Final Report

Fusion of Tomography Tests for DNAPL Source Zone Characterization: Technology Development and Validation

SERDP Project ER-1365

July 2011

Walter Illman
University of Waterloo

Tian-Chyi Jim Yeh
University of Arizona

This document has been cleared for public release

SERDP
DOD • EPA • DOE

TABLE OF CONTENTS **Page Number**

Title page i
Table of Contents ii
List of Tables xii
List of Figures xv
List of Acronyms xxvii
Acknowledgements xxviii

ABSTRACT 1

1. OBJECTIVES 2
1.1 Task 1: Algorithm Development 2
1.2 Task 2: Laboratory Experimental and Numerical Modeling Studies 2
1.3 Task 3: Hydraulic and tracer tomography algorithms that considers the effects of the unsaturated zone 2
1.4 Task 4: Validation of algorithms developed in Task 3 2

2. BACKGROUND 4
2.1 Direct characterization techniques 4
2.2 Indirect characterization techniques 4
2.3 Hydrological inversion for detecting hydraulic heterogeneity 5
2.4 Partitioning tracer inversion for detecting DNAPL distribution 6
2.5 Research Plan 7
2.4 Report organization 7

TASK 1: ALGORITHM DEVELOPMENT
3. CHARACTERIZATION OF AQUIFER HETEROGENEITY USING TRANSIENT HYDRAULIC TOMOGRAPHY 8
3.1 Introduction 8
3.2 Method 10
3.2.1 Groundwater Flow in Three-dimensional Saturated Media 10
3.2.2 Sequential Successive Linear Estimator (SSLE) 10
3.3 Numerical Examples 14
3.3.1 One-Dimensional Flow 14
3.3.2 Three-Dimensional Heterogeneous Aquifer 21

4. ANALYSIS OF HYDRAULIC TOMOGRAPHY USING TEMPORAL MOMENTS OF DRAWDOWN-RECOVERY DATA 26
4.1 Introduction 26
4.2 Derivation of Moment Equations 27
4.3 Sequential Inverse Algorithm 30
4.4 Evaluation of Moment Sensitivities 31
4.5 Numerical Examples 33
4.6 Results and Discussion 38

5. A NEW ESTIMATOR AND A GUIDE FOR HYDRAULIC TOMOGRAPHY ANALYSIS 40
5.1 Introduction 40
5.2 Methodology 40
5.2.1 Governing Groundwater Flow Equations 40
5.2.2 Simultaneous Successive Linear Estimator 41
5.2.3 Preprocessing Data for SimSLE 44
5.2.3.1 Diagnosis of Bias 44
5.2.3.2 Wavelet Denoising 44
5.2.4 Signal to Noise Ratio 45
5.2.5 Estimation of Effective Properties and Variances 46
5.3 Applications to a synthetic aquifer 46
5.3.1 Description of the Aquifer and HT Test 46
5.3.2 Performance Assessment 48
5.3.3 Results 49
5.4 Application to a laboratory sandbox experiment 57
5.5.1 Preprocessing Data 57
5.5.2 Results 58
5.6 Discussion 63

6. HYDRAULIC/PARTITIONING TRACER TOMOGRAPHY FOR CHARACTERIZATION OF DENSE NONAQUEOUS PHASE LIQUID SOURCE ZONES 64
6.1 Introduction 64
6.2 Derivation of Inverse Algorithm 66
6.2.1 Governing equations 66
6.2.2 SSLE Inverse Method 67
6.2.3 Evaluation of Covariance and Cross Covariance 69
6.2.4 Calculation of Sensitivities 70
6.3 Numerical Examples 71

7. ANALYSIS OF TRACER TOMOGRAPHY TEMPORAL MOMENTS OF TRACER BREAKTHROUGH CURVES 79
7.1 Introduction 79
7.2 Methodology 80
7.2.1 Temporal moment equations for a partitioning tracer 80
7.2.1.1 Governing equations 81
7.2.1.2 Temporal moment equations for tracer BTC 81
7.2.2 Numerical method for moment equations 82
7.2.3 Inverse method 83
7.2.3.1 Implementation of temporal moments of tracer in SSLE 83
7.2.3.2 Calculation of sensitivities 85
7.3 Synthetic case study 86
7.3.1 Comparison with inversion using tracer concentrations 86
7.3.2 Tracer first moments for estimating hydraulic conductivity 89

7.3.3 Cross correlation of first temporal moment of tracer to parameters 91

TASK 2: LABORATORY EXPERIMENTAL AND NUMERICAL MODELING STUDIES

8. LABORATORY SANDBOX VALIDATION OF STEADY STATE HYDRAULIC TOMOGRAPHY 94
8.1 Introduction 94
8.2 Sandbox Description 95
8.3 Description of the Forward and Inverse Models 96
8.3.1 Forward Model 96
8.3.2 Inverse Model 96
8.4 Estimates of Hydraulic Conductivity for Validation of K Tomograms 96
8.5 Data Diagnostics 99
8.5.1 Examination of Drawdown Records and their Animations 99
8.5.2 Diagnostic Forward Modeling 99
8.6 Inputs to the Inverse Model 100
8.7 Results 102
8.7.1 Inverse Modeling of Synthetic Cross-Hole Tests in Sandbox 102
8.7.2 Inverse Modeling of Real Cross-Hole Tests in Sandbox 104
8.8 Multi-Method and Multiscale Validation of *K* Tomograms 105
8.8.1 Multi-Method Validation Approach 1: Visual Comparison 105
8.8.2 Multi-Method Validation Approach 2: Validation of *K* Tomogram with Additional Cross-Hole Hydraulic Tests not used in the Inversion 106
8.8.3 Multi-Method Validation Approach 3: Comparison of Statistical Moments 106
8.8.4 Multi-Method Validation Approach 4: Comparison of Local *K* Values from the Real *K* Tomogram to the Reference *K* Tomogram 107
8.8.5 Multi-Method Validation Approach 5: Comparison of Local *K* Values from the Real *K* Tomogram to *K* from Core and Single-Hole Tests 108
8.9 Discussion 110

9. PRACTICAL ISSUES IN IMAGING HYDRAULIC CONDUCTIVITY THROUGH HYDRAULIC TOMOGRAPHY
9.1 Introduction 112
9.2 Numerical Simulation Methods 113
9.2.1 Inverse modeling approach 113
9.2.2 Inputs to the inverse model 113
9.3 Generation of synthetic & laboratory hydraulic test data 114
9.3.1 Synthetic hydraulic test data 114
9.3.2 Laboratory hydraulic test data 115
9.4 Estimation of *K* from synthetic and laboratory experimental data sets 116
9.4.1 Estimation of synthetic *K* data 116
9.4.2 Estimation of laboratory *K* data 117

9.5 Generation of reference *K* tomograms using synthetic and real pumping test data 118
9.6 Effects of pumping rate, order of test data included in SSLE and conditioning on synthetic and real *K* tomograms 119
9.6.1 The effect of pumping rate and signal-to-noise ratio 119
9.6.2 The effect of order of test data included sequentially in the SSLE algorithm 121
9.6.3 The effect of conditioning using core *K* data 123
9.6.4 The effect of conditioning using slug *K* data 123
9.6.5 The effect of conditioning using single-hole *K* data 125
9.7 How good are the *K* tomograms? 126

10. LABORATORY SANDBOX VALIDATION OF TRANSIENT HYDRAULIC TOMOGRAPHY 129
10.1 Introduction 129
10.2 Methods for Characterization of the Sandbox 129
10.3 Inverse Model Description 130
10.4 Inverse Model Parameters 130
10.4.1 Hydraulic Parameters *K* and S_s 131
10.4.2 Variance and Correlation Scales 131
10.4.3 Transient Hydraulic Head Data 131
10.5 Results from Transient Hydraulic Tomography 132
10.6 Comparisons of *K* and S_s Fields from Different Tests and Analyses 134
10.6.1 Visual Comparisons of Patterns of Heterogeneity of Different Tests and Analyses 134
10.6.2 Comparison of Statistical Moments 135
10.6.3 Comparison of Local Values 137
10.6.4. Comparison of *K* Tomogram obtained from Steady State Hydraulic Tomography 139
10.7 Validation of *K* and S_s Tomograms 139

11. LABORATORY SANDBOX VALIDATION OF HYDRAULIC TOMOGRAPHY THAT USES THE TEMPORAL MOMENTS OF DRAWDOWN-RECOVERY DATA
11.1 Introduction 142
11.2 Numerical experiments 142
11.2.1 Purpose of numerical experiments 142
11.2.2 Inverse model description and computational approach/platform 143
11.2.3 Description of numerical experiments and inverse model parameters 144
11.3 Results from inverse modeling of synthetic cross-hole pumping test data 146
11.3.1 Results from THT 146
11.3.2 Results from HT-m 147
11.3.3 Comparison of tomograms and simulation times 149
11.3.4 Quantitative evaluation of synthetic inversion results 149
11.4 Laboratory sandbox data and its inverse modeling 151

11.4.1 Description of laboratory sandbox and instrumentation used to conduct real cross-hole pumping tests 151
11.4.2 Description of real cross-hole pumping tests used for hydraulic tomography in the sandbox 151
11.4.3 Pre-processing of drawdown-recovery data 152
11.4.4 Inverse model and parameters 152
11.5 Results from inverse modeling of cross-hole pumping tests conducted in a laboratory sandbox 153
11.5.1 K and S_s tomograms from the inversion of laboratory sandbox data 153
11.5.2 Visual evaluation of computed tomograms 155
11.5.3 Joint validation of K and S_s tomograms 155
11.6 Discussion 157
11.6.1 Comparison of K tomogram from HT-m to SSHT 157
11.6.2 Comparison of K and S_s tomograms from HT-m to THT 158

12. COMPARISON OF AQUIFER CHARACTERIZATION APPROACHES THROUGH STEADY-STATE GROUNDWATER MODEL VALIDATION: A CONTROLLED LABORATORY SANDBOX STUDY 160
12.1 Introduction 160
12.2 Sandbox and synthetic heterogeneous aquifer construction 162
12.2.1 Sandbox construction 162
12.2.2 Synthetic aquifer construction through cyclic flux of sediment laden water 164
12.3 Characterization of the synthetic heterogeneous aquifer treating the medium to be homogeneous 166
12.3.1 Constant head permeameter analysis of core samples 167
12.3.2 Single-hole pumping tests 167
12.3.3 Cross-hole pumping tests 167
12.3.4 Bidirectional flow-through tests 168
12.3.5 Statistical analysis of test data 168
12.4 Geostatistical analysis of small scale K estimates 169
12.5 Steady-state hydraulic tomography analysis of selected cross-hole tests 170
12.5.1 Inverse model description 170
12.5.2 Input parameters and cross-hole tests used 171
12.5.3 Computation of K tomograms and its visual assessment 172
12.6 Forward simulations of independently conducted cross-hole pumping and flow-through tests 174
12.6.1 Use of independent cross-hole tests to validate homogeneous and heterogeneous K fields 176
12.6.2 Use of bidirectional flow-through tests to validate various homogeneous and heterogeneous K fields 178
12.7 Discussion 180
12.7.1 On model validation 180
12.7.2 On the accuracy of K estimates and its impact on model validation 181

12.7.3 Why does hydraulic tomography yield improved K fields over traditional methods 181

13. CAPTURING HETEROGENEITY IN GROUNDWATER FLOW PARAMETERS: COMPARISON OF APPROACHES THROUGH CONTROLLED SANDBOX EXPERIMENTS 183
13.1 Introduction 183
13.1.1 Characterization methods of subsurface heterogeneity 183
13.1.2 Methods for capturing spatial heterogeneity 184
13.1.3 Goal of this study 185

13.2 Experimental Methods 186
13.2.1 Sandbox and synthetic heterogeneous aquifer description 186
13.2.2 Characterization of synthetic heterogeneous aquifer (single-hole pumping tests) 188
13.2.3 Characterization of synthetic heterogeneous aquifer (cross-hole pumping tests) 189
13.3 Geostatistical analysis of single-hole pumping test data 189
13.4 Construction of a geological model 190
13.5 Cross-hole pumping tests and transient hydraulic tomography analysis 191
13.5.1 Input parameters and cross-hole tests used 191
13.5.2 Computation of *ln-K* and *ln-*S_s tomograms 192
13.5.3 Statistical summary of results 194
13.5.4 Visual comparison of ln-K tomogram to the deposits 195
13.5.5 Comparison of *ln-K* tomograms: transient vs. steady state hydraulic tomography 196
13.6 Performance assessment of results 197
13.6.1 Performance assessment by simulating individual tests 197
13.6.2 Predictability of transient drawdown curves 202

14. TEMPORAL MOMENT ANALYSIS OF PARTITIONING TRACERS IN SELECTED LABORATORY SANDBOX AQUIFERS 205
14.1 Introduction 205
14.2 Temporal moment analysis of partitioning tracer tests 206
14.3 Experimental Methods 209
14.3.1 Sandbox descriptions 209
14.3.2 TCE source zone creation in SBs 3-1, 3-2 and 1-2 209
14.3.3 Determination of tracer partitioning coefficients through batch tests 210
14.3.4 Determination of partitioning tracer concentrations 211
14.3.5 Determination of conservative tracer concentrations 212
14.3.6 Partitioning Tracer Tests on SBs 3-1, 3-2 and 1-2 212
14.3.7 Extraction of cores and its analyses 215
14.4 Results 219
14.4.1 Batch test results 219
14.4.2 Partitioning tracer test results 221

14.4.3 Temporal Moment Analysis (MOM)-Estimations of S_N and DNAPL volumes 222
14.4.3.1 SB 3-1 222
14.4.3.2 SB 3-2 227
14.4.4.3 SB 1-2 231
14.4.4 Core analysis-estimations of S_N and DNAPL volumes 236
14.4.4.1 SB 3-2 236
14.4.4.2 SB 1-2 238
14.4 Discussion 241
14.4.1 Influence of partitioning coefficient on saturation estimates 241
14.4.2 Influence of DNAPL architecture on saturation estimates 243
14.4.2.1 SB 3-1 243
14.4.2.2 SB 3-2 244
14.4.2.3 SB 1-2 244
14.4.3 Influence of flow and sampling on saturation estimates 245
14.4.4 Influence of TCE mass loss due to dissolution on saturation estimates 245

15. HYDRAULIC AND PARTITIONING TRACER TOMOGRAPHY FOR TRICHLOROETHYLENE SOURCE ZONE CHARACTERIZATION: SMALL-SCALE SANDBOX EXPERIMENTS 247
15.1 Introduction 247
15.2 Construction of heterogeneous aquifer and its hydraulic characterization
15.2.1 Sandbox design and construction 249
15.2.2 Porous media used to pack the sandbox 250
15.2.3 Hydraulic characterization prior to TCE spill 251
15.3 Creation of TCE source zone 251
15.4 Characterization of TCE Source zone 252
15.4.1 Determination of partitioning coefficient through batch tests 252
15.4.2 Partitioning tracer tests 252
15.4.3 Direct sampling through coring 254
15.5 Inverse modeling of dipole cross-hole and partitioning tracer tests 255
15.5.1 Description of algorithms 255
15.5.2 Inverse modeling of dipole cross-hole tests 256
15.4.3 Inverse modeling of dipole partitioning tracer tests 256
15.6 Results 257
15.6.1 *K* tomogram and its validation 257
15.6.2 Computation of TCE saturation tomograms using the *K* tomogram 259
15.7 Discussion 261

16. COMPARISON OF HETEROGENEITY CHARACTERIZATION METHODS FOR IMPROVED PREDICTIONS OF SOLUTE TRANSPORT: LABORATORY SANDBOX EXPERIMENTS 264
16.1 Introduction 264
16.2 Experimental methods 266
16.2.1 Sandbox and synthetic aquifer construction 266

16.2.2 Dipole tracer test
16.2.2.1 Description of dipole tracer test 268
16.2.2.1 Description of water sample analysis 269
16.2.3 Aquifer characterization methods 269
16.2.3.1 Permeameter analysis of core samples 269
16.2.3.1 Dipole cross-hole tests 269
16.3 Heterogeneity characterization methods 270
16.3.1 Geostatistical analysis and kriging of core *K* estimates 270
16.3.2 Estimates of effective *K* and macrodispersivity 271
16.3.3 Hydraulic tomography analysis of dipole cross-hole tests 272
16.3.3.1 Inverse modeling approach 272
16.3.3.2 Validation of the *K* tomograms using additional tests 275
16.4 Results: Dipole tracer test 276
16.5 Prediction of dipole tracer test 279
16.5.1 Description of tracer transport simulations 279
16.5.2 Groundwater flow results 279
16.5.3 Dipole tracer transport results 280
16.5.3.1 Simulated versus observed breakthrough curves 282
16.5.3.2 Temporal moment analysis 283
16.6 Discussion 286

17. HYDRAULIC AND PARTITIONING TRACER TOMOGRAPHY EXPERIMENTS IN AN INTERMEDIATE SCALE SANDBOX
17.1 Introduction 288
17.2 Sandbox and aquifer description 288
17.3 Dipole hydraulic tomography survey results 289
17.4 Description of PCE spill 291
17.5 Determination of partitioning coefficient through batch tests 291
17.6 Description of partitioning tracer tests 291
17.7 Direct sampling of PCE through excavation 295
17.8 Plans for future analysis 296

TASK 3: DEVELOPMENT OF HYDRAULIC AND TRACER TOMOGRAPHY ALGORITHMS THAT CONSIDERS THE EFFECT OF THE UNSATURATED ZONE
18. A REVISIT OF DRAWDOWN BEHAVIOR DURING PUMPING IN UNCONFINED AQUIFERS
18.1 Introduction 297
18.2 Methodology 299
18.3 Results and Discussions 302
18.3.1 Flow to a Well in a Homogeneous Aquifer 302
18.3.1.1 Drawdown-Time Curve 302
18.3.1.2 Rate of Change in Storage 303
18.3.1.3 Early Times 303
18.3.1.4 Intermediate Times 304
18.3.1.5 Late Times 305
18.3.2 1-D Vertical Variably Saturated Flow 305

18.3.2.1 Effects of the Exponential Model vs. the VGM model 307
18.3.2.2 Effects of Hydraulic Conductivity Curve vs. Moisture Retention Curve 309
18.3.2.3 Role of an initially unsaturated zone 310
18.3.3 Flow to a Well in a Heterogeneous Aquifer 312
18.3.3.1 Drawdown-Time Curve 312
18.3.3.2 Cross-correlation analysis 315

19. HYDRAULIC TOMOGRAPHY TO CHARACTERIZE THE HETEROGENEITY OF UNCONFINED AQUIFERS 319
19.1 Introduction 319
19.2 Methodology 321
19.3 Numerical Examples 323
19.3.1 One-Dimensional Aquifer 325
19.3.2 Sensitivity Analysis 326
19.3.3 Three Dimensional Numerical Example 328
19.4 Discussion 332

TASK 4: VALIDATION OF ALGORITHMS DEVELOPED IN TASK 3 USING A NEW SANDBOX WITH MONITORING IN BOTH THE UNSATURATED AND SATURATED ZONES

20. FLOW TO A WELL IN A HETEROGENEOUS UNCONFINED AQUIFER: INSIGHTS FROM AN INTERMEDIATE SCALE SANDBOX 334
20.1 Introduction 334
20.2 Experimental setup 335
20.2.1 Sandbox design and instrumentation 335
20.2.2 Packing of sandbox and types of sands used 336
20.3 Characterization of aquifer heterogeneity 337
20.3.1 Characterization of K and S_s 337
20.3.2 Estimation of Effective K and S_s 338
20.3.3 Estimation of Heterogeneous K and S_s 339
20.3.4 Estimation of θ_s, θ_r, α and n 343
20.4 Unconfined pumping test 344
20.5 Variably saturated flow modeling of pumping test in an unconfined aquifer 348
20.5.1 Case 1: Homogeneous saturated and unsaturated zone parameters 349
20.5.2 Case 2: Heterogeneous saturated parameters and homogeneous unsaturated zone parameters 351
20.5.3 Case 3: Heterogeneous saturated parameters and heterogeneous unsaturated zone parameters 353
20.6 Discussion 355

21. FINDINGS AND CONCLUSIONS 357
21.1 Findings and Conclusions from Task 1: Algorithm Development 357

21.2 Findings and Conclusions from Task 2: Laboratory Experimental and Numerical Modeling Studies 362

21.3 Findings and Conclusions from Task 3: Development of Hydraulic and Tracer Tomography Algorithms That Considers the Effect of the Unsaturated Zone 373

20.4 Findings and Conclusions from Task 4: Validation of Algorithms Developed in Task 3 Using a New Sandbox with Monitoring in Both the Unsaturated and Saturated Zones 375

LITERATURE CITED 377

APPENDICES 394

Appendix A: White Paper - Cost Estimate for Implementing the Hydraulic, Conservative Tracer and Partitioning Tracer Tomography Tests to Delineate DNAPL Source Zones in a Sandy Aquifer (submitted on 7/15/2004). 394

Appendix B: White Paper - Definition of Success, Uncertainty Evaluation, and Field Protocol for Implementing the Hydraulic, Conservative Tracer and Partitioning Tracer Tomography Tests to Delineate DNAPL Source Zones (submitted on 7/15/2004) 401

Appendix C. Derivation of Moment Sensitivities using the Adjoint State Method 408

Appendix D. Partitioning coefficients for PCE and selected tracers 411

Appendix E. List of publications 415

LIST OF TABLES

Table 3.1: Comparison of statistical properties of the 3-D aquifer. 25
Table 4.1: Mean and variance of estimation errors. 34
Table 5.1: Means and variances of the generated *lnK* and *lnSs* field in each unit of the synthetic aquifer and over the entire aquifer. 47
Table 5.2: Estimated effective hydraulic properties, and estimated spatial variances of the properties in the synthetic aquifer, using noise-free, noisy, and denoised hydrographs. The number of iterations required to reach to the convergence of the solution is listed in the last column. 50
Table 5.3: Unconditional L1 and L2 norms of the head based on equation (5.18) for the noise-free, noisy, and denoised cases of the synthetic aquifer experiments. 52
Table 5.4: The performance assessment statistics of results from the synthetic aquifer. 55
Table 5.5: Means and variances of true and estimated hydraulic properties of each zone using noise-free, noisy, and denoised hydrographs. 56
Table 8.1: Summary of hydraulic properties determined from core, slug, cross-hole pumping test data and flow-through experiments. 98
Table 8.2: L1 and L2 norms before and after data modifications. 101
Table 8.3: Input data for inverse modeling of 6 pumping tests in sandbox. 102
Table 8.4: Conditional mean and variance estimates for the inversions of the real data with the number of cross-hole tests used in the analysis. 107
Table 9.1: Geometric mean values of K (cm/s) determined from core samples taken from the sandbox. 114
Table 9.2: Summary of pumping rates and duration of cross-hole pumping tests used for hydraulic tomography and its validation. Note that tests with pumping taking place at ports 2, 5, 14, 17, 32, 35, 44, and 47 were used in hydraulic tomography, while the test at port 46 was used for validation. 115
Table 9.3: Summary of hydraulic properties determined from core, slug, and single-hole pumping test data through synthetic simulations. 116
Table 9.4: Summary of hydraulic properties determined from core, slug, single-hole, cross-hole pumping test data and flow-through experiments conducted in the laboratory. 118
Table 10.1: Summary of hydraulic properties determined from core, slug, single-hole, cross-hole pumping test data and flow-through experiments. 129
Table 10.2: Mean $\left(\overline{\ln K}\right)$ and variance $\left(\sigma^2_{\ln K}\right)$ of the log-transformed K estimates for the real inversions with number of cross-hole tests used in the analysis. 137
Table 10.3: Mean $\left(\overline{\ln S_s}\right)$ and variance $\left(\sigma^2_{\ln S_s}\right)$ of the log-transformed S_s estimates for the real inversions with number of cross-hole tests used in the analysis. 137
Table 11.1: Summary of simulation times (min) using the THT and HT-m algorithms and its ratio to obtain the K and S_s tomograms (Figures 11.2 – 11.5). 149
Table 11.2: Summary of the mean and variance of the estimation errors from the comparison of observed to simulated drawdowns. The K and S_s fields are from the THT and HT-m algorithms. The correlation coefficients shown in this table are for the scatter plots shown in Figures 11.7a-b. 151
Table 11.3: Summary of the mean and variance of the estimation errors from the comparison of observed to simulated drawdowns. The K and S_s fields for the heterogeneous case

are from the inversion of the laboratory sandbox data using the HT-m algorithm, while the homogeneous case represents equivalent K and S_s obtained from *Liu et al*. [2007]. The correlation coefficients shown in this table are for the scatter plots shown in Figures 11.9a-b. 156
Table 12.1: Characteristics of porous media used to create a synthetic heterogeneous aquifer. 164
Table 12.2: The sand type and deposited volume for each layer in the sandbox aquifer. 166
Table 12.3: Summary of K estimates from various approaches. 167
Table 12.4: Geostatistical model parameters for kriging core and single-hole ln K data. 169
Table 12.5: Statistical properties of the estimated K tomograms. 174
Table 12.6: Slope and intercept of fitted linear model, coefficient of determination (R^2) and standard errors (L_1 and L_2 statistics) of various homogeneous and heterogeneous cases. 176
Table 12.7: Relative error computed for the simulated flow rates with respect to the measured values. 179
Table 13.1: Characteristics of each layer used to create a synthetic heterogeneous aquifer (after *Illman et al*., 2010). 187
Table 13.2: Geostatistical model parameters for kriging single-hole ln-K and ln-S_s data. 190
Table 13.3: Statistical properties of the estimated ln-K tomograms. 195
Table 13.4: Statistical properties of the estimated ln-S_s tomograms. 195
Table 13.5: Statistics of the linear model fit and correlation of determination (R^2). 199
Table 13.6: L1 norms of observed versus simulated drawdowns from cross-hole tests 8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42. 201
Table 13.7: L2 norms of observed versus simulated drawdowns from cross-hole tests 8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42. 201
Table 13.8: Statistics of the linear model fit, correlation of determination (R^2), and L_1 and L_2 norms for the ensemble analysis of all cases (including transient hydraulic tomography after the inclusion of 2, 4, 6, and 8 tests). 202
Table 14.1: TCE - equilibrium partitioning coefficient (K_{NW}). 220
Table 14.2: PCE - equilibrium partitioning coefficient (K_{NW}). 221
Table 14.3: Summary of PTT 1 – sandbox 3-1. 224
Table 14.4: Summary of PTT 2 – sandbox 3-1. 224
Table 14.5: Summary of PTT 3 – sandbox 3-1. 225
Table 14.6: Summary of PTT 4 – sandbox 3-1. 225
Table 14.7: Summary of PTT 5 – sandbox 3-1. 226
Table 14.8: Summary of PTT 6 – sandbox 3-1. 226
Table 14.9: Overall results of PPTs in SB 3-1. 227
Table 14.10: Summary of PTT 1 – sandbox 3-2. 228
Table 14.11: Summary of PTT 2 – sandbox 3-2. 229
Table 14.12: Summary of PTT 3 – sandbox 3-2. 229
Table 14.13: Summary of PTT 4 – sandbox 3-2. 230
Table 14.14: Overall results of PPTs in SB 3-2. 231
Table 14.15: Summary of PTT 1 – sandbox 1-2. 232
Table 14.16: Summary of PTT 2 – sandbox 1-2. 232
Table 14.17: Summary of PTT 3 – sandbox 1-2. 233
Table 14.18: Summary of PTT 4 – sandbox 1-2. 233

Table 14.19: Overall results of PPTs (2M1B) in SB 1-2. 234
Table 14.20: Summary of PTT 5 (2, 4DM3P_Br). 236
Table 14.21: Summary of PTT 5 (1-HEX_Br). 236
Table 14.22: Experimental values from core analysis - SB 3-2. 237
Table 14.23: Experimental values from core analysis from SB 1-2. 239
Table 14.24: K_{NW}-PTT vs. K_{NW} equilibrium. 242
Table 14.25: TCE mass loss computations. 246
Table 15.1: Characteristics of porous media used to create a synthetic heterogeneous aquifer. 250
Table 15.2: Injection, extraction, and water sampling locations during partitioning tracer test 1 - 4. 253
Table 15.3: Mean TCE saturation and volume estimated through the inversion of various partitioning tracer tests using the *K* tomogram from steady state hydraulic tomography and a known *K* distribution. 260
Table 16.1: Characteristics of porous media used to create a synthetic heterogeneous aquifer. 268
Table 16.2: Results of linear model fit to simulated versus observed drawdowns for coarse and fine grids. Simulated values were obtained during the inverse modeling of these tests (i.e., calibration). 275
Table 16.3: Results of linear model fit to simulated versus observed drawdowns for coarse and fine grids. Simulated values were obtained by modeling independent dipole cross-hole tests not used in the construction of the tomograms (i.e., validation). 275
Table 17.1: Injection, extraction, and water sampling locations during tests 1 - 5. 292

LIST OF FIGURES

Figure 1.1: Flow chart showing the tasks, subtasks, and the corresponding sections of this report. 3

Figure 3.1: Cross correlation between *h* at x=9.5 m and *f*'s at different locations for three selected times during a pumping test. 15

Figure 3.2: Cross correlation between *h* at x=9.5 m and *s*'s at different locations for three selected times during a pumping test. 16

Figure 3.3: Temporal correlation of transient heads at x =7.5 m during a pumping test. 17

Figure 3.4: Comparison of estimated hydraulic conductivity with true hydraulic conductivity in a deterministic case. 18

Figure 3.5: Comparison of estimated specific storage with true specific storage in a deterministic case. 18

Figure 3.6: Estimated hydraulic conductivity from transient hydraulic tomography: (a) estimates from the first pumping test; (b) estimates from the additional second pumping test; (c) estimates from the additional third pumping test; (d) estimates from the additional fourth pumping test. 19

Figure 3.7: Estimated specific storage from transient hydraulic tomography: (a) estimates from the first pumping test; (b) estimates from the additional second pumping test; (c) estimates from the additional third pumping test; and (d) estimates from the additional fourth pumping test. 20

Figure 3.8: Comparison between estimated hydraulic conductivity with the true field in a three dimensional aquifer. Estimated hydraulic conductivity field after (a) two, (b) four, (c) six, (d) and eight pumping tests, and (e) the synthetic true hydraulic conductivity field. 22

Figure 3.9: Comparison between estimated specific storage with the true field in a three dimensional aquifer. Estimated specific storage field after (a) two, (b) four, (c) six, (d) and eight pumping tests, and (e) the synthetic true specific storage field. 22

Figure 3.10: Frequency distributions of estimation errors: (a) hydraulic conductivity field and (b) specific storage field of a three-dimensional aquifer. 23

Figure 3.11: Comparison of variograms between (a) estimated hydraulic conductivity with true field; (b) estimated specific storage with true field in a three-dimensional aquifer. 24

Figure 4.1: True synthetic fields: a) transmissivity ; b) storage coefficient. 34

Figure 4.2: Estimated transmissivity fields using a sparse well field. a) Using transient heads (Case 1); b) using temporal moments (Case 2); c) the scatterplot of estimated versus true ln*T* fields of Case 1; d) the scatterplot of estimated versus true ln*T* field of Case 2. 35

Figure 4.3: Estimated storage coefficient fields using a sparse well field. a) Using transient heads (Case 1); b) using temporal moments (Case 2); c) the scatterplot of estimated versus true ln*S* fields of Case 1; d) the scatterplot of estimated versus true ln*S* field of Case 2. 36

Figure 4.4: Estimated transmissivity fields using a dense well field. a) Using transient heads (Case 3); b) using temporal moments (Case 4); c) the scatterplot of estimated versus true ln*T* fields of Case 3; d) the scatterplot of estimated versus true ln*T* field of Case 4. 37

Figure 4.5: Estimated storage coefficient fields using a dense well field. a) Using transient heads (Case 3); b) using temporal moments (Case 4); c) the scatterplot of estimated versus true ln*S* fields of Case 3; d) the scatterplot of estimated versus true ln*S* field of Case 4. 38
Figure 5.1: Typical hydrographs and corresponding denoised hydrographs (smooth solid line) in the sandbox experiments (Liu et al., 2007) a) SNR: 17.75; b) SNR: 4.41. Notice vertical exaggeration. 45
Figure 5.2: a) and e) are the true *K* (cm/s) and S_s (1/cm) fields of the synthetic aquifer. The dashed lines are the boundaries between units (1, 2, 3, and 4); estimated *K* and S_s fields using noise free data (b and f); using noisy data (c and g); using denoised data (d and h). 48
Figure 5.3: Effectiveness of wavelet denoising. A plot of 480 pairs of heads a) before b) after denoising and simulated heads (red circles) based on effective *K* and S_s of the synthetic aquifer; Symbols + denote the simulated heads plus and minus one standard deviation of head variation induced by heterogeneity. 51
Figure 5.4: The conditional L2 norms of the head as a function of iteration for noise-free, noisy, and denoised data sets for the synthetic aquifer case. 53
Figure 5.5: Variances of the estimated a) lnK and b) $\ln S_s$ fields vs. the number of iterations for different scenarios associated with the synthetic aquifer. 54
Figure 5.6: Comparison of head, and stream lines at t=1.5 s after pumping in the synthetic aquifer with a) true, and b) effective; c) estimated *K* and S_s fields based on c) noisy and d) denoised hydrographs. The statistical metrics are based on the true head field in a). 57
Figure 5.7: Schematic setup and discretization of the sandbox used in the experiment in Liu et al. (2007) and Illman et al. (2007). Open circles are pumping ports; both solid and open circles are used as observation ports. The solid square denotes the pumping port for the validation purpose. The open rectangles shown are the low permeability zones. 58
Figure 5.8: The variances of estimated ln*K* and lnS_s as a function of iteration for the sandbox experiment. 59
Figure 5.9: The conditional L2 norm of the head as a function of iteration for the sandbox experiment. 60
Figure 5.10: The estimated hydraulic conductivity (cm/s) and specific storage (1/cm) fields for the sandbox experiment. 61
Figure 5.11: Validation results: observed vs. estimated heads (cm) at the 47 observation ports at 4 different times in the sandbox. Two simulated heads were considered: one using estimated effective *K* and S_s fields from the equivalent homogeneous domain and the other using the estimated heterogeneous *K* and S_s fields derived from the analysis of HT. 62
Figure 6.1: True Synthetic Fields: a) Hydraulic Conductivity (*K*, m/d); b) Water Saturation (θ_w); and c) NAPL Saturation (θ_n). 73
Figure 6.2: Estimated NAPL Saturation from: a) *K* and θ_w are Known (case 1); b) *K* is Known but θ_w is Unknown (case 2); c) Use Estimated *K* from Heads and θ_w is Unknown (case 3); d) *K* and θ_w are Unknown (case 4); and e) Kriging (case 5). 74

Figure 6.3: Estimated Hydraulic Conductivity from a) Heads; b) Heads and Tracer Data (case 3); and c) Tracer Data (case 4). 76

Figure 6.4: Scatter Plots of True ln NAPL Versus Estimated NAPL from: a) K and θ_w are Known (case 1); b) K is Known but θ_w is Unknown (case 2); c) Use Estimated K from Heads and θ_w is Unknown (case 3); d) K and θ_w are Unknown (case 4); and e) Kriging (case 5). 77

Figure 6.5: Scatter Plots of True ln$K$ Versus Estimated ln$K$ from a) Heads; b) Heads and Tracer Data (case 3); and Tracer Data (case 4). 78

Figure 7.1: Comparison between true and estimated DNAPL content fields. (a) true field; (b) estimated from discrete concentration data; (c) estimated from tracer first moment data. 88

Figure 7.2: Scatter plot of true DNAPL content field versus estimate fields from: (a) discrete concentration data; (b) tracer first moment data. 89

Figure 7.3: Comparison between true and estimated hydraulic conductivity fields: (a) true field; (b) estimated from head data; (c) estimated from tracer first moment data. 90

Figure 7.4: Scatter plot of true hydraulic conductivity field versus estimated field from (a) head data; (b) tracer first moment data. 91

Figure 7.5: Cross correlations between the first moment perturbation and perturbations of: a) ln$K$; b) ln θ_w; and c) ln θ_n and d) cross correlation between perturbation of head and perturbation of ln$K$. Where the filled red circle represents tracer/water injection location and open red circle is the tracer observation location, blue arrows denote the velocity field (not scaled to magnitude), and the black curves with arrows are the streamlines. 93

Figure 8.1: Computer Aided Design (CAD) drawing of sandbox used for the validation of hydraulic tomography. Low K blocks are shown as dotted lines and circles around the numbers indicate the ports used for pumping. 95

Figure 8.2a-i: a) The synthetic true K distribution used to generate synthetic cross-hole test data; b) (47 - case 1) c) (47, 44 - case 2); d) (47, 44, 35 - case 3); e) (47, 44, 35, 32 - case 4); f) (47, 44, 35, 32, 17 - case 5); g) (47, 44, 35, 32, 17, 14 - case 6); h) (47, 44, 35, 32, 17, 14, 5 - case 7); and i) 47, 44, 35, 32, 17, 14, 5, 2 - case 8) show the resulting K tomograms by sequentially inverting the synthetic cross-hole test data. Numbers in parentheses from b) – i) indicate the port numbers used as the pumped well for each cross-hole test. 103

Figure 8.3a-d: Tomogram with: a) original data set without removing bias/errors (case 9); b) after removal of pumped well data affected by skin effect (case 10); c) hydraulic head measurements corrected for variations in the offset (case 11); d) accounting for varying boundary conditions for each pumping test (case 12). All results shown here are after including 8 pumping tests in the inversion as in case 8 described in figure 8.3i. 104

Figure 8.4: Scatter plot of simulated hydraulic head (h_s) versus measured hydraulic head (h_m). The h_s values were obtained by simulating the cross-hole test conducted at port 46 using the K tomogram (Figure 8.3d). The h_m values are from an actual cross-hole test conducted at port 46. 106

Figure 8.5: Scatter plot of K values from the real (case 12) and the reference (case 8) K tomograms. 108

Figure 8.6: Scatter plot of *K* values from the real (case 12) *K* tomogram at the observation point and core *K* estimates. 109
Figure 8.7: Scatter plot of *K* values from the real (case 12) *K* tomogram at the observation point and single-hole *K* estimates. 110
Figure 9.1: Synthetic *K* distribution with port locations for pumping, observation, and conditioning. 114
Figure 9.2: Reference *K* tomograms generated by sequentially inverting 8: a) synthetic and b) real cross-hole pumping tests (modified after Illman et al., 2007). 120
Figure 9.3: *K* tomograms obtained from low Q pumping test data with lower signal-to-noise ratio for the: a) synthetic and b) real cases. 121
Figure 9.4: *K* tomograms obtained by varying the order of pumping test data included in the SSLE algorithm for the: a) synthetic and b) real cases. 122
Figure 9.5: *K* tomograms obtained by conditioning with core *K* values for the: a) synthetic and b) real cases. 124
Figure 9.6: *K* tomograms obtained by conditioning with slug *K* values for the: a) synthetic and b) real cases. 125
Figure 9.7: *K* tomograms obtained by conditioning with single-hole *K* values for the: a) synthetic and b) real cases. 126
Figure 9.8: Scatter plot of simulated drawdown versus observed drawdown at steady state for cross-hole hydraulic tests at port 46. The head values were obtained by simulating cross-hole pumping tests on the computer using the synthetic *K* tomograms (Figures 9.2a, 9.3a, 9.4a, 9.5a, 9.6a, and 9.7a). The observed head values from the true *K* field (Figure 9.1) were also simulated on the computer. 127
Figure 9.9: Scatter plot of simulated drawdown versus observed drawdown at steady state for cross-hole hydraulic tests at port 46. In both cases, the simulated values were obtained through synthetic simulations of the cross-hole test conducted using the real *K* tomograms (Figures 9.2b, 9.3b, 9.4b, 9.5b, 9.6b, and 9.7b). The observed drawdown values are from an actual cross-hole pumping test conducted at port 46. 128
Figure 10.1a-f: a) Sketch of the sandbox with a deterministic heterogeneous aquifer; b) (14, 17 - case 1); c) (14, 17, 32 - case 2); d) (14, 17, 32, 35 - case 3); e) (14, 17, 32, 35, 44 - case 4) and f) (14, 17, 32, 35, 44, 47 - case 5) show the resulting *K* tomograms by sequentially inverting the experimental cross-hole test data. Numbers in parentheses from b) – f) indicate the port numbers used as the pumped well for each cross-hole test. 133
Figure 10.2a-i: a) Sketch of the sandbox with a deterministic heterogeneous aquifer; b) (14, 17 - case 1); c) (14, 17, 32 - case 2); d) (14, 17, 32, 35 - case 3); e) (14, 17, 32, 35, 44 - case 4) and f) (14, 17, 32, 35, 44, 47 - case 5) show the resulting S_s tomograms by sequentially inverting the real cross-hole test data. Numbers in parentheses from b) – f) indicate the port numbers used as the pumped well for each cross-hole test. 134
Figure 10.3: Contour plots of: a) core *K* values, b) slug *K* values, c) cross-hole *K* values with pumping taking place at port 28, and d) cross-hole S_s values with pumping taking place at port 28. 135
Figure 10.4: Scatter plot of *K* values from the real *K* tomogram at the observation point and core, single-hole, and cross-hole *K* estimates. 138

Figure 10.5: Scatter plot of S_s values from the real S_s tomogram at the observation point and single-hole and cross-hole S_s estimates. 139
Figure 10.6: Scatter plot of K values from the transient hydraulic tomography of this paper and those from the steady state hydraulic tomography in *Illman et al*. [2007]. 139
Figure 10.7: Scatter plot of simulated drawdown versus observed drawdown at 3, 10, and 20 seconds after pumping begins for cross-hole hydraulic tests at a) port 47 and b) port 46. In both cases, the simulated values were obtained by through synthetic simulations of the cross-hole test conducted using the K and S_s tomogram (Figures 10.1f and 10.2f). The observed drawdown values are from an actual cross-hole test conducted at each port. 141
Figure 10.8: Simulated drawdown from homogeneous field versus observed head values at the same locations as well as those simulated using the K and S_s tomograms at time equal to: a) 3, b) 10, and c) 20 seconds after pumping. The solid line represents a 45 degree line, and the dashed lines represent the upper and lower bounds of the drawdown. 141
Figure 11.1: a) K and b) S_s distributions in the MMOC3 model used to generate synthetic cross-hole pumping test data used for THT and HT-m analyses (port 28 is shown in yellow). 145
Figure 11.2: K tomograms computed using THT from synthetic pumping test data included sequentially. a) Ports 47, 44. b) Ports 47, 44, 35, 32. c) Ports 47, 44, 35, 32, 17, 14. and d) Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate those used as the pumped well for each cross-hole test. 146
Figure 11.3: S_s tomograms computed using THT from synthetic pumping test data included sequentially. a) Ports 47, 44. b) Ports 47, 44, 35, 32. c) Ports 47, 44, 35, 32, 17, 14. and d) Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate those used as the pumped well for each cross-hole test. 147
Figure 11.4: K tomograms computed using HT-m from synthetic pumping test data included sequentially. a) Ports 47, 44. b) Ports 47, 44, 35, 32. c) Ports 47, 44, 35, 32, 17, 14. and d) Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate those used as the pumped well for each cross-hole test. 148
Figure 11.5: S_s tomograms computed using HT-m from synthetic pumping test data included sequentially. a) Ports 47, 44. b) Ports 47, 44, 35, 32. c) Ports 47, 44, 35, 32, 17, 14. and d) Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate the ports used as the pumped well for each cross-hole test. 148
Figure 11.6: Scatter plot of simulated versus observed drawdown at 3, 10, and 20 seconds after pumping begins for synthetic cross-hole hydraulic tests at port 28 for a) THT and b) HT-m algorithms. In both cases, the simulated values were obtained through synthetic simulations of the cross-hole tests conducted using the K and S_s tomograms (Figures 10.2d and 10.3d for THT; Figures 11.4d and 11.5d for HT-m). The observed drawdown values are from the synthetic cross-hole test conducted using MMOC3 with pumping taking place at port 28 with K and S_s distributions from Figure 11.1a-b. 150
Figure 11.7: K tomograms computed using HT-m from real cross-hole pumping test data included sequentially. a) Ports 44, 47. b) Ports 44, 47, 35, 32. c) Ports 44, 47, 35, 32, 17, 14. d) Ports 44, 47, 35, 32, 17, 14, 5, 2. Port numbers (see Figure 6.1) indicate those used as the pumped well for each cross-hole test. 154

Figure 11.8: S_s tomograms computed using HT-m from real cross-hole pumping test data included sequentially. a) Ports 44, 47. b) Ports 44, 47, 35, 32. c) Ports 44, 47, 35, 32, 17, 14. d) Ports 44, 47, 35, 32, 17, 14, 5, 2. Port numbers (see Figure 6.1) indicate those used as the pumped well for each cross-hole test. 154

Figure 11.9: Scatter plot of simulated drawdown versus observed drawdown at 3, 50, and 150 seconds after pumping begins for cross-hole hydraulic tests at port 46. In this figure, the simulated values were obtained through synthetic simulations of the cross-hole test conducted using: a) the K and S_s tomograms (Figures 11.8d and 11.9d) and b) equivalent K and S_s values from *Liu et al*. [2007]. The observed drawdown values are from an actual cross-hole test conducted at port 46. 157

Figure 11.10: Scatter plot of K values from the HT-m of this paper and those from the SSHT in *Illman et al*. [2007]. 158

Figure 11.11: Scatter plot of a) K and b) S_s values from the HT-m approach of this paper and those from the THT approach described in *Liu et al*. [2007]. 159

Figure 12.1: Schematic diagram of synthetic heterogeneous aquifer used for validation of steady state groundwater flow models. Numbers next to solid squares indicate port numbers, open squares around numbers indicate the 8 ports (2, 5, 14, 17, 32, 35, 44, 47) used for hydraulic tomography, open ovals around numbers indicate the 2 ports (17 and 21) pumped and data analyzed using VSAFT2 by treating the medium to be homogeneous (i.e., traditional pumping test analysis), and the dashed open squares around the 16 other ports (8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42) indicate the pumping locations for the independent cross-hole pumping tests used for validation purposes. 163

Figure 12.2: Schematic diagram of the sediment transport system for the creation of the synthetic heterogeneous aquifer. 165

Figure 12.3: Photograph of synthetic heterogeneous aquifer created via cyclic flux of sediment-laden water. 166

Figure 12.4: K distributions obtained via kriging of: a) core K estimates and b) single-hole K estimates. 170

Figure 12.5: K tomograms computed using: a) 2 pumping tests (ports 47 and 44); b) 4 pumping tests (ports 47, 44, 35, and 32); c) 6 pumping tests (ports 47, 44, 35, 32, 17, and 14); and d) 8 pumping tests (ports 47, 44, 35, 32, 17, 14, 5 and 2). 173

Figure 12.6: Overlay of K tomogram computed using 8 pumping tests over the picture of the heterogeneous sandbox aquifer using Adobe Photoshop. 174

Figure 12.7: Arithmetic plots of drawdown predictions of 16 separate cross-hole pumping tests using the a): core K_G estimates; b) single-hole K_G estimates; c) geometric mean of cross-hole test K_{eq} estimates (port 17); d) geometric mean of cross-hole test K_{eq} estimates (port 21) and e) K_{eff} from flow-through tests. 177

Figure 12.8: Arithmetic plots of drawdown predictions of 16 separate cross-hole pumping tests using: a): kriged core K estimates; b) kriged single-hole K estimates; c) K tomogram from SSHT (2 tests - ports 47 and 44); d) K tomogram from SSHT (4 tests - ports 47, 44, 35 and 32); e) K tomogram from SSHT (6 tests - ports 47, 44, 35, 32, 17 and 14); and f) K tomogram from SSHT (8 tests - ports 47, 44, 35, 32, 17, 14, 5 and 2). 178

Figure 12.9: Contour plot of ln K variance with 8 cross-hole tests included into SSLE algorithm. 180

Figure 13.1: Schematic diagram of synthetic heterogeneous aquifer used for validation of transient groundwater flow models. Numbers next to solid squares indicate port numbers, open squares around numbers indicate the 8 ports (2, 5, 14, 17, 32, 35, 44, 47) used for hydraulic tomography, the open oval indicates the port (21) at which pumping test data was analyzed using VSAFT2 by treating the medium to be homogeneous (i.e., traditional pumping test analysis), and the dashed open squares around the 16 other ports (8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42) indicate the pumping locations for the independent cross-hole pumping tests used for validation purposes [after *Illman et al*., 2010]. 186

Figure 13.2: Photograph of synthetic heterogeneous aquifer created via cyclic flux of sediment-laden water [after *Illman et al*., 2010]. 188

Figure 13.3: Distribution of a) ln-K and b) ln-S_s values for the calibrated geological model. 191

Figure 13.4: ln-K tomograms computed using: a) 2 pumping tests (ports 47 and 44); b) 4 pumping tests (ports 47, 44, 35, and 32); c) 6 pumping tests (ports 47, 44, 35, 32, 17, and 14); d) 8 pumping tests (ports 47, 44, 35, 32, 17, 14, 5 and 2) and e) 8 pumping tests (ports 47, 44, 35, 32, 17, 14, 5 and 2) but interpreted with the steady state hydraulic tomography algorithm (after *Illman et al*., 2010). 193

Figure 13.5: ln-S_s tomograms computed using: a) 2 pumping tests (ports 47 and 44); b) 4 pumping tests (ports 47, 44, 35, and 32); c) 6 pumping tests (ports 47, 44, 35, 32, 17, and 14); and d) 8 pumping tests (ports 47, 44, 35, 32, 17, 14, 5 and 2). 194

Figure 13.6: Scatterplot of ln-K from transient hydraulic tomography (this paper) to ln-K obtained using steady state hydraulic tomography (*Illman et al*., 2010). The dashed line is a 1:1 line indicating a perfect match. The solid line is a best fit line, and the parameters describing this line are only each plot. 196

Figure 13.7: Scatterplots of observed vs simulated drawdowns from independent cross-hole tests 18, 23, 40, and 42 at t = 0.5, 2, 5, and 10 secs. Simulated drawdown values were obtained through numerical simulations with a calibrated geological model of K and S_s. The dashed line is a 1:1 line indicating a perfect match. The solid line is a best fit line, and the parameters describing this line are on each plot. 198

Figure 13.8: Scatterplots of observed vs simulated drawdowns from independent cross-hole tests 8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42 at t = 0.5, 2, 5, and 10 secs. Simulated drawdown values were obtained through numerical simulations with K and S_s tomograms from transient hydraulic tomography. The dashed line is a 1:1 line indicating a perfect match. The solid line is a best fit line, and the parameters describing this line are on each plot. 199

Figure 13.9: Observed (small dots) and simulated (curves) records of drawdown versus time (seconds) during cross-hole pumping test with pumping at port 40. Simulated values are obtained from the forward simulation of the pumping test using the calibrated geological model. 203

Figure 13.10: Observed (small dots) and simulated (curves) records of drawdown versus time (seconds) during cross-hole pumping test with pumping at port 40. Simulated values are obtained from the forward simulation of the pumping test using the K and S_s tomograms from transient hydraulic tomography. 204

Figure 14.1: Microscopic view of the NAPL/Tracer interaction and breakthrough curves. 206

Figure 14.2: TCE source zone created in sandbox 3-1. 209
Figure 14.3: TCE source zone created in sandbox 3-2. 210
Figure 14.4: TCE source zones created in sandbox 1-2. Clockwise from top left to bottom right: photo taken after 15 min; photo taken after 30 min; photo taken after 2 hours; photo taken after 24 hours. 210
Figure 14.5: PTT configurations in SB3-1. From top left, clockwise, PPT1, PPT2, PTT3, PTT4, PTT5 and PTT6, in that order. 213
Figure 14.6: PTT configurations in SB 3-2. From PTT1 (top left), PTT2 (top right), PTT3 (bottom left), and PTT4 (bottom right). 214
Figure 14.7: Sandbox 1-2 PTT designs: PTT1 (top left), PTT2 (top right), PTT3 (bottom left) and PTT5 (bottom right). 215
Figure 14.8: Picture of the telescopic tubing arrangement for SB 1-2 (similar process for SB 3-2 was applied). 216
Figure 14.9: Fully inserted coring wells (SB 1-2). 217
Figure 14.10: Partitioning of core samples in SB 3-2. 217
Figure 14.11: Coring scheme applied on SB 1-2. 218
Figure 14.12: Picture of a completed core sample. 218
Figure 14.13: Equilibrium partitioning of 2,4-dimethyl-3-pentanol with TCE. 220
Figure 14.14: Breakthrough curves of Bromide and 2M1B from port 12 in PTT 2 conducted in SB3-1. 221
Figure 14.15: Breakthrough curves of Bromide and 2M1B from port 12 in PTT 2 conducted in SB3-2. 222
Figure 14.16: Breakthrough curves of Bromide, 2M1B, 2,4DM3P, and 1-Hex from port 47 in PTT 2 conducted in SB1-2. 222
Figure 14.17: Breakthrough curves from the multi-tracer test (PTT5, Port 5). 234
Figure 14.18: Kriged core TCE saturation estimates using GMS. 238
Figure 14.19: Linear interpolation of core TCE saturation estimates using GMS. 238
Figure 14.20: Inverse distance weighted interpolation of core TCE saturation values. 240
Figure 14.21: Linear interpolation of core TCE saturation values. 240
Figure 14.22: Picture of the TCE distribution in SB 1-2. 240
Figure 14.23: Effective partition coefficient (Kpe) of tracers with DNAPL (PCE). 243
Figure 14.24: Penetration of dyed TCE into the silicone seals of SB 1-2. 245
Figure 15.1: Schematic diagram of the sandbox showing sand types, water sampling, and pressure monitoring ports. Two rectangular boxes indicate locations of sand lenses consisting of #14 and F-75 sands. 250
Figure 15.2: Photograph of the TCE source zone after TCE migration has stopped. 252
Figure 15.3: *K* tomogram generated using drawdown data from 9 dipole tests. 258
Figure 15.4: Scatter plot of simulated (cm) versus observed drawdown (cm) under steady state flow conditions. 259
Figure 15.5: Estimated TCE saturation tomograms with inversion of: a) test 1 only; b) test 2 only; c) test 3 only; d) test 4 only; e) tests 1 - 4; and f) tests 1 - 4 with loop iteration. In this case, the *K* tomogram estimated using steady state hydraulic tomography (Figure 15.3) was used for partitioning tracer tomography. 261
Figure 15.6: Estimated TCE saturation tomograms with inversion of: a) test 1 only; b) test 2 only; c) test 3 only; d) test 4 only; e) tests 1 - 4; and f) tests 1 - 4 with loop iteration. In this case, the *K* values of the 3 sand types were obtained from permeameter

measurements and deterministically assigned for partitioning tracer tomography. 262

Figure: 15.7: Kriged TCE distribution based on the interpolation of 30 core samples. Solid circles show the approximate centroid location of the core samples. 263

Figure 16.1: Photograph of the sandbox showing the synthetic heterogeneous aquifer in which the dipole cross-hole and conservative tracer tests were conducted. Large black numbers indicate layer numbers, solid circles indicate port locations, and small blue numbers indicate port numbers. Layer 1 = 20/30; layer 2 = 4030; layer 3 = F85; layer 4 = 20/40; layer 5 = mix; layer 6 = mix; layer 7 = #12; layer 8 = F32; layer 9 = 20/40; layer 10 = F65; layer 11 = #12; layer 12 = 16/30; layer 13 = 20/30; layer 14 = F75; layer 15 = 20/40; layer 16 = mix; layer 17 = F85; layer 18 = 20/30. Note: The layers labelled "mix" consisted of equal volumes of #14, F75, and 16/30 sands. 267

Figure 16.2: Schematic diagram showing the injection, extraction, and sampling ports during the dipole conservative tracer test. 268

Figure 16.3: *K* distribution obtained through kriging of core scale K data from permeameter analysis. 270

Figure 16.4: *K* tomograms computed using the steady state hydraulic tomography algorithm of Yeh and Liu (2000). 274

Figure 16.5: Photographs of sandbox during tracer test 1 at various times: a) t = 12 min; b) t = 27 min; c) t = 62 min; d) t = 100 min; e) t = 175 min; and f) t = 286 min. 276

Figure 16.6: Normalized concentration breakthroughs at the 12 sampling ports (2, 10, 12, 14, 19, 22, 27, 30, 35, 37, 39, 47) and the extraction port (7). The arrival times (t_a), peak concentration, (C_p) and the time when the normalized concentration reaches background levels (t_b) is also noted. 278

Figure 16.7: Simulated versus observed drawdowns at 48 ports during the dipole tracer test. Simulated values were computed using: a) K_{eff} from Gelhar and Axness [1983] solution with statistics of permeameter analysis of core samples; b) kriged *K* field; and e) *K* tomogram computed using fine mesh. 280

Figure 16.8: Concentration distributions from tracer transport simulation with: a) K_{eff} and A_{11} computed using Gelhar and Axness [1983 solution with statistics of permeameter analysis of core samples (case 1); b) kriged *K* field (case 2); and c) *K* tomogram (case 3). 281

Figure 16.9: Breakthrough curves from the dipole tracer test and corresponding match of forward simulations. Tracer transport simulations were performed using various *K* fields. 283

Figure 16.10: Total mass computed by the M_0 for each sampling port: a) case 1; b) case 2; c) case 3; d) case 4; and e) case 5. 284

Figure 16.11: Mean arrival time of the center of mass (μ) computed for each sampling port: a) case 1, b) case 2, c) case 3, d) case 4, e) case 5, and e) case 6. 285

Figure 16.12: The variance of the breakthrough curve (σ) at each sampling port: a) case 1; b) case 2; c) case 3; d) case 4; e) case 5; and f) case 6. 286

Figure 17.1: Photograph of the sandbox showing the synthetic heterogeneous aquifer in which the dipole cross-hole and conservative tracer tests were conducted. Large black numbers indicate layer numbers, solid circles indicate port locations, and small blue numbers indicate port numbers. Layer 1 = 20/30; layer 2 = 4030; layer 3 = F85;

layer 4 = 20/40; layer 5 = mix; layer 6 = mix; layer 7 = #12; layer 8 = F32; layer 9 = 20/40; layer 10 = F65; layer 11 = #12; layer 12 = 16/30; layer 13 = 20/30; layer 14 = F75; layer 15 = 20/40; layer 16 = mix; layer 17 = F85; layer 18 = 20/30. Note: The layers labelled "mix" consisted of equal volumes of #14, F75, and 16/30 sands. 289

Figure 17.2: *K* tomogram computed using the steady state hydraulic tomography algorithm of Yeh and Liu [2000]. 290

Figure 17.3: Photograph of the sandbox showing the distribution of PCE prior to conducting the partitioning tracer tests. 291

Figure 17.4: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 1. 292

Figure 17.5: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 2. 293

Figure 17.6: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 3. 293

Figure 17.7: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 4. 294

Figure 17.8: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 5. 294

Figure 17.9: Photograph of PCE distribution in the sandbox at various stages of partitioning tracer tests. 295

Figure 17.10: Preliminary PCE saturations from kriging excavated samples after conduct of 5 partitioning tracer tests. 296

Figure 18.1: Log-log drawdown time curves at five different elevations (z =1.5, 3.0, 6.0, 7.1, 7.5 m) at a) r = 5 m and b) r = 30 m. Solid lines denote the results based on the solution by Hantush at three observation elevations in the saturated zone z = 1.5, z = 3.0, z = 6.0m. In Figure a, the difference between the Hantush solution at z=1.5 m and z=3.0 m is not distinguishable. In Figure b, the solutions are the same at all three elevations. 302

Figure 18.2: Ω as a function of elevation a) at five early times (t = 0.02, 0.07, 0.3, 1.0, and 3.0 min) corresponding to five different radii (r = 2.5, 5, 10, 20 and 30 m); b) at an intermediate stage (10 minutes after pumping) and c) a late stage (1000 minutes after pumping) at the five radii. The water table is located at the sharp discontinuity. 304

Figure 18.3: Comparison of drawdown-time curves at three different elevations z=10cm, z=40cm and z=70cm in a fully saturated column and those in a soil column with an unsaturated zone of 40 cm with the exponential constitutive relationships, and those with VGM constitutive relationships. 306

Figure 18.4: The rate change in storage as a function of the elevation in a 1-D soil column for a) exponential model and b) VGM models at time 0.01 min, 3.0 min and 100 min after drainage started. WT denotes the water table. The initial water table before draining is at 110cm. b) and d) are the cumulative volume of water released from 1) compaction of the aquifer, 2) drainage from the initial unsaturated zone, and 3) drainage from pores during falling of the water table over the column at a function of time for exponential and VGM constitutive models, respectively. 308

Figure 18.5: a) Drawdown-time curve at z=20cm with different α values when β=0.15/cm and b) Different β values while fixing α=0.15/cm. 310
Figure 18.6: The drawdown-time curve at two elevations z=20cm and 50cm showing the role of an initial unsaturated zone. 311
Figure 18.7: The standard deviation (σ_p) of the head for every parameter as a function of time with variances 1.0 and correlation scale 3m. a) to f) represent the results for observation point at two different radius r=5m and 30m and three different elevations z=3m, z=6m and z=7.1m. The red solid lines show the combined effect of all the six parameters. 314
Figure 18.8: Deviations of drawdown-time curve in an aquifer with heterogeneous K_s, S_s and θ_s parameters from the drawdown-time curve derived with the assumption of aquifer homogeneity; the upper and lower bounds based on the first-order stochastic moment analysis that quantify the deviation. 315
Figure 18.9: Spatial distributions of the cross-correlation between head at observation point(Obs. x=95.0 m, y=100.0 m and z=3.0m) and K_s along a vertical plane running through both the pumping well (Pw) and the observation point at a) the early time, b) the intermediate time, and c) the late time of an S-shaped drawdown-time curve. 316
Figure 18.10: Spatial distributions of the cross-correlation between head at observation point (Obs. x=95.0 m, y=100.0 m and z=3.0m) and S_s along a vertical plane running through both the pumping well (Pw) and the observation point at the early time of an S-shaped drawdown-time curve. The cross-correlation diminishes as the time progresses. 317
Figure 18.11: Spatial distributions of the cross-correlation between head at observation point (Obs. x=95.0 m, y=100.0 m and z=3.0m) and θ_s along a vertical plane running through both the pumping well (Pw) and the observation point at the late time of an S-shaped drawdown-time curve. 317
Figure 19.1: (a) heterogeneous K_s field and (b) α field for 1D synthetic aquifer. 324
Figure 19.2: Simulated pressure heads of two infiltrations for 1D aquifer. 325
Figure 19.3: Estimated: a) K_s and b) α fields for 1D aquifer. 326
Figure 19.4: Axial radially symmetric aquifer. 327
Figure 19.5: Sensitivities of transient pressure to a) K_s and; b) S_s, and c) α. 328
Figure 19.6: Domain and grid of a 3D unconfined aquifer (green columns representing wells). 329
Figure 19.7: Comparison between a) "true" and b) estimated K_s fields. 330
Figure 19.8: Comparison between a) "true" and b) estimated S_s fields. 331
Figure 19.9: Comparison between a) "true" and b) estimated α fields. 332
Figure 20.1: Photograph of the sandbox showing all sensor locations (● = pressure transducers; ● without pressure transducers; ○ = tensiometers; and **x** = water content sensors) and the various layers that were packed. Sensor locations are approximate. 337
Figure 20.2: Schematic diagram showing the ports that were pumped for the cross-hole pumping tests. Solid black circles represent locations of pressure transducers and solid grey circles represent ports that were not instrumented. Solid squares are pumping locations used for the estimation of heterogeneous K and S_s fields. Dashed squares are pumping locations for the validation of the results from transient hydraulic tomography. 338

Figure 20.3: Simulated versus observed drawdown for the cross-hole test at port 22 using effective K and S_s values. 339

Figure 20.4: a) K and b) S_s tomograms computed using the transient hydraulic tomography algorithm of *Zhu and Yeh* [2005] SSLE with 8 cross-hole pumping tests. Solid squares are pumping locations used for transient hydraulic tomography dashed squares are pumping locations for the validation of the results from transient hydraulic tomography. 341

Figure 20.5: a) Simulated versus observed drawdowns from 8 cross-hole pumping tests used for calibration purposes and b) additional tests used for the validation of the K and S_s tomograms. 342

Figure 20.6: Moisture characteristic curves determined through the hanging column method of: a) F-35; b) F-45; c) F110; d) Sil-co-Sil 53; e) Sil-co-Sil 106; and f) F35 and F45 matched simultaneously. The squares represent F35 and gradient symbols represent F35. 343

Figure 20.7: Schematic diagram of the sandbox showing an array of sensors (● = pressure transducers; ● = port without pressure transducers; ○ = tensiometers; and **x** = water content sensors) utilized to monitor the pumping test in a heterogeneous unconfined aquifer. The box indicates the port at which the unconfined pumping test was performed. 344

Figure 20.8: Observed drawdown during the unconfined pumping test at a) 9 pressure transducer ports; and, b) 6 tensiometers. 347

Figure 20.9: Spatial distribution of: a) pressure head and b) volumetric water content in the upper half of the sandbox with time during the unconfined aquifer pumping test at port 3. Symbols indicate the position of various sensors used to monitor the pumping test (● = pressure transducers; ○ = tensiometers; and **x** = water content sensors)). 348

Figure 20.10: Simulated (solid line) and observed (dashed line) drawdown from selected: a) pressure transducers; and, b) tensiometers during the unconfined aquifer pumping test (Case 1). 351

Figure 20.11: Simulated (solid line) and observed (dashed line) drawdown from selected: a) pressure transducers; and, b) tensiometers during the unconfined aquifer pumping test (Case 2). 353

Figure 20.12: Simulated (solid line) and observed (dashed line) drawdown from selected: a) pressure transducers; and, b) tensiometers during the unconfined aquifer pumping test (Case 3). 355

LIST OF ACRONYMS

CTT	Conservative tracer tomography
DNAPL	Dense nonaqueous phase liquids
GC	Gas Chromatograph
GUI	Graphical user interface
HT	Hydraulic tomography
HT-m	Hydraulic tomography that uses the temporal moment of drawdown-recovery data
IC	Ion chromatograph
K	Hydraulic conductivity
PCE	Tetrachloroethene
PTT	Partitioning tracer tomography/Partitioning tracer test
SERDP	Strategic Environmental Research and Development Program
SLE	Successive linear estimator
SSLE	Sequential successive linear estimator
SSHT	Steady state hydraulic tomography
TCE	Trichloroethene
THT	Transient hydraulic tomography
UA	University of Arizona
UI	University of Iowa
UW	University of Waterloo
VSAFT2	Variably Saturated Flow and Transport model in 2 dimensions

ACKNOWLEDGEMENTS

Principal support for this project came from the Strategic Environmental Research & Development Program (SERDP). We thank Dr. Andrea Leeson for her support of this project which we think has led to the development of several new technologies that will expedite site characterization and cleanup at DoD sites. We also thank the support personnel from the SERDP office which made the project possible.

Additional support for the project was provided by the U.S. National Science Foundation (NSF), Natural Sciences and Engineering Resources Council of Canada (NSERC), Ontario Research Foundation (ORF), and Canada Foundation for Innovation (CFI).

This project supported the following undergraduate, graduate students, and research staff from the University of Iowa (UI), University of Waterloo (UW) and the University of Arizona (UA). All of them have contributed to this report in one form or another. Their names are given in alphabetical order: Matthew Alexander (UW), Steven Berg (UW), Tim Corley (UA), Andrew Craig (UI), Pelin Idil (UI), Taehong Kim (UI), Kris Kuhlman (UA), Yanhwa Li (UA), Xiaoyi Liu (UI), Katia Lorenzetti (UW), Deqiang Mao (UA), Antonio Massi (UI), Phoolendra Kumar Mishra (UA), William Neumann (UI), Andrew Price (UW), Kara Prior (UI), Raghavendra M Suribhatla (UA), Martha Whitaker (UA), Brett Woods (UI), Jianwei Xiang (UA), Tintin Yan (UA), Danting Yin (UI), Junfeng Zhu (UA).

ABSTRACT

Dense Nonaqueous Phase Liquids [DNAPLs] are prevalent at a large number of sites throughout the world. The variable release history and geologic heterogeneity make the spatial distribution of DNAPLs in the source zone complex. This causes difficulties in cleanup and can contribute to long-term groundwater contamination for decades to centuries. Therefore, the spatial distribution, mass, and composition of DNAPLs present in the source zone need to be characterized in sufficient detail so that efficient remediation schemes can be designed. Effective characterization of DNAPL source zones is also critical for long-term monitoring and management decisions that are critical to DoD/DOE/EPA's environmental mission of efficient site cleanup and closure.

Objectives: The objectives of this project are: 1] to develop algorithms that fuses different types of information using a stochastic approach to provide a cost-effective characterization, monitoring, and predictive technology for the DNAPL source zone, 2] to conduct laboratory experiments to test and verify this proposed technology, and 3] to distribute the results of the research through a web-based virtual tomography laboratory to assist scientists, engineers, and managers to solve DNAPL contamination problems.

Technical Approach: We present here the modeling and experimental results of a cost-effective technology that images DNAPL source zones in 3-D without extensive invasive sampling. This new technology based on stochastic methods, assimilates results of hydraulic and partitioning tracer tomography surveys to derive the best estimate of the DNAPL distribution and its uncertainty. Specifically, it first analyzes the information derived from hydraulic tomography to identify the 3-D heterogeneity in hydraulic conductivity (K) and specific storage (S_s) of the aquifer. The knowledge of heterogeneity is then used to design conservative tracer and partitioning tracer tomography tests to accurately depict the spatial distribution of DNAPL residual saturation (S_N) in the source zone.

Results: Building on the concept of scan technologies developed in medical sciences and geophysics, this effort has developed a subsurface characterization tool comprised of three "fused" field methodologies: Hydraulic Tomography (HT), Conservative Tracer Tomography (CT), and Partitioning Tracer Tomography (PTT). The results of these methodologies are fused via an innovative algorithm that provides an improved, more accurate depiction of subsurface heterogeneities – subsequently allowing greater certainty in identifying and treating contaminant source zones. Preliminary calculations suggest that the fused tomography technology becomes markedly more cost effective over conventional characterization approaches at sites with suspected investigation areas larger than 2,500 sq. ft. The degree of cost savings increases dramatically in conjunction with the increasing size of the area being characterized. While the resolution of the heterogeneity patterns are dependent on the density of the monitoring well network, the developed algorithm still yields improved estimates of K, S_s, and S_N in comparison to traditional interpretive techniques. A draft deployment protocol is provided.

Benefits: The project will provide a new generation technology that images the three-dimensional spatial distribution of hydraulic heterogeneity in the subsurface, which allows for the improved characterization of DNAPL source zones compared to existing technologies. Effective characterization is critical for its efficient remediation and long-term monitoring. The proposed technology is superior to traditional characterization techniques as it requires less invasive sampling and a far fewer number of boreholes leading to substantial cost savings to obtain the same level of accuracy. The cost savings is found to increase in larger and deeper DNAPL source zones where drilling becomes very expensive. It is also non-destructive and can

be applied repeatedly [e.g., pre- and post-remediation]. More importantly, it also provides uncertainty estimates that can facilitate better decision making.

1. OBJECTIVES:

The overall project objectives are: 1] to develop a software/hardware package that fuses different types of information using a stochastic approach to provide a cost-effective characterization, monitoring, and predictive tool for the DNAPL source zone, 2] to conduct laboratory experiments to test and verify this proposed technology, and 3] to distribute the results of the research to assist scientists, engineers, and managers to solve DNAPL contamination problems. Figure 1.1 is a flow chart that illustrates how the various chapters of this report (indicated in numbers in boxes) relate to the various tasks and subtasks.

1.1 TASK 1: ALGORITHM DEVELOPMENT

The main objective of this task is to develop algorithms for analyzing hydraulic tomography [HT] tests, conservative tracer tomography [CTT] tests, partitioning tracer tomography [PTT] tests, as well as methods to integrate these different tomography experiments to characterize DNAPL source zones.

1.2 TASK 2: LABORATORY EXPERIMENTAL AND NUMERICAL MODELING STUDIES

The main objective of this task is to validate the algorithms developed under task 1 through experimental studies conducted in synthetic aquifers built within laboratory sandboxes.

1.3 TASK 3: DEVELOPMENT OF HYDRAULIC AND TRACER TOMOGRAPHY ALGORITHMS THAT CONSIDERS THE EFFECTS OF THE UNSATURATED ZONE

The current algorithms developed under Task 1 of ER-1365 do not consider the effects of the unsaturated zone. The effects of the unsaturated zone could be important in DNAPL source zones in shallow unconfined aquifers because of the delayed drainage effect. In this task, we will consider the presence of the unsaturated zone in the estimation of the hydraulic parameters and DNAPL saturations in the saturated zone, by explicitly building in the unsaturated zone.

1.4 TASK 4: VALIDATION OF ALGORITHMS DEVELOPED IN TASK 3

In Task 4, we will validate the algorithms developed in Task 3 by building a new sandbox [i.e., UZ/SZ sandbox] with monitoring in both the unsaturated [UZ] and saturated zones [SZ] at the University of Waterloo [UW]. To date, hydraulic tomography experiments have been completed and a preliminary interpretation of those results is presented in this final report.

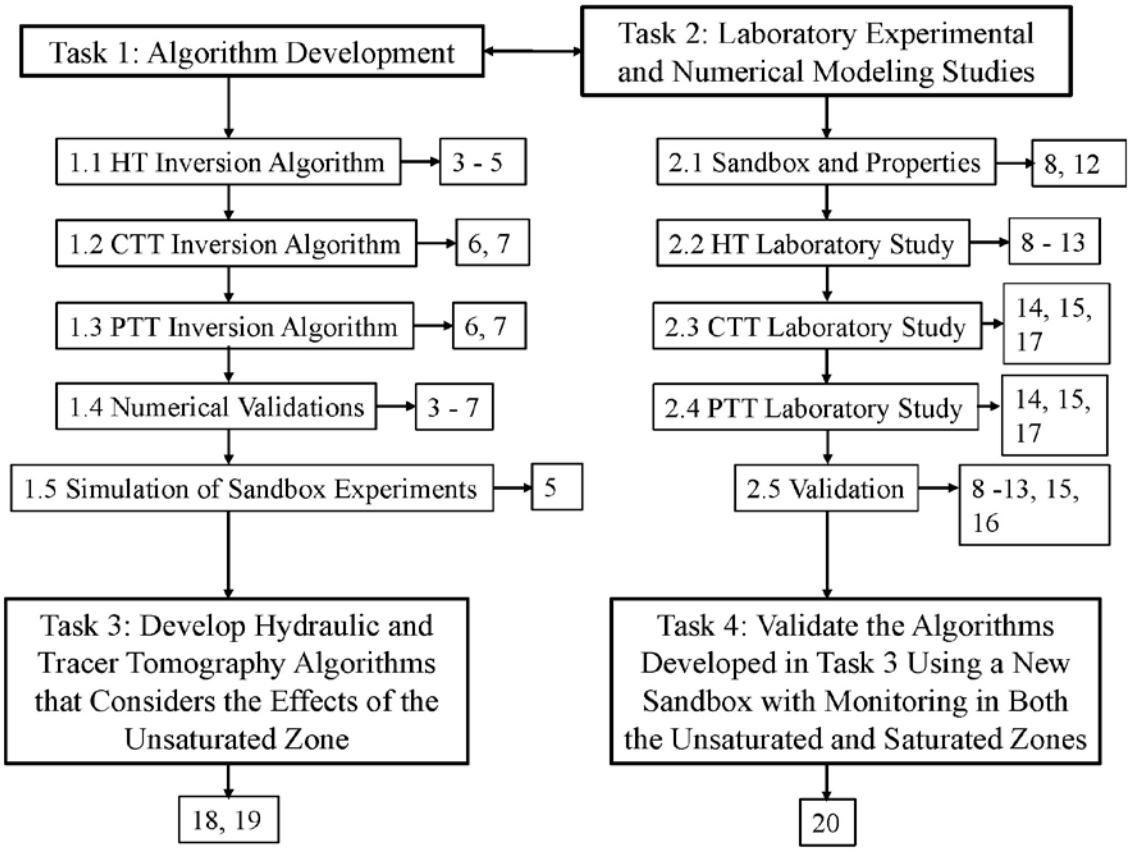


Figure 1.1: Flow chart showing the tasks, subtasks, and the corresponding sections of this report.

2. BACKGROUND:

Dense Nonaqueous Phase Liquids [DNAPLs] are prevalent at a large number of sites throughout the world. The high densities, low interfacial tensions, and low viscosities of halogenated solvents can lead to deep DNAPL penetration [Pankow and Cherry, 1996]. In porous media, much of the DNAPL mass remains in the groundwater as persistent source zones. The variable release history and geologic heterogeneity make the distribution of DNAPL in the source zone complex, where DNAPLs exist as residuals, pools of pure phase, or ganglia. Without remediation, these source zones can contribute to long-term groundwater contamination for decades to centuries. Therefore, the spatial distribution, mass, and composition of DNAPLs present in the source zone need to be characterized in great detail so that efficient remediation schemes can be designed.

The technology for characterizing such contamination is still in the early stages of development. Our knowledge of DNAPL dissolution and remediation issues nevertheless has matured during the last few years [Khachikian and Harmon, 2000]. Many innovative tracer techniques have been introduced for the enhanced characterization of DNAPL source zones [Rao et al., 2000]. While these tracer techniques allow for an in situ estimation of volume-averaged values of DNAPL saturation, there is an urgent need for the development of a cost-effective technology to support field-scale characterization of DNAPL location and strength as a dissolving source [Khachikian and Harmon, 2000].

2.1. DIRECT CHARACTERIZATION TECHNIQUES

To determine the extent of DNAPLs contamination within groundwater, the most common and standard methods are performed with the use of groundwater monitoring wells [Cohen and Mercer, 1993; Pankow and Cherry, 1996] and soil sampling [Taylor and Serafini, 1988]. The use of fully screened monitoring wells however may only imply the likely presence of DNAPLs because ground water contaminant concentrations near regions of DNAPLs are usually less than their aqueous solubility. Furthermore, the usage of fully screened monitoring wells can result in errors on orders of magnitude in assessments of groundwater contamination because the measurement is integrated over the length of the screened interval. Multi-level wells may eliminate some of the problems.

Once the presence of high DNAPL concentrations and sediment sorbed contaminants have been implicated by monitoring wells, soil contaminant surveys are often performed to determine the total level and distribution of contamination. In general, organic contaminants are extracted from the soil with organic solvents such as methylene chloride. The concentration of the VOC and SVOC present in the soil is then determined with a gas chromatograph.

With information from the direct sampling, a three-dimensional distribution of contaminants can then be determined through geostatistical interpolation techniques, such as kriging. However, applications of the method to real-world problems are limited because the requirements for relatively large number of samples [Barcelona and Jaglowski, 1999]. Collection of such a large number of invasive samples is often cost-prohibitive and practically impossible.

2.2. INDIRECT CHARACTERIZATION TECHNIQUES

A variety of techniques including visual inspections of drill cores, soil vapor analysis, geophysical surveys, use of radon abundance data, and partitioning tracer tests have been developed to characterize the DNAPL source zone. A comprehensive review of DNAPL

characterization methods and approaches is provided in Kram et al. [2001] and its cost comparisons in Kram et al. [2002].

Partitioning tracer tests have been increasingly used as a tool to provide a spatially integrated measure of residual DNAPL volume in the flow without causing disturbances to the source zone domain [Jin et al., 1995; Nelson and Brusseau, 1996; Annable et al., 1998]. These tests are performed through the comparison of the transport behavior of conservative and DNAPL partitioning tracers. Solutions of conservative or non-partitioning tracers such as methanol [Annable et al. 1998], bromide [Hayden and Linnemeyer, 1999; Cain et al. 2000], or others are injected and flushed through an aquifer to determine the swept volume of an aquifer. DNAPL partitioning tracers such as 2,2-dimethyl-3-pentanol or other alcohols are delayed in their transport through the aquifer by partitioning into and then out of DNAPL residuals. If the DNAPL partitioning coefficient between the DNAPL and the tracer is known then the saturation of DNAPL within the swept region can be determined. These conventional DNAPL partitioning tracer tests are not capable of delineating the three dimensional distribution of DNAPL residuals. This can be attributed to the fact that tracer breakthrough curves [BTCs] are collected at a well that integrates the arrivals of tracers from all directions. In addition, the interpretation of BTCs assumes homogeneity of the aquifer. Because of these limitations, the conventional partitioning tracer tests also are incapable of assessing DNAPL residuals held within less permeable regions of aquifers [Nelson et al. 1999].

Recently, Rao et al. [2000] and Khachikian and Harmon [2000] provided in-depth assessments of current technologies for source characterization and problems. Both have concluded that our knowledge of DNAPL dissolution and remediation issues has matured. Many innovative tracer techniques have been introduced and demonstrated for the enhanced characterization of DNAPL source zones. More importantly, they all emphasized the importance of heterogeneity in DNAPL spatial distributions, variability of hydraulic conductivity, moisture content, geochemistry, and the relation between hydraulic properties and DNAPL distributions. In particular, Khachikian and Harmon [2000] concluded that "*inverse modeling efforts aimed at exploiting the previous developments should be expanded to support field-scale characterization of DNAPL location and strength as a dissolving source*." Therefore, there is an urgent need for the development of cost-effective tools that can combine with the inverse algorithm for characterizing the hydraulic heterogeneity and spatial distribution of DNAPLs.

2.3. HYDROLOGICAL INVERSION FOR DETECTING HYDRAULIC HETEROGENEITY

Effects of hydraulic heterogeneity on flow and solute transport have been long recognized. Traditional approaches for characterizing the heterogeneity rely on the direct measurement of hydraulic properties [i.e., saturated hydraulic conductivity, porosity, or specific storage referred to as the **primary variables or information**]. The direct measurement of the properties is a difficult, time-consuming and costly task and requires invasive sampling. Characterization of the subsurface in great detail using the direct sampling technology therefore is rarely employed. On the other hand, the pressure head, water level, or concentration of tracers [referred to as **secondary variables or information]** is frequently collected from existing wells and this information is relatively abundant. Consequently, seeking an inverse method that takes advantage of abundant secondary information to estimate hydraulic properties of the aquifer is a logical step.

Hydrological inversion has been a major focus of groundwater hydrology during the last three decades [see Yeh, 1986; Sun, 1994 and McLaughlin and Townley, 1996 for a comprehensive review]. In general, the minimum-output-error [MOE] approach or its variation with Gauss-Newton's searching algorithm is a popular method. Despite its popularity, MOE is limited to identifying only a few zones of heterogeneity, providing images of heterogeneity only at a low resolution due to its inherently inefficient computational algorithm [Kitanidis, 1997].

In recent years, the geostatistics-based inverse technique [cokriging] has received increasing attention. Different from the MOE approach, it produces the first and second statistical moments of both primary and secondary variables, representing their most likely estimates and their uncertainty, respectively, conditioned on available observations. Cokriging relies on the classical linear predictor theory that considers spatial correlation structures of flow processes [such as pressure head and velocity] and the subsurface hydraulic property, and cross-correlation between the flow processes and the hydraulic property. In the past few decades, many researchers [e.g., Kitanidis and Vomvoris, 1983; Hoeksema and Kitanidis, 1984 and 1989; Rubin and Dagan, 1987; Gutjahr and Wilson, 1989; Sun and Yeh, 1992; Harvey and Gorelick, 1995; Yeh et al., 1995 and 1996] have demonstrated its ability to estimate hydraulic conductivity, head, velocity, and concentration of pollutants in heterogeneous aquifers.

The geostatistically-based approach is robust and its success has also been documented even for complex unsaturated flow in heterogeneous media. Yeh and Zhang [1996] developed an inverse technique based on cokriging to identify unsaturated hydraulic parameters in heterogeneous vadose zones under steady state and non-uniform flow conditions. They showed that unsaturated hydraulic parameters in the heterogeneous vadose zone could be identified if sufficient information on pressure and water saturation were available.

Cokriging is useful, but it is a linear predictor. The relation between primary and secondary variables of the subsurface however is highly nonlinear. Therefore, cokriging cannot fully exploit available secondary information. To overcome this limitation, Zhang and Yeh [1997] adapted a successive linear estimator technique [SLE] [Yeh et al., 1996] to the vadose zone inverse problem. They showed that the new approach could yield more detailed images of hydraulic parameters, vital to the prediction of contaminant transport than those produced by cokriging.

Because of the robustness of the SLE approach, Hughson and Yeh [2000] used it to develop a three-dimensional, transient flow model with the widely used moisture/pressure and unsaturated hydraulic conductivity/pressure relationships described by the van Genuchten-Mualem formulas. In addition, an efficient technique was developed to allow soil-water pressure and water content data observed at different times to be sequentially included in the inversion. This technique eliminates numerical difficulties associated with simultaneous inclusion of a large number of data sets. It increases the efficiency of the inverse model such that inverse modeling of a three-dimensional, heterogeneous vadose zone with a large number of primary variables becomes possible.

2.4. PARTITIONING TRACER INVERSION FOR DETECTING DNAPL DISTRIBUTION

To our knowledge, published work on estimation of DNAPL saturations in groundwater applications is limited. Jin et al. [1995] and Wilson and Mackay [1995] used partitioning tracers to estimate the average residual NAPL saturation over the length of the one-dimensional column. Sciortino et al. [2000] developed an inverse model to locate DNAPL pools. A recent study by

James et al. [1997] developed a stochastic method to estimate spatial distribution of NAPL residual content from tracer breakthrough curve moments.

Extending the work by James et al. [1997], Zhang and Graham [2001] presented a distributed-parameter extended Kalman filter for estimating spatially distributed residual saturation of NAPL and Darcy flux and for predicting site-specific movement of a partitioning tracer plume in a three-dimensional heterogeneous aquifer. Their results are exciting and the algorithm is promising even though the Kalman filter approach has inherent computational disadvantages. As both James et al. [1997] and Zhang and Graham [2001] stated, the Kalman filter approach requires significant computational effort and data storage. McLaughlin and Townley [1996] further pointed out that "*The extended Kalman filter is most useful in applications where the number of unknowns is relatively small… Although it has some attractive features, we believe that the extended Kalman filter should not be considered a practical inverse method until its convergence properties are better understood and it has been tested more extensively.*" Despite these issues, using 420 sampling locations [84 multilevel samplers with 5 samples at five evenly spaced vertical locations] out of a total of 686 elements [98 element in the horizontal plane and 7 elements in the vertical] of the simulation domain [see Zhang and Graham, 2001] is highly idealized. The use of such a large number of multilevel samplers is practically impossible for most field applications. Furthermore, a direct analysis of core samples from such a high-density borehole array might provide an accurate estimate of residual DNAPL distribution even without resorting to inverse modeling. Nevertheless, Zhang and Graham [2001] demonstrated that estimation of the residual DNAPL distribution using an inverse approach is theoretically feasible and has great potential.

2.5. RESEARCH PLAN

The above discussion and assessment of current technology for DNAPL source zone characterization leads to several facts. First, partitioning tracers are ideal indicator tracers to detect residual DNAPLs [Rao et al, 2000]. Second, the inverse modeling of partitioning tracers is possible for delineating residual DNAPL distributions. However, an inversion that delivers a detailed DNAPL distribution requires closely spaced multi-level sampler [MLS] to collect a vast amount of secondary information [Sillan et al., 1998]. Such a high-density invasive sampling array is cost-prohibitive and practically impossible in a real-world problem. In addition, the inversion of a field-scale problem demands a computationally efficient and effective inversion methodology. Based on these facts, we may therefore conclude that the development of a cost-effective monitoring technology that does not require high-density invasive samplings and the development of effective and efficient inverse algorithms are of great urgency for characterizing the DNAPL source zone. The hydraulic/partitioning tracer tomography [**HPTT**] and stochastic fusion of information proposed here is therefore the approach to meet the need.

2.6 REPORT ORGANIZATION

Two main tasks have been proposed in the initial version of the proposal and two additional tasks were added later to the project. The report has been organized by grouping the studies that have been completed to date under each task. Therefore, chapters 3 – 7 fall under Task 1, chapters 8 – 17 under Task 2, chapters 18 and 19 under Task 3, and chapter 20 fall under Task 4. Chapter 21 then summarizes the findings and conclusions that we have learned from all the studies completed under project ER-1365.

3. CHARACTERIZATION OF AQUIFER HETEROGENEITY USING TRANSIENT HYDRAULIC TOMOGRAPHY

3.1 INTRODUCTION

Detailed spatial distributions of hydraulic parameters are imperative to improve our ability to predict water and solute movement in the subsurface [e.g., Yeh, 1992, 1998]. Traditional aquifer tests like pumping tests and slug tests only yield hydraulic parameters integrated over a large volume of geologic media [e.g., Butler and Liu, 1993; Beckie and Harvey, 2002]. On the other hand, Wu et al. [2005] reported that the classical analysis for aquifer tests yields spurious transmissivity estimates and storage coefficient estimates that reflect local geology. For characterizing detailed spatial distributions of hydraulic parameters, a new method, hydraulic tomography [Gottlieb and Dietrich, 1995; Renshaw, 1996; Yeh and Liu, 2000; Liu et al., 2002; McDermott et al., 2003], which evolved from the CAT [Computerized Axial Tomography] scan concept of medical sciences and geophysics, appears to be a viable technology.

Hydraulic tomography is, in the most simplified terms, a series of cross-well interference tests. In other words, an aquifer is stressed by pumping water from or injecting water into a well, and monitoring the aquifer's response at other wells. A set of stress/response yields an independent set of equations. Sequentially switching the pumping or injection location, without installing additional wells, results in a large number of aquifer responses caused by stresses at different locations and, in turn, a large number of independent sets of equations. This large number of sets of equations makes the inverse problem [i.e., using aquifer stress and response relation to estimate the spatial distribution of hydraulic parameters] better posed, and the subsequent estimate approaches reality.

Interpreting data from hydraulic tomography presents a challenge, however. The abundance of data generated during tomography can lead to information overload, and cause substantial computational burdens and numerical instabilities [Yeh, 1986, Hughson and Yeh, 2000]. Moreover, the interpretation can be non-unique. Yeh and Liu [2000] developed a sequential successive linear estimator [SSLE] to overcome these difficulties. The SSLE approach eases the computational burdens by sequentially including information obtained from different pumping tests; it resolves the non-uniqueness issue by providing the best unbiased conditional mean estimate. That is, it conceptualizes hydraulic parameter fields as spatial stochastic processes and seeks their mean distributions conditioned on the information obtained from hydraulic tomography, as well as directly measured parameter values [such as from slug tests, or core samples]. Using sand box experiments, Liu et al. [2002] demonstrated that the combination of hydraulic tomography and SSLE is a propitious, cost-effective technique for delineating heterogeneity using a limited number of invasive observations. The work by Yeh and Liu [2000], nonetheless, is limited to steady state flow conditions, which may occur only under special field conditions. Because of this restriction, their method ignores transient head data before flow reaches steady state conditions. Transient head data, although influenced by both hydraulic conductivity and specific storage, are less likely to be affected by uncertainty in boundary conditions. The development of a new estimation procedure thus becomes essential so that all datasets collected during hydraulic tomography surveys can be fully exploited.

Few researchers have investigated transient hydraulic tomography. Bohling et al. [2002] exploited the steady-shape flow regime of transient flow data to interpret tomographic surveys. Under steady-shape conditions at late time of a pumping test before boundary effects take place,

the hydraulic gradient changes little with time -- a situation where sensitivity of head to the specific storage is small. As a consequence, the steady-shape method is useful for estimating hydraulic conductivity but not specific storage.

Their steady-shape method relies on the classical least-squares optimization method and the Levenberg-Marquardt algorithm [Marquardt, 1963] for controlling convergence issues [see Nowak and Cirpka, 2004]. This optimization method is known to suffer from non-uniqueness of the solutions if the inverse problem is ill posed and regularization [Tikhonov and Arsenin, 1977] or prior covariance of parameters [Nowak and Cirpka, 2004] is not used. The least-squares approach is also computationally inefficient if every element in the solution domain [in particular, three-dimensional aquifers with multiple, randomly distributed parameters] is to be estimated. This inefficiency augments if the sensitivity matrices required by the optimization are not evaluated using an efficient algorithm, such as the adjoint state approach.

These shortcomings may be the reasons that test cases in Bohling et al. [2002] were restricted to unrealistic, perfectly stratified aquifers, where the heterogeneity has no angular variations, and specific storage is constant and known a priori. The assumption of a spatially constant and known specific storage value for the entire aquifer makes the inverse problem almost the same as the steady hydraulic tomography as explored by Yeh and Liu [2000]. Perhaps inversion of the transient tomography by Bohling et al. [2002] is less affected by unknown in boundary conditions. Nonetheless, for perfectly horizontal layered aquifers, many traditional hydraulic test methods, without resorting to hydraulic tomography, can easily estimate hydraulic properties of each layer using just one borehole.

Similar to Vasco et al. [2000], Brauchler et al. [2003] developed a method that uses the travel time of a pneumatic pressure pulse to estimate air diffusivity of fractured rocks. Similar to X-ray tomography, their approach relies on the assumption that the pressure pulse travels along a straight line or a curve path. Thus, an analytical solution can be derived for the propagation of the pressure pulse between a source and a pressure sensor. Many pairs of sources and sensors yield a system of one-dimensional analytical equations. A least-squares based inverse procedure developed for seismic tomography can then be applied to the system of equations to estimate the diffusivity distribution. The ray approach avoids complications involved in numerical formulation of the three-dimensional forward and inverse problems, but it ignores interaction between adjacent ray paths and possible boundary effects. Consequently, their method requires an extensive number of iterations and pairs of source/sensor data to achieve a comparable resolution to that achieved from inverting a three-dimensional model. Vesselinov et al. [2001] applied an optimization technique and geostatistics to pneumatic cross-borehole tests in fractured rocks. Because of the baseline of the pneumatic properties is unknown, it is difficult to assess the accuracy of their results.

To our knowledge, few researchers have developed an inverse method for transient hydraulic tomography to estimate both hydraulic conductivity and specific storage of aquifers. For general groundwater inverse problems other than hydraulic tomography, Sun and Yeh [1992] assumed a specific storage field that was homogeneous and known a priori. They then developed a stochastic inverse method to estimate the spatial distribution of transmissivity using only transient head information. For transient hydraulic tomography, Vasco et al. [2000] and Brauchler et al. [2003] estimated diffusivity, the ratio of hydraulic conductivity to specific storage, without any attempt to separate the two parameters.

In this project, we extended the SSLE developed by Yeh and Liu [2000] to transient hydraulic tomography for estimating randomly distributed hydraulic conductivity and specific

storage in 3-D aquifers. This report begins with the derivation of the SSLE for use with transient hydraulic heads. We introduce a loop iteration scheme to improve the accuracy of sequential usage of head data. We then verify our new approach by applying it to a synthetic one-dimensional heterogeneous aquifer. During this one-dimensional test, temporal variation of cross-correlation between transient heads and parameters, as well as temporal correlation of transient heads, is investigated. Results of this investigation lead to a better understanding of effects of conditioning using head measurements on estimates of hydraulic conductivity and specific storage, and an effective sampling strategy, as opposed to developing an entire drawdown-time history, for efficient inversion of the transient hydraulic tomography data. Finally, the new SSLE is applied to a hypothetical three-dimensional, heterogeneous aquifer to demonstrate the robustness of our new approach.

3.2 METHOD

3.2.1 Groundwater Flow in Three-dimensional Saturated Media

In the following analysis, we assume that groundwater flow in three-dimensional, saturated, heterogeneous, porous media can be described by the following equation:

$$\nabla \cdot [K(\mathbf{x})\nabla H] + Q(\mathbf{x}_p) = S_s(\mathbf{x})\frac{\partial H}{\partial t} \qquad [3.1]$$

subject to boundary and initial conditions:

$$H\big|_{\Gamma_1} = H_1, \quad [K(\mathbf{x})\nabla H]\cdot \mathbf{n}\big|_{\Gamma_2} = q, \quad \text{and } H\big|_{t=0} = H_0 \qquad [3.2]$$

where in equation [3.1], H is total head [L], $\mathbf{x}$ is the spatial coordinate [$x = \{x_1, x_2, x_3\}$, [L], and x_3 represents the vertical coordinate and is positive upward], $Q[\mathbf{x}_p]$ is the pumping rate [1/T] at the location $\mathbf{x}_p$, $K[\mathbf{x}]$ is the saturated hydraulic conductivity [L/T], and $S_s[\mathbf{x}]$ is the specific storage [L^{-1}]. In equation [3.2], H_1 is the prescribed total head at Dirichlet boundary Γ_1, q is the specific flux [L/T] at Neumann boundary Γ_2, $\mathbf{n}$ is a unit vector normal to the union of Γ_1 and Γ_2, and H_0 represents the initial total head. The equations are solved by a 3-D finite element approach developed by Srivastava and Yeh [1992] in the following analysis.

3.2.2 Sequential Successive Linear Estimator [SSLE]

The SSLE approach is an extension of the SLE [Successive Linear Estimator] approach [Yeh et al., 1996; Yeh and Zhang, 1996; Zhang and Yeh, 1997; Hanna and Yeh, 1998; Vargas-Guzman and Yeh, 1999, 2002; Hughson and Yeh, 2000]. The SLE approach is essentially cokriging [Yeh et al., 1995] -- Bayesian formalism [Kitanidis, 1986] -- that seeks mean parameter fields conditioned on available point data as well as geologic and hydrologic structures [i.e., spatial covariance functions of parameters and hydraulic heads, and their cross-covariance functions]. Different from cokriging, SLE uses a linear estimator based on differences between observed and simulated hydraulic heads successively to update both conditional means and covariances of the estimates such that the nonlinear relation between information and parameters is considered. As a stochastic estimator analogous to the direct method of the deterministic approach [Yeh, 1986], SLE is conceptually the same as but

methodologically different from the maximum a posterior [McLaughlin and Townley, 1996] and the quasi-linear geostatistical inverse approach [Kitanidis, 1995].

The SSLE approach relies on the SLE concept to sequentially include data sets and update covariances and cross-covariances in the estimation process. The sequential method avoids solving huge systems of equations and therefore reduces numerical difficulties. The approach has been successfully applied to parameter estimations in variably saturated media [e.g., Zhang and Yeh, 1997; Hanna and Yeh, 1998; Hughson and Yeh, 2000], steady hydraulic tomography [Yeh and Liu, 2000; Liu et al., 2002], electrical resistivity tomography [Yeh et al., 2002]; and stochastic information fusion [Yeh and Šimůnek, 2002; Liu and Yeh, 2004]. In this study, we extend this inverse approach to incorporate transient hydraulic head data to estimate both hydraulic conductivity and specific storage fields. As the majority of the SSLE method used in this study remains similar to that in our previous works, we present only a brief summary, but a sensitivity analysis for transient flow, and a new loop iteration scheme are given in detail below.

To characterize the heterogeneity of geologic formations, the SSLE algorithm treats the natural logs of saturated hydraulic conductivity and specific storage as stochastic processes. We therefore assume $\ln K = \overline{K} + f$ and $\ln S_s = \overline{S} + s$, where $\overline{K}$ and $\overline{S}$ are mean values, and f and s denote the perturbations. The transient hydraulic head response to a pumping test in transient hydraulic tomography is represented by $H = \overline{H} + h$, where $\overline{H}$ is the mean and h is the perturbation. Substituting these stochastic variables into [4.1], taking the conditional expectation, and conditioning with some observations of head and parameters generates the mean flow equation as

$$\nabla \cdot [\overline{K}_{con}(\mathbf{x}) \nabla \overline{H}_{con}] + Q(\mathbf{x}_p) = \overline{S}_{con}(\mathbf{x}) \frac{\partial \overline{H}_{con}}{\partial t} \qquad [3.3]$$

where $\overline{K}_{con}$, $\overline{H}_{con}$, and $\overline{S}_{con}$ are conditional effective hydraulic conductivity, hydraulic head and specific storage, respectively [Yeh et al., 1996]. Similar to our previous work, we seek the conditional effective fields of hydraulic conductivity and specific storage, conditioned on the information from transient hydraulic tomography and some direct measurements of K and S_s.

The estimation procedure starts with a weighted linear combination of direct measurements of the parameters and transient head data at different locations to obtain the first estimate of the parameters. The weights are calculated based on statistical moments [namely, means, and covariances] of parameters, the covariances of heads in space and time, the cross-covariances between heads and parameters. The first estimate is then used in the mean flow equation [3.3] to calculate the heads at observation locations and sampling times [i.e., forward simulation]. Differences between the observed and simulated heads are determined subsequently. A weighted linear combination of these differences is then used to improve the previous estimates. Iterations between the forward simulation and estimation continue until the improvement in the estimates diminishes to a prescribed value.

a] Sensitivity analysis of transient flow

In the above estimation procedure, the head covariance in space and time and its cross-covariances with parameters are evaluated using a first-order approximation, which involves evaluation of sensitivity matrices of the governing flow equation. The sensitivity matrices are

evaluated as follows. Transient hydraulic heads are expanded in a Taylor series about the mean values of parameters. After neglecting second and higher order terms, the transient hydraulic head is:

$$H(\mathbf{x},t) = \bar{H}(\mathbf{x},t) + f(\mathbf{x})\frac{\partial H(\mathbf{x},t)}{\partial \ln K(\mathbf{x})}\Big|_{\bar{K},\bar{S}} + s(\mathbf{x})\frac{\partial H(\mathbf{x},t)}{\partial \ln S_s(\mathbf{x})}\Big|_{\bar{K},\bar{S}} \qquad [3.4]$$

The sensitivity terms $\frac{\partial H(\mathbf{x},t)}{\partial \ln K(\mathbf{x})}\Big|_{\bar{K},\bar{S}}$ and $\frac{\partial H(\mathbf{x},t)}{\partial \ln S_s(\mathbf{x})}\Big|_{\bar{K},\bar{S}}$ in equation [3.4] are calculated by the adjoint state method [Sykes, et al. 1985; Li and Yeh, 1998]. We briefly describe the method here [refer to Li and Yeh [1998, and 1999], Sun and Yeh [1992] for a detailed derivation]. The marginal sensitivity of a performance measure P to a parameter χ is defined as

$$\frac{dP}{d\chi} = \int_T \int_\Omega \left(\frac{\partial G}{\partial \chi} + \frac{\partial G}{\partial H}\frac{\partial H}{\partial \chi}\right) d\Omega dt \qquad [3.5]$$

where T and Ω represent time and spatial domain, respectively. The first term of the integral in equation [3.5] indicates the direct dependence of P on χ, while the second term indicates the implicit dependence of P on χ through the heads [Sykes et al., 1985]. In this case,

$$G = H\delta(x - x_k)(t - t_l) \qquad [3.6]$$

representing the hydraulic head at location x_k and time t_l, where δ is Kronecker delta-function which equals unity if x equals x_k and t equals t_l , and equals zero otherwise. We choose an arbitrary function ϕ^* that satisfies

$$S\frac{\partial \phi^*}{\partial t} + \nabla \cdot (K\nabla \phi^*) - \delta(x - x_k)(t - t_l) = 0 \qquad [3.7]$$

with boundary and final conditions:

$$\phi^*\Big|_{\Gamma_1} = 0, \qquad [K(\mathbf{x})\nabla \phi^*]\cdot \mathbf{n}\Big|_{\Gamma_2} = 0, \qquad \phi^*\Big|_{t=T_e} = 0 \qquad [3.8]$$

[note that equations [3.7] and [3.8] are called adjoint state equations]; we further assume that the initial condition is known a priori, such that $\phi\big|_{t=0} = 0$, and hydraulic conductivity and specific storage are not correlated to each other. Thus, the sensitivities of the hydraulic head at location $\mathbf{x}_k$ and time t_l to f and s at location $\mathbf{x}_n$ are given by

$$\frac{\partial H(\mathbf{x}_k,t_l)}{\partial \ln K(\mathbf{x}_n)} = \int_T \int_\Omega \left\{ \frac{\partial K(\mathbf{x})}{\partial \ln K(\mathbf{x}_n)} \frac{\partial \phi^*}{\partial x_i}\frac{\partial H}{\partial x_i} \right\} dt d\Omega \qquad [3.9]$$

$$\frac{\partial H(\mathbf{x}_k, t_l)}{\partial \ln S_s(\mathbf{x}_n)} = \int_T \int_\Omega \left\{ \frac{\partial S(\mathbf{x})}{\partial \ln S_s(\mathbf{x}_n)} \phi^* \frac{\partial H}{\partial t} \right\} dt d\Omega \qquad [3.10]$$

where $\ln K[\mathbf{x}_n]$ and $\ln S_s\ [\mathbf{x}_n]$ are the $\ln K$ and $\ln S_s$ at element n , respectively, when the study domain is discretized.. Note that the adjoint state equations are also transient problems and need to be solved backwardly in time. Also, the mean transient hydraulic heads must be derived beforehand in order to evaluate the sensitivities. The mean flow equation is given by equation [3.3]. After ϕ^* and the mean head are calculated, the sensitivities obtained from equations [3.9] and [3.10] can be used to calculate head covariances and its cross-covariances with parameters, using a first-order approximation [Hughson and Yeh, 2000].

b] Loop iteration scheme

As indicated by Vargas-Guzman and Yeh [2002] and Yeh and Šimůnek [2002] in previous SSLE approaches, the method of adding different data sets sequentially works best for linear systems. The relations between transient head and hydraulic parameters, however, are nonlinear; the sequential approach cannot fully exploit the head information. For instance, assume two datasets, A and B, are used in an inversion problem. The B dataset is added after the A dataset reaches convergence. The SSLE then stops after the B dataset converges. While the final estimates meet the convergence criteria for the B dataset, they may not now meet the convergence criteria for the A dataset. In addition, adding datasets in different sequences may lead to different results. Therefore, we introduced a new loop iteration scheme.

In this loop iteration scheme, the next dataset is added after all the datasets already incorporated meet the converge criteria within one loop. Specifically, a dataset is fed into SSLE first, and SSLE then iterates until this dataset meets a converge criterion. A new dataset is added afterwards, and SSLE again iterates until the new estimate convergences. Instead of adding the next new dataset, the scheme goes back to check the convergence for the first dataset. If the converge criterion is not met, the program starts a loop iteration in which the iteration involves both the first and second datasets. That is, the first dataset is iterated once, and then the second dataset is incorporated and iterated once also; we call this process a loop. The loop iteration continues until both datasets meet the converge criterion within one loop. Then, the next new dataset is added. The algorithm treats this new dataset similarly to the second dataset, except the loop iteration now involves three datasets. Additional datasets are added in a similar way. As a consequence, our inverse approach improves estimates throughout the loops, maximizes the usefulness of datasets, and alleviates the problems associated with our previous SSLE approach.

During a transient pumping test, one can record a large number of head observations at different times. As stated by Sun and Yeh [1992], simultaneous inclusion of transient head data at different times improves the estimates and decreases the head misfit because simultaneous inclusion considers the temporal correlation of transient heads. In our approach, we included in the estimation some selected observed heads at different times during a pumping activity. The head responses from different pumping tests are included sequentially.

3.3 NUMERICAL EXAMPLES

3.3.1 One-Dimensional Flow

To test our inverse approach, a hypothetical, one-dimensional, horizontal, heterogeneous, confined aquifer was used. The aquifer was 20 meters long and was discretized into twenty elements. Each element was one meter long. The left and right sides of the aquifer were set as prescribed head conditions with hydraulic heads of 100 m. Each element was assigned a hydraulic conductivity value and a specific storage value using a stochastic random field generator [Gutjahr, 1989]. The geometric mean of hydraulic conductivity was 0.0026 m/s and the geometric mean of specific storage was 0.0001 m^{-1}. The variance of ln$K$ was 0.5 and the variance of lnS_s was 0.2. The correlation scales for both parameters are 5 m and ln$K$ and lnS_s are assumed to be independent from each other, representing the worst scenario.

Using this one-dimensional aquifer, a pumping test was simulated at location x = 9.5 m with a pumping rate of 0.005 m^3/s. The flow approached a steady state condition after 19 seconds of pumping; about 95% of total drawdown occurred in the first 8 seconds of the pumping test. The cross correlation between head and parameters during the pumping test was evaluated using a first-order approximation and then examined. Figure 3.1 depicts behaviors of the cross correlation between the observed h at x=9.5 m and f's at different locations in the aquifer at three selected times [2, 4, and 6 seconds]. Likewise, Figure 3.2 depicts behaviors of the cross correlation between the h and s's at the three times. The cross correlation between the h and f's generally decreased with the distance away from the head observation location [x=9.5 m] but the cross correlation over the entire aquifer increased with time. And the number of f's having significant cross correlation [say, cross correlation values greater than 0.4] with the head at the observation location increased. Shapes of the cross-correlation functions are different from those in uniform flow [Mizell et al., 1980] due to converging flow and boundary conditions. During a uniform flow, a head is negatively correlated with the hydraulic conductivity values down gradient and positively correlated with the hydraulic conductivity up gradient. Figure 3.2 shows that the cross correlation between the h and the s field decreased with time. At early time, strong cross correlations between the h and s's are confined to the vicinity of the observed head location. These cross correlations, nevertheless, dropped drastically at late time. Such results suggest that a head measurement in a well at late time can provide good estimates of f's over a large portion of the aquifer. On the other hand, head measurements in a well can only yield information of the s nearby and only early time data are useful for the estimate of s. This finding supports the conclusion by Wu et al. [2005] that the storage coefficient estimate from a traditional aquifer test based on the drawdown-time data in an observation well, induced by pumping at another well, is dominated by the local geology between the pumping well and the observation well. Furthermore, to obtain good estimates of f and s during hydraulic tomography tests, head information, encompassing the entire pumping process -- including early time and late time, should be used. The resolution of the estimated f field will be better than that of the s field because of the localized influence of a head measurement on the estimate of s field.

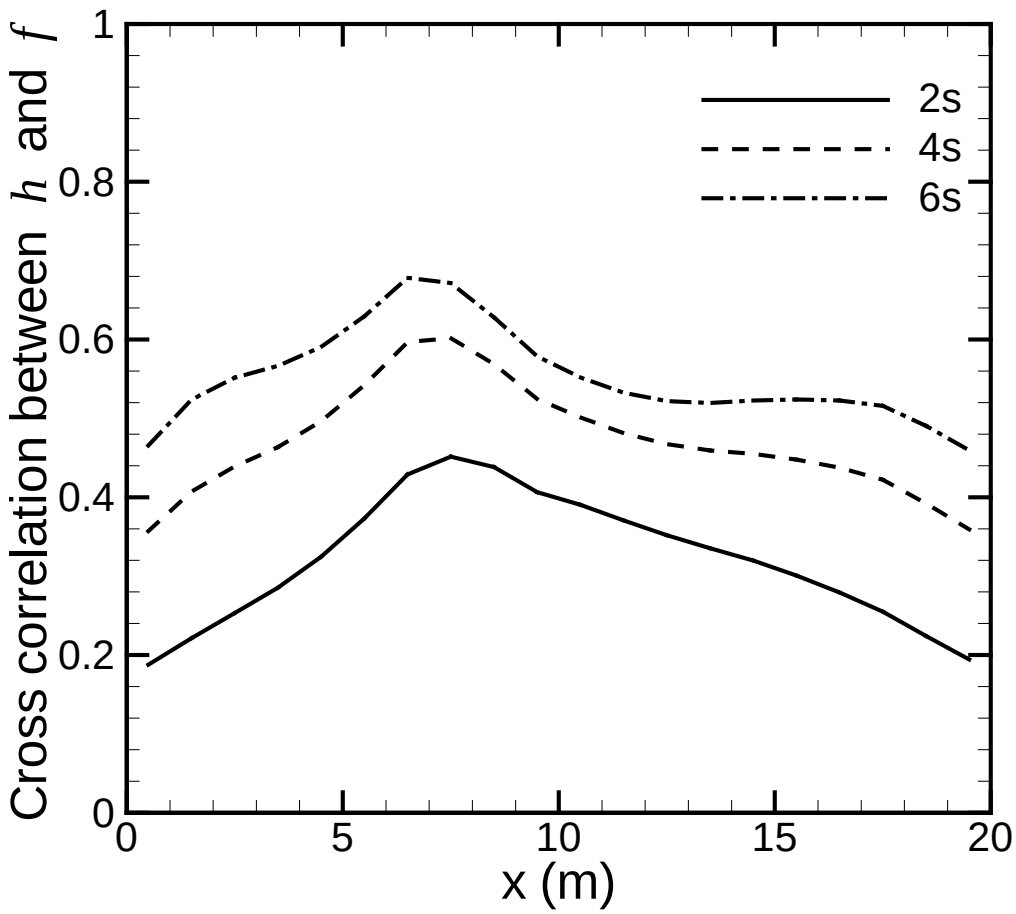


Figure 3.1: Cross correlation between *h* at x=9.5 m and *f*'s at different locations for three selected times during a pumping test.

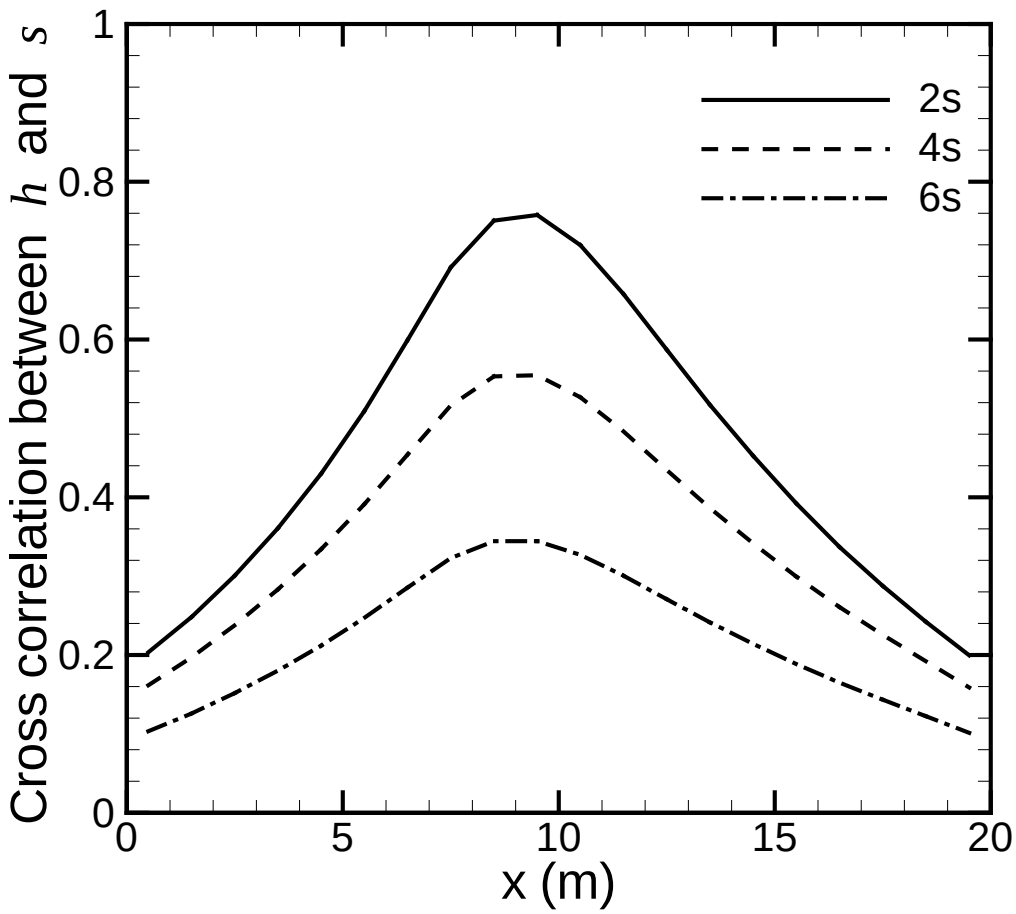


Figure 3.2: Cross correlation between *h* at x=9.5 m and *s*'s at different locations for three selected times during a pumping test.

The temporal correlation of transient heads was also evaluated. Figure 3.3 shows the contours of the temporal correlation of the head at x= 7.5 m from the beginning of the pumping test to 8 seconds. As indicated in the figure, the heads at different times were highly correlated, especially at later time. The high correlation suggests that the heads at a given observation location at different times provide overlapping information. In particular, inclusion of heads at all time steps would be very computational time consuming for our estimator because the adjoint equations [3.7] and [3.8] must be solved once for each head observation in time. Because of the overlapping head information, choosing heads at several time steps instead of using heads at all time steps would significantly reduce the computation burdens and keep the usefulness of head information.

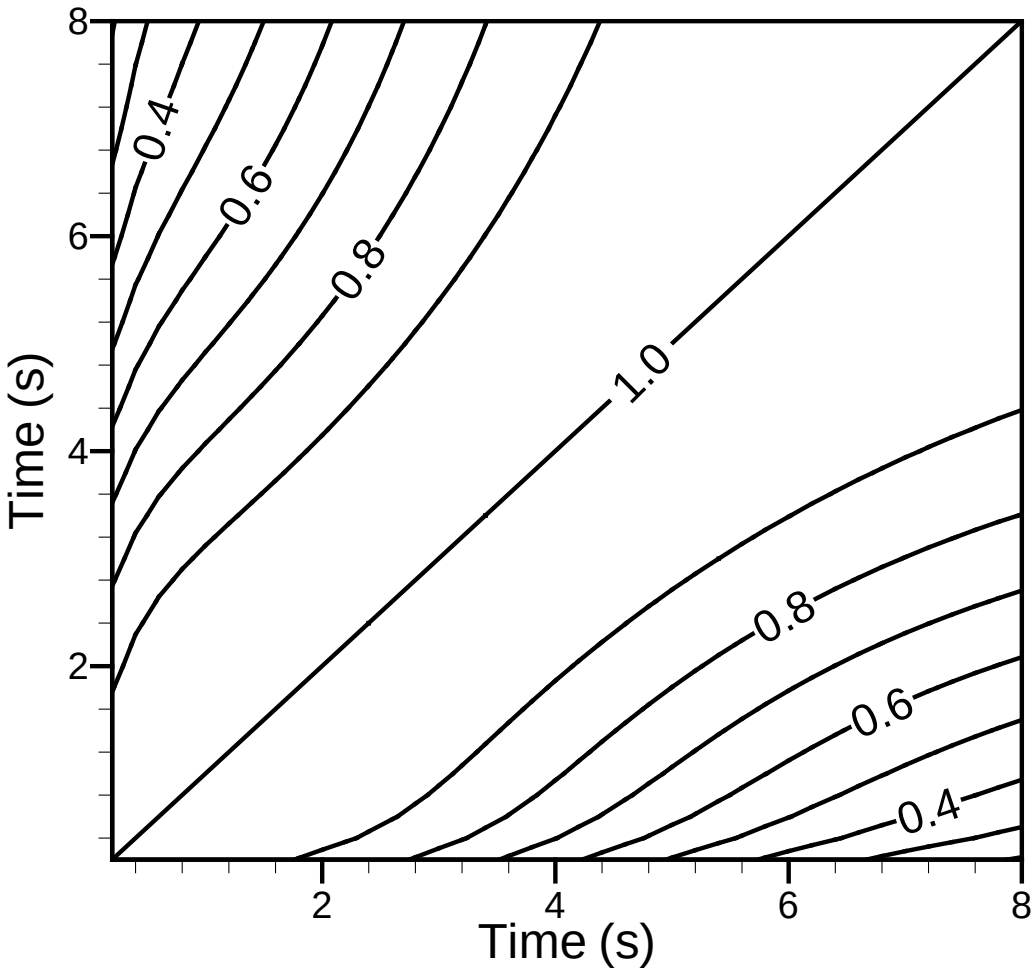


Figure 3.3: Temporal correlation of transient heads at x =7.5 m during a pumping test.

Based on the cross correlation and temporal correlation analysis, we thereafter tested our inverse approach for a well-posed inverse problem [deterministic inverse problems, Yeh et al., 1996]. The head responses of all elements were collected at 2, 4, and 6 seconds, representing early, middle, and late times of the pumping test, respectively. One direct hydraulic conductivity measurement and one specific storage measurement were also assumed to be known at element #20 [i.e., the boundary fluxes are known]. Therefore, the necessary and sufficient conditions for inverse modeling [i.e., the transient head responses of all elements at two time steps, as well as boundary conditions] are fully specified [Sun, 1996 and Yeh and Šimůnek, 2002]. The inverse problem thus becomes well posed and both parameter fields can be uniquely determined. Figures 3.4 and 3.5 compare the true hydraulic conductivity field and specific storage with estimates, respectively. The comparisons indicate that our new algorithm produces accurate estimates for both parameter fields for the deterministic case, and the accuracy of our SSLE method is thus ensured.

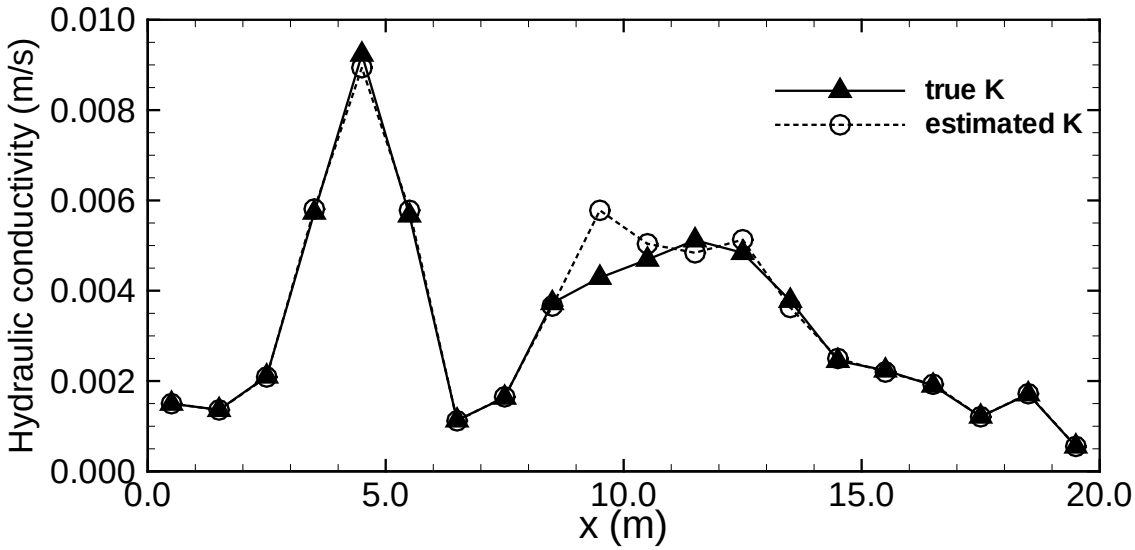


Figure 3.4: Comparison of estimated hydraulic conductivity with true hydraulic conductivity in a deterministic case.

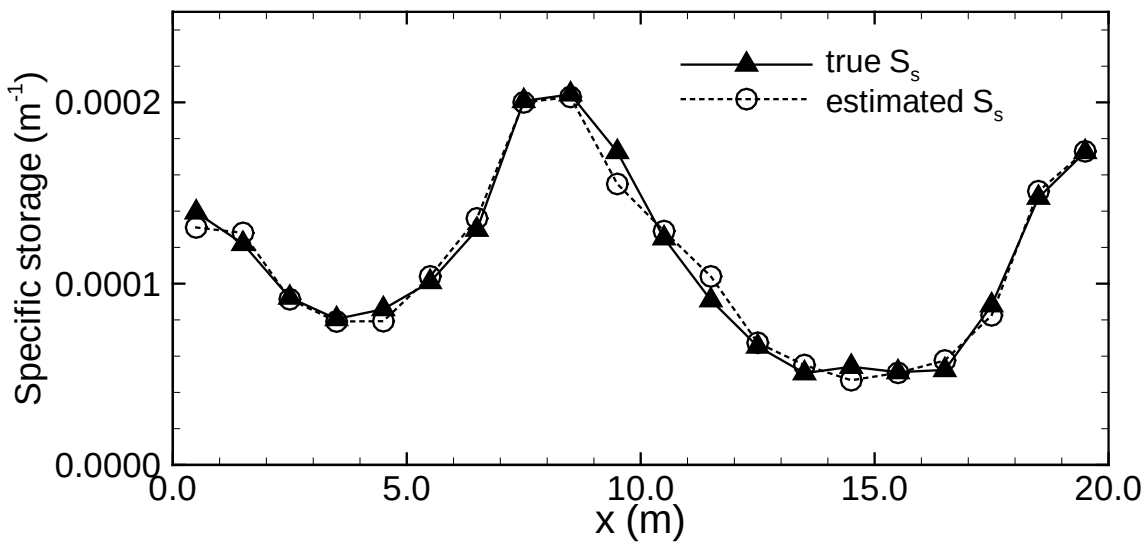


Figure 3.5: Comparison of estimated specific storage with true specific Storage in a deterministic case.

Next, we applied transient hydraulic tomography to the one-dimensional heterogeneous aquifer to demonstrate the benefit of a transient hydraulic tomography test. Four locations in the one dimensional aquifer were selected as pumping and observation wells. These four wells were located at x=3.5 m, 7.5 m, 11.5 m, and 15.5 m. The first pumping activity was initiated at x = 3.5 m, and the corresponding head responses at all four wells were recorded. The pumping rate, pumping time, and observation times were the same as the pumping test of the previous deterministic case. The three additional pumping tests had the same configuration as the first one, except the pumping was initiated at x = 7.5 m, x = 11.5 m, and x = 15.5 m for the second, third, and fourth pumping test, respectively. As a result, a total of 48 head responses were collected to estimate both parameters. Comparisons of the estimated hydraulic conductivity and

specific storage with true parameters are shown in Figures 3.6 and 3.7, respectively. The two figures show that, with only four head observation locations out of a total of twenty elements of the entire aquifer, the hydraulic tomography with our SSLE approach produces close estimates of the true spatial patterns for both parameters. As demonstrated in Figures 3.6a, b, c, and d, and Figure 3.7a, b, c, and d, the estimates progressively improved as more head responses from tomographic pumping tests were incorporated into our SSLE approach. However, the improvement of estimates from the third to the fourth pumping test was small, which indicates that excessive pumping tests only offer negligible improvements for the given number of observation wells. These findings are similar to those reported by Yeh and Liu [2000].

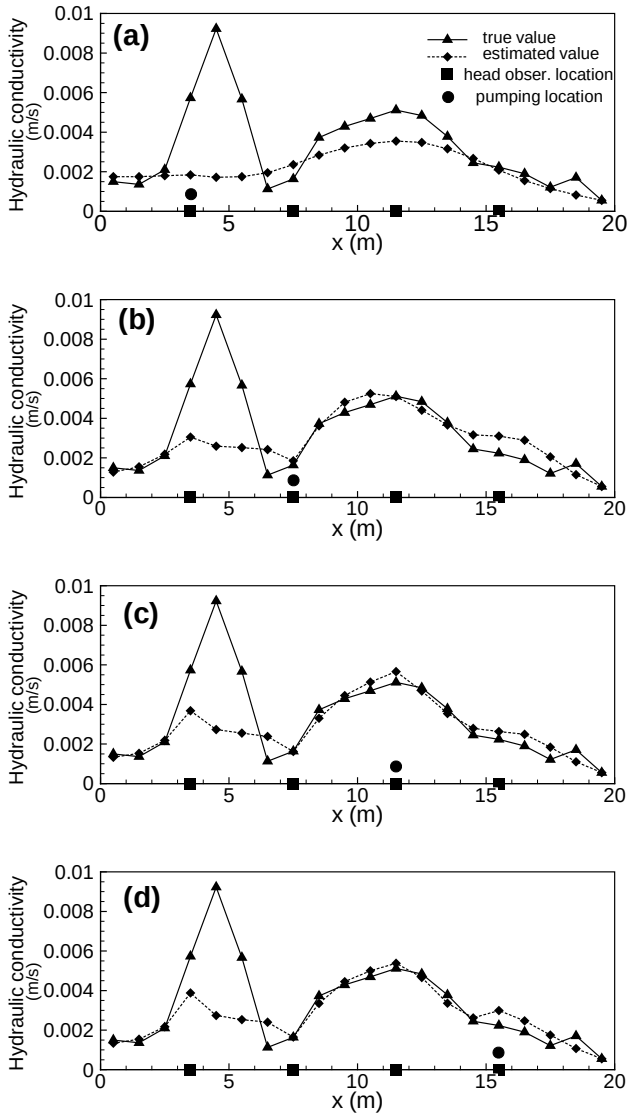


Figure 3.6: Estimated hydraulic conductivity from transient hydraulic tomography: [a] estimates from the first pumping test; [b] estimates from the additional second pumping test; [c] estimates from the additional third pumping test; [d] estimates from the additional fourth pumping test.

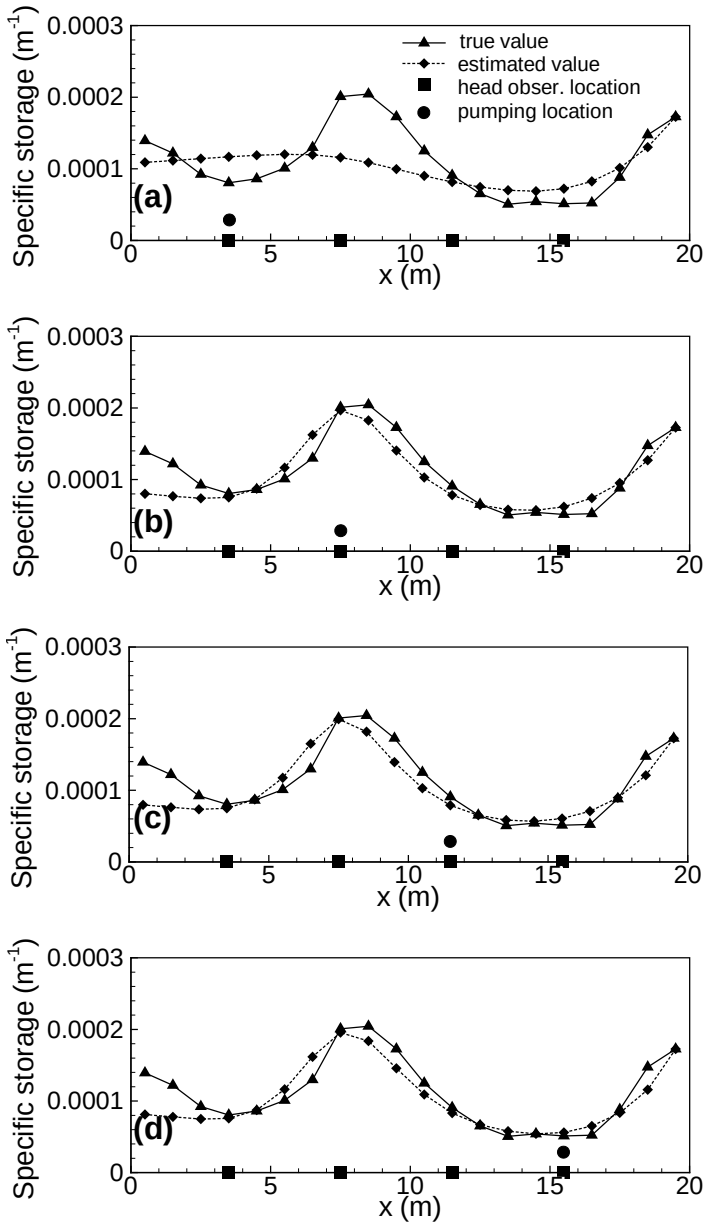


Figure 3.7: Estimated specific storage from transient hydraulic tomography: [a] estimates from the first pumping test; [b] estimates from the additional second pumping test; [c] estimates from the additional third pumping test; and [d] estimates from the additional fourth pumping test.

3.3.2 Three-Dimensional Heterogeneous Aquifer

We subsequently applied our SSLE to transient hydraulic tomography in a synthetic three-dimensional heterogeneous confined aquifer. The geometry of this synthetic heterogeneous aquifer had dimensions of 15 m × 15 m × 15 m, and was discretized into 3375 elements. Each element had a uniform size of 1 m × 1 m × 1 m. The bottom and the top boundaries were set as no-flow, and the remaining four sides were assumed to be a prescribed hydraulic head of 100 m. A three-dimensional Cartesian coordinate system was used for spatial references. The coordinates of the bottom corner at the inner center of the cube [see Figure 3.8] were assigned to be [0, 0, 0] and the upper corner opposite to the bottom corner was assigned as [15, 15, 15]. The heterogeneous parameter fields again were generated by the spectral method [Gutjahr, 1989]. The geometric mean of K was 0.34 m/d and the variance of ln$K$ was 0.5, while the geometric mean of S_s was 0.0002 m^{-1} and the variance of lnS_s was 0.1. The correlation scales in the x, y, and z directions were 20 m, 20 m, and 2 m, respectively.

Four fully penetrating, multi-level wells were placed vertically in the aquifer. The horizontal coordinates for the four wells were [3.5, 3.5], [11.5, 3.5], [3.5, 11.5], and [11.5, 11.5]. Each well had seven head observation ports whose vertical coordinates were 1.5 m, 3.5 m, 5.5 m, 7.5 m, 9.5 m, 11.5 m, and 13.5 m, respectively. Each well also had two pumping ports whose vertical coordinates were 4.5 m and 10.5 m, respectively. One direct hydraulic conductivity measurement and one specific storage measurement were assumed to be known at location [3.5, 3.5, 8.5]. A pumping test was performed at one of the pumping ports with a constant pumping rate of 150 m^3/d. The pumping test was simulated for 0.01 day with a time step of 0.0005 day. The head responses at all 28 observation points were monitored at time 0.002 day, 0.006 day, and 0.01 day. Seven additional pumping tests were simulated, using the same pumping rate and pumping time period, but different pumping ports. A total of 672 head observations were used in our SSLE approach to simultaneously estimate hydraulic conductivity and specific storage.

The SSLE was implemented on a parallel computing platform using the LINUX operating system; the interpretation of the hydraulic tomography tests was carried out using a 10-node PC cluster [Pentium 4 2.8 GHz CPU each]; the total computing time for the interpretation was 610 minutes.

Figures 3.8a, b, c, and d plot the estimated hydraulic conductivity after two, four, six, and eight pumping tests, respectively, and the true hydraulic conductivity field is shown in Figure 3.8e. The estimated specific storage fields after two, four, six, and eight pumping tests are illustrated in Figures 3.9a, b, c, and d with the true field shown in Figure 3.9e. Both figures 3.8 and 3.9 show that the general pattern of heterogeneity of the aquifer was already captured just from the first two pumping tests; after eight pumping tests greater details were revealed, but the improvement rate diminished as more pumping tests were conducted.

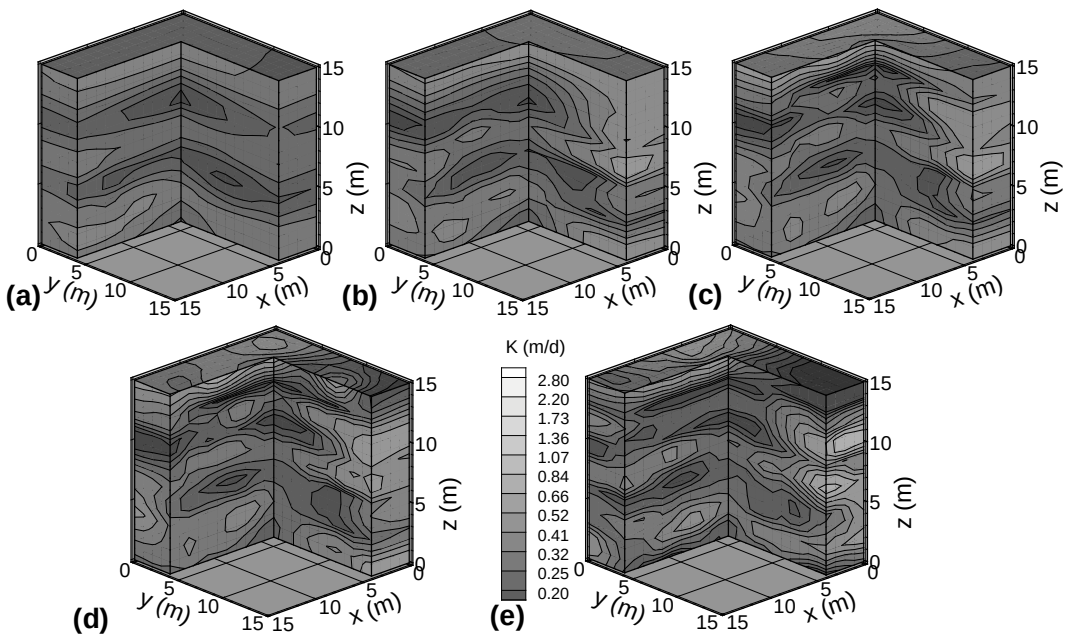


Figure 3.8: Comparison between estimated hydraulic conductivity with the true field in a three dimensional aquifer. Estimated hydraulic conductivity field after [a] two, [b] four, [c] six, [d] and eight pumping tests, and [e] the synthetic true hydraulic conductivity field.

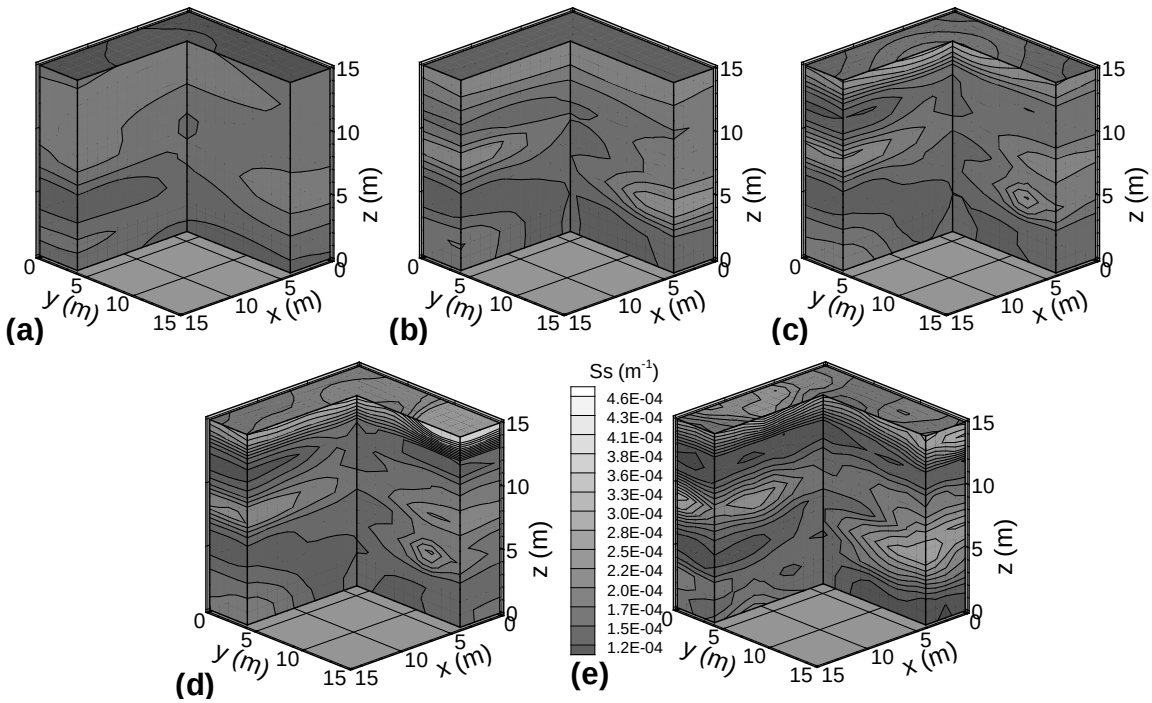


Figure 3.9: Comparison between estimated specific storage with the true field in a three dimensional aquifer. Estimated specific storage field after [a] two, [b] four, [c] six, [d] and eight pumping tests, and [e] the synthetic true specific storage field.

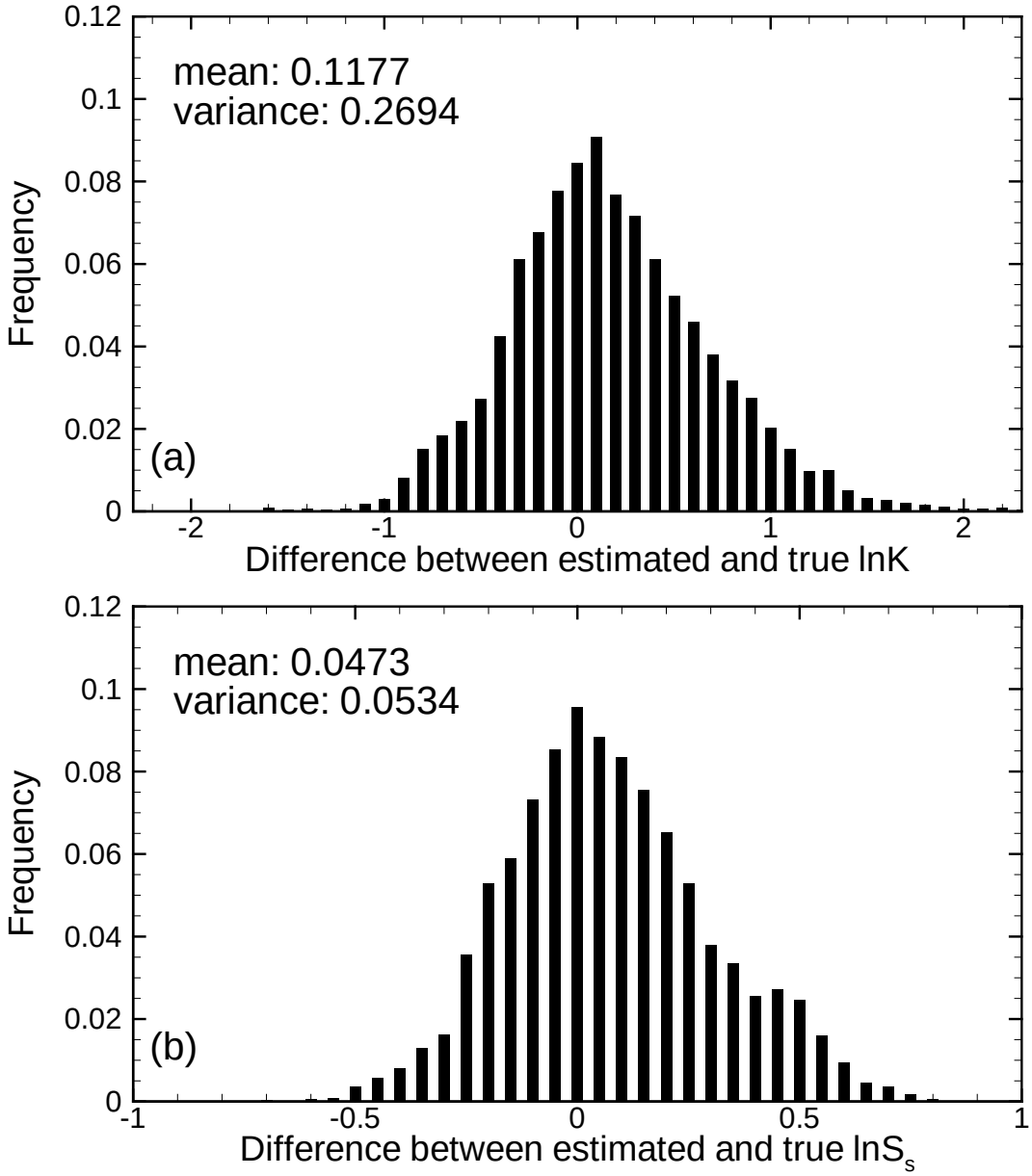


Figure 3.10: Frequency distributions of estimation errors: [a] hydraulic conductivity field and [b] specific storage field of a three-dimensional aquifer.

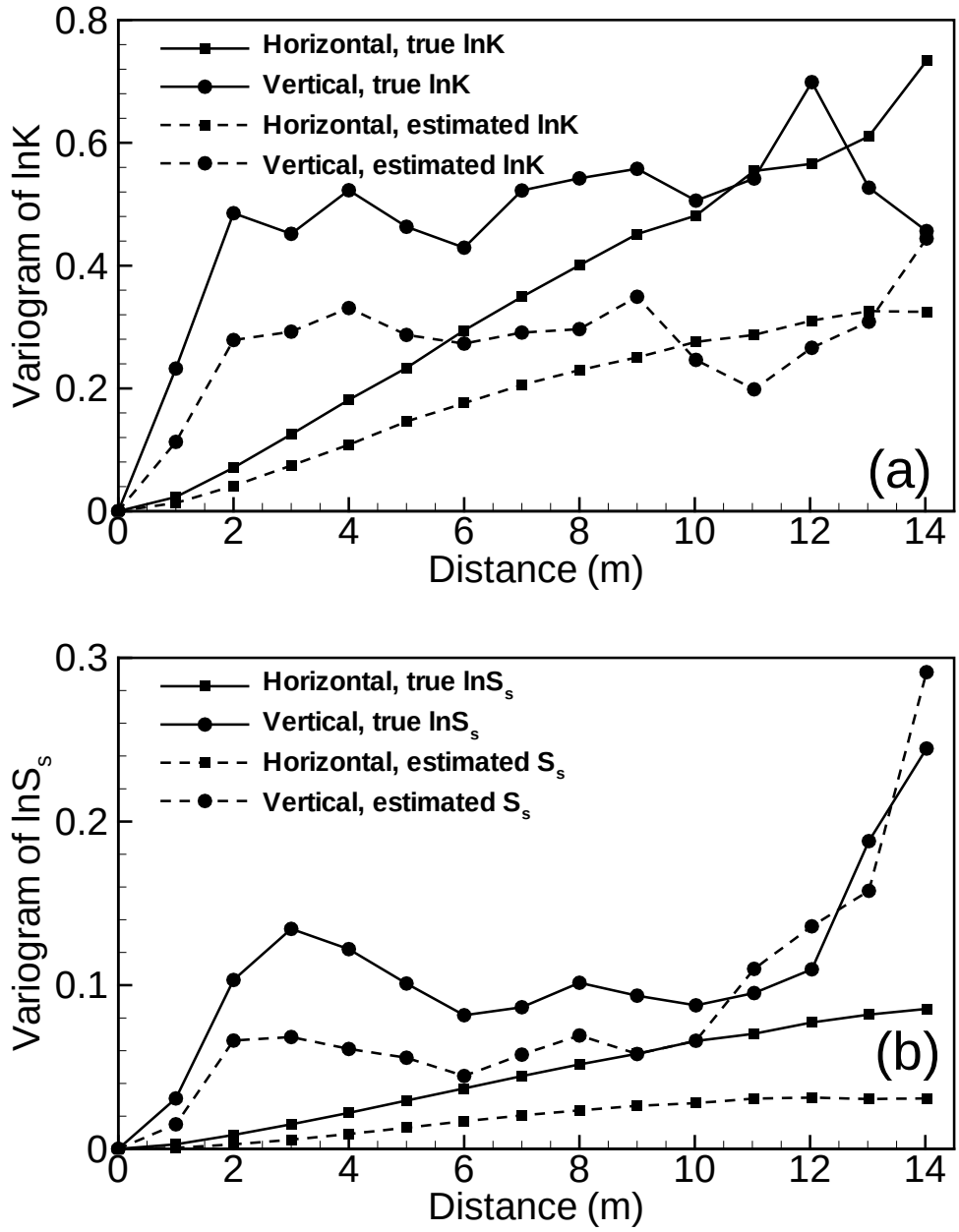


Figure 3.11: Comparison of variograms between [a] estimated hydraulic conductivity with true field; [b] estimated specific storage with true field in a three-dimensional aquifer.

Figure 3.10a shows a frequency distribution with the mean and variance of the difference between the true log hydraulic conductivity field and that estimated [i.e., estimation errors] after eight pumping tests and the distribution of the estimation error of log specific storage and their mean and variance are illustrated in Figure 3.10b. The error distributions are approximately normal, indicative of unbiasness of our estimator. The slight bias in the estimates can be attributed to the effective nature of the estimated parameters [Yeh et al., 1996; Hanna and Yeh, 1998]. The horizontal and vertical variograms of estimated and true hydraulic conductivity and specific storage fields are depicted in figures 3.11a and b, respectively. Generally speaking,

variograms of the estimates have similar spatial patterns as those of the true fields, in both horizontal and vertical directions. The variances of the estimates are expected to be lower and their correlation scales were longer than true ones [see Table 3.1]. This difference is due to the conditional expectation approach embedded in the SSLE method and insufficient data.

Table 3.1: Comparison of statistical properties of the 3-D aquifer.

	true lnK	estimated lnK	true $\ln S_s$	estimated $\ln S_s$
mean	-1.079	-0.96	-8.52	-8.47
variance	0.50	0.30	0.10	0.06
horizontal correlation scale[m]	20	30	20	35
vertical correlation scale [m]	2	4	2	6

Robust as they are, neither the hydraulic tomography nor our SSLE is a perfect method. The more head observations are collected, the higher the resolution of the estimates will be [i.e., there is no optimal]. Inaccurate head observations and hydraulic property measurements [i.e., noises] during hydraulic tomography unequivocally can lead to an inaccurate estimate or instability of the estimate. While our SSLE can overcome the impacts of noise, the estimates become smooth, which means there is a loss of effectiveness of information. These issues have been discussed in Yeh and Liu [2000].

4. ANALYSIS OF HYDRAULIC TOMOGRAPHY USING TEMPORAL MOMENTS OF DRAWDOWN-RECOVERY DATA

4.1 INTRODUCTION

Multi-scale heterogeneity of geologic media is a rule rather than the exception. The knowledge of detailed spatial distributions of hydraulic properties is imperative to predict water and solute movement in the subsurface at high resolution [e.g., Yeh, 1992, 1998]. Traditional aquifer tests [e.g., pumping and slug tests] have been widely employed for estimating hydraulic properties of the subsurface for the last few decades. Besides their costly installation and invasive natures, Beckie and Harvey [2002] reported that slug tests can yield dubious estimates of the storage coefficient of aquifers. Validity of classical analysis for aquifer tests was also questioned by Wu et al. [2005]. They reported that the storage coefficient, S, value obtained from the traditional Theis analysis primarily represents a weighted average of S values over the region between the pumping and observation wells. In contrast to the S estimate, the transmissivity, T, estimate is a weighted average of all T values in the entire domain with relatively high weights near the pumping well and the observation well. In concordance with the finding by Oliver [1993], Wu et al. [2005] concluded that the T estimate can be influenced by any large-sized or strong anomaly within the cone of depression. Thus, interpretation of the meaning of T estimates can be highly uncertain. As a result, previous assessments of transmissivity distributions of aquifers may be subject to serious doubt.

Hydraulic tomography [Gottlieb and Dietrich, 1995; Yeh and Liu, 2000; Liu et al., 2002; Bohling et al., 2002; Zhu and Yeh, 2005], based on the CAT [Computerized Axial Tomography] scan concept of medical sciences, is potentially a viable technology for characterizing detailed spatial distributions of the hydraulic properties. Hydraulic tomography, in a simple term, is a series of cross-well interference tests. More precisely, hydraulic tomography involves stressing an aquifer by pumping water from or injecting water into a well, and monitoring the aquifer's response at other wells. A set of stress and response data yields an independent set of equations. Sequentially switching the pumping or injection location, without installing additional wells, results in a large number of aquifer responses induced by the stresses at different locations and thus, a large number of independent sets of equations. This large number of sets of equations makes the inverse problem [i.e., using aquifer stress and response relation to estimate the spatial distribution of hydraulic parameters] better posed, and the subsequent estimates approach reality.

Interpretation of hydraulic tomography surveys however is a numerical challenge. The large number of well hydrographs generated during tomography often leads to information overload, substantial computational burdens, and numerical instabilities [Hughson and Yeh, 2000]. To overcome these difficulties, Yeh and Liu [2000] developed a sequential successive linear estimator [SSLE] approach. This approach eases the computational burdens by sequentially including information obtained from different pumping tests; it resolves the non-uniqueness issue by providing an unbiased mean of an abstracted stochastic parameter rather than the actual parameter. That is, it conceptualizes hydraulic parameter fields as spatial stochastic processes and seeks their conditional effective parameter distributions. The conditional effective parameters are conditioned on the data obtained from and governing physical principles of hydraulic tomography, as well as our prior knowledge of the geologic structure, and directly measured parameter values [such as from slug tests, or core samples]. The SSLE approach in essence is the Bayesian formalism. Sand box experiments by Liu et al. [2002] and Illman et al. [2005] proved that the combination of hydraulic tomography and SSLE is a

viable, cost-effective technique for delineating heterogeneity using a limited number of invasive observations. The work by Yeh and Liu [2000], nonetheless, is limited to steady state flow conditions, which may occur only under special field conditions. Because of this restriction, their method ignores transient head data before flow reaches steady state conditions.

Several researchers have investigated THT. Bohling et al. [2002] exploited the steady-shape flow regime of transient flow data to interpret tomographic surveys. Similar to Vasco et al. [2000], Brauchler et al. [2003] developed a method that uses the travel time of a pneumatic pressure pulse to estimate air diffusivity of fractured rocks. As in X-ray tomography, their approach relies on the assumption that the pressure pulse travels along a straight line or a curved path. Thus, an analytical solution can be derived for the propagation of the pressure pulse between a source and a pressure sensor. Many pairs of sources and sensors yield a system of one-dimensional analytical equations. A least-squares based inverse procedure developed for seismic tomography can then be applied to the system of equations to estimate the diffusivity distribution. The ray approach avoids complications involved in numerical formulation of the three-dimensional forward and inverse problems, but it ignores interaction between adjacent ray paths and possible boundary effects. Consequently, their method requires an extensive number of iterations and pairs of source/sensor data to achieve a comparable resolution to that achieved from inverting a three-dimensional model. Vesselinov et al. [2001] applied an optimization technique and geostatistics to interpret pneumatic cross-borehole tests in fractured rocks.

Recently, Zhu and Yeh [2005] extended the SSLE approach to THT to estimate both spatially varying hydraulic conductivity and specific storage fields in 3-D random media. They demonstrated that it is possible to obtain detailed hydraulic conductivity and specific storage fields of aquifers using few wells with THT surveys. But as the size of the field site to be characterized increases and the demands of resolution increases, the computational burden increases significantly. A computationally efficient algorithm therefore must be developed for speedy analysis of the THT surveys. For basin-scale naturally recurrent tomographic surveys [such as river-stage tomography, Yeh et al. [2004]], development of such a technology is imperative.

In this study, inspired by the moment generating function approach by Harvey and Gorelick [1995], we develop the temporal moment-generation equation for impulse pumping tests, similar to the recent work by Li et al. [2005]. While Li et al. [2005] focus on applying temporal moments to a single impulse pumping test, we apply the temporal moments to transient hydraulic tomography. Specifically, we incorporate the zeroth and first temporal moments of well hydrographs into the SSLE inverse approach [Yeh and Liu, 2000] for THT. In addition, we implement a loop iteration scheme [Zhu and Yeh, 2005] to avoid the effects of sequential addition of moment information. By directly comparing the estimation using the temporal moments with that using transient heads, we thereafter investigate the temporal moment approach in terms of computational efficiency and accuracy of estimation. Last, limitations of the moment approach are discussed.

4.2 DERIVATION OF MOMENT EQUATIONS

Groundwater flow induced by pumping in three-dimensional, saturated, heterogeneous geologic media is assumed to be described by the following equation:

$$\nabla \cdot [K(\mathbf{x})\nabla H] + Q(\mathbf{x}_p, t) = S_S(\mathbf{x})\frac{\partial H}{\partial t} \qquad [4.1]$$

subject to boundary conditions:

$$H\Big|_{\Gamma_1} = H_1 \text{ and } [K(\mathbf{x})\nabla H]\cdot\mathbf{n}\Big|_{\Gamma_2} = q\,,$$

and initial condition:

$$H(\mathbf{x})\big|_{t=0} = H_0(\mathbf{x}) \qquad [4.2]$$

In Eq. [4.1], H is the total head [L], $\mathbf{x}$ is the spatial coordinate [$\mathbf{x}$ = {x, y, z}, [L], and z represents the vertical coordinate and is positive upward], Q[$\mathbf{x}_p$, t] is the pumping rate per unit volume at location $\mathbf{x}_p$, K[$\mathbf{x}$] is the saturated hydraulic conductivity [L/T], and S_s[$\mathbf{x}$] is the specific storage [L^{-1}]. In Eq. [4.2], H_1 is the prescribed total head at Dirichlet boundary Γ_1, q is the specific flux [L/T] at Neumann boundary Γ_2, $\mathbf{n}$ is a unit vector normal to the boundary, and H_0[$\mathbf{x}$] represents the initial total head distribution under a steady state condition.

If we define drawdown as $s = H\text{-}H_0$, then the drawdown form of the three-dimensional saturated flow equation is given by

$$\nabla\cdot[K(\mathbf{x})\nabla s] + \nabla\cdot[K(\mathbf{x})\nabla H_0] + Q(\mathbf{x}_p,t) = S_s(\mathbf{x})\frac{\partial s}{\partial t} \qquad [4.3]$$

with boundary and initial conditions:

$$s\Big|_{\Gamma_1} = 0\,, [K(\mathbf{x})\nabla s]\cdot\mathbf{n}\Big|_{\Gamma_2} = q - K\cdot\nabla H_0\,, \text{ and } s\big|_{t=0} = 0\,. \qquad [4.4]$$

In Eq. [4.3], $\nabla\cdot[K(\mathbf{x})\nabla H_0]$ represents the divergence of the regional flow prior to pumping tests and is zero as long as it is steady. Applying a moment generating function approach [Harvey and Gorelick, 1995] to the drawdown-time data, the nth temporal moments of drawdown at location $\mathbf{x}_i$ are given by

$$M_n(\mathbf{x}_i) = \int_0^\infty t^n s(\mathbf{x}_i,t)dt \qquad [4.5]$$

where M_n[x_i] is the nth temporal moment of drawdown at location $\mathbf{x}_i$. Notice that $Y = M_1/M_0$ is a characteristic time depicting the arrival time of the center of the area under a drawdown-recovery curve. Multiplying Eq. [4.3] with t^n and integrating over time from 0 to infinite gives

$$\int_0^\infty t^n\nabla\cdot[K(\mathbf{x})\nabla s]dt + \int_0^\infty t^n\nabla\cdot[K(\mathbf{x})\nabla H_0]dt + \int_0^\infty t^n Q(\mathbf{x}_p,t)dt = \int_0^\infty t^n S_s(\mathbf{x})\frac{\partial s}{\partial t}dt \qquad [4.6]$$

Substituting Eq. [4.5] into Eq. [4.6], assuming that the regional flow is steady [i.e., the second term on the left-hand side of Eq. [4.6] is zero], and that the duration of pumping at location $\mathbf{x}_p$

with a constant rate of Q is τ, and using integration by parts for the right-hand term, we obtain the moment equation:

$$\nabla \cdot [K(\mathbf{x})\nabla M_n] + Q(\mathbf{x}_p)\frac{\tau^{n+1}}{n+1} = S_s(\mathbf{x}) s t^n \Big|_{t=0}^{t=\infty} - S_s(\mathbf{x}) n M_{n-1} \qquad [4.7]$$

Applying the same procedure to Eq. [4.4], the boundary conditions for Eq. [4.7] become

$$M_n\Big|_{\Gamma_1} = 0 \quad \text{and} \quad K(\mathbf{x})\frac{\partial M_n}{\partial x_i}\cdot \mathbf{n}\Big|_{\Gamma_2} = \int_0^\infty t^n (q - K\nabla H_0)dt \qquad [4.8]$$

Letting n=0 in Eq. [4.7] leads to the zeroth moment equation

$$\nabla \cdot [K(\mathbf{x})\nabla M_0] + Q(\mathbf{x}_p)\tau = S_S(\mathbf{x}) s \Big|_{t=0}^{t=\infty} \qquad [4.9]$$

The zeroth moment represents the area under the drawdown-recovery curve. The solution to Eq. [4.9] requires our knowledge of drawdown at every point in the solution domain at both initial and final time. While drawdown at t=0 is generally zero everywhere, the spatial distribution of the drawdown in a real-world aquifer after pumping starts is difficult to know with a limited number of wells and current technologies. The final drawdown throughout the aquifer however will be zero after pumping is stopped and enough time is allowed for aquifer to recover. If this is the case, then, Eq. [4.9] becomes

$$\nabla \cdot [K(\mathbf{x})\nabla M_0] + Q(\mathbf{x}_p)\tau = 0 \qquad [4.10]$$

The associated boundary conditions are

$$M_0\Big|_{\Gamma_1} = 0 \quad \text{and} \quad K(\mathbf{x})\frac{\partial M_0}{\partial x_i}\cdot \mathbf{n}\Big|_{\Gamma_2} = \int_0^\infty (q - K\nabla H_0)dt \qquad [4.11]$$

Now, setting n=1 in Eq. [4.7], the governing equation for the first moment is

$$\nabla \cdot [K(\mathbf{x})\nabla M_1] + Q(\mathbf{x}_p)\frac{\tau^2}{2} + S_S(\mathbf{x}) M_0 = 0 \qquad [4.12]$$

The associated boundary conditions are

$$M_1\Big|_{\Gamma_1} = 0 \quad \text{and} \quad K(\mathbf{x})\frac{\partial M_1}{\partial x}\cdot \mathbf{n}\Big|_{\Gamma_2} = \int_0^\infty t(q - K\nabla H_0)dt \qquad [4.13]$$

Eqs. [4.10], [4.11], [4.12], and [4.13] are essentially the governing equations for the zeroth and first temporal moments of drawdown anywhere in the aquifer induced by a single pumping during a hydraulic tomographic survey. They are identical to those equations developed by Li et

al. [2005] for an aquifer test that involves one pumping location in an aquifer. By using the moments of the drawdown, a parabolic equation [i.e., Eq. 4.1] has been transformed to two Poisson's equations [Eq. [4.10] and Eq. [4.12]]. In other words, the governing transient groundwater flow equation is replaced by two steady equations which can be solved directly without using any time-march scheme.

4.3 SEQUENTIAL INVERSE ALGORITHM

Our estimation technique using drawdown moments from hydraulic tomography is based on the sequential successive linear estimator [SSLE] developed by Yeh et al. [1996], Zhang and Yeh [1997], Li and Yeh [1999], Hughson and Yeh [2000], Vargas and Yeh [1999 and 2002], Yeh and Liu [2000], Liu and Yeh [2004], in particular, Zhu and Yeh [2005]. We first assume that the natural logs of saturated hydraulic conductivity and specific storage are stochastic processes. One advantage of using the natural logarithm is that it avoids negative values of the parameters during the estimation. We then assume $\ln K = \overline{K} + f$ and $\ln S_s = \overline{S} + b$, where ln denotes natural logarithm; $\overline{K}$ and $\overline{S}$ are mean values; f and b denote the perturbations. Similarly, the zeroth moment of drawdown induced by a pumping test during transient hydraulic tomography can be decomposed into two parts: $M_0 = \overline{M}_0 + m_0$, where $\overline{M}_0$ is the mean and m_0 is the perturbation. The characteristic time Y can also be decomposed as $Y = \overline{Y} + y$, where $\overline{Y}$ is the mean and y is the perturbation.

Expanding the zeroth moment and characteristic time in a Taylor series about the mean values of parameters, and neglecting second and higher order terms, the perturbations at location i can be expressed as:

$$m_{0i} = f_j \frac{\partial M_{0i}}{\partial \ln K_j}\Big|_{\overline{K},\overline{S}} = \mathbf{f}\mathbf{J}_{M_0 \ln K} \quad y_i = f_j \frac{\partial Y_i}{\partial \ln K_j}\Big|_{\overline{K},\overline{S}} + b_j \frac{\partial Y_i}{\partial \ln S_j}\Big|_{\overline{K},\overline{S}} = \mathbf{f}\mathbf{J}_{Y \ln K} + \mathbf{b}\mathbf{J}_{Y \ln S} \qquad [4.14]$$

where a repeated subscript implies summation of its range. In Eq. [4.14], f_j and b_j are perturbation of $\ln K$ and $\ln S$ at location j and j=1, …N, which is the total number of elements in the domain, and they are denoted by matrices **f** and **b**. The sensitivity matrices of M_0 and Y at location i with respect to $\ln K$ and $\ln S$ perturbation at location j are given by $\mathbf{J}_{M_0 \ln K} = \frac{\partial M_{0i}}{\partial \ln K_j}$, $\mathbf{J}_{Y \ln K} = \frac{\partial Y_i}{\partial \ln K_j}$ and $\mathbf{J}_{Y \ln S} = \frac{\partial Y_i}{\partial \ln S_j}$. Assuming K and S_s are independent from each other, the covariances of m_0 and y, and the cross-covariance between m_0 and f, between y and f, and between y and b can be obtained by a first-order analysis and is expressed as

$$
\begin{aligned}
&\mathbf{R}_{m_0 m_0}(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{J}_{M_0 \ln K}\mathbf{R}_{ff}(\mathbf{x}_i,\mathbf{x}_j)\mathbf{J}^T_{M_0 \ln K} \\
&\mathbf{R}_{yy}(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{J}_{Y \ln K}\mathbf{R}_{ff}(\mathbf{x}_i,\mathbf{x}_j)\mathbf{J}^T_{Y \ln K} + \mathbf{J}_{Y \ln S}\mathbf{R}_b(\mathbf{x}_i,\mathbf{x}_j)\mathbf{J}^T_{Y \ln S} \\
&\mathbf{R}_{m_0 f}(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{J}_{M_0 \ln K}\mathbf{R}_{ff}(\mathbf{x}_i,\mathbf{x}_j) \\
&\mathbf{R}_{yb}(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{J}_{Y \ln S}\mathbf{R}_b(\mathbf{x}_i,\mathbf{x}_j) \\
&\mathbf{R}_{yf}(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{J}_{Y \ln K}\mathbf{R}_{ff}(\mathbf{x}_i,\mathbf{x}_j)
\end{aligned}
\qquad [4.15]
$$

where $\mathbf{R}_{ff}(\mathbf{x}_i,\mathbf{x}_j)$ and $\mathbf{R}_{bb}(\mathbf{x}_i,\mathbf{x}_j)$ are covariance matrices of *f* and *b* between location $\mathbf{x}_i$ and $\mathbf{x}_j$, respectively. $\mathbf{R}_{m_0 f}(\mathbf{x}_i,\mathbf{x}_j)$, $\mathbf{R}_{yf}(\mathbf{x}_i,\mathbf{x}_j)$, $\mathbf{R}_{yb}(\mathbf{x}_i,\mathbf{x}_j)$ are cross-covariance matrices between m_0 and f, y and *f*, and y and *b*, respectively. $\mathbf{R}_{m_0 m_0}(\mathbf{x}_i,\mathbf{x}_j)$ and $\mathbf{R}_{yy}(\mathbf{x}_i,\mathbf{x}_j)$ are covariance matrices of m_0 and *y*, respectively. The superscript *T* represents the transpose of the matrix.

The estimate of *f* at location $\mathbf{x}_i$, $\hat{f}(\mathbf{x}_i)$, is then derived using a linear multi-variate estimator:

$$
\hat{f}(\mathbf{x}_i) = \alpha_j f_j^* + \lambda_k m_{0k}^* + \beta_\ell y_\ell^* \qquad j = 1,\ldots N_f; \quad k = 1,\ldots N_{m_0}; \quad \ell = 1,\ldots N_y \qquad [4.16]
$$

where α_j, λ_k and β_l are weights. N_f, N_{m_0} and N_y are the total number of *f*, zeroth moment, and characteristic time measurements, respectively. Measurements for hydraulic conductivity, zeroth moment, and characteristic time are denoted by f^*, m_0^*, and y^*, respectively. Similarly, the estimate of *b* at location $\mathbf{x}_i$, $\hat{b}(\mathbf{x}_i)$, is given by

$$
\hat{b}(\mathbf{x}_i) = \varepsilon_n b_n^* + \mu_\ell y_\ell^* \qquad n = 1,\ldots N_b; \quad \ell = 1,\ldots N_y \qquad [4.17]
$$

where ε_n and μ_k are weights, and N_b is the total number of b^*'s, which are specific storage measurements. All weights are evaluated based on the covariance and cross covariances [Eq. 4.15] [see Hughson and Yeh, 2000]. Our inverse algorithm then successively updates parameter estimates and residual covariances and sequentially includes data sets from different pumping tests. The updating and sequentially inclusion procedures are not presented here. Detailed derivations are described by Hughson and Yeh [2000] for inversion of unsaturated flow, Yeh and Liu [2000] for hydraulic tomography, and Yeh et al., [2002] for electrical resistivity tomography.

4.4 EVALUATION OF MOMENT SENSITIVITIES

The SSLE approach requires evaluation of the sensitivity matrices. To reduce the computational cost, these sensitivities are evaluated by an adjoint state method [Sykes et al. 1985; Sun and Yeh, 1992]. A detailed derivation of the sensitivities is given at Appendix D. The sensitivity of M_0 to ln*K* is given by

$$
\frac{\partial M_0^k}{\partial \ln K_\ell} = \int_\Omega -\frac{\partial \phi_0^*}{\partial \mathbf{x}} K_\ell \frac{\partial M_0}{\partial \mathbf{x}} d\Omega . \qquad [4.18]
$$

The sensitivity of M_1 to ln$K$ is given by

$$\frac{\partial M_1^k}{\partial \ln K_\ell} = \int_\Omega \left[-\frac{\partial \phi_0^*}{\partial \mathbf{x}} K_\ell \frac{\partial M_0}{\partial \mathbf{x}} - \frac{\partial \phi_1^*}{\partial \mathbf{x}} K_\ell \frac{\partial M_1}{\partial \mathbf{x}} \right] d\Omega \,, \qquad [4.19]$$

and the sensitivity of of M_1 to ln$S$ is given as

$$\frac{\partial M_1^k}{\partial \ln S_\ell} = \int_\Omega \phi_1^* M_0 S_{s\ell} d\Omega \,, \qquad [4.20]$$

where ϕ_0^* and ϕ_1^* are arbitrary functions which are defined in Appendix D; the superscript k denotes the observation at location $\mathbf{x}_k$ and the subscript ℓ denotes the location of the parameter in the domain Ω. To evaluate Eqs. [4.18], the arbitrary functions ϕ_0^* must be known a priori. ϕ_0^* is obtained by solving the adjoint equation [D16]. On the other hand, to calculate the sensitivity of M_1 to ln$K$ or lnS_s [Eqs. [4.19] and [4.20]], the adjoint equations Eqs. [D9] and [D10] are solved sequentially. That is, we derive ϕ_1^* first and then use it in Eq. [D10] to obtain ϕ_0^*.

Both moment Eq. [4.10] and Eq. [4.12] are Poisson's equations, which need to be solved only once, as opposed to the parabolic equation of the governing equation for transient groundwater flow, which must be solved for each time step. As a consequence, significant computational costs can be reduced during evaluation of sensitivity matrices. Consider that the hydrograph at an observation well has been recorded at a one-second interval for a period of 100 seconds during an impulse pumping test. Suppose that this hydrograph is used directly to estimate the hydraulic properties and the adjoint state method is employed to evaluate the associated sensitivity matrices. For each drawdown measurement at a given time, an associated adjoint equation must be solved once. The adjoint equation for transient flow is a parabolic equation [see Zhu and Yeh, 2005]. For each given time, the solution to the adjoint equation must be obtained by a backward time-marching scheme to derive the sensitivity of drawdown with respect to a parameter. If the computational time-step in the evaluation of the sensitivity is assumed to be the same as the sampling interval [1 second], the total number of times that the adjoint matrix equations need to be solved for the 100 measurements will be 5050 [i.e., a sum of the numbers from 1 to 100]. Since drawdowns are highly correlated in time, the number of drawdown measurements used in the estimation can be reduced [Zhu and Yeh, 2005]. Say, if only three samples at 40s, 60s, and 80s from the 100 measurements of the hydrograph are selected for the sensitivity analysis, then the total number of times the system equation for the arbitrary function needs to be solved is reduced to 180.

Now, suppose that we use the moment approach. Instead of using the 3 transient drawdown measurements, only the zeroth and first moments of the hydrograph are needed. Subsequently, one only has to solve the adjoint equation for the zeroth moment Eq. [D16] once and then the two adjoint equations for first moment equation, Eq. [D9] and Eq. [D10]. That is, only three system equations have to be solved for the sensitivity analysis. Here, the example considers just one observation well. As the number of observation wells increases, the

difference becomes even more significant. The temporal moments approach thus is computationally efficient, compared to the drawdown-time approach.

Notice that the zeroth moment equation [Eq. 4.10] only depends on the hydraulic conductivity. As a result, the influence of the specific storage is avoided during estimation of the hydraulic conductivity even under transient flow conditions.

4.5 NUMERICAL EXAMPLES

We created a 2-D heterogeneous confined aquifer to compare computational cost between temporal moments and transient heads approaches for the hydraulic tomography analysis. Since temporal moments are integrated forms of a well hydrograph, they may lose some information that reflects effects of heterogeneity at different parts of an aquifer as the cone of depression evolves. This loss of information may be significant, especially since only the first two moments are considered. Therefore, we also investigated the impacts of information loss on our estimates of hydraulic properties.

The aquifer considered is 100 m long and 100 m wide and is discretized into 2500 elements with a dimension of 2 m by 2 m. The aquifer has spatially varying transmissivity [T] and storage coefficient [S] fields. Both T and S fields are generated by a spectral random field generator [Gutjahr 1989]. The geometric mean of the T field is $0.0035 m^2/s$ and the variance of $\ln T$ for the field is 0.6. Meanwhile, the geometric mean of the S field and its variance in term of $\ln S$ are 0.00023 and 0.2, respectively. The correlation scales for materials in both the x direction and y direction are assumed to be 20 meters. Figure 4.1 plots the generated T and S fields. All boundaries of the aquifer as well as its initial head distribution are assigned to be a constant head of 100 m. Four cases are simulated. Case 1 involves nine wells uniformly distributed on a grid with the distance between two adjacent wells 1.5 times the correlation length of the aquifer properties [see Figure 4.2 for well locations]. This case represented a sparse monitoring network. Nine pumping tests were sequentially simulated at these wells with a pumping rate of $0.1 m^3/s$. The pumping time for each pumping test lasted 50 seconds and then the pumping stopped to allow a full recovery. During each pumping test, head responses were obtained from the nine wells, and we thus have 81 hydrographs after the hydraulic tomography survey. According to Zhu and Yeh [2005], a few selected transient heads in a hydrograph are needed to obtain accurate estimates of hydraulic properties during hydraulic tomography. Based on their findings, we selected the transient heads collected at 10 s, 30 s, 50 s, 70 s, 100 s, and 200 s to estimate T and S simultaneously. These sampling times were selected to capture the behaviors of the rising and falling limbs of the hydrograph.

The well field of Case 1 was used in Case 2 but zeroth moments and characteristic times of the drawdown at the nine wells due to nine sequential pumpings are used to estimate T and S. Since determining the time to fully reach recovery is difficult, calculating moments from the drawdown-recover curve could introduce some measurement error. Introducing measurement errors could obscure the comparison. Consequently, these moments in Case 2 are simulated directly from moment Eqs [4.10] and [4.12] for this case. In practice, estimating moments from observed hydrographs would inevitably involve some errors, but they are difficult to quantify in a simple manner.

An “optimal” monitoring network based on the results by Yeh and Liu [2000] is considered in Case 3. In this case, 25 observation wells [see figure 4 for well locations] were added to the aquifer, which are distributed with a distance [20 m] between two adjacent wells [i.e., one correlation length]. The pumping wells are the same as the nine used in Case 2. As a

result, we had 25 well hydrographs for each pumping test and a total of 225 well hydrographs after the nine sequential pumping tests. Similar to Case 1, we selected the transient heads collected at 10 s, 30 s, 50 s, 70 s, 100 s, and 200 s for estimating *T* and *S*. The corresponding moment approach of Case 3 is presented as Case 4. Instead of using the head data, 225 pairs of the zeroth moments and characteristic times from the 25 wells were simulated for the nine sequential pumping. Then, the moments were used in the estimation. The four numerical cases were executed on a PC-Cluster of 4 processors [2.8GHz, 1G memory each] platform. The simulation times for the four cases were 423.8, 64.4, 766.1, and 179.2 minutes, respectively. Figure 4.2 compares the estimated transmissivity fields of Case 1 and Case 2. Comparison of the estimated storage coefficient fields for these two cases is shown in Figure 4.3. Estimated transmissivity fields of Case 3 and Case 4 are shown in Figure 4.4, and Figure 4.5 compares estimated storage coefficient fields for these two cases.

Table 4.1: Mean and variance of estimation errors.

	Case 1		Case 2		Case 3		Case 4	
	ln*T*	ln*S*	ln*T*	ln*S*	ln*T*	ln*S*	ln*T*	ln*S*
Mean	0.1496	0.0756	0.1642	0.1360	0.0264	-0.0211	0.0499	-0.0124
Variance	0.2435	0.2563	0.1485	0.1575	0.1835	0.1962	0.1410	0.1918

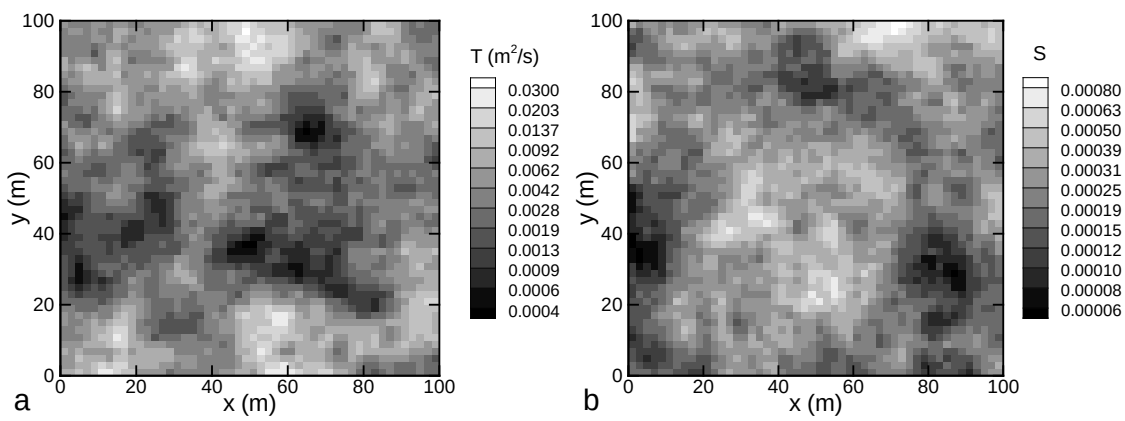


Figure 4.1: True synthetic fields: a] transmissivity ; b] storage coefficient.

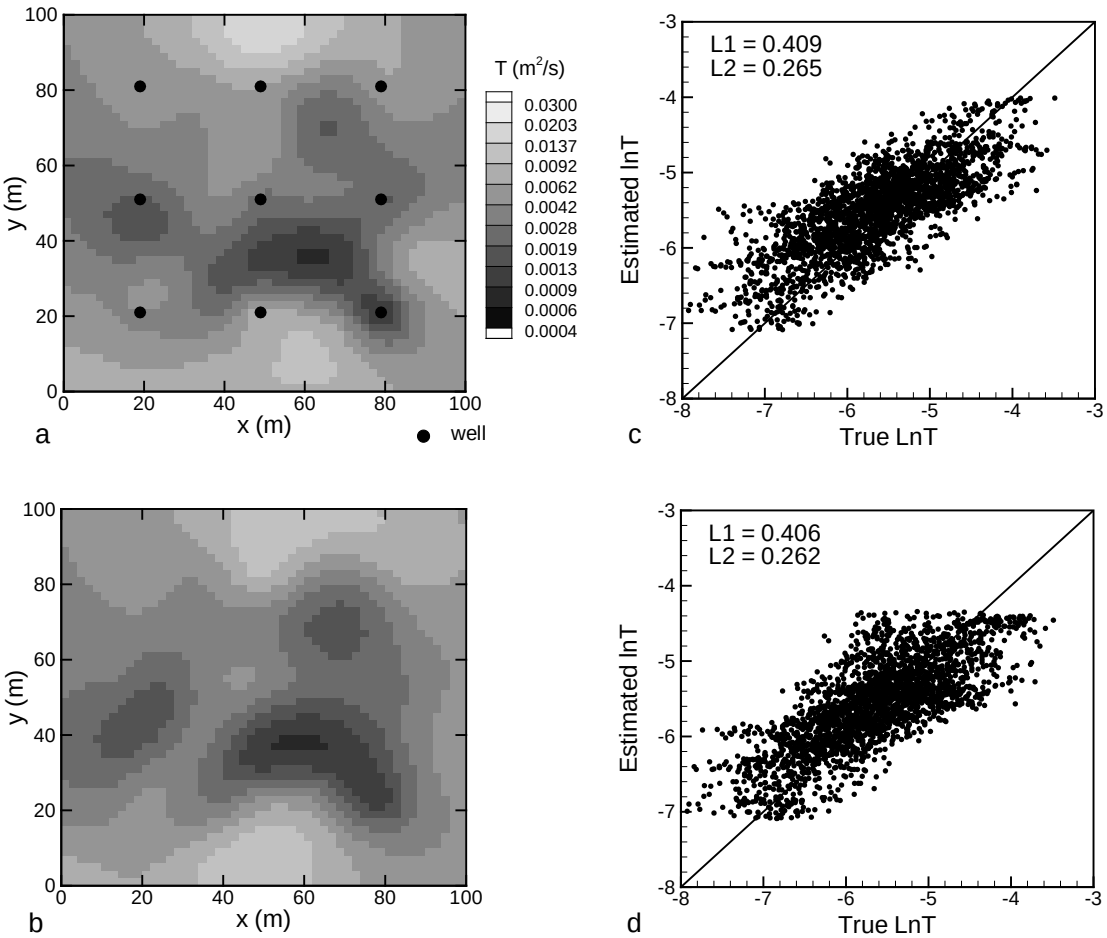


Figure 4.2: Estimated transmissivity fields using a sparse well field. a] Using transient heads [Case 1]; b] using temporal moments [Case 2]; c] the scatterplot of estimated versus true ln*T* fields of Case 1; d] the scatterplot of estimated versus true ln*T* field of Case 2.

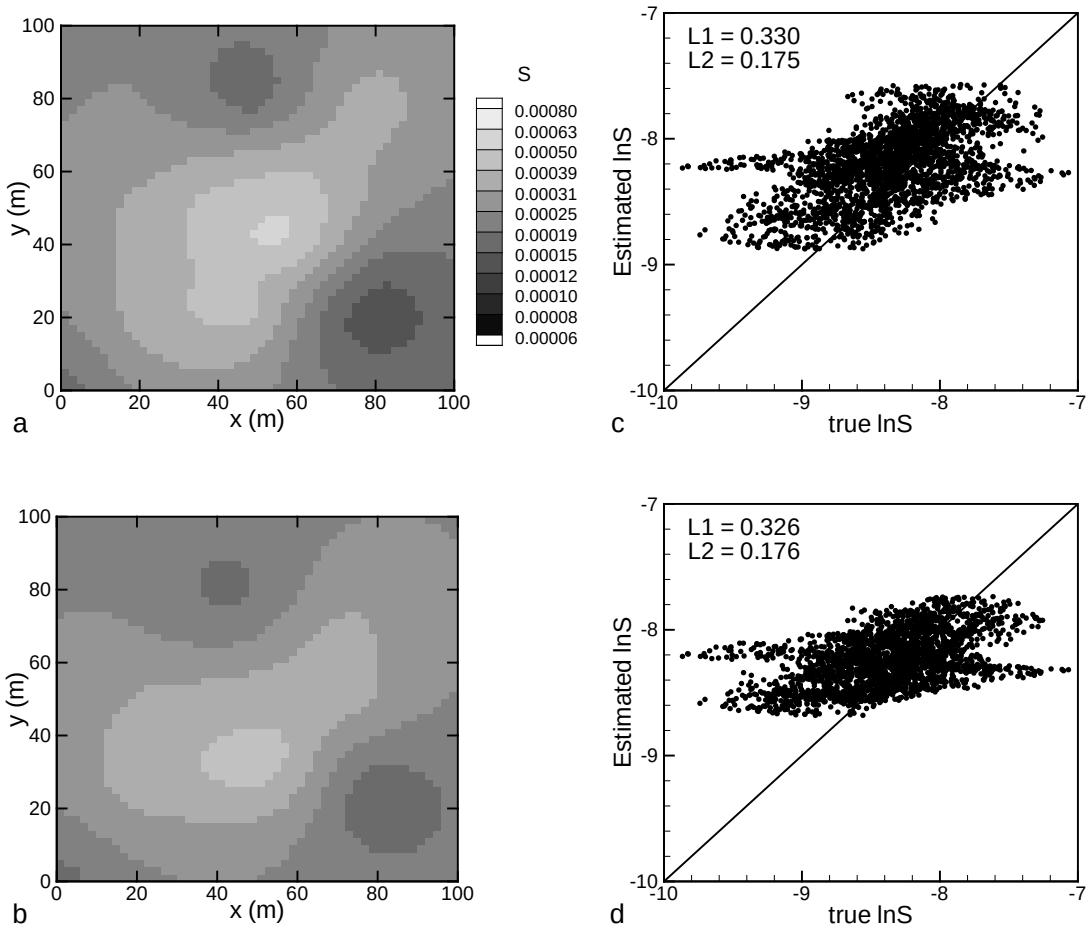


Figure 4.3: Estimated storage coefficient fields using a sparse well field. a] Using transient heads [Case 1]; b] using temporal moments [Case 2]; c] the scatterplot of estimated versus true ln*S* fields of Case 1; d] the scatterplot of estimated versus true ln*S* field of Case 2.

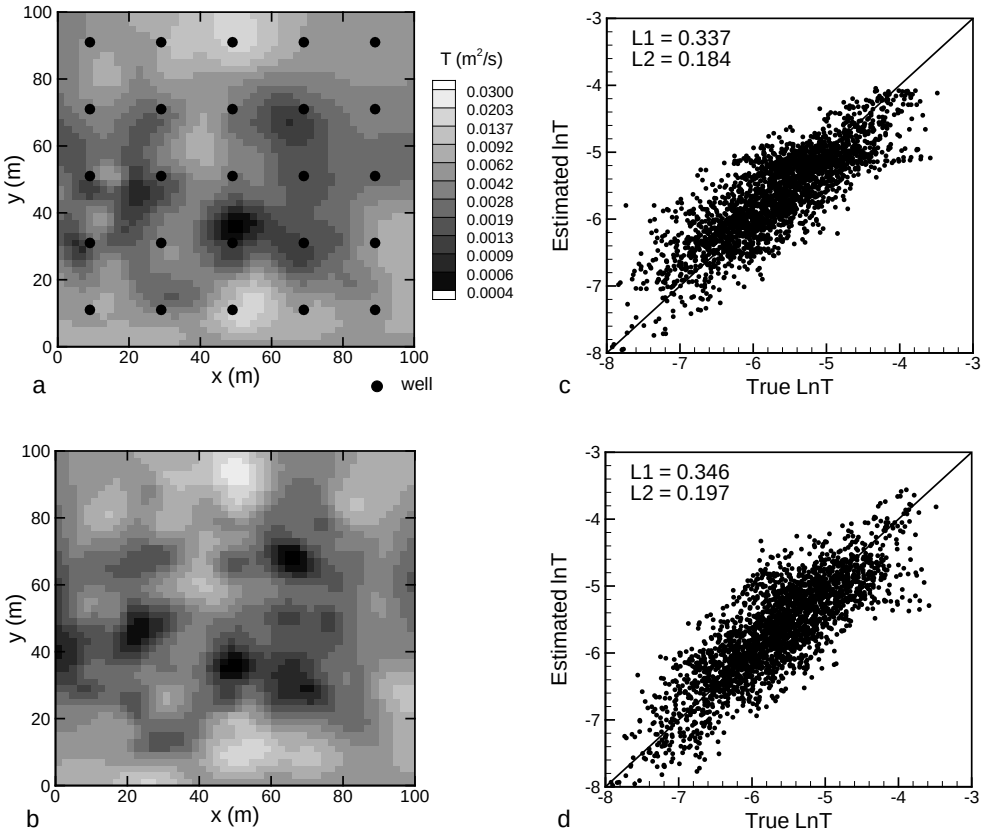


Figure 4.4: Estimated transmissivity fields using a dense well field. a] Using transient heads [Case 3]; b] using temporal moments [Case 4]; c] the scatterplot of estimated versus true ln*T* fields of Case 3; d] the scatterplot of estimated versus true ln*T* field of Case 4.

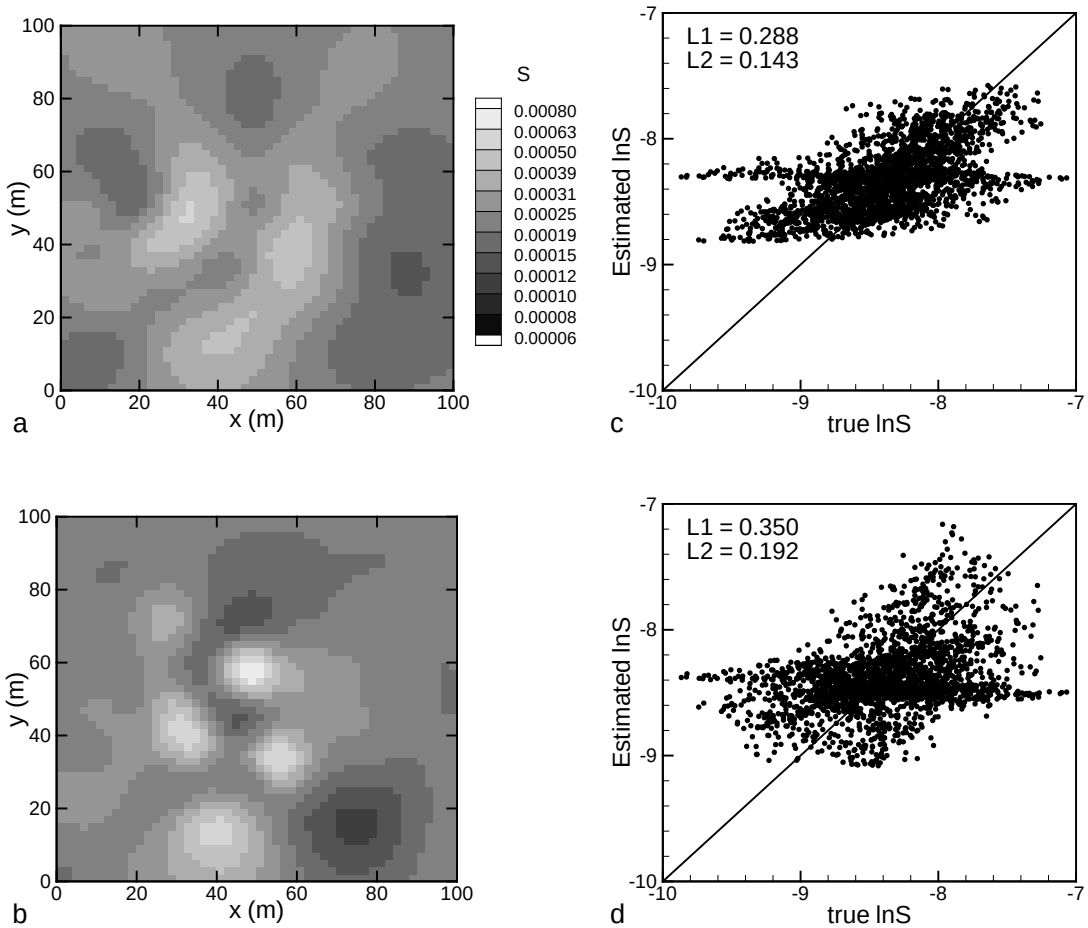


Figure 4.5: Estimated storage coefficient fields using a dense well field. a] Using transient heads [Case 3]; b] using temporal moments [Case 4]; c] the scatterplot of estimated versus true ln*S* fields of Case 3; d] the scatterplot of estimated versus true ln*S* field of Case 4.

4.6 RESULTS AND DISCUSSION

Results of the 2-D synthetic case shows that Case 2 [the moment approach] uses only 15.2% of the computation time needed for Case 1 [the head approach]; likewise, Case 4 [the moment approach] uses only 23.4% of the time needed for Case 3. These results substantiate our speculation about the computational efficiency of the temporal moments approach. In order to quantitatively compare the estimation errors, we plot scatter-plots of estimates versus true fields for all four cases in Figure 4.2 through Figure 4.5, along with mean absolute error normal L1 and mean square error normal L2, which are defined as

$$L1 = \frac{1}{n}\sum_{i=1}^{n}\left|\chi_i - \hat{\chi}_i\right| \qquad \text{and} \qquad L2 = \frac{1}{n}\sum_{i=1}^{n}\left(\chi_i - \hat{\chi}_i\right)^2 \qquad [4.21]$$

where χ_i and $\hat{\chi}_i$ represent the true and estimates of the log-transformed parameter [either ln*T* or ln*S*], respectively, *i* indicates the element number, and *n* is the total number of elements. According to Figures 4.3 through 4.5, the moment approach based on the first two temporal

moments of drawdown can yield similar estimates of transmissivity and storage coefficient fields as those based on transient heads. We also test biasness of our estimations by calculating mean and variance of estimation error, which are defined as

$$\mu = \frac{1}{n}\sum_{i=1}^{n}(\hat{\chi}_i - \chi_i) \quad \text{and} \quad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left((\hat{\chi}_i - \chi_i) - \mu\right)^2 \qquad [4.22]$$

where μ and σ^2 are mean and variance of difference between true and estimated parameter [either ln$T$ or ln$S$]. The mean and variance of estimation errors are listed in Table 4.1. They show that SSLE yields unbiased estimations in all cases. As expected, a dense monitoring network can produce much better results than a sparse network. It is clear however that the results based on the temporal moments are not identical to those based on the transient heads.

The use of the dense and sparse monitoring networks in the above synthetic cases is aimed at investigating the impact of information loss in the moments approach on the inverse results. Transient well hydrographs intuitively bear signatures of the heterogeneity encountered by the drawdown as the cone of depression evolves. These signatures, however, are likely weak because the head recorded at an observation well is highly correlated with different heterogeneities within the cone of depression [Wu et al., 2005]. Because of the integrative nature of the zeroth and first moments of a hydrograph, these moments likely lose these signatures. We therefore anticipate that the interpretation of a hydraulic tomography survey using the temporal moments will yield fewer details about the heterogeneity. Plots of estimated T and S fields and their associated scatter plots in Figures 4.3 and 4.4, and statistics of the estimation errors in Table 4.1 seem to support our hypothesis. The variances of the estimation errors are consistently smaller for the estimated fields based on drawdown than those based on temporal moments. On the other hand, in Case 3 and Case 4 where the monitoring networks are dense, the differences between the estimated fields based on the two approaches are small. In other words, the information loss in the temporal moments is largely compensated for by the dense spatial information.

The storage coefficient estimates from both transient head and temporal moment approaches in both cases are clearly less satisfactory than transmissivity estimates. This may be attributed to the fact that high cross-correlation between head and storage coefficient is limited to a narrow region in between pumping and observation wells; the cross-correlation between head and transmissivity on the other hand is strong over the entire cone of depression at late time [Wu et al., 2005].

Notice that in all cases examined here, the temporal moments at observation wells were directly obtained from moment equations [4.10] and [4.12]. Such a direct evaluation has omitted possible errors in estimating the moments from well hydrographs. Effects of errors in the moments on the estimates of transmissivity and storage coefficient have been studied by Li et al. [2005] during a single aquifer test. According to their study, the effects are not significant. Obviously, rigorous analysis of impacts of errors in moments on hydraulic tomography is necessary but it is beyond the scope of this report.

5. A NEW ESTIMATOR AND A GUIDE FOR HYDRAULIC TOMOGRAPHY ANALYSIS

5.1 INTRODUCTION

Classical aquifer test which involves one pumping well and an observation well has been shown to yield ambiguously averaged hydraulic properties of an aquifer, which vary with the location of observation and pumping wells, and heterogeneity [Wu et al., 2005, Liu et al., 2007, Straface et al., 2007, and Kuhlman et al., 2008]. To avoid obtaining the ambiguously averaged estimates and to provide high-resolution aquifer characterization, a new aquifer characterization technology, known as hydraulic tomography [HT], has recently been developed [e.g., Tosaka et al., 1993; Gottlieb and Dietrich, 1995; Vasco et al., 2000; Yeh and Liu, 2000; Bohling et al., 2002; Brauchler et al., 2003; Zhu and Yeh 2005, 2006; and others]. Although the ability of HT remains to be fully assessed under field conditions, results from sandbox experiments by Liu et al. [2002], Illman et al. [2007 and 2008], and Liu et al. [2007] are encouraging. These studies showed that the transient HT can identify not only the pattern of the heterogeneous hydraulic conductivity [K] field, but also the variation of specific storage [S_s] in the sandbox. Moreover, these estimated K and S_s fields from the HT sandbox experiments further predicted the drawdown evolution caused by a pumping test that was not used in the HT analysis. Likewise, a recent application of HT to a well field at Montalto Uffugo Scalo, Italy, produced an estimated transmissivity field that is deemed to be consistent with the geology of the site [Straface et al., 2007]. Bohling et al. [2007] and Li et al. [2007] also showed promising results of HT in their field experiments.

Most of HT analyses in the past have used the sequential successive linear estimator [SSLE] of Yeh and Liu [2000] or Zhu and Yeh [2005], which includes data sets from HT surveys sequentially. Illman et al. [2008] reported that the order of test data included in SSLE affected the final estimates. In addition, little efforts in the past have focused on investigation of effects of noise in well hydrographs on the analysis of HT and development of methods for removing the noise. In this paper, we thereby develop a geostatistically based method for identifying the subsurface heterogeneity pattern using all the data collected from a HT survey simultaneously, similar to the approaches by Vesselinov et al. [2001], or Li et al. [2007], Fienen et al. [2008], Li et al. [2008], which are built upon the quasi-linear geostastical approach [Kitanidis, 1995]. Moreover, a criterion for the nonlinear estimator is developed to determine the appropriate level of improvement of estimation when hydrographs are infested with noise. We also propose a procedure for preprocessing HT data before application of the estimator. This estimator and procedure were tested in a synthetic aquifer with hierarchical heterogeneity [Barrash and Clemo, 2002, and Ye et al., 2005], and then applied to a sandbox experiment, which involves unknown measurement errors and where our mathematical model may not correctly describe the flow process—model errors.

5.2 METHODOLOGY

5.2.1 Governing Groundwater Flow Equations

A strategy for modeling ground water flow in saturated, heterogeneous, porous media with incomplete specification of aquifer characteristics is the stochastic conditional mean

approach [e.g., Yeh et al., 1996]. That is, the flow process is governed by the following partial differential equation involving conditional means:

$$\nabla \cdot [K(\mathbf{x})\nabla h] + Q(\mathbf{x}_p) = S_s(\mathbf{x})\frac{\partial h}{\partial t} \qquad [5.1]$$

subject to boundary and initial conditions:

$$h\Big|_{\Gamma_1} = h_1, \;\; [K(\mathbf{x})\nabla h]\cdot \mathbf{n}\Big|_{\Gamma_2} = q, \qquad \text{and } h\Big|_{t=0} = h_0 \qquad [5.2]$$

where in equation [5.1], h is conditional effective total head [L], $\mathbf{x}$ is the spatial coordinate [$\mathbf{x}$ = {x, y}, [L], and y represents the vertical coordinate and is positive upward] in the two-dimensional, cross-sectional aquifers examined in this paper, $Q[\mathbf{x}_p]$ is the pumping rate [1/T] at the location $\mathbf{x}_p$, $K[\mathbf{x}]$ is the conditional effective saturated hydraulic conductivity [L/T], and $S_s[\mathbf{x}]$ is the conditional effective specific storage [1/L]. In equation [5.2], h_1 is the prescribed total head at Dirichlet boundary Γ_1, q is the specific discharge [L/T] at Neumann boundary Γ_2, $\mathbf{n}$ is a unit vector normal to Γ_2, and h_0 represents the initial total head. The definitions of variables in the conditional mean equations are identical to those in a deterministic groundwater flow equation if all the parameters, boundary and initial conditions are fully specified [Yeh et al., 1996]. In this chapter, these governing equations are used to simulate the flow field during the HT survey and are solved by a 2-D finite element model [VSAFT2] developed by Yeh et al. [1993].

5.2.2 Simultaneous Successive Linear Estimator

Instead of incorporating data sequentially into the estimation as is done in SSLE, a simultaneous successive linear stochastic estimator [SimSLE] is developed to include all drawdown data from different pumping tests during a HT survey simultaneously to estimate hydraulic properties of aquifers. Simultaneous inclusion of the data offers some advantages over the SSLE approach [see discussion section]. Below is a brief description of the SimSLE.

With given unconditional mean and spatial covariance functions of the hydraulic properties [prior joint probability distribution, implicitly Gaussian], the SimSLE starts with cokriging [a stochastic linear estimator] to estimate the conditional expected value of the property conditioned on $f^*(\mathbf{x}_i)$ [i.e., perturbations of log hydraulic property, K and/or S_s] measured at i^{th} location [i =1,… N_f, where N_f is the total number of f measurements] and the observed head at location $\mathbf{x}_j$ at time t_ℓ during k^{th} pumping test, denoted by $h^*(k,\mathbf{x}_j,t_\ell)$. The linear estimator is,

$$\hat{f}^{(1)}(\mathbf{x}_0) = \sum_{i=1}^{N_f}\lambda_{0i} f^*(\mathbf{x}_i) + \sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)} \mu_{0kj\ell}\left(h^*(k,\mathbf{x}_j,t_\ell) - h_e(k,\mathbf{x}_j,t_\ell)\right) \qquad [5.3]$$

where $\hat{f}^{(1)}(\mathbf{x}_0)$ is the cokriged f value at location $\mathbf{x}_0$; $h_e(k,\mathbf{x}_j,t_\ell)$ is the simulated head at the observation location and time of the pumping test, based on effective properties of an equivalent homogeneous medium; N_p is the total number of pumping tests; $N_h(k)$ is the total number of observation locations for k^{th} pumping test; $N_t(k,j)$ is the total number of head measurements in time at j^{th} observation location during k^{th} pumping test. The cokriging weight [λ_{0i}] represents contribution of measurement f^* at i^{th} location to the estimate at location, $\mathbf{x}_0$. The contribution to the estimate from the observed head $h^*(k,\mathbf{x}_j,t_\ell)$ is denoted by $\mu_{0kj\ell}$. These weights are obtained by solving the following system of equations:

$$\sum_{i=1}^{N_f}\lambda_{0i}R_{ff}(\mathbf{x}_m,\mathbf{x}_i)+\sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)}\mu_{0kj\ell}R_{hf}\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_m\right)=R_{ff}(\mathbf{x}_0,\mathbf{x}_m)$$

$$\sum_{i=1}^{N_f}\lambda_{0i}R_{hf}\left((p,\mathbf{x}_r,t_q),\mathbf{x}_i\right)+\sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)}\mu_{0kj\ell}R_{hh}\left((p,\mathbf{x}_r,t_q),(k,\mathbf{x}_j,t_\ell)\right)=R_{hf}\left((p,\mathbf{x}_r,t_q),\mathbf{x}_0\right) \quad [5.4]$$

in which $m=1,...N_f$; $p=1,...N_p$, $r=1,...N_h(p)$ and ℓ and $q=1,...N_t(p,r)$. Our prior knowledge of the spatial structure [the unconditional covariance function] of *f* is given by R_{ff}. R_{hh} and R_{hf} are the unconditional covariance of h and the unconditional cross covariance of *f* and h, respectively, which are determined by a first order analysis with the given R_{ff}. That is,

$$R_{hf}\left((k,\mathbf{x}_i,t_\ell),\mathbf{x}_m\right)=\sum_{j=1}^{N_e}J\left((k,\mathbf{x}_i,t_\ell),\mathbf{x}_j\right)R_{ff}(\mathbf{x}_j,\mathbf{x}_m)$$

$$k=1,...N_p;i=1,...N_h(k);\ \ell=1,...N_t(k,i);\ m=1,...N_e. \quad [5.5]$$

$$R_{hh}\left((u,\mathbf{x}_i,t_q),(k,\mathbf{x}_j,t_\ell)\right)=\sum_{m=1}^{N_e}R_{hf}\left((u,\mathbf{x}_i,t_q),\mathbf{x}_m\right)J\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_m\right)^T \quad [5.6]$$

$$u\ and\ k=1,...N_p;\ i\ and\ \ j=1,...N_h(u\ or\ k);\ q\ and\ \ \ell=1,...N_t(u\ or\ k,i);$$

where $J\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_m\right)$ is the sensitivity of head at j^{th} location at time t_ℓ for k^{th} pumping test with respect to the change of parameter at m^{th} location; N_e is the number of elements in the study domain. The sensitivity matrix is evaluated using an adjoint state approach [see Zhu and Yeh, 2005 for details].

After obtaining the new estimate for all the elements using cokriging, the conditional covariance of f, $\varepsilon_{ff}^{(1)}(\mathbf{x}_m,\mathbf{x}_q)$, is then determined by

$$\varepsilon_{ff}^{(1)}(\mathbf{x}_m,\mathbf{x}_q)=R_{ff}(\mathbf{x}_m,\mathbf{x}_q)-\sum_{k=1}^{N_f}\lambda_{mk}R_{ff}(\mathbf{x}_k,\mathbf{x}_q)-\sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)}\mu_{mkj\ell}R_{hf}\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_q\right) \quad [5.7]$$

where *m and q*=1, … N_e. The conditional covariance reflects the effect of data on the reduction of uncertainty in the estimated parameter field. Subsequently, the estimated log perturbations of the property fields are added to the log of the effective properties, F[**x**], then converted to the arithmetic scale, and used to solve equation [5.1] with boundary and initial conditions for the conditional effective head fields, $h^{(1)}(k,\mathbf{x}_j,t_\ell)$, of each pumping test.

Following cokriging, a linear estimator of the following form,

$$\hat{f}^{(r+1)}(\mathbf{x}_0)=\hat{f}^{(r)}(\mathbf{x}_0)+\sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)}\omega_{0kj\ell}^{(r)}\left[h^*(k,\mathbf{x}_j,t_\ell)-h^{(r)}(k,\mathbf{x}_j,t_\ell)\right] \quad [5.8]$$

is used to improve the estimate for iteration *r* >1, where $\omega_{0kj\ell}^{(r)}$ is the weight term, representing the contribution of the difference between the observed and simulated conditional heads [i.e., $h^*(k,\mathbf{x}_j,t_\ell)$ and $h^{(r)}(k,\mathbf{x}_j,t_\ell)$, respectively] at iteration *r* at location $\mathbf{x}_j$ of the kth pumping test at time t_ℓ to the estimate at location $\mathbf{x}_0$. The weights are determined by solving the following system of equations:

$$\sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)}\omega_{0kj\ell}^{(r)}\left[\varepsilon_{hh}^{(r)}\left((p,\mathbf{x}_m,t_q),(k,\mathbf{x}_j,t_\ell)\right)+\Theta^{(r)}\delta_{kj\ell}\right]=\varepsilon_{hf}^{(r)}\left((p,\mathbf{x}_m,t_q),\mathbf{x}_0\right) \qquad [5.9]$$

where $p=1,...N_p$, $m=1,...N_h(p)$ and $q=1,...N_t(p,m)$. The terms $\varepsilon_{hh}^{(r)}$ and $\varepsilon_{fh}^{(r)}$ are the conditional covariance and the conditional cross covariance at iteration [r], which are evaluated using equations [5.5] and [5.6] using the conditional covariance of *f* [i.e., $\varepsilon_{ff}^{(r)}$ which is obtained from equation [7] for the first iteration]. A dynamic stabilizer, $\Theta^{(r)}$, is added to the diagonal elements of $\varepsilon_{hh}^{(r)}$ [δ_{kjl} is a Dirac delta, equal to 1 when *k=j=l* and zero otherwise] to stabilize the solution to equation [5.9]. The dynamic stabilizer at iteration, *r*, is the maximum value of the diagonal elements of $\varepsilon_{hh}^{(r)}$ at that iteration times a user-specified multiplier [see Yeh et al., 1996]. After completion of the estimation using equation [5.8] for all elements in the domain, the conditional covariance of *f* is updated subsequently as given below.

$$\varepsilon_{ff}^{(r+1)}(\mathbf{x}_m,\mathbf{x}_n)=\varepsilon_{ff}^{(r)}(\mathbf{x}_m,\mathbf{x}_n)-\sum_{k=1}^{N_p}\sum_{j=1}^{N_h(k)}\sum_{\ell=1}^{N_t(k,j)}\omega_{mkj\ell}\varepsilon_{hf}^{(r)}\left((k,\mathbf{x}_j,t_\ell),\mathbf{x}_n\right) \qquad [5.10]$$

where *n* and *m* =1, … N_e.

The iteration steps of SimSLE are the same as those in the SLE algorithm used in Yeh et al. [1996]. For noise-free hydrographs, the convergence is achieved if 1] change in variances that represent spatial variability of the estimated hydraulic properties between current and last iterations is smaller than a specified tolerance [i.e., the spatial variance of the estimates stabilizes], implying that the SimSLE cannot improve the estimation any further; 2] change of simulated heads between successive iterations is smaller than the tolerance, indicating that the estimates will not significantly improve the head field. If one of the two criteria is met, the estimates are considered to be optimal and the iterations are terminated.

Head observations from laboratory or field experiments often contain noise [i.e., signals caused by processes not modeled by the governing flow equation, such as Earth tide and others including measurement errors] in addition to effects of hierarchical heterogeneity. Such unresolved noises can lead to divergence of inverse solutions [i.e., unrealistic estimates]. As a consequence, an important issue that ought to be addressed is to what degree should the observed head be used to improve estimates of the hydraulic properties. Stablization of mean square error of the simulated head during iteration should address the issue. Consider the mean square error of the head:

$$\begin{aligned}E\left[\left(h^*(k,\mathbf{x},t)-\hat{h}^{(r)}(k,\mathbf{x},t)\right)^2\right]&=E\left[\left((h\ (k,\mathbf{x},t)+\tau)-\hat{h}^{(r)}(k,\mathbf{x},t)\right)^2\right]\\&=E\left[\left(h\ (k,\mathbf{x},t)-\hat{h}^{(r)}(k,\mathbf{x},t)\right)^2\right]+E\left[\tau^2\right] \qquad [5.11]\\&=\varepsilon_{hh}^{(r)}(k,\mathbf{x},t)+\sigma_\tau^2\end{aligned}$$

in which the observed head at location **x** and time *t* during k^{th} pumping event is denoted by $h^*(k,\mathbf{x},t)$; the corresponding simulated head based on the estimated parameters at r^{th} iteration is given by $\hat{h}^{(r)}(k,\mathbf{x},t)$. Equation [5.11] assumes that the observed head consists of the noise free head, $h(k,\mathbf{x},t)$, and random noise, τ, with variance σ_τ^2. The term $\varepsilon_{hh}^{(r)}(k,\mathbf{x},t)$ denotes the theoretical residual head variance for the noise free case at r^{th} iteration [i.e., the diagonal term in equation [5.6]], which should decrease and approach zero with iterations due to improvement of

the parameter estimates. Therefore, the mean square error for cases with noise should asymptotically converge to σ_τ^2 as the number of iterations increases. More importantly, equation [5.11] suggests that once $\varepsilon_{hh}^{(r)}(k,\mathbf{x},t)$ becomes smaller than σ_τ^2 [i.e., the mean square error stabilizes], the iteration should stop to avoid over-usage of imperfect head data [i.e., updating the estimate with noise]. Consequently, for cases where the observed head is not noise free, we use the stabilization of the L2 norm of the conditional heads to terminate the iteration, i.e.,

$$L2_{cond}(r) = \frac{1}{N}\sum_{i=1}^{N}\left(h_i^* - \hat{h}_i^{(r)}\right)^2 \qquad [5.12]$$

where h_i^* and $\hat{h}_i^{(r)}$ are observed and simulated heads, respectively; i is the index denoting the observation in a given time and location from a pumping test; N is the total number of head observations from all the pumping tests. Hereafter, we will refer equation [5.12] as to the conditional L2 norm.

5.2.3 Preprocessing Data for SimSLE

5.2.3.1 Diagnosis of Bias. The first step in preprocessing HT data for the SimSLE analysis is to qualitatively check for any bias or inconsistency in the data. We suggest this be accomplished by a simple rule-of-thumb approach: plotting hydrographs [drawdown-time data] for each pumping test. For example, arrival time of a given drawdown should increase with distance from the pumping well unless there are physically explainable anomalies [e.g., fracture zones or other fast flow paths]. Plots of evolution of contour surface of the drawdown induced by each pumping test generally should follow behaviors of drawdown in homogeneous aquifers [e.g., Bakr, et al., 1978] although details are different. Data with significant anomalies should be examined for possible operational causes [e.g., mislabeling monitoring ports or wells, leakage between packers, malfunctioning of equipment, or other factors]. If the operational causes can be identified, the data sets should be corrected or excluded. Repeating the HT experiment also facilitates a viable diagnosis. Drawdown-log time plots should also reveal possible boundary effects and they are useful in assigning the type of boundary conditions of the modeling domain.

5.2.3.2 Wavelet Denoising. Next, we tackle data noise or signals or perturbations caused by factors other than aquifer heterogeneity [i., fluctuations in pressure transducer voltage signals, diurnal fluctuations, etc.]. To eliminate the noise, a denoising method based on wavelet analysis [e.g., Mallat, 1999; Zhang et al., 2006] was developed. A wavelet analysis is similar to Fourier analysis in the sense that it breaks a signal down into its constituent parts for analysis. In contrast to the Fourier transform, the wavelet transform allows exceptional localization in both the time domain via translations of the mother wavelet, and the scale [frequency] domain via dilations, when analyzing signals of a non-stationary nature. Software from [http://www.mathworks.com] was used. Different families of wavelets were tested in this study and we found that the Daubechies 4 wavelet is effective for our cases.

The wavelet denoising procedure used in this study comprises of the following steps: 1] Applying the wavelet transform to the noisy hydrograph to produce wavelet coefficients; 2] Selecting an appropriate threshold limit and a threshold method to remove the noise; 3] Inversing the wavelet transform of thresholded wavelet coefficients to generate a denoised hydrograph; 4] Calculating the variance of the difference between the denoised hydrographs and those of the equivalent homogenous medium. This variance should be approximately equal to the theoretical head variance from the first-order analysis. If this criterion is not met, repeat steps 2] and 3]. Generally speaking, distinguishing noise from the effects of heterogeneity in a hydrograph can

be highly subjective unless characteristics of noise or heterogeneity are known a priori.

5.2.4 Signal to Noise Ratio

Following wavelet denoising, we examined the signal to noise ratio [SNR] of each well hydrograph. The SNR is an electrical engineering concept, which is defined as the ratio of signal amplitude to noise amplitude. In our study, it is defined as the ratio between the maximum drawdown of the denoised hydrograph and the standard deviation of noise. That is,

$$SNR = \frac{signal}{noise} = \frac{\left(\left|h_d(\mathbf{x},0) - h_d(\mathbf{x},t)\right|\right)_{\max}}{\Sigma_n} \qquad [5.13]$$

where $|\ |$ denotes the absolute value; h_d denotes the head after denoising; Σ_n represents the standard deviation of noise, which is estimated from the wavelet denoising procedure. In signal processing, signals with SNR of 100 or greater are considered to be good signals. For hydrologic processes, we found that hydrographs with an average SNR of 7.13 or greater are effective for our synthetic case. Figure 1a shows an original hydrograph from Illman et al. [2007 and 2008] and Liu et al. [2007] and the corresponding denoised hydrograph; the SNR is 17.75. The denoised hydrograph therefore is considered useful and it improves the estimates. On the other hand, a hydrograph with low SNR data [4.41] [Figure 5.1b], especially near the falling and rising limbs of the graph [reflecting effects of pumping and recovery, respectively], was discarded since it did not improve the estimates. Low SNR data at early time and recovery can lead to erroneous estimates of the S_s field. As a rule of thumb, the SNR should be much greater than 1 such that the trend of drawdown induced by pumping is evident after denoising. However, a more in-depth analysis is needed to establish a rigorous criterion.

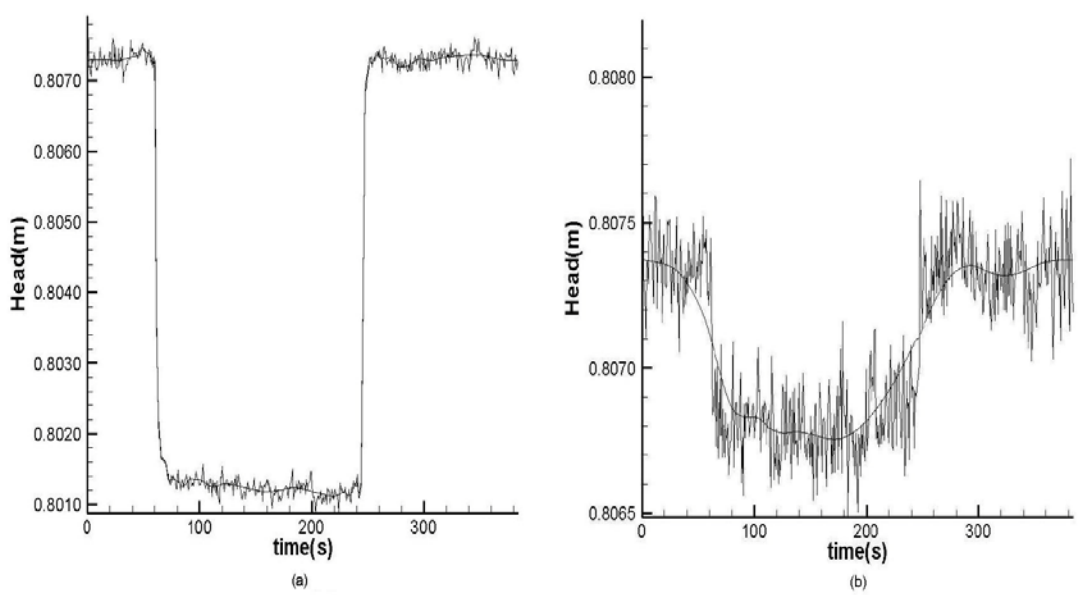


Figure 5.1: Typical hydrographs and corresponding denoised hydrographs [smooth solid line] in the sandbox experiments [Liu et al., 2007] a] SNR: 17.75; b] SNR: 4.41. Notice vertical exaggeration.

5.2.5 Estimation of Effective Properties, and Variances

Afterwards, desirable hydrographs from the HT were selected to estimate unconditional effective K and S_s of an equivalent homogeneous medium with the known pumping rates. This task was accomplished by minimizing the squared difference between the observed head values and those obtained from simulations based on the homogeneity assumption using the VSAFT2:

$$\sum_{i=1}^{n_p}\sum_{j=1}^{n_h}\sum_{k=1}^{n_t}\left(h_{ijk}^{*}-\hat{H}_{ijk}\right)^2=\min_{K,S_s}. \qquad [5.14]$$

In the above equation h_{ijk}^{*} denotes observed head at k^{th} time, j^{th} observation location, and i^{th} pumping test, and the corresponding simulated head using the unconditional effective K and S_s is indicated by $\hat{H}_{ijk}$. Minimization of equation [5.14] was performed using a standard nonlinear least squares [i.e., Gauss-Newton] approach in conjunction with the Levenberg-Marquardt algorithm [Press et. al., 1992]. Sensitivity matrices were evaluated by solving the sensitivity equation [i.e., differentiation of equation [5.1] with respect to the parameter].

Subsequently, square of the difference between the head observed and the head in the equivalent homogeneous medium is used to compute the sample variance of the observed head at the given location and time. This sample head variance is in turn used to estimate the variance of the hydraulic properties by minimizing the following objective function:

$$\sum_{i=1}^{n_p}\sum_{j=1}^{n_h}\sum_{k=1}^{n_t}\left(\sigma_{ijk}^2-\hat{\sigma}_{ijk}^2\right)^2=\text{minimum} \qquad [5.15]$$

where σ_{ijk}^2 is the sample head variance and $\hat{\sigma}_{ijk}^2$ is the theoretical head variance at k^{th} time, j^{th} observation location, and i^{th} pumping test. This theoretical head variance is evaluated using the first-order approximation [i.e., equation [5.6]] with given variance of the hydraulic properties. The Levenberg-Marquardt algorithm was used to seek the variances of the properties that minimize equation [5.15]. Relative change of equation [5.15] between successive iterations is used as convergence criterion. These estimated effective hydraulic properties and variances are used as prior information required by SimSLE for estimating the spatially varying K and S_s fields.

5.3 APPLICATIONS TO A SYNTHETIC AQUIFER

5.3.1 Description of the Aquifer and HT Test

To test our SimSLE and the data preprocessing procedure for the HT analysis, a two-dimensional, cross-sectional, synthetic aquifer of the same length and height as the sandbox experiment conducted by Illman et al. [2007 and 2008] and Liu et al. [2007] was used. The dimensions of the sandbox were 193.0 cm in length, 82.6 cm in height, and 10.2 cm in depth. Twenty four locations were selected [solid circles in Figure 5.2a] to serve as pressure monitoring ports during four pumping tests at four locations [open circles in Figure 5.2a]. Both sides of the aquifer were set to the same constant boundary condition with a total head of 200 cm, while the bottom and top boundaries were set to be no-flux boundaries. The initial total head was assigned to be 200 cm. This aquifer was discretized into 741 elements and 800 nodes with element dimensions of 4.10 cm × 4.13 cm.

The synthetic aquifer was created to imitate a geologic formation of hierarchical heterogeneity, consisting of four units with a bedding dip angle of 20 degrees. The lnK and lnS_s fields within each unit were assumed to be normally distributed random fields with different

means and variances [see Table 5.1]. An exponential model was used as the spatial covariance functions of the fields with correlation scales of 200 cm in the bedding direction and 12 cm in the direction perpendicular to bedding. Different random seed numbers were used to create the ln$K$ and lnS_s fields [Figures 5.2a and 5.2e, respectively] so that they are independent of each other.

Four pumping tests in this aquifer were simulated using a time step of 0.25 seconds for a period of 15 seconds at which the drawdown reached a steady state condition due to the small size of the aquifer. Each pumping test had a pumping rate of 0.3 cm^3/s and drawdown at all 24 wells were recorded. These simulated drawdown-time data sets were regarded as noise-free hydrographs. These noise-free hydrographs were subsequently corrupted by adding normally distributed white noise with a standard deviation of 0.07 cm to represent measurement errors and these corrupted hydrographs are denoted as noisy hydrographs. Finally, the wavelet denoising procedure was applied to these noisy hydrographs to obtain the so-called denoised hydrographs. The SNR values of the 96 corrupted hydrographs were found much greater than 1 and all the hydrographs were included in the following HT analysis.

To investigate the ability of SimSLE to estimate the heterogeneous K and S_s fields, drawdowns at 5 times [four early times at 0.5s, 1.75s, 2.25s and 3s; and one later time at 15s] from the noise-free, noisy, and denoised HT hydrographs were used. Such a choice of sampling times stems from the finding by Wu et al. [2005] that the drawdown at early time is highly correlated with the S_s field and only weakly and negatively correlated with the K field in the area between the pumping and the observation locations. At large time, the drawdown is correlated at various degrees with K values in the area within the cone of depression but not the S_s field. In the estimation, K and S_s values at the top observation port in the left column in Figure 5.2a were assumed to be known and they were used as the hard data for conditioning the estimation in all the cases.

Table 5.1: Means and variances of the generated *lnK* and *lnSs* field in each unit of the synthetic aquifer and over the entire aquifer.

		Unit 1	Unit 2	Unit 3	Unit 4	Overall
lnK	Mean	2.03	0.85	2.46	-0.67	1.39
	Variance	1.01	0.17	2.00	0.46	2.18
lnSs	Mean	-6.06	-4.82	-6.43	-5.06	-5.61
	Variance	0.05	0.24	0.01	0.08	0.66

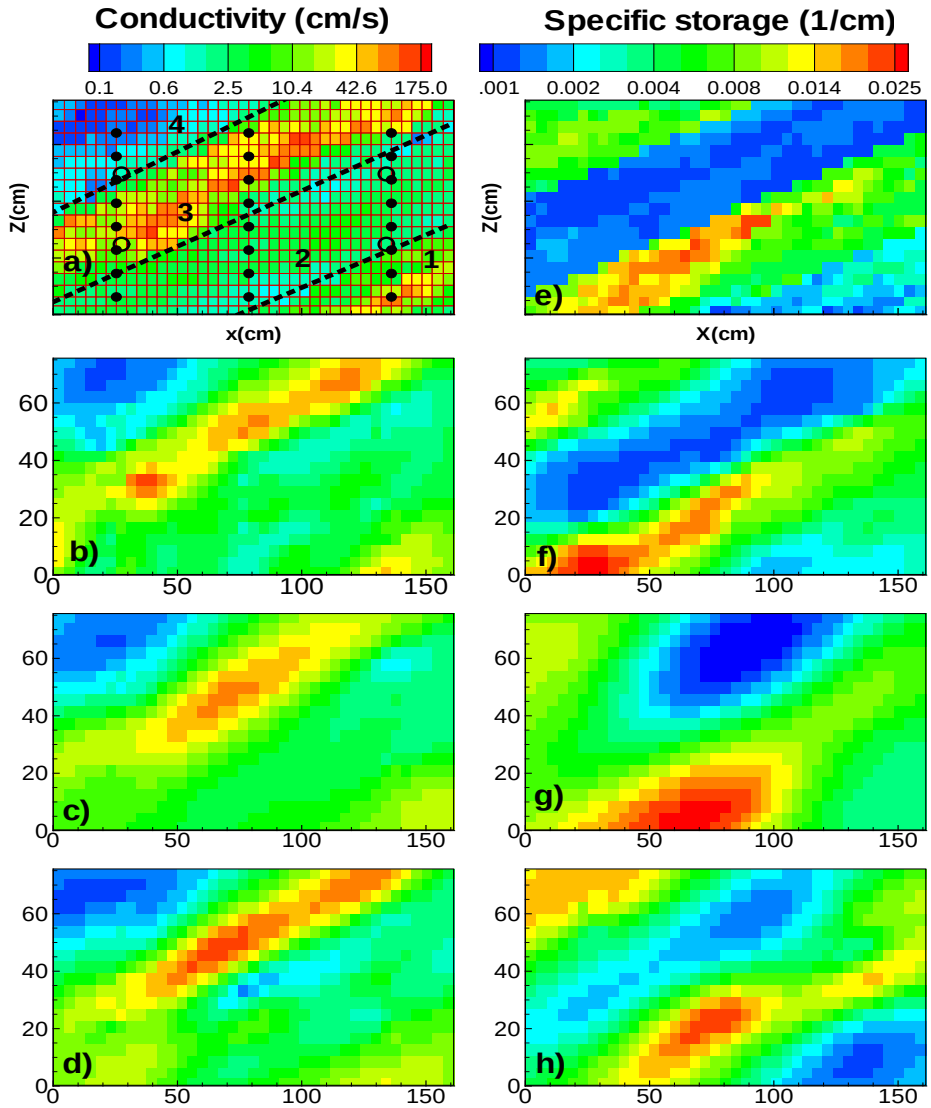


Figure 5.2: a] and e] are the true K [cm/s] and S_s [1/cm] fields of the synthetic aquifer. The dashed lines are the boundaries between units [1, 2, 3, and 4]; estimated K and S_s fields using noise free data [b and f]; using noisy data [c and g]; using denoised data [d and h].

5.3.2 Performance Assessment

Performance of our estimator for the synthetic aquifer case was evaluated using the standard correlation coefficient $[1 \geq |r| \geq 0]$ which measures the similarity between the true and the estimated fields. A correlation coefficient close to 1 means the two fields are similar in pattern, even though the mean value of the two fields may be different. Thereby, in addition to the correlation, mean absolute error [L1 norm] and mean square error [L2 norm] of the estimated field were used to judge the performance.

Besides the correlation analysis, L1, and L2 norms, similarity between the true and estimated hydraulic property fields was also determined using a fuzzy similarity comparison method, which has been applied to the task of comparing spatial patterns [Hagen, 2003]. This method allows a user to specify weights for both location and value matching. Specifically, the

method computes similarity between the true property at a given element and the estimated property at the corresponding element and neighboring elements. The similarity is based on a locational and an error membership value, which are defined below.

An exponential function based on the statistical spatial model for the hydraulic properties was used to define a locational membership

$$V_l = \exp\left(-\sqrt{\left(\frac{l_x}{\lambda_x}\right)^2 + \left(\frac{l_y}{\lambda_y}\right)^2} \right) \qquad [5.16]$$

In which l_x and l_y are the separation distance between two compared elements in the x and y directions; λ_x and λ_y are the correlation scale of the true property field in the x and y directions, respectively. Equation [5.16] implies that the similarity decreases with an increase of the ratio of the separation distance to the correlation scale.

The standard deviation of the true property [STD] is used as an error limit and to define the value of the error membership:

$$V_v = \begin{cases} 1.0 - \dfrac{|v|}{STD} & \dfrac{|v|}{STD} < 1 \\ 0 & \dfrac{|v|}{STD} \geq 1 \end{cases} \qquad [5.17]$$

where v is the difference between the true property at a given element and the estimated property at a selected element. That is, if the difference of two properties is greater than the standard deviation, the two properties are judged to be completely different. Otherwise, a linear decay is used to determine the error membership value.

The true property of an element in the domain is then compared with the estimated property of the same element and with the estimated property of neighboring elements [within the correlation scales]. Equations [5.16] and [5.17] are employed to determine the locational and error membership values. For each pair of the true and the estimated property, a similarity value is then computed by multiplying the locational membership with the error membership. Only the maximum value among the similarity values of all pairs is retained for the selected element. Next, the maximum similarity between the estimated property at that element with the true property at the corresponding element and at neighboring elements is determined. The similarity value between the true and the estimated properties of this element is then defined by the average of the two maximum similarities. This procedure is applied to all the elements in the synthetic domain. Finally, a domain similarity is defined as the average element similarity of all elements in the domain.

5.3.3 Results

Table 5.2 tabulates the estimated effective K and S_s values for an equivalent homogeneous and isotropic porous medium of the synthetic aquifer, using the noise-free, noisy, and denoised hydrographs. Estimated variances of ln$K$ and lnS_s over the entire aquifer are also listed in the table, which were obtained using equation [5.15] with visually estimated correlation scales [100 cm and 33 cm, in the direction parallel and perpendicular to bedding, respectively] and the known dip angle. The table shows that the estimated effective properties based on the noise-free, noisy, and denoised hydrographs are similar. Estimated effective K values are found to be slightly greater than the geometric mean while the effective S_s values are in agreement with

the arithmetic mean of the corresponding values of the four units. Apparently, noise does not have significant effects on the estimates of the effective properties since this is an over-determined inverse problem [i.e., more data than parameters need to be estimated and the problem is well-posed, Yeh et al., 2007]. The least squares approach is expected to be sufficient and effective.

Table 5.2: Estimated effective hydraulic properties, and estimated spatial variances of the properties in the synthetic aquifer, using noise-free, noisy, and denoised hydrographs. The number of iterations required to reach to the convergence of the solution is listed in the last column.

Cases	Effective lnK [cm/s]	Variance lnK	Effective $\ln S_s$ [1/cm]	Variance $\ln S_s$	Iterations
Synth [noise free]	1.422	0.679	-5.114	0.442	16
Synth Noise	1.414	0.520	-5.181	0.154	17
Synth Denoised	1.431	0.571	-5.084	0.329	23

As shown in Table 5.2, estimated variances of ln$K$ and lnS_s with equation [5.15] are smaller than the true ones, likely due to insufficient head data. The nonstationary nature of the drawdown [i.e., its variance depends on the spatially and temporally varying mean gradient] demands a large number of head data at the same radial distance from the pumping wells to obtain a representative sample head variance in equation [5.15].

To illustrate the effectiveness of wavelet denoising, 480 pairs of observed heads before and after denoising at the 24 ports at 5 sampling times during the 4 pumping tests and corresponding simulated heads based on effective K and S_s are plotted as red circles in Figures 3a and b. respectively. Plus symbols in the figure correspond to the simulated heads plus and minus one standard deviation of the theoretical head perturbation caused by heterogeneity, $\varepsilon_{hh}^{(0)}(k,\mathbf{x},t)$. A comparison of Figures 5.3a and 5.3b shows that the wavelet denoising removed most of perturbations of the total heads between 199.5 cm and 200 cm, which are not significantly affected by the pumping test. Head data after denoising are generally within one standard deviation of the theoretical head perturbation caused by heterogeneity, suggesting the remaining perturbations are likely effects of aquifer heterogeneity.

Table 5.3 lists values of unconditional L1 and L2 norms of the hydraulic head, which are defined as

$$L1_{un} = \frac{1}{N}\sum_{i=1}^{N}\left|h_i - \hat{h}_i\right| \text{ and } L2_{un} = \frac{1}{N}\sum_{i=1}^{N}\left|h_i - \hat{h}_i\right|^2 \qquad [5.18]$$

where $\hat{h}$ is the simulated head in the equivalent homogeneous aquifer using the effective parameter values derived from the noise-free hydrographs; h is the noise-free, noisy, or denoised heads of the heterogeneous aquifer, at the 5 selected times during the four pumping tests [i.e., N=480]. These statistics aim to measure effects of the hierarchical heterogeneity [i.e., spatial variability or structured noise], measurement noise, and noise residuals in hydrographs after wavelet denoising. As expected, the noise-free hydrograph has the smallest unconditional L1

and L2 norms, reflecting the effect of heterogeneity only. L1 and L2 norms are highest for the noisy hydrographs where head perturbations are results of both heterogeneity and the white noise. Values of L1 and L2 norms in the last row of Table 5.3 for the denoised hydrographs are between those of noise-free and noisy hydrographs, indicative of only partial removal of noise by the wavelet denoising procedure. Based on the L2 values, 92% of the perturbations in the noisy hydrographs are effects of hierarchical heterogeneity and the remaining 8% are random noise. The wavelet denoising procedure removed 80% of the random noise.

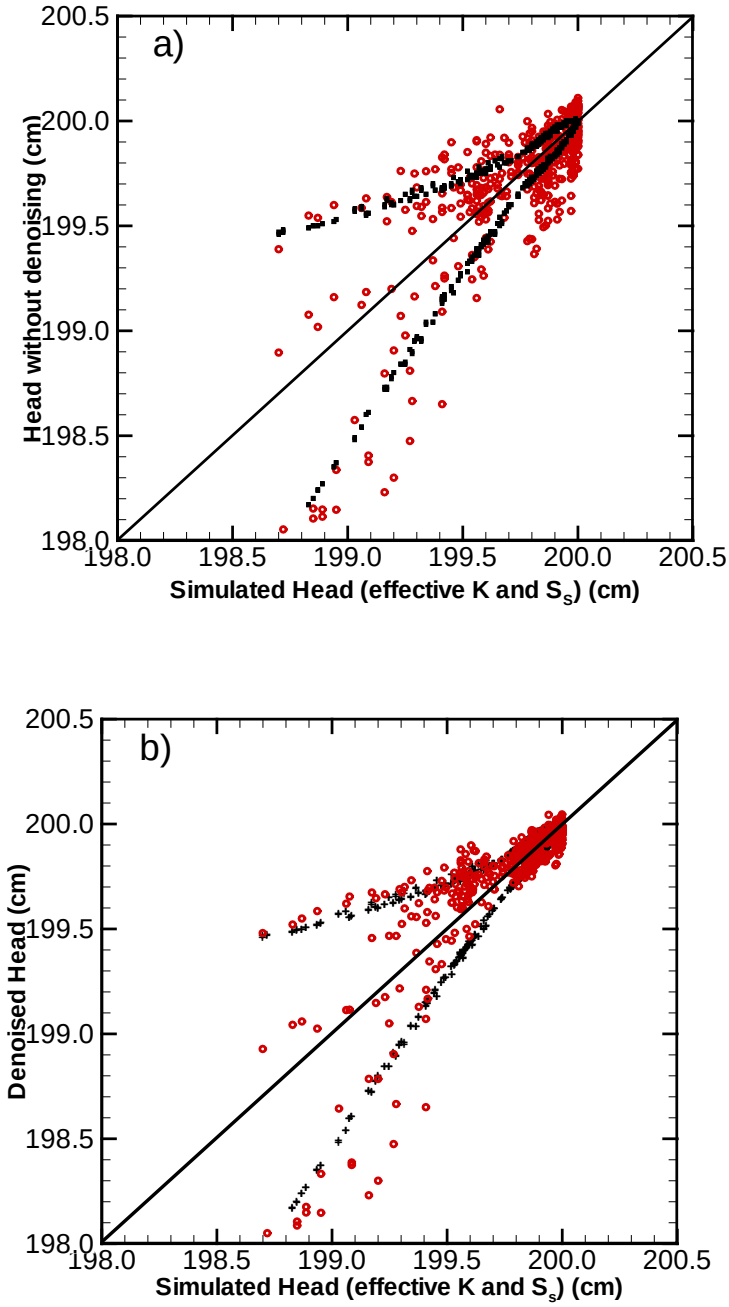


Figure 5.3: Effectiveness of wavelet denoising. A plot of 480 pairs of heads a] before b] after denoising and simulated heads [red circles] based on effective K and S_s of the synthetic aquifer;

Symbols + denote the simulated heads plus and minus one standard deviation of head variation induced by heterogeneity.

Table 5.3: Unconditional L1 and L2 norms of the head based on equation [5.18] for the noise-free, noisy, and denoised cases of the synthetic aquifer experiments.

	L1	L2
Noise-free	0.102	0.0374
noisy	0.124	0.0405
denoised	0.109	0.0380

Estimating 741 pairs of K_s and S_s of the synthetic aquifer based on the 480 drawdowns and one pair of K_s and S_s hard data set is an under-determined [i.e., over-parameterized or ill-posed, see Yeh et al. 2007] inverse problem. That is, the number of the parameters to be estimated is larger than the number of data sets available. While SLE aims to seek the estimate of the conditional effective parameters [Yeh et al., 1996] for under-determined problems, the noise or unresolved noise may lead to anomalously high or low estimates. To circumvent this problem, the criterion based on the stabilization of the conditional L2 norm of the head [i.e., equation [5.12]] was employed.

The conditional L2 norms for noise free, noisy and denoised cases are shown in Figure 5.4 as a function of the number of iterations, while corresponding behaviors of the $\ln K$ and $\ln S_s$ estimates in terms of their spatial variances are illustrated in Figures 5.5a and 5.5b, respectively. As expected, in the noise-free case, the L2 decreased from $0.0374\,cm^2$ [the unconditional L2, representing effects of heterogeneity only, see Table 3] exponentially as expected for the theoretical head variance [equation [5.11] with σ_τ^2 = zero]. Such a decrease suggests continuous improvements of the estimates and thus the predicted heads during iteration. Improvement of the estimates can also be seen in Figures 5.5a and 5.5b. According to these two figures, the variances, based on noise-free hydrographs, started from zero since the first guesses were the effective K and S_s, and then increased as point measurements of K and S_s and head information were included via cokriging. These variances continued to increase after the first iteration due to the successive linear estimation, which successively approximates the nonlinear relation between the head information and the hydraulic properties. Subsequently, these variances approached some stable values, indicating that usefulness of the observed head information was exhausted as the L2 decreased to a very small value. Notice that the spatial variances of the final estimates are smaller than the true variances, suggesting that the estimated fields are smoother than the true fields. This is expected since SimSLE seeks the conditional effective properties. The estimates at the 10th iteration were chosen as our final estimates based on change in variances of the estimates.

Figure 5.4 shows that for the case where hydrographs were noisy, the L2 norm decreased from $0.0405\,cm^2$ [representing effects of heterogeneity plus noise] and stabilized at the 4th iteration to the value of $0.0033\,cm^2$, which is close to the noise level we imposed [i.e., 0.0049 cm^2]. The difference between the true noise level and the L2 norm is expected since L2 norm represents only a sample variance of the noise. Likewise, the variance of the estimated ln$K$ [Figure 5.5a] fluctuated at the 4th iteration, then increased and exceeded the true variance for the

noise hydrographs. The variance of the estimated lnS_s generally increased continuously and rapidly, indicative of divergence of the solution [Figure 5.5b] due to inclusion of noise in the estimation. Therefore, the final estimate was obtained at the 4th iteration where the L2 norm stabilized and before the estimates were "over-improved".

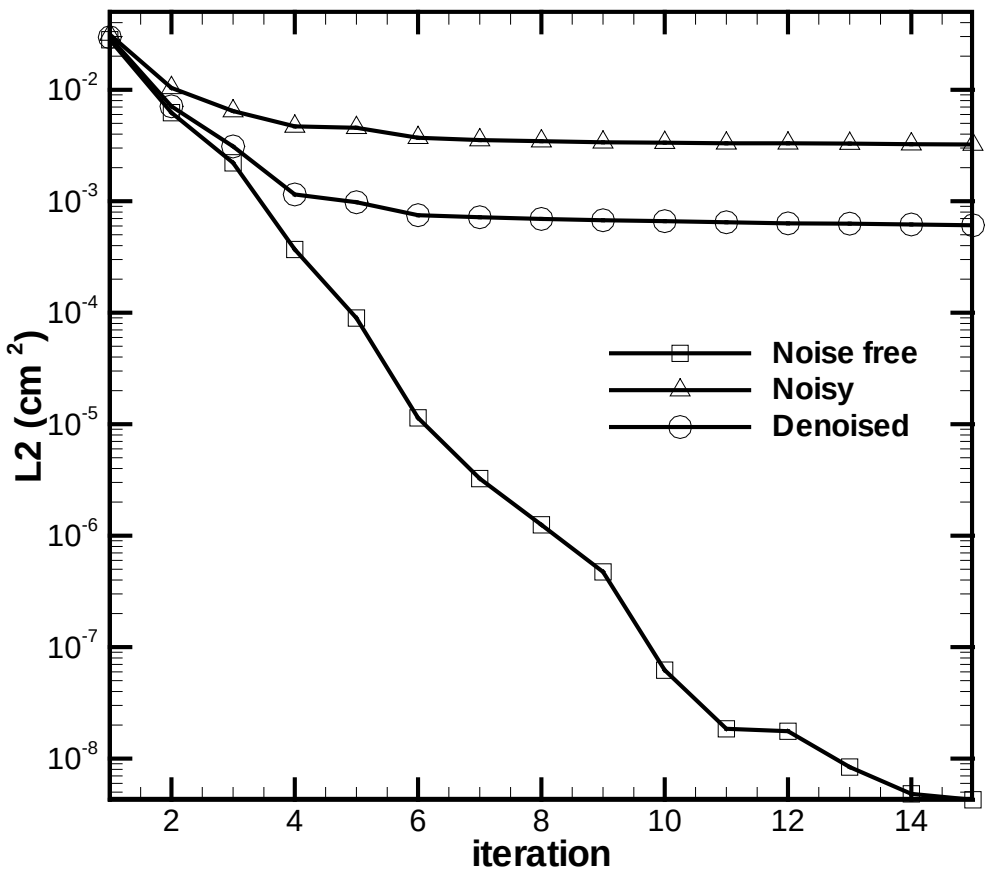


Figure 5.4: The conditional L2 norms of the head as a function of iteration for noise-free, noisy, and denoised data sets for the synthetic aquifer case.

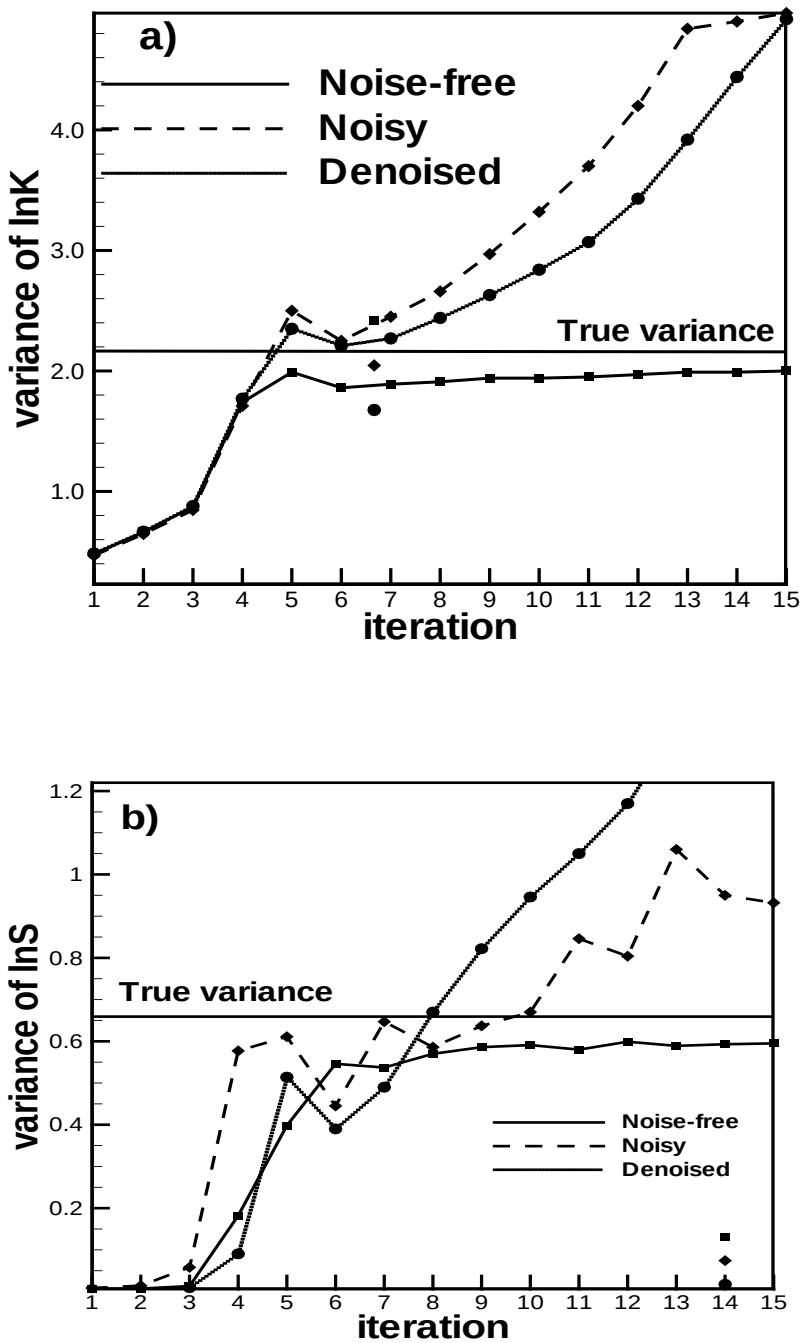


Figure 5.5: Variances of the estimated a] lnK and b] $\ln S_s$ fields vs. the number of iterations for different scenarios associated with the synthetic aquifer.

As illustrated in Figure 5.4, the conditional L2 norm for the denoised case diminished from the unconditional L2 norm, $0.038\,cm^2$ [comprising of the effects from both heterogeneity and noise residuals after wavelet denoising]. Then, it stabilized around the 5th iteration at a value

of $0.0006\,cm^2$, which is smaller than that for the noisy hydrographs since noise was partially removed by the wavelet denoising procedure. The variances of the estimated fields however still increased with iterations similar to those based on noisy hydrographs. According to the L2 stopping criterion, the final estimates were those at the 6th iteration. The variance of the estimated $\ln K$ for this case is slightly greater than its true variance and that of the estimated $\ln S_s$ is smaller than its true variance.

A visual comparison between the true heterogeneous K and S_s fields [Figures 5.2a and 5.2e, respectively] and the final estimated fields based on the noise-free hydrographs at the 10th iteration [Figure 5.2b for K and Figure 5.2f for S_s] suggests that SimSLE depicts hierarchical spatial variation of hydraulic parameters [i.e., variation between units as well as that within a unit]. The estimated K and S_s fields at the 4th iteration using the noisy head values are shown in Figures 5.2c and 5.2g, respectively, while the estimated K and S_s fields based on denoised hydrographs at the 6th iteration are shown in Figures 5.2d and 5.2f, respectively.

Performance metrics for these cases are reported in Table 5.4. According to these metrics, the estimated S_s field can be as good as the estimated K field if the noise-free hydrographs were used. However, it is much worse than the estimated K field if the noisy hydrographs were used. This result confirms that estimation of the S_s field is more prone to effects of noise in hydrographs [e.g., Li et al., 2007]. In comparison to results of the noise-free case, the estimated K and S_s fields using the noisy data and the conditional L2 norm as the stopping criterion are smoother but still retain the general pattern of heterogeneity. The performance metrics also indicate that denoised data can improve the estimates, revealing more details of heterogeneity. But the estimated fields are still inferior to those obtained from noise-free data, suggesting that the wavelet denoising procedure apparently is useful but cannot restore the true hydrograph. Difficulties in estimating S_s field can be attributed to the fact that early time drawdown needed for estimating the S_s field is often small in magnitude. Its SNR is very small once noise is imposed.

Table 5.4: The performance assessment statistics of results from the synthetic aquifer.

		L1	L2	Correlation	Similarity	Iterations
Noise-free	K	0.469	0.405	0. 909	0.889	10
	S_s	0.279	0.140	0.900	0.859	
noised	K	0.643	0.686	0.856	0.833	4
	S_s	0.635	0.666	0.589	0.757	
denoised	K	0.635	0.678	0.857	0.851	6
	S_s	0.536	0.421	0.722	0.762	

Table 5.4 also shows that for evaluation of the estimates over the entire domain, the similarity analysis yields a similar result as other metrics. Nevertheless, we believe that the similarity analysis would have been useful had we targeted the analysis at some specific feature in the domain.

Table 5.5 tabulates means and variances of true and estimated hydraulic properties of each zone using noise-free, noisy, and denoised hydrographs. It further corroborates the early conclusion that the hydraulic tomography and the SimSLE can depict the hierarchical [nonstationary] heterogeneity satisfactorily.

Table 5.5: Means and variances of true and estimated hydraulic properties of each zone using noise-free, noisy, and denoised hydrographs.

	Mean							
	Log Hydraulic conductivity				Log Specific Storage			
Unit	True	Noise Free	Noisy	Denoised	True	Noise Free	Noisy	Denoised
1	2.03	1.86	1.95	2.13	-6.06	-5.78	-6.34	-6.59
2	0.85	1.04	1.43	1.28	-4.82	-4.78	-5.18	-5.08
3	2.46	2.60	2.64	2.60	-6.43	-6.20	-5.85	-6.07
4	-0.67	-0.40	-0.40	-0.64	-5.06	-5.05	-5.16	-4.53

	Variance							
	Log Hydraulic conductivity				Log Specific Storage			
Unit	True	Noise Free	Noisy	Denoised	True	Noise Free	Noisy	Denoised
1	1.01	0.82	0.31	0.67	0.05	0.06	1.60	0.44
2	0.17	0.59	0.50	1.04	0.24	0.38	0.58	0.94
3	2.00	1.47	1.22	1.55	0.01	0.09	0.27	0.46
4	0.46	0.73	0.71	0.57	0.08	0.15	1.46	0.39

Finally, we compare snapshots of simulated heads and stream lines in the aquifer with true and estimated parameter fields at 1.5 seconds after pumping [i.e., early time at which S_s plays an important role] at the center of the aquifer. These simulated fields using true K and S_s parameters are shown in Figure 5.6a; the fields resulting from the effective K and S_s fields are plotted in Figure 5.6b; Figure 5.6c illustrates those fields derived from the estimated K and S_s fields using noisy data; results based on the parameter fields derived from denoised hydrographs are demonstrated in Figure 5.6d. A visual comparison of these figures and the correlation, L1,

and L2 values of the head field listed in the figures further substantiate the usefulness of the HT analysis and denoising procedure.

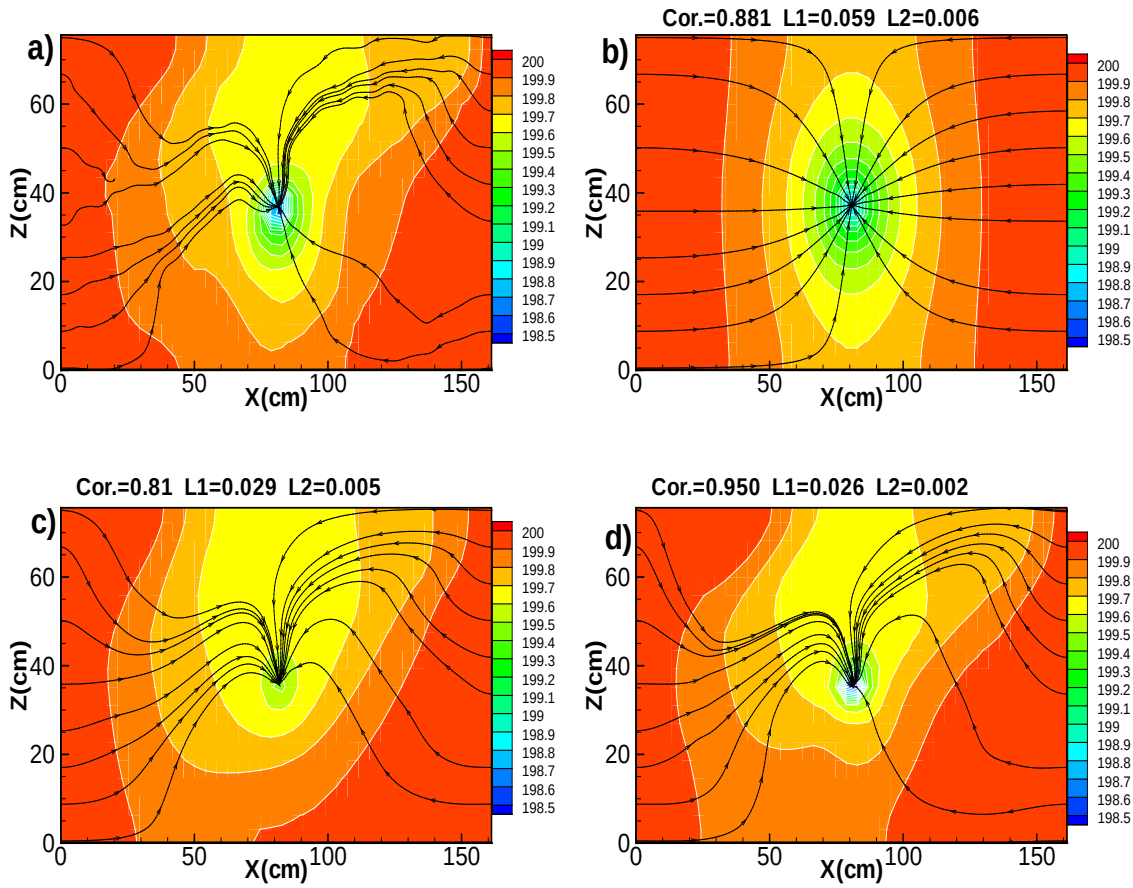


Figure 5.6: Comparison of head, and stream lines at t=1.5 s after pumping in the synthetic aquifer with a] true, and b] effective; c] estimated K and S_s fields based on c] noisy and d] denoised hydrographs. The statistical metrics are based on the true head field in a].

5.4 APPLICATION TO A LABORATORY SANDBOX EXPERIMENT

5.4.1 Preprocessing Data

The laboratory sandbox experiment [see Figure 5.7] conducted by Liu et al. [2007] and Illman et al. [2007 and 2008] involved 8 pumping tests. Drawdown data from the 47 ports excluding the pumping port were collected for each pumping test and they were denoised using the wavelet denoising procedure. Data from two pumping tests [at ports 2 and 5] which have small SNRs that yielded anomalous drawdown contours and did not improve the estimates but caused their divergence were discarded. Only the remaining 6 pumping test data sets were used for the HT analysis. To condition the estimation, one K and S_s values from in situ slug test measurements at port #1 [Figure 5.7] [Liu et al., 2007] were used as the hard data; effective K and S_s [0.1268 cm/s and 8.73x10^{-4} 1/cm] were derived from minimization of equation [5.14]; the variances of ln$K$ and lnS_s were estimated using equation [15] to be 2.0 and 0.1, respectively. The correlation scales were assessed subjectively based on the heterogeneity pattern of the laboratory sandbox [namely, 70 cm and 20 cm in the horizontal and vertical directions, respectively].

During the HT experiment, the total head at the two side boundaries of the sandbox varied slightly [maximum 0.13cm]. To eliminate the effect of time varying boundary condition, the actual drawdown at observation locations minus the observed boundary drawdown was used as correction to the observed drawdowns. These observed drawdowns were subtracted from an assigned initial total heads [200 cm] to obtain the observed total heads for the HT analysis. Five observed total heads at 0.75s, 1.50s, 2.25s, 3.00s, and 15.00s from each observation port during each pumping test were selected for the HT analysis.

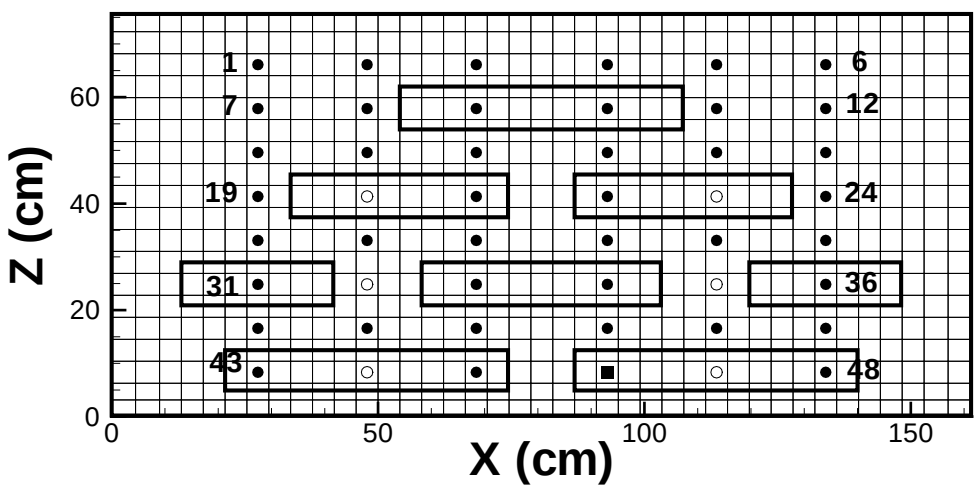


Figure 5.7: Schematic setup and discretization of the sandbox used in the experiment in Liu et al. [2007] and Illman et al. [2007]. Open circles are pumping ports; both solid and open circles are used as observation ports. The solid square denotes the pumping port for the validation purpose. The open rectangles shown are the low permeability zones.

5.4.2 Results

A plot of variances of the estimated ln$K$ and LnS_s fields as a function of the iteration number is illustrated in Figure 5.8. The variances increased continuously indicating effects of unresolved noise. The final estimates were chosen based on the stabilization of the conditional L2 norm of the head during iteration [Figure 5.9], which suggest that the best estimates are at the 6th iteration.

Figure 5.10 shows the distributions of the estimated K and S_s fields, respectively. In these figures, six low K zones in the sandbox [Figure 5.7] are vividly portrayed by the estimated K field, but the low K zones close to the bottom are fuzzy. The low resolution at the bottom is due to the no-flux boundary at the bottom where the flow generally follows the boundary, and where the pressure excitations were not sampled due to absence of monitoring ports between the low K zones and the bottom boundary--consistent with findings by Illman et al. [2007, 2008] and Liu et al. [2007].

The estimated S_s field in Figure 5.10 on the other hand does not reflect the pattern of the K field. Rather, the field reflects an overall trend that the S_s values of the medium at the bottom

are smaller than those on the top. This pattern appears to be physically correct: sands at the bottom are compressed more due to greater overlying materials. The result is also consistent with the estimated S_s field from the analysis of cross-hole aquifer tests by Liu et al. [2007].

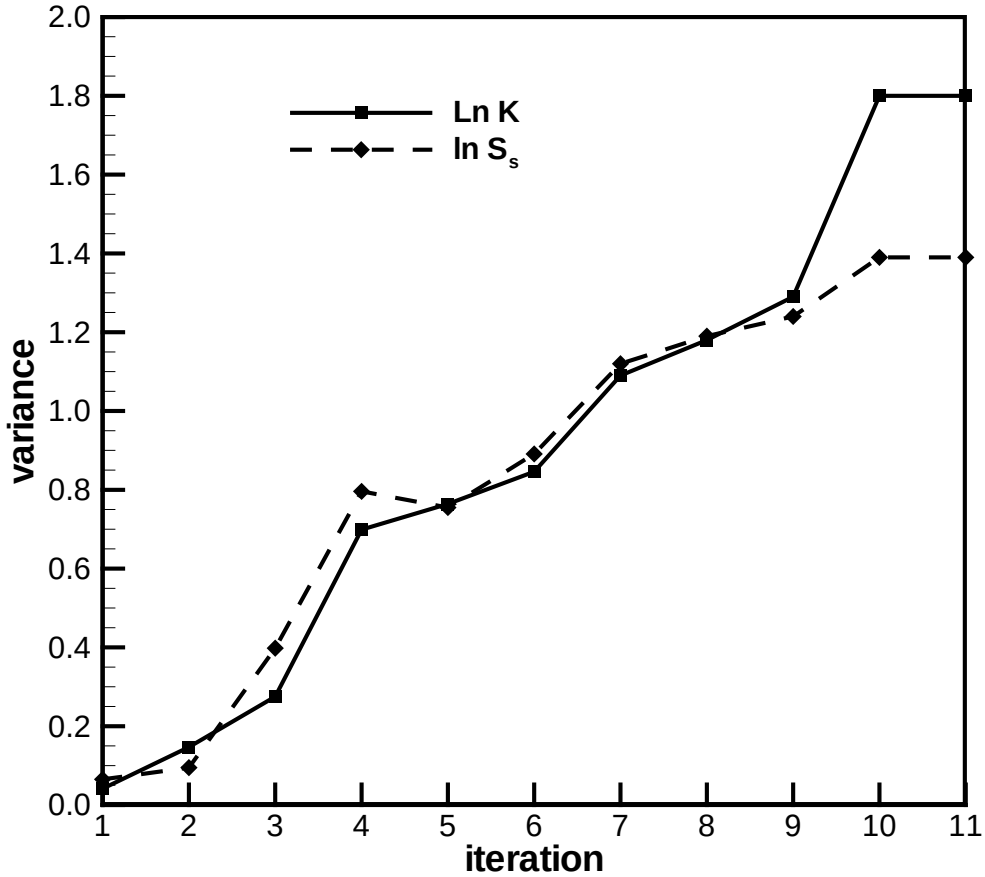


Figure 5.8: The variances of estimated ln$K$ and lnS_s as a function of iteration for the sandbox experiment.

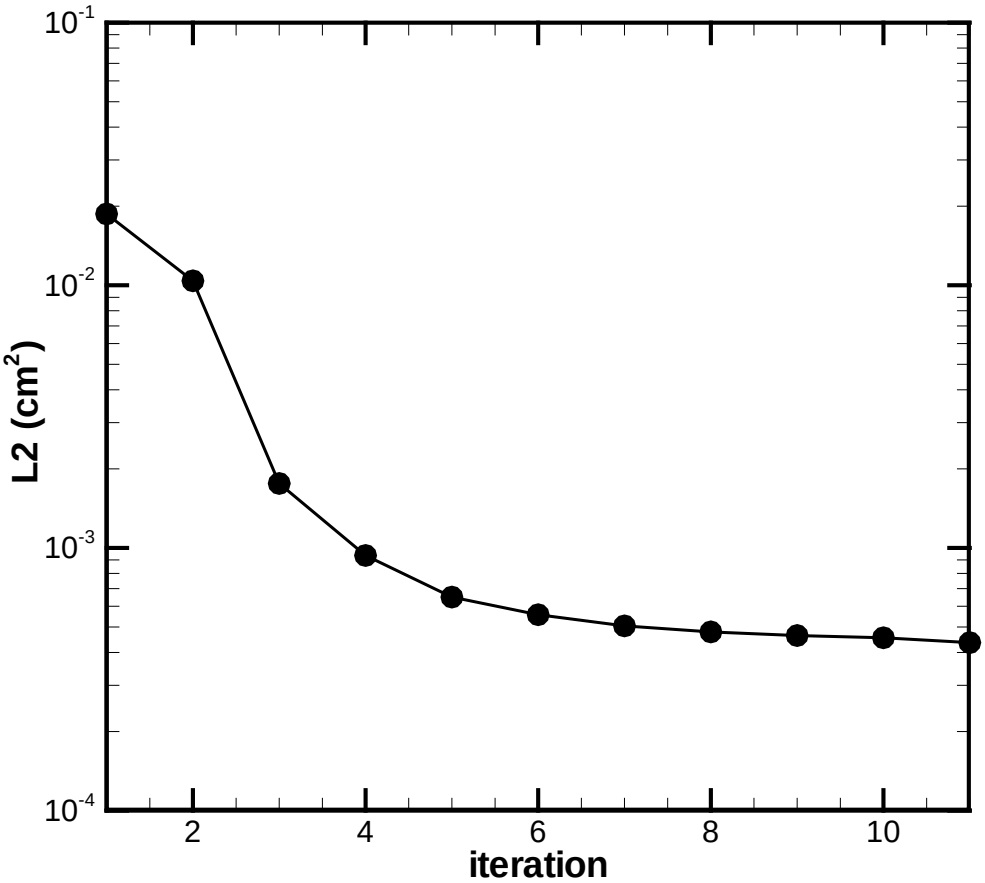


Figure 5.9: The conditional L2 norm of the head as a function of iteration for the sandbox experiment.

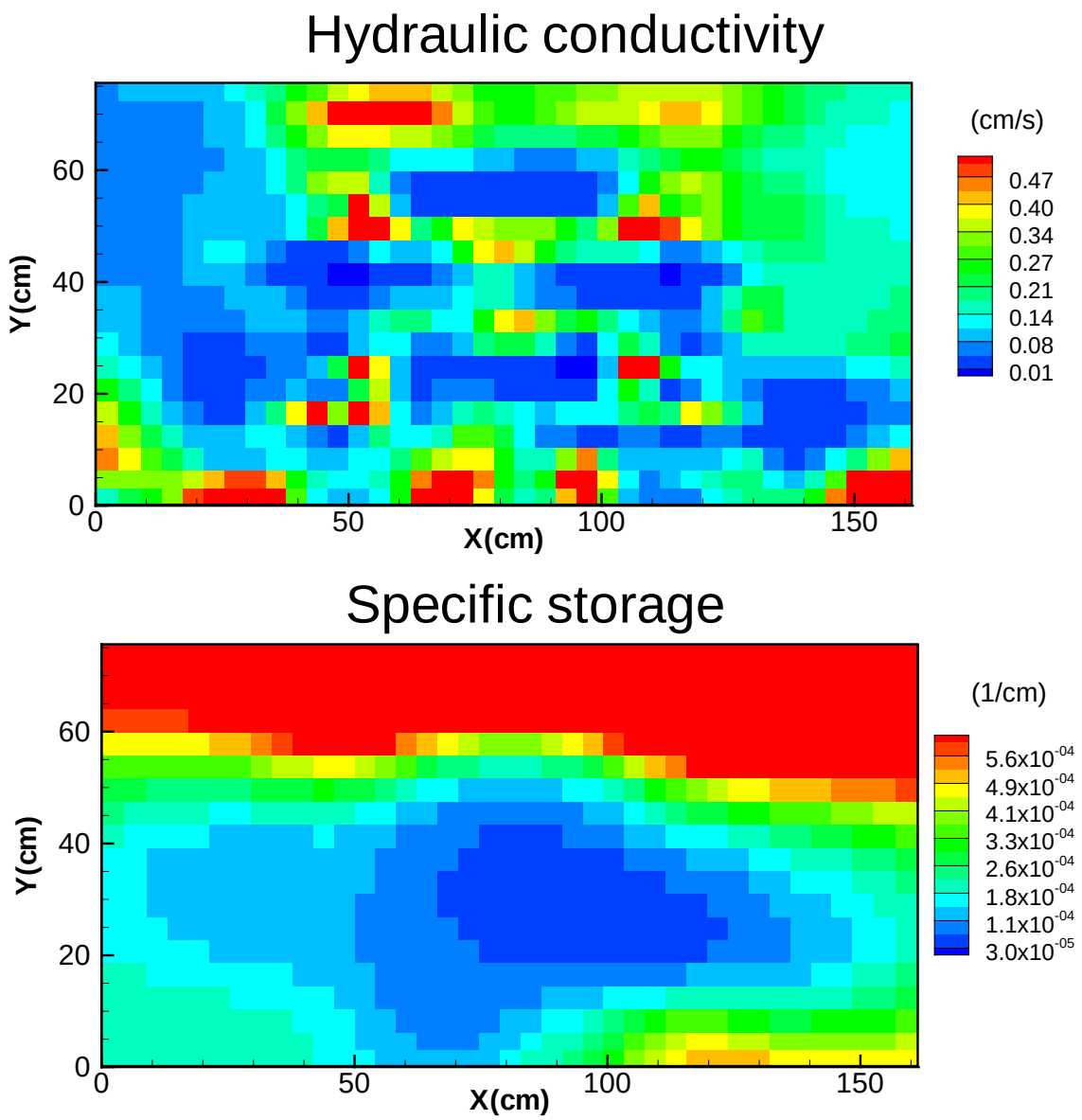


Figure 5.10: The estimated hydraulic conductivity [cm/s] and specific storage [1/cm] fields for the sandbox experiment.

In this sandbox experiment, the true K and S_s fields are unknown and the performance metrics therefore cannot be evaluated. In order to validate these estimates, we followed the approach by Liu et al. [2007] which uses the estimated K and S_s fields as input to the forward flow model to simulate a pumping test conducted at port #46 that was not used in the HT experiment. If the estimated K and S_s fields are representative of the true fields in the sandbox, the temporal and spatial distributions of the simulated head due to the pumping should closely predict the observed ones.

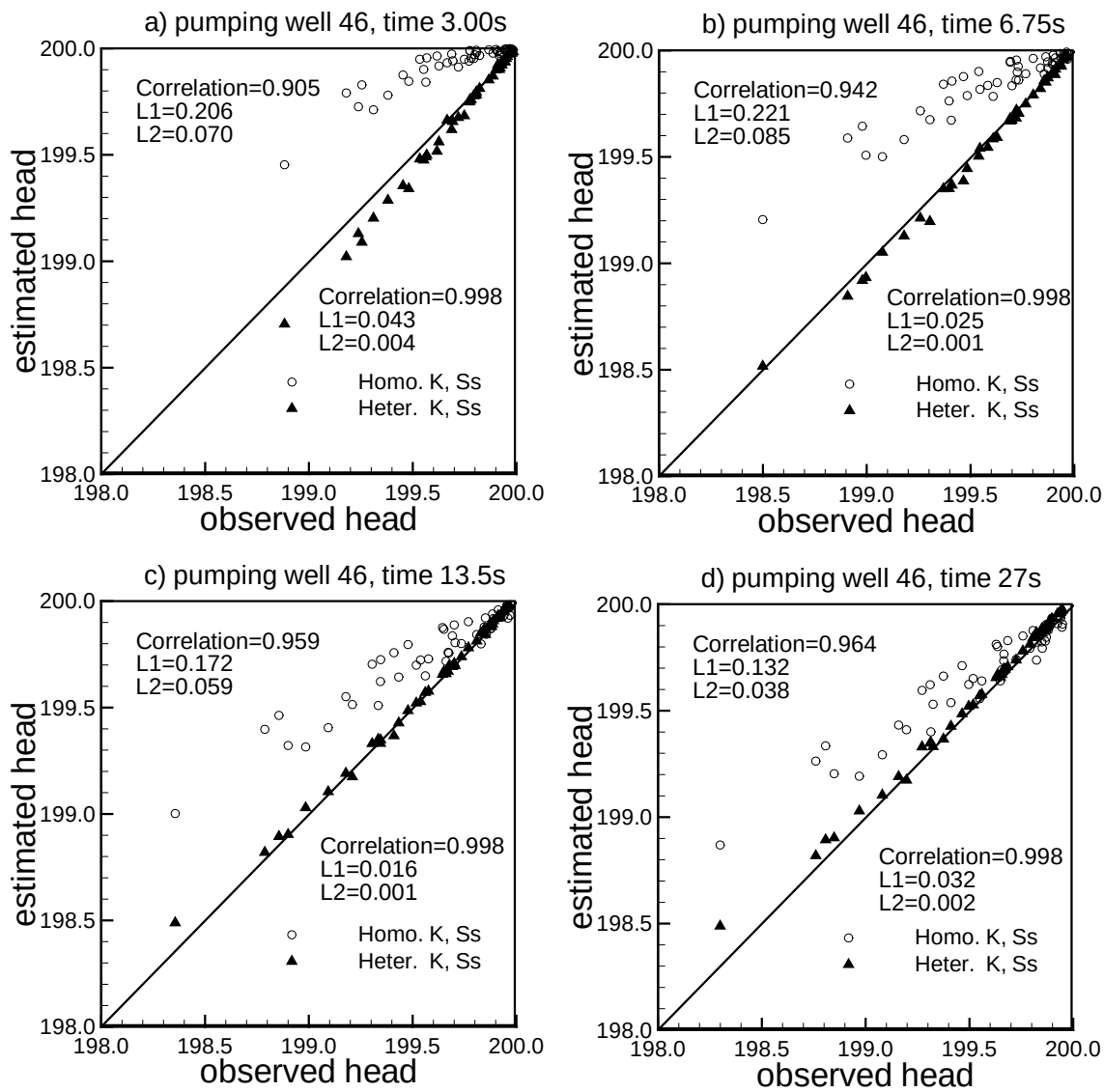


Figure 5.11: Validation results: observed vs. estimated heads [cm] at the 47 observation ports at 4 different times in the sandbox. Two simulated heads were considered: one using estimated effective K and S_s fields from the equivalent homogeneous domain and the other using the estimated heterogeneous K and S_s fields derived from the analysis of HT.

As illustrated in Figure 5.11, the estimated K and S_s fields yielded heads close to the observed ones at all of the 47 ports at 3.0s, 6.75s, 13.5s, and 27.0s [early to late time]. The large discrepancy at 3.0s may be attributed to the numerical discretization error as well as inaccuracy

of the estimated S_s field, which controls the early time behavior of the drawdown.

This comparison suggests that 1] the effective parameters of the sandbox scale obtained by simultaneously fitting drawdowns from the pumping tests at 6 different locations failed to satisfactorily reproduce drawdowns caused by pumping at another location in the sandbox. Perhaps, they can do better for a stress area [e.g., a production well field] that covers an area much greater than many correlation scales of the heterogeneity. 2] HT in conjunction with SimSLE characterizes the heterogeneity of aquifers sufficiently such that drawdown evolution due to a different pumping event is predicted. We believe these results are significant. First, not only do the results reinforce the validity of HT, SimSLE and our data processing approach, but they also demonstrate a need for fine-resolution mapping of K and S_s fields to overcome the phenomenological nature associated with the domain-scale effective parameters. Besides, these results confirm that the classical governing ground water flow equation can yield excellent predictions of drawdowns in a heterogeneous sandbox when the K and S_s fields are adequately characterized and the initial and boundary conditions as well as source/sink terms are fully prescribed. The result supports a similar conclusion reached by Liu et al. [2007].

5.5 DISCUSSION

Distinguishing noise from effects of heterogeneity in a hydrograph can be highly subjective unless characteristics of noise or heterogeneity are known a priori. Complete removal of noise from hydrographs is difficult and unresolved noise residuals can impact the estimation. The impact is manifested through a continuous increase in the variances of estimated hydraulic properties but their spatial pattern remains almost constant. As a result, the head field changes continuously but its L2 norm stabilizes. Stabilization of the conditional L2 norm of the head thus works well as the convergence criterion for our SimSLE when the data are infested with noise.

Finally, although no explicit comparison with SSLE has been presented, we believe that SimSLE has several advantages over the SSLE. 1] SimSLE needs to evaluate the adjoint state equation only once for a given observation location using newly estimated hydraulic property fields from all pumping tests since the adjoint state equation is independent of the pumping rate and pumping location. On the other hand, using SSLE, one must solve the adjoint state equation for each pumping test because the parameters in the adjoint state equation are modified for each pumping test. 2] SimSLE avoids the loop iteration of SSLE, and the computational effort is thus reduced. 3] Adding data sets in different sequences in SSLE may lead to a slightly different final result, which is more sensitive to the last data set [see Illman et al., 2008]. This problem does not exist in SimSLE. 4] SimSLE uses all observations simultaneously, providing more constrains for the inverse problem and thus converges faster than SSLE.

The disadvantages of SimSLE are: 1] since all head data sets are used simultaneously and the same convergence criteria are applied to reach the final estimate in SimSLE, one bad data set may affect the overall quality of the estimate. Note that both SimSLE and SSLE result in the same estimate if the data sets are free of errors. 2] The memory requirement is greater because the sizes of covariance matrix of h and the cross covariance matrix of f and h are larger than the corresponding matrices in SSLE.

6. HYDRAULIC AND PARTITIONING TRACER TOMOGRAPHY FOR CHARACTERIZATION OF DENSE NONAQUEOUS PHASE LIQUID SOURCE ZONES

6.1 INTRODUCTION

Non-aqueous phase liquids [NAPL's] are long-term sources of soil and groundwater contamination. Accurate risk assessment and effective remediation of the sites contaminated with NAPL's require detailed characterization of source zone [i.e., the region wherein NAPL contamination is present]. The source zone controls the size of shape of the associated aqueous phase contamination plume. Source zone characterization often involves characterizing hydrogeology [e.g., spatial variation of hydraulic conductivity, specific storage, and porosity] and delineating source zone [e.g., contaminant mass location, contaminant strength] [NRC, 2004].

Charactering source zone hydrology, especially the heterogeneity of hydrologic parameters, is critical to accurately depict groundwater flow and contaminant plume movement as well as design remediation schemes. Many approaches have been developed to estimate aquifer heterogeneity using both scattered direct measurements and indirect measurements [e.g., hydraulic head, solute concentration]. Generally speaking, direct measurement methods [e.g., core sampling] are costly and invasive to obtain detailed 3-D heterogeneity for field scale problems. Recently developed hydraulic tomography [Gottlieb and Dietrich, 1995; Renshaw, 1996; Yeh and Liu, 2000; Liu et al., 2002; McDermott et al., 2003; Zhu and Yeh, 2005], which evolved from the CAT scan concept of medical sciences, has been proposed as a cost-effective tool to delineate detailed physical heterogeneity of subsurface. Hydraulic tomography is a series of cross-well interference tests. Specifically, an aquifer is stressed by pumping or injecting water through a well; at the same time, the aquifer's responses [i.e., hydraulic heads] are monitored in other wells. If we sequentially switch the pumping or injecting location, stressing the aquifer and monitoring its responses, without installing additional wells, would result in a large number of aquifer responses. These responses are then used to estimate spatial variation of hydraulic parameters. To efficiently interpret the abundant data obtained from hydraulic tomography, Yeh and Liu [200] developed a sequential successive linear estimator [SSLE]. The SSLE approach seeks effective parameter fields conditioned on the data obtained from hydraulic tomography. The SSLE considers the nonlinear relation between hydraulic heads and parameters through an iterative way, and overcomes the difficulties of solving huge system of equations by sequentially including data. Their work shows that combination of SSLE and hydraulic tomography is a cost-effective technique for delineating heterogeneity using a limited number of invasive observations [Liu et al., 2002]. While the work by Yeh and Liu [2000] and Liu et al., [2002] is limited to interpret steady state head responses from hydraulic tomography, Zhu and Yeh [2005] extended the SSLE to interpret transient hydraulic heads for estimating hydraulic conductivity and specific storage fields in 3-D aquifers.

Besides the head data as major information used for mapping physical heterogeneity, solute concentration data have also been used to estimate hydraulic conductivity. For example, Harvey and Gorelick [1995] estimated hydraulic conductivity using measurements of hydraulic conductivity, heads and solute arrival time. They concluded that solute arrival time and heads provide different information about the conductivity field. Li and Yeh [1999] developed a cokriging method to estimate hydraulic conductivity field conditioned on pressure head, solute

transport, and solute arrival time for variable saturated media. They concluded that steady state head measurements are most useful for estimating hydraulic conductivity among three types of measurements. But they showed that adding solute concentration data can improve the estimates than using head data alone. Cirpka and Kitanidis [2001] used the first two temporal moments of solute data to estimate hydraulic conductivity. They indicated that the use of tracer data alone may lead to convergence problems, and hereby recommended to combine head and tracer data for inversion. Solute concentration data are also used to map chemical parameters. For example, Huang et al. [2004] extended a sequential self-calibrating method to estimate both hydraulic conductivity and sorption partitioning coefficient fields conditioning on nonreactive and reactive tracer data.

The tracer concentration data have been used to delineate source zones in the past. Jin et al. [1995] developed a partitioning interwell tracer tests [PITT] to detect and characterize NAPL's. During a PITT test, several different tracers are injected into aquifer through one or more injection wells. The tracers usually consist of non-reactive tracer, partitioning tracers with different partitioning coefficients. NAPL's in the subsurface generally control differences in behaviors of the nonreactive and partitioning tracers breakthroughs. Thus, the breakthroughs can be used to estimate possible locations and concentrations of NAPL's. Using these breakthroughs, Jin et al. [1995] estimated the average NAPL residual saturations by using a nonlinear least squares regression method. James et al. [1997] developed a conditioning algorithm to derive spatial variation of NAPL field from the first two temporal moments of tracer breakthroughs. Sciortino et al. [2000] estimated the location and dimensions of a single DNAPL pool using a least squares minimization based on analytical solutions. Zhang and Graham [2001] directly used tracer concentration measurements to estimate spatial heterogeneity of NAPL saturation by an extended Kalman filter technique. However, Inverse methods by James et al. [1997], Sciortino et al. [2000], and Zhang and Graham [2001] rely on dense sampling wells to reveal major features of NAPL fields, therefore is costly on field problems. Datta-Gupta et al. [2002] proposed a streamline-based inverse method analyzing PITT tests to estimate 3-D spatial variation NAPL saturation. This method uses steam-line simulator for both forward simulation and sensitivity calculation, therefore, providing significant computation efficiency. However, similar to Vasco et al. [2000] and Brauchler et al. [2003], the streamline-based method ignores interaction between adjacent ray paths and possible boundary effects. Consequently, their method requires an extensive number of iterations and pairs of source/sensor data to achieve a comparable resolution to that achieved from inverting a three-dimensional model. Also, the stream-line method has difficulties in solving transport problems involving local physical dispersion [Huang et al., 2004].

Technologies using partitioning tracers to detect NAPL's have been developed in the past. These technologies generally have focused on detecting the presence of NAPL's but not their spatial distribution in a field. Depicting high-resolution spatial distributions of NAPL's in a field using current technologies demands a dense multi-level sampler network. Such a network involves costly invasive drilling operations and is practically impossible. As a result, an innovative cost-effective technology remains to be developed.

In this study, we developed a new technology to meet the demand. We extend the concept of hydraulic tomography to image the spatial distribution of NAPL's in the subsurface—partitioning tracer tomography. Moreover, we adopted the stochastic information fusion concept by Yeh and Šimůnek [2002] to take advantage of the ability of hydraulic tomography to enhance the ability of partitioning tracer tomography.

The concept of tracer tomography is analogous to that of hydraulic tomography. During a tracer tomography survey, a forced gradient steady flow condition is established by injecting water into aquifer from one well. An impulse solution of tracers is released into the aquifer from the injection wells afterward. The tracer concentration data subsequently are sampled from observation wells. After complete breakthrough curves for all observation wells are recorded, a new test is repeated by changing the injection location. Again, the resulting concentration breakthroughs are monitored. By sequentially switching the injection location with a limited number of wells, a large number of concentration breakthrough data sets are obtained. This tomography survey approach in essence is a novel experimental design to "create" more independent data sets from a limited number of wells than any traditional experiment. Finally, an estimation or inverse modeling technique exploits these breakthrough data and injection information to determine the spatial distribution of NAPL's as well as aquifer hydraulic properties. Hydraulic head responses during the tracer injection tests can also be used to estimate spatial distribution of hydraulic properties [i.e., hydraulic tomography]. The head information is directly related to the hydraulic conductivity while the breakthrough data are controlled by many different hydraulic and transport properties [Li and Yeh, 1999]. Estimation of the hydraulic heterogeneity using the information from hydraulic tomography prior to the interpretation of the tracer tomography thus can improve the estimate of the distribution of NAPL's.

While the hydraulic/tracer tomography is a novel experimental design, efficiently and effectively processing the huge amount of data from the experiment is a challenge. We take on this challenge by extending the SSLE inverse approach by Yeh and Liu [2000] and Zhu and Yeh [2005] to map both physical and chemical heterogeneity of the NAPL source zone.

6.2 DERIVATION OF INVERSE ALGORITHM

6.2.1 Governing Equations

During a tracer test, the force gradient flow is assumed to be steady and it can be described by

$$\nabla \cdot [K(\mathbf{x})\nabla H] + Q(\mathbf{x}_k) = 0 \qquad [6.1]$$

subject to boundary conditions:

$$H\big|_{\Gamma_1} = H_1, \;\; [K(\mathbf{x})\nabla H]\cdot \mathbf{n}\big|_{\Gamma_2} = q_h, \qquad [6.2]$$

where H is the total head [L], $\mathbf{x}$ is the spatial coordinates [$\mathbf{x} = \{x_1, x_2, x_3\}$, [L], and x_3 represents the vertical coordinate and is positive upward], $Q[\mathbf{x}_k]$ is the pumping rate [1/T] per unit volume of the aquifer at the location $\mathbf{x}_k$, $K[\mathbf{x}]$ is the spatially varying saturated hydraulic conductivity field [L/T]. In equation [6.1], H_1 is the prescribed total head at Dirichlet boundary Γ_1, q_h is the specific flux [L/T] at Neumann boundary Γ_2, $\mathbf{n}$ is a unit vector normal to the boundary.

Transport of the tracer is assumed to be described by the advection-dispersion-retardation equation

$$\frac{\partial}{\partial t}(\theta_w c + K_N \theta_n c) = -\nabla \cdot (\mathbf{q}c) + \nabla \cdot (\theta_w \mathbf{D} \nabla c) + Q(\mathbf{x}_k) c_n N(\mathbf{x} - \mathbf{x}_k)(t - t_b) \qquad [6.3]$$

Subject to following initial and boundary conditions:

$$c\big|_{t=0} = c_0 \qquad [6.4]$$

$$c\big|_{\Gamma_1} = c_1 \qquad [6.5]$$

$$\mathbf{q}c - \theta \mathbf{D} \nabla c\big|_{\Gamma_2} = \mathbf{q}_s \qquad [6.6]$$

where c is the solute concentration [M/L^3]; θ_w is the volumetric water saturation [water saturated porosity, we call it water saturation thereafter in this study for simplicity]; K_N is the partitioning coefficient and θ_n is the volumetric NAPL saturation [we call it NAPL saturation thereafter in this study for simplicity]; **q** is Darcy velocity vector [L/T] given by

$$q_i = -K \frac{\partial H}{\partial x_i}. \qquad [6.7]$$

D is dispersion tensor given by

$$D_{ij} = \frac{v_i v_j}{|\mathbf{v}|}(\alpha_L - \alpha_T) + \delta_{ij}(\alpha_T v + D_d^*) \qquad [6.8]$$

where α_L and α_T are the longitudinal and transverse dispersivity [L], respectively; v_i is the seepage velocity [L/T] defined as q_i / θ; $|\mathbf{v}|$ is the magnitude of the seepage velocity; δ_{ij} is the Kronecker delta function which equals unity if the indices are identical and zero otherwise, and D_d^* is the molecular diffusivity [L^2/T]. In addition, Q is the injection rate [1/T] per unit volume of the porous medium, and c_n is the solute concentration of the injecting solution; $N(\mathbf{x} - \mathbf{x}_k)(t - t_b)$ is a Dirac delta function which equals unity at location $\mathbf{x}_k$ at time earlier than t_b and zero otherwise. Moreover, c_0 is the initial solute concentration; c_1 is the prescribed concentration at the Dirichlet boundary Γ_1; q_s is solute flux at the Neumann boundary, Γ_2. These flow and solute transport equations are solved by a 3-D finite-element and modified method of characteristic approach [MMOC3] developed by Srivastava and Yeh [1992].

6.2.2 SSLE Inverse Method

Sequential Successive Linear Estimator [SSLE] [Yeh et at., 1996; Yeh and Zhang, 1996; Zhang and Yeh, 1997; Hanna and Yeh, 1998; Varas-Guzman and Yeh, 1999, 2002; Hughson and Yeh, 2000; Yeh and Liu, 2000; Liu et al., 2002; Yeh et al., 2002; Zhu and Yeh, 2005] is a stochastic estimator that seeks effective mean parameter fields conditioned on available primary information [i.e., point measurements of parameters], secondary information [i.e., measurements

of hydraulic heads, water contents, solute concentrations], and spatial covariance functions of parameters [i.e., geologic structures]. The SSLE algorithm is adopted and extended in this study to estimate hydraulic conductivity, water saturation, and NAPL saturation fields using data sets collected from hydraulic/tracer tomography. The major steps of SSLE method for interpreting head and concentration data from hydraulic/tracer tomography are summarized below. Detailed description was given by Yeh and Liu [2000] and Hughson and Yeh [2000].

First, a geologic medium under investigation is defined as a system. The spatial hydraulic conductivity, water saturation, or NAPL saturation distributions of the geologic medium are then regarded as the parameter fields of the system. Due to spatial variability of the parameter fields, as well as their uncertainty, the natural logarithms of the parameter fields, χ, are often represented as stochastic processes. Notice that the use of the logarithm merely is mathematical convenience [e.g., avoiding negative values of parameter estimates]. Each stochastic parameter field is expressed as a combination of a mean part $\overline{\chi}$ and a perturbation part ζ [i.e., $\ln \chi = \overline{\chi} + \zeta$]. Accordingly, the responses of the system, ϕ, due to a tomographic test are considered to be stochastic processes. Each of them is expressed in term of a mean $\overline{\phi}$ and a perturbation ε, such that $\phi = \overline{\phi} + \varepsilon$. The system response referred here are either hydraulic heads or tracer concentrations in the geologic medium.

Starting with a data set obtained from one hydraulic/tracer tomography test, SSLE first calculates mean responses of the geologic medium, given its mean parameter fields $\overline{\chi}$ [or guessed parameter fields]. The differences between the simulated mean system responses and observed responses are then evaluated at sample locations. These differences, in conjunction with available measurements of the parameters, are subsequently employed in classical cokriging [a multivariate linear estimator] to yield linear estimates of conditional mean parameter perturbation field. Evaluation of the conditional [residual] covariances of the parameters then follows. The cokriging weights are determined based on the covariances of parameters in space, temporal and spatial covariance of responses, as well as the cross covariance between parameters and responses. A first-order analysis is used to derive the covariance of responses and their cross covariance with parameters. The first-order analysis involves the use of the sensitivity matrices of the responses to parameters, at given mean parameter values and the covariance of the parameters.

With the linear estimates of the conditional mean parameter fields, the governing flow and solute transport equations are then solved for the conditional mean system responses, given the observed responses and parameters of the system. Notice that cokriging is a linear estimator but the relation between system parameters and responses is nonlinear. As a consequence, the estimates of the conditional mean system parameters are not optimum or the benefit of data are not fully exploited. In other words, the simulated responses at sample locations do not honor the observations. To overcome this problem and to improve the parameter estimates, a new linear estimator is employed that uses the linear combination of weighted differences between the simulated and observed responses at the sample locations. The weights of the new estimator are determined from the covariances of system responses and their cross-covariances with system parameters, similar to cokriging. Instead of using the unconditional [original] covariances of the parameters as in cokriging, the new estimator makes use of the residual [conditional] covariances of parameters from the previous estimate and the sensitivity matrices evaluated with the previously estimated parameter fields. Therefore, as differences between the simulated and observed responses at the sample locations become small, the improvement diminishes. The

above linear procedure for improving the estimates is then repeated until the improvement diminishes to a prescribed value. This is the logic behind the successive linear estimator [SLE]. SLE can include all the measurements collected from a tracer tomography survey simultaneously. However, this approach can lead to numerical problems because of extremely large system of equations [Hughson and Yeh, 2000] to be solved. To avoid this numerical problem, the system responses collected from hydraulic/tracer tomography is added into the inverse process sequentially, such as one tracer test by one tracer test. This approach therefore is called the sequential successive linear estimator [SSLE].

6.2.3 Evaluation of Covariance and Cross Covariance

As discussed previously, covariances of system responses and cross covariances between responses and parameters are the heart of the SSLE method. They are determined by using a first order analysis, which is explained below.

Consider that a system domain is discretized into N elements; each element has a uniform value for a given parameter but the value varies from element to element. Supposed that the response of the system is governed by equations [6.1] and [6.2] and associated boundary and initial conditions. Then, expanding the response at the center of element *i* in a Taylor series about the mean value of parameter fields and neglecting second- and higher-order terms, the response perturbation can be expressed as

$$\varepsilon_i = \frac{\partial \phi_i}{\partial \ln \chi_k^j} \zeta_k^j \qquad [6.9]$$

where ε_i is the response perturbation at location *i* ; ζ_k^j is perturbation of $\ln \chi_k$ at location *j* [*j*=1, 2, …, N, where N is the total number of elements in the domain and *k* is the parameter index: hydraulic conductivity, water saturation, and NAPL saturation]. In equation [4.41], the Einstein summation convention applies. For head responses, *k*=1 and $\ln \chi_k$ represent log hydraulic conductivity field [*k*=1]; for concentration responses, k=1,2,3 and $\ln \chi_k$ represent hydraulic conductivity [*k*=1], water saturation [*k*=2], or NAPL saturation [*k*=3]; $\frac{\partial \phi_i}{\partial \ln \chi_k^j}$ is the sensitivity of ϕ at location *i* to $\ln \chi_k$ at location *j*. Multiplying equation [6.9] by a parameter perturbation at a location *m*, ζ_k^m, and taking expected values gives the cross covariance between the system response, ε_i, and the parameter, ζ_k^m:

$$R(\varepsilon_i, \zeta_k^m) = R(\zeta_k^m, \zeta_k^j) \frac{\partial \phi_i}{\partial \ln \chi_k^j} \qquad j=1, 2,\ldots,N \qquad [6.10]$$

where $R(\zeta_k^m, \zeta_k^j)$ is the covariance of the parameter ζ_k between location *m* and *j*, which is assumed to follow certain covariance functions and known a priori. Similarly, assuming the parameters are un-correlated to each other, the spatial covariance between the system responses at locations *i* and *m* is ε_i and ε_m given by

$$R(\varepsilon_i, \varepsilon_m) = R(\zeta_k^p, \zeta_k^q) \frac{\partial \phi_i}{\partial \ln \zeta_k^p} \frac{\partial \phi_m}{\partial \ln \zeta_k^q}; \quad p \text{ and } q = 1, 2, \ldots, \text{N} \qquad [6.11]$$

6.2.4 Calculation of Sensitivities

The evaluation of covariance and cross covariance requires sensitivities of responses to parameters. Similar to previous SSLE method, we utilize the adjoint state method to minimize computation costs for evaluating sensitivities of responses to parameters. The adjoint method has been described in many literatures in details [Sykes et al., 1985; Sun, 1994; Li and Yeh, 1998]. The sensitivity of head to hydraulic conductivity is given by Li and Yeh [1998] and Zhu and Yeh [2005]. Below a short summary of the derivation of sensitivity of partitioning tracer concentration to parameters is presented.

Given a set of concentration measurements obtained from one tracer test, the adjoint state method first solves for mean heads and concentrations throughout the domain using equation [6.1] and equation [6.3] with associated boundary conditions. Next, for each concentration measurement, two adjoint states ψ_c and ψ_h are obtained through solving the following adjoint state equations [Sun, 1994; Li and Yeh, 1998]:

$$-\frac{\partial \phi_c}{\partial x} \bar{q}_i - \frac{\partial}{\partial x_i}(\bar{\theta}_w D_{ij} \frac{\partial \psi_c}{\partial x_j}) - \bar{\theta}_w \frac{\partial \psi_c}{\partial t} - \bar{\theta}_n K_N \frac{\partial \psi_c}{\partial t} + \frac{\partial R}{\partial c} = 0 \qquad [6.12]$$

$$\frac{\partial}{\partial x_i}(\bar{K} \frac{\partial \psi_h}{\partial x_i}) + \frac{\partial}{\partial x}(\frac{\partial \bar{c}}{\partial x} \bar{K} \psi_c) + \frac{\partial R}{\partial H} = 0 \qquad [6.13]$$

subject to the boundary and final time conditions:

$$\psi_c = 0 \quad \text{at } \Gamma_1$$

$$\bar{q}_i \psi_c + D_{ij} \frac{\partial \psi_c}{\partial x_j} = 0 \quad \text{at } \Gamma_2$$

$$\psi_h = 0 \quad \text{at } \Gamma_1 \qquad [6.14]$$

$$-\bar{K} \frac{\partial \psi_h}{\partial x_j} = \psi_c \bar{K} \frac{\partial \bar{c}}{\partial x_j} \quad \text{at } \Gamma_2$$

$$\psi_c \Big|_{t=T_e} = 0$$

where the overhead denotes the mean; $R = c\delta(x - x_k)(t - t_l)$ represents the concentration at location x_k and time t_l, where δ is Dirac delta function, which equals unity if x equals x_k and t

equals t_l, zero otherwise. Then, the sensitivity of tracer concentration at location x_k and time t_l to parameters at location $\mathbf{x}_n$ are given by:

$$\frac{\partial c(\mathbf{x}_k, t_l)}{\partial \ln K(\mathbf{x}_n)} = \int_T \int_{\Omega_n} [-\bar{K} \frac{\partial \psi_h}{\partial \mathbf{x}} \frac{\partial \bar{H}}{\partial \mathbf{x}} - \psi_c \frac{\partial \bar{c}}{\partial \mathbf{x}} \bar{K} \frac{\partial \bar{H}}{\partial \mathbf{x}}] d\Omega dt \qquad [6.15]$$

$$\frac{\partial c(\mathbf{x}_k, t_l)}{\partial \ln \theta_w(\mathbf{x}_n)} = \int_T \int_{\Omega_n} [\frac{\partial \psi_c}{\partial \mathbf{x}} \bar{\theta}_w D_{ij} \frac{\partial \bar{c}}{\partial \mathbf{x}} + \bar{\theta}_w \psi_c \frac{\partial \bar{c}}{\partial t}] d\Omega dt \qquad [6.16]$$

$$\frac{\partial c(\mathbf{x}_k, t_l)}{\partial \ln \theta_n(\mathbf{x}_n)} = \int_T \int_{\Omega_n} K_N \bar{\theta}_n \psi_c \frac{\partial \bar{c}}{\partial t} d\Omega dt \qquad [6.17]$$

where Ω_n is the volume of the element where the observation at $\mathbf{x}_n$ locates. Notice that in our adjoint methods, we assume the tracer transport induced by a tracer test is convection dominated, and therefore we assume the local dispersion coefficient is small and can be neglected. Adding dependence of dispersion on parameter variation in adjoint method is straightforward, which is given by Sun and Yeh [1990].

6.3 NUMERICAL EXAMPLES

To illustrate our method, we created a synthetic three-dimensional heterogeneous aquifer, which is contaminated with a single component NAPL. This 3-D aquifer was 40 m long in horizontal direction, 0.5 m in width, and 10 m deep in vertical direction. A fixed head of 100 m was specified at left-hand and right-hand boundaries whereas no flow was specified at top and bottom boundaries. No tracer was presented in the domain and all the boundaries before each partitioning tracer test. The aquifer was discretized into 1600 elements with a uniform element size of 0.5 m × 0.5 m × 0.5 m. The log-normal random fields of hydraulic conductivity, water saturation, and NAPL saturation fields were generated by a spectral method by Gutjahr [1989]. The geometric mean of K was 0.86 m/d with a variance of ln$K$ of 0.8. The geometric mean of the porosity of the aquifer was 0.4 and the variance of natural log of porosity is 0.02. The geometric mean of θ_n was 0.1 and the variance of $\ln \theta_n$ was 0.1. The water saturation, θ_w, was then obtained by subtracting θ_n from the generated porosity field. The geometric mean of the resultant θ_w was 0.291 and the variance of $\ln \theta_w$ was 0.057. These random fields were illustrated in Figure 6.1. The correlation scales for the hydraulic conductivity, porosity, and NAPL saturation fields was 3.5 m in the horizontal direction and 1.5 m in the vertical direction. Here, we assumed the transport process induced by a tracer test is convection dominated, and subsequently we neglected the local dispersion. As a result, the longitudinal and transversal dispersivities were set to 0.

Four fully penetrating, multilevel wells were placed vertically in the aquifer. Each well had nine observation ports and two injecting ports [see Figure 6.2]. Notice that the distances between wells are 10 m, which is about 2.86 times longer than horizontal correlation scale. Yeh and Liu [2000] showed that the "optimum" distance of wells is one correlation scale. We assumed these observation ports can measure both hydraulic heads and tracer concentrations. One direct measurement for each parameter was assumed known at x = 4.75m, y = 0.25m, z =

1.25m. Eight tracer tests during the hydraulic/tracer tomography were simulated. Prior to each tracer test, water was continuously injected into the aquifer through one of the eight injecting ports with a constant injection rate of 20 m^3/d. When the flow reached steady state, a partitioning tracer was then released through the same port into the aquifer for 1.5 days. The partitioning coefficients of the partitioning tracers were assumed to be known a priori with a value of 9.0. This value corresponds to TCE-Water partitioning coefficient value of a 4-mehtyl-2-pentanol [4M2P] tracer [Dugan et al., 2003] [they reported a value of 10.3 in their study]. The tracer concentrations at all the observation ports were recorded at every 0.3 days for 6 days. The steady state heads at these observation ports were also recorded. Totally, 288 steady state head measurement and 5184 tracer concentration measurements were collected.

From the head and tracer concentration measurements created by the hydraulic/tracer tomography survey, we used SSLE approach to estimate the heterogeneity of the aquifer. First the steady state head measurements were used to estimate hydraulic conductivity. The estimated *K* is showed in Figure 6.3.

Among 5184 tracer measurements, many of them had zero tracer concentration, which occurred when tracer did not reach or already passed the observation point. The zero-concentration measurements contained very limited information about the movement of tracer, therefore were removed them from the inverse processes to reduce computational burden. As a result, a total of 2350 tracer measurements were included. As being postulated previously, estimation of NAPL saturation fields from partitioning tracer tomography can be improved by our knowledge of heterogeneity of *K* and θ_w. To test this hypothesis and to illustrate the importance of our knowledge of *K* and θ_w fields on imaging NAPL saturations, four cases were considered. In case 1, we assumed that *K* and θ_w fields were perfectly known and in case 2 only *K* was known precisely. Estimated *K* from heads from hydraulic/tracer tomography was as our knowledge of the *K* field but θ_w was unknown in case 3. In case 4, both *K* and θ_w fields were assumed unknown. Case 1 represents an idealized scenario where only spatial varying NAPL saturation field is unknown. The objective of the case is to test the validity of our stochastic estimation technique for tracer tomography. Case 2 investigates the effect of variation in θ_w on θ_n estimation. Case 3 represented a more realistic scenario where hydraulic conductivity and θ_w were unknown and we used all the data available from hydraulic/tracer tomography to image the distribution of NAPL saturation. Case 3 also served as an example to test the improvement of *K* estimates from hydraulic tomography by adding observed tracer concentration data from the tracer tomography. Case 4 was served to test the advantage of using hydraulic/tracer tomography over using tracer tomography alone. In addition, to compare our joint inverse method with traditional geostatistical method, we assumed that NAPL saturation at the 36 observation points were measured and subsequently were used in a traditional kriging method to estimate NAPL distribution in the entire aquifer [case 5]. GEOEAS [http://www.epa.gov/ada/csmos/models/geoeas.html] was used to perform kriging estimation. Based on the parameters used for generating θ_n, a exponential variogram model with a sill of 0.1, horizontal range 3.5 m, vertical range 1.5 m was used.

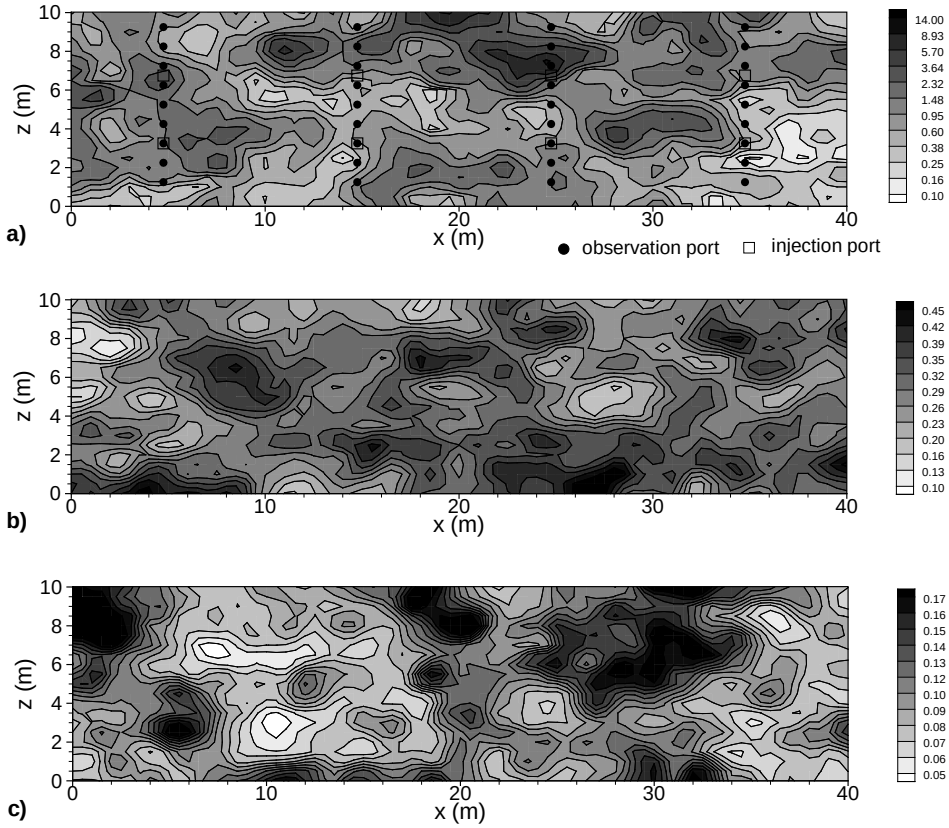


Figure 6.1: True Synthetic Fields: a] Hydraulic Conductivity [K, m/d]; b] Water Saturation [θ_w]; and c] NAPL Saturation [θ_n].

To evaluate the performance of our approach for all three cases, we plotted the scatter-plots for each case to illustrate the estimates, along with mean absolute error norm L1 and mean square error L2.

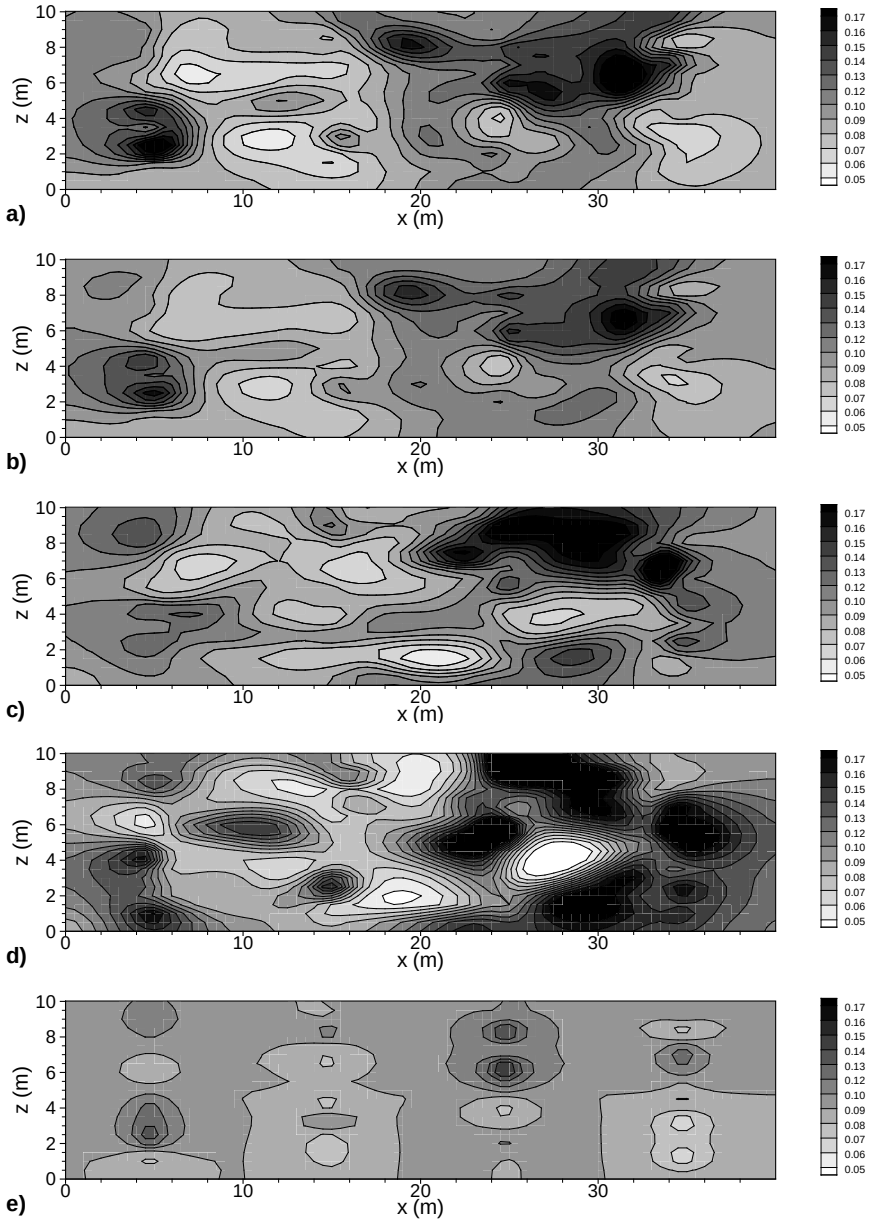


Figure 6.2: Estimated NAPL Saturation from: a] K and θ_w are Known [case 1]; b] K is Known but θ_w is Unknown [case 2]; c] Use Estimated K from Heads and θ_w is Unknown [case 3]; d] K and θ_w are Unknown [case 4]; and e] Kriging [case 5].

Figure 6.2a shows that with perfect knowledge of K and θ_w fields, tracer tomography in conjunction of our stochastic estimation technique can identify the locations and NAPL distribution very well with only four wells. With the exact knowledge of K field but θ_w field [Case 2], the tracer tomography with our stochastic method yields similar results as case 1 [see Figure 6.2b]. This result is of no surprise due to the implicit relation between θ_w and θ_n. But the estimated NAPL saturation field is smoother than that in case 1. Notice that in both cases 1

and 2, our approach is quite capable of detecting presence of NAPL and its saturation distribution between wells. This ability is beyond the capability of current technology.

Instead of knowing the exact *K* field, we used hydraulic tomography to estimate the *K* field first and then use tracer tomography to estimate θ_n distribution [Case 3]. Figure 6.2c shows that under this scenario, our approach still is capable of identifying patterns of major high saturation areas, but anomalies start to appear. Without any prior knowledge of *K* and θ_w [case 4], figure 6.2d shows that tracer tomography alone images some high NAPL areas correctly but creates several anomalies. The above results clearly suggest that knowing spatial variation of hydraulic conductivity is critical to estimation of NAPL distributions.

Finally, we illustrate the effectiveness of traditional direction sampling approach. Using direct measurements of NAPL from the four wells and the kriging method, the NAPL distribution was estimated and is shown in Figure 6.2e. In comparison with cases 1, 2, 3, and 4, this approach provided the worst estimate; it fails to capture high NAPL saturation areas between observation wells. Therefore, hydraulic/tracer tomography can provide more information than the traditional approach using the same well facility although additional hydraulic and tracer tests are involved. The scatter plots and L1, L2 norms in Figure 6.4a, b, c, d, and e also verified our findings.

Lastly, we focus on the effectiveness of hydraulic, tracer, and hydraulic/tracer tomography on estimation of hydraulic conductivity field. Figure 6.3a shows the estimated *K* using hydraulic head only while Figure 6.3b depicts the estimated *K* after adding tracer concentration data. Both figures showed very similar K patterns; the scatter plots and the L1 and L2 norms in Figure 6.5 for two cases are also very similar. In fact, the norms even showed that adding tracer concentration data slightly increases estimation errors. The results illustrated that adding tracer concentration data may not improve *K* estimation. The tracer data sets are merely redundant in our study cases. Intuitively, tracers are highly sensitivity to small-scale but high permeability paths, which are generally undetectable using hydraulic head data. Addition of tracer data should increase the resolution of hydraulic tomography. This intuition rests upon the fact that propagation of hydraulic head is a highly diffusive process while the tracer migration is a convection-dispersion process. The result reported here is likely attributed to the way we generated the heterogeneous aquifer: one K value for each finite element. This issue deserves further investigation.

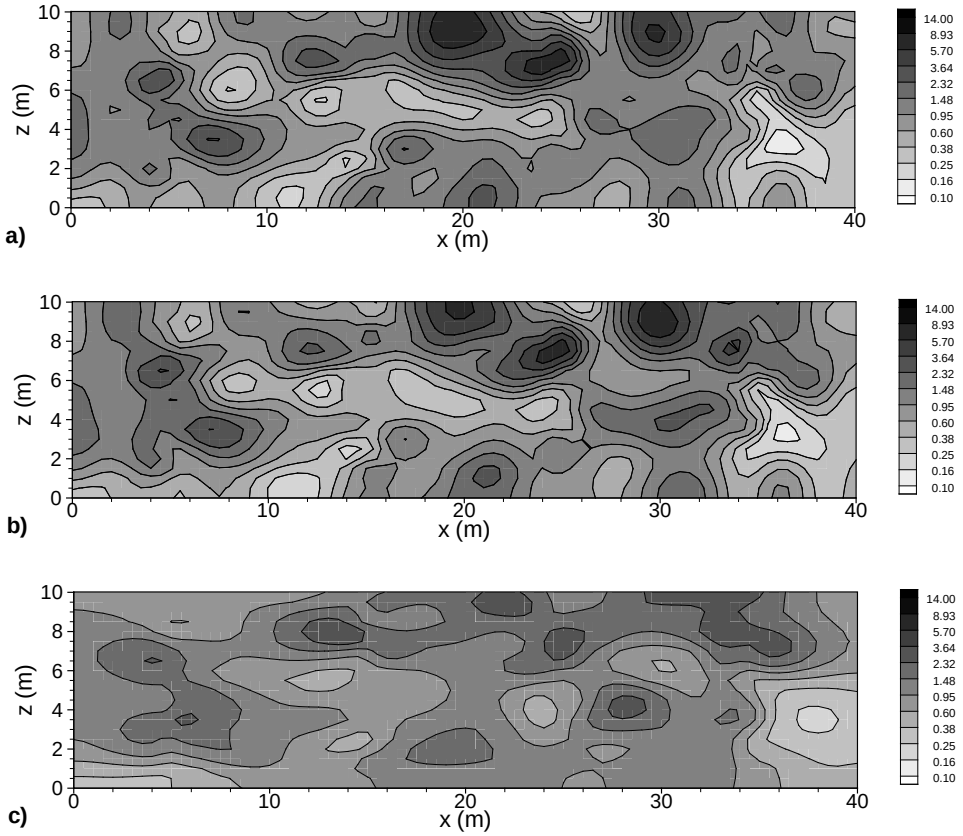


Figure 6.3: Estimated Hydraulic Conductivity from a] Heads; b] Heads and Tracer Data [case 3]; and c] Tracer Data [case 4].

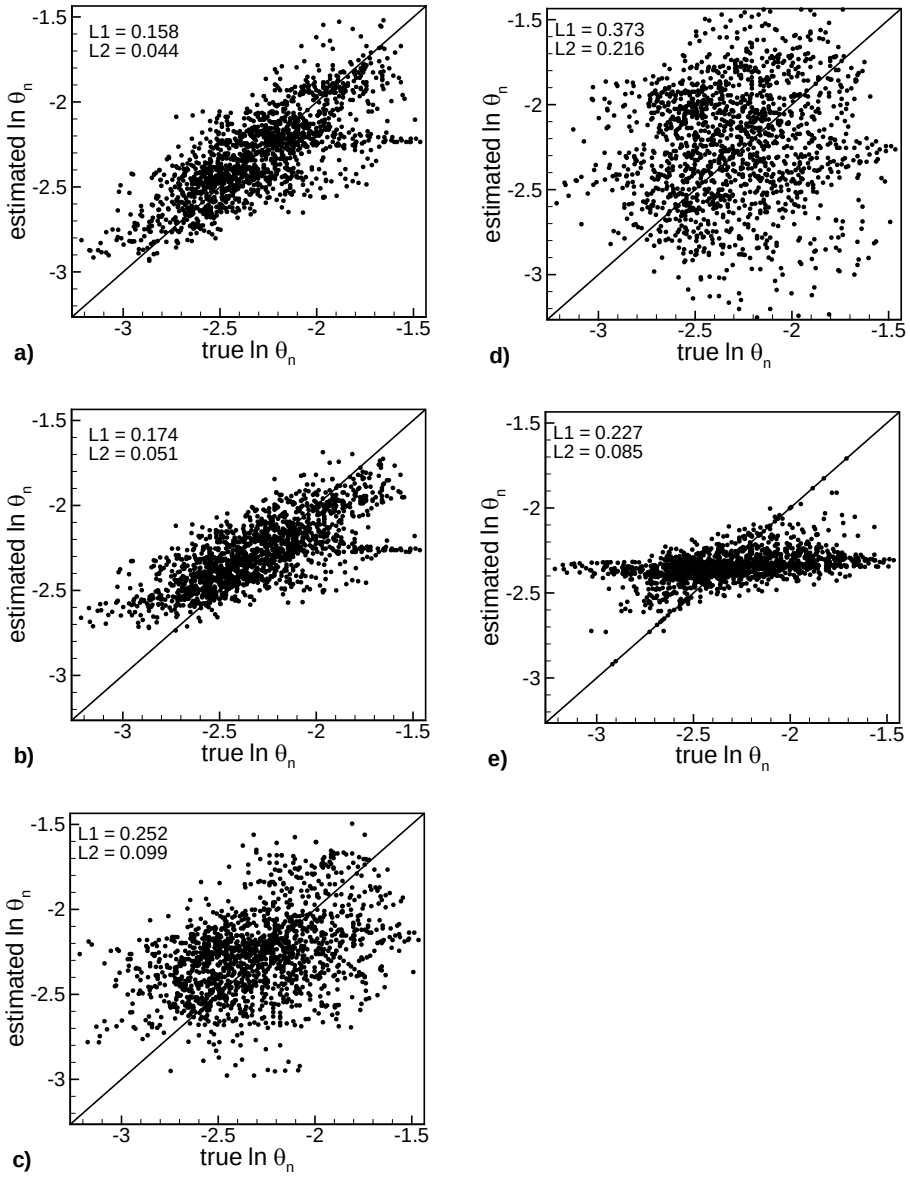


Figure 6.4: Scatter Plots of True ln NAPL Versus Estimated NAPL from: a] K and θ_w are Known [case 1]; b] K is Known but θ_w is Unknown [case 2]; c] Use Estimated K from Heads and θ_w is Unknown [case 3]; d] K and θ_w are Unknown [case 4]; and e] Kriging [case 5].

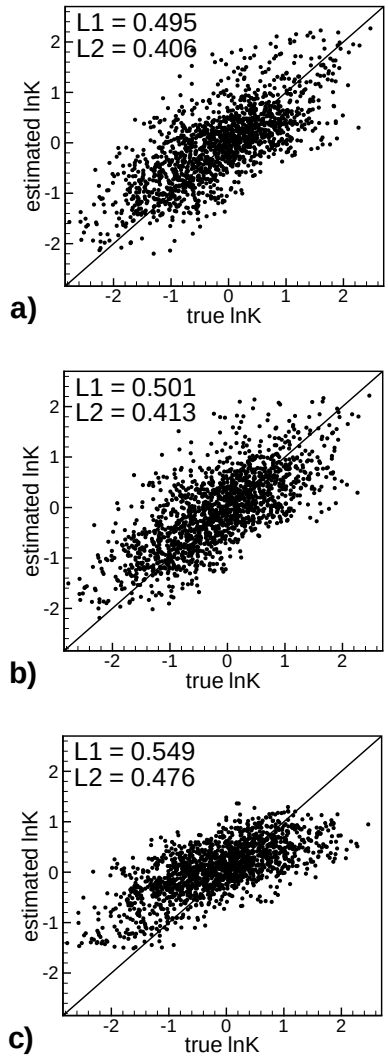


Figure 6.5: Scatter Plots of True ln*K* Versus Estimated ln*K* from a] Heads; b] Heads and Tracer Data [case 3]; and Tracer Data [case 4].

Figure 6.3c shows that using tracer data alone can capture general pattern of the true field but with much less details than that obtained from head data and head/tracer data. The scatter plots and L1, L2 norms in Figures 6.5a, b, and c also demonstrated our findings. This result further supports the conclusions by Li and Yeh [1999] that a concentration breakthrough measurement in a tracer test is a result of both convection and dispersion process. While local dispersion is ignored in our study—numerical dispersion exists, the convection transport of tracers is controlled by porosity and specific discharge [i.e., linear groundwater velocity]. The specific discharge is a function of both hydraulic gradient and the conductivity. Without knowledge of hydraulic head [in turn, its gradient], it is unlikely that a tracer test or tomography can provide accurate estimate of K field. Overall, we believe that head information is better than tracer data on mapping hydraulic conductivity field. In addition, tracer test data can reveal more detailed hydraulic heterogeneity at small scales, although this conjecture remains to be proven.

7. ANALYSIS OF TRACER TOMOGRAPHY USING TEMPORAL MOMENTS OF TRACER BREAKTHROUGH CURVES

7.1 INTRODUCTION

Effective remediation of sites contaminated with Dense Nonaqueous Phase Liquids [DNAPLs] relies on adequate characterization of the size, nature, and distribution of DNAPL source zones [NRC, 2004]. Because of the various human activities of the contaminated sites and complex subsurface environment, characterization of DNAPL source zones is a difficult task. Since 1980s, a variety of methods have been developed and applied to the DNAPL source zone characterization. Some of the methods yield localized point measurements of DNAPLs [i.e., core retrieval and analysis] at a limited number of locations, while others produce average saturation over a relatively large volume of geologic medium [i.e., partitioning tracer tests] [NRC, 2004]. Very few techniques have been developed to estimate detailed spatial distribution of DNAPL saturation in the subsurface.

The concept of tomography, originally developed in medical sciences for detailed imaging of a human body, has been recently introduced into subsurface hydrology. Tomographic surveys allow one to image an object from different angles and perspectives, and therefore provide high resolution images of the object. Utilizing the tomography concept, hydrologists have developed the hydraulic tomography [HT] technique to estimate detailed spatial variation of hydraulic properties of the subsurface. Specifically, in hydraulic tomography, a series of aquifer tests are sequentially conducted at different locations with head responses being collected from other locations during each test. Head responses from all the tests are then used in an inverse method to estimate hydraulic parameters [see Yeh and Liu, 2000; Zhu and Yeh, 2005; and Zhu and Yeh, 2006]. Although the ability of HT remains to be fully assessed under field conditions, results from sandbox experiments by Liu et al. [2002], Illman et al. [2007, 2008], and Liu et al. [2007] are encouraging. These studies showed that transient HT can identify not only the pattern of the heterogeneous hydraulic conductivity [K] field, but also the variation of specific storage [S_s] in the sandbox. Moreover, these estimated K and S_s fields from the HT sandbox experiments accurately predicted the drawdown evolution caused by a pumping test that was not used in the HT analysis. Likewise, a recent application of HT to a well field at Montalto Uffugo Scalo, Italy, produced an estimated transmissivity field that is deemed to be consistent with the geology of the site [Straface et al., 2007]. Bohling et al. [2007] and Li et al. [2007] also showed promising results of HT in their field experiments.

Success of HT has led Yeh and Zhu [2007] to the development of a joint hydraulic and tracer tomography for characterization of DNAPL source zones in the subsurface. Yeh and Zhu [2007] advocated that migration and distribution of DNAPLs are highly controlled by hydraulic heterogeneity of the subsurface. Thereby, HT is a must prior to a partitioning tracer test for detecting DNAPLs. HT delivers high-resolution imaging of the hydraulic heterogeneity which controls tracer migration path and thus the likely locations of DNAPLs. Conducting partitioning tracer tests in a tomographic fashion afterwards can facilitate detailed delineation of the spatial distribution of DNAPLs.

The proposed joint hydraulic/tracer tomography survey consists of a series of steady-state water injection and tracer tests in the DNAPL source zones at different locations, and monitoring head changes and tracer breakthrough curves [BTCs] at many observation locations. Using the

sequential successive linear estimator, the head responses are used first to provide detailed estimates of hydraulic conductivity spatial variation, and the BTCs from tracer tomography are then used to estimate spatial variations of water content and DNAPL content.

Generally speaking, estimating DNAPL spatial variation using discrete tracer concentration data in partitioning tracer tests imposes significant computational cost [e.g., Zhang and Graham, 2001, Yeh and Zhu, 2007]. This computational burden arises from the requirement of fine spatial and temporal discretizations of the advection-dispersion equation, which is solved repeatedly during the estimation procedure. Instead of using discrete concentration values at different times, many researchers have used temporal moments of tracer BTCs to estimate hydraulic properties or DNAPL content in the subsurface. Approaches based on the temporal moments solve equations that are functions of spatial coordinates only. Thus the approaches avoid time marching procedures and offer computational advantages over those based on the discrete concentration values. For instance, Jin et. al. [1995] used temporal moments of BTCs to estimate the average DNAPL saturation for 1D and 2D problems. Harvey and Gorelick [1995] developed temporal moment generating equations that directly solve temporal moments without solving the entire BTCs. Using the generating equations, James et. al., [1997] calculated the covariances of temporal moments and cross-covariances between temporal moments and parameters. James et. al. [2000] further used a Gauss-Newton minimization method to estimate DNAPL spatial variation from temporal moments. Cirpka and Kitanidis [2001] combined hydraulic heads and the first moments of tracer BTCs to infer hydraulic conductivity spatial distribution. Jawitz et. al., [2003] implemented partitioning tracer higher moments into binary models to estimate statistical parameters for NAPL source zones. In spite of these applications, the tracer moment approach has not been applied to tracer tomography.

In this work, the tracer temporal moment method is applied to the joint hydraulic/tracer tomography approach. The aim of the application is to reduce the computational cost associated with the joint approach such that the approach can be appealing for practitioners. In this study, we implement tracer temporal moments in sequential successive linear estimator [SSLE] to interpret the joint hydraulic/tracer tomography data for estimating DNAPLs. We first derive temporal moment equations for the transport of a partitioning tracer. We next apply a streamline upwind/Petrov-Galerkin formulation to reduce numerical oscillation due to discretization in solving convection dominated moment equations. Then, we incorporate the first temporal moment of tracer into the SSLE inverse algorithm, which requires evaluation of sensitivity of temporal moments to hydraulic conductivity, water content, and DNAPL content, along with calculation of covariance of moments and cross-covariance between moments and the above three parameters. Finally, we test the temporal moment method using the same synthetic cases in Yeh and Zhu [2007] and compare the computational cost. The results from temporal moments are also compared with those from inversion of discrete concentration data. The usefulness of temporal moments of tracer BTCs collected in a tomographic way in estimating hydraulic conductivity is also investigated.

7.2 METHODOLOGY

7.2.1 Temporal moment equations for a partitioning tracer

As mentioned previously, the joint hydraulic/tracer tomography survey involves a series of steady-state water injection and tracer tests in the DNAPL source zones at different locations, and monitoring head changes and BTCs at many observation locations. Interpreting this joint

hydraulic/tracer tomography survey requires both forward and inverse modeling efforts. The forward modeling solves steady state flow equation and tracer temporal moment equations, which are discussed below.

7.2.1.1 Governing equations

We assume that the steady state flow field induced by each tracer test in the joint hydraulic/tracer tomography is described by

$$\nabla \cdot [K(\mathbf{x})\nabla H] + Q = 0 \qquad [7.1]$$

subject to boundary conditions:

$$H\big|_{\Gamma_1} = H_1, \;\; [K(\mathbf{x})\nabla H]\cdot \mathbf{n}\big|_{\Gamma_2} = q_h, \qquad [7.2]$$

where H is the total head [L], $\mathbf{x}$ is the spatial coordinates [$\mathbf{x} = \{x_1, x_2, x_3\}$, [L], and x_3 represents the vertical coordinate and is positive upward]. The hydraulic conductivity field is denoted by $K[\mathbf{x}]$ [L/T]. In addition, Q is the source term [1/T]. In equation [16.2], H_1 is the prescribed total head at the Dirichlet boundary Γ_1, q_h is the specific flux [L/T] at the Neumann boundary Γ_2, and $\mathbf{n}$ is a unit vector normal to the boundary.

The transport of a partitioning tracer in the joint hydraulic/tracer tomography is assumed to be governed by the advection-dispersion-retardation equation:

$$\frac{\partial}{\partial t}(\theta_w c + K_N \theta_n c) = -\nabla \cdot (\mathbf{q}c) + \nabla \cdot (\theta_w \mathbf{D}\nabla c) + Q c_n N(\mathbf{x} - \mathbf{x}_a)(t - t_b) \qquad [7.3]$$

subject to following initial and boundary conditions:

$$c\big|_{t=0} = c_0, \quad c\big|_{\Gamma_1} = c_1, \;\; \mathbf{q}c - \theta \mathbf{D}\nabla c\big|_{\Gamma_2} = \mathbf{q}_s \qquad [7.4]$$

where c is the partitioning tracer concentration [M/L^3]; θ_w is the volumetric water content and θ_n is the volumetric DNAPL content. Notice that porosity θ is the sum of θ_w and θ_n; K_N is the partitioning coefficient, $\mathbf{q}$ is Darcy velocity vector [L/T], which is assumed to be time invariant and is calculated from equations [7.1] and [7.2]; $\mathbf{D}$ is dispersion tensor and is given as

$$D_{ij} = \frac{v_i v_j}{|\mathbf{v}|}(\alpha_L - \alpha_T) + \delta_{ij}(\alpha_T v + D_d^*) \qquad [7.5]$$

where α_L and α_T are the longitudinal and transverse dispersivities [L], respectively; v_i is the seepage velocity [L/T] defined as q_i / θ_w; $|\mathbf{v}|$ is the magnitude of the seepage velocity; δ_{ij} is the Kronecker delta function which equals unity if the indices are identical and zero otherwise, and D_d^* is the molecular diffusivity [L^2/T]. In addition, c_n is the concentration of the injecting tracer solution; $N(\mathbf{x} - \mathbf{x}_a)(t - t_b)$ is a step function which equals unity at location $\mathbf{x}_a$ at time earlier than t_b [the time when the partitioning tracer is released] and zero otherwise. c_0 is the initial tracer concentration, which is typically zero for a tracer test; c_1 is the prescribed concentration at the Dirichlet boundary Γ_1; q_s is tracer flux at the Neumann boundary Γ_2.

7.2.1.2 Temporal moment equations for a tracer BTC

The kth temporal moments of a tracer BTC at location $\mathbf{x}$ can be defined as

$$m_k = \int_{t=0}^{t=\infty} t^k c(\mathbf{x}, t)\, dt \qquad [7.6]$$

where k is the order of moment. Multiplying equation [7.3] by t^k, integrating over time, assuming $\mathbf{q}$ is time invariant, and substituting equation [7.6] into equation [7.3] yield moment-generating equation

$$(\theta_w + K_N\theta_n)t^k c\Big|_{t=0}^{t=\infty} - (\theta_w + K_N\theta_n)km_{k-1} = -\nabla\cdot(\mathbf{q}m_k) + \nabla\cdot\left(\theta_w\mathbf{D}\nabla m_k\right) + Qc_n N(\mathbf{x}-\mathbf{x}_a)(t-t_b)\int_{t=0}^{t=\infty} t^k dt \qquad [7.7]$$

For an impulse input tracer test, the initial concentration is commonly zero and the final concentration is also zero after the entire BTC passes. Therefore, the first term on the left hand side disappears. Then, equation [7.7] is simplified to [James, et. al., 2000, Cirpka and Kitanidis, 2001]

$$\nabla\cdot(\mathbf{q}m_k) - \nabla\cdot\left(\theta_w\mathbf{D}\nabla m_k\right) - Qc_n N(\mathbf{x}-\mathbf{x}_a)(t-t_b)\int_{t=0}^{t=\infty} t^k dt = (\theta_w + K_N\theta_n)km_{k-1} \qquad [7.8]$$

subject to the boundary conditions:

$$m_k\Big|_{\Gamma_1} = \int_{t=0}^{t=\infty} t^k c_1 dt\,, \qquad \mathbf{q}m_k - \theta\mathbf{D}\nabla m_k\Big|_{\Gamma_2} = \int_{t=0}^{t=\infty} t^k \mathbf{q}_s dt \qquad [7.9]$$

Notice that the temporal moment equation [7.8] eliminates time dependence in equation [7.3]. In theory, equations [7.8] and [7.9] can be applied to temporal moments of any order. However, our study will focus only on the zeroth moment m_0 [mass under the BTC] and first moment m_1. Notice that m_1/m_0 is mean arrival time of the BTC. Setting k=0 gives the zeroth moment equation,

$$\nabla\cdot(\mathbf{q}m_0) - \nabla\cdot\left(\theta_w\mathbf{D}\nabla m_0\right) - Qc_n N(\mathbf{x}-\mathbf{x}_a)(t-t_b)\int_{t=0}^{t=\infty} dt = 0 \qquad [7.10]$$

Setting k=1 gives the first moment equation,

$$\nabla\cdot(\mathbf{q}m_1) - \nabla\cdot\left(\theta_w\mathbf{D}\nabla m_1\right) - Qc_n N(\mathbf{x}-\mathbf{x}_a)(t-t_b)\int_{t=0}^{t=\infty} tdt = (\theta_w + K_N\theta_n)m_0 \qquad [7.11]$$

In order to obtain the first moment, equation [7.10] must be solved first for the zeroth moment and the result is then used in equation [7.11] to solve for the first moment.

7.2.2 Numerical method for moment equations

The moment generating equation [7.8] is in the form of steady state convection dispersion equation. Application of the traditional Galerkin finite element method to the moment equations can suffer from oscillations if these equations are convection dominated [i.e., Peclet number is higher than 1]. To reduce the oscillations, we use a Streamline-Upwind Petrov-Garlekin [SUPG] method which in principle adds an artificial dispersion [also called stabilization factor], which is proportional to the velocity, to the streamline direction proportional to the velocity to reduce the oscillations [Brooks and Hughs, 1982]. While the general idea of implementing this method is the same, the stabilization factor varies for different applications. Equivalently, the convection dispersion equation [7.8] can remain as is, while the weighting function of the finite element method is perturbed. More specifically, we use

$$W_i = N_i + \tau^* q\cdot\nabla N_i \qquad [7.12]$$

as the weighting function, where the value of τ^* was chosen elementwise as

$$\tau^* = \frac{\bar{\varepsilon} q_x L_x + \bar{\eta} q_y L_y + \bar{\varsigma} q_z L_z}{2(q_x^2 + q_y^2 + q_z^2)} \qquad [7.13]$$

where L_x, L_y, and L_z denote the element length in x-, y-, and z-directions, respectively. Moreover $\bar{\varepsilon}$ is defined as

$$\bar{\varepsilon} = \begin{cases} \dfrac{\alpha_\varepsilon}{3} & \textbf{if } |\alpha_\varepsilon| < 0.5 \\ \coth \alpha_\varepsilon - \dfrac{1}{\alpha_\varepsilon} & \textbf{otherwise} \end{cases} \quad \text{where } \alpha_\varepsilon = \frac{q_x L_x}{2 d_L} \qquad [7.14]$$

The value of $\bar{\eta}$ and $\bar{\varsigma}$ are defined similarly.

For implementation of the convection dispersion solver, we use the Diffpack framework [Langtangen, 2003], which is a collection of object-oriented C++ libraries for solving partial differential equations. In addition to different stabilization techniques, such as those defined in equations [7.12] through [7.14], Diffpack also has an extensive collection of different element types and iterative solvers of linear systems. The resulting convection dispersion solver is programmed as an efficient and flexible plug-in, which fits into an in-house FORTRAN code for hydraulic/tracer tomography.

7.2.3 Inverse Method

7.2.3.1 Implementation of temporal moments of tracer in SSLE inverse method

The sequential successive linear estimator [SSLE] is a geostatistical inverse approach that conceptualizes the parameter fields to be estimated as spatial stochastic processes, and seeks effective parameter fields conditioned on available state information [i.e., measurements of aquifer responses] and parameter measurements. The estimator utilizes statistical moments of parameter fields [i.e., mean and variance] as a priori information and uses partial differential equations of physical process to relate state information to parameter fields. The SSLE has been extensively discussed and applied to a variety of inverse problems in hydrogeology and geophysics. For example, Yeh and Liu [2000] used SSLE to estimate saturated K using steady state head information. Zhu and Yeh [2005, 2006] applied SSLE to estimate K and S_s using transient head and head moment information. Most recently, Yeh and Zhu [2007] extended SSLE to jointly interpret head and tracer concentration data from hydraulic/tracer tomography surveys. Liu and Yeh [2004], Yeh et al. [2002] and Yeh et al. [2006] applied it to analysis of electric resistivity tomography. In this paper, we apply SSLE to interpret tracer tomography by using tracer first temporal moment instead of using tracer concentration measurements directly.

Here, we will only briefly present the SSLE approach. Similar to Yeh and Zhu [2007], the natural logarithms of K, θ_w, and θ_N fields are treated as stochastic processes, Y_l [i.e., $Y_1 = \ln K$; $Y_2 = \ln \theta_w$; $Y_3 = \ln \theta_N$], which are represented by a mean $\bar{Y}_l$ and perturbation y_l component, such that $Y_l = \bar{Y}_l + y_l$. Likewise, the first tracer moment is conceptualized as a stochastic process with a mean $\bar{m}_1$ and a perturbation ε, such that $m_1 = \bar{m}_1 + \varepsilon$. Note that Cirpka and Kitanids [2001] reported that the sensitivities of the zeroth moment to parameters were essentially zero, resulting in very limited information about the spatial variability of parameters. As a result, we will focus on using the first moment for the DNAPL estimation.

The SSLE starts with the classic cokriging method which uses direct measurements of the parameter and measurements of the first moment at some sample locations to obtain a linear estimate of the parameters in the entire domain, that is

$$\hat{y}_l = \boldsymbol{\alpha} y_l^* + \boldsymbol{\beta}\boldsymbol{\varepsilon}^* \qquad [7.15]$$

where y_l^* and $\boldsymbol{\varepsilon}^*$ are $N_y \times 1$ and $N_\varepsilon \times 1$ vectors of measurements of parameters and measurements of the first moments, respectively; N_y is the number of parameter measurements and N_ε is number of the first moment measurements obtained from one tracer test. $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $N_y \times N$ and $N_\varepsilon \times N$ weight matrices and N is the total number of finite element cells in the simulation domain. The weights represent contributions of each measurement to the estimates. The weights are calculated based on unconditional spatial moments [means and covariances] of Y_l, the spatial covariances of m_1, and spatial cross covariances between m_1 and Y_l. The statistical moments of Y_l are input parameters. Using the moments of Y_l, statistical moments of m_1 are subsequently determined by a first order approximation.

Cokriging is a linear estimator. Therefore, it cannot fully exploit the usefulness of m_1 because the relation between m_1 and Y_l is not linear. SSLE uses an iterative approach to include the nonlinear relation and to subsequently improve the estimates. That is, the estimates of parameter fields at rth iteration $\hat{y}_l^r$ are used in the forward models [equations [7.10] and [7.11]] to approximate the conditional mean of m_1. The differences between observed first moments ε^* and calculated first moments ε^r are subsequently used in conjunction with associated weights $\boldsymbol{\omega}^r$ to improve the estimate $\hat{y}_l^{r+1}$,

$$\hat{y}_l^{r+1} = \hat{y}_l^r + \boldsymbol{\omega}^r(\boldsymbol{\varepsilon}^* - \boldsymbol{\varepsilon}^r) \qquad [7.16]$$

These weights are obtained by solving a system of equations that involve conditional covariances and cross-covariances, which are derived from the previous iteration. After improved estimates are obtained, the conditional covariance of Y_l, $\mathbf{R}_{y_l y_l}^{r+1}$, is updated through the following equation

$$\mathbf{R}_{y_l y_l}^{r+1} = \mathbf{R}_{y_l y_l}^{r} - \boldsymbol{\omega}^r \mathbf{R}_{\varepsilon y_l}^{r}. \qquad [7.17]$$

where $\mathbf{R}_{\varepsilon y_l}^{r}$ denotes the conditional cross-covariance between Y_l and the m_1 measurements at iteration r. The updated conditional covariance is then used in the next iteration. The aforementioned steps continue until the improvement in the estimates diminishes to a prescribed value. Specifically, SSLE stops when both the variances of the estimated parameter fields and the differences between the observed and calculated first moments stabilize. SSLE can simultaneously include all the measurements collected from a tracer tomography survey jointly; however, this can lead to numerical and computer memory problems because of an extremely large system of equations [Hughson and Yeh, 2000]. To avoid this problem, the m_1 measurements collected from a tracer tomography survey are added into the inverse process sequentially, such as one tracer test is followed by another tracer test. By adding secondary information in a sequential way, the procedure of successive improvement as described above is used for every tracer test.

7.2.3.2 Calculation of Sensitivities

As indicated above, to use tracer moments in SSLE, one needs to calculate the covariance of m_1 measurements and cross covariances between m_1 measurements and Y_l. They are evaluated by a first order approximation as described in Yeh and Zhu [2007]. The first order approximation method requires the sensitivities of m_1 to Y_l. While the process of calculating covariance and cross covariance is essentially the same as that in Yeh and Zhu [2007], evaluation of sensitivities is different because the equations describing the relation between the m_1 to Y_l are different from the equations relating tracer concentrations to Y_l. We use the adjoint state method to calculate the sensitivity of the first moments to parameters. The purpose of using the adjoint state method is to improve the computational efficiency. The adjoint state method is most efficient when the number of measurements is far smaller than the number of parameters. Detailed description of adjoint state method can be found in Sykes et al. [1985], Sun [1994], and Cirpka and Katanidis [2001]. Here, we briefly describe the adjoint state method for calculating sensitivity of m_1 to $\ln K$, $\ln\theta_w$, and $\ln\theta_n$. The adjoint state method can be described as three steps. The first step is to solve for mean head $\bar{H}$ using equations [16.1] and [16.2] and to solve mean zeroth and first tracer moments $\bar{m}_0$ and $\bar{m}_1$ using equations [16.10] and [16.11] with associated boundary conditions. The second step is to solve three adjoint equations for a first moment measurement m_1^* at location $\mathbf{x}_e$ to obtain three adjoint states, ϕ_1, ϕ_0, and ϕ_h:

$$-\bar{q}_i\nabla\phi_1 - \nabla\cdot(\bar{\theta}_w D_{ij}\nabla\phi_1) = -\frac{\partial P}{\partial m_1} \qquad [7.18]$$

$$-\bar{q}_i\nabla\phi_0 - \nabla\cdot(\bar{\theta}_w D_{ij}\nabla\phi_0) = -\frac{\partial P}{\partial m_0} + \phi_1\bar{\theta}_w + \phi_1 K_N\bar{\theta}_n \qquad [7.19]$$

$$\nabla\cdot(\bar{K}\nabla\phi_h) = -\nabla\cdot(\bar{K}\phi_0\nabla\bar{m}_0) - \nabla\cdot(\bar{K}\phi_1\nabla\bar{m}_1) - \frac{\partial P}{\partial H} \qquad [7.20]$$

subject to the following boundary conditions

$$\phi_h\big|_{\Gamma_1} = 0, \quad \phi_k\big|_{\Gamma_1} = 0, \ (\phi_k\bar{\mathbf{q}} + \mathbf{D}\nabla\phi_k)\cdot n\big|_{\Gamma_2} = 0$$

$$(-\bar{K}\nabla\phi_h - \bar{K}\phi_0\nabla\bar{m}_0 - \bar{K}\phi_1\nabla\bar{m}_1)\cdot n\big|_{\Gamma_2} = 0 \qquad [7.21]$$

In the above equations $P = m_1\delta(\mathbf{x}-\mathbf{x}_e)$ represents the first moment measurement at location $\mathbf{x}_e$ and δ is Kronecker's delta function, which equals one if $\mathbf{x}=\mathbf{x}_e$ and zero otherwise; k is the adjoint state index, equals either 0 or 1. The third step is then to calculate the sensitivities of m_1^* to parameters at location j by:

$$\frac{\partial m_1}{\partial \ln K_j} = \int_{\Omega_j}(-\bar{K}\nabla\phi_h\nabla\bar{H} - \phi_0\nabla\bar{m}_0\bar{K}\nabla\bar{H} - \phi_1\nabla\bar{m}_1\bar{K}\nabla\bar{H})d\Omega \qquad [7.22]$$

$$\frac{\partial m_1}{\partial \ln\theta_{wj}} = \int_{\Omega_j}(\bar{\theta}_w D_{ij}\nabla\phi_0\nabla\bar{m}_0 + \bar{\theta}_w D_{ij}\nabla\phi_1\nabla\bar{m}_1 - \phi_1\bar{\theta}_w\bar{m}_0)d\Omega \qquad [7.23]$$

$$\frac{\partial m_1}{\partial \ln\theta_{nj}} = \int_{\Omega_j} -\phi_1 K_N\bar{\theta}_n\bar{m}_0 d\Omega \qquad [7.24]$$

where Ω_j is the sub-volume of location *j*. In the adjoint state method, equations [7.1], [7.2], [7.11], and [7.12] only need to be solved once while equations [7.18] to [7.21] are solved once for each m_1 measurement.

7.3 SYNTHETIC CASE STUDY

7.3.1 Comparison with Inversion using tracer concentrations

To test the computational efficiency of the proposed inverse method for tracer tomography, we apply the method to the synthetic quasi-three dimensional confined aquifer that was used by Yeh and Zhu [2007]. The synthetic aquifer is 40 m long in horizontal direction, 0.5 m in width, and 10 m in vertical direction. The parameter fields of *K*, θ_w, and θ_n in the heterogeneous aquifer are assuming to be lognormal and are generated using a spectral method. The *K* field has a geometric mean of 0.86 m/d with variance of 1.68 [0.8 for ln*K*]; the θ_w field has a geometric mean of 0.291 with variance of 0.004 [0.057 for $\ln\theta_w$]; and the θ_n field has a geometric mean of 0.1 with variance of 0.012 [0.1 for $\ln\theta_n$]. The longitudinal and transverse dispersivities are constant values of 0.05 m and 0.01 m, respectively. The molecular diffusivity is 10^{-7} m^2/d everywhere. The partitioning coefficient is equal to 9.0. The aquifer is discretized in to 1600 elements with a uniform size of 0.5 m × 0.5 m × 0.5 m. The discretization resulted in a Peclet number about 10.

A hydraulic/partitioning tracer survey is applied to the synthetic aquifer. To conduct HPTT, four fully penetrating wells are emplaced in the aquifer. Each well has nine pressure and tracer dual-purpose sampling ports and two injection ports [Figure 7.3a]. The distance between two adjacent wells is 10 m. The survey consists of eight tracer injection tests. Each test starts with the simulation of a steady state flow field resulting from a continuous injection of water at a rate of 20 m^3/d through one of the eight injection ports. Then the migration of a slug of conservative and partitioning tracers is then simulated. The survey collects 288 steady state head measurements, 288 conservative tracer BTCs, and 288 partitioning tracer BTCs from eight tracer injection tests. The first moments can be either calculated from equation [7.6] after entire BTCs are calculated from equations [7.3] and [7.4] or directly calculated through solving equations [7.10] and [7.11]. In reality, the only way to obtain temporal moments is through BTCs, which however are subject to truncation errors [Luo et al., 2007] as well as noise. As the primary goal of this work is to demonstrate the computational efficiency of employing temporal moments for interpreting BTCs from partitioning tracer tomography, we use equations [7.10] and [7.11] to obtain the zeroth and first moments and ignore effects of the potential truncation error on the BTC temporal moment estimates. The process of interpreting hydraulic/partitioning tracer tomography proposed by Yeh and Zhu [2007] starts with estimating *K* using steady state head data collected from hydraulic/tracer tomography; then, conservative tracer concentration data are used to estimate θ_w and to improve the estimation of *K*. Finally, the partitioning tracer concentration data collected from partitioning tracer tomography are used to estimate θ_n and to again improve the estimate of *K*. In this work, we follow the same steps of the above interpretation process; but during the last step of the interpretation, we use partitioning tracer first moments instead of discrete concentration data. The total CPU time for the inverse simulation of partitioning tracer first moment measurements is about 50 minutes on a single PC [Pentium 4 2.8GHz], while the CPU time for the inverse modeling of partitioning tracer

concentration measurements is about 200 minutes using an eight node PC cluster with the same processors [Pentium 4 2.8GHz] . The saving of computational cost is dramatic. More importantly, using the moment approach makes the inverse modeling of the joint hydraulic/tracer tomography feasible using a single PC and thus more attractive for a rapid analysis of the HPTT survey.

Figure 7.1 compares the estimated θ_n field from discrete concentration data [Figure 7.1b] with the true field [Figure 7.1a] and that [Figure 7.1c] using the first moment data. These figures show that, similar to the estimate from discrete concentration data, the estimates from first moments reveal the general pattern of the spatial variation of synthetic θ_n field but the pattern appears to be smoother than the field estimated from the concentration data. Figure 7.2 presents scatterplots of the estimated and true θ_n fields from these two approaches. The error norms L1 and L2 of the estimated θ_n field and the correlation ρ between the estimated and the true θ_n fields [listed in Figure 7.2] show that the first moments can produce estimates of quality comparable to those from concentration data. L1 and L2 norms are defined as:

$$L1 = \frac{1}{N}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \qquad \text{and} \qquad L2 = \frac{1}{N}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2, \qquad [7.25]$$

And the correlation ρ is defined as

$$\rho = \frac{\frac{1}{N}\sum_{i=1}^{N}(Y_i - \mu_Y)(\hat{Y}_i - \mu_{\hat{Y}})}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - \mu_Y)^2 \cdot \frac{1}{N}\sum_{i=1}^{N}(\hat{Y}_i - \mu_{\hat{Y}})^2}} \qquad [7.26]$$

Where μ_Y and $\mu_{\hat{Y}}$ are means for true and estimated log parameter fields, respectively. However, the estimated field based on the first moment approach has higher values of L1 and L2, and smaller correlation comparing with those based on the concentration-time data, indicative of loss of some information using the first moment of the BTCs.

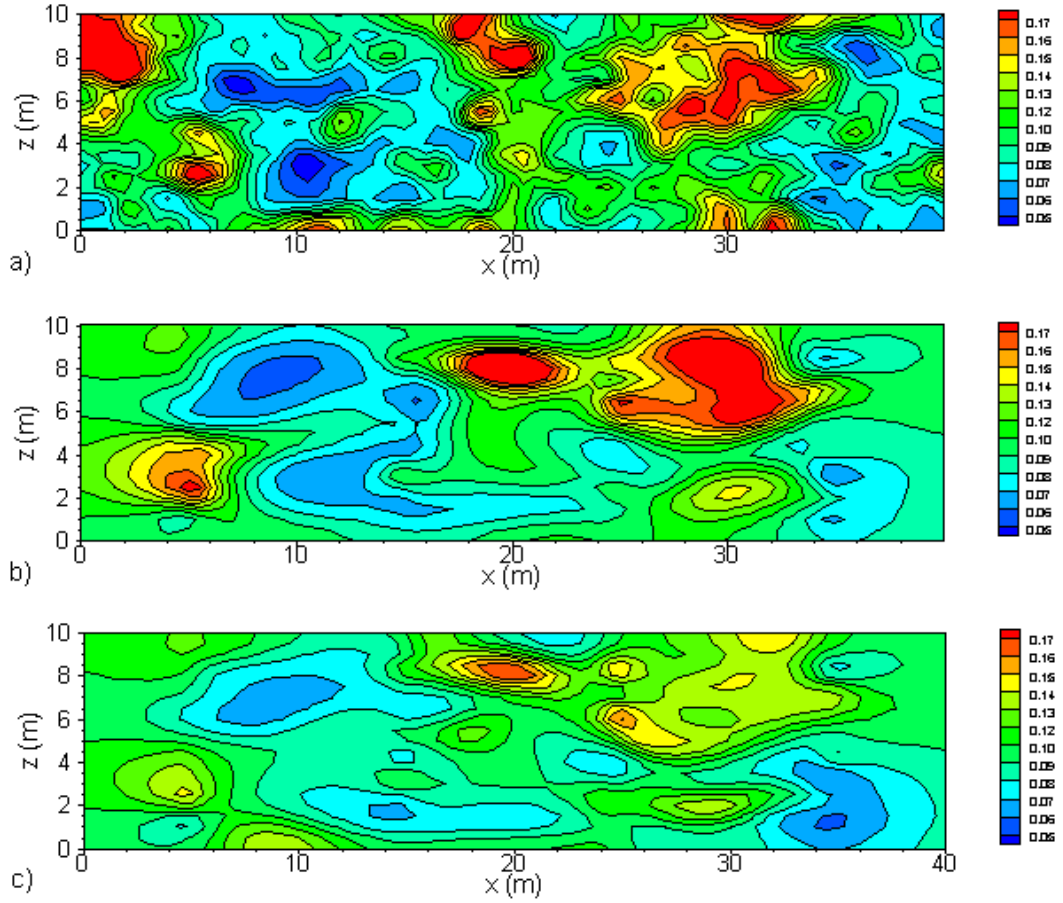


Figure 7.1: Comparison between true and estimated DNAPL content fields. [a] true field; [b] estimated from discrete concentration data; [c] estimated from tracer first moment data

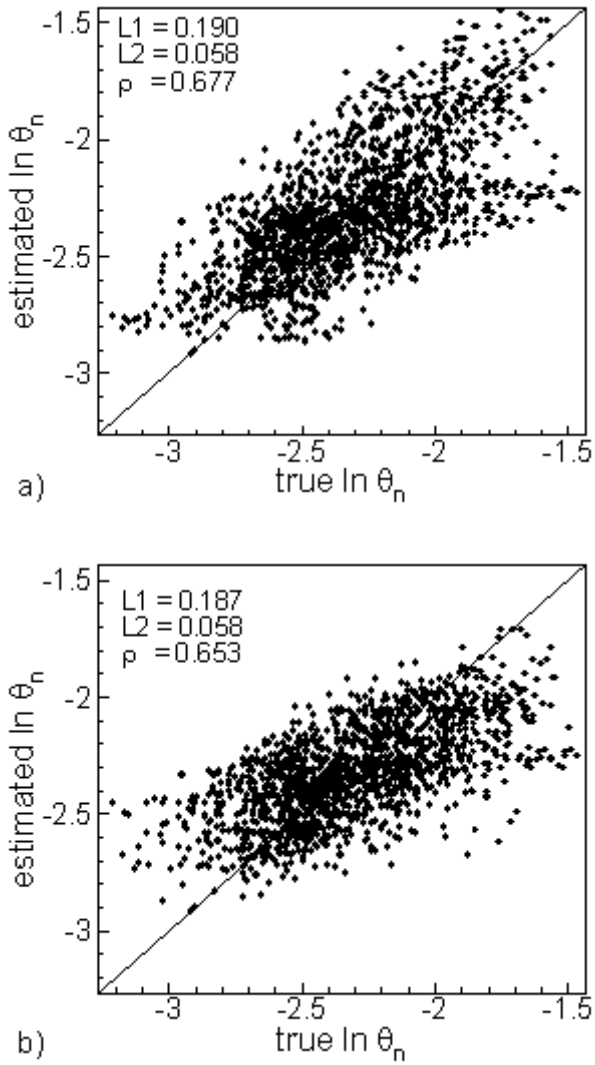


Figure 7.2: Scatter plot of true DNAPL content field versus estimate fields from: [a] discrete concentration data; [b] tracer first moment data.

7.3.2 Tracer first moments for estimating hydraulic conductivity

While tracer arrival-time data generally are considered to be less informative than head data in estimating the *K* field [Li and Yeh, 1999; Cirpka and Kitanidis, 2001], this result has not been confirmed for tomographic surveys. Tomography provides significant amount of arrival time information originated from different flow scenarios, which can be effective in estimating hydraulic conductivity in cases where the high density head data are not available. Nevertheless, Cirpka and Kitanidis [2001] noticed that using arrival time alone may cause numerical instability during the estimation of the hydraulic conductivity field. To test the usefulness of tracer arrival time data collected in a tomographic survey for estimating *K* field, the conservative tracer arrival

time data were used to estimate K and θ_w simultaneously as the conservative tracer arrival time is also affected by θ_w.

Figure 7.3a shows the true hydraulic conductivity field of the synthetic aquifer, which is used to compare with the estimated field based on the head information [Figure 16.3b] and that based on the tracer first moment [Figure 7.3c]. A visual comparison of these figures seems to suggest that the estimated spatial variation of K from head data is a closer approximation of the true in comparison with that from arrival time of conservative tracer. To confirm this finding, scatter plots of the true and estimated fields based on the head and that based on the first moment are illustrated in Figure 7.4a and 7.4b, respectively, along with the statistical measures [L1, L2, and R]. The results substantiate that the estimated ln$K$ field from the head is clearly superior to that estimated field based on the tracer alone, in agreement with the conclusions by Li and Yeh [1999] and Yeh and Zhu [2007]. Notice that using partitioning tracer arrival time would involve additional unknowns [θ_n] and thus it is expected that it will not yield better results than the conservative tracer arrival times.

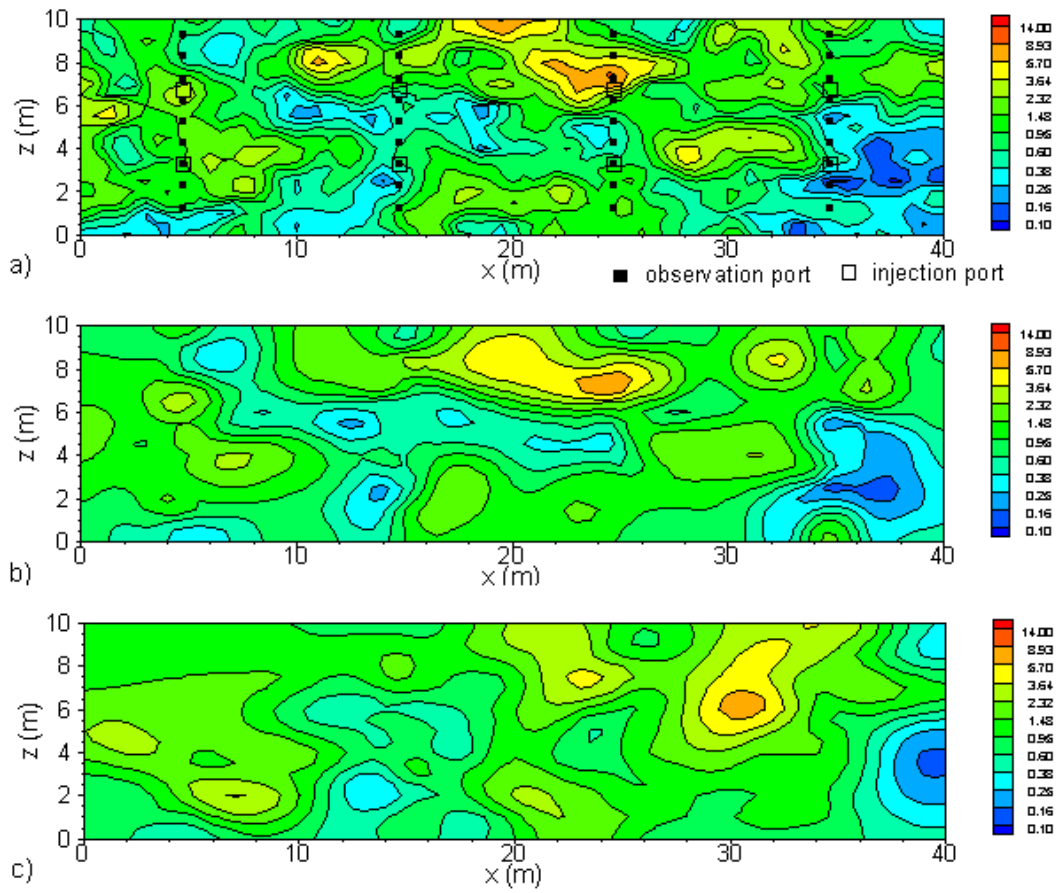


Figure 7.3: Comparison between true and estimated hydraulic conductivity fields: [a] true field; [b] estimated from head data; [c] estimated from tracer first moment data.

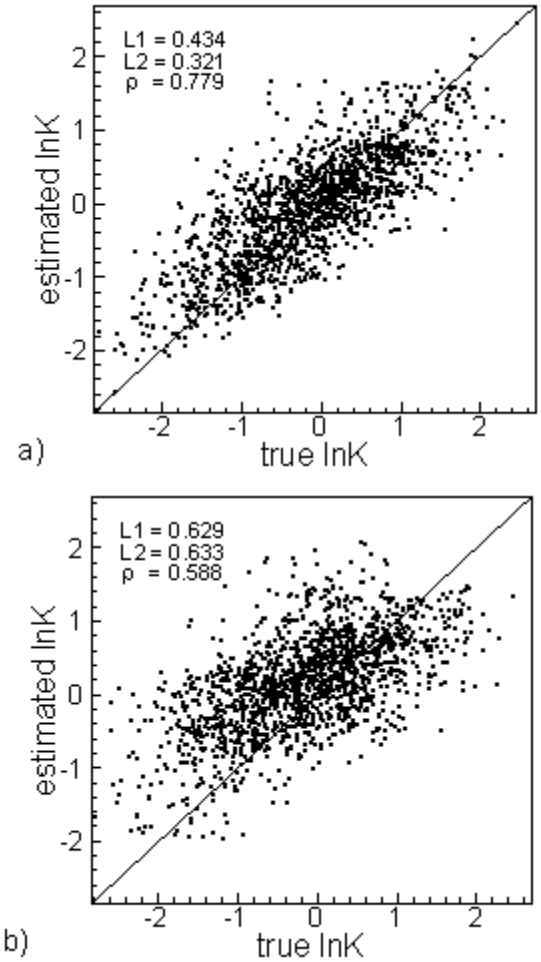


Figure 7.4: Scatter plot of true hydraulic conductivity field versus estimated field from [a] head data; [b] tracer first moment data.

7.3.3 Cross Correlation of first temporal moment of tracer to parameters

To understand how tracer temporal moments can be used to estimate different parameters, we calculate the cross correlation of the first temporal moments of partitioning tracer BTCs to selected parameters [i.e., K, θ_w, and θ_n]. The cross correlation between the first moment perturbation ε at location $\mathbf{x}_i$ and a parameter perturbation y_l at location $\mathbf{x}_j$, $C_{\varepsilon y}(\mathbf{x}_i,\mathbf{x}_j)$, is calculated through the following equation:

$$C_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_j) = \frac{R_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_j)}{\sqrt{R_{\varepsilon\varepsilon}(\mathbf{x}_i,\mathbf{x}_i)R_{y_l y_l}(\mathbf{x}_j,\mathbf{x}_j)}} \qquad [7.27]$$

Where $R_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_j)$ is the cross covariance between ε and y_l at locations $\mathbf{x}_i$ and $\mathbf{x}_j$; $R_{\varepsilon\varepsilon}(\mathbf{x}_i,\mathbf{x}_i)$ and $R_{y_l y_l}(\mathbf{x}_j,\mathbf{x}_j)$ are variances of ε at $\mathbf{x}_i$ and y_l at $\mathbf{x}_j$, respectively. $R_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_j)$ and $R_{\varepsilon\varepsilon}(\mathbf{x}_i,\mathbf{x}_i)$ are calculated through a first order approximation similar to that described in Yeh and Zhu [2007]:

$$R_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_j)=\sum_{m=1}^{N}J_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_m)R_{y_l y_l}(\mathbf{x}_j,\mathbf{x}_m)$$

$$R_{\varepsilon\varepsilon}(\mathbf{x}_i,\mathbf{x}_i)=\sum_{l=1}^{3}(\sum_{m=1}^{N}J_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_m)R_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_m)) \qquad [7.28]$$

Where $J_{\varepsilon y_l}(\mathbf{x}_i,\mathbf{x}_m)$ is the sensitivity of the first moment at location $\mathbf{x}_i$ with respect to the parameters y_l at locations $\mathbf{x}_m$; $R_{y_l y_l}$ is specified a priori at the first iteration and is evaluated through equation [7.17] at subsequent iterations. SSLE is an iterative estimator and $R_{\varepsilon y}$ and $R_{\varepsilon\varepsilon}$ are revaluated at each iteration, so are the cross correlations. Here we will examine only the correlation at the first iteration. In this analysis, one of the eight tracer tests in the hydraulic/tracer tomographic survey in the synthetic aquifer is selected for this analysis. The tracer/water injection location is selected at x=14.75 m, z= 7.25 m and an observation port is located at x= 24.75 m, z = 4.25 m. The spatial statistics [i.e., mean, correlation scales] for the 'true' parameter fields are used for the calculation of the cross correlations, except the variances where are assumed to be the same for all three parameter fields. For comparison, the cross correlation between a steady state head at the observation port and *K* is also calculated.

Figures 7.5a, 7.5b, and 7.5c show the cross correlations between the perturbation of the first moment at the observation location and perturbations of ln*K*, $\ln\theta_w$, and $\ln\theta_n$ throughout the aquifer, respectively. In all these figures, the areas with very low correlation values [i.e., the absolute value less than 0.1] are blanked and velocity vectors of the steady-state flow induced by the injection are plotted. These figures demonstrate that the correlation between the first moment and three parameters exhibit similar pattern. That is, the tracer first moment is only correlated to parameters along a narrow stripe between the injection location and the BTC measurement location, which follows the water flow line indicated in the figures. The area with absolute correlation value greater than 0.3 between the first moment and *K* and between the first moment and θ_n is confined to a small region along the stripe. The highest absolute correlation value for *K* is close to 0.4 and for θ_n is close to 0.5. On the other hand, the correlation between the first moment and θ_w is small with the maximum absolute value less than 0.15. Figure 7.5d shows the cross correlation between a steady head at the observation location and *K* during the same injection test. It is apparent that the high correlation value covers much of the downstream area of the observation location and most of area left of the injection location. The result is consistent with that by Wu et. al. [2005]. A comparison between Figures 7.5a and 7.5d indicates that a steady head measurement at a sampling location is correlated to *K* field over a much larger area than a tracer moment measurement at the same location. Interestingly, Figure 7.5a shows a very small region with absolute correlation value greater than 0.4 whereas Figure 7.5d shows no region with absolute correlation higher than 0.4. These results explain the reason why hydraulic tomography can yield high resolution mapping of *K* distribution over a large area and also why tracer data can complement the estimate of *K* from hydraulic tests. Note that the velocity vectors are plotted merely for showing the flow direction and are not scaled to the magnitude of flow velocity. These figures also indicate that, during a water/tracer injection test, the first moment is negatively related to *K* and positively related to θ_w and θ_n. That is, the smaller the first moment

[earlier arrival time] is, the greater the K value but smaller θ_w and θ_n values are in the stripe.

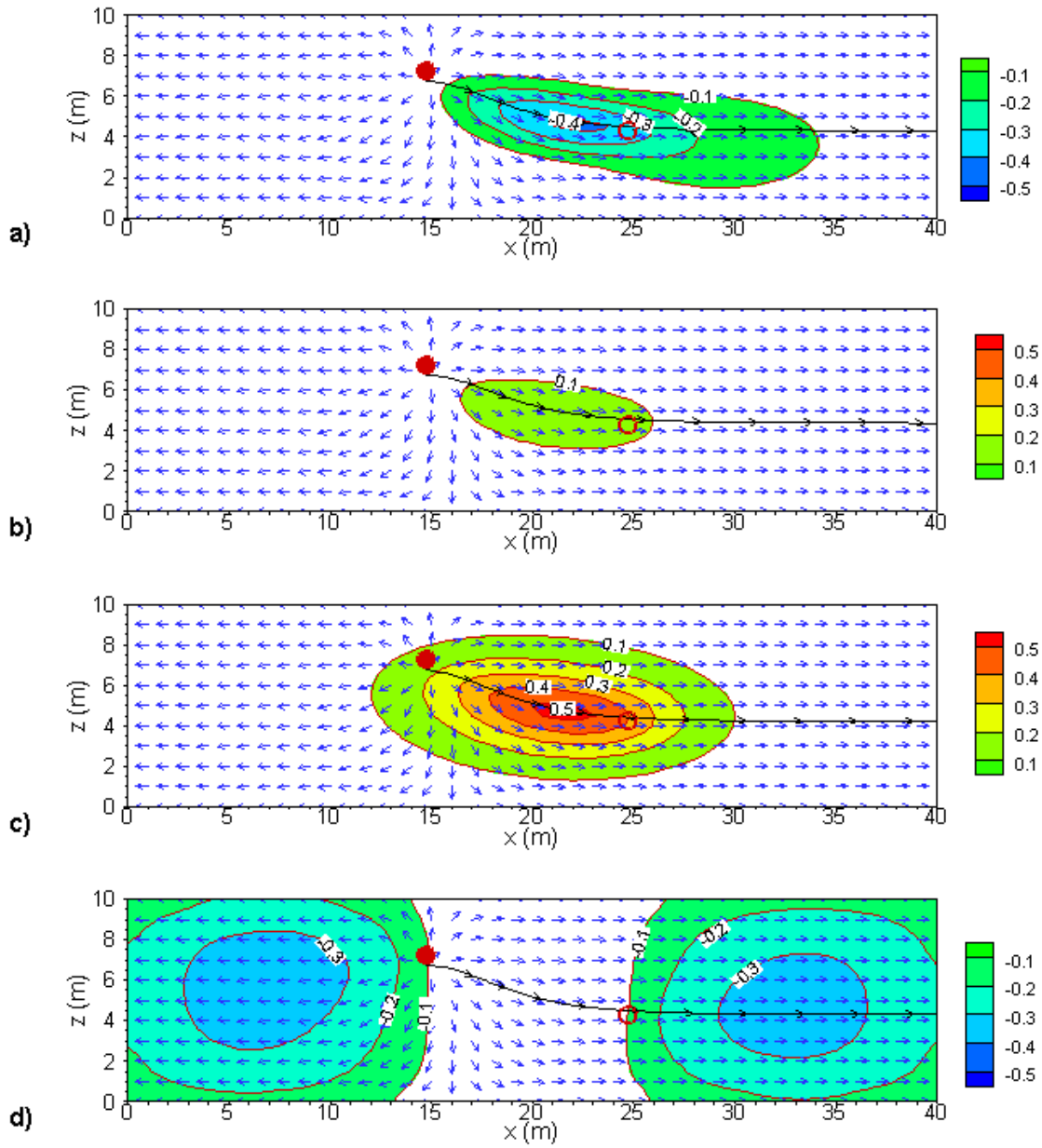


Figure 7.5: Cross correlations between the first moment perturbation and perturbations of: a] ln$K$; b] ln θ_w; and c] ln θ_n and d] cross correlation between perturbation of head and perturbation of ln$K$. Where the filled red circle represents tracer/water injection location and open red circle is the tracer observation location, blue arrows denote the velocity field [not scaled to magnitude], and the black curves with arrows are the streamlines.

8. LABORATORY SANDBOX VALIDATION OF STEADY STATE HYDRAULIC TOMOGRAPHY

8.1 INTRODUCTION

Yeh and Liu [2000] developed a sequential geostatistical inverse method which can be applied to hydraulic tomography for the interpretation of cross-hole hydraulic tests under steady state conditions. The main advantage of sequentially including pumping tests is its computational efficiency. The method is based on the Successive Linear Estimator [SLE] and these authors conducted synthetic simulations for 2- and 3- dimensional cases to test their approach. Validation of the steady state hydraulic tomography was limited to error-free cases of synthetic simulations.

Liu et al. [2002] conducted a laboratory sandbox study to evaluate the performance of hydraulic tomography in characterizing aquifer heterogeneity. This was the first validation study of hydraulic tomography, but the *K* tomograms were only visually compared to the distribution of sand types and to results from synthetic simulations. The *K* tomograms were not compared to small scale estimates of *K* directly and the authors explicitly state that the true *K* distributions were not available for either one of the sandboxes used in the study. The authors mentioned that errors and biases have an effect on their *K* tomograms, but they did not examine the role of errors and biases directly by isolating their causes.

Other researchers [*Gottlieb and Dietrich*, 1995; *Butler et al*., 1999; *Bohling et al*., 2002; *Brauchler et al*., 2003; *McDermott et al*., 2003; *Zhu and Yeh*, 2005, 2006] have developed methods for interpreting hydraulic and pneumatic tomography but none of them conducted a detailed validation of the *K* tomograms. Therefore, the main objective of this task is to examine the accuracy of the *K* tomograms obtained from the steady-state hydraulic tomography algorithm developed by *Yeh and Liu* [2000]. We first obtain a reference *K* tomogram through the inversion of synthetic cross-hole test data generated through numerical simulations. The purpose of reference *K* tomogram generation is to examine the ability of the algorithm to image the heterogeneity pattern under optimal conditions without experimental errors and with full control of forcing functions [initial and boundary conditions as well as source/sink terms]. Parallel to the generation of synthetic data, we conduct hydraulic tests at multiple scales in a laboratory aquifer with deterministic heterogeneity to generate data that are used to validate the *K* tomogram from hydraulic tomography. Measurements include multiple *K* estimates from core, slug, single-hole and cross-hole tests as well as several unidirectional, flow-through experiments conducted upon the sandbox under steady-state flow conditions. Validation of *K* tomograms involved a multi-method and multiscale approach which included: 1] visual comparisons of experimental *K* tomograms [from now on *K* tomogram] to the true sand distributions and the reference *K* tomogram generated using synthetic pumping test data via numerical simulations [from now on reference *K* tomogram]; 2] testing the ability of the *K* tomogram to predict the hydraulic head distribution of an independent cross-hole test not used in the computation of the *K* tomogram; 3] comparison of the conditional mean and variance of local *K* from the *K* tomograms to the sample mean and variance of results from other measurements; 4] comparison of local *K* values from *K* tomograms to those from the reference *K* tomogram; and 5] comparison of local *K* values from *K* tomograms to those obtained from cores and single-hole tests. We also examine the influence of errors and biases on inversion results using forward and inverse simulations of cross-hole tests. Errors and biases arising from the conduct of experiments are usually nonexistent in synthetic simulations as implemented by *Yeh and Liu* [2000] and *Zhu and Yeh* [2005, 2006], so one cannot directly examine their influence. On the other hand, in the field, biases and errors are generally

unknown, so their influence on the *K* tomograms cannot be quantified. Therefore, laboratory sandbox studies can be very important in quantifying the role of errors and biases on the *K* tomograms because all the forcing functions [initial/boundary conditions; sources/sink terms] can be controlled.

We discuss the sandbox used in the study, provide descriptions of various hydraulic tests conducted in the sandbox for characterization, and discuss methods used to obtain data that will be later used to validate the *K* tomograms. We then discuss the forward and inverse analyses of cross-hole tests, including descriptions of data diagnostic tools, inverse modeling results with and without experimental bias and a multi-method/multiscale approach in validating the *K* tomograms.

8.2 SANDBOX DESCRIPTION

The synthetic heterogeneous aquifer constructed in the sandbox was designed to validate various fluid flow and solute transport algorithms and in particular, the hydraulic tomography algorithm. For this validation study we utilized hydraulic test data collected in sandbox 1-1 [see section 10 for details]. Figure 8.1 is a Computer Aided Design [CAD] drawing of the sandbox frontal view, showing the 48 port and pressure transducer locations as well as water reservoirs for controlling hydraulic head.

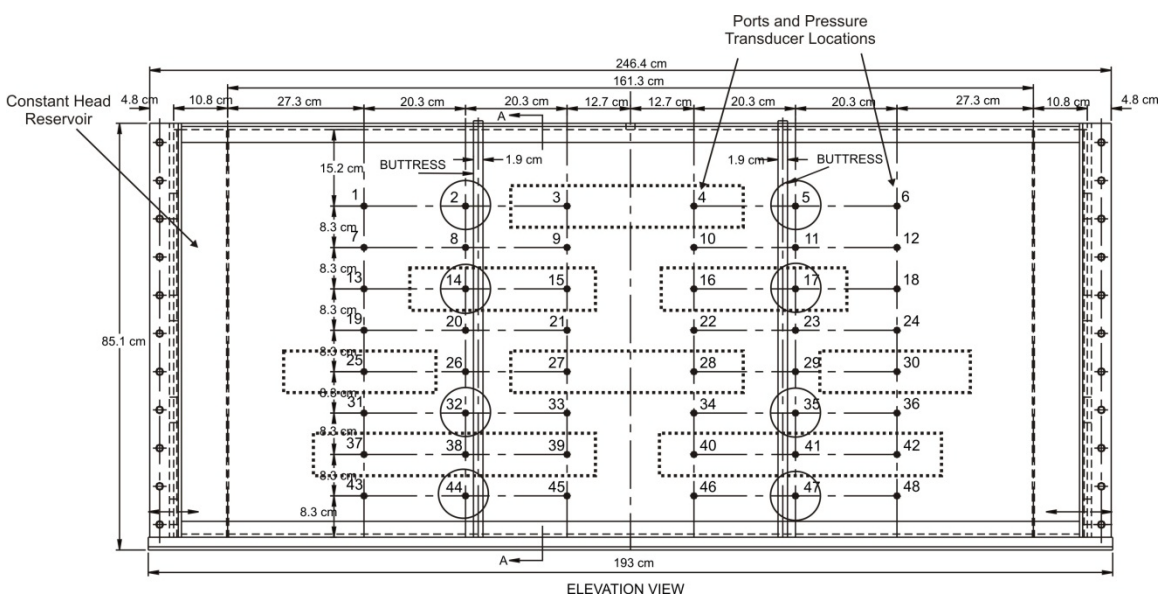


Figure 8.1: Computer Aided Design [CAD] drawing of sandbox used for the validation of hydraulic tomography. Low *K* blocks are shown as dotted lines and circles around the numbers indicate the ports used for pumping.

The flow system for the sandbox is driven by two constant-head reservoirs, one at each end of the sandbox. The adjoining reservoirs are capable of supplying water throughout the length and thickness of the sandbox quickly and efficiently compared to the situation if we had to maintain an external reservoir alone. The boundary head levels can be easily adjusted to be equal or to create a desired hydraulic gradient. We also utilized an intermediate overflow device to maintain equal constant heads on both boundaries in the experiments. The device consisted of a reservoir with an overflow pipe and tubing connecting it to the inlets at the bottom of the constant head reservoirs. The developed system is also capable of maintaining 3 constant head

boundaries simultaneously by ponding water at the top in addition to fixing the hydraulic heads in the 2 constant head reservoirs.

8.3 DESCRIPTION OF THE FORWARD AND INVERSE MODELS

8.3.1 Forward Model

Forward modeling of various hydraulic tests in the sandbox were done using the two-dimensional water flow and solute transport code MMOC2 as well as its three-dimensional version MMOC3 [*Yeh et al.*, 1993]. The forward model is capable of simulating water flow and chemical transport through variably saturated porous media. The flow equation is solved using the Galerkin finite element technique with either the Picard or Newton-Raphson iteration scheme.

We also used VSAFT2, a Graphical User Interface [GUI] program based on MMOC2, to obtain equivalent estimates of hydraulic parameters through manual calibration from the slug, single-hole, and cross-hole tests.

8.3.2 Inverse Model

Inverse modeling of cross-hole tests in the sandbox were conducted using a sequential geostatistical inverse approach developed by *Yeh and Liu* [2000]. We only provide a brief description of the inversion approach here. The inverse model assumes a steady flow field and the natural logarithm of K [ln K] is treated as a stationary stochastic process. The model additionally assumes that the mean and correlation structure of the K field is known *a priori*. The algorithm essentially is composed of two parts. First, the Successive Linear Estimator [SLE] is employed for each cross-hole test. The estimator begins by cokriging the initial K value determined and observed h collected in one pumping test during the tomographic sequence to create a cokriged, mean removed ln K [f, i.e., perturbation of ln K] map. We select an initial K obtained from the traditional analysis of pumping test treating the medium to be homogeneous. Cokriging does not take full advantage of the observed h values because it assumes a linear relationship [*Yeh and Liu*, 2000] between h and K, while the true relationship is nonlinear. To circumvent this problem, a linear estimator based on the differences between the simulated and observed h values is successively employed to improve the estimate.

The second step of *Yeh and Liu* [2000]'s approach is to use the h data sets sequentially instead of simultaneously including them in the inverse model. In essence, the sequential approach uses the estimated K field and covariances, conditioned on previous sets of h measurements as prior information for the next estimation based on a new set of pumping data. This process continues until all the data sets are fully utilized. Modifications made to the code for the present study include its ability to account for variations in the boundary conditions with each pumping test as they are sequentially included and implementing the modified loop scheme described in *Zhu and Yeh* [2005].

8.4 ESTIMATES OF HYDRAULIC CONDUCTIVITY FOR VALIDATION OF K TOMOGRAMS

We first determined the K of the four types of sands from the horizontal cores obtained during the completion of wells and port placement. The extracted cores had dimensions of 1.28 cm in diameter and 10.16 cm in length. These cores were then attached to a custom-made constant head permeameter [*Klute and Dirksen*, 1986] for determination of K. Details to the core

extraction method and the design of the constant head permeameter is provided in *Craig* [2005]. The *K* values from cores are calculated using Darcy's law.

We also conducted slug tests at each of the 48 ports. Due to the small size and configuration of the ports on the sandbox, an external well was attached to the ports instead of boring vertical wells into the sandbox. A slug was introduced to perturb the water level in the horizontal well connected to the port and the corresponding recovery was monitored using a pressure transducer. Because existing analytical solutions cannot be used to interpret the slug tests with our current setting, we analyzed the data by manually calibrating VSAFT2 [*Yeh et al*., 1993], available at www.hwr.arizona.edu/yeh, by treating the model domain to be a two-dimensional, homogeneous medium [*Craig*, 2005]. A fine numerical grid [1.64 cm by 1.64 cm] was developed for the slug test analysis. The numerical simulations were conducted by raising the initial head at the elements corresponding to the slugged port and monitoring the corresponding decay in the head profile. VSAFT2 was chosen to analyze the test data for consistency because the code contains the forward model used later for hydraulic tomography. We report the geometric mean of 40 values that we deem to match the observed data well in Table 8.1. Results obtained revealed that the *K* values are several orders of magnitude smaller than the core values. We suspected that the data are affected by skin effects and wellbore storage. In fact, we investigated the issue further by conducting additional experiments to examine the effects of the number of cuts on the head response to slug tests. In particular, slug tests were conducted in a separate flow cell with tubes consisting of different number of cuts [2 – 8]. This effort revealed that the head response stabilizes after 6 cuts were made on the well. Therefore, all wells in the sandbox discussed in this paper were made by making 6 cuts. Despite these efforts, the *K* values determined using slug tests are very low, thus we question their reliability and do not use them for validation purposes.

We then conducted pumping tests at each of the 46 out of available 48 ports. Ports 36 and 38 have been damaged so we do not pump from these ports. During the pumping tests, the top and two sides of the aquifer served as constant head boundaries, as described earlier, while the bottom remained a no-flow boundary. Pumping rates ranged from 150 to 190 mL/min in most cases. For each test, data collection started before the pump was activated to obtain the initial hydraulic head in the sandbox and the data were collected from all ports every 0.75 sec set to be constant throughout the duration of each experiment. A data collection interval of 0.75 sec was selected to allow for the expected rapid transient change in hydraulic head at the monitoring ports. A peristaltic pump was then activated at the pumping port and allowed to run until the development of steady state flow conditions. The pump was then shut off to collect recovery data until the hydraulic head recovered fully. During each pumping test, pressure heads were collected at all 48 ports.

The data sets were analyzed in several ways. First, we analyzed the 48 drawdown-time data sets induced by pumping at port 22 and those caused by pumping at port 28 by manually calibrating VSAFT 2 and assuming the aquifer is homogeneous. The numerical setup for the manual calibration is identical to the slug test analysis. For the pumping test at port 28 [located in 20/30 sand], all 47 cross-hole intervals were matched and 1 single-hole match was made which yielded a total of 48 estimates for that pumping test. The pumping test at port 22 [located in F-75 sand] yielded also 47 cross-hole matches for observed and simulated drawdown. The single-hole match in this case was unattainable due to the very large drawdown, which VSAFT2 could not simulate. Analysis of the two pumping tests thus yielded 95 estimates of K and S_s for

the equivalent homogeneous medium. These two tests will be denoted as cross-hole tests hereafter.

Out of the 46 pumping tests, we analyzed the drawdown-time data at selected 9 pumping ports [2, 5, 14, 17, 32, 35, 44, 46, and 47] using VSAFT2 to yield local or single-hole estimates of K. These results are denoted as the single-hole results.

We also conducted 9 flow-through experiments through the entire sandbox to obtain the effective hydraulic conductivity [K_{eff}] of the entire sandbox under steady-state unidirectional flow conditions. Specifically, each of these nine experiments was conducted by changing the height of the reservoirs on the both sides of the sandbox. After the flow reached a steady state condition, we measured discharge from one side of the sandbox. We also measured the difference between the heights of the water column in the two constant head reservoirs to determine the hydraulic gradient. The nine pairs of gradient and discharge were computed using Darcy's law to obtain the K_{eff}.

Table 8.1 summarizes the results from all these tests. The mean estimates were obtained by computing the arithmetic mean of the natural logarithm transformed data. The variance was likewise computed using the natural logarithm transformed data set. We also calculated a volume-weighted mean and variance of the core values which are also listed in Table 8.1. The purpose of computing the volume-weighted mean and variance of the core K values was so that these values are upscaled to the size of the finite element grid used for the inversion so that we can compare them later.

In Table 8.1, we see that, in general, the mean values of the cross-hole and flow-through values coincide in this sandbox; however, core, slug, and single-hole test values are noticeably smaller suggesting a scale effect. As mentioned earlier, the slug test values are considerably smaller, so we conclude that near well effects and/or borehole storage dominate the response, causing K estimates to be less reliable in comparison to other measurements. However, K estimates from cross-hole tests in the observation well are very close to the overall K value derived from the flow-through experiments suggesting that these estimates are less affected by near well effects. Therefore, we conclude that the cross-hole observation well data are reliable and we retain them in our analysis.

Table 8.1: Summary of hydraulic properties determined from core, slug, cross-hole pumping test data and flow-through experiments.

Test type	N	$LN-K$	
		$\overline{LN-K}$ [$K \sim$ cms^{-1}]	σ^2_{LN-K}
Core	48	-2.166 [1.146×10^{-1}]†	1.498†
Slug	40	-10.692 [2.272×10^{-5}]	0.431
Single-hole	9	-3.174 [4.182×10^{-2}]	0.570
Cross-hole	96	-1.757 [1.726×10^{-1}]	0.074
Flow-through	9	-1.757 [1.725×10^{-1}]	0.002

† The volume-weighted mean and its corresponding variance are -1.920 [1.467×10^{-1} cms^{-1}] and 1.560 respectively.

Examination of Table 8.1 also shows that $\sigma^2_{\overline{\ln K}}$ varies from one type of test to the next with variance decreasing with the increasing scale. This is because the support volume of each estimate increases from the core, slug, single-hole, cross-hole and flow-through experiments. As the sample volume increases, K is averaged over the investigated volume. We note that the

calculated K values when the medium is treated to be homogeneous are useful, but provide a very limited resolution of the spatial variability in K. In addition, estimation of equivalent parameters treating the medium to be homogenous may yield biased values. This is one important reason why we conduct hydraulic tomography to determine how the K values vary spatially.

8.5 DATA DIAGNOSTICS

8.5.1 Examination of Drawdown Records and their Animations

Prior to inverse modeling of cross-hole hydraulic tests, we conducted a detailed diagnostic study of the data. Such diagnostic tests of data used in forward and inverse models are rarely discussed in the literature, but we found that it should be an integral component to all phases of numerical forward and inverse modeling of cross-hole tests as the use of data corrupted by noise can have a profound effect on model results.

We first plotted the transient head records in all 48 ports including the 2 pressure transducers placed in the constant head reservoirs to examine the propagation of a pressure pulse throughout the aquifer. Plotting of h records in this manner also allowed us to examine whether the pressure transducers were functioning properly as well as to identify the magnitude of noise during a given cross-hole test. The initial diagnosis of the data revealed that pressure transducers can be subject to different noise sources including electrical interference, barometric effects, and minute uncontrollable variations of the water supply. In general, the noise can be removed through signal conditioning and de-trending procedures applied to raw data.

We also contoured the initial head distribution within the sandbox using *Tecplot version 10* [*Amtec Engineering Inc.*, 2003] to identify whether there was any water flow prior to the cross-hole tests. This also allowed us to study the presence of any drift in pressure transducers. Because the transient head record at a given monitoring port provides only limited information about the evolution of the pressure pulse through the aquifer during a given pumping test, we also made animations of head contours using all head records from all 48 ports during a given cross-hole test by plotting successive frames of head distributions over the duration of each test. This process ensured that each test was conducted correctly and gave a pictorial representation of the pressure propagation throughout the sandbox during a given test.

8.5.2 Diagnostic Forward Modeling

We next conducted forward modeling of cross-hole tests to further diagnose the available data. We assumed that the pumping rate [Q] was deterministic and was accurately measured, the pressure transducers were properly calibrated and the drift was removed, the boundary conditions remained stable throughout the experiments and there was no noise to affect the experimental data. The K values used in the forward model are those obtained from taking the mean value from the core for each type of sand. With the forward model, we then simulate each cross-hole test and compare the results from the synthetic to the real data through a scatter plot.

The forward modeling of the cross-hole tests showed that the simulated hydraulic head [h_s] values are in general higher than the measured hydraulic head [h_m] values near the top and bottom of the sandbox. The head value at the pumped port also differs considerably from the simulation results, consistently throughout the sandbox. As we are not aware of the cause[s] of these biases, we conducted a detailed study to determine their cause. A bias can be introduced due to the collection of inaccurate data or through the misapplication of the forward model.

We first examined all head records carefully. This showed that the initial heads are inconsistent indicating the presence of drift in pressure transducers, so we modified the test data to reduce this bias. In particular we accounted for pressure transducer drift by calculating the drawdown [s_i] at port *i*, in the following manner:

$$s_i = h_{i,0} - h_i \qquad [8.1]$$

where $h_{i,0}$ was the initial head at port *i* during a given cross-hole test and h_i is head at port *i* at time *t*. We then averaged the starting h_i by taking the arithmetic mean

$$\overline{h_0} = \frac{1}{n}\sum_{i=1}^{n} h_{i,0} \qquad [8.2]$$

and used this value as the initial head for all ports. We then subtracted the s_i to this starting head to get the modified head for each of the pumping tests.

We also see that the h_s are considerably higher than the h_m values at the pumping port. The extra head drop can be due to inertial effects due to a high pumping rate and/or the development of a low *K* region at the pumped port [i.e., skin effect]. As discussed earlier, a series of diagnostic tests not shown here showed that the number of cuts made in the brass tube can affect the *K* measured at the pumped port and can be responsible for the increased drawdown. Therefore, we tried to account for the skin effect by locally changing the *K* of the element to the value determined from the slug tests, which was several orders of magnitude lower than the values determined through other methods. Results of the simulations accounting for the skin effect at the pumped port showed that the h_s values are closer to the h_m values at the pumping port. We also modified the *K* of all observation ports to account for this skin effect by artificially lowering the *K* to the value determined from the slug tests. However, the bias could not be removed completely so we made a decision not to use the data from the pumping port, but use all original data from the observation intervals without the adjustment of local *K* in the inverse model.

To evaluate the goodness-of-fit between the simulated and measured hydraulic head responses, we calculate the mean squared error norm [L2]:

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}\left(h_{s,i} - h_{m,i}\right)^2 \qquad [8.3]$$

where $h_{s,i}$ is the simulated value of hydraulic head at port *i* and $h_{m,i}$ is the measured value of hydraulic head at port *i*. The smaller the L2 norm, the better the estimate is expected to be. Table 8.2 summarizes the L2 norm for each cross-hole test analyzed showing that forward simulations using modified data [modified only for initial head] reveal a large reduction in the L2 norm in most of the tests. In summary, diagnostic plots, animations of head contours, and forward modeling runs have helped us identify and remove errors/biases in our experiments.

8.6 INPUTS TO THE INVERSE MODEL

To obtain a *K* tomogram from multiple cross-hole tests, we solve a 3D inverse problem for steady flow conditions. The sandbox was discretized into 741 elements and 1600 nodes with

element dimensions of 4.1 cm × 10.2 cm × 4.1 cm. Both sides and the top boundary were set to the same constant head boundary condition, while the bottom boundary of the sandbox was a no-flow boundary. We solve the inverse problem using a consistent grid for both the synthetic and real cases. Here, the synthetic case means that we generate a set of pumping test data by running a series of steady-state forward simulations using MMOC3 with the grid blocks representing the discontinuous blocks of sands in the sandbox. We then use these head and discharge records at the pumping point and observation points in the steady-state hydraulic tomography code of *Yeh and Liu* [2000]. For the real case, we mean the inverse modeling of data collected from the actual cross-hole tests conducted in the sandbox.

Table 8.2: L1 and L2 norms before and after data modifications.

Pumping port number	Original data	After data modification
	L2	L2
2	6.23E-04	1.31E-04
5	2.94E-04	1.49E-04
14	5.15E-04	1.47E-03
17	7.94E-04	6.15E-04
32	1.26E-02	1.99E-02
35	2.95E-02	7.87E-03
44	3.90E-03	8.96E-03
46	5.49E-02	1.53E-02
47	7.95E-03	1.10E-02

Inputs to the inverse model include the initial estimate of K_{eff}, the variance $\left(\sigma^2_{\ln K}\right)$ and the correlation scales of hydraulic conductivity $\left(\lambda_x, \lambda_y, \lambda_z\right)$, h_i, volumetric discharge [Q_n] where n is the test number, and available point [small-scale] measurements of hydraulic conductivity [K_c, K_{SH}]. Here, we do not use the available point scale measurements of K to test the ability of the algorithm to delineate the heterogeneity patterns.

We obtained the initial estimate of K_{eff} separately for the synthetic and real cases. For the synthetic case, we simply take the geometric mean of the small scale or local estimates of the material from the core experiments at the sampling ports. For the inversion of the real cross-hole hydraulic test data, a number of approaches can be used to obtain the initial estimate of K_{eff} for the medium. One is to calculate the geometric mean of the available small scale data [i.e., core, slug, and single-hole data]. The second approach is to do the same for the K_{eq} estimates obtained through the analysis of data from monitoring ports during cross-hole tests by treating the heterogeneous medium to be homogeneous. We also have the results from the flow-through experiments to obtain the K_{eff}. We select the K_{eq} obtained through the traditional analysis of cross-hole tests by treating the medium to be homogeneous because in practice, this is the most logical approach in obtaining a large scale K that should be representative of a large portion of the flow and simulation domain.

Estimation of $\sigma^2_{\ln K}$ always involves uncertainty. A previous numerical study conducted by *Yeh and Liu* [2000], however, has demonstrated that $\sigma^2_{\ln K}$ has negligible effects on the estimated K using the inverse model. Therefore, we obtain $\sigma^2_{\ln K}$ estimates from the available

small scale data and use this as our input $\sigma^2_{\ln K}$ in the inverse model for the real data set. For the inversion of the synthetic data set, we calculate the $\sigma^2_{\ln K}$ from the local values input into the forward model.

Correlation scales represent the average size of heterogeneity that is critical for analyzing the average behavior of aquifers. Correlation scales of any geological formation are difficult to determine. The effects of uncertainty in correlation scales on the estimate based on the tomography are negligible because the tomography produces a large number of head measurements, reflecting the detailed site-specific heterogeneity [*Yeh and Liu*, 2000]. Therefore, the correlation scales were approximated based only on the average thickness and length of the discontinuous sand bodies. Table 8.3 lists values of these statistical parameters [$\sigma^2_{\ln K}, \lambda_x, \lambda_y, \lambda_z$] used in our inverse analysis of cross-hole hydraulic tests for the inversion of synthetic and real cross-hole tests.

Table 8.3: Input data for inverse modeling of 6 pumping tests in sandbox.

Test type	K_{eff} [cms^{-1}]	σ_f^2	λ_x [cm]	λ_y [cm]	λ_z [cm]	Covariance model for f	Q [cm^{3}s^{-1}]
Synthetic	0.19	2.0	30	10	10	exponential	2.92 ~ 3.17
Real	0.17	2.0	30	10	10	exponential	2.92 ~ 3.17

8.7 RESULTS

8.7.1 Inverse Modeling of Synthetic Cross-Hole Tests in Sandbox

We begin the inversion of cross-hole tests by first generating synthetic hydraulic tomography data and inverting them to generate a reference *K* tomogram of the sandbox so that we can compare the results of the inversion of the real data later. This case illustrates the ability of the steady state hydraulic tomography algorithm of *Yeh and Liu* [2000] to obtain a *K* tomogram under idealized conditions where model and measurement errors are excluded and when forcing functions are fully controlled.

All 48 ports [see Figure 8.1 for locations] were used for steady-state hydraulic tomography. The steady state head data were collected at these wells. Figure 8.2a is the true synthetic *K* field used to generate the cross-hole test data employed for hydraulic tomography. Figures 8.2b - i [cases 1 - 8] then shows the results of the estimation of the *K* tomogram from the successive inclusion of test data from pumping tests at ports 47, 44, 35, 32, 17, 14, 5, and 2, conducted in that order. These results clearly show that the inversion algorithm is capable of capturing the low *K* blocks, and other details of aquifer heterogeneity such as windows in low *K* strata that could provide continuous pathways for contaminant transport. It is of interest to note that the synthetic aquifer has a *K* distribution that is non-Gaussian and nonstationary. Because *Yeh and Liu* [2000]'s algorithm assumes a Gaussian and a stationary field, one would not expect this approach to be applicable to this *K* distribution. However, the statistical assumptions inherent in the algorithm become less important as we include large number of observation well data sets in the hydraulic tomography algorithm. This fact is evident from our results. Therefore, *Yeh and Liu* [2000]'s SSLE is not limited to Gaussian and stationary random *K* fields.

The results were also quantitatively evaluated using the L2 norm, defined earlier, but in this case for the estimated and true parameters for the entire computational domain. Figure 8.2b-i

shows that the L2 norm decreases as more pumping tests are added, but the rate of reduction diminishes and stabilizes through cases 6 - 8.

We use the best synthetic result [case 8] as a reference *K* tomogram in which we later compare our tomograms from the inversion of real cross-hole test data.

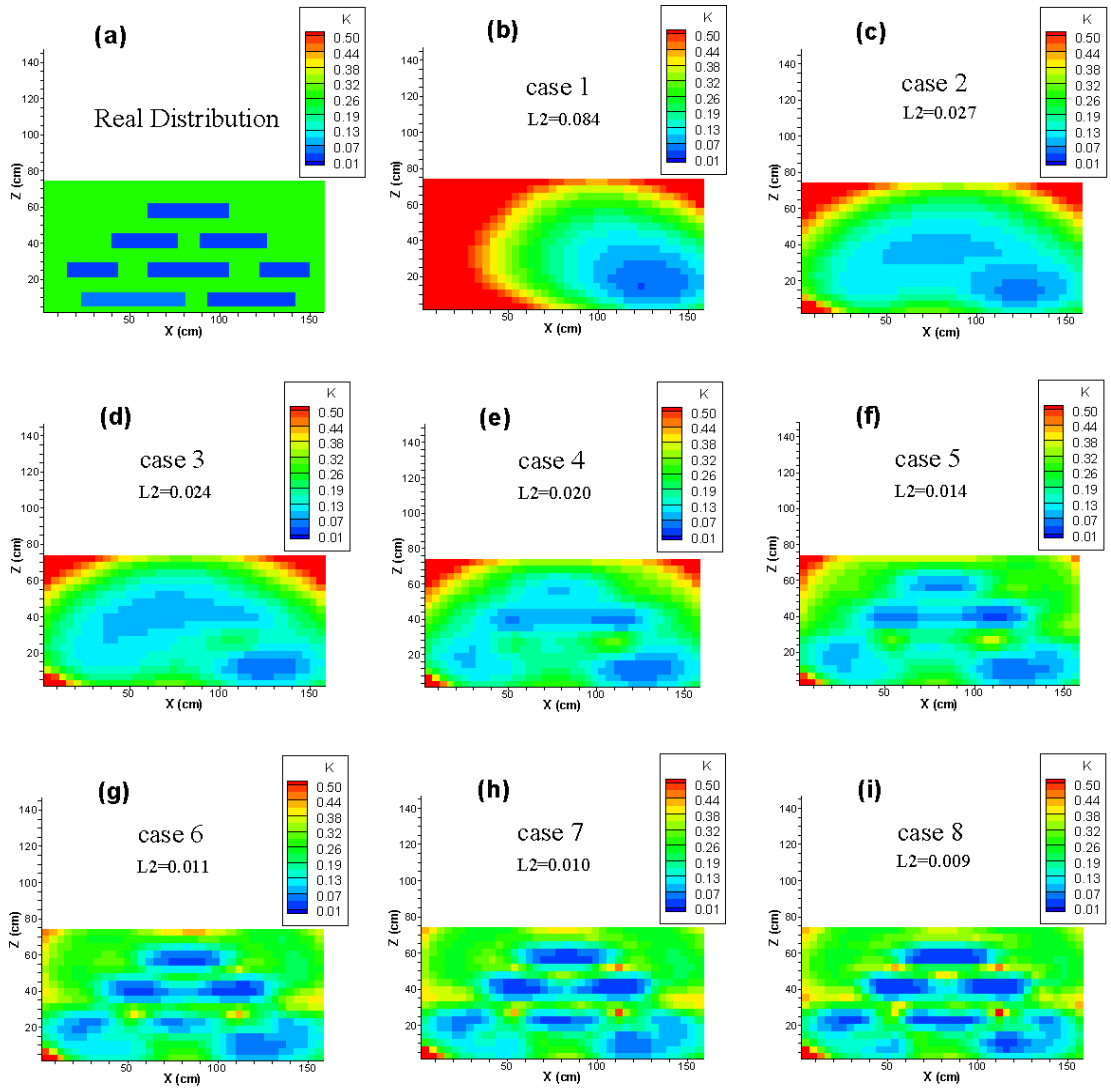


Figure 8.2a-i: a] The synthetic true K distribution used to generate synthetic cross-hole test data; b] [47 - case 1] c] [47, 44 - case 2]; d] [47, 44, 35 - case 3]; e] [47, 44, 35, 32 - case 4]; f] [47, 44, 35, 32, 17 - case 5]; g] [47, 44, 35, 32, 17, 14 - case 6]; h] [47, 44, 35, 32, 17, 14, 5 - case 7]; and i] 47, 44, 35, 32, 17, 14, 5, 2 - case 8] show the resulting K tomograms by sequentially inverting the synthetic cross-hole test data. Numbers in parentheses from b] – i] indicate the port numbers used as the pumped well for each cross-hole test.

8.7.2 Inverse Modeling of Real Cross-Hole Tests in Sandbox

We next examine the results from the inversion of real cross-hole test data obtained in identical fashion to the synthetic case. In reality we do not know the true *K* distribution as in the synthetic case, so we do not show a true *K* distribution as in Figure 8.2a. Instead, we examine how the computed *K* tomogram improves after the identified errors and biases are sequentially removed. Also, the L2 norm is computed only at the ports in which the estimated *K* is compared to the true *K* from available cores. Figure 8.3a-d shows the sequential improvement of the computed tomograms as we remove the errors and biases as discussed in the previous section. Specifically, figure 8.3a is the tomogram of the original data set from pumping taking place at 8 ports 47, 44, 35, 32, 17, 14, 5, and 2 without removing the errors and biases [case 9]. Clearly, these results do not compare favorably with the reference *K* tomogram [case 8].

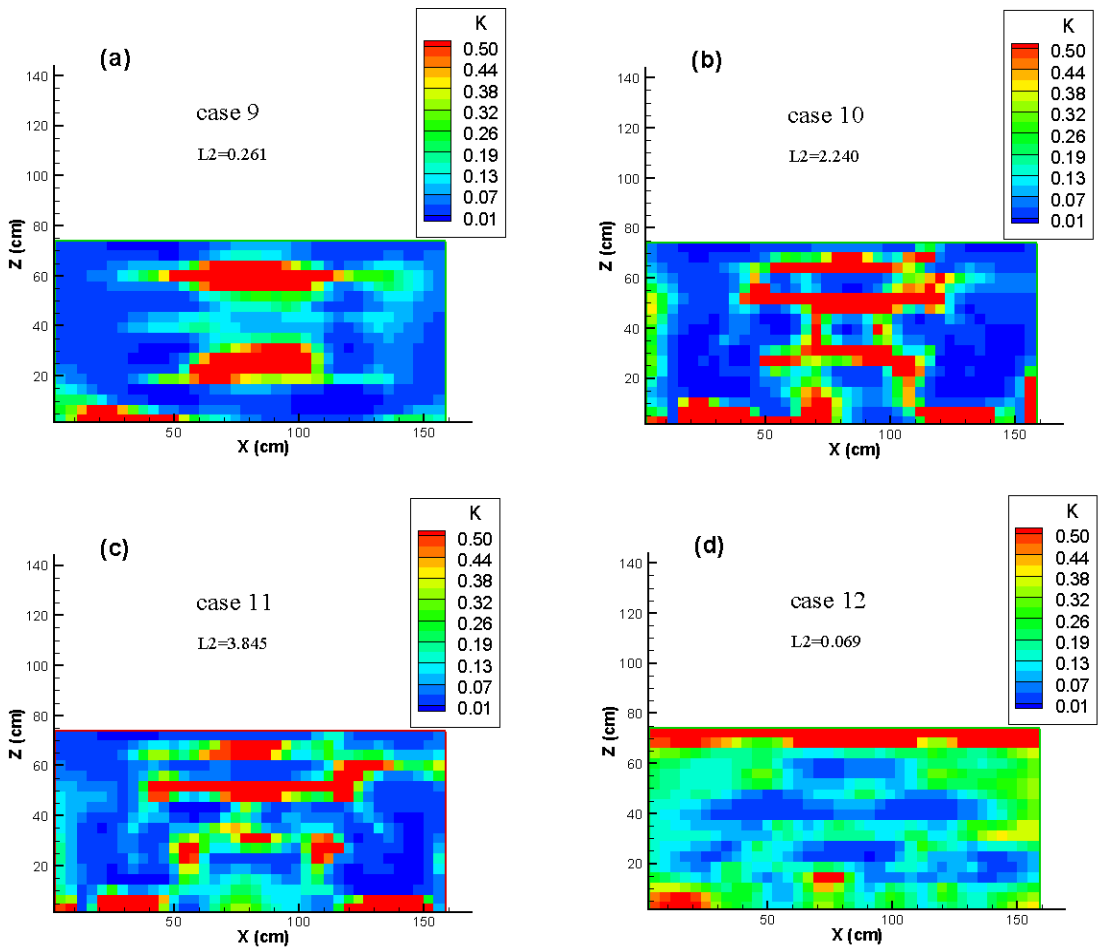


Figure 8.3a-d: Tomogram with: a] original data set without removing bias/errors [case 9]; b] after removal of pumped well data affected by skin effect [case 10]; c] hydraulic head measurements corrected for variations in the offset [case 11]; d] accounting for varying boundary conditions for each pumping test [case 12]. All results shown here are after including 8 pumping tests in the inversion as in case 8 described in figure 8.3i.

Figure 8.3b shows the tomogram after removal of pumped well data affected by skin effect [case 10] and Figure 8.3c shows the tomogram after the *h* measurements are corrected for variations in the drift or offset in the pressure transducers [case 11]. We see a progressive improvement in the computed tomograms as the low *K* blocks begin to emerge. Therefore removal of the pumped well data from the inverse model improves the quality of the computed tomogram. In contrast, the skin effect at the observation wells does not have a significant impact on the results of the steady state hydraulic tomography. The effect of initial head variations due to the drift in pressure transducers also does not have a large effect on the inversion results for this particular case.

The largest effect on the inversion results seems to come from the slight variations in the boundary condition [~1-2 mm differences in water level in the constant head reservoirs] from one pumping test to the next. When we account for the varying boundary condition for each test, the results improve dramatically [figure 8.3d – case 12]. This suggests that steady state hydraulic tomography is very sensitive to the constant head boundary conditions that are input into the model when hydraulic heads are used as input data. If drawdowns are used instead of heads in the inverse modeling effort, the effects of the boundary can be minimized.

8.8 MULTI-METHOD AND MULTISCALE VALIDATION OF *K* TOMOGRAMS

Model validation in the traditional sense means that the accuracy and predictive capability [of a model] can be shown to lie within acceptable limits of error by test data that are independent of the calibration data [*Konikow*, 1978]. In other words, model validation is done through the comparison of simulated and measured heads for both forward and inverse modeling efforts. The comparison of head/K data alone may not be a robust approach in validating the inverse model, so we instead utilize a multi-method and multiscale approach developed herein. The multi-method validation approach consists of validation through different, yet complementary approaches in validating the resulting *K* tomogram. One approach is to visually examine whether the *K* tomogram resembles the true *K* distribution from the sandbox. Another validation approach is to simulate an independent cross-hole test not used in the construction of the *K* tomogram. The most direct validation is achieved through the comparison of the local *K* values obtained from the *K* tomograms to those obtained by independent experimental results such as *K* values from cores and single-hole tests. The cross-hole and flow-through experiments provide equivalent [K_{eq}] and effective [K_{eff}] values, which can also be compared by averaging the local *K* values from the tomograms. Because the independent experimental data are collected at different spatial scales, we consider the comparison of local and averaged *K* values from the *K* tomograms to independent data sets to be validation conducted at different spatial scales. We consider the *K* tomogram to be validated when it can be validated by employing all of these methods utilizing various data sets collected at multiple spatial scales.

8.8.1 Multi-Method Validation Approach 1: Visual Comparison

A visual comparison of the *K* tomogram [Figure 8.3d] obtained by sequentially including 8 tests into the inversion algorithm and the photograph of the sand bodies shows that the steady state hydraulic tomography is able to delineate the major low and high hydraulic conductivity features that comprise the aquifer heterogeneity. This includes the correct delineation of the morphology and positions of low *K* blocks/layers and stratigraphic windows which can allow contaminant migration. We also visually compare the *K* tomogram resulting from the real data sets [Figure 8.3d] to the synthetic *K* tomogram [Figure 8.2i]. This comparison shows that the *K*

near the top of the aquifer for Figure 8.3d is higher than that in Figure 8.2i. This is likely due to the lack of sand compaction in the sandbox, artificially inflating the *K* values in Figure 8.3d. In general, it is evident from the comparisons that the approach is able to correctly depict major heterogeneity features using the hydraulic tomography technology.

8.8.2 Multi-Method Validation Approach 2: Validation of *K* Tomogram with Additional Cross-Hole Hydraulic Tests not used in the Inversion

Another method of validation is to simulate an independent cross-hole test that has not been previously used to calculate the *K* tomogram and examine whether the hydraulic head of this independent test can be predicted accurately. For this, we utilize the *K* tomogram obtained from the inversion of 8 cross-hole tests [Figure 8.3d] and simulate a cross-hole test with pumping taking place at port 46. Figure 8.4 shows the results of comparing the h_s measurements obtained from a synthetic cross-hole test conducted at port 46 and the h_m from these tests. Results show that the h_s is slightly higher than the h_m. However, the comparison is very good considering that this is an independent cross-hole test not used in the construction of the *K* tomogram.

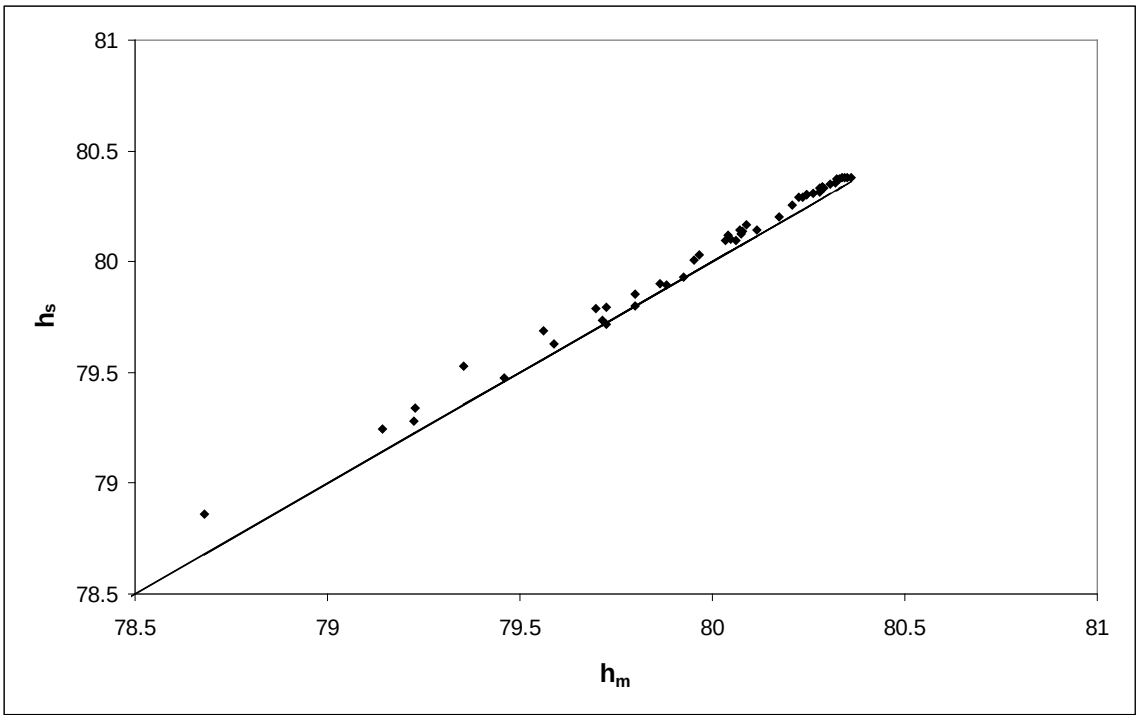


Figure 8.4: Scatter plot of simulated hydraulic head [h_s] versus measured hydraulic head [h_m]. The h_s values were obtained by simulating the cross-hole test conducted at port 46 using the K tomogram [Figure 8.3d]. The h_m values are from an actual cross-hole test conducted at port 46.

8.8.3 Multi-Method Validation Approach 3: Comparison of Statistical Moments

We next validate the *K* tomogram by comparing the conditional mean log-transformed hydraulic conductivity $\overline{\ln K}$ and conditional log-transformed hydraulic conductivity variance $\left(\sigma^2_{\ln K}\right)$ obtained from the entire *K* field in the tomogram to the corresponding statistical moments obtained from other validation data sets. We first compare the conditional $\overline{\ln K} = -1.73$ obtained from our best *K* tomogram [figure 8.3d] [Table 8.4] to the ln K_{eff} [-1.76] obtained from the flow-through experiment [Table 8.1]. This comparison is excellent showing that the inverse

model is able to accurately and uniquely obtain the conditional mean value after including 8 pumping tests in the inversion. We note that the difference between the conditional mean value and the actual mean value decreases quickly from the addition of one cross-hole test to the next [Table 8.4]. We find that this result holds even when the initial value of *K* used in the SLE is varied.

The same cannot be said about the conditional $\sigma^2_{\ln K}$. It takes a larger number of pumping tests for a reasonable value to be estimated. To compare the conditional variance obtained from the inversion, we use the estimates of sample variance obtained from the available core *K* data [1.50]. The latter may be slightly higher than the actual value for the entire population [1.37] because we only use 48 values to compute the sample variance as opposed to the conditional $\sigma^2_{\ln K}$ from the *K* tomogram where we use all values in the computational domain. This comparison shows that the inverse approach is able to estimate the conditional variance quite well after the sequential analysis of 8 cross-hole tests. We consider this comparison to be quite good based on the fact that we do not have the variance of the true *K* field and that the conditional variance approaches the sample variance as more cross-hole test data are included [see Table 8.4]. Other estimates of variances listed in Table 8.1 are representative of the medium at the larger scale, thus do not provide a fair comparison. Therefore, we restrict the comparison to the variance estimate from core samples.

Table 8.4: Conditional mean and variance estimates for the inversions of the real data with the number of cross-hole tests used in the analysis.

Number of cross-hole tests included in analysis [test number]	$\overline{\ln K}$ [$K \sim$ cms^{-1}]	$\sigma^2_{\ln K}$
1 [47]	-1.71 [0.18]	0.34
2 [47+44]	-1.81 [0.16]	0.44
3 [47+44+35]	-1.79 [0.17]	0.60
4 [47+44+35+32]	-1.78 [0.17]	0.80
5 [47+44+35+32+17]	-1.76 [0.17]	0.96
6 [47+44+35+32+17+14]	-1.74 [0.18]	1.16
7 [47+44+35+32+17+14+5]	-1.72 [0.18]	1.35
8 [47+44+35+32+17+14+5+2]	-1.73 [0.18]	1.37

8.9.4 Multi-Method Validation Approach 4: Comparison of Local *K* Values from the Real *K* Tomogram to the Reference *K* Tomogram

We next compare the *K* tomogram to those obtained from the synthetic tests that we consider to be the reference *K* tomogram [case 8 – Figure 8.2i]. As mentioned earlier, the reference *K* tomogram is computed from synthetic cross-hole tests on the computer. The reference *K* tomogram is our best *K* tomogram that can be obtained from the steady state hydraulic tomography approach of *Yeh and Liu* [2000] under optimal conditions without experimental errors and with full control of forcing functions [initial and boundary conditions as well as source/sink terms]. Therefore, one approach to validate the real *K* tomogram is to compare it against the reference *K* tomogram. Figure 8.5 shows a scatter plot of the local *K* estimates from the two cases. We see that the comparison between the two is very good and the differences between the two arise from some errors in the experiments as well as the imperfect packing of the sandbox.

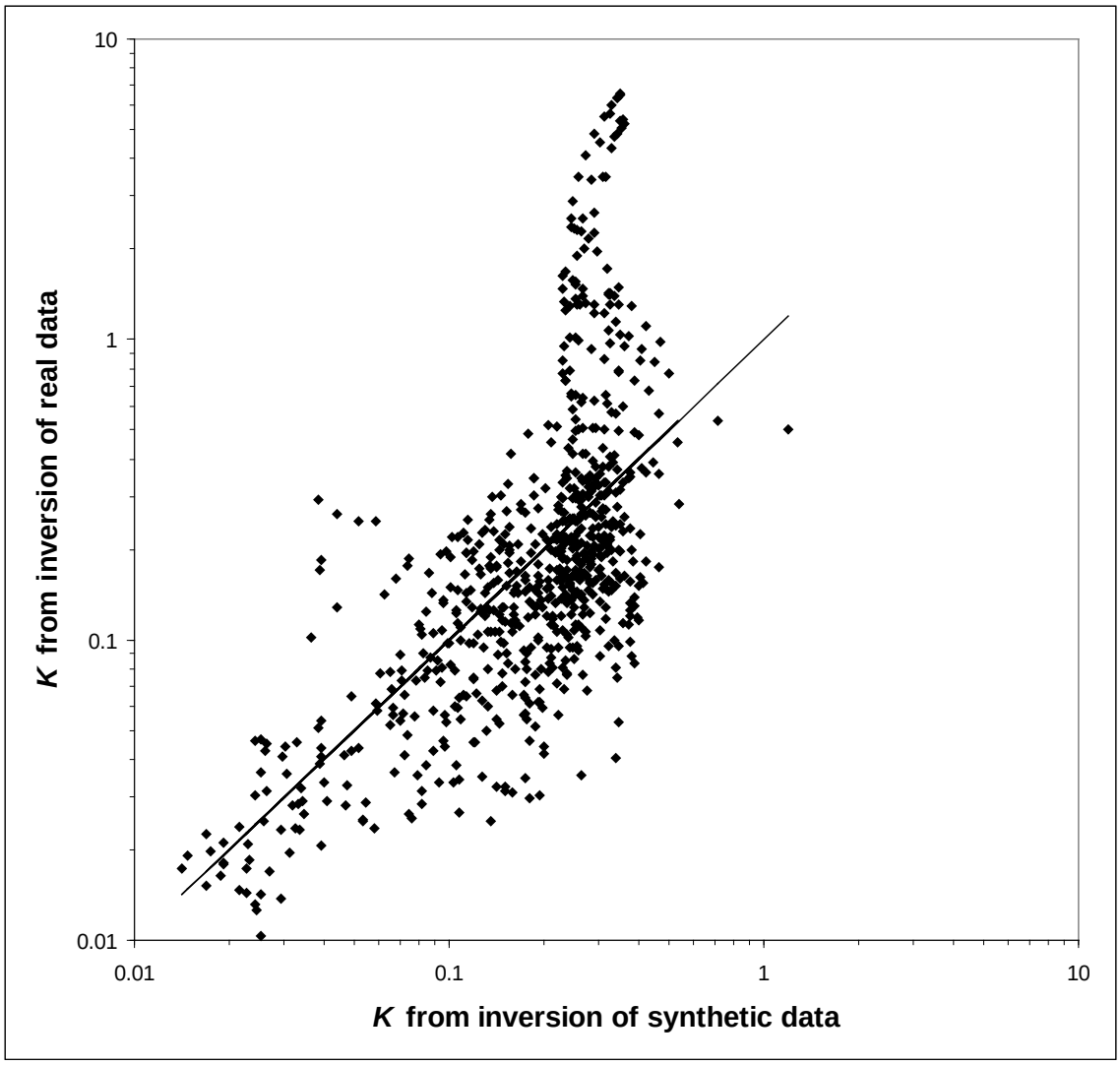


Figure 8.5: Scatter plot of *K* values from the real [case 12] and the reference [case 8] *K* tomograms.

8.9.5 Multi-Method Validation Approach 5: Comparison of Local *K* Values from the Real *K* Tomogram to *K* from Core and Single-Hole Tests

To examine the performance of the algorithm in greater detail, we next compare local *K* values from the *K* tomogram to the *K* estimates from the cores [Figure 8.6] and single-hole tests [Figure 8.7]. Both figures show that there is some scatter as well as bias in results, with the latter that is especially evident in Figure 8.7. Robust as it is, neither the HT nor the SSLE is a perfect method. The more head observations are collected, the higher the resolution of the estimates will be [i.e., there is no optimum].

Likewise, *K* estimates from core and single-hole tests are not devoid of errors contributing to the scatter in Figures 8.6 and 8.7. In addition, inaccurate head observations and hydraulic property measurements [i.e., noise] during HT can unequivocally lead to an inaccurate estimate or cause the estimates to become unstable. The SSLE can overcome the impacts of noise through loosening of convergence criteria, but this causes the estimates to become

smoother, which effectively results in a loss of information gained from the hydraulic head records.

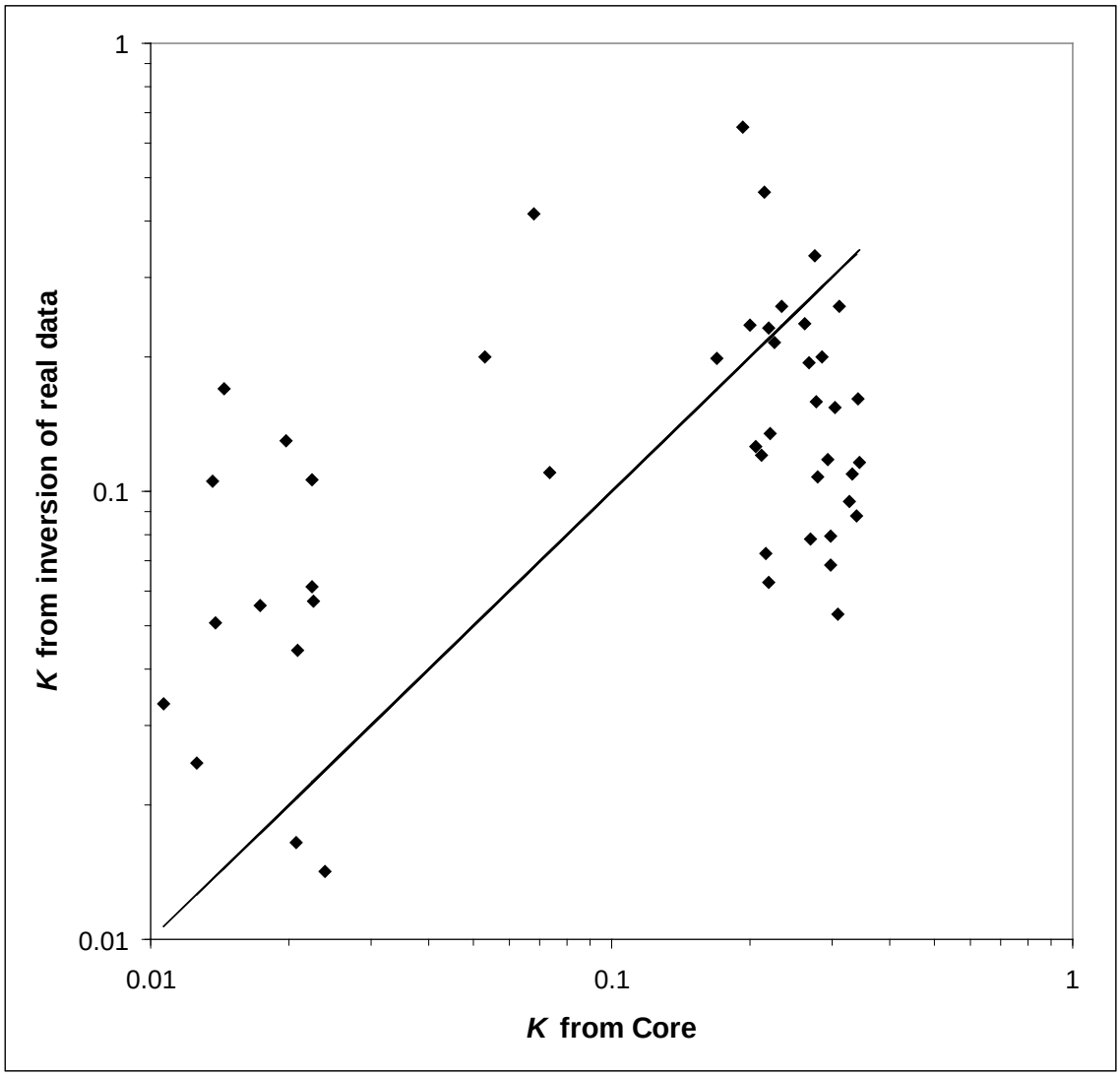


Figure 8.6: Scatter plot of *K* values from the real [case 12] *K* tomogram at the observation point and core *K* estimates.

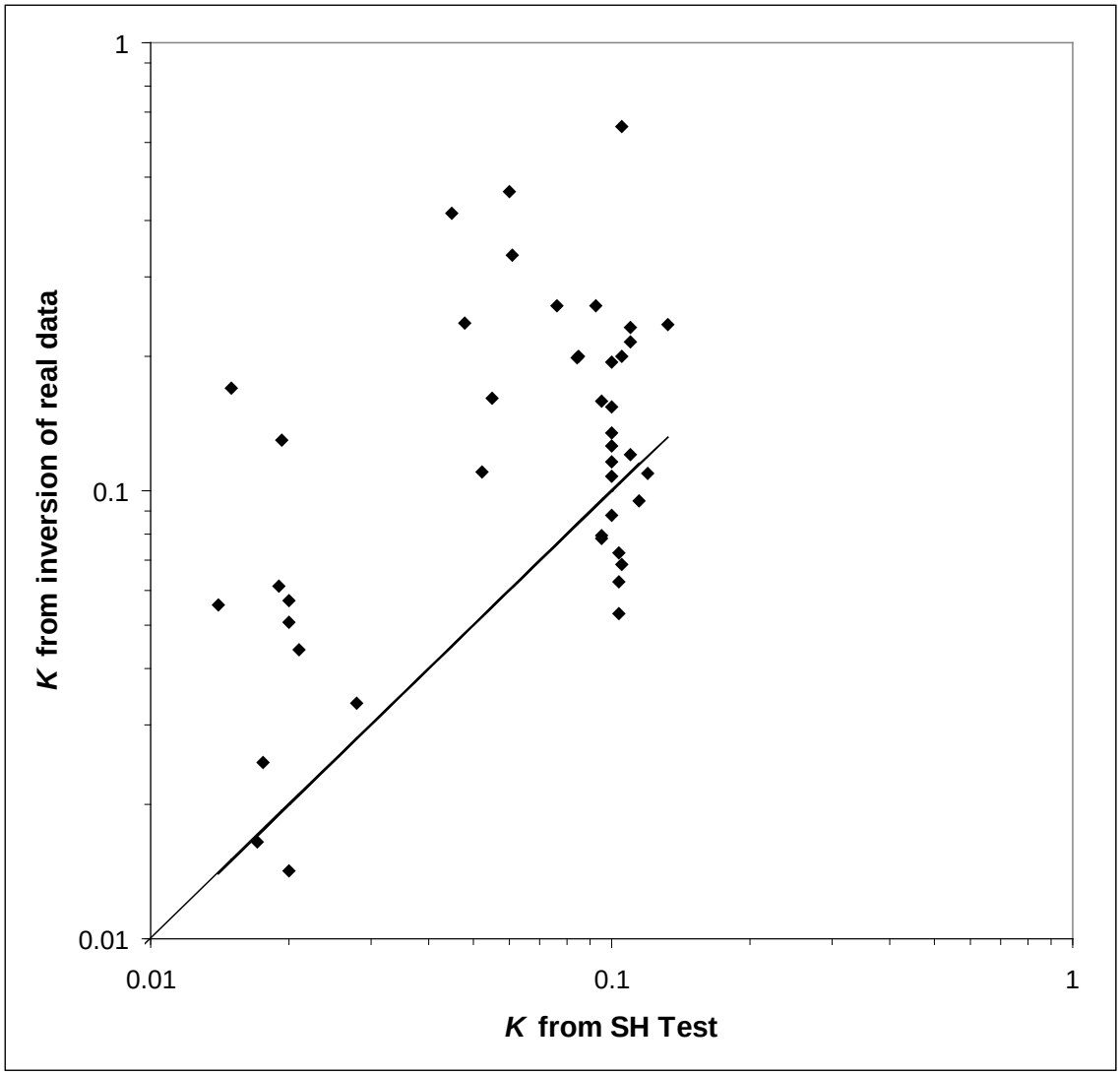


Figure 8.7: Scatter plot of *K* values from the real [case 12] *K* tomogram at the observation point and single-hole *K* estimates.

8.10 DISCUSSION

While the validation of hydraulic and pneumatic tomography under field conditions is our ultimate goal, the validation of hydraulic tomography and other tomography technologies in the laboratory is very important and a necessary step. In the laboratory, we are better able to control the forcing functions fully and quantify the errors. We can also pack a heterogeneous structure that is almost fully prescribed. However, the true *K* field remains unknown because of packing variations. In addition, there are only a finite number of small scale samples that can be used for validation purposes. Therefore, we emphasize that even in the laboratory setting, a direct and complete validation of results is generally difficult.

In this study, we have shown that the *K* tomogram can be validated using multiple methods. The tomograms can also be validated at multiple scales from the smaller scale, when the local *K* values are compared to cores, slug, and single-hole *K* estimates to the larger-scale *K*

estimates from other cross-hole and flow-through experiments. Such a multi-faceted approach in validation adds more confidence on the ability of the algorithm to tackle field scale problems.

One form of model validation involves the establishment of greater confidence in a given model by conducting simulations of data sets that have not been used for calibration purposes. For example, one can calibrate a model using one set of pumping test data. If the calibrated model from this first pumping test can predict system response accurately in a second pumping test [e.g., conducted using another well], one can have greater confidence in the calibrated model. On the other hand, if the parameters need to be adjusted to match the response of the 2nd pumping test, the process becomes a second calibration and additional data sets are needed to continue with the validation exercise. Model validation is complete when the validation data sets are matched against simulated values resulting from the previously calibrated parameter values.

This is precisely the essence of hydraulic tomography conducted with sequential inclusion of data. Therefore, it amounts to a repeated validation of the estimated *K* field with new data sets that are sequentially added. The method is robust, but it is not the panacea technology. This is because the computed *K* tomogram is non-unique as there are an infinite number of solutions to the steady state inverse problem for a heterogeneous *K* field, even when all of the forcing functions are fully specified. Only when data are available at all estimated locations will the inverse problem be well-posed and ultimately lead to a unique solution [e.g., *Yeh et al*., 1996; *Yeh and Liu*, 2000; *Liu et al*., 2002; and *Yeh and Simunek*, 2002]. This is not the case here. However, it is important to recognize that we have obtained a solution to the inverse problem that is consistent with the heterogeneity patterns that we can visualize and directly compare against the experimentally packed sand distributions [Fig. 10.7]. In addition, we were able to validate the resulting *K* tomogram using multiple methods and at multiple scales so our approach provides more confidence in the solution of the inverse problem.

Earlier, we saw that errors and biases can be very important in the result of the *K* tomogram and the blind addition of new data does not mean that it will automatically generate better results. Therefore, more effort should be expended on collecting accurate data and additional research should be conducted on improving hydraulic tomography technology both in sandbox experiments and in the field. This study also emphasizes the importance of reducing experimental errors and biases during validation of traditional groundwater flow and contaminant transport models.

9. PRACTICAL ISSUES IN IMAGING HYDRAULIC CONDUCTIVITY THROUGH HYDRAULIC TOMOGRAPHY

9.1 INTRODUCTION

In section 8, we examined the accuracy of the *K* tomograms obtained from the steady-state hydraulic tomography algorithm developed by Yeh and Liu [2000]. They obtained multiple *K* estimates from core, slug, single-hole and cross-hole tests as well as several uni-directional, flow-through experiments conducted upon the sandbox under steady-state conditions. We also examined the influence of errors and biases on inversion results using forward and inverse simulations of cross-hole tests. They found out that the pressure transducer offsets, skin effect at the pumped well, among other sources of errors can have a large impact on the quality of the inverse modeling results. Likewise in section 10, we present the results from the laboratory validation of the transient hydraulic tomography to estimate both *K* and S_s tomograms simultaneously, but in both cases, we have not examined practical factors such as the effects of signal-to-noise ratio and the role of conditioning on the quality of the inversion results.

There are several issues that need to be further examined. One important issue in applying hydraulic tomography in the field is how much noise a given data set can contain in obtaining an accurate *K* tomogram without having to apply smoothing and/or signal processing techniques, which may result in the loss of information on the parameters contained in the data set. Another issue is that SSLE incorporates pumping test data sequentially, but no studies have been published to date that examines the role of varying the order of pumping test data included in the SSLE algorithm. Finally, during site characterization, test data other than cross-hole hydraulic test data such as from direct push technologies, core, slug, geophysical, and geochemical data may be collected. In some cases, one could utilize some of these data to condition the inverse modeling results. However, there are no studies to our knowledge, in which the role of conditioning on the estimated *K* tomogram by means of hydraulic/pneumatic tomography was systematically studied.

To address these practical yet important issues, in this section we continue our study on hydraulic tomography using synthetic data generated through numerical simulations and real data collected in the laboratory. The main objectives of this section are: 1] to further study the validity of steady state hydraulic tomography through synthetic pumping test data obtained through forward numerical simulations and real data obtained using a laboratory sandbox aquifer; 2] to investigate the effect of varying the pumping rate, which affects the signal-to-noise ratio, and its impact on *K* tomograms; 3] to investigate the effect of varying the order of test data included into the SSLE inversion algorithm; and 4] to investigate the effect of conditioning on *K* tomograms. We assess the validity of the synthetic and real *K* tomograms by simulating an independently conducted pumping test and comparing the head values obtained from the simulated and observed cases.

We first discuss the numerical simulation approach that underlies hydraulic tomography and approaches used to generate synthetic hydraulic test data and corresponding laboratory data sets. We then briefly discuss the reference *K* tomograms from synthetic and real data generated through laboratory sandbox experiments by Illman et al. [2007]. These reference *K* tomograms will be used to compare against the new results presented in this paper. In all cases, we first compare our results visually to the reference *K* tomograms obtained as a benchmark result for both the synthetic and real cases. Here, a synthetic *K* tomogram is one that is generated by inverting pumping test data generated on the computer. A real *K* tomogram is obtained by

inverting actual pumping test data obtained through laboratory experiments. All K tomograms, whether synthetic or real, are generated based on the experimental setup that was described previously in section 8. Further details of the laboratory aquifer, its specifications, and its capabilities are described in Craig [2005], Illman et al. [2007], and Liu et al. [2007].

We emphasize that the synthetic experiments conducted on the computer are necessary to test the SSLE algorithm under optimal conditions in which the experimental errors are neglected and the forcing functions [boundary condition and source/sink terms] are fully controlled. In addition, through numerical simulations, hydraulic tomography and the K tomograms obtained can be tested rigorously. The laboratory experiments described below are also required to test the SSLE algorithm under controlled conditions which is a necessary step towards its field applications.

9.2 NUMERICAL SIMULATION METHODS

9.2.1 Inverse modeling approach

Inverse modeling of all synthetic and real pumping tests are conducted using a sequential geostatistical inverse approach developed by Yeh and Liu [2000], in which all details to the algorithm are provided. Details on Yeh and Liu's [2000] approach are also provided in section 8.3.2.

9.2.2 Inputs to the inverse model

To obtain a K tomogram from multiple cross-hole pumping tests, we solve a 3D inverse problem for steady flow conditions. The sandbox was discretized into 741 elements and 1600 nodes with element dimensions of 4.1 cm × 10.2 cm × 4.1 cm. Both sides and the top boundary were set to the same constant head boundary condition, while the bottom boundary of the sandbox was set to be a no-flow boundary. We solve the inverse problem using a consistent grid for both the synthetic and real cases. Here, the synthetic case means that we generate a set of pumping test data by running a series of steady-state forward simulations using a finite element flow model MMOC3 [Yeh et al., 1993]. We then use these head and discharge records at the pumping point and observation points in the steady-state hydraulic tomography code of Yeh and Liu [2000]. For the real case, we mean the inverse modeling of data collected from the real cross-hole tests conducted in the sandbox.

Inputs to the inverse model include the initial estimate of effective hydraulic conductivity [K_{eff}], variance $\left(\sigma^2_{\ln K}\right)$, correlation scales of hydraulic conductivity$\left(\lambda_x, \lambda_y, \lambda_z\right)$, volumetric discharge [$Q_n$] where n is the test number, and available point [small-scale] measurements of K. Results which do not use available point scale measurements of K to test the ability of the algorithm to delineate the heterogeneity patterns are described in Illman et al. [2007] as well as in this paper. Later in this paper, we examine the effect of using available point scale K data [i.e., conditioning] on the computation of K tomograms.

We obtained the initial estimate of K_{eff} by averaging the K data from monitoring ports during cross-hole tests by treating the heterogeneous medium to be homogeneous. We also have the results from the flow-through experiments [described later] to obtain the K_{eff}. However, we select the K_{eff} obtained through the traditional analysis of cross-hole tests by treating the medium to be homogeneous, because in practice, such values are most readily available through type curve, straight-line, or asymptotic analysis [Illman and Tartakovsky, 2006] of pumping test data.

9.3 GENERATION OF SYNTHETIC & LABORATORY HYDRAULIC TEST DATA

9.3.1 Synthetic hydraulic test data

We generated synthetic slug, single-hole, and cross-hole pumping test data on the computer using the MMOC3 forward model. There were two purposes to generate these synthetic data. One purpose was to generate synthetic cross-hole pumping test data, which are later used to generate the reference *K* tomogram. The reason for generating the slug and single-hole test data are so that we can interpret them and obtain small-scale *K* estimates, which we can in turn use to condition the *K* tomograms.

Figure 9.1 shows the *K* distribution of 4 sand types in the synthetic sandbox that is used to generate synthetic data on the computer. The *K* values for the 4 sand types are listed in Table 9.1 and are obtained through the analysis of core samples through a constant head permeameter which we describe later. We first conducted slug tests at the 48 port locations by raising the initial head and recording the corresponding decay in the head using MMOC3. We then conducted synthetic cross-hole tests by running steady-state forward simulations using MMOC3 at each of the 48 ports by setting a constant pumping rate and recording the hydraulic heads at the other ports. The pumping rate was set between 2.92 and 3.17 ml/s depending on the cross-hole pumping test. Boundary conditions for the simulations were equal constant head conditions at the top and the two side boundaries, while the bottom boundary remained a no-flow boundary.

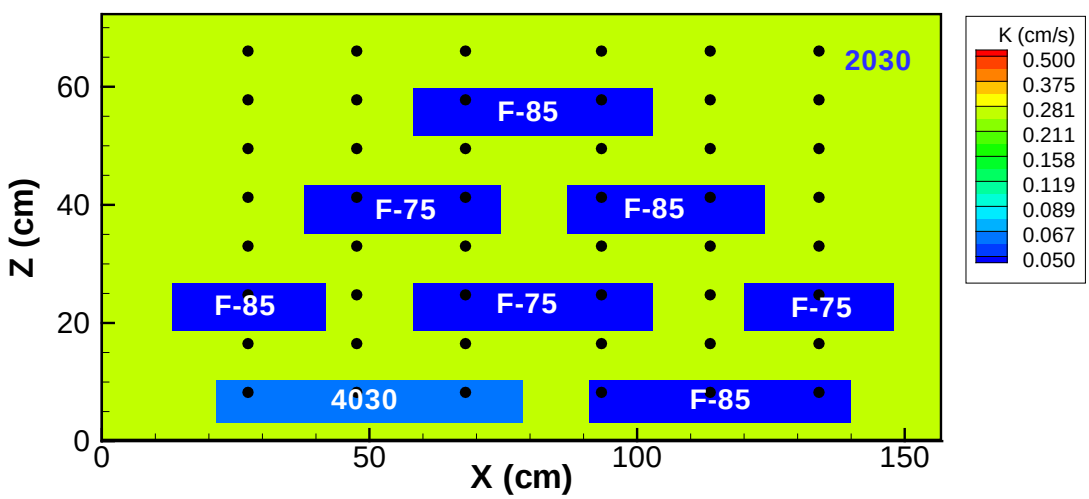


Figure 9.1: Synthetic *K* distribution with port locations for pumping, observation, and conditioning.

Table 9.1: Geometric mean values of *K* [cm/s] determined from core samples taken from the sandbox.

Sand type	Manufacturer	*n*	*K* [cm/s]
20/30	U.S. Silica	32	2.60×10^{-1}
4030	U.S. Silica	3	6.42×10^{-2}
F-75	Unimin Corp.	5	1.99×10^{-2}
F-85	Unimin Corp.	8	1.61×10^{-2}

9.3.2 Laboratory hydraulic test data

Parallel to the generation of synthetic hydraulic test data, we conducted different hydraulic tests in the sandbox to characterize the real aquifer and to generate conditioning data. We first determined the *K* of the four types of sands from the extracted core using a constant head permeameter [Klute and Dirksen, 1986]. We also conducted slug tests at each of the 48 ports that utilized an external well connected to the port.

After completing the slug tests, cross-hole pumping tests were conducted at 46 ports. Ports 36 and 38 experienced minor well screen damage so we do not pump from these ports. Rather, they are only used for head observations. For the cross-hole tests, pumping rates remained constant during the test duration. We used a pumping rate that is consistent with the synthetic case. Table 9.2 provides the pumping rate and duration for the tests interpreted in this paper. Out of the 46 pumping tests, we selected 8 tests [pumping at ports 2, 5, 14, 17, 32, 35, 44, and 47] for hydraulic tomography that we describe later. One test conducted by pumping at port 46 was reserved for validation purposes. We selected the 8 pumping test data for analysis in this paper as pumping took place in 2 vertical columns which represent vertical wells, a situation which could be readily replicated in the field.

Table 9.2: Summary of pumping rates and duration of cross-hole pumping tests used for hydraulic tomography and its validation. Note that tests with pumping taking place at ports 2, 5, 14, 17, 32, 35, 44, and 47 were used in hydraulic tomography, while the test at port 46 was used for validation.

Pumping Port #	Pumping Rate [ml/s]	Pumping Duration [s]
2	3.07	165
5	3.10	205
14	3.17	242
17	3.17	242
32	2.97	243
35	2.92	181
44	3.12	212
46	3.00	254
47	3.12	166

We also conducted 9 flow-through experiments through the entire sandbox to obtain the effective hydraulic conductivity [K_{eff}] of the entire sandbox under steady-state unidirectional flow conditions. Specifically, each of these 9 experiments was conducted by changing the height of the reservoirs on the both sides of the sandbox. After the flow reached steady state, we measured volumetric discharge from one side of the sandbox. The difference between the heights of the water column in the two constant head reservoirs was measured to determine the hydraulic gradient. Further details of the hydraulic experiments conducted in the sandbox can be found in Craig [2005], Illman et al. [2007] and Liu et al. [2007].

9.4 ESTIMATION OF *K* FROM SYNTHETIC AND LABORATORY EXPERIMENTAL DATA SETS

9.4.1Estimation of synthetic *K* data

We estimate *K* from synthetically generated core, slug, and single-hole data for conditioning the *K* tomograms later. The core *K* estimates were obtained by simply reading off the *K* value assigned to the synthetic *K* distribution [Figure 9.1] at the locations where the actual cores were extracted. This assumes that there is no disturbance to the cores and that core *K* estimates are completely accurate [i.e., no experimental error].

We also conducted synthetic slug tests at each of the 48 ports on the computer using MMOC3 and analyzed the data by manually calibrating MMOC3 by treating the model domain to be a three-dimensional, homogeneous medium. We also considered existing analytical solutions to interpret the data but decided against using them for consistency. That is, the synthetic slug test data were generated with the forward model MMOC3, so it would be best to interpret the data through manual calibration using MMOC3. The numerical grid used for the interpretation of synthetic data was identical to the one used for hydraulic tomography. Boundary conditions for the simulations involving manual calibrations were identical to the numerical simulations for synthetic data generation and to the real experiments. The numerical simulations were conducted by raising the initial head at the elements corresponding to the slugged port and monitoring the corresponding decay in the head profile.

We then conducted synthetic pumping tests at each of the 48 ports on the computer using MMOC3. For each steady state simulation, hydraulic head data were collected from all ports. We analyzed the steady state head records at the pumping ports by manually calibrating MMOC3 and assuming the aquifer is homogeneous. The numerical setup for the calibration is identical to the slug test analysis. The *K* values obtained in this manner using MMOC3 yielded local or single-hole estimates of *K*. These results are denoted as the single-hole results.

Table 9.3 summarizes the results from all synthetic data sets. The mean estimates were obtained by computing the arithmetic mean of the natural logarithm transformed data. The variance was likewise computed using the natural logarithm transformed data set. In Table 9.3, we see that, in general, the mean values of the slug and single-hole test values are larger than that of the core values, which suggests a scale effect [e.g., Illman and Neuman, 2001; 2003 and Illman, 2006]. Examination of Table 9.3 also shows that the variance of ln *K* [$\sigma^2_{\ln K}$] varies from one type of test to the next with variance decreasing with the increasing scale. This is because the support volume of each estimate increases from the core, slug, and single-hole tests. As the sample volume increases, *K* is averaged over the investigated volume.

Table 9.3: Summary of hydraulic properties determined from core, slug, and single-hole pumping test data through synthetic simulations.

Test type	*N*	$\ln K$	
		$\overline{\ln K}$ [$K \sim$ cms^{-1}]	$\sigma^2_{\ln K}$
Core	48	-2.166 [1.15× 10^{-1}]	1.456
Slug	48	-1.906 [1.49× 10^{-1}]	0.161
Single-hole	48	-1.953 [1.42× 10^{-1}]	0.083

9.4.2 Estimation of laboratory *K* data

We then determined the *K* of the four types of sands from the horizontal cores in the laboratory. The extracted cores had dimensions of 1.28 cm in diameter and 10.16 cm in length. These cores were then attached to a custom-made constant head permeameter for determination of *K*. Details of the core extraction method and the design of the constant head permeameter are provided in Craig [2005]. The *K* values from cores are calculated using Darcy's law. We report the arithmetic mean of 48 natural logarithm transformed *K* values in Table 9.4.

We also obtained *K* estimates from slug tests conducted at each of the 48 ports. Due to the small size and configuration of the ports on the sandbox, an external well was attached to the ports instead of boring vertical wells into the sandbox. A slug was introduced to perturb the water level in the horizontal well connected to the port and the corresponding recovery was monitored using a pressure transducer. Originally, Craig [2005] analyzed the slug test data in an identical fashion described for the synthetic slug test data, but the interpretation was done using VSAFT2, a GUI version of the MMOC3 code, available for free at http://tian.hwr.arizona.edu/yeh/. Results from these simulations are summarized in Table 9.4. Results obtained revealed that the *K* values were several orders of magnitude smaller than the core values. We suspected that the data are affected by skin effects and wellbore storage. In fact, we investigated the issue further by conducting additional experiments to examine the effects of the number of cuts on the head response to slug tests. In particular, slug tests were conducted in a separate flow cell with tubes consisting of different number of cuts [2–8]. This effort revealed that the head response stabilizes after six cuts were made on the well. Therefore, all wells in the sandbox discussed in this paper were made by making six cuts. Despite these efforts, the *K* values determined from slug tests analyzed with VSAFT2 were very low, thus we questioned their reliability.

Due to the very low *K* values obtained from the analysis of the slug tests with VSAFT2, the interpretation technique was re-evaluated. The slug test data was fitted automatically to the Bouwer-Rice [1976] solution using AQTESOLV software [http://www.aqtesolv.com/]. This software allows the user to accurately model the screened length and radius of the horizontal well and the radius of the external well [casing radius]. This allows the solution to account for the volume of water displaced by the slug in the external well. We report the results from 48 matches in Table 9.4. The results obtained show that the *K* values are on the same order of magnitude as those from the other measurement methods, thus we deem them to be reliable and use them in this paper.

We then estimated *K* from 2 pumping tests. The data sets were analyzed in several ways. First, we analyzed the 48 drawdown-time data sets induced by pumping at port 22 and those caused by pumping at port 28 by manually calibrating VSAFT2 and assuming the aquifer is homogeneous. The numerical setup for the calibration is identical to the slug test analysis. For the pumping test at port 28 [located in 20/30 sand], all 47 cross-hole intervals were matched and 1 single-hole match was made which yielded a total of 48 estimates for that pumping test. The pumping test at port 22 [located in F-75 sand] also yielded 47 cross-hole and 1 single-hole match for observed and simulated drawdown. Analysis of the two pumping tests thus yielded 96 estimates of *K* for the equivalent homogeneous medium. These two tests will be denoted as cross-hole tests hereafter.

We analyzed the drawdown-time data at all 48 pumping ports using VSAFT2 to yield local or single-hole estimates of *K*. These results are denoted as the single-hole results.

The 9 unidirectional flow-through experiments were analyzed by applying Darcy's law to obtain the K_{eff}. Table 9.4 summarizes the results from all these tests computed in a similar manner to Table 9.3. In Table 9.4, we see that, in general, the mean values of the cross-hole and flow-through values coincide in this sandbox. However, core, slug, and single-hole test values are noticeably smaller suggesting a scale effect. As mentioned earlier, the slug test values from analysis with VSAFT2 are considerably lower, so we conclude that this analysis method is not reliable.

We note that pumped port data during cross-hole pumping tests [i.e., single-hole data] and slug test data could be subjected to borehole storage, skin, and other nonlinear effects which could complicate the analysis. However, we find that the observation port data are not subjected to these complications, thus they could be utilized for analysis. This is because the K estimates from cross-hole tests in the observation well are very close to the overall K value derived from the flow-through experiments suggesting that these estimates are less affected by near well effects. Therefore, we conclude that the cross-hole observation well data are reliable and we retain them in our analysis. Examination of Table 9.4 also shows that $\sigma^2_{\ln K}$ varies from one type of test to the next with variance decreasing with the increasing scale. This behavior is similar to the results from the synthetic hydraulic test data [Table 9.3] and is attributed again to the increase in support volume as the scale of the test increases.

Table 9.4: Summary of hydraulic properties determined from core, slug, single-hole, cross-hole pumping test data and flow-through experiments conducted in the laboratory.

Test type	N	$\ln K$	
		$\overline{\ln K}$ [$K \sim$ cms^{-1}]	$\sigma^2_{\ln K}$
Core	48	-2.166 [1.146×10^{-1}]	1.498
Slug [VSAFT2]	40	-10.692 [2.273×10^{-5}]	0.431
Slug [AQTESOLV]	46	-1.910 [1.481×10^{-1}]	0.521
Single-hole	48	-2.835 [5.872×10^{-2}]	0.589
Cross-hole	96	-1.757 [1.726×10^{-1}]	0.074
Flow-through	9	-1.757 [1.726×10^{-1}]	0.002

9.5 GENERATION OF REFERENCE *K* TOMOGRAMS USING SYNTHETIC AND REAL PUMPING TEST DATA

Here, we briefly describe the reference K tomograms computed using synthetic and real data by Illman et al. [2007]. The purpose of constructing a reference K tomogram using synthetic data is to examine the ability of the SSLE algorithm to image the heterogeneity pattern under optimal conditions without experimental errors and with full control of forcing functions [initial and boundary conditions as well as source/sink terms]. The reference K tomogram is then used to assess the effects of signal-to-noise ratio and conditioning. We also obtain a reference K tomogram using real data so that we can later compare these results to those generated with noisy data and those conditioned to available core, slug, and single-hole K estimates.

The reference K tomogram [Figure 9.2a] using synthetic data was computed through the inversion of 8 synthetic cross-hole test data generated via numerical simulations. The computation was done sequentially by including pumping tests conducted at ports 47, 44, 35, 32, 17, 14, 5, and 2, in that order. All 48 ports were used in steady-state hydraulic tomography. This result clearly shows that the SSLE algorithm is capable of capturing the correct position of the low K blocks, its morphology, dimensions, and other details of aquifer heterogeneity such as

windows in low *K* strata that could provide continuous pathways for contaminant transport. However, the low *K* blocks and their dimensions at the bottom of the sandbox [Figure 9.2a] are not captured as clearly as the other blocks positioned higher in the sandbox when compared to Figure 9.1. This could perhaps be due to the fact that there are no observation ports beneath the low *K* blocks.

The reference *K* tomogram [Figure 9.2b] using real data was computed through the inversion of 8 sets of cross-hole pumping test data collected in the laboratory sandbox aquifer. We utilized the pumping tests taking place at the same ports as in the synthetic case and in the same order [i.e., 47, 44, 35, 32, 17, 14, 5, 2,]. It is important to note that no conditioning data were used to generate this reference *K* tomogram. Figure 9.2b shows the best *K* tomogram obtained by Illman et al. [2007] after various error and bias reduction schemes were applied to the raw data set collected in the sandbox. The main reasons for applying the error and bias reduction schemes were to remove outliers and excessively noisy data. Briefly, the error reduction scheme consisted of: 1] removal of pumped well data thought to be affected by a skin effect at the pumped port; 2] correcting for drift or offset in the pressure transducers; and 3] accounting for slight variations in boundary conditions from one pumping test to the next in the SSLE algorithm. Figure 9.3b shows that the heterogeneity pattern consisting of low *K* blocks are mostly captured, except for the bottom 2 blocks. There is also a thin and continuous high *K* zone at the top of the image that is not visible on the reference *K* tomogram [Figure 9.2a]. This is likely due to the lack of compaction of sands near the top of the sandbox.

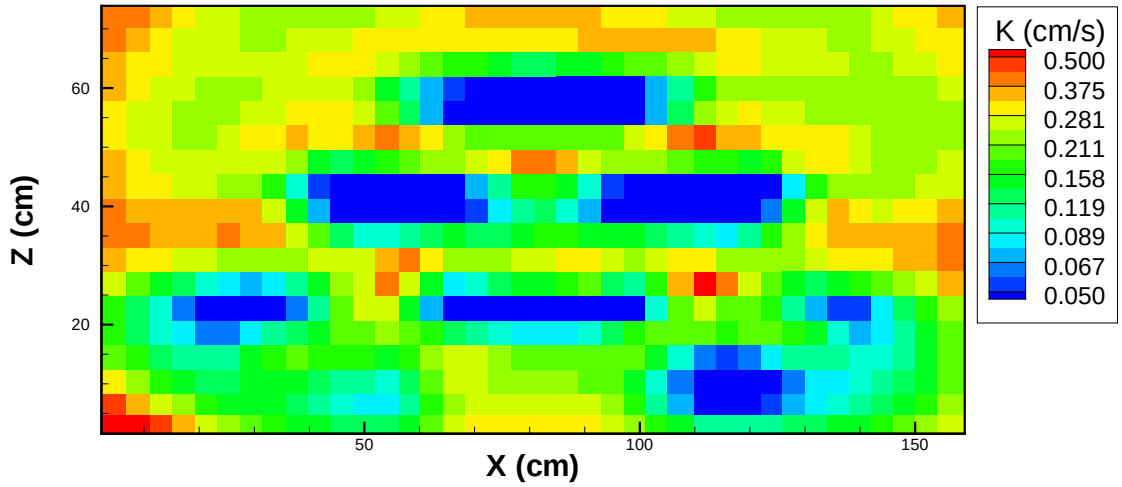
9.6 EFFECTS OF PUMPING RATE, ORDER OF TEST DATA INCLUDED IN SSLE AND CONDITIONING ON SYNTHETIC AND REAL *K* TOMOGRAMS

We next describe the effects of: 1] varying the pumping rate, which affects the signal-to-noise ratio of head data collected during hydraulic tomography; 2] the order of test data included in the SSLE algorithm, and 3] conditioning on both synthetic and real *K* tomograms. The reference *K* tomograms from the synthetic [Figure 9.2a] and real [Figure 9.2b] cases will serve as our baseline results for purposes of comparison.

9.6.1 The effect of pumping rate and signal-to-noise ratio

We first examine the effect of varying the signal-to-noise ratio in the cross-hole pumping test data on the resulting *K* tomograms. A larger signal-to-noise ratio data can be generated by increasing the pumping rate and/or decreasing the noise level through signal conditioning techniques. Here, we vary the signal-to-noise ratio by conducting pumping tests at 2 different rates. The high pumping rate case was already presented by Illman et al. [2007] and shown here as Figure 9.2a-b. We present below the low flow rate case. The flow rate used for the synthetic simulations was 1.6 ml/s, while for the real case it varied for each test. It is evident from the synthetic case [Figure 9.3a] that the pumping rate and hence the signal-to-noise ratio has very little effect on the quality of the synthetic *K* tomogram when compared to the reference *K* tomogram [Figure 9.2a]. However, for the real case [Figure 9.3b], it has a noticeable detrimental impact when compared to Figure 9.2b, because of the larger noise level in the hydraulic head data associated with a lower pumping rate. We emphasize that the SSLE algorithm can overcome the effects of the noise level through loosening of the convergence criteria. However, loosening of the convergence criteria results in a smoother result that approaches the effective *K* value, which means that the heterogeneities become less well-defined.

a]



b]

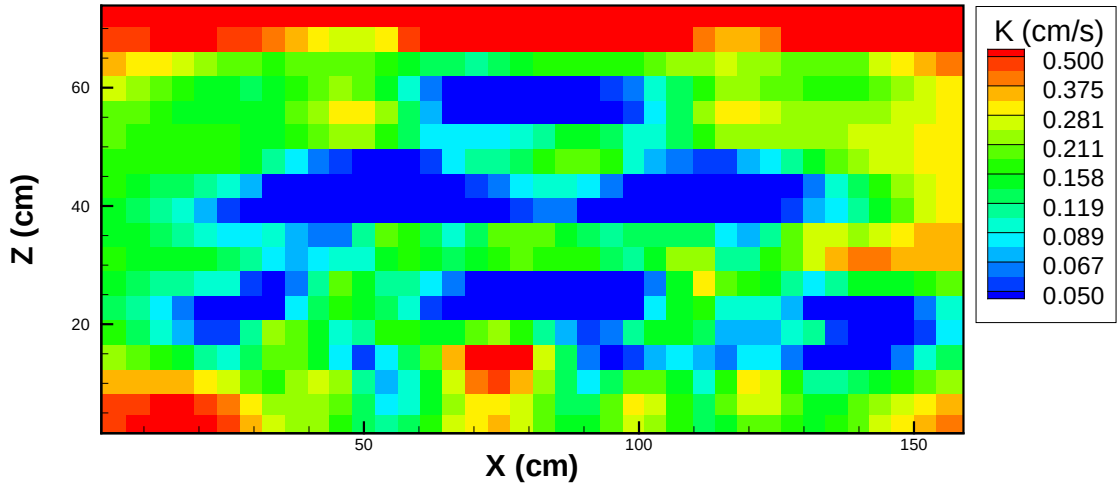
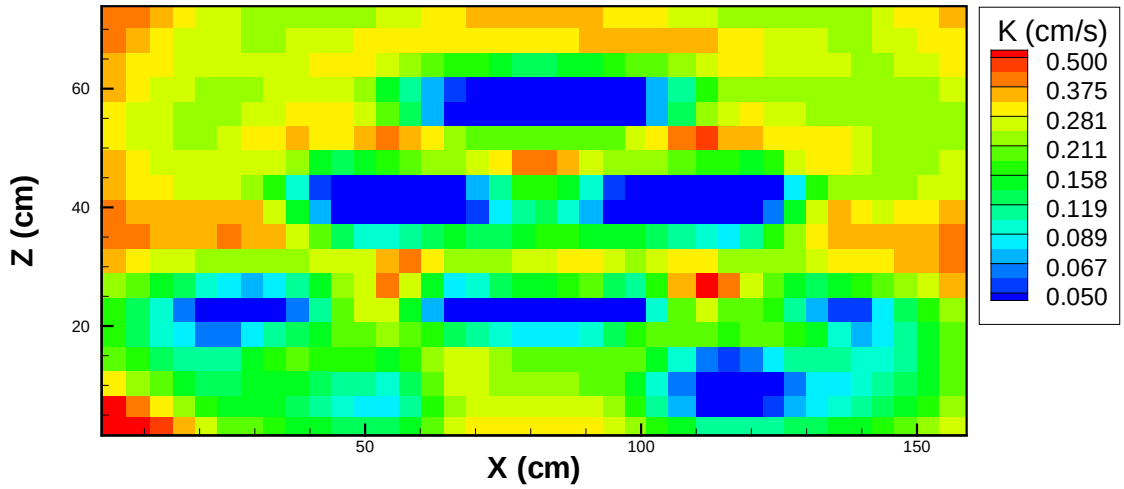


Figure 9.2: Reference *K* tomograms generated by sequentially inverting 8: a] synthetic and b] real cross-hole pumping tests [modified after Illman et al., 2007].

a]



b]

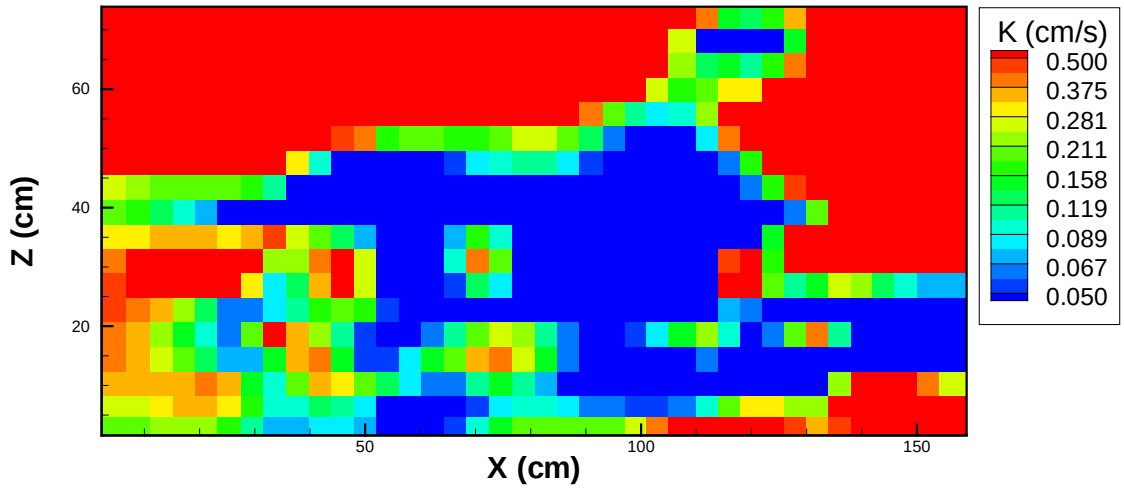


Figure 9.3: *K* tomograms obtained from low Q pumping test data with lower signal-to-noise ratio for the: a] synthetic and b] real cases.

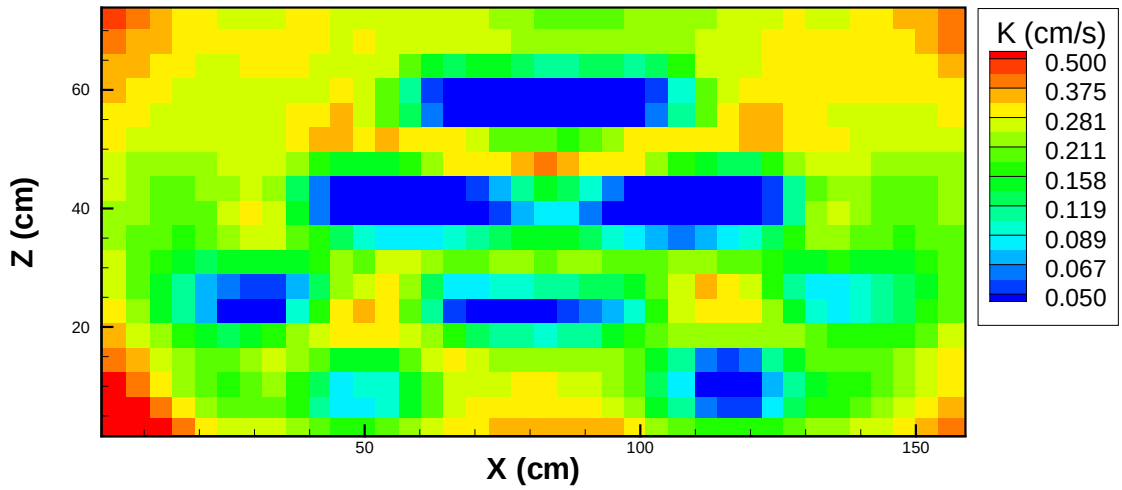
9.7.2 The effect of order of test data included sequentially in the SSLE algorithm

We next vary the order of test data included into the SSLE algorithm. For the synthetic case [Figure 9.4a], we included the tests with pumping taking place in the order of ports 2, 5, 14, 17, 32, 35, 44, and 47. Figure 9.4a shows that this has little effect on the computed *K* tomogram when compared to Figure 9.2a. We tried various combinations and found that the order of test data included has very little effect on the *K* tomograms.

In contrast, the order of test data included has a large impact on the computation of the real *K* tomogram [Figure 9.4b] when compared to Figure 9.2b. This is because some head records from each pumping test are noisier than others. We found that the data most devoid of noise were found near the bottom of the sandbox, while the noisiest data were usually located near the top of the sandbox. This is due to the fact that a higher water column sits on the pressure transducers near the bottom of the sandbox, which causes the pressure transducers to be less

affected by noise. Also, a stronger signal is generated near the bottom of the sandbox because of the no-flow boundary and due to the superposition principle.

a]



b]

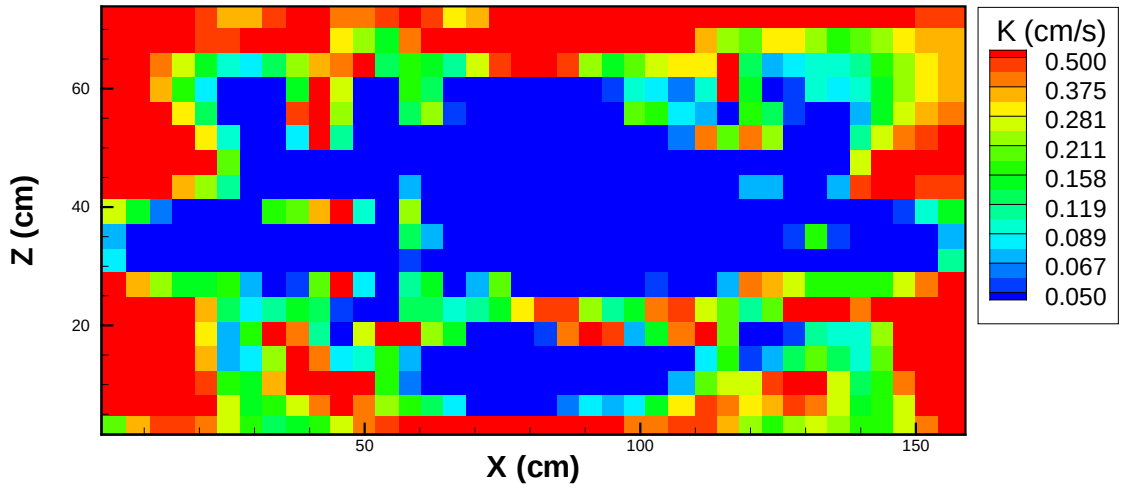


Figure 9.4: *K* tomograms obtained by varying the order of pumping test data included in the SSLE algorithm for the: a] synthetic and b] real cases.

We tried various combinations of test data included in the SSLE algorithm and found that the order of test data included has a large impact on *K* tomogram generation. Our investigation showed that including the cleanest data first in the SSLE code and including the noisiest data later appeared to improve the results. The main reason for this is because SSLE uses a weighted linear combination of the differences between the estimated heads and measured heads to improve the tomogram. Therefore, an unrefined *K* distribution obtained in the beginning of inversion process will generate larger differences between the estimated and measured heads than a refined *K* distribution causing the inverse solution to become unstable as additional data

are included into SSLE. Therefore, including the cleanest data with higher signal-noise ratio in the beginning of the inversion process tends to improve the results.

9.6.3 The effect of conditioning using core *K* data

Conditioning is generally thought to improve results of conditional stochastic simulations and hydraulic tomography [e.g., Yeh and Liu, 2000]. We investigate this observation through the use of various conditioning data with the SSLE algorithm. For this investigation, 48 conditioning points consisting of various data types [core, slug, and single-hole *K* estimates] are placed at the ports shown on Figure 9.1.

We first include core *K* estimates into the computation of the *K* tomogram. Figure 9.5a shows that conditioning of the *K* tomogram with core *K* data dramatically improves the quality of the synthetic tomogram when compared to Figure 9.2a. In particular, the low *K* blocks throughout the tomogram appear clearer compared to the background aquifer material. We also see that the dimensions of the low *K* blocks approach the true case [Figure 9.1] and more importantly, the low *K* blocks at the bottom of the sandbox become clearer.

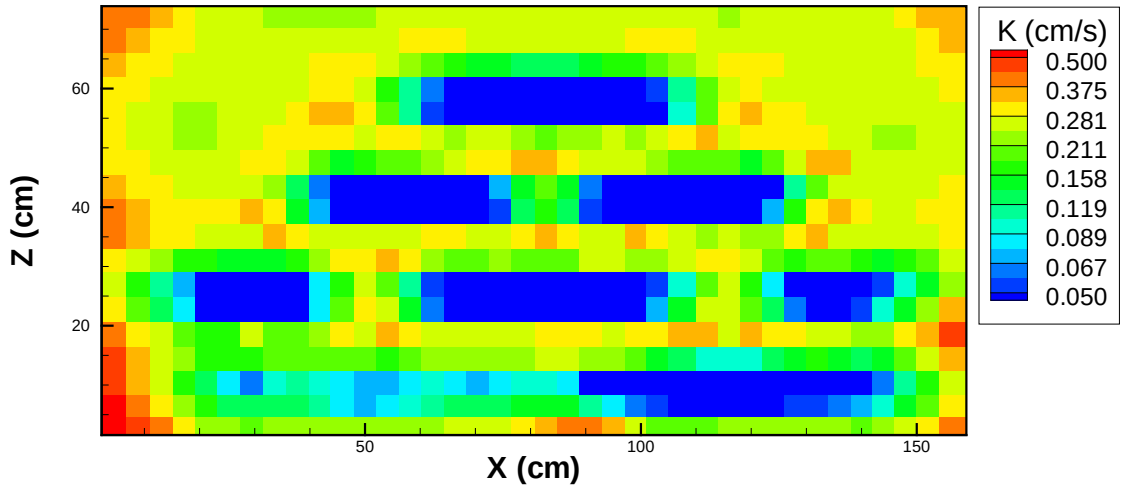
We next condition the real *K* tomogram with real core *K* data obtained in the laboratory at 48 port locations. Figure 9.5b again shows a dramatic improvement of the *K* tomogram in comparison to the reference *K* tomogram [Figure 9.2b]. In particular, we see the low *K* block appearing clearly at the bottom of the sandbox.

9.6.4 The effect of conditioning using slug *K* data

We next condition the *K* tomogram with available synthetic and real, slug *K* data. For the synthetic case [Figure 9.6a], we see a slight deterioration in the *K* tomogram. This is because the slug *K* estimates represent a larger support volume than the core *K* estimate. In fact, results from forward simulation of slug tests using VSAFT2 not shown here confirmed this finding. That is, the head distribution resulting from the synthetic slug test generated on the computer revealed a volume of influence that is larger than the numerical grid used in the computation of the *K* tomogram. Therefore, inclusion of slug *K* data as conditioning points causes smoothing of the *K* tomogram.

Conditioning of the real *K* tomogram [Figure 9.6b] using 48 slug test *K* data leads to a slight deterioration in the tomogram when compared to Figure 9.2b. In particular, we see the smoothing of the low *K* blocks throughout the aquifer as in the synthetic case.

a]



b]

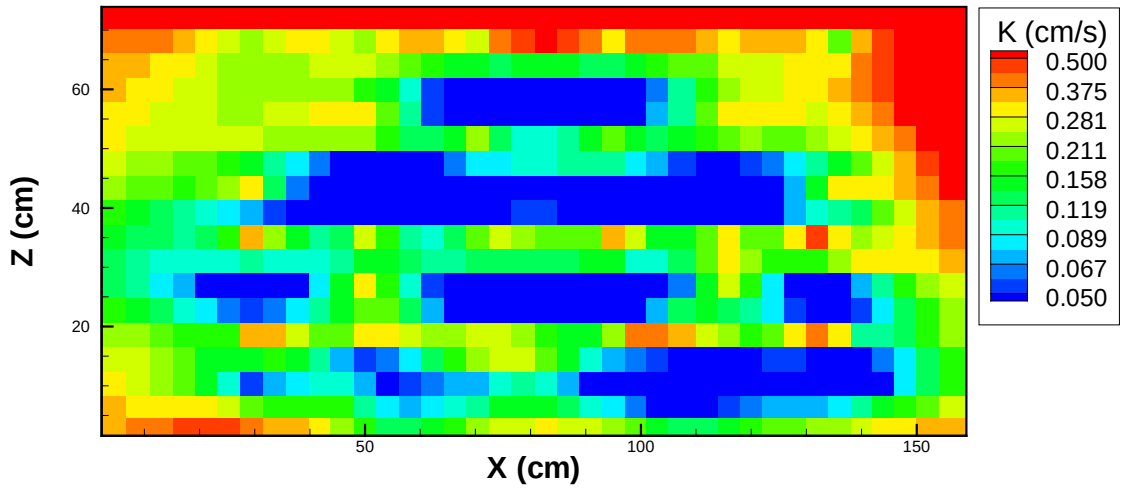
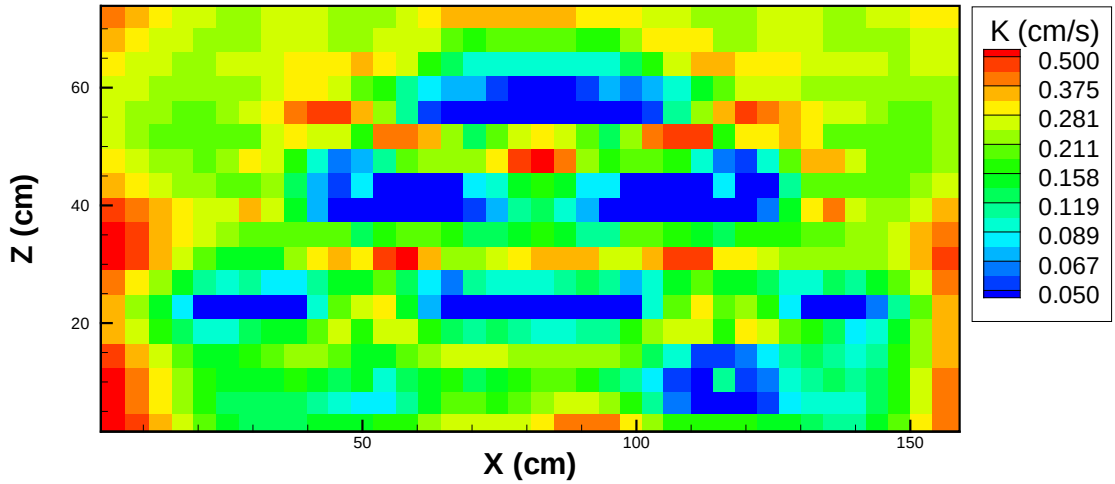


Figure 9.5: *K* tomograms obtained by conditioning with core *K* values for the: a] synthetic and b] real cases.

a]



b]

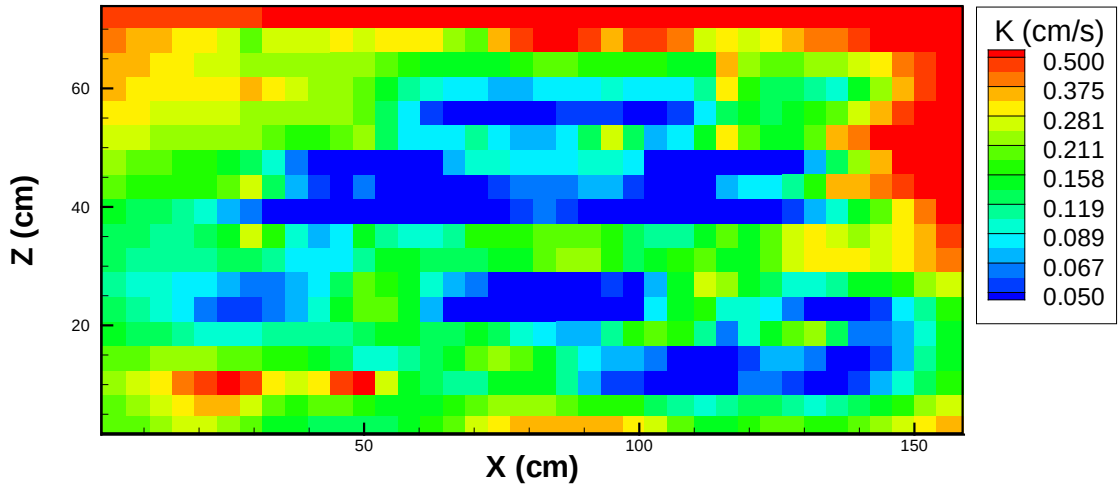


Figure 9.6: *K* tomograms obtained by conditioning with slug *K* values for the: a] synthetic and b] real cases.

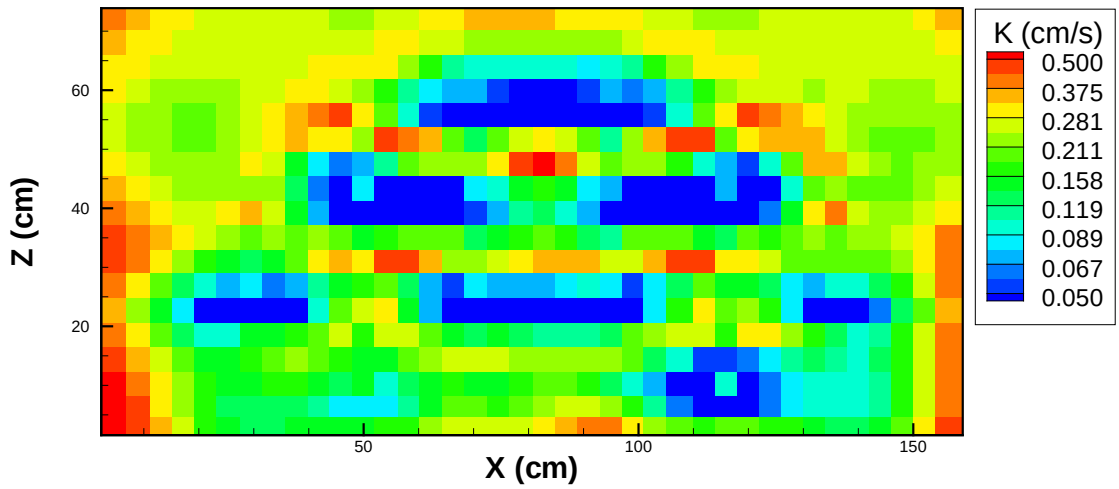
9.6.5 The effect of conditioning using single-hole *K* data

We next condition the *K* tomogram with available synthetic and real single-hole *K* data. As in the case of conditioning with slug *K* data, the synthetic result [Figure 9.7a] shows a slight deterioration and smoothing of the *K* tomogram when compared to Figure 9.2a. This is because single-hole *K* represents a much larger support volume. VSAFT2 simulations not shown here reveal that the cone of depression essentially reaches all boundaries at steady state. Therefore, single-hole *K* estimates represent a considerably larger area around the pumping port. Therefore, inclusion of single-hole *K* data as conditioning points, causes smoothing of the *K* tomogram. It is of interest to note that the *K* tomograms from Figures 9.6a and 9.7a look similar.

Conditioning of the real *K* tomogram [Figure 9.7b] using single-hole *K* data leads to a slight deterioration in the tomogram when compared to Figure 9.2b. In particular, we see the

smoothing of the low *K* blocks throughout the aquifer. These results are similar to those obtained through conditioning with slug *K* test data [Figure 9.6b].

a]



b]

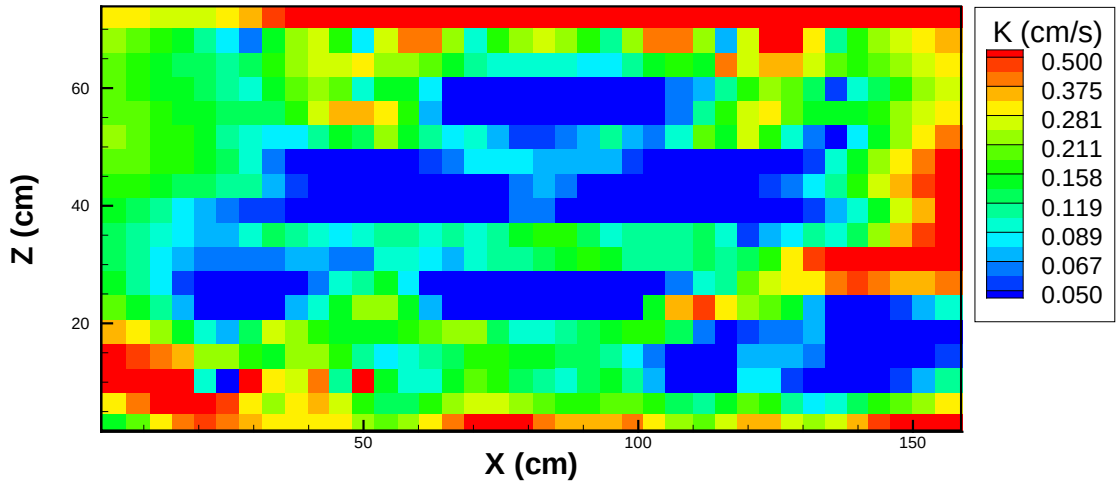


Figure 9.7: *K* tomograms obtained by conditioning with single-hole *K* values for the: a] synthetic and b] real cases.

9.7 HOW GOOD ARE THE *K* TOMOGRAMS?

As shown in the previous section, one could assess the effects of varying signal-to-noise ratio and conditioning on the computed *K* tomograms by comparing the results with the reference *K* tomograms [Figure 9.2a-b]. However, visual comparison of the tomograms is qualitative. A more rigorous and quantitative approach to validate the computed *K* tomograms is necessary.

Illman et al. [2007] and Liu et al. [2007] suggested that an appropriate validation approach is to test the predictability of the hydraulic head estimates under different flow scenarios. In order to do this, we validate the *K* tomograms by simulating an additional

pumping test that was not used in the inversion. The forward simulation of this independent test will yield heads at the monitoring ports, which are then compared to the actual head data through a scatter plot. We adopt this approach to validate the synthetic and real *K* tomograms. In particular, for the synthetic case, we simulate another pumping test using the *K* tomograms [Figure 9.2a, 9.3a, 9.4a, 9.5a, 9.6a, and 9.7a] and the true *K* distribution [Figure 9.1]. In all simulations, pumping takes place at port 46. A comparison of the steady state head values through a scatter plot then should validate the *K* tomograms.

Figures 9.8a-f shows the results of this comparison for the synthetic case. It shows that the comparison is excellent for all cases for monitoring ports far away from the pumping port, where the drawdown is smaller. We notice that simulated head values are larger than the observed ones, when the head values are small for all cases. However, the *K* tomogram obtained by conditioning using core *K* data yields the smallest discrepancy between the simulated and observed head values near the pumping port. The comparison of the heads near the pumped port is worse for the other cases. This is consistent with the fact that uncertainty in the predicted drawdown grows with the mean gradient according to stochastic analysis and confirms the earlier finding by Liu et al. [2007].

Two criteria, the average absolute error norm [L1] and the mean squared error norm [L2], were used to quantitatively evaluate the goodness-of-fit between the simulated and observed hydraulic head responses. The smaller the L1 and L2 norms are, the better the estimate is. Specifically, when one compares the results [Figures 9.8b-f] to the unconditioned case [Figure 9.8a]: 1] Figure 9.8b shows that the effect of low pumping rate has a negligible impact on results; 2] Figure 9.8c shows that changing the order of pumping tests slightly deteriorates the results; and c] Figures 9.8d-f show that conditioning tends to improve the results for the synthetic cases.

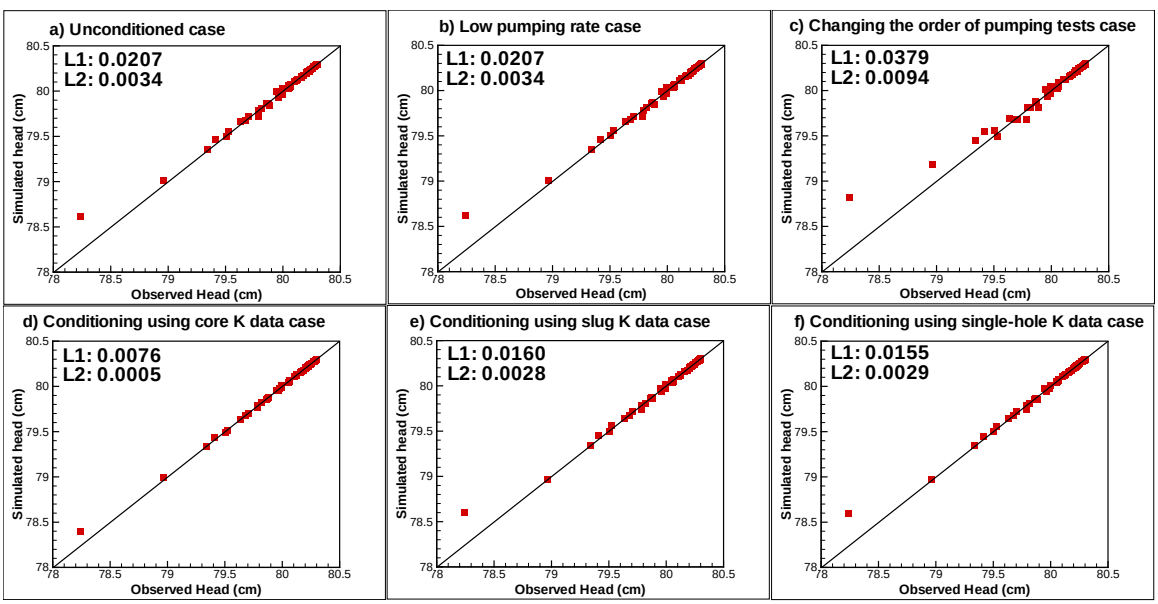


Figure 9.8: Scatter plot of simulated drawdown versus observed drawdown at steady state for cross-hole hydraulic tests at port 46. The head values were obtained by simulating cross-hole pumping tests on the computer using the synthetic *K* tomograms [Figures 9.2a, 9.3a, 9.4a, 9.5a, 9.6a, and 9.7a]. The observed head values from the true *K* field [Figure 9.1] were also simulated on the computer.

We also utilize this validation approach on the real *K* tomograms. For this, we utilize the various real *K* tomograms [Figures 9.2b, 9.3b, 9.4b, 9.5b, 9.6b, and 9.7b] and simulate an independent cross-hole pumping test with pumping taking place at port 46. Pumping test at port 46 was chosen for validation purposes because it was not used in the construction of the *K* tomograms and it is also a pumping test with the cleanest data. Figure 9.9a-f shows the results of this comparison. According to this figure, the data pairs are scattered along the 45 degree line, indicating predicted head distributions generally are statistically unbiased in comparison with the observed except for those generated with a low pumping rate [i.e., the results with low signal-to noise ratio] [Figure 9.9b]. This plot and the quantitative measures [L1 and L2 norms] show that the comparison is consistently very good providing us with further confidence that SSLE can provide an unbiased estimation of the head distribution. Note that the simulated heads will not necessarily match the observed ones perfectly as they should, due to the fact that the tomograms are conditional effective *K* fields and also due to noise in the observations.

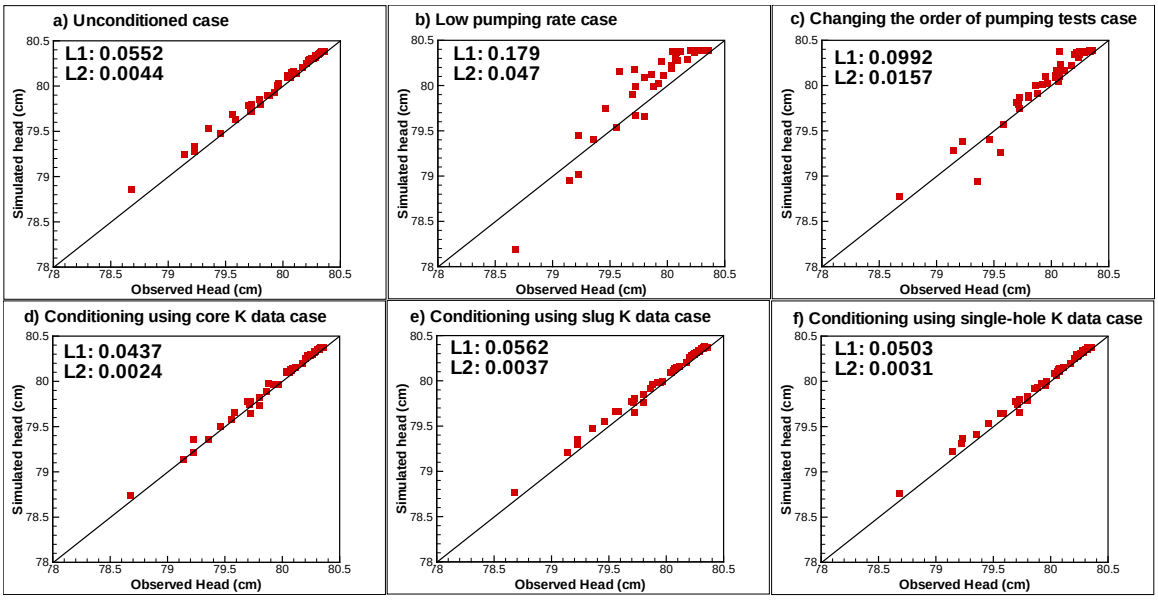


Figure 9.9: Scatter plot of simulated drawdown versus observed drawdown at steady state for cross-hole hydraulic tests at port 46. In both cases, the simulated values were obtained through synthetic simulations of the cross-hole test conducted using the real *K* tomograms [Figures 9.2b, 9.3b, 9.4b, 9.5b, 9.6b, and 9.7b]. The observed drawdown values are from an actual cross-hole pumping test conducted at port 46.

10. LABORATORY SANDBOX VALIDATION OF TRANSIENT HYDRAULIC TOMOGRAPHY

10.1 INTRODUCTION

Zhu and Yeh [2005] developed an algorithm for transient hydraulic tomography through the use of the Sequential Successive Linear Estimator [SSLE]. Their approach combines the traditional geostatistical approach and governing flow physical principles to interpolate and extrapolate at locations where samples are not available. As a consequence, the SSLE as implemented in hydraulic tomography yields more realistic estimates than kriging and deterministic/zone-based inverse modeling approaches that consider principles of flow and use one pumping or injection data set only. They showed that K and S_s distributions can be obtained through data sets from numerically simulated pumping tests in synthetic heterogeneous aquifers. They suggested that the transient hydraulic tomography is a potentially cost-effective and high-resolution technique for mapping spatial distributions of the K and S_s in aquifers.

While various algorithms for hydraulic tomography have been developed and some of them have been validated in sand box experiments [Liu et al., 2002; and Illman et al., 2007, 2008], to date, comprehensive validation of the transient hydraulic tomography has not been done either in the laboratory or the field setting. A field validation is the ultimate goal, but prior to that, laboratory validations are necessary in which the baseline heterogeneity is largely known, and all forcing functions and errors can be controlled as opposed to field applications. The main objectives of this section are: 1] to sequentially invert cross-hole pumping test data obtained in a synthetic aquifer with deterministic heterogeneity in sandbox experiments to obtain K and S_s tomograms using the transient hydraulic tomography algorithm of Zhu and Yeh [2005] and 2] to validate the K and S_s tomograms using various independent data and methods.

10.2 METHODS FOR CHARACTERIZATION OF THE SANDBOX

Different hydraulic tests were performed to characterize the hydraulic parameters in the sandbox, including determination of K using core samples, in-situ slug tests, and in-situ pumping tests. Details to each test were described previously and Table 10.1 summarizes the results from all these tests. The mean estimates were obtained by computing the arithmetic mean of the natural logarithm transformed data. The variance was likewise computed using the natural logarithm transformed data set. We also calculated a volume-weighted mean and variance of the core values which are also listed in Table 10.1. The purpose of computing the volume-weighted mean and variance of the core K values was so that these values are upscaled to the size of the finite element grid used for the inversion so that we can compare them later.

Table 10.1: Summary of hydraulic properties determined from core, slug, single-hole, cross-hole pumping test data and flow-through experiments.

Test type	N	$LN-K$		$LN-S_s$	
		$\overline{LN-K}$ [K ~ cms^{-1}]	σ^2_{LN-K}	$\overline{LN-S_s}$ [S_s ~ cms^{-1}]	$\sigma^2_{LN-S_s}$
Core	48	-2.166 [1.146×10^{-1}]†	1.498†	N/A	N/A
Slug	40	-10.692 [2.272×10^{-5}]	0.431	N/A	N/A
Single-hole	9	-3.174 [4.182×10^{-2}]	0.570	-7.960 [3.492×10^{-4}]	1.897
Cross-hole	95	-1.757 [1.726×10^{-1}]	0.074	-8.378 [2.300×10^{-4}]	0.047
Flow-through	9	-1.757 [1.725×10^{-1}]	0.002	N/A	N/A

† The volume-weighted mean and its corresponding variance are -1.920[1.467×10^{-1}] and 1.560 respectively.

10.3 INVERSE MODEL DESCRIPTION

Mapping of the spatial distribution of hydraulic properties in the sandbox were carried out using a transient hydraulic tomography algorithm developed by Zhu and Yeh [2005] and hydraulic head data from the six pumping tests. A brief description of the algorithm is given below.

The inverse model assumes a transient flow field and the natural logarithm of K [$ln\ K$] and S_s [$ln\ S_s$] are both treated as multi-Gaussian, second-order stationary, stochastic processes. The model additionally assumes that the mean and correlation structure of the K and S_s fields are known *a priori*.

The estimation procedure starts with a weighted linear combination of parameter measurements and transient hydraulic head data at different locations to obtain the first estimate of the parameters. The weights are calculated using the means and covariances of parameters, the covariances of hydraulic heads in space and time, and the cross-covariances between heads and parameters. The first estimate is then used in the mean flow equation to calculate the heads at observation locations and sampling times through a forward simulation. At the end of this forward simulation, the differences between the observed and simulated hydraulic heads are calculated and a weighted linear combination of these differences is then used to improve the previous estimates. Iterations between the forward simulation and estimation continue until the improvement in the estimates diminishes to a prescribed value.

SSLE can handle measurement error through the specification of two convergence criteria in the algorithm. The criteria are the change of variance and pressure heads between two consecutive iterations, which are both set to 0.01 in our model.

The transient hydraulic tomography algorithm developed by *Zhu and Yeh* [2005] allows for the sequential inclusion of pumping test data. Some modifications were made to the code for the present study to account for variations in the constant head boundary conditions from one pumping test to the next, as they are sequentially included.

10.4 INVERSE MODEL PARAMETERS

To obtain the K and S_s tomograms, the synthetic aquifer was discretized into 741 elements and 1600 nodes with element dimensions of 4.1 cm × 10.2 cm × 4.1 cm. The numerical grid is somewhat smaller [161.3 cm by 75.6 cm by 10.2 cm] than the actual dimensions of the sandbox [193.0 cm by 82.6 cm by 10.2 cm] as we only model the portion of the sandbox that contains the porous medium. Both sides and the top boundary were set to be constant head boundary conditions, while the bottom boundary of the sandbox was considered a no-flow boundary. This grid setup is consistent with that was used in the steady state hydraulic tomography by Illman et al. [2007, 2008].

Inputs to the inverse model include initial guesses for the K and S_s, estimates of variances and the correlation scales for both parameters, volumetric discharge [Q_n] from each pumping test where n is the test number, available point [small-scale] measurements of K and S_s, as well as head data at various times selected from the head-time curve. Although available point [small-scale] measurements of K and S_s can be input to the inverse model, we do not use these measurements to condition the estimated parameter fields to test the inversion algorithm.

10.4.1 Hydraulic Parameters K and S_s

A number of methods can be used to obtain the initial guess of K and S_s. One can set an arbitrary value that is reasonable for the geologic medium considered or to estimate the average or effective hydraulic conductivity [K_{eff}] and specific storage [S_{seff}] for an equivalent homogeneous sandbox. If there are small scale data available, then a geometric mean of the available small scale data [i.e., core, slug, and single-hole data] can be calculated. An alternative to this is to use the equivalent hydraulic conductivity and storage estimates obtained through the analysis of cross-hole test data by treating the medium to be homogeneous. Yet, another approach is to use the results from the flow-through experiments to obtain the K_{eff}. We elect to utilize the mean value of the K and S_s obtained from the analysis of cross-hole tests treating the medium to be homogeneous as these estimates are commonly and readily available in real field situations.

10.4.2 Variance and Correlation Scales

The variances and correlation scales of the K and S_s fields are also required inputs to the inverse model. However, estimation of variance always involves uncertainty. A previous numerical study conducted by Yeh and Liu [2000] has shown that the variance has negligible effects on the estimated K using the inverse model. We expect the same for both K and S_s for transient hydraulic tomography. Therefore, we obtain variance estimates from the available small scale data and use this as our input variance in the inverse model for the real data set.

Correlation scales represent the average size of heterogeneity, which is difficult to determine accurately without a large number of data sets in the field. The effects of uncertainty in correlation scales on the estimate based on the tomography are negligible because the tomography produces a large number of head measurements, reflecting the detailed site-specific heterogeneity [Yeh and Liu, 2000]. Therefore, the correlation scales were approximated based only on the average thickness and length of the discontinuous sand bodies.

10.4.3 Transient Hydraulic Head Data

Transient hydraulic head records are required observation data for transient hydraulic tomography. These were obtained from ports that yielded data that were not too noisy. The remaining data then were treated with various error reduction schemes discussed in Illman et al. [2006]. Briefly, the error reduction schemes consisted of accounting for pressure transducer drift and removal of data at the pumping port affected by skin effects, and averaging the drawdown data at steady state. We then calculated drawdown for each port during a pumping test. Once the drawdown is computed, a 6th–order polynomial curve was fit to each transient drawdown record. We then extracted 4 to 5 points that are evenly spaced to capture the transient head record thoroughly. Drawdown curves that could not be properly fitted were manually excluded from the analysis. The average R^2 for the fitted curves picked out for inverse modeling is 0.9963, which means the curves were well fitted. In total, we utilized 6 independent cross-hole tests for the analysis. There were other cross-hole tests available, but those 6 were considered to be the best among the data set. In addition, we selected the pumping tests from these ports as they are evenly distributed along two vertical profiles throughout the sandbox. More specifically, we utilized 4 drawdown data from 19 ports totaling 76 in cross-hole tests #14 and #17, 4 drawdown data from 30 ports totaling 120 in cross-hole test #32, 4 drawdown data from 26 ports totaling 104 points for pumping test #35, 4 drawdown data at 36 ports totaling 144 points for pumping test #44 and 5 drawdown data at 26 ports totaling 130 points for pumping test #47. In total, we utilized 650

drawdown records from 6 different tests in our transient inversions. *Illman et al*. [2006] made use of 2 additional tests for steady state hydraulic tomography. Here we do not use these data as the transient drawdown records are noisy.

10.5 RESULTS FROM TRANSIENT HYDRAULIC TOMOGRAPHY

Figure 10.1a is a drawing of the sandbox showing the synthetic aquifer with each sand type marked. Figures 10.1b-f are the *K* tomograms obtained by inverting head data induced by the six pumping tests [14, 17, 32, 35, 44, 47] conducted in that order. Results from using the first two pumping tests [Figure 10.1b] reveal detailed heterogeneity patterns near the top of the sandbox, where pumping took place. However, little detail to the heterogeneity pattern is revealed near the bottom of the sandbox. This is alleviated with the heterogeneity structure for the entire aquifer appearing when additional cross-hole tests are included sequentially. Upon inclusion of all of the 6 pumping tests in the inversion, a vivid image [Figure 10.1f] of the heterogeneity structure appears. However, the 2 low *K* blocks near the bottom boundary still are not well resolved. In addition, a high *K* zone is apparent near the top of the sandbox. This may be attributed to the lack of compaction of sands near the top.

Despite of the lack of resolution near the bottom, the results collectively show that the inversion algorithm is capable of capturing the pattern of the *K* distribution, which is critical for an analysis of contaminant migration. Another interesting observation that one may make qualitatively is that the transient hydraulic tomography reaches the same quality of result with fewer pumping tests than the steady state approach, which required 8 pumping tests [Illman et al., 2007]. We also visually compared the *K* tomogram resulting from transient hydraulic tomography to that obtained from the steady state hydraulic tomography [figure 10.1g]. This comparison revealed that the tomograms are very similar.

Figure 10.2b - f shows the corresponding S_s tomogram that was estimated simultaneously. In contrast to Figure 10.1b - f, the structure consisting of variable size sand bodies visible in the *K* tomogram is not visible for the S_s tomogram. This can be attributed to the fact that sands of relatively low compressibility [of various sizes] were used to construct the synthetic aquifer. However, a decreasing trend in S_s with depth in the synthetic aquifer is apparent. Physically speaking, this makes sense because the sands in the upper portion are less compressed, while the deeper sands are more compressed due to the stress exerted by the overlying material. This finding perhaps suggests that the *K* values are not significantly correlated with the S_s values in this sand box.

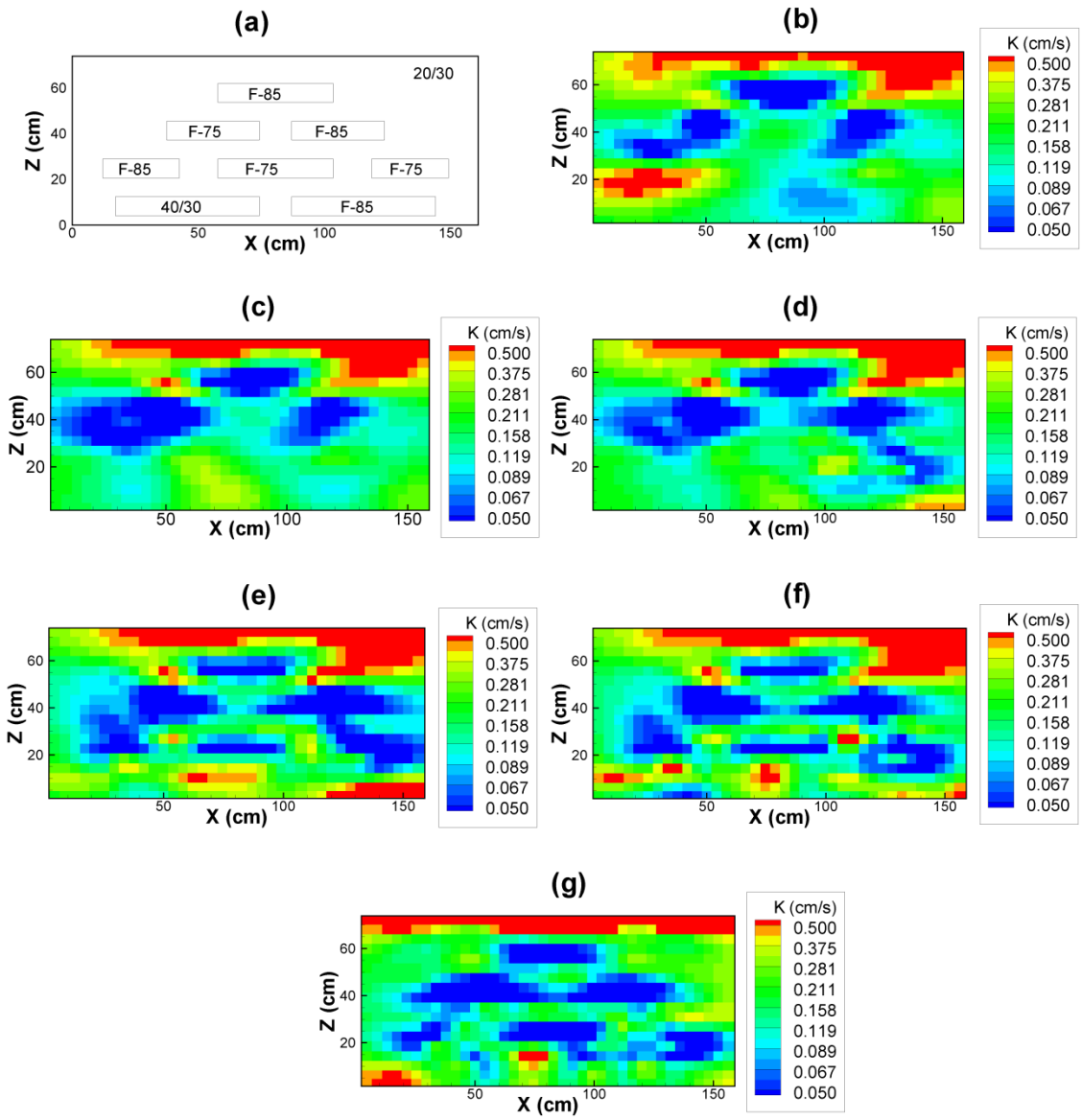


Figure 10.1a-f: a] Sketch of the sandbox with a deterministic heterogeneous aquifer; b] [14, 17 - case 1]; c] [14, 17, 32 - case 2]; d] [14, 17, 32, 35 - case 3]; e] [14, 17, 32, 35, 44 - case 4] and f] [14, 17, 32, 35, 44, 47 - case 5] show the resulting *K* tomograms by sequentially inverting the experimental cross-hole test data. Numbers in parentheses from b] – f] indicate the port numbers used as the pumped well for each cross-hole test.

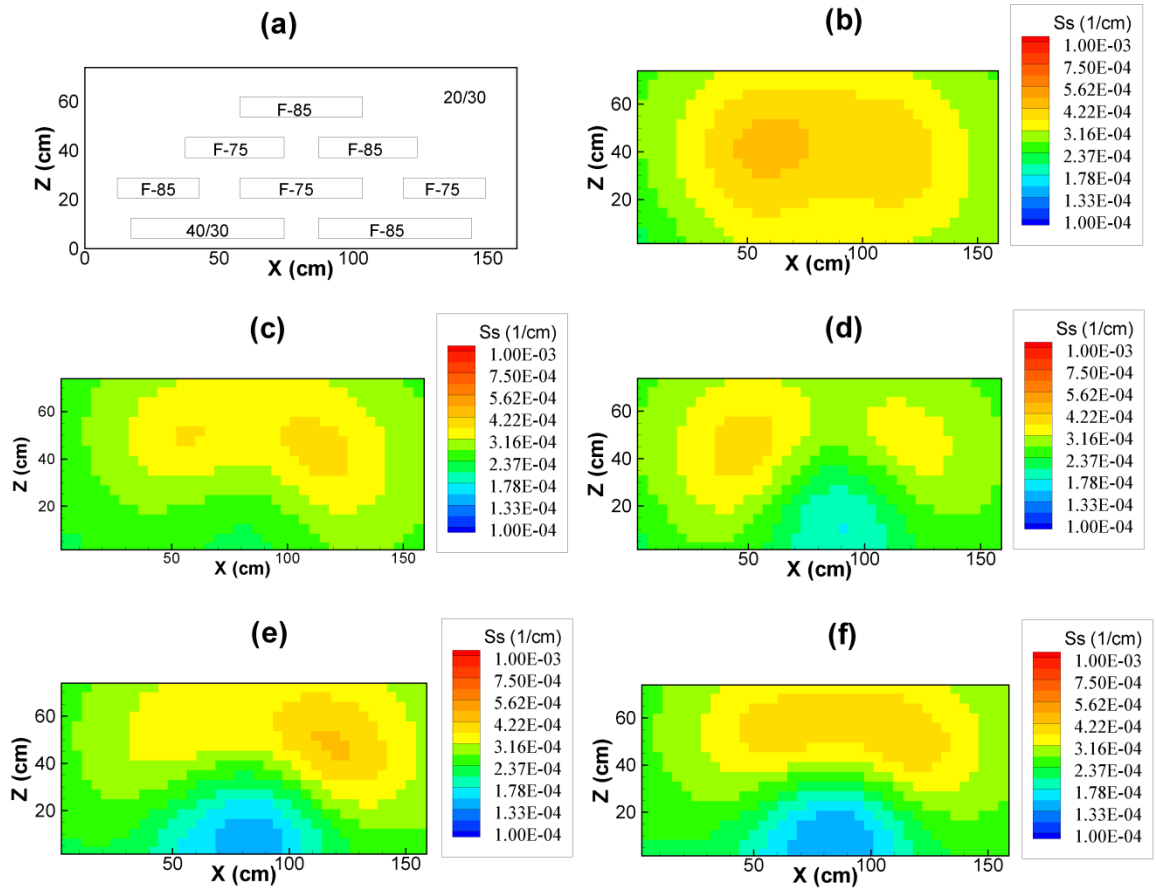


Figure 10.2a-i: a] Sketch of the sandbox with a deterministic heterogeneous aquifer; b] [14, 17 - case 1]; c] [14, 17, 32 - case 2]; d] [14, 17, 32, 35 - case 3]; e] [14, 17, 32, 35, 44 - case 4] and f] [14, 17, 32, 35, 44, 47 - case 5] show the resulting S_s tomograms by sequentially inverting the real cross-hole test data. Numbers in parentheses from b] – f] indicate the port numbers used as the pumped well for each cross-hole test.

10.6 COMPARISONS OF *K* AND S_S FIELDS FROM DIFFERENT TESTS AND ANALYSES

10.6.1 Visual Comparisons of Patterns of Heterogeneity of Different Tests and Analyses

Figure 10.3a shows the contour map of the *K* values estimated from the 48 core samples. The map as expected outlines the distribution of the blocks of low conductivity values, indicating the distribution of these core measurements. Similarly, the contour map of the *K* values estimated from the 40 slug tests also reveals a similar pattern [see Figure 10.3b]. Finally, a map of the *K* estimates based on the 48 hydrographs induced by pumping at port 28 [the cross-hole test] is shown in Figure 10.3c. As suggested by *Wu et al*. [2005], each estimated *K* represents some kind of average of the hydraulic conductivities over the cone of depression, but it is influenced by the hydraulic conductivity values near the pumping well and the observation well. As a result, the distribution of these estimates is smooth and does not necessarily show the

pattern of the heterogeneity in the sand box. The distribution of the K estimates based on the cross-hole tests with pumping at port 22 shows a very similar pattern.

Next, we plot in Figure 10.3d the spatial distribution of the estimated S_s values from the cross-hole tests using port 28 as the pumping well. It is interesting to observe that this spatial distribution is in some agreement with that resulting from the transient hydraulic tomography [see figure 10.2f]. That is, higher specific storage values are at the upper portion of the sand box. This result appears to support the finding by *Wu et al*. [2005] that the S_s estimates from the cross-hole analysis with the assumption of homogeneity reflect the S_s values between the pumping well and the 48 observation wells that spread over the entire sand box.

According to the above visual comparisons, we may conclude that the measurements using core samples and slug tests can satisfactorily map the heterogeneity pattern, but not the actual values of K, in the sand box if the number of tests or samples is sufficient. On the contrary, the cross-hole tests that utilize the homogeneous medium assumption produce only inconsistent average values of the hydraulic conductivity in the sand box. However, they apparently can be used to estimate the spatial distribution of S_s values in our sand box.

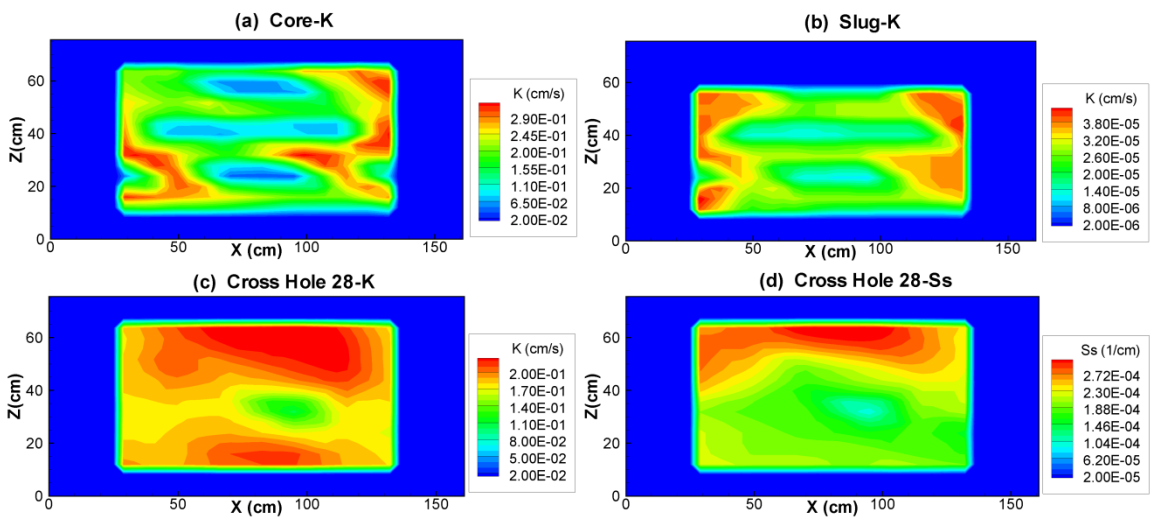


Figure 10.3: Contour plots of: a] core K values, b] slug K values, c] cross-hole K values with pumping taking place at port 28, and d] cross-hole S_s values with pumping taking place at port 28.

10.6.2 Comparison of Statistical Moments

After the visual evaluations of spatial patterns of K and S_s estimates from different tests and analyses, we quantitatively evaluate the K tomogram by comparing its sample and population means of estimated log-transformed hydraulic conductivity [$\overline{\ln K}$] and its sample and population variances $\left(\sigma^2_{\ln K}\right)$ to the corresponding sample statistical moments obtained from other tests. The sample mean of the tomogram is computed by taking the geometric mean of the hydraulic conductivity estimates at 48 elements corresponding to port locations, while the population mean is computed from the hydraulic conductivity estimates at all 741 elements.

First, we compare the sample mean [-1.860] and the population mean [-1.729] of K values obtained from transient hydraulic tomography using 6 pumping tests [Table 10.2] to the ln K_{eff} [-1.757] obtained from the flow-through experiment [Table 10.1]. This comparison shows that the estimate means of the tomogram approach the mean K value of an effective homogeneous medium after including the 6 pumping tests in the inversion for this particular

experimental setup. We note the difference between the mean values and the ln K_{eff} value determined from the flow-through experiment decreases quickly from the addition of one pumping test to the next and the values stabilize [Table 10.2]. These mean estimates are also in agreement with the mean of the estimates from the two cross-hole analyses. On the other hand, the means of the estimates from core samples, slug tests, and single-hole tests are smaller than those from the tomography, flow-through tests, and the cross-hole analysis.

We next compare the sample and population variance [0.768 and 0.906, respectively] of the K tomogram to the estimates of variance obtained from the available core K data [1.498]. This comparison shows that both the sample variance [0.768] and population variance [0.906] are much smaller than of the variance of the K values of core samples. This difference likely can be attributed to several factors: 1] the core value represents an averaged value over a smaller volume [V = 13.04 cm^3] of the sand in comparison with the averaging volume of the element [V = 171.46 cm^3] for the K estimates from the hydraulic tomography--disparity in scale between the two types of estimates; 2] each core value is an effective K for a one-dimensional flow situation, whereas the K values of the tomograms represents conditional "effective" K values from six multi-dimensional flow fields; 3] the core K values also may be subjected to higher variation due to the fact that we have disturbed the geometry and/or removed the compressive forces when extracting these sediments. Therefore, the results are consistent with our expectation.

Likewise, without any surprises, the variance of the tomograms is greater than those of cross-hole analyses and flow-through experiments since each K estimate from the cross-hole analyses and flow-through experiments represents a spatially averaged value over the cone of depression or the entire sandbox. In comparison with the variances from slug tests and single-hole analyses, the variance of the tomograms is slightly higher but is of the same order of magnitude of those of the slug tests and single-hole analyses.

Since no estimates of S_s for the cores were derived, we compare the S_s tomogram through the estimates of statistical moments from available single-hole and cross-hole pumping tests. We first compare the sample mean [-8.144] and population mean [-8.047] of log-transformed specific storage $\overline{\ln S_s}$ of the S_s tomogram obtained from the 6 sets of pumping tests [Table 10.3] to the mean of the available single-hole estimates [-7.960] as well as the mean of the equivalent estimates [-8.378] derived from the two cross-hole tests [Table 10.1]. The results are in some agreement, indicative of some usefulness of the cross-hole analysis that assumes homogeneity of the sand box—perhaps when the number of observations is sufficiently large.

We next compare the sample [0.214] and population [0.190] variance of log-transformed specific storage $\overline{\ln S_s}$ from the S_s tomogram [Table 10.3] to variance estimates obtained from single-hole [1.897] and the 2 cross-hole [0.047] tests [Table 10.1]. Apparently, there is a large discrepancy between these variances of the tomogram and the variance of S_s estimates from the single-hole tests. We expect that the variance of the estimates from the single-hole test should be greater than that from the tomogram but not by a factor of ten. This may be due to the relatively small number of single-hole S_s estimates.

The sample and population variance of the estimates from hydraulic tomography, as we anticipated, is about a factor of four larger than the variance of the equivalent S_s obtained by the cross-hole analyses that treat the medium to be homogeneous.

Table 10.2: Mean $\left(\overline{\ln K}\right)$ and variance $\left(\sigma^2_{\ln K}\right)$ of the log-transformed K estimates for the real inversions with number of cross-hole tests used in the analysis.

Number of cross-hole tests included in analysis [test number]	Sample		Population	
	$\overline{\ln K}$ [K ~ cms^{-1}]	$\sigma^2_{\ln K}$	$\overline{\ln K}$ [K ~ cms^{-1}]	$\sigma^2_{\ln K}$
1 [14]	-2.042 [0.130]	0.763	-1.906 [0.149]	0.696
2 [14+17]	-1.642 [0.194]	0.566	-1.490 [0.225]	0.559
3 [14+17+32]	-1.873 [0.154]	0.703	-1.678 [0.187]	0.721
4 [14+17+32+35]	-1.953 [0.142]	0.720	-1.769 [0.171]	0.762
5 [14+17+32+35+44]	-1.909 [0.148]	0.822	-1.727 [0.178]	0.947
6 [14+17+32+35+44+47]	-1.860 [0.156]	0.768	-1.729 [0.177]	0.906

Table 10.3: Mean $\left(\overline{\ln S_s}\right)$ and variance $\left(\sigma^2_{\ln S_s}\right)$ of the log-transformed S_s estimates for the real inversions with number of cross-hole tests used in the analysis.

Number of cross-hole tests included in analysis [test number]	Sample		Population	
	$\overline{\ln S_s}$ [S_s ~ cm^{-1}]	$\sigma^2_{\ln S_s}$	$\overline{\ln S_s}$ [S_s ~ cm^{-1}]	$\sigma^2_{\ln S_s}$
1 [14]	-7.978 [3.430E-04]	0.018	-8.032 [3.250E-04]	0.021
2 [14+17]	-7.749 [4.311E-04]	0.027	-7.831 [3.973E-04]	0.038
3 [14+17+32]	-8.010 [3.321E-04]	0.058	-8.061 [3.157E-04]	0.055
4 [14+17+32+35]	-8.071 [3.123E-04]	0.080	-8.111 [3.002E-04]	0.074
5[14+17+32+35+44]	-8.014 [3.309E-04]	0.208	-8.066 [3.141E-04]	0.192
6 [14+17+32+35+44+47]	-7.996 [3.368E-04]	0.214	-8.047 [3.201E-04]	0.190

10.6.3 Comparison of Local Values

To examine the performance of different tests in greater detail, we next compare local K values from the K tomogram to the K estimates at sample locations from the core measurements, single-hole tests as well as from cross-hole analyses at each observation port induced by pumping at port 22 and port 28 [Figure 10.4]. A correlation coefficient is used to quantify the spatial correspondence of log-transformed estimates from tomography, χ_i, and the estimates from other tests, $\hat{\chi}_i$,

$$R = \frac{\frac{1}{N}\sum_{i=1}^{N}(\chi_i - \mu_\chi)(\hat{\chi}_i - \mu_{\hat{\chi}})}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\chi_i - \mu_\chi)^2 \cdot \frac{1}{N}\sum_{i=1}^{N}(\hat{\chi}_i - \mu_{\hat{\chi}})^2}} \qquad [10.1]$$

where N is the total number of elements; μ_χ and $\mu_{\hat{\chi}}$ are means for the estimates from tomography and the estimates from other tests and analyses, respectively. The R values are also given in Figure 10.4.

This figure shows that there is quite a bit of scatter but the bias is not too large when the local K estimates from the tomography are compared to the core values and the correlation value of the two types of estimates is 0.57. On other hand, there is a noticeable bias when the local K values from the tomogram are compared to the single-hole test data but they correlation value is high [0.85], indicating a similar spatial pattern between the two estimates. This bias [i.e.

systemically lower values for the estimates by the single hole tests] can be explained by near well effects, that is, the skin effect discussed by *Illman et al*. [2006] that can cause local *K* estimates from single-hole tests to be slightly smaller. Or possibly, it can be attributed to scale disparity of these estimates.

Estimates of *K* values at the monitoring ports from the cross-hole analyses for the two pumping locations [22 and 28] appear as vertical, narrow clusters in Figure 10.4. The narrow clusters suggest the estimates are smooth [averaged values] in comparison with those from tomogram at the same locations.

We next compare local S_s values from the S_s tomogram to the S_s estimates from the single-hole tests [Figure 10.5]. Here, numbers associated with each data point indicates the port number in which the pumping took place. Overall, we see that the S_s estimates are larger near the top of the sandbox and it decreases as we move deeper into the sandbox as noted earlier for the S_s tomograms [Figures 10.3a - f]. The correlation between the two estimates is 0.9, indicating a very similar spatial pattern of the two estimates. However, variability of S_s from the single-hole tests treating the medium to be homogeneous is much larger than that of the S_s estimates from hydraulic tomography.

The plot of the estimates from the cross-hole analyses for pumping ports 22 and 28 in Figure 10.5 yields a narrow vertical cluster, indicative of similar mean values but small variation in the estimates from the two cross-hole analyses and great variation in the S_s tomogram.

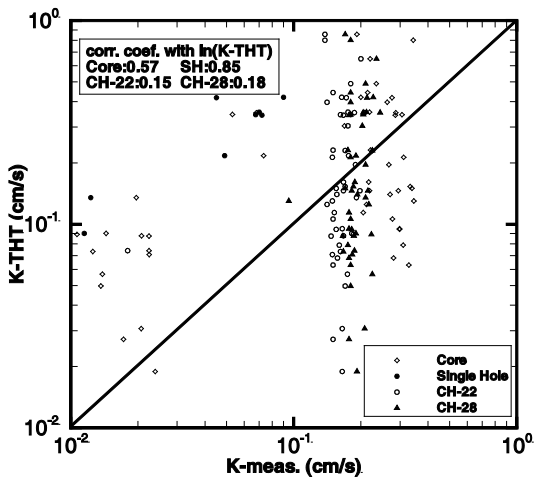


Figure 10.4: Scatter plot of *K* values from the real *K* tomogram at the observation point and core, single-hole, and cross-hole *K* estimates.

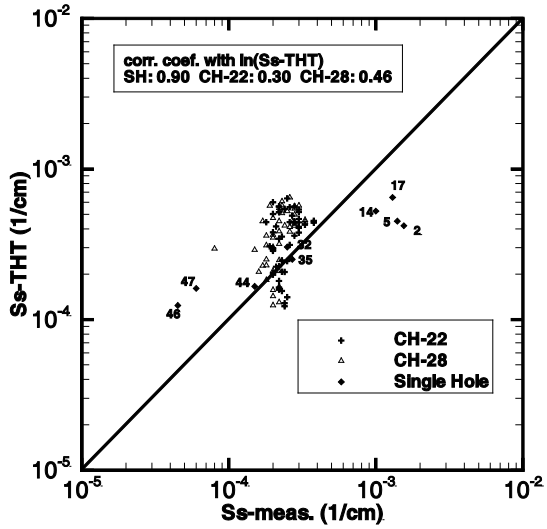


Figure 10.5: Scatter plot of S_s values from the real S_s tomogram at the observation point and single-hole and cross-hole S_s estimates.

10.6.4. Comparison of *K* Tomogram obtained from Steady State Hydraulic Tomography

We also compare the *K* tomogram obtained from the transient hydraulic tomography to that obtained from its steady-state counterpart [Figure 10.6]. Here, the *K* data came from Figure 4d in *Illman et al*. [2007]. The results show that the estimated *K* values using transient head data are almost identical to those based on the steady head data with a correlation value of 0.83 [i.e., the patterns are very similar]. The difference between the two perhaps reflects the influence of S_s parameters on the estimation of *K*. The dashed line in the figure indicates the 95% confidence interval.

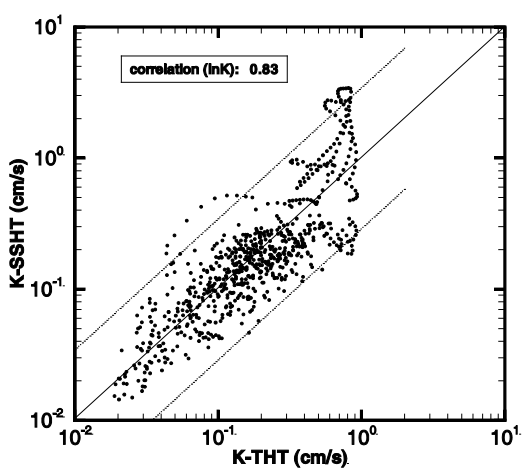


Figure 10.6: Scatter plot of *K* values from the transient hydraulic tomography of this paper and those from the steady state hydraulic tomography in *Illman et al*. [2007].

10.7 VALIDATION OF *K* AND S_S TOMOGRAMS

As shown in previous section, some of the tests and analyses yield similar *K* and S_s statistics and patterns. It is however difficult to validate each other due to scale disparity, different flow conditions, etc. As a consequently, an appropriate validation approach is to test

the predictability of the estimates under different flow scenarios. In order to do this, we first verify the K and S_s tomograms by simulating the last pumping test at port 47 used in the construction of the tomogram. We then compare the simulated and measured transient drawdown at early [3 sec], intermediate [10 sec] and late time [20 sec] periods [Figure 10.7a] at all ports except for the pumped port. The correlation values between the simulated and observed drawdowns at the three times as well as the means and variances of their differences are reported in the figure. This plot and the quantitative measures show that the comparison is excellent providing us with further confidence that SSLE can provide an unbiased estimation of the drawdown distribution. Note that the simulated drawdowns will not necessarily match the observed ones perfectly as they should, due to the fact that the tomograms are conditional effective K and S_s fields and also due to noise in the observations. Notice that simulated drawdown values are larger than the observed ones when the values are large. This is consistent with the fact that uncertainty in the predicted drawdown grows with the mean gradient according to stochastic analysis [to be explained later].

A better validation of the K and S_s tomograms is to simulate an additional pumping test that was not used in the inversion and to examine whether the drawdown at various sampling ports of this independent test can be predicted accurately at various times. For this, we utilize the K and S_s tomograms obtained from the inversion of the 6 pumping tests [Figures 10.1f and 10.2f] and simulate a test with pumping taking place at port 46. Pumping test at port 46 was chosen for validation purposes because it was one of the pumping tests with the cleanest data devoid of external factors. We note that the fact that port 46 is close to port 47 does not make it easier to simulate the observed behavior.

Figure 10.7b shows the results of this comparison. In this figure, the observed drawdowns are plotted against the predicted drawdowns at times 3, 10, and 20 seconds at the 48 observation ports. It also shows the mean and variance of the difference between the observed and predicted drawdowns as well as the correlations between the observed and predicted. According to this figure, the data pairs are scattered along the 45 degree line, indicating predicted drawdown distributions generally are statistically unbiased in comparison with the observed. The high correlation values for the three different times [0.995] suggest the predicted drawdown distribution is almost identical to the observed distribution—at least the drawdowns at the three times at the observation ports. This is an exciting result because it indicates that using the K and S_s fields derived from hydraulic tomography, one can yield an excellent prediction of the drawdown behavior in the sandbox. Again, the predicted drawdown values using the tomograms are slightly greater than the observed values at all the sampling ports at the three times.

Figure 10.8a-c presents another way to validate the tomograms. In the figure, the horizontal axis denotes the simulated drawdown values at the sampling ports based on the mean [or effective] K and S_s values from their tomograms. The vertical axis represents the observed head values at the same locations as well as those simulated using the K and S_s tomograms from figures 10.1f and 10.2f. In addition, a 45 degree line indicating perfect drawdowns corresponding to those resulting from the mean parameters is plotted along with an upper and a lower bound of the drawdown that denote the variability in downdowns due to heterogeneity of K and S_s ignored by the homogeneous assumption. The upper and lower bounds were constructed by adding and subtracting twice the square root of the predicted drawdown variance at each observation location, to and from the simulated drawdown based on the mean K and S_s tomograms, respectively. The drawdown variances were calculated from the SSLE algorithm using our initial guess values for the correlation scales of ln K and ln S_s and the mean and

variance values of the estimated K and S_s tomograms. Note that the true correlation scales, means, and variances are unknown; the variances of the tomogram should be smaller than the true ones; the means are close to the true as illustrated earlier. Therefore, the upper and lower bounds, in effect, are uncertain themselves but the mean behavior is likely reasonable. This figure thus suggests that using effective parameters, a flow model assuming homogeneity predicts quite different drawdowns than the observed at the 48 ports in the heterogeneous aquifer [i.e., the sand box]. That is, the predicted drawdowns are consistently smaller near the pumping port and greater away from the pumping location than the observed ones. On the contrary, using the K and S_s fields resulting from hydraulic tomography, a classic governing equation for groundwater flow can yield an excellent prediction of the drawdowns at three different times at all the 48 observation ports in the heterogeneous sand box. This result is exciting in view of inherent non-uniqueness of the estimation and errors in the measurements. Certainly, more independent pumping events would make this validation more significant.

Lastly, we believe that these results manifest the utility of the hydraulic tomographic surveys and the robustness of the SSLE algorithm. More importantly, these results implicitly validate the classical governing equations for groundwater flow in heterogeneous porous media [at least in our sand box].

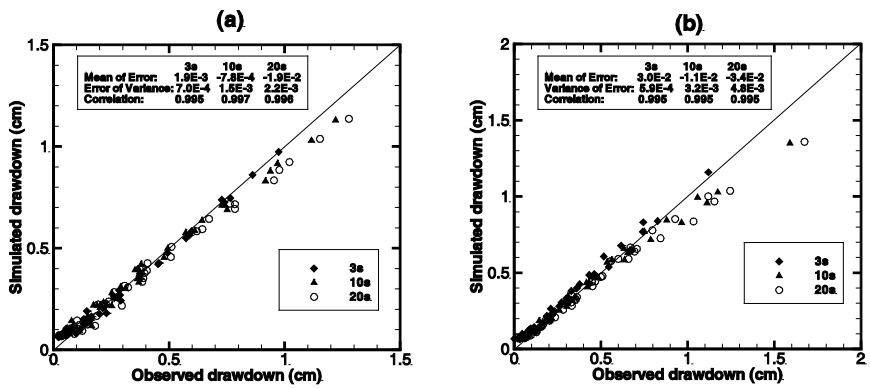


Figure 10.7: Scatter plot of simulated drawdown versus observed drawdown at 3, 10, and 20 seconds after pumping begins for cross-hole hydraulic tests at a] port 47 and b] port 46. In both cases, the simulated values were obtained by through synthetic simulations of the cross-hole test conducted using the K and S_s tomogram [Figures 10.1f and 10.2f]. The observed drawdown values are from an actual cross-hole test conducted at each port.

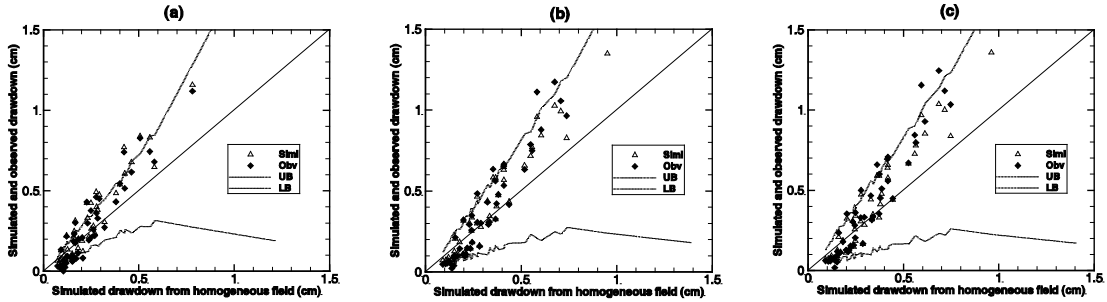


Figure 10.8: Simulated drawdown from homogeneous field versus observed head values at the same locations as well as those simulated using the K and S_s tomograms at time equal to: a] 3, b] 10, and c] 20 seconds after pumping. The solid line represents a 45 degree line, and the dashed lines represent the upper and lower bounds of the drawdown.

11. LABORATORY SANDBOX VALIDATION OF HYDRAULIC TOMOGRAPHY THAT USES THE TEMPORAL MOMENTS OF DRAWDOWN-RECOVERY DATA

11.1. INTRODUCTION

In section 10 we have shown that THT is an effective technique in imaging K and S_s, but is computationally costly. To overcome the computational challenge for THT analysis, *Zhu and Yeh* [2006] developed an approach that utilizes the 0th and 1st temporal moments of well hydrographs, instead of drawdown data itself. The governing equations for the temporal moments are Poisson's equations. These equations demand less computational resources as opposed to the parabolic equation that governs drawdown evolution. Likewise, the adjoint equations for evaluating sensitivities of the moments for parameter estimation also take the same forms. Therefore a HT which uses the temporal moments of drawdown-recovery data [HT-m] expedites the interpretation of THT surveys. Results of the investigation revealed that the HT-m yielded estimates comparable to those using transient heads, but at significantly less computational costs.

Previously, *Zhu and Yeh* [2006] postulated that the calculation of the temporal moments, which involves integration in time of the drawdown-recovery data, may cause loss of information on the parameters that are being estimated by the inverse procedure. They conducted numerical experiments using Gaussian and stationary random K and S_s fields to test this conjecture. Results revealed that the HT-m yields less accurate K and S_s estimates, due to the integrated nature of temporal moments, which lose some information about heterogeneity borne in the transient hydrograph. However, they found that this loss of information can be overcome by adding more observation data at various points.

While various algorithms for SSHT and THT have been developed and some of them have been tested numerically, several sandbox studies have been conducted to evaluate the performance of both SSHT [*Liu et al*., 2002; *Illman et al*., 2007a,b] and THT [*Liu et al*., 2007]. To date, the evaluation of the HT-m has not been accomplished either in the laboratory or the field setting. A field validation is the ultimate goal, but prior to that, laboratory validations are necessary in which the baseline heterogeneity is largely known, and all forcing functions and errors can be controlled as opposed to field applications.

In this section, we: 1] investigate the “effectiveness” of the HT-m approach in comparison to the THT through synthetic simulations using nonstationary K and S_s fields; 2] invert cross-hole pumping test data obtained in a laboratory aquifer with deterministic heterogeneity to obtain K and S_s tomograms using the HT-m algorithm; and 3] to investigate the performance of HT-m in comparison to other hydraulic tomography approaches, that is, the SSHT [*Yeh and Liu*, 2000] and THT [*Zhu and Yeh*, 2005] algorithms.

11.2. NUMERICAL EXPERIMENTS

11.2.1 Purpose of numerical experiments

The main purpose of conducting numerical experiments is to rigorously test the THT [*Zhu and Yeh*, 2005] and HT-m [*Zhu and Yeh*, 2006] algorithms. In these experiments, we generate synthetic experimental data through forward simulations of cross-hole pumping tests and use those data in the inverse models to generate K and S_s tomograms. The inversion of the

synthetic data illustrates the ability of the THT and HT-m algorithms to obtain K and S_s tomograms simultaneously under idealized conditions. In this idealized condition, model and measurement errors are excluded and forcing functions are fully controlled. That is, synthetic measurements are derived from the same physical principles as those used for test data interpretation. Our intention is to further establish the credibility of the inverse approach. We can then interpret cross-hole pumping test data obtained from laboratory sandbox experiments to further test the inverse techniques under controlled environments, where measurements/model errors may exist, but can be controlled.

Zhu and Yeh [2006] conducted a similar synthetic study to evaluate the performance of the THT and HT-m algorithms. However, their numerical experiments involved multi-Gaussian and stationary K and S_s fields generated using the spectral algorithm of *Gutjahr* [1989]. In contrast, our synthetic studies are different, yet complementary to theirs' because ours' involve non-Gaussian and nonstationary K and S_s fields to test the feasibility of utilizing these codes for data collected in the laboratory sandbox that we describe later.

11.2.2 Inverse model description and computational approach/platform

The estimation techniques using transient head [*Zhu and Yeh*, 2005] and temporal moments [*Zhu and Yeh*, 2006] are both based on the sequential successive linear estimator [SSLE]. We utilize the algorithms to map the spatial distribution of K and S_s in a synthetic and real sandbox aquifer. A brief description of the THT and HT-m algorithms is given below.

For THT, the inverse model assumes a transient flow field and the natural logarithm of K [$ln\ K$] and S_s [$ln\ S_s$] are both treated as stochastic processes. One advantage of using natural logarithms of K and S_s is that it avoids negative values of parameters during the estimation process [*Zhu and Yeh*, 2006]. The model additionally assumes that the mean and correlation structure of the K and S_s fields are known *a priori*. Similarly, the 0th-moment of drawdown-recovery data induced by a pumping test during HT-m as well as the characteristic time [ratio of 1st and 0th-moments] are treated as stochastic processes consisting of the mean and perturbation components.

The estimation procedure for THT starts with a weighted linear combination of parameter measurements and head at different locations to obtain the first estimate of the parameters. The weights are calculated using the means and covariances of parameters, the covariances of head in space, and the cross-covariances between the head and parameters. The first estimate is then used in the mean flow equation to calculate the head at observation locations and sampling times through a forward simulation. At the end of this forward simulation, the differences between the observed and simulated heads are calculated and a weighted linear combination of these differences is then used to improve the previous estimates. Iterations between the forward simulation and estimation continue until the improvement in the estimates diminishes to a prescribed value. We note that the HT-m uses temporal moment data instead of heads used in the THT algorithm. Details to the computational algorithms for both THT and HT-m are provided in *Zhu and Yeh* [2005, 2006]. We emphasize that the THT algorithm is based on the parabolic equation, the governing equation for transient groundwater flow, which must be solved for each time step. The HT-m algorithm, on the other hand, is based on the moment equations [eqns. 10 and 12 of *Zhu and Yeh*, 2006], which are Poisson's equations, thus it needs to be solved only once.

All numerical simulations were executed using 8 of 16 processors on a PC-cluster consisting of [1 master and 15 slaves each with Pentium IV 3.6 GHz with 1 GB of RAM] located at IIHR-Hydroscience & Engineering at the University of Iowa.

11.2.3 Description of numerical experiments and inverse model parameters

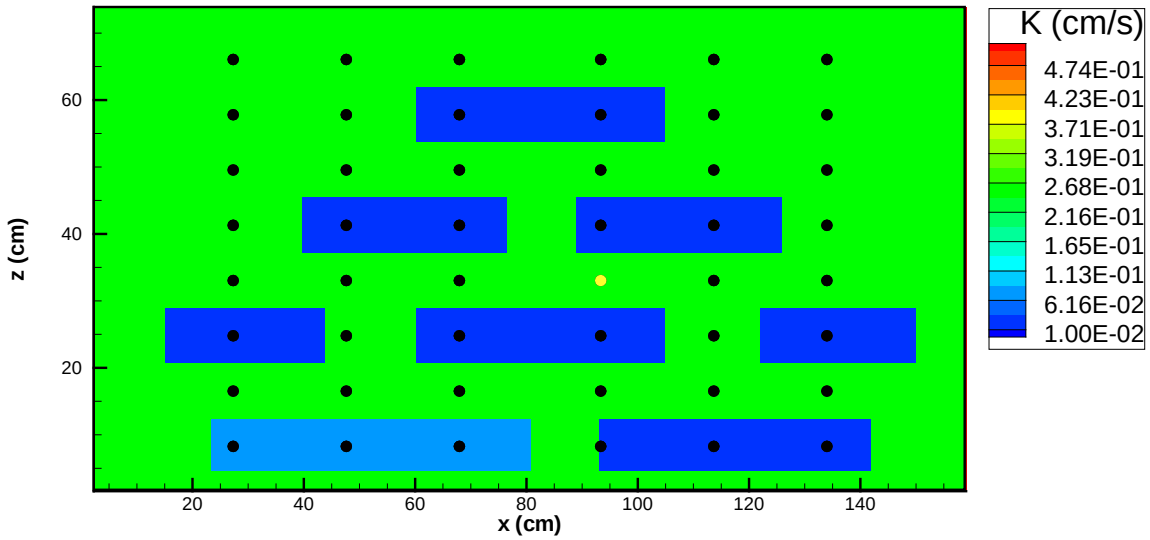
We generated synthetic cross-hole pumping test data on the computer using the MMOC3 forward model [*Yeh et al*., 1993] using the same experimental setup for sandbox described in previous sections. The synthetic aquifer was discretized into 741 elements and 1600 nodes with element dimensions of 4.1 cm × 10.2 cm × 4.1 cm. Both sides and the top boundary were set to be constant head boundary conditions, while the bottom boundary of the sandbox was considered a no-flow boundary. This grid setup is consistent with that was utilized for SSHT by *Illman et al*. [2007, 2008] and for THT by *Liu et al*. [2007] to analyze the laboratory experiments.

The synthetically generated data are used to conduct inverse modeling with both the THT and HT-m codes. Figure 11.1a-b shows the K and S_s distributions in the synthetic sandbox that is used to generate synthetic cross-hole pumping test data on the computer. We conducted synthetic cross-hole tests by running transient forward simulations using MMOC3 at 8 ports [2, 5, 14, 17, 32, 35, 44, and 47] individually by setting a constant pumping rate and recorded the hydraulic head. The pumping rate for each test varied between 2.92 – 3.17 cm^3/s, while the pumping time for each test varied between 51 to 56 s. Pumping was then stopped to allow for full recovery of the drawdown. During each pumping test, hydraulic head responses were obtained from all 48 ports yielding 48 × 8 = 384 drawdown-time curves after the HT survey involving multiple cross-hole tests. The inversion codes for THT and HT-m do not require the entire drawdown-recovery curve in the estimation procedure. Therefore, we selected few data to conduct the inversion. Here, we selected the transient heads collected at 1, 2, 5, 10, and 20 seconds to estimate K and S_s simultaneously using the THT algorithm. For the interpretation of synthetic data using the HT-m algorithm, we obtain the moments directly from the forward algorithm [i.e., the moment equations 10 and 12 in *Zhu and Yeh*, 2006] and use these in the inverse code. This is because according to *Zhu and Yeh* [2006], the time to reach full recovery in drawdown is difficult to determine, thus there is a possibility that calculating moments from the drawdown-recovery curve could introduce some measurement error. Introduction of this measurement error can obscure the comparison of results from THT and HT-m.

We also generated additional synthetic pumping test data at port 28. Pumping test conducted by pumping at port 28 represents a cross-hole pumping test with pumping taking place in a high K zone. We recorded the entire drawdown and recovery responses. These data are later used for evaluating the validity of the computed K and S_s tomograms using both the THT and HT-m algorithms.

Inputs to the inverse model include initial guesses for the K and S_s, estimates of variances and the correlation scales for both parameters, volumetric discharge [Q_n] from each pumping tests, where n is the test number, available point [small-scale] measurements of K and S_s, as well as head-time information for the THT algorithm. For the HT-m algorithm, inputs are similar to the THT algorithm, but instead of the head-time information, the 0th moment and characteristic time [ratio of 1st and 0th moments] computed from the drawdown-recovery curve are included. Although available point [small-scale] measurements of K and S_s can be input to both inversion algorithms, we do not use these measurements to condition the estimated parameter fields to test the inversion algorithm.

a]



b]

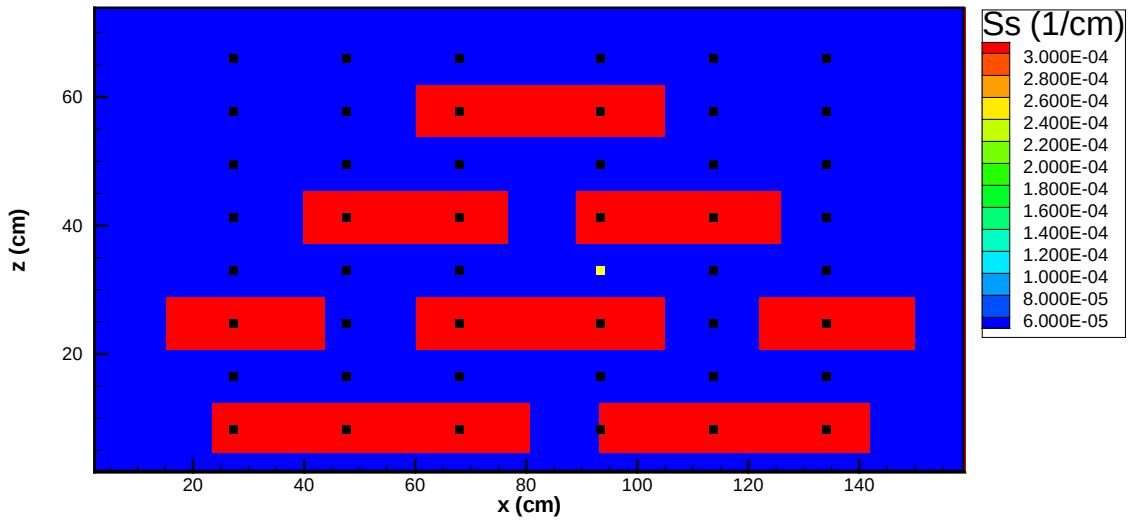


Figure 11.1: a] K and b] S_s distributions in the MMOC3 model used to generate synthetic cross-hole pumping test data used for THT and HT-m analyses [port 28 is shown in yellow].

11.3 RESULTS FROM INVERSE MODELING OF SYNTHETIC CROSS-HOLE PUMPING TEST DATA

11.3.1 Results from THT

We conduct the inversion of 8 synthetic cross-hole pumping tests with pumped ports shown on Figure 11.1. All 48 ports [see Figure 11.1 for locations] were used to collect drawdown responses for the inverse modeling. Figures 11.2a - d show the results of the estimation of the *K* tomogram from the successive inclusion of test data from pumping tests at ports 47, 44, 35, 32, 17, 14, 5, and 2, conducted in that order. A visual comparison of the *K* tomogram [Figure 11.2d] obtained by sequentially including 8 tests into the inversion algorithm and the original *K* distribution [Figure 11.1a] shows that the THT is able to delineate the major low and high *K* features that comprise the aquifer heterogeneity. This includes the correct delineation of the positions and shapes of low *K* blocks and openings between blocks that could allow contaminant migration.

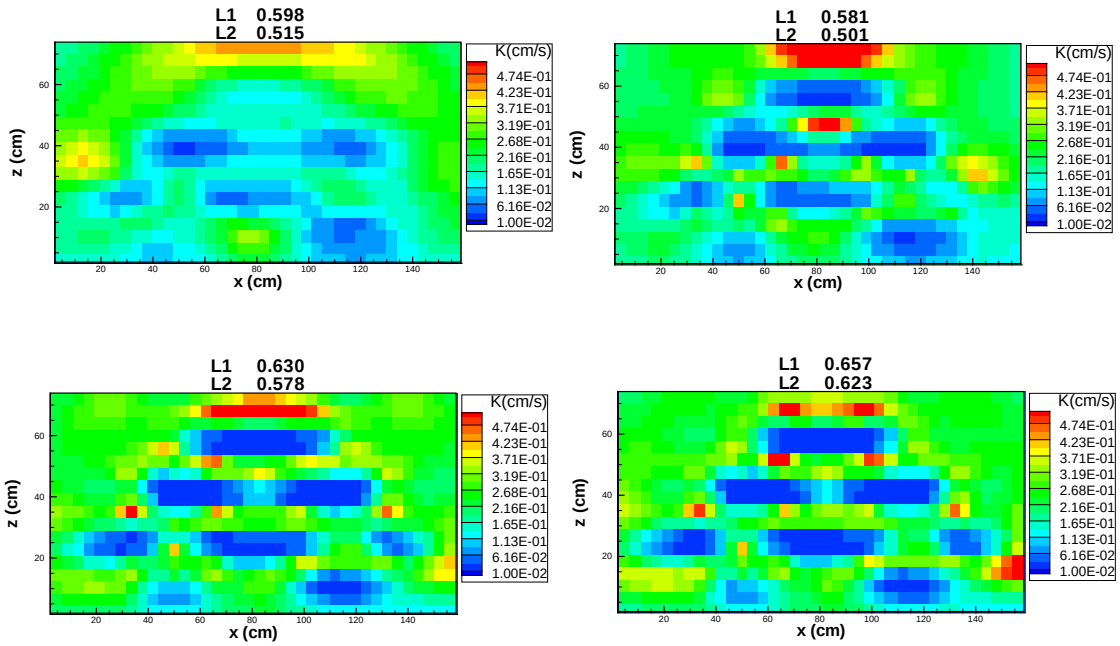


Figure 11.2: *K* tomograms computed using THT from synthetic pumping test data included sequentially. a] Ports 47, 44. b] Ports 47, 44, 35, 32. c] Ports 47, 44, 35, 32, 17, 14. and d] Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate those used as the pumped well for each cross-hole test.

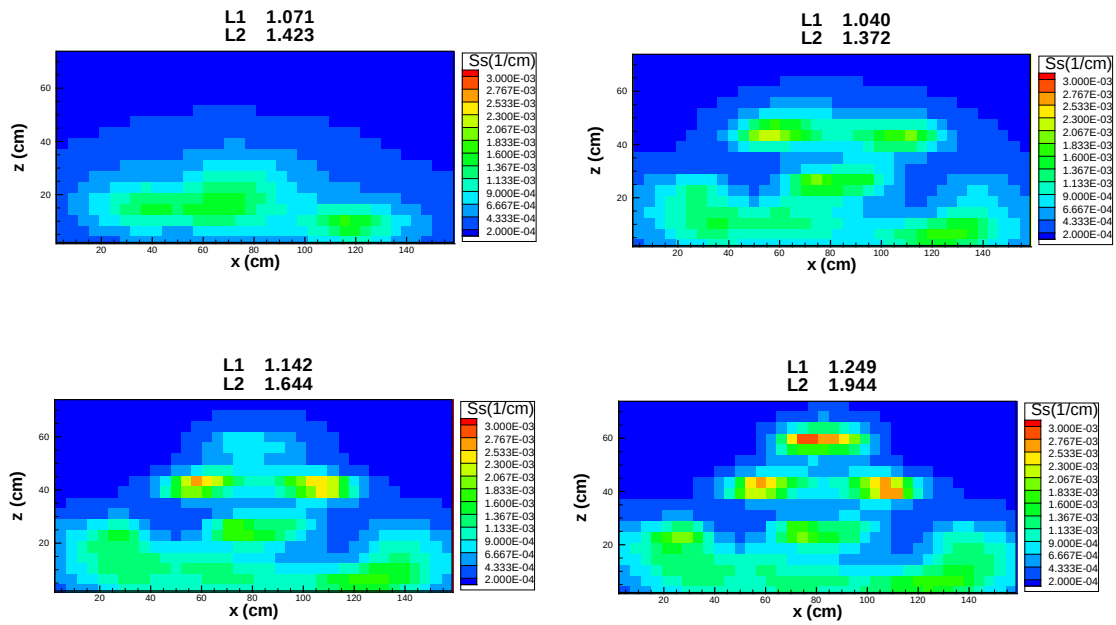


Figure 11.3: S_s tomograms computed using THT from synthetic pumping test data included sequentially. a] Ports 47, 44. b] Ports 47, 44, 35, 32. c] Ports 47, 44, 35, 32, 17, 14. and d] Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate those used as the pumped well for each cross-hole test.

It is of interest to note that the synthetic aquifer has a K distribution that is non-Gaussian and nonstationary. Because *Zhu and Yeh* [2005]'s algorithm assumes a Gaussian and a stationary field, one would not expect this approach to be applicable to this K distribution. However, the statistical assumptions inherent in the algorithm become less important as we include a large number of observation well data sets in the hydraulic tomography algorithm. This fact is evident from our results. Therefore, *Zhu and Yeh* [2005]'s SSLE is not limited to Gaussian and stationary random K fields.

Figure 11.3a-d shows the corresponding S_s tomogram that was estimated simultaneously. In contrast to Figure 11.2a-d, the structure consisting of variable size sand bodies is visible for the S_s tomogram near the top and center of the tomogram, but the clarity of S_s heterogeneity decreases as the depth increases. In general, we see that the S_s tomogram does not capture the blocks as well as the K tomograms near the bottom of the sandbox. Examination of Figures 11.2a-d and 11.3a-d show that it takes more pumping tests to delineate the heterogeneity pattern of S_s in comparison to K using the THT algorithm.

Both Figures 11.2a-d and 11.3a-d also show the mean absolute error norm L_1 and mean square error norm L_2 defined previously. We note that the L_1 and the L_2 statistics do not vary significantly from one case to the next in Figures 11.2a-d and 11.3-d. However, the statistics are in general lower for the K tomograms in comparison to the S_s tomograms, which suggest that the estimation of K is more accurate than S_s.

11.3.2 Results from HT-m

We also conducted the inversion of the same 8 synthetic cross-hole pumping tests in the same order using the HT-m algorithm of *Zhu and Yeh* [2006] to obtain K and S_s tomograms simultaneously. As in the THT case, all 48 ports were used for HT-m. Figure 11.4a-d shows the

results of the estimation of the *K* tomogram from the HT-m analysis. These results are similar in quality to those obtained from through THT algorithm showing that the HT-m is capable of estimating *K* fields accurately even if the synthetic aquifer has a *K* distribution that is non-Gaussian and nonstationary. Therefore, as in *Zhu and Yeh* [2005]'s algorithm for THT, the SSLE for the HT-m developed by *Zhu and Yeh* [2006] is not limited to Gaussian and stationary random *K* fields.

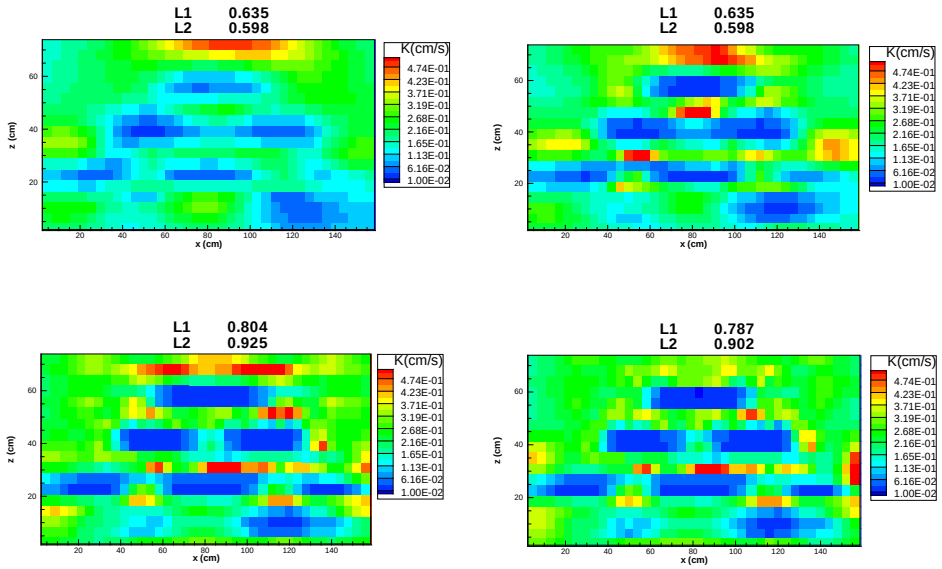


Figure 11.4: *K* tomograms computed using HT-m from synthetic pumping test data included sequentially. a] Ports 47, 44. b] Ports 47, 44, 35, 32. c] Ports 47, 44, 35, 32, 17, 14. and d] Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate those used as the pumped well for each cross-hole test.

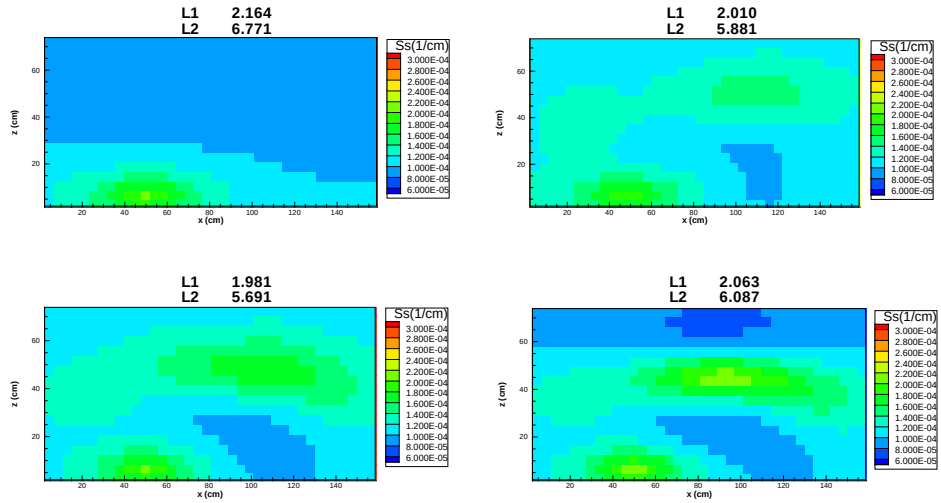


Figure 11.5: S_s tomograms computed using HT-m from synthetic pumping test data included sequentially. a] Ports 47, 44. b] Ports 47, 44, 35, 32. c] Ports 47, 44, 35, 32, 17, 14. and d] Ports 47, 44, 35, 32, 17, 14, 5, 2. Port numbers indicate the ports used as the pumped well for each cross-hole test.

11.3.3 Comparison of tomograms and simulation times

Figure 11.5a-d shows the corresponding S_s tomogram that was estimated simultaneously. In contrast to Figure 11.4a-d, the structure consisting of variable size sand bodies visible in the K tomogram is not as clear for the S_s tomogram. These results collectively suggest that the HT-m approach is not as sensitive to estimating the S_s heterogeneity pattern in comparison to the THT approach. This suggestion was initially made by *Zhu and Yeh* [2006], but here we independently confirm their findings and find that their conclusion also holds true for non-Gaussian and nonstationary K and S_s fields.

We find that the quality of K tomograms does not vary considerably between the THT and HT-m approaches but the quality of the S_s tomogram is considerably better for the THT approach in comparison to the HT-m approach. This is evident by examining the L_1 and the L_2 statistics from one case to the next.

However, better estimation of the S_s tomogram through THT comes with an increase in simulation time. We list in Table 11.1 the simulation times for the 2 approaches. Results show that the simulation times of the THT is approximately 20 times longer when compared to the HT-m approach.

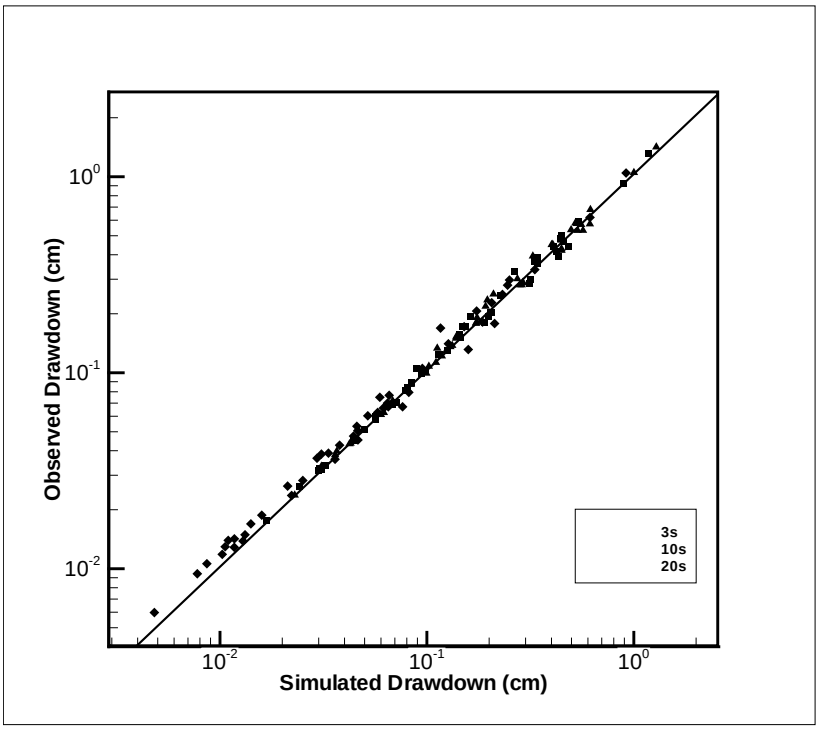
Table 11.1: Summary of simulation times [min] using the THT and HT-m algorithms and its ratio to obtain the K and S_s tomograms [Figures 11.2 – 11.5].

Pumping test #	THT [min]	HT-m [min]	THT [min]/HT-m [min]
2	97.81	4.74	20.64
4	192.15	9.25	20.77
6	287.21	14.05	20.44
8	380.29	19.12	19.89

11.3.4 Quantitative evaluation of synthetic inversion results

We evaluate the K and S_s tomograms obtained from the THT and HT-m approaches further by simulating a cross-hole pumping test not used in the construction of the tomograms. Figures 11.6a-b show scatter plots of observed drawdown versus simulated drawdown at 3, 10, and 20 seconds after pumping begins for synthetic cross-hole hydraulic tests at port 28. Table 11.2 summarizes the corresponding means and variances of the differences between the 2 results and the correlation coefficients. Port 28 is located within a high K zone. In both cases, the simulated values were obtained via synthetic forward simulations of cross-hole tests using MMOC3 with K and S_s tomograms computed from the THT [Figures 10.2d and 10.3d] and HT-m algorithms [Figure 11.4d and 11.5d]. The observed drawdown values are from the synthetic cross-hole test conducted using MMOC3 with pumping taking place at port 28 with the "true" K and S_s distributions from Figure 11.1a-b. These results show that the THT algorithm is able to delineate the K and S_s fields very accurately allowing for the prediction of drawdowns from an independently conducted pumping test from early to late times. Likewise, the HT-m approach yields K and S_s tomograms that allow for pretty good predictions of drawdowns from an independent pumping test. However, the accuracy of the drawdown prediction is not as good at early times, but improves with time. This is because the S_s estimate obtained from HT-m is not as good as that obtained from THT. These results collectively suggest that both approaches could be applied successfully in a nonstationary and non-Gaussian K and S_s fields, but the THT is more

robust than HT-m. We next describe the inversion of cross-hole pumping tests conducted in the laboratory using the HT-m approach.



b]

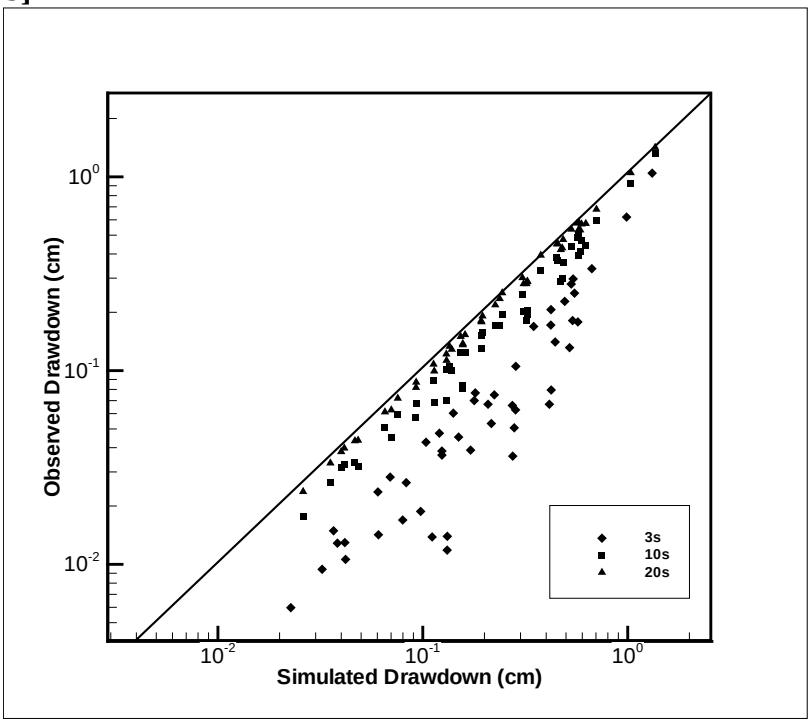


Figure 11.6: Scatter plot of simulated versus observed drawdown at 3, 10, and 20 seconds after pumping begins for synthetic cross-hole hydraulic tests at port 28 for a] THT and b] HT-m algorithms. In both cases, the simulated values were obtained through synthetic simulations of

the cross-hole tests conducted using the K and S_s tomograms [Figures 10.2d and 10.3d for THT; Figures 11.4d and 11.5d for HT-m]. The observed drawdown values are from the synthetic cross-hole test conducted using MMOC3 with pumping taking place at port 28 with K and S_s distributions from Figure 11.1a-b.

Table 11.2: Summary of the mean and variance of the estimation errors from the comparison of observed to simulated drawdowns. The K and S_s fields are from the THT and HT-m algorithms. The correlation coefficients shown in this table are for the scatter plots shown in Figures 11.7a-b.

	THT algorithm			HT-m algorithm		
Time [s]	Mean	Variance	Correlation	Mean	Variance	Correlation
3	0.107	0.027	0.995	0.283	0.067	0.927
10	0.226	0.052	0.996	0.308	0.074	0.984
20	0.282	0.067	0.996	0.308	0.074	0.997

11.4 LABORATORY SANDBOX DATA AND ITS INVERSE MODELING

11.4.1 Description of laboratory sandbox and instrumentation used to conduct real cross-hole pumping tests

The previous sections described the effectiveness of HT-m in relation to the more computationally intensive THT using experimental error-free synthetic cases. We next test the HT-m approach using cross-hole pumping tests conducted in a synthetic heterogeneous aquifer built in the laboratory.

The synthetic heterogeneous aquifer constructed in the sandbox was designed to test various HT algorithms under a controlled setting and a CAD drawing of this sandbox was previously shown. Details to the sandbox and data acquisition system can be found in *Craig* [2005], *Illman et al*. [2007], and *Liu et al*. [2007], but here, we include information that we considered to be most salient for the description of the inverse analysis.

The sandbox was packed with 4 different commercially sieved sands [20/30 and 40/30, U.S. Silica; F-75 and F-85, Unimin Corporation] by *Craig* [2005]. In particular, we packed 8 rectangular sand bodies consisting of lower K material [40/30, F-75, and F-85] within high K sand [20/30]. Each rectangular sand body consists of the material of different Ks and the locations are indicated using dashed rectangles on Figure 8.1.

The flow system for the cell is provided by two constant-head reservoirs, one at each end of the sandbox. Three constant head boundaries can be developed by ponding water over the top of the sand, effectively connecting the two reservoirs. We chose this boundary condition configuration for all cross-hole tests used in hydraulic tomography as they were the most stable during each cross-hole pumping test.

11.4.2 Description of real cross-hole pumping tests used for hydraulic tomography in the sandbox

To generate data for hydraulic tomography, we conducted cross-hole pumping tests at 9 separate pumping ports. During the pumping tests, the top and two sides of the aquifer served as constant head boundaries, while the bottom remained a no-flow boundary. Pumping rates ranged from 2.50 - 3.17 cm^3/sec in those tests. For each test, data collection started before the pump was activated to obtain the initial hydraulic head in the sandbox. A peristaltic pump was then activated at the pumping port and allowed to run until the development of steady state flow

conditions. The pump was then shut off to collect recovery data until the hydraulic head recovered fully. During each pumping test, hydraulic head data were collected at all 48 ports and in the two boundary reservoirs using pressure transducers.

We selected 8 out the 9 pumping tests and the drawdown-time observations at the rest of 47 ports during each test as the data sets for the hydraulic tomographic survey. The eight pumping/drawdown data sets are then used for HT-m analysis. The remaining one test was reserved for validation purposes.

11.4.3 Pre-processing of drawdown-recovery data

Illman et al. [2008] found from the analysis of HT data obtained in a laboratory sandbox aquifer that the single-to-noise ratio can be critical in inverse modeling of cross-hole pumping test data. To overcome this issue of signal-to-noise ratio on the quality of hydraulic tomography surveys, *Xiang et al.* [2008] developed a wavelet transform approach based on the work by *Zhang et al*. [2006] to remove noise from experimental data prior to implementing the data in their inverse code. We utilized the wavelet transform tool implemented in VSAFT2 by *Xiang et al.* [2008] to de-noise all drawdown-recovery curves prior to calculating the temporal moments.

11.4.4 Inverse model and parameters

Inputs to the inverse model for analyzing the experimental data are the same as those described earlier for the synthetic case, except for the volumetric discharge from each pumping test and the temporal moment data. The temporal moments of the drawdown are calculated using the moment generation function. The nth temporal moments [$M_n[x_i]$] of drawdown at location x_i are given by

$$M_n(x_i) = \int_0^\infty t^n s(x_i,t)dt \qquad [11.1]$$

where t is time, $s(x_i,t)$ is drawdown defined as $s(x_i,t) = H_0 - h(x_i,t)$, H_0 is the initial hydraulic head treated to be constant everywhere in the domain, and $h(x_i,t)$ is hydraulic head. One can compute the 0th temporal moment using the moment generating function by setting $n = 0$:

$$M_0(x_i) = \int_0^\infty s(x_i,t)dt \qquad [11.2]$$

The 0th moment represents the area under the drawdown-recovery curve. Here, we calculate the 0th moment of the laboratory pumping test data using the following approximation:

$$M_0 = \sum_{i=1}^{k} s_i \Delta t_i = s_1(t_1 - t_0) + s_2(t_2 - t_1) + s_3(t_3 - t_2) + ... + s_k(t_k - t_{k-1}) \qquad [11.3]$$

The 1st moment can be computed by setting $n = 1$ in eqn. 11.1:

$$M_1(x_i) = \int_0^\infty ts(x_i,t)dt \qquad [11.4]$$

Likewise, the 1st moment of the laboratory pumping test data is computed using the following approximation:

$$M_1 = \sum_{i=1}^{k} t_i s_i \Delta t_i = t_1 s_1 (t_1 - t_0) + t_2 s_2 (t_2 - t_1) + t_3 s_3 (t_3 - t_2) + ... + t_k s_k (t_k - t_{k-1}) \qquad [11.5]$$

The 1st moment represents the arrival time of the center of mass for the drawdown-recovery curve. The characteristic time is then given by

$$Y(x_i) = M_1(x_i) / M_0(x_i) \qquad [11.6]$$

Physically, the characteristic time corresponds to the arrival time of the center of the mass under the drawdown-recovery curve normalized by the area.

The HT-m code can also incorporate available small scale measurements of K and S_s, but we do not use these measurements to condition the estimated parameter fields to test the inversion algorithm.

Illman et al. [2007] found that the variations in constant head boundary conditions from one test to the next can be critical in conducting a SSHT survey. However, for THT and HT-m surveys, this is not critical because the analysis is done using drawdowns and moments of drawdown-recovery curves instead of actual head values.

11.5 RESULTS FROM INVERSE MODELING OF CROSS-HOLE PUMPING TESTS CONDUCTED IN A LABORATORY SANDBOX

11.5.1 *K* and S_s tomograms from the inversion of laboratory sandbox data

Figures 11.7a-d are the K tomograms obtained by inverting moment data induced by the eight pumping tests [44, 47, 35, 32, 17, 14, 5, 2] conducted in that order. The result [Figure 11.7a] using the first two pumping tests reveals little detail to the heterogeneity pattern throughout the sandbox. This is alleviated with the heterogeneity structure for the entire aquifer appearing when additional cross-hole tests are included sequentially [Figures 11.7b through 11.7d], although the results look smoothed in comparison to those obtained previously using SSHT by *Illman et al*. [2007] and THT by *Liu et al*. [2007]. The K tomogram appears smoothed because the calculation of the temporal moments used in the algorithm involves the integration of the drawdown-recovery data. This, in effect, results in the smoothing of the drawdown-recovery data set.

Despite the lack of resolution, the results show that the HT-m algorithm is capable of capturing the pattern of the K distribution. We also visually compared the K tomogram resulting from HT-m to that obtained from THT [figure 10.2g] by *Liu et al*. [2007]. This comparison revealed that the tomograms are similar.

Figure 11.8a-d shows the corresponding S_s tomogram that was estimated simultaneously. In contrast to Figure 11.7a - d, the structure consisting of variable size sand bodies visible in the K tomogram is not visible for the S_s tomogram. This can be attributed to the fact that sands of relatively low compressibility [of various sizes] were used to construct the synthetic aquifer. However, a decreasing trend in S_s with depth is apparent. Physically speaking, this makes sense because the sands in the upper portion are less compressed, while the deeper sands are more

compressed due to the stress exerted by the overlying material. This finding perhaps suggests that the K values are not significantly correlated with the S_s values in this sandbox.

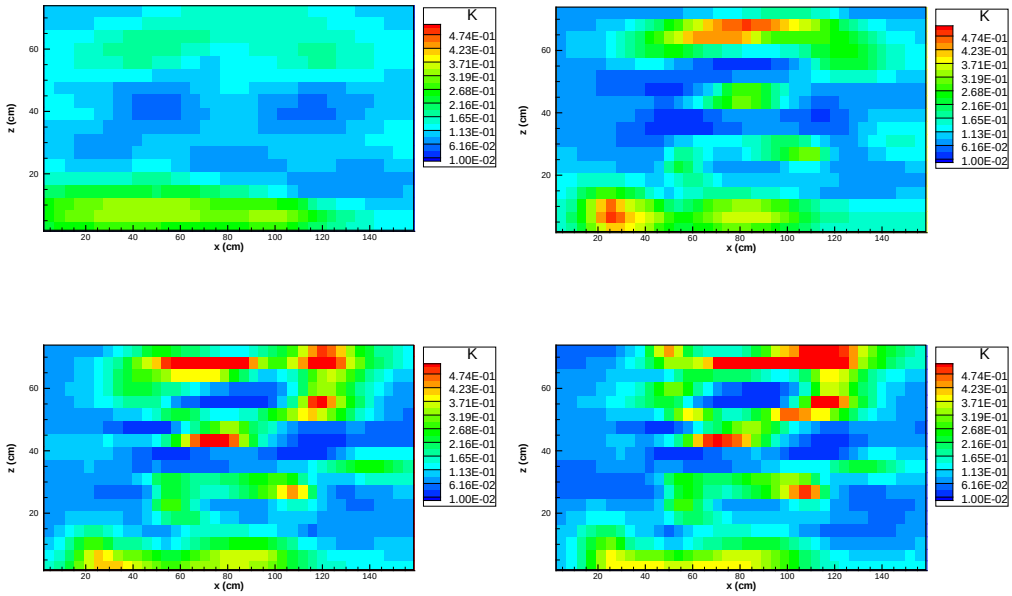


Figure 11.7: K tomograms computed using HT-m from real cross-hole pumping test data included sequentially. a] Ports 44, 47. b] Ports 44, 47, 35, 32. c] Ports 44, 47, 35, 32, 17, 14. d] Ports 44, 47, 35, 32, 17, 14, 5, 2. Port numbers [see Figure 6.1] indicate those used as the pumped well for each cross-hole test.

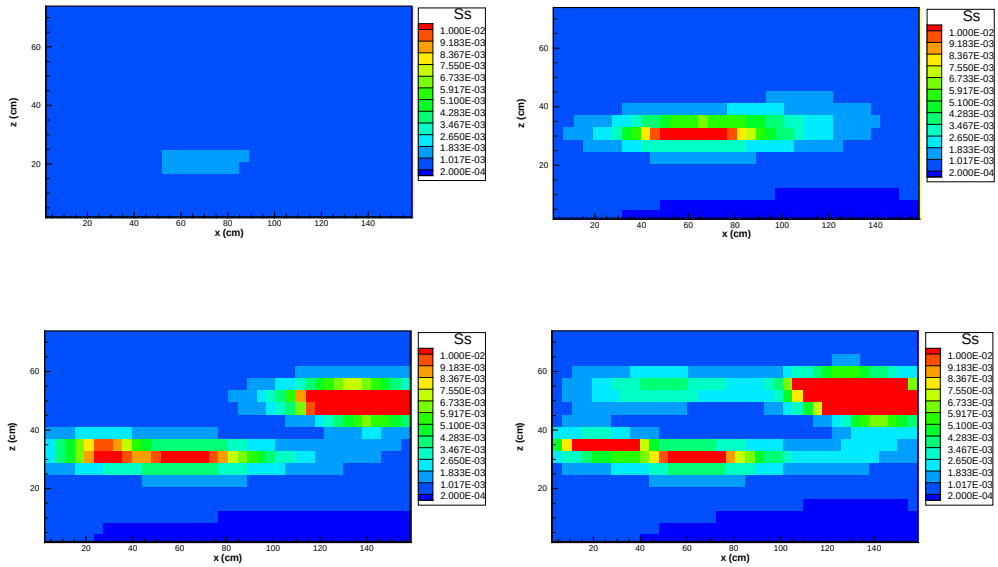


Figure 11.8: S_s tomograms computed using HT-m from real cross-hole pumping test data included sequentially. a] Ports 44, 47. b] Ports 44, 47, 35, 32. c] Ports 44, 47, 35, 32, 17, 14. d] Ports 44, 47, 35, 32, 17, 14, 5, 2. Port numbers [see Figure 6.1] indicate those used as the pumped well for each cross-hole test.

11.5.2 Visual evaluation of computed tomograms

A visual comparison of the K tomogram [Figure 11.7d] obtained by sequentially including 8 tests into the inversion algorithm and CAD drawing showing the location of low K sand bodies [Figure 11.1] shows that the HT-m is in general, able to delineate the major low and high K features that comprise the aquifer heterogeneity. We also visually compare the K tomogram resulting from the real data sets [Figure 11.7d] to the synthetic K tomogram obtained via the HT-m algorithm [Figure 11.4d]. This comparison shows that the tomogram from the real data shows some semblance to the results from the synthetic data, but the real results [Figure 11.7d] appear considerably smoother than the synthetic results [Figure 11.4d].

We do not make a direct visual comparison of the S_s tomogram computed using the HT-m algorithm to the blocks shown on the CAD drawing, as the S_s values do not vary considerably from one sand type to the next. We also note that a visual comparison cannot be made with the synthetic tomograms [Figures 11.3d and 11.5d] because the latter are synthetic. The S_s tomogram which we can compare Figure 11.8d is to that obtained by *Liu et al.* [2007] and this is done later through a scatter plot.

11.5.3 Joint validation of K and S_s tomograms

As shown in previous sections, a visual comparison of the tomograms to the CAD drawing and results from synthetic simulations is one approach to validate the quality of the computed tomograms. However, a more robust approach to validate is to test the predictability of the estimates under different flow scenarios.

In order to do this, we validate the K and S_s tomograms by simulating an additional pumping test that was not used in the inversion and to examine whether the drawdown at various sampling ports of this independent test can be predicted accurately at various times. For this, we utilize the K and S_s tomograms obtained from the inversion of the 8 pumping tests [Figures 11.7d and 11.8d] and simulate a test with pumping taking place at port 46. Pumping test at port 46 was chosen for validation purposes because it was one of the pumping tests with the cleanest data devoid of external factors. We then compare the simulated and measured transient drawdown at early [3 sec], intermediate [50 sec] and late time [150 sec] periods at all ports except for the pumped port.

Figure 11.9b was obtained in a similar fashion, but the K and S_s fields are homogeneous and these equivalent values [$K = 1.73 \times 10^{-1}$ cms^{-1} & $S_s = 2.30 \times 10^{-4}$ cm^{-1}] were obtained previously by *Liu et al*. [2007]. The correlation values between the simulated and observed drawdowns at the three times as well as the means and variances of their differences are reported in Table 11.3.

Figure 11.9a shows that the comparison using early time data are heavily biased. However, the comparison improves with time considering the results from the later times 50 secs and 150 secs. The late time results show that the bias becomes negligible [i.e., data become evenly scattered along the 45 degree line], indicating that predicted drawdown distributions generally are statistically unbiased in comparison with the observed.

The means and variances of the differences between the simulated and observed drawdown values shown in Table 11.3 also indicate that the results improve over time. In addition, the relatively high correlation values of 0.857 for 50 secs and 0.854 for 150 secs suggest the predicted drawdown distribution is almost identical to the observed distribution—at least the drawdowns at late times at the observation ports. This is an encouraging result because it indicates that using the K and S_s fields derived from HT-m approach, one can obtain a very

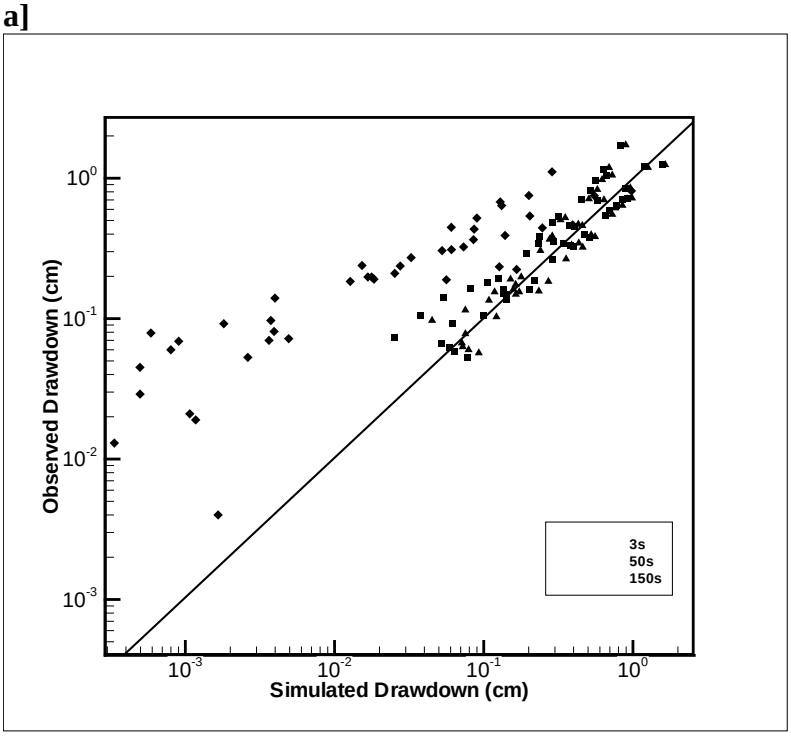
good prediction of the drawdown behavior in the sand box. However, we note that these results are not as robust as those obtained for the THT approach presented in *Liu et al*. [2007], in which those authors obtained correlation coefficients of 0.995 over the time periods of 3, 10, and 20 secs using the same cross-hole pumping test data.

The results from simulating pumping test with pumping taking place at port 46 using the homogeneous K and S_s fields [Figure 11.9b] shows that there is a slight bias in results at all times. It is also evident from Table 11.3 that the means and variances of the difference between the simulated and observed drawdowns increase with time, while the correlation coefficient does not vary significantly.

Table 11.3: Summary of the mean and variance of the estimation errors from the comparison of observed to simulated drawdowns. The K and S_s fields for the heterogeneous case are from the inversion of the laboratory sandbox data using the HT-m algorithm, while the homogeneous case represents equivalent K and S_s obtained from *Liu et al*. [2007]. The correlation coefficients shown in this table are for the scatter plots shown in Figures 11.9a-b.

	Heterogeneous fields			Homogeneous fields		
Time [sec]	Mean	Variance	Correlation	Mean	Variance	Correlation
3	0.187	0.029	0.716	0.108	0.021	0.739
50	0.125	0.026	0.857	0.161	0.052	0.782
150	0.127	0.024	0.854	0.166	0.054	0.775

a]



b]

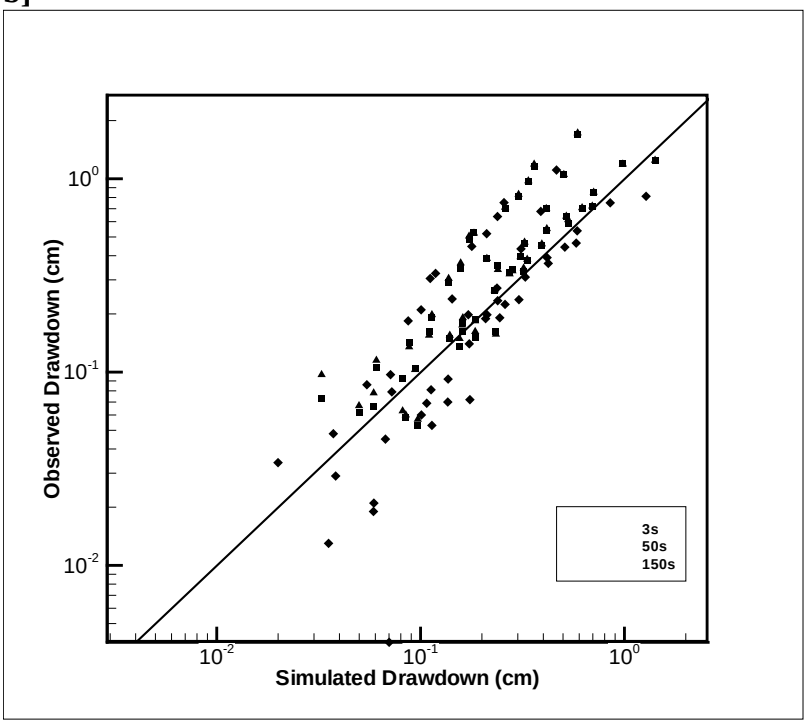


Figure 11.9: Scatter plot of simulated drawdown versus observed drawdown at 3, 50, and 150 seconds after pumping begins for cross-hole hydraulic tests at port 46. In this figure, the simulated values were obtained through synthetic simulations of the cross-hole test conducted using: a] the K and S_s tomograms [Figures 11.8d and 11.9d] and b] equivalent K and S_s values from *Liu et al*. [2007]. The observed drawdown values are from an actual cross-hole test conducted at port 46.

11.6. DISCUSSION

11.6.1 Comparison of *K* tomogram from HT-m to SSHT

We next compare the local values of K obtained from the K tomogram computed using the HT-m algorithm to those from the SSHT algorithm in Figure 11.10. Here, the K values from the SSHT were obtained by *Illman et al*. [2007] using the same 8 pumping tests. The result shows that the estimated K values using the temporal moments of the drawdown-recovery curve are similar to those based on the steady head data but the scatter is very large which is reflected in a low correlation value of 0.104. The difference between the two perhaps reflects the influence of S_s parameters on the estimation of K.

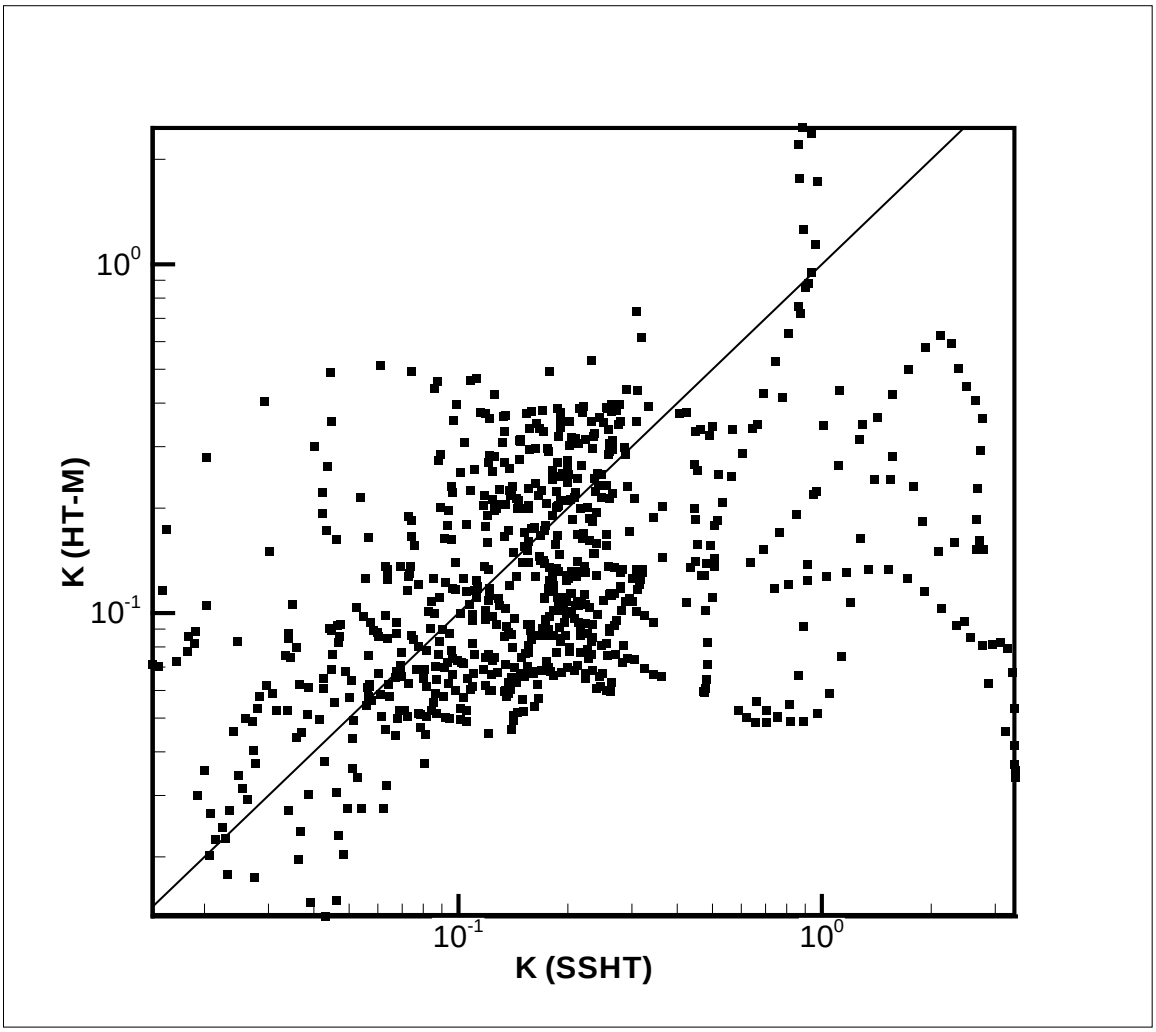


Figure 11.10: Scatter plot of *K* values from the HT-m of this paper and those from the SSHT in *Illman et al*. [2007].

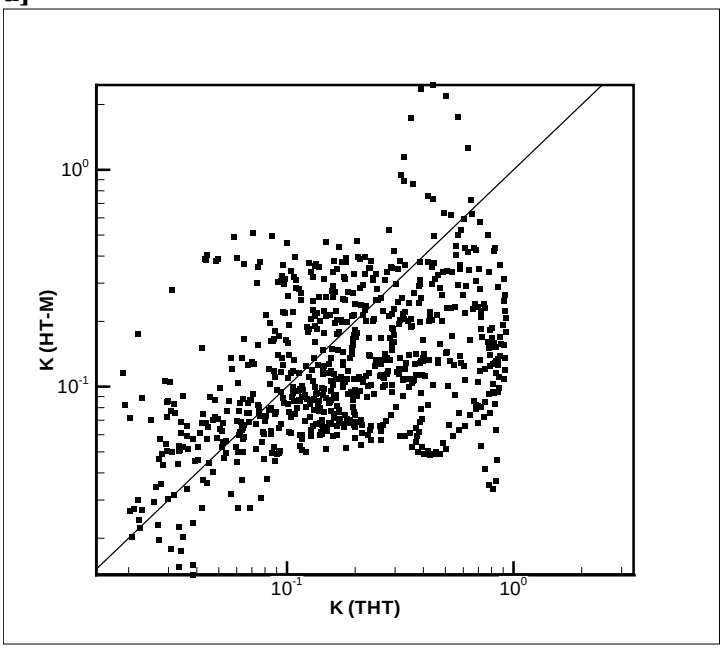
11.6.2 Comparison of K and S_s tomograms from HT-m to THT

We next compare the local values of K and S_s obtained from the K and S_s tomograms computed using the HT-m algorithm to those from the THT algorithm in Figures 11.11a-b. Here, the local K and S_s data from the THT algorithm are from Figures 2f and 3f in *Liu et al.* [2007]. The results [Figure 11.11a] show that the estimated K values using the HT-m approach are similar to those based on the transient head data with a moderately high correlation value of 0.587. The difference between the two perhaps reflects the effect of smoothing rendered by the computation of the temporal moments on the estimation of K.

Examination of Figure 11.11b shows that the estimation of S_s from the HT-m approach appears to be less reliable. This could be due to three possibilities. One possibility is that the HT-m approach relies on the computation of the temporal moments of the drawdown-recovery data, which involves integration. Data integration causes smoothing and perhaps loss of information on aquifer heterogeneity. In particular, the computation of moments and integration could cause significant loss of information on S_s which is contained in the early-time data. Previous research by *Zhu and Yeh* [2006] and others suggest that this may be the case.

Another possible explanation is that the use of the temporal moments, which requires the use of entire drawdown-recovery curves, may result in a loss of sensitivity to the inversion of S_s. We also found that the characteristic time was considerably noisier than the 0th moments contributing to the ineffectiveness in estimating S_s. Therefore, the estimation of S_s spatial variability using the HT-m approach was found to be more challenging than the THT approach.

a]



b]

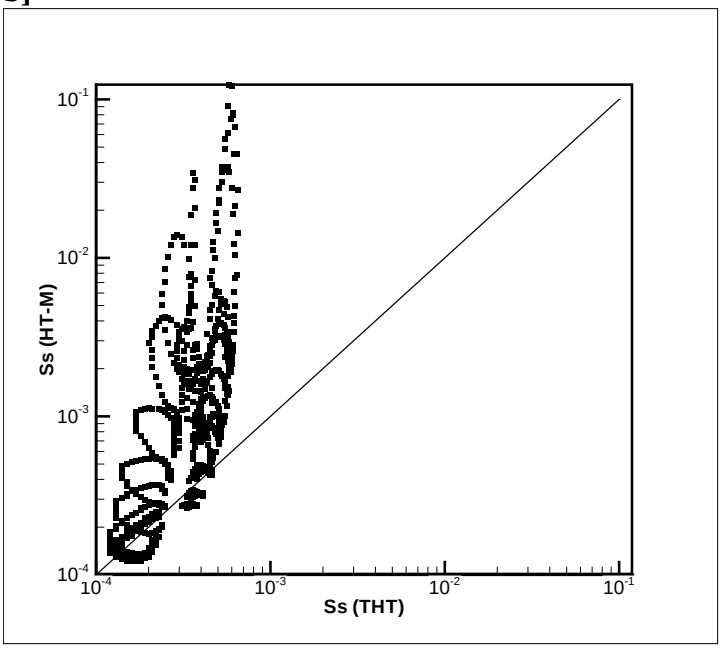


Figure 11.11: Scatter plot of a] K and b] S_s values from the HT-m approach of this paper and those from the THT approach described in *Liu et al*. [2007].

12. COMPARISON OF AQUIFER CHARACTERIZATION APPROACHES THROUGH STEADY-STATE GROUNDWATER MODEL VALIDATION: A CONTROLLED LABORATORY SANDBOX STUDY

12.1 INTRODUCTION

The characterization of subsurface heterogeneity of hydraulic parameters is of paramount importance to water supply assessment or protection, and the remediation of groundwater throughout the world. Many communities rely on groundwater and in the case of Kitchener-Waterloo region of southern Ontario in Canada, over 80% of the water supply comes from groundwater. Unfortunately in Walkerton, Ontario, Canada, over 2,300 people have fallen ill and 7 of them have passed due to the bacterial contamination of municipal wells [*Walkerton Inquiry*, 2002]. Since this tragedy, a large effort has gone into delineating groundwater protection areas for water supply wells in Canada. The accurate delineation of such areas relies on detailed and reliable information on subsurface heterogeneities of hydraulic parameters such as hydraulic conductivity [K] and specific storage [S_s].

A number of aquifer characterization approaches to delineate the heterogeneity in hydraulic parameters have been developed and tested over the last several decades [e.g., *Sudicky*, 1986, *Boggs et al*., 1990; *Wolf et al*., 1992; *Hess et al*., 1992; *Sudicky et al.*, 2009]. The classical approach in delineating K heterogeneity is to collect a large number of core samples from multiple boreholes and to conduct grain size or permeameter analyses. Other approaches include the slug testing of large numbers of piezometers and/or monitoring wells [e.g., *Rehfeldt et al*., 1992], flow meter [e.g., *Hufschmied*, 1986; *Molz et al*., 1989; *Boman et al*., 1997], steady-state dipole flow [e.g., *Kabala*, 1993; *Zlotnik et al*., 2001] or single-hole pumping or injection tests conducted at successive intervals in multiple boreholes to obtain a large number of point scale K estimates, that are subsequently analyzed using geostatistical methods. More recently, geophysical methods [e.g., *Hyndman and Gorelic*k, 1996; *Hubbard and Rubin*, 2000] and hydraulic tomography [e.g., *Neuman*, 1987; *Gottlieb and Dietrich*, 1995; *Yeh and Liu*, 2000; *Bohling et al.*, 2002; *Zhu and Yeh*, 2005, 2006] have been used to characterize subsurface heterogeneity. In particular, *Sudicky* [1986] obtained over 1,279 K estimates from cores taken at the Borden site in Canada through a falling head permeameter and conducted geostatistical analysis to estimate the macrodispersion observed during a long-term tracer test at the site. He found good agreement between the macrodispersion coefficient determined from the long-term tracer test and those computed independently from stochastic theory using the statistical properties of K. This study was the first to validate the stochastic theory of *Gelhar and Axness* [1983] through the independent collection of K heterogeneity data and its application to predict macrodispersion of tracers at the Borden site.

Several researchers have compared the validity of various characterization approaches in the field [e.g., *Lee et al*., 1985; *Zlotnik and Zurbuchen*, 2003; *Butler*, 2005]. For example, *Butler* [2005] compared the K results from permeameter analysis of core samples, traditional slug tests, dipole flow tests, multilevel slug tests, borehole flowmeter tests, direct-push slug tests, pumping tests and, hydraulic tomography at the Geohydrologic Experimental and Monitoring Site [GEMS] in Kansas. He found that different characterization approaches yield reliable estimates of K estimates along the boreholes at the site. However, he found that most of the techniques examined cannot provide information on subsurface heterogeneity between wells.

Recently, hydraulic tomography has been developed to obtain information on subsurface heterogeneity of K and S_s through sequential pumping tests. Hydraulic tomography is similar in concept to the Computerized Axial Tomography [CAT] scan technology, but the energy source is a change [decrease or increase] in hydraulic head achieved through water pumping or injection, and the sensors [pressure transducers] measure the propagation of head change at multiple locations throughout the aquifer. These hydraulic head data are then used to interpret the spatial distribution of K, S_s, connectivity of highly permeable zones, and to quantify uncertainties associated with the interpretations. The result is a highly resolved three-dimensional depiction of hydraulic parameters and their uncertainties.

Various inverse methods have been developed for hydraulic tomography, which utilize pumping test data simultaneously or sequentially [e.g., *Gottlieb and Dietrich,* 1995; *Yeh and Liu,* 2000; *Bohling et al.,* 2002; *Brauchler et al.,* 2003; *McDermott et al.,* 2003; *Zhu and Yeh,* 2005, 2006; *Li et al*., 2005; *Fienen et al*., 2008; *Castagna and Bellin*, 2009; *Xiang et al*., 2009]. Pneumatic tomography is similar in concept to hydraulic tomography, but the cross-hole tests are conducted with air in the unsaturated zone [*Illman and Neuman*, 2001, 2003] and interpreted using a stochastic inverse model [*Vesselinov et al*., 2001a-b] which interprets multiple tests simultaneously or sequentially.

In the quest of quantitatively evaluating the performance of hydraulic tomography, various researchers [e.g., *Liu et al.,* 2002; *Illman et al.,* 2007, 2008; *Liu et al.,* 2007; and *Yin and Illman,* 2009] have conducted laboratory scale studies in which the heterogeneity patterns can be prescribed and forcing functions controlled. In particular, *Illman et al*. [2007, 2008] showed that steady state hydraulic tomography [*Yeh and Liu*, 2000] is a promising technique in imaging subsurface heterogeneity patterns of K [from now on K tomograms] through laboratory sandbox studies. *Liu et al*. [2007] did the same, but used transient hydraulic tomography [*Zhu and Yeh*, 2005] to image K and S_s heterogeneity patterns [from now on K and S_s tomograms]. Most recently, *Yin and Illman* [2009] showed that the hydraulic tomography based on the temporal moments of drawdown-recovery data [*Zhu and Yeh*, 2006] yield a satisfactory K tomogram, but an unreliable S_s tomogram. In these previous studies, different methods and data collected at different scales were utilized to evaluate the tomograms. Among the various methods for tomogram evaluation, both *Illman et al*. [2007, 2008] and *Liu et al*. [2007] found that the best method was to simulate an independent cross-hole pumping test not used in the construction of the tomograms. In particular, *Illman et al*. [2007, 2008] were successful in predicting the steady state head response from a single cross-hole pumping test using the K tomogram from steady state hydraulic tomography. Likewise, *Liu et al*. [2007] were successful in predicting the transient head response of a single, independent cross-hole pumping test using the K and S_s tomograms from transient hydraulic tomography. These results were encouraging in that an independently conducted test can be predicted accurately if the heterogeneity pattern and forcing functions are all known. However, the experimental finding could have been construed to be weak and fortuitous because there was only one independent test from the experiments conducted by these authors previously.

In the field, transient hydraulic tomography has also been applied in unconsolidated media [*Straface et al*., 2007], while the hydraulic tomography based on the steady shape analysis [*Bohling et al*., 2002] was reported for unconsolidated materials by *Bohling et al*. [2007]. More recently, *Li et al*. [2008] utilized their hydraulic tomography approach to image floodplain deposits. Likewise, *Cardiff et al*. [2009] conducted a 2D depth-averaged hydraulic tomography analysis under steady state conditions in an unconfined aquifer. Most recently, *Illman et al*.

[2009] used transient hydraulic tomography to image heterogeneity patterns of K and S_s in fractured granite at the kilometre scale, which appear to resemble connective fault zones and low K barriers. The numerical study by *Hao et al*. [2008] supports the notion that hydraulic tomography can be effective in imaging the connectivity of fractured rocks.

There is a critical need to assess the capabilities of various characterization approaches in a controlled environment. There are a number of methods to assess these approaches, but we feel that this is accomplished best through the prediction of independently conducted drawdown inducing stresses induced by pumping tests or through the careful measurement of flow rates through the aquifer volume under question. The field assessment of various characterization techniques is of course most desirable, but field experiments can be affected by various external forcings, uncontrollable noise and can be costly. In contrast, laboratory experiments can be efficient in assessing hydraulic tomography and other characterization approaches because both forcing functions and experimental errors can be controlled. The main objective of this paper is to assess K estimates obtained by various approaches by predicting independent cross-hole pumping tests and total flow rates from flow-through tests in a synthetic heterogeneous aquifer. Specifically, we: 1] characterize the synthetic heterogeneous aquifer built in the sandbox through various techniques [permeameter analyses of core samples, single-hole pumping tests, cross-hole pumping tests, and flow-through tests], 2] obtain the mean K fields through traditional analysis of test data by treating the medium to be homogeneous, 3] obtain heterogeneous K fields through kriging and steady state hydraulic tomography, and 4] conduct forward simulations of 16 cross-hole pumping tests and 6 flow-though tests using these homogeneous and heterogeneous K fields and comparing them to actual drawdown and flow rate data from the laboratory experiments. The direct comparison of simulated and measured drawdowns as well as its statistical analysis yields a quantitative performance measure for each of these K fields. Moreover, the direct comparison of simulated and measured flow rates from the flow-through experiments allow for an additional test to quantify the performance of the various K fields. The comparison also amounts to the validation of various steady state groundwater models with different K fields obtained by various aquifer characterization methods.

12.2 SANDBOX AND SYNTHETIC HETEROGENEOUS AQUIFER CONSTRUCTION

12.2.1 Sandbox construction

A synthetic heterogeneous aquifer was constructed in a sandbox to validate various fluid flow and solute transport algorithms. The sandbox is 193.0 cm in length, 82.6 cm in height, and has a depth of 10.2 cm. All materials used inside the sandbox are made of 316 stainless steel, brass, or Viton®. Forty eight ports, 1.3 cm in diameter, have been cut out of the stainless steel wall to allow coring of the aquifer as well as installation of horizontal wells. Each well was constructed by making 6 cuts spaced 1.46 cm apart in sections of brass tubing. The cuts were then covered with a stainless steel mesh that was bonded to the tubing with corrosion resistant epoxy. Extreme care was taken to avoid the epoxy filling the mesh which could impede water flow. The wells fully penetrate the thickness of the synthetic aquifer. This allowed each location to be monitored by a pressure transducer, used as a pumping or an injection port and as a water sampling port. Figure 12.1 is a schematic drawing of the frontal view of the synthetic aquifer, showing the 48 port and pressure transducer locations.

The flow system for the sandbox is driven by two constant-head reservoirs, one at each end of the sandbox. The adjoining reservoirs are capable of supplying water throughout the

length and thickness of the sandbox quickly and efficiently. In this sandbox, a series of 4-inch perforated plate/mesh combination was installed at each end of the sandbox to serve as a porous media/water interface and provide hydraulic control. The boundary head levels can be easily adjusted to be equal or to create a desired hydraulic gradient. The developed system is capable of maintaining 3 constant head boundaries simultaneously by ponding water at the top in addition to fixing the hydraulic heads in the 2 constant head reservoirs.

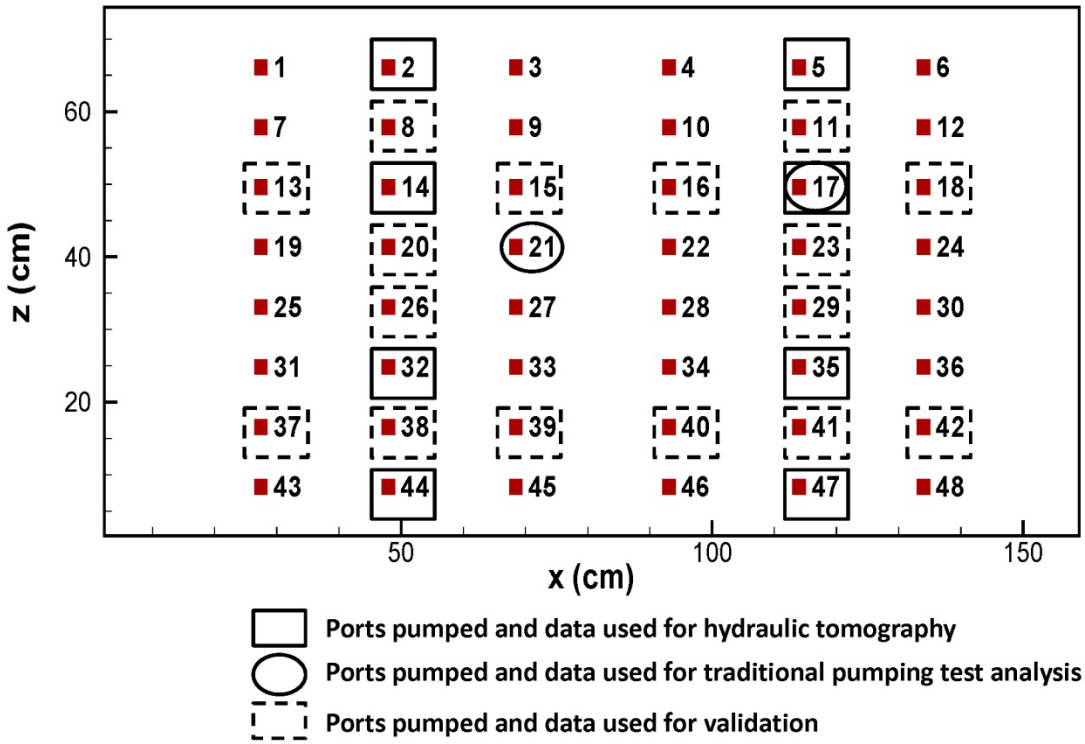


Figure 12.1: Schematic diagram of synthetic heterogeneous aquifer used for validation of steady state groundwater flow models. Numbers next to solid squares indicate port numbers, open squares around numbers indicate the 8 ports [2, 5, 14, 17, 32, 35, 44, 47] used for hydraulic tomography, open ovals around numbers indicate the 2 ports [17 and 21] pumped and data analyzed using VSAFT2 by treating the medium to be homogeneous [i.e., traditional pumping test analysis], and the dashed open squares around the 16 other ports [8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42] indicate the pumping locations for the independent cross-hole pumping tests used for validation purposes.

The data acquisition system used for the laboratory experiments consisted of three major components. Pressure measurements were made with 50 *Setra* model 209 gage pressure transducers with a range of 0 to 1 psi, 48 of which measured hydraulic head in the aquifer and one in each constant head reservoir. These pressure transducers were installed at each of the 48 data acquisition ports in the stainless steel wall of the sandbox. The second component was a 64-channel data acquisition board from *National Instruments.* A hub that separates excitation and output currents for the transducers was also assembled. The third component was a dedicated PC with *National Instruments LabVIEW* software for automated data acquisition. Further details of the sandbox construction can be found in *Craig* [2005].

12.2.2 Synthetic aquifer construction through cyclic flux of sediment laden water

Sandbox experiments have been conducted for a variety of purposes using different methods for aquifer construction. Early studies utilized uniform packing of sands to create a homogeneous medium and uniformly heterogeneous packing [e.g., *Silliman and Simpson*, 1987; *Schincariol and Schwartz* 1990, *Illangasekare et al*., 1995]. More recently, complex heterogeneity patterns have been packed by various researchers [e.g., *Welty and Elsner*, 1997; *Silliman et al*., 1998; *Chao et al*., 2000; *Barth et al*., 2001; *Silliman*, 2001; *Danquigny et al*., 2004; *Fernàndez-Garcia et al*., 2005] in two and three-dimensional small and intermediate scale sand tanks to mimic the more complex heterogeneity patterns with statistical properties including correlation lengths that are representative to those found in natural geologic media. The process involves generating a correlated *K* field using a random field simulator and carefully packing the resulting *K* distribution with different size sand blocks. *Jose et al*. [2004] created more natural deposits consisting of lenses rather than blocks through a settling procedure in standing water. This approach produced small scale heterogeneities such as laminations, cross lamination and rippled like structures that are similar to those found in natural fluvial deposits.

In this study, the synthetic heterogeneous aquifer was created through the cyclic deposition of sediments under varying water flow and feed rates of sediments. Previous sandbox aquifer studies by our group [*Illman et al*., 2007; 2008; *Liu et al*., 2007; *Yin and Illman*, 2009] involved careful packing of heterogeneity patterns in a prescribed fashion. Packing of sand bodies by hand is a laborious procedure and complex heterogeneity patterns are hard to create. Our goal in relying on sediment transport was to create a more realistic heterogeneity pattern with various scales of heterogeneity in an efficient manner. The one drawback of this approach is that the heterogeneity pattern in the sandbox cannot be controlled precisely as in the case of packing a sandbox aquifer by hand.

Table 12.1: Characteristics of porous media used to create a synthetic heterogeneous aquifer.

Sand type	d_{50} [mm]	*K* [cm/s] Shepherd	*K* [cm/s] Darcy
16/30	0.872	1.32×10^{-1}	3.84×10^{-1}
20/30	0.750	1.03×10^{-1}	3.12×10^{-1}
20/40	0.578	6.68×10^{-2}	2.05×10^{-1}
#12	0.525	5.70×10^{-2}	2.05×10^{-1}
F32	0.504	5.33×10^{-2}	1.45×10^{-1}
#14	0.457	4.53×10^{-2}	1.21×10^{-1}
4030	0.355	2.99×10^{-2}	5.79×10^{-2}
F55	0.242	1.59×10^{-2}	2.80×10^{-2}
F65	0.204	1.20×10^{-2}	1.83×10^{-2}
F75	0.174	9.22×10^{-3}	1.73×10^{-2}
F85	0.151	7.29×10^{-3}	1.35×10^{-2}

Various commercially available sands of variable grain sizes [Table 12.1] [Unimin Corporation and U.S. Silica] were used to create the synthetic aquifer. Examination of the grain size distribution [not shown here, but available in *Craig*, 2005] shows that the sands are well sorted and the d_{50}, which is the particle diameter for which 50% of the weight is finer, ranges between 0.151 mm to 0.872 mm. Based on *Shepherd* [1989]'s empirical power-law relationship relating the d_{50} to the *K*, the estimated *K* for each class of sands ranges from 7.29×10^{-3} to 1.32×10^{-1} cm/s from the smallest to the largest d_{50}. We also estimated *K* for each sand type using a constant head permeameter yielding a range of 1.35×10^{-2} to 3.84×10^{-1} cm/s from the smallest

to the largest d_{50} showing that the K estimates are in general consistent, but differences do arise from the two approaches.

The sediment layers were deposited using a sediment suspension tank that lies in series with the sandbox and a 1/3 horsepower pump [Figure 12.2]. The pump was operational in one direction only. Therefore, the plumbing was designed to accommodate flow in either direction through the sandbox. The design of how the different sands of varying grain sizes were deposited heavily relied on flow rates into the sandbox from the suspension tank. The flow rates were determined for each sand type empirically by determining the mean velocities for under which a particular sediment type is transported, deposited, and scoured within the sandbox [*Craig*, 2005]. For each layer, a specific sediment type was chosen and introduced into the suspension tank [Figure 12.2]. To introduce heterogeneity among the layers, care was taken to introduce sediments with variable d_{50} from one layer to the next. Deposition of each layer was accomplished by adjusting the flow rate of the sediment-laden water into the sandbox. Under uniform flow conditions, the coarser particles appeared to drop off first and progressively finer particles deposited on top of the coarser ones creating small scale heterogeneity within each deposited layer. To create heterogeneity at the larger scale and inter-fingering of different sediment layers, the flow of water into the tank was alternated after the deposition of each layer. In total, 17 layers were deposited to mimic an inter-fingering fluvial deposit with the lowest deposit in the sandbox designated as layer 1 with additional deposits overlying labelled in an ascending numerical order. Table 12.2 summarizes the layers deposited, its sand type, and the volume of sand deposited for a particular layer. Because the deposition of the final layer was not possible as the height of the synthetic aquifer exceeded the top most orifice in the sandbox used for sediment transport, 20/30 sand was placed over layer 17 to fill the tank to the top. Figure 12.3 is a photograph of the sandbox showing each of the deposited layers. Further details to the creation of the synthetic heterogeneous aquifer are summarized in *Craig* [2005].

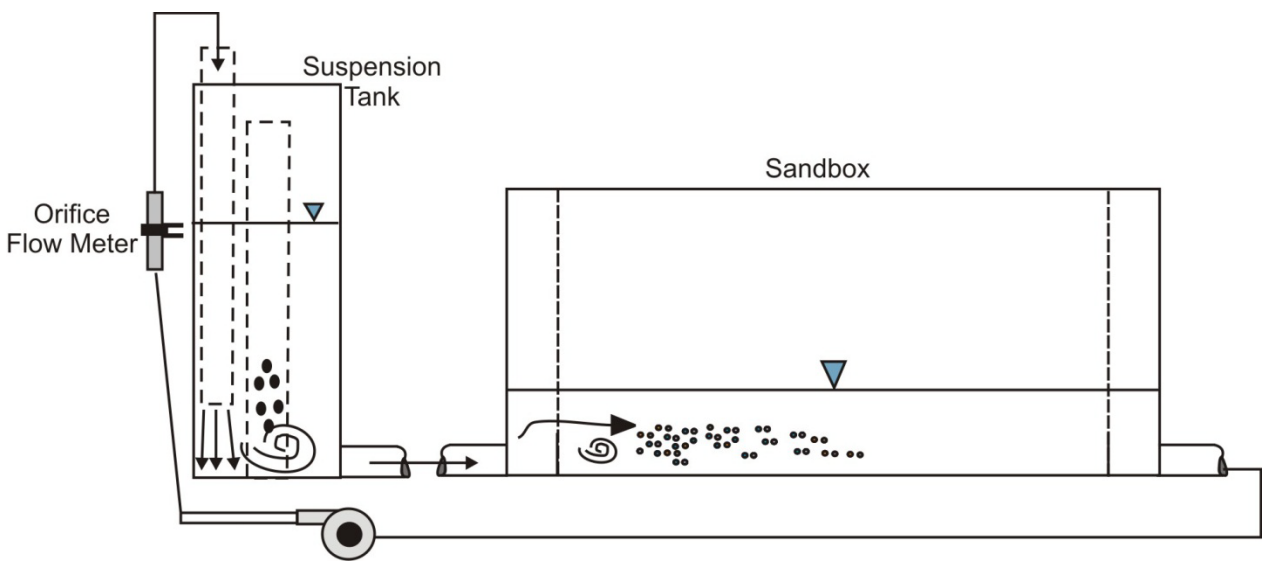


Figure 12.2: Schematic diagram of the sediment transport system for the creation of the synthetic heterogeneous aquifer.

Table 12.2: The sand type and deposited volume for each layer in the sandbox aquifer.

Layer	Sand type	Volume [m^3]
1	20/30	0.0051
2	4030	0.0034
3	F85	0.0031
4	20/40	0.0051
5	Mix	0.0034
6	Mix	0.0034
7	#12	0.0043
8	F32	0.0034
9	20/40	0.0031
10	F65	0.0043
11	#12	0.0019
12	16/30	0.0024
13	20/30	0.0051
14	F-75	0.0043
15	20/40	0.0043
16	Mix	0.0043
17	F-85	0.0043
18	20/30	0.0051

Note: The layers labelled "mix" consisted of equal volumes of #14, F75, and 16/30 sands.

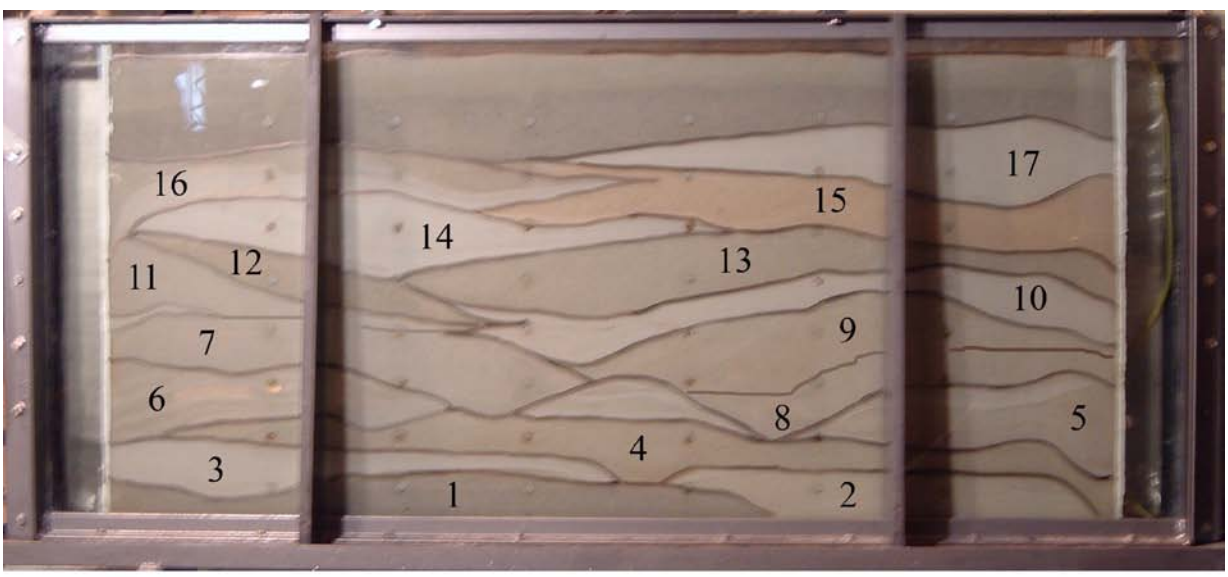


Figure 12.3: Photograph of synthetic heterogeneous aquifer created via cyclic flux of sediment-laden water.

12.3 CHARACTERIZATION OF THE SYNTHETIC HETEROGENEOUS AQUIFER TREATING THE MEDIUM TO BE HOMOGENEOUS

The characterization of the synthetic heterogeneous aquifer was accomplished by various established aquifer characterization methods and with hydraulic tomography. Our main goal was to test the ability of these approaches to characterize the aquifer at different scales. Below, we describe the various characterization approaches and report the geometric mean of K values [K_G]

determined from each test type as well as corresponding variance [$\sigma^2_{\ln K}$] estimates in Table 12.3. The K_G and $\sigma^2_{\ln K}$ were computed using:

$$K_G = \exp\left(\frac{1}{n}\sum_{i=1}^{n}\ln K_i\right) \qquad [12.1]$$

$$\sigma^2_{\ln K} = \frac{\sum_{i=1}^{n}\left(\ln K_i - \overline{\ln K}\right)^2}{(n-1)} \qquad [12.2]$$

where n is the sample size, i is the index number, ln K is the natural logarithm of K, and $\overline{\ln K}$ is the arithmetic mean of ln K values.

Table 12.3: Summary of K estimates from various approaches.

Test type	n	Min K [cm/s]	Max K [cm/s]	K_G [cm/s]	$\sigma^2_{\ln K}$
Core	48	0.0111	0.298	0.077	0.868
Single-hole	48	0.014	0.320	0.059	0.384
Cross-hole [Port 17]	48	0.021	0.315	0.076	0.301
Cross-hole [Port 21]	48	0.054	0.420	0.111	0.224
Flow-through	6	0.131	0.141	0.136	0.001

12.3.1 Constant head permeameter analysis of core samples

Core samples from the synthetic aquifer were obtained by inserting a horizontal core tube through the 48 wells. The extracted cores had dimensions of 1.28 cm in diameter and 10.16 cm in length. After the collection of the core, a cotton cloth with permeability much higher than the core samples was placed on each end and prepared for their analysis in a constant head permeameter [*Klute and Dirksen*, 1986]. Because the extracted core was very small, a special permeameter was built for testing purposes. Inspection of cored sands revealed that sands are quite uniform, thus we do not expect significant anisotropy in K measurements. Details to the core extraction method and the design of the constant head permeameter are provided in *Craig* [2005].

12.3.2 Single-hole pumping tests

We then conducted single-hole pumping tests at each of the 48 ports. The tests were conducted by pumping water at each port and monitoring the transient head change within the pumped well. A constant pumping rate [Q = 1.25 cm^3/sec] was set for each single-hole pumping test. For each test, data collection started without the pump running in order to obtain the initial hydraulic head in the sandbox at all measurement ports. A peristaltic pump was then activated at the pumping port and allowed to run at a constant rate until the development of steady state flow conditions. The pump was then shut off to collect recovery data. The steady state head data were analyzed using VSAFT2 [*Yeh et al*., 1993] through manual calibration by treating the aquifer to be homogeneous. Details to the numerical modeling and calibration effort are provided in *Craig* [2005].

12.3.3 Cross-hole pumping tests

We then conducted cross-hole pumping tests at each port along columns 2 [ports 2, 8, 14, 20, 26, 32, 38, and 44] and 5 [ports 5, 11, 17, 23, 29, 35, 41, and 47] and 9 additional pumping

tests at various ports outside of these two columns [ports 13, 15, 16, 18, 21, 37, 39, 40, and 42] [see Figure 12.1]. The cross-hole tests were conducted by pumping at rates ranging from 2.50 - 3.17 cm^3/sec at 25 separate ports indicated by open and dashed squares on Figure 12.1. Prior to each cross-hole pumping test, all pressure transducers were calibrated to ensure accurate data collection. We next collected hydraulic head data for several minutes in all pressure transducers to establish a static, initial condition. After establishment of static conditions, we pumped from each port using a peristaltic pump, while taking head measurements in all 48 ports and a pressure transducer in each of the constant head boundaries. For each test, pumping continued until the development of steady state conditions, which was determined by observing the stabilization of all head measurements within the aquifer. The pump was then shut off to collect recovery head data until its full recovery. We note that we only utilized steady-state head data for subsequent analyses to estimate K by treating the medium to be homogeneous and through steady state hydraulic tomography to estimate a K tomogram.

An estimate of K obtained between the pumped and observation intervals when the medium is treated to be homogeneous is considered to be an equivalent hydraulic conductivity [K_{eq}] [*Renard and de Marsily*, 1997; *Neuman*, 2005]. To obtain an estimate of K_{eq} at each of the observation ports, we selected 2 cross-hole pumping tests among the 25 tests. One pumping test took place in port 17 and the other in port 21. Both ports were completed in layer 13. The core samples from ports 17 and 21 yielded K values of 1.64×10^{-2} and 2.46×10^{-1} cm/s, respectively.

The observation head data from the two tests were analyzed by manually calibrating VSAFT 2 and assuming the aquifer is homogeneous. The numerical setup for the calibration was identical to the single-hole test analysis. Analysis of the two pumping tests yielded 96 estimates of K for the equivalent homogeneous medium [K_{eq}]. Values of K_{eq} from each cross-hole pumping test were then averaged using eqn. [12.1] to estimate the K_G for the entire synthetic aquifer.

12.3.4 Bidirectional flow-through tests

We also conducted 6 bidirectional flow-through tests through the entire synthetic aquifer to obtain a large scale K estimate under steady-state flow conditions. Specifically, each of these 6 tests was conducted by changing the height of the constant head reservoirs on both sides of the sandbox. After the flow reached a steady state condition, we measured discharge from one side of the sandbox. We also measured the difference between the heights of the water column in the two constant head reservoirs to determine the hydraulic gradient. We then applied Darcy's law to obtain the effective hydraulic conductivity [K_{eff}] of the synthetic aquifer. This estimate is considered to be the K_{eff} value of the synthetic aquifer as flow takes place throughout the entire synthetic heterogeneous aquifer, the medium is treated to be homogeneous, and is independent of boundary conditions [*Renard and de Marsily*, 1997]. An average of the six different K_{eff} estimates was taken by applying eqn. [12.1] to obtain the K_G.

The 6 flow-through tests were conducted by alternating the flow direction yielding three tests with flow taking place from the left to right and three additional tests with flow taking place from the right to left through the synthetic heterogeneous aquifer. In previous efforts [*Illman et al*., 2007, 2008; *Liu et al*., 2007] flow-through tests were conducted in only one direction using a different sandbox aquifer with a prescribed heterogeneity pattern.

12.3.5 Statistical analysis of test data

Table 12.3 provides a summary of key statistics from the various data collected in this sandbox. Results show that the K_G from various characterization approaches ranges from 0.059 – 0.136 cm/s, while the $\sigma^2_{\ln K}$ ranges from 0.001 to 0.868. Each test type yields a different K_G and $\sigma^2_{\ln K}$ indicating the uncertain nature of estimating these parameters in a heterogeneous aquifer. It is of interest to note that the K_G from the 2 cross-hole tests depends on the pumping location in this sandbox. A similar observation was noted by *Wu et al*. [2005] who used numerical simulations to study the behaviour of effective hydraulic parameters during pumping tests, while *Illman and Neuman* [2001, 2003], *Vesselinov et al*. [2001a], and *Illman* [2006] observed a similar behaviour, in which large scale permeability estimates varied spatially at a field site consisting of unsaturated fractured tuffaceous rocks.

12.4 GEOSTATISTICAL ANALYSIS OF SMALL SCALE *K* ESTIMATES

Geostatistical analysis of core and single-hole *K* data were conducted using the Surfer 8 software developed by Golden Software, Inc [www.goldensoftware.com]. All 48 data were used for the kriging of core and single-hole data. We developed a grid that is compatible with the steady state hydraulic tomography analysis described later so that the estimated *K* fields from both approaches could be compared. The exponential variogram model was used to fit the experimental variograms in both horizontal and vertical directions, resulting in an anisotropic variogram model which can be readily incorporated into the Kriging function in Surfer 8 for interpolation.

Table 12.4 lists the variogram parameters fit to the experimental variograms. Figures 12.4a and 12.4b show the kriged *K* fields derived from the core and single-hole *K* data, respectively. The results, in general, reveal smoother *K* fields in comparison to the interfingering layers shown in Figure 12.3 which is expected considering that there are only 48 data points used for kriging.

Table 12.4: Geostatistical model parameters for kriging core and single-hole ln *K* data.

Data	Model	Nugget	Range [cm]	Sill	Anisotropy ratio
Core	Exponential	0	17	0.90	2.2
Single-hole	Exponential	0	16	0.45	2.0

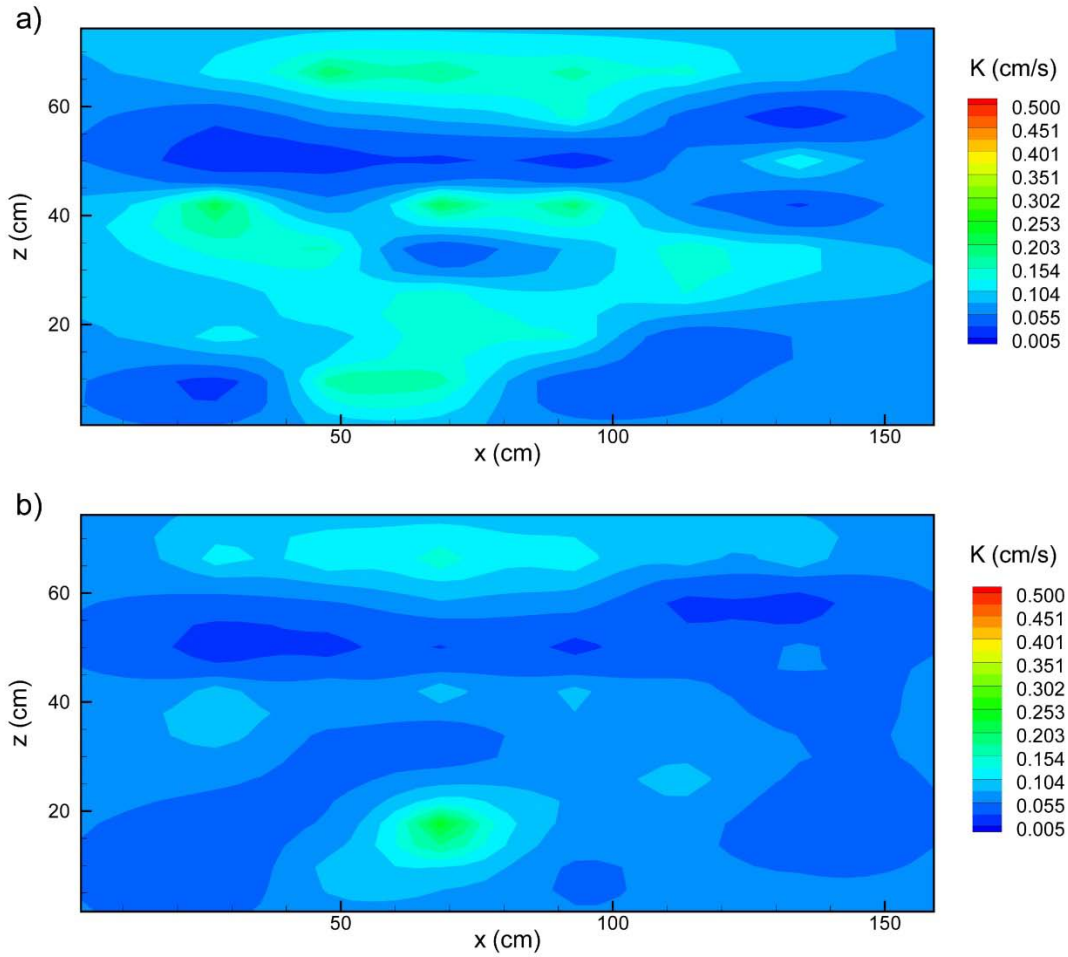


Figure 12.4: *K* distributions obtained via kriging of: a] core *K* estimates and b] single-hole *K* estimates.

12.5. STEADY STATE HYDRAULIC TOMOGRAPHY ANALYSIS OF SELECTED CROSS-HOLE PUMPING TESTS

12.5.1 Inverse model description

The steady state hydraulic tomography analysis of cross-hole pumping tests in the sandbox was conducted using a sequential geostatistical inverse approach developed by *Yeh and Liu* [2000]. We only provide a brief description of the inversion approach here. The inverse model assumes a steady flow field and the natural logarithm of *K* [ln *K*] is treated as a stationary stochastic process. The model additionally assumes that the mean and correlation structure of the *K* field is known *a priori*. The algorithm essentially is composed of two parts. First, the Successive Linear Estimator [SLE] is employed for each cross-hole test. The estimator begins by cokriging the initial estimate of K_{eff} and observed heads collected in one pumping test during the tomographic sequence to create a cokriged, mean removed ln *K* [*f*, i.e., perturbation of ln *K*] map.

Cokriging does not take full advantage of the observed head values because it assumes a linear relationship [*Yeh and Liu*, 2000] between head and *K*, while the true relationship is

nonlinear. To circumvent this problem, a linear estimator based on the differences between the simulated and observed head values is successively employed to improve the estimate.

The second step of *Yeh and Liu* [2000]'s approach is to use the hydraulic head data sets sequentially instead of including them simultaneously in the inverse model thus the term, "sequential successive linear estimator" [SSLE] is used to describe the inverse algorithm hereon. In essence, the sequential approach uses the estimated K field and covariances, conditioned on previous sets of head measurements as prior information for the next estimation based on a new set of pumping data. This process continues until all the data sets are fully utilized. Modifications made to the code for this study include its ability to account for variations in the boundary conditions with each pumping test as they are sequentially included and implementing the modified loop scheme described in *Zhu and Yeh* [2005].

To obtain a K tomogram from the available cross-hole pumping tests, we solve an inverse problem for steady-state flow conditions. The synthetic aquifer was discretized into 741 elements and 1600 nodes with element dimensions of 4.1 cm × 10.2 cm × 4.1 cm. A similar grid was also used previously for steady state hydraulic tomography analysis of a different sandbox aquifer by *Illman et al*. [2007, 2008], transient hydraulic tomography by *Liu et al*. [2007], and hydraulic tomography based on the temporal moments of drawdown-recovery data by *Yin and Illman* [2009].

Illman et al. [2007] found that the variations in constant head boundary conditions from one test to the next can be critical in conducting a steady state hydraulic tomography survey. Therefore, we vary the boundary head values from one test to the next as they are included into the inverse model.

12.5.2 Input parameters and cross-hole tests used

Input data to the inverse model include initial guesses for the K_{eff}, $\sigma^2_{\ln K}$ and the correlation scales [λ_x, λ_y, and λ_z], discharge rate [Q] from each pumping test, and steady-state heads at constant boundaries, as well as available point [small-scale, i.e., core, slug, and single-hole tests] measurements of K. In all previous studies of laboratory sandbox evaluations of hydraulic tomography using a different sandbox aquifer [*Illman et al*., 2007, 2008; *Liu et al*., 2007; *Yin and Illman*, 2009], we excluded small scale data for the inverse modeling effort to test the algorithms without conditioning. For this study, we also exclude core and single-hole K estimates and conduct the steady state hydraulic tomography analysis without conditioning.

We obtained the initial estimate of K_{eff} by taking the geometric mean of all K_{eq} estimates obtained through the analysis of data from observation ports during the cross-hole pumping tests by treating the heterogeneous medium to be homogeneous. This is a reasonable estimate of K_{eff} considering that in practice, a value of K_{eff} is not readily available as a large number of small scale K estimate needs to be collected and a suitable stochastic theory applied to compute the K_{eff}. The traditional analysis of cross-hole pumping tests by treating the medium to be homogeneous [if they are all averaged] should yield a representative value of K for a large portion of the flow and simulation domain. However, our results from Table 12.3 show that cross-hole pumping tests conducted at different sections of the aquifer by pumping at a different port could yield different mean K estimates rendering the estimation of K_{eff} by relying solely on averaging K_{eq} values to be unreliable. Therefore, care must be taken in estimating the K_{eff}.

Estimation of $\sigma^2_{\ln K}$ always involves uncertainty. A previous numerical study conducted by *Yeh and Liu* [2000], however, has demonstrated that $\sigma^2_{\ln K}$ has negligible effects on the

estimated K using the inverse model. Therefore, we use an initial value of 3.0, which is larger than that obtained from the available small scale data [0.87] and set this as our input $\sigma^2_{\ln K}$ in the inverse model.

Correlation scales represent the average size of heterogeneity that is critical for analyzing the average behaviour of aquifers. Correlation scales of any geological formation are difficult to determine. The effects of uncertainty in correlation scales on the estimate based on the tomography are negligible because the tomography produces a large number of head measurements, reflecting the detailed site-specific heterogeneity [*Yeh and Liu*, 2000]. Therefore, the correlation scales were approximated based only on the average thickness and length of the discontinuous sand bodies, which are set as λ_x = 50 cm, λ_y = 10.2 cm, and λ_z = 10.0 cm.

For the steady state hydraulic tomography analysis, we selected 8 pumping tests at ports 2, 5, 14, 17, 32, 35, 44, and 47 and the corresponding steady-state head observations at the rest of 47 ports during each test as data sets. We elected to not use the head data from the pumped port from each test because the pumped port could be affected by skin effects [*Illman et al*., 2007]. The remaining 16 tests with pumping taking place at ports [8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42] were reserved for the validation of the K tomogram.

Prior to the computation of the K tomogram with the SSLE algorithm, we pre-processed the hydraulic head data. This is because *Illman et al*. [2008] found from the analysis of cross-hole pumping test data obtained in a laboratory sandbox aquifer that the signal-to-noise ratio can be critical in inverse modeling of cross-hole pumping test data. Despite calibrating the pressure transducer prior to the start of each pumping test, there is a minute level of drift always present in each of the pressure transducers. Therefore, we removed this drift by shifting the head value to a common one for all transducers in the sandbox prior to each test. The collected data were additionally processed by taking the average of the steady state head values collected over a period of time in a given monitoring port.

Illman et al. [2008] previously found that the order in which the pumping test data is included into the SSLE algorithm is critical to the final result. This is because each cross-hole pumping test data set contains different levels of noise. Our experiments in the laboratory were conducted as uniformly as possible, but there are noises that we could not control from one test to the next. Our findings include the importance of examining test data carefully and using the cleanest data [with the highest signal-to-noise ratio] first and progressively including noisier data [with lower signal-to-noise ratio] into the SSLE algorithm. This is because the SSLE algorithm in its current form [*Yeh and Liu*, 2000; *Zhu and Yeh*, 2005, 2006] is more sensitive to noise during the beginning stages of K tomogram computation. This sensitivity is due to the uniform convergence criteria used for sequentially inverting all pumping test data. Careful examination of head data showed that the data most devoid of noise were found near the bottom of the sandbox, while the noisiest data were usually located near the top of the sandbox. This is due to the fact that a higher water column sits on the pressure transducers near the bottom of the sandbox, which causes the pressure transducers to be less affected by noise. Also, the signal is magnified near the bottom no-flow boundary due to the superposition principle. Therefore, we included the steady-state head data from cross-hole tests with pumping taking place in the order of ports 47, 44, 35, 32, 17, 14, 5 and 2 into the SSLE algorithm.

12.5.3 Computation of *K* tomograms and its visual assessment

All computations for steady state hydraulic tomography analyses were executed using 8 of 16 processors on a PC-cluster consisting [of 1 master and 15 slaves each with Pentium IV 3.2

GHz with 1 GB of RAM] at the University of Arizona. The total computational time was about one minute indicating the efficiency of the SSLE algorithm. Figure 12.5a-d are the *K* tomograms obtained by inverting the steady state head data from 2, 4, 6, and 8 pumping test, respectively. Figure 12.5a shows that with only 2 pumping tests, a coarse picture of the heterogeneity pattern emerges and as more tests are included into the SSLE algorithm, more detail to the heterogeneity structure emerges. The final *K* tomogram obtained [Figure 12.5d] using 8 pumping tests revealed considerable detail to the heterogeneity structure.

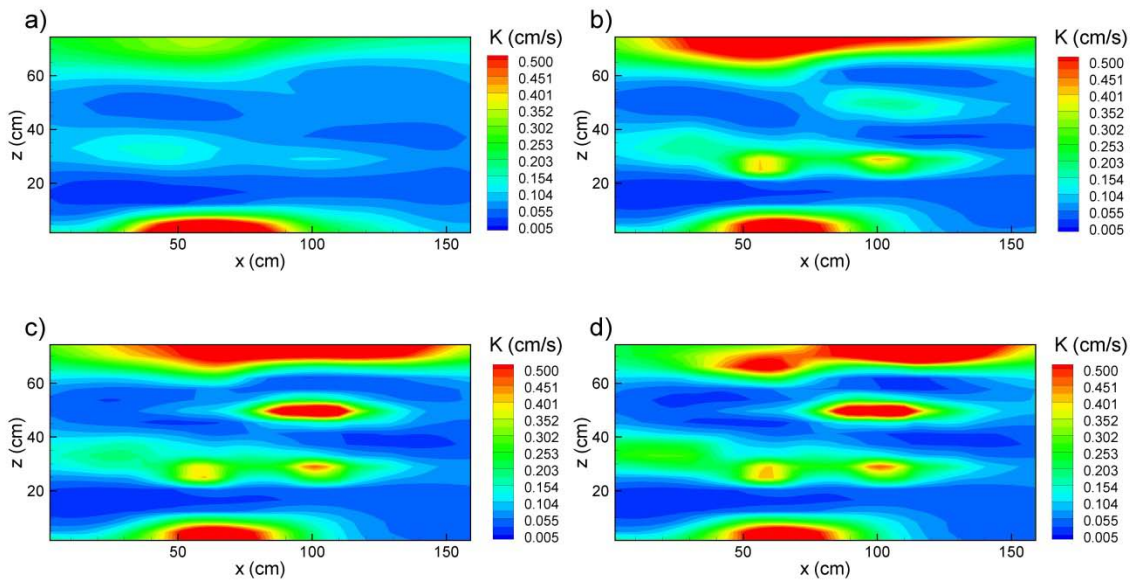


Figure 12.5: *K* tomograms computed using: a] 2 pumping tests [ports 47 and 44]; b] 4 pumping tests [ports 47, 44, 35, and 32]; c] 6 pumping tests [ports 47, 44, 35, 32, 17, and 14]; and d] 8 pumping tests [ports 47, 44, 35, 32, 17, 14, 5 and 2].

Table 12.5 summarizes the $K_G, \sigma^2_{\ln K}$, and correlation lengths of the resulting *K* tomograms. The estimated K_G of the *K* tomogram after including data from 8 tests was 1.0×10^{-1} cm/s while the estimated $\sigma^2_{\ln K}$ was 1.12. The K_G was in the midrange of the values from other characterization approaches, while $\sigma^2_{\ln K}$ was highest among all approaches [Table 12.3]. It is of interest to note that there is little change in the K_G and the correlation lengths of the *K* distribution as more tests were included. On the other hand, the $\sigma^2_{\ln K}$ increased as more tests was included into the SSLE algorithm. These results imply that with as few as 2 pumping tests, one could reliably estimate the K_G and the correlation lengths of the *K* distribution in the synthetic aquifer, however the accurate estimation of $\sigma^2_{\ln K}$ requires more cross-hole pumping tests.

Table 12.5: Statistical properties of the estimated *K* tomograms.

Case	K_G [cm/s]	$\sigma^2_{\ln K}$	λ_x [cm]	λ_z [cm]
Hydraulic tomography [Sequential – 2 tests]	0.090	0.542	16	7
Hydraulic tomography [Sequential – 4 tests]	0.101	0.890	18	9
Hydraulic tomography [Sequential – 6 tests]	0.099	1.036	14	6
Hydraulic tomography [Sequential – 8 tests]	0.100	1.121	18	7

To compare the *K* tomogram [Figure 12.5d] to the location of the sand deposits or lithofacies [Figure 12.3] we overlay Figure 12.5d over Figure 12.3 using Adobe Photoshop. Examination of Figure 12.6 shows a good correspondence between the high *K* zones of *K* tomogram to layer 1 [20/30], 7 [#12], 9 [20/40], 11 [#12], 15 [20/40], and 18 [20/30]. Despite the good correspondence of the general locations of these features, it is also apparent from Figure 12.6 that the boundaries of the lithofacies from the deposits to the boundaries of high *K* features from the tomogram are not in close agreement. We also note that layers 4 [20/40], 12 [16/30], and 13 [20/30], all expected to be high *K* features are not shown on the *K* tomogram.

We also compare the *K* tomogram [Figure 12.5d] to the kriged *K* estimates of small scale data [Figure 12.4a and 12.4b]. A visual comparison of the estimated *K* tomogram to the kriged *K* fields from the small scale estimates of *K* reveals a weak correspondence implying that steady state hydraulic tomography yields results that are significantly different from the standard geostatistical analysis of small scale *K* data.

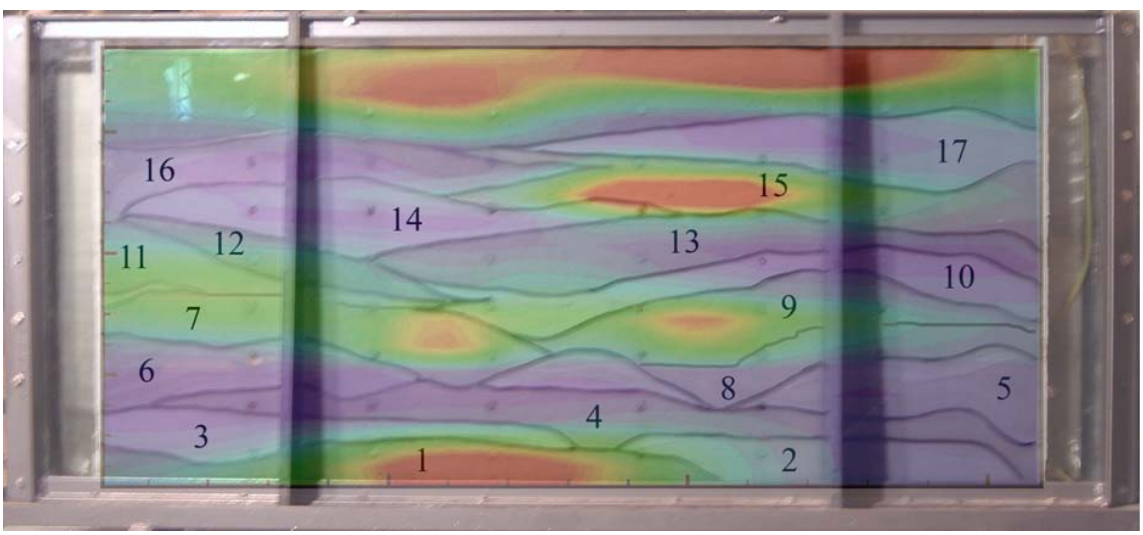


Figure 12.6: Overlay of *K* tomogram computed using 8 pumping tests over the picture of the heterogeneous sandbox aquifer using Adobe Photoshop.

12.6. FORWARD SIMULATIONS OF INDEPENDENTLY CONDUCTED CROSS-HOLE PUMPING AND FLOW-THROUGH TESTS

A number of approaches have been proposed in the literature to validate forward and inverse groundwater flow models. For example, *Foglia et al*. [2007] tested alternative forward groundwater models using the corrected Akaike information criterion, Bayesian information criterion, and generalized cross-validation. *Poeter and Anderson* [2005] utilized a Kullback-Leibler information criterion that selects parsimonious groundwater models that provided a more realistic measure of precision than evaluation of any one model or evaluation based on other

commonly referenced model selection criteria. In terms of the validation of inverse modeling results, *Painter et al*. [2007] compared three separate transmissivity estimation approaches and compared the resulting heads from forward simulation and actual head data from the Edwards aquifer in Texas. *Illman et al*. [2007] and *Liu et al*. [2007] used various data [core, slug, and single-hole test results] and statistical moments of these small scale data, as well as independent cross-hole pumping test data to validate their K and S_s tomograms. In both studies, *Illman et al*. [2007] and *Liu et al*. [2007] both concluded that the best approach to validate a K or S_s tomogram was to simulate an independently conducted test not used in the construction of the tomograms. We consider the cross-hole tests to be the most appropriate testing approach for validation purposes because the tests can be conducted at different locations within the aquifer with a large number of observation points as implemented in our sandbox study. Pumping at different locations will induce flow throughout the aquifer yielding drawdowns along the layers as well as across them. Other methods could potentially involve the discrete collection of aquifer samples, small scale hydraulic tests, and flux measurements all of which yield localized information that may not be entirely suitable for large scale validation of groundwater flow models.

Another method to test the performance of the various K fields is to simulate the flow-through experiments using the various homogeneous and heterogeneous K fields. If the estimated K field is accurate, it should, in principle, yield accurate estimates of total flow rate on the outflow end of the sandbox resulting from the flow-through tests.

Here, to quantitatively assess the validity of the various homogeneous and heterogeneous K estimates, we simulate 16 additional cross-hole pumping and 6 bidirectional flow-though tests using the forward groundwater model MMOC3 developed by *Yeh et al*. [1993]. For the validation using cross-hole tests, the simulated and measured drawdown values for each case are plotted on separate scatter plots and a linear model is fitted to each case without forcing the intercept to zero. The linear model fit and the coefficient of determination [R^2] provide indications of scatter and bias. The R^2 is a statistic that provides a quantitative measure of similarity between the simulated and measured drawdown values. A high R^2 value means that the simulated and measured drawdown values are linearly correlated, even though the mean values could be different. Other measures of correspondence between the simulated and observed drawdown values are the mean absolute error [L_1] and the mean square error [L_2] norms. The L_1 and L_2 norms are computed as:

$$L_1 = \frac{1}{n}\sum_{i=1}^{n}\left|\chi_i - \hat{\chi}_i\right| \qquad [12.3]$$

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}\left(\chi_i - \hat{\chi}_i\right)^2 \qquad [12.4]$$

where n is the total number of drawdown data, i indicates the data number, and χ_i and $\hat{\chi}_i$ represent the estimates from the simulated and measured drawdowns, respectively.

The slope and intercept of the linear model fit, the R^2 values, and the L_1 and L_2 norms are summarized in Table 12.6. These statistics collectively provide quantitative measures of the performance of each steady-state forward groundwater model with different K fields in simulating the 16 cross-hole pumping tests. The use of more than one cross-hole pumping test

conducted at different parts of the synthetic heterogeneous aquifer ensures a more credible validation of each of these models.

For the validation of the various K fields using the flow-through tests, we simply compare the total flow rates obtained through the forward simulation of the flow-through tests using MMOC3 to those obtained from the actual experimental data.

Table 12.6: Slope and intercept of fitted linear model, coefficient of determination [R^2] and standard errors [L_1 and L_2 statistics] of various homogeneous and heterogeneous cases.

Case	slope	intercept	R^2	L_1	L_2
Homogeneous K field					
K_G [core]	1.02	0.20	0.83	0.33	0.22
K_G [single-hole]	1.32	0.26	0.83	0.60	0.70
K_G [cross-hole – port 17]	1.03	0.21	0.83	0.34	0.23
K_G [cross-hole – port 21]	0.71	0.15	0.83	0.31	0.17
K_G [flow through]	0.58	0.13	0.83	0.39	0.28
Heterogeneous K field					
Kriged core K	1.10	0.14	0.89	0.30	0.18
Kriged single-hole K	1.45	0.13	0.89	0.58	0.67
Hydraulic tomography [Sequential – 2 tests]	1.07	-0.02	0.93	0.18	0.07
Hydraulic tomography [Sequential – 4 tests]	1.05	0.00	0.96	0.13	0.04
Hydraulic tomography [Sequential – 6 tests]	1.08	0.05	0.97	0.15	0.05
Hydraulic tomography [Sequential – 8 tests]	1.07	0.02	0.97	0.13	0.04

12.6.1 Use of independent cross-hole tests to validate homogeneous and heterogeneous *K* fields

Figure 12.7 shows a series of scatter plots that compare the simulated drawdown values from various K_G fields from Table 3 against the measured values from 16 separate cross-hole pumping tests. The figure includes a 45 degree line which indicates a perfect correlation. In particular, Figure 12.7a shows that the K_G from the permeameter analysis of cores yields reasonable predictions of 16 separate cross-hole tests over a large range of drawdown values. Our assessment is based on the near unit slope [1.02] of the linear model fit to the data. However, the intercept of the linear model is 0.20 suggests that the predictions are biased and a R^2 value of 0.83 reveals that there is some scatter. Figure 12.7b shows that the geometric mean of the K_G from single-hole tests yields biased predictions of 16 separate cross-hole tests. The simulated drawdowns are in general higher than the measured ones, which is a consequence of the lower K_G estimated using the single-hole tests conducted in this aquifer.

The K_G values from the 2 cross-hole tests yielded mixed results. For the K_G value estimated from the cross-hole pumping test with pumping taking place at port 17, the prediction of the 16 cross-hole pumping tests are reasonable [Figure 12.7c] and is in similar quality as the core case [Figure 12.7a]. However, the prediction based on the K_G from cross-hole pumping test with pumping taking place at port 21 is biased [Figure 12.7d]. For this case, the measured drawdowns are in general higher than the simulated ones, which is a consequence of the higher K_G estimated through cross-hole pumping tests. The flow-through test estimate of K_G, likewise yielded biased predictions [Figures 12.7e], similar to cross-hole pumping test case where pumping took place at port 21. This suggests that K_G estimated by flow-through tests is higher than the true mean K of this sandbox aquifer. The latter makes sense considering that flow during the flow-through tests takes place horizontally primarily through the high K layers. This should

result in a mean K value that should be closer to the arithmetic mean K, which is higher than the geometric mean.

Figure 12.8 shows a series of scatter plots which compare the simulated drawdown values from various heterogeneous K fields against the measured values from 16 separate cross-hole pumping tests. Figure 12.8a shows that the kriged K field based on permeameter tests on cores yields slightly improved predictions of 16 separate cross-hole pumping tests over a large range of drawdown values in comparison to the K_G case. This is supported by the slight improvement of the R^2 value as well as the decrease in the L_1 and L_2 norms [Table 12.6] when the kriged case is compared to the K_G case. Likewise, Figure 12.8b illustrates that the kriged K field from single-hole pumping test data shows little improvement from its homogeneous counterpart based on the linear model fit, the R^2 value as well as the lower L_1 and L_2 norms.

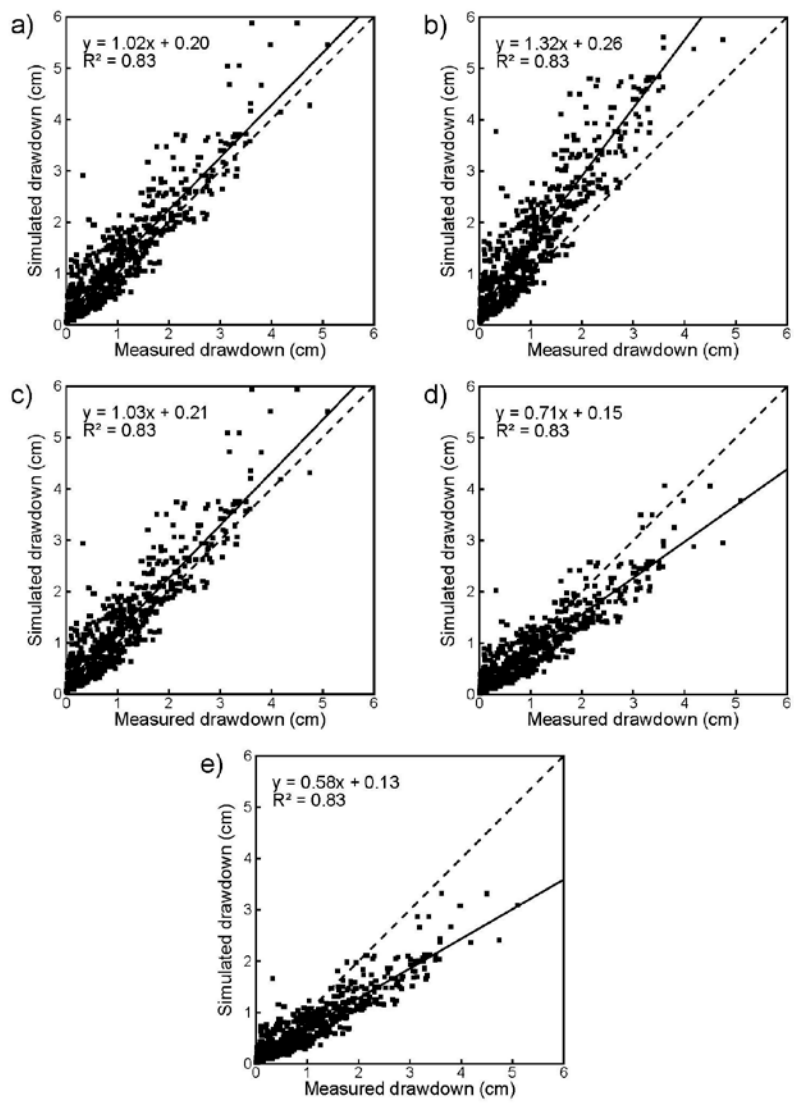


Figure 12.7: Arithmetic plots of drawdown predictions of 16 separate cross-hole pumping tests using the a]: core K_G estimates; b] single-hole K_G estimates; c] geometric mean of cross-hole test K_{eq} estimates [port 17]; d] geometric mean of cross-hole test K_{eq} estimates [port 21] and e] K_{eff} from flow-through tests.

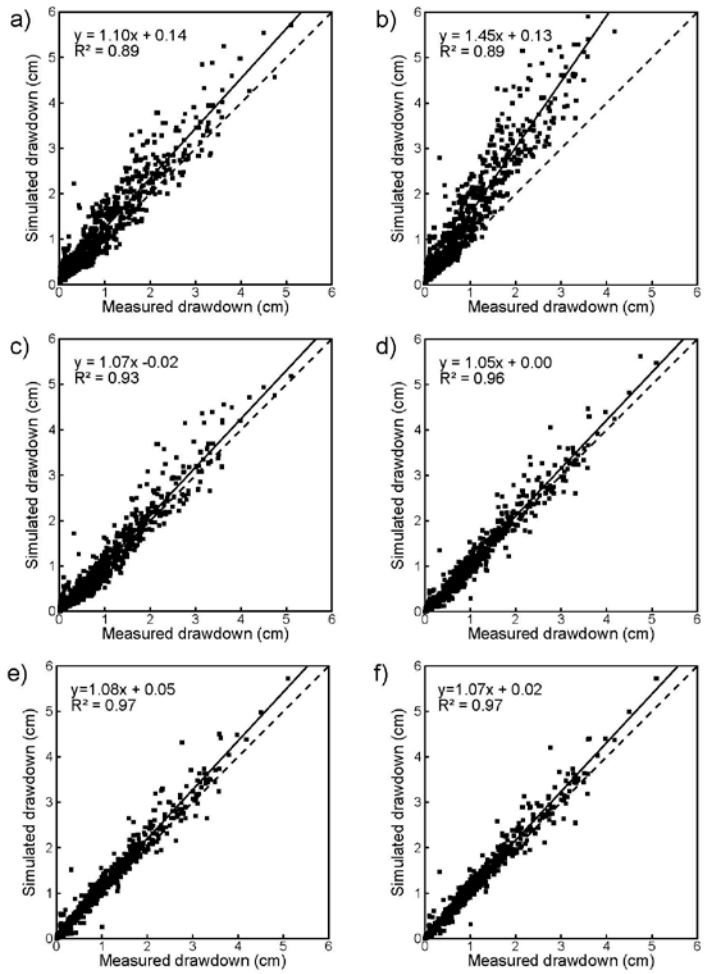


Figure 12.8: Arithmetic plots of drawdown predictions of 16 separate cross-hole pumping tests using: a]: kriged core *K* estimates; b] kriged single-hole *K* estimates; c] *K* tomogram from SSHT [2 tests - ports 47 and 44]; d] *K* tomogram from SSHT [4 tests - ports 47, 44, 35 and 32]; e] *K* tomogram from SSHT [6 tests - ports 47, 44, 35, 32, 17 and 14]; and f] *K* tomogram from SSHT [8 tests - ports 47, 44, 35, 32, 17, 14, 5 and 2].

In contrast, the use of the *K* tomograms from hydraulic tomography yields a drastically improved prediction of measured drawdowns using 16 independent cross-hole pumping tests not used in the construction of the tomogram. Figure 12.8c – 12.8f show these comparisons for *K* tomograms constructed using 2, 4, 6, and 8 cross-hole pumping tests, respectively. It is evident from these figures that the predictability of independently conducted cross-hole pumping tests improves with a larger number of tests included into the SSLE algorithm.

12.6.2 Use of bidirectional flow-through tests to validate various homogeneous and heterogeneous *K* fields

The various homogeneous and heterogeneous *K* fields were additionally validated by simulating the actual flow-through tests using MMOC3. This was done by prescribing hydraulic head values at the left and right boundaries of the sandbox aquifer and conducting forward

simulations. The flow rates from the outflow boundary were computed by first calculating flow for every element along the outflow column and second summing the flows from these elements to obtain the total flow rate. In the first step, the flow rate for each outflow element was calculated using Darcy's law. The computed total flow rate, is then compared against the measured total flow rate from the actual flow-through tests.

Table 12.7 summarizes the relative error of the simulated flow rate with respect to the measured flow rates from 6 separate, bidirectional flow-through tests. In general, we see that the homogeneous *K* fields from the core, single-hole and cross-hole tests underestimate the flow rates of the various flow-through tests resulting in large relative errors [i.e., -15.8 ~ -58.6%]. This implies that the total outflow through the synthetic heterogeneous aquifer cannot be accurately predicted using the various homogeneous *K* fields. However, this is not the case when we use the K_G determined through the flow-through experiments. This is an expected result because the simulation of the flow-through tests was done with K_G determined from the flow-through tests. We speculate that the slight difference in flow rates results from experimental errors and the differences in how total flow is measured for the simulation and the actual experiments.

Table 12.7: Relative error computed for the simulated flow rates with respect to the measured values.

Test	Relative error [%]							
	Homogeneous *K* fields					Heterogeneous *K* fields		
	FT	Core	SH	CH [17]	CH [21]	Kriged [core]	Kriged [SH]	*K* tomogram
LR1	3.2	-41.6	-55.2	-42.3	-15.8	-37.9	-52.6	11.1
LR2	1.6	-42.5	-55.9	-43.2	-17.1	-38.9	-53.4	9.3
LR3	0.1	-43.3	-56.6	-44.1	-18.3	-39.8	-54.1	7.7
RL4	-4.7	-46.0	-58.6	-46.7	-22.2	-42.7	-56.2	2.6
RL5	-2.0	-44.5	-57.5	-45.2	-20.0	-41.0	-55.0	5.5
RL6	-2.3	-44.7	-57.6	-45.4	-20.3	-41.2	-55.1	5.2

Note: LR1 stands for flow-through test number one with flow from left to right. FT = flow-through test; SH = single-hole tests, and CH [17] = cross-hole test with pumping at port 17.

The simulated flow rates from the kriged *K* fields from cores and single-hole tests also yielded inaccurate estimates of total flow as shown by the high relative error [-37.9 ~ -56.2%]. The relative error was slightly less for kriged core and single-hole *K* fields in comparison to their homogeneous counterpart, which suggests that accounting for the heterogeneous nature of the strata can lead to improved estimates of flow.

In contrast, the relatively error is significantly reduced when we use the *K* tomogram from steady-state hydraulic tomography to predict total outflow. In particular, the relative error ranges from 2.6 to 11.1 %, which is significantly less than the previous cases examined. The main reason for the slight inaccuracy is due to the uncertain nature of *K* values away from the observation ports. Figure 12.9 is a contour map of ln *K* variance for the case in which 8 cross-hole tests are included into the SSLE algorithm [Figure 12.8d]. This figure shows that the variance is highest outside the interior of the synthetic heterogeneous aquifer, where observation ports are sparse. Given the uncertain nature of the *K* estimates in these regions, it is not surprising to us that there is a slight discrepancy in the simulated and measured outflow rates. The prediction of total outflow rate will likely improve with additional pressure transducers placed in the high ln *K* variance areas. Based on these results, it leads us to believe that the SSLE

algorithm yields the most accurate *K* field among all the characterization approaches validated here.

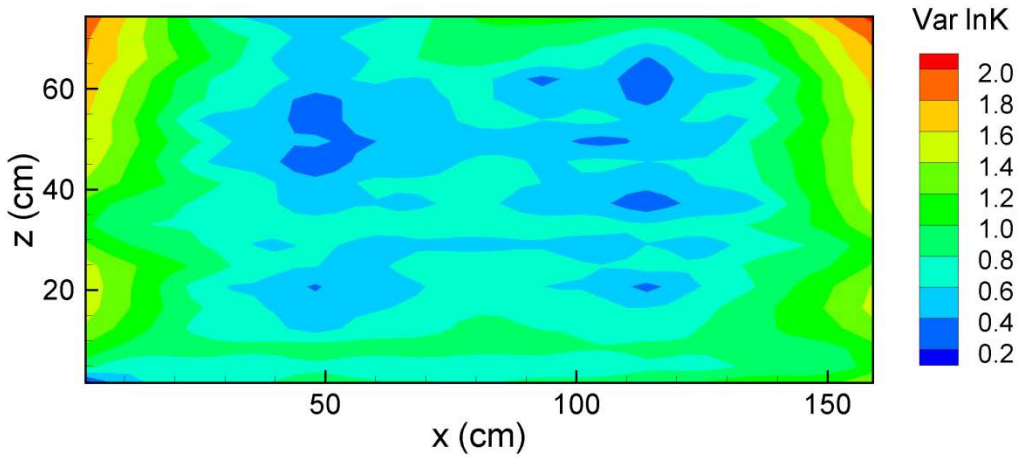


Figure 12.9: Contour plot of ln *K* variance with 8 cross-hole tests included into SSLE algorithm.

12.7 DISCUSSION

12.7.1 On model validation

The validation of groundwater models has become an important issue [e.g., *Hassanizadeh and Carrera*, 1992; *Konikow and Bredehoeft*, 1992; *Oreskes et al*., 1994; *Refsgaard and Knudsen*, 1996; *Refsgaard*, 1997; *Refsgaard and Henriksen*, 2004] over the last several decades. Model validation involves the establishment of greater confidence in the model by conducting simulations of the system under conditions in which the data has not been used for calibration purposes. For example, one can calibrate a groundwater flow model using one set of pumping test data. If the calibrated model from this first pumping test can predict system response accurately in a second pumping test [e.g., conducted using another well], one can have greater confidence in the calibrated model. On the other hand, if the parameters need to be adjusted to match the response of the 2nd pumping test, the process becomes a second calibration and additional datasets are needed to continue with the validation exercise.

We emphasize that successful model validation implied in this paper involves the establishment of greater confidence in the model by conducting simulations of the system under conditions in which the data has not been used for calibration purposes and obtaining excellent predictions of not only one, but multiple independent tests. It does not mean that the model, *per se*, is invincible and predicts the truth at all occasions. It is also important to recognize that "absolute" validation of a model is likely never possible. In this sense, we agree with *Refsgaard and Henriksen* [2004]'s definition of model validation: it is the "*substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model*."

12.7.2 On the accuracy of *K* estimates and its impact on model validation

The laboratory experimental results clearly showed that the K_G estimates from various established characterization methods yielded biased predictions of independently conducted cross-hole pumping tests. This, in turn resulted in erroneous predictions of total outflow through our synthetic heterogeneous aquifer. Therefore, steady state groundwater models constructed using the K_G estimates obtained via traditional characterization approaches are considered to be *invalid* based on the definition of model validation given in the previous section. In fact, our results illustrate that small differences in the K_G values among various characterization methods result in large differences in steady state drawdown and flow rate predictions of independent tests. Inaccurate predictions of drawdowns and flow rates will undoubtedly and negatively impact contaminant transport predictions.

A surprising finding from this study is that heterogeneous *K* fields from kriging did not cure this problem and yielded biased predictions. However, the *K* tomogram from steady-state hydraulic tomography was able to predict drawdowns of all independently conducted cross-hole hydraulic tests with considerable accuracy in comparison to the steady state groundwater model based on other *K* estimates. The *K* tomogram was also able to provide better predictions of flow rates in comparison to other approaches. These results lead us to conclude that the steady-state groundwater flow models based on the *K* tomogram can be considered to be validated sufficiently based on the definition of model validation given before.

12.7.3 Why does hydraulic tomography yield improved *K* fields over traditional methods?

One reason why hydraulic tomography yields improved *K* fields in comparison to other traditional methods is because the sequential or simultaneous inclusion of multiple pumping test data into the SSLE algorithm amounts to a repeated calibration and validation of the estimated *K* field with multiple datasets that are analyzed by the inverse algorithm. Another reason is that hydraulic tomography relies on multiple pumping tests conducted at different locations in the tested medium. The signals generated through multiple pumping events at different points in the aquifer are captured with pressure transducers placed in neighboring boreholes and synthesized using a suitable inverse approach. The large amount of data generated using hydraulic tomography surveys leads to improved estimates of *K* in the aquifer. Of course, the accuracy of the *K* tomogram computed will depend on the inverse modeling strategy used and the number of pumping tests included into the inverse algorithm and the pumped locations as well as the number of observation points available for the inverse analysis.

Potential reasons why established approaches may not perform as well as hydraulic tomography include: 1] experimental errors arising from packing of core samples and anisotropy in *K* resulting from the 1D nature of permeameter measurements; 2] small sampling volume; 3] near well effects that could affect single-hole or other small scale tests that involve injection or withdrawal of fluids; and 4] lack of adequate sampling between boreholes that could affect the estimation of connectivity of *K* heterogeneity. Furthermore, it appears that the homogeneous *K* value obtained by averaging of small scale *K* data or taking the geometric mean of the K_{eq} from cross-hole tests is simply not adequate enough to predict drawdowns of independent cross-hole pumping tests in a heterogeneous aquifer because the computed mean represents a volume smaller than the representative elementary volume [REV] [e.g., *Jankovic et al*., 2003]. However, as pointed out by *Neuman* [1987] and *Tartakvosky and Neuman* [1998], there is generally no guarantee that a REV can be defined for a given geologic medium. Even if such an REV can be defined, it will be so large that estimation of *K* using available experimental method becomes

impractical. In the real world, estimates of *K* or their distributions are often obtained at a scale smaller than the REV or what we refer to as the sub-REV scale. This is partially due to the lack of guidelines in estimating the size of a REV suitable for a given site leading to greater uncertainty in *K* estimation using traditional characterization approaches. This is another reason why we consider hydraulic tomography to be a robust approach in delineating the subsurface *K* heterogeneity as it does not rely on any REV assumptions for accurate estimation of the *K* field.

Our laboratory experiments have shown the powerful capability of steady state hydraulic tomography to obtain an accurate *K* tomogram. However, as robust as hydraulic tomography may be, we acknowledge that it is by no means a panacea technology. This is because the computed *K* tomogram is non-unique as there are an infinite number of solutions to the steady-state inverse problem for a heterogeneous *K* field, even when all of the forcing functions are fully specified. Only when data are available at all estimated locations will the inverse problem be well-posed and ultimately lead to a unique solution [e.g., *Yeh et al*., 1996; *Yeh and Liu*, 2000; and *Yeh and Simunek*, 2002]. This, obviously, is not the case here. For example, we noted earlier that some of the layers in the synthetic heterogeneous aquifer were not identified by steady state hydraulic tomography. We suspect that some of these features were not identified because of the limited number of pressure transducers installed along each column of wells [Figure 1]. A larger number of pressure transducers or conditioning the *K* tomogram with higher resolution measurements of *K* along the columns would likely improve the resolution of the *K* tomogram. In any case, it is important to note that the misidentification of some of these layers could have large impacts on predicting solute transport, thus further improvement in hydraulic tomography technology is desirable. This improvement will require technological advancements that will lower the cost of downhole pressure measurements. If monitoring costs can be substantially lowered, a denser pressure measurement network becomes more feasible increasing the ability of hydraulic tomography to capture finer scale details of geologic heterogeneity.

Despite some shortcomings, it is important to recognize that we have obtained a solution to the inverse problem [i.e., the *K* tomogram] that is consistent with the heterogeneity patterns that yield accurate predictions of steady state drawdowns from multiple, independent cross-hole pumping tests not used in the construction of the *K* tomogram. The robustness of our results is further supported by the accurate prediction of flow rates through the synthetic heterogeneous aquifer. Results presented here show that accurate steady-state groundwater models can be developed, and that the degree of its validation will in general improve when there is more accurate information on subsurface heterogeneity and forcing functions. Therefore, we contend that more emphasis should be placed on accurate site characterization to delineate subsurface heterogeneity patterns and their connectivity, developing improved methods for defining source/sinks [other than the pumping/injection rates which are typically known], and developing improved methods for increasing the accuracy in specifying boundary conditions. If this is done adequately, we predict that steady state groundwater model will become more accurate and its predictability of future scenarios will improve. Of course, this prediction is based on the caveat that no fundamental changes occur to the aquifer [e.g., deformation] after the calibration process has been completed.

13. CAPTURING HETEROGENEITY IN GROUNDWATER FLOW PARAMETERS: COMPARISON OF APPROACHES THROUGH CONTROLLED SANDBOX EXPERIMENTS

13.1. INTRODUCTION

13.1.1 Characterization methods of subsurface heterogeneity

Subsurface characterization for groundwater investigations relies on the determination of the distribution of hydraulic parameters such as hydraulic conductivity [K] and specific storage [S_s]. These values are then used to build groundwater models of various complexities to obtain quantitative estimates of hydraulic heads, groundwater fluxes, and the distribution and concentration of contaminants. Commonly, hydraulic parameters are estimated by collecting cores and subjecting them to permeameter tests and grain size analysis in a lab, or conducting slug, single-hole and/or pumping tests *in situ*. These *in situ* methods rely on analytical solutions that treat the geological medium to be homogeneous. These simplified solutions and the resulting estimated parameters have been utilized in a variety of real world applications and academic studies, despite the fact that the subsurface is heterogeneous at multiple scales. In particular, the knowledge of detailed three-dimensional distributions of K is critical for the prediction of contaminant transport, delineation of well catchment zones, and quantification of groundwater fluxes including surface-water/groundwater exchange. Even though many studies treat S_s to not vary significantly, in some formations, where the aquifer compressibilities vary significantly from one material type to the next [e.g., sands versus clays], S_s could vary several orders of magnitude. The characterization of subsurface heterogeneity is fraught with difficulties as numerous samples are required to delineate the variability of hydraulic parameters as well as their spatial correlations and connectivity. Using soil cores to accurately characterize the K heterogeneity of a site requires a large number of samples to be tested in the laboratory using a constant or a falling head permeameter [e.g., *Sudicky*, 1986; *Sudicky et al*., 2010]. Alternatively, these samples are sieved to obtain grain size distributions, which can then be analyzed using various empirical relations to estimate K.

Characterizing the heterogeneity in S_s is seldom done as the parameter is considered to be less variable than K [e.g., *Gelhar*, 1993 and others]. In fact, data in the literature suggests this to be the case in both porous and fractured geologic media [e.g., *Meier et al*., 1998; *Sanchez-Vila et al*., 1999; *Illman and Neuman*, 2001; *Vesselinov et al*. 2001; *Martinez-Landa and Carrera*, 2005; *Illman and Tartakovsky*, 2007; *Liu et al*., 2007; *Willmann et al*., 2007; *Illman et al*. 2009], although the estimated variance is known to be dependent on the estimation method. In particular, *Sanchez-Vila et al*. [1999] showed that when one uses the Jacob's method to infer transmissivity [T] and storage coefficient [S], one obtains some "apparent" T which is constant and an "apparent" S which is highly variable in an aquifer, but in reality, T is variable and S is constant. In a different study, *Wu et al*. [2005] using numerical experiments and a first-order correlation analysis showed that the effective T and S for an equivalent homogeneous aquifer of Gaussian random T and S fields vary with time as well as the principal directions of the effective T. These studies suggest that the traditional interpretation of pumping tests by treating the medium to be homogeneous could potentially lead to biased estimates of hydraulic parameters and new methods for capturing subsurface heterogeneity are therefore necessary.

13.1.2 Methods for capturing spatial heterogeneity

Common approaches when mapping K [and less so S_s] heterogeneity are to utilize geostatistical or stochastic estimation techniques or more sophisticated interpolation methods. In particular, these approaches are considered to be the *de facto* standards which assume that a user-specified covariance function is valid and hydrogeologic parameters are log-normal and stationary. These assumptions are difficult to satisfy in many geologic settings. Because of these assumptions, and when data are not abundant, stochastic estimation techniques may provide a smooth image of the spatial heterogeneity and may not represent the true distribution accurately. Although a variety of stochastic simulation techniques [e.g., *Deutsch and Journel*, 1998] exist that can overcome this issue of smoothing, it still does not address the preservation of many geological features. This is due to the fact that traditional geostatistical methods are based on variograms computed using two-point statistics. To overcome this shortcoming, multiple point geostatistics [e.g., *Guardiano and Srivastava*, 1993; *Caers*, 2001; *Strebelle*, 2002; *de Vries et al*., 2009] have been developed through the use of more complex point configurations, whose statistics are retrieved from training images that represent the geological facies distributions obtained from outcrop mappings and/or geophysical imaging.

Other approaches used to model subsurface heterogeneity include the Transition Probability Markov Chain method [*Carle and Fogg*, 1997; *Carle*, 1999; *Weissmann et al*., 1999] and the indicator kriging approach [*Journel*, 1983; *Journel and Isaaks*, 1984; *Journel and Alabert*, 1990; and *Journel and Gomez-Hernandez*, 1993]. Both of these approaches allow one to construct discontinuous facies models. However, the Markov model is better able to account for spatial cross-correlation, such as juxtapositional relationships, including the fining-upward tendencies of different facies, than the indicator approach [*De Marsily et al*., 2005].

More recently, geostatistical and stochastic inversion methods have received increasing attention. The approach produces the first and second statistical moments of hydrogeologic variables, representing their most likely estimates and their uncertainty, respectively, conditioned on available observations. Cokriging relies on the classical linear predictor theory that considers spatial correlation structures of flow processes [such as hydraulic head and velocity] and the subsurface hydraulic property, and cross-correlation between the flow processes and the hydraulic property. In the past few decades, many researchers [e.g., *Kitanidis and Vomvoris*, 1983; *Hoeksema and Kitanidi*s, 1984 and 1989; *Rubin and Dagan*, 1987; *Gutjahr and Wilson*, 1989; *Sun and Yeh*, 1992; *Harvey and Gorelick*, 1995; *Yeh et al.*, 1995 and 1996] have demonstrated its ability to estimate K, head, velocity, as well as solute concentrations in heterogeneous aquifers.

Recently, hydraulic tomography has been developed to obtain information on subsurface heterogeneity of K and S_s through sequential pumping tests. To our knowledge, *Neuman* [1987] was the first to suggest the approach using geophysical tomography as an analogy. Since then, various inverse methods have been developed for hydraulic tomography, which utilize pumping test data simultaneously or sequentially [e.g., *Gottlieb and Dietrich*, 1995; *Yeh and Liu*, 2000; *Bohling et al.*, 2002; *Brauchler et al.*, 2003; *McDermott et al.*, 2003; *Zhu and Yeh*, 2005, 2006; *Li et al*., 2005; *Fienen et al*., 2008; *Castagna and Bellin*, 2009; *Xiang et al*., 2009; *Liu and Kitanidis*, 2011]. Numerous laboratory [e.g., *Liu et al*., 2002; *Illman et al*., 2007, 2008; *Liu et al*., 2007; *Yin and Illman*, 2009; *Illman et al*., 2010] and field experiments [e.g., *Bohling et al*., 2007; *Straface et al*., 2007; *Illman et al*., 2009; *Cardiff et al*., 2009] have been conducted to show the utility of hydraulic tomography, but a rigorous study which compares the results to other more traditional characterization methods is generally lacking.

In the laboratory, *Illman et al*. [2010] recently assessed the performance of various methods for characterizing K estimates by predicting the hydraulic response observed in cross-hole pumping tests in a synthetic heterogeneous aquifer, and total flow rates obtained via flow-through tests. Specifically, they characterized a synthetic heterogeneous sandbox aquifer using various techniques [permeameter analyses of core samples, single-hole, cross-hole, and flow-through testing]. They then obtained mean K estimates through traditional analysis of test data by treating the medium to be homogeneous. Heterogeneous K fields were obtained through kriging and steady state hydraulic tomography. To assess the performance of the each characterization approach, *Illman et al*. [2010] conducted forward simulations of 16 independent pumping tests and 6 steady-state, flow-through tests using these homogeneous and heterogeneous K fields. The results of these simulations were then compared to the observed data. Results showed that the mean K and heterogeneous K fields estimated through kriging of small scale K data [core and single-hole tests] produced biased predictions of drawdowns and flow rates under steady-state conditions. In contrast, the heterogeneous K distribution or "K tomogram" estimated via steady state hydraulic tomography yielded excellent predictions of drawdowns of pumping tests not used in the construction of the tomogram and very good estimates of total flow rates from the flow-through tests. Based on these results, *Illman et al*. [2010] suggested that steady-state groundwater model validation is possible, if the heterogeneous K distribution and forcing functions [boundary conditions and source/sink terms] are characterized sufficiently.

13.1.3 Goal of this study

This study extends the work of *Illman et al*. [2010] who examined only various K characterization approaches and their performance in predicting independent test data under steady state conditions. In particular, the main goal of this study is to extend the work of *Illman et al*. [2010] to the transient case. Using the same sandbox aquifer as *Illman et al*. [2010], we jointly assess the performance of various characterization and modeling techniques that treat the aquifer to be either homogeneous or heterogeneous through the prediction of independent, transient cross-hole pumping tests not used in the characterization effort. Specifically, we characterize the 2D heterogeneous aquifer using both single-hole and cross-hole pumping tests. These data are then used to construct various forward groundwater models with homogeneous and heterogeneous K and S_s estimates. Two homogeneous or effective parameter models are constructed: 1] by averaging local scale K and S_s estimates from single-hole pumping tests and treating the medium to be homogeneous; and 2] using MMOC3 [*Yeh et al*., 1993] coupled with PEST [*Doherty*, 1994] to estimate K and S_s by simultaneously matching the transient drawdown data from all ports during a cross-hole pumping test. Three heterogeneous models are constructed and consist of spatially variable K and S_s fields obtained via: 1] kriging single-hole K and S_s data; 2] accurately capturing the layering and calibrating the K and S_s values for these layers using a parameter estimation program [i.e., a calibrated geological model]; and 3] conducting transient hydraulic tomography. The performance of these homogeneous and heterogeneous K and S_s fields are then quantitatively assessed by simulating 16 independent cross-hole pumping tests and comparing the simulated drawdowns to the observed drawdowns.

13.2 EXPERIMENTAL METHODS

13.2.1 Sandbox and synthetic heterogeneous aquifer description

A two-dimensional synthetic heterogeneous aquifer was constructed in a sandbox measuring 193.0 cm in length, 82.6 cm in height, and 10.2 cm deep. Forty eight ports, 1.3 cm in diameter, were cut out of the stainless steel wall to allow coring of the aquifer as well as installation of fully penetrating horizontal wells [Figure 13.1]. Each port is instrumented with a 0 to 1 psig *Setra model 209* pressure transducer.

The synthetic heterogeneous aquifer was created through the cyclic deposition of sediments under varying water flow and sediment feed rates. Our goal in relying on sediment transport was to create a more realistic heterogeneity pattern with various scales of heterogeneity in an efficient manner. Table 13.1 summarizes the grain size characteristics, K estimates, and S_s estimates of the sands used to create the heterogeneous aquifer, and the layers in which these sands occur. Figure 13.2 is a photograph of the synthetic heterogeneous aquifer showing the interfingering nature of the deposits and layer numbers. Further details to this synthetic heterogeneous aquifer and its construction approach are provided in *Illman et al*. [2010].

The aquifer system was bounded by three connected constant head boundaries [one situated at the top of the tank, and one and each end]. The remaining boundaries [front, back, and bottom] were all no flow boundaries.

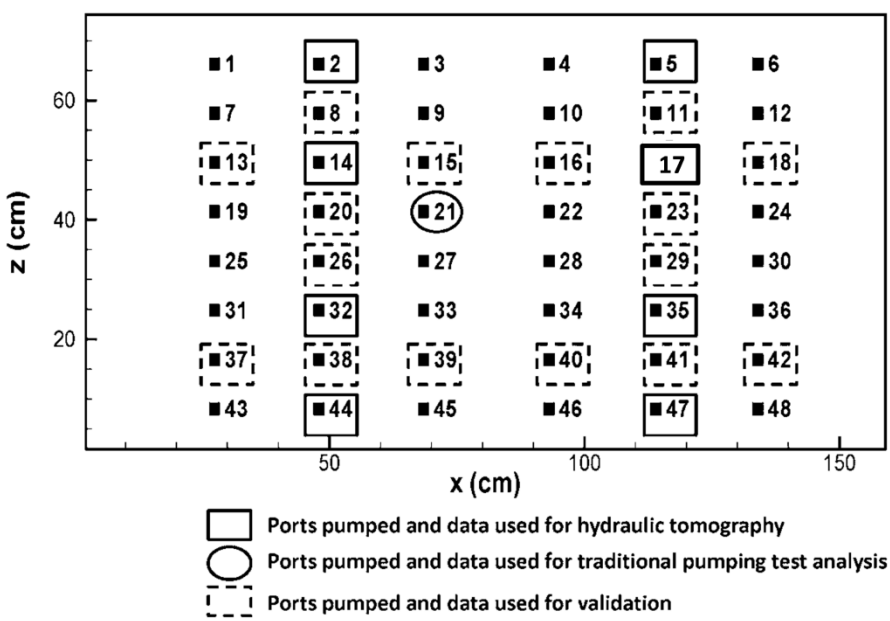


Figure 13.1: Schematic diagram of synthetic heterogeneous aquifer used for validation of transient groundwater flow models. Numbers next to solid squares indicate port numbers, open squares around numbers indicate the 8 ports [2, 5, 14, 17, 32, 35, 44, 47] used for hydraulic tomography, the open oval indicates the port [21] at which pumping test data was analyzed using VSAFT2 by treating the medium to be homogeneous [i.e., traditional pumping test analysis], and the dashed open squares around the 16 other ports [8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42] indicate the pumping locations for the independent cross-hole pumping tests used for validation purposes [after *Illman et al*., 2010].

Table 13.1: Characteristics of each layer used to create a synthetic heterogeneous aquifer [after *Illman et al*., 2010].

Layer	Sand	d_{50} [mm]	K [cm/s] Shepherd	Core Permeameter K [cm/s]*	Single-hole K [cm/s]*	Single-hole S_s [cm/s]*
1	20/30	0.75	1.03×10^{-1}	3.20×10^{-2}	5.32×10^{-2}	2.12×10^{-4}
2	4030	0.35	2.99×10^{-2}	5.29×10^{-2}	5.67×10^{-2}	2.60×10^{-4}
3	F-85	0.15	7.29×10^{-3}	7.14×10^{-2}	5.70×10^{-2}	5.00×10^{-4}
4	20/40	0.58	6.68×10^{-2}	5.68×10^{-2}	5.10×10^{-2}	2.22×10^{-4}
5	mix	0.46	N/A	N/A	N/A	N/A
6	mix	0.46	N/A	8.16×10^{-2}	5.00×10^{-2}	4.00×10^{-4}
7	#12	0.52	5.70×10^{-2}	1.27×10^{-1}	7.35×10^{-2}	4.20×10^{-4}
8	F32	0.5	5.33×10^{-2}	1.34×10^{-1}	4.50×10^{-2}	1.75×10^{-4}
9	20/40	0.58	6.68×10^{-2}	8.69×10^{-2}	4.60×10^{-2}	2.15×10^{-4}
10	F-65	0.2	1.20×10^{-2}	1.13×10^{-1}	8.25×10^{-2}	1.14×10^{-3}
11	#12	0.52	5.70×10^{-2}	1.37×10^{-1}	2.05×10^{-1}	2.15×10^{-4}
12	16/30	0.87	1.32×10^{-1}	3.40×10^{-2}	4.95×10^{-2}	6.32×10^{-4}
13	20/30	0.75	1.03×10^{-1}	2.60×10^{-1}	1.05×10^{-1}	9.80×10^{-4}
14	F-75	0.17	9.22×10^{-3}	9.79×10^{-2}	5.70×10^{-2}	9.80×10^{-4}
15	20/40	0.58	6.68×10^{-2}	8.58×10^{-2}	7.50×10^{-2}	2.00×10^{-3}
16	mix	0.46	N/A	4.16×10^{-2}	2.69×10^{-2}	7.11×10^{-4}
17	F-85	0.15	7.29×10^{-3}	4.51×10^{-2}	4.47×10^{-2}	1.14×10^{-3}
18	20/30	0.75	1.03×10^{-1}	1.45×10^{-1}	1.16×10^{-1}	3.38×10^{-3}

*** If multiple ports are in the same layer then the geometric mean is presented**

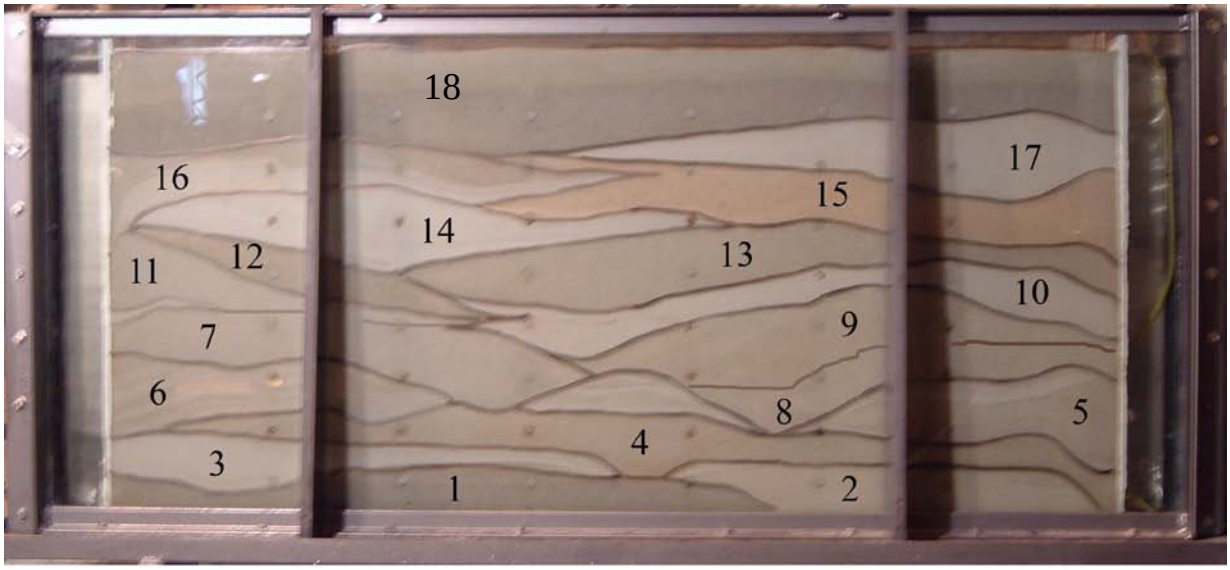


Figure 13.2: Photograph of synthetic heterogeneous aquifer created via cyclic flux of sediment-laden water [after *Illman et al*., 2010].

13.2.2 Characterization of synthetic heterogeneous aquifer [single-hole pumping tests]

The synthetic heterogeneous aquifer was characterized using single-hole pumping tests to obtain K and S_s estimates at each of the 48 ports. Since the support scale of parameters estimated via single-hole tests is unknown, we assume the length of the well screen open to the aquifer is representative [e.g., *Guzman et al*., 1996, *Illman and Neuman*, 2000; *Illman*, 2005] of the support scale. The tests were conducted by pumping water at each port at a constant rate and monitoring the transient head change within the pumped well using a pressure transducer. A constant pumping rate [Q = 1.25 cm^3/sec] was set for each single-hole pumping test. For each test, data collection started without the pump running in order to obtain the initial hydraulic head in the sandbox at all measurement ports. A peristaltic pump was then activated at the pumping port and allowed to run at a constant rate until the development of steady state flow conditions. The entire transient head data response was matched using VSAFT2 [*Yeh et al*., 1993] through manual calibration by treating the aquifer to be homogeneous. VSAFT2 was chosen for the analysis as opposed to traditional type curve models because the numerical model is able to more accurately describe the sandbox geometry and boundary conditions. Details to the numerical modeling and calibration effort are provided in *Craig* [2005]. The single-hole K estimates ranged from 0.01 cm/s to 0.32 cm/s with a geometric mean of 0.06 cm/s and a variance [$\sigma^2_{\ln K}$] of 0.38, and single-hole S_s estimates ranged from 1.0 x 10^{-4} cm^{-1} to 5.5 x 10^{-3} cm^{-1} with a geometric mean of 6.1 x 10^{-4} cm^{-1} and a variance [$\sigma^2_{\ln S_s}$] of 0.97. Here, we consider the geometric mean of the 48 local K and S_s estimates from the single-hole pumping tests to represent an effective K and S_s for the entire aquifer. We also considered the use of arithmetic and harmonic means as alternatives to the geometric mean. However, the arithmetic mean is representative of flow along stratification and the harmonic mean is representative of flow across layers. Since these single-hole tests induce flow along and across the layers we feel the geometric mean provides the most reasonable estimate of effective parameters for K and S_s.

13.2.3 Characterization of the synthetic heterogeneous aquifer [cross-hole pumping tests]

Twenty five cross-hole tests were also performed in the sandbox for the purposes of effective parameter characterization, calibration of a geological model [presented in section 13.4], transient hydraulic tomography [presented in section 13.5] and model validation [presented in section 13.6]. The tests were conducted at each port along columns 2 [ports 2, 8, 14, 20, 26, 32, 38, and 44] and 5 [ports 5, 11, 17, 23, 29, 35, 41, and 47] and 9 additional pumping tests at various ports outside of these two columns [ports 13, 15, 16, 18, 21, 37, 39, 40, and 42] [see Figure 13.1]. The cross-hole tests were conducted by pumping at rates ranging from 2.50 - 3.17 cm^3/sec at 25 separate ports indicated by open and dashed squares on Figure 13.1. During each test, head measurements in all 48 ports [and the constant head reservoirs] were recorded. Pumping continued until the development of steady state conditions, which was determined by observing the stabilization of all head measurements within the aquifer.

The cross-hole test at Port 21 was used to estimate effective homogeneous values of K and S_s. Analogous to the analysis of single-hole data, the observation head data from this test were analyzed by manually calibrating each observation port data by VSAFT 2 and treating the aquifer to be homogeneous. Estimates of K and S_s obtained between the pumping and observation intervals, when the medium is treated to be homogeneous, are considered to be an equivalent hydraulic conductivity [K_{eq}] or specific storage [S_{seq}] [*Renard and de Marsily*, 1997; *Neuman*, 2005]. Analysis of the cross-hole test yielded 48 estimates of K and S_s for the equivalent homogeneous medium. The K_{eq} estimates ranged from 0.054 cm/s to 0.42 cm/s with a geometric mean of 0.11 cm/s and a variance [$\sigma^2_{\ln K}$] of 0.22. The corresponding S_{seq} estimates ranged from 8.5 x 10^{-5} cm^{-1} to 7.5 x 10^{-3} cm^{-1} with a geometric mean of 3.3 x 10^{-4} cm^{-1} and a variance [$\sigma^2_{\ln S_s}$] of 1.03.

The cross-hole pumping test at Port 20 is used to estimate effective homogeneous values of K and S_s. The parameter estimation program PEST [*Doherty*, 1994] was coupled with the groundwater flow model MMOC3 [*Yeh et al.*, 1993] to simultaneously match the transient data recorded at all ports. The synthetic aquifer used for the parameter estimation was discretized into 741 elements and 1600 nodes with element dimensions of 4.1 cm × 10.2 cm × 4.1 cm. Both sides and top boundaries were set to the same constant head, while the bottom, front, and back boundaries of the sandbox were considered no-flow boundaries. Estimated K and S_s values are 6.76 x 10^{-3} cm/s and 6.80 10^{-4} cm^{-1} respectively.

13.3 GEOSTATISTICAL ANALYSIS OF SINGLE-HOLE PUMPING TEST DATA

Kriging and other simplified interpolation methods are commonly used to characterize subsurface heterogeneity [e.g., *Sudicky*, 1986; *Adams and Gelhar*, 1992; *Chen et al.*, 2000; and *Sudicky et al.*, 2010]. As such, we used kriging of single-hole K and S_s data to generate heterogeneous distributions that could be used for forward modeling of cross-hole pumping tests. The exponential variogram model was fit to the experimental variograms in both horizontal and vertical directions, resulting in an anisotropic variogram model. Examination of the experimental variogram of the ln-S_s data revealed an increasing variogram with lag distance. We attributed this to a trend of ln-S_s values declining from high to low values from the top to the bottom of the sandbox. Such a trend was also observed in the S_s values from a separate sandbox packed in an entirely different way [*Liu et al*., 2007]. We detrended the data [e.g., *Chen et al*., 2000] by fitting an anisotropic exponential model to the residuals of the original experimental variograms for ln-S_s data.

Table 13.2 lists the variogram parameters fit to the experimental variograms. For both model fits, the anisotropy ratio was determined to be 3 with the horizontal correlation length larger than the vertical correlation length due the layered nature of the deposits. The results [not shown here], in general, reveal smoother K and S_s fields in comparison to the interfingering layers shown in Figure 13.2, which is expected considering that there are only 48 data points used for kriging.

Table 13.2: Geostatistical model parameters for kriging single-hole ln-K and ln-S_s data.

Data	Model	Nugget	Range [cm]	Sill	Anisotropy ratio
Single-hole ln-K	Exponential	0	30	0.40	3
Single-hole ln-S_s	Exponential	0	30	0.45	3

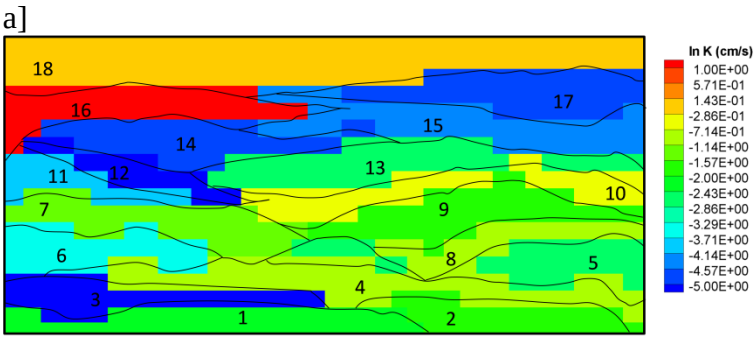
13.4. CONSTRUCTION OF A GEOLOGICAL MODEL

Groundwater flow and transport models are commonly built using various hydrogeologic data and often deterministically incorporate the knowledge of site geology. To compare the performance of groundwater flow models based on the knowledge of geology to other models, we constructed a numerical model by using Figure 13.2 as a reference to construct a parameter field that closely resembles the stratigraphy of the synthetic heterogeneous aquifer.

To construct the geological model, we assumed that the stratification is known for the entire simulation domain. In practice, a perfect knowledge of stratigraphy is not available, thus we consider this to be a best case scenario in terms of having information on stratigraphy. We further assumed that the stratification shown on the glass [Figure 13.2] was uniform throughout the thickness of the sandbox.

The parameter estimation program PEST [*Doherty*, 1994] coupled with the groundwater flow model MMOC3 [*Yeh et al.*, 1993] was used to estimate K and S_s values for each layer using the data collected during the cross-hole test at Port 20. In total 38 parameters were estimated [K and S_s for 19 layers]. Only 18 layers are identified in Figure 13.2, however, layer 5 which is discontinuous because of erosion that occurred during the deposition of layer 8 is treated as two separate layers in the PEST estimation. The model domain used for this estimation is identical to that described in section 2.3. Figure 13.3 shows the ln K and ln S_s [Figures 13.3a and 13.3b respectively] distribution for the calibrated geological model.

a]



b]

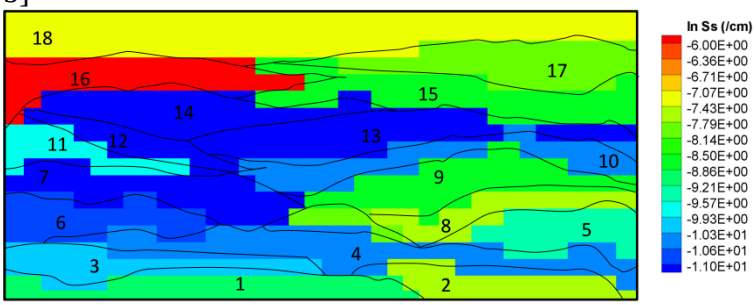


Figure 13.3: Distribution of a] ln-K and b] ln-S_s values for the calibrated geological model.

13.5 CROSS-HOLE PUMPING TESTS AND TRANSIENT HYDRAULIC TOMOGRAPHY ANALYSIS

The transient hydraulic tomography analysis of cross-hole pumping tests in the sandbox was conducted using the Sequential Successive Linear Estimator [SSLE] code developed by *Zhu and Yeh* [2005]. The inverse model assumes a transient flow field and the natural logarithm of K [*ln-K*] and S_s [*ln-*S_s] are both treated as multi-Gaussian, second-order stationary, stochastic processes. The model additionally assumes that the mean and correlation structure of the K and S_s fields are known *a priori*. Further details to the SSLE code can be found in *Zhu and Yeh* [2005].

13.5.1 Input parameters and cross-hole tests used

The model domain used to obtain K and S_s tomograms with transient hydraulic tomography is identical to that described in section 13.2.3 for the calibration of the geological mode.

Inputs to the inverse model include initial guesses for the K and S_s, estimates of variances and the correlation scales for both parameters, volumetric discharge [Q_n] from each pumping test where n is the test number, available point [small-scale] measurements of K and S_s, as well as head data at various times selected from the head-time curve. Although available point [small-scale] measurements of K and S_s can be input to the inverse model, we do not use these measurements to condition the estimated parameter fields.

For this THT analysis, the initial parameter fields are homogeneous and represented by the mean value of the K and S_s obtained from the analysis described in section 13.2.3 of a cross-hole test at port 21 which treats the medium to be homogeneous. Estimates of variance for K and S_s which have been shown to have negligible effects on the resulting tomogram [*Yeh and Liu*, 2000] were based on estimates from the available small scale data and used as our input variance in the inverse model. Correlation scales represent the average size of heterogeneity, which is difficult to determine accurately without a large number of data sets in the field. The effects of uncertainty in correlation scales on the estimate based on the tomography are negligible because the tomography produces a large number of head measurements, reflecting the detailed site-specific heterogeneity [*Yeh and Liu*, 2000]. Therefore, the correlation scales were approximated based only on the average thickness and length of the discontinuous sand bodies.

Prior to the incorporation into the inverse model the transient head records were treated with various error reduction schemes discussed in *Illman et al*. [2007], while data from pumped ports were not included into the inverse model because of excessive noise resulting from the use of a peristaltic pump. Each drawdown curve was then fit with a 5th or 6th–order polynomial curve following *Liu et al*. [2007]. A 5th or 6th-order polynomial was found to best capture the

overall drawdown behaviour for the majority of the data. We then manually extracted 5 points representative of the entire transient record. We found that selection of 5 points from the drawdown curve generally captured the salient features of the overall drawdown behavior. We can, of course, increase the number of drawdown data, however, this will increase computational effort for the inverse modeling.

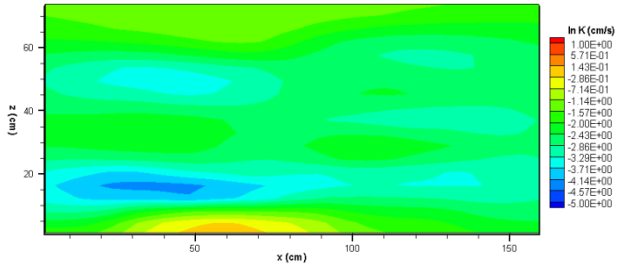
Data curves that could not be properly fit due to excessive noise were manually excluded from the analysis. In total, we utilized 8 independent cross-hole tests with pumping taking place at ports 47, 44, 35, 32, 17, 14, 5 and 2 for the analysis. More specifically, we utilized 5 data points from 47 ports totalling 235 data in all tests, except for the pumping test in port 2. At port 2, 5 data points were obtained from 43 ports totalling 215 data. Some of the data points were excluded from this particular test as the data were excessively noisy. In total, we utilized 1860 data points from 8 different tests in our transient inversions.

13.5.2 Computation of ln-K and ln-S_s tomograms

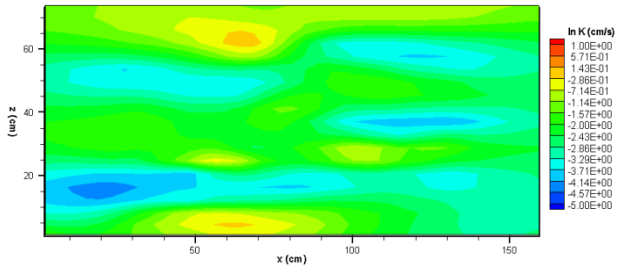
All computations for transient hydraulic tomography analyses were executed using 44 of 48 processors on a PC-cluster consisting [of 1 master and 12 slaves each with Intel Q6600 Quad Core CPU running at 2.4 GHz with 16 GB of RAM per slave] at the University of Waterloo. The operating system managing the cluster was CentOS 5.3 based on a 64-bit system. The total computational time for inverting data from 8 pumping tests was about 14 minutes. Figure 13.4a-d are the ln-K tomograms obtained by inverting the transient head data from 2, 4, 6, and 8 pumping test, respectively. Figure 13.4a shows that with only 2 pumping tests, a coarse picture of the heterogeneity pattern emerges, although the distribution is still pretty smooth and many details of the heterogeneity and in particular, details to the stratification are missing. As more tests are included into the SSLE algorithm, we see that more detail of the heterogeneity structure emerges. In particular, the final ln-K tomogram obtained [Figure 13.4d] using 8 pumping tests reveals considerable detail to the heterogeneity structure including the connectivity of various high and low K layers. Figure 13.4e is a ln-K tomogram computed using the steady state hydraulic tomography algorithm of *Yeh and Liu* [2000] by *Illman et al*. [2010] and is included for comparison purposes.

Figures 13.5a through 13.5d show the corresponding ln-S_s tomograms that were estimated simultaneously. In contrast to Figures 13.4a–4d, the layering structure visible in the ln-K tomogram is not visible for the ln-S_s tomogram. However, a decreasing trend in ln-S_s with depth in the synthetic aquifer is apparent. Physically speaking, this makes sense because the sands in the upper portion are less compressed, while the deeper sands are more compressed due to the stress exerted by the overlying material. This finding suggests that ln-K values are not significantly correlated with the ln-S_s values in this sandbox and is in agreement with those found by *Liu et al*. [2007] for a different sandbox packed with a considerably different heterogeneity pattern. The only data available for comparison are the single-hole estimates presented in Table 13.1 and shown as a kriged distribution. While not identical, the pattern is similar, and the values are in a similar range, suggesting the S_s-tomogram is physically based.

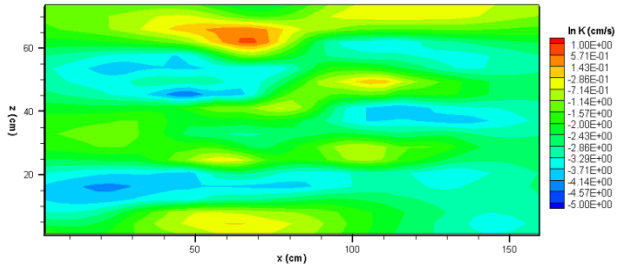
a]



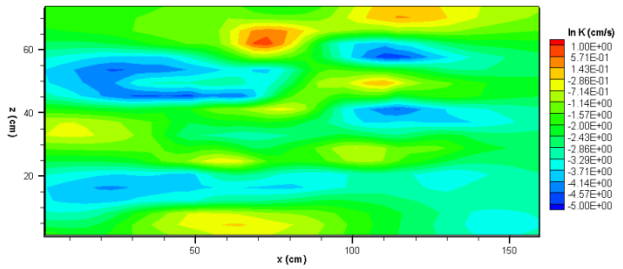
b]



c]



d]



e]

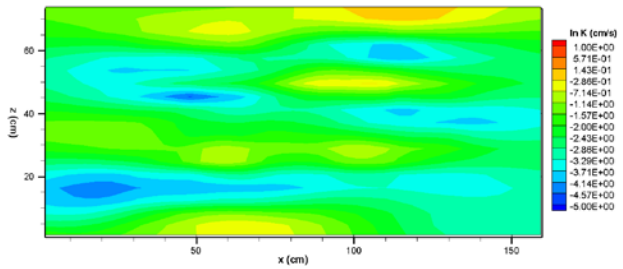
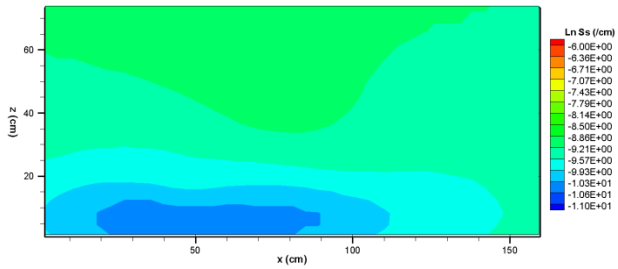


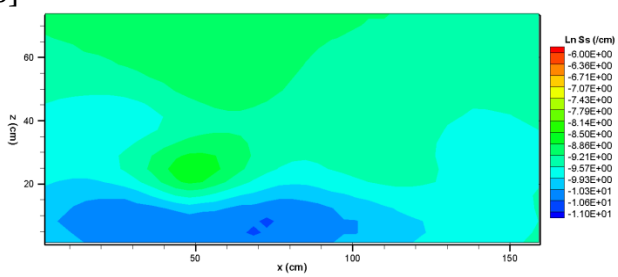
Figure 13.4: ln-*K* tomograms computed using: a] 2 pumping tests [ports 47 and 44]; b] 4 pumping tests [ports 47, 44, 35, and 32]; c] 6 pumping tests [ports 47, 44, 35, 32, 17, and 14]; d]

8 pumping tests [ports 47, 44, 35, 32, 17, 14, 5 and 2] and e] 8 pumping tests [ports 47, 44, 35, 32, 17, 14, 5 and 2] but interpreted with the steady state hydraulic tomography algorithm [after *Illman et al*., 2010].

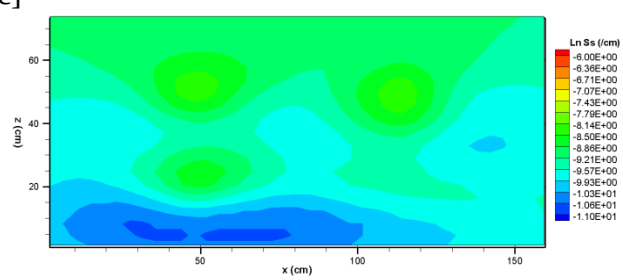
a]



b]



c]



d]

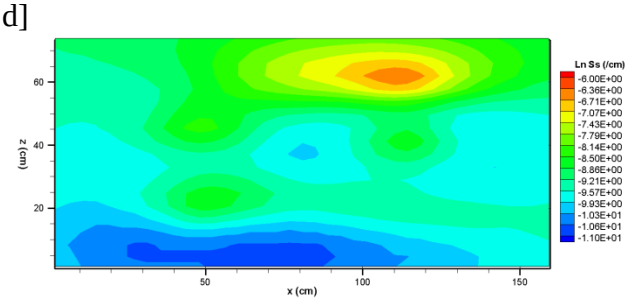


Figure 13.5: ln-S_s tomograms computed using: a] 2 pumping tests [ports 47 and 44]; b] 4 pumping tests [ports 47, 44, 35, and 32]; c] 6 pumping tests [ports 47, 44, 35, 32, 17, and 14]; and d] 8 pumping tests [ports 47, 44, 35, 32, 17, 14, 5 and 2].

13.5.3 Statistical summary of results

Table 13.3 summarizes the mean, variance, and correlation lengths of the resulting ln-K tomogram. The estimated geometric mean [K_G] of the ln-K tomogram after including data from 8 cross-hole tests was 1.0×10^{-1} cm/s, while the estimated variance [$\sigma^2_{\ln K}$] was 1.32. We note that the value of K_G is identical to that estimated using steady state hydraulic tomography by *Illman et al*. [2010], but the $\sigma^2_{\ln K}$ is slightly higher compared to a value of 1.12. The estimated K_G from

transient hydraulic tomography is somewhat higher than the estimate of K_G obtained by taking the geometric mean [0.06 cm/s] of the 48 local K values from single-hole tests. In contrast, the estimate of $\sigma^2_{\ln K}$ from transient hydraulic tomography [$\sigma^2_{\ln K} = 1.32$] is considerably higher than that estimated from the 48 single-hole K data [$\sigma^2_{\ln K} = 0.38$].

Table 13.3: Statistical properties of the estimated ln-K tomograms.

Case	$\overline{\ln K}$ [K ~ cms^{-1}]	$\sigma^2_{\ln K}$	λ_x [cm]	λ_z [cm]
Hydraulic tomography [2 tests]	-2.42 [0.09]	0.62	25	10
Hydraulic tomography [4 tests]	-2.28 [0.10]	0.95	20	10
Hydraulic tomography [6 tests]	-2.28 [0.10]	1.17	24	12
Hydraulic tomography [8 tests]	-2.30 [0.10]	1.32	24	12

It is of interest to note that there is little change in the K_G and the correlation lengths of the ln-K tomograms as more tests are included in the inverse analysis. On the other hand, $\sigma^2_{\ln K}$ increases as more cross-hole tests are included into the inverse model. These results imply that with as few as 2 pumping tests, one could reliably estimate the K_G and the correlation lengths of the K distribution in the synthetic aquifer, however the accurate estimation of $\sigma^2_{\ln K}$ requires more cross-hole. A similar trend in the improvement of estimates of geostatistical parameters was observed in their hydraulic tomography analysis of steady state head data by *Illman et al*. [2010].

Table 13.4 summarizes the mean, variance, and correlation lengths of the ln-S_s tomogram. The estimated geometric mean [S_{sG}] of the ln-S_s tomogram after including data from 8 tests was 9.08×10^{-5} cm^{-1}, while the estimated variance [$\sigma^2_{\ln S_s}$] was 0.76. The estimate of S_{sG} obtained by taking the geometric mean of 48 single-hole S_s estimates yields 6.1 x 10^{-4} cm^{-1} which is somewhat higher than the estimate obtained through transient hydraulic tomography. The $\sigma^2_{\ln S_s}$ estimate obtained from the single-hole data [$\sigma^2_{\ln S_s} = 0.97$] is close to that estimated through transient hydraulic tomography. The estimated correlation lengths in the horizontal [λ_x] and vertical directions [λ_z] appear to stabilize as additional tests are included in the analysis.

Table 13.4: Statistical properties of the estimated ln-S_s tomograms.

Case	$\overline{\ln S_s}$ [Ss ~cm^{-1}]	$\sigma^2_{\ln S_s}$	λ_x [cm]	λ_z [cm]
Hydraulic tomography [2 tests]	-9.49 [7.55E-05]	0.15	15	15
Hydraulic tomography [4 tests]	-9.57 [6.98E-05]	0.19	40	20
Hydraulic tomography [6 tests]	-9.53 [7.30E-05]	0.27	34	17
Hydraulic tomography [8 tests]	-9.31 [9.08E-05]	0.76	38	19

13.5.4 Visual comparison of ln-*K* tomogram to the deposits

A visual comparison of the ln-K tomogram [Figure 13.4e] to the deposits [Figure 13.2] shows that many of the features are captured, although due to the intralayer variability in K, we do not expect a 1:1 correlation of the ln-K tomogram and the stratification seen in Figure 13.2. The comparison of the ln-K tomogram from transient hydraulic tomography to the kriged K distribution [available online as supplementary Figure S1a] shows a marked difference in the K distribution. We notice that many of the features captured by the ln-K tomogram are captured by

the kriged map, but the latter is distinctively smoother. In addition, the connectivity of the layers captured in the ln-*K* tomogram are not visible in the kriged ln-*K* field. Finally, we point out that the kriged ln-*K* distribution covers the midrange values of the ln-*K* tomogram revealing the kriged ln-*K* field produces a reasonable ln-*K* field in an average sense, but lacks the details in the heterogeneity pattern that the ln-*K* tomogram reveals.

13.5.5 Comparison of ln-*K* tomograms: transient vs. steady state hydraulic tomography

We next compare the ln-*K* tomogram obtained using transient hydraulic tomography [Figure 13.4d] to the one obtained using steady state hydraulic tomography [Figure 13.4e] computed previously by *Illman et al*. [2010]. In both cases, 8 cross-hole tests are used for the inverse analysis. Comparison of the two figures shows that overall ln-*K* distributions from the two approaches are similar, however, more details are visible in the ln-*K* tomogram from transient hydraulic tomography. To facilitate a pixel-by-pixel comparison, we include a scatterplot [Figure 13.6] of ln-*K* values from Figure 13.4d and 13.4e. The dashed line indicates a perfect 1:1 correlation; the solid line is a linear model fit to the data; and, R^2 is the coefficient of determination. Results show that the data cluster around the 1:1 line indicating agreement between the two cases, however, we acknowledge that there is some scatter and bias in the estimates.

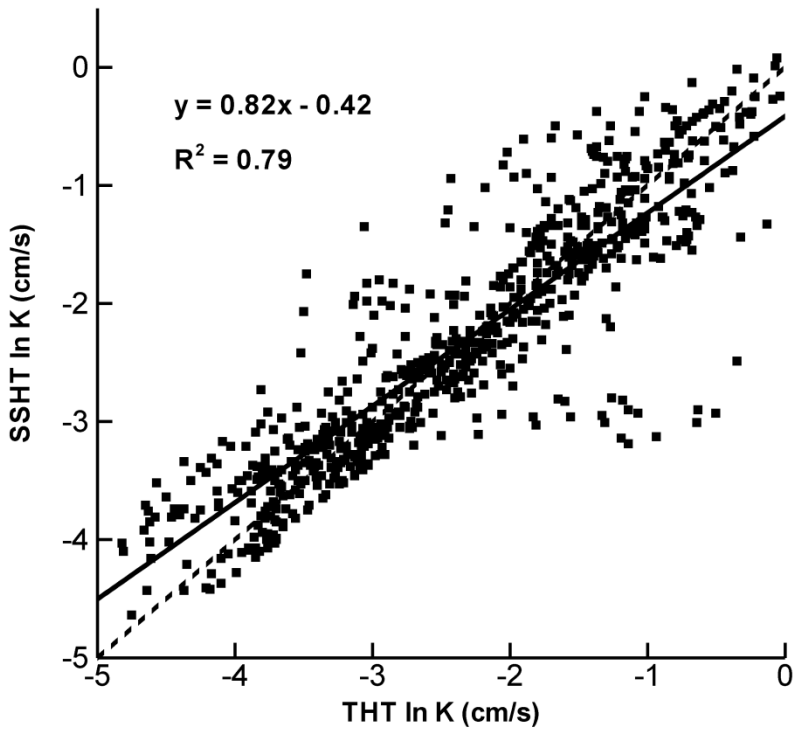


Figure 13.6: Scatterplot of ln-*K* from transient hydraulic tomography [this paper] to ln-*K* obtained using steady state hydraulic tomography [*Illman et al*., 2010]. The dashed line is a 1:1 line indicating a perfect match. The solid line is a best fit line, and the parameters describing this line are only each plot.

13.6 PERFORMANCE ASSESSMENT OF RESULTS

13.6.1 Performance assessment by simulating individual tests

The various effective and heterogeneous K and S_s fields are assessed by simulating 16 independent cross-hole pumping tests. These cross-holes tests are considered independent as they were not used by the various methods described earlier to characterize the aquifer [the exception is the cross-hole test performed at Port 20 which was used to calibrate the geological model]. If a given characterization technique can capture the salient features of the true heterogeneity of the aquifer, then the resulting prediction of the independent pumping tests should be accurate. That is, the discrepancy between observed and simulated drawdown values should be small. In contrast, if the predicted drawdown values are inaccurate, then we consider the approach used to idealize the heterogeneity to be poor. In particular, we construct forward numerical models using: a] the effective K and S_s estimates from single-hole tests; b] the PEST estimated effective parameters of K and S_s calibrated to the cross-hole test at port 20; c] the kriged K and S_s fields from single-hole tests; d] the PEST calibrated geological model; and e] the K and S_s tomograms from transient hydraulic tomography. For comparison purposes only the results from the best performing cases [d and e] are illustrated with figures in the manuscript. Tables 13.5 to 13.8 summarizing various performance metrics are provided to allow for a direct comparison of all five cases.

Figure 13.7 shows scatterplots of observed versus simulated drawdowns from independent cross-hole tests 18, 23, 40, and 42. The simulated drawdown values were obtained through numerical simulations using the calibrated geological model. Figure 13.7 includes a dashed line indicating a perfect 1:1 correlation, a solid line which is the linear model fit to the data, and the coefficient of determination [R^2]. Data plotted on the Figure 13.7 are drawdown values from 0.5, 2, 5, and 10 seconds since the pumping test began. We see that for most cases the points cluster around the 1:1 line with some positive or negative bias as indicated by the slope of the linear model fit. This same pattern is seen for most of the other tests except for those performed near the top of the aquifer, where the simulated drawdown tends to be smaller than the observed suggesting that the estimated K of the upper layer [layer 18] is too high. The R^2 values for the 16 cases range from 0.002 to 0.86 with an arithmetic mean value of 0.65. The slope and the intercept of the linear model fit also provide an indication of bias. The improvement of groundwater flow model prediction accuracy seen through the incorporation of lithofacies in conjunction with model calibration is in agreement with the findings of *Sakaki et al*. [2009].

In contrast, Figure 13.8 shows a significant improvement in the predictions of drawdowns for the 4 selected independent cross-hole pumping tests. The K and S_s distributions for all the forward models presented in this figure were obtained by transient hydraulic tomography. The R^2 values for the 16 cases range from 0.82 to 0.99 with an arithmetic mean of 0.94 indicating a marked improvement over the other modeling approaches. In addition, the comparison of the scatterplots in Figure 13.8 to those from Figure 13.7 clearly shows that transient hydraulic tomography is able to better predict independent cross-hole tests.

The results of this comparison clearly show that transient hydraulic tomography yields considerably better parameter estimates and thus is better able to predict the drawdown for 16 independent cross-hole pumping tests.

Table 13.5 summarizes the minimum, maximum, and the mean values of the slope and intercept of the linear model fit as well as the R^2 for all 16 tests for cases a] to e]. This table shows that the slope of the linear model fit is quite variable ranging from 0.01 to 3.27 for all characterization approaches and the mean values range from 0.75 to 1.00. Likewise, the intercept of the linear model fit ranges from -0.76 to 0.27 with a mean ranging between -0.44 to 0.02. Finally, the R^2 values range from 0.002 to 0.99 with a mean ranging from 0.25 to 0.94. Examination of the slope, intercept, as well as the R^2 values across all the characterization methods shows that in an average sense over 16 independent cross-hole pumping tests, the K and S_s tomograms computed by transient hydraulic tomography yields the best predictions of drawdown data with least bias and scatter.

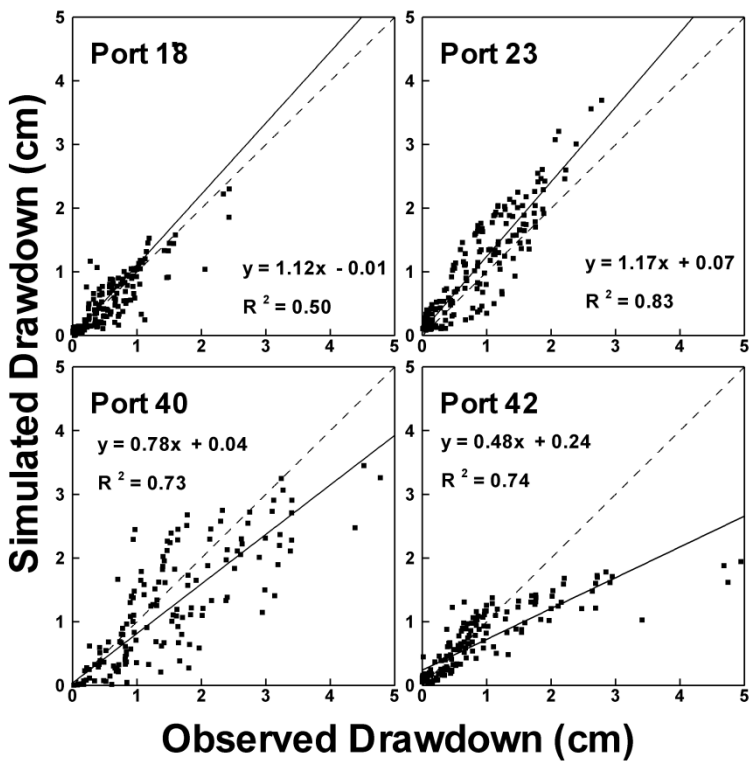


Figure 13.7: Scatterplots of observed vs simulated drawdowns from independent cross-hole tests 18, 23, 40, and 42 at t = 0.5, 2, 5, and 10 secs. Simulated drawdown values were obtained through numerical simulations with a calibrated geological model of K and S_s. The dashed line is a 1:1 line indicating a perfect match. The solid line is a best fit line, and the parameters describing this line are on each plot.

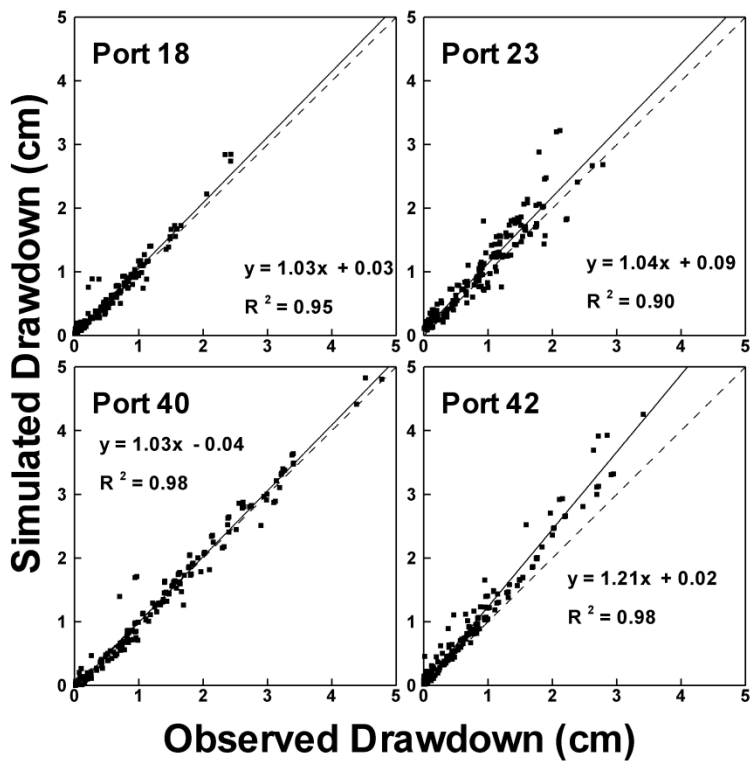


Figure 13.8: Scatterplots of observed vs simulated drawdowns from independent cross-hole tests 8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42 at t = 0.5, 2, 5, and 10 secs. Simulated drawdown values were obtained through numerical simulations with K and S_s tomograms from transient hydraulic tomography. The dashed line is a 1:1 line indicating a perfect match. The solid line is a best fit line, and the parameters describing this line are on each plot.

Table 13.5: Statistics of the linear model fit and correlation of determination [R^2].

Case	Slope			Intercept			R^2		
	min	max	mean	min	max	mean	min	max	mean
Geometric mean K/S_s [single hole]	0.51	1.50	0.79	-0.25	0.09	-0.01	0.14	0.66	0.53
PEST Effective K/S_s	0.72	3.27	1.28	-0.76	-0.24	-0.44	0.08	0.38	0.25
Kriged K & S_s	0.62	1.50	0.87	-0.27	0.09	-0.13	0.17	0.73	0.58
Calibrated Geological Model	0.01	1.44	0.96	-0.25	0.27	0.02	0.002	0.86	0.65
Transient Hydraulic Tomography	0.87	1.21	1.00	-0.04	0.09	0.02	0.82	0.99	0.94

To further quantitatively assess the correspondence between the simulated and observed drawdown values, we compute the mean absolute error [L_1] and the mean square error [L_2] norms of all cases examined. The L_1 and L_2 norms are computed as:

$$L_1 = \frac{1}{n}\sum_{i=1}^{n}\left|\chi_i - \hat{\chi}_i\right| \qquad [13.1]$$

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}\left(\chi_i - \hat{\chi}_i\right)^2 \qquad [13.2]$$

where n is the total number of drawdown data, i indicates the data number, and χ_i and $\hat{\chi}_i$ represent the estimates from the simulated and measured drawdowns, respectively. The L_1 and L_2 norms were calculated for each case by evaluating the observed and simulated drawdowns at 4 times [0.5, 2, 5, and 10 seconds] at each port, except for the port that was pumped due to excessive noise. Thus, each L_1 and L_2 norm represents 188 observations.

Table 13.6 summarizes the L_1 norm, while Table 13.7 summarizes the L_2 norm calculated for all the cases. The cells of each entry in the table are color-coded. The minimum value in the table is assigned a color of dark green, the maximum value a color of dark red, and the median value a color of yellow. Values intermediate to these anchor points are assigned appropriate intermediate colors. Both Tables 13.6 and 13.7 show that forward simulations using K and S_s tomograms from transient hydraulic tomography, consistently yields the lowest L_1 and L_2 norms for all 16 independent cross-hole pumping tests, suggesting that the approach yields the best predictions.

The scatterplots for cases a] to e] were also analyzed in an ensemble sense to see if the homogeneous cases were able to estimate the average behaviour of all of the 16 cross-hole tests. This comparison is presented for Table 13.8 and includes; L_1 norm, L_2 norm, slope, intercept, and R^2 for all cases when the data from all 16 cross-hole tests are analyzed collectively. The trends seen in the individual scatterplots are also seen in these ensemble scatterplots. The results for transient hydraulic tomography after 2, 4, 6, and 8 tests are also presented in Table 13.8. While not changing significantly with the inclusions of additional cross-hole tests, the matches do improve slightly.

Table 13.6: L1 norms of observed versus simulated drawdowns from cross-hole tests 8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42.

	Port 8	**Port 11**	**Port 13**	**Port 15**	**Port 16**	**Port 18**	**Port 20**	**Port 23**	**Port 26**	**Port 29**	**Port 37**	**Port 38**	**Port 39**	**Port 40**	**Port 41**	**Port 42**
Geometric mean K/S_s [single hole]	0.20	0.31	0.23	0.35	0.40	0.30	0.61	0.39	0.62	0.57	0.51	0.66	0.58	0.76	0.65	0.42
PEST Effective K/S_s	0.39	0.62	0.48	0.65	0.78	0.57	0.91	0.72	0.91	0.89	0.81	0.99	0.93	1.18	0.97	0.67
Kriged K & S_s	0.19	0.33	0.23	0.34	0.38	0.30	0.56	0.37	0.55	0.50	0.48	0.57	0.48	0.65	0.53	0.36
Calibrated Geological Model	0.17	0.22	0.38	0.31	0.35	0.22	0.27	0.32	0.48	0.36	0.44	0.39	0.39	0.44	0.39	0.30
Transient Hydraulic Tomography	0.06	0.11	0.12	0.11	0.11	0.07	0.13	0.17	0.17	0.16	0.07	0.13	0.15	0.10	0.10	0.19
Max	1.18															
Min	0.06															

Table 13.7: L2 norms of observed versus simulated drawdowns from cross-hole tests 8, 11, 13, 15, 16, 18, 20, 23, 26, 29, 37, 38, 39, 40, 41, and 42.

	Port 8	**Port 11**	**Port 13**	**Port 15**	**Port 16**	**Port 18**	**Port 20**	**Port 23**	**Port 26**	**Port 29**	**Port 37**	**Port 38**	**Port 39**	**Port 40**	**Port 41**	**Port 42**
Geometric mean K/S_s [single hole]	0.14	0.18	0.13	0.26	0.25	0.16	0.58	0.25	0.61	0.57	0.50	0.81	0.61	1.00	0.78	0.37
PEST Effective K/S_s	1.97	2.06	1.92	2.26	1.99	1.89	2.25	1.97	2.09	2.24	2.08	2.50	2.28	2.82	2.25	1.71
Kriged K & S_s	0.12	0.18	0.14	0.24	0.24	0.16	0.49	0.24	0.49	0.45	0.43	0.61	0.43	0.75	0.53	0.27
Calibrated Geological Model	0.09	0.19	0.83	0.15	0.23	0.31	0.37	0.17	0.39	0.25	0.47	0.32	0.34	0.35	0.38	0.31
Transient Hydraulic Tomography	0.01	0.02	0.04	0.03	0.04	0.02	0.03	0.06	0.05	0.05	0.01	0.03	0.06	0.02	0.02	0.10
Max	2.82															
Min	0.01															

Table 13.8: Statistics of the linear model fit, correlation of determination [R^2], and L_1 and L_2 norms for the ensemble analysis of all cases [including transient hydraulic tomography after the inclusion of 2, 4, 6, and 8 tests]

Case	Slope	Intercept	R^2	L_1	L_2
		Homogeneous K Field			
Geometric mean K/S_s [single hole]	0.61	-0.02	0.44	0.48	0.53
PEST Effective K/S_s	0.86	-0.25	0.20	0.78	2.14
		Heterogeneous K Field			
Kriged K & S_s	0.73	-0.06	0.52	0.43	0.45
Calibrated Geological Model	0.99	0.18	0.67	0.34	0.32
Transient hydraulic tomography [Sequential - 2 tests]	1.06	0.01	0.92	0.17	0.07
Transient hydraulic tomography [Sequential - 4 tests]	1.03	0.04	0.95	0.13	0.04
Transient hydraulic tomography [Sequential - 6 tests]	1.05	0.06	0.96	0.14	0.04
Transient hydraulic tomography [Sequential - 8 tests]	1.01	0.02	0.95	0.12	0.04

13.6.2 Predictability of transient drawdown curves

Finally, to further illustrate the robustness of transient hydraulic tomography, we also plot simulated and observed drawdown curves for a cross-hole test performed at port 40. Again only the cases d] and e] are included in the manuscript. In particular, Figures 13.9 and 13.10 show double logarithmic plots of observed [small dots] and simulated [curves] drawdown records at 16 selected ports. In Figure 13.9, the simulated drawdown values are obtained through numerical simulations using the calibrated geological model, while in Figure 13.10, simulated drawdown values are obtained through numerical simulations with the K- and S_s-tomograms.

Figure 13.9 shows that the calibrated geological model does reasonable job of predicting drawdown at most of the ports, however, drawdown is significantly under estimated in the upper ports. This again suggests that the upper layer [layer 18] is not accurately characterized. In contrast, Figure 13.10 shows a drastically improved result when the simulated values are obtained from the forward simulation of the pumping test using the K and S_s tomograms from transient hydraulic tomography. Examination of Figure 13.10 shows that the match is not perfect, but overall, the drawdown curves are captured very well throughout the duration of the pumping test, which the other approaches failed to accomplish.

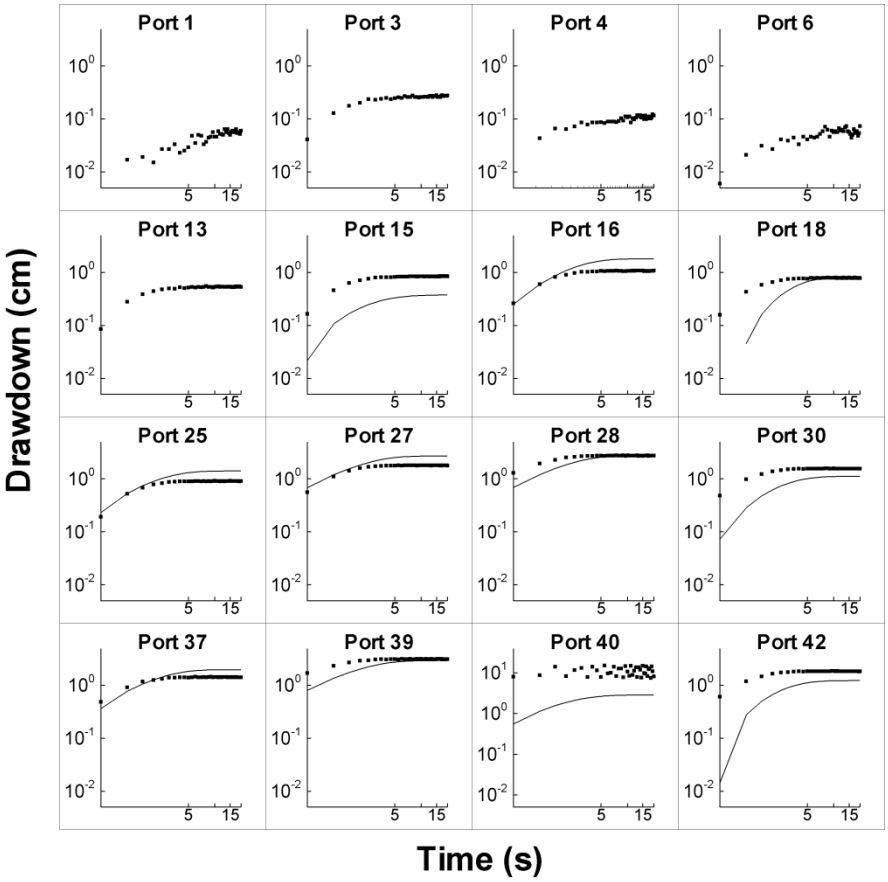


Figure 13.9: Observed [small dots] and simulated [curves] records of drawdown versus time [seconds] during cross-hole pumping test with pumping at port 40. Simulated values are obtained from the forward simulation of the pumping test using the calibrated geological model.

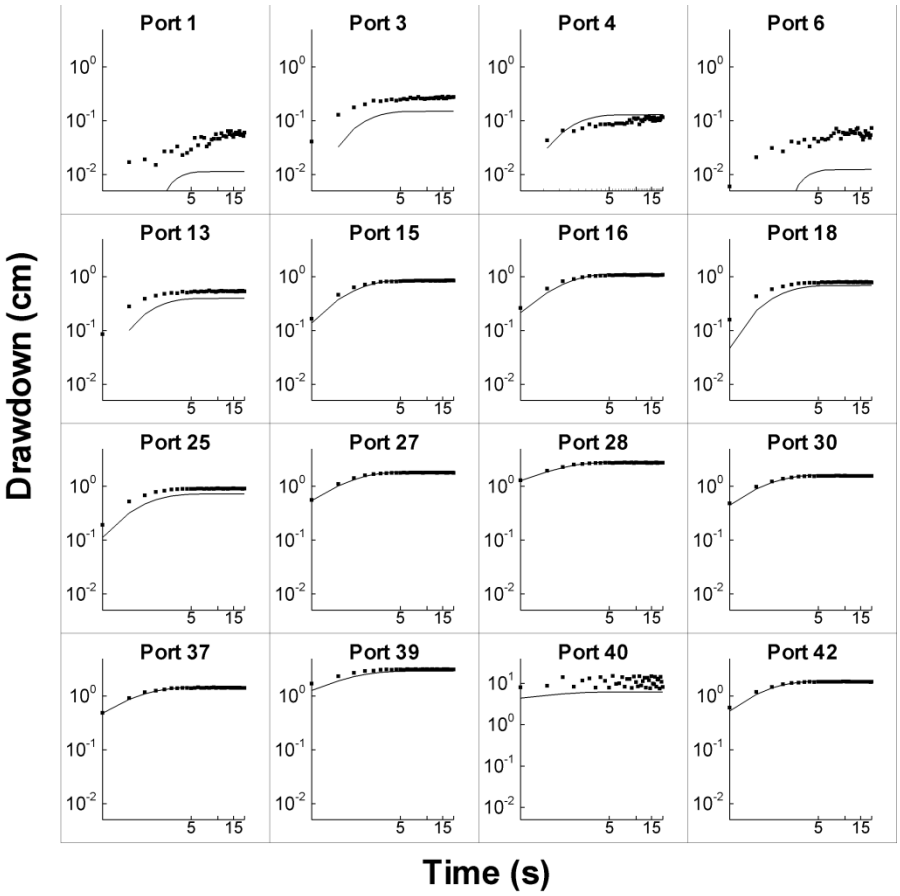


Figure 13.10: Observed [small dots] and simulated [curves] records of drawdown versus time [seconds] during cross-hole pumping test with pumping at port 40. Simulated values are obtained from the forward simulation of the pumping test using the K and S_s tomograms from transient hydraulic tomography.

14. TEMPORAL MOMENT ANALYSIS OF PARTITIONING TRACERS IN LABORATORY SANDBOX AQUIFERS

14.1 INTRODUCTION

There are a number of approaches to characterize the presence of DNAPLs in the source area. In general, methods to characterize DNAPL source zones are based on core sampling, cone penetrometer testing, geophysical logging and partitioning tracer methods [Kram et al., 2001]. Kram et al. [2001] provided a literature survey on the strengths and weaknesses of various characterization methods.

Baseline methods [DNAPL samples obtained from observation wells, chemical analysis of soil, water and rock samples among others] consist of sample collection during drilling. Samples are collected using drilling equipment and are analyzed by means of EPA-approved methods for identifying Volatile Organic Carbons [VOCs]. Due to the fact that VOC can be readily set free during handling and transport related operations, the loss of VOCs could result in the underestimation of actual concentrations. Another important issue is that DNAPLs may be found as tiny globules, and the probability of detecting them is low unless the sampling frequency is high. There are also uncertainties related to the soil-water partitioning coefficient which depends on the fraction of organic carbon content present in the subsurface, sorption and other phenomena [Kram et al., 2001].

Soil gas surveys can also be used to screen DNAPLs. This is done by inserting soil-vapor collection devices, and capturing the samples by means of applying a vacuum in the soil. Once the sample is collected, the VOCs can be measured using gas chromatography analysis. The data obtained is used in order to detect a DNAPL release area. Another method to detect DNAPLs is by measuring the concentrations of Radon-222. This element [which is often present as a dissolved, inert gas in subsurface fluids] occurs in nature, and results from the decaying of Uranium-238. Since this gas has a strong affinity to organic fluids compared to water, the loss of Rn-222 concentrations measured at monitoring wells is a signal of DNAPL presence [Semprini et al. 1998]. However, since this technique relies on a best-guess approach for sampling, it is used more as an evaluation technique of remedial effectiveness than a detection tool [Kram et al., 2001].

Another method to characterize DNAPL source zone is through partitioning tracers. The partitioning tracer test was first used in the petroleum industry as a way to measure oil saturations prior to the beginning of enhanced oil recovery [Cooke, 1971; Deans, 1971; Tomich, 1973] by measuring residual oil. Later Jin et al., 1995, used PTTs to detect NAPLs. Partitioning Tracer Tests [PTTs] are a technique that can be used or applied in place in order to estimate the volume and saturation of the DNAPLs including chlorinated solvents such as TCE and PCE in the subsurface. The theoretical basis of partitioning tracer tests is summarized by Annable et al. [1998] based on the method proposed by Jin et al. [1995]. The technique consists of injection and extraction of a solution of non-partitioning and partitioning tracers into a porous media. The concentrations of the injected tracers are known *a priori* and the concentrations of the extracted tracers are monitored over time. A non-partitioning tracer is injected in conjunction with the partitioning tracer in order to separate the effects of the non-partitioning and partitioning tracers.

The difference between the arrival times of the tracers is monitored by analyzing the breakthrough curve [BTC]. The “delay” [Figure 10.1] in the BTCs will lead to what is called the retardation factor [R]. The retardation factor is proportional to the water-NAPL partition

coefficient, K_{NW}, and the saturation of the DNAPL [S_N]. Each partitioning tracer will experience different retarded arrival times at the extraction wells. The time necessary for a tracer to arrive at the recovery well is a function of its partitioning coefficient and the average DNAPL saturation, S_N, in the section tested [Jin et al. 1995; Annable et al 1998]. As the partitioning coefficient increases, the time required for the tracer to reach the extraction well increases. By monitoring the concentrations at the wells, BTCs for each tracer will be generated and the separation between the partitioning and the non partitioning tracers will allow an estimation of the DNAPL volume in the subsurface. If there is no DNAPL present in the subsurface, all the curves [the ones for non-partitioning and partitioning tracers] should be superimposed. Prior to the conduct of a partitioning tracer test, it may be necessary that the general location of the DNAPL is identified through other characterization methods *a priori* [Kram et al. 2001].

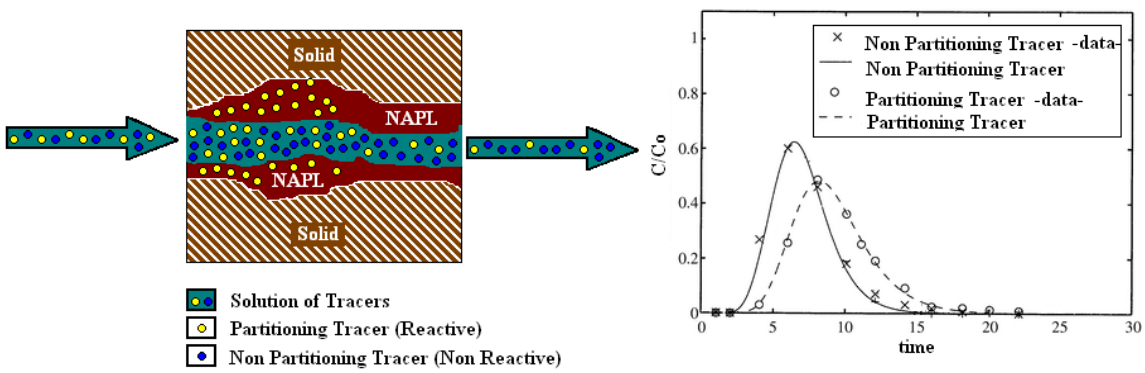


Figure 14.1: Microscopic view of the NAPL/Tracer interaction and breakthrough curves.

To date, many studies have been conducted on the use of PTTs as an innovative tool to assess NAPL saturations in the subsurface at the field scale [e.g., Annable et al., 1998; Brooks et al., 2005], in laboratory column experiments [e.g., Maloszewki et al. 1994; Jin et al., 1995; Pang et al. 1998; Yu et al., 1999], and in flow cells with heterogeneous media and diverse sampling configurations [e.g., Nelson et al., 1999] while others have studied the effect of DNAPL architecture on the performance of partitioning tracer tests [e.g., Moreno-Barbero et al., 2007].

The objectives of this investigation are: 1] to investigate the effect of hydraulic conductivity heterogeneity on TCE saturation estimates from the temporal moment analysis of partitioning tracers and 2] to investigate the performance of various partitioning tracers in heterogeneous aquifers. To accomplish the objectives, we conducted a set of PTTs in three different heterogeneous artificial aquifers [SB3-1, SB-3-2 and SB 1-2] with a TCE source zone. We describe the laboratory experimental methods and provide corresponding analyses of data. Then we provide an interpretation for the results and conclusions and recommendations are made for future investigations.

14.2 TEMPORAL MOMENT ANALYSIS OF PARTITIONING TRACER TESTS

The analysis of BTCs have been carried out principally through two approaches: approaches: calibration of an analytical or numerical model or through the Method of Moments [MOM]. Our discussion here will be limited to the MOM approach.

The MOM relies on the concept of DNAPL-water partitioning coefficient, K_{NW} which linearly relates the concentration of the tracer in the DNAPL phase to the concentration in the aqueous phase.

$$K_{NW} = \frac{C_N}{C_W} \qquad [14.1]$$

Under steady state flow conditions, the retardation of the partitioning tracer with respect to the non-reactive tracer results due to the presence of DNAPLs and the delayed average travel time is a function of the average DNAPL saturation, S_N, within the tracer flow field [Pope et al. 1994]

$$S_N = \frac{R-1}{K_{NW} + (R-1)} \qquad [14.2]$$

where R is the retardation factor, which accounts for the difference of average characteristic times for both the partitioning, t_p and the non partitioning, t_n, tracers

$$R = \frac{t_p}{t_n} \qquad [14.3]$$

The average travel times [t_p and t_n] are computed through the MOM approach [Annable et al, 1998].

We have the dimensionless relative concentration of the tracer, $C_e(x_i, t)$, in space and time which is expressed as:

$$C_e(x_i, t) = \frac{C(x_i, t) - C_i}{C_0 - C_i} \qquad [14.4]$$

where: $C(x_i, t)$ is the sampled concentration at the port, at any given time "t" during the experiment; C_0, is the initial concentration of the tracer [partitioning or non-partitioning] at the tracer reservoir and C_i is the initial concentration [background concentration] of the tracer [partitioning or non-partitioning] in the SB. The nth moment equation is given by:

$$M_n(x_i) = \int_0^\infty t^n \cdot C_e(x_i, t) \cdot dt \qquad [14.5]$$

where t^n is time, n is an exponent with values of 0,1,2,… and $C_e(x_i, t)$ is the dimensionless relative concentration in space and time. Equation 14.5 allows one to compute the 0th and the 1st moments as:

$$M_0(x_i) = \int_0^{\infty} C_e(x_i, t) \cdot dt \qquad [14.6]$$

$$M_1(x_i) = \int_0^{\infty} C_e(x_i, t) \cdot t \cdot dt \qquad [14.7]$$

The moments can be approximated by a summation [Eikens and Carr 1989], thus the 0th and 1st moment at each well can be obtained through the following expressions:

$$M_o = \int_{T_1}^{T_2} C_e(t) \cdot dt \approx \sum_{i=1}^{N} C_{ei} \cdot \Delta t \qquad [14.8]$$

$$M_1 = \int_{T_1}^{T_2} C_e(t) \cdot t \cdot dt \approx \sum_{i=1}^{N} C_{ei} \cdot t.\Delta t \qquad [14.9]$$

Here, $C_{e,i}(t)$ is the relative concentration of the sample at the time of its acquisition [t] and Δt is the time difference between the sample time t_i and the previous sample time t_{i-1}. The average arrival time for each tracer is then obtained by the division of the 1st and the 0th moment as shown below

$$t_p = \frac{M_{1,p}}{M_{0,p}} \qquad [14.10]$$

$$t_n = \frac{M_{1,n}}{M_{0,n}} \qquad [14.11]$$

Assuming that the there are no DNAPL mass losses and that with the contribution of each well the source zone is swept, the estimation of the DNAPL volume in the sandbox per sampling interval is given by the following expression:

$$V_{NAPL,i} = V_{void} \cdot S_{N,i} \qquad [14.12]$$

where, $S_{N,i}$ is the DNAPL saturation per port and V_{void} is the void volume of the swept zone. Finally, the average DNAPL volume can be computed by averaging the DNAPL volume estimate from all sampled wells as follows:

$$\overline{V_{NAPL}} = V_{void} \cdot \overline{S_{N,i}} \qquad [14.13]$$

14.3 EXPERIMENTAL METHODS

14.3.1 Sandbox descriptions

Four sandboxes [SB 3-1, SB 3-2, SB 1-2 and SB 1-3] were built at the IIHR-Hydroscience & Engineering at the University of Iowa in order to conduct the laboratory investigation of partitioning tracer tests. Sandboxes 3-1 and 3-2 were to be used as a pilot study for DNAPL release and partitioning tracer tests [PTTs]. Details to each of the sandboxes and hydraulic tests conducted in them are described in section 10.

14.3.2 TCE source zone creation in SBs 3-1, 3-2 and 1-2

TCE source zones were created in sandboxes 3-1, 3-2 and 1-2 [figures 14.2 – 14.4] by injecting known volumes of TCE into the porous media just below the soil surface. In sandbox 3-1, 25 mL of TCE previously dyed with [Red Sudan IV, Sigma-Aldrich] was injected to create a source zone which constituted a model-averaged DNAPL saturation, S_N, around 1%. In SB 3-2, the amount of dyed TCE injected was 125mL which constituted a model-averaged TCE saturation around 5%. The saturation estimates are calculated with the following expression:

$$S_{Ni} = \frac{V_{NAPL}}{V_{voids}} \qquad [14.14]$$

where V_{NAPL} is the volume of TCE present and V_{voids} is the volume of voids. For SB 1-2, 2000 ml of TCE were spilled in the sandbox yielding an average DNAPL saturation around 5%.

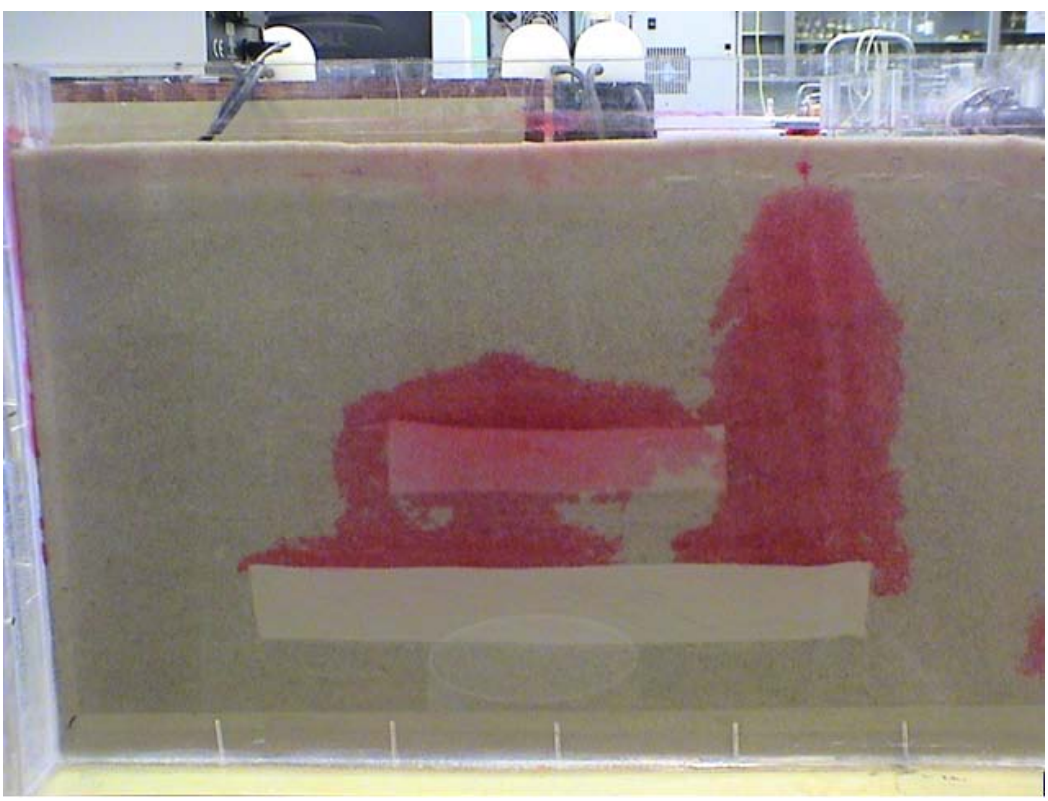


Figure 14.2: TCE source zone created in sandbox 3-1.

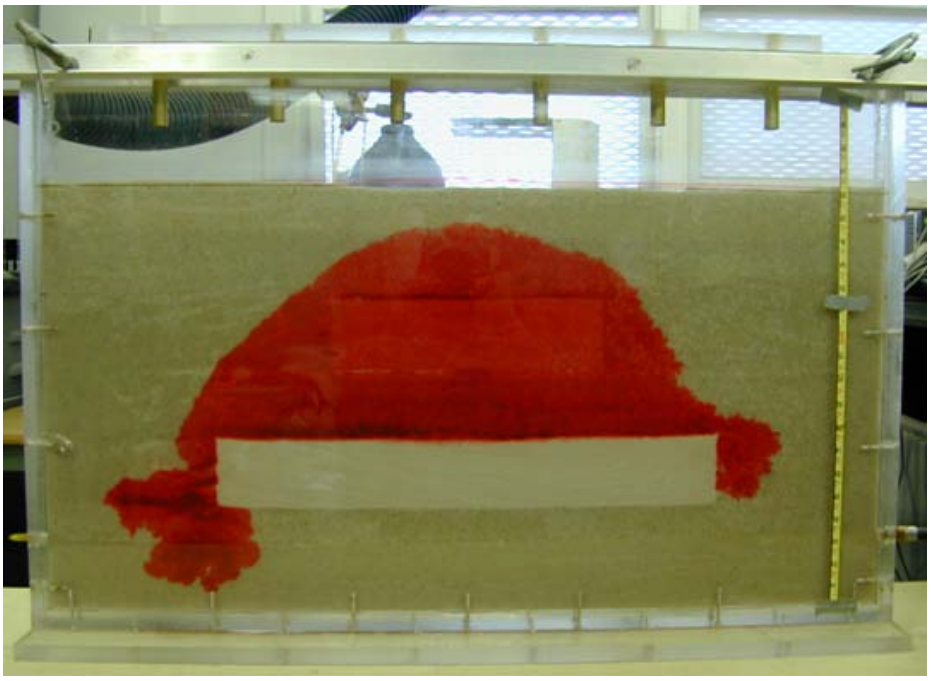


Figure 14.3: TCE source zone created in sandbox 3-2.

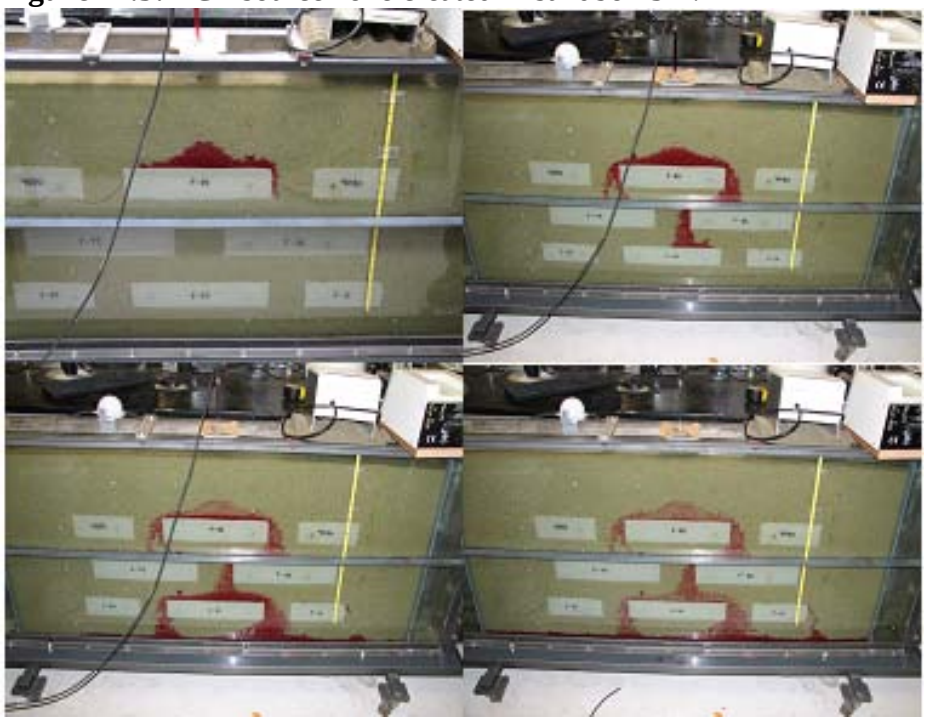


Figure 14.4: TCE source zones created in sandbox 1-2. Clockwise from top left to bottom right: photo taken after 15 min; photo taken after 30 min; photo taken after 2 hours; photo taken after 24 hours.

14.3.3 Determination of tracer partitioning coefficients through batch tests

In order to determine the DNAPL saturation, knowledge of the partitioning tracer coefficients for each site and each type of DNAPL needs to be known. Methods to estimate

partitioning coefficients have been developed using UNIFAC [Wang et al., 1998] and on defining equivalent alkane carbon number for each contaminant in a DNAPL and the tracers [Dwarakanath and Pope, 1998]. We propose to follow the example of other researchers [Brusseau et al., 1999; Hayden and Linnemeyer, 1999; Cain et al., 2000] who determined specific partitioning coefficients for individual tracers with the specific samples of the DNAPLs from their sites.

Four alcohols were chosen to be used as partitioning tracers for this study. These tracers were chosen because of their frequent use in partitioning tracer tests and because they exhibit a wide range of K_{NW} values. The NAPL-phase used in this study was trichloroethylene [TCE] because it is such a common DNAPL pollutant in the United States. The purities of the alcohols and TCE purchased from Sigma-Aldrich and Fluka are as follows: TCE > 99.5%; IPA > 99.5%; 2M1B > 99%; 2E1H > 99%; 2, 4DM3P = 99%; and Hex ≥ 99.9%.

Seven sets of batch tests were conducted to determine NAPL – water partition coefficients for each tracer in water-tracer-NAPL solutions as well as to examine the effects of the presence or organic carbon [f_{oc}] on partition coefficient estimation. Six partitioning-tracer stock solutions [levels] were prepared volumetrically with Hamilton gas-tight syringes in 250-mL volumetric flasks. The levels ranged from 20mg/L to 640 mg/L except for 2E1H. 2E1H, due to its low solubility limit of approximately 850 mg/L [*Istok et al.* [2002]], was prepared at concentrations ranging from 12.5mg/L to 400mg/L. Purified water [Barnstead Nanopure, Diamond] was used to dilute the aqueous tracer stock solutions to the six different concentrations: 50, 100, 200, 400, 800, and 1600 mg/L. The 6 different levels were prepared in pre-cleaned 20mL EPA/ VOA glass vials with Teflon-lined septa from EnviroTech. For 2E1H the 6 batch test levels were 25, 50, 75, 150, 300, and 600 mg/L.

The batch tests to determine NAPL-water partition coefficients for each tracer contained six sample groups. The batch tests were prepared in pre-cleaned 25mL EPA/ VOA glass vials with Teflon-lined septa from EnviroTech. Each sample set consisted of triplicate vials containing TCE and the aqueous tracer solution and a control group of triplicate vials containing aqueous tracer solution only [initial concentration]. 20 mL of aqueous tracer solution was dispensed into each vial using a 10 mL pipette followed by 2.5 mL injection of TCE with a gas-tight syringe. The vials were immediately capped to avoid losses to the atmosphere.

The 25mL vials were allowed to equilibrate on a platform shaker [Innova 2000] at 150 rpm and room temperature for 24 hours, which is sufficient to obtain equilibrium [*McCray and Dugan*, 2002]. 1.5 mL aliquots of the aqueous phase were withdrawn from the 25mL vials with minimal disturbance using a gas-tight syringe and transferred to 1.5 mL glass GC auto sampler vials with Teflon septa.

14.3.4 Determination of partitioning tracer concentrations

Concentrations of alcohols used as partitioning tracers were determined using a gas chromatograph [Shimadzu GC 2010] with an auto sampler [AOC-20i auto injector] equipped with a flame ionization detector [FID]. The GC/FID method consisted of a Restek Stabilwax® capillary column [30m, 0.32 mm ID, with film thickness of 1µm], helium as the carrier gas, zero-grade air, an injection volume of 1 µL, and injector and FID temperatures of 200°C. Also, a split ratio of 35:1 to vent excess water [steam] in the sample and a linear velocity of 50 cm/s were incorporated into the method. A different temperature program was used for each alcohol. This provided more efficient way to estimate the concentration of the collected samples [the 5 different temperature gradient programs used are listed in Appendix F]. Four external calibration

standards were prepared for 2M1B in deionized water at levels of 1, 10, 100, and 1000 mg/L. The calibration plots were created by injecting triplicates of each calibration level and were linear [$R^2 > 0.99$]. Calibrations were made in each analysis sequence consisting of 12 calibration injections [triplicates from each one of the 4 standards], 120 tracer samples, and 3 quality control samples. The quality control samples were vials filled with 100 mg/L standard and was analyzed after each set of 40 samples to ensure that the signal response of the machine remained accurate. Methanol injections [blanks] were made in between sets of samples to prevent carryover. We note that previous researchers [Dugan et al. 2003] prepared calibration standards composed of the tracers in methanol. We used water instead of methanol to analyze the sample set more accurately since our samples are aqueous and we need calibration curves that fully represent the actual conditions of our samples.

The same procedure was applied to obtain the partitioning coefficients of the tracers with PCE. PCE was selected to create a source zone in SB1-3 due to its lower solubility compared with TCE.

14.3.5 Determination of conservative tracer concentrations

Conservative tracer [Br] concentrations were determined using an ion chromatograph [IC] [Dionex ICS 2000] equipped with an AS40 auto sampler. The IC analysis method consisted of an AS-18 4 mm column, EGC II KOH eluent generator with an eluent concentration of 23mM, SRS-Ultra II 4mm suppressor with a current of 57 mA, DS6 heated conductivity cell with a cell temperature of 35° C, and a flow rate of 1 mL/min. Four external calibration standards were prepared for bromide with 1000mg/L Fluka IC standard diluted with de-ionized water at levels of 1, 10, 100, and 200 mg/L. The calibration plots were created by injecting triplicates of each calibration level and were linear [$R^2 > 0.99$]. Calibrations were made in each analysis sequence consisting of 12 calibration injections [triplicates of the 4 standards] followed by 80 tracer samples. Clean up runs were made in between each standard and sample injection to prevent carryover by increasing the eluent strength to 50 mM and increasing the suppressor current to 124 mA for 8 minutes.

14.3.6 Partitioning tracer tests in SBs 3-1, 3-2 and 1-2

PTTs were conducted on SB 3-1, 3-2 and 1-2 to quantify DNAPL saturations in the heterogeneous aquifers. In each PTT, a dipole flow field was created by injecting water at one port at a rate of about 500 mL/min and extracting at another port at a rate of about 450 mL/min. A mass balance is then achieved by setting 12 sampling lines at a cumulative rate of about 50 mL/min [~4.17 mL/min/line]. Achieving a mass balance is crucial because the 4 boundaries of the sandboxes had no-flow boundary conditions. Deviations in the mass balance would show up as a rising or falling water level.

Once steady state flow conditions were reached, a valve was switched on the injection line from water to a tracer solution consisting of bromide [non-partitioning tracer ~100mg/L] and 2-methyl-1-butanol [2M1B] [partitioning tracer ~500mg/L] for 10 minutes and then the valve was switched back to water. Separate sample sets were collected for bromide and 2M1B throughout the duration of the test and a time stamp was assigned to each sample.

The data collection from each one of the sampling lines was possible with a manifold created in the lab. In this fashion we obtained 12 simultaneous samples with vials placed in sampling racks. Thirty 2M1B samples and 20 bromide samples were collected from each of the 12 sampling lines during each PTT for a total of 600 samples per PTT.

A series of six PTTs were conducted in sandbox 3-1. In each PTT, a dipole flow field is created by injecting water at one port at a rate of about 500 mL/min and extracting at another port at a rate of about 450 mL/min. A mass balance is then achieved by setting 12 sampling lines at a cumulative rate of about 50 mL/min [~4.17 mL/min per line]. Figure 14.5 illustrates the injection, extraction, and sampling ports of all PTTs carried out in sandbox 3-1.

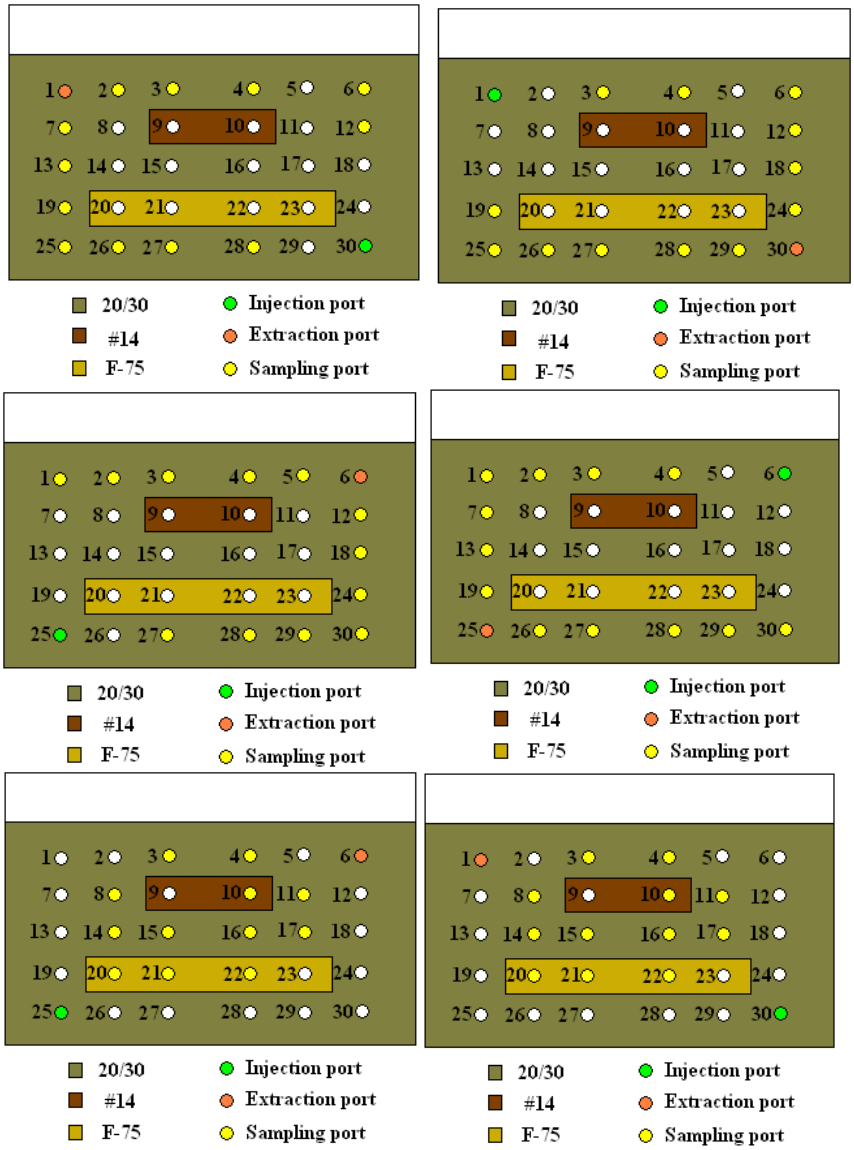


Figure 14.5: PTT configurations in SB3-1. From top left, clockwise, PPT1, PPT2, PTT3, PTT4, PTT5 and PTT6, in that order.

A series of four PTTs were conducted in sandbox 3-2. In each PTT, a dipole flow field was created by injecting water at one port at a rate of about 500 mL/min and extracting at another port at a rate of about 450 mL/min. A mass balance is then achieved by setting 12 sampling lines at a cumulative rate of about 50 mL/min [~4.17 mL/min per line]. Figure 14.6 illustrates the design of all PPTs carried out in sandbox 3-2.

A series of four PTTs were conducted in sandbox 1-2. In each PTT, a dipole flow field is created by injecting water at one port at a rate of about 500 mL/min and extracting at another port at a rate of about 450 mL/min. A mass balance is then achieved by setting 12 sampling lines at a cumulative rate of about 50 mL/min [~4.17 mL/min per line]. PTTs 1 through 3 were conducted with bromide [non-partitioning tracer] and 2M1B [partitioning tracer]. PTT 5 was conducted with bromide and three partitioning tracers: 2M1B, 2-4-dimethyl-3-pentanol [24DM-3P], and 1-hexanol [1-HEX]. A schematic representation of the configuration of each PTT test in SB 1-2 is presented in Figure 14.7.

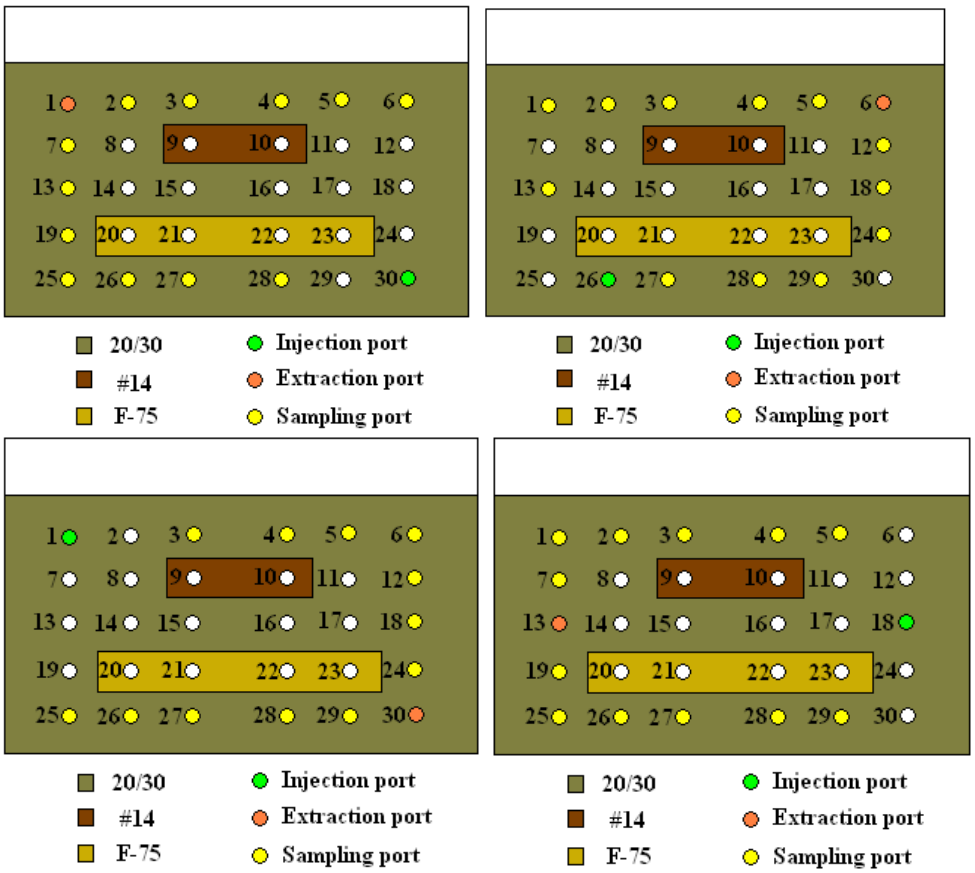


Figure 14.6: PTT configurations in SB 3-2. From PTT1 [top left], PTT2 [top right], PTT3 [bottom left], and PTT4 [bottom right].

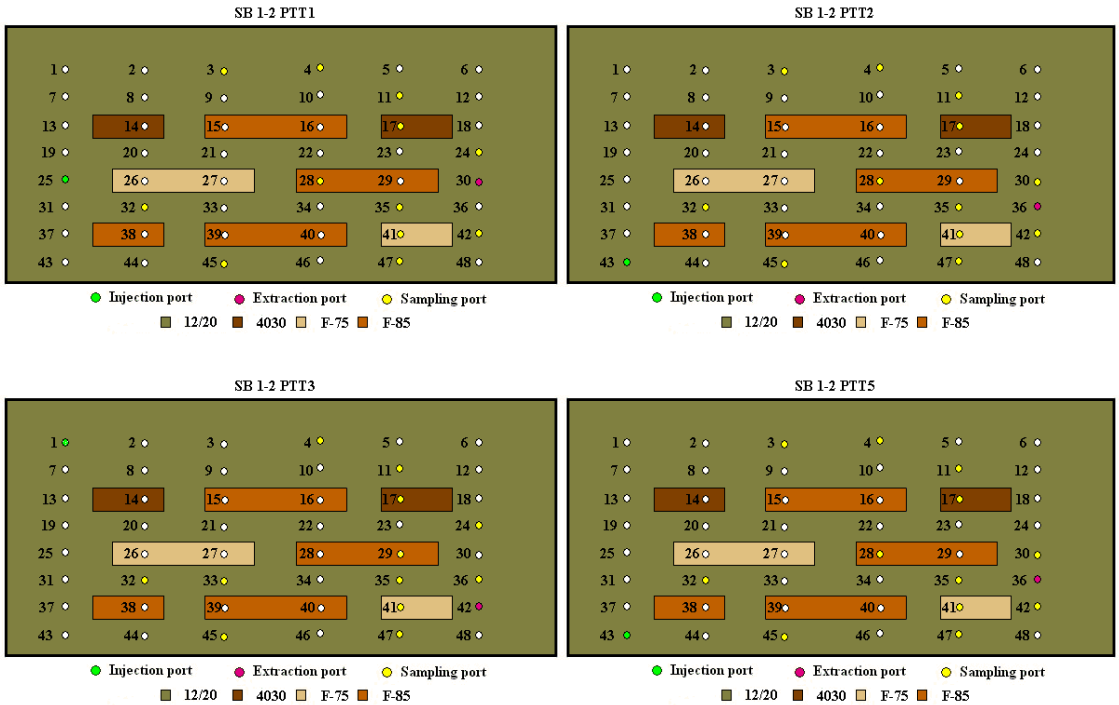


Figure 14.7: Sandbox 1-2 PTT designs: PTT1 [top left], PTT2 [top right], PTT3 [bottom left] and PTT5 [bottom right].

14.3.7 Extraction of cores and its analyses

As soon as PTTs were completed on SB 3-2 and SB 1-2, cores were taken from those sandboxes. The purpose of obtaining cores was to obtain an independent data set of TCE saturations so that we can compare to those obtained by means of MOM analysis of PTTs. For SB 3-2, six brass tubes were inserted [vertically downward] into the sandbox in fully saturated conditions until they reached the floor of the model. A short piece of larger diameter tubing served as a guide where another brass tube was inserted vertically [Figure 14.8; see also Figure 14.9]. Then, bentonite was added to the top of the core to eliminate the empty space and the core tubes were plugged at the end with epoxy putty. Once the epoxy putty was dried, the rest of the model was. The cores were then removed individually and cut into segments with a pipe cutter following the schematic presented on Figure 14.10 for SB3-2 and Figure 14.11 for SB1-2. Each of the segments was capped on each end with a silicone plug [Figure 14.12] and placed in a refrigerator at 5° C.

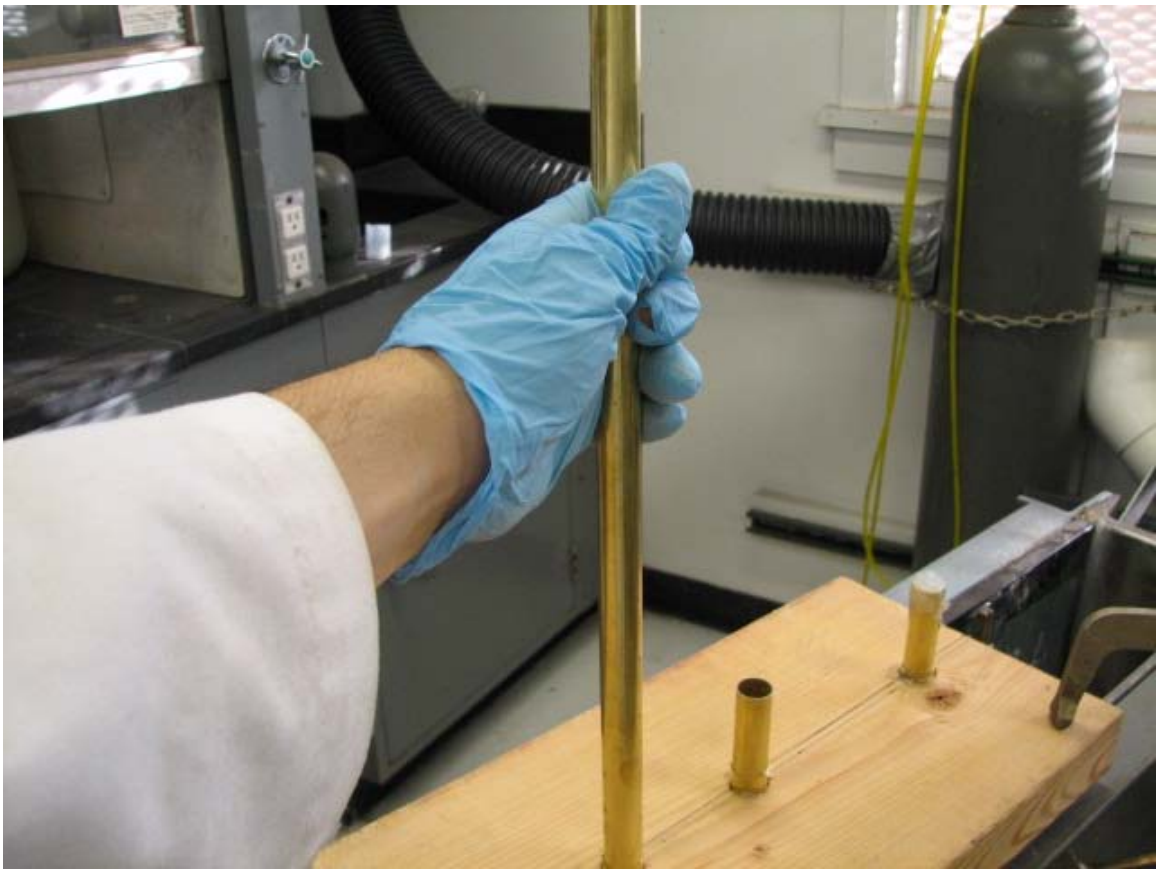


Figure 14.8: Picture of the telescopic tubing arrangement for SB 1-2 [similar process for SB 3-2 was applied].

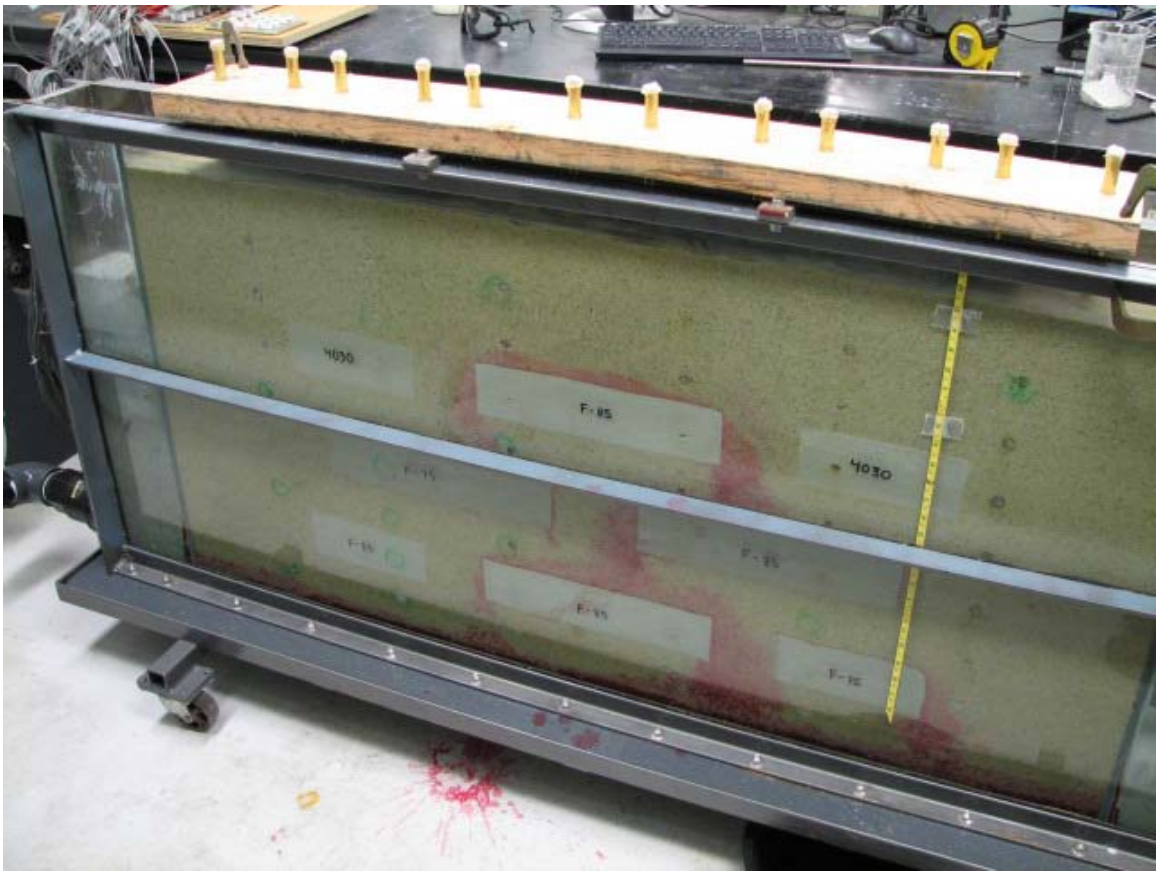


Figure 14.9: Fully inserted coring wells [SB 1-2].

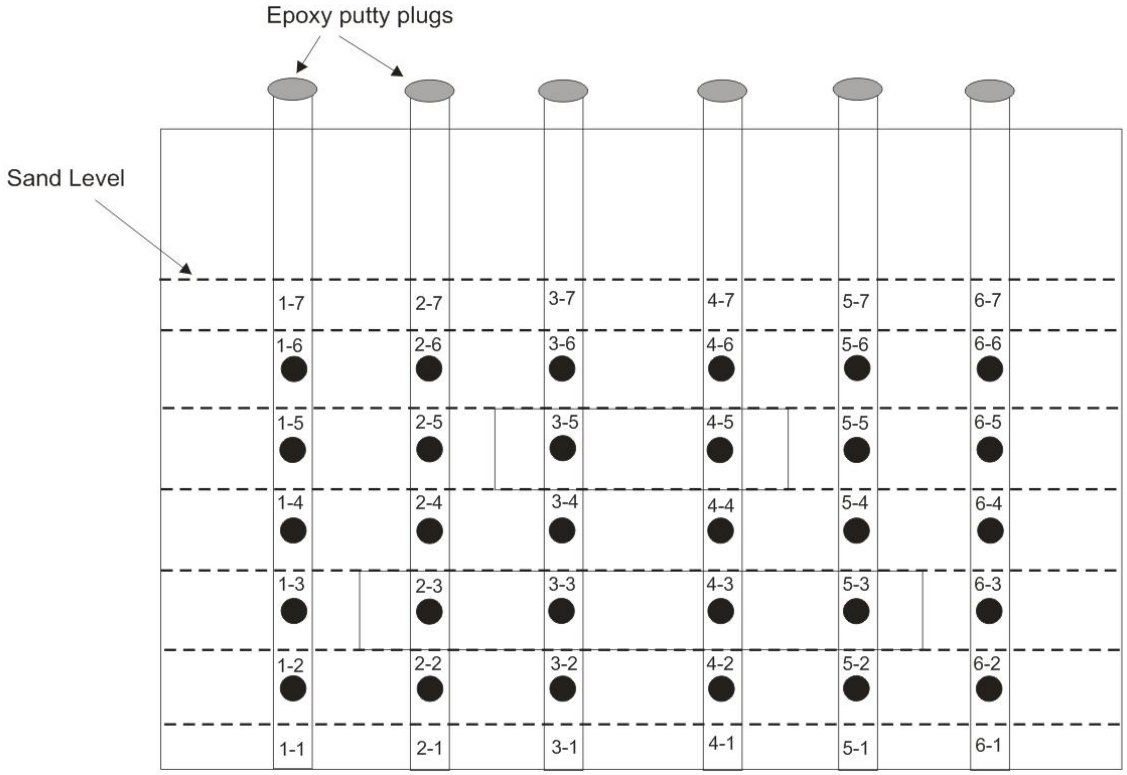


Figure 14.10: Partitioning of core samples in SB 3-2.

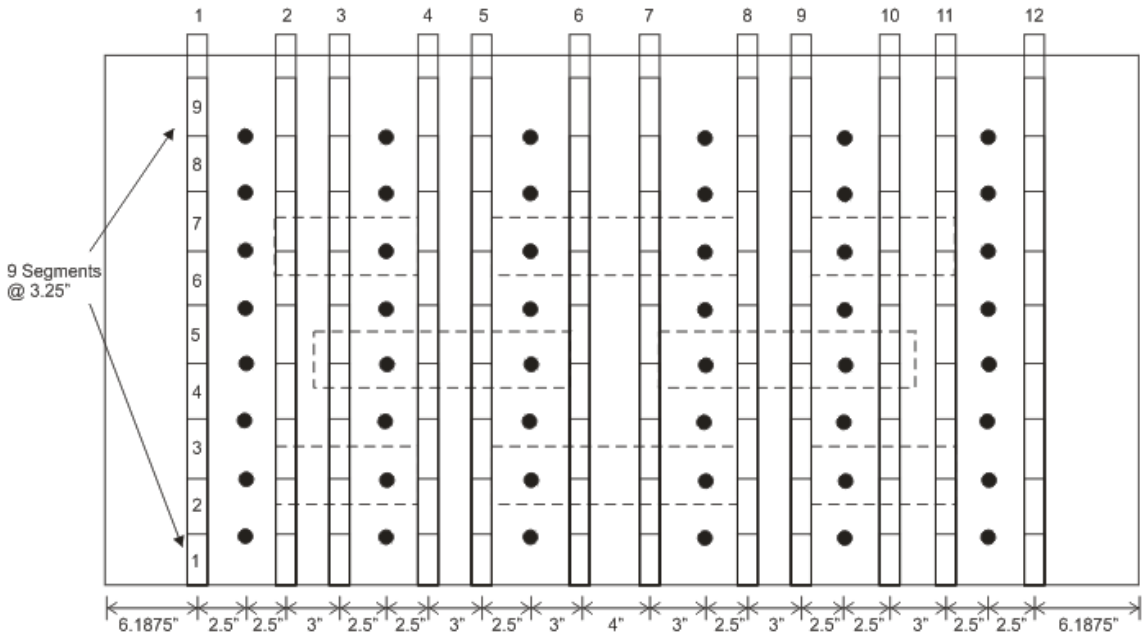


Figure 14.11: Coring scheme applied on SB 1-2.

Figure 14.12: Picture of a completed core sample.

The core samples were collected from the core tubing segments and deposited into pre-labeled, pre-weighed 40 mL vials and capped with Teflon lined septa caps. After this step, thee vials were then weighed again and the mass of the vial containing the cap and sample was recorded. The vials were then filled to the top with Methylene Chloride [MeCl] and weighed again. We then recorded the mass of the vial, along with the cap, sample and the MeCl. The volume of MeCl was calculated based on the mass differences and density of MeCl. The vials were placed on an Innova 2000 orbital shaker at 150 rpm for 24 hours. After shaking, the vials were weighed again and the volume of MeCl was corrected if any mass loss was observed. Approximately 5 mL was extracted from each core sample vial and transferred to three 1.5 mL GC vials.

The GC samples were then analyzed via GC-FID with a 7 level triplicate calibration method consisting of standards of TCE in MeCl at levels of 1, 10,100, 500, 1,000, 5,000 and 10,000 mg/L. The sample TCE concentrations were recorded and the average and standard deviations from each set of 3 samples were calculated. The mass of TCE in each core sample was calculated by multiplying the average sample concentration by the volume of liquid in the sample [assumed to be the volume of MeCl]. The mass was then converted to volume by dividing by the density of TCE. The core samples were then dried thoroughly by baking at 50° C to evaporate the MeCl and then at 105° C for 24 hours to remove any remaining water in the core sample. Total volume of the dried sample was determined by measuring in a 25 mL graduated cylinder with 0.2 mL graduations.

The volume of solids in each sample was determined by placing the sample in a 25 mL graduated cylinder pre-filled with 12 mL of de-ionized water and measuring the displaced volume. The void volume is then calculated by the total volume minus the solid volume. The TCE saturation of each sample was calculated using equation 14.2.

14.4 RESULTS

14.4.1 Batch test results

In order to estimate the saturations of the DNAPL [TCE for SBs 3-1, 3-2, 1-2 and PCE for SB1-3] one must determine the equilibrium partitioning coefficient between the DNAPL and the alcohols involved. More details of these results are provided in Appendix E.

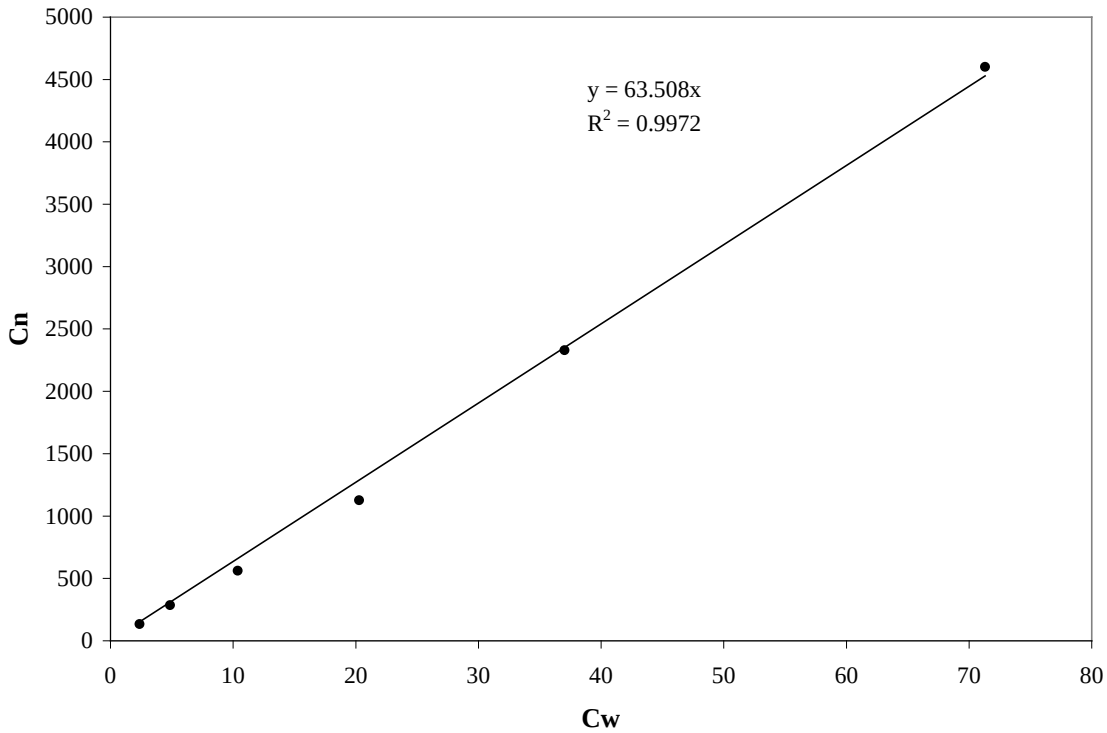


Figure 14.13: Equilibrium partitioning of 2,4-dimethyl-3-pentanol with TCE.

An example of a linear model fit to the equilibrium partitioning of 2,4-dimethyl-3-pentanol with TCE [Trichloroethylene] is shown in Figure 14.13. In this figure each point represents an arithmetic mean of 5 separate batch tests. A linear model fit with a R^2>0.99 provides confidence in these test data. The slope is the equilibrium partitioning coefficient determined from the batch test.

Batch test results for the partitioning tracers utilized and the TCE are summarized in Table 14.1 while partitioning coefficient for PCE with the tracers are summarized in Table 14.2. In both tables, we compare our values to those obtained by others. We notice partitioning coefficients can vary from one laboratory to the next. All the values resulting from our work had linear fits with high correlation coefficients [R^2 > 0.99], ensuring the quality of the values. In Appendix E, we present additional plots and details of these values that were obtained to study the effects of tracer interaction by conducting batch tests for each alcohol on its own, by pairs of alcohols and in a multi-tracer cocktail [three alcohols and PCE].

Table 14.1: TCE - equilibrium partitioning coefficient [K_{NW}].

Alcohol Tracer	Dimensionless TCE-Water K_{NW} values
Isopropyl alcohol [IPA]	0.35^a 0.36^b 0.1^c
2-methyl-1butanol [2M1B]	3.37^a 3.71^b
2-ethyl-1hexanol [2E1H]	221^a 202^b 227^d
2,4-dimethyl-3-pentanol [2, 4DM3P]	63.3^a 71.3^b 38^e

Source: a] this study; b] Dugan et al., [2003]; c] Wang et al. [1998]; d] Istok et al. [2002]; and e] Young et al. [1999].

Table 14.2: PCE - equilibrium partitioning coefficient [K_{NW}].

Alcohol Tracer	Dimensionless PCE-Water K_{NW} values
2-methyl-1butanol [2M1B]	1.3[a]
4-methyl-2pentanol [4M2P]	4.56[a] ;4.22[a]; 4.33[a]; 4.96[a]; 10.3[b]; 3.6[c]; 5.63[d]
1-hexanol [1-Hex]	7.88[a]; 7.72[a]; 8.00[a]; 6.80[a]; 10.00[b]; 8.49[c] ;6.56[c]
2,4-dimethyl-3-pentanol [2, 4DM3P]	28.77[a]; 28.16[a]; 29.64[a]; 27.80[a]; 26.30[a]; 31.6[b] ; 30[f]

Source: a] this study; b] Dugan, P.J. et al. [2007]; c] Cho J. et al. [2003]; d] Moreno-Barbero and Illangasekare [2006]; e] Dwarakanath and Pope [1998]; and e] Ramsburg et al. [2005].

14.4.2 Partitioning tracer test results

The concentration records of both non-partitioning and partitioning tracers obtained from the sampling ports of each PTT displays the history of the solute transport inside the synthetic aquifer. In this section we describe some of the breakthrough curves [BTCs] from the SBs studied. The breakthrough curves show [Figures 14.14-14.16] the propagation of the tracers in terms of Ce pulse inside the SBs, depicting bell shaped curves for both tracers. This behavior was experienced in all partitioning tracer tests. From these figures, it is possible to see a shift between the arrival times of the non-partitioning and partitioning tracer. For example, Figure 14.14 shows the normalized BTCs for partitioning [2M1B] and non-partitioning [Br] tracers in PTT2 at port 12 for SB 3-1. The graph illustrates the retardation of 2M1B relative to bromide. Figure 14.15 shows the BTCs at port 12 for PTT2 in SB 3-2. Here the delay between 2M1B and bromide is more visible than that seen in Figure 14.14. The greater retardation is due to the fact that the TCE volume spilled in SB3-2 was 5 times that spilled in SB3-1.

Figure 14.16 plots the breakthrough curves from the multi-tracer experiment [PTT 5, SB 1-2]. Here, 4 curves are shown. One is for the non-partitioning tracer [bromide] and the others for the suite of partitioning tracers [2M1B, 1-Hex and 2,4DM 3P]. It is possible to see that all three partitioning tracers depict bell shapes, but they have their own characteristics in terms of maximum concentration and tail based on their range of partitioning coefficients.

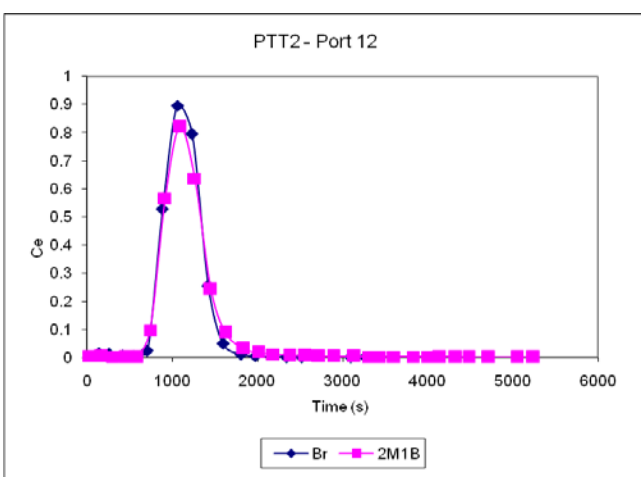


Figure 14.14: Breakthrough curves of Bromide and 2M1B from port 12 in PTT 2 conducted in SB3-1.

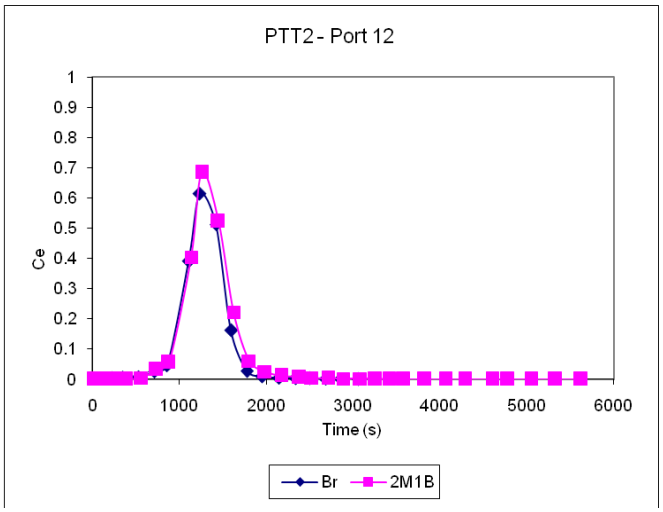


Figure 14.15: Breakthrough curves of Bromide and 2M1B from port 12 in PTT 2 conducted in SB3-2.

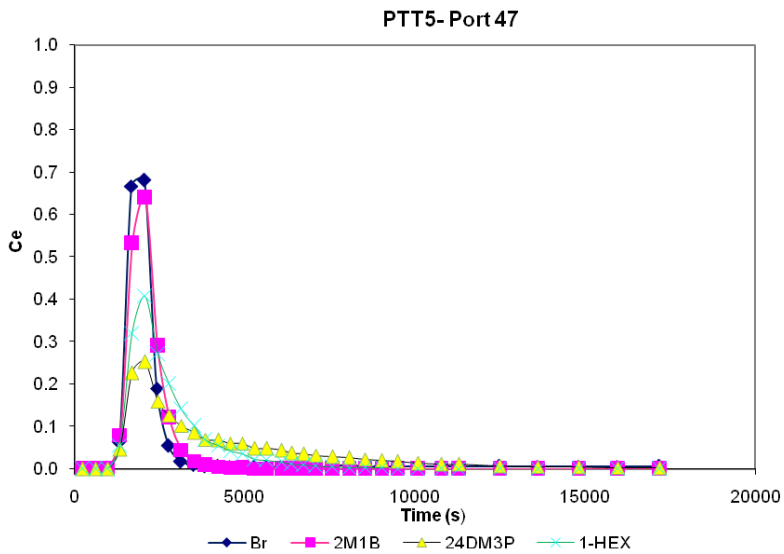


Figure 14.16: Breakthrough curves of Bromide, 2M1B, 2,4DM3P, and 1-Hex from port 47 in PTT 2 conducted in SB1-2.

14.4.3 Temporal Moment Analysis [MOM]-Estimations of S_N and DNAPL volumes

14.4.3.1 Results from SB3-1

As we mentioned earlier, 6 PTTs were carried out in SB3-1. Sandbox 3-1 had a total volume of 9,593 cm^3 and a corresponding void volume of 3,135 cm^3 computed with an average porosity of 0.33 [which is the overall average between the arithmetic mean and geometric mean for fine and coarse sands porosity range reported by Sanders, 1998]. For each test, concentrations of partitioning and non-partitioning tracers were monitored at 12 sampling ports throughout all tests. The data analyzed via MOM for PTT 1-6 of SB 3-1 are presented in Tables 14.3 – 14.8.

For each table we list the port number, estimated retardation coefficient, the TCE saturation estimated at the particular port, and the average TCE volume computed in the sandbox from the data from the port using a weighting scheme that we describe below.

A weighted average assuming that the weights are linear function of the distance from the injection point to the monitoring ports [WDI] is also applied to compute the TCE saturation and volumes per each port studied and a total estimate is computed by adding the contribution of the ports involved on each PTT. The WDI approach is based weights the total distance [TD] from the injection port to the monitoring or sampling ports, d_{I,P_i} ,

$$TD = \sum_{i=1}^{N} d_{I,P_i} \qquad [14.14]$$

while the weights [$weight_{I,P_i}$] are computed using

$$weight_{I,P_i} = \frac{d_{I,P_i}}{TD} \qquad [14.15]$$

The volume of TCE is computed by multiplying the weight of every port times the volume of TCE

$$V_{TCE,i} = weight_{I,P_i} \times S_{N,i} \times V_{voids} \qquad [14.16]$$

We note that the sum of all weights is equal to 1 and that if the distance between the injection to the sampling ports were the same we would be computing the arithmetic mean.

Table 14.3 summarizes the results from PTT1. It reveals that TCE volume detected in each port varies considerably from one port to the next. The potential causes of this behavior will be discussed later.

Table 14.4 summarizes the results from PTT2. It reveals that the average TCE volume detected through PTT2 is lower than PTT1 [for all methods]. The potential causes of this behavior will be discussed later.

Table 14.5 summarizes the results from PTT3. It reveals that the average TCE volume detected through PTT3 is lower than PTT2 [for all methods]. Table 14.5 also shows that there is one negative value of volume [-5.58L] and saturation [-0.018] due to a retardation factor that is lower than 1. Negatives values of retardation [R] are not included in the computations since they are not physically correct. Negative values are likely due to non-equilibrium interaction between the tracer and a DNAPL pool, or the degradation of the tracer.

Table 14.6 summarizes the results from PTT4. It reveals that the average TCE volume detected through PTT4 is approximately the same as PPT3 and the variance [440.98] found to be more than 2 times smaller for this test in comparison to the variance on PTT3.

Table 14.7 summarizes the results from PTT5. It shows that the average TCE volume detected through PTT5 is approximately 42.64 ml which reverses the tendency shown on the first four PTTs, in which the estimated TCE volumes decreased from test to test. The detected TCE volume from PTT 5 is similar to that obtained from PTT1.

Table 14.8 summarizes the results from PTT6. The estimates are again similar to the PTT5. There was an overestimation of the DNAPL saturation and volume in the aquifer [39.4 ml].

From all the results described for SB 3-1 it can be seen that the MOM analysis does not provide an accurate estimate of TCE saturation and the corresponding volume.

Table 14.3: Summary of PTT 1 – sandbox 3-1.

PTT1	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL]- WDI
	2	1.0415	0.0122	38.16	3.94
	3	1.1129	0.0324	101.61	8.68
	4	1.1519	0.0431	135.21	9.02
	6	1.0177	0.0052	16.39	0.84
	7	1.0368	0.0108	33.85	3.98
	12	1.0470	0.0138	43.16	1.66
	13	1.0268	0.0079	24.69	2.81
	19	1.0150	0.0044	13.90	1.55
	25	1.0154	0.0046	14.29	1.59
	26	1.0277	0.0082	25.55	2.29
	27	1.0336	0.0099	30.94	2.11
	28	1.0560	0.0164	51.28	2.19
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0485	0.0141	44.09	40.66
	Geomean	1.0484	0.0108	34.13	Variance
	Variance	0.0018	0.0001	1387.55	7.37

Table 14.4: Summary of PTT 2 – sandbox 3-1.

PTT2	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL]- WDI
	3	1.0074	0.0022	6.84	0.30
	4	1.0275	0.0081	25.40	1.80
	6	1.0107	0.0032	9.96	1.15
	12	1.0741	0.0215	67.49	7.81
	18	1.0956	0.0276	86.45	10.20
	19	1.0151	0.0045	14.01	0.56
	24	1.0819	0.0237	74.39	9.05
	25	1.0035	0.0010	3.21	0.17
	26	1.0183	0.0054	16.94	0.97
	27	1.0190	0.0056	17.57	1.21
	28	1.0207	0.0061	19.12	1.69
	29	1.0138	0.0041	12.78	1.37
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0323	0.0094	29.51	36.27
	Geomean	1.0319	0.0061	19.16	Variance
	Variance	0.0010	0.0001	839.26	13.58

Table 14.5: Summary of PTT 3 – sandbox 3-1.

PTT3	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL]-WDI
	1	1.0063	0.0019	5.85	0.29
	2	1.0201	0.0059	18.56	1.01
	3	1.0127	0.0037	11.74	0.77
	4	1.0059	0.0017	5.48	0.46
	5	1.0154	0.0045	14.24	1.45
	12	1.0132	0.0039	12.19	1.41
	18	1.0055	0.0016	5.14	0.58
	24	1.0117	0.0035	10.83	1.19
	27	1.1289	0.0368	115.48	4.85
	28	1.0245	0.0072	22.64	1.52
	29	0.9940	-0.0018	-5.58	-0.49
	30	1.0109	0.0032	10.06	1.10
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0232	0.0067	21.11	14.63
	Geomean	1.0226	0.0041	12.99	Variance
	Variance	0.0013	0.0001	1008.85	1.67

Table 14.6: Summary of PTT 4 – sandbox 3-1.

PTT4	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	1	1.0095	0.0028	8.83	0.96
	2	1.0121	0.0036	11.21	0.99
	3	1.0943	0.0272	85.31	5.73
	4	1.0326	0.0096	30.01	1.26
	7	1.0188	0.0055	17.37	1.91
	13	1.0202	0.0060	18.67	2.09
	19	1.0273	0.0080	25.19	2.91
	26	1.0283	0.0083	26.15	2.66
	27	1.0234	0.0069	21.62	1.82
	28	1.0119	0.0035	11.07	0.73
	29	1.0072	0.0021	6.71	0.37
	30	1.0138	0.0041	12.80	0.65
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0250	0.0073	22.91	22.06
	Geomean	1.0261	0.0061	19.14	Variance
	Variance	0.0005	0.0000	440.98	2.15

Table 14.7: Summary of PTT 5 – sandbox 3-1.

PTT5	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	3	1.0445	0.0130	40.90	3.59
	4	1.0308	0.0091	28.41	3.19
	8	1.0797	0.0231	72.48	4.19
	10	1.0365	0.0107	33.63	3.47
	11	1.0308	0.0091	28.40	3.64
	14	1.0824	0.0239	74.86	3.28
	15	1.0477	0.0139	43.71	2.86
	16	1.0322	0.0095	29.67	2.85
	17	1.0277	0.0082	25.58	3.14
	20	1.0650	0.0189	59.31	1.94
	21	1.0494	0.0145	45.32	2.66
	22	1.0319	0.0094	29.37	2.69
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0466	0.0136	42.64	37.50
	Geomean	1.0466	0.0127	39.70	Variance
	Variance	0.0004	0.0000	303.59	0.34

Table 14.8: Summary of PTT 6 – sandbox 3-1.

PTT6	Port	R	Sn	V_{TCE} [mL]	V_{TCE} [mL]-WDI
	3	1.0285	0.0084	26.25	1.55
	4	1.0509	0.0149	46.62	4.39
	8	1.0256	0.0075	23.61	0.81
	10	1.0585	0.0171	53.54	5.13
	11	1.0938	0.0271	84.90	10.60
	14	1.0224	0.0066	20.74	0.95
	15	1.0300	0.0088	27.64	1.90
	16	1.0408	0.0120	37.47	3.77
	17	1.0909	0.0263	82.39	10.59
	20	1.0389	0.0114	35.77	2.17
	21	1.0381	0.0112	35.05	2.78
	22	1.0525	0.0153	48.08	5.20
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0473	0.0126	39.40	49.85
	Geomean	1.0491	0.0130	40.89	Variance
	Variance	0.0006	0.0000	453.45	11.32

Examination of Tables 14.3 through 14.8 all show that the TCE saturations and volumes vary considerably from one port to the next as well as from one test to the next. We also computed the means of all ports for each test and present those results in Table 14.9. We see a significant variation of TCE volumes from one test to the next. Table 14.9 also lists the overall mean from all the tests as well as its relative error. From this analysis we find that this average value [33.28 ml] overestimates the original amount spilled [25.00 ml] in SB 3-1 by 8.28 ml resulting in a relative error of 33.11%. Similarly, the overall WDI estimate over-predicted the TCE volume.

The best overall result for SB3-1 was computed with the geometric mean, 25.33 mL [1.30% error].

Table 14.9: Overall results of PPTs in SB 3-1.

	Arith Mean	Geomean	WDI
SB 3-1	V_{TCE} [mL]	V_{TCE} [mL]	V_{TCE} [mL]
PTT1	44.09	34.13	40.66
PTT2	29.51	19.16	36.94
PTT3	21.11	12.99	14.63
PTT4	22.91	19.14	22.06
PTT5	42.64	39.70	37.50
PTT6	39.40	40.89	49.85
Overall TCE Detected [mL]	33.28	25.33	33.49
Volume spilled [mL]	25.00	25.00	25.00
% Relative Error [RE]	33.11	1.30	33.98

In summary we found that, estimation of TCE volume and/or saturations via MOM varies from sampling port to sampling port for each test. The overall average saturations and volumes vary from test to test resulting in under and overestimations of the saturation and volume of TCE.

14.4.3.2 SB 3-2

Four PTTs were conducted in SB 3-2 which had an actual TCE source zone volume of 125 mL. Sandbox 3-2 had a total volume of 7,985.02 cm^3 and a corresponding void volume of 2,635.06 cm^3computed with an average porosity of 0.33 [same as SB 3-1] For each test, concentrations of partitioning and non-partitioning tracers were monitored in 12 ports throughout all tests. Tables 14.10 through 14.13 provide the results of PPT1 - 4 conducted in SB 3-2.

From PTT1 [Table 14.10], the estimates of DNAPL saturation and detection vary from sampling port to port. The DNAPL saturation [S_N] ranges from 0.0033 [0.33%] to 0.119 [11.9%], with an arithmetic mean of 4.38% and a variance of 0.0014. Based on this, the detected volume of TCE ranges from approximately 8.68 mL to 313 mL, with an arithmetic mean of 115 mL.

The injected volume of TCE was 125 ml thus the average volume obtained from PTT1 underestimates the actual introduced volume by 8%. The geometric mean was 74.47 mL [-40.48%].The WDI estimate of 110.7 mL [-12%], was found to be similar to the arithmetic mean value.

Some ports [26 and 28] did not show retardation values larger than 1 and therefore gave negatives values of TCE. The negative saturation and volumes of these ports were not taken in consideration into the overall average and variance.

PTT2 is summarized in Table 14.11. The saturation values vary from 0.0002 [0.02%] to 0.0964 [9.64%] with an arithmetic mean of 2.87% and a variance of 0.0008. from the range of detected volumes varies from 0.45 to 254.14 mL with an arithmetic mean of 75.63 mL, a geometric mean of 36.48 mL and a variance of 5,254.31; The volume estimated by WDI was 68.24 mL with a variance of 17.66. Computations with all methods displayed underestimations of 39.5%, 70.82% and 47.2% respectively.

Table 14.12 summarizes the results for PTT3. The TCE saturations vary from 0.0047 [0.47%] to 0.0435 [4.35%] with an arithmetic mean of 0.0195 and a variance of 0.0002. In terms

of detected volume the range goes from 6.23 to 104 mL with an arithmetic mean of 46.78 mL, a geometric mean of 37.32 ml and a variance of 1179.08. The weighted method reported 52.97 mL and a variance of 16.04. All values underestimate the TCE by 62.6% [arithmetic], 57.62% [WDI], and 70.14% [geometric].

Table 14.13 summarizes the results of PTT4. The DNAPL saturations vary from 0.0007 [0.07%] to 0.0348 [3.48%] with an arithmetic mean of 0.0095 and a variance of 0.0002. In terms of detected volume, the range covers from 1.02 to 91.58 mL with an arithmetic mean of 24.97 mL a geometric mean of 12.23 mL and a variance of 764.49. The IDW approach resulted in 23.99 mL with a variance of 26.37. All methods underestimated the amount of TCE by 80%, 90.22% and 78.90%.

Table 14.10: Summary of PTT 1 – sandbox 3-2.

PTT1	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	6	1.0796	0.0231	60.82	3.06
	5	1.1299	0.0371	97.79	5.34
	4	1.0727	0.0211	55.68	3.65
	3	1.2709	0.0744	196.09	16.47
	2	1.0111	0.0033	8.68	0.88
	7	1.2169	0.0605	159.36	18.41
	13	1.1459	0.0415	109.36	12.25
	19	1.4556	0.1191	313.80	34.48
	25	1.1898	0.0533	140.51	15.34
	26	0.9956	-0.0013	-3.47	-0.31
	27	1.0156	0.0046	12.12	0.81
	28	0.9566	-0.0131	-34.40	-1.44
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.1284	0.0438	115.42	110.70
	Geomean	1.1522	0.0283	74.47	Variance
	Variance	0.0199	0.0014	8598.51	113.65

Table 14.11: Summary of PTT 2 – sandbox 3-2.

PTT2	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	1	1.0721	0.0209	55.20	3.99
	5	1.1145	0.0328	86.56	9.62
	4	1.1704	0.0481	126.83	11.47
	3	1.1851	0.0521	137.17	9.91
	2	1.1041	0.0300	78.95	5.26
	12	1.0509	0.0149	39.18	4.97
	18	1.0203	0.0060	15.74	1.91
	13	1.3597	0.0964	254.14	11.03
	29	1.1193	0.0342	90.10	8.01
	24	1.0145	0.0043	11.32	1.33
	27	1.0006	0.0002	0.45	0.01
	28	1.0153	0.0045	11.88	0.73
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.1022	0.0287	75.63	68.24
	Geomean	1.0982	0.0138	36.48	Variance
	Variance	0.0104	0.0008	5254.31	17.66

Table 14.12: Summary of PTT 3 – sandbox 3-2.

PTT3	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	25	1.0778	0.0226	54.06	3.00
	5	1.0088	0.0026	6.23	0.60
	4	1.0195	0.0058	13.78	1.02
	3	1.0301	0.0088	21.18	0.98
	6	1.0161	0.0047	11.36	1.36
	12	1.0843	0.0244	58.46	7.07
	18	1.0415	0.0122	29.17	3.59
	26	1.1318	0.0376	90.15	5.41
	29	1.1371	0.0391	93.66	10.46
	24	1.1531	0.0435	104.10	13.23
	27	1.0763	0.0222	53.07	3.83
	28	1.0372	0.0109	26.19	2.42
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0678	0.0195	46.78	52.97
	Geomean	1.0667	0.0142	37.32	Variance
	Variance	0.0026	0.0002	1179.08	16.04

Table 14.13: Summary of PTT 4 – sandbox 3-2.

PTT4	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	25	1.0265	0.0078	20.56	0.92
	5	0.9322	-0.0205	-54.10	-8.76
	4	1.0082	0.0024	6.41	0.81
	3	1.0176	0.0052	13.69	1.18
	2	1.0013	0.0004	1.02	0.06
	1	1.0468	0.0137	36.11	1.61
	7	1.0049	0.0014	3.81	0.08
	26	1.0485	0.0142	37.41	2.17
	29	1.1213	0.0348	91.58	14.84
	19	1.0668	0.0194	51.19	1.14
	27	1.0141	0.0042	10.97	0.95
	28	1.0024	0.0007	1.87	0.24
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0242	0.0095	24.97	23.99
	Geomean	1.0320	0.0046	12.23	Variance
	Variance	0.0021	0.0002	764.49	26.37

The values estimated for the PTTs on SB 3-2 are presented in Table 14.14. The arithmetic values ranged from 115.42 to 24.97 mL, the geometric values from 74.47 to 12.23 mL, whereas the WDI vary from 110.70 to 23.99 mL. The overall values for the three methods employed are 65.70 mL, 40.13 mL, and 65.33 mL. An underestimation of the TCE by 47.44% [arithmetic], 67.9% [geometric] and 45.74 % [WDI] was found. From all the results described for SB 3-2 it can be seen that the temporal moment analysis method does not provide an accurate estimate of TCE saturation and the corresponding volume.

In summary we found that:

- Estimation via MOM varies from sampling port to sampling port from individual tests. [Tables 14.10 to 14.13]
- The overall average saturations and volumes vary from test to test [tables 14.10 to 14.13] resulting in underestimations of the saturation and volume of TCE.
- In spite of injecting a TCE volume 5 times larger than in SB 3-1, the overall average value from all PTTs showed a relative errors of from 47.44% to 67.9% compared with the actual volume spilled on SB 3-2 [Table 14.14]. The arithmetic mean was the best estimation method.
- In spite of having a TCE saturation of 5% [5 times the saturation in SB 3-1], the architecture of DNAPL contained large pools and this could potentially affect the performance of the method.

Table 14.14: Overall results of PPTs in SB 3-2.

	Arith Mean	Geomean	WDI
SB 3-2	V_{TCE} [mL]	V_{TCE} [mL]	V_{TCE} [mL]
PTT1	115.42	74.47	110.70
PTT2	75.63	36.48	68.24
PTT3	46.78	37.32	52.97
PTT4	24.97	12.23	23.99
Overall TCE Detected [mL]	65.70	40.13	63.98
Volume spilled [mL]	125.00	125.00	125.00
% Relative Error [RE]	47.44	67.90	48.82

14.4.3.3 SB 1-2

A set of four PTTs were carried out in sandbox 1-2. The first 3 tests of this set, PTTs 1-3, were conducted with bromide [non-partitioning tracer] and 2M1B [partitioning tracer]. For PTT 4, a suite of three partitioning tracers were injected with the same non - partitioning tracers used in this work, bromide [Br]. The partitioning tracers used in these experiments are 2M1B, 2-4-dimethyl-3-pentanol [24DM-3P], and 1-hexanol [HEX].

Sandbox 1-2 had a total volume of 122,227 cm^3 and a corresponding void volume of 46,935.17 cm^3 computed with an experimental average porosity of 0.384 [for this work]. Following are summary tables and illustrations of each PTT from sandbox 1-2 which had an actual TCE source zone volume of 2,000 mL.

Table 14.15 presents the estimations of the saturations and volumes of TCE for PTT1. The saturations of the DNAPL ranged from 0.36% to 3.72% with arithmetic mean of 1.43%, a geometric mean of 1.23% and a variance of 0.0001. The volumes span from 184.01 mL to 1917.12 mL with an arithmetic mean of 736.97 mL and a geometric mean 579.24 mL, experiencing a variance of 296701.62. WDI provided an estimate of 711.81 mL with a variance of 2150.82.

Table 14.16 shows the results for PTT2. The detected DNAPL volume ranged from 111.81 to 1977.75 mL with an arithmetic mean of 825.08 mL and a variance of 296609.42. The TCE saturation reported reveals an arithmetic mean of 0.0176 varying from 0.24% to 4.21%.

As in previous cases, some ports reported retardation values lower than 1 therefore these ports [3 and 45] were not taken into consideration for the computations of the arithmetic mean and WDI. The detected arithmetic mean of TCE increased by almost 100 mL compared with the amount detected on PTT1 [736.97 mL], the geometric mean was 629.61 mL and the WDI estimate was 731.92 mL with a variance of 3042.21. All methods [arithmetic, geometric and WDI] underestimated the actual amount of the TCE by 58.75%, 68.52 % and 63.15%, respectively.

Table 14.17 shows the results of PTT3. The arithmetic mean of TCE volume is 939.48 mL, while the corresponding geometric mean is 160.87 mL, and a WDI value of 710.61 mL. The arithmetic mean saturation of TCE is 1.44% with a range from 0.02% to 11.86% and a variance 0.0019. The geometric mean value for TCE saturation is 0.34%.

The three methods underestimated the original spill by 53.03% [arithmetic], 91.96% [geometric], and 64.47% [WDI]. It is interesting to note is that the low value of the geometric estimation is influenced by the drastic reduction on the positive saturations carried by MOM.

Table 14.18 shows the results of PTT5 based on 2M1B and Br. Here, the arithmetic mean of TCE saturation is 0.0178 with a variance of 0.0001 and an arithmetic mean of TCE volume is 834.48 mL with a variance of 324,481.03.

The geometric mean of TCE volume was 690.29 mL and the corresponding WDI mean value of 843.95 mL. All methods, arithmetic, geometric and WDI, underestimate the injected amount of TCE in 58.28%, 65.47% and 57.80%, respectively.

Table 14.15: Summary of PTT 1 – sandbox 1-2.

PTT1	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	41	1.1435	0.0372	1917.12	167.24
	3	1.0287	0.0077	396.75	20.55
	4	1.0133	0.0036	184.01	13.43
	11	1.0362	0.0097	499.04	44.36
	42	1.0544	0.0145	745.60	79.67
	17	1.0592	0.0157	810.18	70.52
	24	1.0309	0.0082	425.77	45.07
	28	1.1224	0.0319	1645.34	107.75
	35	1.0641	0.0170	875.95	75.24
	47	1.0584	0.0155	799.30	71.09
	45	1.0144	0.0039	200.11	9.42
	32	1.0249	0.0067	344.53	7.47
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0542	0.0143	736.97	711.81
	Geomean	1.0535	0.0123	579.24	Variance
	Variance	0.0017	0.0001	296701.62	2150.82

Table 14.16: Summary of PTT 2 – sandbox 1-2.

PTT2	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	41	1.0262	0.0077	361.90	32.42
	3	0.9902	-0.0029	-136.32	-9.94
	4	1.0080	0.0024	111.81	10.13
	11	1.0152	0.0045	210.14	21.60
	42	1.0844	0.0244	1146.62	126.67
	17	1.0669	0.0195	913.38	90.26
	30	1.0739	0.0215	1007.69	113.94
	28	1.0436	0.0128	599.15	43.63
	35	1.0748	0.0217	1019.33	92.53
	47	1.1483	0.0421	1977.75	176.34
	45	0.9936	-0.0019	-89.08	-3.74
	32	1.0661	0.0192	903.08	24.41
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.0493	0.0176	825.08	731.92
	Geomean	1.0600	0.0134	629.61	Variance
	Variance	0.0021	0.0001	296609.42	3042.21

Table 14.17: Summary of PTT 3 – sandbox 1-2.

PTT3	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	41	1.0336	0.0099	462.77	44.69
	36	1.4535	0.1186	5566.69	617.71
	4	0.9499	-0.0151	-708.91	-45.42
	11	1.0055	0.0016	75.95	6.39
	29	0.9091	-0.0277	-1301.26	-116.71
	17	1.0213	0.0063	295.18	25.18
	24	1.0095	0.0028	132.35	14.06
	33	1.0023	0.0007	32.01	1.80
	35	0.9749	-0.0075	-351.96	-32.68
	47	0.9895	-0.0031	-147.11	-14.83
	45	1.0008	0.0002	11.39	0.78
	32	0.9826	-0.0052	-243.57	-10.87
	Port	R	S_N	V_{TCE} [mL]	Σ VTCE [mL]
	Average	1.0545	0.0144	939.48	710.61
	Geomean	1.0658	0.0034	160.87	Variance
	Variance	0.0280	0.0019	4189239.67	52049.59

Table 14.18: Summary of PTT 5 – sandbox 1-2.

PTT5	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	41	1.0250	0.0074	345.90	30.98
	3	1.0407	0.0119	560.74	40.90
	4	1.0356	0.0104	490.33	44.42
	11	1.0152	0.0045	210.41	21.63
	42	1.0603	0.0176	825.72	91.22
	17	1.1716	0.0484	2273.62	224.68
	30	1.0841	0.0243	1142.73	129.21
	28	1.0401	0.0118	552.58	40.24
	35	1.0783	0.0227	1065.37	96.71
	47	1.0812	0.0235	1104.32	98.46
	45	1.0442	0.0129	607.61	25.49
	32	0.9849	-0.0045	-211.85	-5.73
	Port	R	S_N	V_{TCE} [mL]	Σ V_{TCE} [mL]
	Average	1.0615	0.0178	834.48	843.95
	Geomean	1.0542	0.0147	690.29	Variance
	Variance	0.0019	0.0001	324481.03	3713.93

Finally, the overall amount detected within the 4 tests is presented on Table 14.19. This table shows the arithmetic mean, geometric and WDI TCE estimates for each one of the 4 PTTs analyzed in SB 1-2.

Table 14.19: Overall results of PPTs [2M1B] in SB 1-2.

	Arith Mean	Geomean	IDW
SB 1-2	VTCE [mL]	VTCE [mL]	VTCE [mL]
PTT1	736.97	579.24	711.81
PTT2	825.08	629.61	731.92
PTT3	939.48	160.87	710.61
PTT5	834.48	690.29	843.95
Overall TCE Detected [mL]	834.00	515.00	749.57
Volume spilled [mL]	2000.00	2000.00	2000.00
% Relative Error [RE]	58.30	74.25	62.52

The overall amount is 834.48 mL [arithmetic], 515.00 mL [geometric] and 749.57 mL [WDI]. If we compare this amount with the spilled TCE [2000 mL] we have underestimations of 58.3%, 74.25% and 62.25% with the actual spill.

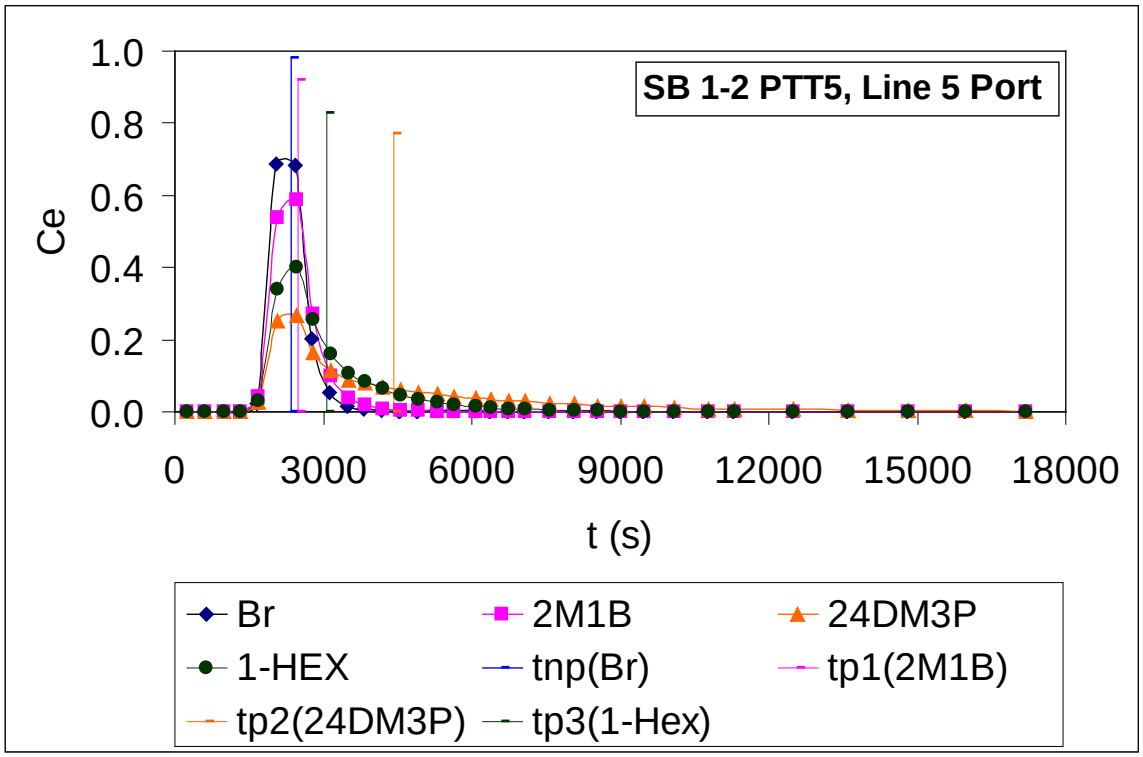


Figure 14.17: Breakthrough curves from the multi-tracer test [PTT5, Port 5].

PTT5 incorporated a suite of three partitioning tracers: 2M1B, 24DM3P, and 1-Hex and a non-partitioning tracer, bromide. Based on figure 14.17, we can see the 3 different estimations using the 3 partitioning tracers. It is possible to observe that every tracer has a different retardation factor, which it is based on the fact that every tracer has a different arrival time [vertical solid lines]. The tracer that reported the largest retardation [compared with the Br-BTC] was 24DM3P.

Tables 14.20 and 14.21 report the estimations based on every pair of tracers by analyzing the data of one PT and the bromide separately. Table 14.20 shows the estimates of S_N and TCE volume based on 24DM3P and Br. Here the amount of TCE detected is more than 3 times lower [269.11 mL] than the arithmetic mean TCE detected by using 2M1B [834.48].

A similar result was obtained by using 1-Hex [Table 14.21], the alcohol with the largest K_{NW} and therefore the greatest retardation. However, non-equilibrium is reflected in the under estimation. The average TCE volume was 332.16 mL which is almost 3 times lower than the amount detected by 2M1B.

The saturations range from 0.003% to 1.57% with the arithmetic mean of 0.71%, a geometric mean 00049.with a variance of 4E-5. The volumes detected via arithmetic, geometric and WDI are 332.16, 231.12 and 325.75 mL with variances of 80,871.76 for the arithmetic and geometric mean approaches and 656.64 for the WDI. All of them underestimate the spill by 83.39 %, 88.44% and 83.71%.

The 1-Hex estimations were calculated using the K_{NW} of Dugan et al [2003] since batch test for 1-Hex and TCE were not conducted for this work. From all the results described for SB 1-2 it can be seen that the temporal moment analysis method does not provide an accurate estimate of TCE saturation and the corresponding volume. In summary we found that:

- Estimation via MOM varies from sampling port to sampling port from individual tests [Tables 14.15 to 14.18].
- The overall average saturations and volumes vary from test to test [Tables 14.15 to 14.18] resulting in underestimations of the saturation and volume of TCE.
- The overall average value from all PTTs, based in bromide and 2M1B, showed a relative error [based on the arithmetic approach] of 58.30% [Table 14.19].
- For PTT5, the estimations based on the three different tracers vary substantially [Table 14.19, 14.20 and 14.21].
- The better results were found for the lowest partitioning tracer [2M1B], Table 14.19 [834 mL], followed by the 1-Hex estimation, the middle partitioning tracer [332.16 mL] ending with the biggest partitioner, 24DM-3P [231.12 mL].
- The cause of this low detection was mainly due to the fact that TCE migrated to the bottom of SB 1-2 and therefore was out of the swept zone of the partitioning tracer test.

Table 14.20: Summary of PTT 5 [2, 4DM3P_Br].

PTT5	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] - WDI
	41	1.2244	0.0035	165.79	14.85
	3	1.2179	0.0034	161.02	11.74
	4	1.2181	0.0034	161.16	14.60
	11	1.0228	0.0004	16.88	1.73
	42	1.8912	0.0139	651.61	71.98
	17	1.0832	0.0013	61.61	6.09
	30	1.2562	0.0040	189.22	21.40
	28	1.2662	0.0042	196.59	14.32
	35	1.3048	0.0048	227.75	20.41
	47	2.1257	0.0175	820.10	73.12
	45	1.4160	0.0065	308.47	12.86
	32	1.0127	0.0002	1.67	0.25
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.3667	0.0057	269.11	263.11
	Geomean	1.3344	0.0038	179.53	Variance
	Variance	0.1143	0.00003	60549.72	609.18

Table 14.21: Summary of PTT 5 [1-HEX_Br].

PTT5	Port	R	S_N	V_{TCE} [mL]	V_{TCE} [mL] – WDI
	41	1.0947	0.0048	226.82	20.55
	3	1.0913	0.0047	218.64	15.95
	4	1.0905	0.0046	216.84	19.65
	11	1.0152	0.0008	36.49	3.31
	42	1.3101	0.0157	734.82	66.57
	17	1.0832	0.0042	199.40	18.07
	30	1.0212	0.0011	50.88	4.61
	28	1.1151	0.0059	275.47	24.96
	35	1.1429	0.0073	341.36	30.93
	47	1.4160	0.0209	980.39	87.41
	45	1.1561	0.0079	372.63	33.76
	32	1.0005	0.0000	1.32	0.12
	Port	R	S_N	V_{TCE} [mL]	ΣV_{TCE} [mL]
	Average	1.1397	0.0071	332.16	325.75
	Geomean	1.1343	0.0049	231.12	Variance
	Variance	0.0146	0.00004	80871.76	656.64

14.4.4 Core analysis-estimations of S_N and DNAPL volumes

14.4.4.1 SB 3-2

TCE saturation estimates were obtained from the cores and summarized in Table 14.22. We then interpolated the local TCE core saturations using the Groundwater Modeling System [GMS] software in an equally spaced grid of 25 x 50 elements. The input file was the saturation

estimates from the core analysis. Figure 14.18 shows a kriged estimate of TCE saturation in SB 3-2 whereas Figure 14.19 shows a linear interpolation.

By inspecting the results, of the TCE saturations we could see that the values along the coring array are very low giving an arithmetic mean saturation of 0.00087. The values reported were introduced in the GMS file in order to conduct a clearer spatial distribution. Figures 14.18 and 14.19 show how the TCE in SB 3-2 is located mostly in the middle of the SB, where the fine lens of F-75 is located. Visually, the coring results coincide with the photographs as far as the location of the highest concentrations of TCE. However, the saturations estimated from coring are very low.

Even if our TCE estimations were too low to make a comparison with the assessed amount via PTTs, the information gathered gave an important insight on how the DNAPL was distributed inside the porous media.

Table 14.22: Experimental values from core analysis - SB 3-2.

Core	x [in]	y [in]	TCE S_N	Core	x [in]	y [in.]	TCE S_N
1_1	5	3.5	1.1646E-03	4_1	21	3.5	2.3061E-06
1_2	5	6.75	5.1777E-05	4_2	21	6.75	1.2630E-03
1_3	5	10	3.2545E-06	4_3	21	10	8.1497E-06
1_4	5	13.25	3.6038E-06	4_4	21	13.25	1.4596E-05
1_5	5	16.5	1.1180E-05	4_5	21	16.5	1.3910E-05
2_1	10	3.5	1.8786E-06	5_1	26	3.5	6.2327E-06
2_2	10	6.75	1.0859E-02	5_2	26	6.75	2.8763E-04
2_3	10	10	1.7998E-06	5_3	26	10	1.8142E-06
2_4	10	13.25	1.5805E-07	5_4	26	13.25	1.7264E-06
2_5	10	16.5	7.7776E-06	5_5	26	16.5	1.4064E-05
3_1	15	3.5	2.9508E-06	6_1	31	3.5	6.2803E-06
3_2	15	6.75	1.2408E-02	6_2	31	6.75	1.4805E-05
3_3	15	10	1.9712E-05	6_3	31	10	2.3784E-06
3_4	15	13.25	3.1509E-06	6_4	31	13.25	6.0959E-06
3_5	15	16.5	3.0145E-06	6_5	31	16.5	1.2669E-05

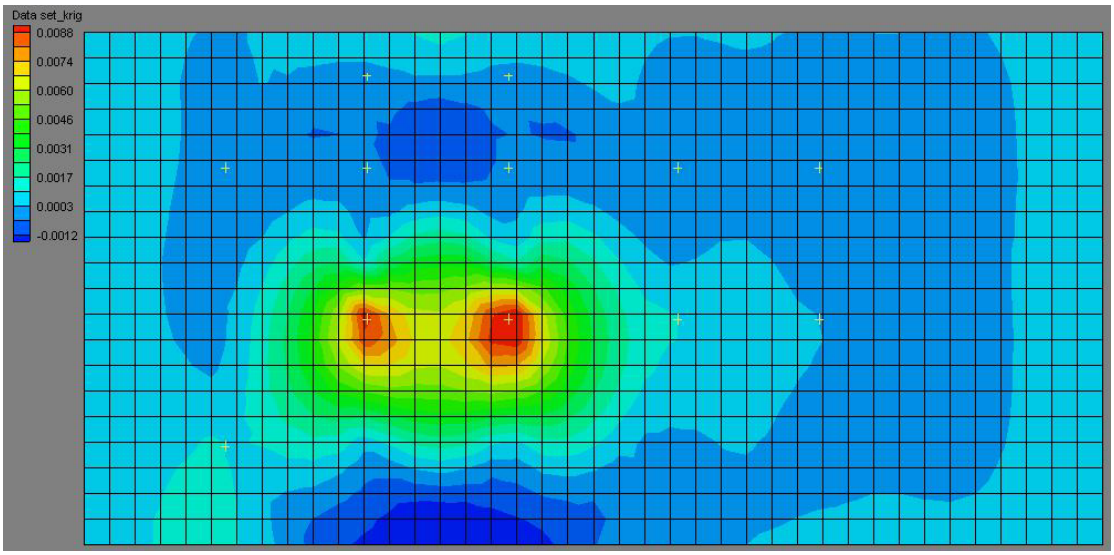


Figure 14.18: Kriged core TCE saturation estimates using GMS.

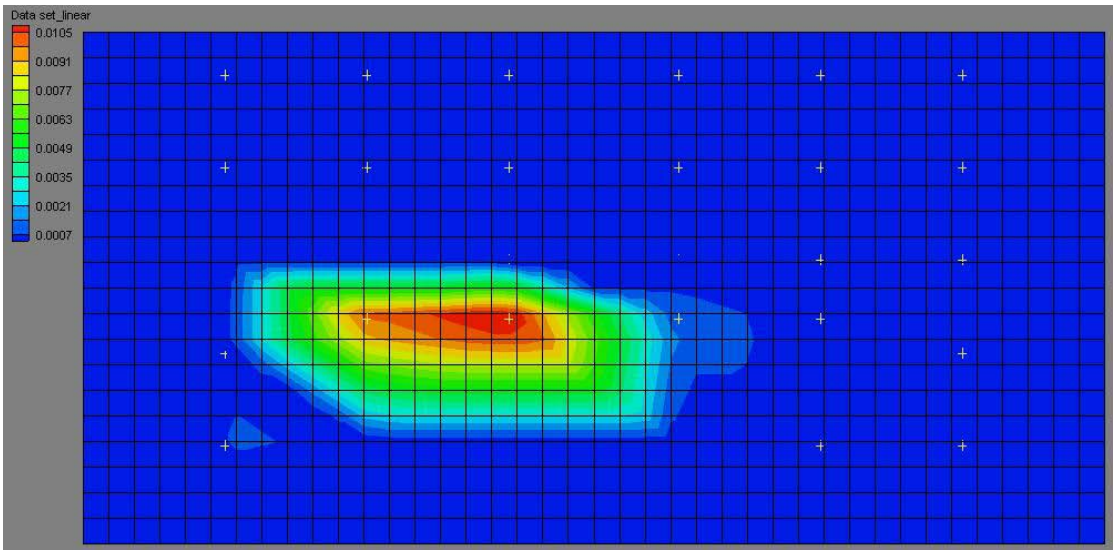


Figure 14.19: Linear interpolation of core TCE saturation estimates using GMS.

14.4.4.2 SB 1-2

TCE saturations were estimated from the cores [Table 14.23] and interpolated contour plots of the TCE core saturations on SB 3-2 were created with GMS in an equally spaced grid of 25x50 elements. In the same spirit of the previous section Figures 14.20 [Inverse distance weighted interpolation of the TCE saturation] and 14.21 [linear interpolation of TCE saturation] shows the location of the mass of TCE in the box. By inspecting the TCE mass counter plots generated for SB 1-2, it is possible to see that the TCE migrated to the bottom of SB 1-2. A photograph of the TCE architecture in SB 1-2 is shown in Figure 14.22.

The average saturation from the interpolation based on the core samples is 0.0028 which is lower than the saturation estimate from PTTs. More work has to be developed refining the coring technique for future work. Regardless, these results collectively suggest that high resolution coring, even if the core holes are relatively closely spaced and the standard interpolation of those results could lead to severe underestimations of DNAPL saturations.

Table 14.23: Experimental values from core analysis from SB 1-2.

Core	x [cm]	z [cm]	TCE S_N	Core	x [cm]	z [cm]	TCE S_N
1 2	15.71625	12.3825	3.18E-06	7 2	80.48625	12.3825	0.000217
1 3	15.71625	20.6375	1.11E-06	7 3	80.48625	20.6375	0.006331
1 4	15.71625	28.8925	1.11E-06	7 4	80.48625	28.8925	0.00128
1 5	15.71625	37.1475	1.22E-06	7 7	80.48625	53.6575	6.63E-06
1 6	15.71625	45.4025	1.00E-06	7 8	80.48625	61.9125	1.14E-05
2 1	28.41625	4.1275	2.25E-05	8 1	94.45625	4.1275	0.012636
2 2	28.41625	12.3825	6.95E-07	8 2	94.45625	12.3825	3.85E-05
2 3	28.41625	20.6375	8.58E-07	8 3	94.45625	20.6375	3.20E-05
2 4	28.41625	28.8925	6.73E-07	8 4	94.45625	28.8925	4.16E-05
2 5	28.41625	37.1475	5.27E-07	8 5	94.45625	37.1475	0.007091
2 6	28.41625	45.4025	6.36E-07	8 6	94.45625	45.4025	1.15E-05
3 1	36.03625	4.1275	5.16E-05	9 1	102.0763	4.1275	0.019093
3 2	36.03625	12.3825	6.11E-06	9 2	102.0763	12.3825	0.003568
3 3	36.03625	20.6375	1.68E-06	9 3	102.0763	20.6375	2.47E-06
3 4	36.03625	28.8925	7.76E-07	9 4	102.0763	28.8925	0.000105
3 5	36.03625	37.1475	1.44E-06	9 5	102.0763	37.1475	3.02E-05
3 6	36.03625	45.4025	1.28E-06	9 6	102.0763	45.4025	3.68E-06
4 1	48.73625	4.1275	0.013373	9 7	102.0763	53.6575	2.39E-06
4 2	48.73625	12.3825	2.31E-06	10 1	114.7763	4.1275	0.006814
4 3	48.73625	20.6375	1.44E-06	10 2	114.7763	12.3825	0.000489
4 4	48.73625	28.8925	8.44E-07	10 3	114.7763	20.6375	0.003702
4 5	48.73625	37.1475	1.23E-06	10 4	114.7763	28.8925	3.08E-05
4 6	48.73625	45.4025	2.19E-06	10 5	114.7763	37.1475	1.11E-05
4 7	48.73625	53.6575	1.46E-06	10 6	114.7763	45.4025	1.24E-05
5 1	56.35625	4.1275	0.005145	11 1	122.3963	4.1275	0.074713
5 2	56.35625	12.3825	3.84E-05	11 2	122.3963	12.3825	0.008387
5 3	56.35625	20.6375	4.41E-06	11 3	122.3963	20.6375	0.010679
5 4	56.35625	28.8925	3.96E-06	11 4	122.3963	28.8925	1.37E-05
5 5	56.35625	37.1475	3.11E-06	11 5	122.3963	37.1475	7.49E-06
5 6	56.35625	45.4025	3.56E-06	11 6	122.3963	45.4025	3.05E-06
6 1	70.32625	4.1275	0.004832	12 1	135.0963	4.1275	0.01644
6 2	70.32625	12.3825	1.99E-05	12 2	135.0963	12.3825	3.18E-05
6 3	70.32625	20.6375	0.000637	12 3	135.0963	20.6375	4.01E-06
6 4	70.32625	28.8925	5.61E-06	12 4	135.0963	28.8925	2.12E-05
6 5	70.32625	37.1475	3.60E-06	12 5	135.0963	37.1475	2.46E-05
6 6	70.32625	45.4025	0.000112	12 6	135.0963	45.4025	1.02E-05
6 7	70.32625	53.6575	2.12E-06				

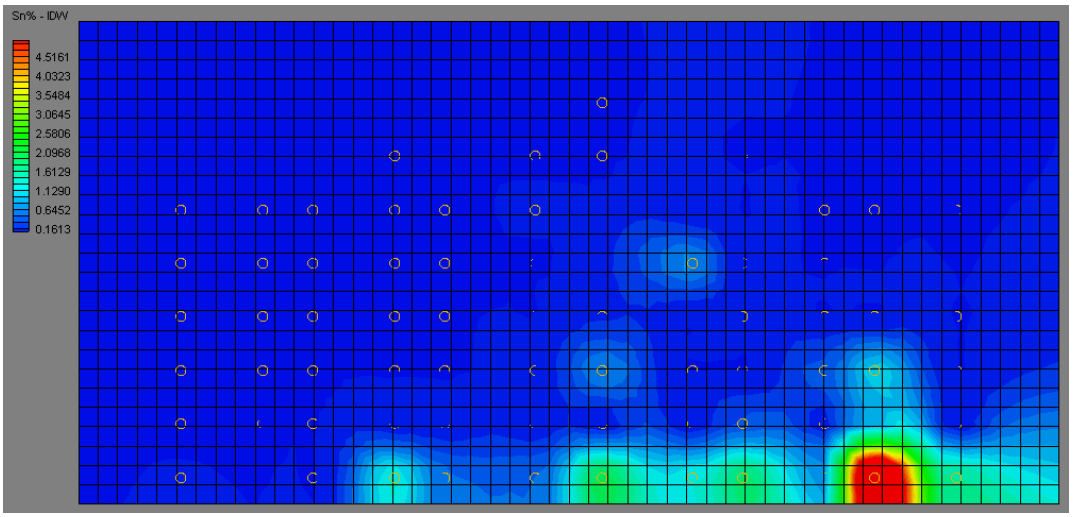


Figure 14.20: Inverse distance weighted interpolation of core TCE saturation values.

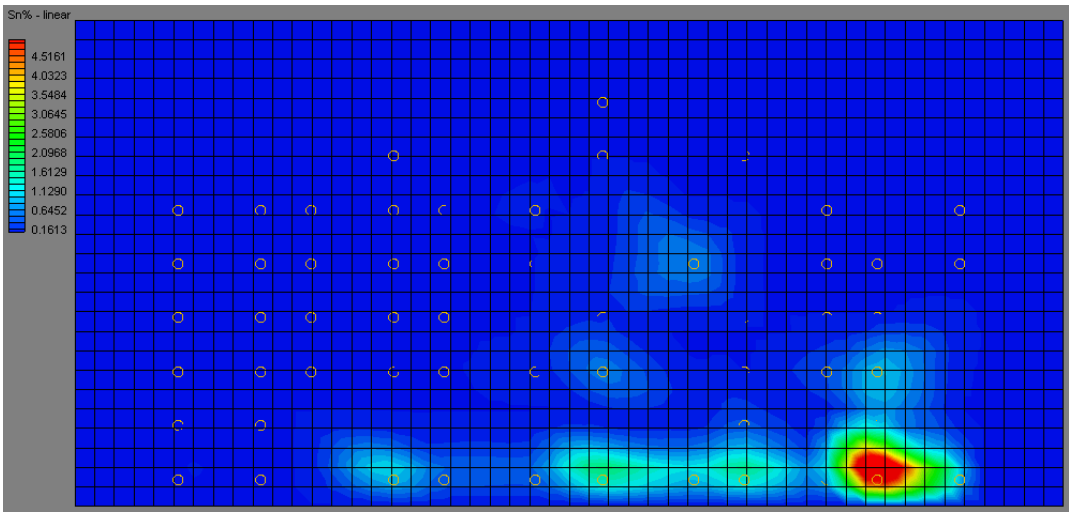


Figure 14.21: Linear interpolation of core TCE saturation values.

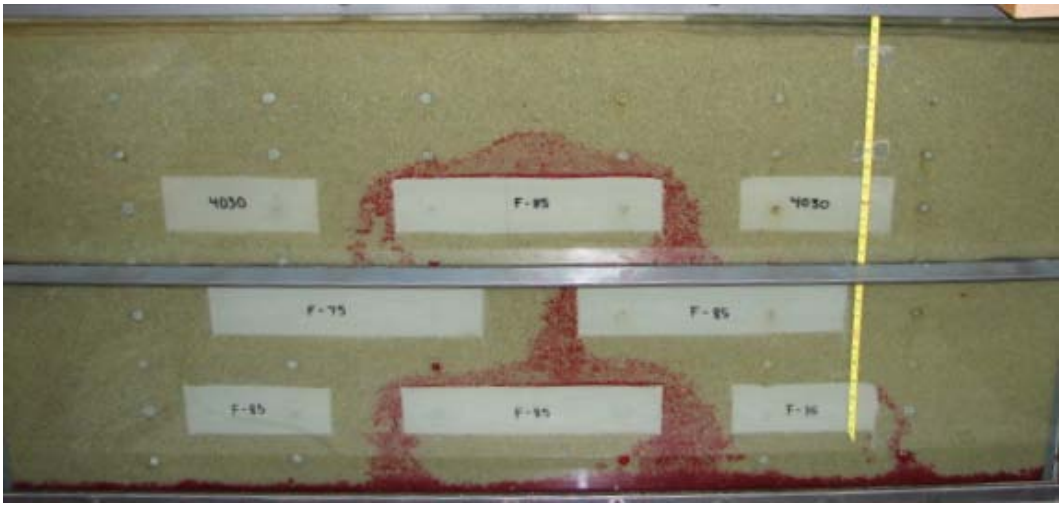


Figure 14.22: Picture of the TCE distribution in SB 1-2.

14.5 DISCUSSION

14.5.1 Influence of partitioning coefficient on saturation estimates

The estimates from the multi tracer experiment in SB 1-2 resulted in different TCE saturation estimates. Similar results were observed by Moreno-Barbero et al. [2006], Brooks et al. [2002] and Imhoff et al. [2003]. These researchers found that smaller partitioning coefficient yielded more accurate results.

This is shown in the reported arithmetic TCE volumes: 834.48 ml [2M1B-TCE] for the lowest K_{NW}, 332.16 ml for the middle partitioning alcohol [1-Hex-TCE] and the 263.11 ml [24DM-3P] for the highest partitioning tracer detected. Intuitively, we would think that with an increase in equilibrium partitioning coefficient [K_{NW}], the saturation estimates would be closer to the actual amount since the retardation is proportional to K_{NW}. However, this is not the case under non-equilibrium conditions. The MOM approach assumes an equilibrium partitioning coefficient.

At early times, when the tracers are injected, the partitioning coefficients are lower than the values obtained for the equilibrium batch tests exhibiting a kinetic behavior. Figure 14.23 [after Moreno-Barbero et al, 2006], shows how at early times, the tracer with the highest partitioning coefficient yields very low values compared with its equilibrium value. However, with time the partitioning coefficient approaches the equilibrium value. We also see that this behavior is accentuated when K_{NW} is larger.

The use of the equilibrium coefficient could introduce significant errors in the mass estimation when the partitioning coefficient is high and the contact time between the tracer pulse and DNAPL is less than the time required to achieve equilibrium partitioning. In general, the results of the MOM in terms of the detected DNAPL volume [with the exception of the overall results on SB 3-1], were low.

Table 14.24 reports the partitioning coefficients [K_{NW}] computed using equation 14.2 from the average values of the TCE saturations [S_N] and the retardation [R] obtained on every PTT conducted in the three SBs studied. The results give us an indication of whether equilibrium conditions are achieved or not for each tracer test. For SB 3-1, reported values were very similar to the equilibrium coefficient for 2M1B. For SB 3-2, the K_{NW} values vary more than those determined from SB 3-1. However the K_{NW} values from the PTT are close to the equilibrium value of 3.71. For SB 1-2, the results do not vary as much as SB 3-2 for 2M1B. Likewise, the 1-Hex tracer showed an average experimental partitioning coefficient [19.87] that is close to the equilibrium [19.5] value reported in Dugan et al. [2003]. In contrast the partitioning tracer with the highest K_{NW}, 24DM3P, revealed the largest discrepancy [63.4] from its equilibrium value [71.3], which signifies that the tracer partitioning did not reach equilibrium conditions. These computations are not conclusive but imply the need for further research.

Table 14.24: K_{NW}-PTT vs. K_{NW} equilibrium.

Tracer	2M1B	
SB 3-1	K_{NW} -PTT	K_{NW} equilibrium
PTT1	3.84	3.71
PTT2	3.80	
PTT3	3.81	
PTT4	3.78	
PTT5	3.81	
PTT6	4.19	
SB 3-2	K_{NW} -PTT	
PTT1	3.33	
PTT2	4.00	
PTT3	3.88	
PTT4	2.83	
SB 1-2	K_{NW} -PTT	
PTT1	3.51	
PTT2	3.52	
PTT3	3.83	
PTT5	3.52	
Tracer	1-Hex	
SB 1-2	K_{NW} -PTT	K_{NW} equilibrium
PTT5	19.87	19.5
Tracer	24DM3P	
SB 1-2	K_{NW} -PTT	K_{NW} equilibrium
PTT5	64.31	71.3

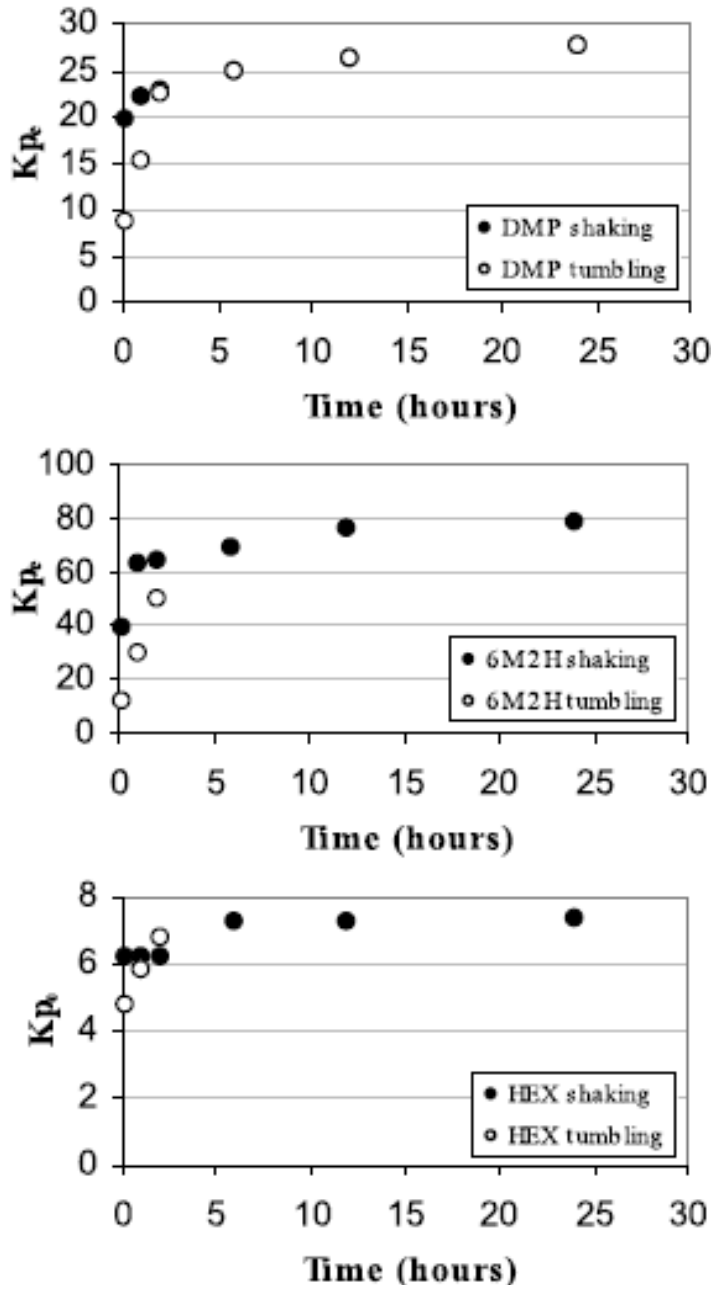


Figure 14.23: Effective partition coefficient [Kpe] of tracers with DNAPL [PCE].

14.4.2 Influence of DNAPL architecture on saturation estimates

14.4.2.1 SB 3-1

The architecture of the spill plays an important role on saturation estimates. For example, for SB 3-1 the spill had two important pathways, one near the middle of the SB and the other at the right half of the SB. This distribution could help to achieve a greater interaction between the tracers and the DNAPL when changing the flow directions from test to test. SB3-1 had the lowest amount of TCE released but it resulted in a residual phase source zone. According to Moreno-Barbero et al. [2006], the PTT approach can determine TCE at residual saturation but

does not provide reliable estimates when it exists as pools. This is likely the reason of the high overall saturation estimates with the three approaches [arithmetic, geometric and IDW]. In this sandbox, the geometric mean provided the best estimate of DNAPL volume.

14.4.2.2 SB 3-2

A more uniform spill was achieved but the formation of large pools jeopardized the results. This also was found by other researchers. Rao et al. [2000] and Moreno-Barbero et al. [2006] mentioned that DNAPL pools compromises the reliability of the PTTs and this is reflected in the outcome of the MOM estimates. DNAPL pools are not homogeneous. These arrangements contain various zones that have associated diverse saturations [Moreno-Barbero et al. 2006]. Therefore, our estimates can severely under or overestimate the DNAPL volume in the majority of the cases that we have examined. Pools have high saturations on their lower portions, and the tracer interaction is limited by the interface thickness between the DNAPL and the tracer. This results in the tracer traveling on the outer portions of the DNAPL [Moreno-Barbero et al. 2006].

14.4.2.3 SB 1-2

In this case, the TCE reached the bottom of the SB and the DNAPL was undetectable because it resided outside any possible sampling intervals [below the bottom row of ports]. The 12/20 sand triggered the migration to the bottom of the SB. The consequence was that a significant amount of TCE was not accounted for with the MOM estimates. Figure 14.24 shows how the TCE penetrated the silicone seals of SB 1-2. This picture was taken when we were refurbishing SB 1-2 in order to create SB 1-3. This also indicates that TCE mass losses occurred in SB 1-2. Some TCE was trapped in the silicone.

Figure 14.24: Penetration of dyed TCE into the silicone seals of SB 1-2.

14.4.3. Influence of flow and sampling on saturation estimates

The variability of the estimates between ports can be explained by the diverse flow paths experienced by the tracers within the porous media. Since we have low hydraulic conductivity blocks inside our porous media, the flow can be deflected or by-passed by the tracers resulting in some ports having preferential locations for sample collection.

When the sampling and flowing scheme were repeated, PTT2 and PTT5 in SB 1-2, the experimental estimates did vary in less than 2% which tells us: 1] that the detection was repeatable and 2] there were not important TCE mass losses. The sampling scheme [Nelson et al., 1999] is critical and the experience gained in these tests was of incredible value.

14.4.4 Influence of TCE mass loss due to dissolution on saturation estimates

In order to estimate the TCE mass loss during each PTT conducted, a simplified computation was carried out by assuming that the TCE [and its concentration] is uniformly distributed within the synthetic aquifer. The spilled TCE mass [volume spilled times the TCE density] for each test was divided by the void volume of the synthetic aquifer in order to compute the TCE concentration. The TCE mass loss is then computed as:

$$TCEloss = \frac{C_{TCE} * t * Q}{\rho_{TCE}} \qquad [14.17]$$

where C_{TCE} is the TCE concentration, t is the length of each PTT, Q is pumping rate, ρ_{TCE} is TCE density. By summing the TCE mass from one PTT to the next, one can then compute the cumulative TCE loss as:

$$TCE_{CummulativeLoss} = TCE_{initial} - \sum_i^n TCEloss_i \qquad [14.18]$$

where $TCEloss_i$ is computed using equation 14.17. Table 14.25 summarizes the results of these calculations. It reveals that TCE mass loss for PTT in SB3-1 was on the order of 2 mL, while for SB 3-2 it was 7.5 mL and for SB 1-2, it was 15 mL.

Table 14.25: TCE mass loss computations.

SB 3-1			
PTT	time[s]	*TCEloss* [mL] per PTT	Initial TCE [mL] per PTT
1	6000	0.39868	25.00
2	6000	0.39868	24.60
3	6000	0.39868	24.20
4	6000	0.39868	23.80
5	5000	0.33224	23.41
6	5000		23.07
SB 3-2			
PTT	time[s]	*TCEloss* [mL] per PTT	Initial TCE [mL] per PTT
1	6400	2.52999	125.00
2	6200	2.45093	122.47
3	6400	2.52999	120.02
4	6400		117.49
SB 1-2			
PTT	time[s]	*TCEloss* [mL] per PTT	Initial TCE [mL] per PTT
1	14000	4.85755	2000.00
2	14000	4.85755	1995.14
3	14000	4.85755	1990.28
4	18000		1985.43

15. HYDRAULIC AND PARTITIONING TRACER TOMOGRAPHY FOR TRICHLOROETHYLENE SOURCE ZONE CHARACTERIZATION: SMALL-SCALE SANDBOX EXPERIMENTS

15.1. INTRODUCTION

Dense Nonaqueous Phase Liquids [DNAPLs] are prevalent at a large number of sites throughout the world. The high densities, low interfacial tensions, and low viscosities of halogenated solvents can lead to deep DNAPL penetration [*Pankow and Cherry*, 1996]. In porous media, much of the DNAPL mass remains in the groundwater as persistent source zones. The variable release history, unstable flow, and geologic heterogeneity make the distribution of DNAPL in the source zone complex, where DNAPLs exist as residuals or as pools of pure phase. Without remediation, these source zones can contribute to long-term groundwater contamination for decades to centuries. The spatial distribution of DNAPLs in the source zone needs to be characterized so that efficient active or passive remediation schemes [*Alvarez and Illman*, 2006; *Illman and Alvarez*, 2009] can be designed.

To determine the extent of DNAPL contamination, the most common methods are the use of groundwater monitoring wells and soil sampling [*Pankow and Cherry*, 1996; *Cohen and Mercer*, 1993]. The use of fully screened monitoring wells, however, may only imply the likely presence of DNAPLs because groundwater contaminant concentrations near DNAPLs are usually less than their aqueous solubility. Furthermore, the usage of fully screened monitoring wells can result in errors on orders of magnitude in assessing DNAPL distributions because the measurement is integrated over the length of the screened interval.

One alternative is to directly sample groundwater and soil materials by obtaining cores using split spoon samplers or direct push methods [*Kram et al*., 2001]. If a sufficient number of samples are collected, the approach could potentially produce high resolution records of DNAPL concentrations from groundwater and soil samples down each borehole. The organic contaminants are commonly extracted from the soil with organic solvents such as methylene chloride and the contaminants present are then determined with a gas chromatograph. If a sufficient number of samples are available in multiple boreholes, a three-dimensional distribution of DNAPL saturations can then be obtained through various interpolation techniques. However, applications of the method to real-world problems are limited because the large number of samples is cost-prohibitive.

Other techniques to characterize DNAPL source zones include visual inspections of drill cores, soil vapor analysis, geophysical surveys, use of radon abundance data, and partitioning tracer tests [*Kram et al*., 2001]. In particular, partitioning tracer tests have been used as a tool to provide a spatially integrated measure of residual DNAPL volume in the flow without causing disturbances to the source zone domain [*Jin et al*., 1995; *Nelson and Brusseau*, 1996; *Annable et al*., 1998]. Partitioning tracer tests are appealing from the standpoint of minimizing drilling in the source zone because the tracer injection and samples collection wells can be located outside of the DNAPL source zone. These tests are performed through the comparison of the transport behavior of conservative and partitioning tracers injected across a suspected DNAPL source zone.

The partitioning tracer method relies on the reaction of tracers with DNAPLs. In particular, transport of tracers through the source zone causes partitioning into and out of

DNAPLs which results in the retardation of tracer arrival at monitoring wells in comparison to conservative tracers. The resulting breakthrough curves are then typically analyzed using a temporal moment method to compute the average DNAPL saturation. Despite the success of the approach demonstrated through several field studies [*Nelson and Brusseau*, 1996; *Annable et al*., 1998; *Young et al*., 1999; *Cain et al*., 2000; *Brooks et al*., 2002], the temporal moment approach is not capable of obtaining information on the distribution of DNAPLs because the approach does not consider the flow paths of the tracer. Despite its shortcomings, the second and third temporal moments of the breakthrough curves can be used to estimate the statistical parameters characterizing the DNAPL spatial distribution [*Jawitz et al*., 2003]. Recent research also suggests that the partitioning tracer tests may be of limited utility in detecting high saturation zones [i.e., pools], and can be affected by nonequilibrium effects if the partitioning tracer utilized has a high partitioning coefficient [*Moreno-Barbero and Illangasekare*, 2006].

To our knowledge, there are only few studies that have shown the estimation of DNAPL saturation distributions through interpretation of partitioning tracer tests. For example, a stochastic method was developed to estimate spatial distribution of NAPL residual content from tracer breakthrough curve moments [*James et al*., 2000], and an inverse method was developed to locate DNAPL pools under steady flow conditions [*Sciortino et al*., 2000]. A distributed-parameter extended Kalman filter approach was then developed for estimating spatially distributed residual saturation of NAPL and Darcy flux to predict the movement of a partitioning tracer plume in a three-dimensional heterogeneous aquifer [*Zhang and Graham*, 2001]. Moreover, a streamline-based inverse method for analyzing partitioning tracer tests was developed to estimate the three dimensional spatial variation of NAPL saturation through the analogy between streamlines and seismic ray tracing [*Datta-Gupta et al*., 2002]. Most recently, an analytical-numerical inverse modeling approach that utilizes concentration profiles obtained from a TCE plume to predict the distribution of the TCE in a vertical plane within the source zone has been developed [*Dridi et al*., 2009]. As our literature survey shows, interpretive techniques developed are promising in delineating residual DNAPL distributions in the subsurface. However, existing techniques have not incorporated potentially useful hydraulic head data during the test and prior knowledge of hydraulic heterogeneity.

Recently, the Hydraulic/Partitioning Tracer Tomography [HPTT] was developed to characterize DNAPL source zones [*Yeh and Zhu*, 2007]. The HPTT technology relies on the fusion of data from hydraulic tomography [HT] and partitioning tracer tomography [PTT]. HT is similar in concept to the Computerized Axial Tomography [CAT] scan technology, but the energy source is a change [decrease or increase] in hydraulic head achieved through water pumping or injection, and the sensors [pressure transducers] detect the arrival of head change at multiple locations throughout the source zone. Hydraulic head monitored at numerous observation ports are then used to quantify the spatial distribution of hydraulic parameters, their connectivity, and uncertainties associated with the Sequential Successive Linear Estimator [SSLE] algorithm developed for steady state [SSHT] and transient hydraulic tomography [THT]. To date, the technique has been tested using synthetic simulations [*Yeh and Liu*, 2000; *Zhu and Yeh*, 2005; *Zhu and Yeh*, 2006; *Ni et al*., 2009], laboratory sandbox experiments [*Liu et al*., 2002; *Illman et al*., 2007; 2008; *Liu et al*., 2007; *Yin and Illman*, 2009; *Illman et al*., 2010] and in the field [*Straface et al*., 2007; *Illman et al*., 2009; *Berg and Illman*, 2009].

PTT is similar in concept to HT, but the source is a partitioning tracer that is introduced at some injection point and the signal is the corresponding breakthrough curve that is obtained at multiple sampling points. Additional partitioning tracer tests are then sequentially conducted by

injecting the tracers at other locations and the corresponding breakthrough curves are obtained at various sampling points. Because the partitioning tracers sweep through the DNAPL source zone from several different directions, there is an opportunity to see the source zone from different directions yielding additional information on saturation distributions that cannot be obtained from a single partitioning tracer test or through its interpretation using a temporal moment method.

The main advantage of conducting HT first and then conducting PTT is that the *K* heterogeneity can first be captured through HT. This then allows one to conduct a partitioning tracer test that can be better controlled hydraulically leading to an improved sweep of the DNAPLs in the source zone. In particular synthetic simulations have shown [*Yeh and Zhu*, 2007] that the accurate knowledge of *K* distribution within the source zone is critical in obtaining an accurate estimate of DNAPL saturation distribution. In addition, when the partitioning tracer tests are conducted in a tomographic manner, more detailed information on DNAPL saturations can be obtained in comparison to a traditional partitioning tracer test.

Our ultimate goal is to validate the HPTT approach in the field to image DNAPL source zones for hydraulic heterogeneity and DNAPL saturations as well as their uncertainties. However, it is necessary that the approach is first tested in the laboratory in which the source zone characteristics are well known and experimental conditions can be fully controlled. Therefore, the main goal of this study is to independently evaluate the performance of the HPTT algorithm [*Yeh and Zhu*, 2007] in a synthetic aquifer constructed in a laboratory sandbox. Specific objectives of the study include imaging of *K* heterogeneity of a synthetic sandbox aquifer through hydraulic tomography and to characterize the TCE source zone through HPTT and direct sampling of cores. We compare the *K* and TCE saturation distributions [or *K* and TCE saturation tomograms from now on] to the actual position of the TCE source zone visually obtained and to results from geostatistical analysis of core samples.

15.2 CONSTRUCTION OF HETEROGENEOUS AQUIFER AND ITS HYDRAULIC CHARACTERIZATION

15.2.1 Sandbox design and construction

A synthetic TCE source zone was constructed in a small sandbox to validate the hydraulic and partitioning tracer tomography algorithm. The sandbox is 91.44 cm in length, 61 cm in height, and has a thickness of 1.9 cm. The height of the sand in the tank is 45.7 cm. The small thickness of the sandbox was to enable the visual comparison of dyed TCE spills to resulting TCE saturation tomograms. The sandbox was constructed with ¾” thick Plexiglas. The sandbox has solid “no-flow” boundaries on the left and right ends and at the bottom. The top boundary is open to the atmosphere in which water was allowed to pond on top of the sand.

The sandbox was equipped with 30 water sampling and pressure monitoring ports that could also be utilized for water injection and extraction [Figure 15.1]. Unlike previous sandboxes built by our group [*Illman et al*., 2007; 2010], there were no wells completed in the synthetic aquifer. Instead, a circular piece of stainless steel mesh was placed into a recessed circle on the inside wall covering the port hole. The screen held the sand in place and allowed water flow.

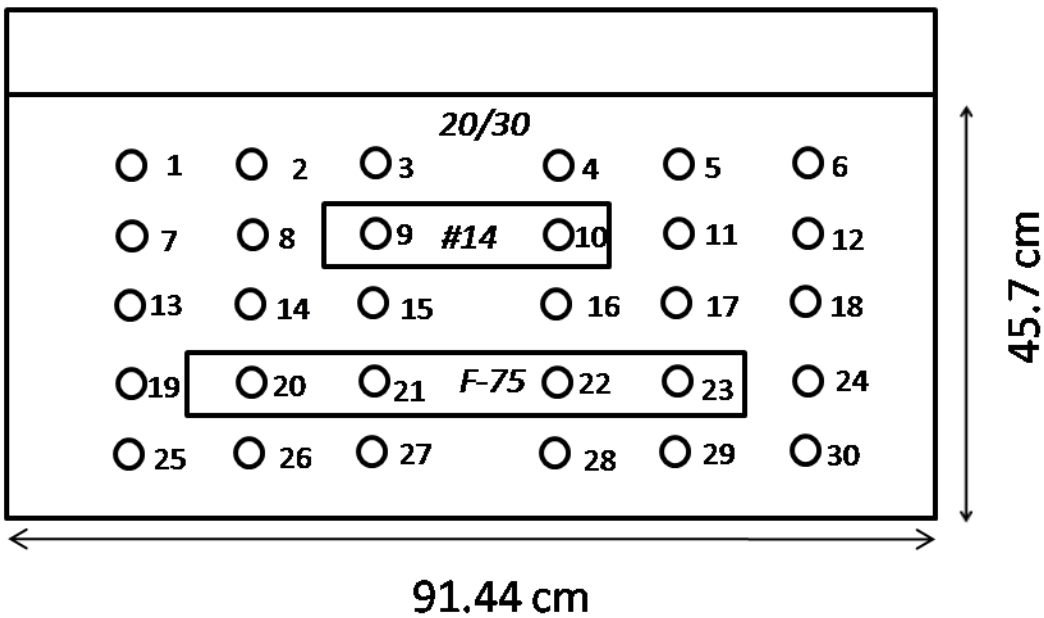


Figure 15.1: Schematic diagram of the sandbox showing sand types, water sampling, and pressure monitoring ports. Two rectangular boxes indicate locations of sand lenses consisting of #14 and F-75 sands.

15.2.2 Porous media used to pack the sandbox

Porous media chosen for packing the sandbox was based on two criteria: 1] to create a simplified heterogeneity pattern with sufficient contrast *K* to allow hydraulic tomography to image the various *K* zones and 2] to create a low *K* zone in the central part of the source zone to cause infiltrating TCE from the top to pool on this layer. Our goal was to create a realistic looking TCE source zone through infiltration in the central part of the sandbox so that partitioning tracer tests can be conducted across the central zone. Based on these criteria, we selected three commercially available sands [F-75, #14, and 20/30]. The F-75 sand is the finest sand [d_{50} = 0.174 mm] among the three sand types and was carefully packed as a single horizontal lens within a coarse matrix of 20/30 sand [d_{50} = 0.750 mm] [Figure 15.1]. The #14 sand is a medium grained sand [d_{50} = 0.457 mm] and was packed as short layer above the F-75 layer. The synthetic aquifer has a total volume of 7,985.02 cm^3 and a corresponding void volume of 2,635.06 cm^3 yielding an average porosity of 0.33.

A schematic diagram of the 30 sampling ports and pressure monitoring ports as well as the sand distribution is shown on Figure 15.1. Table 15.1 summarizes the *K* of the sands previously determined [*Illman et al*., 2007; *Liu et al*., 2007; *Illman et al*., 2010] using an empirical model [*Shepherd*, 1989] and a constant head permeameter [*Klute and Dirksen*, 1986].

Table 15.1: Characteristics of porous media used to create a synthetic heterogeneous aquifer.

Sand type	d_{50} [mm]	*K* [cm/s] Shepherd	*K* [cm/s] Darcy
20/30	0.750	1.03×10^{-1}	3.12×10^{-1}
#14	0.457	4.53×10^{-2}	1.21×10^{-1}
F75	0.174	9.22×10^{-3}	1.73×10^{-2}

15.2.3 Hydraulic characterization prior to TCE spill

To characterize the synthetic heterogeneous aquifer in the sandbox, we conducted dipole cross-hole tests prior to spilling the TCE. In the dipole cross-hole test, a mass balance of water injected and extracted was maintained using a peristaltic pump by connecting the injection and extraction ports in a single loop. Nine pairs of ports consisting of tests 1 through 9 [e.g., test 1: injection at port 1 and extraction at port 30, from now on I1/E30; test 2: I2/E29; test 3: I3/E28; test 4: I4/E27; test 5: I5/E26; test 6: I6/E25; test 7: I7/E24; test 8: I12/E19; and test 9: I13/E18] were chosen for these dipole cross-hole tests based on symmetry and coverage of the aquifer features [Figure 15.1]. Nine additional tests [tests 10-18] with the injection and extraction ports reversed were later used to evaluate the computed *K* tomogram [i.e., test 10: I30/E1; test 11: I29/E2; test 12: I28/E3; test 13: I27/E4; test 14: I26/E5; test 15: I25/E6; test 16: I24/E7; test 17: I19/E12; and test 18: I18/E13]. Injection and extraction rates for all dipole cross-hole tests averaged 150 mL/min. Prior to each cross-hole pumping test, all pressure transducers were calibrated to ensure accurate data collection. We then collected hydraulic head data for several minutes in all pressure transducers to establish a static, initial condition. After establishment of static conditions, we pumped from each port using a peristaltic pump, while taking head measurements in all 30 ports. For each test, pumping continued until the development of steady state conditions, which was determined by observing the stabilization of all head measurements on the data logger connected to a computer. After reaching steady state, the pump was shut off to collect recovery head data until its full recovery. Transient head data were also collected, but in this study, we only utilized steady-state head data for steady state hydraulic tomography to estimate a *K* tomogram.

15.3 CREATION OF TCE SOURCE ZONE

A TCE source zone was then created in the synthetic aquifer by injecting 125 mL of TCE [> 99.5% purity, Sigma-Aldrich] dyed with Sudan IV [Sigma-Aldrich] using a syringe approximately 5 mm beneath the soil surface as shown on Figure 15.1. To achieve a constant injection rate, we used a peristaltic pump and the TCE was injected a rate of 12.5 ml per minute over 10 minutes. After commencing the injection, the TCE quickly infiltrated through the 20/30 sand, slowly moved through the #14 sand layer, and ponded on the F-75 sand as shown on Figure 15.2. The presence of the #14 sand layer apparently caused some spreading of the TCE resulting in a wider zone of contamination than initially anticipated. Figure 15.2 is photograph of the TCE distribution after no additional movement of the TCE infiltration front was observed. This figure also shows that some of the TCE moved downwards off the F-75 layer on the left side approximately 5 cm into the 20/30 sand. The injection of TCE in this sandbox resulted in a sandbox aquifer-averaged TCE saturation [S_N] of approximately 5% by assuming a porosity of 0.33 typical of these sands, which is equivalent to a volumetric TCE content of 0.0165.

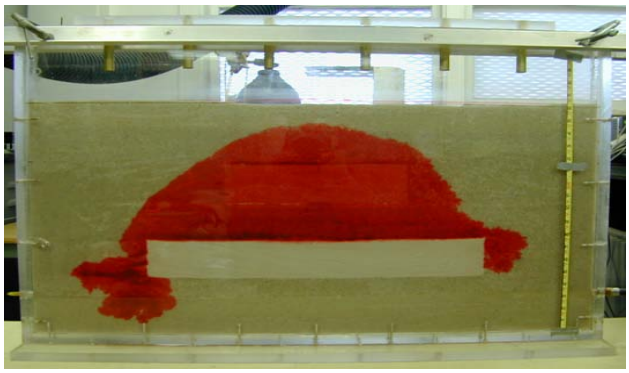


Figure 15.2: Photograph of the TCE source zone after TCE migration has stopped.

15.4 CHARACTERIZATION OF TCE SOURCE ZONE

15.4.1 Determination of partitioning coefficient through batch tests

In order to determine TCE saturation from partitioning tracer tests, the partitioning coefficient [K_{NW}] of the tracer and TCE is needed. Methods to estimate partitioning coefficients have been developed using UNIFAC [*Wang et al*., 1997] and on defining equivalent alkane carbon number for each contaminant in a DNAPL and the tracers [*Dwarakanath and Pope*, 1998]. Here, we estimated the partitioning coefficients through batch tests. The partitioning tracer selected for this study was 2-methyl-1-butanol [2M1B, purity > 99%; Sigma-Aldrich].

Six partitioning-tracer stock solutions of varying concentrations were prepared volumetrically with Hamilton gas-tight syringes in 250-mL volumetric flasks. Purified water [Barnstead Nanopure] was used to dilute the aqueous tracer stock solutions to the six different concentrations: 50, 100, 200, 400, 800, and 1600 mg/L of 2M1B. The 6 different levels were prepared in pre-cleaned 20mL EPA/ VOA glass vials with Teflon-lined septa [EnviroTech]. Each sample set consisted of triplicate vials containing TCE and the aqueous 2M1B tracer solution and a control group of triplicate vials containing aqueous tracer solution only. Twenty mL of aqueous tracer solution was dispensed into each vial using a 10 mL pipette followed by 2.5 mL injection of TCE with a gas-tight syringe. The vials were immediately capped to avoid volatilization to the atmosphere.

The 25mL vials were allowed to equilibrate on a platform shaker [Innova 2000] at 150 rpm and room temperature for 24 hours, sufficient to obtain equilibrium. Aliquots with an aqueous phase volume of 1.5 mL were withdrawn from the 25mL vials with minimal disturbance using a gas-tight syringe and transferred to 1.5 mL glass GC auto sampler vials with Teflon septa and analyzed for 2M1B concentrations using a gas chromatograph [Shimadzu GC 2010] with an auto sampler [AOC-20i auto injector] equipped with a flame ionization detector [FID]. The GC-FID method consisted of a Restek Stabilwax® capillary column [30 m, 0.32 mm ID, with film thickness of 1µm], helium as the carrier gas, zero-grade air, injection volume of 1 µL, and injector and FID temperatures of 200°C. Also, a split ratio of 35:1 to vent excess water [steam] in the sample and a linear velocity of 50 cm/s were incorporated into the method. Four external calibration standards were prepared for 2M1B in deionized water at levels of 1, 10, 100, and 1000 mg/L. The calibration curve [R^2 > 0.99] was created by injecting triplicates of water samples containing different concentrations of 2M1B. The batch test results were plotted on an

arithmetic plot and the partitioning coefficient [K_{NW}] for TCE-2M1B was estimated to be 3.54. This value compares favorably to those obtained by another research group [*Dugan et al*., 2003].

15.4.2 Partitioning tracer tests

Upon settlement of the TCE in the source area, we conducted 4 separate partitioning tracer tests. In each partitioning tracer test, a dipole flow field was created by injecting tap water at one port at a rate of 500 mL/min and extracting at another port at 450 mL/min. A water mass balance was then achieved by setting 12 sampling lines at a cumulative rate of 50 mL/min. Achieving a mass balance was crucial because the 3 boundaries of the sandboxes had no-flow boundary conditions, while the top boundary was open to the atmosphere. Deviations in the mass balance would result in a rising or falling water level which we monitored with pressure transducers. Table 15.2 summarizes the injection, extraction, and sampling ports for each of the 4 partitioning tracer tests.

Table 15.2: Injection, extraction, and water sampling locations during partitioning tracer test 1 - 4.

Partitioning tracer test	Injection/extration ports	Sampling port
1	30/1	2, 3, 4, 5, 6, 7, 13, 19, 25, 26, 27, 28
2	26/6	1, 2, 3, 4, 5, 12, 13, 18, 24, 27, 28, 29
3	1/30	3, 4, 5, 6, 12, 18, 24, 25, 26, 27, 28, 29
4	18/13	1, 2, 3, 4, 5, 7, 19, 25, 26, 27, 28, 29

Once steady state flow conditions were reached in which the injection, extraction, and sampling rates all stabilized, a valve was switched on the injection line from water to a solution containing 2M1B as the partitioning tracer with a concentration of 500mg/L. Bromide was used as a conservative tracer with a concentration of 100 mg/L. The tracer solution was injected for 10 minutes and then the valve was switched back to tap water. Separate sample sets were collected for bromide and 2M1B throughout the duration of the test from the 12 sampling ports. Crude estimates of the tracer travel time were obtained using Darcy's law to estimate sampling times at each port. Water sampling from each one of the sampling lines was made possible with a manifold that fed the water samples into twelve 1.5 ml vials simultaneously. Thirty 2M1B samples and 20 bromide samples were collected from each of the 12 sampling lines during each partitioning tracer test for a total of 600 samples. Each tracer test lasted approximately 2 hours.

The 4 tracer tests were designed to maximize the contact of the tracer solution with the TCE source area. Each of the tracer tests was performed with different injection/extraction and water sampling schemes to sweep the tracers through the source zone from different directions in a tomographic fashion. During these experiments, nonaqueous phase TCE was observed to be stable. This means that we did not witness mobilization of TCE during the tracer tests despite the fact that water was injected and extracted during each tracer test.

Water samples taken during each tracer test were separated and chilled in a refrigerator at 4°C until sample analysis. Concentrations of 2M1B were determined using a procedure described earlier for the analysis of the batch test samples using the GC-FID. One notable difference is the calibration sequence. Calibrations were made in each analysis sequence consisting of 12 calibration injections [triplicates from each one of the 4 standards], 120 tracer samples, and 3 quality control samples. The quality control samples were vials filled with 100 mg/L standard and was analyzed after each set of 40 samples to ensure that the signal response of the machine remained accurate. Methanol injections [blanks] were made in between sets of samples to

prevent carryover. We note that other researchers [*Dugan et al.*, 2003] prepared calibration standards composed of the tracers in methanol. We used water instead of methanol to analyze the sample set more accurately, since our samples are aqueous that required calibration curves that fully represent the actual conditions of our samples.

The conservative tracer [Br] concentrations were determined using an ion chromatograph [IC] [Dionex ICS 2000] equipped with an AS40 auto sampler. The IC analysis method consisted of an AS-18 4 mm column, EGC II KOH eluent generator with an eluent concentration of 23mM, SRS-Ultra II 4mm suppressor with a current of 57 mA, DS6 heated conductivity cell with a cell temperature of 35°C, and a flow rate of 1 mL/min. Four external calibration standards were prepared for bromide with 1000mg/L Fluka IC standard diluted with de-ionized water at levels of 1, 10, 100, and 200 mg/L. The calibration plots were created by injecting triplicates of each calibration level and were linear [$R^2 > 0.99$]. Calibrations were made in each analysis sequence consisting of 12 calibration injections [triplicates of the 4 standards] followed by 80 tracer samples. Clean up runs were made in between each standard and sample injection to prevent carryover by increasing the eluent strength to 50 mM and increasing the suppressor current to 124 mA for 8 minutes.

15.4.3 Direct sampling of TCE saturation through coring

As soon as the partitioning tracer tests were completed, six core sampling tubes were inserted from the top of the sandbox to obtain independent estimates of TCE saturations to facilitate the comparison of TCE saturation estimates. In particular, 6 brass tubes were pushed vertically downward slightly offset from each of the six columns of sampling ports into the sandbox under fully saturated conditions until they reached the sandbox bottom. Then, bentonite was added to the top of the core to eliminate the empty space and the core tubes were plugged at the end with epoxy putty. Once the epoxy putty was dried, the cores were then removed individually and cut into segments with a pipe cutter. Each of the segments was capped on each end with a silicone plug and cooled in the refrigerator at 4°C until the samples were ready for sample preparation. The core samples were collected from the core tubing segments and deposited into pre-labeled, pre-weighed 40 mL vials and capped with Teflon-lined septa caps. After this step, the vials were then weighed again and the mass of the vial containing the cap and sample was recorded. The vials were then filled to the top with methylene chloride and weighed again. We then recorded the mass of the vial, along with the cap, sample and the methylene chloride. The volume of methylene chloride was calculated based on the mass differences and density of methylene chloride. The vials were placed on an Innova 2000 orbital shaker at 150 rpm for 24 hours. After shaking, the vials were weighed again and the volume of methylene chloride was corrected if any mass loss was observed. Approximately 5 mL of the solution was extracted from each 40 mL vial and transferred to three 1.5 mL vials for GC analysis.

The GC samples were then analyzed via GC-FID with a 7-level triplicate calibration method consisting of standards of TCE in methylene chloride at levels of 1, 10,100, 500, 1,000, 5,000 and 10,000 mg/L. The sample TCE concentrations were recorded and the average and standard deviations from each set of 3 samples were calculated. The mass of TCE in each core sample was calculated by multiplying the average sample by volume of liquid in the sample [assumed to be the volume of methylene chloride]. The mass was then converted to volume by dividing by the density of TCE. The core samples were then dried thoroughly by baking at 50° C to evaporate the methylene chloride and then at 105° C for 24 hours to remove any remaining

water in the core sample. Total volume of the dried sample was determined by measuring in a 25 mL graduated cylinder with 0.2 mL graduations.

The volume of solids in each sample was determined by placing the sample in a 25 mL graduated cylinder pre-filled with 12 mL of de-ionized water and measuring the displaced volume. The void volume is then calculated by the total volume minus the solid volume. The TCE saturation of each sample was calculated using:

$$S_{Ni} = \frac{V_{NAPL}}{V_{voids}} \qquad [15.1]$$

15.5 INVERSE MODELING OF DIPOLE CROSS-HOLE AND PARTITIONING TRACER TESTS

The hydraulic and partitioning tracer tomography analysis of the dipole cross-hole and partitioning tracer tests were conducted using the Sequential Successive Linear Estimator [SSLE] developed for HPTT [*Yeh and Zhu*, 2007]. The SSLE algorithm first analyzes data from the dipole cross-hole tests to estimate the K field and then the dipole partitioning tracer test data sequentially to delineate both the volumetric water content [θ_w] and volumetric TCE content [θ_N] distributions. Note that the porosity is the sum of water content and TCE content. The general procedure for hydraulic and partitioning tracer tomography analyses is identical and the details to the algorithm are provided in the original publication [*Yeh and Zhu*, 2007]. Here, we provide a brief description of the inversion method.

15.5.1 Description of algorithm

The inverse model assumes a steady flow field and the natural logarithm of K, θ_w, and θ_N [ln K, ln θ_w, and ln θ_N, respectively] are treated as a stationary stochastic process. The model additionally assumes that the mean and correlation structure of the K, θ_w, and θ_N fields are known, *a priori*. The algorithm is composed of two parts. First, the Successive Linear Estimator [SLE] is employed for each dipole cross-hole test to estimate a K tomogram. The estimator begins by cokriging the initial estimate of effective hydraulic conductivity [K_{eff}] and observed heads collected in one pumping test during the tomographic sequence to create a cokriged, mean removed ln K [f, i.e., perturbation of ln K] map for the hydraulic tomography analysis. For the partitioning tracer tomography, the cokriging is done instead using the initial estimates of effective θ_w, and θ_N with the concentration data.

Cokriging does not take full advantage of the observed head values because it assumes a linear relationship [*Yeh and Liu*, 2000] between head and K, while the true relationship is nonlinear. To circumvent this problem, a linear estimator based on the differences between the simulated and observed head values is successively employed to improve the estimate. Analogous to inversion of head data, the partitioning tracer tomography utilizes the differences between the concentration measurements in a successive fashion to overcome the nonlinear relationships between concentrations, θ_w, and θ_N.

The second step is to use the hydraulic head data sets [or concentration data for PTT] sequentially instead of including them simultaneously in the inverse model thus the term, "Sequential Successive Linear Estimator" [SSLE] is used to describe the inverse algorithm hereon. In essence, the sequential approach uses the estimated K [or θ_w, and θ_N for PTT] field and covariances, conditioned on previous sets of head measurements [or concentration measurements for PTT] as prior information for the next estimation based on a new set of pumping test data for HT and tracer test data for PTT. This process continues until all the data

sets are fully utilized. Modifications made to the code for this study include its ability to account for variations in the boundary conditions with each pumping test as they are sequentially included and implementing the modified loop scheme [*Zhu and Yeh*, 2005].

15.5.2 Inverse modeling of dipole cross-hole tests

To obtain a *K* tomogram from the available cross-hole pumping tests, we solve an inverse problem for steady-state flow conditions. The synthetic aquifer was discretized into 612 elements and 1332 nodes with element dimensions averaging 2.54 cm × 1.91 cm × 2.69 cm. The top boundary was set to be a constant head boundary while the other three sides were set as no flow boundaries. Input data to the inverse model include initial guesses for the effective hydraulic conductivity [K_{eff} = 0.3 cm/s], variance [$\sigma^2_{\ln K}$ = 0.1] and the correlation scales [λ_x = 40.0 cm, λ_y = 1.91 cm and λ_z = 5.0 cm], injection/extration rates [*Q*] during each dipole cross-hole test, as well as available point [small-scale, i.e., core, slug, and single-hole tests] measurements of *K*. For the SSHT analysis presented here no small scale data were used for the inverse modeling.

For the steady state hydraulic tomography analysis, we selected 9 dipole cross-hole tests 1- 9 and the corresponding steady-state head observations at the rest of 28 ports during each test as data sets. We elected to not use the head data from the injection and extraction ports from each test because those ports could be affected by skin effects [*Illman et al*., 2007]. Prior to the computation of the *K* tomogram with the SSLE algorithm, we pre-processed the hydraulic head data to remove noise. This is because previous efforts [*Illman et al*., 2008; *Xiang et al*., 2009] have shown that the signal-to-noise ratio can be critical in inverse modeling of cross-hole pumping tests.

15.5.3 Inverse modeling of dipole partitioning tracer tests

After the computation of the *K* tomogram, we used the HPTT code to compute the TCE content tomogram within the laboratory aquifer. The TCE content tomogram was converted to a TCE saturation tomogram by dividing the TCE content tomogram by the porosity [0.33].

The model domain utilized for the inverse modeling of the partitioning tracers was identical to that for the steady state hydraulic tomography. Boundary conditions are also identical with the exception that the upper boundary is treated as a no flow boundary as oppose to a constant head boundary. The selection of no flow boundaries for all sides ensures that no mass is lost across a boundary during the simulations.

Input data to the inverse model for PTT include the hydraulic conductivity distribution [either determined from SSHT or assigned values based on knowledge of the layers within the tank], dispersivity [set to 0 uniformly within the model domain], 2M1B-TCE partitioning coefficient [3.54], variance [for TCE content] [$\sigma^2_{\ln K}$ = 0.01] and the correlation scales for TCE content [λ_x = 20.0 cm, λ_y = 1.91 cm and λ_z = 1.0 cm], injection/extraction, and water sampling rates [*Q*] during each dipole partitioning tracer test, as well as available point scale measurements of TCE content obtained from coring or other means. Similar to the SSHT analysis presented earlier, no small scale data were used to condition the inverse model.

Since bromide was included in the PTT injection solution as a conservative tracer it was possible to assess the ability of the *K*-tomogram to reproduce the transport of the conservative tracer. The observed breakthrough curves [BTCs] at each sample location, for each tracer test were visually compared to a synthetic simulation of the test. In general there was good agreement between the synthetic and observed breakthrough curves. Simulated and observed breakthrough curves were considered to be in good agreement when they were nearly coincident.

When this was the case it meant that the flow path between the injection port and the observation port was accurately represented by the model. Any error between the simulated and observed breakthrough curves (particularly in terms of the tail end) could be result in errors in the estimate of DNAPL location and mass. As such, only those observations ports that had good agreement for the conservative tracer breakthrough curves were included for the estimation of DNAPL residual saturation.

The sampling ports included in the DNAPL saturation estimation were: test 1-Ports 2, 3, 6, 13, 25, 26, 27, 28; test 2 - Ports 4, 5, 12, 13, 18, 24, 27, 28, 29; test 3 - Ports 3, 4, 5, 6, 12, 24, 25, 26, 27; and test 4 - Ports 1, 5, 25, 26, 27, 28, 29. We selected an average of 5 data points to define the breakthrough curves at each port. The first point selected was at approximately 1% of the injected concentration. The second point was approximately half the peak concentration. The third point was the peak concentration and the fourth point was half the peak concentration after the passing of the peak and the fifth point represented approximately 1% of the injected concentration after the passing of the peak. The total number of data points used for each test was 40, 55, 60, and 45 for tests 1, 2, 3 and 4, respectively.

15.6 RESULTS

15.6.1 *K* tomogram and its validation

All computations for steady state hydraulic tomography analyses were executed using 44 of 48 processors on a PC-cluster consisting [of 1 master and 12 slaves each with Intel Q6600 Quad Core CPU running at 2.4 GHz with 16 GB of RAM per slave] at University of Waterloo. The operating system managing the cluster was CentOS 5.3 based on a 64-bit system. The total computational time for inverting data from 9 dipole cross-hole tests was about 1 minute indicating the efficiency of the SSLE algorithm to image the *K* distribution using steady state head data.

Figure 15.3 is the *K* tomogram obtained by inverting the steady state head data from 9 dipole cross-hole pumping tests. The *K* values range between 5.3×10^{-1} and 9.0×10^{-3} cm/s, has a geometric mean of 2.1×10^{-1} cm/s cm/s and a $\sigma^2_{\ln K}$ of 0.91. For comparison purposes, we include the outline of the two blocks embedded within the 20/30 sand. Comparison of the *K* values from the tomogram to those from the permeameter analysis of core samples previously reported elsewhere [24, 28] shows that the values are quite comparable. In particular, the permeameter estimate of *K* for the 20/30 sand is 3.12×10^{-1} cm/s, for #14 sand $K = 1.21\times10^{-1}$ cm/s, and for F75 sand $K = 1.73\times10^{-2}$ cm/s [Table 5.1]. Given the small difference between the *K* values of #14 and 20/30 sands, it is not surprising that the block consisting of #14 sands is not detected. In contrast, the block consisting of F75 appears quite clearly due to the larger contrast in the *K* values between the 20/30 and F75 sands, although the block appear smoother and larger than the outline indicated on the tomogram. This is a direct consequence of having only 30 ports available for pressure measurements. A previous study [*Yeh and Liu*, 2000] has shown that as the number of pressure measurement points increase, the resolution of the *K* heterogeneity increases.

There are a number of approaches to evaluate the *K* tomogram. One is the visual assessment which involves the comparison of the tomogram to the actual sand distributions if this is available as in this laboratory case. Another method may be to compare local estimates of *K* or their statistical moments to those from the tomogram. We previously [*Illman et al*., 2007; *Liu et al*., 2007] concluded that the best way to evaluate a *K* tomogram was to simulate one or more independent pumping tests not used in the construction of the *K* tomogram.

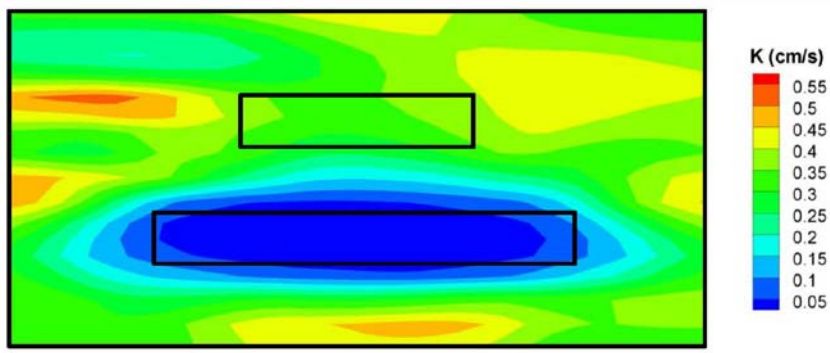


Figure 15.3: *K* tomogram generated using drawdown data from 9 dipole tests.

Here, to quantitatively assess the validity of the *K* tomogram, we simulated 9 additional dipole cross-hole tests [tests 10 - 19] using the forward groundwater model MMOC3 [*Yeh et al*., 1993]. For the validation using dipole cross-hole tests, the simulated and measured drawdown values for each case are plotted on separate scatter plots and a linear model is fitted to each case without forcing the intercept to zero. The linear model fit and the coefficient of determination [R^2] provide indications of scatter and bias. The R^2 is a statistic that provides a quantitative measure of similarity between the simulated and measured drawdown values. A high R^2 value means that the simulated and measured drawdown values are linearly correlated, even though the mean values could be different. Other measures of correspondence between the simulated and observed drawdown values are the mean absolute error [L_1] and the mean square error [L_2] norms. The L_1 and L_2 norms are computed as:

$$L_1 = \frac{1}{n}\sum_{i=1}^{n}\left|\chi_i - \hat{\chi}_i\right| \qquad [15.1]$$

$$L_2 = \frac{1}{n}\sum_{i=1}^{n}\left(\chi_i - \hat{\chi}_i\right)^2 \qquad [15.2]$$

where n is the total number of drawdown data, i indicates the data number, and χ_i and $\hat{\chi}_i$ represent the estimates from the simulated and measured drawdowns, respectively.

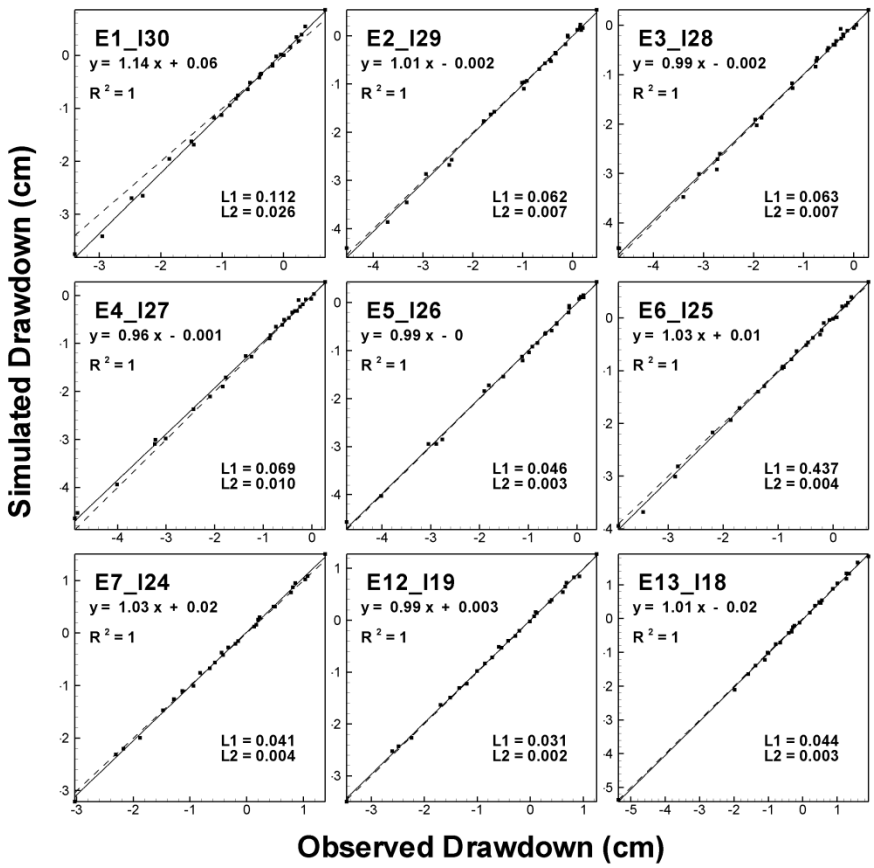


Figure 15.4: Scatter plot of simulated [cm] versus observed drawdown [cm] under steady state flow conditions.

Figure 15.4 shows a series of scatter plots that compare the simulated drawdown values obtained through forward simulations of the 9 independent dipole cross-hole tests using the *K* tomogram from figure 15.3 to the measured drawdown values from the actual tests. The figure includes a 45 degree line which indicates a perfect correlation. In general, we observe that the *K* tomogram yields excellent predictions of 9 dipole cross-hole tests as indicated by the linear model fit, coefficient of determination and L_1 and L_2 norms. One reason why the fits may be so good is that the test configurations are identical except for the switching of the injection and extraction locations. More recently, another study [*Illman et al.*, 2010] showed a similar evaluation in a different sandbox with a considerably more complex heterogeneity pattern created by the cyclic flux of sediment laden water. In those tests, independent cross-hole pumping tests with different source locations were used to construct and validate their *K* tomogram showing the excellent performance of SSHT in imaging the subsurface heterogeneity in *K*.

15.6.2 Computation of TCE saturation tomograms using the *K* tomogram

After the completion of SSHT to obtain the *K* tomogram, we next computed the TCE saturation tomogram using the *K* tomogram as the *K* field for the tracer inversion. As in SSHT, the computations for partitioning tracer tomography were executed using 44 of 48 processors on

the same PC-cluster described above. The total computational time for inverting data from 1 partitioning tracer test data ranged between 5 min to 10 min depending on the test. The sequential inversion of the 4 partitioning tracer tests required approximately 30 minutes and with the looping scheme [described next] the inversion required approximately 60 minutes indicating that more computational resources are required for the inversion of tracer test data.

A looping scheme [*Zhu and Yeh*, 2005] was developed to improve the SSLE algorithm to fully exploit the data as they are included sequentially. In this scheme, as new data sets are added into SSLE, the algorithm iterates until the data set meets a converge criterion. Then, instead of going straight to the next new data set, the scheme goes back to check the convergence for the first data set. If the converge criterion is not met, the program starts a loop iteration in which the iteration involves both the first and second data sets. That is, the first data set is iterated once, and then the second data set is incorporated and iterated once also in a loop. The loop iteration continues until both data sets meet the converge criterion within one loop. Then, the next new data set is added. The algorithm treats this new data set similarly to the second data set, except the loop iteration now involves three data sets. Additional data sets are added in a similar way. As a consequence, the SSLE approach improves estimates throughout the loops, maximizes the usefulness of data sets, and alleviates the problems associated with their previous SSLE approach.

Figures 15.5a - f show results from various cases in inverting the partitioning tracer test data. In particular, Figures 15.5a – d shows inversions of individual partitioning tracer tests 1 through 4, respectively. These results reveal that with a single partitioning tracer test, the existence of TCE in the source area is confirmed although the distribution is not very realistic compared to the photograph of the TCE distribution. This suggests that perhaps one tracer test with a limited number of sampling ports [Test 1: Ports 2, 3, 6, 13, 25, 26, 27, 28; test 2 - Ports 4, 5, 12, 13, 18, 24, 27, 28, 29; test 3 - Ports 3, 4, 5, 6, 12, 24, 25, 26, 27; and test 4 - Ports 1, 5, 25, 26, 27, 28, 29] may not yield accurate distributions of TCE saturation. Interestingly, each TCE saturation tomogram shows a different TCE saturation distribution as well as mean TCE saturations [Table 5.3] suggesting that the inversion of a single partitioning tracer test can yield a different mean TCE saturation value. For example, the inversion of test 1 alone yields a mean TCE saturation of 3.7% while test 4 yields 1.9%. Recalling that 125 ml of TCE was injected into the sandbox, the mean saturation is 5% suggesting that the mean estimate is close but there is definitely some error which could affect remediation designs.

Table 15.3: Mean TCE saturation and volume estimated through the inversion of various partitioning tracer tests using the *K* tomogram from steady state hydraulic tomography and a known *K* distribution.

	K tomogram from SSHT		Known *K* distribution	
Test	Mean TCE saturation	V_{TCE} [mL]	Mean TCE saturation	V_{TCE} [mL]
1	0.0372	98.10	0.0277	71.99
2	0.0496	128.55	0.0476	126.09
3	0.0317	81.39	0.0309	81.09
4	0.0189	46.53	0.0202	51.48
1-4	0.0218	55.46	0.0491	128.73
1-4 [L]	0.0473	130.06	0.0593	156.94
4-1	0.0360	97.09	0.0521	139.60
4-1 [L]	0.0475	128.12	0.0699	189.58

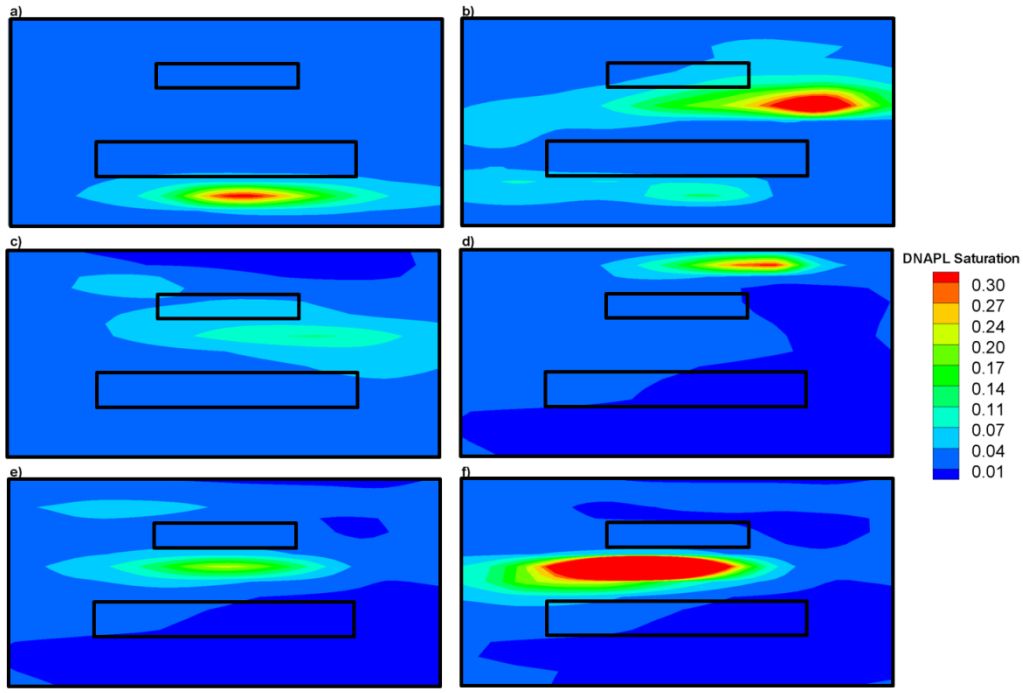


Figure 15.5: Estimated TCE saturation tomograms with inversion of: a] test 1 only; b] test 2 only; c] test 3 only; d] test 4 only; e] tests 1 - 4; and f] tests 1 - 4 with loop iteration. In this case, the *K* tomogram estimated using steady state hydraulic tomography [Figure 15.3] was used for partitioning tracer tomography.

We next sequentially inverted tests 1 through 4 in that order without the loop iteration scheme and the corresponding TCE saturation tomogram is shown in Figure 15.6e. This figure shows a region of elevated TCE saturation above the F75 block suggesting that after including 4 tests, HPTT is able to locate a higher TCE saturation region in the central part of the aquifer, despite the exact position not being correct as it is above the F75 lens. We do note that the mean TCE saturation is approximately 50% less of the true value [5%] at 2.2%. Applying the loop iteration scheme described earlier, our estimate of the mean TCE saturation improves to 4.7% and the, the TCE saturation tomogram looks more realistic with higher TCE saturations in the expected area.

15.7 DISCUSSION

Despite the encouraging results, it is natural for one to question whether the TCE saturation distribution can be further improved if we had a better estimate of *K* heterogeneity. A better estimate of *K* heterogeneity can be obtained with a larger number of measurement points during steady state hydraulic tomography [*Yeh and Liu*, 2000]. As we do not have additional sampling ports, we instead deterministically assign *K* values to the 3 sand types based on *K* measurements from the permeameter tests and then use this *K* field for the inversion of the partitioning tracer tests. The *K* tomogram that we computed earlier can be considered to be an unbiased conditional mean field and because it is based on geostatistical methods, the variability in the heterogeneity is relatively smooth with a limited number of pressure measurement points.

Assigning *K* values as we do here is tantamount to knowing precisely the location of the stratification and its morphology as well as the locations of sharp boundaries between zones of different *K* values.

Figures 15.6a through 15.6f are the estimated TCE saturation tomograms for the various cases that we examined before. In particular, Figures 15.6a-d are those estimated from individual tests 1 through 4, Figure 15.6e is the result from sequentially inverting tests 1 through 4, and Figure 15.6f is the result from sequentially inverting tests 1 – 4 but also incorporates the loop iteration scheme. As in Figure 15.5a-d, the inversion of individual partitioning tracer tests shows that the mean TCE saturation estimates are not that much better than the case when the *K* tomogram was used. However, as Figures 15.6e and 15.6f shows the position of the high TCE saturation zone is considerably better constrained when the positions of the strata are known. This suggests that improved knowledge of K heterogeneity and in particular, precise knowledge of boundaries between strata can be critical in imaging TCE saturations in heterogeneous aquifers.

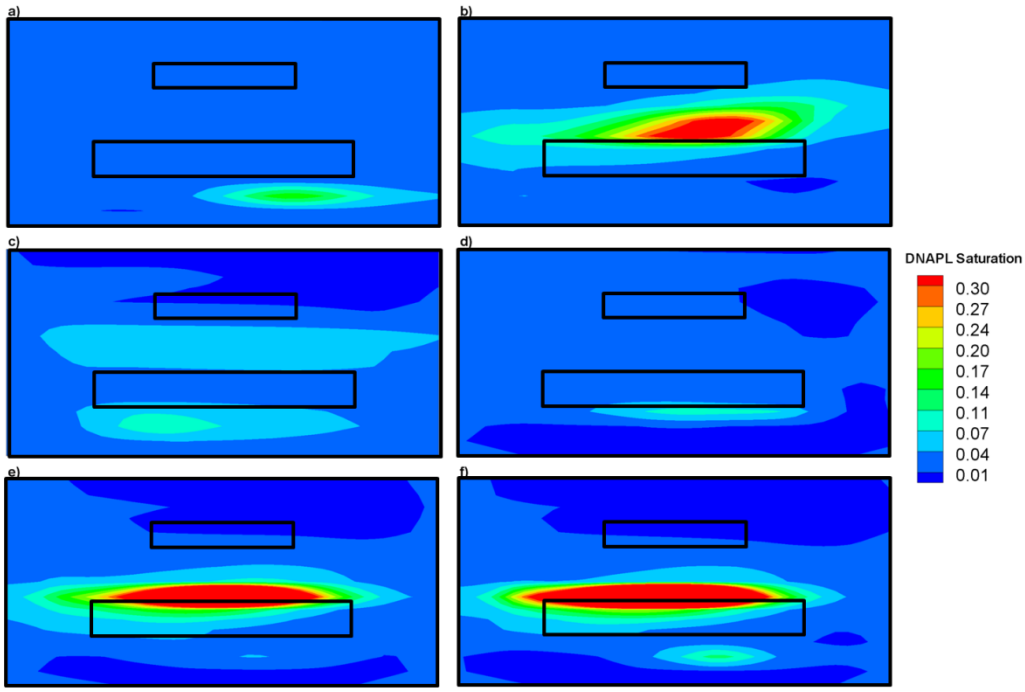


Figure 15.6: Estimated TCE saturation tomograms with inversion of: a] test 1 only; b] test 2 only; c] test 3 only; d] test 4 only; e] tests 1 - 4; and f] tests 1 - 4 with loop iteration. In this case, the *K* values of the 3 sand types were obtained from permeameter measurements and deterministically assigned for partitioning tracer tomography.

The visual comparison of the photograph showing the likely location of the TCE and the TCE saturation tomograms that we have computed show that HPTT is promising as it can location areas of higher saturations. Obviously, a better method to quantify TCE saturation is needed. For example, researchers have utilized X-ray [*Oostrom et al*., 1999; *Moreno-Barbero and Illangasekare*, 2006] to image DNAPL saturations in their flow cells.

However, a direct comparison of saturation estimates to those from HPTT is more desirable. To achieve this, we next compare the TCE saturation tomogram to those obtained from the interpolation of TCE saturation data that we have obtained through coring. In particular,

Figure 15.7 is the kriged TCE saturation distribution showing regions of elevated TCE saturation above and within the upper portion of the F75 layer. A grid by 1 cm by 1cm was used for kriging. A slightly higher TCE saturation region is also visible to the left side of the F75 block that corresponds with the contaminated area evident in the photograph [Figure 15.2].

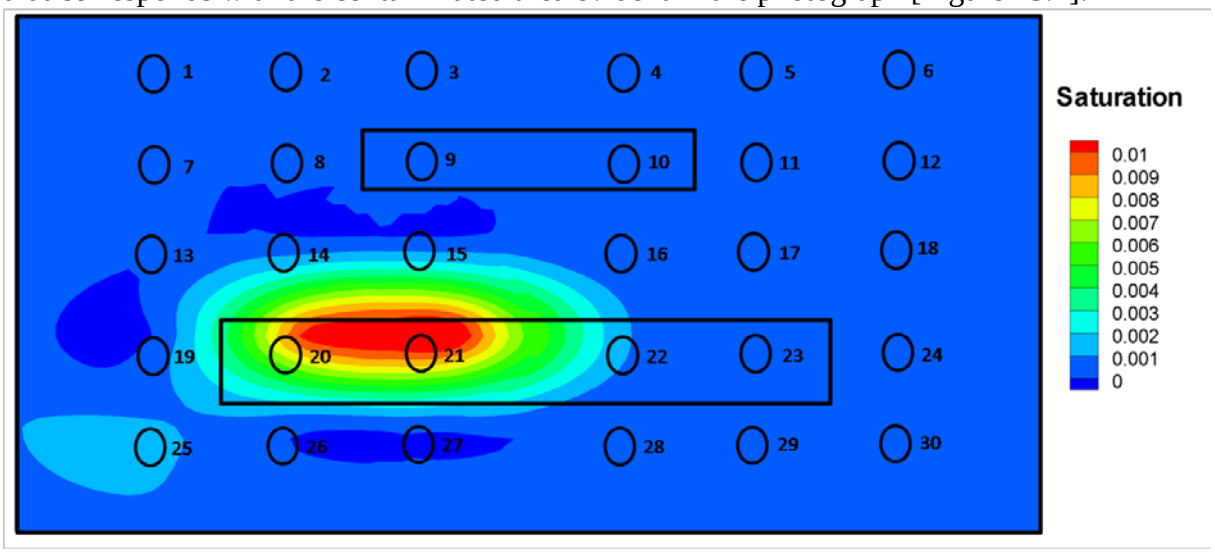


Figure: 15.7: Kriged TCE distribution based on the interpolation of 30 core samples. Solid circles show the approximate centroid location of the core samples.

There are two important features to note in Figure 15.7. One is that the TCE saturation is considerably lower than those estimated by the HPTT technique. The mean TCE saturation estimated through interpolation is 0.088%, which is significantly lower than the actual mean TCE saturation of 5% or those estimated by HPTT. This suggests that direct sampling through coring could potentially contribute to underestimation of TCE saturations. Another observation is that the higher TCE saturation zone is partially located within the F75 block. This is highly unlikely as seen also in Figure 15.2 that nonaqueous phase of TCE cannot penetrate the F75 sands due to the higher displacement pressure to overcome in order for the nonaqueous phase TCE to penetrate into the F75 sand. One likely explanation for observing nonaqueous phase TCE within the block is that the driving of the coring tube may have consolidated the sands causing the location of the high TCE saturation zone to be pushed deeper. This observation highlights the problem with direct sampling in the source zone - that drilling can contribute to mobilization of DNAPLs to greater depths exacerbating the contamination. In addition, coring of unconsolidated sediments can result in sample loss which can additionally contribute to inaccuracies in estimates of TCE saturations.

16. COMPARISON OF HETEROGENEITY CHARACTERIZATION METHODS FOR IMPROVED PREDICTIONS OF SOLUTE TRANSPORT: LABORATORY SANDBOX EXPERIMENTS

16.1 INTRODUCTION

Solute transport in heterogeneous media has been a topic of great interest over the last several decades. This is because the accurate prediction of contaminant transport will improve our ability to safeguard groundwater supplies through wellhead protection programs, and improvements in remediation designs will increase the efficiency of cleanup efforts. The recognition that improved predictions of solute transport results from better accounting for subsurface heterogeneity in hydraulic conductivity [*K*] has led to the development of stochastic theories for groundwater flow and solute transport. The emergence of stochastic theories has led to much optimism about our abilities to predict solute transport behavior both spatially and temporally. In particular, Yeh [1992] reviewed a number of these theories and discussed their advantages and limitations. He concluded that effective parameter approaches based on modern stochastic theories can yield averaged predictions of concentrations; however, more detailed information on the subsurface is needed to describe local scale behavior.

Various field [e.g., MacKay et al., 1986; LeBlanc et al., 1991; Boggs et al., 1992] and laboratory experiments [e.g., Fernàndez-Garcia et al., 2005 and others] have been conducted to show the applicability of these stochastic theories. In particular, Sudicky [1986] characterized the spatial variability of hydraulic conductivity in great detail at the Borden site in Canada, by conducting permeability measurements on a series of cores taken along the longitudinal and transverse directions of a natural gradient tracer test conducted adjacently. He then used the theoretical expressions developed by Gelhar and Axness [1983] and Dagan [1982, 1984] to show a good prediction of dispersion from the injected tracer resulting in optimism that theoretical expressions can provide meaningful estimates of effective transport parameters. More recently, at a landfill site in North Bay, Canada, Sudicky et al. [2010] obtained a large number of hydraulic conductivity estimates at many locations, similar to the study at Borden [Sudicky, 1986]. Using these estimates Sudicky et al. [2010] conducted geostatistical analysis, calculated the effective conductivity and macrodispersivity, and used those estimates in a three-dimensional numerical model to show a reasonably close prediction of the water table position as well as the extent and migration rates of a plume emanating from the landfill. This study demonstrated that the use of statistically derived parameters based on stochastic theories results in reliable large-scale three-dimensional flow and transport models for complex hydrogeological systems. One other important finding was that the length of the plume was relatively insensitive to the value of the longitudinal macrodispersivity under the conditions of steady flow in three-dimensions with a constant source strength. The success of the effective parameter approach along with the macrodispersion concept [e.g., Gelhar and Axness, 1983; Dagan, 1982, 1984] for large scale plumes at the Borden and North Bay sites is very encouraging. This is because of the simplicity and practicality of these solutions in satisfactorily reproducing the spatial moments of the plume. However, this success is largely due to the fact that the plume has travelled a distance much larger than the dominant scales of heterogeneity resulting in a Fickian regime of solute transport. For shorter distances, this approach may be inadequate owing to limitations of the ergodicity assumption embedded in the macrodispersion concept [Yeh, 1992].

This limitation was demonstrated at the Georgetown site in South Carolina, USA, where McCarthy et al. [1996] conducted two-well, forced-gradient tracer experiments over a distance of 5 m in a coastal sandy aquifer. The evolution of three-dimensional chloride plumes during two tracer experiments was observed. A three-dimensional finite element model for flow and transport was then utilized by Yeh et al. [1995] to model the tracer test data. The site was unique because there was a large number of hydraulic conductivity data obtained from slug tests used to predict solute transport in the aquifer. Results showed that their predictive ability was limited to the bulk behavior of the plumes, which was mainly controlled by some 'significant' heterogeneities. It is clear from these field sites that stochastic theories are useful in predicting the average behavior of the solute plume, but solute transport behavior at individual point locations is still difficult to predict based on these theories. In other words, the classical advection-dispersion equation often fails to predict the behavior of solute in the subsurface.

Subsurface heterogeneity is often the cause of this failure and has been attributed to non-Fickian behavior observed both through laboratory [e.g., Levy and Berkowitz, 2003; Cortis et al., 2004] and field experiments [e.g., Sudicky et al., 1983; Mackay et al., 1986]. Research at other sites, most notably the MADE site also suggests that existing stochastic theories may not be able to predict even the average plume behavior when the site is strongly heterogeneous and when the statistical distribution of hydraulic parameters is nonstationary [Rehfeldt et al., 1992] leading to non-Fickian behavior. This has led to a race for more complex mathematical approaches such as space-time nonlocal stochastic advective-dispersive flux theory [e.g., Neuman, 1993; Cushman and Ginn, 1993; Zhang and Neuman, 1996], continuous time random walk [CTRW], [e.g., Scher and Lax, 1973; Berkowitz et al., 2006], and advection dispersion equations based on fractional derivatives to describe solute transport in strongly heterogeneous porous and fractured geological media [e.g., Meerschaert et al., 1999; Benson et al., 2000]. Neuman and Tartakovsky [2009] recently offered their perspective on differences, commonalities, and relative merits of these approaches.

Despite the increasing complexity of mathematical models that have been developed, the goal of accurately predicting solute transport concentrations in a multiscale heterogeneous environment is still elusive, and a much improved approach to efficiently characterize the subsurface is needed. Such accuracy in predicting solute transport is likely needed in contaminant transport forecasting and the design of remediation programs. That is, at many sites, tracer tests and the injection of remediation agents are conducted at a scale smaller than which the effective parameter concept applies.

Hydraulic tomography has been proposed as an alternative to traditional site characterization approaches. The method relies on the inversion of multiple cross-hole pumping or injection tests that are conducted in the saturated zone. To date, various inversion strategies have been developed [e.g., Gottlieb and Dietrich, 1995; Yeh and Liu, 2000; Vesselinov et al., 2001; Bohling et al., 2002; Brauchler et al., 2003; McDermott et al., 2003; Zhu and Yeh, 2005, 2006; Li et al., 2005; Fienen et al., 2008; Castagna and Bellin, 2009; Xiang et al., 2009] that show success in mapping subsurface hydraulic parameters. Laboratory experiments have been conducted by a number of researchers [Liu et al. 2002; Illman et al., 2007, 2008; 2010; Liu et al., 2007; Yin and Illman, 2009] to validate hydraulic tomography under conditions in which experimental errors and forcing functions can be controlled. The success of hydraulic tomography in terms of imaging the heterogeneity patterns in hydraulic parameters is also evident through various field experiments [Bohling et al., 2007; Straface et al., 2007; Illman et al., 2009; Cardiff et al., 2009; Berg and Illman, 2010]. However, whether the estimates of

hydraulic parameters obtained from hydraulic tomography can be used to accurately predict solute transport behavior has not been researched to date in the laboratory or in the field.

Recently, Ni et al. [2009] conducted numerical simulations to show that hydraulic tomography is able to obtain the best estimates of hydraulic parameters that allow for the most accurate simulation of solute transport behavior. In particular, these authors used a synthetic two-dimensional aquifer to assess the accuracy of predicted concentration breakthrough curves [BTCs] on the basis of the *K* fields estimated by geometric mean, kriging, and hydraulic tomography. Such *K* fields represent different degrees of flow resolutions as compared with the synthetically generated one. It was found that without intensive experiments to obtain accurate estimates of dispersivities at sites, the flow field based on the *K* distribution or [*K* tomogram from now on] can yield accurate predictions of breakthrough peaks and phases.

The main objective of this study is to compare three different heterogeneity characterization methods in terms of predicting solute transport. The three methods compared are: 1] the effective parameter approach of Gelhar and Axness [1983]; 2] oridinary kriging; and 3] hydraulic tomography. We conducted our comparison by creating a heterogeneous aquifer through the cyclic flux of sediment-laden water in a laboratory sandbox. The aquifer was first characterized by taking core samples to obtain local scale estimates of hydraulic conductivity [*K*] which were then used to calculate effective parameters, and to perform kriging. We then conduct a hydraulic tomography survey using the code developed by Yeh and Liu [2000] to map the heterogeneous *K* field [*K* tomogram from now on]. Upon completion of the hydraulic tomography survey, we conduct a conservative tracer test in a dipole configuration. Water samples collected at various locations within the synthetic aquifer allowed us to obtain breakthrough curves. The tracer test is then simulated through a forward transport simulation using the effective parameter field, the kriged *K* field, and the *K* tomogram. We compare the evolution of the tracer plume qualitatively by examining the photographs of the dyed plume and comparing those to the various numerical simulation cases based on the effective parameter concept and hydraulic tomography. We then make a direct comparison of the simulated and observed breakthrough curves. Finally, we quantitatively assess the results through the comparison of temporal moments computed from the simulated and observed breakthrough curves.

16.2 EXPERIMENTAL METHODS

16.2.1 Sandbox and Synthetic Aquifer Construction

A synthetic heterogeneous aquifer was constructed in a vertical, laboratory sandbox to validate various fluid flow and solute transport algorithms. The sandbox is 193.0 cm in length, 82.6 cm in height, and has a depth of 10.2 cm. The aquifer was created through the cyclic deposition of sediments under varying water flow and feed rates of sediments [Illman et al., 2010a]. Previous efforts by our group [Illman et al., 2007, 2008; Liu et al., 2007; Yin and Illman, 2009; Illman et al., 2010b] involved packing the sandbox in a relatively simplistic pattern. Our goal in relying on sediment transport was to create a more realistic heterogeneity pattern with various scales of heterogeneity in an efficient manner. Table 16.1 summarizes the sands used to create the synthetic aquifer including d_{50}, which is the particle diameter for which 50% of the grains are finer, hydraulic conductivity [*K]* estimates obtained using Shepherd's empirical model [Shepherd, 1989], and hydraulic conductivity [*K]* estimates obtained using a constant head permeameter device [Klute and Dirksen, 1986]. Figure 16.1 is a photograph of the frontal view

of the synthetic aquifer, showing the interfingering nature of the deposits with numbers indicating the layers. Port locations are also shown on this figure. Further details to this synthetic heterogeneous aquifer and its construction approach are provided in Illman et al. [2010a].

Forty eight ports, 1.3 cm in diameter, have been cut out of the stainless steel wall to allow coring of the aquifer as well as installation of horizontal wells. Each well was constructed by making 6 cuts spaced 1.46 cm apart in sections of brass tubing. The cuts were then covered with a stainless steel mesh that was bonded to the tubing with corrosion resistant epoxy. Extreme care was taken to avoid the epoxy filling the mesh which could impede water flow. The wells, which penetrate the thickness of the synthetic aquifer were installed after the deposition of the layers. This allowed each location to be monitored by a pressure transducer, used as a pumping or an injection port and as a water sampling port.

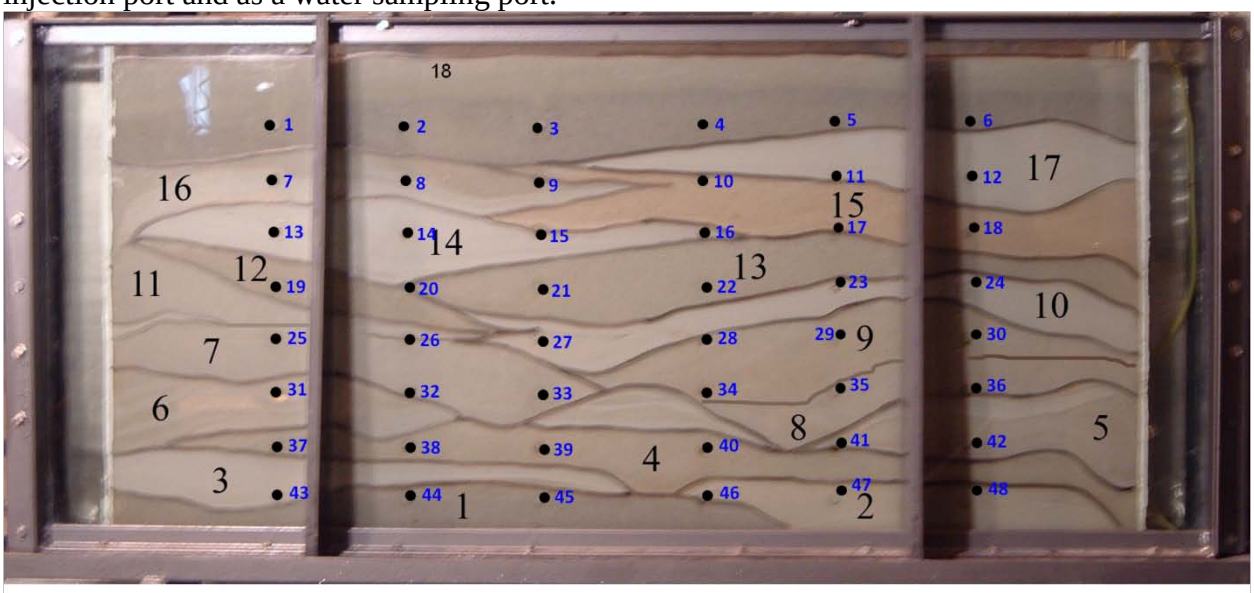


Figure 16.1: Photograph of the sandbox showing the synthetic heterogeneous aquifer in which the dipole cross-hole and conservative tracer tests were conducted. Large black numbers indicate layer numbers, solid circles indicate port locations, and small blue numbers indicate port numbers. Layer 1 = 20/30; layer 2 = 4030; layer 3 = F85; layer 4 = 20/40; layer 5 = mix; layer 6 = mix; layer 7 = #12; layer 8 = F32; layer 9 = 20/40; layer 10 = F65; layer 11 = #12; layer 12 = 16/30; layer 13 = 20/30; layer 14 = F75; layer 15 = 20/40; layer 16 = mix; layer 17 = F85; layer 18 = 20/30. Note: The layers labelled “mix” consisted of equal volumes of #14, F75, and 16/30 sands.

For this particular study, all boundaries around the synthetic aquifer were set as no-flow boundaries to achieve better mass control for our dipole cross-hole and tracer experiments. This was found to be critical during our initial studies as our previous sandboxes had constant head reservoirs which could potentially diminish the tracer signals via dilution.

The data acquisition system used for the laboratory experiments consisted of three major components. Pressure measurements were made with 50 Setra model 209 gauge pressure transducers with a range of 0 to 1 psi, 48 of which measured hydraulic head in the aquifer. These pressure transducers were installed at each of the 48 ports in the stainless steel wall of the sandbox.

Table 16.1: Characteristics of porous media used to create a synthetic heterogeneous aquifer.

Sand type	d_{50} [mm]	K [cm/s] Shepherd	K [cm/s] Darcy
16/30	0.872	1.32×10^{-1}	3.84×10^{-1}
20/30	0.750	1.03×10^{-1}	3.12×10^{-1}
20/40	0.578	6.68×10^{-2}	2.05×10^{-1}
#12	0.525	5.70×10^{-2}	2.05×10^{-1}
F32	0.504	5.33×10^{-2}	1.45×10^{-1}
#14	0.457	4.53×10^{-2}	1.21×10^{-1}
4030	0.355	2.99×10^{-2}	5.79×10^{-2}
F55	0.242	1.59×10^{-2}	2.80×10^{-2}
F65	0.204	1.20×10^{-2}	1.83×10^{-2}
F75	0.174	9.22×10^{-3}	1.73×10^{-2}
F85	0.151	7.29×10^{-3}	1.35×10^{-2}

16.2.2 Dipole tracer test

16.2.2.1 Description of dipole tracer test

After the completion of the dipole cross-hole tests, we conducted a dipole tracer test using bromide as a conservative tracer. Prior to the injection of the tracer, a dipole flow field was established by injecting tap water at port 42 at a rate of 372.4 mL/min and extracting at port 7 at 304 mL/min. A water mass balance was then achieved by setting 12 sampling lines at a cumulative rate of 68.4 mL/min. Figure 2 is a schematic diagram showing the injection, extraction and sampling ports for the dipole conservative tracer test.

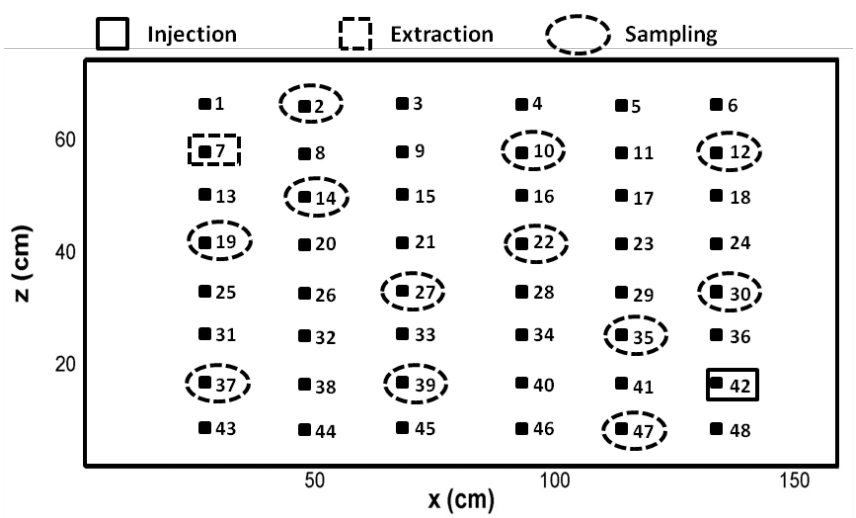


Figure 16.2: Schematic diagram showing the injection, extraction, and sampling ports during the dipole conservative tracer test.

Achieving a mass balance was crucial because the 4 boundaries of the sandboxes had no-flow boundary conditions. Water flow rates in each of the sampling lines were monitored frequently and the pressure transducers were also used to measure the pressure in the tank during

the tracer test. Deviations in the mass balance could result in an over-pressurization of the tank which could cause the glass holding the synthetic aquifer to rupture.

Once steady state flow conditions were reached [injection, extraction, and sampling rates were stable and in equilibrium] using a valve the injection line was switched from water to a solution containing bromide [Fisher Scientific] as the conservative tracer at a concentration of 151.3 mg/L. Food coloring [70 ml of blue 'Tone's' Food Coloring] was added to the tracer solution so that we could observe the movement of the tracer plume in the sandbox through the glass. The tracer solution was injected for 10 minutes and then the injection line was switched back to tap water. Water samples were collected for bromide throughout the duration of the test from the 12 sampling ports and the extraction well at port 7. Crude estimates of the tracer travel time were obtained using Darcy's law to estimate sampling times at each port. Water sampling from each one of the sampling lines was made possible with a manifold that fed the water samples into twelve 1.5 ml vials simultaneously. Sixty water samples were collected from each of the 12 sampling lines [plus the effluent line] during the dipole tracer test for a total of 780 samples. The dipole tracer test lasted approximately 5 hours.

16.2.2.2 Description of water sample analysis

Water samples taken during the tracer test were chilled in a refrigerator at 4°C until sample analysis. The conservative tracer [Br^{-}] concentrations were determined using an ion chromatograph [IC] [Dionex ICS 2000] equipped with an AS40 auto sampler. The IC analysis method consisted of an AS-18 4 mm column, EGC II KOH eluent generator with an eluent concentration of 30mM, SRS-Ultra II 4mm suppressor with a current of 75 mA, DS6 heated conductivity cell with a cell temperature of 35°C, and a flow rate of 1 mL/min. Five external calibration standards were prepared for bromide with 1000mg/L Fluka IC standard diluted with de-ionized water at levels of 1, 5, 10, 50, 100 mg/L. The calibration plots were created by injecting triplicates of each calibration level and were linear [$R^2 > 0.99$]. Calibrations were made in each analysis sequence consisting of 15 calibration injections [triplicates of the 5 standards] followed by 60 tracer samples.

16.2.3 Aquifer Characterization Methods

16.2.3.1 Permeameter analysis of core samples

We first determined the *K* of the sands from the horizontal cores obtained during the placement of ports. The extracted cores had dimensions of 1.28 cm in diameter and 10.16 cm in length. These cores were then attached to a custom-made constant head permeameter [Klute and Dirksen, 1986] for determination of *K*. Details of the core extraction method and the design of the constant head permeameter is provided in Craig [2005]. The *K* values from cores are calculated using Darcy's law.

16.2.3.2 Dipole cross-hole tests

Cross-hole hydraulic tests were conducted by pumping water from one well and injecting the pumped water at another location [dipole cross-hole tests from now on] to characterize the synthetic heterogeneous aquifer in the sandbox. We note that these tests are new and different from the cross-hole pumping tests conducted using a single well that were previously reported by Illman et al. [2010a].

In the dipole cross-hole tests, a mass balance of water injected and extracted was maintained using a peristaltic pump by connecting the injection and extraction ports in a single loop. Eight pairs of ports consisting of tests 1 through 8 [e.g., test 1: extraction at port 2 and injection at port 47, from now on E42/I47; test 2: E42/I7; test 3: E4/I45; test 4: E15/I34; test 5: E24/I25; test 6: E17/I32; test 7: E23/I26; and test 8: E20/I29] were chosen for these dipole cross-hole tests [see Figure 1 for port locations]. Eight additional tests [tests 9-16] were also conducted to collect data to evaluate the hydraulic tomography results that we discuss later. Injection and pumping rates for all dipole cross-hole tests averaged 480 mL/min. Prior to each cross-hole pumping test, all pressure transducers were calibrated to ensure accurate data collection. We then collected hydraulic head data for several minutes in all pressure transducers to establish a static, initial condition. After establishment of static conditions, we pumped from each port using a peristaltic pump, injecting the pumped water into another port, while taking head measurements at all 48 ports. For each test, pumping continued until the development of steady state conditions, which was determined by observing the stabilization of all head measurements on the data logger connected to a computer. After reaching steady state, the pump was shut off to collect recovery head data until its full recovery. Transient head data were also collected, but in this study, we only utilized steady-state head data for steady state hydraulic tomography to estimate *K* tomograms.

16.3 HETEROGENEITY CHARACTERIZATION METHODS

16.3.1 Geostatistical analysis and kriging of core *K* estimates

Geostatistical analysis of core *K* data was conducted using the Surfer 8 software developed by Golden Software, Inc [www.goldensoftware.com]. The exponential variogram model was used to fit the experimental variograms in both horizontal and vertical directions, resulting in an anisotropic variogram model. The variogram parameters fit to the experimental variograms include the geometric mean [K_G = 0.08 cms^{-1}], the variance [$\sigma^2_{\ln K}$ = 0.87] and the correlation lengths [λ_x = 17.00 cm and λ_z = 7.73 cm].

We then kriged the core *K* estimates using the anisotropic variogram model. The kriged domain was discretized into 73 vertical elements and 162 horizontal elements. The search ellipsoid measured 60 cm by 60 cm. Figure 3 shows the kriged *K* field which delineates some of the major layers but definitely has a smoothed appearance.

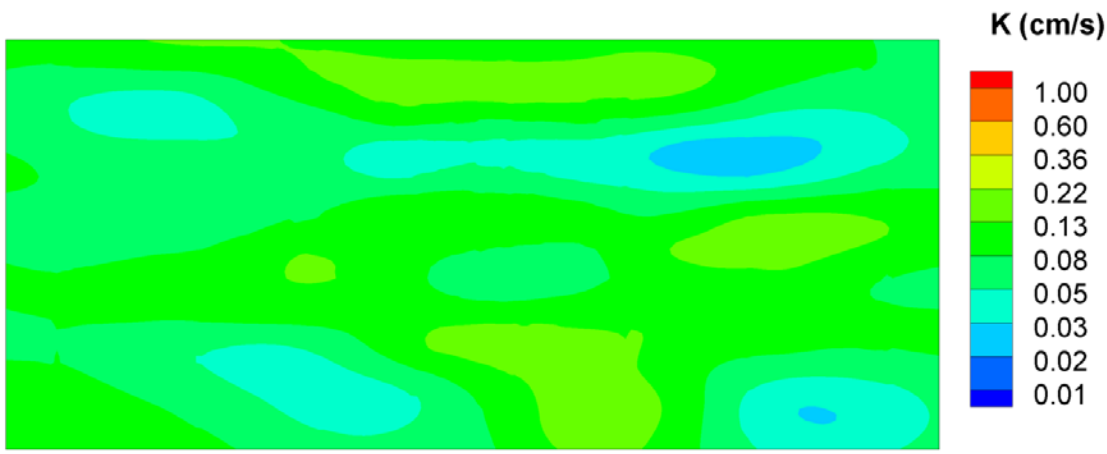


Figure 16.3: *K* distribution obtained through kriging of core scale K data from permeameter analysis.

16.3.2 Estimates of effective hydraulic conductivity and macrodispersivity

The development of stochastic theories for application to solute transport in heterogeneous media [e.g., Gelhar and Axness, 1983; Dagan, 1982, 1984; Neuman et al. 1987] has led to a methodology for predicting the bulk effective flow and transport parameters from a knowledge of the geostatistical parameters describing the spatial variations of the underlying *Ln[K]* process. The use of effective flow and transport parameters should allow the replacement of the real heterogeneous aquifer with a macroscopic homogeneous equivalent for the purposes of making the large-scale predictions of groundwater flow and contaminant migration.

The relationship between the asymptotic macrodispersivity tensor [A_{ii}, i = 1, 2, 3] and the geostatistical parameters describing the spatial variability of the hydraulic conductivity field have been derived for a variety of hydrogeological settings by Gelhar and Axness [1983] under the assumption of uniform flow. For the case of statistical anisotropy in the vertical and horizontal directions where $\lambda_1 > \lambda_2 > \lambda_3$ [Case 2, Gelhar and Axness, 1983], the longitudinal macrodispersivity A_{11} is given by

$$A_{11} = \sigma_Y^2 \lambda_1 \lambda_2 \Big/ \left[\xi^2 \left(\lambda_1^2 \sin^2 \phi + \lambda_2^2 \cos^2 \phi \right)^{1/2} \right] \qquad [16.1]$$

where

$$\xi = \exp\left[\sigma_Y^2 \left(0.5 - g_{22} \right) \right] \Big/ \left(\sin^2 \phi + \bar{K}_{22} / \bar{K}_{11} \cos^2 \phi \right) \qquad [16.2]$$

$$A_{11} = \sigma_Y^2 \lambda_1 \qquad [16.3]$$

and ϕ is the angle in the horizontal plane between the mean flow direction and the longitudinal axis of the effective hydraulic conductivity tensor [$\bar{K}_{ii}$]. It should be noted that a longitudinal macrodispersion coefficient calculated using equation [16.3] must be augmented by the value of the local longitudinal dispersion coefficient. When the mean flow direction coincides with the λ_1 direction [ϕ = 0°], the transverse macrodispersivity values are zero, thus indicating that the transverse macrodispersion process is controlled by local transverse dispersion. The effective hydraulic conductivity tensor used above is given by

$$\bar{K}_{ii} = K_g \exp\left[\sigma_Y^2 \left(0.5 - g_{ii} \right) \right] \qquad [16.4]$$

where the flow integrals, g_{ii} are functions of the correlation lengths defined by Gelhar and Axness [1983].

Using the best-fit estimate values for λ_1 and λ_2 equal to 17.00 cm and 7.73 cm, respectively, and a variance σ_Y^2= 0.87, as estimated from the two-dimensional variogram analysis, the principal values of the effective hydraulic conductivity tensor are given as K_{11} = 0.09 cm/s and K_{22} = 0.07 cm/s for the permeameter estimates of K. The computed value of the longitudinal macrodispersivity using equation [16.3] equals 14.76 cm on the basis of the geostatistical parameters derived from the two-dimensional variogram analysis.

Because our tracer experiment was conducted as a dipole test, strictly speaking, the Gelhar and Axness [1983] solution is not applicable as the solution was derived under the assumption of uniform flow. Tiedeman and Hsieh [2004] have shown that equal strength two well tracer test produced dispersivity estimates that are comparable to those from a natural gradient tracer test when the heterogeneity in hydraulic conductivity was low [i.e., $\sigma_Y^2 = 1.0$ and the separation distance between the injection and extraction wells was relatively large so that sufficient mixing of the tracers can take place. Available estimates of σ_Y^2 from the geostatistical analyses of core

and single-hole K estimates yielded values of 0.87 and 0.38, respectively [see Table 16.2]. Likewise, Illman et al. [2010a] obtained a variance estimate of 1.12 from hydraulic tomography after inverting 8 cross-hole pumping tests. Therefore, based on these σ_Y^2 estimates, we anticipate that effective parameters determined using the Gelhar and Axness [1983] solution will be appropriate even under our experimental conditions.

In terms of the distance between the injection and extraction points, Tiedeman and Hsieh [2004] also provided comparisons between those determined using dipole and natural gradient tracer tests. Their study showed that as the distance between the injection and extraction points increased with respect to the correlation scale of the heterogeneity, the closer the estimate of the dispersivities were from the dipole and natural gradient tests.

The straight-line distance between the injection point [port 42] and the extraction point [port 7] for our synthetic aquifer is 113.8 cm. Table 16.2 provides estimates of correlation lengths obtained via geostatistical analysis of core samples and single-hole test data. Likewise, Illman et al. [2010a] also presented estimates of λ_1 = 18.00 cm and λ_2 = 7.00 cm which were obtained through the inversion of 8 cross-hole pumping tests. The estimates of correlation lengths from all three data sets are quite similar.

Because the tracers traverses along and across the layers, we compute an average correlation scale by taking the arithmetic mean of λ_1 = 18.00 cm and λ_2 = 7.00 cm obtained y Illman et al. [2010a]. This yields an averaged value of 12.5 cm. This value is approximately one-tenth of the straight line distance between the injection and extraction points. Based on the numerical experiments conducted by Tiedeman and Hsieh [2004], the dispersivity estimates will be somewhat smaller due to the short separation distance between the injection and extraction wells. We also anticipate that the boundary effects may have an impact on the dispersivity estimates as it restricts flow and tracer transport. To overcome this difficulty, we later show numerical simulation results with macrodispersivity estimates lowered by 30% from those computed by the Gelhar and Axness solution to examine the sensitivity of the macrodispersivity on tracer transport results.

16.3.3 Hydraulic tomography analysis of dipole cross-hole tests

16.3.3.1 Inverse modeling approach

The steady state hydraulic tomography analysis of dipole cross-hole tests in the sandbox was conducted using a sequential geostatistical inverse approach developed by Yeh and Liu [2000]. We only provide a brief description of the inversion approach here. The inverse model assumes a steady flow field and the natural logarithm of K [ln K] is treated as a stationary stochastic process. The model additionally assumes that the mean and correlation structure of the K field is known *a priori*. The algorithm essentially is composed of two parts. First, the Successive Linear Estimator [SLE] is employed for each cross-hole test. The estimator begins by cokriging the initial estimate of K_{eff} and observed heads collected in one pumping test during the tomographic sequence to create a cokriged, mean removed ln K [f, i.e., perturbation of ln K] map.

Cokriging does not take full advantage of the observed head values because it assumes a linear relationship [Yeh and Liu, 2000] between head and K, while the true relationship is nonlinear. To circumvent this problem, a linear estimator based on the differences between the simulated and observed head values is successively employed to improve the estimate.

The second step of Yeh and Liu [2000]'s approach is to use the hydraulic head data sets sequentially instead of including them simultaneously in the inverse model thus the term,

"sequential successive linear estimator" [SSLE] is used to describe the inverse algorithm hereon. In essence, the sequential approach uses the estimated *K* field and covariances, conditioned on previous sets of head measurements as prior information for the next estimation based on a new set of pumping data. This process continues until all the data sets are fully utilized.

To obtain a *K* tomogram from the available dipole cross-hole tests, we solve an inverse problem for steady-state flow conditions. Boundary conditions were set to be no-flow for all sides. We created a fine grid with the synthetic aquifer discretized into 11,826 elements and 24,124 nodes with element dimensions of 1.0 cm × 10.2 cm × 1.0 cm.

Inputs to the inverse model include initial guesses for the *K*, estimates of variances and the correlation scales for both parameters, volumetric discharge [Q_n] from each pumping test where *n* is the test number, available point [small-scale] measurements of *K*, as well as steady-state head data. Although available point [small-scale] measurements of *K* can be input to the inverse model, we do not use these measurements to condition the estimated parameter fields to test the inversion algorithm.

A number of methods can be used to obtain the initial guess of *K*. One can set an arbitrary value that is reasonable for the geologic medium considered or to estimate the average or effective hydraulic conductivity [K_{eff}] for an equivalent homogeneous sandbox. Here, we elect to utilize the mean value of the *K* obtained from the analysis of core *K* estimate given earlier.

The variances and correlation scales of the *K* field are also required inputs to the inverse model. However, estimation of variance always involves uncertainty. A previous numerical study conducted by Yeh and Liu [2000] has shown that the variance has negligible effects on the estimated *K* using the inverse model. Therefore, we obtain variance estimates from the available small scale data given earlier and use this as our input variance in the inverse model for the real data set.

Correlation scales represent the average size of heterogeneity, which is difficult to determine accurately without a large number of data sets in the field. The effects of uncertainty in correlation scales on the estimate based on the tomography are negligible because the tomography produces a large number of head measurements, reflecting the detailed site-specific heterogeneity [Yeh and Liu, 2000]. Therefore, the correlation scales were approximated based only on the average thickness and length of the discontinuous sand bodies. For this analysis, we approximated the λ_1 and λ_2 to be 40.00 cm and 5.00 cm as an initial guess, respectively, based on the average thickness and length of the discontinuous sand bodies in this sandbox.

For the steady state hydraulic tomography analysis, we selected 8 dipole cross-hole tests [test 1: E2/I47; test 2: E42/I7; test 3: E4/I45; test 4: E15/I34; test 5: E24/I25; test 6: E17/I32; test 7: E23/I26; and test 8: E20/I29] and the corresponding steady-state head observations at the rest of 46 ports during each test as data sets. The steady–state head value from each port was obtained by averaging the steady-state portion of the record. We elected to not use the head data from the injection and extraction ports from each test because these ports could be affected by skin effects [Illman et al., 2007].

Prior to the computation of the *K* tomogram with the SSLE algorithm, we pre-processed the hydraulic head data. This is because Illman et al. [2008] found from the analysis of cross-hole pumping test data obtained in a laboratory sandbox aquifer that the signal-to-noise ratio can be critical in inverse modeling of cross-hole pumping test data. Despite calibrating the pressure transducer prior to the start of each pumping test, there is a minute level of drift always present in each of the pressure transducers. Therefore, we removed this drift by shifting the head value to a common one for all transducers in the sandbox prior to each test. The collected data were

additionally processed by taking the average of the steady state head values collected over a period of time in a given monitoring port.

All computations for steady state hydraulic tomography analyses were executed using 44 of 48 processors on a PC-cluster consisting [of 1 master and 12 slaves each with Intel Q6600 Quad Core CPU running at 2.4 GHz with 16 GB of RAM per slave] at the University of Waterloo. The operating system managing the cluster was CentOS 5.3 based on a 64-bit system. The total computational time for inverting data from 8 pumping tests was about 330 minutes [5.5 hours].

Figure 16.4 is the *K* tomogram obtained by inverting the steady state head data from 8 dipole cross-hole tests. Noteworthy statistical properties of this *K* tomogram include the geometric mean [K_G = 0.15 cms^{-1}], the variance [$\sigma^2_{\ln K}$ = 1.70] and the correlation lengths [λ_x = 60 cm and λ_z = 20 cm]. These values are somewhat higher than those presented in a previous study using the same sandbox [Illman et al 2010a]. Some key differences in experimental design may be responsible for this difference: 1] This study uses a dipole cross-hole test configuration where the previous study used single well cross-hole tests; 2] the pumping locations used in this study are different than the previous study resulting in different regions of the aquifer being stressed to different degrees; and 3] all of the boundary conditions in this study are no flow, whereas three of the boundary conditions [top, and both ends] were constant head in the previous study. The most notable difference in the *K*-tomogram from the two studies is that the tomograms generated in this study have a large high *K* zone near the top of the aquifer. This appears to be dominating the geostatistical analysis resulting in longer estimated correlation lengths and a higher geometric mean [K_G]. The tomograms from both this study and the previous study were tested against independent dipole cross-hole tests and the tomogram generated from this study performed the best indicating that for this particular case/configuration this is the most accurate tomogram to use.

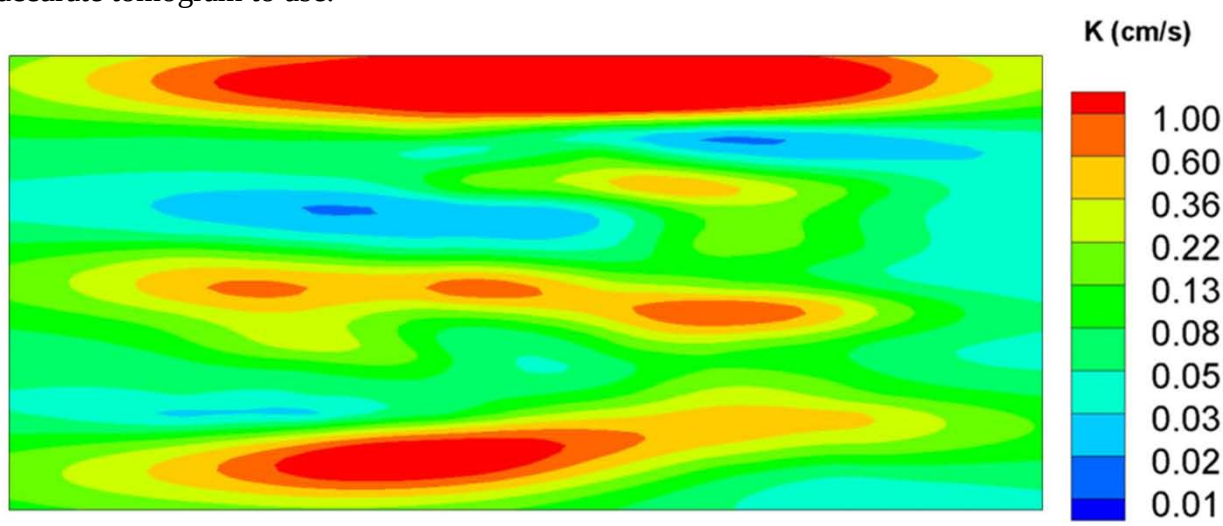


Figure 16.4: *K* tomograms computed using the steady state hydraulic tomography algorithm of Yeh and Liu [2000].

To quantitatively assess the results, we plotted the simulated and measured drawdown responses for each dipole cross-hole test used for the inverse modeling effort [i.e., calibration] on separate scatter plots and a linear model was fit to each case without forcing the intercept to zero.

The linear model fit and the coefficient of determination [R^2] provide indications of scatter and bias. The R^2 is a statistic that provides a quantitative measure of similarity between the simulated and measured drawdown values. A high R^2 value means that the simulated and measured drawdown values are linearly correlated, even though the mean values could be different. The slope and intercept of the linear model fit, and the R^2 values are summarized in Table 16.2.

Table 16.2: Results of linear model fit to simulated versus observed drawdowns for coarse and fine grids. Simulated values were obtained during the inverse modeling of these tests [i.e., calibration].

Test	Slope	Intercept	R^2
E2_I47	1.00	-0.11	0.96
E42_I7	0.84	0.11	0.93
E4_I45	1.01	-0.02	0.97
E15_I34	1.13	0.16	0.89
E24_I25	1.01	0.05	0.96
E17_I32	1.04	0.03	0.95
E23_I26	1.03	0.04	0.95
E20_I29	0.98	0.05	0.91

16.3.3.1 Validation of the *K* tomograms using additional tests

Illman et al. [2007, 2008, 2010a] and Liu et al. [2007] previously found that the best way to test the results from hydraulic tomography was to use the hydraulic parameter distributions in predicting one or more independent drawdown inducing events. Therefore, to further assess reliability of the *K* tomograms, we simulated 8 additional dipole cross-hole tests [no. 9 – 16] [i.e., test 9: E8/I41; test 10: E9/I40; test 11: E10/I39; test 12: E11/I38; test 13: E12/I37; test 14: E16/I33; test 15: E18/I31; and test 16: E30/I19] using the forward groundwater model MMOC3 [Yeh et al., 1993]. The results were then assessed by plotting individual scatterplots of simulated and measured drawdown values and fitting a linear model as described earlier. Results summarized in Table 16.3 suggest that the K tomogram is able to accurately predict independent cross-hole dipole tests. These statistics collectively provide quantitative measures of the performance of each steady-state forward groundwater model with different *K* fields in simulating the 8 dipole cross-hole tests. The use of more than one dipole cross-hole test conducted at different parts of the synthetic heterogeneous aquifer ensures a more credible validation of each of these models.

Table 16.3: Results of linear model fit to simulated versus observed drawdowns for coarse and fine grids. Simulated values were obtained by modeling independent dipole cross-hole tests not used in the construction of the tomograms [i.e., validation].

Test	Slope	Intercept	R^2
E8_I41	1.01	0.98	0.94
E9_I40	1.05	0.35	0.95
E10_I39	1.12	0.64	0.93
E11_I38	1.03	0.43	0.93
E12_I37	1.02	0.31	0.94

E16_I33	1.08	0.90	0.91
E18_I31	1.03	0.28	0.95
E30_I19	0.95	0.02	0.95

16.4 RESULTS: DIPOLE TRACER TEST

We next report on the results from the dipole tracer test by first examining the evolution of the tracer plume throughout the experiment. To achieve this, numerous photographs were taken during the dipole tracer test. Figure 16.5 shows six such photos taken at t = 12 min, 27 min, 62 min, 100 min, 175 min, and 286 min. At t = 0 min, the injection of the tracer solution began and at t = 10 min, the injection was switched from tracer solution to tap water.

The photographs are utilized for qualitative assessment of the tracer migration. Figure 16.5a was taken at t = 12 min, 2 minutes after the injection of the tracer solution at port 42 completed in layer 5 ended. Layer 5 consisted of a mixture of equal volumes of #14, F75, and 16/30 sands. We notice from this figure that the tracer solution uniformly distributes in layer 5. However, this figure also shows that the tracer solution is preferentially transported along layer 4 [20/40] which has a K of 2.05×10^{-1} cm/s.

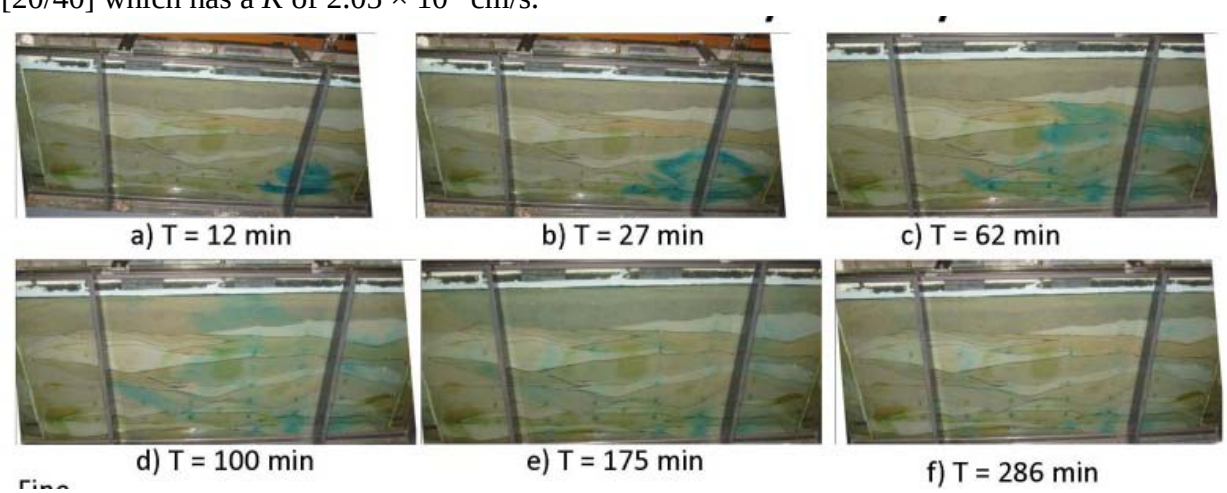


Figure 16.5: Photographs of sandbox during tracer test 1 at various times: a] t = 12 min; b] t = 27 min; c] t = 62 min; d] t = 100 min; e] t = 175 min; and f] t = 286 min.

At 27 min. [Figure 16.5b], 17 minutes after the injection was switched from tracer solution to tap water, we notice that the blue dye becomes cleared around the injection port. Figure 16.5b shows that the tracer continues to move rapidly through layer 4. The tracer solution also moves laterally and upwards through layer 8 [F32], layer 9 [20/40] and into layer 10 [F65].

At 62 min. [Figure 16.5c], we notice that the tracer solution becomes more diluted through the injection of tap water and the visual observation that the hue of the blue dye lessening. We also notice that the tracer has migrated into layers 13 [20/30] and 15 [20/40]. It also begins to migrate into layer 17 [F85] which has the lowest K in the aquifer at 1.35×10^{-2} cm/s.

At 100 min. [Figure 16.5d], the tracer appears to breakthrough layer 17 [F85] vertically and reaches layer 18 [20/30]. Because of the high K of the 20/30 sand [3.12×10^{-1} cm/s], the tracer then rapidly migrates laterally through layer 18. The tracer also appears to migrate rapidly through layer 7 [#12] which has a K of 2.05×10^{-1} cm/s.

By 175 min. [Figure 16.5e], the tracer appears to reach port 7 which is the extraction port that is completed in layer 14 [F75]. We notice that despite the fact the tracer solution appears to be diluted, the tracer plume has reached most parts of the aquifer and all of the layers between the injection and extraction ports.

By 286 min. [Figure 16.5f], we observe that the tracer has been removed from the high *K* layers. However, there is a noticeable amount of blue dye in low *K* layers including: layer 3, layer 8, layer 10, layer 14, and layer 17. This suggests that the low *K* layers can contribute to store tracers for a long period requiring a long period to flush the tracers out of the synthetic aquifer.

Figure 16.6 shows normalized concentration breakthrough curves of the bromide tracer at the 12 sampling ports [2, 10, 12, 14, 19, 22, 27, 30, 35, 37, 39, 47] and at the extraction port [7]. On Figure 16.6, we also note the arrival times [t_a], peak concentration [C_p], and the time when the normalized concentration reaches background levels [t_b].

The breakthrough curves obtained from different ports [Figure 16.6] illustrate the differences in them. For example, breakthrough curves at ports [30, 35, and 47] that are a short distance away from the injection port, show early arrival of tracers, sharp peaks and a rapid return to background concentrations. The normalized peak concentrations range from 0.82 – 0.99 and qualitatively, the breakthrough curves look similar.

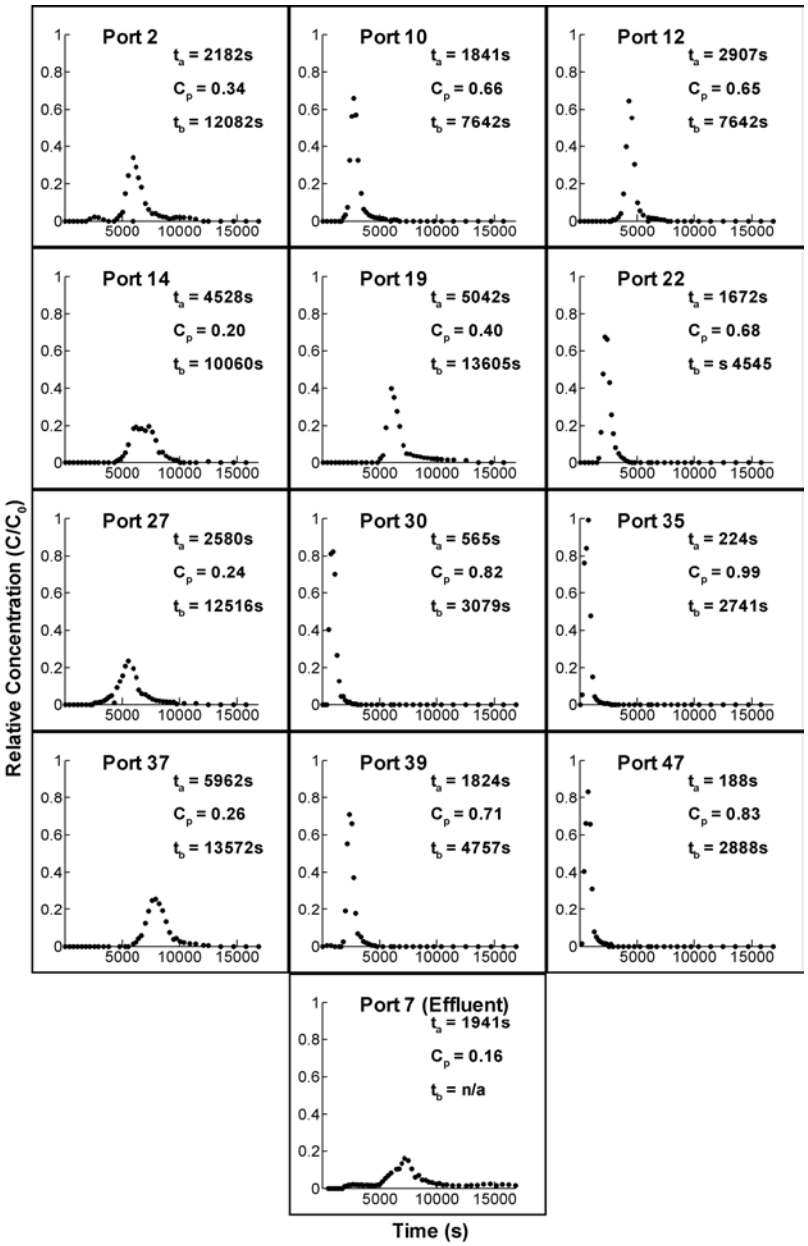


Figure 16.6: Normalized concentration breakthroughs at the 12 sampling ports [2, 10, 12, 14, 19, 22, 27, 30, 35, 37, 39, 47] and the extraction port [7]. The arrival times [t_a], peak concentration, [C_p] and the time when the normalized concentration reaches background levels [t_b] is also noted.

For the ports [10, 12, 22, 27, 37, and 39] that can be considered to be at intermediate distances, the breakthrough curves can be separated into two groups. In one group of ports [10, 12, 22, and 39], the normalized peak concentrations are generally higher with values ranging between 0.65 and 0.71 in comparison to the other group of ports [27 and 37] with values ranging between 0.24 and 0.26. The breakthrough curve from ports 27 and 37 have a lower peak, but is more spread out. In both groups, the breakthrough curves all appear to have a Gaussian

distribution. We also note that the recovered mass is less at these intermediate distance ports because sampling is taking place continuously.

Examination of breakthrough curves from ports [2, 14, and 19] that are farthest away from the injection port shows a different behaviour at least for ports 2 and 19. In particular, the breakthrough curve at port 2 suggests the presence of 3 peaks with the first one arriving approximately at 2,563 sec, the second at about 5,962 sec, and the last one at about 9,738 sec. Given the low noise level of the concentration measurements, we think that these are real peaks. This type of multi peak behaviour has been observed in fractured rocks and in highly heterogeneous porous media.

Port 19 also shows a different breakthough curve in which there is only a single peak, but it exhibits a long tail with concentrations reaching background levels at approximately 13,605 sec. This suggests a non-Gaussian transport behaviour at the sampling port.

The extraction port also reveals multiple peaks, one arriving at 2,661 sec, the second one at 7,174 sec, and a third one at 14,625 sec. We notice a considerable tailing of concentrations after the second peak and the concentration does not reach background levels at the conclusion of the tracer experiment [~18,000 sec].

16.5 PREDICTION OF DIPOLE TRACER TEST

16.5.1 Description of tracer transport simulations

A two-dimensional, saturated flow and transport model of the synthetic aquifer was developed using the finite-element code MMOC3 [Yeh et al., 1993]. The purpose of this modeling effort was to compare drawdowns and solute concentrations during the dipole tracer tests at various sampling points. We conducted the simulation using an effective parameter approach and a heterogeneous K field using results from hydraulic tomography. For the effective parameter approach, we utilized the effective hydraulic conductivity [K_{eff}] to simulate groundwater flow during the dipole tracer test, and a longitudinal macrodispersivity [A_{11}] to simulate the migration of the conservative tracer. Both the K_{eff} and A_{11} were estimated from the stochastic theory of Gelhar and Axness [1983]. To address the issue that dipole tracer tests conducted in a bounded aquifer could potentially yield smaller dispersivity estimates, we also conducted an additional simulation with macrodispersivity reduced by 30% to examine its sensitivity to tracer transport. We also simulated groundwater flow and plume migration using the kriged K field and the K tomogram. For the transport simulation using the kriged K field and the K tomogram, we set the dispersivity value equal to zero, while the effective porosity was set to 0.36 for all cases. In total, 3 different cases were considered.

For all cases, flow conditions were established by simulating steady state groundwater flow with injection, extraction, and sampling ports. A tracer solution with a bromide concentration of 151.3 mg/L is then injected for 10 minutes and the solute transport is simulated under transient conditions. As in the experiment, all boundaries are set as no-flow and no-flux boundaries.

16.5.2 Groundwater flow results

We first modeled the steady-state drawdown values during the dipole tracer test at the sampling ports for the effective and heterogeneous K distributions. Figure 16.7 shows scatter plots of simulated versus observed drawdowns for the 3 cases that we examined. In particular, Figure 16.7a shows the comparison when the K_{eff} computed using the Gelhar and Axness [1983] solution is used to simulate the drawdown behaviour. Likewise, Figure 16.7b shows the same

when the kriged K field is used for the forward simulation. These comparisons show that the K_{eff} and kriged K field both yield biased predictions of drawdowns at various locations. In addition, we observe a significant scatter suggesting that the K_{eff} and kriged K field may not be suitable for predicting drawdowns within this synthetic aquifer. In contrast, Figures 16.7c shows a much improved prediction of steady state drawdowns using the K tomogram. There is some scatter and bias, but it is evident that the K tomogram yields considerably improved results in terms of drawdown predictions in comparison to the effective parameter approach.

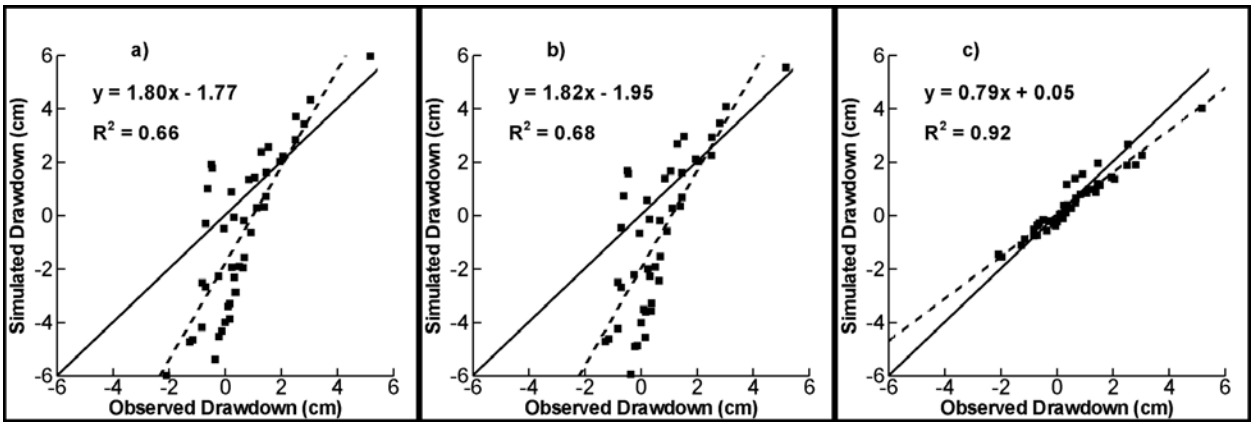


Figure 16.7: Simulated versus observed drawdowns at 48 ports during the dipole tracer test. Simulated values were computed using: a] K_{eff} from Gelhar and Axness [1983] solution with statistics of permeameter analysis of core samples; b] kriged K field; and e] K tomogram computed using fine mesh.

16.5.3 Dipole tracer transport results

We next modeled the tracer migration using the various effective and heterogeneous parameter fields. Figure 16.8 shows the concentration distributions from the various tracer transport simulations. In particular, Figure 16.8a shows the spatial distribution of tracer concentrations at various times using the macrodisperisivity estimates obtained via the Gelhar and Axness [1983] solution. The results reveal that the concentration distribution evolves quite uniformly with significant spreading of the tracer concentration. Not shown here are results of simulations in which we reduced the macrodispersivity by 30%. Surprisingly, reducing macrodispersivity by 30% did not influence our results confirming the findings by Sudicky et al. [2010].

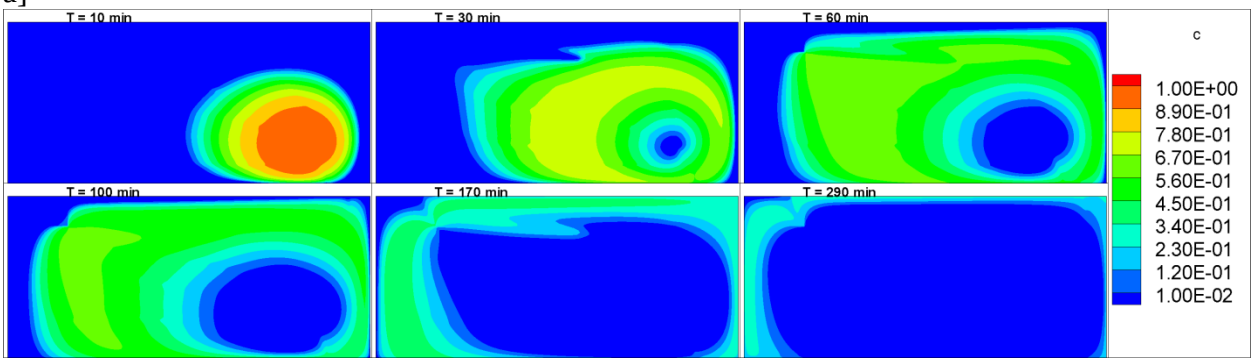
Accounting for heterogeneity in K results in nonuniform concentration distributions. In particular, Figure 16.8b is the result using the kriged K field while Figure 16.8c is the result based on the K tomogram. We note that dispersivity was set to zero for both cases. Both Figures 16.8b and 16.8c show that due to the mapping of the heterogeneity, the concentration distributions are less smooth in comparison to the effective parameter case [Figure 16.8a]. These figures reveal a much more heterogeneous distribution of tracer concentrations with tracers following preferential pathways with hydraulic tomography providing a more heterogeneous concentration distribution in comparison to kriging.

We next compare the results from Figure 16.8a through 16.8c with Figure 16.5a through 16.5c. The time when the photographs were taken and the transport simulation results are not exactly coincident, but the two figures nonetheless can be compared qualitatively. The comparison shows that the migration of the tracer represented by the migration of the dye is

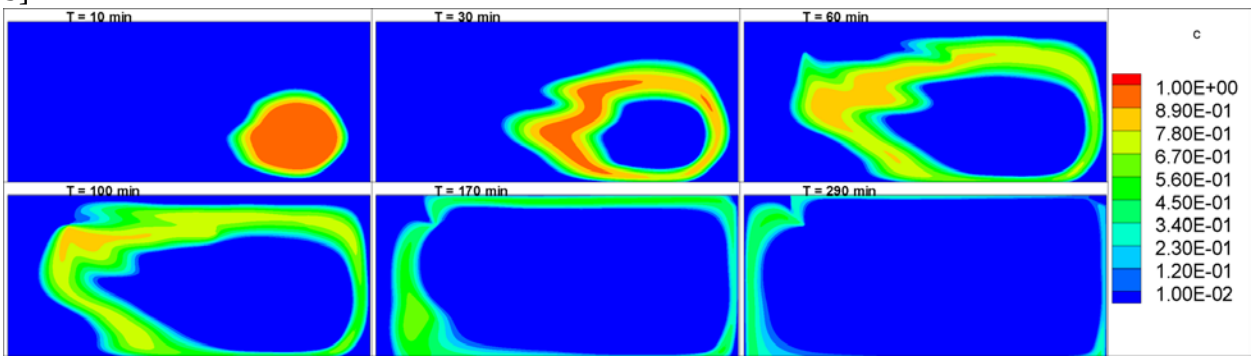
better represented by Figure 16.8b and best represented by Figure 16.8c. This is expected as the transport simulations involving effective parameters do not consider the *K* heterogeneity, while kriging and steady state hydraulic tomography yields increasingly realistic distributions of *K* heterogeneity that can better represent the migration of tracers.

One significant point to make is that while the effective parameter approach requires an effective hydraulic conductivity and macrodispersivity coefficient obtained from the geostatistical analysis of a large number of core or single-well hydraulic test, hydraulic tomography can yield a *K* distribution that does not require the collection of a large number of small scale samples as well as costly tracer experiments to obtain estimates of dispersivity.

a]



b]



c]

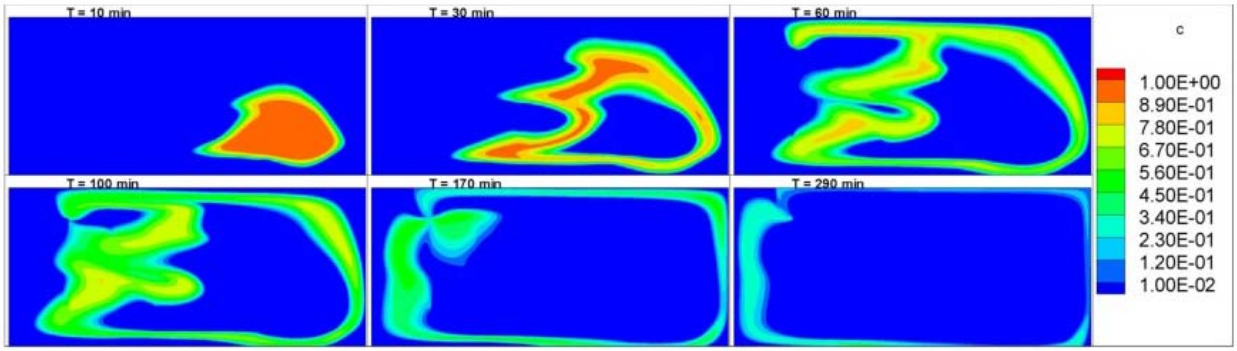


Figure 16.8: Concentration distributions from tracer transport simulation with: a] K_{eff} and A_{11} computed using Gelhar and Axness [1983 solution with statistics of permeameter analysis of core samples [case 1]; b] kriged *K* field [case 2]; and c] *K* tomogram [case 3].

16.5.3.1 Simulated versus observed breakthrough curves

We next make a direct comparison of the breakthrough curves obtained through numerical simulations and the actual tracer data [Figure 16.9]. The matches of the breakthrough curves obtained from ports [30, 35, and 47] close to the injection port is quite good for all cases. In particular, the arrival time, peak concentrations, and time for tracer concentrations to reach background levels is quite consistent for all cases.

Differences in the quality of matches begin to emerge when the breakthrough curves from the intermediate distance ports [10, 12, 22, 27, 37, and 39] are examined. We observe that case 1 [effective parameters] consistently underpredicts the peak concentration at all ports in this category and show an earlier arrival of tracers. In contrast, cases 2 [kriging] and 3 [*K* tomogram] show a marked improvement in the quality of the fits with case 3 performing the best visually, on the average.

Examination of the furthest ports [2, 14, 19] from the injection port shows mixed results. In general, we observe that case 3 performs better than cases 1 and 2 in terms of better predicting the peak concentrations and arrival times of tracers. However, we also note that the matches are far from perfect.

At the extraction port [7], the results are again mixed, although we find case 3 to capture the first 2 peaks that arrive at different times. Case 1, on the other hand, because it is based on the effective parameter approach, cannot capture the multiple peak behavior. Likewise, case 2 with the smoother *K* field does not capture this behavior. Because the visual comparisons of the simulated versus observed break through curves are qualitative, we next compute and compare the temporal moments.

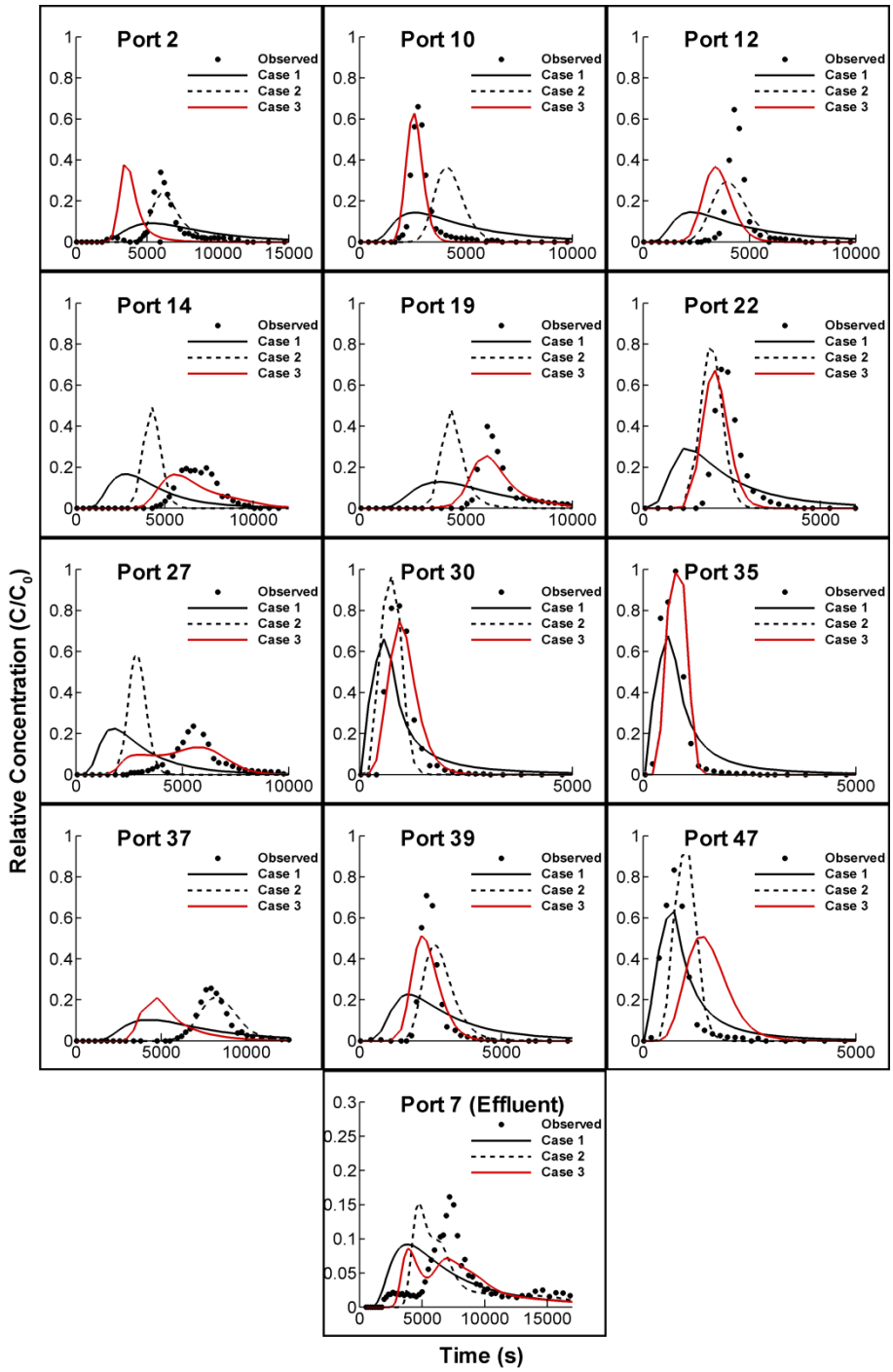


Figure 16.9: Breakthrough curves from the dipole tracer test and corresponding match of forward simulations. Tracer transport simulations were performed using various *K* fields.

16.5.3.2 Temporal moment analysis

The method of moments for temporal concentration distribution was used to characterize the bromide breakthrough data at all wells. The nth temporal moments [M_n] of concentration [C] at location [x, y] at time [t] are given by:

$$M_n = \int_0^\infty t^n C(x, y, t) dt \qquad [16.5]$$

where t is time, $C(x_i, t)$ is tracer concentration. One can compute the zeroth [M_0], first [M_1], and second [M_2] temporal moments using the moment generating function [eqn. 5] by setting $n = 0$, 1 and 2 respectively. The temporal moments were obtained through numerical integration of the breakthrough data using the trapezoidal rule.

The total mass of solute passing through the sandbox at each sampling point is obtained by computing the M_0 from the breakthrough curves. The first normalized moment of breakthrough curves at each sampling ports were used to estimate the mean arrival time of the center of bromide mass [μ]:

$$\mu = \frac{M_1}{M_0}. \qquad [16.6]$$

The variance of the breakthrough curve is then calculated by,

$$\sigma = \frac{M_2}{M_0} - \left(\frac{M_1}{M_0}\right)^2. \qquad [16.7]$$

In general, the variance represents the spread of the concentration distribution and is influenced by mechanical dispersion and molecular diffusion.

Figures 16.10 through 16.12 show the M_0, μ, and σ computed from the temporal moment analyses of simulated and observed breakthrough curves. In particular, Figure 16.10 reveals the estimates of M_0 are, on average, higher for the simulated breakthrough curves for case 1. This suggests that there are discrepancies of total mass estimated at many of the water sampling points. One reason behind this is that in uniform media, the tracer mass spreads uniformly causing tracer mass to increase throughout the domain. However, in reality, the synthetic aquifer is heterogeneous which causes more tracer mass to arrive in some sampling intervals than others depending on the flow path and the heterogeneities that the tracer cloud encounters. Cases 2 and 3 suggest that when the heterogeneity in K is considered, the estimation of the total mass at a given sampling point improves.

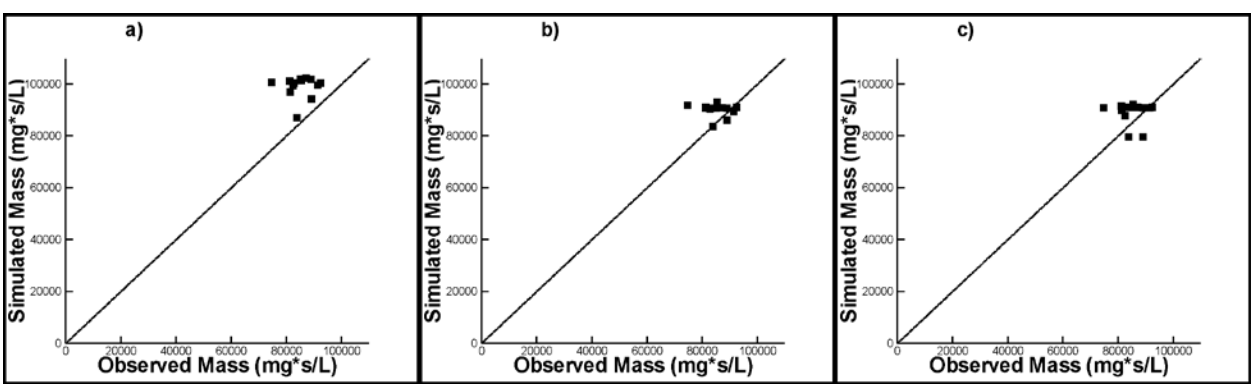


Figure 16.10: Total mass computed by the M_0 for each sampling port: a] case 1; b] case 2; and c] case 3.

The mean arrival time of the center of mass [μ] for the 6 cases is plotted on Figure 16.11. Case 1 shows that the arrival times simulated using the effective parameters for the ports close to the injection port is very similar to those observed. As the tracer cloud moves away from the injection point, we see the discrepancy between the simulated and observed values to increase. Similar behaviours are observed for cases 2 and 3 which suggest that the more detailed heterogeneity pattern included does not affect the prediction of the arrival time of the center of mass at large times. What is encouraging is that the scatter is centered around the 1:1 line for cases 1 and 2. Case 3, on the other hand, shows considerably less scatter. There are two sampling points 2 and 37 that causes the simulated versus observed relationship to be biased at larger times. These two ports are at a large travel distance and near boundaries. If we choose to disregard these two points, then the relationship between the simulated and observed values becomes quite similar. This suggests that the mapping of heterogeneity via hydraulic tomography yields more accurate estimates of the arrival times of the center of mass for this synthetic aquifer.

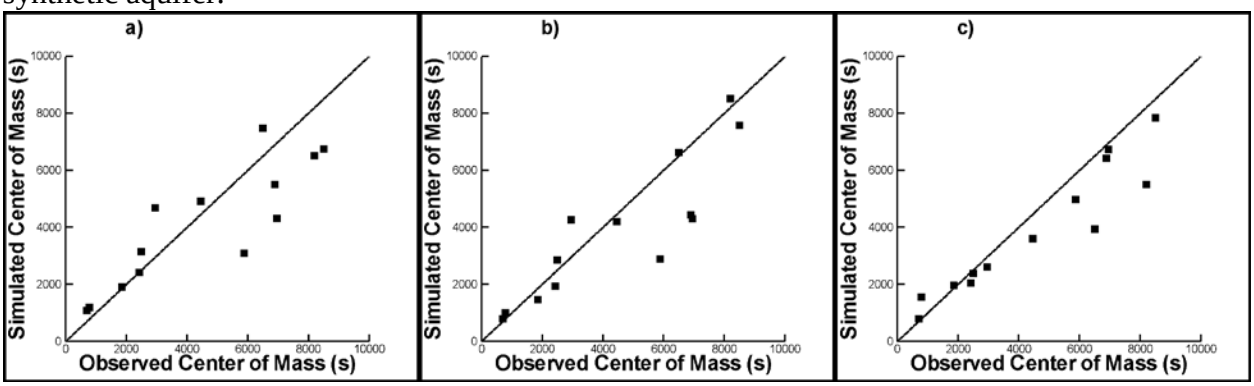


Figure 16.11: Mean arrival time of the center of mass [μ] computed for each sampling port: a] case 1, b] case 2, and c] case 3.

The variance of the breakthrough curves [σ] was also calculated at each sampling point and plotted on Figure 16.12. It is evident from case 1 that the simulated variances are considerably higher than the observed ones closer to the injection point. This suggests that the effective parameters used in the transport simulation may be overpredicting the temporal spreading of the plume. This may be the case as the transport regime maybe in a non-Fickian regime close to the injection point. The variance estimates at the extraction point for case 1 are, however, very good, suggesting that effective parameters can yield accurate estimates of the variance at the extraction point if the travel distance/time is long enough reaching a Fickian regime. When we examine case 2 in which the kriged *K* field is used, the bias in the variance estimates diminishes at sampling intervals close to the injection well. In contrast, a remarkable improvement is seen in case 3 in which both the bias and scatter diminish considerably. This suggests that hydraulic tomography is the best approach for predictions of tracer migration.

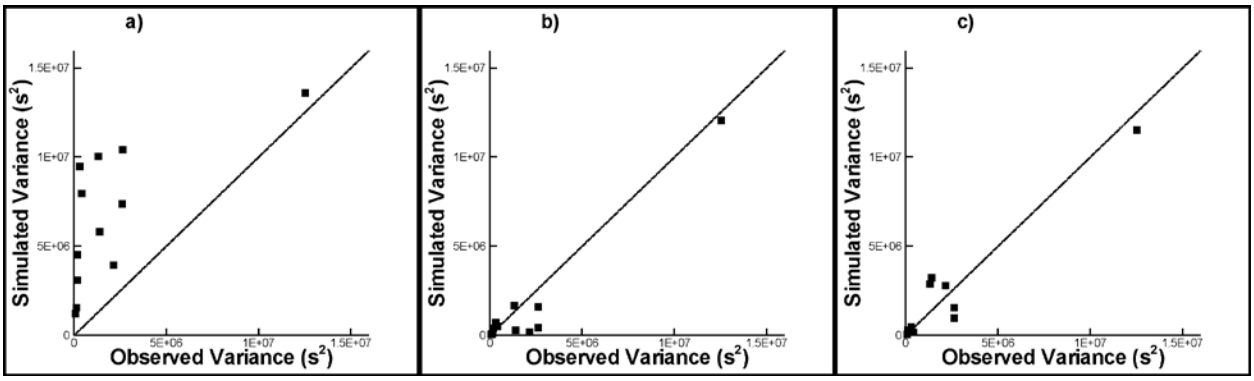


Figure 16.12: The variance of the breakthrough curve [σ] at each sampling port: a] case 1; b] case 2; and c] case 3.

16.6 DISCUSSION

The main purpose of this paper was to compare the accuracy of solute transport predictions based on numerical simulations conducted using various heterogeneity characterization methods. Results clearly showed the large improvement of predictions in terms of drawdown behavior and tracer transport when the *K* distribution obtained from hydraulic tomography was used. As stated by Yeh [1992] and demonstrated through field studies [e.g., Sudicky, 1986; Sudicky et al., 2010], the effective parameter concept is likely applicable in a large enough domain where the pumping or tracer tests are conducted so that the approach can be used to predict the bulk behavior of the system. However, we found that in this sandbox through forward numerical simulations that the effective hydraulic conductivity failed to provide accurate predictions of drawdowns during the dipole tracer test and the macrodispersivity used in the forward transport simulations could not predict solute transport behavior well. In addition, lowering of the macrodisperisvity by 30% did not affect our results. This suggests that macrodispersivity is relatively insensitive to the arrival time, peak concentration, and time to which concentration drops to a background level. This is in line with the sensitivity analysis of macrodispersivity conducted by Sudicky et al. [2010] based on data at the North Bay site in Canada.

The overall implication of our findings is that in a smaller flow and transport domain, these effective parameter approaches may not be applicable in predicting local scale or even large scale behaviors of drawdowns and solute transport. This necessitates the development of heterogeneity characterization and modeling approaches. Results from hydraulic tomography showed that the prediction of drawdowns during the dipole tracer test was quite good, and that both qualitative and quantitative evaluations of tracer transport revealed that the transport behavior can be captured quite well. However, we do note that the *K* distribution obtained from hydraulic tomography still could not capture all the details of the tracer breakthrough. This is likely due to the fact that the hydraulic tomography was done at a coarser scale than what an accurate prediction of tracer transport requires. That is, it appears that a finer characterization of the aquifer through hydraulic tomography or some other method is necessary in obtaining more accurate predictions of solute transport at the scale of the sandbox that we investigated. Finer scale characterization can be achieved in a number of ways. One approach is to place additional monitoring points along each of the boreholes. This will result in hydraulic tomography conducted at a higher resolution. Another approach may be to condition hydraulic tomography with other types of information such as geology, high resolution core and slug test data, as well as through geophysical tests. Yet another approach may be to utilize the tracer test data in the

inverse modeling effort to further refine the *K* distribution. This is obviously beyond the current manuscript but is a topic for further research in the future.

Finally, our study suggests that accurate predictions of solute transport can become possible without collecting a large number of small scale samples to estimate effective parameters or ascribing dispersivity estimates that are costly to obtain at the field scale through tracer tests. All one has to conduct are well designed hydraulic tomography surveys to capture the *K* heterogeneity pattern at a given site.

17. HYDRAULIC AND PARTITIONING TRACER TOMOGRAPHY EXPERIMENTS IN AN INTERMEDIATE SCALE SANDBOX

17.1 Introduction

Trichloroethene [TCE] and other dense nonaqueous phase liquids [DNAPL] are prevalent at a large number of sites throughout the world. The variable release history, unstable flow, and geologic heterogeneity make the spatial distribution of DNAPLs complex. This causes difficulties in site remediation contributing to long-term groundwater contamination for decades to centuries. One important reason for long cleanup times is the large uncertainty in DNAPL saturations in source zones. Significant improvement in remediation designs requires the accurate knowledge of hydraulic conductivity [K] heterogeneity and DNAPL saturation [S_N] distributions.

Illman et al. [2010b] presented small-scale laboratory experiments to demonstrate the efficacy of Sequential Successive Linear Estimator [SSLE] algorithm that images DNAPL source zones. The algorithm relies on the fusion of hydraulic and partitioning tracer tomography [HPTT] to derive the best estimate of the K heterogeneity, DNAPL saturation [S_N] distribution and their uncertainty. The HPTT approach is a non-destructive approach and can be applied repeatedly prior to and post-remediation. More importantly, it also provides uncertainty estimates that can facilitate better decision making on source zone characterization, remediation, and long-term monitoring. Results from their laboratory experiments showed that S_N distributions compare favorably with DNAPL distributions observed in the sandbox, but not so with local S_N estimates from core samples. They also found that the delineation of K heterogeneity can have a large impact on computed S_N distributions emphasizing the importance of accurate delineation of hydraulic heterogeneity.

Despite the success of these experiments, additional work was deemed necessary as only one tracer was used in the study of Illman et al. [2010b] and the heterogeneity pattern was very simple. Therefore, the main objective of this study was to conduct additional HPTT experiments in an intermediate scale sandbox with a considerably more complex heterogeneity pattern that was previously characterized by Illman et al. [2010b] and Berg and Illman [in review].

17.2 Sandbox and aquifer description

A synthetic heterogeneous aquifer was constructed in a sandbox to validate the HPTT algorithm of Yeh and Zhu [2007]. The sandbox is 193.0 cm in length, 82.6 cm in height, and has a depth of 10.2 cm. Forty eight ports, 1.3 cm in diameter, have been cut out of the stainless steel wall to allow coring of the aquifer as well as installation of horizontal wells. The wells fully penetrate the thickness of the synthetic aquifer. This allowed each location to be monitored by a pressure transducer, used as a pumping or an injection port and as a water sampling port. Figure 17.1 is a schematic drawing of the frontal view of the synthetic aquifer, showing the 48 port and pressure transducer locations.

In this study, the synthetic heterogeneous aquifer was created through the cyclic deposition of sediments under varying water flow and feed rates of sediments. Previous sandbox aquifer studies by our group [Illman et al., 2007, 2008; Liu et al., 2007; Yin and Illman, 2009] involved careful packing of heterogeneity patterns in a prescribed fashion. Packing of sand bodies by hand is a laborious procedure and complex heterogeneity patterns are hard to create.

Our goal in relying on sediment transport was to create a more realistic heterogeneity pattern with various scales of heterogeneity in an efficient manner. The one drawback of this approach is that the heterogeneity pattern in the sandbox cannot be controlled precisely as in the case of packing a sandbox aquifer by hand. Further details to the sandbox and the synthetic heterogeneous aquifer can be found in Illman et al. [2010a]

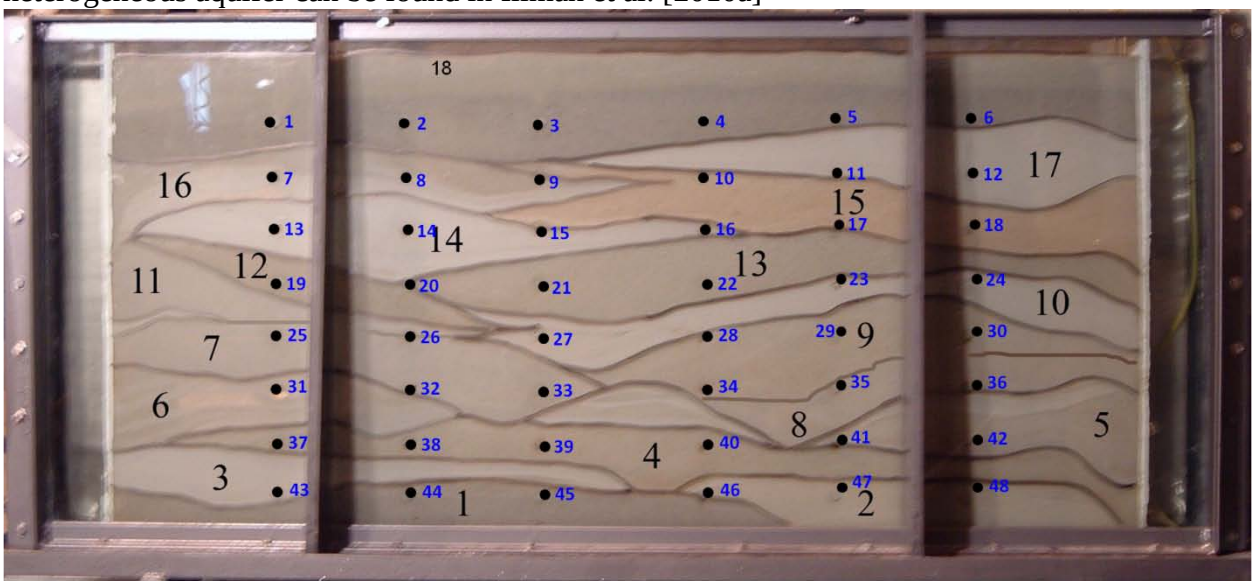


Figure 17.1: Photograph of the sandbox showing the synthetic heterogeneous aquifer in which the dipole cross-hole and conservative tracer tests were conducted. Large black numbers indicate layer numbers, solid circles indicate port locations, and small blue numbers indicate port numbers. Layer 1 = 20/30; layer 2 = 4030; layer 3 = F85; layer 4 = 20/40; layer 5 = mix; layer 6 = mix; layer 7 = #12; layer 8 = F32; layer 9 = 20/40; layer 10 = F65; layer 11 = #12; layer 12 = 16/30; layer 13 = 20/30; layer 14 = F75; layer 15 = 20/40; layer 16 = mix; layer 17 = F85; layer 18 = 20/30. Note: The layers labelled "mix" consisted of equal volumes of #14, F75, and 16/30 sands.

17.3 Dipole hydraulic tomography survey results

Cross-hole hydraulic tests were conducted by pumping water from one well and injecting the pumped water at another location [dipole cross-hole tests from now on] to characterize the synthetic heterogeneous aquifer in the sandbox. We note that these tests are new and different from the cross-hole pumping tests conducted using a single well that were previously reported by Illman et al. [2010a].

In the dipole cross-hole tests, a mass balance of water injected and extracted was maintained using a peristaltic pump by connecting the injection and extraction ports in a single loop. Eight pairs of ports consisting of tests 1 through 8 [e.g., test 1: extraction at port 2 and injection at port 47, from now on E42/I47; test 2: E42/I7; test 3: E4/I45; test 4: E15/I34; test 5: E24/I25; test 6: E17/I32; test 7: E23/I26; and test 8: E20/I29] were chosen for these dipole cross-hole tests [see Figure 1 for port locations]. Eight additional tests [tests 9-16] were also conducted to collect data to evaluate the hydraulic tomography results that we discuss later. Injection and pumping rates for all dipole cross-hole tests averaged 480 mL/min. Prior to each cross-hole pumping test, all pressure transducers were calibrated to ensure accurate data collection. We then collected hydraulic head data for several minutes in all pressure transducers to establish a static,

initial condition. After establishment of static conditions, we pumped from each port using a peristaltic pump, injecting the pumped water into another port, while taking head measurements at all 48 ports. For each test, pumping continued until the development of steady state conditions, which was determined by observing the stabilization of all head measurements on the data logger connected to a computer. After reaching steady state, the pump was shut off to collect recovery head data until its full recovery. Transient head data were also collected, but in this study, we only utilized steady-state head data for steady state hydraulic tomography to estimate *K* tomograms.

The steady state hydraulic tomography analysis of dipole cross-hole tests in the sandbox was conducted using a sequential geostatistical inverse approach developed by Yeh and Liu [2000]. Details to the analysis are described in Chapter 8. Figure 17.2 is the *K* tomogram obtained by inverting the steady state head data from 8 dipole cross-hole tests. Noteworthy statistical properties of this *K* tomogram include the geometric mean [K_G = 0.15 cms^{-1}], the variance [$\sigma^2_{\ln K}$ = 1.70] and the correlation lengths [λ_x = 60 cm and λ_z = 20 cm]. These values are somewhat higher than those presented in a previous study using the same sandbox [Illman et al 2010a]. Some key differences in experimental design may be responsible for this difference: 1] This study uses a dipole cross-hole test configuration where the previous study used single well cross-hole tests; 2] the pumping locations used in this study are different than the previous study resulting in different regions of the aquifer being stressed to different degrees; and 3] all of the boundary conditions in this study are no flow, whereas three of the boundary conditions [top, and both ends] were constant head in the previous study. The most notable difference in the *K*-tomogram from the two studies is that the tomograms generated in this study have a large high *K* zone near the top of the aquifer. This appears to be dominating the geostatistical analysis resulting in longer estimated correlation lengths and a higher geometric mean [K_G]. The tomograms from both this study and the previous study were tested against independent dipole cross-hole tests and the tomogram generated from this study performed the best indicating that for this particular case/configuration this is the most accurate tomogram to use.

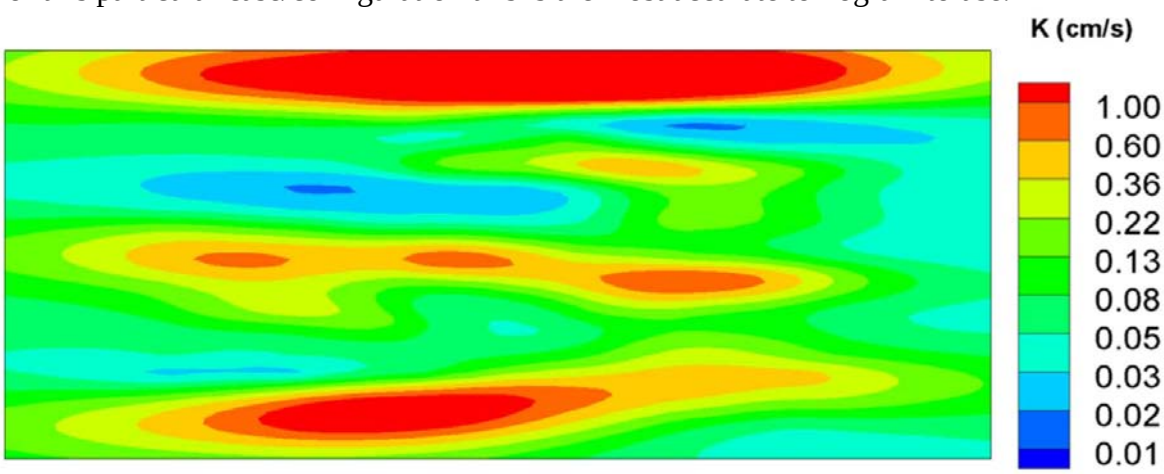


Figure 17.2: *K* tomogram computed using the steady state hydraulic tomography algorithm of Yeh and Liu [2000].

17.4 Description of PCE spill

A PCE source zone was then created in the synthetic aquifer by injecting 1.25 L of PCE [> 99.9% purity, Sigma-Aldrich] dyed with Sudan IV [Sigma-Aldrich] at the top center of the sandbox. To achieve a constant injection rate, we used a peristaltic pump and the PCE was injected a rate of 50 ml per minute over 25 minutes. After commencing the injection, the PCE quickly infiltrated through the 20/30 sand, slowly moved through layers 15, 16 and 17, and ponded on the F-65 sand [layer 10] as shown on Figure 17.3. Figure 17.3 is photograph of the PCE distribution after no additional movement of the PCE infiltration front was observed prior to the partitioning tracer tests for test 1. The injection of PCE in this sandbox resulted in a sandbox aquifer-averaged TCE saturation [S_N] of approximately 3% by assuming a porosity of 0.33 typical of these sands, which is equivalent to a volumetric TCE content of 0.0105.

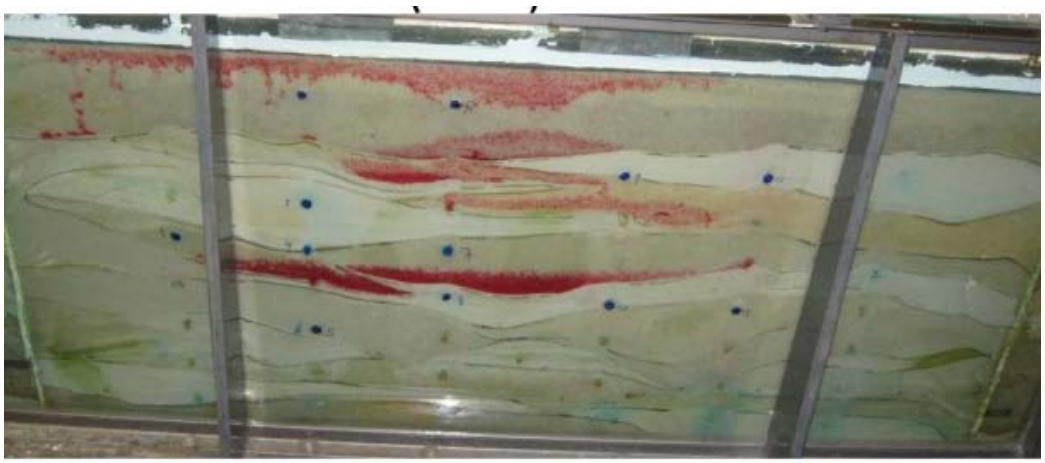


Figure 17.3: Photograph of the sandbox showing the distribution of PCE prior to conducting the partitioning tracer tests.

17.5 Determination of partitioning coefficient through batch tests

In order to determine PCE saturation from partitioning tracer tests, the partitioning coefficient [K_{NW}] of the tracer and PCE is needed. Methods to estimate partitioning coefficients have been developed using UNIFAC [Wang et al., 1997] and on defining equivalent alkane carbon number for each contaminant in a DNAPL and the tracers [Dwarkanath et al., 1998]. Here, we estimated the partitioning coefficients through batch tests and follow the work of Illman et al. [2010b]. The partitioning tracers selected for this study were 2-methyl-1-butanol [2M1B, purity > 99%; Sigma-Aldrich], 1-hexanol [1-Hex, purity > 99%; Sigma-Aldrich], 4-Methyl-2-Pentanol [4M2P, purity > 99%; Sigma-Aldrich], and 2-4-Dimethyl-Pentanol [24-DMP, purity > 99%; Sigma-Aldrich]. Details to the batch tests to determine the partitioning coefficient between the tracers and the PCE are provided in Chapter 14.

17.6 Description of partitioning tracer tests

Upon settlement of the PCE in the source area, we conducted 5 separate partitioning tracer tests. In each partitioning tracer test, a dipole flow field was created by injecting tap water at one port at a rate of 400 mL/min and extracting at another port at 340 mL/min. A water mass balance was then achieved by setting 12 sampling lines at a cumulative rate of 60 mL/min. Achieving a mass balance was crucial because the 3 boundaries of the sandboxes had no-flow boundary conditions, while the top boundary was open to the atmosphere. Deviations in the mass

balance would result in a rising or falling water level which we monitored with pressure transducers. Table 1 summarizes the injection, extraction, and sampling ports for each of the 5 partitioning tracer tests.

Table 17.1: Injection, extraction, and water sampling locations during tests 1 - 5.

Partitioning tracer test	Injection/extration ports	Injected tracers	Sampling port
1	37/12	Bromide, 2M1B, 1-Hex	3, 4, 5, 10, 11, 17, 21, 27, 28, 29, 33, 35
2	42/7	Chloride, 4M2P, 24-DMP	2, 3, 10, 11, 14, 19, 20, 21, 27, 28, 29, 32
3	24/25	Bromide, 2M1B, 1-Hex	2, 3, 10, 11, 14, 19, 20, 21, 27, 28, 29, 32
4	25/24	Chloride, 4M2P, 24-DMP	10, 11, 14, 20, 21, 27, 28, 29, 34, 35, 41, 47
5	37/12	Bromide, 2M1B, 1-Hex	4, 9, 10, 11, 20, 21, 28, 29, 34, 35, 41, 47

Once steady state flow conditions were reached in which the injection, extraction, and sampling rates all stabilized, a valve was switched on the injection line from water to a solution containing various partitioning tracers depending on the tracer test with a concentration of approximately 500 mg/L. Bromide [Fisher Scientific] was used as a conservative tracer with a concentration of approximately 100 mg/L. The tracer solution was injected for 10 minutes and then the valve was switched back to tap water. Separate sample sets were collected for bromide and the partitioning tracers throughout the duration of the test from the 12 sampling ports. Crude estimates of the tracer travel time were obtained using Darcy's law to estimate sampling times at each port. Water sampling from each one of the sampling lines was made possible with a manifold that fed the water samples into twelve 1.5 ml vials simultaneously. Sixty samples for analysis of partitioning tracers and 60 bromide samples were collected from each of the 12 sampling lines during each partitioning tracer test for a total of 1440 samples per tracer test. Each tracer test lasted approximately 5 hours.

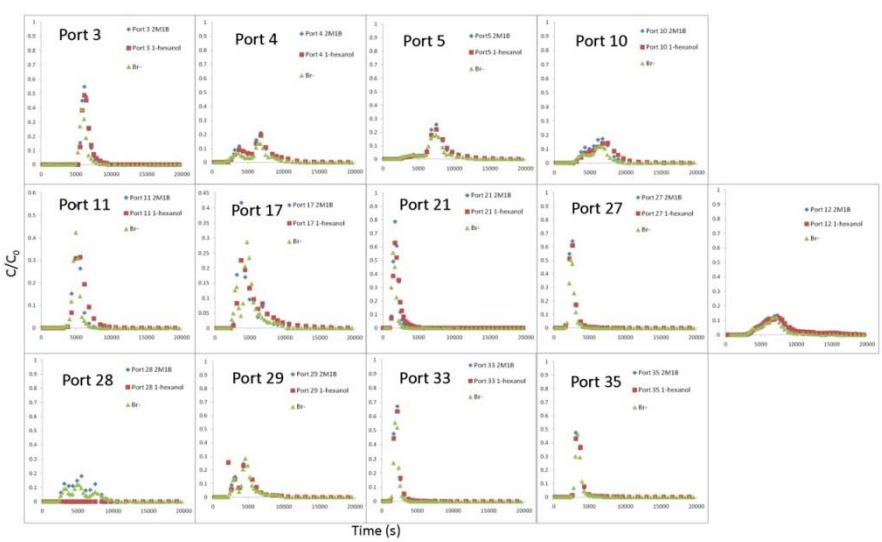


Figure 17.4: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 1.

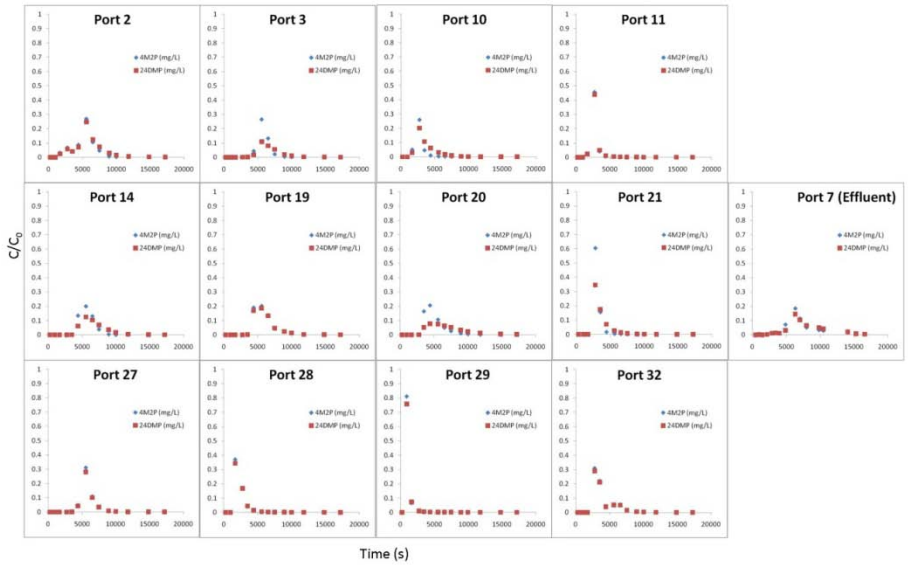


Figure 17.5: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 2.

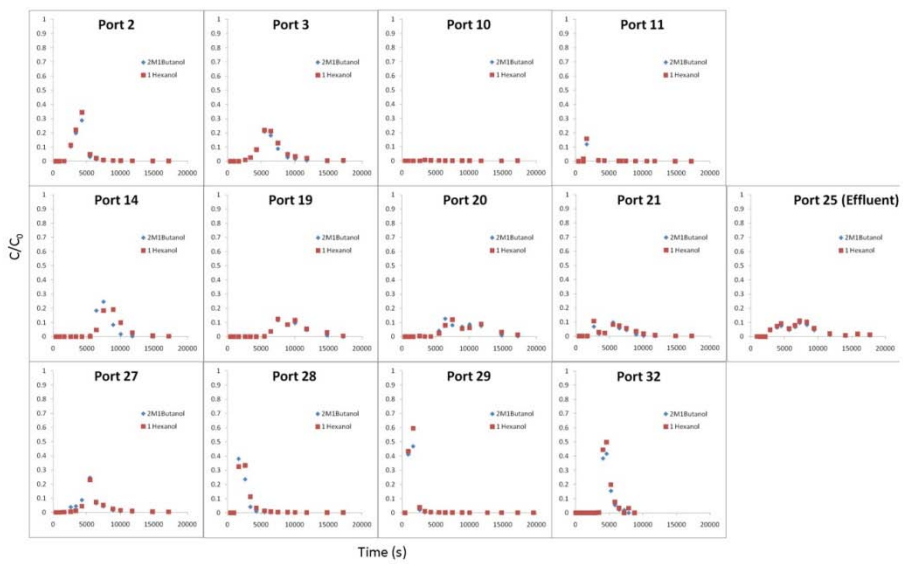


Figure 17.6: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 3.

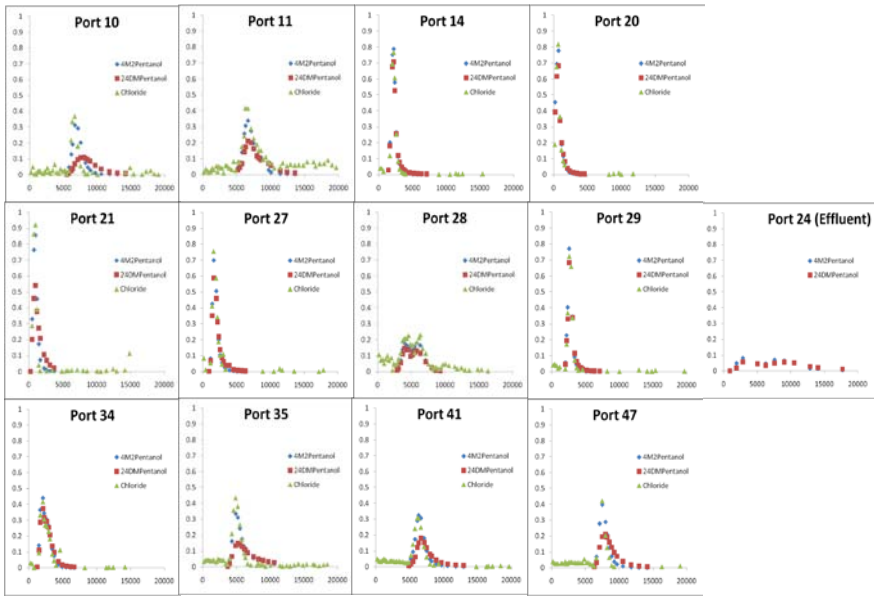


Figure 17.7: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 4.

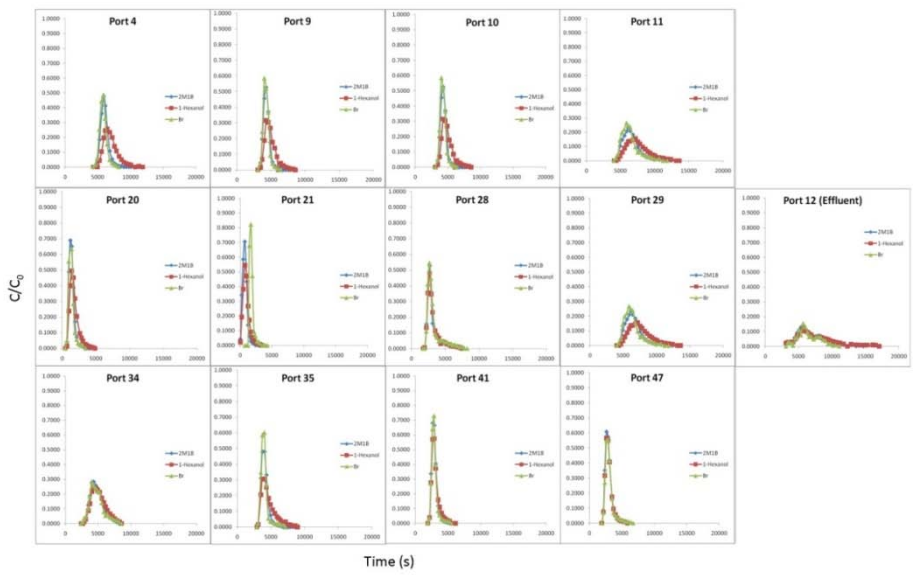


Figure 17.8: Breakthrough curves from the various sampled ports and the effluent port during partitioning tracer test 5.

The 5 tracer tests were designed to maximize the contact of the tracer solution with the PCE source area. Each of the tracer tests was performed with different conservative and partitioning tracer as well as different injection/extraction and water sampling schemes to sweep the tracers through the source zone from different directions in a tomographic fashion. Figures 17.4 – 17.8 are breakthrough curves from partitioning tracer test 1 – 5.

The PCE appeared to be relatively stable after injection and during the first two partitioning tracer tests. However, during the third partitioning tracer test, the PCE was mobilized [Figure 17.9]. At the beginning of partitioning tracer test 3 the PCE was located as centrally in the tank as a pool on top of sand layers #10 and #11. During the third partitioning tracer test the PCE broke through these layers and continued to migrate downwards, eventually pooling on top of sand layers #1 and #2. There are several possibilities that can provide explanations for this: 1] the PCE even though it looked like it was in equilibrium was in fact still moving very slowly; 2] that the tracer tests caused enhanced dissolution of PCE from the nonaqueous to the aqueous phase due to the flushing which then caused the spontaneous movement of the PCE [e.g., Roy et al. 2004]. Regardless of the mechanism behind the mobilization, this causes additional challenges to the interpretation of these tests. The mobilization of the DNAPL was not witnessed in the experiments of Illman et al. [2010b] who used TCE for their experiments.

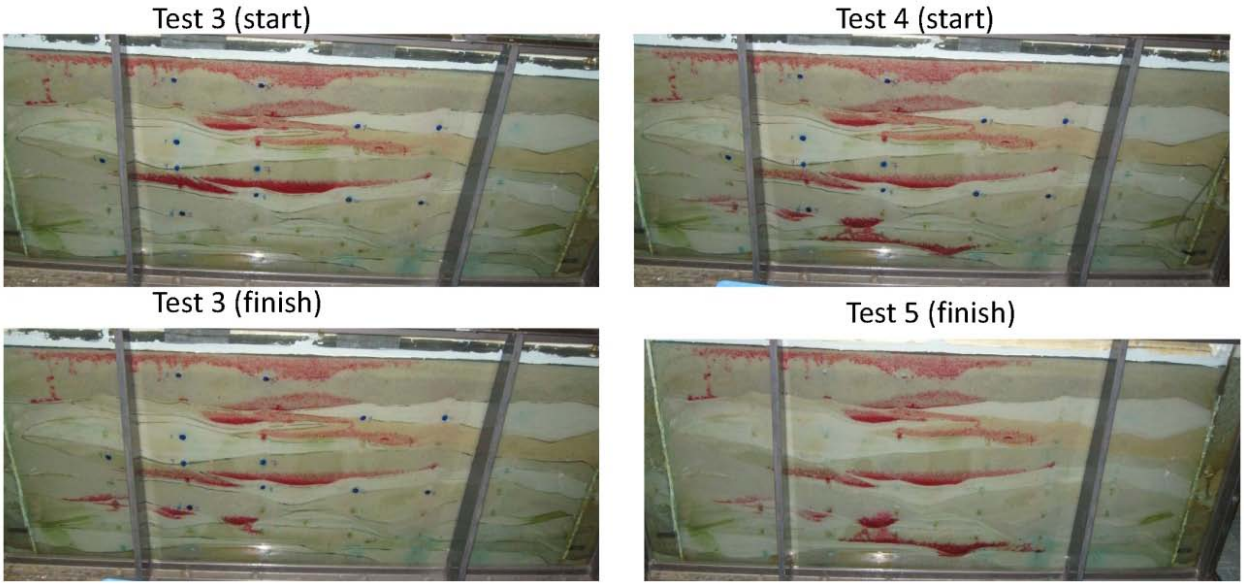


Figure 17.9: Photograph of PCE distribution in the sandbox at various stages of partitioning tracer tests.

17.8 Direct sampling of PCE through excavation

Upon completion of the partitioning tracer tests, the top cover was removed to excavate the aquifer in order to obtain direct samples to facilitate the comparison of PCE saturation estimates from HPTT and the actual distribution in the sandbox aquifer. In particular, a grid of 10 cm by 10 cm blocks was drawn on the sandbox and the excavation took place row by row by inserting a divider. To ensure the samples were representative of the entire grid block, the sand within each block was mixed prior to sampling. Excavation took place under fully saturated conditions and the water level was lowered to the next row after the completion of the excavation of a given row until they reached the sandbox bottom. The excavated samples were then deposited into pre-labeled, pre-weighed 40 mL vials and capped with Teflon-lined septa caps. After this step, the vials were then weighed again and the mass of the vial containing the cap and sample was recorded. The vials were then filled to the top with methylene chloride [Sigma Aldrich] and weighed again. We then recorded the mass of the vial, along with the cap, sample and the methylene chloride. The volume of methylene chloride was calculated based on the mass

differences and density of methylene chloride. The samples were then cooled in the refrigerator at 4°C until they were ready for sample preparation and analysis for PCE mass.

Initially, PCE Standards in the range of 1mg/L thru 1000 mg/L were prepared in DCM (Dichloromethane). Then, 5 uL of the PCE in DCM standards and samples were added to 2 ml of internal standard solution (500 ug/L of 1,2-Dibromoethane in pentane) in a 5mL screw cap glass vials with Teflon-faced septum. The sample was then placed on an orbital shaker for 15 minutes at 300 RPM. The sample was then transferred to a 2 mL glass crimp-top GC vial and put on a Hewlett Packard 5890 Series II gas chromatograph equipped with a Ni63 ECD detector. Using a HP 7673 liquid auto sampler a 1uL sample was injected onto a J&W DB-624 (30m x 0.52mm x 3u) capillary column. The detector temperature was 300°C, injection temperature was 200°C and the column temperature ramp is 50°C to 150°C at a rate of 15°C/min and then held for 1 min. The carrier gas is ultrapure helium with a total flow rate of 25 mL/min and the make-up gas is 5% methane, 95% argon. The method detection limit was 0.46 mg/L.

The mass of PCE in each core sample was calculated by multiplying the average sample by volume of liquid in the sample [assumed to be the volume of methylene chloride]. The mass was then converted to volume by dividing by the density of PCE. To determine the volume of soil in each vial we backcalculated it since we had information on grain density, PCE mass, and methylene chloride mass. We then assumed a porosity of 0.33 to estimate the saturation.

Saturation estimates were available at 106 locations within the sandbox. We kriged these saturation estimates by fitting an exponential variogram to the experimental variogram. The vertical correlation length was 15.4 cm and the with an anisotropy ratio of 2. Figure 17.10 shows the saturation distribution in the sand tank.

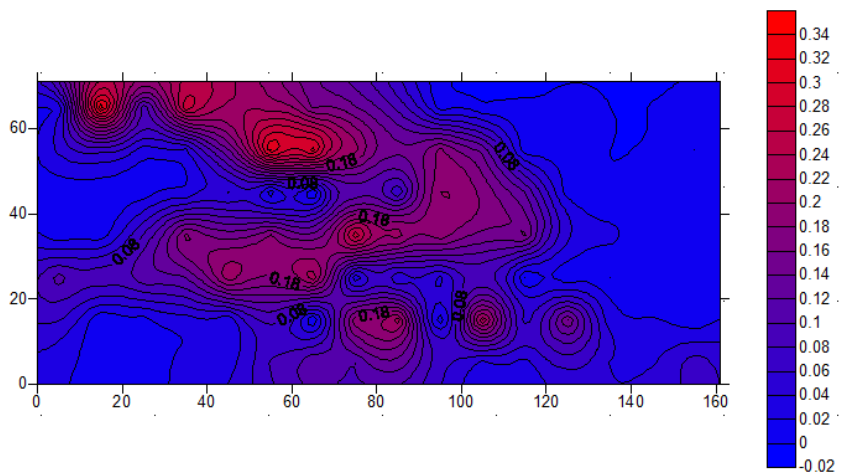


Figure 17.10: Preliminary PCE saturations from kriging excavated samples after conduct of 5 partitioning tracer tests.

17.9 Plans for future analysis

Funding has run out for the analysis of these tracer tests. Additional funding will be sought through SERDP in the future through a separate project to continue the analysis of these data sets.

18. A REVISIT OF DRAWDOWN BEHAVIOR DURING PUMPING IN UNCONFINED AQUIFERS

18.1 Introduction

The response of an unconfined aquifer during a pumping test often is characterized by an S-shaped drawdown-time curve on a log-log scale graph. More specifically, when an unconfined aquifer is pumped at a constant rate, the depth-averaged drawdown appears to follow the Theis [1935] solution for flow in a confined aquifer at early times; at intermediate times, the drawdown becomes less than that predicted by the Theis model; the drawdown again appears to conform to the Theis solution at late times. Such a phenomenon has often been referred to as the delayed yield or delayed gravity response of the water table. The explanation and analysis of this response have been a topic for debate over the years.

Flow processes induced by pumping in unconfined aquifers are generally complex due to presence of saturated and unsaturated zones, and the time-varying interface of the two zones [the water table], in addition to the heterogeneous nature of geologic media. As a consequence, hydrologists over the past few decades have adopted simplified conceptual models such that the mathematical analysis of the drawdown caused by pumping in an unconfined aquifer is mathematically tractable.

Generally speaking, two categories of conceptual mathematical models based on analytical solutions have been developed over the years to quantify and explain the delayed yield phenomenon. The first category is the depth-averaged, radial flow model with a delayed drainage source term [Boulton, 1954, 1963]. As stated by Boulton [1963], "the water-bearing material through which the water table has fallen during the early stages of pumping does not yield up its water immediately." As a result, the Boulton model includes an empirical delay coefficient to the specific yield term to represent the slow water-release process, and to account for the intermediate stage of the S-shaped well hydrograph [the log-log drawdown-time data plot].

The other category includes those based on a concept of instantaneous and complete drainage at the water table [e.g., Dagan, 1967; Brutsaert, 1970; Streltsova, 1972a, 1972b; Neuman, 1972; and Lakshminarayana and Rajagopalan, 1978]. These are radial flow models which consider 2-D flow field in the vertical plane along the radial distance, and the delayed drawdown region of the time-drawdown curve is considered to be caused by the downward hydraulic-head gradient below the water table [i.e., gravity delayed yield]. They neglect the influence of the unsaturated zone above the water table and assume instantaneous drainage of the initially saturated pores.

Application of both types of models to time-drawdown data collected from field pumping tests generally yields specific-yield values that are substantially below those that would be expected on the basis of other methods of measurement [see Nwankwor et al., 1984; Endres et al., 2006]. Nwankwor et al. [1984] attributed the low values of specific yield obtained from the type-curve methods to an inadequate representation of the drainage processes occurring near the water table.

Based on the field data, Nwankwor et al. [1992] explained the S-shaped time-drawdown behavior of unconfined sand aquifers as a consequence of changes in vertical hydraulic gradients and water-content profile above the water table. Numerical modeling of variably-saturated flow by Akindunni and Gillham [1992] and analysis of drainage from a soil column by Narasimhan and Zhu [1993] supported the explanation by Nwankwor et al. [1992] about the

importance of the initially unsaturated zone above the water table. Narasimhan and Zhu [1993] concluded that a simple exponential release function used by Boulton does not accurately simulate drainage from above the water table. To obtain a more general mathematical approximation of the drainage process, Moench et al. [2001] developed an analytical model that used a linear combination of exponential functions to simulate release from above the water table.

Endres et al [2006] compared bulk vadose zone response predicted by the analytical models [i.e., Boulton, 1954, 1963; Neuman, 1972; Moench et al. 2001] and inferred from field measurements using hydraulic head data and soil moisture content profiles obtained during a seven day pumping test at CFB Borden, Ontario. They concluded that the water table boundary conditions used in these analytical models do not adequately replicate the mechanisms controlling the vadose zone behavior during a pumping test.

Tartakovsky and Neuman [2007] developed a semi-analytical model that includes flow through unsaturated zone above the water table. The analytical nature of the model forced them to ignore unsaturated flow induced by lowering the water table during the pumping test, and to assume that the unsaturated hydraulic conductivity and moisture capacity term vary with elevation instead of capillary pressure which varies with flow at different locations. Based on this model, they concluded that unsaturated flow in the vadose zone above the initial water table has significant impact on dimensionless log-log drawdown-time behavior in the saturated zone when the aquifer has large retention capacity and/or small initial saturated thickness. Moreover, as horizontal saturated hydraulic conductivity increases relative to the vertical, the effect of unsaturated flow on drawdown in the aquifer diminishes.

Although effects of heterogeneity on variably saturated flow have been investigated by Li and Yeh [1998], few have investigated the effects of heterogeneity on drawdown-time curves during pumping in unconfined aquifers. Akindunni and Gillham [1992] mentioned possible effects of heterogeneity when analyzing data from pumping tests in the Borden Aquifer. Bunn et al. [2010] conducted a Monte Carlo analysis to investigate the effects of the spatial variability of the saturated hydraulic conductivity in the unconfined aquifer on the capillary fringe extension observed in the field [Bevan et al., 2005].

In this study, numerical experiments of pumping in an unconfined aquifer are conducted using a finite element numerical model which solves the governing equations for flow through variably saturated media. A first-order stochastic moment approach is then developed to quantify the effects of spatial variability of parameters for both saturated and unsaturated zones on the drawdown-time curve induced by the pumping in the aquifer. By examining the rate of change in storage at various locations in the aquifer including the saturated zone, the vadose zone above the initial water table, and the porous medium after lowering of the water table in the aquifer, we demonstrate that the S-shaped log-log drawdown time behavior during pumping in an unconfined aquifer is merely a manifestation of the transition of the water release mechanism from compaction of porous media and expansion of water to actual dewatering of porous media under vertical flows. We subsequently employ the stochastic moment approach to explore the role of heterogeneity in the development of the S-shaped drawdown-time curves and the spatial cross-correlation between the drawdown at a given location and heterogeneity at various part of the aquifer. Finally, implications of the results of this study are discussed.

18.2 METHODOLOGY

A finite-element numerical model, VSAFT3 [Variably Saturated Flow and Transport in 3D] by Srivastava and Yeh [1992] is employed in this investigation. This program solves the partial differential equation that describes flow in 3-D, variably saturated geologic media:

$$\nabla \cdot [K(h,\mathbf{x})\nabla(h+z)] = \omega S_s(\mathbf{x})\frac{\partial h}{\partial t} + \frac{\partial \theta}{\partial t} = \left(\omega S_s(\mathbf{x}) + C(h,\mathbf{x})\right)\frac{\partial h}{\partial t} \qquad [18.1]$$

where ∇ is the spatial gradient, t is time, θ represents the volumetric moisture content, and z is the elevation, which is positive upward. h is the pressure head and is positive when the medium is fully saturated and negative when unsaturated. The saturation index ω is equal to one if the medium is saturated and zero if the medium is unsaturated. The term $S_s[\mathbf{x}]$ represents the specific storage, $C[h,\mathbf{x}]$ is the soil moisture capacity, which can be derived from the moisture/pressure constitutive relationship, and $K[h,\mathbf{x}]$ is the hydraulic conductivity-pressure constitutive function. The program employs the Newton-Raphson iteration scheme to solve the nonlinear finite element approximation of equation [18.1].

We use VSAFT3 to simulate flow to a well due to pumping in a 3-D unconfined aquifer. The dimension of the aquifer is 200 m×200 m in the horizontal plane and 9 m in the vertical and is discretized into 380,880 finite elements. A variable mesh is used for the discretization. In the vertical plane, a vertical interval of 0.2m is used from z=6.0 m to 7.5 m [i.e., about the water table] and a vertical interval of 0.5 m is employed from z=0.0 m to 6.0 m and from z=7.5m to 9.0m. In the horizontal plane, the interval of the mesh is 0.5m from x=76m to 124m and y=76m to 124m [around the pumping well] and 4.0m otherwise.

A no-flux boundary is assigned to the top [no infiltration or evaporation] and bottom of the aquifer; no-flux boundaries are imposed on the other four sides of the aquifer. The initial pressure head distribution in the aquifer is set to be hydrostatic with the water table at z = 6.7 m, representing a static condition. A well is represented by a line source of length 4.0 m from the bottom center of the aquifer [x = 100 m, and y = 100 m] with a constant rate of discharge [0.06 m^3/min], in which the borehole storage is neglected.

The hydraulic conductivity-pressure head and moisture-pressure head constitutive relationship of the aquifer are described by an exponential model [Gardner, 1958]:

$$K(h) = K_s \exp(\alpha h) \qquad [18.2]$$

And

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\exp(\beta h) \quad , \qquad [18.3]$$

respectively. In equations [18.2] and [18.3], α and β are pore-size distribution parameters for the unsaturated hydraulic conductivity/pressure head function and the moisture/pressure head relationship, respectively; θ_s and θ_r denote saturated and residual water content.

For the cases where the aquifer is assumed homogeneous, values of the parameters in equations [18.2] and [18.3] are referred from the study by Akindunni and Gillham [1992]: K_s= 0.00396 m/min, S_s= 0.000325/m, θ_s= 0.37 and θ_r= 0.07, α = 4.0/m and β = 4.0/m, respectively.

The thickness of the capillary fringe [or air entry value] is assumed to be $1/\beta$. The pumping test is simulated for 2,000 minutes.

In order to investigate the effect of parameter heterogeneity on the drawdown-time curve, a first-order stochastic moment analysis developed by Li and Yeh [1998] and Hughson and Yeh [2000] is employed. Specifically, natural logarithms of the parameters in the constitutive relationships of the unsaturated hydraulic properties are treated as stochastic processes in space, which are characterized by their means and spatial covariance functions. Subsequently, the head is expanded in a Taylor series about the mean values of hydraulic parameters. After neglecting the second-order and high-order terms, the first-order approximation of the total head can be written as

$$H(\mathbf{x},t) = \overline{H}(\mathbf{x},t) + p(\mathbf{x},t) = \overline{H}(\mathbf{x},t) + f(\mathbf{x}) \cdot \frac{\partial H}{\partial \ln K_S(\mathbf{x})} + s(\mathbf{x}) \cdot \frac{\partial H}{\partial \ln S_S(\mathbf{x})} + a(\mathbf{x}) \cdot \frac{\partial H}{\partial \ln \alpha(\mathbf{x})} + b(\mathbf{x}) \cdot \frac{\partial H}{\partial \ln \beta(\mathbf{x})} + t_s(\mathbf{x}) \cdot \frac{\partial H}{\partial \ln \theta_s(\mathbf{x})} + t_r(\mathbf{x}) \cdot \frac{\partial H}{\partial \ln \theta_r(\mathbf{x})} \qquad [18.4]$$

where $f[\mathbf{x}]$, $s[\mathbf{x}]$, $a[\mathbf{x}]$, $b[\mathbf{x}]$, $t_s[\mathbf{x}]$ and $t_r[\mathbf{x}]$ are the perturbations of the logarithms of parameters, $K_s[\mathbf{x}]$, $S_s[\mathbf{x}]$, $\alpha[\mathbf{x}]$, $\beta[\mathbf{x}]$, $\theta_s[\mathbf{x}]$ and $\theta_r[\mathbf{x}]$. These perturbations are dimensionless. $\overline{H}$ is the mean total head, evaluated using the mean parameters and it represents the head in an equivalent homogeneous aquifer; p denotes the head perturbation around the mean head, resulting from spatial variability of the parameters. The partial derivatives are the sensitivity matrices, which are evaluated at the mean parameters $\overline{K_s}, \overline{S_s}, \overline{\alpha}, \overline{\beta}, \overline{\theta_s}$ and $\overline{\theta_r}$. According to equation [18.4], the head perturbation can be written in a matrix form as

$$\mathbf{p} = \mathbf{J}_{pf}\mathbf{f} + \mathbf{J}_{ps}\mathbf{s} + \mathbf{J}_{pa}\mathbf{a} + \mathbf{J}_{pb}\mathbf{b} + \mathbf{J}_{pt_s}\mathbf{t}_s + \mathbf{J}_{pt_r}\mathbf{t}_\mathbf{r} \qquad [18.5]$$

Here, bold characters denote either matrices or vectors. $\mathbf{J}_{pf}$ is the sensitivity of $\mathbf{p}$ to change in parameter $\boldsymbol{f}$ or Jacobian matrix, which is calculated by the adjoint method [see Skyes et al., 1985; Li and Yeh, 1998 and 1999; Hughson and Yeh, 2000]. Multiplying equation [18.5] with itself, taking the expectation, and assuming that the perturbation of the different parameters are statistically independent from each other lead to the covariance of the head:

$$\mathbf{R}_{pp} = \mathbf{J}_{pf}\mathbf{R}_{ff}\mathbf{J}_{pf}^{T} + \mathbf{J}_{ps}\mathbf{R}_{ss}\mathbf{J}_{ps}^{T} + \mathbf{J}_{pa}\mathbf{R}_{aa}\mathbf{J}_{pa}^{T} + \mathbf{J}_{pb}\mathbf{R}_{bb}\mathbf{J}_{pb}^{T} + \mathbf{J}_{pt_s}\mathbf{R}_{t_st_s}\mathbf{J}_{pt_s}^{T} + \mathbf{J}_{pt_r}\mathbf{R}_{t_rt_r}\mathbf{J}_{pt_r}^{T} \qquad [18.6]$$

$\mathbf{R}_{ff}$, $\mathbf{R}_{ss}$, $\mathbf{R}_{aa}$, $\mathbf{R}_{bb}$, $\mathbf{R}_{t_st_s}$ and $\mathbf{R}_{t_rt_r}$ are the covariance function matrices for the parameters, which are assumed to follow an exponential covariance model [Gelhar, 1993]. The superscript T denotes the transpose. Each diagonal element of $\mathbf{R}_{pp}$ is the head variance [σ_p^2] at location $\mathbf{x}$ and at time t, representing the mean square deviation of head in a heterogeneous aquifer from the head calculated using the mean parameters for the homogeneous aquifer. The corresponding standard deviation [σ_p] is added to and subtracted from mean drawdown-time curves at a given

observation location calculated by the mean parameters to yield an upper and a lower bounds of the drawdown-time curve. Such lower and upper bounds delineate the influence of heterogeneity of different parameters. In other words, the log-log drawdown-time curve of a heterogeneous unconfined aquifer will likely fall in between these bounds. Greater differences between the upper and lower bounds suggest that the log-log drawdown-time curve of the heterogeneous aquifer may be quite different from the perfect S shape obtained using the mean parameters, representing an equivalent homogeneous aquifer. While the first-order analysis only yields approximate upper and lower bounds and Monte Carlo simulations with a large number of realizations would produce a more accurate upper and lower bounds, the result of the first-order analysis efficiently illustrates general effects of heterogeneity on the S shaped drawdown-time curve.

To evaluate the mean head in equation [18.4], the values of the mean parameters are assumed to be the same as the parameter values in the homogeneous case and VSAFT3 is used. Two levels of heterogeneity [variances of the logarithm of parameters equal to 0.1 and 1] for all parameters are examined. In addition, the spatial structure of all parameters is assumed to be described by a 3-D exponential correlation function with some specified correlation scales in horizontal and vertical directions, which represent the spatial structure of the heterogeneity.

In addition to the first-order stochastic moment analysis, a single realization of each parameter field is generated using a stochastic field generator with the given mean and variance and correlation scales of the parameters of the exponential constitutive model. These fields are then input to VSAFT3 to simulate the response of a heterogeneous aquifer. The response of the heterogeneous aquifer is subsequently compared to the response derived from mean parameters [homogeneous aquifer] to show the effect of heterogeneity on the drawdown-time curve.

Subsequently, a cross-correlation analysis is carried out to investigate how the head at a given location in the saturated zone is affected by the parameter value at different locations in the aquifer. This cross-correlation analysis again is based on the first-order stochastic moment analysis. For example, the cross-correlation between p at location $\mathbf{x}_i$ and f at $\mathbf{x}_j$ can be expressed as

$$\rho_{pf}(\mathbf{x}_i,\mathbf{x}_j,t) = \frac{\mathbf{R}_{pf}(\mathbf{x}_i,\mathbf{x}_j,t)}{\sigma_p(\mathbf{x}_i,t)\sigma_f} = \frac{\mathbf{J}_{pf}(\mathbf{x}_i,\mathbf{x}_j,t)\mathbf{R}_{ff}(\mathbf{x}_i,\mathbf{x}_j)}{\sigma_p(\mathbf{x}_i,t)\sigma_f} \qquad [18.7]$$

where j = 1, N and N is the total number of elements in the domain and σ_f is the standard deviation of f. Here, we also assume independence of each different parameter. The cross-correlation between p at location $\mathbf{x}_i$ and the other parameters at $\mathbf{x}_j$ can be written in a similar expression.

18.3. RESULTS AND DISCUSSIONS

The results and discussions will be grouped into 1] flow to a well in a homogeneous aquifer, 2] 1-D variably saturated flow, and 3] flow to a well in a heterogeneous aquifer.

18.3.1. Flow to a Well in a Homogeneous Aquifer

18.3.1.1 Drawdown-Time Curve

Simulated drawdowns in the unconfined aquifer as a function of time at five elevations [1.5 m, 3.0 m, 6.0 m, 7.1 m, and 7.5 m] at radial distance 5.0 m and 30.0 m from the pumping well are plotted using log scales in Figures 1a, and 1b, respectively. Notice that elevations 7.1 m and 7.5 m are above the water table, and drawdowns at these elevations denote change of pressures in the vadose zone due to pumping in the saturated zone. These two figures show that drawdowns in the saturated zone exhibit the characteristic S shape, which is most pronounced at bottom of the aquifer. The flat part of the S curve at z = 6.0 m occurs earlier than at the bottom of the aquifer [i.e., z = 1.5 m and z = 3.0 m]. On the other hand, the drawdown-time curves at elevations 7.1 m and 7.5 m in the unsaturated zone do not exhibit the S shape. These simulated results are consistent with field observations [e.g., Bevan et al., 2005].

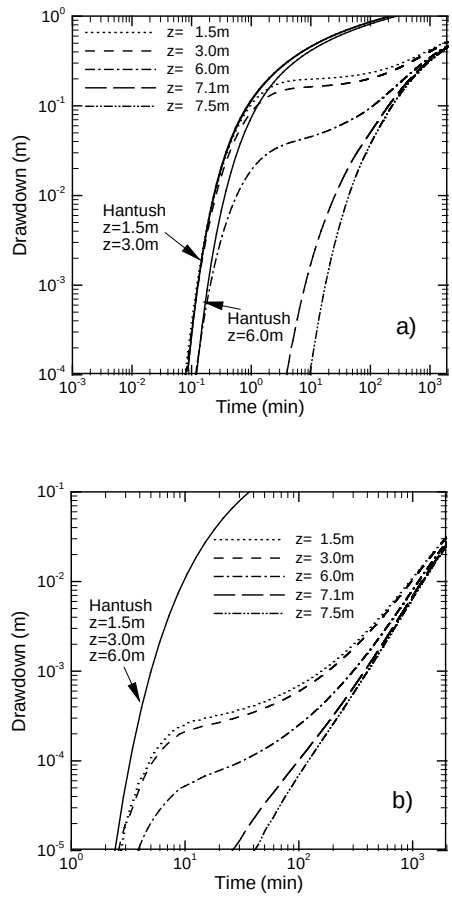


Figure 18.1: Log-log drawdown time curves at five different elevations [z =1.5, 3.0, 6.0, 7.1, 7.5 m] at a] r = 5 m and b] r = 30 m. Solid lines denote the results based on the solution by Hantush at three observation elevations in the saturated zone z = 1.5, z = 3.0, z = 6.0m. In Figure a, the difference between the Hantush solution at z=1.5 m and z=3.0 m is not distinguishable. In Figure b, the solutions are the same at all three elevations.

Figures 18.1a and 18.1b also show drawdown-time curves at different elevations [z = 1.5 m, 3.0 m, and 6.0 m] in a confined aquifer that has an infinite lateral extent, with the same saturated thickness [6.7m] and the saturated hydraulic properties of the unconfined aquifer. These curves are obtained by using the solution of Hantush [1961a and 1961b], which considers the partial perforation of the pumping well in the aquifer. According to Figure 18.1a, drawdowns in the confined aquifer at r = 5 m, at higher elevation z=6.0 m are smaller that those at z = 1.5 m and z = 3.0 m, indicative of vertical flow. At r = 30.0 m, far from pumping well, identical drawdowns at the three elevations suggest that the effect of the partial perforation of the pumping well [i.e., vertical flow] is negligible [Figure 18.1b]. As expected, no S-shaped curve is observed.

Notice that in both figures the drawdown-time curves in the unconfined aquifer departs from its initial steep drawdown portion and deviates from the corresponding drawdowns of the confined aquifer in a descending elevation order [i.e., z = 6.0 m, z = 3.0 m and z = 1.5 m]. This sequential departure [i.e., smaller drawdowns than the confined aquifer] implies that an "additional" source of water is flowing from the upper to the lower part of the saturated zone of the aquifer as recharge, in comparison with the flow process in the confined aquifer. As will be explained below, such an "additional" source of water is mainly attributed to a different water release mechanism at this stage of the flow process in the unconfined aquifer.

18.3.1.2 Rate of Change in Storage

Different from previous works, our study uses the evolution of the "rate of change in storage" per volume of the porous medium [Ω] at various distances and depths during a pumping test in an unconfined aquifer to explain the so called "delayed yield" phenomenon. Mathematically, this rate of change in storage is the right-hand side of equation [18.1], which represents the net mass flux in a unit volume of the porous medium:

$$\Omega = \left(\omega S_s(\mathbf{x}) + C(h,\mathbf{x})\right)\frac{\partial h}{\partial t} \qquad [18.8]$$

In VSAFT3, each node links a number of elements depending on the discretization. The rate of change in storage per unit volume of around a node between two successive time steps is calculated by dividing the change in storage over the elements associated with the node by the time step and total volume of the elements.

18.3.1.3 Early times

In Figure 18.2a, the quantity Ω is plotted as a function of elevation at different radial distances from the pumping well at various times of the early stage of the S-shaped drawdown-time curve. Specifically, these distances and times are r = 2.5 m at t = 0.02 min; r = 5.0 m at t = 0.07 min; r = 10 m at t = 0.3 min; r = 20 m at t = 1.0 min, r = 30 m at t = 3.0 min. We use different times for different radii to illuminate the effect. At this early stage, the Ω value at r = 2.5 m is the largest at the bottom of the aquifer over the elevation from 0 m to 4 m where the pumping well screen is located. This value then decreases with elevation and reaches the minimum at the water table; it then increases sharply at the water table and subsequently decreases with elevation in the unsaturated zone. Overall, the Ω values below the water table exhibit a strong nonlinear pattern with greater Ω values at the bottom of the aquifer than those near the water table. This pattern is consistent with the fact that the greatest change in pressure

is at the bottom of the aquifer near the pumping location. The Ω value in unsaturated zone at these times is virtually zero except close to the water table. Similar behaviors are also observed at the other radial distances. This behavior suggests that water flow into the pumping well is mainly from expansion of water and compaction of the aquifer in the saturated zone, which is characterized by the specific storage of the aquifer. This time period corresponds to the early time segment in the log-log drawdown-time plot. The water table position can only be approximated from the plot of Ω as a function of elevation in Figure 18.2, as different finite element meshes are used at different radial distances.

18.3.1.4 Intermediate Times

During the intermediate stage [the flat portion] of the log-log drawdown-time curve [t = 10 min after pumping], the behavior of Ω as a function of elevation at r = 2.5 m, 5.0 m, 10.0 m, 20.0 m and 30.0 m from the pumping well is shown in Figure 18.2b. In comparison with Figure 18.2a, the Ω values in the vadose zone above the water table [near the sharp break in curve around z=6.5 m] at all radii begin to increase. The Ω values in the saturated region below the water table decrease at all radial distances. The closer to the pumping well the location, the greater reduction in Ω in comparison with Figure 18.2a. Furthermore, the Ω values at locations close to the pumping well are almost constant with elevation. All these behaviors indicate that the flow field near the well is close to the steady state [quasi steady state] in which the head field changes at the same rate at all elevations.

They also suggests that pumping has begun to induce significant amount of water released from the vadose zone to supply the well discharge, whereas the elastic storage of the aquifer near the pumping location is almost exhausted. In other words, the drainage of water from pores in the unsaturated zone increasingly becomes the primary source of water for the well discharge. It yields more water to the saturated region below than the elastic storage of the region itself and contributes to an "additional" source of water compared with the similar flow in the confined aquifer. As a result, the drawdown-time curve deviates from that based on the Hantush solution [1961a, 1961b]. The transition of water release mechanism from compaction of the porous medium and expansion of water to dewatering of pores thus explains the "additional" source of water and in turn, the flat part of the intermediate stage of the S-shaped drawdown-time curve.

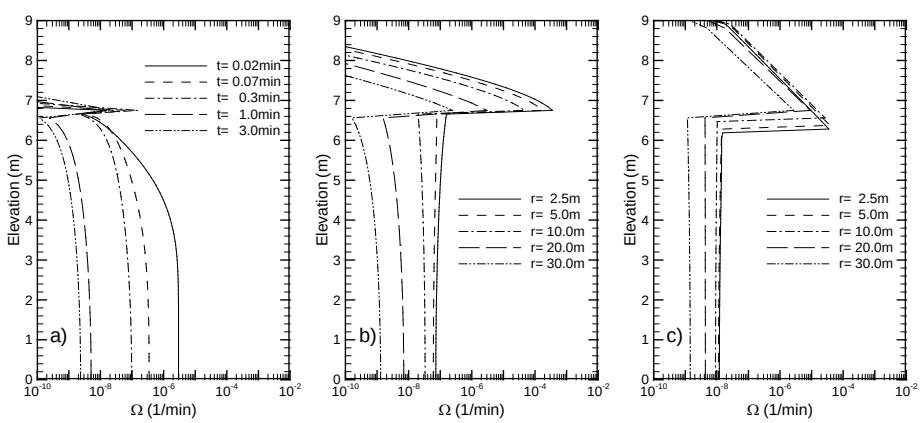


Figure 18.2: Ω as a function of elevation a] at five early times [t = 0.02, 0.07, 0.3, 1.0, and 3.0 min] corresponding to five different radii [r = 2.5, 5, 10, 20 and 30 m]; b] at an intermediate stage [10 minutes after pumping] and c] a late stage [1000 minutes after pumping] at the five radii. The water table is located at the sharp discontinuity.

18.3.1.5 Late Times

The behavior of Ω as a function of elevation at the five radii at t = 1000 min after pumping [i.e., the late stage of the S curve] are plotted in Figure 18.2c. At this stage the water table at all radii starts to fall noticeably; the Ω values in the unsaturated zone are much greater at high elevations at this late stage than the early or the intermediate stage. The Ω values in the saturated region at all radii are very small and do not vary with the elevation. These behaviors of the Ω values indicate that flow in this region apparently approaches steady-state conditions; water released from the unsaturated zone and from previously saturated pores due to falling of the water table is equal to the amount of water withdrawn by the pumping well. Thus, little change in storage occurs in the saturated region of the aquifer. Note that although there is still water released from the aquifer due to compaction of the aquifer and expansion of water, its amount is much smaller than the amount of water released from falling of the water table and the initially unsaturated zone.

18.3.2. 1-D Vertical Variably Saturated Flow

To further emphasize the importance of the transition of water release mechanism from compaction of the aquifer to drainage of porous media in the formation of the S-shaped hydrograph, the drainage process of a vertical soil column is simulated. The column is 150 cm in length and 0.1 cm in width and thickness. The column is discretized into 1500 square elements with 0.1cm in length, width, and thickness. Such a small element size is selected to ensure accurate tracking the location of water table and correct mass balance of the experiments.

No-flux boundaries are assigned to the top, bottom, and all sides of the column. The initial pressure distribution in the domain is assumed to be hydrostatic. The initial water table is located at elevation 110 cm from the bottom of the vertical domain, and an initial thickness of the unsaturated zone is 40cm. At time greater than zero, a constant discharge q=0.08 cm/min is imposed at the bottom boundary of the column. The simulation lasts 150 min and drawdown data [difference between the initial head and head after the drainage starts] are recorded at elevations z = 10 cm, 40 cm, 70 cm.

The saturated hydraulic parameters for the soil column are K_s= 0.495 cm/min, S_s= 0.0001/cm. The experiment is conducted for two different unsaturated hydraulic properties: the exponential model [equations 2 and 3] and Van Genuchten model [1980].

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\left(1 + \gamma |h|^n\right)^{-m} \qquad [18.9]$$

where | | is absolute value, θ_s is the saturated moisture content, θ_r is the moisture content at residual saturation and γ, n, and m are shape-fitting parameters with m=1-1/n. We further assume that the $K[h]$ follows Mualem's [1976] pore-size distribution model expressed as:

$$K(h) = K_s(1 - (\gamma|h|)^{n-1}[1 + (\gamma|h|)^n]^{-m})^2 / [1 + (\gamma|h|)^n]^{m/2} \qquad [18.10]$$

where K_s is the locally isotropic saturated hydraulic conductivity; and γ, n, and m are assumed to be the same as those in equation [18.9]. Hereafter, equations [18.9] and [18.10] are referred to

as the VGM model [hereafter]. The parameters for the exponential model are α = 0.25/cm, β = 0.15/cm, and γ =0.15/cm, n=2.5 for the VGM model. For both models, θ_s and θ_r are set to 0.35 and 0.045, respectively. Notice that VGM model has a slightly more distinct capillary fringe than the exponential model.

An analytical solution describing saturated flow in a semi-infinite vertical domain [Ozisik, 1993] is also used to describe drawdown behavior due to release of water from the column caused by compaction of the porous medium and expansion of water only. Using the definition of hydraulic diffusivity $D=K/S_s$, the analytical solution to the governing equation:

$$\frac{\partial^2 H}{\partial z^2} = \frac{1}{D}\frac{\partial H}{\partial t} \qquad [18.11]$$

with initial and boundary conditions:

$$H\big|_{t=0} = H_0 \quad \text{and} \quad K\frac{\partial H}{\partial z}\bigg|_{z=0} = q \quad H\big|_{z=\infty} = H_0 \qquad [18.12]$$

is

$$s(z,t) = H_0 - H = H_0 - \frac{q}{K}\left[2\sqrt{\frac{Dt}{\pi}}\exp\left(-\frac{z^2}{4Dt}\right) - z\left(1 - erf\left(\frac{z}{2\sqrt{Dt}}\right)\right)\right] \qquad [18.13]$$

where $s(z,t)$ is the drawdown. The parameters and discharge in the equation are kept the same as those in the 1-D variably saturated experiment, i.e., K_s= 0.495cm/min, S_s= 0.0001/cm, and q = 0.08 cm/min.

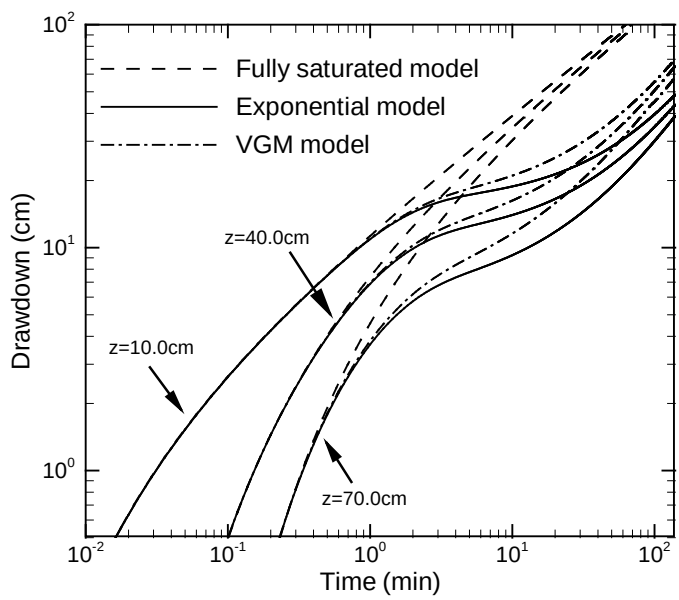


Figure 18.3: Comparison of drawdown-time curves at three different elevations z=10cm, z=40cm and z=70cm in a fully saturated column and those in a soil column with an unsaturated zone of 40 cm with the exponential constitutive relationships, and those with VGM constitutive relationships.

Figure 18.3 depicts simulated drawdowns as a function of time at elevations z = 10 cm, z = 40 cm, and z = 70 cm during the drainage process in the soil column with the exponential and the VGM models as well as during the drainage process under the fully saturated condition. According to the figure, 1-D vertical variably saturated flow can yield S-shaped log-log drawdown-time curves similar to those drawdown time curves in the 3-D unconfined aquifer. Furthermore, the drawdown from the variably saturated model deviates from the drawdown of the fully saturated model sequentially at z = 70 cm first, then at z = 40 cm and z = 10 cm. This sequential occurrence of the deviation [smaller drawdown] indicates that the water released by desaturation of the unsaturated zone above the water table arrives at z = 70cm first and then progressively migrates downward. In other words, the transition of water release mechanism progresses from the location near the water table to the bottom of the column. These results are the same as those in the case of pumping in an unconfined aquifer as indicated in Figure 18.1.

18.3.2.1 Effects of the Exponential Model vs. the VGM model

Compared with the VGM model, the exponential model does not describe the unsaturated hydraulic properties of real soils well near the air entry pressure value. Nevertheless, according to Figure 18.3, the overall shapes of the drawdown-time curves resulting from the two models are similar. Drawdowns based on the VGM model are however consistently larger than those based on the exponential model once the effects of the drainage process become effective. This can be attributed to the greater water holding capacity of the moisture retention curve of the VGM model. The Ω values in the 1-D vertical column as a function of elevation at t = 0.01 min, 3.0 min, and 100.0 min after drainage started are plotted in Figures 18.4a and 18.4c for the exponential and the VGM model, respectively. Both figures show that at the early time [t =0.01 min], change in storage takes place below the water table; it is the greatest at the bottom of the column; it decreases nonlinearly towards zero near the elevation of 50 cm. The discharge from the column at this time is the water released due to compaction of the porous medium and expansion of the water. As expected, the rates of change in storage in these two figures for the two models at the early time are identical.

At t = 3.0 min, the nonlinear behavior of the Ω value in the saturated zone becomes less pronounced; contributions from the unsaturated zone above the water table begins to increase. Differences in the rate of change in storage resulting from these two models become apparent. In the case of the VGM model, the Ω value decreases sharply right above the water table and then increases with the elevation [see circled areas in Figure 18.4c]. The sharp decrease in the rate of change in storage occurs at the location corresponding to the capillary fringe depicted by the model. The capillary fringe is under negative pressure and does not release water unless its air entry pressure is exceeded by downward forces. The same plot based on the exponential model [Figure 18.4a] however does not show such a sharp change near the water table. This is attributed to the fact the exponential model does not have a distinct air entry cut-off pressure [or an abrupt capillary fringe]. Above the water table, both models show that the rate change of storage decreases with the height but the VGM model results in a smoother pattern than the exponential model.

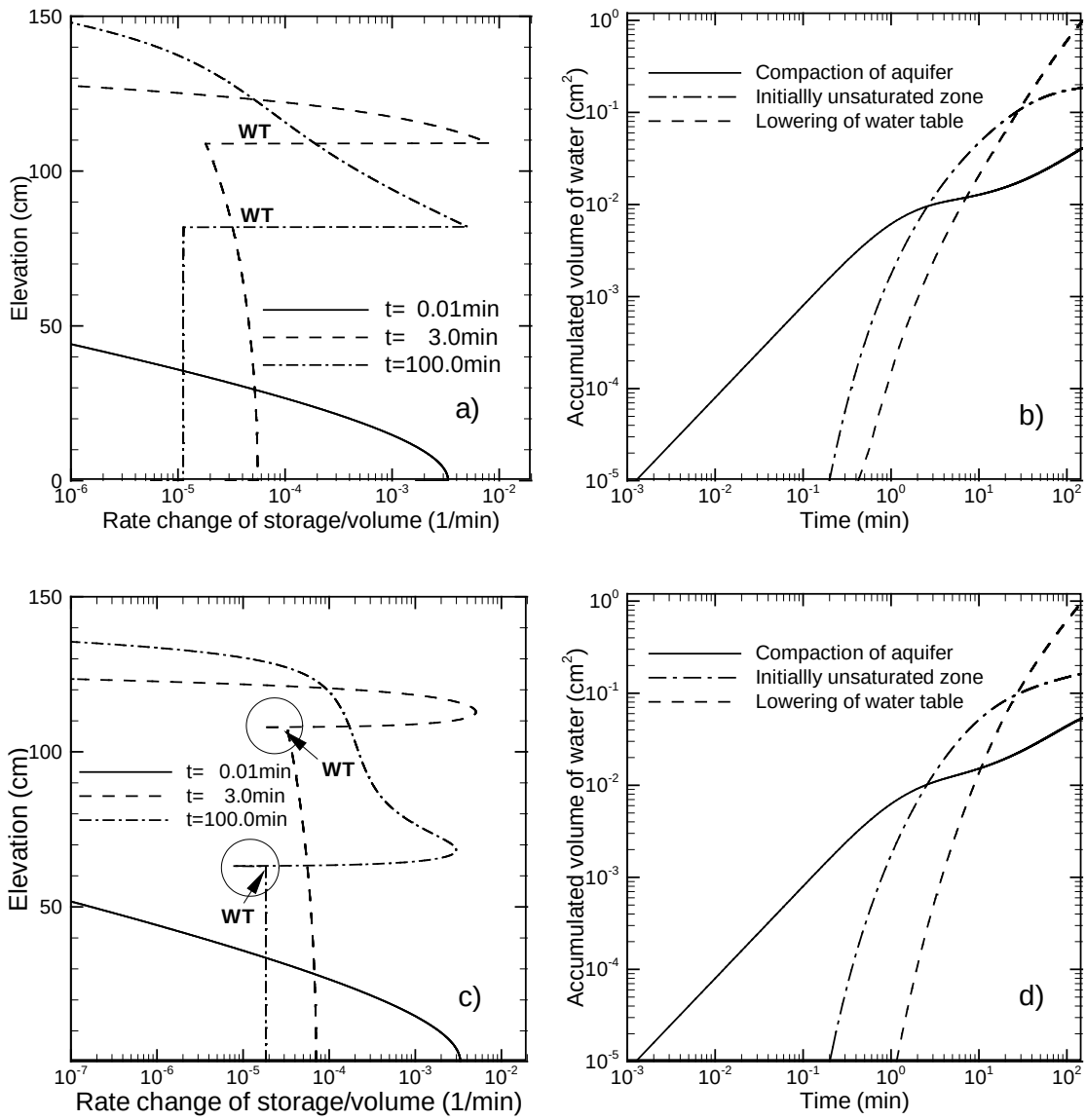


Figure 18.4: The rate change in storage as a function of the elevation in a 1-D soil column for a] exponential model and b] VGM models at time 0.01 min, 3.0 min and 100 min after drainage started. WT denotes the water table. The initial water table before draining is at 110cm. b] and d] are the cumulative volume of water released from 1] compaction of the aquifer, 2] drainage from the initial unsaturated zone, and 3] drainage from pores during falling of the water table over the column at a function of time for exponential and VGM constitutive models, respectively.

At time equal to 100 min, the water table has dropped significantly. The rate of change in storage of the unsaturated zone above the water table becomes very large. Conversely, the Ω values at all locations below the water table become very small and remains constant over the depth, indicating that flow in the saturated zone below the water table is close to a steady state. That is, discharge from desaturation of previously saturated pores due to falling of the water table and desaturation of the unsaturated zone above are transmitted to the bottom discharge boundary without causing a noticeable change in storage in the region below the water table. At

this time, the discharge from the column is mainly from the drainage of pores above the water table. Notice that the behavior of the Ω value in the 1-D experiment is very similar to that in the 3-D experiments [Figure 18.2].

Figures 18.4b and 18.4d depict the cumulative volume of water released over the length of the column as a function of time simulated using the exponential and the VGM models, respectively. The volume of water released is classified into three groups 1] from compaction of the aquifer, 2] from drainage of the unsaturated zone above the initial water table and 3] from drainage of the initially saturated medium during falling of the water table. The pattern of these curves from the exponential and the VGM models generally are almost identical. The water coming from the initially unsaturated zone of exponential model is slightly higher than the water from VGM model. The contribution from falling of the water table simulated by the exponential model begins a little earlier than that by the VGM model [0.3 vs. 1.3 min] due to the distinct air entry value and greater water holding capacity of the VGM. At time greater than 20 min, drainage of initially saturated media due to falling of the water table becomes the dominant water release mechanism.

These results suggest that the use of the VGM or the exponential model does not significantly change the result in general. The exponential model will be used in the rest of analysis.

18.3.2.2 Effects of Hydraulic Conductivity Curve vs. Moisture Retention Curve

The comparison of the simulated drawdown-time curves based on the VGM and the exponential models indicates that the moisture/pressure constitutive relation seems to have a greater influence on the drawdown than the hydraulic conductivity/pressure constitutive relation. To confirm this speculation, the 1-D column drainage problem is investigated with the exponential model with different α and β. The parameter α dictates the rate of reduction in unsaturated hydraulic conductivity, while β controls the rate of reduction in moisture content as the medium becomes less saturated. Figures 18.5a and 18.5b illustrate the influence of parameters α and β, respectively, on the log-log drawdown-time behavior at elevation z = 20 cm in the column. With an increase of α value, drawdown becomes greater and the S shape of the drawdown-time curve becomes less obvious. Conversely, when the value of β increases, the drawdown becomes smaller, and the S shape becomes more pronounced. That is, a porous medium which has a smaller air entry value [a large β value] in the moisture retention curve can release more water from pores than a medium with a larger air entry value [a small β value]. This finding is consistent with the previous findings of the effects of the exponential and VGM models on the S-shaped drawdown-time curve. As indicated by these results, the influence of the two parameters on drawdown is entirely different. Therefore, it is inappropriate to treat these two parameters as identical as has been done in the past [e.g., Tartakovsky and Neuman, 2007]. In effects, laboratory experiments by Yeh and Harvey [1990], Moench [2008], and others have demonstrated that the two parameters are not necessarily the same.

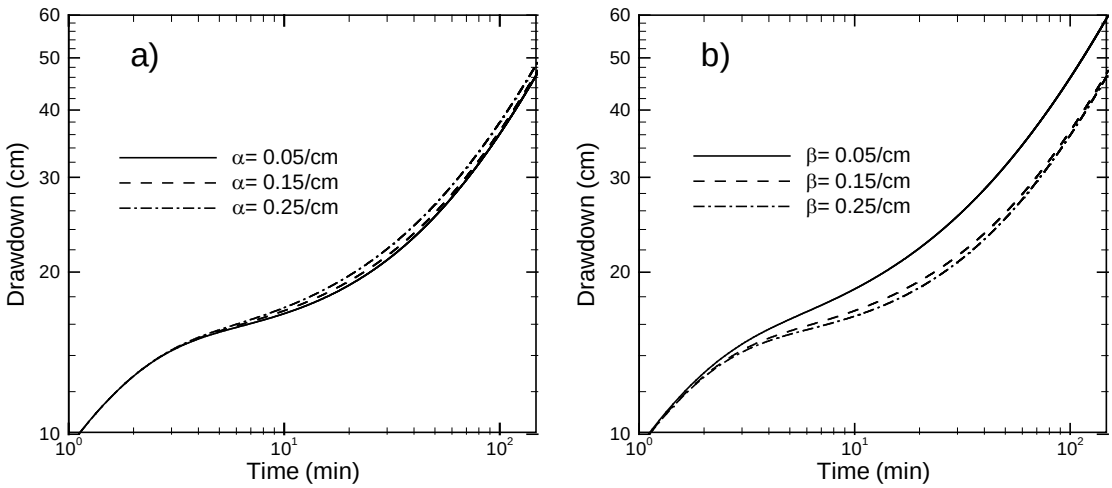


Figure 18.5: a] Drawdown-time curve at z=20cm with different α values when β=0.15/cm and b] Different β values while fixing α=0.15/cm.

18.3.2.3. Role of an Initially Unsaturated Zone

While the important role of the vadose zone [initially unsaturated zone] has been emphasized by many previous studies as discussed in the introduction, our study here advocates the importance of transition of the water release mechanisms in variably saturated flow process during an aquifer pumping test. That is, we argue that even in a fully saturated aquifer without an initially unsaturated zone, the transition between the two water release mechanisms may occur when the discharge is greater than recharge and thus initially fully saturated pores are desaturated. As such, S-shaped drawdown curves likely will occur during pumping in such an aquifer.

In order to test this hypothesis, we conducted a new numerical column experiment by relocating the upper impermeable boundary to the location of the initial water table of the previous column experiment. As a result, the length of the new column is reduced to 110 cm from the length of 150 cm of the previous experiment. The new column is initially fully saturated with a hydrostatic pressure head distribution with zero pressure at the top. We shall call the previous experiment as Case 1, which involves the vadose zone, and the new experiment as Case 2, in which the vadose zone does not exist. Figure 18.6 shows the log-log drawdown-time curves at two elevations [z = 20 cm and 50 cm] for the two cases. In both cases, the log-log drawdown-time curves follow an S shape. At the beginning, the drawdown curves are identical for the two cases since the water is released strictly from the specific storage of the column of the two cases. Afterwards, water is released mainly from unsaturated zone and desaturation of the initially fully saturated pores due to falling of the water table for Case 1. On the other hand, in Case 2, water is released mainly from the latter. While log-log drawdown-time curves of both cases are similar, a major difference between the two cases is that for the case without an initial unsaturated zone, the drawdown is slightly greater since no water from initially unsaturated zone is available as in Case 1. At large times, behaviors of the two drawdown-time curves appear to be the same although the drawdown in Case 2 is slightly higher.

Results of this experiment further confirm that the “additional” source of water causing the S-shaped drawdown curve is not limited to the additional water from the initial vadose zone above the water table of an unconfined aquifer. It also includes the drainage of initially fully

saturated pores in the saturated zone. Therefore, the transition of water release mechanism during vertical flows is the cause of the S-shaped drawdown observed in the saturated zone during pumping in an unconfined aquifer. The concept of the expansion of capillary fringe proposed by Nwankwor et al [1992] and the delayed yield concept are phenomena created by this transition of water release mechanisms.

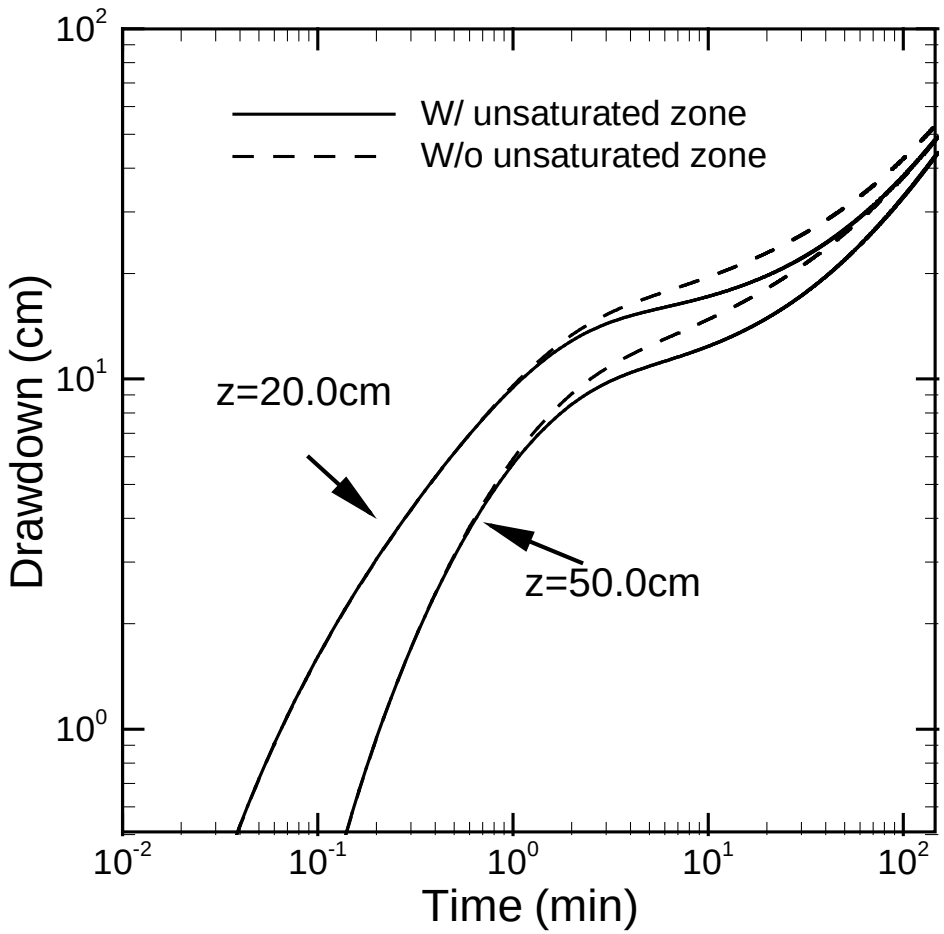


Figure 18.6: The drawdown-time curve at two elevations z=20cm and 50cm showing the role of an initial unsaturated zone.

18.3.3. Flow to a Well in Heterogeneous Aquifers

18.3.3.1 Drawdown-time Curve

As mentioned in section 2, effects of spatial variability of aquifer parameters on the drawdown-time curve will be quantified by an upper bound and a lower bound of the drawdown around its mean as a function of time at an observation point in the aquifer. The upper bound is the mean drawdown plus the standard deviation of the head [σ_p] at the given time and location and the lower bound is the mean minus σ_p.

The mean drawdown-time curves for the aquifer considered have been presented in Figure 18.1. The temporal evolutions of σ_p at three elevations z = 3m, z = 6m, and z = 7.1m at radius r = 5m from the pumping well in the aquifer due to spatial variability of all six hydraulic parameters [K_s, S_s, α, β, θ_s, and θ_r] are plotted as solid red lines in Figure 18.7 a, b, and c. The three elevations represent the bottom, the middle of the saturated zone and the initially unsaturated zone of the unconfined aquifer. The contribution to σ_p from variability of each of the six parameters is shown as a line with a different symbol in the same figures. Similar plots for locations at r = 30m at elevations z=3m, z=6m, and z=7.1m are illustrated in Figures 18.7d, e, and f respectively. The aquifer considered in these figures is considered to be statistically isotropic with correlation scale 3m in all directions. The variances of the logarithm of parameters are intentionally set to the same variance: 1.0 such that the effect of the variance can be isolated.

According to Figures 18.7a and 18.7d, the head standard deviation in the initially unsaturated zone [note that initial water table is located at z=6.7m] at r=5m and r=30m generally increases with time. However, it decreases slightly at the location close to pumping well [r=5m, Figure 18.7a] at large times. This reduction can be attributed to the fact that the flow in the unsaturated zone close to the well approaches the hydrostatic condition earlier. Overall, variations in K_s and θ_s are shown to have greater effects on the head variation in the unsaturated zone than those of α, β, and θ_r. and the effect of the variation in S_s is minimal.

For the head in the saturated region, the impact of the variation in K_s is most profound and it persists over time [see Figures 18.7b, c, e, and f]. The effect of variability in S_s is as significant as that of K_s at the early time but it decreases rapidly and stabilizes when the transition of water release mechanism from elastic compaction and expansion to drainage of pores commences or when the variability of unsaturated parameters [α, β, θ_s, and θ_r] takes effect. At large times, the effect of the variability of θ_s becomes as significant as those of K_s and greater than effects of variability of α, β, and θ_r.

Comparisons of Figures 18.7a, 18.7b, and 18.7c with Figures 18.7d, 18.7e, and 18.7f indicate that the temporal evolution of the head standard deviation due to variability of each parameter at the same elevation at different radii from the pumping well follows similar patterns. In the saturated zone, the head standard deviation increases rapidly at the early time, stabilizes slightly at the intermediate time, and increases again at the large time [Figure 18.7b, 18.7c, 18.7e, and 18.7f]. This seems to suggest that the head standard deviation is dependent on the mean hydraulic head. The magnitude of the head standard deviation generally is smaller at larger distance [r=30m] from the pumping well.

In general, smaller variances of the parameters are found to yield a smaller σ_p while the temporal evolution of σ_p remains the same as those shown in Figure 18.7. Similarly, as the statistical anisotropy increases [i.e., the correlation scales in the horizontal directions become greater than the vertical one], the magnitude of σ_p increases and the temporal behavior of σ_p

remains similar to as those of aquifers with statistical isotropy [Figure 18.7]. This implies that effects of variability of parameters on head increase if the geologic formation is stratified.

The above analysis shows that even the variance is assumed to be the same, the variability of of K_s, S_s or θ_s has greater consequential impacts on the drawdown-time curve in the saturated zone than the variability of α, β, and θ_r. While few studies have quantified the variance of the parameters of the exponential model, studies of the variance of the VGM parameters at several field sites [Russo and Bouton, 1992] have shown that the variability of the parameters for saturated media is generally greater than that of parameters for unsaturated media. Therefore, the spatial variability of parameters of media under unsaturated conditions do not have significant impacts on the drawdown-time curve in the saturated zone during pumping tests in unconfined aquifers.

These results may also suggest that identifying the spatial variability of α and β values of the exponential model for the vadose zone using the head observed in the saturated zone may be difficult unless head and moisture content measurements in the unsaturated zone are used.

To demonstrate the utility of the head standard deviation, a single realization of heterogeneous unconfined aquifers is created. In this example, we only consider the combined effects of spatial variability of K_s, S_s and θ_s parameters since other parameters have minor impacts on the drawdown as indicated in Figure 18.7. To make this example realistic, vertical and horizontal correlation scales of the heterogeneity are set to be 0.15m and 3.0 m, respectively, the same as the statistics of the Borden aquifer [see Sudicky, 1986]. Variance for K_s specified as 0.25 according to the work by Woodbury and Sudicky [1991] and the same for S_s. A variance of 0.05 is assigned for θ_s. The mean values of these parameters, pumping rate, and setup are the same as values used in the previous 3D homogeneous model.

Figure 18.8 shows the simulated drawdown as a function of time at r = 5m and z = 3m in this heterogeneous aquifer, the corresponding mean drawdown-time curve predicted based on aquifer homogeneity assumption, and upper and lower bounds associated with the mean curve. Apparently, the drawdown-time curve of this realization of heterogeneous aquifers is different from the one based on a homogeneous model using the mean parameter values and is bounded by its upper and lower bounds. In other words, the upper and lower bounds are a measure of uncertainty associated with prediction based on the homogeneous model [or likely deviation of drawdown in a heterogeneous aquifer from the predicted one assuming aquifer homogeneity]. The deviation reported by Akindunni and Gillham [1992] of the observed drawdown in the Borden aquifer from that predicted by an equivalent homogeneous model thus is likely due to local heterogeneity neglected by the model. That is, one should not force an equivalent homogenous model to reproduce exactly the observed drawdown-time curve in a field aquifer. The upper and lower bounds can serve as the calibration target to avoid over-fitting of the model.

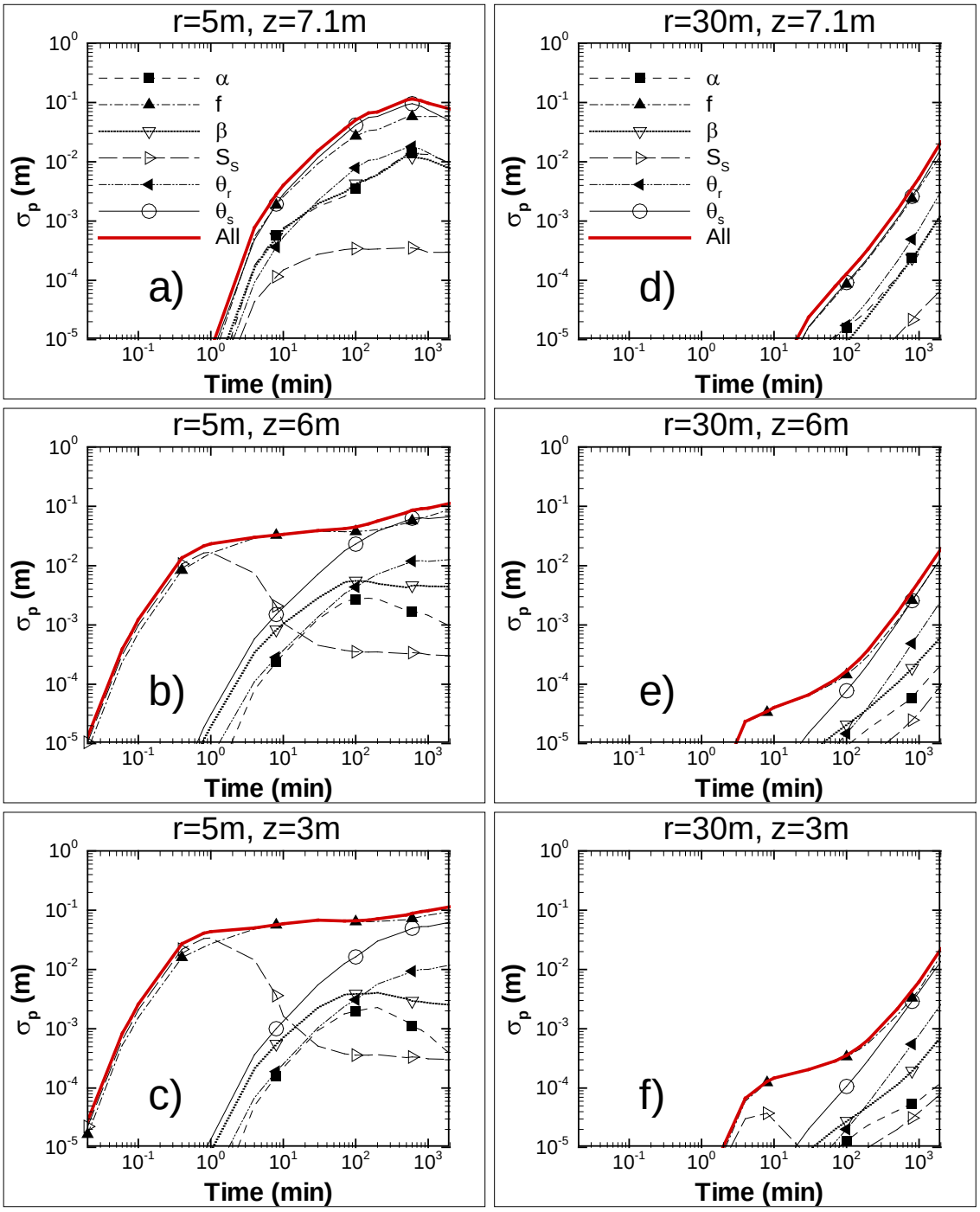


Figure 18.7: The standard deviation [σ_p] of the head for every parameter as a function of time with variances 1.0 and correlation scale 3m. a] to f] represent the results for observation point at two different radius r=5m and 30m and three different elevations z=3m, z=6m and z=7.1m. The red solid lines show the combined effect of all the six parameters.

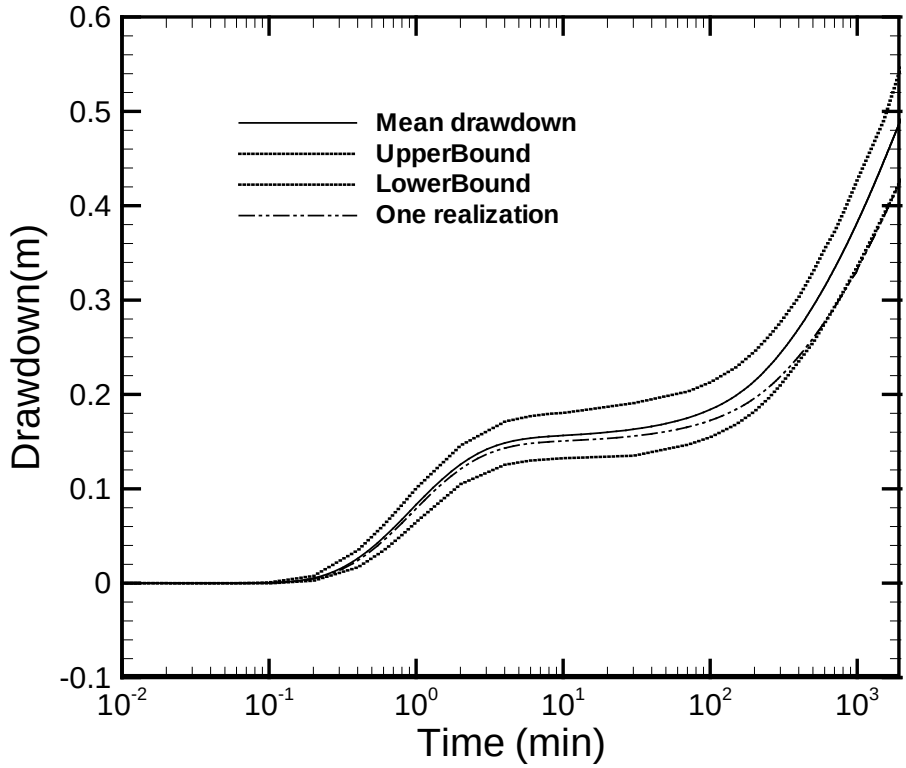


Figure 18.8: Deviations of drawdown-time curve in an aquifer with heterogeneous K_s, S_s and θ_s parameters from the drawdown-time curve derived with the assumption of aquifer homogeneity; the upper and lower bounds based on the first-order stochastic moment analysis that quantify the deviation.

18.3.3.2 Cross-correlation Analysis

The degree of influence on the head at a given location in the aquifer by different properties at every part of the aquifer of the above example at different times is discussed next. Figure 18.9 shows the spatial distributions of cross-correlation value between the head observed at r=5m and z=3m and K_s in a cross-sectional vertical plane running through the pumping well and the observation location. At the early time [0.08 min] when the water table remains static, the head at the observation location is negatively correlated with K_s values in the region between the pumping well and the observation location [Figure 18.9a]. At intermediate time [t=10 min] when the water table starts to fall, the head becomes positively correlated with K_s values in the area above the observation point and the pumping well, and is negatively correlated with K_s values in the region between the observation location and pumping well, below the positive correlation zone [Figure 18.9b]. This pattern of correlation [from upper positive correlation region to the negative correlation region] follows the direction of the vertical flow. This intermediate time period corresponds to the time when the flat portion of the drawdown-time curve occurs [Figure 18.1]. At the large time [1000 min], the lowering of the water table near the pumping well becomes noticeable. The positive correlation region at t = 10 min breaks into two: one at the top left of the observation and the other on the right of the pumping well [figure

18.9c]. Such changes in the correlation pattern seem to follow the change of the flow field [i.e., from mostly vertical flow to greater horizontal flow]. The overall non-symmetrical pattern is consistent with results from Mizell et al. [1982], Li and Yeh [1999], Wu et al. [2005] and Zhu and Yeh [2006] for confined aquifers. That is, the head at a given location is positively correlated with the K_s of the up gradient region, negatively correlated with K_s in the down gradient region with respect to the head location along the stream line toward the pumping well. It is slightly positively correlated with the K_s along stream lines on the other side of the pumping well.

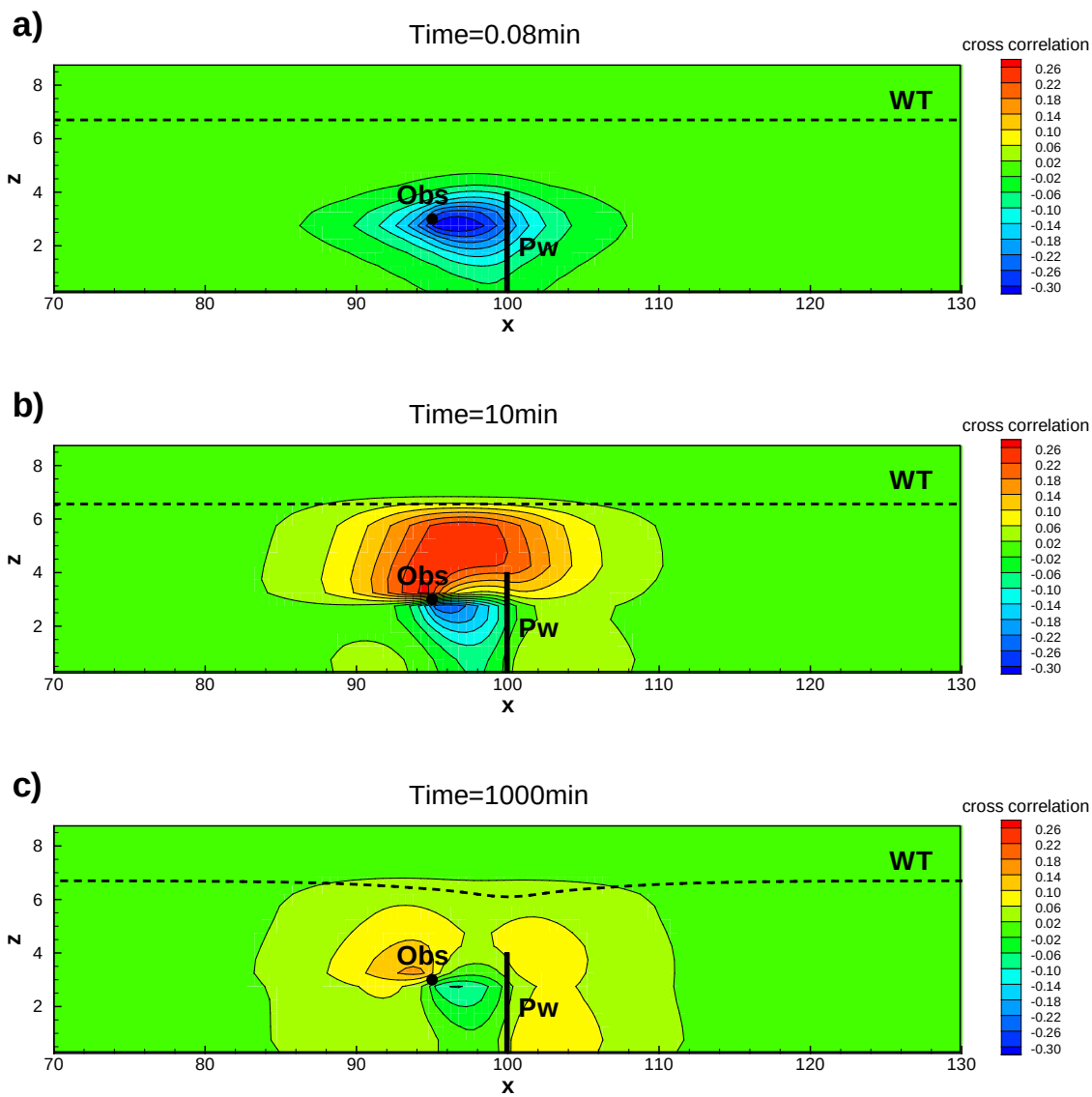


Figure 18.9: Spatial distributions of the cross-correlation between head at observation point[Obs. x=95.0 m, y=100.0 m and z=3.0m] and K_s along a vertical plane running through both the pumping well [Pw] and the observation point at a] the early time, b] the intermediate time, and c] the late time of an S-shaped drawdown-time curve.

Similar to previous studies [Wu et al., 2005], positive correlation values between the head and S_s field are found to be confined to a narrow region between the pumping and the

observation location at early times [Figure 18.10]; this correlation diminishes to zero as flow near the pumping well approaches a quasi-steady state during the late times.

Very small positive cross correlation of the observed head and θ_s appears at the early time t = 10 min and is confined to the region above the observation location and the water table. This positive correlation then continues to increase and the size of the correlated area expands with the expansion of cone of depression as time progresses [Figure 18.11]. That is, the head at the observation location at the late time is highly influenced by the θ_s values in this region.

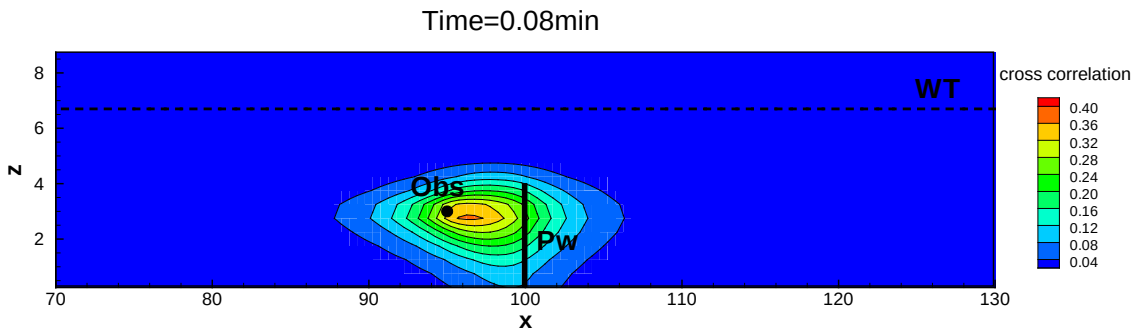


Figure 18.10: Spatial distributions of the cross-correlation between head at observation point [Obs. x=95.0 m, y=100.0 m and z=3.0m] and S_s along a vertical plane running through both the pumping well [Pw] and the observation point at the early time of an S-shaped drawdown-time curve. The cross-correlation diminishes as the time progresses.

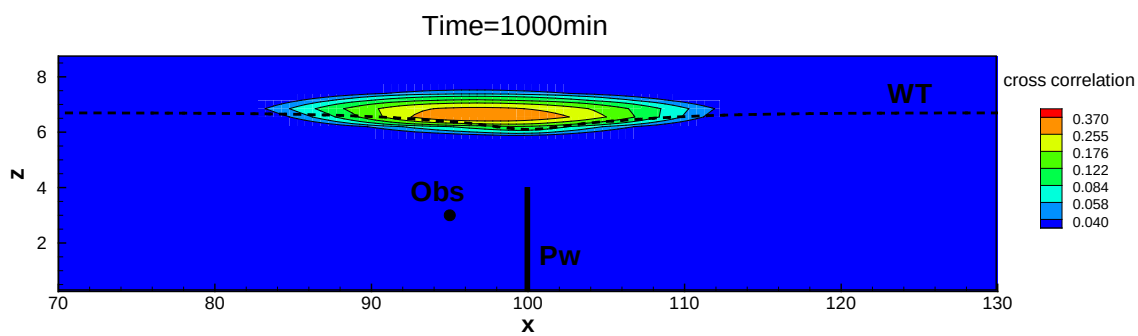


Figure 18.11: Spatial distributions of the cross-correlation between head at observation point [Obs. x=95.0 m, y=100.0 m and z=3.0m] and θ_s along a vertical plane running through both the pumping well [Pw] and the observation point at the late time of an S-shaped drawdown-time curve.

These results demonstrate that the drawdown at an observation location in the aquifer is not equally influenced by heterogeneity everywhere in the aquifer. Such a finding thus raise the same salient question about the representativeness of the estimates using models that assume homogeneity and a limited number of observations as did by Wu et al [2005]. Wu et al. [2005], Straface et al. [2007], and Wen et al. [2010] demonstrated that estimated effective T and S of an equivalent homogeneous, confined aquifer based on a limited number observations vary with locations of pumping and observation wells. Likewise, the estimated T anisotropy of the equivalent homogeneous aquifer varies in magnitude and directions unless a large number of observation wells are used such that ergodicity assumption [i.e., equivalence of ensemble and

spatial average] embedded in the homogeneous flow equations is satisfied. We believe the same issues confront the estimation of parameters for unconfined aquifers by application of a model based on the homogeneity assumption without using a large number of observation wells [e.g., Tartakovsky and Neuman, 2007]. These issues certainly deserve a rigorous analysis, which is however beyond the scope of this paper.

Finally, results of our analysis point to a simple fact: a multi-dimensional variably saturated flow model such as equation [18.1], which considers the transition of water release mechanisms and heterogeneity would provide a more realistic representation of flow processes in the unconfined aquifer during a pumping test. In addition, our results show that variability of unsaturated constitutive relationships in the unsaturated zone does not play a significant role in the head in the saturated zone except θ_s, which does not vary significantly in general. As such, recently developed hydraulic tomography [Tosaka et al.,1993; Gottlieb and Dietrich, 1995; Vasco et al., 2000; Yeh and Liu, 2000; Bohling et al., 2002; Liu et al., 2002; Brauchler et al., 2003; and Zhu and Yeh, 2005 and 2006; Li et al., 2007; Liu et al.; 2007; Illman et al., 2007, 2008; Fienen et al., 2008; Castagna and Bellin, 2009; Cardiff et al., 2009; Yin and Illman, 2009; Xiang et al., 2009; Illman et al., 2010] coupling with equation [1] using spatially averaged unsaturated constitutive relationships for the unsaturated zone [i.e., neglecting their spatial variability] may foster a practical way for delineating K_s and S_s heterogeneity in the saturated region of the unconfined aquifer.

19. HYDRAULIC TOMOGRAPHY TO CHARACTERIZATION OF HETEROGENEITY OF UNCONFINED AQUIFERS

19.1. INTRODUCTION

While detailed characterization of the subsurface is critical to groundwater resources management and remediation, classical cross-well aquifer test and analysis has been shown to yield spatially averaged hydraulic properties of an aquifer. The estimated values depend on the location of observation and pumping wells, and heterogeneity [Wu et al., 2005, Liu et al., 2007, and Straface et al., 2007]. To overcome this problem and to provide high-resolution aquifer characterization, a new aquifer characterization technology, known as hydraulic tomography [HT], has recently been developed [e.g., Neuman 1987a; Gottlieb and Dietrich 1995; Vasco et al., 2000; Yeh and Liu, 2000; Bohling et al., 2002; Brauchler et al., 2003; Zhu and Yeh, 2005 and 2006; and others]. In simple terms, HT is a sequential cross-hole hydraulic test followed by inversion of all the data to map the spatial distribution of aquifer hydraulic properties. Although the ability of HT remains to be fully assessed under field conditions, results from sandbox experiments by Liu et al. [2002], Illman et al. [2007 and 2008], and Liu et al. [2007] are encouraging. These studies showed that the transient HT can identify not only the pattern of the heterogeneous hydraulic conductivity [K_s] field, but also the variation of specific storage [S_s] in the sandbox. These estimated K_s and S_s fields from the HT sandbox experiments accurately predicted the drawdown evolution caused by a pumping test that was not used in the HT analysis. A field application of HT to a well field at Montalto Uffugo Scalo, Italy, produced an estimated transmissivity field that is deemed to be consistent with the geology of the site [Straface et al., 2007]. Bohling et al. [2007] and Li et al. [2007] also showed promising results of HT in their field experiments.

Although HT has been successfully applied to confined aquifers to address issues related to heterogeneity, few attempts have been focused on its application to unconfined aquifers. Some of the main reasons are the complexity of an unconfined flow system due to its interface with the unsaturated zone, the difficulties inherent in solving the unsaturated flow problem, and effects of heterogeneity. Over the past few decades, attempts have been made to simplify the problem with idealized approximations so that the problem is made tractable [i.e., principle of parsimony]. Besides the aquifer homogeneity assumption, a key feature that these idealizations must account for is the inflection frequently observed in the log time-drawdown plot of data [S shape drawn-time curve or well hydrograph] from unconfined aquifer pumping tests. That is, when a known unconfined aquifer is pumped at a constant rate, the water-level drawdown follows the Theis [1935] non-equilibrium response at early time; at intermediate times the drawdown is less than that predicted by the Theis model; while at late time, the drawdown again conforms with the Theis response.

Briefly, two categories of simplifications have been used in this regard. Both these simplifications restrict attention to flow in the saturated zone but account for the drainage of water from above the moving water table through the use of a source term in the governing differential equation. The first category is the one-dimensional, depth-averaged, radial flow model with delayed drainage of Boulton [1954, 1963]. As expressed by Boulton [1963], "the water-bearing material through which the water table has fallen during the early stages of pumping does not yield up its water immediately." The Boulton model includes an empirical delay coefficient to represent the water-release process, accounting for the delayed-drawdown region of the time-drawdown graph.

The other category includes those based on a concept of instantaneous and complete drainage at the water table [Dagan, 1967; Neuman, 1972]. These are radial flow models which consider 2-D flow field in the vertical plan along the radial distance, and the delayed¬ drawdown region of the time-drawdown curve is considered to be caused by the downward hydraulic-head gradient below the water table [i.e., gravity delay yield]. Similar approaches and assumptions of instantaneous and complete drainage at the water table were used by Brutsaert [1970], Streltsova [1972a, b], and Lakshiminarayna and Rajagopalan [1978]. Both types of models consider water released during the rapid drawdown at very early time to be the result of elastic effects of the aquifer.

Application of the type curves developed from both types of models to time-drawdown data collected from field pumping tests generally give specific-yield [effective porosity] values that are substantially below those that would be expected on the basis of other methods of measurement [see Nwankwor et al. [1984], Endres et al., [2007]]. For example, most nonindurated sand or gravel aquifers are expected to have values of specific yield in the range of about 0.2 to 0.3. However, applications of Boulton type curves to time-drawdown data measured in a sand- gravel aquifer yielded the values between 0.02 and 0.09 [Boulton, 1963]. Similarly, for glaciofluvial sand aquifers Prickett [1965] obtained values ranging from 0.03 to 0.13. Bonnet et al. [1970] reported values of 0.04 and 0.08 for a predominantly medium-grain sand aquifer. For the same aquifer studied by Bonnet et al., and using the Neuman type curves, Neuman [1975] reported specific-yield values between 0.04 and 0.07.

Nwankwor et al. [1984] speculate that the low values of specific yield obtained from the type-curve methods are the result of an inadequate representation of the drainage processes occurring near the water table. Neuman [1987b] argued that whereas specific yields obtained in the laboratory may be useful for the evaluation of ground-water reserves that may be potentially recoverable over long time periods; they are generally not relevant to the problem of relating ground-water level fluctuations to pumpage.

Numerical modeling of variably-saturated flow by Narasimhan and Zhu [1993] and Akindunni and Gillham [1992] confirmed that a simple exponential release function used by Boulton does not accurately simulate drainage from above the water table. To obtain a more general mathematical approximation of the drainage process, Moench et al. [2001] developed a model that used a linear combination of exponential functions to simulate release from above the water table. Moench [2004] found that this model improved the match between the observed and predicted hydraulic head time-drawdown relationships for unconfined aquifer pumping tests. Further, Moench [2003] suggested that parameters from this model could be used as the basis for estimation of soil-moisture characteristics on a field scale from pumping test data.

Endres et al [2007] compared bulk vadose zone response predicted by the analytical models and inferred from field measurements using hydraulic head data and soil moisture content profiles obtained during a seven day pumping test at CFB Borden, Ontario. They concluded that the water table boundary conditions used in these analytical models do not adequately replicate the mechanisms controlling the vadose zone behavior during a pumping test.

Recently, Tartakovsky and Neuman [2007] presented a semi-analytical model that avoids the simplifying assumptions by Boulton [1954] and Neuman [1972]. However, because of the analytical nature of the model, they had to assume the aquifer and vadose zone are homogeneous and have uniform thickness. Moreover, hydraulic properties of the vadose zone are limited to a highly simplified exponential model. While their study is a step forward in the analysis of unconfined aquifer tests, accuracy of the estimates by their simplified, homogeneous approach

for real-world, heterogeneous aquifers remains questionable, see Wu et al. [2005] and Straface et al., [2007].

Moreover, Narasimhan [2007] contended that "the advent of the digital computers during the 1960s, and their subsequent phenomenal growth have significantly improved our abilities to quantitatively study transient flow of groundwater under field conditions of practical interest. Numerical models, in conjunction with powerful desk-top computers and graphics capabilities enable us to efficiently solve transient flow of groundwater under far more general conditions than could ever be handled analytically. In concluding their paper, Endres et al. [2007] ask, ''. . . what are the mechanisms that need to be incorporated into an unconfined aquifer model to accurately simulate the bulk vadose zone response to pumping?''. The answer includes saturated-unsaturated flow in general three dimensions, well-bore storage effects, actual thickness of the vadose zone, variations in flow rates, effect of external loads, and non-hydrostatic boundary conditions. These process details can now be readily handled with available numerical models. From a practical, data interpretation perspective, the time has come to exploit the power of numerical models to minimize model approximations and give increased credence to physical attributes of aquifer systems in interpreting field data. We are not constrained any more by the mathematical difficulties inherent in obtaining analytical solutions to partial differential equations.

This optimistic statement about the utility of numerical models should be moderated by the recognition that even the most sophisticated mathematics available to us can describe natural earth systems only as approximate idealizations. Outputs from mathematical models serve us best when they are appreciated with judgment stemming practical field experience."

In reply to Narasimhan [2007], Endres et al. [2007] suggested that "…, we can envision the eventual use of joint geophysical-hydraulic data sets for pumping test analysis in unconfined aquifers where geophysical images will be the basis for vadose zone characterization."

We share the view of Narasimhan [2007], Endres et al. [2007], Nwankwor et al. [1992], and others about the limitation of analytical models, the need for better understanding of processes in the vadose zone, and the necessity of exploiting the power of numerical models and hydrogeophysics. Moreover, we contend that aquifer heterogeneity neglected in analytical or numerical analysis of aquifer tests also has significantly impaired our understanding of flow behavior in unconfined aquifers [e.g., Wu et al. [2005] for confined aquifers]. Similarly, we believe that geophysical images will be the basis for vadose zone characterization but we caution the danger of ignoring spatial variability of constitutive relation between geophysical and vadose zone hydrologic attributes [Yeh et al., 2002].

In this study, we first describe the equation used for solving the groundwater flow in a heterogeneous unconfined aquifer. We then derive a geostatistical inverse method for estimating hydraulic parameters of an unconfined aquifer using pressure heads from a HT survey. We subsequently test the method using a well-posed case and explore the relation between pressure heads and different parameters. As a proof of concept, we finally apply the method to a synthetic three-dimensional heterogeneous unconfined aquifer.

19.2. METHODOLOGY

We assume groundwater flow in a heterogeneous unconfined aquifer can be described by the Richards equation as follows:

— , [19.1]

where ψ is pressure head, which is positive when the aquifer is fully saturated and is negative when the aquifer is unsaturated, and z is the vertical coordinate positive in upward direction; the term is hydraulic conductivity, which equals saturated hydraulic conductivity K_s when and is pressure depended when ; The term S_s represents specific storage; β is a saturation index, which is one when ψ is non-negative and is zero when ψ is negative; t is time; and is the moisture capacity. To describe the saturation-pressure relationship in an unconfined aquifer, we use Gardner-Russo model:

, [19.2]

, [19.3]

where α is the a parameter related the soil pore size distribution; and are saturated water content and residual water content, respectively; and m is a parameter related to soil tortuosity and is chosen to be zero for mathematical simplicity. Eq. [19.1] is solved by the finite element method. Since it is a nonlinear equation, Eq. [19.1] is solved by the Newton-Raphson method.

To estimate hydraulic parameters for an unconfined aquifer, we use the successive linear estimator [SLE]. The SLE is a geostatistical inverse method that seeks mean parameter fields conditioned on available point data [i.e, pressure observations and parameter measurements] as well as geologic and hydrologic structures [i.e., spatial covariance functions of parameters]. Originated from the traditional cokriging algorithm, the SLE uses an iterative process to successively update the estimated parameter values using differences between observed and simulated pressure heads, such that the nonlinear relation between pressure heads and parameters is considered. During the iteration process, both conditional means and covariances of the estimates are also updated successively to achieve fast convergence and to quantify uncertainties. The general idea of the SLE can be summarized as

[19.4]
[19.5]

where in Eq. [19.4], is the first estimation of parameter [i.e., K_s, S_s, and α] at location **x**; is the observed parameter value, represents the observed pressure heads from the first test; and are weights calculated using statistic moments [i.e., covariances of parameters, covariances of pressure head observations, and the cross-covariances between pressure head observations and parameters] by minimizing the estimation variance. Those statistic moments are evaluated by a first order approximation of the Taylor series expansion. In Eq. [19.5], and are the ith and [i-1]th estimation of the parameter; is the weights calculated from residual covariances of observed pressure heads; is the observed pressure heads from the kth test, and is the calculated pressure heads from the [i-1]th estimated parameters. The pressure head observations from different tests can be added into the inversion process either sequentially or simultaneously. The method of sequential inclusion of data is referred as sequential SLE or SSLE, which reduces the computation burden of an inverse problem by keeping the covariance matrix of pressure head observations in a management size. The SSLE is particularly useful for large data sets. The method of simultaneous inclusion of data is referred

as simultaneous SLE or SimSLE, which provides faster convergence than SSLE and avoids the possible different estimated parameters in SSLE if adding noisy data in different sequences. More detailed description of SLE can be found in Zhang and Yeh [1997], Hughson and Yeh [2000], Zhu and Yeh [2005], and Xiang et al. [2009]. In this study, the pressure head data were interpreted sequentially.

19.3 NUMERICAL EXAMPLES

We illustrate the method by three numerical examples. The purpose of the first example is to test the capability of the inverse method if sufficient data are collected. The second example is to explore the relation between pressure head and different parameters through sensitivity analysis. In the final example, we apply the hydraulic tomography concept to a three-dimensional heterogeneous unconfined aquifer.

19.3.1 One Dimensional Aquifer

Gardner-Russo model [Eq. [19.2]] indicates that, if two $K[\psi]$ $-\psi$ data pairs are collected for a homogeneous soil block, K_s and α of this block can be uniquely solved. Based on above consideration, we tested the SLE method in a steady state unsaturated flow case. The case was a 2 m vertical soil column with 20 soil blocks of same size. Both K_s and α were considered as heterogeneous with values varying from block to block. Figs. 19.1a and 19.1b show the K_s and α field, respectively. Two infiltrations were simulated. The first one set the bottom as constant pressure head of zero and an infiltration rate of 0.01 m/d on top. The second case set the bottom constant pressure head as -1 m and an infiltration rate of 0.001 m/d on top. For each test, 19 steady state pressure observations, one for each block, were collected. K_s and α for the top block were also collected [Fig. 19.2]. Then all the collected data were used in the SLE inverse method to estimate K_s and α for the remaining 19 soil blocks. Fig. 19.3 shows that estimated K_s and α replicate the "true" fields near perfectly. This case demonstrates that SLE can accurately estimate unsaturated flow parameters.

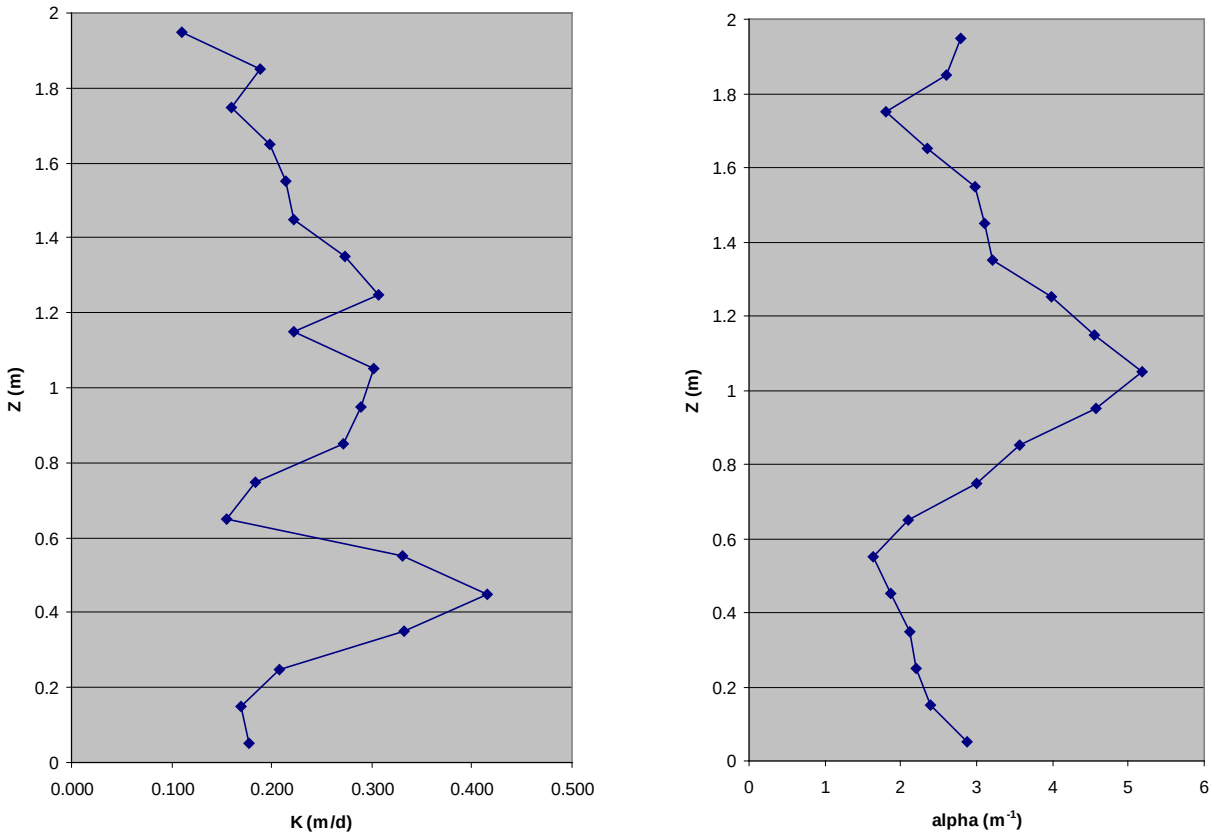


Figure 19.1: [a] heterogeneous K_s field and [b] α field for 1D synthetic aquifer.

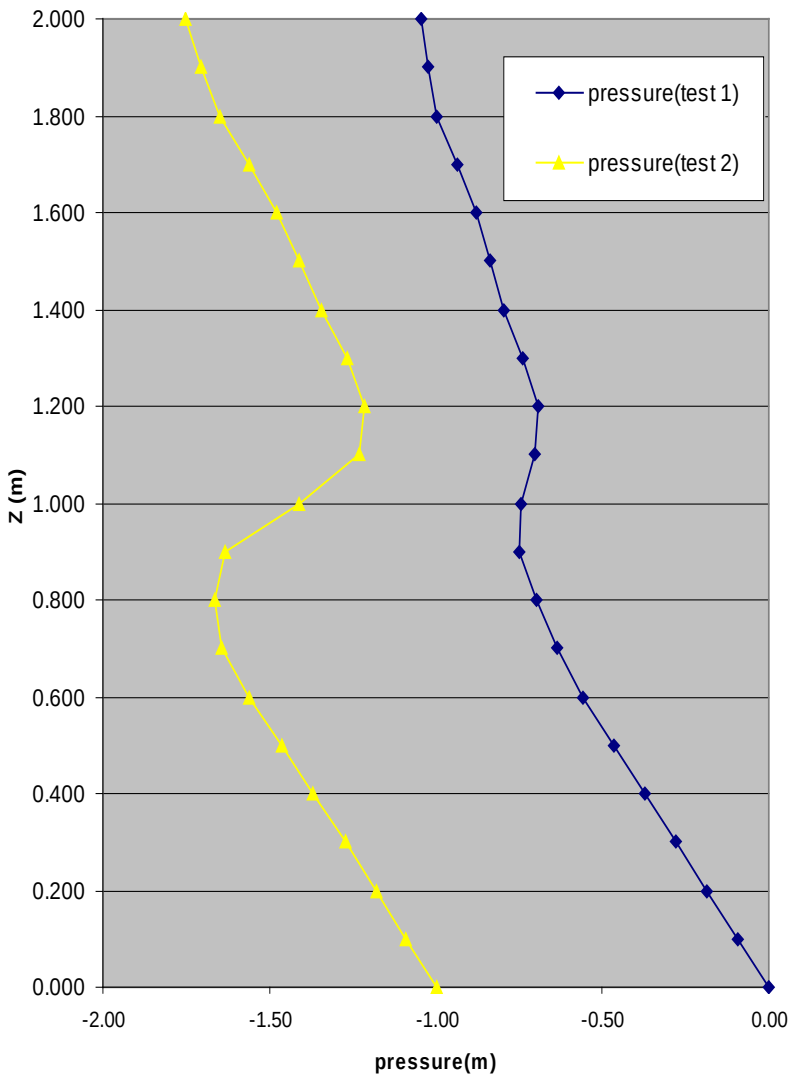


Figure 19.2: Simulated pressure heads of two infiltrations for 1D aquifer.

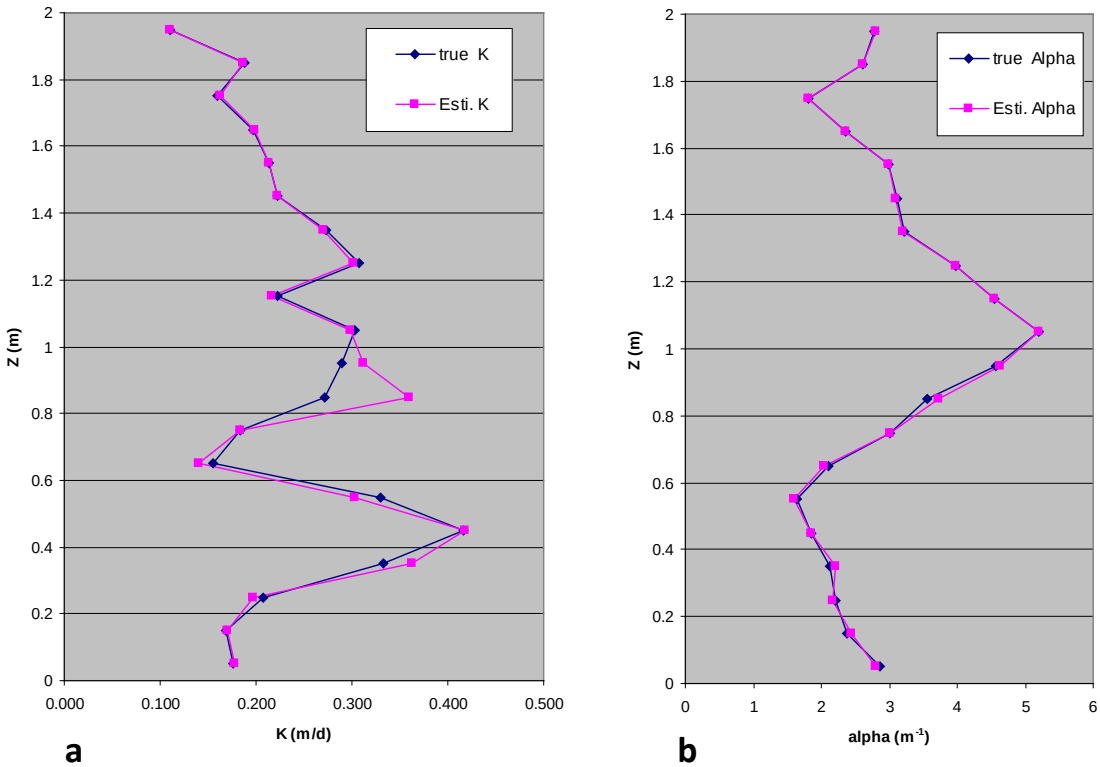


Figure 19.3: Estimated: a] K_s and b] α fields for 1D aquifer.

19.3.2 Sensitivity Analysis

A typical hydraulic tomography survey consists of multiple pumping tests with each test injecting or pumping water from different locations and depths. The transient pressure changes [i.e., drawdown-time curve] during each pumping test are recorded from multiple observation wells. Each drawdown-time curve can have a large number of data points, especially if recorded with automatic pressure transducers. Including all the data points is computational prohibitive and unnecessary. Zhu and Yeh [2005] showed that transient head data at one observation location during a pumping test are correlated to each other. As a result, by choosing observations at several time steps can significantly reduce the computational burden. On the other hand, through sensitivity and cross-covariance analysis, Zhu and Yeh [2005] and Wu et al. [2005] showed, for a confined aquifer, transient heads at different stages of a pumping test have different correlation to K_s and S_s. To keep the usefulness of heads, the chosen time steps should encompass the entire pumping stages, including early time and late time. To understand the relation between transient pressure responses due to a pumping test in an unconfined aquifer and different parameters, a three-dimensional axial radical symmetric aquifer was generated. The homogeneous cylindrical aquifer was 80 m in radius, 10 m in height. The values for K_s, S_s, α, and were 5.70 m/d, 0.000325 m^{-1}, 0.8 m^{-1}, 0.37, and 0.07, respectively. The water table was 3 m below the ground surface. The pumping well was located at the center of the aquifer with screened depth from bottom to 5 m below the ground surface and the pumping rate was 0.001 m^3/s. The observation well was located at depth of 5 m from the bottom and 1 m horizontally from the pumping well [Fig. 19.4].

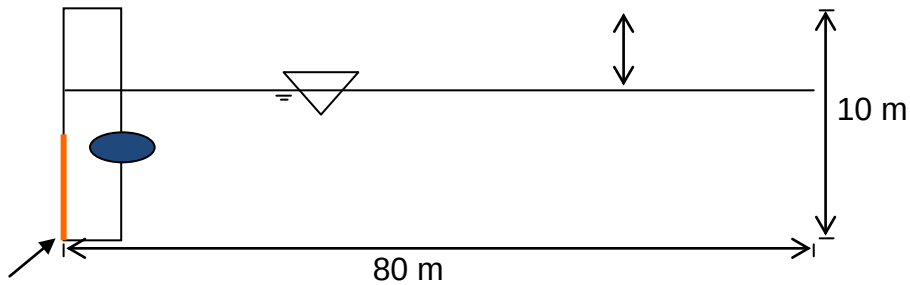
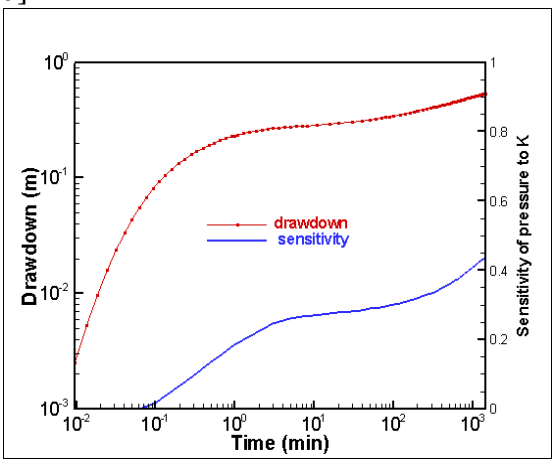
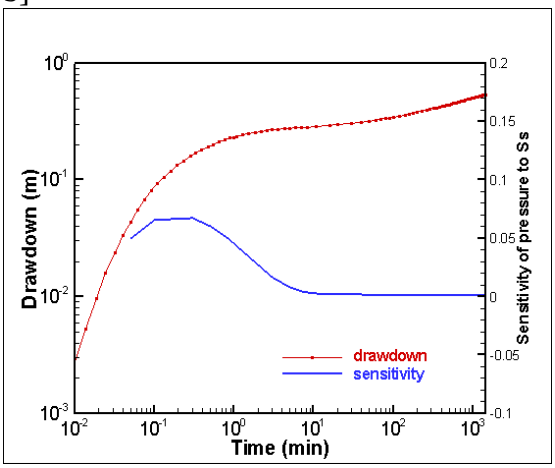


Figure 19.4: Axial radially symmetric aquifer.

a]



b]



c]

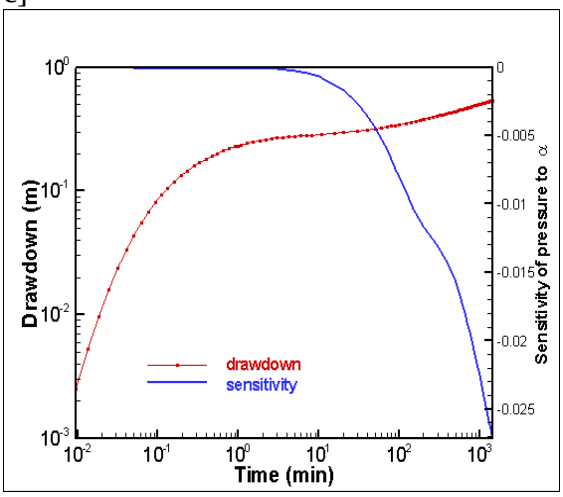


Figure 19.5: Sensitivities of transient pressure to a] K_s and; b] S_s, and c] α

Fig. 19.5 shows the simulated drawdown-time curve and the sensitivities of the transient pressure to the three parameters, K_s, S_s, and α. The simulated drawdown-time curve shows apparent S shape in the middle, which is an indication that the 3-D variably saturated flow model [Eq. 19.1] is appropriate for simulating flow in an unconfined aquifer. Fig. 19.5a shows that the drawdowns at early time are slightly related to K_s and the sensitivity increases with time. Fig. 19.5b shows that only drawdowns at very early time are sensitivity to S_s. Fig. 19.5c shows that drawdowns at early time are not sensitivity to α at all and the value increases with. Notice that the sensitivity of the pressure to α is negative. Fig. 19.5 indicates that, to accurately estimate the three parameters simultaneously, drawdown data at early time, at intermediate time, and at later time should be included.

19.3.3 Three Dimensional Numerical Example

In this section, we applied hydraulic tomography to a synthetic three dimensional heterogeneous unconfined aquifer. This aquifer was 60 m long, 60 m wide, and 18 m in depth [elevation was zero at the bottom and positive upward]. The elevation of water table is 14 m. The heterogeneous K_s, S_s, and α fields were generated by a spectral method [Gutjahr, 1989]. The values of homogeneous and were 0.45 and 0.045, respectively. Total five wells were placed in the center of the aquifer [shown by green lines in Fig. 19.6]. Each well was served as fully-penetrating pumping well and multi-level observation well while not pumping. The five wells were used to simulate a hydraulic tomography survey consisting of five pumping tests. A constant rate of 0.0096 m^3/sec was used for all pumping tests and each pumping test was simulated for 3600 seconds. For each drawdown-time curve, four time steps at 80, 400, 2000, and 3600 seconds were selected. Totally 800 drawdown observations were used for estimating K_s, S_s, and α simultaneously. Chosen data for each pumping test were treated as one data set and data sets from different pumping tests were added sequentially.

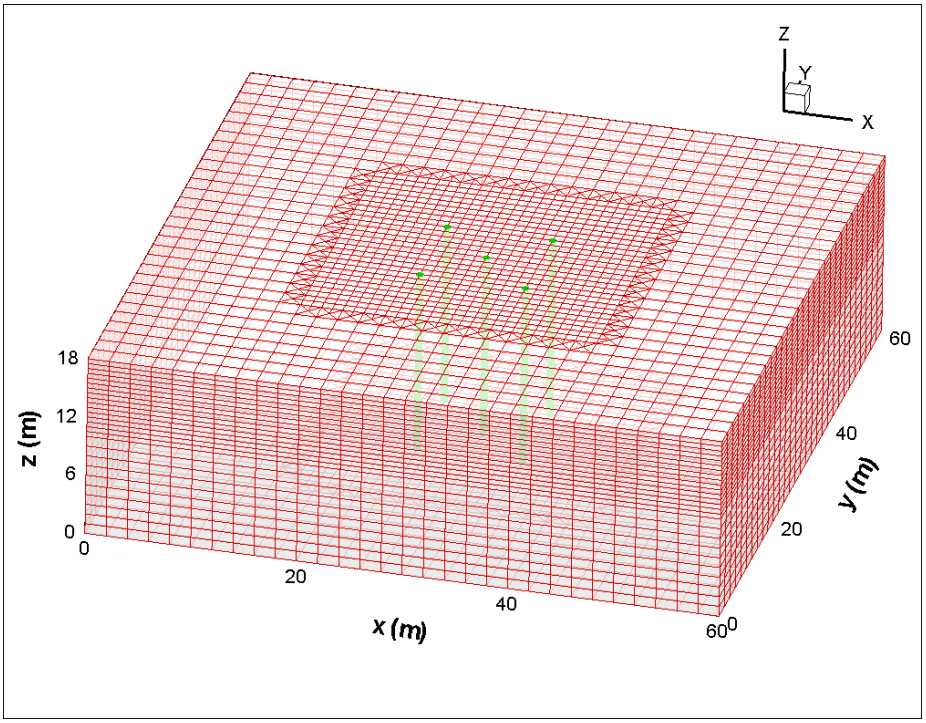


Figure 19.6: Domain and grid of a 3D unconfined aquifer [green columns representing wells].

This example was simulated using multi-grid method, using brick and prism elements in difference sizes [see Fig. 19.6]. The base element is a brick of 2 m × 2 m × 1 m with 1m size in vertical direction. To accurately simulate the transient flow near the water table and in the unsaturated zone due to a pumping test, the vertical grid size was cut in half for the field between elevation12 m to the top. The horizontal grid size for the area encompassing the well field was also cut in half in both x and y directions. The finer brick elements were connected to the coarse brick elements through prism elements. It also needs to point out that although the forward simulation used multi-grid, the inversion process used uniform grid of the size of the base element. As a result, the forward problem had 44512 elements and the inverse problem had 8100 elements. The data interpretation was carried out on a PC-cluster, using eight processors [Intel Pentium 2.8 GHz each]. The total run time was about 3200 minutes.

Figs 19.7 through 19.9 compare the estimated fields with the "true" fields. Fig. 19.7 shows that the estimated K_s field reveals the general pattern of the "true" field near the central area. However, the estimated heterogeneities were limited within the central field due to the short period of pumping time and location of well fields. Fig. 19.8 illustrates that estimated S_s field only shows a large S_s zone near the bottom of the aquifer. Notice that the S_s can only be inferred from saturated pressure data; it is impossible to infer any value of S_s for unsaturated and drainage areas. Fig. 19.9 indicates that estimated field only shows a large α zone near the central top of the aquifer. Again, α is a model-related parameter and can only be inferred from unsaturated pressure data. Nevertheless, as a proof of concept, this case shows the potential of applying hydraulic tomography to characterize unconfined aquifers.

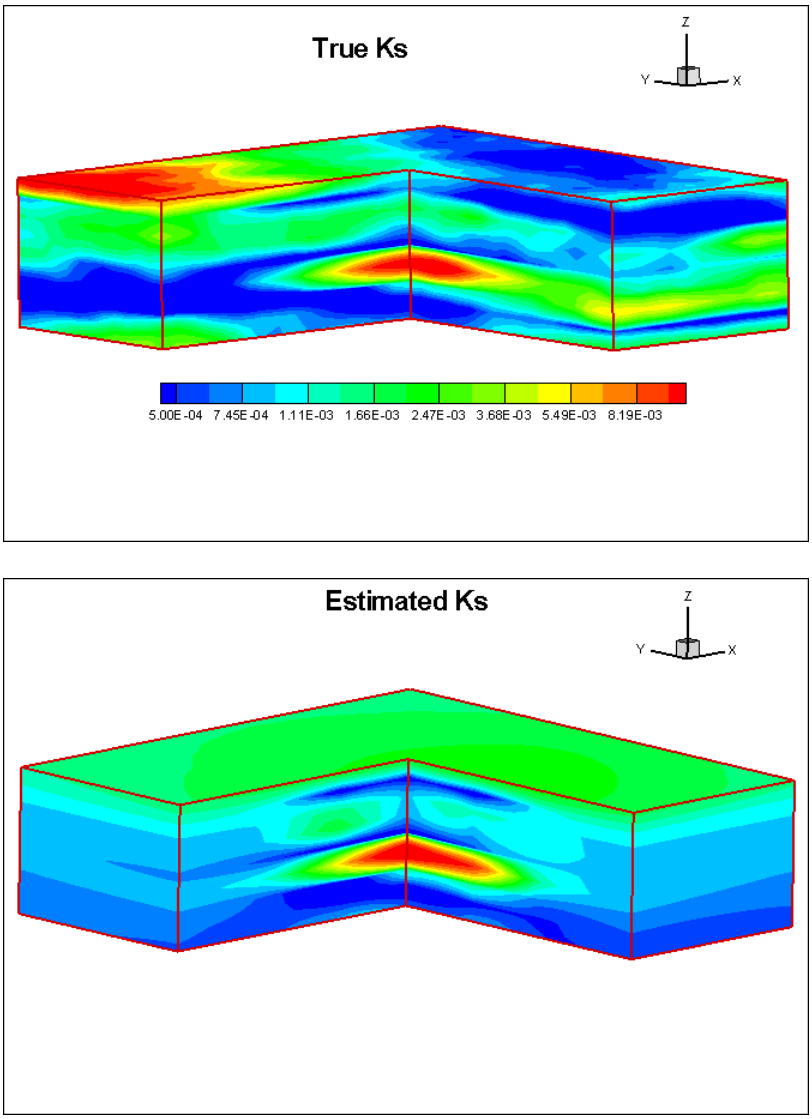


Figure 19.7: Comparison between a] "true" and b] estimated K_s fields.

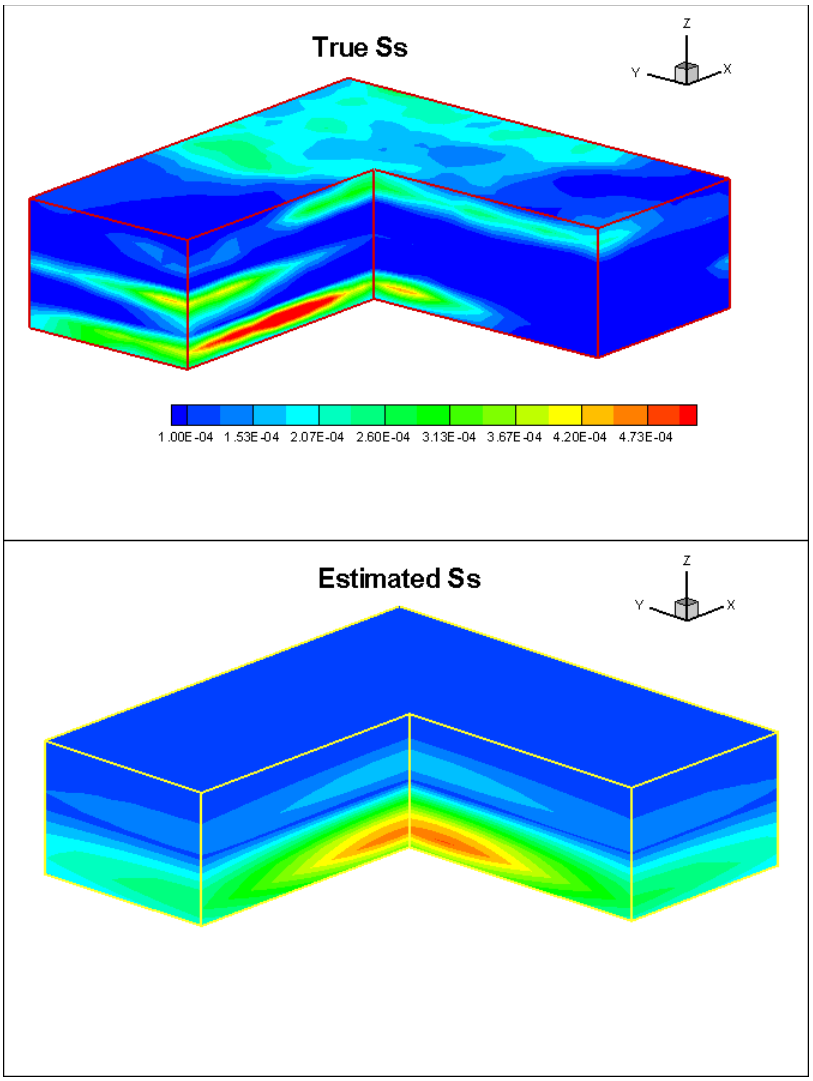


Figure 19.8: Comparison between a] "true" and b] estimated S_s fields.

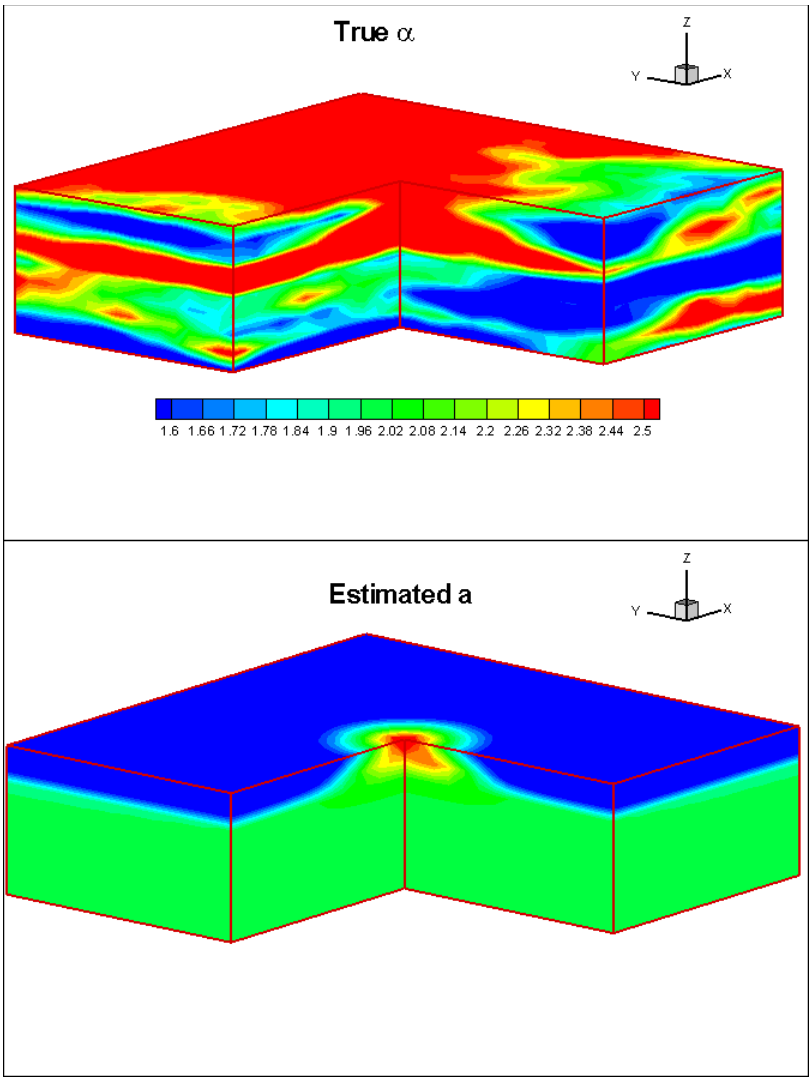


Figure 19.9: Comparison between a] "true" and b] estimated α fields.

19.4 DISCUSSION

In this study, we applied the hydraulic tomography concept to characterize heterogeneities of unconfined aquifers. A three dimensional variably saturated flow model was used to simulate groundwater flow in unconfined aquifers due to the complex nature of groundwater movement near the water table and in the vadose zone. The spatial variations of major hydraulic parameters [K_s, S_s, and α] were estimated by the SSLE method using transient pressure data collected from a hydraulic tomography survey. The method was tested by two numerical examples and the relation between unconfined transient pressure heads and the three parameters were also explored. The synthetic examples demonstrate that a fully variably saturated flow model is appropriate for simulating groundwater flow in unconfined aquifers and hydraulic tomography is a promising tool for characterizing unconfined aquifer. The examples also show that SSLE can handle highly nonlinear problems, providing unbiased estimates of multiple parameters simultaneously. The sensitivity analysis shows that the transient pressure head changes due to a pumping test are related to different parameters non-equally at different stage of the pumping

test. As a result, the pressure data from all stages of the pumping tests should be used to maximize information utilization.

This particular study of developing a hydraulic tomography algorithm that fully considers the effect of the unsaturated zone is merely a proof of concept. This is because the joint estimation of saturated and unsaturated zone parameters from pumping tests in unconfined aquifers is complex considering the nonlinear relationships between unsaturated hydraulic conductivity and water content (as well as the nonlinear relationship between water content and pressure head). Below we discuss some of the limitations and also some potential areas for future research.

First, we only considered three hydraulic parameters [K_s, S_s, and α] as spatial varying parameters in this study. Other parameters, including and , were considered as homogeneous and known a priori. In reality, these parameters are also heterogeneous in nature. However, including too many parameters in the inversion process can potentially cause numerical instability and less reliable estimates. A correlation analysis of pressure responses to all potential spatial varying parameters may be the first logical step to identify a few most responsible parameters that will be estimated through inversion. The remaining parameters will be considered as constant.

Second, the parameter α for Gardner-Russo model is apparently model-dependent. If one other saturation-pressure model was used, new parameters need to be estimated. The parameters for any saturation-pressure model are likely to be spatial varying [Yeh et al., 2002] and should be treated as stochastic processes [Hughson and Yeh, 2000].

Third, A unique aspect of characterizing an unconfined aquifer is that some parameters [such as S_s] only relate to positive pressure heads [i.e., where the flow is in saturated condition] whereas other parameters [such as α in this study] only relate to negative pressure heads [i.e., where the flow is unsaturated condition]. The full characterizations of these parameters for the entire field of interest are typically unobtainable. The aspect can limit the application of HT for unconfined aquifer. For example, if a HT was conducted when the water table is high [i.e, a very wet year], the estimated α will be limited to small portion of the aquifer and will not be reliable for simulating the groundwater flow when the water table is much lower [i.e., a very dry year].

Finally, the variably saturated flow model is nonlinear in nature and requires fine spatial and temporal resolutions for accurate representation of unsaturated flow. The interpretation of field HT for an unconfined aquifer using any inverse method will be likely computational intensive. In this study, although a multi-grid method and parallel computing were used for the fully three dimensional case, the computational cost is still too high for a typical field application. While more sophisticated computing technologies should be pursued to further reduce the computation cost, the future research should also focus on establishing a practical, field-oriented protocol for HT application of unconfined aquifers similar to that developed by Xiang et al. [2009] for confined aquifers.

20. FLOW TO A WELL IN A HETEROGENEOUS UNCONFINED AQUIFER: INSIGHTS FROM AN INTERMEDIATE SCALE SANDBOX

20.1 Introduction

In recent decades the topic of flow to wells during pumping in unconfined aquifers has been both a topic of great interest, and, great debate. The need to better understand and interpret the observations from these tests has led to the development of various analytical solutions [e.g., *Boulton*, 1954, 1963; *Dagan*, 1967; *Kroszynski and Dagan*, 1975; *Streltsova*, 1973; *Neuman*, 1972, 1974; *Moench*, 1997; *Mathias and Butler*, 2006]. *Mishra and Neuman* [2010] provided a comprehensive review of these solutions offering their perspectives on the advancement of the theoretical development of unconfined aquifer analysis. *Mishra and Neuman* [2010] also presented an analytical solution for flow to a partially penetrating well in a compressible unconfined aquifer that infers saturated and unsaturated hydraulic properties from drawdowns recorded in the saturated and/or unsaturated zone. This work is an extension of the solution developed by *Tartakovsky and Neuman* [2007] and adds: i] a more flexible representation of unsaturated constitutive properties; and ii] a finite thickness for the unsaturated zone. The type curve approach was used in conjunction with PEST [*Doherty*, 1994] to simultaneously analyze seven observed drawdown records from a pumping test conducted at the Cape Cod site by *Moench et al*. [2001]. The analysis yielded comparable estimates of hydraulic conductivity [K] and specific storage [S_s] and somewhat higher values of specific yield [S_y] when compared to those obtained by *Moench et al*. [2001] and *Tartakovsky and Neuman* [2007]. Their estimates of the van Genuchten-Mualem parameters were also found to be comparable to laboratory estimates obtained for similar materials in the area. Despite the completeness and usefulness of *Mishra and Neuman's* [2010] solution, analytical solutions that treat the medium to be homogeneous cannot address the issue of heterogeneity, which is the rule rather than the exception. Recently, the homogeneity assumption required by analytical solutions for heterogeneous saturated aquifers have come into question by *Wu et al*. [2005]. *Wu et al*. [2005] used random Gaussian fields [of transmissivity and storativity] to demonstrate that parameter estimates made using the *Theis* [1935] solution varied throughout the duration of the pumping test. Transmissivity estimates approached the geometric mean and storativity estimates were dominated by the material between the pumping well and the observation point. It is reasonable to expect similar limitations for analytical solutions used to estimate parameters from unconfined pumping tests.

To date, a limited number of investigations have explicitly examined the effect of heterogeneity on pumping tests performed in unconfined aquifers [e.g., *Bunn et al*., 2010; *Mao et al*., in review]. For example, using multiple realizations of heterogeneous fields, *Bunn et al*. [2010] examined the connection between K heterogeneity and the capillary fringe extension phenomenon observed during a pumping test conducted by *Bevan et al*. [2005] at the Borden site in Canada. The ensemble mean hydraulic heads were able to reproduce the field observations quite well, however, these simulations were unable to reproduce the capillary fringe extension observed in the field data.

More recently, *Mao et al*. [in review] utilized the stochastic moment approach to better understand the role of heterogeneity in the development of the S-shaped drawdown curves and a cross-correlation analysis to show that drawdowns at locations in a heterogeneous unconfined aquifer are mainly affected by local heterogeneity near the pumping and observation wells. *Mao*

et al. [in review] also looked at the sensitivity of the S-shaped curve to the spatial variability of hydraulic parameters and concluded that it is most sensitive to hydraulic conductivity [K], specific storage [S_s] and saturated water content [θ_s] and insensitive to the variability of unsaturated hydraulic parameters.

Mapping of K heterogeneity in unconfined aquifers was recently reported by *Cardiff et al*. [2009] under steady-state conditions. These authors analyzed multiple drawdown data collected from fully-screened wells completed in an unconfined aquifer located at the Boise Hydrogeophysical Research Site to obtain depth-averaged K distributions through their hydraulic tomography algorithm. Various inverse methods have been developed for hydraulic tomography, which utilize pumping test data simultaneously or sequentially [e.g., *Gottlieb and Dietrich*, 1995; *Yeh and Liu*, 2000; *Bohling et al*., 2002; *Brauchler et al*., 2003; *Zhu and Yeh*, 2005, 2006; *Li et al*., 2005; 2007; *Fienen et al*., 2008; *Cardiff et al.*, 2009; *Castagna and Bellin*, 2009; *Xiang et al*., 2009; *Liu and Kitanidis*, 2011], but to our knowledge, there are no hydraulic tomography algorithms that can interpret the transient drawdown records from a heterogeneous, unconfined aquifer. One potential reason for this is that pumping in unconfined aquifers induces flow regimes that are considerably more complex than in confined aquifers, as it involves flow in both the saturated and unsaturated zones with the water table acting as a time-varying interface [*Mao et al*., in review]. In fact, *Mao et al*. [in review] advocates the use of a variably saturated flow equation for the interpretation of drawdown responses in an unconfined aquifer.

In this study, we ask the question: "if we have the necessary parameters and forcing functions to populate a variably saturated flow model, can we accurately model the responses to pumping in a heterogeneous unconfined aquifer in both the saturated and unsaturated zones?" To answer this question, we designed an intermediate-scale laboratory sandbox containing a heterogeneous aquifer that was instrumented pressure transducers, tensiometers, and water content sensors. A number of cross-hole pumping tests were performed while the tank was fully saturated for characterization purposes. These cross-hole tests were used to estimated both homogeneous and heterogeneous K and S_s distributions. Unsaturated parameters of the different material types in the sandbox were measured using hanging water column experiments.

We next conduct another pumping test in which we allow the water table to fall, while the drawdown responses are monitored at multiple locations in the saturated zone with pressure transducers. Concurrently, we monitor the unsaturated zone with tensiometers equipped with pressure transducers and water content sensors. After this, we model the pumping test using MMOC3 [*Yeh et al*., 1993], a variably saturated flow and transport model to evaluate the ability of different homogeneous and heterogeneous representations of the sandbox aquifer in both the saturated and unsaturated zones to reproduce the observed drawdown responses. By comparing various idealizations of saturated and unsaturated zone parameters, we are able to find out what parameters have the greatest impacts on making accurate predictions of responses in the unsaturated and saturated zones.

20.2 EXPERIMENTAL SETUP

20.2.1 Sandbox design and instrumentation

A synthetic heterogeneous unconfined aquifer was constructed in a vertical, laboratory sandbox to investigate how the saturated and unsaturated zones respond to a pumping test that results in drainage. The sandbox was 244 cm in length, 122 cm in height, and had a depth of 9.4 cm. A glass plate covers the front of the sandbox and a stainless steel plate covers the back. A

total of one hundred four ports were drilled into the stainless plate to allow the sandbox to be instrumented with sensors [pressure transducers, tensiometers, and water content sensors].

Fifty eight of these ports were used for the installation of fully penetrating horizontal wells. Each well was constructed by drilling fourteen 0.5-cm diameter holes along a section of brass tubing. The holes were then covered with a stainless steel mesh that was bonded to the tubing with corrosion resistant epoxy. These wells, which penetrate the thickness of the synthetic aquifer were installed after the packing of the sandbox aquifer. Additionally, 47 of these wells contain 0-2 psig pressure transducers [Model S35] from BHL Instruments. Eleven wells do not contain pressure transducers but can still be used for pumping purposes.

Twenty two of the ports [located in the upper portion of the sandbox] were used for the installation of column tensiometers [CL-029B Flow Cell Tensiometer, Soil Measurement Systems] equipped with Microswitch pressure transducers. The remaining 24 ports [also located in the upper portion of the sandbox] were instrumented with EC5 water content sensors [Decagon Devices Inc.]. All sensors were connected to a National Instruments Compaq Data Acquisition System [NIDAQ] [Model #9178] which allows for real-time monitoring of signals during pumping tests. Figure 20.1 is a photograph of the sandbox showing the locations of the sensors.

For this particular study, all boundaries except the upper boundary are no-flow boundaries. The top boundary was set as a constant head boundary for the saturated cross-hole tests performed for characterization purposes, and left open to the atmosphere for the unconfined pumping test.

20.2.2 Packing of sandbox and types of sands used

Prior to the installation of the wells and tensiometers, and concurrent with the installation of the water content sensors, the sandbox was packed with commercially sieved silica sand [six types; F35, F45, F65, F75, F85, and F110] and silt [four types; Sil-co-sil 45, Sil-co-sil 53, Sil-co-sil 106, and Sil-co-sil 250] from U.S. Silica. This material was packed to match a heterogeneity pattern drawn on the glass side of the sandbox. The sand was slowly added to the sandbox in thin layers [approximately 0.5 to 1 cm thick] and packed between each addition to achieve uniform compaction.

After the sand box was filled, CO_2 gas was injected into the sandbox to displace the air trapped during packing. Over the period of 3 days, CO_2 was injected at a low pressure [2 - 5 psi] from a number of different ports working from the bottom of the sandbox to the top. The top of the sandbox was sealed except for a small hole that allowed the gas to exit. After all of the CO_2 was pushed through the sandbox de-aired water was added at port 3 under a gradient of approximately 10 - 20 cm. This low gradient meant that any air above the water level would be displaced out the top of the sandbox. Since the air in the sandbox was replaced with CO_2, any trapped gas would readily dissolve into the water phase. This procedure minimized the risk of air entrapment and ensured the sandbox was fully saturated.

The predominant materials in the sandbox are the medium sands F35 and F45. Within these sands are discontinuous layers of finer grained material that average 0.8 m to 1 m in length and average 10 to 15 cm thick. The most notable features of this heterogeneity pattern are the three discontinuous lenses near the top of the sandbox [Sil-co-sil 106 in the top left, Sil-co-sil 53 in the upper middle, and F110 in the top right]. These lenses [in particular the silt lenses] are expected to have significantly different saturated and unsaturated properties than the surrounding F35 and F45 sand. Near the bottom of the sandbox, the most salient feature is a layer of silt [Sil-

co-sil 250] that extends along the length of the sandbox expect for near the middle where a window is present. This window connects the F35 sand layer below with the F35 layer above.

Figure 20.1: Photograph of the sandbox showing all sensor locations [● = pressure transducers; ● without pressure transducers; ○ = tensiometers; and **x** = water content sensors] and the various layers that were packed. Sensor locations are approximate.

20.3 CHARACTERIZATION OF AQUIFER HETEROGENEITY

The analysis of pumping tests performed in unconfined aquifers requires knowledge of both saturated and unsaturated parameters. Our analysis will require knowledge K, S_s, θ_s [saturated moisture content], θ_r [residual moisture content], and the *van Genuchten* [1980] fitting parameters describing the primary drainage curve [α and n].

20.3.1 Characterization of K and S_s

While the sandbox was fully saturated a number of cross-hole pumping tests were performed for characterization purposes. During these tests the top boundary of the sandbox was maintained as a constant head boundary by having 3 cm of standing water in the top portion of the sandbox [above the sand]. The extraction line was connected the desired well and the pump was operated long enough to fill the effluent line [which discharged into the top reservoir, creating a loop and ensuring mass balance]. Once the line was filled the pump was turned off and the system was allowed to return to equilibrium [static].

Sixteen ports were chosen for these cross-hole tests [Figure 20.2]. During these tests, one port was pumped, while the pressure responses within the sandbox were simultaneously monitored at 47 points sampling at a rate of 4 Hz. The pumping rates for all cross-hole tests was 240 mL/min. Prior to the cross-hole pumping tests, all pressure transducers were calibrated to

ensure accurate data collection. We then collected hydraulic head data for several minutes in all pressure transducers to establish a static, initial condition. After establishment of static conditions, we pumped from each port using a peristaltic pump, while taking head measurements at all 47 ports. For each test, pumping continued until the development of steady state conditions [approximately 1-2 minutes], which was determined by observing the stabilization of all head measurements on the data logger connected to a computer. After reaching steady state, the pump was shut off and recovery data were collected.

Analysis methods for interpreting pumping test data, whether for confined, semi-confined, or unconfined aquifers often assume the aquifer to be homogeneous, however, this is rarely the case. As such, the cross-hole data collected in this sandbox is analyzed in two manners; 1] A single cross-hole test is used to estimate effective K and S_s values [section 20.3.2] and, 2] eight cross-hole tests are used create heterogeneous K and S_s estimates using hydraulic tomography [section 20.3.3]. The remaining eight cross-hole tests are then used to validate these transient hydraulic tomography generated heterogeneous fields.

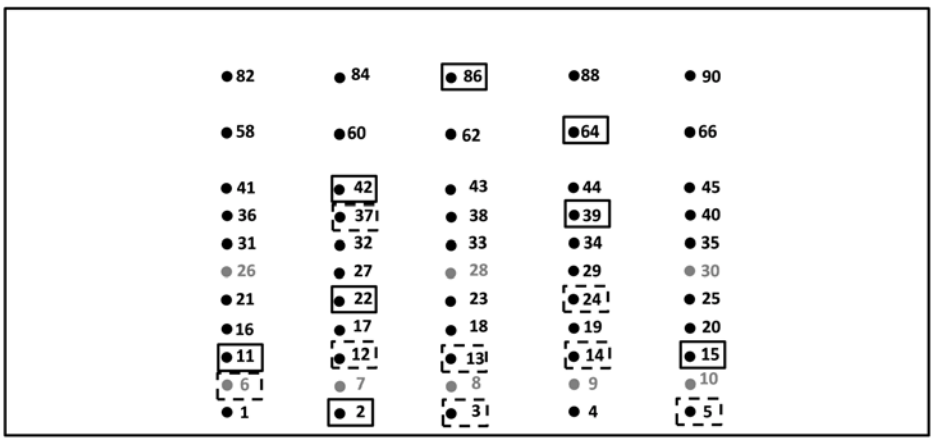


Figure 20.2: Schematic diagram showing the ports that were pumped for the cross-hole pumping tests. Solid black circles represent locations of pressure transducers and solid grey circles represent ports that were not instrumented. Solid squares are pumping locations used for the estimation of heterogeneous K and S_s fields. Dashed squares are pumping locations for the validation of the results from transient hydraulic tomography.

20.3.2 Estimation of Effective K and S_s

Using a cross-hole pumping test performed at port 22, the parameter estimation program PEST [Doherty, 1994] was coupled with MMOC3 [*Yeh et al.*, 1993] to estimate effective K and S_s values representing the entire domain by fitting all the drawdown curves simultaneously. The synthetic aquifer was discretized into 3,645 elements and 7,544 nodes with average element dimensions of 3.0 cm × 9.4 cm × 2.5 cm. The top boundary was set to be a constant head boundary and the remaining boundaries of the sandbox were no-flow boundaries. Six data points [at t = 1, 5, 10, 20, 40, and 80 seconds] from 46 ports totalling 276 data points were used for parameter estimation. The initial estimates of K and S_s were 0.1 cm/s and 1 x 10^{-4} cm^{-1} respectively.

The estimated effective K was 1.85 x 10^{-2} cm/s [with lower and upper 95% confidence bounds of 1.76 x 10^{-2} cm/s and 1.94 x 10^{-2} cm/s respectively], and the estimated effective S_s was 3.94 x 10^{-5} cm^{-1} [with lower and upper 95% confidence bounds of 3.40 x 10^{-5} cm^{-1} and 4.57 x 10^{-5} cm^{-1} respectively]. The estimate parameters were then used to simulate the cross-hole test at

port 22, essentially testing the calibration of the effective parameters. Figure 20.3 is a calibration scatter plot [observed vs. simulated drawdown] for the cross-hole test at port 22. This plot shows the ability of these effective values to reproduce the observed drawdown. The dashed line represents the 1:1 line which would be a perfect match between simulated and observed. The linear model fit [solid line] has a slope near 1 suggesting that the effective values capture the average response in the sandbox for this cross-hole test. The scatter as indicated by the coefficient of determination [R^2] suggests that the details of the heterogeneity are not accurately captured. This is expected since homogeneous values of K and S_s were used for this simulation.

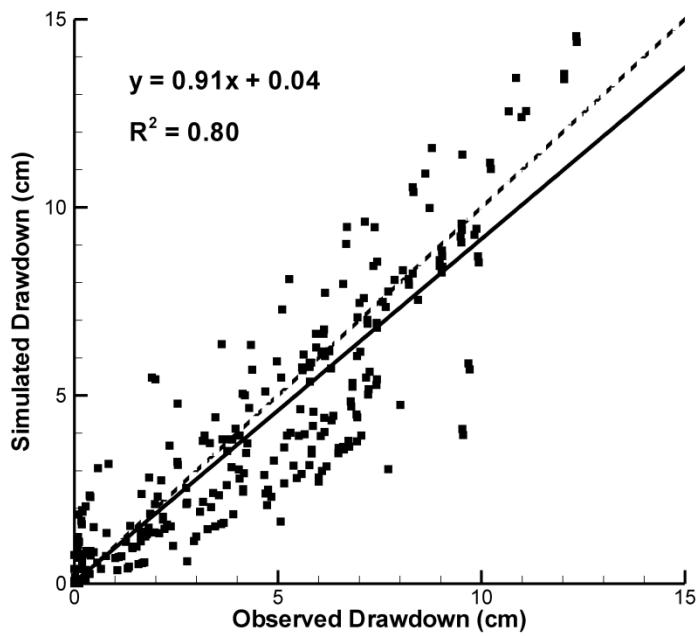


Figure 20.3: Simulated versus observed drawdown for the cross-hole test at port 22 using effective K and S_s values.

20.3.3 Estimation of Heterogeneous K and S_s

The use of hydraulic tomography for the characterization of heterogeneity in sandbox aquifers has been shown to be a robust technique that outperforms other methods which often rely on the interpolation of point scale data [*Illman et al*., 2007; *Liu et al*., 2007; *Illman et al.* 2010; and *Berg and Illman*, in review]. Thus, we elect to use transient hydraulic tomography [*Zhu and Yeh*, 2005] for the characterization of K and S_s within this sandbox.

The Sequential Successive Linear Estimator [SSLE] code developed by *Zhu and Yeh* [2005] was used to conduct transient hydraulic tomography with multiple cross-hole tests performed under fully saturated conditions to obtain a heterogeneous distribution of K and S_s [from now on referred to as K and S_s tomograms]. The model domain used for this inversion is identical to the one used for estimation of effective K and S_s described in section 20.3.1.

Inputs to the inverse model include initial guesses for the K [0.1 cm/s] and S_s [1 x 10^{-4} cm^{-1}], estimates of variances [σ^2_{lnK} = 3.0 and σ^2_{lnSs} = 3.0] and estimated correlation scales for both K [λ_x = 150 cm and λ_z = 20 cm] and S_s [λ_x = 150 cm and λ_z = 20 cm], volumetric discharge [Q_n] from each pumping test where n is the test number, available point [small-scale] measurements of K and S_s, as well as head data at various times selected from the head-time curve. Although

available point [small-scale] measurements of K and S_s can be input to the inverse model, we do not use these measurements to condition the estimated parameter fields to test the inversion algorithm.

In total, we utilized 8 independent cross-hole tests with pumping taking place at ports 2, 11, 15, 22, 39, 42, 64 and 86 for the transient hydraulic tomography analysis. More specifically, we utilized 6 data points [at t = 1, 5, 10, 20, 40, and 80 seconds] from 46 ports totalling 276 data for each test. We excluded the data from the pumped port because of possible skin effects. In total, we utilized 2208 data points from 8 different tests in our transient inversions.

All computations for transient hydraulic tomography analyses were executed using 44 of 48 processors on a PC-cluster consisting [of 1 master and 12 slaves each with Intel Q6600 Quad Core CPU running at 2.4 GHz with 16 GB of RAM per slave] at the University of Waterloo. The operating system managing the cluster was CentOS 5.3 based on a 64-bit system. The total computational time for inverting data from 8 pumping tests was approximately 24 hours.

The resulting K and S_s tomograms are shown on Figure 20.4a and 20.4b. A visual comparison of the K tomogram to the deposits [Figure 20.1] shows that many of the important features are captured. Of particular note are the low K zones which compare very favourably with the real distribution. The two low K layers extending towards the middle from both sides near the bottom of the sandbox are present, as well as the two low K zones near the top left [there is even some indication in the tomogram that these may not be connected, which is the case]. Additionally, the Sil-co-Sil 106 layer near the middle right of the sandbox is captured, however, it does not extend all the way to the boundary. This is not surprising considering there are no observation points in this region, thus, estimates here have greater uncertainty. The K tomogram also preserves the connectivity of the high K units, particularly the window in the silt [Sil-co-sil 250] layer near the bottom middle, and the large window in the silt near the top right. Thus, we feel confident that the K tomogram reflects the heterogeneity of the sandbox.

Examination of the S_s tomogram [Figure 20.4b] reveals a trend of increasing S_s values with increasing elevation within the sand box. Such a trend was also observed in the S_s values from separate sandbox studies packed using different methods, and containing different heterogeneity patterns [*Liu et al*., 2007; *Berg and Illman*, in review]. A likely explanation for this pattern being witnessed in several sandbox studies is that the material near the bottom of the sandbox is much less compressible due to the weight of the overlying material, thus, resulting in low S_s values near the bottom of the sandbox and higher S_s values near the top. Additionally some heterogeneity in the distribution of S_s is present within this trend from top to bottom, however, it does not appear to be strongly correlated with any of the features seen in the K tomogram [Figure 4a] or the picture of the sandbox [Figure 20.1]. This suggests that the vertical trend in S_s is the dominant feature controlling the drawdown response in the sandbox.

To quantitatively assess the results, we plot the simulated and observed drawdown responses [Figure 20.5a] for each cross-hole test used for the inverse modeling effort [i.e., calibration] on separate scatter plots. A linear model [solid line] was fit to each case without forcing the intercept to zero. All 276 points included [6 data points at t = 1, 5, 10, 20, 40, and 80 seconds from 46 ports] are presented on the scatter plots. The dashed line represents the 1:1 line which would be a perfect match between simulated and observed. The linear model fit [solid line0 and the coefficient of determination [R^2] provided on each plot give an indication of scatter and bias. The R^2 is a statistic that provides a quantitative measure of similarity between the simulated and measured drawdown values. A high R^2 value means that the simulated and

observed drawdown values are linearly correlated, even though the mean values could be different.

Illman et al. [2007, 2008, 2010] and *Liu et al*. [2007] previously found that the best way to test the results from hydraulic tomography was to use the hydraulic parameter distributions in predicting one or more independent drawdown events. Therefore, to further assess the reliability of the K and S_s tomograms, we simulated 8 additional cross-hole tests performed at ports 3, 5, 6, 12, 13, 14, 24, and 37 using the forward groundwater model MMOC3 [*Yeh et al*., 1993]. The results were then assessed by plotting individual scatterplots of simulated and observed drawdown [Figure 20.5b] and fitting a linear model as described earlier. The scatter plots in Figure 5b show that the K and S_s tomograms were able to accurately predict independent cross-hole tests. The largest error is seen for the independent cross-hole test at port 5. There are 6 points that are significantly below the 1:1 line were measure at port 4 [adjacent to port 5]. This suggest that the heterogeneity in this portion of the sandbox may not be accurately captured. Despite this, the remaining ports for the cross-hole test at port 5 are matched accurately suggesting that the rest of the sandbox is fairly accurately characterized. In addition, the other validation tests in Figure 20.5b show good agreement between observed and simulated suggesting that the heterogeneity pattern of the sandbox is accurately captured.

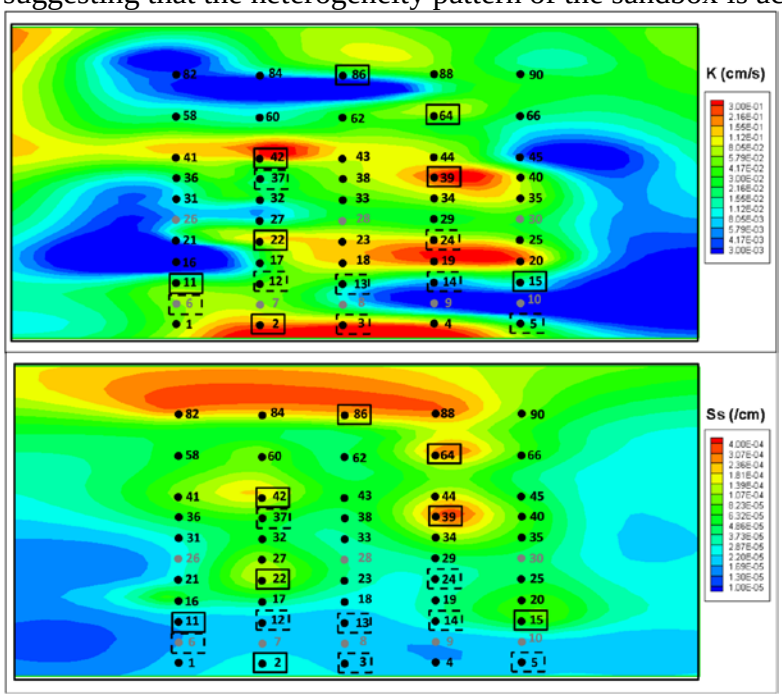
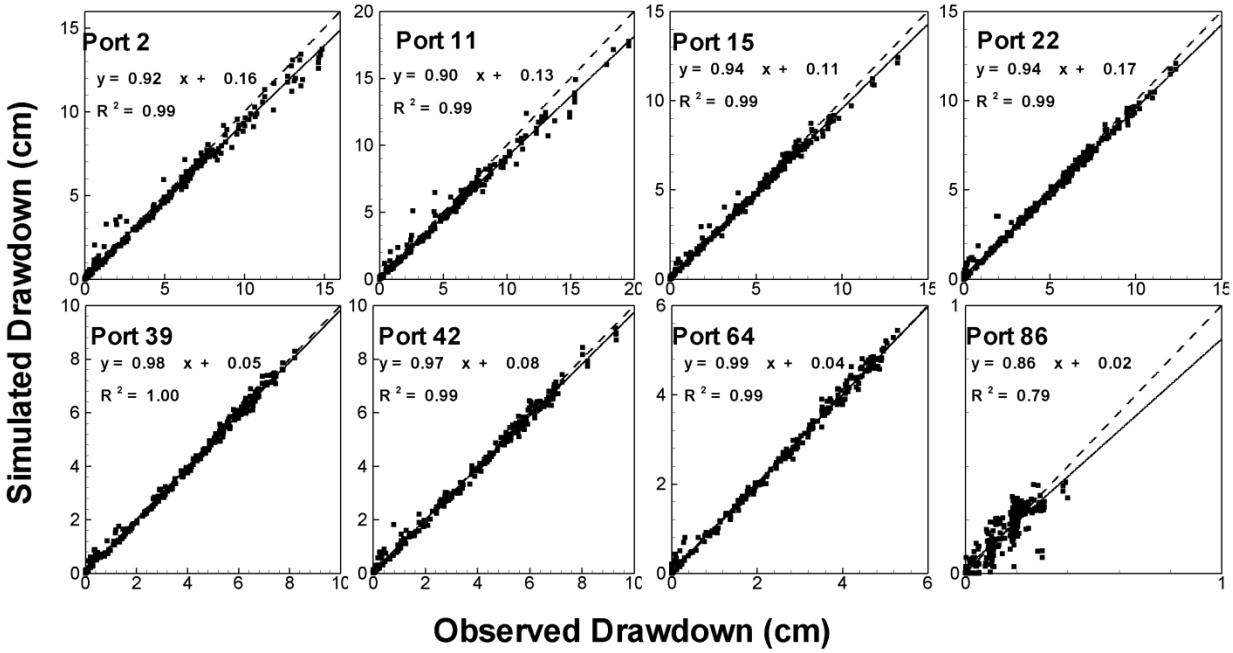


Figure 20.4: a] K and b] S_s tomograms computed using the transient hydraulic tomography algorithm of *Zhu and Yeh* [2005] SSLE with 8 cross-hole pumping tests. Solid squares are pumping locations used for transient hydraulic tomography dashed squares are pumping locations for the validation of the results from transient hydraulic tomography.

a]



b]

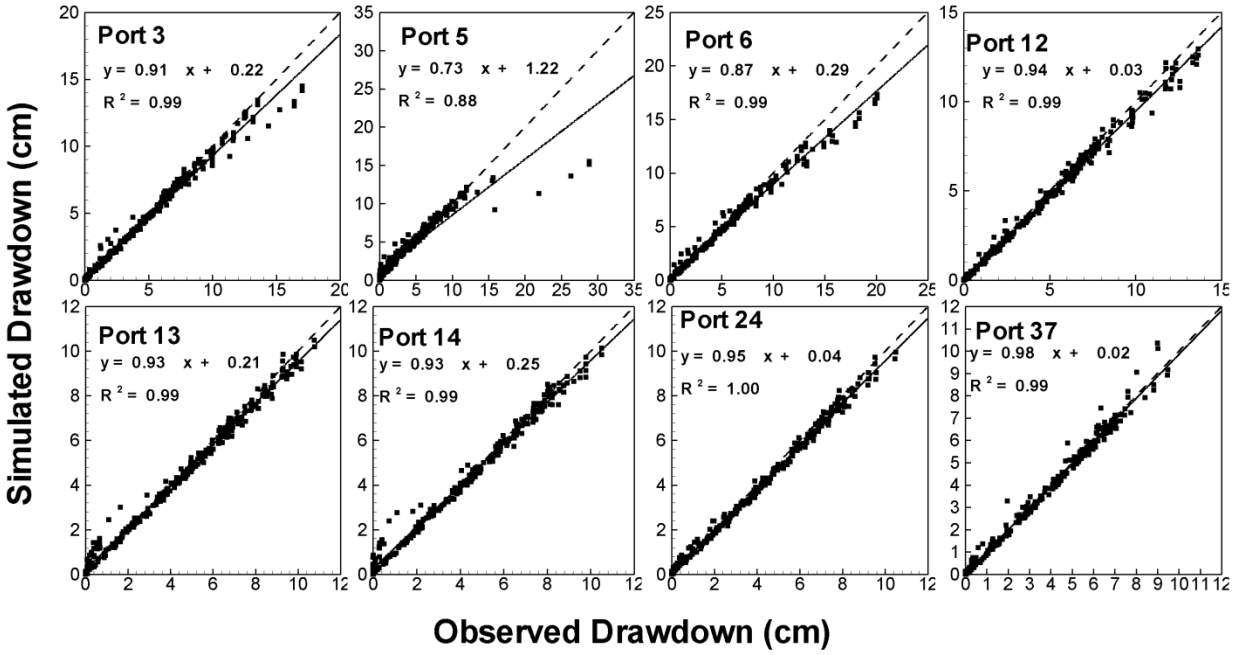


Figure 20.5: a] Simulated versus observed drawdowns from 8 cross-hole pumping tests used for calibration purposes and b] additional tests used for the validation of the K and S_s tomograms.

20.3.4 Estimation of θ_s, θ_r, α and n

Using the hanging water column method [*Stephens*, 1995] the primary drainage curve was measured for five of the materials packed in the sandbox [F35, F45, F110, Sil-co-sil 53, and Sil-co-sil 106]. These materials were selected because they are located in the upper portion of the sandbox and are likely to experience negative pressure heads [and as a result a possible release of water] when the water table is lowered during drainage or and unconfined pumping test.

Negative pressure head was increased incrementally until either residual water content was reached, or a value of -100 cm was achieved. A value of -100 cm was selected because it is unlikely that values greater than this will be seen during unconfined pumping or drainage tests in the sandbox. The *van Genuchten* [1980] model was fit to the primary drainage curves for each material by varying the α and n parameters using a generalized reduced gradient nonlinear routine implemented in Excel [except for Sil-co-Sil 53 and 106 which were fit manually because the solver failed to converge for these two cases]. In this analysis, the residual water content was estimated from the moisture characteristic curve and fixed during the regression. The model fit to the moisture characteristic curve, optimized parameters and the associated coefficient of determination [R^2] are presented on Figure 20.6. The parameter values obtained for each of these materials [Figure 20.6a-e] can be used to construct a heterogeneous case for the model domain, however, this level of detail is rarely available and a single 'effective' drainage curve would be used to describe the entire domain.

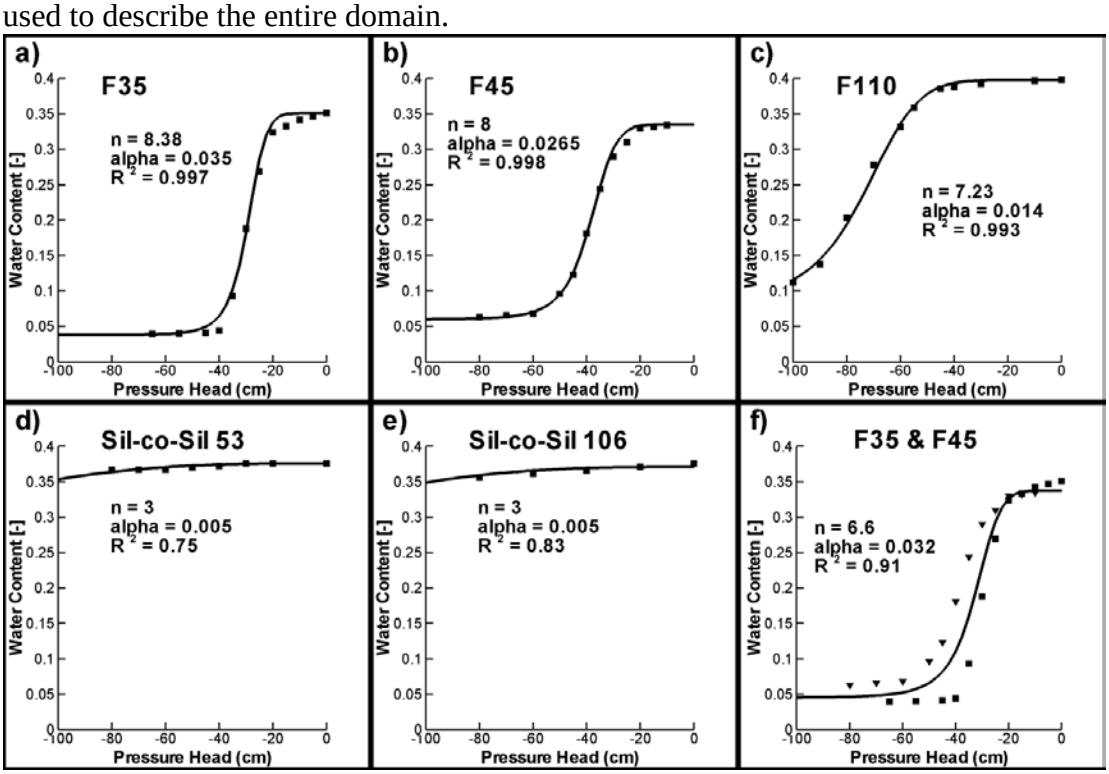


Figure 20.6: Moisture characteristic curves determined through the hanging column method of: a] F-35; b] F-45; c] F110; d] Sil-co-Sil 53; e] Sil-co-Sil 106; and f] F35 and F45 matched simultaneously. The squares represent F35 and gradient symbols represent F35.

To obtain an effective α and n representing the entire aquifer domain, we fit the *van Genuchten* [1980] model to the primary drainage curves for both F35 and F45 simultaneously [Figure 20.5f]. These two sands were selected because they comprise the majority of the material in the unsaturated zone and will therefore largely control the drainage response. The other materials present in the unsaturated zone [F110, Sil-co-Sil 53, and Sil-co-Sil 106] were not included in this simultaneous match as they are only present in minor portions of the unsaturated zone and are present as discontinuous lenses. Since they are present as discontinuous lenses, it is possible that when the coarse F35 and F45 material drains these lenses will become disconnected from the main aquifer and will not be able to effectively release their water. Thus, only the two dominant material types in the upper portion of the sandbox are selected for this effective parameter estimation.

20.4 UNCONFINED PUMPING TEST

After the sandbox was characterized using cross-hole tests while fully saturated the water table was allowed to fall under unconfined conditions. The objective of this test was to collect data from both the saturated and unsaturated zones during a pumping test so that we can compare the observations to variably saturated flow simulations constructed from our characterization efforts. Figure 20.7 is a schematic diagram showing the locations of pressure transducers, tensiometers, and water content sensors installed in the sandbox as well as the port pumped for this test.

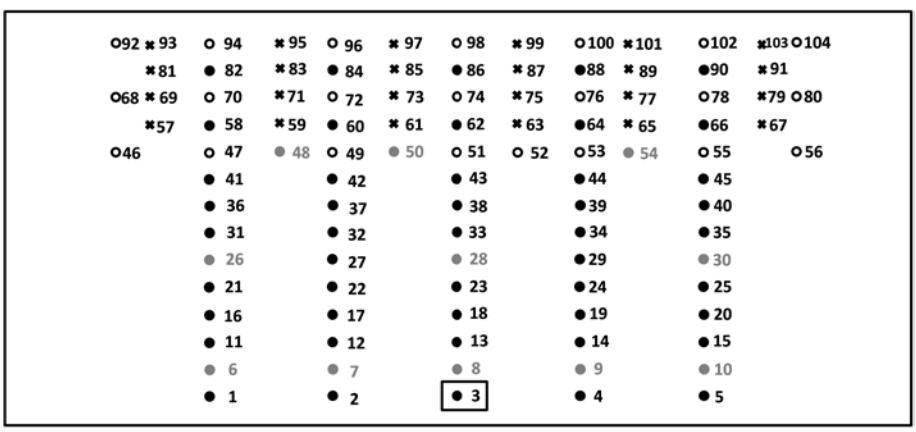


Figure 20.7: Schematic diagram of the sandbox showing an array of sensors [● = pressure transducers; ● = port without pressure transducers; ○ = tensiometers; and **x** = water content sensors] utilized to monitor the pumping test in a heterogeneous unconfined aquifer. The box indicates the port at which the unconfined pumping test was performed.

The pumping test started with the sandbox fully saturated and the water table at the same elevation as the top of the sand. This was confirmed visually using a manometer to get an independent reading of the static water level. Prior to the beginning of the test all sensors were monitored for several minutes to check for sensor stability and to record baseline data. This baseline data is then use as the t = 0 reference to which the data collected during the test can be compared. Port 3 [located at the bottom middle of the sandbox] was then pumped at a rate 60 ml/min for 7.5 hrs. The signal from all of the sensors [pressure transducers, tensiometers, and water content sensors] was recorded for the duration of the test.

Figure 20.8 shows the drawdown as measured at selected pressure transducers [Figure 20.8a] and tensiometers [Figure 20.8b]. The ports in these figures are organized such that their relative positions to each other reflect their relative positions within the sandbox. For example, Port 58 is located left of Port 62 and above Port 31. Figure 20.7 can be used as a reference to determine the absolute location of each port within the sandbox. The response measured at all ports is available in the supplementary section.

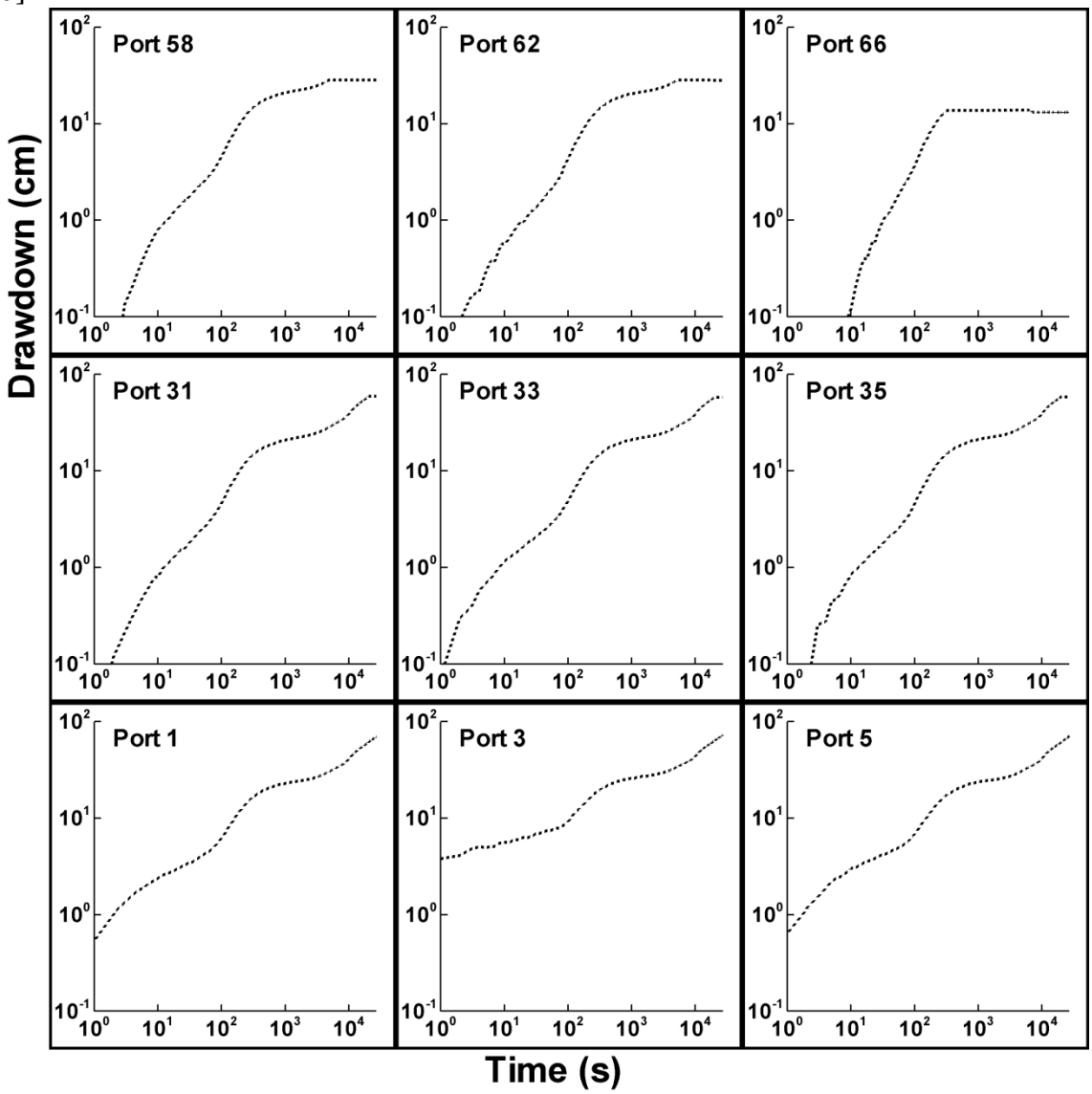
The pressure transducers [Figure 20.8a] are only able to measure the pressure response within the saturated zone. As the test progresses and the water table lowers some of the upper wells drain and the pressure transducer is no longer able to measure a change in pressure. The plateau in the signal for the upper ports [Ports 58, 62, and 66] in Figure 20.8a reflects this process. The shape of the drawdown curves are very different from what would be expected for a fully saturated system and reflect the unconfined nature of this test. The early portion of the curves display characteristic S-shaped behaviour of pumping tests performed in an unconfined aquifer. The S-shape is more pronounced for deep ports [further below the water table] than for shallow ports. This observation is consistent with that of *Nwankwor et al.,* [1992] and *Bevan et al.,* [2005]. At late time the observed drawdown increases again, deviating from the classical S-shaped unconfined drawdown curve. We attribute this to the lack of a constant head boundary which means at late time the no flow boundaries dominate the observed response and additional drawdown occurs.

The tensiometer responses measured during this test are shown in Figure 20.8b. Since the water table at the start of the test was coincident with the top of the sand in the sandbox the tensiometers were located below the water table at the start of the test. By calibrating the tensiometers for this range, these tensiometers record a continuous signal as the water table drops past them [i.e. they record both positive and negative pressure head]. Thus, unlike the pressure transducers installed in wells [Figure 20.8a], the tensiometers are able to record a signal for the entire duration of the test. Initial inspection of Figure 20.8b seems to suggest that the S-shaped curve is present in the tensiometer response. However, upon closer inspection the S-shape at early time is missing and it is the additional late time drawdown caused by the no flow boundaries that creates the impression of an S-shape at late time. Some of the bottom tensiometers [Ports 46 to 52] appear to have a slight S-shape at early time, however, it is not very prominent and sensor noise makes it difficult to make any concrete statements about this. Pressure transducers at similar elevations [Figure 20.8a: Ports 58 - 66] do not have a very pronounced S-shape at early time. Tensiometer data collected by *Nwankwor et al.,* [1992] during a pumping test at CFB Borden shows a very slight S-shaped response at early time only for ports near the pumping well. Thus, the absence of an S-shaped response in the tensiometer data here is not surprising. Ports 92, 98, and 104 [located in the top row of the sandbox] appear to plateau at late time. This likely indicates that the soil they are in is now disconnected from the main body of water, thus, the continued drop of the water table no longer influences the soil in this region.

Figure 20.9 shows spatial distributions of pressure head [Figure 20.9a] and water content [Figure 20.9b] at t = 400 s, 2,000s, 8,000s, and 27,000 s, after pumping began at port 3. The pressure and water content data were interpolated through kriging. We point out that the curvature of pressure head seen along the side boundaries are an artifact of the interpolation procedure. Figure 4a reveals that based on the pressure head distributions, an unsaturated zone begins to develop as early as t = 400 s. Subsequent pressure head distributions reveal that the water table drops with time and at t = 27,000 s, we notice the impact of the heterogeneity [i.e.,

the low *K* zone] on the contours of the pressure head. Based on this figure the water table is approximately 65 cm from the top of the aquifer.

a]



b]

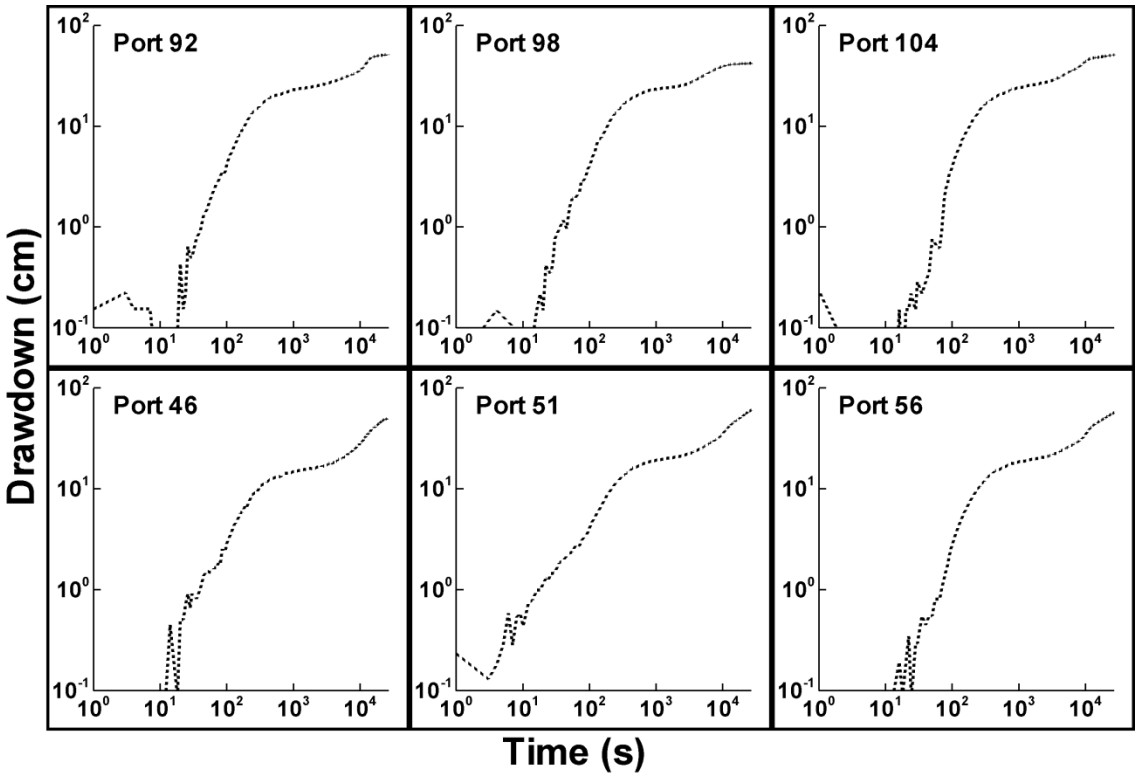
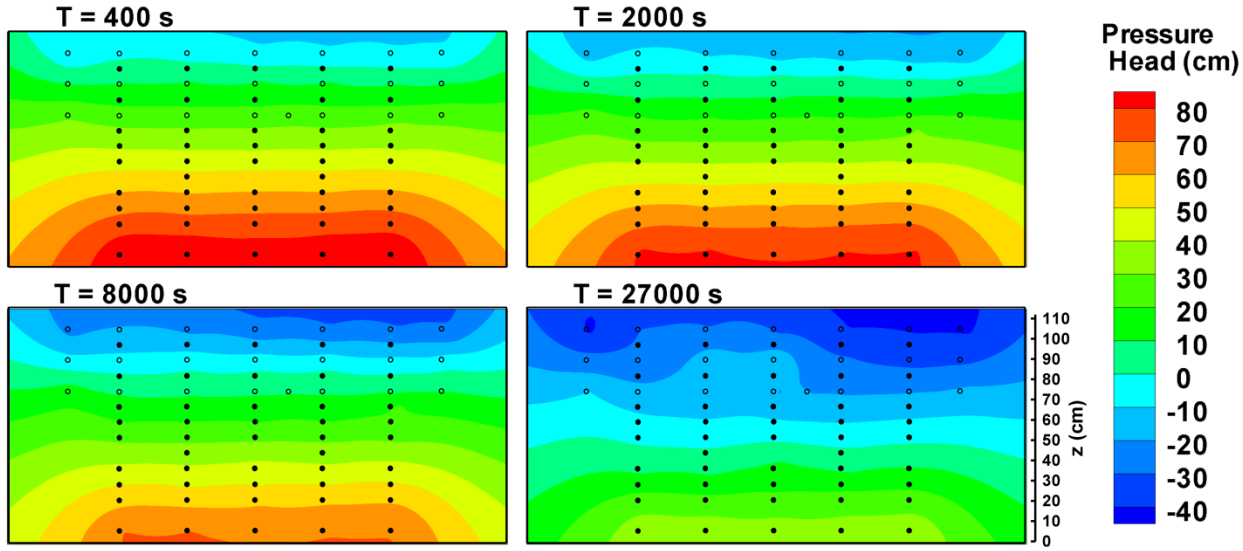


Figure 20.8: Observed drawdown during the unconfined pumping test at a] 9 pressure transducer ports; and, b] 6 tensiometers.

a]



b]

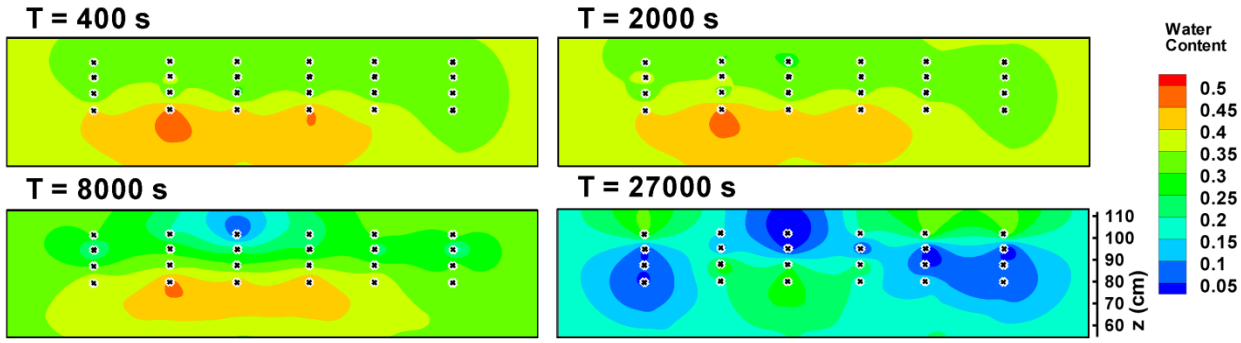


Figure 20.9: Spatial distribution of: a] pressure head and b] volumetric water content in the upper half of the sandbox with time during the unconfined aquifer pumping test at port 3. Symbols indicate the position of various sensors used to monitor the pumping test [● = pressure transducers; ○ = tensiometers; and **x** = water content sensors]].

Figure 20.9b shows the spatial distribution of the volumetric water content profiles in the upper half of the sandbox. Volumetric water content is generally uniform at the beginning of the test, but we see the impacts of pumping at t = 400 s in which the water content begins to decrease near the top of the aquifer. A slightly higher region of water content is visible near the top of the aquifer indicating the presence of a low *K*, high porosity layer. As the pumping test continues, the pore space drains throughout the aquifer except for the low *K* lens which retains some of its water due to capillarity. By t = 27,000 s, the upper portion of the aquifer is at or near residual water content [shown as dark blue] although some water is still retained in the low *K* lens.

20.5. VARIABLY SATURATED FLOW MODELING OF PUMPING TEST IN AN UNCONFINED AQUIFER

We next utilize the homogeneous and heterogeneous values of K, S_s, n, and α for predicting the response of the unsaturated and saturated zones during the unconfined pumping test that we described in section 4. A finite-element numerical model, MMOC3 developed by *Yeh et al*. [1993] is used for these calculations. This program solves the partial differential equation that describes flow in 3-D, variably saturated geologic media:

$$\nabla \cdot [K(\psi, \mathbf{x}) \nabla(\psi + z)] = \omega S_s(\mathbf{x}) \frac{\partial \psi}{\partial t} + \frac{\partial \theta}{\partial t} = \left(\omega S_s(\mathbf{x}) + C(\psi, \mathbf{x})\right) \frac{\partial \psi}{\partial t} \qquad [20.1]$$

where ∇ is the spatial gradient, $K(\psi, \mathbf{x})$ is the unsaturated hydraulic conductivity function, ψ is the pressure head where it is positive in the saturated zone, while negative in the unsaturated zone, $\mathbf{x}$ is spatial coordinates, z is the elevation, ω is the saturation index, $S_s(\psi, \mathbf{x})$ is specific storage, t is time, θ is volumetric moisture content, and $C(\psi, \mathbf{x})$ is the soil moisture capacity derived from the moisture/pressure constitutive relationship. The saturation index is equal to 1 when the porous medium is fully saturated and zero when it is unsaturated. The program employs the Newton-Raphson iteration scheme to solve the nonlinear finite element approximation of equation [1].

We use MMOC3 [*Yeh et al.*, 1993] to simulate flow to a well due to pumping in a 2-D unconfined aquifer. The dimension of the aquifer is 2.44 m × 0.094 m in the horizontal plane and 1.12 m in the vertical and is discretized into 3,645 finite elements and 7,544 nodes. A

nonuniform mesh is used for the discretization in which some elements are larger than others so that the elements line up properly with the ports.

A no-flux boundary is assigned to the top [no infiltration or evaporation] and bottom of the aquifer; no-flux boundaries are imposed on the other five sides of the aquifer. The initial pressure head distribution in the aquifer is set to be hydrostatic with the water table at z = 1.12 m, representing a static condition. Pumping takes place at Port 3, located 7.3 cm from the bottom center of the aquifer with a constant pumping rate 60 cm^3/min. Because pumping takes place at a point, we neglected borehole storage which we consider to be a minor effect.

The hydraulic conductivity-pressure head and moisture-pressure head constitutive relationship of the aquifer are described by the model by *van-Genuchten* [1980]:

[20.2]

where | | is absolute value, θ_s is the saturated moisture content, θ_r is the moisture content at residual saturation and α, *n*, and *m* are shape-fitting parameters with m=1-1/n. We further assume that the $K[h]$ follows *Mualem's* [1976] pore-size distribution model expressed as:

$$K_r(\psi) = \frac{\left\{1-\left(\alpha|\psi|\right)^{n-1}\left[1+\left(\alpha|\psi|\right)^{n}\right]^{-m}\right\}^2}{\left[1+\left(\alpha|\psi|\right)^{n}\right]^{m/2}} \qquad [20.3]$$

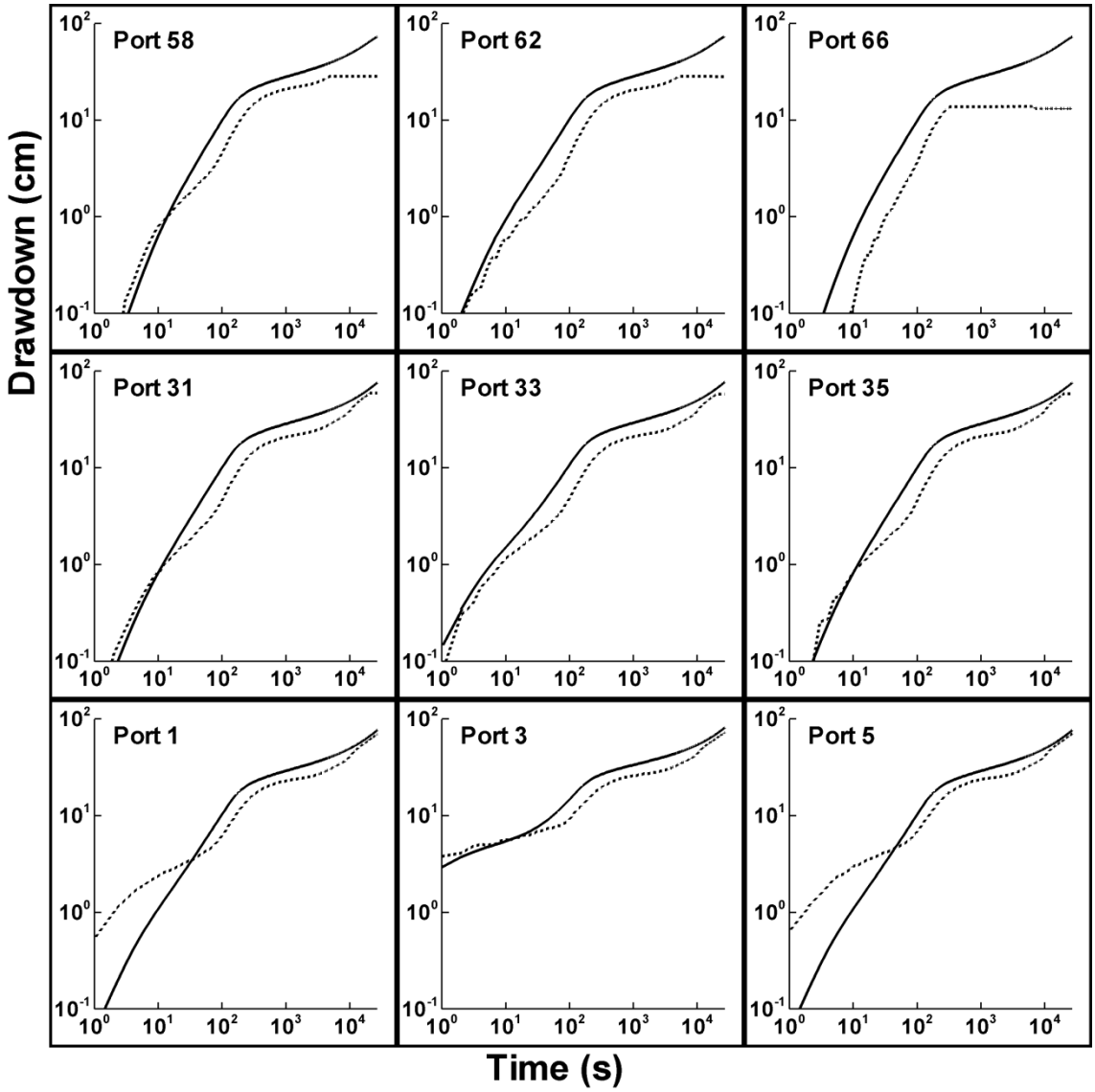
where K_s is the locally isotropic saturated hydraulic conductivity; and α, *n*, and *m* are assumed to be the same as those in equation [2]. The K_s and S_s were estimated earlier from the cross-hole pumping tests described in section 3.1.2, while the unsaturated zone parameters for the van Genuchten-Mualem model were obtained through the hanging water column [3.1.3].

20.5.1 Case 1: Homogeneous saturated and unsaturated zone parameters

In Case 1, we obtained effective K and S_s estimates for the cross-hole test performed at Port 22 by coupling MMOC3 with the parameter estimation PEST. The resulting PEST estimated K and S_s values were 1.85 x 10^{-2} cm/s and 3.94 x 10^{-5} cm^{-1} respectively. We also used the van-Genuchten model parameters [$\alpha = 0.032$ and n = 6.6] estimated from a simultaneous match of sands F35 and F45 [Figure 20.5f] to obtain a homogeneous estimate for the unsaturated zone parameters.

Forward simulation results for the pumping test under unconfined conditions are shown on Figure 20.10. In particular, Figure 10a shows the simulated [solid line] versus the observed [dashed line] drawdowns from selected pressure transducers in the saturated zone, while Figure 20.10b shows a similar comparison for pressure transducers installed in the unsaturated zone. Figures 20.10a and 20.10b reveal that the comparison between the simulated and observed drawdowns in both the saturated [Figure 20.10a] and unsaturated [Figure 20.10b] zones are poor. The simulated drawdown is generally greater than the observed, especially at intermediate and late time. Early time drawdown at pressure transducers are overestimated at some ports and underestimated at others. This is attributed to the homogeneous K and S_s field which, captures the heterogeneity in an average sense, and thus, is not expected to accurately simulate the response at all ports.

a]


Port 58
Port 62
Port 66
Port 31
Port 33
Port 35
Port 1
Port 3
Port 5
Drawdown (cm)
Time (s)

b]

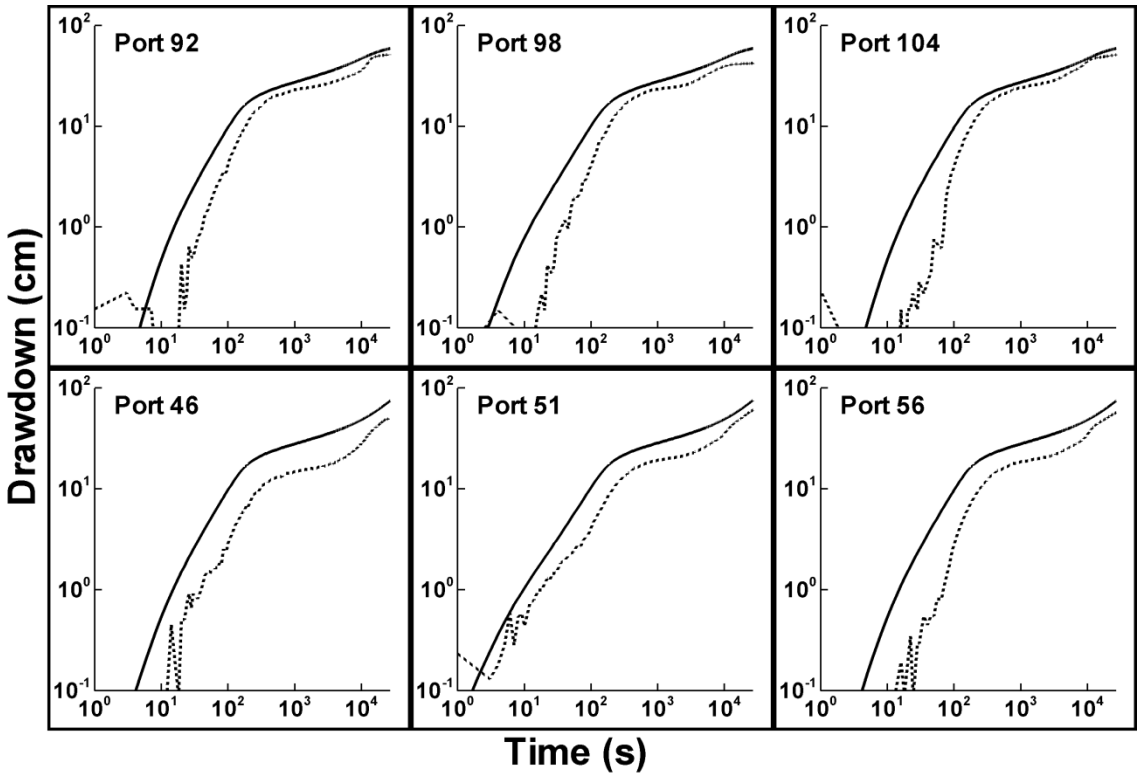
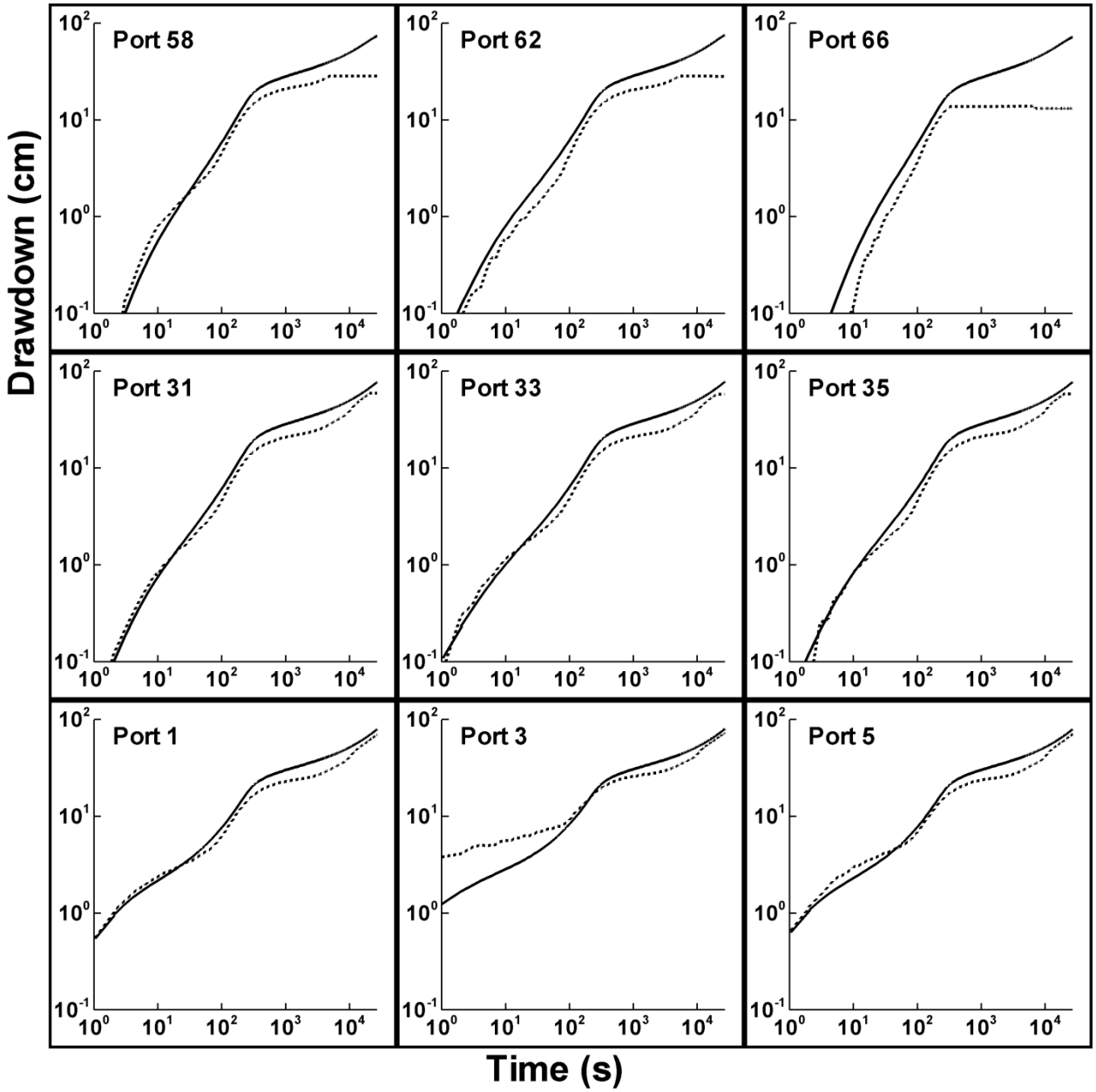


Figure 20.10: Simulated [solid line] and observed [dashed line] drawdown from selected: a] pressure transducers; and, b] tensiometers during the unconfined aquifer pumping test [Case 1].

20.5.2 Case 2: Heterogeneous saturated parameters and homogeneous unsaturated zone parameters

In Case 2, we obtained the K and S_s fields through transient hydraulic tomography of 8 separate cross-hole tests. The treatment of unsaturated zone parameters is identical to Case 1 [effective homogeneous values]. Forward simulation results shown on Figures 20.11a for the saturated zone reveals a marked improvement in the prediction of drawdowns at all times with the variably saturated flow model. On the other hand, we observe that the drawdown predictions in the unsaturated zone [Figure 20.11b] are slightly improved but still not very good showing that the simulated values generally exceeding the observed values especially at early time. This discrepancy could be attributed to a number of factors such as heterogeneity in unsaturated zone parameters and also due to a delay in response of the tensiometers.

a]


Port 58
Port 62
Port 66
Port 31
Port 33
Port 35
Port 1
Port 3
Port 5
Drawdown (cm)
Time (s)

b]

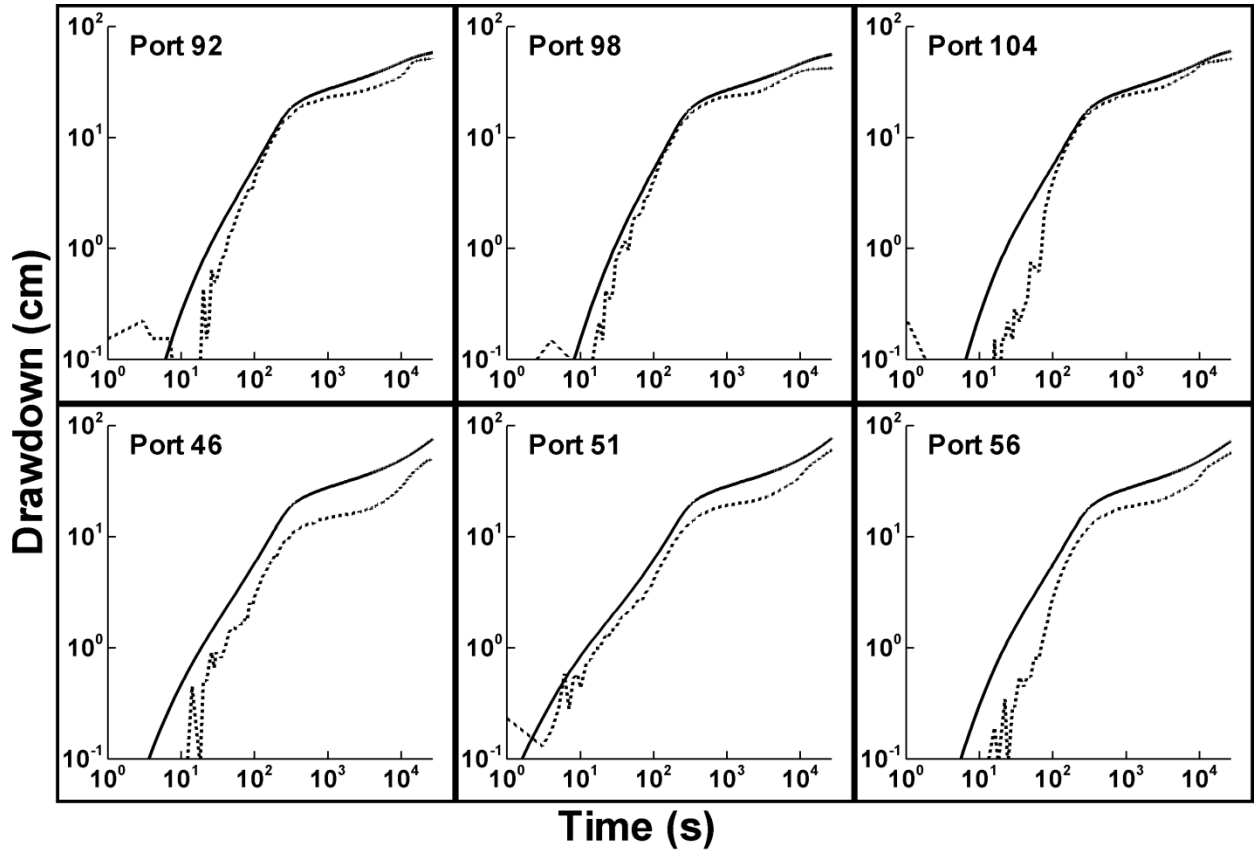
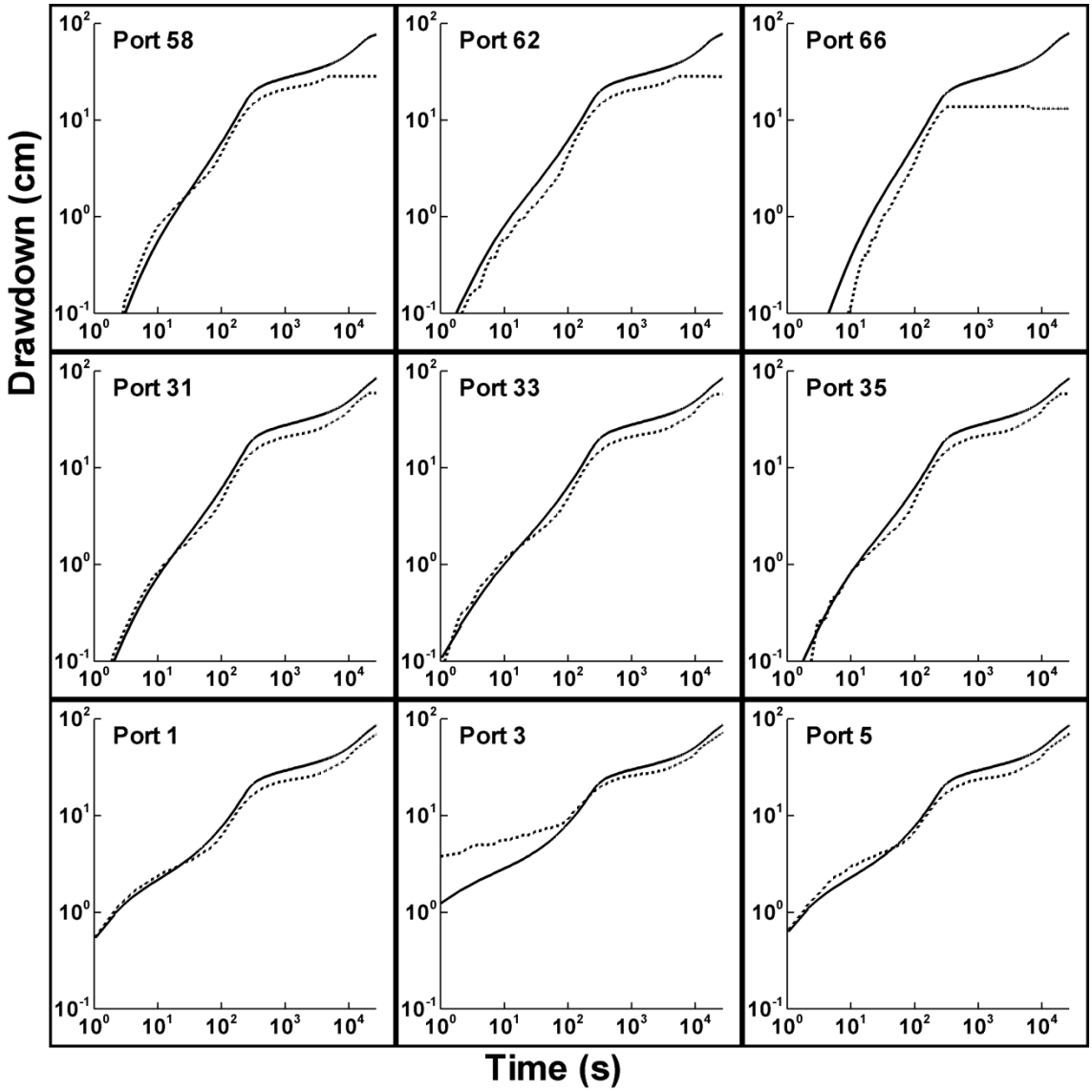


Figure 20.11: Simulated [solid line] and observed [dashed line] drawdown from selected: a] pressure transducers; and, b] tensiometers during the unconfined aquifer pumping test [Case 2].

20.5.3 Case 3: Heterogeneous saturated parameters and heterogeneous unsaturated zone parameters

To investigate the cause of discrepancy of drawdowns observed in the unsaturated zone, we also considered in Case 3, the variability of the unsaturated zone parameters by assigning homogeneous values for each layer. Figure 12a reveals that the quality of matches has virtually not changed in comparison to Figure 20.11a perhaps suggesting that heterogeneity in unsaturated zone parameters may not be that critical in responses in the saturated zone. This observation is in agreement with the findings from a numerical study conducted by *Mao et al*. [in review]. Predictions of drawdown in the unsaturated zone show only a slight improvement, suggesting that as long as the K and S_s are accurately know and the unsaturated parameters are represented accurately [in an effective sense], it is possible to predict unconfined pumping tests using a variably saturated flow model.

a]


Port 58
Port 62
Port 66
Port 31
Port 33
Port 35
Port 1
Port 3
Port 5
Drawdown (cm)
Time (s)

b]

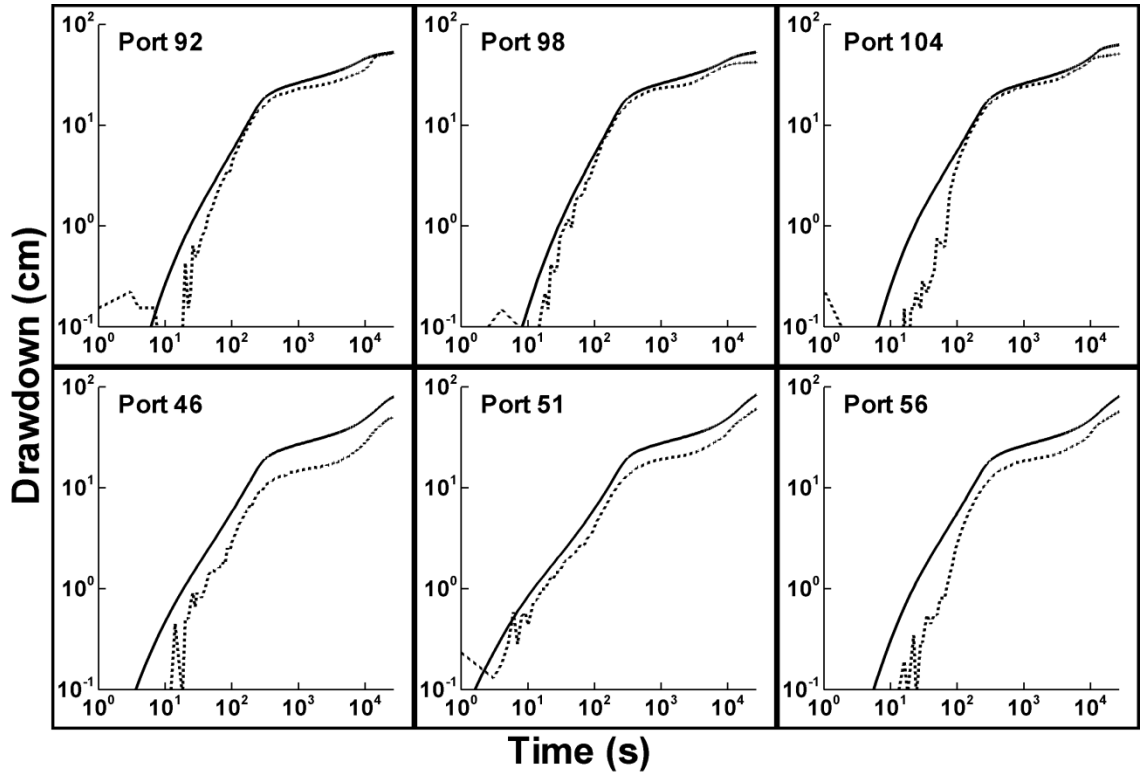


Figure 20.12: Simulated [solid line] and observed [dashed line] drawdown from selected: a] pressure transducers; and, b] tensiometers during the unconfined aquifer pumping test [Case 3].

20.6. DISCUSSION

Flow to wells due to the pumping of unconfined aquifers has been a topic of great interest for many decades. The majority of the studies described in the literature have focused on analytical solutions which treat the medium to be homogeneous. However, it is the rule rather than the exception that aquifers are heterogeneous.

We study the impact of the unsaturated zone and effects of heterogeneity on unconfined aquifer flow in a synthetic aquifer packed in an intermediate-scale laboratory sandbox. The synthetic aquifer was characterized initially with eight cross-hole pumping tests under fully saturated conditions and the hydraulic data are interpreted using the transient hydraulic tomography code developed by *Zhu and Yeh* [2005]. The resulting hydraulic conductivity [K] and specific storage [S_s] tomograms were validated using eight additional tests not used in the inverse modeling. Effective parameters of K and S_s were also estimated through the inverse modeling of individual cross-hole pumping tests. Laboratory hanging column experiments of sands in the unsaturated zone were also to determine the van Genuchten parameters. We then conducted another cross-hole pumping test in which a port near the bottom of the aquifer was pumped at a constant rate. During this test, the water table was allowed to freely move downwards in response to pumping. The saturated zone was monitored via pressure transducers and the unsaturated zone with tensiometers and water content sensors.

Forward models of various complexities in saturated and unsaturated parameters were then built using the variably saturated code, MMOC3 [*Yeh et al*., 1993] to examine the sensitivity of

the homogeneous and heterogeneous parameter estimates on the predictability of the unconfined aquifer test. Results show that an average or effective K and S_s determined through the averaging of equivalent K and S_s obtained from the inverse modeling of a pumping test cannot yield accurate predictions of drawdown responses in the saturated and unsaturated zones. Our predictions of drawdown responses in the saturated zone improved dramatically when the K and S_s distributions were utilized in the forward simulation of the unconfined aquifer pumping test. Because we monitored the pressure changes in the unsaturated zone during this experiment, we also attempted to evaluate the predictions of pressure responses in the unsaturated zone. Results show that slightly improved predictions of drawdowns in the unsaturated zone could be achieved if heterogeneous zones were assigned to the model. However, our results showed that predictions of drawdowns in the saturated zone were relatively insensitive to whether we conceptualized the unsaturated zone to be homogeneous or heterogeneous which confirms a conclusion reached by *Mao et al*. [in review] using numerical simulations.

Based on this study it is possible to accurately predict the response of a heterogeneous unconfined aquifer to pumping as long as the saturated parameters [K and S_s] are accurately characterized and an accurate effective value of the unsaturated parameters is known. This study was unique in that the heterogeneity pattern was exactly known and it was possible to select effective unsaturated parameters by assessing would material would dominate the drainage response. In the field however, this is not as straight forward and the selection of unsaturated parameters could pose a challenge. As such, interpreting pumping tests in heterogeneous aquifers will benefit from inverse modelling of both the unsaturated and saturated zones. This will require the use of a variably saturated model coupled with an inverse algorithm such as the Sequential Successive Linear Estimator [e.g., *Hughson and Yeh*, 2000] for the proper interpretation of flow to wells in a heterogeneous unconfined aquifer.

21. FINDINGS AND CONCLUSIONS

21.1 FINDINGS AND CONCLUSIONS FROM TASK 1: ALGORITHM DEVELOPMENT

21.1.1 Characterization of aquifer heterogeneity using transient hydraulic tomography methodology

- The synthetic cases show that transient hydraulic tomography is a promising and viable tool for detecting detailed spatial variations of hydraulic parameters with a limited number of wells. Our SSLE can provide unbiased estimates of multiple parameters simultaneously, and reveal their detailed spatial distributions. In addition, our SSLE permits sequential inclusion of head data from different pumping tests, such that the size of the covariance matrix is small and can be solved with relative ease. By using a loop-iteration scheme, our new SSLE improves estimates throughout the loops and maximizes the usefulness of head information.

- The cross correlation analysis shows that the correlation between head and specific storage is high at early time, diminishes rapidly with time, and is confined to the vicinity of the head observation location. On the contrary, the correlation between head and hydraulic conductivity increases and the area with high correlations broadens with time. To simultaneously estimate hydraulic conductivity and specific storage parameters, head data at both early and late times thus should be used.

- The transient heads are highly temporally correlated, especially at later times. Such a temporal correlation structure allows our SSLE to use only a few selected heads at some time steps, instead of all available heads at all time steps, to reduce computational cost, while keeping the usefulness of the head information.

- Our SSLE approach involves backward calculation of adjoint equations during the sensitivity analysis for transient flow. For the same number of observation locations, a transient pumping test generates much more head information than a steady state pumping test. Even when head data are used for only a few selected time steps, instead of all time steps, the computational burden of transient hydraulic conductivity is significantly greater than steady state hydraulic tomography. More computationally efficient methodologies must be developed to improve the analysis of transient hydraulic tomographic surveys. Finally, a 2-D version of SSLE for the transient hydraulic tomography is available at http://tian.hwr.arizona.edu/yeh/download.

21.1.2 Analysis of hydraulic tomography using temporal moments drawdown-recovery data

- The temporal moment approach significantly reduces computational cost for interpreting transient hydraulic tomography. The cost reduction is attributed to the fact that the

governing equations for the temporal moments are Poisson's equations. As a consequence, the forward modeling required for improving new estimates of transmissivity and storage coefficients does not have to solve the parabolic equation that governs groundwater flow. The parabolic equation in general has to be solved by a time marching scheme, implying that a system of equations must be evaluated at each time step. Avoiding solving the system of equations for each time step thus reduces computational efforts. Furthermore, the adjoint equations for evaluating sensitivity matrices of temporal moments are also Poisson's equations. In our SSLE and some other inverse models, the adjoint equations are solved for each measurement during each iteration of the estimation process. Again, without evaluating time-dependent adjoint equations, computational burdens during evaluation of sensitivity are reduced. This reduction is particularly significant when the number of temporal and spatial observations is large.

- The temporal moment approach for interpreting hydraulic tomography is unequivocally more efficient than the approach using transient head data directly. The differences in results of the moment and head approaches are small if a dense network is used. Furthermore, for the situations where only hydraulic conductivity is of interest, the zeroth temporal moment of transient well hydrograph can be used to estimate transmissivity without involving the estimation of the storage coefficient. This unique characteristic of the temporal moment approach makes the approach highly attractive for practical applications.

21.1.3 A new estimator and a guide for hydraulic tomography analysis

- Results of this study show that in spite of noise in hydrographs from a HT survey, the least squares approach can satisfactorily estimate effective hydraulic properties of the synthetic aquifer with hierarchical heterogeneity because the inverse problem is well posed.

- Accurate estimation of spatial variances of K and S_s from HT data is difficult due to the nonstationary nature of the flow field, which demands a large amount of head data to obtain representative sample head variances.

- For ill-posed problems, HT in conjunction with our SimSLE yield satisfactory estimates of the hierarchical K and S_s fields of the synthetic aquifer. If hydrographs are corrupted with noise, the SNR is a useful measure of reliability of a corrupted hydrograph, and wavelet denoising the hydrographs is a viable means to improve the estimate. In addition, the use of stabilization of the conditional L2 norm of head as a convergence criterion in SimSLE avoids overexploitation of noisy data.

- HT surveys delineate detailed hydraulic heterogeneity in aquifers, which can be used to predict different flow scenarios. That is, the estimate hydraulic properties do not suffer from the phenomenological nature associated with the domain-scale effective properties.

- Finally, simultaneous inclusion of hydrographs from all pumping tests in the analysis

offers some advantages over the previous sequential approach but it suffers from the requirement of huge computational resources.

21.1.4 Hydraulic/partitioning tracer tomography for characterization of dense nonaqueous phase liquid source zones

- We have developed a hydraulic/tracer tomography technique for characterizing NAPL source zone. The hydraulic/tracer tomography generates a large number of head and tracer data, which is interpreted by a joint SSLE inverse approach. The synthetic cases show that our technique can reveal major spatial distributions of NAPL using only four wells. Although field situations are more complicated than the synthetic case, the primary results from the synthetic case demonstrate that the hydraulic/tracer tomography combining with SSLE method potentially is a viable and cost-effective tool for subsurface heterogeneity and contaminant characterization.

- Under saturated condition, the variation of θ_w [the porosity in this case] is usually small. For example, the porosity of sandy material ranges from 0.25 to 0.50 [Freeze and Cherry, 1979]. Due to the naturally small variability of porosity, the variation of water saturation in a NAPL source zone has limit impacts on NAPL estimates, especially when tracers with high partitioning coefficient between water and NAPL are used. In our synthetic case, the porosity was intentionally created with higher variability than most natural porous media. As a result, the water saturation in field problems will have even less influences on NAPL estimations.

- Heterogeneity of hydraulic conductivity, on the other hand, has great impacts on NAPL estimation. Without detailed characterization of hydraulic conductivity, it would be extremely difficult to provide a reliable characterization of NAPL source zone. Hydraulic tomography is a viable tool for characterizing hydraulic conductivity distribution and can facilitate better NAPL estimation.

- Our synthetic cases show that adding tracer data my not improve *K* estimation. The cases also show using head data alone generates better results than using tracer data alone. Similar findings were also reported by Li and Yeh [1999] and Cirpka and Kitanidis [2001]. However, Datta-Gupta et al. [1997] claimed that tracer data are better than transient pressure in estimating large-scale permeability variations. Huang et al. [2004] concluded that tracer breakthroughs are important indicators of the flow paths and flow barriers and therefore should be used in heterogeneity characterization whenever they are available. These issues need additional investigation and are beyond the scope of this study.

- In our four inverse cases [case 1 to case 4], we assume only one NAPL measurement is available for inversion. Our SSLE method can include all direct NAPL measurements whenever they are available. These direct measurements will definitely increase the accuracy on NAPL estimations [Liu and Yeh, 2004].

- Directly interpreting concentration data using SSLE adds significant computational burden than interpreting head data. One head observation involves solving one adjoint equation for the evaluation of sensitivity while one concentration observation needs two adjoint equations. In this study, we used 8 processors [Pentium 2.8GHz each] in a PC cluster. The running time for one case is about 4000 minutes. As a result, more computationally efficient methodologies must be developed to expedite the analysis of tracer tomographic surveys.

21.1.5 Analysis of tracer tomography temporal moments of tracer breakthrough curves

- Using hydraulic/tracer tomography for characterizing DNAPL source zone is new but it faces a few challenges. One of them is the computational cost, which is mainly arises from the usage of a large number of discrete concentration measurements during the estimation procedure, in which the solution to the hyperbolic advection-dispersion equation demands fine spatial and temporal discretizations. Even with computationally efficient adjoint state method for evaluating sensitivities, the computational cost remains to be very high since a tracer tomographic survey is designed to collect a large number of concentration measurements during tracer releases at different locations. To reduce the computational cost associated with the new technology, the temporal moments of tracer BTCs are employed for parameter estimation in this study instead of point concentration measurements at some discrete times. Temporal moments are calculated using moment-generating equations. Solving these moment-generating equations is computationally more efficient than solving transient advection dispersion equation because the temporal moments are time independent. Furthermore, the adjoint state equations have similar form as the corresponding forward equation, thus there is huge reduction in computational cost during the sensitivity calculation. The synthetic case study indicates that inverting the first moment on a single PC only needs ¼ of the time used in inverting direct concentrations on an 8-node PC cluster. Thus, using temporal moments makes hydraulic/tracer tomography approach feasible for applications with large-scale domains. A similar computational cost comparison between direct head measurements and head temporal moments was given by Zhu and Yeh [2006].

- Solving advection dominated moment-generating equations is prone to numerical oscillations. The oscillations occur during evaluation of both forward and adjoint state equations, which can lead to numerical instability during inversion. As the SUPG method is commonly used to reduce the oscillation, choosing an appropriate stabilization factor is critical. Our numerical simulations indicate that the one we used in conjunction with a solver from Diffpack works very well. With a Peclet number of approximately 10 in the synthetic cases, we do not experience convergence problems using the first temporal moment alone. However, small oscillations still remain with SUPG method. Whether the oscillations could cause numerical instability for other cases has not been tested. Cirpka and Kitanidis [2001] used a stabilization factor different than our choice for solving moment generating equations, but they reported that inverse modeling using arrival time alone could lead to convergence problems. More elegant solving techniques that further reduce oscillation, such as adaptive mesh refinement, may be needed for more general applications.

- Our results show that the estimates of θ_n from the first moments are not as good as those using discrete concentration data of the BTCs. This can be attributed to the fact that the first temporal moment of a BTC is an average value of the arrival times of concentration values of different parts of a BTC. The arrival time of each concentration value of a BTC reflects the travel path and heterogeneity that a volume of a tracer encountered between the injection and sampling points. Thus, the first moment only contains a portion of information in the BTC. Although higher moments potentially contain more information, the computational cost for using higher moments however will be significantly greater than the first moment since all moments have to be solved recursively. Nevertheless, estimates from the first moment reveal major high and low DNAPL content zones in the synthetic case, suggesting that the first temporal moments are useful for estimation of DNAPL source zones.

- The cross correlation analysis indicates that a tracer first moment measurement at one location contains information about parameters along a narrow stripe between the injection location and the tracer measurement location while a steady head at the same location is correlated with hydraulic conductivity over a large area. Transient head measurements at later time of a pumping test exhibit similar cross correlation with hydraulic conductivity as steady state heads. One tracer test with a limited number of sampling locations therefore can only provide limited area coverage, making tracer moment measurements less efficient than hydraulic head measurements in estimating hydraulic conductivity. Furthermore, with all the simplifications and assumptions in this work for describing tracer movements through a DNAPL source zone, a tracer first moment measurement is still correlated not only to hydraulic conductivity but also to water content [conservative and partitioning tracers] and DNAPL content [partitioning tracer]. A tracer measurement can also be affected by parameters such as diffusivity, partitioning coefficient and others, potentially further reducing the correlation between tracer first moment and hydraulic conductivity. Consequently, to yield a reasonable estimate of hydraulic conductivity using tracer measurements, one should consider conducting multiple tracer tests in a tomographic fashion, where each test provides different areal coverage.

- Notice that cross correlation analysis is different from sensitivity analysis. Equations [8.27] and [8.28] show that calculating cross correlation between a state variable [i.e. tracer first moment] and a parameter [i.e. hydraulic conductivity] is the cross-covariance normalized by the variance of the state variable and the variance of the parameter. The variance of the state variable in turn is affected by both the sensitivities of the state variables to all parameters and the covariances [or uncertainties] of all parameters [i.e., hydraulic conductivity, water content, and DNAPL content in this study]. The cross-correlation as demonstrated in this study is therefore an appropriate criterion for assessing the effectiveness of a measurement as well as different types of measurements on parameter estimation.

- Finally, other issues related to the applicability of the new approach to real-world problems have yet to be addressed. These include: 1] the assumption that tracer only

reacts with DNAPLs; 2] the assumption that the partitioning process is in equilibrium; 3] the assumption that DNAPLs are not mobilized by tracer tests; 4] memory requirement for large-scale problems. Nevertheless, our study demonstrates that using tracer temporal moments can dramatically reduce the computational cost in interpreting tracer tomography and can yield reasonable estimates of DNAPL content, making a step forward for applying hydraulic/tracer tomography to real-world problems.

21.2 FINDINGS AND CONCLUSIONS FROM TASK 2: LABORATORY EXPERIMENTAL AND NUMERICAL MODELING STUDIES

21.2.1 Laboratory sandbox validation of steady-state hydraulic tomography

- Hydraulic tomography is a technology that facilitates the imaging of subsurface heterogeneity in hydraulic parameters. To date, a comprehensive validation of the hydraulic conductivity [*K*] tomogram has not been done either at the laboratory or field scales. Previous laboratory investigations assumed that packing was perfect and in general, small scale data were not available for a direct comparison. This study provides the first such examination using small-scale *K* data obtained from cores and single-hole tests as well as large-scale *K* estimates obtained from flow-through experiments in a sandbox with deterministic heterogeneity in hydraulic parameters.

- Prior to inverse modeling of data, we conducted a detailed diagnostic study to investigate the magnitude and cause of errors and biases in data through scatter plots, contour plots and data animations. Such diagnostic tests of data used in forward and inverse models are rarely discussed in the literature, but we found that it should be an integral component of all phases of numerical forward and inverse modeling of cross-hole tests as the use of data corrupted by noise can have a profound effect on both forward and inverse model results.

- Validation of the *K* tomograms involved a multi-method and multiscale approach which included: 1] visual comparisons of *K* tomograms to the true sand distributions as well as to the reference *K* tomogram; 2] testing the ability of the *K* tomogram to predict the hydraulic head distribution of an independent cross-hole test not used in the computation of the *K* tomogram; 3] comparison of the conditional mean and variance of local *K* from the *K* tomograms to the sample mean and variance of results from other measurements; 4] comparison of local *K* values in *K* tomograms to those from the reference tomogram; and 5] comparison of local *K* values in *K* tomograms to those obtained from cores and single-hole tests. The multi-method and multiscale validation approach proposed herein further illustrates the robustness of hydraulic tomography in subsurface heterogeneity delineation.

- Previously, the effects of errors and biases on the *K* tomograms have not been investigated in detail. The steady state inversion of cross-hole tests in a synthetic laboratory aquifer showed that the approach is sensitive to errors and biases. Data diagnostics combined with forward modeling provided valuable insight into identifying the cause of such errors and biases. Specifically, the errors identified include drift in

pressure transducer readings, a skin effect influencing hydraulic head at the pumped well, and inaccurate treatment of boundary conditions, among others. We found that accurate modeling of boundary conditions is essential in conducting steady-state hydraulic tomography and obtaining accurate *K* tomograms. In real field situations, the boundary conditions of the field site need to be studied carefully through forward modeling and better site characterization. Further research is clearly needed to improve hydraulic tomography technology both in the laboratory and under field conditions.

21.2.2 Practical issues in imaging hydraulic conductivity through hydraulic tomography

- We first investigated the mean values of the synthetic and real hydraulic tests conducted on the computer and in the laboratory sandbox aquifer. Results show that the mean *K* in general increases with the measurement scale for the synthetic simulations when the mean *K* from cores is compared to the mean *K* from both slug and single-hole tests. There is also a scale effect in *K* for the real data, when mean *K* from cores is compared to mean *K* values from cross-hole and flow-through tests. The same holds true for the comparison between mean *K* from both slug and single-hole tests to the larger scale tests. The variance, however, decreases with the measurement scale as the larger scale measurements average the porous medium.

- We investigated the effect of pumping rate which affects the signal-to-noise ratio of observed test data on the quality of the computed *K* tomogram. We found that the signal-to-noise ratio is not as important for the synthetic case, but very important for the real case. That is, the computed *K* tomogram was clearer [and closer to the reference *K* tomogram] when the pumping test was conducted at a higher rate, yielding higher signal-to-noise ratio data, when analyzing real data. This is because different cross-hole pumping test data sets contains different levels of noise and its magnitude depends on a number of factors including: 1] quality of pressure transducers [i.e., accuracy and precision of pressure transducers]; 2] calibration of pressure transducers and its repetition prior to pumping tests; 3] proximity of the monitoring ports to the pumping location; 4] the pumping rate which affects the signal-to-noise ratio of observed test data; 5] presence or absence of constant head or no-flow boundaries; 6] external stresses; and 7] aquifer diffusivities and their heterogeneity. Therefore, we recommend that in practice, more attention should be placed in designing cross-hole pumping tests to maximize the information content of the data sets and to minimize noise through signal processing and/or noise reduction techniques.

- We also examined the order of pumping tests included in the inversion algorithm and found that the order has a minimal impact on the inversion results for synthetic data, but greatly affects the *K* tomogram using real data. This is because each cross-hole pumping test data set contains different levels of noise. Our experiments in the laboratory were conducted as uniformly as possible, but there are noises that we could not control from one test to the next. Our findings include the importance of examining test data carefully and utilizing the cleanest data [with the highest signal-to-noise ratio] first and progressively including noisier data [with lower signal-to-noise ratio] into the SSLE algorithm. This is because SSLE in its current form is more sensitive to noise during the

beginning stages of K tomogram computation. This sensitivity is due to the uniform convergence criteria used for sequentially inverting all pumping test data. One improvement that could be made to the SSLE algorithm is to include an option for variable convergence criteria to account for different pumping tests with different noise levels. Furthermore, when hydraulic tomography is conducted using field data with SSLE in its current form, we recommend that the cleanest data be included into the SSLE algorithm first and progressively including noisier data.

- Conditioning has been thought to help constrain inverse modeling results. We found that this is certainly the case for synthetic test data when utilizing the SSLE algorithm. However, this study shows that the conditioning data itself can be subject to errors and so conditioning may not necessarily help in obtaining an improved solution. In particular, conditioning with data that are corrupted with noise can actually worsen the quality of the K tomograms. Therefore, we recommend that more attention should be paid to collection of better conditioning data and minimization of its errors.

- We also examined the type of data used to conduct the conditioning of K tomograms. In this paper, we utilized synthetic and real core, slug, and single-hole K estimates. Our study showed that core data [assuming that they can be accurately obtained] improved both synthetic and real K tomograms. Conditioning of the K tomogram with slug and single-hole K estimates smoothed the K tomograms for this sandbox due to the larger support volume associated with these tests in comparison to the numerical grid used to compute the K tomograms. Therefore, we recommend scrutinizing the data type used to do the conditioning as not all data are created equally. Furthermore, we recommend that one should consider the support volume of data used to condition the K tomogram as it can have an impact on the resolution of the K tomogram.

21.2.3 Laboratory sandbox validation of transient hydraulic tomography

- The main objective of the study is to validate the recently developed hydraulic tomography concept and an analysis algorithm [i.e., SSLE] in a heterogeneous sand box. In order to accomplish this goal, we investigated the ability of various hydraulic tests and analyses to characterize the heterogeneous sand box. These tests and analyses include determination of the K from core samples, slug tests, single-hole analyses, cross-hole analyses, and unidirectional flow-through test.

- Based on the results of this investigation, we draw the following major conclusions: 1] with the number of samples and test locations used in our experiment, the estimated K values from core samples and slug tests can delineate the heterogeneity pattern of the sand box, but the hydraulic tomography provides considerably more details; 2] the average of the 95 K estimates from the cross-hole tests and analyses yield an averaged K value that is close to the effective K determined from the uniform flow experiment or the mean value of the estimates from hydraulic tomography; 3] estimates of S_s from the single-hole analysis and those from hydraulic tomography exhibit physically plausible S_s distributions; 4] while some of the tests and analyses yield similar K and S_s statistics and patterns, it is difficult to validate each other due to scale disparity, different flow

conditions, etc. As a consequently, an appropriate validation approach is to test the predictability of the estimates under different flow scenarios.

- This is the approach we chose to validate the hydraulic tomography concept and the SSLE algorithm. Based on this approach, we demonstrate that using the estimated K and S_s fields from the hydraulic tomography, a classic governing flow equation predicts drawdown distributions caused by an independent pumping event in close agreement with the observed distributions at three different times. We thereby conclude that the hydraulic tomography concept and the analysis algorithm [SSLE] is a viable tool for characterizing aquifers at high resolutions although field tests are needed to further substantiate this claim for a real-world problem**.**

21.2.4 Laboratory sandbox validation of hydraulic tomography that uses the temporal moments of drawdown-recovery data

- We find that it is possible to analyze multiple cross-hole pumping test data with hydraulic tomography which uses the moments of the drawdown-recovery data sets [HT-m]. The HT-m approach is considerably faster than the THT approach with the latter approach utilizing transient drawdown records. This is because the HT-m approach simplifies the governing equation from a single parabolic equation to two Poisson's equations for the 0th moment and characteristic time defined as the ratio between the 1st and 0th moments.

- Based on the synthetic simulations and analysis of the real data and its validation we conclude that the HT-m approach is a factor of 20 faster than the THT approach for the case at hand, but there is a definite tradeoff in the quality of results. The interpretation of synthetic cross-hole test data obtained on the computer using the HT-m and THT algorithms revealed that the quality of the K tomograms is comparable for both HT-m and THT algorithms. However, this is not the case for the S_s tomograms. We find that the S_s tomogram computed using the HT-m algorithm deteriorated in comparison to the results obtained from the THT algorithm for the synthetic case which we examined here.

- The interpretation of the real cross-hole pumping tests conducted in a laboratory sandbox aquifer using the HT-m approach showed that the K tomogram is smoother than the one computed using the synthetic data set in this paper. The computed tomogram also appeared smoother in comparison to the K tomograms computed previously by *Illman et al*. [2007] using the SSHT algorithm and also by *Liu et al*. [2007] using the THT approach.

- The interpretation of the real cross-hole test data using the HT-m approach showed that the S_s tomogram appears to be less reliable. This could be due to three possibilities. One possibility is that the HT-m approach relies on the computation of the temporal moments of the drawdown-recovery data, which involves integration. Data integration causes smoothing and perhaps loss of information on aquifer heterogeneity. This could cause significant loss of information on S_s which is contained in the early-time data. Previous research by *Zhu and Yeh* [2006] and others suggest that this may be the case. Another possible explanation is that the use of the temporal moments, which requires the entire drawdown-recovery curve may cause loss of sensitivity to the inversion S_s. We also

found that the characteristic time defined as the ratio between the 1st and the 0th moments was considerably noisier than the 0th moments, which can perhaps contribute to the difficulties in estimating S_s using the HT-m approach.

21.2.5 Comparison of aquifer characterization approaches through steady-state groundwater model validation: A controlled laboratory sandbox study

- The traditional characterization by means of permeameter analysis of cores, single-hole and cross-hole tests of this laboratory heterogeneous aquifer yielded a range of geometric mean of hydraulic conductivity [K_G] and variance of ln K [$\sigma^2_{\ln K}$]. The K_G ranged from 0.059 – 0.136 cm/s, while the $\sigma^2_{\ln K}$ ranged from 0.001 to 0.868. The implication of these results is that no one method could be relied upon to characterize the aquifer accurately to obtain a reliable K_G or $\sigma^2_{\ln K}$ because of differences in scales in which the data are collected as well as experimental and interpretive errors that arise from various methods.

- Cross-hole pumping tests conducted at different locations within the aquifer [with different pumping locations] yielded different K_G values of approximately a factor of three difference. We emphasize that this small difference in the K_G could be interpreted to be small or lie within a margin of error given the uncertainties associated with each pumping test. In this sandbox aquifer, the factor of three difference in K caused a large variation in terms of model validation. A similar finding that the mean K from cross-hole tests differed from various cross-hole tests was reported by *Vesselinov et al*. [2001a-b], and *Illman* [2006] in the context of cross-hole pneumatic injection tests conducted in unsaturated fractured tuff at the Apache Leap Research Site [ALRS], and by *Illman and Tartakovsky* [2006] in the context of cross-hole hydraulic tests in fractured granite at the Grimsel site.

- Steady state hydraulic tomography of cross-hole pumping tests on the synthetic heterogeneous aquifer yielded the highest $\sigma^2_{\ln K}$ among the various characterization approaches. A similar finding was reported by *Vesselinov et al*. [2001b] who analyzed 3 cross-hole pneumatic injection tests conducted in unsaturated fractured tuffaceous rock at the ALRS in a tomographic manner. The cause of the high $\sigma^2_{\ln K}$ estimates from hydraulic and pneumatic tomography result from multiple factors, including but not limited to the fact that the approach is more suitable in capturing the heterogeneity patterns, but can also be affected by conceptual, numerical and estimation errors.

- Steady state hydraulic tomography analysis of cross-hole pumping tests yielded a K tomogram that was qualitatively similar looking to the photograph of the synthetic aquifer deposits. In particular, a number of high K features and their locations estimated using the steady state hydraulic tomography approach corresponded quite well with certain deposits known to have high K values in the aquifer. However, this correspondence was far from perfect with a few strata expected to have high K layers

did not appear in the *K* tomogram. Despite the good correspondence of the general locations of high *K* features, the comparison showed that the boundaries of the lithofacies from the deposits to the boundaries of high *K* features from the tomogram are not in close agreement. The similarity or difference between the lithofacies from geological investigations and hydrofacies determined from hydraulic tomography and other methods should be examined with greater scrutiny in future studies.

- The geometric mean *K* value determined from small scale data and homogeneous interpretation of cross-hole tests were used in a forward model to simulate 16 cross-hole pumping tests conducted in the synthetic aquifer. The forward simulations revealed that the predictions of 16 pumping tests using geometric and effective *K* estimates from various methods showed biased results in terms of predicting drawdowns from independent cross-hole tests. In particular, when the estimated geometric mean value was higher, the measured drawdown values were found to be higher in comparison to the simulated values. The converse was also found to be true.

- Kriging of *K* estimates from small scale tests [core, slug, and single-hole] was conducted. Forward simulations of the 16 pumping tests showed biased results in terms of predicting drawdowns from independent cross-hole tests. This suggests that the traditional geostatistical analysis of small scale *K* data may not yield accurate *K* distributions for the prediction of heads from large scale cross-hole pumping tests.

- Steady state hydraulic tomography analysis of 8 cross-hole tests yielded a more accurate heterogeneous distribution of *K* in comparison to the kriged fields. Our observation is based on the forward simulation of 16 cross-hole tests not used in the construction of the tomogram which revealed an excellent correspondence between the measured and simulated drawdowns.

- The various homogeneous and heterogeneous *K* estimates were then further tested by simulating 6 bidirectional flow-through tests. The homogeneous *K* estimates by taking the geometric mean of core, single-hole and cross-hole estimates of *K* significantly underestimated total flow from the outflow boundary. Likewise, forward simulations using heterogeneous *K* distributions obtained by kriging of core and single-hole *K* also underestimated total flow, although the predictions were slightly better than their homogeneous counterpart. In contrast, total flow estimates obtained using the *K* tomogram from steady state hydraulic tomography analysis of 8 cross-hole tests was considerably more accurate providing further support for the robustness of hydraulic tomography in accurately characterizing the subsurface.

- Based on our experimental work and modeling efforts, we conclude that steady-state groundwater models can be validated at least in this laboratory sandbox aquifer. Of course, absolute validation will likely never be possible. This conclusion comes with the caveat that groundwater model validation is possible only when the subsurface heterogeneity patterns and forcing functions are captured to sufficient accuracy. Finally, this study suggests that more effort should go into the accurate characterization of the

subsurface if the objective is to better predict subsurface head distributions, which then should yield better predictions of the fate of subsurface contaminants. Our results lend strong support to *Yeh and Lee* [2007]'s opinion that it is time to change the way we collect and analyze data for subsurface characterization.

21.2.6 Capturing heterogeneity in groundwater flow parameters: Comparison of approaches through controlled sandbox experiments

- Groundwater modeling has become a vital component to water supply and contaminant transport investigations. These models require representative hydraulic conductivity [K] and specific storage [S_s] estimates, or sets of estimates that idealize subsurface heterogeneity. Currently, there are a number of modeling approaches to idealize K and S_s heterogeneity in varying degrees of detail, but there is no consensus in terms of which approach results in groundwater models that have the best predictive capability. The main goal of this study is to find out which heterogeneity modeling approach [e.g. effective parameters, geostatistics, geological model, and hydraulic tomography] performs the best when input into a forward groundwater model in predicting 16 independent cross-hole pumping tests not used in the characterization effort. The study is conducted in a synthetic heterogeneous, aquifer built in the laboratory under controlled conditions.

- We first characterized the heterogeneous aquifer through single- and cross-hole pumping tests. Single-hole pumping tests conducted at 48 locations throughout the aquifer were interpreted by assuming that the medium is uniform. The values were then averaged to obtain effective values of K and S_s. A cross-hole test with pumping taking place near the central portion of the aquifer [port 21] yielded 48 drawdown curves. Similar to the single-hole data, 48 drawdown data were interpreted by treating the medium to be uniform, which yielded equivalent estimates of K and S_s. A geometric mean of these 48 data was then taken to obtain an effective K and S_s estimates for the entire aquifer. Comparison of the effective K and S_s estimates from the single- and cross-hole pumping tests show that they are quite different, suggesting that effective values depend on the test type and the scale at which they are conducted.

- The single- and cross-hole pumping test data were then used to construct various forward groundwater models with homogeneous and heterogeneous K and S_s estimates. Homogeneous models were constructed by using the effective parameters from the single- and cross-hole pumping tests. Heterogeneous models consisted of spatially variable K and S_s fields obtained via: 1] kriging single-hole data; 2] accurately capturing the stratigraphy and assigning single-hole K and S_s estimates corresponding to a given layer; and 3] conducting transient hydraulic tomography to obtain ln-K and ln-S_s tomograms.

- Visual comparisons of the ln-K tomogram to the deposits show that many of the features are captured although we do not expect a direct correlation of the ln-K distribution and the stratification due to the intralayer variability in K and the fact that there are a limited number of measurement points along each column of ports. We anticipate that additional details to the ln-K tomogram will emerge as more pressure measurements are made along

each column. The comparison of the ln-K tomogram obtained using transient hydraulic tomography to the one obtained using steady state hydraulic tomography by *Illman et al*. [2010] showed overall similarity. In both cases, 8 pumping tests were used for the inverse analysis. However, we observe that more details in the heterogeneity pattern are visible in the ln-K tomogram obtained using transient hydraulic tomography. This is the direct result of using a larger number of drawdown data in transient hydraulic tomography to capture the entire drawdown curve. In contrast, steady state hydraulic tomography only uses a single steady state head value from each drawdown record.

- We next compared the ln-S_s tomogram computed using transient hydraulic tomography to the kriged ln-S_s field obtained from the geostatistical analysis [not shown here] of the 48 single-hole K and S_s estimates. The comparison revealed that the pattern in which ln-S_s is distributed is, in general, similar. In particular, ln-S_s values are higher near the top of the aquifer while it decreases towards the bottom. However, the actual values of ln-S_s differed quite significantly when the results from kriging and transient hydraulic tomography are compared. The differences in estimated values can have significant impacts on the rate at which drawdown propagates through the aquifer and more importantly on estimates of water stored in a given aquifer.

- The performance of these homogeneous and heterogeneous K and S_s fields were then tested through the forward numerical simulation of cross-hole pumping tests that were not used in the characterization effort. Drawdown values from 16 cross-hole pumping tests conducted in the sandbox aquifer were then directly compared through scatter plots. The comparison was done for individual tests and also for all 16 tests together. A linear model was fit to each of the scatter plots and the coefficient of determination [R^2] computed to quantitatively assess the goodness of fit between the observed and simulated drawdown values. The slope and intercept of the linear model fit provides information on prediction bias.

- We found that the forward numerical simulations using the effective K and S_s values computed by averaging the single- and cross-hole test K and S_s estimates could not predict the 16 cross-hole pumping tests accurately for almost all tests and significant bias was shown when the observed and simulated drawdowns were compared through a series of scatter plots. The scatter and bias were evident in not just individual tests, but also when the data from all 16 tests were plotted together.

- The heterogeneous models were also tested in a similar fashion. We found that the forward simulations of 16 independent cross-hole pumping tests using the kriged K and S_s fields and the geological model, which accurately represents the stratification, both failed to accurately predict the transient drawdown responses from all 16 tests whether they were examined individually or collectively. This was surprising as one would expect that both heterogeneous models should better capture the variability in hydraulic parameters which should then lead to better predictions of independent cross-hole tests. This was certainly not the case in our synthetic heterogeneous aquifer. In contrast, the forward numerical modeling of the 16 tests using the ln-K and ln-S_s tomograms

computed via transient hydraulic tomography convincingly showed that these distributions led to the smallest discrepancy between observed and simulated drawdowns.

- Our sandbox experiments suggest the robustness of transient hydraulic tomography conducted in a controlled environment. Perhaps the robustness of hydraulic tomography results from the repeated calibration of multiple cross-hole pumping test data, which leads to the improved characterization of connectivity of hydraulic parameters. The accurate estimation of hydraulic parameters and their connectivity are critical in the accurate prediction of independent cross-hole pumping tests which the traditional geostatistical approach failed to do so in our laboratory sandbox aquifer. We are optimistic of the capabilities of hydraulic tomography, but are also cautious in generalizing our sandbox results to the field scale. It remains to be seen whether hydraulic tomography will yield robust results under field conditions. We are currently conducting a comprehensive field assessment of hydraulic tomography and other heterogeneous characterization methods at the University of Waterloo. These results will be reported in the near future.

21.2.7 Temporal moment analysis of partitioning tracers in selected laboratory sandbox aquifers

- Based on the results analyzed and presented in this work it is possible to conclude the following: The Partitioning Tracer Technique [PTT] is highly dependent on several factors: 1] the tracer test design including the injection and sampling points 2] the heterogeneities of the porous media play a crucial role on the advective transport of solutes, even if they are geometrically simple and built in highly controlled artificial aquifers [SBs].

- The architecture [morphology of the spill] of the DNAPL spill was found to affect PTT results significantly. Uneven DNAPL distributions and entrapped DNAPL in unreachable flow and transport pathways were found to be important in the underestimation of DNAPL saturations as they are not fully accounted by the method of temporal moments.

- In every PTT design, the interactions between the flow regime and the heterogeneities will result in ports that will preferentially sample the tracers. Some low K blocks will deflect the currents of the flow field and therefore the interaction between the tracers and the entrapped DNAPL is reduced. This behavior leads to underestimation of the DNAPL volume. Conversely, part of the flow may travel to other ports increasing the chances of the interactions between the entrapped DNAPL and the tracers. This behavior will result in over estimations of the DNAPL volume.

- The strength of the partitioning coefficient [K_{NW}] was also found to have an impact of DNAPL saturation estimates. We found that tracers with smaller K_{NW} tended to yield more accurate DNAPL saturation estimates that those with large K_{NW} values. Moreno-Barbero et al. [2006] suggested that the estimates will improve by applying non-equilibrium models instead of the application of equilibrium models. Also, an accurate

estimate of the interaction time between tracer and DNAPL would provide a means to choose a more appropriate K_{NW} value to apply to the method of temporal moments.

- The saturation and volume estimates from the method of temporal moments vary from port to port and from experiment to experiment due to the conclusions made in point 1. The best way to use and estimate DNAPL saturations and volume is by doing an overall average for each set of tests.

- Based on our laboratory experiments, we found that the method of temporal moments over/underestimates by up to 50% the volume & mass of DNAPL. However, the estimates based on method of temporal moments are more accurate than the estimates given by the core analysis.

- The multiple tracer test [SB 1-2], showed that the partitioning coefficient plays an important role and that the best estimates of DNAPL are obtained from the tracer with the lowest partitioning coefficient. The performance of the method of temporal moments may be improved by using non-equilibrium partitioning coefficients.

- The use of method of temporal moments is quite simple and useful in spite of all the variables involved within a PTT: DNAPL architecture, sampling scheme, nature of tracers and possible mass losses from the SB.

- In order to improve the partitioning tracer testing method, it may be necessary to address the following issues: 1] PTTs conducted in the laboratory should account for the outflow during tracer sampling. By doing so, one should be able to account for DNAPL mass losses during tracer sampling. The results reported in this section points to the potential need for proper accounting of DNAPL mass loss. Our results show that the mass losses estimates are low; 2] One potential way to improve the performance of PTTs is to lower the injection rate in order to maximize the contact of the partitioning tracers with the DNAPLs in the source zone.

21.2.8 Hydraulic and partitioning tracer tomography for trichloroethylene source zone characterization: Small-scale sandbox experiments

- Trichloroethene [TCE] and other dense nonaqueous phase liquids [DNAPL] are prevalent at a large number of sites throughout the world. The variable release history, unstable flow, and geologic heterogeneity make the spatial distribution of DNAPLs complex. This causes difficulties in site remediation contributing to long-term groundwater contamination for decades to centuries. We presented small-scale laboratory experiments to demonstrate the efficacy of Sequential Successive Linear Estimator [SSLE] algorithm [*Yeh and Zhu*, 2007] that images DNAPL source zones. The algorithm relies on the fusion of hydraulic and partitioning tracer tomography [HPTT] tests to derive the best estimate of the hydraulic conductivity [*K*] heterogeneity, DNAPL saturation distribution and their uncertainty. This study leads to the following major conclusions:

- HPTT is a non-destructive method that yields reliable estimates of mean TCE saturations. The approach utilizes dipole cross-hole tests to image the *K* heterogeneity through hydraulic tomography [HT]. The *K* tomogram then is employed in partitioning tracer tomography [PTT] to image the spatial distribution of TCE saturation.

- The HPTT approach is a non-destructive approach and can be applied repeatedly prior to and post-remediation. More importantly, it also provides uncertainty estimates that can facilitate better decision making on source zone characterization, remediation, and long-term monitoring.

- Results from our laboratory experiments show that the mean TCE saturation estimated through the interpretation of individual tracer tests can potentially be erroneous. This is because the sweeping of the source zone is dependent on the heterogeneity pattern and tracer test design. Our findings cast doubt on estimating the saturation and volumes of TCE through the interpretation of single tracer test. However, the mean TCE saturation was very close to the actual TCE saturation when multiple tracer tests were sequentially analyzed using the HPTT method.

- TCE saturation distributions compare favorably with TCE distributions observed in the sandbox. The estimates improved when the loop iteration scheme [*Zhu and Yeh*, 2005] developed originally for hydraulic tomography was applied to the interpretation of multiple partitioning tracer tests.

- The TCE saturation estimates obtained through HPTT did not compare favorably with local saturation estimates from core samples. Direct sampling of core samples to obtain point scale estimates of TCE saturations can potentially produce inaccurate maps of TCE saturations. This is due to our observation that driving of core sampling tube can cause consolidation of sediments and sample loss during extraction of core samples.

- Delineation of *K* heterogeneity prior to the conduct of partitioning tracer tests can improve test designs and maximize sweeping of the tracers through the source zone. The interpretation of the partitioning tracer tests through inverse modeling is significantly improved because the heterogeneity patterns are imaged a priori. This emphasizes the importance of accurate characterization of subsurface heterogeneity as this has a significant impact on tracer transport and its corresponding inverse modeling.

- The dipole configuration of cross-hole hydraulic and partitioning tracer tests assists in ensuring the mass balance of water and tracers as well as aids in the minimization of the impacts of boundary conditions. It also has the benefit of inducing a larger signal in terms of drawdowns observed in observation ports. The larger signal detected in the observation ports results in a higher signal in comparison to data noise that are inherent in laboratory or field experiment. Furthermore, the dipole configuration has a practical implication in which pumped water that is likely contaminated is immediately injected back into the aquifer which helps in cost savings because an above ground treatment facility of contaminated water is not required. Through our laboratory experiments, we conclude that the dipole partitioning tracer tests does not mobilize the nonaqueous phase

TCE. Further studies are needed to ascertain the stability of nonaqueous phase TCE in the source zone through multiphase flow modeling of dipole partitioning tracer tests under different levels of saturations and heterogeneity patterns.

- Maps of TCE saturations imaged with the HPTT technology have multiple benefits. The TCE saturation tomogram can be used to design efficient remediation schemes. In addition, the *K* tomogram imaged also can be used to optimize the injection of various amendments such as chemical oxidation agents and nanoiron for nonaqueous phase destruction and microbes to induce bioremediation. Because the HPTT technology is based on stochastic methods, it provides uncertainty estimates in *K*, TCE saturation, and their connectivity. Maps of uncertainties are useful because these can assist in guiding additional data collection.

21.2.9 Hydraulic and partitioning tracer tomography experiments in an intermediate scale sandbox

- For this study, research is still ongoing thus findings and conclusions will be reported in the future.

21.3 FINDINGS AND CONCLUSIONS FROM TASK 3: DEVELOPMENT OF HYDRAULIC AND TRACER TOMOGRAPHY ALGORITHMS THAT CONSIDERS THE EFFECT OF THE UNSATURATED ZONE

21.3.1 A revisit of drawdown behavior during pumping in unconfined aquifers

- Results of our numerical experiments suggest that the transition between the two water release mechanisms [i.e., expansion of water and compaction of porous media and drainage of the porous media] along with vertical flow are the cause of the characteristic S-shaped drawdown curve commonly observed in the unconfined aquifer pumping test.

  At the early stage of the S-shaped drawdown-time curve, contribution to the well discharge is mainly from the compaction of the aquifer and expansion of the water. At the intermediate stage, not only the drainage from the unsaturated zone above the initial water table but also that from initially saturated pores during falling of the water table play an important role in the development of the S-shape hydrograph. At the late time stage the contribution from falling of the water table [i.e., drainage of initially saturated pores] becomes the dominant source of water. We believe the widely accepted terms "delayed yield" or "delayed drainage" that suggests a lag in water release from the initially unsaturated zone does not elucidate the transition of the two water release mechanisms and perhaps these terms are misleading.

- The first-order stochastic moment analysis shows that the S-shaped drawdown-time curve during a pumping test in unconfined aquifers is sensitive to the spatial variability of hydraulic conductivity, specific storage, and saturated moisture content and insensitive to the variability of other parameters for unsaturated hydraulic properties. This may suggest that in order to identify these unsaturated hydraulic properties, measurements of negative pressure heads and moisture contents are necessary. Furthermore, a cross-correlation

analysis reveals that an observed drawdown at a given location in a heterogeneous unconfined aquifer is mainly influenced by local heterogeneity near the pumping and observation wells. As such, applications of a model assuming homogeneity to estimation of parameters may require a large number of spatial observations in order to yield representative parameter values. This requirement of a large number of observations in space may support an inverse modeling effort to estimate the spatially distributed parameters.

- In conclusion, we believe a multi-dimensional variably saturated flow model, which considers the transition of water release mechanisms and accounts for heterogeneity, would provide a more realistic representation of flow processes in unconfined aquifer during a pumping test. Hydraulic tomography would be a viable approach for delineating heterogeneity in unconfined aquifers.

21.3.2 Hydraulic tomography to characterize the heterogeneity of unconfined aquifers

- In this study, we applied the hydraulic tomography concept to characterize heterogeneities of unconfined aquifers. A three dimensional variably saturated flow model was used to simulate groundwater flow in unconfined aquifers due to the complex nature of groundwater movement near the water table and in the vadose zone. The spatial variations of major hydraulic parameters [K_s, S_s, and α] were estimated by the SSLE method using transient pressure data collected from a hydraulic tomography survey. The method was tested by two numerical examples and the relation between unconfined transient pressure heads and the three parameters were also explored. The synthetic examples demonstrate that a fully variably saturated flow model is appropriate for simulating groundwater flow in unconfined aquifers and hydraulic tomography is a promising tool for characterizing unconfined aquifer. The examples also show that SSLE can handle highly nonlinear problems, providing unbiased estimates of multiple parameters simultaneously. The sensitivity analysis shows that the transient pressure head changes due to a pumping test are related to different parameters non-equally at different stage of the pumping test. As a result, the pressure data from all stages of the pumping tests should be used to maximize information utilization.

- This study is merely a proof of concept; a workable field application of hydraulic tomography to unconfined aquifers remains to be developed. Below we discuss some of the limitations and also some potential areas for future research.

- First, we only considered three hydraulic parameters [K_s, S_s, and α] as spatial varying parameters in this study. Other parameters, including and , were considered as homogeneous and known a priori. In reality, these parameters are also heterogeneous in nature. However, including too many parameters in the inversion process can potentially cause numerical instability and less reliable estimates. A correlation analysis of pressure responses to all potential spatial varying parameters may be the first logical step to identify a few most responsible parameters that will be estimated through inversion. The remaining parameters will be considered as constant.

- Second, the parameter α for Gardner-Russo model is apparently model-dependent. If one other saturation-pressure model was used, new parameters need to be estimated. The parameters for any saturation-pressure model are likely to be spatial varying [Yeh et al., 2002] and should be treated as stochastic processes [Hughson and Yeh, 2000].

- Third, A unique aspect of characterizing an unconfined aquifer is that some parameters [such as S_s] only relate to positive pressure heads [i.e., where the flow is in saturated condition] whereas other parameters [such as α in this study] only relate to negative pressure heads [i.e., where the flow is unsaturated condition]. The full characterizations of these parameters for the entire field of interest are typically unobtainable. The aspect can limit the application of HT for unconfined aquifer. For example, if a HT was conducted when the water table is high [i.e, a very wet year], the estimated α will be limited to small portion of the aquifer and will not be reliable for simulating the groundwater flow when the water table is much lower [i.e., a very dry year].

- Finally, the variably saturated flow model is nonlinear in nature and requires fine spatial and temporal resolutions for accurate representation of unsaturated flow. The interpretation of field HT for an unconfined aquifer using any inverse method will be likely computational intensive. In this study, although a multi-grid method and parallel computing were used for the fully three dimensional case, the computational cost is still too high for a typical field application. While more sophisticated computing technologies should be pursued to further reduce the computation cost, the future research should also focus on establishing a practical, field-oriented protocol for HT application of unconfined aquifers similar to that developed by Xiang et al. [2009] for confined aquifers.

21.4 CONCLUSIONS FROM TASK 4: VALIDATION OF ALGORITHMS DEVELOPED IN TASK 3 USING A NEW SANDBOX WITH MONITORING IN BOTH THE UNSATURATED AND SATURATED ZONES

21.4.1 Flow to a well in a heterogeneous unconfined aquifer: Insights from an intermediate scale sandbox

- Flow to wells due to the pumping of unconfined aquifers has been a topic of great interest for many decades. The majority of the studies described in the literature have focused on analytical solutions which treat the medium to be homogeneous. However, it is the rule rather than the exception that aquifers are heterogeneous.

- We study the impact of the unsaturated zone and effects of heterogeneity on unconfined aquifer flow in a synthetic aquifer packed in an intermediate-scale laboratory sandbox. The synthetic aquifer was characterized initially with eight cross-hole pumping tests under fully saturated conditions and the hydraulic data are interpreted using the transient hydraulic tomography code developed by *Zhu and Yeh* [2005]. The resulting hydraulic conductivity [K] and specific storage [S_s] tomograms were validated using eight additional tests not used in the inverse modeling. Effective parameters of K and S_s were also estimated through the inverse modeling of individual cross-hole pumping tests. Laboratory hanging column experiments of sands in the unsaturated zone were also to

determine the van Genuchten parameters. We then conducted another cross-hole pumping test in which a port near the bottom of the aquifer was pumped at a constant rate. During this test, the water table was allowed to freely move downwards in response to pumping. The saturated zone was monitored via pressure transducers and the unsaturated zone with tensiometers and water content sensors.

- Forward models of various complexities in saturated and unsaturated parameters were then built using the variably saturated code, MMOC3 [*Yeh et al*., 1993] to examine the sensitivity of the homogeneous and heterogeneous parameter estimates on the predictability of the unconfined aquifer test. Results show that an average or effective K and S_s determined through the averaging of equivalent K and S_s obtained from the inverse modeling of a pumping test cannot yield accurate predictions of drawdown responses in the saturated and unsaturated zones. Our predictions of drawdown responses in the saturated zone improved dramatically when the K and S_s distributions were utilized in the forward simulation of the unconfined aquifer pumping test. Because we monitored the pressure changes in the unsaturated zone during this experiment, we also attempted to evaluate the predictions of pressure responses in the unsaturated zone. Results show that slightly improved predictions of drawdowns in the unsaturated zone could be achieved if heterogeneous zones were assigned to the model. However, our results showed that predictions of drawdowns in the saturated zone were relatively insensitive to whether we conceptualized the unsaturated zone to be homogeneous or heterogeneous which confirms a conclusion reached by *Mao et al*. [in review] using numerical simulations.

- Based on this study it is possible to accurately predict the response of a heterogeneous unconfined aquifer to pumping as long as the saturated parameters [K and S_s] are accurately characterized and an accurate effective value of the unsaturated parameters is known. This study was unique in that the heterogeneity pattern was exactly known and it was possible to select effective unsaturated parameters by assessing would material would dominate the drainage response. In the field however, this is not as straight forward and the selection of unsaturated parameters could pose a challenge. As such, interpreting pumping tests in heterogeneous aquifers will benefit from inverse modelling of both the unsaturated and saturated zones. This will require the use of a variably saturated model coupled with an inverse algorithm such as the Sequential Successive Linear Estimator [e.g., *Hughson and Yeh*, 2000] for the proper interpretation of flow to wells in a heterogeneous unconfined aquifer.

LITERATURE CITED

Adams, E. E. and L. W. Gelhar [1992], Field study of dispersion in a heterogeneous aquifer: 2. Spatial moments analysis, *Water Resour. Res.*, 28[12], 3293–3307.

Akindunni, F.F., and R.W. Gillham [1992], Unsaturated and saturated flow in response to pumping of an unconfined aquifer: numerical investigation of delayed drainage, *Ground water*, *30*, 873–884.

Alexander, M., S. J. Berg, and W. A. Illman [2010], Field study of hydrogeologic characterization methods in a heterogeneous aquifer, *Ground Water*, now available online.

Alvarez, P.; Illman, W.A., 2006 *Bioremediation and Natural Attenuation: Process Fundamentals and Mathematical Models*, John Wiley & Sons, 609 pp.

Annable, M. D.; Rao, P. S. C.; Hatfield, K.; Graham, W. D.; Wood, A. L.; Enfield, C. G. [1998] Partitioning tracers for measuring residual NAPL: Field-scale test results, *J. Environ. Eng*., *124*[6], 498-503.

Bakr, A. A., L. W. Gelhar, A. L. Gutjahr, and J. R. MacMillan. 1978. Stochastic Analysis of Spatial Variability in Subsurface Flows. 1. Comparison of One- and Three-Dimensional Flows. Water Resour. Res. 14[2]:263-271.

Barrash, W., and T. Clemo [2002], Hierarchical geostatistics and multifacies systems: Boise Hydrogeophysical Research Site, Boise, Idaho, Water Resour. Res., 38[10], 1196, doi:10.1029/2002WR001436.

Barth, G., M. Hill, T. Illangasekare, and H. Rajaram [2001], Predictive Modeling of Flow and Transport in a Two-Dimensional Intermediate-Scale, Heterogeneous Porous Medium, *Water Resour. Res*., 37[10], 2503-2512.

Beckie, R., and C. F. Harvey [2002], What does a slug test measure: An investigation of instrument response and the effects of heterogeneity, Water Resour. Res., 38[12], 1290, doi:10.1029/2001WR001072.

Benson DA, Wheatcraft SW, Meerschaert MM. [2000], The fractional-order governing equation of Levy motion. Water Resour. Res., 36[6]: 1413–23.

Berg, S. J. and W. A. Illman [2011], Three-dimensional transient hydraulic tomography and comparison to other heterogeneity imaging methods, Eos Trans. AGU, Fall Meet. Suppl., Abstract. H13D-0996 [poster presentation].

Berg, S. J. and W. A. Illman [2011], Capturing heterogeneity in groundwater flow parameters: Comparison of approaches through controlled sandbox experiments, *Water Resources Research*, in review.

Berkowitz, B., A. Cortis, M. Dentz, H. Scher [2006], Modeling non-Fickian transport in geological formations as a continuous time random walk. Rev. Geophys., 44:RG2003.

Bevan, M.J., A.L. Endres, D.L. Rudolph and G. Parkin [2005], A field scale study of pumping-induced drainage and recovery in an unconfined aquifer, *J. Hydrol.*, *315*, 52-70.

Boggs, J. M., S. C. Young, L. M. Beard, L. W. Gelhar, K. R. Rehfeldt, and E. E. Adams [1992], Field study of dispersion in a heterogeneous aquifer: 1. Overview and site description, Water Resour. Res., 28[12], 3281–3291, doi:10.1029/92WR01756.

Bohling, G. C., X. Zhan, J. J. Butler Jr., and L. Zheng [2002], Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities, *Water Resour. Res*., 38[12], 1324, doi:10.1029/2001WR001176.

Bohling, G. C., J. J. Butler Jr., X. Zhan, and M. D. Knoll [2007], A field assessment of the value of steady shape hydraulic tomography for characterization of aquifer heterogeneities, *Water Resour. Res*., 43, W05430, doi:10.1029/2006WR004932.

Boman, G.K., Molz, F.J., and K.D. Boone [1997], Borehole flowmeter application in fluvial sediments: Methodology, results, and assessment, *Ground Water*, 35[3], 443-450.

Bonnet, M., Forkasiewicz, and P. Peaudecerf [1970], Méthodes d'interprétation de pumpages d'essai en nappe libre. Bur. *Rech. Geol. Min. Rep*. 70SGN 359 HYD, Orleans, France.

Boulton, N. S. [1954], Unsteady radial flow to a pumped well allowing for delayed yield from storage, *Int. Ass. Sci. Hydrol*., *2*: 472-477.

Boulton, N. S. [1963], Analysis of data from nonequilibrium pumping tests allowing for delayed yield from storage, *Proc. Inst. Civil Eng.*, *26*, 469-482.

Bouwer, H., and R.C. Rice. 1976. A slug test for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. Water Resources Research 12, no. 3: 423–428.

Brauchler, R., R. Liedl, and P. Dietrich [2003], A travel time based hydraulic tomographic approach, *Water Resour. Res.*, *39*[12], 1370, doi:10.1029/2003WR002262.

Brooks, A.N. and Hughes, T.J.R., [1982]. Streamline upwind/Petrov–Galerkin formulations for convection dominated flow with particular emphasis on the incompressible Navier–Stokes equations. *Computer Methods in Applied Mechanics Engineering* 32, pp. 199–259.

Brooks, M.C.; Annable, M.D.; Rao, P.S.C.; Hatfield, K.; Jawitz, J.W.; Wise, W.R.; Wood, A.J.; Enfield, C.G. [2002], Controlled release, blind tests of DNAPL charcterization using partitioning tracers. *J. Contam. Hydrol*. *59*, 187–210.

Brutsaert, W. [1970], Immiscible multiphase flow in ground water hydrology: a computer analysis of the well flow problem, PhD Thesis, Colorado State University, Fort Collins.

Bunn M.I., J. P. Jones, A.L. Endres and D. L. Rudolph [2010], Effects of hydraulic conductivity heterogeneity on vadose zone response to pumping in an unconfined aquifer, *J. Hydrol.*, *387*, 90-104.

Butler, J. J., Jr., andW. Z. Liu [1993], Pumping tests in nonuniform aquifers: The radially asymmetric case, Water Resour. Res., 29[2], 259–269.

Butler, J. J., Jr. [2005], Hydrogeological Methods for Estimation of Spatial Variations in Hydraulic Conductivity, In: *Hydrogeophysics*, 23-58, Water Science and Technology Library, Vol. 50, Springer Netherlands, doi: 10.1007/1-4020-3102-5.

Butler, J. J., Jr., P. Dietrich, V. Wittig, and T. Christy [2007], Characterizing hydraulic conductivity with the direct-push permeameter, *Ground Water*, 45[4], 409– 419.

Cacas, M. C., E. Ledoux, G. de Marsily, A. Barbreau, P. Calmels, B. Gaillard, and R. Margritta, [1990a], Modeling fracture flow with a stochastic discrete fracture network: Calibration and validation, 1, The flow model, *Water Resour. Res*., *26*[3], 479-489.

Cacas, M. C., E. Ledoux, G. de Marsily, A. Barbreau, P. Calmels, B. Gaillard, and R. Margritta, [1990b], Modeling fracture flow with a stochastic discrete fracture network: Calibration and validation, 2, The transport model, *Water Resour. Res*., *26*[3], 491-500.

Caers J [2001] Geostatistical reservoir modelling using statistical pattern recognition. *J Petroleum Sci Eng* 29:177–188.

Cain, R.B.; Johnson, G. R.; McCray, J. E.; Blanford, W. J.; Brusseau, M. L. [2000] Partitioning tracer tests for evaluating remediation performance, *Ground Water*, *38*, 752-761.

Cardiff, M., W. Barrash, P. K. Kitanidis, B. Malama, A. Revil, S. Straface, and E. Rizzo. [2009], A potential-based inversion of unconfined steady-state hydraulic tomography, *Ground Water*, 47[2], 259-270.

Carle, S. F., and G. E. Fogg [1997], Modeling spatial variability with one and multi-dimensional continuous Markov Chains, *Math Geol*., 29[7], 891-918.

Carle, S. F. [1999], *T-PROGS: Transition Probability Geostatistical Software*, Version 2.1.

Castagna, M., and A. Bellin [2009], A Bayesian approach for inversion of hydraulic tomographic data, *Water Resour. Res.*, 45, W04410, doi:10.1029/2008WR007078.

Chao, H. C., H. Rajaram, T. Illangasekare [2000], Intermediate-scale experiments and numerical simulations of transport under radial flow in a two-dimensional heterogeneous porous medium, *Water Resour. Res*., 36[10], 2869-2884.

Chen, G., W. A. Illman, D. L. Thompson, V. V. Vesselinov, and S. P. Neuman [2000], Geostatistical, type-curve and inverse analyses of pneumatic injection tests in unsaturated fractured tuffs at the Apache Leap Research Site near Superior, Arizona, in Dynamics of Fluids in Fractured Rocks, Geophys. Monogr. Ser., vol. 122, edited by B. Faybishenko, P. A. Witherspoon, and S. M. Benson, pp. 73–98, AGU, Washington, D. C..

Cirpka, O. A., and P. K. Kitanidis [2001], Sensitivity of temporal moments calculated by the adjoint-state method and joint inversing of head and tracer data, *Adv. Water Resour*. *24*[1], 89-103.

Cohen, R. M. and Mercer, J. W. [1993], *DNAPL Site Characterization*. Boca Raton, Florida: Smoley.

Cooper, H. H., Jr., and C. E. Jacob [1946], A generalized graphical method for evaluating formation constants and summarizing well-field history, Eos Trans. AGU, 27, 526– 534.

Cortis A, Berkowitz B. Anomalous transport in ‘‘classical” soil and sand columns. Soil Sci Soc Am J 2004;68[5]:1539–48.

Craig, A. J., [2005], *Measurement of hydraulic parameters at multiple scales in two synthetic heterogeneous aquifers constructed in the laboratory*, M.S. thesis, Department of Civil and Environmental Engineering, The University of Iowa.

Cushman, J.H. and T. R. Ginn [1993], Nonlocal dispersion in media with continuously evolving scales of heterogeneity. Transp Porous Media,13[1]: 123–38.

Dagan, G.A. [1967], Method of determining permeability and effective porosity of unconfined anisotropic aquifers, *Water Resour. Res.*, *3[4]*, 1059-1071.

Dagan, G. [1982], Stochastic modeling of groundwater flow by unconditional and conditional probabilities: 2. The solute transport, Water Resour. Res., 18[4], 835– 848, doi:10.1029/WR018i004p00835.

Dagan, G. [1984], Solute transport in heterogeneous formations, J. Fluid Mech., 145, 151– 177, doi:10.1017/S0022112084002858.

Danquigny, C., P. Ackerer, J. P. Carlier [2004], Laboratory tracer tests on three-dimensional reconstructed heterogeneous porous media, *J. Hydrol.*, 294[1-3], 196-212.

Datta-Gupta, A., Yoon, S., Vasco, D. W., Pope, G. A. [2002], Inverse modeling of partitioning interwell tracer tests: A streamline approach, *Water Resources Research*, *38*[6], 1079, doi:10.1029/2001WR000597.

de Marsily, G., 1978. De l’Identification des systemes en hydrogeologiques, these de docteur es sciences. Univ. Pierre et Marie Curie-Paris VI, Paris.

De Marsily, G., F. Delay, J. Gonçalvès, Ph. Renard, V. Teles, and S. Violette, Dealing with spatial heterogeneity, *Hydrogeol. J.*, 13:161-183, 2005.

de Vries, L. M., J. Carrera, O. Falivene, O. Gratacos, and L. J. Slooten [2009], Application of multiple point geostatistics to non-stationary images, *Mathematical Geosci*., 41[1], 29-42.

Dershowitz, W., G. Lee, J. Geier, S. Hitchcock, and P. LaPointe, [1997], Fracman Interactive Discrete Feature Data Analysis, Geometric Modeling, and Exploration Simulation, User Documentation, Version 2.6, Golder Associates Inc. Seattle, WA.

Deutsch, C. V., and A. G. Journel [1998], *GSLIB Geostatistical Software Library and User's Guide*, 2nd ed., 369 pp.

Di Federico, V., and S. P. Neuman [1998], Transport in multiscale log conductivity fields with truncated power variograms, Water Resour. Res., 34[5], 963–973, doi:10.1029/98WR00221.

Dietrich, P., and C. Leven [2005], Direct Push Technologies, In Groundwater Geophysics, edited by R. Kirsch, Springer, 321– 340.

Dietrich, P., J. J. Butler Jr., and K. Faiß [2008], A rapid method for hydraulic profiling in unconsolidated formations, Ground Water, 46[2], 323– 328, doi:10.1111/j.1745-6584.2007.00377.x.

Doherty, J. [1994], *PEST: Model-Independent Parameter Estimation*, 122 pp., Watermark Comput., Corinda, Qld., Australia.

Dridi, L.; Pollet, I.; Razakarisoa, O.; Schäfer G. [2009] Characterization of DNAPL source zone in a porous aquifer using the Partitioning Interwell Tracer Test and in inverse modeling approach, *Journal of Contaminant Hydrology*, *107*, 22-44.

Dugan, P. J., McCray, J. E., Thyne G. D. [2003], Influence of a solubility-enhancing agent [cyclodextrin] on NAPL-water partition coefficients, with implications for partitioning tracer tests, *Water Resour. Res*., *39*[5], 1123, doi:10.1029/2002WR001672.

Dverstorp, B., J. Andersson, and W. Nordqvist, [1992], Discrete fracture network interpretation of field tracer migration in sparsely fractured rock, *Water Resour. Res*., *28*[9], 2327-2343.

Dwarakanath, V.; Pope, G.A. [1998], New approach for estimating alcohol partition coefficients between nonaqueous phase liquids and water. *Environ. Sci. Technol*., *32*, 1662–1666.

Endres, A. L., J. P. Jones and E.A. Bertrand [2006], Pumping-induced vadose zone drainage and storage in an unconfined aquifer: A comparison of analytical model predictions and field measurements, *J. Hydrol.*, *335*, 207- 218, doi:10.1016/j.jhydrol.2006.07.018.

Fernàndez-Garcia, D., H. Rajaram, and T. H. Illangasekare [2005], Assessment of the predictive capabilities of stochastic theories in a three-dimensional laboratory test aquifer: Effective hydraulic conductivity and temporal moments of breakthrough curves, Water Resour. Res., 41, W04002, doi:10.1029/2004WR003523.

Fienen, M. N., T. Clemo, and P. K. Kitanidis [2008], An interactive Bayesian geostatistical inverse protocol for hydraulic tomography, *Water Resour. Res.*, *44*, W00B01, doi:10.1029/2007WR006730.

Foglia, L., S.W. Mehl, M.C. Hill, P. Perona, P. Burlando [2007], Testing Alternative Ground Water Models Using Cross-Validation and Other Methods, *Ground Water*, 45[5], 627-641.

Garabedian, S. P., D. R. Leblanc, L. W. Gelhar, and M. C. Celia [1991], Large-scale natural-gradient test in sand and gravel, Cape Cod, Massachusetts: 2. Analysis of spatial moments for a nonreactive tracer, Water Resour. Res., 27[5], 911– 924, doi:10.1029/91WR00242.

Gardner, W.R. [1958], Some steady-state solutions of unsaturated moisture flow equations with application to evaporation from a water table, *Soil Sci.*, *85*:228–232.

Gelhar, L. W., and C. L. Axness [1983], Three-dimensional analysis of macrodispersion in a stratified aquifer, Water Resour. Res., 19, 161– 180, doi:10.1029/WR019i001p00161.

Gelhar, L. W., [1993], *Stochastic Subsurface Hydrology*. Prentice-Hall. Englewood Cliffs, New Jersey.

Gottlieb, J., and P. Dietrich [1995], Identification of the permeability distribution in soil by hydraulic tomography, *Inverse Probl*., 11[2], 353–360.

Guardiano, F. and R. M. Srivastava[1993], Multivariate geostatistics: beyond bivariate moments. In: Soares A [ed] *Geostatistics-Troia*. Kluwer, Dordrecht, pp 133–144

Gutjahr, A. [1989], Fast Fourier transforms for random field generation, project report, contract 4-R58-2690 R, N. M. Inst. of Min. and Technol., Socorro.

Gutjahr, A. L., and J. L. Wilson [1989], Cokriging for stochastic models, *Transp. Porous Media*, 4[6], 585-598.

Guzman, A. G., A. M. Geddis, M. J. Henrich, C. F. Lohrstorfer, and S. P. Neuman [1996], Summary of air permeability data from single-hole injection tests in unsaturated fractured tuffs at the Apache Leap Research Site: Results of steady-state test interpretation, Rep. NUREG/CR-6360, U.S. Nucl. Regul. Comm., Washington, D. C.

Hagen, A. 2003. Fuzzy set approach to assessing similarity of categorical maps, Int. j. Geographical Information Science,Vol. 17, NO. 3, 235-249.

Haldorsen H. H. and D. M. Chang [1986], Notes on stochastic shales from outcrop to simulation models. In: Lake LW, Carol HB Jr [eds] Reservoir characterization. Academic, New York, pp 152–167.

Hanna, S., and T.-C. J. Yeh [1998], Estimation of co-conditional moments of transmissivity, hydraulic head, and velocity fields, Adv. Water Resour., 22[1], 87– 95.

Hantush, M.S. [1961a], Drawdown around a partially penetrating well, *Jour. of the Hyd. , Div., Proc. of the Am. Soc. of Civil Eng., vol. 87*, no. HY4, pp: 83-98.

Hantush, M.S. [1961b], Aquifer tests on partially penetrating wells, *Jour. of the Hyd. , Div., Proc. of the Am. Soc. of Civil Eng., vol. 87*, no. HY5, pp: 171-194.

Hao, Y., T.-C. J. Yeh, W. A. Illman, K. Ando, K.-C. Hsu [2008], Hydraulic tomography for detecting fracture connectivity, *Ground Water*, 46[2], 183-192.

Harter, Th., and T.-C. J. Yeh [1996], Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils, Water Resour. Res., 32[6], 1597– 1610.

Harvey, C. and S. M. Gorelick [1995], Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity, *Water Resour. Res.*, 31[7], 1615-1626.

Harvey, C. F., and S. M. Gorelick [1995], Temporal moment-generating equations: Modeling transport and mass transfer in heterogeneous aquifers, *Water Resour. Res.*, 31[8], 1895-1912, 10.1029/95WR01231.

Hassanizadeh, S. M. and J. Carrera [1992], Editorial, *Adv. Water Resour*., 15[1]: 1-3.

Heinz, J., S. Kleineidam, G. Teutsch, and T. Aigner [2003], Heterogeneity patterns of Quaternary glaciofluvial gravel bodies [SW-Germany]: Application to hydrogeology, *Sediment. Geol*., 158, 1–23, doi:10.1016/S0037-0738[02]00239-7.

Hess, K. M., S. H. Wolf, and M. A. Celia [1992], Large-Scale Natural Gradient Tracer Test in Sand and Gravel, Cape Cod, Massachusetts 3. Hydraulic Conductivity Variability and Calculated Macrodispersivities, *Water Resour. Res*., 28[8], 2011–2027.

Hoeksema, R. J., and P. K. Kitanidis [1984], An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, *Water Resour. Res.*, 20[7], 1003-1020.

Hoeksema, R. J., and P. K. Kitanidis [1989], Prediction of transmissivities, heads, and seepage velocities using mathematical modeling and geostatistics, *Adv. Water Resour.*, 12, 90-101.

Hubbard, S. S., and Y. Rubin [2000], Hydrogeological parameter estimation using geophysical data: A review of selected techniques, *J. Contam. Hydrol*., 45, 3 –34, 2000.

Hufschmeid, P. [1986], Estimation of three-dimensional statistically anisotropic hydraulic conductivity field by means of single well pumping tests combined with flowmeter measurements, *Hydrogeologie*, 2, 163-174.

Hughson, D. L., and T.-C. J. Yeh [2000], An Inverse Model for Three-Dimensional Flow in Variably Saturated Porous Media, *Water Resour. Res.*, *36[4]*, 829–839.

Hyndman, D. W., and S. M. Gorelick [1996], Estimating lithological transport properties in three dimensions using seismic and tracer data: The Kesterton Aquifer, *Water Resour. Res*., 32[9], 2659– 2670.

Illangasekare, T. H., E. J. Armbruster, and D. N. Yates [1995], Non-aqueous phase fluids in heterogeneous aquifers – Experimental study, *ASCE J. of Env. Eng*., 121[8], 571-579.

Illman, W.A., 1999. Single- and cross-hole pneumatic injection tests in unsaturated fractured tuffs at the Apache Leap Research Site near Superior, Arizona, Ph.D. dissertation, Department of Hydrol. and Water Resour., University of Arizona, Tucson.

Illman, W. A. [2005], Type curve analyses of pneumatic single-hole tests in unsaturated fractured tuff: Direct evidence for a porosity scale effect, *Water Resour. Res*., 41, W04018, doi:10.1029/2004WR003703.

Illman, W. A. [2006], Strong field evidence of directional permeability scale effect in fractured rock, *Journal of Hydrology*, 319 [1-4], 227-236.

Illman, W. A., and S. P. Neuman [2000], Type-curve interpretation of multirate single-hole pneumatic injection tests in unsaturated fractured rock, *Ground Water*, 38[6], 899–911, doi:10.1111/j.1745-6584.2000.tb00690.x.

Illman, W. A., and S. P. Neuman [2001], Type-curve interpretation of a cross-hole pneumatic test in unsaturated fractured tuff, *Water Resour. Res*., 37[3], 583–604, doi:10.1029/2000WR900273.

Illman, W. A. and S. P. Neuman [2003], Steady-state analyses of cross-hole pneumatic injection tests in unsaturated fractured tuff, *Journal of Hydrology*, 281[1-2], 36-54.

Illman, W. A., and D. M. Tartakovsky [2006], Asymptotic analysis of cross-hole hydraulic tests in fractured granite, *Ground Water*, 44[4], 555–563, doi:10.1111/j.1745-6584.2006.00201.x.

Illman, W.A., Thompson, D.L., Vesselinov, V.V., Chen, G., Neuman, S.P., 1998. Single- and Cross-Hole Pneumatic Tests in Unsaturated Fractured Tuffs at the Apache Leap Research Site: Phenomenology, Spatial Variability, Connectivity and Scale, NUREG/CR-5559, U.S. Nucl. Regul. Comm., Washington D.C.

Illman, W. A., X. Liu, and A. Craig [2007], Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms, *J. of Hydrol.*, 341[3-4], 222-234.

Illman, W. A., A. J. Craig, and X. Liu, [2008], Practical issues in imaging hydraulic conductivity through hydraulic tomography, *Ground Water*, 46[1], 120-132.

Illman, W. A.; Alvarez, P. J. [2009] Performance assessment of bioremediation and natural attenuation, *Critical Reviews in Environmental Science and Technology*, *39*[4], 209-270, doi: 10.1080/10643380701413385.

Illman, W. A., X. Liu, S. Takeuchi, T. J. Yeh, K. Ando, and H. Saegusa [2009], Hydraulic tomography in fractured granite: Mizunami Underground Research site, Japan, *Water Resour. Res*., 45, W01406, doi:10.1029/2007WR006715.

Illman, W. A., J. Zhu, A. J. Craig, and D. Yin [2010a], Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study, Water Resour. Res., 46, W04502, doi:10.1029/2009WR007745.

Illman, W. A., S. J. Berg, X. Liu, A. J. Craig and A. Massi [2010b], Hydraulic/partitioning tracer tomography for trichloroethene source zone characterization: Small-scale sandbox experiments, Environ. Sci. Technol., 44[22], pp. 8609-8614, doi: 10.1021/es101654j.

James, A., W. Graham, K. Hatfield, P. Rao, and M. Annable [1997], Optimal Estimation of Residual Non–Aqueous Phase Liquid Saturations Using Partitioning Tracer Concentration Data, Water Resour. Res., 33[12], 2621-2636.

James, A. I., W. D. Graham, K. Hatfield, P. S. C. Rao, and M. D. Annable [2000], Estimation of Spatially Variable Residual Nonaqueous Phase Liquid Saturations in Nonuniform Flow Fields Using Partitioning Tracer Data, Water Resour. Res., 36[4], 999–1012.

Jankovic, I., A. Fiori and G. Dagan, [2003], Effective conductivity of an isotropic heterogeneous medium of lognormal conductivity distribution, *Multiscale Model. Simul*., 1[1], 40-56.

Jawitz, J.W., M. D. Annable, G. G. Demmy, and P. S. C. Rao, Estimating nonaqueous phase liquid spatial variability using partitioning tracer higher temporal moments, *Water Resour. Res.*, 39[7], 1192, doi:10.1029/2002WR001309, 2003.

Jin, M., M. Delshad, V. Dwarakanath, D. C. McKinney, G. A. Pope, K. Sepehrnoori, and C. E. Tilburg [1995], Partitioning tracer test for detection, estimation, and remediation performance assessment of subsurface nonaqueous phase liquids, *Water Resour. Res*., *31*[5], 1201–1211.

Jose, S. C., M. A. Rahman, and O. A. Cirpka [2004], Large-scale sandbox experiment on longitudinal effective dispersion in heterogeneous porous media, *Water Resour. Res.*, *40*, W12415, doi:10.1029/2004WR003363.

Journel, A. G. [1983], Nonparametric estimation of spatial distribution. *Math. Geol.* 15[3]: 445-468.

Journel, A. G. and F. G. Alabert [1990]. New method for reservoir mapping., *J. Pet. Technol.* February 42[2]:212-218.

Journel, A. G. and J. Gomez-Hernandez [1993]. Stochastic imaging of the Wilmington clastic sequence. *Soc. Pet. Eng. Form. Eval*. March 8[1]: 33-40.

Journel, A. G. and E. K. Isaacs [1984]. Conditional indicator simulation: application to a Saskatcheouan uranium deposit. *Math. Geol.* 16]7]:685-718.

Kabala, Z. J. [1993], The dipole flow test: A new single-borehole test for aquifer characterization, *Water Resour. Res*., 29[1], 99–107.

Khachikian, C., and T. C. Harmon [2000], Nonaqueous phase liquid dissolution in porous Media: Current state of knowledge and research needs, Transp. Porous Media, 38, 3 –28.

Kitanidis, P. K., and E. G. Vomvoris [1983], A geostatistical approach to the inverse problem in groundwater modeling and one-dimensional simulations, *Water Resour. Res.*, 19[3], 677--690.

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

Kitanidis, P. K. 1995. Quasi-linear geostatistical theory for inversing, Water Resour. Res., 31[10], 2411-2420, 10.1029/95WR01945.

Klingbeil, R., S. Kleineidam, U. Asprion, T. Aigner, and G. Teutsch [1999], Relating lithofacies to hydrofacies: Outcrop-based hydrogeological characterization of Quaternary gravel deposits, *Sediment. Geol*., 129, 299– 310, doi:10.1016/S0037-0738[99]00067-6.

Klise, K. A., G. S. Weissmann, S. A. McKenna, E. M. Nichols, J. D. Frechette, T. F. Wawrzyniec, and V. C. Tidwell [2009], Exploring solute transport and streamline connectivity using lidar-based outcrop images and geostatistical representations of heterogeneity, *Water Resour. Res*., 45, W05413, doi:10.1029/2008WR007500.

Klute, A., and C. Dirksen [1986], Hydraulic conductivity and diffusivity: Laboratory methods, in *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, 2nd ed., edited by A. Klute, chap. 28, 687–734, Am. Soc. of Agron., Madison, Wisc.

Koltermann, C. E., and S. M. Gorelick [1992], Paleoclimatic signature in terrestrial flood deposits. *Science* 256:1775-1782.

Koltermann, C. E., and S. M. Gorelick [1996], Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches, *Water Resour. Res.*, *32*[9], 2617–2658.

Konikow, L., 1978. Calibration of ground-water models. In: Verification of Mathematical and Physical Models in Hydraulic Engineering. American Society of Civil Engineers, New York, pp. 87–93.

Konikow L. F. and J. D. Bredehoeft [1992], Ground-water models cannot be validated, *Adv. Water Resour*., 15[1]: 75-83.

Kram, M. L.; Keller, A. A.; Rossabi, J.; Everett, L. G. [2001] DNAPL Characterization Methods and Approaches, Part 1: Performance Comparisons, *Groundwater Monitoring Review*, 109-123.

Kroszynski, U. I., and G. Dagan [1975], Well pumping in unconfined aquifers: The influence of the unsaturated zone, *Water Resour. Res*., 11, 479–490, doi:10.1029/WR011i003p00479.

Kuhlman, K. L, A. C. Hinnell, P. K. Mishra, T-C. J. Yeh. 2008, Basin-scale transmissivity and storativity estimation using hydraulic tomography. Ground Water. Sep-Oct;46[5]:706-15.

Lakshminarayana, V. and S. P. Rajagopalan [1978], Type-curve analysis of time-drawdown data for partially penetrating wells in unconfined anisotropic aquifers, *Ground Water*, *16*, 328-333.

Langtangen, H. P. [2003], Computational Partial Differential Equations; Numerical Methods and Diffpack Programming, 2nd edition, volume 1 in Texts in Computational Science and Engineering, Springer Verlag.

LeBlanc, D. R., S. P. Garabedian, K. M. Hess, L. W. Gelhar, R. D. Quadri,vK. G. Stollenwerk, and W. W. Wood [1991], Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts: 1. Experimental design and observed tracer movement, Water Resour. Res., 27[5], 895– 910, doi:10.1029/91WR00241.

Lee, D. M., W. D. Reynolds, and D. E. Elrick [1985], A comparison of 3 field methods for measuring saturated hydraulic conductivity. *Canadian Journal of Soil Science*, 65[3]: 563-573.

Levy M, Berkowitz B. Measurement and analysis of non-Fickian dispersion in heterogeneous porous media. J Contam Hydrol 2003;64[3–4]:203–26.

Li, B., T.-C. J. Yeh [1998], Sensitivity and moment analyses of head in variably saturated regimes, *Adv. Water Res.*, *21*, 477–485, 1998.

Li, B., and T.-C. J. Yeh [1999], Cokriging estimation of the conductivity field under variably saturated flow conditions, *Water Resour. Res.*, *35[12]*, 3663-3674.

Li, W., W. Nowak, and O. A. Cirpka [2005], Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown, *Water Resour. Res.*, *41*, W08403, doi:10.1029/2004WR003874.

Li, W., A. Englert, O. A. Cirpka, J.Vanderborght, and H.Vereecken [2007],Two dimensional characterization of hydraulic heterogeneity by multiple pumping tests, Water Resour. Res., 43, W04433, doi:10.1029/2006WR005333.

Li, W., A. Englert, O. A. Cirpka, and H. Vereecken [2008], Three-dimensional geostatistical inversion of flowmeter and pumping test data, *Ground Water*, 46[2], 193 – 201.

Liu, G., J. J. Butler Jr., G. C. Bohling, E. Reboulet, S. Knobbe, and D. W. Hyndman [2009], A new method for high-resolution characterization of hydraulic conductivity, *Water Resour. Res*., 45, W08202, doi:10.1029/2009WR008319.

Liu, S., T. -C. J. Yeh, and R. Gardiner [2002], Effectiveness of hydraulic tomography: sandbox experiments. *Water Resour. Res.* 38[4]: 10.1029/2001WR000338.

Liu S., and T –C. J. Yeh [2004], An Integrative Approach for Monitoring Water Movement in the Vadose Zone , *Vadose Zone J.* 3: 681-692.

Liu, X., and P. K. Kitanidis [2011], Large-Scale Inverse Modeling with an Application in Hydraulic Tomography, *Water Resour. Res*., doi:10.1029/2010WR009144, in press.

Liu, X., W. A. Illman, A. J. Craig, J. Zhu, and T.-C. J. Yeh [2007], Laboratory sandbox validation of transient hydraulic tomography, *Water Resour. Res.*, *43*, W05404, doi:10.1029/2006WR005144.

Long, J.C.S., P. Gilmour, and P.A. Witherspoon [1985], A model for steady fluid flow in random 3-D network of disc shaped fractures, *Water Resour. Res*., *21*[8], 1105-1115.

Luo, J., O. A. Cirpka, and P. K. Kitanidis [2006], Temporal-moment matching for truncated breakthrough curves for step or step-pulse injection, *Adv. Water Resour.*, 29, 1306-1313.

Mackay, D. M., D. L. Freyberg, P. V. Roberts, and J. A. Cherry [1986], A Natural Gradient Experiment on Solute Transport in a Sand Aquifer 1. Approach and Overview of Plume Movement, Water Resour. Res., 22[13], 2017–2029.

Mallat, S. [1999], A wavelet tour of signal processing, ISBN : 0-12-466606-X, pp577, Academic Press.

Mao, D., L. Wan, T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen [2011], A revisit of drawdown behavior during pumping in unconfined aquifers, *Water Resour. Res*., [in review].

Marquardt, D. W. [1963], An algorithm for least squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Math., 2, 431–441.

Martinez-Landa, L., and J. Carrera [2005], An analysis of hydraulic conductivity scale effects in granite [Full-scale Engineered Barrier Experiment [FEBEX], Grimsel, Switzerland], *Water Resour. Res*., 41, W03006, doi:10.1029/2004WR003458.

Mas-Pla, J., T.-C. J. Yeh, J. F. McCarthy, and T. M. Williams [1992], Numerical simulation of a two-well tracer experiment, Ground Water, 30[6], 958– 964.

Mas-Pla, J., T.-C. J. Yeh, J. F. McCarthy, and T. M. Williams [1992], A forced gradient tracer experiment in a coastal sandy aquifer, Georgetown site, South Carolina, Ground Water, 30, 958–964.

Mathias, S. A., and A. P. Butler [2006], Linearized Richards' equation approach to pumping test analysis in compressible aquifers, *Water Resour. Res*., 42, W06408, doi:10.1029/2005WR004680.

McCarthy, J., B. Gu, L. Liang, J. Mas-Pla, T. Williams, and T.-C. Yeh [1996], Field Tracer Tests on the Mobility of Natural Organic Matter in a Sandy Aquifer, Water Resour. Res., 32[5], 1223-1238.

McDermott C. I., M. Sauter, and R. Liedl [2003], New experimental techniques for pneumatic tomographical determination of the flow and transport parameters of highly fractured porous rock samples, *J. of Hydrol.*, *278* [1-4], 51-63.

McLaughlin, D., and L. R. Townley [1996], A reassessment of the groundwater inverse problem, Water Resour. Res., 32[5], 1131– 1162.

Meerschaert, M. M., D. A. Benson, and B. Baumer [1999], Multidimensional advection and fractional dispersion. Phys. Rev. E., 59[5]: 5026–8.

Meier, P. M., J. Carrera, and X. Sánchez-Vila [1998], An evaluation of Jacob's method for interpretation of pumping tests in heterogeneous formations, *Water Resour. Res*., 34[5], 1011– 1025.

Mishra, P. K., and S. P. Neuman [2010], Improved forward and inverse analyses of saturated-unsaturated flow toward a well in a compressible unconfined aquifer, *Water Resour. Res*., 46, W07508, doi:10.1029/2009WR008899.

Mizell, S. A., L. W. Gelhar, and A. L. Gutjahr [1980], Stochastic analysis of spatial variability in two-dimensional groundwater flow with implications for observation-well-network design, Rep. H-6, Geophys. Res. Cent., N. M. Inst. of Min. and Technol., Socorro.

Mizell, S., A. Gutjahr, and L. Gelhar [1982], Stochastic Analysis of Spatial Variability in Two-Dimensional Steady Groundwater Flow Assuming Stationary and Nonstationary Heads, *Water Resour. Res.*, *18[4]*, 1053-1067.

Moench, A. F. [1997], Flow to a well of finite diameter in a homogeneous, anisotropic water table aquifer, *Water Resour. Res*., 33, 1397–1407, doi:10.1029/97WR00651.

Moench, A.F., and S.P. Garabedian, and D.R. LeBlanc [2001], Estimation of hydraulic parameters from an unconfined aquifer test conducted in a glacial outwash deposit, Cape Cod, Massachusetts. US Geological Survey Professional Paper 1629.

Moench, A.F. [2003], Estimation of hectare-scale soil-moisture characteristics from aquifer-test data. *Journal of Hydrology* 281, 82–95.

Moench, A.F. [2004], Importance of the vadose zone in analyses of unconfined aquifer tests. *Ground Water* 42, 223–233.

Moench, A. F. [2008], Analytical model and numerical analyses of unconfined aquifer test considering unsaturated zone characteristics, *Water Resour. Res.*, *44*, W06409, doi: 10.1029/2006WR005736.

Molz, F.J., Morin, R.H., Hess, A.E., Melville, J.G., and O. Guven [1989], The impeller meter for measuring aquifer permeability variations: Evaluation and comparison with other tests, *Water Resour. Res*., 25[7], 1677-1683.

Moreno-Barbero, E.; Illangasekare, T. H. [2006], Influence of dense nonaqueous phase liquid pool morphology on the performance of partitioning tracer tests: Evaluation of the equilibrium assumption, *Water Resour. Res*., 42, W04408, doi:10.1029/2005WR004074.

Mualem, Y. [1976], A new model for predicting the hydraulic conductivity of unsaturated porous media, *Water. Resour. Res.*, *12*.

Narasimhan, T. N., and M. Zhu [1993], Transient flow of water to a well in an unconfined aquifer: Applicability of some conceptual models, *Water Resour. Res.*, *29*, 179-191, doi: 10.1029/92WR01959.

Narasimhan T.N. [2007], Comment on 'Pumping-induced vadose zone drainage and storage in an unconfined aquifer: A comparison of analytical model predictions and field measurements', *Journal of Hydrology*, 335 [1-2], 219-220.

National Research Council [2004], *Contaminants* in the Subsurface: Source Zone Assessment and Remediation, *The National Academies Press*, Washington, DC.

Nelson, N.T.; Brusseau, M. L. [1996], Field study of the partitioning tracer method for detection of dense nonaqueous phase liquid in a trichloroethene-contaminated aquifer, *Environmental Science & Technology*, 30, 2859-2863.

Neuman, S. P. [1972], Theory of flow in unconfined aquifers considering delayed response of the water table, *Water Resour. Res.*, *8*, 1031-1044, doi: 10.1029/WR008i004p01031.

Neuman, S. P. [1974], Effects of partial penetration on flow in unconfined aquifers considering delayed aquifer response, *Water Resour. Res.*, 10, 303–312, doi:10.1029/WR010i002p00303.

Neuman, S. P. [1975], Analysis of pumping test data From anisotropic unconfined aquifers considering delayed gravity response, *Water Resour. Res.*, 11[2], 329–342.

Neuman, S. P. [1987a], Stochastic continuum representation of fractured rock permeability as an alternative to the REV and fracture network concepts, pp. 533-561, Rock Mechanics: Proceedings of the 28th U.S. Symposium, Tucson, Arizona, edited by I.W. Farmer, J.J.K. Daemen, C.S. Desai, C.E. Glass, and S.P. Neuman, A.A. Balkema, Rotterdam, 1240 p.

Neuman, S. P. [1987b], On methods of determining specific yield, *Ground Water*, 25, 679-684.

Neuman, S. P., C. L. Winter, and C. M. Newman [1987], Stochastic theory of field-scale Fickian dispersion in anisotropic porous media, Water Resour. Res., 23[3], 453– 466, doi:10.1029/WR023i003p00453.

Neuman, S. P. [1993], Eulerian–Lagrangian theory of transport in space–time nonstationary velocity fields: exact nonlocal formalism by conditional moments and weak approximation. Water Resour Res., 29[3]: 633–45.

Neuman, S. P. [2005], Trends, prospects and challenges in quantifying flow and transport through fractured rocks, *Hydrogeol. J.*, 13: 124-147.

Neuman, S. P. and D. M. Tartakvosky [2009], Perspective on theories of non-Fickian transport in heterogeneous media, *Adv. Water Resour*. 32[?], 670-680.

Ni, Chuen-Fa, T.-C. J. Yeh, and J.-S. Chen [2009], Cost-effective hydraulic tomography surveys for predicting flow and transport in heterogeneous aquifers, Environ. Sci. Technol., 43[10], 3720-3727.

Nowak, W., and O. A. Cirpka [2004], A modified Levenberg-Marquardt algorithm for quasi-linear geostatistical inversing, Adv. Water Resour., 27[7], 737– 750.

Nwankwor G.I., J.A. Cherry, and R.W. Gillham [1984], A comparative study of specific yield determinations for a shallow sand aquifer, *Ground Water*, *22*, 764–772.

Nwankwor, G. I., R. W. Gillham, G. van der Kamp, and F.F. Akindunni [1992], Effect of the capillary fringe during pumping of unconfined aquifers: Field evidence of delayed drainage, *Ground Water*, *30[5]*, 690-700.

O'Connor, D. R. [2002], Report of the Walkerton Inquiry, Ont. Minist. Of the Attorney Gen., Toronto, Ont., Canada.

Oldenborger, G. A., R. A. Schincariol, and L. Mansinha [2002], Spacelocal spectral texture segmentation applied to characterizing the heterogeneity of hydraulic conductivity, *Water Resour. Res.*, 38[8], 1154, doi:10.1029/2001WR000496.

Oliver, D. S. [1993], The influence of nonuniform transmissivity and storativity on drawdown,

Water Resour. Res., 29[1], 169– 178.
Oostrom, M.; Hofstee, C.; Walker, R. C.; Dane, J. H. [1999], Movement and remediation of trichloroethylene in a saturated heterogeneous porous medium. 1. Spill behavior and initial dissolution, *Journal of Contaminant Hydrology*, *37*, 159-178.
Oreskes, N., K. Shraderfrechette, and K. Belitz [1994], Verification, Validation, and confirmation of numerical-models in the earth-sciences, *Science*, 263 [5147]: 641-646.
Ozisik, M.N. [1993], Heat Conduction, 2nd edition, John Wiley & Sons, New York.
Painter, S. L. A. D. Woodbury, and Y. Jiang [2007], Transmissivity estimation for highly heterogeneous aquifers: comparison of three methods applied to the Edwards Aquifer, Texas, USA, *Hydrogeol. J.*, 15: 315-331.
Pankow, J. F.; Cherry J. A., [1996], *Dense Chlorinated Solvents and other DNAPLs in Ground water: History, Behavior, and Remediation*, Waterloo Press, **1996**, 522 pp.
Poeter, E. P. and D. R. Anderson [2005], Multi-model ranking and inference in ground water modeling, *Ground Water*, 43[4]: 597–605.
Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Numerical Recipes in Fortran 77: The Art of scientific Computing. Cambridge University Press. New York. pp933.
Prickett, T. A. [1965], Type-curve solution to aquifer tests under water-table conditions. *Ground Water*, 3, 3,5-14.
Rao, P. S. C., M. D. Annable, and H. Kim [2000], NAPL source zone characterization and remediation technology performance assessment: recent developments and applications of tracer techniques, J. Contam. Hydrol., 45, 63– 78.
Refsgaard, J. C., and H. J. Henriksen [2004], Modelling guidelines-terminology and guiding principles, *Adv. Water Resour*., 27[1], 71-82.
Refsgaard, J. C. [1997], Parameterisation, calibration and validation of distributed hydrological models, *J. Hydrol*., 198 [1-4]: 69-97.
Refsgaard, J. C., and J. Knudsen [1996], Operational validation and intercomparison of different types of hydrological models, *Water Resour. Res.*, *32*[7], 2189–2202.
Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar [1992], Field study of dispersion in a heterogeneous aquifer, 3. Geostatistical analysis of hydraulic conductivity, *Water Resour. Res*., 28[12] 3309-3324.
Renard, P. and G. De Marsily [1997], Calculating equivalent permeability: a review, *Adv. Water Resour*., 20[5-6] 253-278.
Renshaw, C. E. [1996], Estimation of fracture zone geometry from steadystate hydraulic head data using iterative sequential cokriging, Geophys. Res. Lett., 23[19], 2685– 2688.
Roy, W., J. E. Smith, R. W. Gillham [2004], Laboratory evidence of natural remobilization of multicomponent DNAPL pools due to dissolution, Journal of Contaminant Hydrology, 74[1-4], 145-161.
Rubin, Y., and G. Dagan [1987], Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1, Theory, *Water Resour. Res.*, 23[7], 1185-1192.
Russo, D., and M. Bouton [1992], Statistical analysis of spatial variability in unsaturated flow parameters, *Water Resour. Res.*, 28, 1911-1925,
Sakaki, T., C. C. Frippiat, M. Komatsu, and T. H. Illangasekare [2009], On the value of lithofacies data for improving groundwater flow model accuracy in a three-dimensional laboratory-scale synthetic aquifer, *Water Resour. Res.*, 45, W11404, doi:10.1029/2008WR007229.

Sánchez-Vila, X., P. M. Meier, and J. Carrera [1999], Pumping test in heterogeneous aquifers: An analytical study of what can be obtained from their interpretation using Jacob's method, *Water Resour. Res*., 35[4], 943– 952.

Scher H. and M. Lax [1973], Stochastic transport in a disordered solid: 1. Theory. Phys. Rev. B., 7:4491–502.

Schincariol, R., and F. Schwartz [1990], An Experimental Investigation of Variable Density Flow and Mixing in Homogeneous and Heterogeneous Media, *Water Resour. Res*., 26[10], 2317-2329.

Sciortino, A.; Harmon, T. C.; Yeh, W. W.-G. [2000] Inverse modeling for locating dense nonaqueous pools in groundwater under steady flow conditions, *Water Resour. Res*., *36*[7], 1723– 1736.

Shepherd, R.G. [1989], Correlations of permeability and grain size. *Ground Water*, 27[5]: 633-638.

Sillan, R. K., M. D. Annable, P. C. C. Rao, D. P. Dai, K. Hatfield, W. D. Graham, A. L. Wood, and C. G. Enfield [1998], Evaluation of in situ cosolvent flushing dynamics using a network of multi-level samplers, Water Resour. Res., 34[9], 2191–2202.

Silliman, S. E., and E. S. Simpson [1987], Laboratory Evidence of the Scale Effect in Dispersion of Solutes in Porous Media, *Water Resour. Res*., 23[8], 1667–1673.

Silliman, S. E., L. Zheng and P. Conwell [1998], The use of laboratory experiments for the study of conservative solute transport in heterogeneous porous media, *Hydrogeol. J*., 6[1], 166-177.

Silliman S. E. [2001], Laboratory study of chemical transport to wells within heterogeneous porous media, *Water Resour. Res*., 37[7], 1883-1892.

Srivastava, R., and T.-C. J. Yeh [1992], A three-dimensional numerical model for water flow and transport of chemically reactive solute through porous media under variably saturated conditions, *Adv. Water Resour.*, *15*[5], 275–287.

Stephens, D.B. [1995], *Vadose Zone Hydrology*. Boca Raton, Florida: CRC Press.

Straface S., T.-C. J. Yeh, J. Zhu, S. Troisi, C. H. Lee [2007], Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy, *Water Resour. Res*., 43, W07432, doi:10.1029/2006WR005287.

Strebelle S [2002] Conditional simulation of complex geological structures using multiple-point statistics. Math Geol 34:1–22.

Streltsova T. D. [1972a], Unsteady radial flow in an unconfined aquifer, *Water Resour. Res.*, *8[4]*, 1059.

Streltsova, T.D [1972b], Unconfined aquifer and slow drainage, *J. Hydrol.*, *16*, 117-123.

Streltsova, T. [1973], Flow near a Pumped Well in an Unconfined Aquifer under Nonsteady Conditions, *Water Resour. Res*., 9[1], 227-235.

Sudicky EA, Cherry JA, Frind EO [1983]. Migration of contaminants in groundwater at a landfill: a case study. 4. A natural-gradient dispersion test. J Hydrol., 63[1–2]:81–108.

Sudicky, E. A. [1986], A Natural Gradient Experiment on Solute Transport in a Sand Aquifer: Spatial Variability of Hydraulic Conductivity and Its Role in the Dispersion Process, *Water Resour. Res*., 22[13], 2069–2082.

Sudicky, E. A., W. A. Illman, I. K. Goltz, J. J. Adams, and R. G. McLaren [2010], Heterogeneity in hydraulic conductivity and its role on the macroscale transport of a solute plume: From measurements to a practical application of stochastic flow and transport theory, *Water Resour. Res*., 46, W01508, doi:10.1029/2008WR007558.

Sun, N. –Z. [1994], Inverse Problems in Groundwater Modeling, pp. 103-140, Springer, New York.

Sun, N.-Z., and W. W.-G. Yeh [1992], A stochastic inverse solution for transient groundwater flow: parameter identification and reliability analysis, *Water Resour. Res.*, 28[12], 3269-3280.

Sykes, J.-F., J. L. Wilson, and R. W. Andrews [1985], Sensitivity analysis of steady state groundwater flow using adjoint operators, *Water Resour. Res.*, *21[3]*, 359–371.

Tartakovsky, D. M. and S. P. Neuman [1998], Transient effective hydraulic conductivities under slowly and rapidly varying mean gradients in bounded three-dimensional random media, *Water Resour. Res.*, 34[1], 21-32.

Tartakovsky, G. D., and S. P. Neuman [2007], Three-dimensional saturated-unsaturated flow with axial symmetry to a partially penetrating well in a compressible unconfined aquifer, *Water Resour. Res*., *43*, W01410, doi: 10.1029/2006WR005153.

Teles V., J. P. Bravard, G. de Marsily, and E. Perrier [2001], Modelling of the construction of the Rhône alluvial plain since 15,000 years BP. *Sedimentology* 48: 1209-1224.

Teles V., F. Delay, and G. de Marsily [2004], Comparison between different methods for characterizing the heterogeneity of alluvial media: groundwater flow and transport simulations. *J. Hydrol.* 294[1-3]: 103-121.

Theis, C.V. [1935], The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage, *Am. Geophys. Union Trans.*, *16*, 519-524.

Tiedeman, C. R., and P. A. Hsieh [2004], Evaluation of longitudinal dispersivity estimates from simulated forced- and natural-gradient tracer tests in heterogeneous aquifers, *Water Resour. Res*., 40, W01512, doi:10.1029/2003WR002401.

Tikhonov, A. N., and V. Y. Arsenin [1977], Solutions of Ill-Posed Problems, John Wiley, Hoboken, N. J.

Tosaka, H., K. Masumoto, and K. Kojima [1993], Hydropulse tomography for identifying 3-D permeability distribution, in *Proceedings of the 4th Annual International Conference on High Level Radioactive Waste Management*.

Townley LR, Wilson JL. Computationally efficient algorithms for parameter estimation and uncertainty propagation in numerical models of groundwater flow. Water Resour Res 1985;21[12]:1851–60.

van Genuchten M.Th. [1980], A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Sci. Soc. Am. J.*, *44*, 892-898.

Vargas-Guzman, A. J., and T.-C. J. Yeh [1999], Sequential kriging and cokriging: Two powerful geostatistical approaches, Stochastic Environ. Res. Risk Assess., 13, 416– 435.

Vargas-Guzman, A. J., and T.-C. J. Yeh [2002], The successive linear estimator: A revisit, Adv. Water Resour., 25, 773–781.

Vasco, D.W., H. Keers, and K. Karasaki [2000], Estimation of reservoir properties using transient pressure data: An asymptotic approach, *Water Resour. Res.*, *36 [12]*, 3447-3465.

Vesselinov, V. V., S. P. Neuman, and W. A. Illman [2001a], Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 1. Methodology and borehole effects, *Water Resour. Res.*, 37[12], 3001-3018.

Vesselinov, V. V., S. P. Neuman, and W. A. Illman [2001], Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff: 2. Equivalent

parameters, high-resolution stochastic imaging and scale effects, *Water Resour. Res.*, 37[12], 3019-3042.

Walkerton Inquiry [2002], Queen's Printer for Ontario, ISBN 0-7794-2559-6.

Wang, P., Searakanath, V. Rouse, B. A. Pope, G. A. Sepehrnoori, K. [1997], Partition coefficients for alcohol tracers between nonaqueous-phase liquids and water from UNIFAC-solubility method, Adv. Water Resour., *21*[2], 171– 181.

Webb, E. K. [1995], Simulation of braided channel topology and topography, *Water Resour. Res.* 31[10], 2603-2612.

Webb, E. K. and M. P. Anderson [1996], Simulation of preferential flow in three-dimensional, heterogeneous conductivity fields with realistic internal architecture, *Water Resour. Res.*, 32[3], 533-546

Welty, C. and M. M. Elsner [1997], Constructing correlated random fields in the laboratory for observations of fluid flow and mass transport, *J. of Hydrol.*, 202[1-4], 192-211.

Weissmann, G., S. Carle, and G. Fogg [1999], Three-Dimensional Hydrofacies Modeling Based on Soil Surveys and Transition Probability Geostatistics, *Water Resour. Res*., 35[6], 1761-1770.

Wen, J.C., C.M Wu, T.-C. J. Yeh and C.M. Tseng [2010], Estimation of effective aquifer hydraulic properties from an aquifer test with multi-well observations, *Hydrogeology Journal*, DOI: 10.1007/s10040-010-0577-1.

Whittaker, J., and G. Teutsch [1999], Numerical simulation of subsurface characterization methods: Application to a natural aquifer analogue, *Adv. Water Resour*., 22[8], 819–829, doi:10.1016/S0309-1708[98]00056-6.

Willmann, M., J. Carrera, X. Sánchez-Vila, and E. Vázquez-Suñé [2007], On the meaning of the transmissivity values obtained from recovery tests, *Hydrogeology Journal*, 15[5], 833-842.

Wolf, S. H., M. A. Celia, and K. M. Hess [1992], Evaluation of hydraulic conductivities calculated from multiport-permeameter measurements, *Ground Water*, 29[4], 516-525.

Woodbury, A.D., and Sudicky, E.A. [1991], The geostatistical characteristics of the Borden aquifer, *Water Resour. Res.*, *27*, 533-546.

Wu, C.-M., T.-C. J. Yeh, J. Zhu, T. H. Lee, N.-S. Hsu, C.-H. Chen, and A. F. Sancho [2005], Traditional analysis of aquifer tests: Comparing apples to oranges?, *Water Resour. Res*., 41, W09402, doi:10.1029/2004WR003717.

Xiang, J., T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen [2009], A simultaneous successive linear estimator and a guide for hydraulic tomography analysis, *Water Resour. Res*., 45, W02432, doi:10.1029/2008WR007180.

Ye, M., R. Khaleel, and T.-C. J. Yeh [2005], Stochastic analysis of moisture plume dynamics of a field injection experiment, Water Resour. Res., 41, W03013, doi:10.1029/2004WR003735.

Yeh, T.-C. J., and D. J. Harvey [1990], Effective unsaturated hydraulic conductivity of layered sands, *Water Resour. Res.*, *26*, 1271–1279.

Yeh, T.-C.J., Stochastic modeling of groundwater flow and solute transport in aquifers, Journal of Hydrologic Processes, Vol. 6, p 369-395, 1992.

Yeh, T.-C. J. [1998], Scale issues of heterogeneity in vadose-zone hydrology, in Scale Dependence and Scale Invariance in Hydrology, edited by G. Sposito, Cambridge Univ. Press, New York.

Yeh, T.-C. J., R. Srivastava, A. Guzman, and T. Harter [1993], A numerical model for water flow and chemical transport in variably saturated porous media, *Ground Water*, 31[4], 634-644.

Yeh, T.-C. J., A.L. Gutjahr, and M. Jin [1995], An iterative cokriging-like technique for groundwater flow modeling, *Ground Water*, 33[l], 33-41.

Yeh, T.-C.J., J. Mas-Pla, T.M. Williams, and J.F. McCarthy [1995], Observation and three-dimensional simulation of chloride plumes in a sandy aquifer under forced-gradient conditions, *Water Resources Research*, 31[9], 2141-2157.

Yeh, T.-C., M. Jin, and S. Hanna [1996], An iterative stochastic inverse method: conditional effective transmissivity and hydraulic head fields, *Water Resour. Res.*, 32[l], 85-92.

Yeh, T.-C., and J. Zhang [1996], A geostatistical inverse method for variably saturated flow in the vadose zone, Water Resour. Res., 32[9], 2757–2766.

Yeh, T.-C. J., and S. Liu [2000], Hydraulic tomography: Development of a new aquifer test method, *Water Resour. Res.*, 36[8], 2095– 2105.

Yeh, T.-C. J., S. Liu, R. J. Glass, K. Baker, J. R. Brainard, D. Alumbaugh, and D. LaBrecque [2002], A geostatistically based inverse model for electrical resistivity surveys and its applications to vadose zone hydrology, *Water Resour. Res.*, 38[12], 1278, doi:10.1029/2001WR001204.

Yeh, T.-C. J., and J. Simunek [2002], Stochastic Fusion of Information for Characterizing and Monitoring the Vadose Zone, *Vadose Zone Journal*, Vol. 1, 207-221.

Yeh, T.-C. J., K. Hsu, C. Lee, J.Wen, and C. Ting [2004], On the possibility of using river stage tomography to characterize the aquifer properties of the Choshuishi Alluvial Fan, Taiwan, Eos Trans. AGU, 85[47], Fall Meet. Suppl., Abstract H11D-0329.

Yeh, T.-C. J., C. H. Lee, K-C Hsu, and Y-C Tan [2007], Fusion of Active and Passive Hydrologic and Geophysical Tomographic Surveys: The Future of Subsurface Characterization, in Data Integration in Subsurface Hydrology Edited by D. W. Hyndman, F. D. Day-Lewis, and K. Singha, AGU geophysical monograph 171.

Yeh, T.-C. J., and J. Zhu [2007], Hydraulic/partitioning tracer tomography for characterization of dense nonaqueous phase liquid source zones, *Water Resour. Res.*, 43, W06435, doi:10.1029/2006WR004877.

Yeh T.-C. J., C.-H. Lee [2007], Time to change the way we collect and analyze data for aquifer characterization. *Ground Water* 45[2]: 116-118.

Yeh, W. W.-G. [1986], Review of parameter-identification procedures in groundwater hydrology: The inverse problem, Water Resour. Res., 22[2], 95– 108.

Yin, D., and W. A. Illman [2009], Hydraulic tomography using temporal moments of drawdown recovery data: A laboratory sandbox study, *Water Resour. Res*., 45, W01502, doi:10.1029/2007WR006623.

Young, C.M.; Jackson, R.; Jin, M.; Londergan, J.T.; Mariner, P.E.; Pope, G.A.; Anderson, F.J.; Houk, T. [1999], Characterization of a TCE DNAPL zone in alluvium by partitioning tracers. *Groundwater Monitor. Remediat. 19*[1], 84–94.

Zhang D. and S. P. Neuman [1996], Effect of local dispersion on solute transport in randomly heterogeneous media. Water Resour. Res., 32[9]: 2715–23.

Zhang, Y. and W. D. Graham [2001], Partitioning tracer transport in a hydrogeochemically heterogeneous aquifer, *Water Resour. Res., 37*[8], p. 2037 [2000WR900378].

Zhang, Y. and Graham, W. D. [2001], Spatial characterization of a hydrogeochemically heterogeneous aquifer using partitioning tracers: Optimal estimation of aquifer parameters, *Water Resour. Res*., *37*[8], 2046-2063, doi: 10.1029/2000WR900377.

Zhang, Q, R. Aliaga-Rossel, and P Choi [2006], Denoising of gamma-ray signals by interval-dependent thresholds of wavelet analysis, Meas. Sci. Technol. 17, 731-735.

Zhang, J., and T.-C.J. Yeh [1997], An iterative geostatistical Inverse method for steady flow in the vadose zone, *Water Resour. Res.*, 33[1], 63-71.

Zheng, C. M., and S. M. Gorelick [2003], Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale, Ground Water, 41, 142– 155.

Zhu, J., and T.-C. J. Yeh [2005], Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resour. Res.*, *41*, W07028, doi:10.1029/2004WR003790.

Zhu, J., and T.-C. J. Yeh [2006], Analysis of hydraulic tomography using temporal moments of drawdown recovery data, *Water Resour. Res.*, *42*, W02403, doi:10.1029/2005WR004309.

Zhu, J., X. Cai, and T.-C. J. Yeh [2009], Analysis of tracer tomography using temporal moments of tracer breakthrough curves, *Adv. Water Resour*., 32, 391-400.

Zlotnik, V.A., Zurbuchen, B.R., and T. Ptak [2001], The steady-state dipole-flow test for characterization of hydraulic conductivity statistics in a highly permeable aquifer: Horkheimer Insel site, Germany, *Ground Water*, 39[4], 504-516.

Zlotnik, V.A. and B. R. Zurbuchen [2003], Field study of hydraulic conductivity in a heterogeneous aquifer: Comparison of single-borehole measurements using different instruments. *Water Resour. Res.*, 39[4], 1101, doi:10.1029/2002WR001415.

APPENDICES

Appendix A: Cost Estimate for Implementing the Hydraulic, Conservative Tracer and Partitioning Tracer Tomography Tests to Delineate DNAPL Source Zones in a Sandy Aquifer [submitted on 7/15/2004].

A.1. Summary of White Paper

This white paper addresses one concern raised by the by the review panel during the Spring In-Progress Review [IPR] Meeting held in Washington DC on April 14, 2004. The concern is related to the cost of implementing the technology at a representative DNAPL site and how the costs are affected by increases in aquifer depth, and volume of the source zone. We compare the resulting cost estimates to existing technology that will yield comparable information on DNAPL source zones.

A.2. Technology background

Dense Nonaqueous Phase Liquids [DNAPLs] are prevalent at a large number of sites throughout the world. The variable release history and geologic heterogeneity make the spatial distribution of DNAPLs in the source zone complex and make absolute cleanup difficult. A small amount of residual DNAPL can contribute to long-term groundwater contamination for decades to centuries. Therefore, the spatial distribution, mass, and composition of DNAPLs present in the source zone need to be characterized in great detail so that efficient remediation schemes can be designed. Development of a cost-effective technology that provides detailed images of DNAPL distributions in the source zones thus is essential.

The current technology for characterizing the DNAPL source zone heavily relies on invasive sampling conducted at point-scales. Although it is generally considered to be accurate at the sampling point if proper sample handling protocols are followed, for a source zone covering a large volume, densely distributed point samples are required to depict a detailed 3-D distribution of DNAPLs. Due to the very large number of samples that are required for proper characterization, current technology that relies on invasive sampling can be very costly and destructive if accurate site characterization is required for a large source zone. To overcome the shortcomings of current technology, our new technology exploits recently developed hydraulic/pneumatic tomography and expands the concept to the development of a conservative and partitioning tracer tomography to image the subsurface and DNAPL distributions over a large volume of geologic media without extensive invasive sampling. The technology uses a stochastic fusion of information methodology to assimilate data obtained from currently available indirect and direct methods of DNAPL source zone characterization [Kram et al., 2001] and results of hydraulic, conservative, and partitioning tracer tomography surveys to derive the best estimate of the DNAPL residual distribution and to quantify its uncertainty. As a result, the new technology can characterize a greater volume of geological media at higher resolution at lower cost. Specifically, it first analyzes the information derived from Hydraulic Tomography [HT] by incorporating point samples of hydraulic properties to identify three-dimensional hydraulic heterogeneity. It then improves the estimate of the heterogeneity by incorporating new information acquired from the Conservative Tracer Tomography [CTT]. Afterward, the improved hydraulic heterogeneity is used to simulate the HT such that more detailed information about the response of the subsurface becomes available. This new information again is fed back to the technique to update the estimate of the heterogeneity. The iterative process continues until

the available information and measurements are fully utilized for identifying the heterogeneity that controls the spatial distribution of DNAPLs. Upon completion, the newly derived knowledge of the heterogeneity are then used to design partitioning tracer tomography [PTT] tests and are combined with the hydraulic data and other prior, point-scale information on DNAPLs [and the site hydrology] to depict the 3-D spatial distribution of DNAPL residual saturation in the source zone.

The technology is a new generation technology that images the three-dimensional spatial distribution of hydraulic heterogeneity in the subsurface, which allows for the significantly improved characterization of DNAPL source zones compared to existing technologies listed in Kram et al. [2001]. This new technology also provides the means for real-time monitoring and forecasting of the fate of DNAPLs during remediation. It should be also be beneficial for characterization and monitoring efforts that rely on natural attenuation after a significant amount of DNAPLs have been removed after active remediation. Most importantly, it provides uncertainty estimates that can facilitate better decision making.

A.3. Costs to implement the technology at a representative field site

This white paper presents cost estimates for implementing the HT, CTT, and PTT technologies [assuming that the proposed technology has been validated and is ready for application] in the field by a typical engineering firm in the Northeast. In particular, we examine how the cost of this technology varies with the increasing volume of the DNAPL source zone and compare it against the cost of a baseline approach described in Kram et al. [2002]. Four representative scenarios of different source zone size [Case 1~10ft^3; Case 2~50 ft^3; Case 3~100 ft^3; and Case 4~150 ft^3] are used to illustrate this variation. We assume that the source zone has an area 100 ft^2 for case 1, 2,500 ft^2 for case 2, 10,000 ft^2 for case 3 and 22,500 ft^2 for case 4. Table A.1 summarizes the physical dimensions of the source zones considered.

Table A.1: Physical dimensions of source zone for the 4 cases examined.

	Depth [ft]	Source area [ft^2]	Source Volume [ft^3]
Case 1 [10×10×10]	10	100	1000
Case 2 [50×50×50]	50	2500	125,000
Case 3 [100×100×100]	100	10,000	1,000,000
Case 4 [150×150×150]	150	22,500	3,375,000

In all of the scenarios, we assume that there is a DNAPL source zone consisting of Trichloroethylene [TCE] in an unconfined aquifer consisting of unconsolidated sediment deposits. It is assumed that the site is not well characterized but water samples taken from available monitoring wells show TCE concentrations that suggest the presence of a source zone. Site hydraulic gradient data is known from a few existing monitoring wells installed already at the site.

As excavation of the contaminated soils could be costly considering its volume, and that it has been determined that an in situ treatment of the DNAPL source zone is feasible, a detailed characterization of the source zone is deemed necessary. The underlying clay unit is assumed to be thin and so drilling and disturbance of the source zone is not feasible as the TCE can further migrate downwards into the underlying aquifer which is utilized by this hypothetical Northeast city as the municipal water supply. Considering all the risks involved with the traditional

methods of DNAPL source zone characterization and following the outside-in approach to minimize spreading of contamination [Pankow and Cherry, 1996], it was decided to drill the fewest number of boreholes at the site. These site considerations and risks lead to the selection of HT, CTT, and PTT tests to delineate the source zone so that well-informed decisions can be made with regards to the selection of site remediation that will take place subsequent to the site characterization efforts. The major advantage behind this technique over existing techniques is that one will be able to obtain a 3-dimensional distribution of DNAPL residual saturation and its uncertainty estimates, with minimal invasive drillings. In addition, hydraulic and transport properties such as hydraulic conductivity, specific storage, effective porosity and dispersivity will also be obtained in three-dimensions [with its corresponding uncertainty estimates] making the subsequent selection of a remediation scheme and its design considerably more efficient. Uncertainty estimates of all parameters will help site managers and engineers make well-informed decisions about the next course of action.

For each of the 4 cases, we assume that 5 boreholes are drilled in a 5-spot pattern within the source zone and 25 injection and observation zones are utilized to conduct the HT, CTT, and PTT [Figure A.1]. Table A.2 summarizes the number of HT, CTT, and PTT tests that we estimate will be required to characterize the source zone. These numbers especially for the CTT and PTT tests are rough estimates as we do not have simulation results for the CTT and PTT tests at this time. Our estimates for the number of HT tests are based on our previous [Yeh and Liu, 2000] and more recent simulation results. Here we assume that for each HT test, 2 days are required from setup to completion. Likewise, for the CTT tests, we assume that it will take 3 days to complete each test. On the other hand, we assume that each PTT test will take 5 days to complete, accounting for the retardation of the partitioning tracers. These estimates are based on calculations of steady-state drawdown using the point source solution [Carlslaw and Jaeger, 1959] and travel times assuming an aquifer with an effective hydraulic conductivity of 86.4 ft/day [medium sand] and an effective porosity of 0.25. For example, in case 1, we assume that 5 HT tests, 2 CTT, and 2 PTT tests are conducted sequentially over a period of 26 days [1 day = 10 hours] with a two-man crew to characterize the source zone. We have also budgeted 80 hours for a numerical modeler who is assumed to be fully trained to conduct preliminary simulations and interpretation of test results.

Estimates of variable costs such as drilling, rental, and sample analyses are obtained from RS Means [1998]. As the cost estimates of various items in RS Means [1998] do not reflect the costs in 2004, the estimates are adjusted to reflect the 2004 costs with an inflation rate of 3% per year. We also use an estimate of disposal of drilling wastes obtained from Kram et al. [2002] which we adjust to reflect the costs in 2004. Estimates of project overhead and profit [multiplier = 3.1] are obtained from Unger and Coover [2004].

We note that equipment [such as packer assemblage, tracer tanks, data acquisition system, etc.] costs presented here are based on the assumption that the contractor will rent the necessary equipment for the duration of the field project. Rental of the equipment reduces the costs if the testing duration is short, but the costs can be higher if the test duration is long. We recommend that the contractor purchase the equipment for long duration projects.

We also note that the number of chemical analyses can be significantly reduced if the first sweep of PTT does not detect the presence of DNAPLs. Additional partitioning tracer tests are used to capture the three-dimensional distribution of DNAPLs if it is found to be present from the first PTT test. We note that the sample analysis cost for the tracers can be high [$175/sample for VOC/alcohol analysis] as a large number of samples need to be analyzed for

the CTT and PTT tests. However, this variable cost may be reduced considerably if the samples can be analyzed by the contractor instead of subcontracting the analysis. We also emphasize that a smaller number of CTT and PTT tests may be necessary if the hydraulic heterogeneity is quantified well by the HT tests, thus reducing the overall project cost substantially.

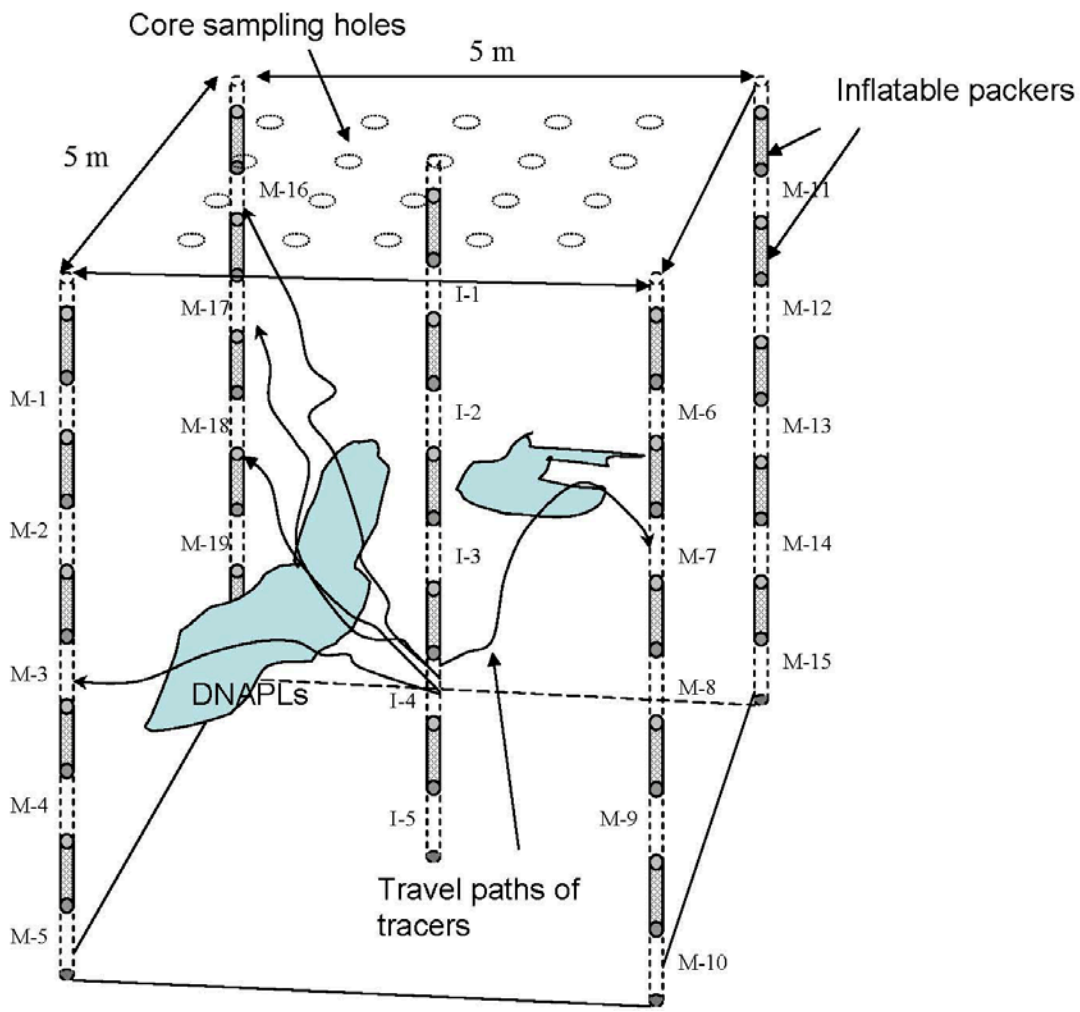


Figure A.1: Conceptual diagram illustrating the tomographic approach to characterize DNAPL source zones.

Table A.2: Estimated number of HT, CTT, and PTT tests required to characterize DNAPL source zones in Cases 1 through 4.

	HT	CTT	PTT	Days/person required	man-hours/person required
Case 1 [10×10×10]	5	2	2	26.0	260
Case 2 [50×50×50]	10	3	3	44.0	440

Case 3 [100×100×100]	15	4	4	62.0	620
Case 4 [150×150×150]	20	5	5	80.0	800

A.4. Costs for the Baseline Approach

To obtain a useful cost comparison with proven existing technology to characterize DNAPL source zones, we compare the costs of our technology to the baseline approach discussed in Kram et al. [2002]. The baseline approach consists of discrete sample collection from the surface at 5 ft depth intervals. Although this approach is commonly used, the likelihood of detecting DNAPL ganglia and microglobules using this approach is very low [Kram et al., 2002]. We also note that the excessive drilling within the DNAPL source zone may exacerbate the contamination problem through short-circuiting of DNAPLs along the boreholes to greater depths.

For the cost calculation, we assume that boreholes are drilled with direct push equipment such as that from Geoprobe. According to the company website [www.geoprobesystems.com], the rig is able to push from 20 to 50 holes up to a depth of 100 ft per day. We take their low end estimate of 20 boreholes for our calculation of the number of days required to complete the investigation. In addition, the website notes that depths of 100 ft can be achieved using this approach; however, larger drilling equipment may be necessary to drill and collect samples at larger depths and in materials that are less than ideal to push in, which could ultimately increase the cost for drilling and sample collection. For simplicity, we assume that the direct push technique is applicable for all cases considered here. Other costs pertaining to the baseline approach are obtained from Table 2 in Kram et al., [2002] and reproduced here as Table A.3.

Table A.3: Generic cost estimates for line item baseline case components [modified after Kram et al., 2002].

Item	Cost [$]	Per Unit
Drill rig	10	Foot
Push rig	3500	Day
Sampling	20	Sample
Grouting	3	Foot
Mobilization-demobilization	1000	Day
Standby labor	170	Hour
Decontamination labor	100	Hour
Drilling waste disposal	40	Cubic feet
Laboratory chemical analyses	150	Sample
Laboratory physical analyses	200	Sample
Reporting	3500	Report [average]

We assume that on a horizontal grid, a borehole is drilled every 5 ft for cases 1 and 2 while for cases 3 and 4, we assume that boreholes are drilled every 10 ft to achieve the degree of characterization that is comparable to the HT, CTT, and PTT approach. Therefore, for case 1 with a 10 ft by 10 ft grid, we assume that [3 × 3 boreholes] 9 boreholes are required. For case 2 with a 50 ft by 50 ft grid, [11 × 11 boreholes] 121 boreholes are required. We assume that sampling takes place at 5 ft intervals as it was suggested in Kram et al. [2002]. They state that this sampling interval is by no means valid for all sites. Rather, Kram et al. [2002] consider this to be a typical sampling increment that is applicable to many sites. Using these parameters, we

calculate the total length for the drilling, number of samples taken, and the time it will take for the completion of drilling. This information is summarized in Table A.4.

Table A.4: Summary of parameters in calculating cost estimates for the baseline case.

	*Number of boreholes	Total length [ft]	**Number of samples	Days for drilling	Total hours for drilling
Case 1	9	90	18	0.5	5.0
Case 2	121	6050	1210	6.1	61.0
Case 3	121	12100	2420	6.1	61.0
Case 4	256	38400	7680	12.8	128.0

Note: *Boreholes are drilled every 5 ft for cases 1 and 2 while for cases 3 and 4 boreholes are drilled every 10 ft. **Samples are collected at every 5 ft for all cases 1-4

A.5. Cost Comparisons

Table A.5 shows the comparison of costs from the two approaches for all the cases. Results show that the costs of characterizing DNAPL source zone rise modestly with increasing source area and volume with our HT, CTT, and PTT technology, but the costs rise exponentially for the baseline approach.

Table A.5: Summary of costs for the HT, CTT, and PTT approach versus the baseline approach.

	2004 Cost for HT, CTT, and PTT tests	2004 Cost for Baseline Approach
Case 1	$ 207,327	$ 27,279
Case 2	$ 357,421	$ 696,098
Case 3	$ 517,404	$ 1,304,913
Case 4	$ 677,387	$ 4,034,755

This is because if a comprehensive characterization of the DNAPL source zone is conducted with the baseline approach, a large number of boreholes and samples are required to obtain direct estimates of DNAPL concentrations. On the other hand, our technology uses a minimum number of boreholes to characterize the source zone. Drilling of a smaller number of boreholes will also lessen the chance for exacerbating the contamination problem through short-circuiting of DNAPLs.

Several factors can affect the costs in characterizing DNAPL source zone with the HT, CTT, and PTT technology. The resolution of the DNAPL residual saturation and the hydraulic/transport parameters are dependent on the number of tests conducted at a given field site. However, a higher resolution image of the DNAPL source zone should translate into a greater cost savings during the remediation phase because more information is available to design a better remediation strategy, which in turn should reduce the overall project costs with less uncertainty. Other factors such as the degree of physical and chemical heterogeneity may also affect the number of tests required. Therefore, the costs will be higher if a larger number of HT, CTT, and PTT tests are required to delineate the DNAPL source zone. Despite some of these uncertainties, our technology should be applicable in all types of terrains including fractured rocks and also in the vadose zone. There is also theoretically no limit to the depth at which the technology can be applied.

On the other hand, drilling costs for the baseline approach can be significantly more in gravelly material and in consolidated deposits including fractured rocks as the direct push

approach can only be applied in soft, unconsolidated sediments. For large depths, conventional drilling technologies need to be used for drilling and sampling collection.

In conclusion, our cost estimates suggest that the DNAPL source zone characterization by the HT, CTT, and PTT technology is considerably more cost effective than traditional approaches to obtain a three-dimensional tomogram that shows the spatial variability of flow and transport parameters as well as DNAPL saturations within the source zone. The costs are considerably lower for medium to larger DNAPL sites, however, the technology could be highly effective in smaller DNAPL sites as well because it will yield information that will make the remediation effort and decision making considerably easier. These results suggest for the expanded research effort into field implementation of the HT, CTT, and PTT technology.

Appendix A References

Carslaw, H. S. and J. C. Jaeger, *Conduction of Heat in Solids*, 2nd ed., Oxford University Press, 1959.

Kram, M. L., A. A. Keller, J. Rossabi, and L. G. Everett, DNAPL Characterization Methods and Approaches, Part 1: Performance Comparisons, *Groundwater Monitoring Review*, 109-123, Fall 2001.

Kram, M. L., A. A. Keller, J. Rossabi, and L. G. Everett, DNAPL Characterization Methods and Approaches, Part 2: Cost Comparisons, *Groundwater Monitoring Review*, 46-61, Winter 2002.

Pankow J. F. and J. A. Cherry, *Dense Chlorinated Solvents and other DNAPLs in Ground water: History, Behavior, and Remediation*, Waterloo Press, 522 pp., 1996.

RS Means, Environmental Remediation Cost Data – Unit Price, 4th Annual Edition, 1998.

Unger, M. and M. Coover, Memorandum to A. Leeson and S. Dockum, Development of the Fusion Tomography DNAPL Characterization Technology White Paper [SERDP CU-1365], March 3, 2004.

Appendix B. Definition of Success, Uncertainty Evaluation, and Field Protocol for Implementing the Hydraulic, Conservative Tracer and Partitioning Tracer Tomography Tests to Delineate DNAPL Source Zones [submitted on 7/15/2004]

B.1. Summary of White Paper

This white paper addresses several concerns raised by the by the review panel during the Spring In-Progress Review [IPR] Meeting held in Washington DC on April 14, 2004. These concerns are: 1] to establish a clear definition of what constitutes the success of this technology with respect to cost-effective characterization and monitoring of DNAPL source zones; 2] to provide plans for uncertainty reduction and parameter estimation in our predictive monitoring; and 3] to provide a field protocol for the technology.

B.2. Technology background

Dense Nonaqueous Phase Liquids [DNAPLs] are prevalent at a large number of sites throughout the world. The variable release history and geologic heterogeneity make the spatial distribution of DNAPLs in the source zone complex and make absolute cleanup difficult. A small amount of residual DNAPL can contribute to long-term groundwater contamination for decades to centuries. Therefore, the spatial distribution, mass, and composition of DNAPLs present in the source zone need to be characterized in great detail so that efficient remediation schemes can be designed. Development of a cost-effective technology that provides detailed images of DNAPL distributions in the source zones thus is essential.

The current technology for characterizing the DNAPL source zone heavily relies on invasive sampling conducted at point-scales. Although it is generally considered to be accurate at the sampling point if proper sample handling protocols are followed, for a source zone covering a large volume, densely distributed point samples are required to depict a detailed 3-D distribution of DNAPLs. Due to the very large number of samples that are required for proper characterization, current technology that relies on invasive sampling can be very costly and destructive if accurate site characterization is required for a large source zone. To overcome the shortcomings of current technology, our new technology exploits recently developed hydraulic/pneumatic tomography and expands the concept to the development of a conservative and partitioning tracer tomography to image the subsurface and DNAPL distributions over a large volume of geologic media without extensive invasive sampling. The technology uses a stochastic fusion of information methodology to assimilate data obtained from currently available indirect and direct methods of DNAPL source zone characterization [Kram et al., 2001] and results of hydraulic, conservative, and partitioning tracer tomography surveys to derive the best estimate of the DNAPL residual distribution and to quantify its uncertainty. As a result, the new technology can characterize a greater volume of geological media at higher resolution at lower cost. Specifically, it first analyzes the information derived from Hydraulic Tomography [HT] by incorporating point samples of hydraulic properties to identify three-dimensional hydraulic heterogeneity. It then improves the estimate of the heterogeneity by incorporating new information acquired from the Conservative Tracer Tomography [CTT]. Afterward, the improved hydraulic heterogeneity is used to simulate the HT such that more detailed information about the response of the subsurface becomes available. This new information again is fed back to the technique to update the estimate of the heterogeneity. The iterative process continues until the available information and measurements are fully utilized for identifying the heterogeneity that controls the spatial distribution of DNAPLs. Upon completion, the newly derived

knowledge of the heterogeneity are then used to design partitioning tracer tomography [PTT] tests and are combined with the hydraulic data and other prior, point-scale information on DNAPLs [and the site hydrology] to depict the 3-D spatial distribution of DNAPL residual saturation in the source zone.

The technology is a new generation technology that images the three-dimensional spatial distribution of hydraulic heterogeneity in the subsurface, which allows for the significantly improved characterization of DNAPL source zones compared to existing technologies listed in Kram et al. [2001]. Effective source zone characterization is critical for its efficient remediation and long-term monitoring for natural attenuation. This new technology also provides the means for real-time monitoring and forecasting of the fate of DNAPLs during remediation. It should be also be beneficial for characterization and monitoring efforts that rely on natural attenuation after a significant amount of DNAPLs have been removed after active remediation. Most importantly, it provides uncertainty estimates that can facilitate better decision making.

B.3. Definition of Success for the Developed Technology

The definition of success for this technology depends on the success of the 3 individual tomography tests [hydraulic, conservative tracer, and partitioning tracer] and their joint interpretation.

We propose three methods to quantify the success of the technology. In the first approach we will utilize numerical experiments to test our technology under conditions where “data" are free of measurement errors and model errors. These computational experiments will define the quality and reliability of information derived from our approaches, and will assist in defining sampling strategies to be used in the field. The second method to quantify the success of the approach utilizes the data from the sandbox experiments. These will provide critical data under controlled environments where measurement/model errors may exist but can be controlled. In addition, we will create different types of heterogeneities. The ultimate test of the technology will be in the field. In this case, measurement/model errors are unknown and heterogeneity is realistic. The following describes the approaches in quantifying the success of each component of our DNAPL source zone characterization technology.

The success of this technology is defined as its ability to depict hydraulic conductivity, specific storage, porosity, dispersivity, and DNAPLs [if any] variations over a large volume of geological media in three-dimensions at less cost and time than the point sampling approach such as those described in Kram et al. [2001]. Several approaches can be used to assess the success of the technology: direct and indirect approaches. A direct approach is to compare the distribution of estimated hydraulic conductivity and specific storage with those obtained from densely collected small scale estimates such as that from cores, slug and/or single-hole tests conducted within the DNAPL source zone. Alternative to the direct approach are indirect approaches, including pumping and tracer tests. Specifically, the estimated hydraulic properties by our technology will provide better predictions of overall 3-D hydraulic head and tracer distributions resulting from these tests than the properties derived from point samples at the same wells for the tomographic surveys. During the tests, the true 3-D hydraulic head and tracer distributions will be obtained by a densely distributed point sampling network. The success also can be assessed by comparing predicted flow and transport processes [i.e., hydraulic heads, flow rates, and tracer concentration breakthroughs] observed at the tomographic survey wells by our estimated properties with those by the point sampling at the same wells.

It is important to remember that none of the direct [e.g., soil and groundwater contamination surveys through drilling] or indirect approaches [e.g., visual inspections of drill cores, soil vapor analysis, geophysical surveys, use of radon abundance data, and partitioning tracer tests] described in Kram et al. [2001] on its own can provide detailed distributions of hydraulic heterogeneity and DNAPL saturations in three-dimensions. Distributions of hydraulic heterogeneity and variability in transport parameters, and distributions of DNAPL saturations may be obtained through the geostatistical analysis of point scale data sets. However, in order to obtain a detailed distribution using the geostatistical analysis of point scale samples, a large number of boreholes and samples obtained from each borehole will be required to characterize a given site. In addition, it is important to remember that the geostatistical analysis of point scale measurements do not factor in the physics of flow and transport of fluids and contaminants and thus may not provide information on medium connectivity, which can be critical in predicting the hydraulic and transport heterogeneity, as well as heterogeneity in DNAPL saturations. For example, slug test results and their corresponding geostatistical analysis may provide a distribution of hydraulic conductivity but this will not provide information on the connectivity of low and high permeability features that can dominate DNAPL infiltration processes and the locations of DNAPLs.

Our technology will also provide estimates of parameter uncertainties such that additional characterization can be designed if deemed necessary.

While the proposed technology produces promising results in numerical and laboratory experiments and we are confident in our logic behind the proposed technology, there are several concerns that need to be resolved. In particular:

1. We do not know if partitioning tracers can react with other unknown factors in the subsurface [other than the DNAPLs] to yield a similar retardation in arrival time under field conditions and to cause false alarms. That is, the idea of using partitioning tracers for detecting DNAPLs has been verified in bench-scale and field-scale tests, but the effects of geochemical heterogeneity in the field and its influence on the partitioning coefficients have not been studied.
2. The accuracy of hydraulic and tracer test data collected during the tomographic surveys is crucial to the results of our analysis. Issues related to the strength of the signal with respect to noise [signal to strength ratio] remain to be resolved under field conditions. Possible inconsistencies in the results of different tomographic surveys may occur and means to screen out inconsistent data sets are still needed.

B.4. Plans for Uncertainty Reduction and Parameter Estimation in Our Predictive Modeling

Our previous experience has shown that more accurate auxiliary information collected can lead to greater uncertainty reduction and parameter estimation in our DNAPL source zone model.

1. We need to develop our ability to include additional information [such as water and chemical fluxes] collected from the packers in the wells used for the tomographic surveys.

2. We need to develop our ability to include results from geophysical surveys [such as cross-bore hole electrical resistivity tomography for tracking three-dimensional tracer plumes and surface ground penetrating radar for detecting groundwater spatial variations during the tests].

3. Repeated hydraulic tomography experiments and geophysical surveys with varying source strengths may alleviate uncertainty associated with signal/noise ratio and data inconsistency issues as we discussed previously.

4. The accuracy of the tomograms for hydraulic conductivity has been demonstrated to be directly related to the accuracy of the data taken [i.e., the resolution of the pressure transducers] using synthetic simulations [Yeh and Liu, 2000; Liu et al., 2002]. Illman et al. [1998; see also Illman, 1999] also found that the quality of the data can be greatly be affected by the quality of the pressure transducer used in the field through their study on pneumatic tomography tests conducted in unsaturated fractured rocks. We suggest to conduct detailed laboratory and field studies to assess this issue to further to reduce the uncertainty in the calculated tomograms.

5. Since the effects of geochemical variability on partitioning tracer are not well understood, and geochemical variability is often site-specific, we recommend the following to address this issue at any given field site. Prior to employing the proposed technology to characterize a DNAPL source zone, we should first conduct column tracer tests using core samples excavated from the site to ensure that local geochemical heterogeneity will not cause false alarms in our interpretation.

In conclusion, the ability of our proposed technology for field DNAPL source characterization remains to be tested. We are certain that our technology will produce robust results in the laboratory experiments since the experimental errors are known and many of the factors that may influence the effectiveness of the technology can be controlled. For field problems, our technology will be significantly superior to the current point sampling technology in terms of delineating 3-D hydrologic heterogeneity, which is critical in detecting DNAPLs even if partitioning tracer technology is not used.

B.5. Field Protocol for the HT, CTT, and PTT Technology

A preliminary field protocol for the technology is provided below. We emphasize that this is by no means comprehensive and that additional surveys may be required or other characterization approaches may be substituted for different geologic terrains and conditions [e.g., fractured rocks].

The major stages for implementing the technology at a particular site consists of the following steps.

1. Planning and Phase 1 investigation
    a. Site visit and reconnaissance to assess equipment and testing needs;
    b. Collection of existing site data including topographic, geologic, and hydrologic maps as well as aerial photographs that may yield information on historical site

operation. Data on water levels, hydraulic conductivity, and information on the presence of DNAPLs are also gathered during this stage;

c. Delineation of source area based on historical data and site operation records;
d. If no data is available for potential areas for DNAPL source zones, then preliminary soil gas surveys or other approaches discussed in Kram et al. [2001] may be needed to delineate the source area;
e. Geophysical surveys to obtain general geological structures—geostatistical analysis of the surveys;
f. Preliminary numerical simulation of the tomographic tests for design of the initial site characterization.

2. Initial Site Characterization
    a. On-site drilling;
    b. Sample collection and analysis [hydraulic conductivity, total porosity, column experiments, bulk density, partitioning coefficients, etc.];
    c. Slug tests using direct push equipment during drilling or slug/single-hole tests using packers after drilling and well completion;
    d. Geostatisical analysis of available site data including previously collected geophysical survey data sets;
    e. Construction of site model that can be used to design HT tests [sampling intervals and frequencies].

3. Hydraulic Tomography [HT]
    a. Conduct HT test and immediately analyze the data using the inversion algorithm while the next HT survey is being conducted;
    b. Switch pumping location and conduct another HT survey; analyze the data sequentially;
    c. Conduct HT tests [the number will depend on degree in heterogeneity quantified by the variance of hydraulic conductivity] until no improvement in hydraulic heterogeneity pattern is observed from sequential inclusion of HT data into the inversion code.

4. Conservative Tracer Tomography [CTT]
    a. Conduct simulation studies to design the optimal injection, monitoring locations for the conservative and partitioning tracer tomography tests as well as their sampling frequencies;
    b. Conduct CTT using bromide as a conservative tracer and analyze samples with on-site ion chromatograph;
    c. Use information gained from CTT survey to refine the hydraulic heterogeneity and obtain estimates of transport parameters [effective porosity and dispersivity and their spatial distribution].

5. Partitioning Tracer Tomography [PTT]
    a. Conduct additional tracer simulation studies to refine the design for the optimal injection, monitoring locations for the tracer tomography tests as well as their sampling frequencies;

b. Conduct PTT tests using partitioning tracers;
c. Analyze partitioning tracers with an on-site Gas Chromatograph [GC] equipped with a Flame Ionization Detector [FID];
d. Analyze the resulting breakthrough curves using the inversion code;
e. Continue with PTT survey until no improvement in the tomogram is obtained.

6. Validation
    a. If validation is necessary, then a large number of samples must be obtained using direct push or conventional drilling technology with a split spoon sampler;
    b. In general, organic contaminants are extracted from the soil with organic solvents such as methylene chloride. The concentration of the VOC and SVOC present in the soil is then determined with a gas chromatograph.

B.6. Applicability of Hydraulic and Partitioning Tracer Tomography (HPTT) technology at existing DoD sites.

SERDP reviewers of the Draft Final Report have raised the following concern regarding this technology, "Final Report offers a wealth of information on groundwater hydraulic behavior and a potentially promising characterization technology; however, at this point in time, the DoD will likely be hesitant to invest in new groundwater wells for the purposes of additional characterization data. A potentially significant cost effective approach entails the use of existing wells to conduct the key tomographic tests as well as existing data to further refine and condition the tomographic output. Is such a scenario possible?"

Response: The resolution of subsurface heterogeneity obtained will depend on the number of wells, monitoring locations, pumping/injection locations as well as hydrogeological conditions at a given DoD site. While we use wells with screened intervals at different depths (i.e., multilevel wells) to explain the concept of hydraulic and partitioning tracer tomography (HPTT), we must emphasize that these techniques are not restricted to this particular type of well facilities. In cases where a field site is limited to a small number of wells (see Figure A.1, p.397, for an illustrative example), using wells with multi-level screened intervals would be necessary to realize the benefit of HPTT technology: 3-D image of hydraulic heterogeneity and DNAPL residual saturation. Otherwise, the tomography is readily applicable to existing well fields which consist of either fully or partially screened wells. For example, if all wells (for instance, 10 wells) are screened over the same depths, HPTT will provide images of the heterogeneity distribution in the lateral direction. In order to obtain 3-D images of heterogeneity over the well field in this case, installing one multi-level pumping well will be adequate to provide excitations at different depths. If all the wells at the site are screened over different depths, HPTT tests can also be carried out to obtain images of the heterogeneity in 3-D, without drilling new wells.

While existing wells can be used for the application of the technology at existing DoD sites, we also hope that DoD is amenable in installing additional multi-level monitoring wells at some sites to improve the resolution of subsurface heterogeneity images. This is because the upfront characterization will more than likely improve the effectiveness of various remediation technologies and will ultimately lead to improved remediation performance, shorter cleanup times, and cost savings.

Appendix B References

Illman W. A., D. L. Thompson, V. V. Vesselinov, G. Chen, and S. P. Neuman, *Single- and Cross-Hole Pneumatic Tests in Unsaturated Fractured Tuffs at the Apache Leap Research Site: Phenomenology, Spatial Variability, Connectivity and Scale*, NUREG/CR-5559, U. S. Nucl. Regul. Comm., Washington D.C., 1998.

Illman, W. A., *Single- and cross-hole pneumatic injection tests in unsaturated fractured tuffs at the Apache Leap Research Site near Superior, Arizona*, Ph.D. dissertation, Dep. of Hydrol. and Water Resour., Univ. of Ariz., Tucson, 1999.

Kram, M. L., A. A. Keller, J. Rossabi, and L. G. Everett, DNAPL Characterization Methods and Approaches, Part 1: Performance Comparisons, *Groundwater Monitoring Review*, 109-123, Fall 2001.

Liu, S., T. -C. J. Yeh and R. Gardiner, Effectiveness of hydraulic tomography: numerical and sandbox experiments, *Water Resour. Res.*, 38[4], 1034, 2002.

Yeh, T.-C. J. and S. Liu, Hydraulic tomography: development of a new aquifer test method, *Water Resour. Res.*, 2095-2105, 36[8], 2000.

Appendix C: Derivation of Moment Sensitivities using the Adjoint State Method

Differentiating the zeroth moment equation,

$$\nabla \cdot [K(\mathbf{x})\nabla M_0] + Q(\mathbf{x}_p)\tau = 0 \qquad \text{[C1]}$$

with respect to a parameter χ [where χ represent ln$K$ or ln$S$ at any location in the spatial domain Ω] give

$$\frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}}\right) + \frac{\partial}{\partial \mathbf{x}}\left(K\frac{\partial \phi_0}{\partial \mathbf{x}}\right) = 0 \qquad \text{[C2]}$$

where $\phi_0 = \partial M_0 / \partial \chi$ is called state sensitivity. Multiplying the resultant equation by an arbitrary function, ϕ_0^*, and integrating the equation over the entire spatial domain, Ω, gives

$$\int_\Omega [\phi_0^* \frac{\partial}{\partial \mathbf{x}}\left(\frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}}\right) + \phi_0^* \frac{\partial}{\partial \mathbf{x}}\left(K\frac{\partial \phi_0}{\partial \mathbf{x}}\right)] d\Omega = 0\,. \qquad \text{[C3]}$$

Applying Green's theorems to both terms of the left hand side of Eq. [C3] yields the following equation,

$$\int_\Omega [-\frac{\partial \phi_0^*}{\partial \mathbf{x}}\frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}} + \phi_0 \frac{\partial}{\partial \mathbf{x}}\left(K\frac{\partial \phi_0^*}{\partial \mathbf{x}}\right)] d\Omega + \int_\Gamma [\phi_0^* \frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}} + \phi_0^* K\frac{\partial \phi_0}{\partial \mathbf{x}} - \phi_0 K\frac{\partial \phi_0^*}{\partial \mathbf{x}}] \cdot \mathbf{n}\, d\Gamma = 0 \qquad \text{[C4]}$$

Applying a similar procedure to the first moment equation,

$$\nabla \cdot [K(\mathbf{x})\nabla M_1] + Q(\mathbf{x}_p)\frac{\tau^2}{2} + S_S(\mathbf{x})M_0 = 0\,, \qquad \text{[C5]}$$

we obtain

$$\begin{aligned} &\int_\Omega [-\frac{\partial \phi_1^*}{\partial \mathbf{x}}\frac{\partial K}{\partial \chi}\frac{\partial M_1}{\partial \mathbf{x}} + \phi_1 \frac{\partial}{\partial \mathbf{x}}\left(K\frac{\partial \phi_1^*}{\partial \mathbf{x}}\right) + \phi_1^* \frac{\partial S_s}{\partial \chi} M_0 + \phi_1^* \phi_0 S_s] d\Omega + \\ &\int_\Gamma [\phi_1^* \frac{\partial K}{\partial \chi}\frac{\partial M_1}{\partial \mathbf{x}} + \phi_1^* K\frac{\partial \phi_1}{\partial \mathbf{x}} - \phi_1 K\frac{\partial \phi_1^*}{\partial \mathbf{x}}] \cdot \mathbf{n}\, d\Gamma = 0 \end{aligned} \qquad \text{[C6]}$$

where ϕ_1^* is another arbitrary function and $\phi_1 = \partial M_1 / \partial \chi$ is another state sensitivity. The marginal sensitivity of a performance measure P is given as

$$\frac{dP}{d\chi} = \int_\Omega \left(\frac{\partial G}{\partial \chi} + \frac{\partial G}{\partial M_0}\phi_0 + \frac{\partial G}{\partial M_1}\phi_1\right) d\Omega \qquad \text{[C7]}$$

where G is the state function. The first term in the right side of Eq. [C7] represents direct dependence of the performance measure on the parameter whereas the second and third terms provide indirect dependence of the performance measure on the parameter through moments. Next, we add Eq. [C4] and Eq. [C6] into Eq. [C7], yielding

$$\begin{aligned}\frac{dP}{d\chi} = \int_{\Omega} [&\frac{\partial G}{\partial \chi} + \frac{\partial G}{\partial M_0}\phi_0 + \phi_0 \frac{\partial}{\partial \mathbf{x}}(K\frac{\partial \phi_0^*}{\partial \mathbf{x}}) + \phi_1^* \phi_0 S_s + \frac{\partial G}{\partial M_1}\phi_1 + \phi_1 \frac{\partial}{\partial \mathbf{x}}(K\frac{\partial \phi_1^*}{\partial \mathbf{x}}) - \\ &\frac{\partial \phi_0^*}{\partial \mathbf{x}}\frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}} - \frac{\partial \phi_1^*}{\partial \mathbf{x}}\frac{\partial K}{\partial \chi}\frac{\partial M_1}{\partial \mathbf{x}} + \phi_1^* \frac{\partial S_s}{\partial \chi} M_0] d\Omega + \\ &\int_{\Gamma} [\phi_0^* \frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}} + \phi_0^* K \frac{\partial \phi_0}{\partial \mathbf{x}} - \phi_0 K \frac{\partial \phi_0^*}{\partial \mathbf{x}}] \cdot \mathbf{n}\, d\Gamma + \\ &\int_{\Gamma} [\phi_1^* \frac{\partial K}{\partial \chi}\frac{\partial M_1}{\partial \mathbf{x}} + \phi_1^* K \frac{\partial \phi_1}{\partial \mathbf{x}} - \phi_1 K \frac{\partial \phi_1^*}{\partial \mathbf{x}}] \cdot \mathbf{n}\, d\Gamma\end{aligned} \quad . \qquad \text{[C8]}$$

If we let the two arbitrary functions ϕ_0^* and ϕ_1^* satisfy the following two adjoint equations:

$$\frac{\partial}{\partial \mathbf{x}}(K\frac{\partial \phi_1^*}{\partial \mathbf{x}}) + \frac{\partial G}{\partial M_1} = 0 \qquad \text{[C9]}$$

$$\frac{\partial}{\partial \mathbf{x}}(K\frac{\partial \phi_0^*}{\partial \mathbf{x}}) + \phi_1^* S_s + \frac{\partial G}{\partial M_0} = 0 \qquad \text{[C10]}$$

and the boundary conditions

$$\phi_n^* = 0 \quad \text{at } \Gamma_1 \qquad \text{[C11]}$$

$$(K\nabla \phi_n^*) \cdot \mathbf{n} = 0 \ \text{ at } \Gamma_2 \quad [n\text{=0,1}], \qquad \text{[C12]}$$

The terms associating with state sensitivities ϕ_0 and ϕ_1 as well as boundary terms in Eq. [C8] disappear. Eq. [C8] is reduced to

$$\frac{dP}{d\chi} = \int_{\Omega} [\frac{\partial G}{\partial \chi} - \frac{\partial \phi_0^*}{\partial \mathbf{x}}\frac{\partial K}{\partial \chi}\frac{\partial M_0}{\partial \mathbf{x}} - \frac{\partial \phi_1^*}{\partial \mathbf{x}}\frac{\partial K}{\partial \chi}\frac{\partial M_1}{\partial \mathbf{x}} + \phi_1^* \frac{\partial S_s}{\partial \chi} M_0] d\Omega \qquad \text{[C13]}$$

If an observation of the moments is made at location, $\mathbf{x}_k$, the state function becomes $G = M_n \delta(\mathbf{x} - \mathbf{x}_k)$ in which n=0, and 1, denoting the zero and the first moment, respectively. Now, let $G = M_1 \delta(\mathbf{x} - \mathbf{x}_k)$ be the first moment at location $\mathbf{x}_k$, the direct dependence term,

$\partial G/\partial\chi$, becomes zero. If K and S_s are uncorrelated, then the term, $\partial S_s/\partial \ln K$ in Eq. [C13] is zero. As a result, we have the sensitivity of M_1^k with respect to $\ln K_\ell$:

$$\frac{\partial M_1^k}{\partial \ln K_\ell} = \int_\Omega \left[-\frac{\partial \phi_0^*}{\partial \mathbf{x}} K_\ell \frac{\partial M_0}{\partial \mathbf{x}} - \frac{\partial \phi_1^*}{\partial \mathbf{x}} K_\ell \frac{\partial M_1}{\partial \mathbf{x}} \right] d\Omega \qquad \text{[C14]}$$

where the superscript k denotes the observation at location $\mathbf{x}_k$ and the subscript ℓ denotes the location of the parameter in the domain. Similarly, the sensitivity of M_1 to ln$S$ is given as

$$\frac{\partial M_1^k}{\partial \ln S_\ell} = \int_\Omega \phi_1^* M_0 S_{s\ell} d\Omega . \qquad \text{[C15]}$$

Next, let $G = M_0 \delta(\mathbf{x} - \mathbf{x}_k)$ be the zeroth moment at location $\mathbf{x}_k$, then term, $\partial G/\partial M_1$, in Eq. [C9] is zero. As a result, the solution ϕ_1^* of adjoint equation [C9] is zero everywhere. The adjoint equation [C10] therefore is reduced to

$$\frac{\partial}{\partial \mathbf{x}}(K \frac{\partial \phi_0^*}{\partial \mathbf{x}}) + \frac{\partial G}{\partial M_0} = 0 , \qquad \text{[C16]}$$

and Eq. [C13] becomes

$$\frac{dP}{d\chi} = \int_\Omega [\frac{\partial G}{\partial \chi} - \frac{\partial \phi_0^*}{\partial \mathbf{x}} \frac{\partial K}{\partial \chi} \frac{\partial M_0}{\partial \mathbf{x}}] d\Omega . \qquad \text{[C17]}$$

The sensitivity of M_0 to ln$K$ is then given as

$$\frac{\partial M_0^k}{\partial \ln K_\ell} = \int_\Omega - \frac{\partial \phi_0^*}{\partial \mathbf{x}} K_\ell \frac{\partial M_0}{\partial \mathbf{x}} d\Omega \qquad \text{[C18]}$$

Appendix D. Partitioning coefficients for PCE and selected tracers

Four alcohols were chosen to be used as partitioning tracers for this study. The alcohols, our measured dimensionless K_{NW} values, and values reported by previous researchers are listed in Table D.1. The NAPL-phase used in this study was tetrachloroethene [PCE]. The purities of the alcohols and PCE purchased from Sigma-Aldrich and Fluka are as follows: PCE > 99.9%; IPA > 99.5%; 2M1B > 99%; 4M2P> 98%; 2, 4DM3P = 99%; and 1- Hex ≥ 99.9%.

Nine sets of batch tests were conducted to determine NAPL – water partition coefficients for each tracer in water-tracer-NAPL solutions as well as to examine the effects of the presence two or more tracers, simultaneously, on partition coefficient estimation. The batch test were designed to estimate one alcohol/PCE partitioning coefficient, then duos of alcohols and PCE and finally three alcohols and PCE. Per each test regardless the alcohols involved, five partitioning-tracer stock solutions were prepared volumetrically with Hamilton gas-tight syringes in 250-mL volumetric flasks. The levels ranged from 50 mg/L to 800 mg/L.

Purified water [Barnstead Nanopure, Diamond] was used to dilute the aqueous tracer stock solutions to the five different concentrations: 50, 100, 200, 400, and 800 mg/L. The 5 different levels were prepared in pre-cleaned 20mL EPA/ VOA glass vials with Teflon-lined septa from EnviroTech.

The batch tests to determine NAPL-water partition coefficients for each tracer contained five sample groups. The batch tests were prepared in pre-cleaned 22.5mL EPA/ VOA glass vials with Teflon-lined septa from EnviroTech. Each sample set consisted of triplicate vials containing TCE and the aqueous tracer solution and a control group of triplicate vials containing aqueous tracer solution only [initial concentration]. 20 mL of aqueous tracer solution was dispensed into each vial using a 10 mL pipette followed by 2.5 mL injection of TCE with a gas-tight syringe. The vials were immediately capped to avoid losses to the atmosphere.

Table D.1: PCE-Equilibrium Partitioning Coefficient K_{NW}, based on single, dual and multiple tracers batch tests.

Alcohol [s]	K_{NW} []	Statistics	
4M2P and 1-Hex	4.56	Mean	St. Dev
4M2P and 24DM3P	4.22	4.517	0.328
4M2P_24DM3P and 1-Hex	4.328		
4M2P	4.96		
1-Hex and 24DM3P	7.877	Mean	St. Dev
1-Hex and 4M2P	7.72	7.598	0.544
1-Hex, 4M2P and 24DM3P	7.996		
1-Hex	6.8		
24DM3P and 1-Hex	28.77	Mean	St. Dev
24DM3P and 4M2P	28.164	28.594	0.804
24DM3P,_4M2P and 1-Hex	29.64		
24DM3P	27.8		

The 22.5mL vials were allowed to equilibrate on a platform shaker [Innova 2000] at 150 rpm and room temperature [22.2°C] for 24 hours, which is sufficient to obtain equilibrium [McCray and Dugan, 2002]. 1.5 mL aliquots of the aqueous phase were withdrawn from the 25mL vials with minimal disturbance using a gas-tight syringe and transferred to 1.5 mL glass GC auto sampler vials with Teflon septa.

Partitioning tracer concentrations were determined using a gas chromatograph [Shimadzu GC 2010] with auto sampler [AOC-20i auto injector] equipped with a flame ionization detector [FID]. The GC/FID method consisted of a Restek Stabilwax® capillary column [30m, 0.32 mm ID, with film thickness of 1μm], helium as the carrier gas, zero-grade air, an injection volume of 1 μL, and injector and FID temperatures of 250°C. A different temperature program was used for each alcohol. This provided a faster and more efficient means of collecting sample concentration estimates. Table D.2 lists the 5 different temperature gradient programs used.

Table D.2: Temperature gradient programs used for each alcohol tracer in GC-FID.

Alcohol Tracer [s]	Rate [°C/min]	Temp [°C]	Hold [min]	Total Run [min]
2M1B	0	50	1	7.71
	100	120	0	
	20	220	1	
4M2B 4M2B/24DM3P	0	50	1	6.72
	100	110	0	
	15	180	0	
	100	225	0	
1-Hex 1-Hex/4M2P 1-Hex/24DM3P	0	50	1	9.06
	100	110	0	
	10	180	0	
	100	225	0	
24DM3P	0	50	1	6.72
	100	110	0	
	15	180	0	
	100	225	0	
24DM3P/1-Hex	0	50	1	9.06
	100	110	0	
	10	180	0	
	100	225	0	
3 Tracer suite [24DM3P/1-Hex/4M2P]	0	50	1	9.06
	100	110	0	
	10	180	0	
	100	225	0	

Four external calibration standards [please find at the end of this appendix an example of the computation and procedure followed] were prepared for each tracer in purified water at levels of 1, 10, 100, and 1000 mg/L. The calibration plots were linear [$R^2 > 0.99$]. Concentrations of the partitioning tracers in the NAPL phase were determined by calculating the difference between

initial and equilibrium tracer concentration in the aqueous phase along with the measured amounts of water and NAPL in the sample vials. Figure D.1 shows the results from a batch test with 2, 4DM3P. The slope of the line is equal to the NAPL-water partitioning coefficient, $K_{NW.}$

The data on Table D.1 reveals that the partitioning coefficient fluctuates. These variations are acceptable based on the findings of the single dual and multiple batch tests, also the range obtained in the experiments is similar to the range of the values reported in the literature. By inspecting the plot and the values listed on Table D.1 it is suggested that the influence of the possible tracer interaction is not significant. However, these conclusions are still preliminary and additional studies are needed.

Figure D.1 is a plot of multiple batch test results for 1-Hex. The coefficients vary slightly more than the ones of the lower partitioning tracer [4M2P], however the variations are acceptable based on the single dual and multiple batch tests and also the range of these experimental values is smaller than the values found in the literature.

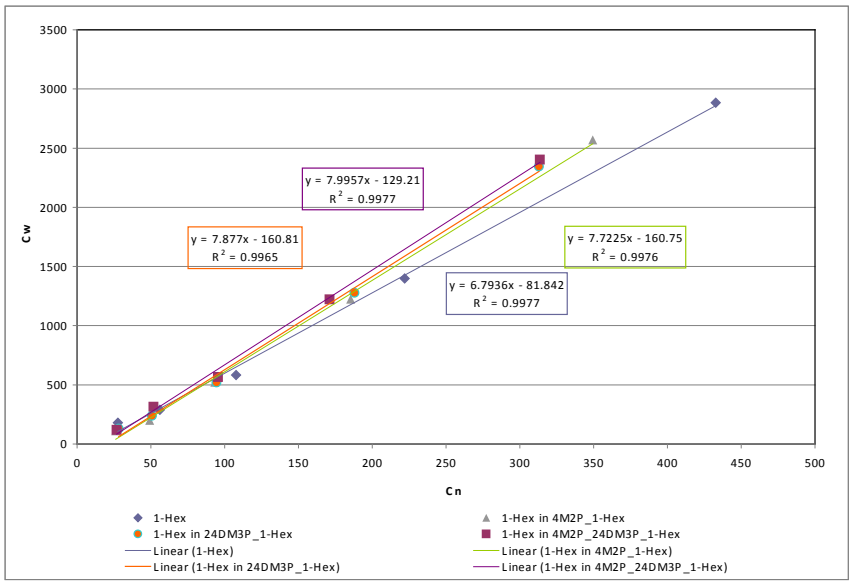


Figure D.1: Equilibrium partitioning of 1-Hex for single, dual, and multiple tracers and PCE.

Figure D.2 is a plot of multiple batch test results for 24DM3P which is the partitioning tracer with the highest partitioning coefficient considered here. The variations of the results appear to be even smaller than those for 1-Hex. This study suggests that the tracer interaction may not be a significant issue for the partitioning tracer and PCE tested as reported in the literature. A more important issue may be the nonequilibrium partitioning effect.

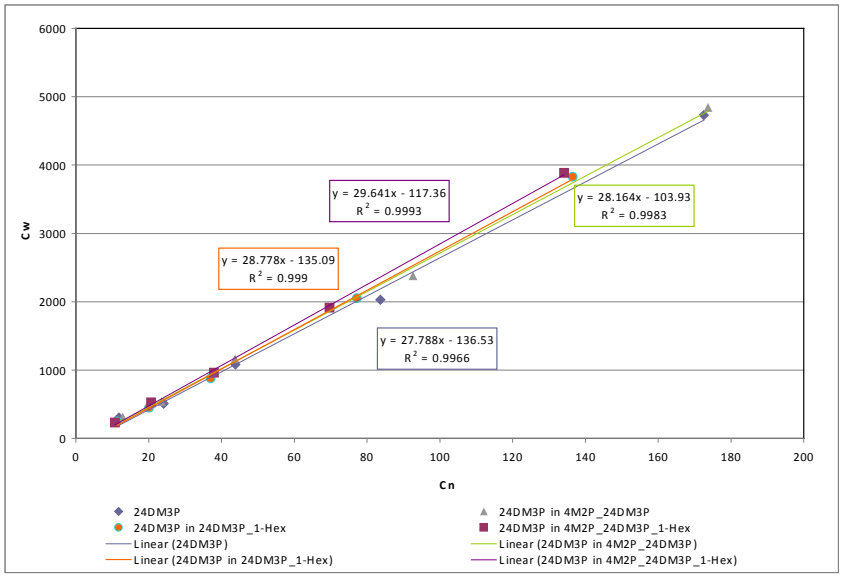


Figure D.2: Equilibrium partitioning of 24DM3P for single, dual and multiple tracers and PCE.

Appendix E: List of publications

Journal articles resulting from ER-1365 that are published, in press, submitted or are under preparation.

Berg, S. J. and W. A. Illman [2011], Capturing heterogeneity in groundwater flow parameters: Comparison of approaches through controlled sandbox experiments, *Water Resources Research*, in revision.

Berg S. J. and W. A. Illman [2011], Flow to a partially penetrating well in a heterogeneous unconfined aquifer: Insights from intermediate scale sandbox experiments, *Water Resources Research*, in revision

Hao, Y., T.-C. J. Yeh, W. A. Illman, K. Ando, K.-C. Hsu [2008], Hydraulic tomography for detecting fracture connectivity, *Ground Water*, 46[2], 183-192.

Kuhlman, K. L., A. C. Hinnell, P. K. Mishra, and T.-C. J. Yeh [2008], Basin-scale transmissivity and storativity estimation using hydraulic tomography, *Ground Water*, 46[5], 706-715.

Illman, W. A., X. Liu, and A. Craig [2007], Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms, *Journal of Hydrology*, 341[3-4], 222-234.

Illman, W. A., X. Liu, S. Takeuchi, T. J. Yeh, K. Ando, and H. Saegusa [2009], Hydraulic tomography in fractured granite: Mizunami Underground Research site, Japan, *Water Resour. Res*., 45, W01406, doi:10.1029/2007WR006715.

Illman, W. A., J. Zhu, A. J. Craig, and D. Yin [2010a], Comparison of aquifer characterization approaches through steady state groundwater model validation: A controlled laboratory sandbox study, Water Resour. Res., 46, W04502, doi:10.1029/2009WR007745.

Illman, W. A., S. J. Berg, X. Liu, A. J. Craig and A. Massi [2010b], Hydraulic/partitioning tracer tomography for trichloroethene source zone characterization: Small-scale sandbox experiments, Environ. Sci. Technol., 44[22], pp. 8609-8614, doi: 10.1021/es101654j.

Illman, W. A., S. J. Berg, and T.-C. J. Yeh [2011], Comparison of heterogeneity characterization methods for improved predictions of solute transport: Laboratory sandbox experiments, *Ground Water*, in revision

Liu, X., W. A. Illman, A. J. Craig, J. Zhu, and T.-C. J. Yeh [2007], Laboratory sandbox validation of transient hydraulic tomography, *Water Resour. Res*., 43, W05404, doi:10.1029/2006WR005144.

Mao, D., L. Wan, T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen [2011], A revisit of drawdown behavior during pumping in unconfined aquifers, *Water Resour. Res*., 47, W05502, doi:10.1029/2010WR009326.

Straface, S., T.-C. J. Yeh, J. Zhu, S. Troisi, and C. H. Lee [2007], Sequential aquifer tests at a well field, Montalto Uffugo Scalo, Italy, *Water Resour. Res.*, 43, W07432, doi:10.1029/2006WR005287.

Wu, C.-M., T.-C. J. Yeh, J. Zhu, T. H. Lee, N.-S. Hsu, C.-H. Chen, and A. F. Sancho [2005], Traditional analysis of aquifer tests: Comparing apples to oranges?, *Water Resour. Res.*, 41, W09402, doi:10.1029/2004WR003717.

Xiang, J., T.-C. J. Yeh, C.-H. Lee, K.-C. Hsu, and J.-C. Wen [2009], A simultaneous successive linear estimator and a guide for hydraulic tomography analysis, *Water Resour. Res*., 45, W02432, doi:10.1029/2008WR007180.

Yeh, T.-C. J., and J. Zhu [2007], Hydraulic/partitioning tracer tomography for characterization of dense nonaqueous phase liquid source zones, *Water Resour. Res*., 43, W06435, doi:10.1029/2006WR004877.

Yeh T.-C. J., C.-H. Lee [2007], Time to change the way we collect and analyze data for aquifer characterization. *Ground Water* 45[2]: 116-118.

Yeh, T.-C. J., C. H. Lee, K-C Hsu, and Y-C Tan [2007], Fusion of Active and Passive Hydrologic and Geophysical Tomographic Surveys: The Future of Subsurface Characterization, in Data Integration in Subsurface Hydrology Edited by D. W. Hyndman, F. D. Day-Lewis, and K. Singha, AGU geophysical monograph 171.

Yeh, T.-C. J., C. H. Lee, K. C. Hsu, W. A. Illman, W. Barrash, X. Cai, J. Daniels, E. Sudicky, L. Wan, G. Li, and C. L. Winter [2008], A view towards the future of subsurface characterization: CAT scanning groundwater basins, *Water Resour. Res.*, 44, W03301, doi: 10.1029/2007WR006375.

Zhu, J., and T.-C. J. Yeh [2005], Characterization of aquifer heterogeneity using transient hydraulic tomography, *Water Resour. Res.*, 41, W07028, doi:10.1029/2004WR003790.

Zhu, J., and T.-C. J. Yeh [2006], Analysis of hydraulic tomography using temporal moments of drawdown recovery data, *Water Resour. Res.*, 42, W02403, doi:10.1029/2005WR004309.

Zhu, J., X. Cai, and T.-C. J. Yeh [2009], Analysis of tracer tomography using temporal moments of tracer breakthrough curves, *Adv. Water Resour*., 32, 391-400.

Technical reports [in print, accepted for publication, or submitted for publication]

- None

Conference/Symposium Proceedings Papers [other than Abstracts] that are scientifically recognizable and referenced

Illman, W. A., Hydraulic tomography for accurate characterization of subsurface hydraulic heterogeneity, Water in Mining conference proceedings, 2nd International Congress on Water Management in the Mining Industry, June 9 - 11, 2010, Santiago, Chile.

Illman, W. A., Imaging of subsurface heterogeneity with hydraulic tomography: Laboratory sandbox investigations, ModelCARE 2007 Conference Proceedings, Copenhagen, Denmark, 2007.

Published Technical Abstracts [e.g., SERDP's Annual Symposium]

Illman, W. A.*, S. J. Berg, T.-C. Jim Yeh, and D. Mao, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, SERDP & ESTCP -Partners in Environmental Technology Technical Symposium & Workshop, November 29 – December 2, 2010, Washington DC [poster presentation].

Illman, W. A.* and S. J. Berg [2010], Comparison of aquifer characterization approaches through transient groundwater model validation: A controlled laboratory sandbox study, American Geophysical Union Western Pacific geophysical Meeting, Taipei, Taiwan, June 22 - 26, 2010 **[invited oral presentation]**.

Illman, W.*, [2009], Hydraulic tomography for accurate characterization of subsurface hydraulic heterogeneity, Water in Mining II International Congress on Water Management in the Mining Industry, June 9 -11, 2010, Santiago, Chile **[oral presentation]**.

Illman, W. A.*, S. J. Berg, T.-C. Jim Yeh, D. Mao, A. Craig, and J. Zhu, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, SERDP & ESTCP - Partners in Environmental Technology Technical Symposium & Workshop, December 1 - 3, 2009, Washington DC [poster presentation].

Illman, W. A.*, K. Lorenzetti, S. J. Berg, M. Alexander, T.-C. Jim Yeh, J. Zhu, and A. Craig, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, SERDP & ESTCP -Partners in Environmental Technology Technical Symposium & Workshop, December 2 - 4, 2008, Washington DC [poster presentation].

Illman, W. A.*, J. Zhu, A. Craig, and D. Yin, Can a groundwater flow model be validated? Encouraging, positive evidence from laboratory hydraulic tomography experiments, Eos Trans. AGU, 89[53], Fall Meet. Suppl., Abstract H33H-04.

Illman, W. A.*, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, American Society of Civil Engineers, World Environmental & Water Resources Congress, May 13-16, 2008, Honolulu, Hawaii **[invited oral presentation]**.

Illman, W. A.*, X. Liu, T.-C. J. Yeh, K. Ando, S. Takeuchi and H. Saegusa, Hydraulic Tomography in Fractured Granite: The Mizunami Underground Research Laboratory Site, Japan, American Geophysical Union 2007 Fall meeting **[invited oral presentation]**.

Illman, W. A.*, T.-C. Jim Yeh, A. Craig, A. Massi, X. Liu, D. Yin, and J. Zhu, Modeling and experimental studies for DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, Groundwater Resources Association of California, DNAPL-2 Symposium: Source Zone Characterization & Remediation, November 14-15, 2007 [oral presentation].

Illman, W. A.*, Evaluation of transient hydraulic tomography and common hydraulic characterization approaches through laboratory sandbox experiments, National Kaohsiung University, Taiwan, 11/22/2007 **[invited oral presentation].**

Illman, W. A.*, Imaging of subsurface heterogeneity with hydraulic tomography: Laboratory sandbox investigations, ModelCARE 2007, Copenhagen, Denmark, 2007 [oral presentation].

Illman, W. A.*, A. Craig, X. Liu, A. Massi, T.-C. J. Yeh, D. Yin, and J. Zhu, A new paradigm in DNAPL source zone characterization: 3D imaging of contaminant distributions through hydraulic and tracer tomography, Spring AGU meeting, Acapulco, Mexico, 2007 **[invited oral presentation].**

X. Liu*, W. A. Illman, A. J. Craig, and A. Massi, Hydraulic and tracer tomography for the characterization of DNAPL source zones: A laboratory sandbox investigation, Spring AGU meeting, Acapulco, Mexico, 2007 [oral presentation].

Yin*, D., W. A. Illman, X. Liu, and A. J. Craig, Hydraulic tomography using temporal moments of drawdown-recovery data: Laboratory sandbox study, Spring AGU meeting, Acapulco, Mexico, 2007 [oral presentation].

Illman, W. A.*, X. Liu, and A. Craig, Steady-state hydraulic tomography: the role of signal-to-noise ratio and conditioning on hydraulic conductivity tomograms, EGU spring meeting, Vienna, Austria, Geophysical Research Abstracts, Vol. 9, 11187, 2007 [oral presentation].

Illman, W. A.*, X. Liu, D. Yin, A. Craig, Multi-method and multiscale validation of hydraulic tomography in a laboratory aquifer with deterministic heterogeneity, Society for Industrial and Applied Mathematics [SIAM] Conference on Mathematical & Computational Issues in Geoscience, March 19-22, 2007, Santa Fe, NM **[invited oral presentation].**

Illman, W. A.*, A.J. Craig, A. Massi, X. Liu, T.-C. J. Yeh, J. Zhu, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography : Experimental and modeling studies, SERDP & ESTCP -Partners in Environmental Technology Technical Symposium & Workshop, November 28 - 30, 2006, Washington DC [poster presentation].

Yin, D.*, W. A. Illman, X. Liu, and A. J. Craig, Laboratory sandbox validation of hydraulic tomography using temporal moments of drawdown recovery data, AGU Western Pacific Geophysical Meeting, Beijing, China, 2006 [oral presentation].

Liu, X.*, W. A. Illman, T.-C. Jim Yeh, K. Ando, and S. Takeuchi, Two- and three-dimensional modeling studies of cross-hole hydraulic tests in fractured granite at Mizunami, Japan, AGU Western Pacific Geophysical Meeting, Beijing, China, 2006 [poster presentation].

Liu, X.*, W. A. Illman, and A. J. Craig, Transient hydraulic tomography in a sandbox with deterministic heterogeneity: Multi-method and multi-scale validation of hydraulic conductivity and specific storage tomograms, AGU Western Pacific Geophysical Meeting, Beijing, China, 2006 [oral presentation].

Illman, W. A.* and T.-C. Jim Yeh, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, AGU Western Pacific Geophysical Meeting, Beijing, China, 2006 **[invited oral presentation].**

Tartakovsky D. M. and W. A. Illman*, Lessons from the asymptotic analysis of three-dimensional pressure interference tests in fractured granite, AGU Western Pacific Geophysical Meeting, Beijing, China, 2006 **[invited oral presentation].**

Illman, W. A.* and T.-C. Jim Yeh, Fusion of tomography tests to characterize DNAPL source zones: conceptual framework and preliminary results from theoretical, computational, and experimental studies ModelCARE2005, 5th International Conference on Calibration and Reliability in Groundwater Modeling: From Uncertainty to Decision Making, The Hague [Scheveningen], The Netherlands, 6-9 June 2005 [oral presentation].

Illman, W. A.* and D. M. Tartakovsky, Asymptotic analysis of three-dimensional pressure interference tests, European Geophysical Society Annual Meeting, Vienna, Austria, 2005, in *Geophysical Research Abstracts*, Vol. 7, 02922, 2005 [oral presentation].

Liu, X.*, W. A. Illman, and A. J. Craig, Transient hydraulic tomography in a sandbox with deterministic heterogeneity: Validation of hydraulic conductivity and specific storage tomograms, American Geophysical Union 2005 Fall meeting [poster presentation; **winner of AGU Hydrology section Outstanding Student Presentation Award**].

Illman, W. A.*, X. Liu, T.-C. J. Yeh, K. Ando, S. Takeuchi and H. Saegusa, Hydraulic Tomography in Fractured Granite: The Mizunami Underground Research Laboratory Site, Japan, American Geophysical Union 2007 Fall meeting **[invited oral presentation]**.

Illman, W. A.*, T.-C. Jim Yeh, A. Craig, A. Massi, X. Liu, D. Yin, and J. Zhu, Modeling and experimental studies for DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, SERDP & ESTCP -Partners in Environmental Technology Technical Symposium & Workshop, December 4 - 6, 2007, Washington DC [poster presentation].

Illman, W. A.*, T.-C. Jim Yeh, A. Craig, A. Massi, X. Liu, D. Yin, and J. Zhu, Modeling and experimental studies for DNAPL source zone characterization by the fusion of hydraulic and tracer tomography tests, Groundwater Resources Association of California, DNAPL-2 Symposium: Source Zone Characterization & Remediation, November 14-15, 2007 [oral presentation].

Illman, W. A.*, X. Liu, D. Yin, A. Craig, Multi-method and multiscale validation of hydraulic tomography in a laboratory aquifer with deterministic heterogeneity, Society for Industrial and Applied Mathematics [SIAM] Conference on Mathematical & Computational Issues in Geoscience, March 19-22, 2007, Santa Fe, NM **[invited oral presentation].**

Illman, W. A.*, A.J. Craig, A. Massi, X. Liu, T.-C. J. Yeh, J. Zhu, DNAPL source zone characterization by the fusion of hydraulic and tracer tomography : Experimental and modeling studies, SERDP & ESTCP -Partners in Environmental Technology Technical Symposium & Workshop, November 28 - 30, 2006, Washington DC [poster presentation].

Yeh, T.-C. J.*, W. A. Illman, A. Kruger, E. Sudicky, and J. Daniels, New technology to improve the understanding of contaminant migration in fractured geological settings, The Geological Society of America, 2006 Philadelphia Annual Meeting, 22-25 October, 2006 **[invited oral presentation].**

Craig, A. J.*, W. A. Illman, and X. Liu, A Cyclic Sediment Transport Approach to Create a Synthetic Aquifer with Multiscale Heterogeneity, American Geophysical Union 2005 Fall meeting [poster presentation].

Daniels, J. J.*, W. A. Illman, T.-C. J. Yeh, M. Parashar, S. A. Hariri, A. Kruger, Fusion of Multiple Levels of Subsurface Information for Imaging and Hydrologic Analysis, American Geophysical Union 2005 Fall meeting [oral presentation].

Kruger, A.*, W. A. Illman, and S. Yang, A Real-Time, Remote-Controlled Experimental Electrical Resistivity Tomography, American Geophysical Union 2005 Fall meeting [oral presentation].

Illman, W. A.*, X. Liu, and A. J. Craig, Steady-state hydraulic tomography in a laboratory aquifer with deterministic heterogeneity: Multi-method and multiscale validation of hydraulic conductivity tomograms, American Geophysical Union 2005 Fall meeting **[invited oral presentation].**

Illman, W. A.*, T.-C. Jim Yeh, J. Zhu, A. Craig, Stochastic fusion of information to characterize DNAPL source zones: conceptual framework and preliminary results from theoretical, computational, and experimental studies, Eos Trans. AGU, 85[47], Fall Meet. Suppl., Abstract H43H-02, 2004 [oral presentation].

Craig, A.* and W. A. Illman, Comparison of hydraulic parameters measured at multiple scales in a synthetic heterogeneous aquifer, Eos Trans. AGU, 85[47], Fall Meet. Suppl., Abstract H21E-1050, 2004 [poster presentation].

Illman, W. A.*, T.-C. Jim Yeh, Characterization of DNAPL source zones by fusion of hydraulic and tracer tomography test data: theoretical and experimental studies SERDP & ESTCP - Partners in Environmental Technology Technical Symposium & Workshop November 30 - December 1, 2004, Washington DC **[invited poster presentation].**

Craig, A. and W. A. Illman*, Measurement of hydraulic parameters at multiple scales in a synthetic heterogeneous aquifer constructed in the laboratory, SERDP & ESTCP -Partners in Environmental Technology Technical Symposium & Workshop November 30 - December 1, 2004, Washington DC [poster presentation].

Illman, W.A.* and T.-C. J. Yeh, DNAPL Source Zone Characterization by the Stochastic Fusion of Information: Blueprint and Preliminary Results, SERDP & ESTCP -*Partners in Environmental Technology Technical Symposium & Workshop* December 2-4, 2003, Washington DC [poster presentation].

Illman, W. A.* and T.-C. J. Yeh, Fusion of Tomography Tests for DNAPL Source Zone Characterization: Technology Development and Validation, Strategic Environmental Research & Development Program, Presentation to the Scientific Advisory Board, June 11, 2003 **[invited oral presentation].**

Published Text Books or Book Chapters

1. Alvarez, P. and W. A. Illman, *Bioremediation and Natural Attenuation: Process Fundamentals and Mathematical Models*, John Wiley & Sons, 2006.