error in the apparent correlation length at a given measurement support increases as the ratio of the correlation length in the direction of increasing measurement support to the correlation length in the direction of calculated apparent correlation length decreases. Hence, as the parameter field becomes more anisotropic the impact on the error in apparent correlation length becomes more extreme, and there is relatively less error for increasing support dimension in the direction of maximum correlation than in the direction of minimum correlation. Furthermore, the true correlation length in the direction of increasing support dimension is inversely related to the rate at which the apparent correlation increases with support, and there is a diminishing effect on the

error in the apparent correlation length at large supports relative to the underlying correlation length.

The results of this study indicate that current analytical expressions for relating an increase in measurement support to the apparent correlation length are incapable of predicting the behavior of more complicated, anisotropic systems. Thus, when integrating measurements collected with different techniques over a range of scales, especially when considering anisotropic property fields such as exist at the field or catchment scales, it is important to consider the impact of the true correlation length on the rate at which error the apparent correlation length increases with support. Further studies are required to develop quantitative relationships between error in the apparent correlation length and measurement support for anisotropic systems and non-equidimensional measurement supports. These relationships are necessary to quantify uncertainty in estimates of the correlation structure from remote sensing or geophysical images so they may be more applicable to hydrologic models of near-surface processes. Knowledge of the measurement support and any *a priori* information on the surface or subsurface correlation should be accounted for in the design of geophysical surveys, the interpretation of data, and in integration of parameter estimates to hydrogeologic models.

CHAPTER 6. CONCLUSIONS

Understanding the spatial distribution of hydrologic properties that govern flow and transport behavior in the vadose zone is critical to developing effective groundwater management strategies. Obtaining the spatial distribution of properties with traditional measurements can be challenging as they often provide very accurate measurements but sparse spatial coverage. Because it is sensitive to water content in the vadose zone, electrical resistivity imaging (ERI) is a geophysical method that is quite useful for creating spatially exhaustive maps of the subsurface properties. ERI provides a continuous map of the electrical conductivity structure, which can be related to the subsurface distribution of hydrologic properties. However, to produce more detailed maps of variations in conductivity structure that better inform hydrologic models requires that we refine methods for monitoring subsurface hydrologic processes with ERI.

The goal of this dissertation is to set forth a framework for informed imaging: the incorporation of all existing information about an experiment into the acquisition, inversion and interpretation of ERI data, to provide better quality maps of conductivity. This is achieved through two proposed methodologies: the first is an algorithm for informed experimental design, and the second implements an application-specific inversion algorithm based on the extended Kalman filter (EKF). The proposed informed experimental design algorithm selects acquisition arrays for ERI experiments that produce greater data sensitivity to an imaged region than can be achieved using standard arrays. I also demonstrate the effective implementation of an application of the EKF for inverting ERI data. The EKF has not previously been applied to processing field data, though it has been shown to be feasible for synthetic and laboratory data. As a precursor to the application of informed experimental design, I demonstrate the development of an *a priori* hydrostratigraphic model integrating geologic, hydrologic, and geophysical data using advanced visualization software. Furthermore, one of the complexities of combining different types of data

with a range of measurement scales into a single hydrologic model, errors in the estimation of the correlation structure, is examined for anisotropic systems and measurements with non-equidimensional support, as are often encountered with geophysical or remote sensing measurements.

The primary contributions of this dissertation are as follows:

- Creation of an *a priori* model of the hydrostratigraphic structure of the Harkins Slough Recharge Pond in an advanced visualization environment. The model combines geologic, hydrologic, and geophysical data into a systematic framework for describing properties of interest for modeling flow behavior. Simulated flow behavior based on the hydrostratigraphic model reveals gaps in understanding and highlights regions of interest for ERI monitoring.
- Development and application of a methodology for informed experimental design. The methodology is consistent with the concept of informed imaging in that it uses the created *a priori* model to guide the selection of acquisition arrays with greater data sensitivity than can be achieved using standard arrays. The proposed experimental design process also has the flexibility to incorporate advances in data acquisition systems, such as multichannel instruments.
- Acquisition of unique ERI data that monitors infiltration beneath a recharge pond with high temporal resolution. The combination of 1D and 2D ERI imaging experiments with short sampling intervals over several months produces a data set that images infiltration processes with more detail than is usually achieved in the field. This data set provides a unique opportunity for advancing the way in which ERI is incorporated into hydrologic monitoring.
- Implementation of an EKF algorithm for inverting time-lapse ERI field data. This is the first time the EKF has been successfully applied to ERI data outside

of synthetic examples, providing a new approach to ERI monitoring. Rather than regarding each time-step as an independent image, the EKF models both the measurement process and the evolution of the monitored physical system to integrate all information acquired from previous time-steps into the current model update.

- Identification of relationships between measurement support and apparent correlation length in anisotropic systems and for non-equidimensional measurements that are not explained by existing analytical models. Analysis of the impact of anisotropy and non-equidimensional measurement on errors in the apparent correlation length are important when combining remote-sensing or geophysical data with conventional hydrologic measurements into a single model of their hydrologic properties that control flow and transport behavior.

The goal of informed imaging is to provide a systematic framework for improving the quality of acquired ERI data and the analysis of that data. The application of informed imaging to field experiments opens many areas of research to exploration. As noted in Chapter 4, only the Wenner data from one surface line was inverted using EKF. However, EKF is capable of incorporating all other data sets into the estimation of the conductivity structure. Unlike most applied time-lapse inversion algorithms that treat each data set as a “snapshot,” the EKF is ideal for time-series signals because it considers both the observation process and an evolving physical system. Further research into merging data sets that use different acquisition arrays into a single time-series signal for continuous monitoring could greatly improve estimates of the conductivity structure through time. For example, if time-lapse data were to be collected with different acquisition arrays designed to image different depth regions throughout time, they could be consolidated into a single data set using an EKF inversion scheme. This is particularly relevant to future work as we move toward better integration of geophysical and hydrological data.

Building on the concept of continuous monitoring, there is an opportunity to expand the informed experimental design methodology to remove temporal constraints that arise from the “snapshot” approach to acquisition of images. The method can be expanded to accommodate a continuous imaging framework while still identifying an imaging experiment with greater data sensitivity to the imaged region. Furthermore, there is an opportunity to adapt the presented method to real-time experimental design, in which acquisition arrays are continuously updated as information about the subsurface structure is obtained. In terms of increasing efficiency of informed experimental design, the proposed design process still relies on a source pair, but the algorithm could be expanded to allow for multiple source and sink electrodes capable of generating “focused” currents. The question of how to identify the most informative combinations of source electrodes, and the relative magnitudes of the current sources and sinks, remains an open avenue of research.

Beyond further development of the proposed methods, I see two unique opportunities for expanding the use of ERI in monitoring vadose-zone processes. The first is identifying the value of time-lapse ERI data to structural and flow models of the subsurface. Chapter 2 identifies regions of interest for monitoring to improve understanding of the infiltration behavior at the Harkins Slough Recharge Pond. How best to incorporate the acquired monitoring data into the hydrostratigraphic model or flow simulation remains unresolved. Likewise, how to evaluate the contribution of the data to generating improved simulations of flow, ones that better reproduce observed behavior, requires further investigation. The second opportunity is similar to the first but relates to the data acquired at the pond. While it was shown that the unique data set collected at the recharge pond gives a detailed reflection of infiltration processes, it also creates an opportunity for expanding the inversion process to approximate hydrologic properties directly from acquired geophysical data without the use of petrophysical relationships. For example, given that electrical conductivity is highly related to water saturation in the vadose zone, it may be possible to estimate the hydraulic conductivity field in the subsurface directly from the rate and direction of changes in the electrical conductivity structure. The estimation of hydraulic

conductivity or other hydrologic parameters directly from changes in geophysical properties could greatly enhance characterization of the spatial distribution of these parameters, improving models of processes dependent on these properties.

This dissertation makes several contributions to the advancement of electrical resistivity imaging for monitoring vadose zone processes. In essence, this work formalizes the concept of informed imaging, providing a framework for incorporating all available information into the acquisition and processing of data that better measure the processes or properties of interest. I have developed an experimental design method that selects acquisition arrays producing high data sensitivity to an imaged region. I have also presented an approach for inverting time-lapse electrical resistivity data that incorporates all prior information about the evolving physical system and measurement error into the estimated conductivity structure at each time-step. Finally, my work has addressed some of the challenges of integrating different sources of data into a single hydrogeophysical model, improving simulations of subsurface processes and interpretations that guide decisions on the management of groundwater resources.

APPENDIX A.

Note on Appendix A:

Schlumberger's PETREL and ECLIPSE software packages are designed for creating structural models and flow simulations using seismic and supplementary well-log data. The primary application is simulation of 3-phase (oil-gas-water) flow in oil and gas reservoirs at depth. The following appendix contains notes on modifications to standard PETREL/ECLIPSE functions and processes that can be used to model 2-phase (gas-water) flow in the near surface. The purpose of this appendix is primarily to address importing and processing unsupported types of data that are often encountered in near-surface monitoring projects in the PETREL environment. It also describes modifications to standard inputs for modeling flow using ECLIPSE that allows for the simulation of two-phase (gas-water) flow at low pressures and saturations. The following is not intended to replace PETREL or ECLIPSE user-guides and tutorials, but as a supplemental guide when using the software for near-surface modeling applications. Furthermore, the descriptions in this appendix assume some familiarity with the options and processes used to create models using PETREL and ECLIPSE.

This appendix is organized according the workflow for building a PETREL structural model to use in an ECLIPSE flow simulation.

Section A1. includes descriptions of the most common data file formats used in PETREL; the pre-processing required to convert data to a PETREL-compatible format; and the workflow for importing different data formats into PETREL.

Section A2. addresses the creation of a structural model from diverse data sets.

Section A3. describes the process for exporting PETREL data to an ECLIPSE-readable data file, and the various sections of the ECLIPSE run file required to create a flow simulation. The most commonly used keywords for near-surface simulations associated with each section are described, but a full list can be found in the ECLIPSE user manual.

A1. Data

A1.1 PETREL Data Formats

PETREL accepts data in a number of formats, which can be found under “File ⇒ Import ⇒ File Selection Type.” Highlighting a given file type will bring up a short description and example file format. For shallow water applications the most common file types are *Simple Well Logs (ASCII)*, for logs and/or cores; *SEG-Y Seismic Data*, for 2d seismic sections with headers in UKOAA format; and *Image*, for processed GPR/ERI cross-sections or other unsupported processed data types that can be converted to depth-section images.

Simple Well Log takes an ASCII file in the standard GSLIB format, including file name, number of columns, column headers, and data in tab-delimited columns. Columns include the (x,y,z)-coordinates of each data point and the corresponding values. This format accepts values as general numbers, in scientific notation, or a combination of both.

An example of the format is:

```
filename
5
X
Y
Z
DATA1
DATA2
X1   Y1   Z1   DATA11   DATA21
X2   Y2   Z2   DATA12   DATA22
.
.
.
```

No end file indicator is needed to import files in this format. No duplicate point locations are allowed, and z-values must be monotonically increasing or decreasing. When importing these data as an on-shore well, z-values must be saved as meters below mean sea level (i.e. as negative values if above sea level). Section I.c.1 contains more information on the process of importing data.

2D Seismic Line accepts seismic sections with header information in a standard UKOAA format. Section I.b includes details on converting headers to this format. UKOAA headers include the line number, shot number, and the (X,Y)-coordinates of each common depth point (cdp) trace. Data can be imported as either time or depth traces for each cdp location. If imported as a depth section, the z-value of the top or base of the section can be changed using the calculator function after the section is imported into PETREL.

The last common file type for near-surface applications is the *Image* file. This is the most straightforward method for importing processed data sets that are not supported by PETREL (e.g. ground-penetrating radar (GPR) or electrical resistivity imaging (ERI) processed sections). Data plots saved as .jpg or .bmp files can be loaded into PETREL as image files. After importing, the coordinate locations of the four corners of the image are input into “Image Properties.” This works best if the images follow a predominantly straight line, and the section has the same top and bottom depths throughout. To change the location of the image a *Surface Imaging Module* license is required. Any reflections or features in the images must be digitized into a PETREL surface to be used as a structural or flow horizon.

A1.2 Pre-processing Data

A1.2.1 Coordinate Transforms

Coordinate transform is required if different coordinate systems are used in locating acquired data. The process below creates a transform for two coordinate systems in the same units. The first step is to develop a rotation, translation, and error calculation for moving between coordinate matrix **M**, given as a [3xn] matrix of (x,y,z)-locations in local coordinates, and **D**, the corresponding matrix in standard coordinates (those used for locating all data in the PETREL model). **M** and **D** must be co-located measurements simply given in different coordinate systems. Coordinates given solely in the local coordinate frame can then be translated to standard coordinates based on the derived relationship.

Before applying the script to derive the rotation matrix and translation vector, **M** must have the same units as the standard coordinate system. In the pseudo-code below the `make_coordinate_transform_matrices` script determines the rotation matrix, **R**, the translation vector, **T**, and associated error, **e**, for converting the (x,y,z) coordinates of the local system to the same position in standard coordinates. The following sample code gives the process for converting an [n x 3] array, **coord**, with the local (x,y,z) locations, respectively, to standard coordinates, where n is the number of points in the local coordinate system:

```
[R,T,e] = make_coordinate_transform_matrices(D,M);
for i=1:length(coord)
    coord_standard(i,:) = (R*[coord(i,1) coord(i,2) coord(i,3)]')'+T';
end;
```

A1.2.2 Finding Elevations of (X,Y)-Surveyed Points

Before calculating the elevations, (x,y)-locations must be converted into the standard coordinates (those used for importing data into PETREL). Section A1.2.1 contains details on coordinate transforms. A topographic model (`topo.dat`), converted to standard coordinates, must also be available and interpolated onto a regular grid. Applying a MATLAB interpolation function to estimate the z-values associated with

the (x,y) coordinates based on the (x₀,y₀,z₀) values of the topographic model will then give an estimate of the z-coordinate at new locations.

This process should only be applied to points that fall on the interior of the topographic model where there are a sufficient number of surveyed points to estimate elevations reliably. On the model edges values are only poorly estimated and should not be used for interpolation until the topographic model is expanded.

A sample code for this calculation is given below. In the code xmax, xmin, ymax, and ymin refer to the minimum and maximum x and y coordinate values in the topographic model. xtopo, ytopo, ztopo are the regularly-gridded coordinates in the topographic model; xgiven, ygiven are the coordinates given without z-values. Zest is the estimated z-coordinate:

```
[Xgrid,Ygrid] = . . .  
meshgrid(linspace(xmin,xmax,numpts),linspace(ymin,ymax,numpts));  
Zgrid = griddata(xtopo,ytopo,ztopo,Xgrid,Ygrid,'cubic');  
Zest = interp2(Xgrid,Ygrid,Zgrid,xgiven,ygiven);
```

A1.2.3 Reading Dates/Times from ASCII Files into MATLAB

When time stamps are associated with temporal data in ASCII files, they must be read into MATLAB as character strings and converted to a date variable. A simple approach is to convert time stamps to serial date numbers, which expresses the time as the number of days since 01 Jan 0000 0:00:00. Subtracting a later reference date from all date numbers will change the reference date for a particular data set. (e.g. Subtracting the serial date number for 01 Jan 2000 0:00:00 will change all values to days since that reference date.) To convert to serial date numbers time stamps must include a year, month, day, hour, min, and second.

A sample code is included below. The if/else statement may be omitted if dates have a consistent number of characters throughout the whole file:

```
fid = open('ASCII_file','r');  
  
% Skip multiple header lines by using this command for each one in  
the % file  
for i=1:n  
    temp=fgetl(fid);  
end  
  
%Read tab-delimited columns as strings (%s) or numbers (%f)
```

```

data = textscan(fid, '%s.. %f...');
fields = {'date', 'time'};
temp = cell2struct(data(:,n), fields, numfields);
    %n = the number columns with string values
date = temp.date; time = temp.time;

% If strings are different lengths use if/else statements.
for i=1:numdates
    if (date(:,end)==' ')
        mo(i) = str2num(date(:,mo_cols));
        day(i) = str2num(date(:,day_cols));
        yr(i) = str2num(date(:,col:end-1)); % if date = 'm/dd/yy'
    else
        mo(i) = str2num(date(:,mo_cols));
        day(i) = str2num(date(:,day_cols));
        yr(i) = str2num(date(:,col:end)); % if date = 'mm/dd/yy'
    end

    hr(i) = time(:,1:2);
    mn(i) = time(:,4:5);
    sec(i) = time(:,7:8); % for time = hh:mm:ss
end

datestamp = datenum(yr, mo, day, hr, mn, sec);

```

A1.2.4 Averaging Transient date from hourly/subhourly to daily values

Often times it will be necessary to convert time-series data sets to a common sampling interval. Once the time stamp for each data point has been converted to a MATLAB serial date number (Date/time converted to days since 01 Jan 0000 0:00:00 or some other reference as previously described) the sample code below can average hourly or sub-hourly measurements to daily values. With slight modifications monthly and yearly averages can be obtained through a similar process.

```

tmin = floor(min(timeseries)); tmax = floor(max(timeseries));
k=1;

for i=tmin:tmax
    sum = zeros(1:num_datacolumns);
    n = 0;
    for j=1:length(timeseries)
        if (floor(timeseries(j))==i)
            sum = sum + data(j,:);
            n = n+1;
        end
    end
    daily_ave(k,1)=i; daily_ave(k,2:length(data)+1) = sum/n;
    k = k+1;
end

```

A1.2.5 Converting coordinates in SEG-Y Headers

If the coordinates associated with processed seismic headers are not in the same coordinate system used to import data into PETREL, the coordinates and header data must be transformed to a standard UKOAA format reflecting coordinate information consistent with the PETREL coordinate system.

This is accomplished by first creating a header file with the line number, shot point number, x-coordinate and y-coordinate (in standard coordinate system values) in the following format:

Line name (left adjusted):	Columns 1-16
Shot point number (justified):	Columns 17-24
X-coordinate (justified):	Columns 45-52
Y-coordinate (justified):	Columns 53-60

In the UKOAA format, X- and Y-coordinates cannot exceed 8 characters. However, values with too few significant figures will appear to overlap and cause distortions in the image when data are loaded into PETREL. The header file should be saved into the same directory as the existing seg-y data with a “.prn” extension.

The headers can then be transformed by launching the SEG-Y Utility function in PETREL. Within the utility, the created UKOAA header file should be loaded into the UKOAA selection tab and the original data should be included as the seg-y file to be imported. It is important to include the line numbers where indicated. The default UKOAA format options can be used, but the trace identifier should be changed to cdp, shotpoint, or trace number as appropriate. The output is saved to a new directory and includes a new *.sgy file with the same name as the original file and a report file. The *.sgy file in the created directory contains the new header information and is the one that should be imported into PETREL.

A1.2.6 Reading Electrical Resistivity data from MPT-ERT output files

Collection of ERI data at a site using a Multi-Phase Technology ERT (MPT-ERT) instrument requires data processing using software from the Environmental Geophysics Group at Stanford prior to input into PETREL. Scripts have been developed by the group to process these data in MATLAB. As a first step, MPT-output data must be converted into a MATLAB-compatible format. The script `datareader.m` (included below) can be used to extract data from an MPT-ERT output file (*.out) into MATLAB. This code only works for complete data files, so data files that stop prior to the end of a run should be eliminated before reading the values into MATLAB, as they will cause an error. These data can be added individually later.

Once data have been read into MATLAB, measurements sent to error files (*.err) can be replaced with NaNs using the script `datafill.m` (also included below). The (x,y,z) locations of each source, sink, positive receiver and negative receiver electrode are recorded for each measurement in the output of `datafill.m`. If data values are available they are recorded with the corresponding coordinates, otherwise columns are filled with NaN.

```
datareader.m
% Script for reading data from MPTERT .out files.
% Create and load a matrix of all folders and files to be read from
MPT
% output. These are 'folder' and 'file', below.

% Read data from files (Sample, R, Rsd, Charge, IP, IPsd, Amps
for i = 1:size(folder,1)
    for j=1:size(file,1)

        % make variable with directory and file name
        infile = sprintf('/Users/vanessa/Documents/Research/Pajaro
            Valley/FieldData0809/Output/%s/out_file_%s_%s.out',
                folder(i,:),folder(i,:), file(j,:));

        %open file
        fid = fopen(infile,'r');

        %move through 17 lines of mpt output file header
        for k=1:17; temp=fgetl(fid);end

        %read in data (this is only sample #, R, Rsd, Charge, IP,
        % IPsd,and Amps; elec locations are saved to another .mat file)
```



```

data = textscan(fid, '%f %*f %*f %*f %*f %*f %*f %*f %*f %f
%f %f %f %f %f %*f %*f %*f');

% make cell matrix into number matrix
outdata = cell2mat(data);

%write to .mat file
outfile = sprintf('/Users/vanessa/Documents/Research/Pajaro
Valley/FieldData0809/ERTMAT/%s_%s.mat',
folder(i,:),file(j,:));
save(outfile,'outdata');

fclose(fid);
end
end

datafill.m
% Script to output electrode locations and data for all ERT data
files
% collected.

% Columns of MPT-ERT output files:
% TX+(x,y,z) TX-(x,y,z) RX+(x,y,z) RX-(x,y,z) Res Res_sd Chgbility IP
% IPsd Current
% TX/RX are in utm modified transformed coordinates (total of 12 col)
% Res =V/I
% Res_sd = sd of res values
% Chgbility = chargeability
% IP
% IPsd = sd of IP meas
% Current = driving current in Amps

% load mat file with elec index, locations, and file names
load matgen090603.mat

% make vector of file labels in terms of minutes
mo = str2num(files(:,1:2)); day = str2num(files(:,3:4));
yr = str2num(files(:,5:6));
hr = str2num(files(:,8:9)); mn = str2num(files(:,10:11));
s = zeros(1937,1);
mns = datenum(yr,mo,day,hr,mn,s); mns = (mns - min(mns))*24*60;
mns = num2str(mns);
clear mo day yr hr mn s;

% define data matrix that will be output with elec locations
data = zeros(4656,18);
data(:,1:12) = [elecloc(elecind(:,2),2:4) elecloc(elecind(:,3),2:4)
elecloc(elecind(:,4),2:4) elecloc(elecind(:,5),2:4)];

for k = 1:size(files,1);

%load data from ERTMAT dir
infile = sprintf('/Users/vanessa/Documents/Research/Pajaro
Valley/FieldData0809/ERTMAT/%s',files(k,:));
load(infile);

```

```

% clear data from last run
data(:,13:end) = zeros(4656,6);

% fill in data matrix
for i = 1:4656
    a = find(outdata(:,1) == i);
    if isempty(a)
        data(i,13:end) = NaN(1,6);
    else
        data(i,13:end) = outdata(a,2:end);
    end
end

% make out file to dir ERT0809 and save new data
outfile = sprintf('/Users/vanessa/Documents/Research/Pajaro
Valley/FieldData0809/ERT0809/%s',mns(k,:));
save(outfile,'data');

% clear outdata variable for next run
clear outdata;
end

```

A1.3 Importing Data into PETREL

A1.3.1 Simple Well Logs

When first importing log data, *Simple Well Logs (ASCII)* is the most straightforward option. Log data are imported by selecting the file to import and the *Simple Well Logs* data type. In the dialogue box, icons can be used to set the number of columns to be imported. For the first file, the “Generate new” log type should be used, and a variable name and template must be defined for each data series. When importing subsequent wells or logs with the same data types, the tabs under “Global Well Logs” will select the log type associated with each column. In all cases, if z-values are given in meters below mean sea level (msl), the “Onshore” option under the “Geometry” tab must be highlighted.

To add data series to existing wells either the *Simple Well Logs* or *Well Logs* file type may be used. If using *Simple Well Log*, the process follows that described above. If using the *Well Logs* data type the z-axis should be given in meters above msl instead of meters below msl.

A1.3.2 SEG-Y with Preset Parameters

The *SEG-Y with Preset Parameters* import selection type is only required if header files are altered to the UKOAA standard format. If the original header files contain the correct coordinate data and no other alterations are required, the original data file may be loaded into PETREL using the more generic *SEG-Y Seismic Data* option.

The *SEG-Y with Preset Parameters* allows the user to import processed seg-y seismic data with altered header files that include the (x,y)-location of traces in the standard coordinate system. To import data under this file type, the *.sgy file in the directory created by the SEG-Y Utility as described in the section A1.2.5 above must be used. The coordinate scaling value should be changed to “1”. All other options may be left at the default values. In the second window, it is possible to alter the name of the seismic line and to change between time and depth on the z-axis.

Once a seismic section has been loaded into PETREL, the depth of the section can be shifted by right-clicking the seismic variable under the “Inputs” tab to the left of the visualization window. Under the “Geometry” tab the shift function moves the section up or down a given number of seconds or meters. To apply changes to the data, it is necessary to click the “Shift” button before exiting the menu. To use this approach, the absolute elevation of the top of the seismic section must be known prior to importing the data.

A1.3.3 Images

For data types, such as GPR or ERI, that result in processed images from data that are not currently supported by PETREL, the simplest solution is to import the data as images (either *.jpg or *.bmp files). The default options for importing this type of file are sufficient. Once the image variable appears in the “Inputs” window, the (x,y,z)-locations of the four corners of the image are added by right-clicking on the image icon and selecting the “Property” tab. The data cannot be added to the visualization window until the coordinates have been added. This process works best for data that were collected in a (predominantly) straight line and for which the depth of the image section is well constrained. If this is not the case, the location of features

identified in the images may be distorted. Adding the coordinate locations to image corners requires the *Surface Imaging Module* license, which is not current part of the Stanford PETREL package but is available from Schlumberger.

A1.3.4 Using Settings/Calculator/Spreadsheet

Even after being imported into PETREL, a data series may be manipulated in a number of ways. Most modifications required can be achieved using the “Settings,” “Calculator,” or “Spreadsheet” options available on the right-click menu of any item on the “Inputs” tab. If an individual item is selected, only the properties of the data series associated with that item will change. To change properties for all data series of a particular type the corresponding options under the “Global” heading should be selected.

The “Settings” option can be used to change the geometry (e.g. the z-locations) and the color bar properties associated with a data series. This tab also contains several options for altering how the data series is labeled and how it appears in the visualization window. The “Calculator” tab is useful for performing conversion within PETREL. For example, the calculator can change permeability values from m/day to Darcy’s by applying the appropriate transform. Although PETREL does keep a history of calculator operations, it does not associate units with a particular data series, so any conversions should be well-documented. “Spreadsheet” contains all locations and property values associated with a given well location. The individual entries can be changed manually within the spreadsheet. If possible, however, this option should be avoided as PETREL has no option to track changes made within the spreadsheet.

A1.3.5 Using Global Well Logs and Templates

All well data is associated with a log type template when it is imported into PETREL. The default option is “generic well log.” When possible, it is convenient to associate a type of data with the corresponding PETREL data template. Right-clicking on the data name under the “Global Well Logs” heading and selecting the “Templates” option can change the template for any data type. The properties of the color bar

corresponding to each data series can be changed in a similar fashion by selecting the “Color bars” option.

Right clicking the “Global Well Logs” heading, and selecting the “Symbols” option can also change the appearance of wells in the visualization window. This will change the size of the wellhead labels in addition the appearance of the well sections.

A1.3.6 Digitizing Surfaces

It is possible to extract features from data input as images by digitizing surfaces along horizons appearing within the image. Surface can be digitized by selecting “Make/Edit Polygon” under the “Utilities” tab; choosing “Make/Edit Points ⇒ Add New Points;” and clicking along the horizon of interest. It should be verified that all points fall on the plane of the image. This procedure generates a new set of points for each image and each horizon. When combining horizons from a number of images, selecting the first set of points created, right clicking, and choosing “Points ⇒ Settings ⇒ Operations ⇒ Append Points” will set the horizon to be augmented. Highlighting another set of points created in the “Inputs” tab, and clicking the “Run” button will add those points to the first series. This process can be repeated for all remaining sets of points. The merged set of digitized points is then converted to a surface using the process described below in section A2.1.2.

A2. Generating Structural Models in PETREL

A2.1 Defining Model Boundaries

The following section describes the process for defining the (x,y,z)-extents of a PETREL model space and cells.

A2.1.1 Defining Polygons

The lateral extents of a Petrel model space are defining by a bounding polygon. Polygons are created using the tool in “Processes ⇒ Utilities ⇒ Make/Edit Polygon.” Double-clicking the last option brings up a dialogue box for defining the (x,y,z) boundaries of a rectangle or some other set polygon. Alternatively, clicking at various points within the visualization window will create a polygon with nodes at selected points. For each model, two polygons with the same (x,y) boundary coordinates must be created at the upper and lower extent of the model space. This is achieved by 1) repeating the process above with 2 different z-values or 2) copying, pasting, and renaming the polygon in the “Inputs” tab and then adjusting the depth of the polygon. The depth is adjusted by right-clicking the copied polygon and selecting “Settings ⇒ Geometry” to change the value of z through variable “A.”

A2.1.2 Creating Surfaces

Polygons are converted to surfaces using “Processes ⇒ Utilities ⇒ Make/Edit Surface.” Before opening this function, a polygon on the Inputs tab should be highlighted. Clicking the arrow icon will set the highlighted polygon as the main input. The same polygon can be inserted for the surface boundary unless it does not cover the full (x,y)-extent of the desired surface. Selecting the “Get Selected Limits” option will prompt PETREL to automatically fill the remaining settings for creating a surface. This process should be used to create a surface at the top and bottom of the model space to define the upper and lower vertical extents, based on the polygons used to create the lateral boundaries of the model space. It should be verified that the resulting limits match the (x,y)-extent of the model space. Again, when creating

multiple surfaces, the process above can be repeated for each surface or the surface can be copied, renamed, and shifted using the same process described for the polygon.

A2.2 Making Grids

A2.2.1 Make a Simple Grid

A simple grid is created using the function under “Processes ⇒ Utilities ⇒ Make Simple Grid.” Highlighting a surface with the correct (x,y)-extents on the “Inputs” tab and using the arrow icon will set the selected surface as the model boundary. Using the “Get Limits from Selected” option will automatically fill required fields for creating a grid.

The number of nodes specified in the x- and y-directions determines cell spacing. There is always one more node than cell in a grid. Hence, the number of nodes should be set to one more than the value of the X- or Y-length divided by the desired dx or dy cell increment. A good starting value is roughly 50 cells (51 nodes) in both the x- and y-directions unless the model space is particularly large or greater detail is required over the entire grid space. If only a small area requires higher resolution it is better to use a local grid refinement in the region of high-resolution. Local grids are discussed in section A2.2.3.

In addition to selecting the boundary, the options should also be set to “Skeleton only” and surfaces or values should be selected for the “Top Limit” and “Base Limit.” If a topography surface is available, it can be used to define the upper boundary of the model space. If no topography is available a flat surface can be used for both the top limit and the bottom limits.

A2.2.2 Defining Horizons and Layers

Created grids appear under the “Models” tab. The process to this point only defines the cell spacing in the (x,y) plane. Before creating the dz cell spacing PETREL requires the creation of horizons. Horizons also appear under the “Models” tab and correspond to surfaces in the Inputs tab. Horizons guide the layering process and provide the boundaries for defining regions within the model; horizons divide the

model space into zones with the number of layers defined for each zone. When running ECLIPSE, mass balance is achieved by measuring flow from one zone to another across horizons. If there are no distinct regions in the model it is possible to define the entire model space as a single zone with equally spaced layers.

Z-cell spacings are created by opening “Processes ⇒ Structural Modeling ⇒ Make Horizons,” and using the “N”-icon to create the desired number of horizons. Highlighting and loading surfaces from the Inputs tab using the arrow icons will create horizons that correspond to those surfaces. The number of layers in the zone should be set such that they give the desired interval spacing for each zone, where the total thickness of the zone divided by the number of layers is the z-cell spacing. Ideally, the z-spacing should be equal to the resolution of the input lithology data, although the z-spacing can be defined on a coarser or finer scale. After running this utility the entire (x,y,z)-grid should be available under the Models tab and can be seen in the visualization window.

A2.2.3 Local Grid Refinement

If a portion of the model requires higher resolution, the grid in that area can be refined to smaller cell sizes using “Processes ⇒ Structural Model ⇒ Make Local Grid.” When the grid is refined within a subregion of the model space PETREL and ECLIPSE divides the model into a local (refined) and global (original) grid. Although these grids are connected all structural (PETREL) and flow (ECLIPSE) properties must be defined for both the local and global grids.

Under the function opened above, selecting “Create New” and loading a polygon defines the boundaries of the local grid. It may be necessary to create a polygon with the boundaries of the local grid using the process described above for creating the global grid. Selecting the new number of (x,y,z)-nodes will create a subgrid with finer discretization; the apply changes option must be applied for local grid refinement to appear.

A2.2.4 Creating Regions within a Model

If trying to calculate mass balance using PETREL or ECLIPSE, the model must be divided into two or more regions as fluxes are determined across the boundaries between regions. Creating two regions in a model with only one zone is accomplished by creating a PETREL property with value 1 in every grid cell under the “Models” tab. Using “Process \Rightarrow Property Modeling \Rightarrow Geometric” the template for the new property can be set to “Region” and a value of 1 will be assigned everywhere in the grid for that property. Bringing up the “Settings” option by right-clicking on the new property and selecting the “Filter” tab will make it possible to select a subspace of the model that will be converted to a second region. Choosing the “Index filter” option and filtering on the variable K (the z-index), the minimum and maximum z-indices of the second region can be specified. After applying changes, right clicking on the new property again and using the calculator to set the region value equal to 2 will change the value of the property within the filtered volume. The first region is associated with the value 1, and the second region is associated with the value 2.

The process described above is only required if defining regions within a model consisting of only one zone. If a model has been segmented into different zones by defining more than one horizon, the above process is not required. The additional horizons in the model interior automatically create multiple zones.

A2.3 Populating a Grid with Properties

A2.3.1 Using Property Modeling

Data variables to be transferred from well logs to the model space must first be upscaled to the grid spacing. Under “Processes \Rightarrow Property Modeling \Rightarrow Scale up Well Logs,” selecting “Create New Property,” the property to be scaled, and all the wells containing the selected data series will upscale data logs in all wells. The settings for upscaling are automatically linked to the layers and grid highlighted in the Models tab.

Once all the needed properties have been upscaled, the model space is populated by opening “Processes \Rightarrow Property Modeling \Rightarrow Petrophysical Modeling,”

and selecting “Use Common Settings” and “Overwrite.” Highlighting the desired upscaled variable and unlocking the properties using the icon in the lower left corner will open the interpolation scheme options. Unless detailed variogram information is available a simple scheme with the default orientation and correlation length is recommended. However, if geostatistical information is available it should be used to provide a better estimation of the spatial variability of properties. If more than one zone has been defined, all zones should be highlighted under the “Zone Filter” on the Models tab before applying the interpolation. Once this modeling process has been applied all structural variables required to run an ECLIPSE model have been defined over the entire model space and values can be exported to run a flow simulation.

A2.3.2 Defining Properties using Neural Networks

Neural networks can be used to predict a data set in wells where certain logs are missing based on correlations between data sets in other wells. For example, the electrical resistivity log in a well with only lithology data can be estimated from the correlation of lithology and resistivity in wells with both sets of data. Ideally at least three types of logs should be used to create a neural network, and two of those logs should be available in the well for which the third data set is estimated. The wells used to create the neural network are called training wells or training logs, and one-half to two-thirds of the data points in these wells are used to create the neural network while the remaining one-half to one-third of the points are used to test the neural network.

Before running the Neural Network Tool, it may be beneficial to detrend the data sets used to train the model. This can be achieved in PETREL by selected a data series within a training well and opening “View Log Editor.” Subtracting a heavily smoothed version of the data set from the original series will give an approximation of the differenced property. It should be noted this is an optional step that may improve results but is not necessary if there are no obvious trends in the data.

To begin, the “Well Manager” function should be opened under “Import ⇒ Wells.” Selecting “Show ⇒ Logs,” the training wells and logs to be used should be copied into a new folder with an obvious label. The Neural Network tool will only

work on this folder and its contents so all training information must be included. When creating the neural network it should be kept in mind that like many other property tools in PETREL, the Neural Network tool is designed in the Input tab and applied under the Models tab.

Creating a neural network for the selected wells and logs uses “Train Estimation Model” function under the “Utilities” tab. Wells copied into the neural network folder and the training logs should be loaded, and the “Estimation” option should be selected instead of the default “Classification” option. The log to be estimated should be highlighted as the model. This log must be available for all training wells but will be the missing log in the wells for which the neural network is created. In the Settings environment, the output log and the training parameters can be set to default parameters unless more information about the how the model should be trained is available. During this step the user should consider that overtraining the model and using correlated logs should be avoided as they will yield sub-optimal results when applied at locations distant from the training well or wells. The created neural network can be viewed by going to “Windows ⇒ Well Selector ⇒ Well Logs ⇒ Created Neural Net.”

The neural network is applied to an existing grid by opening “Property Modeling ⇒ Geometric ⇒ Zone/Increments,” and selecting the created neural network object in the Inputs tab. Using the option “Make General Zone,” and a giving a new name for the object will avoid overwriting the neural network. In “Processes ⇒ Property Modeling ⇒ Upscale Well Log,” choosing create new property, selecting the property to be estimated and the wells for which it should be estimated creates an upscaled property based on correlations in the neural network. Opening the “Petrophysical Modeling” function and selecting the zone created from the trained neural network will populate the property over the model space. After applying this function, the estimated data are integrated with the existing data from which the neural network was created.

A2.4 Incorporating Time-Dependent Data

It is possible to incorporate time-dependent information into PETREL as observed data. These data can be used as direct inputs into ECLIPSE as production information, boundary conditions, or as time-series observation data against which ECLIPSE outputs can be evaluated.

A2.4.1 Production Data

If water is being produced from recovery wells, and there is a time-lapse record of production history, these data can be included within PETREL as a sink to the fluid-flow system.

Before importing the time-series data, the completion data for each producing well must be defined in terms of the well radius, borehole radius, and screened interval. This allows the software to convert production, in terms of flow rate, to pressure gradient information. Well completion data are included by selecting ‘Global Completions ⇒ Import on Selection,’ and importing a *.ev file containing the units, well name, completion date, “perforation”, top and bottom perforation elevations, and well diameter. Alternatively, this can be done by opening a well-selection window, highlighting the desired well under the “Inputs” tab and selecting “Processes ⇒ Well Completion Design ⇒ Add/Edit Casing.” The date under this option should be set to the start of the simulation or before. Then, under the “Add/Edit Perforation” tab, the well casing diameter and perforation zone should be set. Finally, the well-hole diameter can be set under “Completions ⇒ Settings.”

Once the well completion has been defined, the production data are included by simply selecting “Inputs ⇒ Wells ⇒ Global Observed Data ⇒ Import on Selection” and selecting a *.vol file associated with each well. The *.vol file should include a date and time stamp, as well as the production rate in the appropriate units. The software assumes the production rate is constant at the given value until the next time-step and flow rate included in the *.vol file. Hence, if a well stops producing, a line should be included in the *.vol with the time production stops and a flow rate of 0.

A2.4.2 Hydraulic Head Observations

Observations of the hydraulic head at given locations are imported in the same way as production data. PETREL does not differentiate between these types of data. Hence, when exporting time-dependent data from PETREL to ECLIPSE, it is important to be aware of which observed data sets are included within the development strategy. Observed hydraulic head data should not be included in the ECLIPSE run. However, once associated with the PETREL model, the observed hydraulic head values can be used to history-match the data set after the flow simulation has run. Piezometric maps of the hydraulic head can also be created by importing the hydraulic head values observed at a number of wells at a given time as points with attributes and then creating a surface using the same process described for creating a topographic surface.

A2.4.3 Boundary Conditions

Boundary conditions on the edges of the model space can be created using time-varying flux or hydraulic head information using a similar process to that defined above for creating aquifers. However, for allowing flow across the vertical edges of the model, the aquifer should be defined to include only the outer edges of cells along the boundary, and the aquifer connection should be set to 'top-down.' The time-varying hydraulic head or flux data are included by defining the type of aquifer as either constant head or constant flux and changing the values at regular intervals throughout the simulation.

A3. Flow Simulations in ECLIPSE

ECLIPSE is a multi-phase flow simulator integrated with the PETREL software package. It was originally developed for modeling three-phase flow (oil-water-gas) in reservoirs at depth, but can also be use for near-surface, two-phase (gas-water) systems. ECLIPSE uses structural and property models developed in PETREL to solve a mass balance equation based on Darcy's Law. The solution is calculated in terms of pressure and takes minimum inputs of: the porosity and permeability structure in the model space, pressure-saturation curves of the local sediments, fluid properties for each phase, the initial pressure conditions, and grid structure of the model.

ECLIPSE accepts inputs through various keywords under different headings in a single run file with a ".DATA" extension. All properties can be included directly in the run file with the associated keyword or as an independent *.inc file called by the run file. There are a few basic keywords that appear in almost all run files, but numerous optional keywords exist that can simulate almost any hydrologic condition. The ECLIPSE help file contains a glossary of all keywords, the associated inputs, and the various functionalities achieved by using different combinations of inputs. The run file is split into a number of sections that describe spatial and temporal conditions of the simulation. Keywords are associated with these various sections.

A3.1 Exporting Data to ECLIPSE using the PETREL Interface

Before running a fluid model from PETREL, it is useful to import files containing standard fluid properties that are associated with several keywords. Often for near-surface applications, site-specific values for these properties are unavailable and these standard values suffice. However, if site-specific values are available they should always be used over the default values. Fluid property files are imported by selecting "Import ⇒ ECLIPSE Fluid Model" from the PETREL File menu and choosing a model type. For variably saturated flow simulations "Black Oil Model 1" is a good choice; it contains all liquid properties need for the water and gas phases.

Importing this file creates a new property object on the Inputs tab named “Fluids.” This property holds the information about the physical properties of the fluids.

For each type of fluid (i.e. water and air) in the model an aquifer must be created. This is achieved by creating a polygon covering the area of the aquifer using the “Make/Edit Polygon” function under “Utilities” and the “Make New/Deactivate Old” button to highlight the region of the aquifer in the visualization window. Clicking the “Close Selected Polygon” button creates the polygon and all points interior to the polygon are part of the aquifer. The polygon is made into an aquifer by opening “Simulation ⇒ Make Aquifer,” and choosing “Create New” with constant pressure and loading the created polygon into the Region of Interest slot. Highlighting the property tab, the pressure can be set to 1 bar at the elevation of the pond head. Salt concentration is left at 0 unless the salinity of the water is known, and the imported fluid model is included under this tab. Repeating the process will create a second aquifer for the gas phase. However, the gas aquifer should cover the entire surface of the model and the pressure should be set equal to 1 bar on the ground surface.

The next step is to create a “Development Strategy” by editing the end date and selecting the reporting frequency and property values to be exported. The time variable should also be set to the date when the flow simulation begins. A simulation case is defined by loading grid properties, including the structural and hydrologic parameters (porosity and permeability), and the created aquifers. In the simulation case, functions for the saturation curve, initial conditions, and rock compaction must also be loaded. It should be verified under the “Strategies” tab that the required properties are loaded under the “Development Strategy Folder.”

Once all the tabs have been filled with the necessary properties and input files exporting the simulation case creates the ECLIPSE files required to run a simulation. The simulation folder created by PETREL contains the run file and an INCLUDE file with all the keywords required for each section.

A3.2 Creating an ECLIPSE Driver File Outside PETREL

A simulation case can also be defined outside the PETREL environment, though it may still be convenient to export some property and grid information from

the PETREL workspace for large or complex model geometries. A simulation case can be generated through a *.DATA file with the appropriate keywords within each section. The description below for creating such an ECLIPSE run file contains only basic keywords for generating a simple model including an aquifer, pumping wells, and simple boundary conditions. Keywords for creating more complex flow simulations can be found in the ECLIPSE help file.

Included below are the sections of the ECLIPSE run file and a list of the most basic required keywords. Each ECLIPSE keyword takes a number of arguments that give values to the different properties associated with the keyword. Descriptions of the arguments, as well as the order in which arguments appear after the keyword, are included in the definition of each keyword in the ECLIPSE help file. A value must be provided for each argument of the keyword in order for ECLIPSE to compile correctly. To use the ECLIPSE default values for an argument enter “1*” in the argument position. Some general notes and observations for keywords and arguments are also included for each section.

A3.2.1 Section of the ECLIPSE Driver File

RUNSPEC

This section initializes run information for the simulation. Once the units for input variables have been set it is important to verify that all input values are consistent with the specified units. This can be checked under ‘Units’ in the ECLIPSE help file. It should be noted that “MDarcy” are in fact milliDarcy.

TITLE: gives a name to the run for all outputs.

START: gives the time at the start of the run.

METRIC: defines metric units for all input values.

WATER: includes water as a fluid in the simulation.

GAS: includes gas as a fluid in the simulation.

DIMENS: gives the maximum dimensions of the global grid. ECLIPSE allots memory based on the dimension of the grid.

LGR: local grid refinement. Contains information on the local grid if one has been created.

UNIFOUT: puts all output files into one header file instead of individual output files for each time-step.

AQUDIMS: maximum aquifer dimensions; should be set equal to the maximum number of model cells connected to the aquifer. The number of dimensions should be equal to the number of cells in the x-direction multiplied by the number of cells in the y-direction if there is a gas aquifer over the entire model space, as is required if the top of the model space is the ground surface.

EQUILDIMS: equilibration regions allows for a slope in the water table. This value is equal to 1 if the region is flat.

TABDIMENS: defines all the tables in a data set. Each table is independent based on the region to which it is applied. For most applications this is equal to the default inputs. However, if using regions the fifth argument on the input should be set equal to the number of regions in the model.

NOMONITOR: turns off run-time monitoring in the summary files.

NORSSPEC: prevents ECLIPSE from create a new summary file each time it restarts.

NOINSPEC: prevents ECLIPSE from creating an initial index file.

MSGFILE: controls generation of a message file depending on the argument. This file is not required, and often not useful, and can be turned of with an argument of "0/."

GRID

This section contains information on the grid structure. It is typically easiest to export this information from PETREL by creating a simulation case as defined in the previous section and copying this information into the run file or by exporting the grid directly from PETREL by right-clicking the desired model and exporting as a general ECLIPSE file in ASCII format.

GRIDFILE: contains the number of gridfile.

INIT: causes ECLIPSE to output the initial files with the results.

INCLUDE

GRID.inc: tells ECLIPSE to include the file grid.INC, which contains the grid exported from PETREL. The “INCLUDE” command is needed to associate any *.inc file with the run file. The grid.inc file can be created by selecting to export the entire grid under the ‘Models’ tab in PETREL. If created this way, the file will contain both grid information and the property values associated with each grid cell.

EDIT

No keywords are generally required for this section.

PROPS

This section defines matrix properties of the model region. It should be noted that the simulation is very sensitive to the saturation functions included in this section. If possible, site-specific saturation-permeability curves should be used. Otherwise, these curves can be treated as tuning parameters. Furthermore, if the model contains sand and clay regions, it may be worthwhile to define two model regions based on the clay-content and to define different saturation curves for each section.

SGWFN: gas-water saturation functions give the air-water saturation curves. If no other data are available this can be set to 0 or 1 at the boundaries for no saturation or complete saturation, respectively.

HYDRHEAD: converts pressure into hydraulic heads given a reference elevation and density. This is useful if trying to calibrate outputs to actual data.

PVTW: gives water properties such as volume fraction, viscosity and compressibility.

PVDG: gives the above properties for the gas phase.

DENSITY: provides ECLIPSE with the densities of the different fluid phases at 20° C and standard pressure. The density of water here should be equivalent to that used in the hydraulic head reference.

ROCK: values of pressure-compressibility of the rock matrix. The storage estimation and conversion from storage to porosity and compressibility are derived from this keyword.

FILLEPS: takes no arguments but puts saturation endpoints into the output associated with the INIT file.

REGIONS

This section defines subregions of the model space if they exist.

INCLUDE

FIPNUM.inc: includes file FIPNUM.inc, which associates each model cell with the region in which it resides. If there are no regions this keyword is not required. If regions are created the FIPNUM.inc file is created from a PETREL property related to the creation of regions.

SOLUTION

The solution section includes information on the initial and boundary conditions for solving the multiphase flow equations.

Whenever possible, it is significantly easier to define a flat water table when initializing the model. An equilibrium steady-state hydraulic head structure can be achieved by running the simulation with just the boundary conditions for some time before the actual model run begins. Furthermore, the depth units in ECLIPSE are opposite of those used in PETREL. If, for example, the water table exists above the zero datum, the elevation argument of the EQUIL keyword should be a negative value.

If the geometries of the aquifers are at all complex, it is much more reliable to generate the AQUACON keywords through PETREL and then to copy the connection arguments into ECLIPSE. When doing this, the default value for the last argument is defaulted to 'YES' in PETREL and will need to be changed to 'NO' within the driver file. This avoids the possibility of creating sources of water within the subsurface.

EQUIL: gives the initial (flat) water table.

AQUCHWAT: connects the aquifer to the subsurface and includes water properties.

Reviewing this keyword before running is highly recommended. This only gives connection for the water aquifers; another keyword will be required for air if modeling the unsaturated zone. If there is local grid refinement under the aquifer, ECLIPSE classifies these as different, but connected aquifers. If a local grid refinement exists more than one aquifer must be defined using this keyword.

AQUCHGAS: has the same behavior as AQUCHWAT, but gives values associated with the gas or air component in the unsaturated zone. This keyword is also critical for connecting surface aquifers to the subsurface. If the top of the model is the ground surface, the gas aquifer should be defined to cover the entire model surface.

AQUACON: defines the cell interfaces at which global aquifers connect to the rest of the model. Whereas ACUCHWAT and AQUCHGAS are classified by the phase of the aquifer, AQUACON is classified by whether the aquifer is global or local.

AQUACONL: behaves as AQUACON but applies to aquifers connected to the local (refined) grid area. If no grid refinement is used this keyword is not required.

SUMMARY

This section requests desired outputs from the simulation. It is worthwhile to preview the possible outputs within the description of this section in the ECLIPSE help file. It should be noted, however, that not all outputs are available under all licenses and that requesting an unavailable output will cause the simulation to fail.

RUNSUM: requests a summary of the run results.

SEPARATE: produces separate files for each requested output.

RPTONLY: requests output only at specified times.

AAQR: requests all data that was specified to be checked.

BWSAT: gives saturation through time at a grid cell.

RWIP: requests output of the regional water in place.

SCHEDULE

The SCHEDULE section includes all time-dependent conditions of the flow simulation.

LGRLOCK: solves the local grid implicitly with the global grid. This keyword requires the names of the local grids and is only necessary if the model space has been refined in some subregion.

RPTRST: restarts the file output control to give the requested output at each specified time.

DATES: requests output at the date given as an argument. If followed by AQUCHWAT or some other property definition keyword it changes the existing conditions at the date specified in the argument.

A3.2.2 Sample ECLIPSE Driver File

RUNSPEC

TITLE
SAMPLEFILE

START
1 JAN 2006 /

METRIC

WATER
GAS

DIMENS
41 42 55 /

WELLDIMS
7 17 2 7 /

NOMONITOR

UNIFOUT

TABDIMS
2 1 22 20 1* 20 20 5* 1 /

AQUDIMS

1* 1* 1* 1* 29 1320 5 14 /

EQLDIMS

1 /

GRID

INIT

GRIDFILE

2 /

NOECHO

INCLUDE

griddefinition.inc/

ECHO

EDIT

PROPS

PVTW

1 1 4.59E-005 1.005 0 /

PVDG

1 1.0001 0.0185
2 1 0.0185 /

DENSITY

800 998.2 1.204 /

SGFN

--!!! Program generated keyword. Remove this comment to prevent PETREL editing or removing this keyword.

0 0 0
0.0075 0.009087 0
0.04 0.058375 0
0.12 0.20027 0
0.24 0.40827 0
0.36 0.58986 0
0.44 0.73862 0
0.52 0.85601 0
0.6 0.91959 0
0.64 0.97458 0
0.99 1 0
/

SWFN

--!!! Program generated keyword. Remove this comment to prevent PETREL editing or removing this keyword.

0.01	0	1
0.36	1.86E-008	0.8
0.4	1E-007	0.4978
0.48	2.96E-005	0.1661
0.56	0.000349	0.0727
0.64	0.002	0.0372
0.76	0.015982	0.0159
0.88	0.072595	0.0059
0.96	0.23402	0.002
0.9925	0.48519	0.0006
1	1	0

/

HYDRHEAD

-- Reference data for output of hydraulic heads

0 998.2 /

ROCK

1 0.01 /

FILLEPS

REGIONS

SOLUTION

AQUALIST

--!!! Program generated keyword. Remove this comment to prevent PETREL editing or removing this keyword.

BC_HI

4 /

BC_LOW

5 /

ATM

2 /

POND

1 /

AQUCHWAT

4 -0 PRESSURE 2.9 2000000000 1 0.0 1* 1* NO/

5 -0 PRESSURE 2.3 2000000000 1 0.0 1* 1* NO/

/

AQUCHGAS

2 -0 1 2000000000 1 /

/

AQUFLUX
1 0.0 0.0 /
/

AQUANCON

Aquifer name, identification number	LGR name,	Time,	Bin index,	Bin value,	Aquifer
pond,	, Initial properties,	0,	0,		1
atm,	, Initial properties,	0,	1,		2
BC_hi,	, Initial properties,	0,	2.9,		4
BC_low,	, Initial properties,	0,	2.3,		5
End of table					

1 20 22 23 25 1 1 K- 1* 1* NO /
1 23 24 25 25 1 1 K- 1* 1* NO /
1 23 23 24 24 1 1 K- 1* 1* NO /
1 19 19 20 23 1 1 K- 1* 1* NO /
1 20 21 22 22 1 1 K- 1* 1* NO /
2 25 25 9 42 1 1 K- 1* 1* NO /
2 18 18 1 35 1 1 K- 1* 1* NO /
4 10 10 8 8 1 55 I- 1* 1* NO /
4 11 11 7 7 1 55 I- 1* 1* NO /
4 12 12 6 6 1 55 I- 1* 1* NO /
4 13 13 5 5 1 55 I- 1* 1* NO /
4 15 15 4 4 1 55 I- 1* 1* NO /
4 16 16 3 3 1 55 I- 1* 1* NO /
4 17 17 2 2 1 55 I- 1* 1* NO /
4 18 18 1 1 1 55 I- 1* 1* NO /
4 18 18 1 1 1 55 I+ 1* 1* NO /
5 25 25 42 42 1 55 I- 1* 1* NO /
5 24 24 41 41 1 55 I- 1* 1* NO /
5 23 23 40 40 1 55 I- 1* 1* NO /
5 22 22 39 39 1 55 I- 1* 1* NO /
5 21 21 38 38 1 55 I- 1* 1* NO /
5 20 20 37 37 1 55 I- 1* 1* NO /
5 19 19 36 36 1 55 I- 1* 1* NO /
5 18 18 34 35 1 55 I- 1* 1* NO /
5 17 17 33 33 1 55 I- 1* 1* NO /
5 16 16 32 32 1 55 I- 1* 1* NO /
5 15 15 31 31 1 55 I- 1* 1* NO /
5 25 25 42 42 1 55 I+ 1* 1* NO /
/

EQUIL
0 2.6 -16 0 1* 0 0 0 0 /

RPTRST
BASIC=3 FLOWS /

RPTSOL
RESTART=2 FIP /

SUMMARY

RUNSUM
SEPARATE

-- Request results only at report time no results for each time-step

RPTONLY

AAQR
/

AAQT
/

SCHEDULE

AQUCHWAT
4 -0 PRESSURE 2.9 2000000000 1 0.0 /
5 -0 PRESSURE 2.3 2000000000 1 0.0 /
/

DATES
1 JAN 2008 /
/

WELSPECS
--'WELL5' is the simulation well name used to describe flow from 'well5'
--
WELL5 'GROUP 1' 10 21 1* WATER /
/

COMPDAT
WELL5 10 21 31 31 OPEN 1* 0.2094 0.45720 11.15 0.00 1* Z 3.96 /
WELL5 10 21 32 32 OPEN 1* 0.0247 0.45720 1.32 0.00 1* Z 3.96 /
WELL5 10 21 33 33 OPEN 1* 0.0181 0.45720 0.96 0.00 1* Z 3.96 /
WELL5 10 21 34 34 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 35 35 OPEN 1* 0.1933 0.45720 10.29 0.00 1* Z 3.96 /
WELL5 10 21 36 36 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 37 37 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 38 38 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 39 39 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 40 40 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 41 41 OPEN 1* 0.0193 0.45720 1.03 0.00 1* Z 3.96 /
WELL5 10 21 42 42 OPEN 1* 0.1055 0.45720 5.62 0.00 1* Z 3.96 /
/

WCONHIST
WELL5 OPEN WRAT 1* 0.88 /
/

AQUFLUX
1 0.6209 0.0 /
/

AQUCHWAT

4 -0 PRESSURE 2.9 2000000000 1 0.0 /

5 -0 PRESSURE 2.3 2000000000 1 0.0 /

/

DATES

15 JAN 2008 /

/

WCONHIST

WELL5 OPEN WRAT 1* 9.38 /

/

AQUFLUX

1 0.6279 0.0 /

/

DATES

29 JAN 2008 /

/

WCONHIST

WELL5 OPEN WRAT 1* 3.72 /

/

AQUFLUX

1 0.7949 0.0 /

/

DATES

12 FEB 2008 /

/

WCONHIST

WELL5 OPEN WRAT 1* 12.17 /

/

AQUFLUX

1 0.5155 0.0 /

/

DATES

11 MAR 2008 /

/

WCONHIST

WELL5 OPEN WRAT 1* 49.90 /

/

AQUFLUX

1 0.1365 0.0 /

/

AQUCHWAT

4 -0 PRESSURE 2.95 2000000000 1 0.0 /
5 -0 PRESSURE 2.5 2000000000 1 0.0 /
/

DATES
25 MAR 2008 /
/

WCONHIST
WELL5 OPEN WRAT 1* 62.27 /
/

AQUFLUX
1 0.1365 0.0 /
/

DATES
8 APR 2008 /
/

WCONHIST
WELL5 OPEN WRAT 1* 92.31 /
/

AQUFLUX
1 0.1365 0.0 /
/

DATES
3 JUN 2008 /
/

WCONHIST
WELL5 OPEN WRAT 1* 100.97 /
/

AQUFLUX
1 0.1365 0.0 /
/

DATES
17 JUN 2008 /
/

WCONHIST
WELL5 OPEN WRAT 1* 102.56 /
/

AQUFLUX
1 0.0 0.0 /
/

DATES
23 SEP 2008 /

/

WCONHIST

WELL5 SHUT WRAT 1* 0.00 /

/

AQUCHWAT

4 -0 PRESSURE 3 2000000000 1 0.0 /

5 -0 PRESSURE 2.4 2000000000 1 0.0 /

/

DATES

31 DEC 2008 /

/

Appendix B.

Electrode positions for the II arrays selected using informed experimental design. X-coordinates for positive and negative source and potential electrodes are included in the columns below. Z-coordinates of all electrodes are at 0 m depth.

Source		Potential	
x+	x-	x+	x-
-4.25	-0.75	-4	-1
-4.25	-0.75	-4	-0.5
-4.25	-0.75	-3.75	-1
-4.25	-0.75	-4	-1.25
-4.25	-0.75	-4	-0.25
2.25	4.5	2.5	4.25
2.25	4.5	2.5	4.75
2.25	4.5	2.5	4
2.25	4.5	2.75	4.25
2.25	4.5	2.5	5
-4.25	-2.25	-4	-2.5
-4.25	-2.25	-4	-2
-4.25	-2.25	-3.75	-2.5
-4.25	-2.25	-4	-2.75
-4.25	-2.25	-4	-1.75
-1.25	2.75	-1	2.5
-1.25	2.75	-1	3
-1.25	2.75	-1	2.25
-1.25	2.75	-0.75	2.5
-1.25	2.75	-1	3.25
-4.5	-1	-4.25	-1.25
-4.5	-1	-4.25	-0.75
-4.5	-1	-4	-1.25
-4.5	-1	-4.25	-1.5
-4.5	-1	-4.25	-0.5
2	4.25	2.25	4
2	4.25	2.25	4.5
2	4.25	2.25	3.75
2	4.25	2.5	4
2	4.25	2.25	4.75
-4	-2	-3.75	-2.25
-4	-2	-3.75	-1.75
-4	-2	-3.5	-2.25
-4	-2	-3.75	-2.5
-4	-2	-3.75	-1.5
-1	3	-0.75	2.75
-1	3	-0.75	3.25
-1	3	-0.75	2.5
-1	3	-0.5	2.75
-1	3	-0.75	3.5
-3.75	-1.75	-3.5	-2
-3.75	-1.75	-3.5	-1.5
-3.75	-1.75	-3.25	-2
-3.75	-1.75	-3.5	-2.25
-3.75	-1.75	-3.5	-1.25

-0.75	3.25
-0.75	3.25
-0.75	3.25
-0.75	3.25
-0.75	3.25
-4	-3
-4	-3
-4	-3
-4	-3
-4	-3
-1.5	3.25
-1.5	3.25
-1.5	3.25
-1.5	3.25
-1.5	3.25
-3.75	5
-3.75	5
-3.75	5
-3.75	5
-3.75	5
1.5	3
1.5	3
1.5	3
1.5	3
1.5	3
-4.5	-3.5
-4.5	-3.5
-4.5	-3.5
-4.5	-3.5
-4.5	-3.5
-2	2.75
-2	2.75
-2	2.75
-2	2.75
-2	2.75
-3.5	4.75
-3.5	4.75
-3.5	4.75
-3.5	4.75
-3.5	4.75
-0.75	0
-0.75	0
-0.75	0
-0.75	0
-0.75	0
-3.75	1
-3.75	1
-3.75	1
-3.75	1
-3.75	1
2.5	3.5
2.5	3.5
2.5	3.5
2.5	3.5
2.5	3.5

-0.5	3
-0.5	3.5
-0.5	2.75
-0.25	3
-0.5	3.75
-3.75	-2.75
-3.75	-2.5
-3.5	-2.75
-3.75	-2.25
-3.5	-2.5
-1.25	3
-1.25	3.5
-1.25	2.75
-1	3
-1.25	3.75
-3.5	4.75
-3.5	4.5
-3.25	4.75
-3.25	4.5
-3.5	4.25
1.75	2.75
1.75	3.25
1.75	2.5
2	2.75
1.75	3.5
-4.25	-3.25
-4.25	-3
-4	-3.25
-4.25	-2.75
-4	-3
-1.75	2.5
-1.75	3
-1.75	2.25
-1.5	2.5
-1.75	3.25
-3.25	4.5
-3.25	5
-3.25	4.25
-3	4.5
-3	5
-0.5	0.25
-0.5	0.5
-0.5	0.75
-0.25	0.5
-0.5	1
-3.5	0.75
-3.5	1.25
-3.25	0.75
-3.5	0.5
-3.5	1.5
2.75	3.75
2.75	4
3	3.75
2.75	4.25
3	4

REFERENCES

- Archie, G.E., 1942, The electrical resistivity log as an aid in determining some reservoir characteristics: Am. Inst. Min. Metall. Pet. Eng. Tech. Report 1422.
- Bardow, A., 2006, Optimal experimental design for ill-posed problems: Proceedings, 16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering, 173-178.
- Barth, N., and C. Wunsch, 1990, Oceanographic experimental design by simulated annealing: Journal of Physical Oceanography, **20**, 1249-1263.
- Bayless, J.W., and E.O. Brigham, 1970, Application of the Kalman filter to continuous signal restoration: GEOPHYSICS, **35**, no. 1, 2-23.
- Beard, D.C., and P.K. Weyl, 1973, Influence of texture on porosity and permeability of unconsolidated sand: The American Association of Petroleum Geologists Bulletin, **57**, no. 2, 349-369.
- Beckie, R., 1996, Measurement scale, network sampling scale, and groundwater models: Water Resources Research, **32**, no. 1, 65-76.
- Binley, A., G. Cassiani, R. Middleton, and P. Winship, 2002, Vadose zone flow model parameterization using cross-borehole radar and resistivity imaging: Journal of Hydrology, **267**, no. 3-4, 147-159.
- Blöschl, G., and M. Sivapalan, 1995, Scale issues in hydrologic modeling – A review: Hydrologic Processes, **9**, 251-290.
- Boadu, F.K., 2000, Hydraulic conductivity of soils from grain-size distribution: New models: Journal of Geotechnical and Geoenvironmental Engineering, **126**, no. 8, 739-746.
- Boyd, S., and L. Vandenberghe, 2004, Convex optimization: Cambridge University Press.
- Camporese, M., G. Cassiani, R. Deiana, and P. Saladin, 2008, Assessment of local hydraulic properties from electrical resistivity tomography monitoring of tracer test experiments: Eos Trans. AGU, **89**, no. 53, Fall Meet. Suppl., Abstract H44C-05.
- Cardell-Oliver, R., M. Kranz, K. Smettem, and K. Mayer, 2005, A reactive soil moisture sensor network: Design and field evaluation: International Journal of Distributed Sensor Networks, **1**, 149-162, doi: 10.1080/15501320590966422.
- Chambers, J.E., M.H. Loke, R.D. Ogilvy, and P.I. Meldrum, 2004, Noninvasive monitoring of DNAPL migration through a saturated porous medium using electrical impedance tomography: Journal of Contaminant Hydrology, **68**, no. 1-2, 1-22.
- Cherkeva, E., and A.C. Tripp, 1996, Optimal survey design using focused resistivity arrays: IEEE Transactions on Geoscience and Remote Sensing, **34**, no. 2, 358-366.

- Choi, M., and J.M. Jacobs, 2007, Soil moisture variability of root zone profiles within SMEX02 remote sensing footprints: *Advances in Water Resources*, **30**, 883-896. doi: 10.1016/j.advwatres.2006.07.007
- Coleman, T.F., and Y. Li, 1994, On the convergence of interior-reflective Newton methods for nonlinear optimization subject to bounds: *Mathematical Programming*, **67**, 189-224.
- Coles, D.A., and F.D. Morgan, 2009, A method of fast, sequential experimental design for linearized geophysical inverse problems: *Geophysics Journal International*, **178**, 145-158.
- Corwin, D.L, and S.M. Lesch, 2005, Characterizing soil spatial variability with apparent soil electrical conductivity I. Survey protocols: *Computers and Electronics in Agriculture*, **46**, 103-133. doi: 10.1026/j.compag.2004.11.002.
- Corwin, D.L., J. Hopmans, and G.H. de Rooij, 2006, From field- to landscape-scale vadose zone processes: Scale issues, modeling, and monitoring: *Vadose Zone Journal*, **5**, 129-139.
- Curtis, A., 1999, Optimal design of focused experiments and surveys: *Geophysics Journal International*, **136**, 205-215.
- Curtis, A., 2004a, Theory of model-based geophysical survey and experimental design: Part 1 – linear problems: *The Leading Edge*, 997-1004.
- Curtis, A., 2004b, Theory of model-based geophysical survey and experimental design: Part 2 – nonlinear problems: *The Leading Edge*, 1112-1117.
- Cushman, J.H., 1986, On measurement, scale and scaling: *Water Resources Research*, **2**, no. 2, 129-134.
- Dahlin, T., and Z. Bing, 2004, A numerical comparison of 2D resistivity imaging with 10 electrode arrays: *Geophysical Prospecting*, **52**, 379-398.
- Dahlin, T., and M.H. Loke, 1998, Resolution of 2D Wenner resistivity imaging as assessed by numerical modeling: *Journal of Applied Geophysics*, **38**, 237-249.
- Daily, W., A. Ramirez, D. LaBrecque, and J. Nitao, 1992, Electrical resistivity tomography of vadose water movement: *Water Resources Research*, **28**, 1429-1442.
- Day-Lewis, F. D., K. Singha, and A.M. Binley, 2005, Applying petrophysical models to radar traveltime and electrical resistivity tomograms: resolution-dependent limitations: *Journal of Geophysical Research*, **110**, B08206.
- Deutsch, C.V., and A.J. Journel, 1998, *GSLIB Geostatistical Software Library and User's Guide*, 2nd Edition: Oxford University Press.
- Dogan, A., and L.H. Hotz, 2005, Saturated-unsaturated 3D groundwater model. I: Development: *Journal of Hydrologic Engineering*, **10**, no. 6, 492-504.
- Eppstein, M.J., and D.E. Dougherty, 1996, Simultaneous estimation of transmissivity values and zonation: *Water Resources Research*, **32**, no. 11, 3321-3336.

- Eppstein, M. J., and D.E. Dougherty, 1998, Optimal 3-D travelttime tomography: *GEOPHYSICS*, **63**, no. 3, 1053-1061.
- Eppstein, M.J., and D.E. Dougherty, 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**, 1889-1900, doi: 10.1029/98WRR00776.
- Fetter, C.W., 1999, *Applied hydrogeology*, 4th ed. Macmillan, New York, 598 pp.
- Friedl, S., 2003, Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach: *Geophysics Journal International*, **153**, 305-316.
- Fukumori, I. and P. Malanotte-Rizzoli, 1995, An approximate Kalman filter for ocean data assimilation: An example with an idealized gulf stream model: *Journal of Geophysical Research*, **100**, no. C4, 6777-6794.
- Furman, A., A.W. Warrick, and T.P.A. Ferré, 2002, Electrical potential distributions in a heterogeneous subsurface in response to applied current: solution for circular inclusions: *Vadose Zone Journal*, **1**, 273-280.
- Furman, A., T. Ferré, and A. Warrick, 2004, Optimization of ERT surveys for monitoring transient hydrological events using perturbation sensitivity and genetic algorithms: *Vadose Zone Journal*, **3**, 1230-1239.
- Gelhar, W., 1986, Stochastic subsurface hydrology from theory to application: *Water Resources Research*, **22**, no. 9, 135S-145S.
- Gelhar, W., 1993, *Stochastic Subsurface Hydrology*: Prentice Hall, Inc.
- Goes, B.J.M., and J.A.C. Meekes, 2004, An effective electrode configuration for the detection of DNAPLs with electrical resistivity tomography: *Journal of Environmental and Engineering Geophysics*, **9**, no. 3, 127-141.
- Grayson, R.A., A.W. Western, F.H.S. Chiew and G. Bloschl, 1997, Preferred stats in spatial soil moisture patterns: Local and nonlocal controls: *Water Resources Research*, **33**, 2897-2908. doi: 10.1029/97WR02174.
- Haber, E., L. Horesh, and L. Tenorio, 2008, Numerical methods for optimal experimental design of large-scale ill-posed problems: *Inverse Problems*, **24**, 17pp.
- Haines, S.S, A. Pidlisecky, and R. Knight, 2009, Hydrogeologic structure underlying a recharge pond delineated with shear-wave seismic reflection and cone penetrometer data: *Near Surface Geophysics*, **7**, no. 5-6, 315-327.
- Halihan, T., S. Paxton, T. Fenstermaker, and M. Riley, 2005, Post-remediation evaluation of a LNAPL site using electrical resistivity imaging: *Journal of Environmental Monitoring*, **7**, 283-287.
- Huisman, J.A., C. Sperl, W. Bouten, and J.M. Verstraten, 2001, Soil water content measurements at different scales: Accuracy of time domain reflectometry and ground-

penetrating radar: *Journal of Hydrology*, **245**, 45-58. doi: 10.1016/S0022-1694(02)00239-1.

Hyndman, D.W., J.M. Harris, and S.M. Gorelick, 1994, Coupled seismic and tracer test inversion for aquifer property characterization: *Water Resources Research*, **30**, no. 7, 1965-1977.

Ines, A.M. and B.P. Mohanty, 2008, Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 1. Conceptual modeling: *Water Resources Research*, **44**, W06422. doi:10.1029/2007WR005990.

Journel, A.G., and G.J. Huijbregts, 1978, *Mining Geostatistics*: Academic Press.

Kaipio, J.P., P.A. Karjalainen, E. Somersalo, and M. Vauhkonen, State estimation in time-varying electrical impedance tomography: *Annals of the New York Academy of Sciences*, **873**, no. 1, 430-439.

Kemna, A., B. Kulesa, and H. Vereecken, 2002, Imaging and characterization of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models: *Journal of Hydrology*, **267**, no. 3-4, 125-146.

Kim, K.Y., B.S. Kim, M.C. Kim, Y.J. Lee, and M. Vauhkonen, 2001, Image reconstruction in time-varying electrical impedance tomography based on the extended Kalman filter: *Measurement Science Technology*, **12**, 1032-1039.

Kim, K.Y., S.I. Kang, M.C. Kim, Y.J. Lee, and M. Vauhkonen, 2002, Dynamic image reconstruction in electrical impedance tomography with known internal structures: *IEEE Trans. on Magnetics*, **38**, no. 2, 1301-1304.

Knight, R., P. Tercier, and J. Irving, 2004, The effect of vertical measurement resolution on the correlation structure of a ground penetrating radar reflection image: *Geophysical Research Letters*, **31**, L21607. doi: 10.1029/2004GL021112.

Knight, R., and A.L. Endres, 2005, An introduction to rock physics for near-surface applications *in* D. Butler, ed., *Near-surface geophysics, volume 1: Concepts and fundamentals*: Society of Exploration Geophysicists, 31-70.

Kowalksy, M.B., S. Finsterle, J. Peterson, S. Hubbard, Y. Rubin, E. Majer, A. Ward, and G. Gee, 2005, Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrologic data: *Water Resources Research*, **41**, W11425. doi: 10.1029/2005WR004237.

Labrecque, D., and X. Yang, 2001, Difference inversion of ERT data: A fast inversion method for 3-D in situ monitoring: *Journal of Environmental and Engineering Geophysics*, **6**, no. 2, 83-89.

Labrecque, D., G. Heath, R. Sharpe, and R. Versteeg, 2004, Autonomous monitoring of fluid movement using 3-D electrical resistivity tomography: *Journal of Environmental and Engineering Geophysics*, **9**, no. 3, 167-176.

Lear, J., Memo to Rosemary Knight RE: Summary of methods used to calculate average textures to populate PVMF2K, Balance Hydrologics, Monterey, CA, January 20, 2009.

- Lehman, H., 1994, Potential representation by independent configurations on a multi-electrode array: *Geophysics Journal International*, **120**, 331-338.
- Lehikoinen, A., S. Finsterle, A. Voutilainen, M.B. Kowalsky, and J.P. Kaipio, 2009, Dynamical inversion of geophysical ERT data: State estimation in the vadose zone: *Inverse Problems in Science and Engineering*, **17**, no. 6, 715-736.
- Lesmes D.P. and Friedman S.P., 2005. Relationships between the electrical and hydrogeological properties of rocks and soils: *Hydrogeophysics*, **50**, 87-128.
- Loke, M.H., 1999, Electrical imaging surveys for environmental and engineering studies: A practical guide to 2-D and 3-D surveys, accessed 22 January 2009. <http://www.geomatrix.co.uk/appnotes.htm>.
- Looms, M.C., K.H. Jensen, A. Binley, and L. Nielsen, 2008, Monitoring unsaturated flow and transport using cross-borehole geophysical methods: *Vadose Journal*, **7**, 227-237.
- Maurer, H., and D.E. Boerner, 1998, Optimized and robust experimental design: a non-linear application to EM sounding: *Geophysics Journal International*, **132**, 458-468.
- Maurer, H., D.E. Boerner, and A. Curtis, 2000, Design strategies for electromagnetic geophysical surveys: *Inverse Problems*, **16**, 1097-1117.
- Michot, D., Y. Benderitter, A. Dorigny, B. Nicoullaud, D. King, and A. Tabbagh, 2003, Spatial and temporal monitoring of soil water content with an irrigated corn crop cover using surface electrical resistivity tomography: *Water Resources Research*, **39**, no. 5, 1138. doi: 10.1029/2002WR001581.
- Miller, C.R., and P.S. Routh, 2007, Resolution analysis of geophysical images: Comparison between point spread function and region of data influence measures: *Geophysical Prospecting*, **55**, 835-852.
- Mitchell, V., A. Pidlisecky, and R. Knight, 2010, Informed experimental design for electrical resistivity imaging: *Near Surface Geophysics* (submitted).
- Njoku, E.G., and D. Entekahbi, 1996, Passive microwave remote sensing of soil moisture: *Journal of Hydrology*, **184**, 101-129
- Oldenburg, D.W., and Y. Li., 1999, Estimating depth of investigation in dc resistivity and IP surveys: *GEOPHYSICS*, **64**, 403-416.
- Oldenborger, G.A, P.S. Routh, and M.D. Knoll, 2005, Sensitivity of electrical resistivity tomography data to electrode position errors: *Geophysics Journal International*, **163**, 1-9.
- Oldenborger, G.A, P.S. Routh, and M.D. Knoll, 2007, Model reliability for 3D electrical resistivity tomography: Application of the volume of investigation index to a time-lapse monitoring experiment: *GEOPHYSICS*, **72**, no 4., F167-F175.

- Parada, L. M., and X. Liang, 2004, Optimal multiscale Kalman filter for assimilation of near-surface soil moisture into land surface models: *Journal of Geophysical Research*, **109**, D24109. doi:10.1029/2004JD004745.
- Parasnis, D.S., 1986, *Principles of applied geophysics*, 4th ed.: Chapman and Hall.
- Park, S.K., and G.P. Van, 1991, Inversion of pole-pole data for 3-D resistivity structure beneath arrays of electrodes: *GEOPHYSICS*, **56**, 951-960.
- Park, S., 1998, Fluid migration in the vadose zone from 3-D inversion of resistivity monitoring data: *GEOPHYSICS*, **63**, no. 1, 41-51.
- Pazman, A., 1986, *Foundations of Optimum Experimental Design*: D. Reidel Publishing Company.
- Pidlisecky, A., E. Haber, and R. Knight, 2007, RESINVM3D: A 3D resistivity inversion package: *GEOPHYSICS*, **72**, no. 2, H1-H10. doi: 10.1190/1.2402499.
- Pidlisecky, A., and R. Knight, 2008, FW2_5D: A MATLAB 2.5-D electrical resistivity modeling code: *Computers and Geosciences*, **34**, 1645-1654.
- Pidlisecky, A., and R. Knight, 2010, The use of wavelet analysis to derive infiltration rates from 1D resistivity records: *Vadose Zone* (submitted).
- Pukelsheim, F., 1993, *Optimal Design of Experiments*: John Wiley & Sons, Inc.
- Quan, Y., and J.M. Harris, 2008, Stochastic seismic inversion using both waveform and travelttime data and its application to time-lapse monitoring: *SEG Expanded Abstracts*, **27**, 1915- 1919.
- Rabinowitz, N., and D.M. Steinberg, 1990, Optimal configuration of a seismographic network: a statistical approach: *Bulletin of the Seismologic Society of America*, **80**, 1033-1036.
- Rawls, W.J., D.L. Brakensiek, and K.E. Saxton, 1982, Estimation of soil water properties: *Transaction of the American Society of Agricultural Engineers*, **25**, no. 5, 1316-1320.
- Rea, J. and R. Knight, 1998, Geostatistical analysis of ground-penetrating radar data: A means of describing spatial variation in the subsurface: *Water Resources Research*, **34**, no. 3, 329- 339.
- Reichle, R. H., W. T. Crow, and C. L. Keppenne, 2008, An adaptive ensemble Kalman filter for soil moisture data assimilation: *Water Resources Research*, **44**, W03423. doi:10.1029/2007WR006357.
- Robinson, D.A., A. Binley, N. Crook, F.D. Day-Lewis, T.P.A. Ferré, V.J.S. Grauch, R. Knight, M. Knoll, V. Lakshmi, R. Miller, J. Nyquist, L. Pellerin, K. Singha, and L. Slater, 2008, Advancing process-based watershed hydrological research using near-surface geophysics: a vision for, and review of, electrical and magnetic geophysical methods: *Hydrologic Processes* **22**, 3604-3635.
- Robinson, D.A., C.S. Campbell, J.W. Hopmans, B.K Hornbuckle, S.B. Jones, R. Knight, F. Ogden, J. Selker, and O. Wendroth, 2008, Soil moisture measurements for

ecological and hydrological watershed scale observatories: A review: *Vadose Zone Journal*, 358-389. doi: 10.2136/vzj2007.0143.

Rodriguez-Iturbe, I., and J.M Mejia, 1974, On the transformation of point rainfall to areal rainfall: *Water Resources Research*, **10**, no. 4, 729-735.

Rodriguez-Iturbe, I., G.K. Vogel, R. Rigon, D. Entekhabi, F. Castelli, and A. Rinaldo, 1995, On the spatial organization of soil moisture fields: *Geophysical Research Letters*, **22**, 2757-2760. doi:10.1029/95GL02779.

Sheets, K.R., and J.M.H Hendrickx, 1995, Noninvasive soil-water content measurement using electromagnetic induction: *Water Resources Research*, **31**, 2401-2409. doi: 10.1029/95WR01949.

Singha, K., and S. M. Gorelick, 2005, Saline tracer visualized with three-dimensional electrical resistivity tomography: Field-scale spatial moment analysis: *Water Resources Research*, **41**, W05023. doi:10.1029/2004WR003460.

Singha, K., and S. Gorelick, 2006, Effects of spatially variable resolution on field-scale estimates of tracer concentration from electrical inversions using Archie's Law: *GEOPHYSICS*, **71**, no. 3, 83-91.

Skøien, J.O., and G. Blöchl, 2006a, Sampling scale effects in random field and implications for environmental monitoring: *Environmental Monitoring and Assessment*, **114**, 521-552.

Skøien, J.O., and G. Blöchl, 2006b, Scale effects in estimating the variogram and implications for soil hydrology: *Vadose Zone Journal*, **5**, 153-167.

Slater, L., A. Binley, R. Versteeg, G. Cassiani, R. Birken, and S. Sandberg, 2002, A 3D ERT study of solute transport in a large experimental tank: *Journal of Applied Geophysics*, **49**, no. 4, 211-229.

Spitzer, K., 1998, The three-dimensional DC sensitivity for surface and subsurface sources: *Geophysics Journal International*, **134**, 736-746.

Stummer, P., H. Maurer, H. Horstmeyer, and A.G. Green, 2002, Optimization of DC resistivity data acquisition: Real-time experimental design and a new multi-electrode system: *IEEE Transactions on Geoscience and Remote Sensing*, **40**, 2727-2735.

Stummer, P., H. Maurer, and A.G. Green, 2004, Experimental design: Electrical resistivity data sets that provide optimum subsurface information: *GEOPHYSICS*, **69**, 120-139.

Tikhonov, A.N., and V.Y. Arsenin, 1977, *Solutions of ill-posed problems*: W.H. Winston and Sons.

Trigo, F.C., R. Gonzalez-Lima, M.B.P. Amato, 2004, Electrical impedance tomography using the extended Kalman filter: *IEEE Transactions on Biomedical Engineering*, **51**, no. 1, 72-81.

- Van den Berg, J., A. Curtis, and J. Trampert, 2003, Optimal nonlinear Bayesian experimental design: an application to amplitude versus off-set experiments: *Geophysics Journal International*, **155**, 411-421.
- Van Genuchten, M.T., and D.R. Nielsen, 1985, On describing and predicting the hydraulic properties of unsaturated soils: *Annals of Geophysics*, **3**, no. 5, 615-628.
- Vauhkonen, M. P.A. Karjalainen, and J.P. Kaipio, 1998, A Kalman filter approach to track fast impedance changes in electrical impedance tomography: *IEEE Transaction on Biomedical Engineering*, **45**, no. 4, 486-493.
- Vereecken, H., J.A. Huisman, H. Boga, J. Vanderborght, J.A. Vrugt, and J.W. Hopmans, 2008, On the value of soil measurements in vadose zone hydrology: A review: *Water Resources Research*, **44**. doi: 10.1029/2008WRR006829.
- Verron, J., L. Gourdeau, D. T. Pham, R. Murtugudde, and A. J. Busalacchi, 1999, An extended Kalman filter to assimilate satellite altimeter data into a nonlinear numerical model of the tropical Pacific Ocean: Method and validation: *Journal of Geophysical Research*, **104**, no. C3, 5441–5458.
- Vrugt, J.A., P.H. Stauffer, T. Wöhling, B.A. Robinson, V.V. Vesslinov, 2008, Inverse modeling of subsurface flow and transport properties: A review with new developments: *Vadose Zone Journal*, **7**, no. 2, 843-864.
- Webster, R., and M.A. Oliver, 2001, Sample adequately to estimate variograms of soil properties: *Journal of Soil Sciences*, **43**, 177-192.
- Western, A.W., G. Bloschl, and R.B. Grayson, 1999, On the spatial scaling of soil moisture: *Journal of Hydrology*, **217**, 203-224.
- Western, A.W., and R.B. Grayson, 1998, The Tarrawarra data set: Soil moisture patterns, soil characteristics, and hydrological flux measurements: *Water Resources Research*, **34**, 2765-2768. doi: 10.1029/98WR01833.
- Wilkinson, P.B., J.E. Chambers, P.I. Meldrum, R.D. Ogilvy, and S. Caunt, 2006, Optimization of array configurations and panel combinations for the detection and imaging of abandoned mineshafts using 3D cross-hole electrical resistivity tomography: *Journal of Environmental and Engineering Geophysics*, **11**, 213-221.
- Xu, B., and M. Noel, 1993, On the completeness of data sets with multi-electrode systems for electrical resistivity survey: *Geophysical Prospecting*, **41**, 791-801.
- Yeh, T.C.J., M. Ye, and R. Khaleel, 2005, Estimation of effective unsaturated hydraulic conductivity tensor using spatial moments of observed moisture plume: *Water Resources Research*, **41**, W03014-W03026. doi: 10.1029/2004WR003736.
- Zhou, C., L. Liu, and J.W. Lane, Jr., 2001, Nonlinear inversion of borehole-radar tomography data to reconstruct velocity and attenuation distribution in earth materials: *Journal of Applied Geophysics*, **47**, no. 3-4, 271-284.