NOTE TO USERS

Page(s) missing in number only; text follows. Page(s) were scanned as received.

19

Page(s) not included in the original manuscript and are not available from the author or university. The manuscript was scanned as received.

This reproduction is the best copy available.
CONDITIONING NONLOCAL STEADY-STATE FLOW ON HYDRAULIC
HEAD AND CONDUCTIVITY THROUGH GEOSTATISTICAL INVERSION

by

Abel Felipe Hernández Ochoa

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF HYDROLOGY AND WATER RESOURCES
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
WITH A MAJOR IN HYDROLOGY
In the Graduate College
THE UNIVERSITY OF ARIZONA

2002
As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Abel Felipe Hernandez Ochoa entitled Conditioning nonlocal steady-state flow on hydraulic head and conductivity through geostatistical inversion and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Shlomo P. Neuman

Thomas Maddock III

Arthur W. Warrick

Donald E. Myers

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Shlomo P. Neuman
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Signed: [Signature]
ACKNOWLEDGEMENTS

I am indebted to my dissertation advisor, Regent’s Professor Shlomo P. Neuman, for his guidance, support, and patience in the long journey to completing my doctoral program. The lessons I have learned from him in the making of this dissertation are invaluable and will undoubtedly benefit my future career. I particularly appreciated his way to approach research and express findings effectively. I am especially grateful to him for suggesting me to work in the fascinating topic of my dissertation.

I would also like to thank Professors Donald E. Myers, Thomas Maddock III, and Arthur W. Warrick for serving on both my oral comprehensive and defense examination committees as well as providing me with helpful comments on my research. I have particularly benefited from the two courses on geostatistics taught by Professor Myers and from his kind assistance whenever I went to his office looking for advice.

I must thank Professor Jesús Carrera for serving on my oral comprehensive exam and teaching me important concepts on the inverse problem and adjoint state method during his sabbatical semester at The University of Arizona. I would also like to thank Professor Alberto Guadagnini for his kind help in the numerical implementation of the finite-element nonlocal flow equations.

My deepest thanks go to my parents, María Antonieta and Francisco, for their caring love and encouragement over the years; my brothers, Paco, Luis, and Alejandro, for being always my best friends; all my nephews, nieces, and sisters-in-law for their love; and the memory of my aunt Mamá Lila for her caring love.

My especial thankfulness goes to my stepdaughter Ivonne Melissa for her nourishing love and encouragement both in Tucson and from Mexico. I must also acknowledge Ms. Carmen Palacios for her support during part of my doctoral studies.

I would like to thank my best friend Ariel Carrasco and his wife, Flor, for supporting me in so many ways during my studies; Adrián Espinola, Silvia Madrid, Orna Amir, and Maricela Trujillo for their friendship and help during my Ph.D. program; Reynaldo Alanís and Emilia Jiménez for being such good friends in good and bad times; Gema Mercado for her emotional support and encouragement; and my friends at the former Grupo de Estudios Oceanográficos of the Instituto de Investigaciones Eléctricas (IIE) for encouraging me to pursue my doctorate.

I am grateful to my friends in the Department of Hydrology and Water Resources: Yunjung Hyun, Eric Morales, Alexander Tartakovsky, Ming Ye, Feyzan Misirli, and Wenbin Wang for having good times and rewarding discussions with them.

I would also like to acknowledge and thank M.C. Víctor Arellano and Dr. Ángel Fierros at the IIE for their support in extending my leave-of-absence whenever needed.

The majority of my Ph.D. program was supported by scholarships from CONACYT and IIE of Mexico. It was also supported in part by the National Science Foundation under NSF/ITR Grant EAR-0110289 during 2001-2002.
To my Parents

and

Ivonne
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>List of Figures</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>18</td>
</tr>
<tr>
<td>Abstract</td>
<td>21</td>
</tr>
<tr>
<td>Chapter 1. INTRODUCTION</td>
<td>23</td>
</tr>
<tr>
<td>1.1. Background</td>
<td>23</td>
</tr>
<tr>
<td>1.2. Stochastic (Forward) Models of Steady-State Flow</td>
<td>25</td>
</tr>
<tr>
<td>1.3. The Inverse Problem</td>
<td>28</td>
</tr>
<tr>
<td>1.4. Conditioning Stochastic Steady-State Flow on Hydraulic Head</td>
<td>32</td>
</tr>
<tr>
<td>1.4.1. Previous Works</td>
<td>33</td>
</tr>
<tr>
<td>1.4.2. Comparison of Inverse Methods</td>
<td>37</td>
</tr>
<tr>
<td>1.4.3. Conditional Simulations versus Conditional Moment Equations</td>
<td>38</td>
</tr>
<tr>
<td>1.5. Objective</td>
<td>39</td>
</tr>
<tr>
<td>Chapter 2. CONDITIONAL MOMENT EQUATIONS OF STEADY STATE FLOW</td>
<td>41</td>
</tr>
<tr>
<td>2.1. Overview</td>
<td>41</td>
</tr>
<tr>
<td>2.2. Background</td>
<td>42</td>
</tr>
<tr>
<td>2.3. Recursive Conditional Moment Equations</td>
<td>44</td>
</tr>
<tr>
<td>2.3.1. Recursive First Moment Equations</td>
<td>45</td>
</tr>
<tr>
<td>2.3.2. Recursive Second Moment Equations</td>
<td>47</td>
</tr>
<tr>
<td>2.4. Finite Element Moment Equations</td>
<td>49</td>
</tr>
<tr>
<td>2.4.1. Conditional Mean Head and Flux</td>
<td>50</td>
</tr>
<tr>
<td>2.4.2. Conditional Covariance of Hydraulic Head</td>
<td>54</td>
</tr>
<tr>
<td>2.4.3. Conditional Covariance Tensor of Flux</td>
<td>56</td>
</tr>
<tr>
<td>2.5. Computational Approach</td>
<td>57</td>
</tr>
<tr>
<td>2.5.1. Mean Head and Flux</td>
<td>57</td>
</tr>
<tr>
<td>2.5.2. Head and Flux Covariances</td>
<td>58</td>
</tr>
<tr>
<td>2.5.3. Efficiency of the Numerical Algorithm</td>
<td>59</td>
</tr>
<tr>
<td>2.5.4. Predictive Head and Flux Errors upon Inversion</td>
<td>60</td>
</tr>
<tr>
<td>2.6. Localization of Conditional Mean Flow</td>
<td>61</td>
</tr>
<tr>
<td>Chapter 3. GEOSTATISTICAL ESTIMATION OF LOG HYDRAULIC CONDUCTIVITY</td>
<td>64</td>
</tr>
<tr>
<td>3.1. Statistical Structure of Log Hydraulic Conductivity</td>
<td>64</td>
</tr>
<tr>
<td>3.1.1. Variogram and Statistical Parameters</td>
<td>66</td>
</tr>
</tbody>
</table>
Table of Contents—Continued

3.1.2. Prior Statistical Parameters ........................................... 67
3.1.3. Log $K$ Measurements .................................................. 68
3.2. Prior Kriging ...................................................................... 69
  3.2.1. Prior Kriging Algorithm ................................................. 69
3.3. Posterior Kriging ............................................................. 72
  3.3.1. Extended Measurement Vector $Y_E$ ............................... 72
  3.3.2. Correlation among $Y_E$ Errors ..................................... 74
  3.3.3. Posterior Kriging Algorithm .......................................... 76
  3.3.4. Computational Implementation of Posterior Kriging Algorithm 79

Chapter 4. Geostatistical Inverse Method ............................... 81
  4.1. Inverse Problem ............................................................ 81
  4.2. Statistics of Errors ........................................................ 83
  4.3. Nonlinear Parameter Estimation ....................................... 86
    4.3.1. Maximum Likelihood Formulation ............................... 86
    4.3.2. Simplification of Likelihood Function ......................... 87
    4.3.3. Minimization Algorithm ........................................... 89
    4.3.4. Calculation of Jacobian ........................................... 92
    4.3.5. Inverse Estimation of Statistical Parameters ................ 93
    4.3.6. Linearized Error Analysis of Parameter Estimates .......... 95
    4.3.7. Assessment of Statistical Assumptions and Parameter Model 97
  4.4. Identification of Parameter Model .................................... 98
    4.4.1. Discrimination Criteria incorporating Likelihood .......... 98
    4.4.2. Quality Criteria using Statistical and Graphical Analyses 99
    4.4.3. Usage of Discrimination and Quality Criteria ............... 101
  4.5. Sensitivity Analysis of Inverse Model .............................. 101

Chapter 5. Numerical Examples of Geostatistical Inversion .......... 104
  5.1. Reference Fields of $Y$, Head, and Flux ........................... 105
  5.2. Case 1: Conditioning only on $Y$ ...................................... 108
  5.3. Case 2: Conditioning only on Head ................................... 125
  5.4. Case 3: Conditioning on both Head and $Y$ without Priors .... 145
  5.5. Case 4: Conditioning on both Head and $Y$ with Priors ......... 163
  5.6. Case 5: Conditioning on both Head and $Y$, considering $Y$ Measurements as Parameters .............................................. 181
  5.7. Identification of Model .................................................. 198
# Table of Contents—Continued

## Chapter 6. Sensitivity Analysis

- 6.1. Number of Pilot Points \( P \) ......................................................... 203
- 6.2. Spatial Distribution of Pilot Points .................................................. 207
- 6.3. Number of Head Measurements \( J \) ..................................................... 212
  - 6.3.1. Small Number of Head Measurements ......................................... 216
- 6.4. Number of Log Conductivity Measurements \( I \) ................................. 220
- 6.5. Unconditional Variance of Log Hydraulic Conductivity \( \sigma_{y_u}^2 \) .......... 223
- 6.6. (Isotropic) Integral Scale \( \rho \) .......................................................... 227
- 6.7. Presence of a conditioning Pumping Well ......................................... 233
- 6.8. Comparative Analysis across Cases in Previous Sections .................... 234

## Chapter 7. Conclusions ................................................................. 238

## Appendix A. Derivation of Universal Point Kriging Expressions with Correlated Measurements

- A.1. Kriging Estimator ........................................................................... 244
- A.2. Kriging Variance and Covariance .................................................... 245

## References ....................................................................................... 248
LIST OF FIGURES

FIGURE 2.1. Hydrogeologic domain with finite element grid. .................... 49

FIGURE 5.1. Reference (unconditional) realization of \( Y(x) \) with \( \sigma_{Y_u}^2 = 1 \). .... 106

FIGURE 5.2. Reference head field corresponding to the unconditional \( Y(x) \) realization with \( \sigma_{Y_u}^2 = 1 \). ................................................. 106

FIGURE 5.3. Reference field of flux corresponding to the unconditional \( Y(x) \) realization with \( \sigma_{Y_u}^2 = 1 \): (a) longitudinal and (b) transverse components. 107

FIGURE 5.4. Hydrogeological domain with layout of \( Y \) measurement points \((x)\) for Case 1. A pumping well (●) with \( Q = 1 \) is located at the center. 108

FIGURE 5.5. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1. .................................................. 110

FIGURE 5.6. (a) Histogram of all \( Y \) residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1. .................................................. 110

FIGURE 5.7. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1. .................................................. 111

FIGURE 5.8. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1. .................................................. 111

FIGURE 5.9. Fields of conditional (a) \( \langle Y(x) \rangle_c \) and (b) \( \sigma_{Y}^2(x) \) for Case 1. (c) Longitudinal and (d) transverse sections of \( \sigma_{Y}^2(x) \). .......................... 112

FIGURE 5.10. (a) Contours of (forward) conditional mean head \( \langle h(x) \rangle_c \) for Case 1. Sections of forward mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), and (e) \( x_1 = 11 \). ................................. 114

FIGURE 5.11. (a) Contours of up-to second-order approximation of conditional head variance \( \sigma_{h}^2 \) for Case 1; (b) longitudinal and (c) transverse sections. 115

FIGURE 5.12. (a) Contours and (b) sections of conditional head covariance \( C_{h}^{[2]}(x,y) \) for reference point (●) \( y = (9,4) \). (c) Contours and (d) sections of conditional head covariance \( C_{h}^{[2]}(x,y) \) for reference point (●) \( y = (6.8,4.8) \). Case 1. ................................. 116

FIGURE 5.13. (a) Contours of (forward) conditional mean longitudinal flux \( \langle q_{x1}^{[2]}(x) \rangle_c \); sections of forward and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), and (e) \( x_1 = 10.9 \). (f) Contours of (forward) conditional mean transverse flux \( \langle q_{x2}^{[2]}(x) \rangle_c \); sections of forward and reference transverse fluxes at (g) \( x_2 = 3.9 \), (h) \( x_2 = 4.9 \), (i) \( x_1 = 8.9 \), and (j) \( x_1 = 10.9 \). Case 1. ................................. 118
LIST OF FIGURES—Continued

FIGURE 5.13. (continued) ................................................................. 119

FIGURE 5.14. (a) Contours of conditional variance of longitudinal flux $C_{q_{11}}^{(2)}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q_{22}}^{(2)}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q_{12}}^{(2)}(0) = C_{q_{21}}^{(2)}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 1. ................................................................. 120

FIGURE 5.14. (continued) ................................................................. 121

FIGURE 5.14. (continued) ................................................................. 122

FIGURE 5.15. (a)–(d) Sections of conditional flux covariance $C_{q_{ij}}^{(2)}(x,y)$ $(i,j = 1,2)$ with reference to point $y = (8.9,3.9)$. (e)–(h) Sections of conditional flux covariance $C_{q_{ij}}^{(2)}(x,y)$ $(i,j = 1,2)$ with reference to point $y = (6.9,4.9)$. Case 1. ................................................................. 123

FIGURE 5.16. Hydrogeological domain with layout of head measurement (+) and (numbered) pilot (P) points for Case 2. A pumping well (●) with $Q = 1$ is located at the center. ................................................................. 126

FIGURE 5.17. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 2. ................................................................. 128

FIGURE 5.18. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 2. 129

FIGURE 5.19. (a) Measured ($h^*$) and second-order approximation nonlocal mean ($h_m$) heads at measurement points; (b) head residuals ($h^* - h_m$) at the same points. Case 2. ................................................................. 129

FIGURE 5.20. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2. ................................................................. 130

FIGURE 5.21. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2. ................................................................. 130

FIGURE 5.22. (a) Histogram of all $Y$ residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2. ................................................................. 131

FIGURE 5.23. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2. ................................................................. 132

FIGURE 5.24. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2. ................................................................. 132
LIST OF FIGURES—Continued

FIGURE 5.25. Fields of conditional (a) $\langle Y(x) \rangle_c$ and (b) $\sigma^2_c(x)$ for Case 2. (c) Longitudinal and (d) transverse sections of $\sigma^2_c(x)$. 133

FIGURE 5.26. (a) Contours of (inverse) conditional mean head $\langle h^2(x) \rangle_c$ for Case 2. Sections of inverse mean and reference heads at (b) $x_2 = 4$, (c) $x_2 = 5$, (d) $x_1 = 9$, and (e) $x_1 = 11$. 134

FIGURE 5.27. (a) Contours of up-to second-order approximation of conditional head variance $\sigma^2_h$ for Case 2; (b) longitudinal and (c) transverse sections. 135

FIGURE 5.28. (a) Contours and (b) sections of conditional head covariance $C^2_h(x, y)$ for reference point (●) $y = (9,4)$. (c) Contours and (d) sections of conditional head covariance $C^2_h(x, y)$ for reference point (●) $y = (6.8,4.8)$. Case 2. 136

FIGURE 5.29. (a) Contours of (inverse) conditional mean longitudinal flux $\langle q^2_1(x) \rangle_c$, sections of inverse and reference longitudinal fluxes at (b) $x_2 = 3.9$, (c) $x_2 = 4.9$, (d) $x_1 = 8.9$, and (e) $x_1 = 10.9$. (f) Contours of (inverse) conditional mean transverse flux $\langle q^2_2(x) \rangle_c$, sections of inverse and reference transverse fluxes at (g) $x_2 = 3.9$, (h) $x_2 = 4.9$, (i) $x_1 = 8.9$, and (j) $x_1 = 10.9$. Case 2. 137

FIGURE 5.29. (continued) 139

FIGURE 5.30. (a) Contours of conditional variance of longitudinal flux $C^2_{q_11}(x)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C^2_{q_22}(x)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q_12}(x, y) = C^2_{q_11}(x)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 2. 140

FIGURE 5.30. (continued) 141

FIGURE 5.30. (continued) 142

FIGURE 5.31. (a)–(d) Sections of conditional flux covariance $C^2_{q_1j}(x, y)$ ($i, j = 1, 2$) with reference to point $y = (8.9,3.9)$. (e)–(h) Sections of conditional flux covariance $C^2_{q_2j}(x, y)$ ($i, j = 1, 2$) with reference to point $y = (6.9,4.9)$. Case 2. 143

FIGURE 5.32. Hydrogeological domain with layout of head (+) and $Y$ (x) measurement points as well as (numbered) pilot points ([P]) for Case 3. A pumping well (●) with $Q = 1$ is located at the center. 145

FIGURE 5.33. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 3. 147

FIGURE 5.34. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 3. 147
LIST OF FIGURES—Continued

FIGURE 5.35. (a) Measured \( h^* \) and second-order approximation nonlocal mean \( h_m \) heads at measurement points; (b) head residuals \( h^* - h_m \) at the same points. Case 3. 148

FIGURE 5.36. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3. 148

FIGURE 5.37. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3. 149

FIGURE 5.38. (a) Histogram of all head residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3. 149

FIGURE 5.39. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3. 150

FIGURE 5.40. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3. 150

FIGURE 5.41. Fields of conditional (a) \( \langle Y(x) \rangle_c \) and (b) \( \sigma^2_Y(x) \) for Case 3. (c) Longitudinal and (d) transverse sections of \( \sigma^2_Y(x) \). 151

FIGURE 5.42. (a) Contours of (inverse) conditional mean head \( \langle h(x) \rangle_c \) for Case 3. Sections of inverse mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), (e) \( x_1 = 11 \), and (f) \( x_1 = 10 \). 152

FIGURE 5.42. (continued) 153

FIGURE 5.43. (a) Contours of up-to second-order approximation of conditional head variance \( \sigma^2_h \) for Case 3; (b) longitudinal and (c) transverse sections. 153

FIGURE 5.44. (a) Contours and (b) sections of conditional head covariance \( C_h^{[2]}(x,y) \) for reference point (●) \( y = (9,4) \). (c) Contours and (d) sections of conditional head covariance \( C_h^{[2]}(x,y) \) for reference point (●) \( y = (6.8,4.8) \). Case 3. 154

FIGURE 5.45. (a) Contours of (inverse) conditional mean longitudinal flux \( \langle q_{11}(x) \rangle_c \); sections of inverse and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), (e) \( x_1 = 10.9 \), and (f) \( x_1 = 9.9 \). (g) Contours of (inverse) conditional mean transverse flux \( \langle q_{21}(x) \rangle_c \); sections of inverse and reference transverse fluxes at (h) \( x_2 = 3.9 \), (i) \( x_2 = 4.9 \), (j) \( x_1 = 8.9 \), (k) \( x_1 = 10.9 \), and (l) \( x_1 = 9.9 \). Case 3. 155

FIGURE 5.45. (Continued) 156

FIGURE 5.45. (Continued) 157
LIST OF FIGURES—Continued

FIGURE 5.46. (a) Contours of conditional variance of longitudinal flux $C_{q11}^{[2]}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q22}^{[2]}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q12}^{[2]}(0) = C_{q21}^{[2]}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 3. .......................................................... 158

FIGURE 5.46. (Continued) .................................................................................................................. 159

FIGURE 5.46. (Continued) .................................................................................................................. 160

FIGURE 5.47. (a)–(d) Sections of conditional flux covariance $C_{qij}^{[2]}(x,y) (i,j = 1,2)$ with reference to point $y = (8.9,3.9)$. (e)–(h) Sections of conditional flux covariance $C_{qij}^{[2]}(x,y) (i,j = 1,2)$ with reference to point $y = (6.9,4.9)$. Case 3. .......................................................... 162

FIGURE 5.48. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 4. .......................................................... 164

FIGURE 5.49. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 4. 164

FIGURE 5.50. (a) Measured $(h^*)$ and second-order approximation nonlocal mean $(h_m)$ heads at measurement points; (b) head residuals $(h^* - h_m)$ at the same points. Case 4. .......................................................... 165

FIGURE 5.51. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4. .......................................................... 165

FIGURE 5.52. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4. .......................................................... 166

FIGURE 5.53. (a) Rotated prior and estimated In hydraulic conductivity for each eigenvalue of covariance matrix $S_Y$; (b) In $K$ residuals at the same points. Case 4. .......................................................... 166

FIGURE 5.54. (a) Histogram of rotated $Y$ residuals corresponding to the eigenvalues of matrix $S_Y$ and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4. .......................................................... 166

FIGURE 5.55. (a) Histogram of all $Y$ residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4. .......................................................... 167

FIGURE 5.56. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4. .......................................................... 167
LIST OF FIGURES—Continued

FIGURE 5.57. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4. ............................................................. 168

FIGURE 5.58. Fields of conditional (a) \( \langle Y(x) \rangle_c \) and (b) \( \sigma^2_q(x) \) for Case 4. (c) Longitudinal and (d) transverse sections of \( \sigma^2_q(x) \). ............................................................. 169

FIGURE 5.59. (a) Contours of (inverse) conditional mean head \( \langle h^{[1]}(x) \rangle_c \) for Case 4. Sections of inverse mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), (e) \( x_1 = 11 \), and (f) \( x_1 = 10 \). .................. 170

FIGURE 5.59. (continued) ............................................................. 171

FIGURE 5.60. (a) Contours of up-to second-order approximation of conditional head variance \( \sigma^2_h \) for Case 4; (b) longitudinal and (c) transverse sections. 171

FIGURE 5.61. (a) Contours and (b) sections of conditional head covariance \( C^2_h(x,y) \) for reference point \( \bullet \) \( y = (9,4) \). (c) Contours and (d) sections of conditional head covariance \( C^2_h(x,y) \) for reference point \( \bullet \) \( y = (6.8,4.8) \). Case 4. ............................................................. 172

FIGURE 5.62. (a) Contours of (inverse) conditional mean longitudinal flux \( \langle q^{[1]}_l(x) \rangle_c \); sections of inverse and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), (e) \( x_1 = 10.9 \), and (f) \( x_1 = 9.9 \). (g) Contours of (inverse) conditional mean transverse flux \( \langle q^{[2]}_t(x) \rangle_c \); sections of inverse and reference transverse fluxes at (h) \( x_2 = 3.9 \), (i) \( x_2 = 4.9 \), (j) \( x_1 = 8.9 \), (k) \( x_1 = 10.9 \), and (l) \( x_1 = 9.9 \). Case 4. .................. 173

FIGURE 5.62. (Continued) ............................................................. 174

FIGURE 5.62. (Continued) ............................................................. 175

FIGURE 5.63. (a) Contours of conditional variance of longitudinal flux \( C^2_{q_{111}}(0) \); (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux \( C^2_{q_{12}}(0) \); (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance \( C^2_{q_{12}}(0) = C^2_{q_{21}}(0) \) between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 4. ............................................................. 176

FIGURE 5.63. (Continued) ............................................................. 177

FIGURE 5.63. (Continued) ............................................................. 178

FIGURE 5.64. (a)–(d) Sections of conditional flux covariance \( C^2_{q_{ij}}(x,y) \) \( (i,j = 1,2) \) with reference to point \( y = (8.9,3.9) \). (e)–(h) Sections of conditional flux covariance \( C^2_{q_{ij}}(x,y) \) \( (i,j = 1,2) \) with reference to point \( y = (6.9,4.9) \). Case 4. .................. 180

FIGURE 5.65. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 5. .................. 182
Figure 5.66. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 5.

Figure 5.67. (a) Measured \( h^* \) and second-order approximation nonlocal mean \( h_n \) heads at measurement points; (b) head residuals \( h^* - h_n \) at the same points. Case 5.

Figure 5.68. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

Figure 5.69. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

Figure 5.70. (a) Prior \( Y_{PE} \) and estimated \( \hat{Y}_{PE} \) ln hydraulic conductivity at pilot points; (b) ln K residuals \( Y_{PE} - \hat{Y}_{PE} \) at the same points. Case 5.

Figure 5.71. (a) Histogram of Y residuals at pilot points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

Figure 5.72. (a) Histogram of all Y residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

Figure 5.73. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

Figure 5.74. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

Figure 5.75. Fields of conditional (a) \( \langle Y(x) \rangle_c \) and (b) \( \sigma^2_c(x) \) for Case 5. (c) Longitudinal and (d) transverse sections of \( \sigma^2_c(x) \).

Figure 5.76. (a) Contours of (inverse) conditional mean head \( \langle h^{[2]}(x) \rangle_c \) for Case 5. Sections of inverse mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), and (e) \( x_1 = 11 \).

Figure 5.77. (a) Contours of up-to second-order approximation of conditional head variance \( \sigma_h^2 \) for Case 5; (b) longitudinal and (c) transverse sections.

Figure 5.78. (a) Contours and (b) sections of conditional head covariance \( C^{[2]}_h(x,y) \) for reference point (●) \( y = (9,4) \). (c) Contours and (d) sections of conditional head covariance \( C^{[2]}_h(x,y) \) for reference point (●) \( y = (6.8,4.8) \).
LIST OF FIGURES—Continued

FIGURE 5.79. (a) Contours of (inverse) conditional mean longitudinal flux \( \langle q_{x1}^{[2]}(x) \rangle_c \); sections of inverse and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), and (e) \( x_1 = 10.9 \). (f) Contours of (inverse) conditional mean transverse flux \( \langle q_{x2}^{[2]}(x) \rangle_c \); sections of inverse and reference transverse fluxes at (g) \( x_2 = 3.9 \), (h) \( x_2 = 4.9 \), (i) \( x_1 = 8.9 \), and (j) \( x_1 = 10.9 \). Case 5. 192

FIGURE 5.79. (continued) 193

FIGURE 5.80. (a) Contours of conditional variance of longitudinal flux \( C_{q_{x1}^{[2]}}(0) \); (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux \( C_{q_{x2}^{[2]}}(0) \); (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance \( C_{q_{x1}^{[2]}q_{x2}^{[2]}}(0) \) between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 5. 194

FIGURE 5.80. (continued) 195

FIGURE 5.80. (continued) 196

FIGURE 5.81. (a)–(d) Sections of conditional flux covariance \( C_{q_{ij}^{[2]}(x,y)}(i,j = 1, 2) \) with reference to point \( y = (8.9, 3.9) \). (e)–(h) Sections of conditional flux covariance \( C_{q_{ij}^{[2]}(x,y)}(i,j = 1, 2) \) with reference to point \( y = (6.9, 4.9) \). Case 5. 197

FIGURE 6.1. Hydrogeological domain with layout of eight pilot points \( (P) \). A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 204

FIGURE 6.2. Hydrogeological domain with layout of 32 pilot points \( (P) \). A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 204

FIGURE 6.3. Hydrogeological domain with layout \( (\pi_2) \) of 32 pilot points \( (P) \) positioned in zones of either relative maxima, minima, or large gradients of \( \ln K \). A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 208

FIGURE 6.4. Hydrogeological domain with layout \( (\pi_3) \) of pilot points \( (P) \) located at head-measurement points. A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 208

FIGURE 6.5. Hydrogeological domain with layout \( (\pi_4) \) of 32 pilot points \( (P) \) placed randomly in the interiors of evenly spaced subdomains of \( 2 \times 2 \) dimensions each. A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 209

FIGURE 6.6. Hydrogeological domain with layout \( (\pi_5) \) of 32 pilot points \( (P) \) located randomly. A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 209

FIGURE 6.7. Hydrogeological domain with layout of 18 head-measurement \( (+) \) and 16 pilot \( (\bar{P}) \) points. A pumping well \( (\bullet) \) with \( Q = 1 \) is located at the center. 213
LIST OF FIGURES—Continued

FIGURE 6.8. Hydrogeological domain with layout of 72 head-measurement (+) and 16 pilot (P) points. A pumping well (●) with \( Q = 1 \) is located at the center. ................................................................. 213

FIGURE 6.9. Hydrogeological domain with layout of five head-measurement (+) and four pilot (P) points. A pumping well (●) with \( Q = 1 \) and a (hidden) head-measurement point are located at the center. .................. 216

FIGURE 6.10. Hydrogeological domain with layout of nine head-measurement (+) and four pilot (P) points. A pumping well (●) with \( Q = 1 \) and a (hidden) head-measurement point are located at the center. ............... 217

FIGURE 6.11. Hydrogeological domain with layout of nine head-measurement (+) and eight pilot (P) points. A pumping well (●) with \( Q = 1 \) and a (hidden) head-measurement point are located at the center. ............... 217

FIGURE 6.12. Hydrogeological domain with layout of twelve pilot points (P). A pumping well (●) with \( Q = 1 \) is located at the center. .................. 220

FIGURE 6.13. Hydrogeological domain with layout of twelve pilot (P) and 16 Y-measurement (x) points. A pumping well (●) with \( Q = 1 \) is located at the center. ................................................................. 221

FIGURE 6.14. Hydrogeological domain with layout of twelve pilot (P) and 32 Y-measurement (x) points. A pumping well (●) with \( Q = 1 \) is located at the center. ................................................................. 222

FIGURE 6.15. Reference (a) unconditional realization of \( Y \) with \( \sigma^2_Y = 2 \) and (isotropic) \( \rho = 1 \). (b) Associated reference head field considers boundary conditions and pumping well with \( Q = 1 \). ................................. 223

FIGURE 6.16. Reference (a) unconditional realization of \( Y \) with \( \sigma^2_Y = 4 \) and (isotropic) \( \rho = 1 \). (b) Associated reference head field considers boundary conditions and pumping well with \( Q = 1 \). ................................. 224

FIGURE 6.17. Reference (a) unconditional realization of \( Y \) with (isotropic) \( \rho = 4 \) and \( \sigma^2_Y = 1 \). (b) Associated reference head field considers boundary conditions and pumping well with \( Q = 1 \). ................................. 225
LIST OF TABLES

TABLE 5.1. Y “measurements” extracted from the reference field. .......... 109
TABLE 5.2. Grid summary statistics of computed residuals and variances for Case 1. .................................................. 125
TABLE 5.3. Pilot point estimates obtained for Case 2. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates. 127
TABLE 5.4. Grid summary statistics of computed residuals and variances for Case 2. .................................................. 144
TABLE 5.5. Pilot point estimates obtained for Case 3. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates. 146
TABLE 5.6. Grid summary statistics of computed residuals and variances for Case 3. .................................................. 161
TABLE 5.7. Pilot point estimates obtained for Case 4. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates. 163
TABLE 5.8. Grid summary statistics of computed residuals and variances for Case 4. .................................................. 179
TABLE 5.9. Pilot point estimates obtained for Case 5. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates. 181
TABLE 5.10. Grid summary statistics of computed residuals and variances for Case 5. .................................................. 195
TABLE 5.11. Summary statistics of residuals and variances computed on the grid for one forward (Case 1) and four inverse cases. Ranking of each case is shown in parentheses for proper statistics. ......................... 200
TABLE 5.11. (continued) .................................................. 201
TABLE 5.12. Calibration, residual, and likelihood statistics. Ranking of each case is shown in parentheses for proper statistics. All cases consider $J = 36$ and $P = 16$. N.A. stands for “Not Applicable”. ......................... 202
TABLE 6.1. Calibration statistics for three cases with different number of pilot points. .................................................. 205
TABLE 6.2. Grid summary statistics of computed residuals and variances for three cases having different number of pilot points. ......................... 206
TABLE 6.3. Calibration statistics for five cases having different distributions of 32 pilot points. .................................................. 210
TABLE 6.4. Grid summary statistics of computed residuals and variances for five cases having different distributions of 32 pilot points. ......................... 211
TABLE 6.5. Calibration statistics for three cases with different numbers of head measurements. .................................................. 214
TABLE 6.6. Grid summary statistics of computed residuals and variances for three cases with different numbers of head measurements. ......................... 215
### List of Tables—Continued

**Table 6.7.** Calibration statistics for three cases having a small number of head measurements. ........................................ 218

**Table 6.8.** Grid summary statistics of computed residuals and variances for three cases having a small number of head measurements. ................ 219

**Table 6.9.** Calibration statistics for three cases with different numbers of ln $K$ measurements. .................................................. 221

**Table 6.10.** Grid summary statistics of computed residuals and variances for three cases with different numbers of ln $K$ measurements. ................ 222

**Table 6.11.** Calibration statistics for three cases having different variance of $Y$; each case performed twice as in Cases 3 and 4 in Chapter 5. ....... 224

**Table 6.12.** Grid summary statistics of computed residuals and variances for three cases having different variance of $Y$; each case performed twice as in Cases 3 and 4 in Chapter 5. ........................................ 228

**Table 6.12.** (continued) ............................................................................................................................... 229

**Table 6.13.** Calibration statistics for two cases with different $Y$ integral scales. ....................................................... 230

**Table 6.14.** Grid summary statistics of computed residuals and variances for two cases with different $Y$ integral scales. .......................... 232

**Table 6.15.** Calibration statistics for two cases with and without the presence of a pumping well with $Q = 1$ at (9, 4). ..................... 234

**Table 6.16.** Grid summary statistics of computed residuals and variances for two cases with and without the presence of a pumping well with $Q = 1$ at (9, 4). .......................................................... 235
ABSTRACT

Nonlocal moment equations allow one to render optimum predictions of flow in randomly heterogeneous media deterministically conditional on measured values of medium properties and to assess the corresponding predictive uncertainty. I present a geostatistical inverse algorithm for steady-state flow that makes it possible to further condition such predictions and assessments on measured values of hydraulic head and (or) flux. My algorithm is based on recursive finite-element approximations of exact first and second conditional moment equations. Computational efficiency is enhanced through the use of a direct sparse matrix solver. Hydraulic conductivity is parameterized via universal kriging based on unknown values at pilot points and (optionally) measured values at other discrete locations. Correlation among parameter estimates (or priors) is considered in the universal kriging equations. Optimum unbiased inverse estimates of natural log hydraulic conductivity, head and flux are obtained by minimizing a calibration criterion, composed of residuals of head or (and) flux and (possibly) log conductivity, using the Levenberg-Marquardt algorithm. Statistical parameters characterizing the natural variability of hydraulic conductivity can also be estimated using this algorithm. I illustrate the method for superimposed mean uniform and convergent flows in a bounded two-dimensional domain under various conditions for a range of parameters. My examples illustrate how conductivity and head data act separately or jointly to reduce parameter estimation errors and model
predictive uncertainty. Over-parameterization is seen to create zones of high mean conductivity, in which flux prediction is more uncertain than is in other regions. It is found that a regular distribution of pilot points works better than does an irregular layout and that the number of pilot points should be as close as possible to the number of head data while maintaining parameters reasonably uncorrelated. Head and flux predictions are very satisfactory for cases with either log conductivity variance or integral scale between one and four, though prediction quality deteriorates with either larger variances or shorter integral scales. The method may perform satisfactorily in cases with no conductivity measurements and only a few head data.
Chapter 1
INTRODUCTION

1.1 Background

Environmental subsurface studies include the prediction of groundwater flow and transport of contaminants at sites that need to be remediated or where anthropogenic waste is to be stored. Other hydrogeological studies include flow prediction in the vicinity of civil constructions such as dams or in protection zones of drinking water. Models of steady state flow are often used as a first step toward building more complex models such as those describing transport or transient flow. Hydraulic conductivity is the main controlling parameter in subsurface flow. Yet it can be estimated with only high uncertainty in heterogeneous media (i.e., those exhibiting significant spatial variability in hydraulic conductivity). This renders model predictions of head and flux uncertain. Traditional deterministic numerical models [de Marsily, 1986; Bear and Verruijt, 1987] developed in the twentieth century can represent flow under different field conditions in irregularly bounded domains but cannot assess predictive uncertainty. In the 1970s stochastic hydrogeology emerged as the discipline that best fulfills this necessity. Stochastic approaches include analytical and numerical techniques leading to flow solutions in heterogeneous, unbounded or bounded domains [Dagan, 1989; Dagan and Neuman, 1997; Zhang, 2001].
Spatially varying hydraulic conductivity may be "measured" in situ through pumping, injection, packer, or slug tests [de Marsily 1986; Batu, 1998]. In these tests, groundwater is either pumped or injected (steadily or suddenly) and hydraulic head (piezometric) as well as flux rate data are recorded. The conductivity measurements are then inferred indirectly or inversely from these data using graphical (or computational) procedures or semi-empirical relations. The former procedures (such as the Jacob or Theis methods) are based on mathematical flow models of (confined or unconfined) aquifers having effective (homogeneous) conductivities, or transmissivities, and simple geometries; the models may also account for leakage or well effects. A conductivity measurement is assigned to a discrete point (i.e., the center of the test) in the field but represents the average over the test volume.

A hydraulic conductivity field, $K(x)$, can then be spatially estimated from conductivity measurements obtained in such tests using a zonation, interpolation, or geostatistical approach (see page 29). Whereas in the first two procedures $K(x)$ is viewed as a deterministic spatial function, in the geostatistical approach it is considered to be multivariate (autocorrelated) random field. Applying geostatistical methods such as kriging or (un)conditional simulations requires conductivity data in enough number and quality to infer statistical parameters of the variogram or covariance of $K(x)$. Estimating $K(x)$ in any of the above manners entails conditioning the resulting flow predictions on conductivity data.

Regarding $K(x)$ as a multivariate (auto-uncorrelated or autocorrelated) random
field renders the governing flow equations stochastic. These can be cast and solved in terms of statistical moments of ensembles of equally likely solutions (or realizations) for head and flux in probability space. The transformed equations are termed "stochastic moment equations". Only the first two moments are usually considered: the mean and variance-covariance. While the former represents a (possibly, unbiased and optimum) prediction, the latter provides a measure of uncertainty-(spatial) correlation. Another way of solving the stochastic equations is by means of Monte Carlo simulations. In such a method, a large (ideally infinite) number (in the order of hundreds or thousands) of realizations of $K(x)$ are generated and corresponding head and flux solutions, for each, are obtained through the stochastic flow equations. The probability density functions (pdf) for these variables or their statistical (ensemble) moments may then be derived approximately in a straightforward, though computationally expensive, manner. Both approaches may yield unconditional or conditional (on measured conductivity) solutions. If conductivity data are available, conditional solutions may be used to reduce predictive uncertainty throughout the field and especially in the neighborhoods of measurement (or conditioning) points.

1.2 Stochastic (Forward) Models of Steady-State Flow

Next, I briefly describe some models that are detailed in the texts of Dagan [1989], Dagan and Neuman [1997], and Zhang [2001]. The conditional methods of (forward) solution described therein consider conditioning only on hydraulic conductivity data.
In Section 1.4, I describe works that condition parameters and flow solutions on head and/or flux measurements through inversion.

Analytical solutions are commonly obtained by expanding them in terms of $\sigma_Y^2 \ll 1$, the variance of log hydraulic conductivity $Y = \ln K$, with $Y$ constrained to be statistically homogeneous Gaussian with an exponential covariance function [Dagan, 1989]. Only leading moments are usually considered; some works further neglect the moments of cross products between purely random quantities. These flow solutions are given for unbounded domains under a hydraulic gradient of uniform mean; only a few studies contemplate uniform mean flow in bounded domains.

A more powerful and straightforward approach is to perform (unconditional or conditional) Monte Carlo simulations that can yield solutions for strongly heterogeneous ($\sigma_Y^2 \geq 1$) or statistically nonhomogeneous formations in domains of arbitrary geometry. However, Monte Carlo solutions have many sources of potential error and as media heterogeneity increases need both increasingly finer computational grids to insure precision and larger number of realizations to attain convergence (if achievable) [Neuman, 1997; Zhang, 2001].

A numerical alternative to Monte Carlo simulations consists of solutions of moment equations. Zhang [1998] obtained unconditional equations for the mean and variance-covariance of hydraulic head for steady state flow in statistically-nonhomogeneous, heterogeneous, bounded media using perturbation expansions to first order in $\sigma_Y^2$. He solved the the unconditional moment equations on a two-
dimensional finite-difference grid for statistically homogeneous and nonhomogeneous cases of log conductivity under mean uniform flow.

Guadagnini and Neuman [1999a] derived conditional moment equations for the first two statistical moments of head and flux based on a recursive closure approximation of (implicit) exact moment equations through a perturbation expansion to first order in $\sigma^2$. They considered mean steady-state flow in statistically-nonhomogeneous, heterogeneous, bounded media. Both the mean and second-moment flow equations depend on a symmetric system of mean Green’s function equations. These equations are functions of the conditional mean log hydraulic conductivity, $\langle Y(x) \rangle_c$, and conditional variance-covariance of the same property, $\langle Y'(x)Y'(y) \rangle_c$, where $Y'(x) = Y(x) - \langle Y(x) \rangle_c$. The mean flow equations contain a residual-flux term that renders them integro-differential (i.e., nonlocal) and non-Darcian. They solved their moment equations and a localized version of them both unconditionally and conditionally on a finite element grid (their method of solution is not restricted to this scheme). Their numerical examples compared their first and second moment solutions with those derived from Monte Carlo simulations under superimposed mean-uniform and convergent flow regimes in two dimensions [Guadagnini and Neuman, 1999a, b].

Other types of stochastic techniques (less directly related to my work) for deriving and solving moment equations include spectral, vector space-state, and Adomian decomposition methods. Zhang [2001] provides an up-to-date introduction to them.
1.3 The Inverse Problem

In practice hydraulic conductivity measurements are seldom enough in both number and quality for estimating $K(x)$ in heterogeneous formations adequately. However, head and (or) flux data may help accomplish this through an inverse procedure; that is, by solving an inverse problem. This leads to conditioning both hydraulic parameter estimates and state variable predictions on hydraulic head and (or) flux data in addition to (possibly) conductivity measurements. Clifton and Neuman [1982] found that such a conditioning resulted in a very significant reduction in both parameter estimation and head prediction uncertainty for a real aquifer.

The objective for solving an inverse problem is to estimate hydraulic parameters (i.e., hydraulic conductivity, boundary head and/or flux, recharge, etc.) in the field and their associated estimation error structures. According to the classification of Neuman [1973], in this introduction I consider only the class of indirect inverse methods; those that are based on minimizing a norm of flow prediction error. In this class the parameter estimation procedure is usually posed in terms of (1) maximizing a likelihood criterion numerically (e.g., Hoeksema and Kitanidis [1984]) or analytically (e.g., Dagan [1985]), or (2) minimizing a (model-fit) objective function iteratively based on (forward) model predictions computed at each iteration (e.g., Cooley [1982], Carrera and Neuman [1986a]). The objective function is derived by nonlinear regression or maximum likelihood and consists of a weighted sum of squared
residuals. The latter are differences between state-variable measurements and their model prediction counterparts; a weighted sum of differences between prior and posterior estimates of parameters may be potentially included as a regularization term in the objective function. The weights are usually given by the inverse of standard deviations of measurement or prior errors, which may be correlated (e.g., Carrera and Neuman [1986a]). Parameter priors are given by measurements or surrogate values inferred from geological/geophysical information or kriged measurements.

A deterministic or a stochastic approach can be used to solve the inverse problem. In the first one \( K(x) \) is considered deterministic but treated statistically in the inverse formulation. In the stochastic approach \( K(x) \), hydraulic head, and flux are viewed as random; geostatistical methods may thus be used to estimate them spatially.

An inverse problem is usually solved for a small number of hydraulic and/or statistical parameters compared to the total number of available data for reasons given below. These parameters are defined through a parameterization scheme to characterize the spatial variability of \( K(x) \). They may represent (i) zones of uniform conductivity value (i.e., zonation; e.g., Carrera and Neuman [1986a, c]), (ii) coefficients of a low-order interpolation polynomial [Sun, 1994], (iii) node conductivity values on a (bilinear, cubic spline, etc.) interpolation grid [Sun, 1994], (iv) statistical parameters defining the natural \( K(x) \) (e.g., Hoeksema and Kitanidis [1984, Dagan [1985]), or (v) fictitious discrete conductivity measurements. In the latter two cases the hydraulic conductivity field is spatially estimated via kriging or geostatistical
simulation [Deutsch and Journel, 1997; Chilès and Delfiner, 1999]. Parameteriza-
tion of type (v) is known as the pilot point technique, originally proposed by de 
Marsily [1978] and later modified by RamaRao et al. [1995]). The latter version has 
been questioned recently by Cooley [2000] and Cooley and Hill [2000] and defended 
by RamaRao et al. [2000]. This parameterization technique has proven suitable for 
investigations [e.g., Gómez-Hernández et al., 1997; Franssen and Gómez-Hernández, 
2002] and successful in applications [e.g., Xue and Datta-Gupta, 1997; Zimmerman 
et al., 1998; Capilla et al., 1998; Doherty, 2001, 2002b]. It is potentially suitable for 
stochastic inverse methods because it can provide estimation variance-covariance.

Once a parameterization scheme is chosen, determining a parameterization (or 
parameter model) optimally or suboptimally is an important component of the inverse 
problem. This may be accomplished using suitable comparative criteria such as (a) 
those pertaining to model discrimination (e.g., Carrera and Neuman [1986a, c]), those 
proposed by Hill [1998]; or through an optimization procedure such as those related 
to observation network design (e.g., Carrera and Neuman [1986c], Chilès and Delfiner 
[1999]) or the one in the work of LaVenue and Pickens [1992].

Inverse problems are inherently ill-posed to some degree (which does not preclude 
one from attempting to solve them satisfactorily) [Yeh, 1986; Carrera and Neuman, 
1986b; Carrera, 1987; Ginn and Cushman, 1990; Sun, 1994; McLaughlin and Townley, 
1996]. This means that their solutions fail to fulfill at least one of the following condi-
tions: (a) existence or identifiability, (b) uniqueness, and (c) stability. In general, the
degree of ill-posedness depends on the model, available data, and parameterization. Improving any of these (most easily, parameterization) leads to better posed problems. Hill [1998] suggests practical (method-independent) guidelines to help render calibrations well-posed.

Difficulty in identifying an (accurate) inverse solution is caused mainly by ubiquitous head or flux measurement error (noise). This is circumvented if the errors are more-or-less normally distributed and the inverse problem is solved statistically. Including priors in a regularization or plausibility term in the objective function may also help attain a solution by constraining potential solutions in parameter space and, therefore, improving well-posedness. Nonuniqueness means that two or more sets of parameter values minimize the objective function equally; this is explained by the existence of multiple local or global minima in objective-function space. In such a case the solution obtained by an iterative inverse method may depend on the initial values starting the minimization process. Stability means that small variations in the data result in small changes in parameter estimates. Unstable solutions are associated with flat or elongated contours of the objective function, thus causing oscillations and slow convergence in the iterative process. Instability is probably the most common manifestation of ill-posedness in iterative methods and arises especially when there are insufficient number of data relative to the number of parameters. Reducing the number of parameters often eliminates instability. Working with log hydraulic conductivity (instead of hydraulic conductivity) insures positive estimates of $K$ and helps
alleviate the problem.

Ideally, a well-posed problem has a unique global minimum that is insensitive to variations in the data. Ill-posed problems often converge in many iterations and yield numerical solutions that are insensitive to parameters. An eigenanalysis of the covariance matrix of parameter estimation errors and computed linear confidence intervals of parameter estimates may provide information on the quality of the estimates.

Thorough discussions on the inverse problem and reviews of inverse methods can be found in the works by Yeh [1986], Carrera and Neuman [1986a, b], Carrera [1987], Ginn and Cushman [1990], Sun [1994], and McLaughlin and Townley [1996].

1.4 Conditioning Stochastic Steady-State Flow on Hydraulic Head

The first successful attempts to obtain indirect solutions of inverse problems automatically were pursued by using traditional deterministic flow models and minimizing an objective function (e.g., Neuman [1980], Cooley [1982], Carrera and Neuman [1986a]). Yet some of these works utilized kriging to interpolate $K(x)$ smoothly on the field (e.g., de Marsily [1984]) and to get parameter priors (e.g., Neuman and Yakowitz [1979]). A little later, linearized (geostatistical) stochastic inverse methods were introduced in the groundwater literature [Kitanidis and Vonvoris, 1983; Dagan, 1985; Gutjahr and Wilson, 1989]. These are based on a linear approximation of the relationship (i.e., the flow equation) between random hydraulic parameters and
head. Most recently inverse methods have been based on generated unconditional, or conditional, stochastic simulations of $K(x)$ that are at some point conditioned on state variable data through flow equations (e.g., Gutjahr et al. [1994], Ramarao et al. [1995], Gómez-Hernández et al. [1997], Oliver et al. [1997], Hanna and Yeh [1998], Hu [2000]). In conditioning conductivity simulations on head and/or flux data, the former are forced to approximately maintain the spatial variability of the natural $K(x)$ through the statistical (variogram or covariance) parameters.

Next, I briefly describe some of the most representative inverse methods in the groundwater literature.

1.4.1 Previous Works

*Carrera and Neuman* [1986a] introduced an inverse methodology based on maximum likelihood theory that allows one to estimate both hydraulic and statistical parameters. This is a unique characteristic of their method. The latter can be considered as encompassing the methods by Neuman and Yakowitz [1979], Neuman [1980], Cooley [1982], among others, as it reduces to nonlinear regression when statistical parameters are known or fixed in the minimization of the objective function. A parameterization model is identified by means of discrimination criteria, and flow prediction is performed using a traditional flow model. *Carrera and Neuman* [1986c] applied their method using zonation, and *Carrera et al.* [1997] utilized it with priors inferred from kriging.
Kitanidis and Vomvoris [1983] presented a numerical inverse formulation that aims to estimate a few statistical parameters characterizing both the covariance or variogram of statistically homogeneous log conductivity and its drift from hydraulic head and conductivity data in one dimension. Hoeksema and Kitandis [1984] extended his method to two-dimensional flow. The method is formulated using a maximum likelihood approach and a linear relationship between hydraulic head and conductivity. This approximation guarantees its application to heterogeneous fields with only $\alpha^2 < 1$. The latter relationship is used to obtain head covariance and cross covariance between heads and log conductivities in a numerical grid. Upon identifying the covariance model of $Y(x)$ via discrimination criteria, $Y(x)$ is estimated by cokriging all measurements, and head is determined by solving flow equations on a numerical grid. By having only a few parameters to estimate this formulation is numerically inexpensive, as compared to regression-like methods, and tends to be well posed.

Dagan [1985] developed a semianalytical method to estimate hydraulic conductivity, whose approach is similar to (but less numerically-demanding than) the one used by Kitanidis and Vomvoris [1983]. Rubin and Dagan [1987] modified it to account for random, uniform recharge. The differences with the method by Hoeksema and Kitandis [1984] are as follows. Dagan [1985] obtained head covariance and cross covariance between heads and log conductivities analytically, which limits the method’s applicability to domains with simple geometries and boundary conditions. Rubin and Dagan [1987] further assumed that mean head varies quadratically in space and
that log conductivity is statistically-homogeneous, multivariate normal. Mean head is obtained analytically in the domain by means of kriging, which avoids defining a numerical grid and introducing discretization error.

Gutjahr and Wilson [1989] proposed a linearized stochastic formulation in Fourier space, whose inverse mean head predictor is similar to the one given by Dagan [1985]. The head covariance and cross-covariance between estimates of $Y$ and head are obtained as spectral representations. Gutjahr et al. [1994] modified this method to render stochastic simulations of log hydraulic conductivity conditioned on both head and conductivity information. This method can be viewed as an extension of the linearized geostatistical methods [Hoeksema and Kitandis, 1984; Dagan, 1985; Gutjahr and Wilson, 1989] to handle nonlinear cases with $\sigma_Y^2 > 1$ iteratively. The jointly multivariate-normal hydraulic head and conductivity fields are unconditionally simulated through the obtained covariances [Gutjahr and Wilson, 1989]. Then, these fields are conditioned on head and conductivity data via cokriging, the forward problem is solved, and a fit between measured and computed heads is evaluated. Due to the linear approximation between head and conductivity, the covariances must be recalculated and the procedure repeated until a satisfactory match is attained.

RamaRao et al. [1995] modified the original pilot-point inverse method of de Marsily [1978] to obtain realizations of $Y(x)$ conditioned on head and conductivity measurements. This is done by first generating conditional (on $Y$ data) realizations of $Y(x)$ and then calibrating each on hydraulic head data. (In my pilot-point param-
eterization, I allow for priors in the calibration criterion.) Parameters act as fictitious \( Y \) measurements that are added sequentially during calibration and located optimally (on a numerical grid) at pilot points using adjoint state equations. The values of parameters are constrained by a multiple of the prior kriging error or an arbitrary limit. At the end of each iteration, parameters are assessed and the conditional realization is updated via kriging. This method is however computationally demanding; La Venue et al. [1995] used it obtain 70 conditional realizations of log conductivity in a real aquifer.

The method of Gómez-Hernández et al. [1997] was first outlined by Sahuquillo et al. [1992] and next applied by Capilla et al. [1997, 1998] and Franssen and Gómez-Hernández [2002]. It is similar to the pilot-point method of RamaRao et al. [1995] but differs with the latter in that: pilot or master points are arbitrarily located a priori; \( Y \) measurements are considered as parameters whose values are constrained by \( Y \) measurement errors; and convergence is attained in a few iterations based on a simplified calibration criterion, which renders this method faster than the pilot point method. Among all previous methods, this is the most alike to the one I present here. The main difference being in that while Gómez-Hernández et al. [1997] solve for individual (random) conditional realizations of an ensemble, I solve for the conditional mean and second moments of such an ensemble.
1.4.2 Comparison of Inverse Methods

Several analytical, conceptual, and numerical comparisons among inverse methods having different approaches have been presented by Kuiper [1986], Yeh [1986], Carrera and Neuman [1986a], Carrera [1987], Ginn and Cushman [1990], Carrera and Glorioso [1991], Keidser and Rosbjerg [1991], Ahmed and de Marsily [1993], Sun [1994], McLaughlin and Townley, 1996], Zimmerman et al. [1996], and Carrera et al. [1997]. The most recent and objective comparison to date is the one published by Zimmerman et al. [1997], who considered seven of the most representative geostatistical methods. While six of them are described above [Hoeksema and Kitandis, 1984; Rubin and Dagan, 1987; Carrera and Neuman, 1986a; Gutjahr et al., 1994; Ramarao et al., 1995; Gómez-Hernández et al., 1997], the method by Grindrod and Impey [1991], based on fractal simulations, yielded the least favourable results and is not considered here.

Some of the most important conclusions derived from the comparison by Zimmerman et al. [1997] are as follows. The methods by Rubin and Dagan [1987], Carrera and Neuman [1986a], Ramarao et al. [1995], and Gómez-Hernández et al. [1997] performed best, overall, among all methods and similarly among them in four different real-like synthetic test problems displaying variances of $\log_{10} K$ between 1.4 and 2.1. They also found that “... The experience and skill of the modeller and the time and effort spent on the modelling of the problem have been shown to be essential compo-
nents of success”. This indicates that it is difficult to draw definite conclusions about the superiority of any methodology. For example, the method by Rubin and Dagan [1987] was one of the best despite being unclear its assessment capabilities for cases with \( \sigma^2_T > 1 \); a similar conclusion was drawn by Carrera et al. [1997]. Zimmerman et al. [1997] also found that inferring satisfactorily the geostatistical structure of \( K(x) \) from both hydraulic conductivity and head data is at least as important as estimating parameter values adequately.

However, all methods failed to estimate predictive uncertainty satisfactorily. I hope this work contributes to fulfill this necessity.

1.4.3 Conditional Simulations versus Conditional Moment Equations

Parameter estimation using the stochastic simulation methods described above determines random \( K(x) \) conditioned on head and possibly conductivity data for each realization. Then, approximate ensemble mean and variance-covariance for conductivity, head, and flux can be computed as in Monte Carlo methods. Estimating these ensemble quantities with accuracy entails obtaining a large number of inversions. Yet, a net error covariance matrix for mean parameter estimates seems difficult to assess. In principle, the well-posedness of each inversion should be guaranteed as recommended by Hill [1998]; a task that is impractical for the burden of work it implies.

If stochastic moment equations are used in parameter estimation (something that
has not been reported in the literature yet), one needs to estimate both mean log conductivity and its variance-covariance (conditioned on head and possibly conductivity data) by means of a single inversion only. In this manner the covariance matrix of the mean-parameter-estimate errors can be assessed.

It appears that it is easier to assess quality of solutions through conditional moment equations than it is to do it by means of stochastic simulation methods.

1.5 Objective

The main objective of my dissertation is to estimate hydraulic conductivity and predict steady-state head and flux (as well as their corresponding measures of uncertainty) conditionally on hydraulic head and (or) flux data in addition to (possibly) conductivity measurements in randomly heterogeneous bounded domains based on nonlocal conditional moment equations of flow. This is accomplished by means of a geostatistical numerical inverse procedure described herein.

The major issues I address in my dissertation are: (1) conditioning the moment equations of Guadagnini and Neuman [1999a] not only on hydraulic conductivity but also or solely on hydraulic head and (or) flux; (2) establishing a numerical inverse methodology with which to accomplish this purpose; (3) improving the numerical performance of the finite-element flow simulator of Guadagnini and Neuman [1999a]; (4) defining a suitable parameterization to represent the transform of log hydraulic conductivity, $Y(x)$, in inversions; (5) delineating a geostatistical algorithm with which
to estimate conductivity, its transform, and the conditional spatial covariance of the latter prior to or during inversion; (6) assessing parameter estimation error; (7) predicting state variables conditioned on their measured values and estimating their corresponding predictive uncertainties; (8) obtaining or improving geostatistical parameter values through inverse estimates; (9) identifying suboptimal parameter models through discrimination and quality criteria; (10) evaluating and comparing numerical examples having different parameterizations, data types, and availability of hydraulic prior information; and (11) carrying out a sensitivity analysis to test the performance of the geostatistical inverse model under different field conditions or calibration parameter values.

In Chapter 2, I introduce the recursive moment flow equations of Guadagnini and Neuman [1999a], their finite element form, and their localized solution. I also address there issue (3) above. Chapter 3 presents the assumptions I make for the statistical structure of the hydraulic conductivity field, addresses the aforementioned issues (4) and (5), and gives some recommendations on the computational construction of the kriging algorithm. In Chapter 4, I address issues (2), (6), (8), and (9) above under a statistical framework defined therein. Chapters 5 and 6 treat the aforementioned issues (10) and (11), respectively. The most relevant conclusions of this dissertation are presented in Chapter 7.
Chapter 2
CONDITIONAL MOMENT EQUATIONS OF STEADY STATE FLOW

2.1 Overview

Guadagnini and Neuman [1999a, b] developed deterministic recursive expressions and corresponding finite element approximations for first (mean) and second (variance-covariance) statistical (ensemble) moments of steady state hydraulic head \( h(x) \) and Darcy flux vector \( q(x) \) based on an exact theory due to Neuman and Orr [1993a, b] and Neuman et al. [1996]. Stochasticity enters by treating hydraulic conductivity as a correlated random field and the source/sink, as well as boundary terms, as random variables. The hydraulic-head and flux moments are conditional on hydraulic conductivity. The first moments represent optimum unbiased predictions and the second moments provide measures of predictive uncertainty. The conditional moment equations are integro-differential (nonlocal). Conditioning requires inferring the statistical properties of log hydraulic conductivity, \( Y = \ln K \), from discrete measurements using geostatistical methods [Deutsch and Journel, 1998]. The conditional nature of the nonlocal flow equations leads to different solutions for a given physical setting if different sets of conductivity measurements are utilized. Hence, nonlocal flow solutions are nonunique.

The objective of this work is to develop an inverse method that allows conditioning
the above moment equations not only on hydraulic conductivity data but also (or only) on hydraulic head (and/or flux) data. For this purpose, I adopt the iterative maximum-likelihood approach of Carrera and Neuman [1986a, b] in its updated form [Carrera et al., 1997]. When relevant statistical parameters are known, this reduces to nonlinear regression [e.g., Cooley, 1982, 1985a]. Solution of the inverse problem yields parameter estimates conditioned on head (and/or flux) as well as (possibly) log hydraulic conductivity data. It also yields corresponding predictions of head and flux as well as measures of the associated prediction errors. Both the parameter estimates and the head and flux predictions are nonunique due to their dependence on a particular set of measurements.

In this chapter, I describe the recursive and finite-element moment equations of Guadagnini and Neuman [1999a, b], their computational solution, and prediction errors of head and flux.

2.2 Background

According to Guadagnini and Neuman [1999a], \( q(x) \) and \( h(x) \) obey the continuity equation

\[
-\nabla \cdot q(x) + f(x) = 0 \quad x \in \Omega \tag{2.1}
\]

and Darcy's law

\[
q(x) = -K(x)\nabla h(x) \quad x \in \Omega \tag{2.2}
\]
subject to Dirichlet

\[ h(x) = H(x) \quad x \in \Gamma_D \]  \hspace{1cm} (2.3)

and Neumann

\[ -q(x) \cdot n(x) = Q(x) \quad x \in \Gamma_N \]  \hspace{1cm} (2.4)

boundary conditions within a randomly nonuniform domain \( \Omega \). Here log hydraulic conductivity, \( Y(x) = K(x) \), is a random scalar function of space with a given spatial autocovariance function; \( f(x) \) is a random source/sink term (i.e., withdrawal or injection well, areal recharge, etc.); \( H(x) \) is random prescribed head on Dirichlet-type boundaries \( \Gamma_D \); \( Q(x) \) is random prescribed flux across Neumann-type boundaries \( \Gamma_N \); \( n(x) \) is a unit outward vector normal to the boundary \( \Gamma = \Gamma_D \cup \Gamma_N \); and \( f(x), H(x) \), and \( Q(x) \) are prescribed in a statistically independent manner.

All variables and parameters entering into Equations (2.1)–(2.4) are defined on a nonzero support volume \( \omega \), centered about \( x \), that is small compared to \( \Omega \) but large enough for Darcy’s law to be locally valid [Neuman and Orr, 1993a].

Due to the random nature of their coefficient and forcing terms, Equations (2.1)–(2.4) are stochastic. This renders \( h(x) \) and \( q(x) \) random. In general, the solution of Equations (2.1)–(2.4) is nonhomogeneous (nonstationary in space) due to the combined effect of arbitrary source/sink and boundary terms and conditioning.

\( \langle Y(x) \rangle_c \) denotes the conditional ensemble mean of \( Y(x) \). It represents an optimum (minimum estimation-error variance) unbiased (zero estimation-error mean) estimator that is smoother in space (due to stochastic averaging) than the unknown random
function $Y(x)$. $Y'(x)$ is a zero-mean fluctuation about the mean,

$$Y'(x) = Y(x) - \langle Y(x) \rangle_c \quad \langle Y'(x) \rangle_c = 0 \quad (2.5)$$

Similarly, zero-mean prediction errors of head and flux are defined as

$$h'(x) = h(x) - \langle h(x) \rangle_c \quad \langle h'(x) \rangle_c = 0 \quad (2.6)$$

$$q'(x) = q(x) - \langle q(x) \rangle_c \quad \langle q'(x) \rangle_c = 0 \quad (2.7)$$

where $\langle h(x) \rangle_c$ and $\langle q(x) \rangle_c$ represent optimum unbiased predictors of the generally unknown random $h(x)$ and $q(x)$, respectively.

### 2.3 Recursive Conditional Moment Equations

The recursive conditional moment equations of Guadagnini and Neuman [1999a] represent perturbation expansions in terms of a small parameter $\sigma_Y$, which represents a measure of the standard deviation of $Y'(x) = Y(x) - \langle Y(x) \rangle_c$. This nominally restricts their application to mildly heterogeneous fields, or to strongly heterogeneous but well-conditioned fields, with $\sigma_Y \ll 1$. Guadagnini and Neuman [1999b] demonstrate that the equations may in fact work well for $\sigma_Y$ as large as 4.

Moments such as $\langle h(x) \rangle_c$ are expressed as perturbation expansions

$$\langle h(x) \rangle_c = \langle h^{(0)}(x) \rangle_c + \langle h^{(1)}(x) \rangle_c + \langle h^{(2)}(x) \rangle_c + \cdots \quad (2.8)$$

where the superscript $(i)$ represents order, or proportionality to $\sigma_Y^i$. 
2.3.1 Recursive First Moment Equations

Like Guadagnini and Neuman [1999a, b], I concern myself with expressions up to second order in $\sigma_Y$. These authors have shown that the zero-order solution satisfies

$$-\nabla \cdot \langle q^{(0)}(x) \rangle_c + \langle f(x) \rangle = 0 \tag{2.9}$$

$$\langle q^{(0)}(x) \rangle_c = -K_G(x) \nabla \langle h^{(0)}(x) \rangle_c \tag{2.10}$$

subject to corresponding boundary conditions,

$$\langle h^{(0)}(x) \rangle_c = \langle H(x) \rangle \quad x \in \Gamma_D \tag{2.11}$$

$$-\langle q^{(0)}(x) \rangle_c \cdot n(x) = \langle Q(x) \rangle \quad x \in \Gamma_N \tag{2.12}$$

where $\langle f(x) \rangle$, $\langle H(x) \rangle$, and $\langle Q(x) \rangle$ are prescribed unconditional first moments of the statistically independent random source and boundary functions.

Whereas $\langle h^{(1)}(x) \rangle_c$ and $\langle q^{(1)}(x) \rangle_c$ are identical to zero, the second-order solution obeys [Guadagnini and Neuman, 1999a]

$$\nabla \cdot \langle q^{(2)}(x) \rangle_c = 0 \tag{2.13}$$

$$\langle q^{(2)}(x) \rangle_c = -K_G(x) \left[ \nabla \langle h^{(2)}(x) \rangle_c + \frac{\sigma^2_Y(x)}{2} \nabla \langle h^{(0)}(x) \rangle_c \right] + r^{(2)}_c(x) \tag{2.14}$$

subject to

$$\langle h^{(2)}(x) \rangle_c = 0 \quad x \in \Gamma_D \tag{2.15}$$
\[-\langle q^{(2)}(x) \rangle_c \cdot n(x) = 0 \quad x \in \Gamma_N \tag{2.16}\]

where \(r_c^{(2)}(x)\) is second order approximation of the conditional "residual flux" \(r_c(x)\) and given by

\[r_c^{(2)}(x) = \int_{\Omega} a_c^{(2)}(y, x) \nabla_y \langle h^{(0)}(y) \rangle_c \, dy \tag{2.17}\]

here \(a_c^{(2)}(y, x)\) is second order approximation of the second-rank positive semidefinite symmetric tensor \(a_c(y, x)\) and given by

\[a_c^{(2)}(y, x) = K_G(x)K_G(y) \langle Y''(x)Y''(y) \rangle_c \nabla_x \nabla_y^T \langle G^{(0)}(y, x) \rangle_c \tag{2.18}\]

where \(G^{(0)}(y, x)\) is zero order approximation of the symmetric random Green's function \(G(y, x)\) associated with Equations (2.1)-(2.4) (i.e., the solution of these equations when \(f(x) = \delta(x - y)\) subject to homogeneous boundary conditions \(H(x) = Q(x) = 0\) [Neuman and Orr, 1993a]; and \(\langle Y''(x)Y''(y) \rangle_c\) is conditional co-variance of \(Y\) between locations \(x\) and \(y\).

The recursive Equations (2.9)-(2.16) constitute a system of integro-differential equations. Their parameters and state variables depend both on the statistical properties of \(K(x)\) and information (scale, location, type, quantity, and quality) one has about this field [Neuman and Orr, 1993a]. These equations are recursive in the sense that \(i\)th and lower order head and flux approximations are needed to calculate \((i+1)\)th order approximations of these variables.

From Equation (2.17), note that \(r_c^{(2)}(x)\) is not proportional to the local hydraulic gradient. Instead, it depends on global integrals which render the second-order flux
approximation nonlocal and non-Darcian. This means that an effective hydraulic conductivity cannot be defined for the general case [Neuman and Orr, 1993a]. Also, note that the conditional kernel \( a_c \) is a system parameter independent of source or boundary terms [Neuman and Orr, 1993a]. Yet, \( a_c \) is nonlocal and nonunique due to its conditional nature.

It is worthwhile noting that \( K_C(x) \) and \( \langle Y'(x)Y'(y) \rangle_c \) are quantities to be estimated by my inverse method and ultimately conditioned on hydraulic head and/or flux and (possibly) conductivity measurements.

### 2.3.2 Recursive Second Moment Equations

Guadagnini and Neuman [1999a] also used a perturbation expansion to obtain recursive approximations of their exact conditional second moments \( C_h(x, y) \), \( C_{hK}(x, y) \), and \( C_q(x, y) \). These moments constitute measures of predictive uncertainty of head and flux and of spatial correlation between \( h \) and \( K \), respectively. Only second-order approximations are given below; lower-order approximations are zero. The second-order approximation of the conditional covariance for heads satisfies

\[
\nabla_x \cdot \left[ K_C(x) \nabla_x C_h^{(2)}(x, y) + C_{hK}^{(2)}(x, y) \nabla_x \langle h^{(0)}(x) \rangle_c \right] + \int_{\Omega} \langle f'(x)f'(z) \rangle \langle G^{(0)}(z,y) \rangle_c \, dz = 0 \quad x \in \Omega
\]

(2.19)
subject to boundary conditions

\[
C_h^{(2)}(x, y) = - \int_{\Omega} \langle H'(x)H'(y) \rangle \left( [K_G(z) \nabla_z \langle G^0(z, y) \rangle]_c + 1 \right) \cdot n(z) dz
\]

\[x \in \Gamma_D \quad (2.20)\]

\[
\left[ K_G(x) \nabla_x C_h^{(2)}(x, y) + C_{hk}^{(2)}(x, y) \nabla_x \langle h^0(x) \rangle \right]_c \cdot n(x)
= \int_{\Omega} \langle Q'(x)Q'(z) \rangle \langle G^0(z, y) \rangle \langle h^0(z) \rangle_\xi dz \quad x \in \Gamma_N
gn\]

\[x \in \Gamma_N \quad (2.21)\]

The second-order approximation of the conditional cross-covariance \( C_{hk}(x, y) \) is given by

\[
C_{hk}^{(2)}(x, y) = - K_G(x)
\]

\[
\cdot \int_{\Omega} \nabla^T_z \langle h^0(z) \rangle_c \nabla_z \langle G^0(z, y) \rangle_\xi K_G(z) \langle Y(z)Y'(x) \rangle_\xi dz
\]

\[x \in \Gamma_N \quad (2.22)\]

Thus, by solving Equations (2.19)–(2.22) one can determine up-to second-order approximation \( C_h^{(2)}(x, y) = C_h^{(2)}(x, y) \).

The conditional covariance of flux, \( C_q(x, y) = \langle q'(x)q'^T(y) \rangle_\xi \), is a tensor that is generally nonsymmetric [Guadagnini and Neuman, 1999a]. Its approximation up-to second order is

\[
C_q^{(2)}(x, y) = K_G(x)K_G(y) \left[ \nabla_x \nabla^T_y C_h^{(2)}(x, y) \right.
\]

\[
+ \langle Y'(x)Y''(y) \rangle_c \nabla_x \langle h^0(x) \rangle_c \nabla^T_y \langle h^0(y) \rangle_c \left. \right]
\]

\[+ K_G(y) \nabla_x \langle h^0(x) \rangle_c \nabla^T_y C_{hk}^{(2)}(x, y) \]

\[+ K_G(x) \nabla_x C_{hk}^{(2)}(y, x) \nabla^T_y \langle h^0(y) \rangle_c \]

\[x \in \Gamma_N \quad (2.23)\]
FIGURE 2.1. Hydrogeologic domain with finite element grid.

2.4 Finite Element Moment Equations

In order to solve the recursive nonlocal moment equations, Guadagnini and Neuman [1999a] developed a Galerkin finite-element orthogonalization scheme, which I adopt here. For simplicity, I prescribe deterministic source and boundary terms \( \langle f(x) \rangle = f(x), x \in \Omega; \langle H(x) \rangle = H(x), x \in \Gamma_D; \langle Q(x) \rangle = Q(x), x \in \Gamma_N \). I solve the problem on a two-dimensional domain (identical to that of Guadagnini and Neuman [1999a, b]) of length \( L_1 = 18 \) (in arbitrary consistent units) and width \( L_2 = 8 \). The domain is discretized into \( N_E = 3600 \) (40 rows \( \times \) 90 columns) square elements of uniform size \( \Delta x_1 = \Delta x_2 = 0.2 \) (Figure 2.1). Interpolation across each element is done using
bilinear Lagrange shape functions given (for each element node) by

\[
\psi_1^e(x) \equiv \frac{1}{4} \left( 1 + \frac{x_1}{\Delta L} \right) \left( 1 + \frac{x_2}{\Delta L} \right) \\
\psi_2^e(x) \equiv \frac{1}{4} \left( 1 + \frac{x_1}{\Delta L} \right) \left( 1 - \frac{x_2}{\Delta L} \right) \\
\psi_3^e(x) \equiv \frac{1}{4} \left( 1 - \frac{x_1}{\Delta L} \right) \left( 1 - \frac{x_2}{\Delta L} \right) \\
\psi_4^e(x) \equiv \frac{1}{4} \left( 1 - \frac{x_1}{\Delta L} \right) \left( 1 + \frac{x_2}{\Delta L} \right)
\] (2.24)

where local (element) node “1” is located at the right-uppermost element vertex and following nodes are positioned at subsequent vertices in clockwise order; \( e \) denotes “element”; \( x_1 \) and \( x_2 \) are Cartesian (two-dimensional) components of vector \( x \); and \( \Delta L \) \((= \Delta x_1 = \Delta x_2)\) is square element length.

### 2.4.1 Conditional Mean Head and Flux

The zero order component of mean head is given by Equations (2.9)–(2.12). Galerkin orthogonalization of these equations leads to a finite element system similar in form to that of a traditional deterministic flow model:

\[
\sum_{m=1}^{N} A_{nm} h_m^{(0)} = b_{0n}, \quad n = 1, 2, \ldots, N
\] (2.25)

where \( N \) represents the number of grid nodes which are not on \( \Gamma_D \) \((N = 3649, \text{ in my computational grid})\); coefficient \( A_{nm} \) (which leads to a sparse and symmetric matrix) is given in terms of Lagrange interpolation functions, \( \psi_n \), within a global integral on \( \Omega \):

\[
A_{nm} = \int_{\Omega} K_G(x) \nabla \psi_n(x) \cdot \nabla \psi_m(x) \, dx
\] (2.26)
and $b_{mn}$ is the mean (interior and/or boundary) source term for node $n$. These zero-order finite-element equations carry the mean source and boundary terms ($\langle f(x) \rangle$, $\langle H(x) \rangle$, $\langle Q(x) \rangle$) onto the computation of mean head.

Galerkin orthogonalization of the second-order moment Equations (2.13)–(2.16) yields the algebraic system

$$\sum_{m=1}^{N} [A_{nm}h_m^{(2)} + B_{nm}h_m^{(0)}] = P_n + S_n \quad n = 1, 2, \ldots, N$$

(2.27)

where $h_m \equiv \langle h(x_m) \rangle_c$ is conditional mean head at node $m$; coefficient $B_{nm}$ (leading to a sparse and symmetric matrix) is given in terms of Lagrange interpolation functions, $\psi_n$, within a global integral on $\Omega$:

$$B_{nm} = \int_{\Omega} K_G(x) \frac{\sigma_G(x)}{2} \nabla \psi_n(x) \cdot \nabla \psi_m(x) \, dx$$

(2.28)

$P_n$ (a term depending on Dirichlet-boundary heads) and $S_n$ (a global integral term involving the second-order residual flux approximation) are respectively given by

$$P_n = - \sum_{m=1}^{N_D} [A_{nm}h_m^{(2)} + B_{nm}h_m^{(0)}]$$

(2.29)

$$S_n = \int_{\Omega} r_{c}^{(2)}(x) \cdot \nabla \psi_n(x) \, dx$$

$$= \sum_{e=1}^{M_e} k_G(e) \sum_{i=1}^{N_x} \Theta_{ni} \sum_{e'=1}^{M_q} k_G(e') \left\langle Y''(e)Y'(e') \right\rangle_c$$

$$\cdot \sum_{j=1}^{N_y} \sum_{k=1}^{N_y} h_{k}^{(0)} \Delta_{jk}$$

(2.30)

where $N_D$ is number of nodes on $\Gamma_D$; $r_c^{(2)}(x)$ is second-order residual flux approxima-
tion given at element midpoint $x^e$ by

$$r^{(2)}_c(x^e) = K_G(e) \sum_{i=1}^{N_x} \nabla_x \psi_i(x^e) \sum_{e'=1}^{M_y} K_G(e') \langle Y'(e)Y'(e') \rangle_c$$

$$\cdot \sum_{j=1}^{N_x} G_{ij}^{ee'} \sum_{k=1}^{N_y} h^{(0)e'}_k \Delta^{e'e'}_{jk}$$

(2.31)

$K_G(e)$ is regarded as uniform within each element $e$ and estimated via kriging (see Section 3.3); conditional $Y$ covariance $\langle Y'(e)Y'(e') \rangle_c$ between element midpoints is calculated by means of kriging as well; $N_x$ and $N_y$ indicate the number of nodes in elements $e$ and $e'$ within $x$ and $y$ planes, respectively; $M_x$ and $M_y$ are number of elements in the $x$ and $y$ planes; $G_{ij}^{ee'}$ is zero order Green’s function at node $i$ of element $e$ in the $x$ plane due to a unit source at node $j$ of element $e'$ in the $y$ plane; $h^{(0)e'}_k$ is the zero order head, $h^{(0)}(y^e')$, at node $k$ of element $e'$ in the $y$ plane; and $\Theta^{ee}_{ni}$ and $\Delta^{e'e'}_{jk}$ are local integral terms given, respectively by

$$\Theta^{ee}_{ni} = \int_{\Omega(e)} \nabla_x \psi_n(x^e) \cdot \nabla_x \psi_i(x^e) \, dx$$

(2.32)

and

$$\Delta^{e'e'}_{jk} = \int_{\Omega(e')} \nabla_y \psi_j(y^{e'}) \cdot \nabla_y \psi_k(y^{e'}) \, dy$$

(2.33)

where the integrals are computed over elements $e$ and $e'$ in the $x$ and $y$ planes, respectively. The computations of $\Theta^{ee}_{ni}$ and $\Delta^{e'e'}_{jk}$ are identical; thus, calculating either one determines the other. On the other hand, the computation of $S_n$ and $r^{(2)}_c$ require the numerical evaluation of the zero order approximation of the mean Green’s function, $G_{ij}^{ee'}$, at grid nodes. I describe this evaluation below.
Thus, the second order approximations for conditional mean head and flux are given by

\[
\langle h^{[2]}(x) \rangle_c = \langle h^{(0)}(x) \rangle_c + \langle h^{(2)}(x) \rangle_c \tag{2.34}
\]

\[
\langle q^{[2]}(x) \rangle_c = \langle q^{(0)}(x) \rangle_c + \langle q^{(2)}(x) \rangle_c \tag{2.35}
\]

respectively, where the superscript \([i]\) represents terms up to \(i\)th order.

**Matrix Form.** Equations (2.25) and (2.27) can be rewritten in matrix notation as

\[
Ah^{(0)} = b_0 \tag{2.36}
\]

\[
Ah^{(2)} = b_2 \equiv P + S - Bh^{(0)} \tag{2.37}
\]

respectively, where bold quantities are vectors or matrices formed by their corresponding non-bold entries. Thus, the second-order approximation of mean head in matrix form is given by

\[
h \equiv \langle h(x)^{[2]} \rangle \equiv h^{(0)}(x) + h^{(2)}(x) \tag{2.38}
\]

By adding Equations (2.36) and (2.37) and using Equation (2.38), I obtain the matrix equation of second-order mean heads,

\[
Ah = b \equiv b_0 + b_2 \tag{2.39}
\]

Even though this equation system looks similar to a traditional deterministic groundwater-flow equation, it can only be understood in light of a stochastic interpretation. Yet, Equation (2.39) is computationally much more elaborate than a traditional flow model due to the calculation of the nonlocal term \(b_2\).
Computation of Mean Green’s Function. $\langle G^{(0)}(y, x) \rangle_c$ is the zero-order approximation of the mean Green’s function and satisfies the equation

$$\nabla_y \cdot [K(y) \nabla_y \langle G^{(0)}(y, x) \rangle_c] + \delta(y - x) = 0 \quad \text{on } \Omega \tag{2.40}$$

subject to

$$\langle G^{(0)}(y, x) \rangle_c = 0 \quad y \in \Gamma_D \tag{2.41}$$

$$\nabla_y \langle G^{(0)}(y, x) \rangle_c \cdot n(y) = 0 \quad y \in \Gamma_N \tag{2.42}$$

The numerical evaluation of $\langle G^{(0)}(y, x) \rangle_c$ is performed as follows: a unit-strength injection well is placed at node $x$ and the response $\langle G^{(0)}(y, x) \rangle_c$ is computed at every node $y$ on the grid by evaluating Equations (2.40)–(2.42); the well is then located at another node and the process repeated; the computation of the mean Green’s function is complete when all nodes have served as source locations. Note that the procedure for determining the Green’s function implies solving $N$ different systems similar to Equations (2.40)–(2.42).

$G^{ee'}_{ij}$ in Equations (2.30) and (2.31) is equal to the value of $\langle G^{(0)} \rangle_c$ at node $i$ of element $e$ due to a unit source at node $j$ of element $e'$. Additionally, $\langle G^{(0)}(y, x) \rangle_c$ is symmetric (i.e., $G^{ee'}_{ij} = G^{e'e}_{ji}$) and vanishes at Dirichlet nodes.

2.4.2 Conditional Covariance of Hydraulic Head

The recursive second-order approximation of conditional head covariance given by Equations (2.19)–(2.21) reduces to the following system due to the deterministic forc-
ing terms:
\[
\nabla \cdot \left[ K_G(x) \nabla_x C_h^{(2)}(x, y) + C_{hK}^{(2)}(x, y) \nabla_x \langle h^{(0)}(x) \rangle_e \right] = 0 \quad (2.43)
\]
subject to
\[
C_h^{(2)}(x, y) = 0 \quad x \in \Gamma_D \quad (2.44)
\]
\[
\left[ K_G(x) \nabla_x C_h^{(2)}(x, y) + C_{hK}^{(2)}(x, y) \nabla_x \langle h^{(0)}(x) \rangle_e \right] \cdot \mathbf{n}(x) = 0 \quad x \in \Gamma_N \quad (2.45)
\]

Further Galerkin orthogonalization of Equations (2.43)–(2.45) yields the finite-element equations of head covariance,
\[
\sum_{m=1}^{N} \left[ E_{nm} C_{m,j}^{(2)} + F_{nm,j} h_m^{(0)} \right] = R_n \quad n = 1, 2, \cdots, N \quad (2.46)
\]
where \( C_{m,j}^{(2)} \) is head covariance between node \( j \) in the \( y \) plane and node \( m \) in the \( x \) plane; \( E_{nm} \) and \( F_{nm,j} \) are global integral coefficients in terms of functions \( \psi_n \) and given by
\[
E_{nm} = \int_{\Omega} K_G(x) \nabla \psi_n(x) \cdot \nabla \psi_m(x) \, dx \quad (2.47)
\]
\[
F_{nm,j} = \int_{\Omega} C_{hK}^{(2)}(x) \nabla \psi_n(x) \cdot \nabla \psi_m(x) \, dx
= \sum_{e=1}^{M_e} C_{hK}^{(2)}(x^e) \int_{\Omega(e)} \nabla_x \psi_n(x^e) \cdot \nabla_x \psi_m(x^e) \, dx \quad (2.48)
\]
respectively (note that \( E_{nm} \) is identical to \( A_{nm} \); thus, only one of them needs to be computed); \( R_n \) is a term involving prescribed heads at nodes on \( \Gamma_D \),
\[
R_n = -\sum_{m=1}^{N_D} \left[ E_{nm} h_m^{(2)} + F_{nm,j} h_m^{(0)} \right] \quad (2.49)
\]
From Equation (2.22), the finite-element second-order approximation of cross-covariance, \( C_{hJKy}^{(2)}(x) \), between node \( j_y \) in the \( y \) plane and midpoint \( x^e \) in the \( x \) plane is given by

\[
C_{hJKy}^{(2)}(x^e) = -K_G(e) \sum_{e''=1}^{M_x} K_G(e'') \left( Y''(e) Y'(e') \right)_c c \sum_{i_x=1}^{N_x} G_{isjy}^{(0)e''} \sum_{i_y=1}^{N_y} h_{jy}^{(0)e''} \Xi_{isjy}^{e''e''} \tag{2.50}
\]

where \( G_{isjy}^{(0)e''} \) is the zero-order Green's function at node \( i_z \) in the \( z \) plane due to a unit source at node \( j_y \) in the \( y \) plane; and \( \Xi_{isjy}^{e''e''} \) is a term computationally identical to both \( \Theta_{ni}^{e''} \) and \( \Delta_{jk}^{e''e''} \) (Equations (2.32) and (2.33)) and given by

\[
\Xi_{isjy}^{e''e''} = \int_{\Omega(e'')} \nabla_z \psi_{isy}^z(x^e) \cdot \nabla_z \psi_{jy}^z(x^e) dx \tag{2.51}
\]

### 2.4.3 Conditional Covariance Tensor of Flux

By means of the bilinear Lagrange interpolation, the components of the second order tensor of conditional flux covariance between element centers (Equation 2.23) become

\[
\frac{C_{qij}^{(2)}(x^e, y^{e'})}{K_G(x^e)K_G(y^{e'})} = \sum_{n=1}^{N_x} \sum_{m=1}^{N_y} \Psi_{nm}^{ij} \left[ C_{nm}^{(2)} + \left( Y'(x^e) Y'(y^{e'}) \right)_c h_n^{(0)}h_m^{(0)} \right. \\
+ \frac{h_n^{(0)}C_{hKm}^{(2)}(x^e)}{K_G(x^e)} + \frac{h_m^{(0)}C_{hKl}^{(2)}(y^{e'})}{K_G(y^{e'})} \right] \quad i, j = 1, 2 \tag{2.52}
\]

where \( n \) is a node in element \( e \); \( m \) is a node in element \( e' \); indices \( i \) and \( j \) represent vectorial components of flux \( q \); and

\[
\Psi_{nm}^{ij} = \frac{\partial \psi_n(x^e)}{\partial x_i} \frac{\partial \psi_m(y^{e'})}{\partial y_j} \quad i, j = 1, 2 \tag{2.53}
\]
2.5 Computational Approach

In this section, I describe the sequential computational steps needed to solve the nonlocal moment equations and provide some hints for expediting their numerical solution.

2.5.1 Mean Head and Flux

In Section 2.4.1, I introduced the numerically workable finite-element expressions of conditional mean head and flux that will compose the forward nonlocal flow model of my inverse method (see Chapter 4). This forward model will be numerically solved many times (in my examples, up to several thousands) during the course of an inversion; with each forward model run implying a different flow or \( K_G \) field scenario. Therefore, it is highly desirable to optimize the numerical forward-model algorithm.

The following procedure for solving the nonlocal forward model is a slight modification of the original methodology of Guadagnini and Neuman [1999a]. Quantities should be computed in the given order.

1. \( K_G(e) \) and \( \langle Y'(e)Y'(e') \rangle_c \), from posterior (or prior) universal kriging (Chapter 4);

2. \( h_{m}^{(0)} \), by solving Equations (2.25) and (2.26);

3. \( \langle q^{(0)}(x) \rangle_c \), by means of Equation (2.10);
4. $\langle G^{(0)}(y, x) \rangle_c$, as described in Section 2.4.1;

5. $r^{(2)}_c(x^e)$, using Equations (2.31) and (2.33);

6. $h^{(2)}_m$, by solving Equations (2.27)–(2.33);

7. $\langle q^{(2)}(x) \rangle_c$, by means of Equation (2.14); and

8. $\langle h(x)^{(2)} \rangle_c$ and $\langle q(x)^{(2)} \rangle_c$, using Equations (2.34) and (2.35), respectively.

It takes more numerical operations to compute $r^{(2)}_c(x)$ in step 5 than it takes to compute any other quantity in the other steps because the numerical evaluation of $r^{(2)}_c(x)$ requires computing a global integral for each grid element (see Equation (2.31)). One should be able to reduce the computational burden by recognizing that the integrand dies out with distance.

### 2.5.2 Head and Flux Covariances

Computation of second moment quantities should proceed in the following order:

1. $C^{(2)}_{hK_{xy}}(x^e)$, using Equations (2.50) and (2.51);

2. $C^{(2)}_h(x, y) = C^{(2)}_h(x, y)$, by solving Equations (2.46)–(2.49); and

3. $C^{(2)}_{qij}(x^e, y^e)$, by means of (2.52) and (2.53).

Since there are on the order of $N^2$ different terms $C^{(2)}_{hK_{xy}}(x^e)$ to compute, Step 1 is the most costly calculation among all first and second moment quantities.
2.5.3 Efficiency of the Numerical Algorithm

Below, I give a few hints for improving the performance of the mean and second moment algorithms:

1. Since a considerable part of a forward model run is spent on solving the linear systems of the Green’s function, the choice of an adequate linear equation solver can save significant computing effort. Whereas Guadagnini and Neuman [1999a, b] had originally employed an iterative method solver based on an incomplete Choleski conjugate gradient algorithm, I have found that a direct linear solver using symbolic multifrontal Choleski factorization for sparse and symmetric matrices [Liu, 1987] was much (at least one order of magnitude) faster in my numerical examples.

2. Determining the Green’s function (Section 2.4.1) entails solving \( N \) different equation systems under the same \( K_{G} \) field. This allows saving considerable computer time since only one coefficient matrix needs to be calculated for these \( N \) systems. Therefore, only one (common) factorization has to be computed for these equation systems when using a direct solution method.

3. The previous note also applies to the equation systems of zero- and second-order heads and of head covariance as all of them possess coefficient matrices identical to that of the Green’s function system.
4. Compiler optimization, if available, may significantly enhance numerical algorithm efficiency, especially if the original code was programmed without optimization techniques. Otherwise, it might be worthwhile to modify the most busy areas of the code using optimization techniques such as efficient loop nesting and array handling.

5. BLAS (Basic Linear Algebra Subprograms) subroutines (at levels 1, 2, or 3) are designed to perform operations among scalars, vectors, and matrices [Lawson et al., 1979; Dongarra et al., 1988, 1990]. Programming with such subroutines may improve the forward-model algorithm efficiency and allow for parallel computing if needed.

6. Storing the results of the vector-vector or vector-matrix operations that will be reused later (as in the computation of $r_c^{(2)}(x^c)$, $C_{h}^{(2)}(x^c)$, and $C_{qij}^{(2)}(x^c, y^c)$) may save computer time if enough memory is available.

I found that Steps 1, 2, and 4 had the greatest impact on expediting computations.

### 2.5.4 Predictive Head and Flux Errors upon Inversion

$\langle h(x) \rangle_c$ and $\langle q(x) \rangle_c$ represent conditional predictors of head and flux, respectively. I have earlier presented conditional expressions of $C_h(x, y)$ and $C_q(x, y)$, which represent measures of predictive error or uncertainty. These predictors and their associated
errors are all conditioned on the available hydraulic conductivity information via estimates of \( Y(x) \) and \( Y(x)Y'(y) \). In my work, conditioning on this information is done via inversion, using head and/or flux data and possibly conductivity measurements too. Thus, the resulting predictive errors of head and flux contain parameter estimation errors in addition to stochastic averaging errors.

By inspecting values of \( C_h \) and \( C_q \), one can compare the effect that conditioning on different types of data has on predictive uncertainty. For example, if both hydraulic head and conductivity data are available, one can condition the nonlocal moment equations on 1) \( K \) data alone, via forward modeling; 2) head data alone, by means of inversion; and 3) both head and \( K \) information, geostatistically and/or through inverse modeling using prior information. In general, one expects joint conditioning on reliable parameter and head measurements to yield smaller prediction errors than conditioning on only one such set of data.

Neuman and Guadagnini [2000] presented an analysis of errors entering into groundwater model calibration that the reader might find relevant.

### 2.6 Localization of Conditional Mean Flow

Guadagnini and Neuman [1999a, b] developed and illustrated an approximation based on localizing the nonlocal moment flow equations so as to obtain an approximated Darcian flow predictor. An important aspect of this localization is that it allows one to clearly see the conceptual connections of the nonlocal moment flow equations with
traditional fully-deterministic flow equations. A localized solution is simpler to compute than a second-order finite-element solution and might be used as a less expensive, faster (although less accurate) forward model with which to approach the (optimum) inverse solution whenever this localization is reasonably justified. However, localization does not allow computing second moments of head or flux. Following is a brief account of the findings of Guadagnini and Neuman [1999a] on localization of the nonlocal mean moment solution.

The localization approximation is theoretically justified whenever the residual flux can be written as

\[ r_c(x) \approx \int_{\Omega} a_c(y, x) dy \nabla \langle h(x) \rangle_c \quad x \in \Omega \quad (2.54) \]

A case in which this approximation is valid occurs when the conditional conductivity field is nearly homogeneous in the neighborhood of \( x \) and \( \nabla \langle h(x) \rangle_c \) is approximately uniform in this neighborhood.

If the residual flux localization is justified, then the conditional mean flux can be written in Darcian form (see Guadagnini and Neuman [1999a] for details):

\[ \langle q(x) \rangle_c \approx -K_c(x) \nabla \langle h(x) \rangle_c \quad (2.55) \]

where

\[ K_c(x) = \langle K(x) \rangle_c I - \bar{K}_c(x) \quad \bar{K}_c(x) = \int_{\Omega} a_c(y, x) dy \quad (2.56) \]

Here, \( I \) is the identity tensor and \( K_c(x) \) is a conditional (data- and space-dependent) apparent hydraulic conductivity tensor. That is, localization of the nonlocal flow
equations entails a directional, conditional apparent hydraulic conductivity, $K_c(x)$, which is generally anisotropic and local (defined on $\omega$). Nevertheless, $K_c(x)$ is a biased estimate of the actual $\omega$-scale hydraulic conductivity $K(x)$.

From Equation (2.55), the localized moment flow equations are identical in form to the traditional deterministic flow equations which are commonly used for flow prediction under heterogeneous and uncertain field conditions. Therefore, the application of traditional models is valid only as an approximation and their results must be interpreted as a mean solution. In addition, the traditional hydraulic conductivity of deterministic flow models must be regarded as a nonunique (conditional) effective parameter that depends on the quantity, quality, scale, location, and type of available data. While traditional deterministic head and flux should be viewed as conditional predictors of these states, traditional model calibration on available data can be seen as a conditioning process on such information. The continuous variation of the calibrated parameters as more data are added into this process is a consequence of increasing the database and thus changing the data-dependent (conditional) parameters, which the model calibration is to estimate.
To numerically solve the nonlocal mean head equations in the preceding chapter, one needs to specify values of the conditional geometric mean conductivity $K_G(x)$ (through back transformation of $\langle Y(x) \rangle_c$) and $\langle Y'(x)Y''(y) \rangle_c$ throughout $\Omega$. My inverse methodology allows conditioning these values on hydraulic head or flux as well as conductivity. The conditioning is accomplished in part geostatistically and in part through numerical inversion. This chapter describes the geostatistical component of my conditioning approach.

### 3.1 Statistical Structure of Log Hydraulic Conductivity

In aquifer parameter estimation or groundwater model calibration, inverse problems are usually posed in terms of $Y = \ln K$. One reason for using $Y$ instead of $K$ is that its univariate distribution tends to be close to normal (Gaussian).

Though the multivariate distribution of $Y$ may not be Gaussian, it is often assumed to be. The $Y$ field is then fully characterized by its first two moments (mean, variance and covariance) and can be delineated optimally by means of a linear estimator. Working with $Y$ further guarantees that $K$ remains non-negative. I take log hydraulic conductivity to have the following unconditional statistical properties.
1. Y is random.

2. Y is decomposed as $Y(x) \equiv m(x) + Y'(x)$ where $m(x) = \langle Y(x) \rangle$ is a deterministic drift and $Y'(x)$ is a zero-mean random fluctuation about $m(x)$ (i.e., $\langle Y'(x) \rangle \equiv 0$).

3. (a) $\langle Y(x) \rangle = m(x)$ may be unknown.

   (b) The drift is represented as

   $$m(x) = \sum_{r=1}^{R} a_r f_r(x) \quad (3.1)$$

   where $f_r$ are known drift functions, $a_r$ are possibly unknown drift coefficients, and $R$ is a given number of terms in the approximation of $m(x)$.

4. The variance $\sigma^2_{Y_u} \equiv \langle [Y(x) - m(x)]^2 \rangle = \langle [Y'(x)]^2 \rangle = \sigma^2_{Y_u}$ of $Y$ may be unknown, in which case it is estimated by inversion.

5. The covariance $C_Y(x - y) \equiv \langle [Y(x) - m(x)][Y(y) - m(y)] \rangle = C_Y(y - x) = C_Y(y - x)$ of $Y$ may include unknown statistical parameters to be estimated by inversion.

6. The variogram $\gamma_Y(x - y) \equiv \langle [Y(x) - Y(y)]^2 \rangle = \gamma_Y(y - x) = C_Y(0) - C_Y(x - y) = \gamma_Y(x - y)$ of $Y$ may include unknown statistical parameters to be estimated by inversion.

Here $\sigma^2_{Y_u}$ is unconditional and $\sigma^2_Y$ is the conditional variance of $Y$. 
3.1.1 Variogram and Statistical Parameters

Consider a variogram $\gamma_Y$ that may have a geometric anisotropy characterized by two principal integral scales $\rho_1$ and $\rho_2$ parallel to two principal Cartesian axes $x_1$ and $x_2$, respectively (this formulation can be readily extended to three dimensions). The variogram can be expressed as

$$\gamma_Y(s) = n_0 + S_0 \gamma_{Y_R}(s)$$

where $n_0$ and $S_0 = \sigma_Y^2$ are nugget and sill, respectively, and

$$s_R = \sqrt{\left(\frac{s_1}{\rho_1}\right)^2 + \left(\frac{s_2}{\rho_2}\right)^2} = \text{reduced separation magnitude}$$

where $s_1$ and $s_2$ are components of the separation vector $s$ along $x_1$ and $x_2$, respectively. $\gamma_{Y_R}$ has unit sill and may be exponential or spherical, among others [Deutsch and Journel, 1997; Chiles and Delfiner, 1999]. $\gamma_Y$ may have more summands (or structures) such as $S_0 \gamma_{Y_R}$ in Equation (3.2) but I limit my development to $\gamma_Y$ above for simplicity.

Thus, the variogram is defined by statistical parameters $n_0, S_0, \rho_1, \rho_2$, and possibly other, depending on the type and number of variogram models used. Then, I define $\theta$ as the vector of statistical parameters of $Y$

$$\theta = (n_0, S_0, \rho_1, \rho_2, \ldots)^T$$

It has been shown [Zimmerman et al., 1998] that selecting a suitable variogram is crucial for success in geostatistical inverse modeling. If reliable hydraulic-conductivity
data are unavailable or not enough to characterize $\theta$, head, flux, or other types of data may be used to obtain or improve the statistical parameters as described below. The nonlocal flow model requires not only $Y$ values but also estimation error variances (from kriging) throughout $\Omega$ that depend on $\theta$. Hence, good estimation of statistical parameters is important for my inverse method.

I assume that $\theta$ can be: inferred geostatistically [Deutsch and Journel, 1997] from measured values of $Y$ if enough reliable data are available; inferred indirectly by regression or cokriging from other variables (porosity, soil/rock type, petrophysical variables, geophysical signals) if $Y$ data are limited in number or quality; and/or estimated by inversion on the basis of head and/or flux measurements (Section 4.3.5). The latter is possible regardless of whether or not direct or indirect measurements of $Y$ are available.

3.1.2 Prior Statistical Parameters

The statistical parameters ($n_0, S_0, \rho_1, \rho_2$, and possibly others) may be unknown, but estimates of them ($n_0^*, S_0^*, \rho_1^*, \rho_2^*$, and possibly other priors) must be obtained, in the manner described in Section 3.1.1, or guessed a priori in order to perform kriging in the inversion process. (The superscript (*) refers to priors or measurements throughout this dissertation.) Thus, I can define a prior vector $\theta^*$ as

$$\theta^* = (n_0^*, S_0^*, \rho_1^*, \rho_2^*, \ldots)^T$$

(3.5)
If unknown, $\theta^*$ will be estimated by inversion (see Section 4.4). On the other hand, if enough reliable $Y$ measurements are available the statistical parameter priors can be calculated by modeling the sample variogram and no further improvement by inverse estimates will be needed.

### 3.1.3 Log $K$ Measurements

If available, $Y$ measurements at points $x_i$ are allowed to contain errors $\varepsilon^*_Y = \varepsilon^*_Y(x_i)$. Thus, $Y$ measurements can be expressed as

$$Y^*(x_i) = Y(x_i) + \varepsilon^*_Y(x_i)$$

where $Y(x_i)$ is the (unknown) random $Y$ value at measurement point $x_i$. Prior errors $\varepsilon^*_Y$ are assumed to have a zero mean, a known or approximate variance, and to be uncorrelated with both themselves and $Y$; namely,

$$\langle \varepsilon^*_Y(x) \rangle \equiv 0 \quad x \in \Omega$$

$$\langle \varepsilon^*_Y(x)^2 \rangle \equiv \text{Var}[\varepsilon^*_Y(x)] \quad x \in \Omega$$

$$\langle \varepsilon^*_Y(x)\varepsilon^*_Y(y) \rangle \equiv 0 \quad \forall x \neq y \quad x, y \in \Omega$$

$$\langle \varepsilon^*_Y(x)\, Y(y) \rangle \equiv 0 \quad \forall x, y \in \Omega$$
3.2 Prior Kriging

In this section, I present the equations for estimating parameter priors at pilot points (see Section 4.1) and their associated covariances via universal kriging. I term this type of kriging as “prior kriging”, which is performed only once (as compared to “posterior kriging”, presented in Section 3.3, which is done many times in the inverse procedure), before starting the parameter estimation process. However, if there is no log hydraulic conductivity information to perform such kriging, then one needs to determine these priors in a different way. As described above, one may alternatively estimate them via regression or cokriging; otherwise, guessed $Y$ estimates may serve as initial parameter values with which to begin the inverse process. On the other hand, if sufficient $Y$ data are available then parameter priors can be obtained by means of universal kriging using these data and associated errors ($\varepsilon_{Y_i}$).

3.2.1 Prior Kriging Algorithm

Consider $I$ measurement and $P$ pilot points located at points $x_i$ and $x_p$, respectively. (The number and spatial distribution of pilot points are issues that deserve investigation and will be discussed further in Chapter 4.) Fictitious priors $Y_p^*(x_p)$ at pilot points are calculated by kriging the available measurements $Y_M^*(x_i)$:

$$Y_p^*(x_p) = \sum_{i=1}^{I} \lambda_i(x_p) Y_M^*(x_i) \quad (p = 1, 2, \ldots, P)$$

(3.11)
where $\lambda_i(x_p)$ is the $i$th kriging weight at point $x_p$ obtained upon solving the following $(I+R)$ kriging equations either in terms of covariances $C_Y$ or variograms $\gamma_Y$ \[Chilès and Delfiner, 1999,\]

\[
\sum_{j=1}^{I} \lambda_j(x_p)C_Y(x_i - x_j) + \lambda_i(x_p) \cdot \text{Var} [\varepsilon_Y(x_i)] - \sum_{r=1}^{R} \mu_r(x_p)f_r(x_i) = C_Y(x_i - x_p) \quad (i = 1, 2, \ldots, I) \quad (3.12)
\]

or,

\[
\sum_{j=1}^{I} \lambda_j(x_p)\gamma_Y(x_i - x_j) - \lambda_i(x_p) \cdot \text{Var} [\varepsilon_Y(x_i)] + \sum_{r=1}^{R} \mu_r(x_p)f_r(x_i) = \gamma_Y(x_i - x_p) \quad (i = 1, 2, \ldots, I) \quad (3.13)
\]

and

\[
\sum_{j=1}^{I} \lambda_j(x_p)f_r(x_j) = f_r(x_p) \quad (r = 1, 2, \ldots, R) \quad (3.14)
\]

where $\mu_r(x_p)$ is the $r$th Lagrange multiplier at point $x_p$ that is obtained jointly with $\lambda_i(x_p)$; $R$ is a given number of terms in the approximation of $m(x)$ (Section 3.1); and $f_r(x_p)$ is the $r$th drift function at point $x_p$ (Section 3.1). Kriging variances at pilot points can be calculated by means of

\[
\langle Y_p^*(x_p)^2 \rangle = \sigma^2_{Y_u} - \sum_{j=1}^{I} \lambda_j(x_p)C_Y(x_j - x_p) + \sum_{r=1}^{R} \mu_r(x_p)f_r(x_p) \quad (p = 1, 2, \ldots, P) \quad (3.15)
\]

or

\[
\langle Y_p^*(x_p)^2 \rangle = \sum_{j=1}^{I} \lambda_j(x_p)\gamma_Y(x_j - x_p) + \sum_{r=1}^{R} \mu_r(x_p)f_r(x_p) \quad (p = 1, 2, \ldots, P) \quad (3.16)
\]
where, $Y_p^*(x_p) = Y(x_p) - Y_p^*(x_p)$. Kriging covariances are rarely given in geostatistical texts. The prior kriging covariance expressions below are a special case of the extended kriging covariances derived in Appendix A (assuming the same data are used for both sets of expressions):

$$\langle Y_p^*(x_p)Y_p^*(x_q) \rangle = C_Y(x_p - x_q) - \sum_{j=1}^{I} \lambda_j(x_q)C_Y(x_j - x_p)$$

$$+ \sum_{r=1}^{R} \mu_r(x_q) f_r(x_p) \quad (p, q = 1, 2, \ldots, P) \quad (3.17)$$

or

$$\langle Y_p^*(x_p)Y_p^*(x_q) \rangle = -\gamma_Y(x_p - x_q) + \sum_{j=1}^{I} \lambda_j(x_q) \gamma_Y(x_j - x_p)$$

$$+ \sum_{r=1}^{R} \mu_r(x_q) f_r(x_p) \quad (p, q = 1, 2, \ldots, P) \quad (3.18)$$

Thus, the prior kriging algorithm yields (conditional) $Y$ estimates and variances at pilot and measurement points; covariances between estimates at either pilot or measurement points; and cross-covariances between estimates at measurement and pilot points. The resulting estimates $Y_p^*(x)$ are unbiased ($\langle Y_p^*(x) \rangle = \langle Y(x) \rangle$) and optimum (minimum estimation variance, $\langle [Y(x) - Y_p^*(x)]^2 \rangle$). An important difference between priors $Y^*_M(x_i)$ and $Y_p^*(x_p)$ is that whereas the former contain mutually uncorrelated measurement errors, the latter contain estimation errors that are correlated according to Equation (3.17) or (3.18).

The solution and uniqueness of the kriging Equations (3.12)–(3.14) are guaranteed if the functions $f_r$ are linearly independent on the set of measurement points $x_i$. That
This condition is generally satisfied if points $\mathbf{x}_i$ are located irregularly; difficulties in the solution may arise if every point $\mathbf{x}_i$ lies on a regular mesh or along a (geometrically) quadratic curve [Chilès and Delfiner, 1999].

In this chapter, kriging expressions are written in terms of covariances or variograms. Whereas my assumed statistical structure (homogeneous increments) of $Y$ allows us to use the less restricted variogram expressions (where $\sigma_{Y_u}^2$ is not required), the covariance equations are better suited for computation. This is so because in the system of covariance equation the larger matrix entries are on the main diagonal, which allows us using efficient equation solvers.

### 3.3 Posterior Kriging

In this section, I introduce the posterior kriging algorithm and its statistical premises.

#### 3.3.1 Extended Measurement Vector $Y_E$

Once priors, $Y_p^*(\mathbf{x}_p)$, at pilot points have been obtained via Equation (3.11), one can append these fictitious “measurements” to the real measurements $Y_M^*(\mathbf{x}_i)$, if any, so
as to compose an extended vector \((Y^*_E)\) of \(Y\) priors,

\[
Y^*_E \equiv \begin{bmatrix} Y^*_M(x_1), Y^*_M(x_2), \ldots, Y^*_M(x_i), Y^*_P(x_1), Y^*_P(x_2), \ldots, Y^*_P(x_P) \end{bmatrix}^T
\]

or

\[
Y^*_E = \begin{bmatrix} Y^*_E(x_1), Y^*_E(x_2), \ldots, Y^*_E(x_{i+P}) \end{bmatrix}^T
\]

(3.20)

\(Y^*_E\) constitutes an augmented “dataset” with which to estimate \(Y\) values at any location in \(\Omega\) during the first inverse iteration (note that \(\theta^*\) is needed to obtain \(Y^*_E\)). At the end of this iteration, parameter values at pilot and (optionally) measurement points will be updated into \(Y_E^{(1)}\) parameter estimates. In general,

\[
Y_E^{(i)} = \begin{bmatrix} Y_E^{(i)}(x_1), Y_E^{(i)}(x_2), \ldots, Y_E^{(i)}(x_{i+P}) \end{bmatrix}^T
\]

(3.21)

is the updated extended vector at the end of the \(i\)th inverse iteration. The iterative updating process by parameter estimates continues until convergence is attained; only then the final parameter estimates in \(\hat{Y}_E\) will be determined. At the beginning of the inversion process, one sets \(Y_E^{(0)} \equiv Y^*_E\).

Moreover, the choice of hydraulic parameter unknowns can be made optional according to the problem at hand. If \(Y\)-measurement errors are very small, \(Y^*_M(x_i)\) could be (optionally) treated as known quantities and \(Y_P(x_p)\) would be the only unknown hydraulic parameters. If enough \(Y\) measurements with large errors are available, pilot-point parameters may not be needed (i.e., \(P = 0\)), thus avoiding over-parameterization. Furthermore, in the case of \(Y\) measurements having significantly different errors \((\sigma_{y_j})\), it might be suitable to set a combination of (a) known parameters \(Y^*_M(x_i)\) with (relatively) small errors; (b) unknown parameters \(Y^*_M(x_i)\) with
(relatively) large errors; and (c) (optional) unknown pilot-point parameters \( Y_p(x_p) \).

### 3.3.2 Correlation among \( Y_E \) Errors

Prior \( Y_E \) errors are defined as

\[
\varepsilon_{Y_E}^* \equiv Y_E^*(x_k) - Y(x_k) \quad (k = 1, 2, \ldots, I + P)
\]

forming the vector

\[
\varepsilon_{YPE}^* \equiv \left[ \varepsilon_{Y_E1}^*, \varepsilon_{Y_E2}^*, \ldots, \varepsilon_{Y_E(I+P)}^* \right]
\]

They are assumed to have zero mean and be uncorrelated with \( Y \) but mutually correlated,

\[
\langle \varepsilon_{Y_Ek}^* \varepsilon_{Y_El}^* \rangle = 0 \quad \forall k, l = 1, 2, \ldots, I + P
\]

Likewise, updated \( Y_E \) errors are given by

\[
\varepsilon_{Y_Ek} \equiv Y_E(x_k) - Y(x_k) \quad (k = 1, 2, \ldots, I + P)
\]

which compose the vector

\[
\varepsilon_{YPE} \equiv \left[ \varepsilon_{Y_E1}, \varepsilon_{Y_E2}, \ldots, \varepsilon_{Y_E(I+P)} \right]
\]
Errors $\varepsilon_{Yk}$ have zero mean and are uncorrelated with $Y$ but correlated with themselves through inverse parameter estimation,

$$\langle \varepsilon_{Yk} \rangle \equiv 0 \quad \forall k = 1, 2, \ldots, I + P \tag{3.29}$$

$$\langle \varepsilon_{Yk} Y \rangle \equiv 0 \quad \forall k = 1, 2, \ldots, I + P \tag{3.30}$$

$$\langle \varepsilon_{Yk} \varepsilon_{Yl} \rangle \equiv E_{kl} \tag{3.31}$$

$$\equiv \begin{cases} 
\text{Var}[\hat{\varepsilon}_Y(x_k) \cdot \delta_{kl}] ; & \forall k, l = 1, 2, \ldots, I_f \\
0 ; & \forall k = 1, 2, \ldots, I_f ; \ l = I_f + 1, I_f + 2, \ldots, I + P \\
Q_{k-I_f, l-I_f} ; & \forall k, l = I_f + 1, I_f + 2, \ldots, I + P 
\end{cases}$$

where $I_f = I + P - N_Y$ is number of $Y$ measurements treated as fixed; $N_Y$ is number of unknown parameters of $Y$; and $Q_{kl}$ represents error correlation between $Y$ estimates at points $x_k$ and $x_l$, and is an entry of the (symmetric) parameter-estimation covariance matrix $Q$ given by the inverse of the Fisher information matrix (Section 4.3.6). Note that $\langle \varepsilon_{Yk} \varepsilon_{Yl} \rangle$, $E_{kl}$, and $E_{kl}$ are all symmetric.

The premise in Equation (3.30) is only correct up to a first-order (linear) approximation of the nonlinear estimator I employ. If parameter estimates near the optimum show a considerable deviation from first-order behavior (as discussed in Section 4.3.6), then use of the posterior kriging algorithm is not justified. Nevertheless, a change in the parameter model (a reparameterization; see Section 4.4) might yield a more linear behavior which would justify the use of the algorithm.
3.3.3 Posterior Kriging Algorithm

Appendix A contains the derivation of extended kriging equations, variances, and covariances. I include the universal kriging algorithm with-correlated-measurements, and its derivation, because it is given at best partially in standard geostatistical references.

Posterior kriging $Y$ estimates are denoted by $Y_K(x)$ and calculated with

$$Y_K(x) = \sum_{i=1}^{I} \omega_i(x) Y_M(x_i) + \sum_{p=1}^{P} \omega_p(x) Y_P(x_p) = \sum_{l=1}^{I+P} \omega_l(x) Y_{El} \quad (x \in \Omega) \quad (3.32)$$

where $\omega_k(x)$ is the $k$th kriging weight at point $x$, which is determined by solving the next $(I+P+R)$ extended kriging equations [Chilès and Delfiner, 1999] given in terms of both covariances and variograms

$$\sum_{l=1}^{I+P} \omega_l(x) C_Y(x_k - x_l) + \sum_{l=1}^{I+P} \omega_l(x) E_{kl} - \sum_{r=1}^{R} \mu_r(x) f_r(x_k)$$

$$= C_Y(x_k - x) \quad (k = 1, 2, \ldots I+P; \ x \in \Omega) \quad (3.33)$$

or

$$\sum_{l=1}^{I+P} \omega_l(x) \gamma_Y(x_k - x_l) - \sum_{l=1}^{I+P} \omega_l(x) E_{kl} + \sum_{r=1}^{R} \mu_r(x) f_r(x_k)$$

$$= \gamma_Y(x_k - x) \quad (k = 1, 2, \ldots I+P; \ x \in \Omega) \quad (3.34)$$

and

$$\sum_{i=1}^{I} \omega_i(x) f_r(x_i) + \sum_{p=1}^{P} \omega_p(x) f_r(x_p) = \sum_{l=1}^{I+P} \omega_l(x) f_r(x_l) = f_r(x)$$

$$\quad (r = 1, 2, \ldots, R; \ x \in \Omega) \quad (3.35)$$
Here \( C_Y(x - y) \neq C_Y^*(x - y) \), in general (\( C_Y^*(x - y) \) is the prior covariance of \( Y \)); 
\( \mu_r(x) \) is the \( r \)th Lagrange multiplier at point \( x \) that is obtained jointly with \( \omega_k(x) \); and \( R \) as well as \( f_r(x_p) \) were defined in Section (3.1). Posterior kriging variances are 
given by

\[
(Y_K'(x))^2 = \sigma_{Y_u}^2 - \sum_{I+P} \omega_l(x) C_Y(x_l - x) + \sum_{r=1}^R \mu_r(x) f_r(x) \quad (x \in \Omega) \quad (3.36)
\]
or

\[
(Y_K''(x))^2 = \sum_{I+P} \omega_l(y) \gamma_Y(x_l - x) + \sum_{r=1}^R \mu_r(x) f_r(x) \quad (x \in \Omega) \quad (3.37)
\]
where \( Y_K'(x) = Y(x) - Y_K(x) \). Extended kriging covariances are derived in Appendix A and given by

\[
\langle Y_K'(x) Y_K''(y) \rangle = C_Y(x - y) - \sum_{I+P} \omega_l(y) C_Y(x_l - x) + \sum_{r=1}^R \mu_r(y) f_r(x) \quad (x, y \in \Omega) \quad (3.38)
\]
or

\[
\langle Y_K'(x) Y_K''(y) \rangle = -\gamma_Y(x - y) + \sum_{I+P} \omega_l(y) \gamma_Y(x_l - x) + \sum_{r=1}^R \mu_r(y) f_r(x) \quad (x, y \in \Omega) \quad (3.39)
\]
Parameter estimates \( Y^{(i)}_E \) in Equation (3.32) act as \( Y \) “measurements” to be kriged at 
each iteration \( i \) of the inverse method. \( Y \) measurements may be optionally treated as 
unknown parameters in my inverse method because of their measurement uncertainty. 
If one chooses to do so, the new values, \( Y_K(x_i) \), will be, in general, somewhat different
than their measured counterparts $Y^*_m(x_i)$. This is so because including measurement errors in the kriging algorithm leads to a nonexact interpolation.

At the first inverse iteration, one sets $E_{kl}^{(0)} = E_{kl}^*$ in Equations (3.33) and (3.34). Also, note that $Q$, utilized as input in these equations, plays the role of a link between the geostatistical estimation of hydraulic parameters and the parameter estimation process.

Recalling the numerical grid definition in Chapter 2, the posterior kriging algorithm yields updated estimates of $\langle Y(x^c) \rangle_c$ and variances $\langle [Y'(x^c)]^2 \rangle_c$ at grid element centers as well as covariances $\langle Y'(x^c)Y'(y^c) \rangle_c$ between estimates at element midpoints. These quantities are needed at each forward model run of the inverse process. Upon inversion, they will be completely conditioned on all available data (hydraulic head, flux, conductivity).

The estimates determined by the posterior kriging algorithm would have the same properties of conditionality, unbiasedness, and optimality that their prior kriging counterparts have. However, the difference between the prior and posterior kriging algorithms lies in the statistical characteristics of the “measurements” to be kriged. Whereas the former considers only uncorrelated (real) measurements, the latter adds correlated pilot-point parameter estimates to the set of real measurements.

The condition of solution and uniqueness for the posterior kriging algorithm is similar to that of prior kriging (Equation (3.19)).
3.3.4 Computational Implementation of Posterior Kriging Algorithm

To my knowledge, all popular geostatistical computational packages ignore correlation between measurement (in my case inverse estimation) errors and none computes kriging covariances. Therefore, I needed to write a computer program for implementing my prior and posterior kriging algorithms.

Since posterior kriging is performed many times in the course of inversion, it is important to note the following:

- Equations (3.33)-(3.39) must be solved for each grid element. Fortunately, all solutions entail the same coefficient matrix so that Equations (3.33)-(3.35) must be factorized only once. The same factorization is applicable at each inverse iteration as long as the statistical parameters and pilot point configuration remain fixed. If statistical parameters are updated by inversion or if a different parameterization is considered, a new factorization of Equations (3.33)-(3.35) is required.

- Usage of Basic Linear Algebra Subprograms (BLAS) [Lawson et al., 1979; Dongarra et al., 1988, 1990] may improve code efficiency, especially if BLAS routines of levels 2 (vector-matrix operations) and 3 (matrix-matrix operations) are utilized.

- As recommended in Chapter 2, compiler optimization may improve computation speed considerably. If this compiling option is unavailable, one should resort
to efficient loop nesting and adequate handling of vectors and matrices, among
other effective programming techniques.

- Kriging covariance matrices are symmetric and, hence, only a triangular part
  of them needs to be computed.
Chapter 4
GEOSTATISTICAL INVERSE METHOD

This chapter describes the numerical inverse method used to condition (in the calibration sense) the finite-element nonlocal mean flow equations in Chapter 2 on hydraulic head and/or flux.

4.1 Inverse Problem

In this dissertation the mean flow equations (2.10), (2.14), and (2.25)-(2.35) are conditioned on head and flux data in addition to or instead of hydraulic conductivity data by inversion. Then, the second moment flow equations in Sections 2.4.2 and 2.4.3 are conditioned by solving them using parameters and mean flow variables already conditioned on head, flux, and/or conductivity data.

If calibration considers only head data, conditioning of mean flow is achieved through the mean head equations (2.25)-(2.34). If flux measurements are appended to or substituted for head data then all mean flow equations (2.10), (2.14), and (2.25)-(2.35) must be employed.

Let

$$\beta = (Y_{PE}, \theta)^T$$

(4.1)

be a vector of (partially or totally) unknown hydraulic and statistical parameters,
where

$$Y_{PE} = [Y_{PE}(x_1), Y_{PE}(x_2), \ldots, Y_{PE}(x_{Ny})]$$  \hspace{1cm} (4.2)

is a subvector of $Y_E$ formed by all parameter values at pilot points and either some or all $Y$ measurements; $Ny$ is total number of unknown hydraulic parameters ($P \leq Ny \leq I + P$); and the statistical parameter vector $\theta$ is defined in Section 3.1. Nonlocal head ($\langle h^{[2]}(x) \rangle$) and flux ($\langle q^{[2]}_{11}(x) \rangle$, $\langle q^{[2]}_{22}(x) \rangle$) predictions from Equations (2.10), (2.14), and (2.25)–(2.35) depend on parameters $\beta$ in a mathematically implicit manner. My numerical inverse problem consists of obtaining optimum parameters $\beta$ (partially or totally) based on discrete measurements of head, flux, and possibly conductivity data through an iterative inverse solution of the last equations. The solution depends on the parameterization chosen for ln $K$.

My parameterization, via pilot points, is defined in Equations (3.11) and (3.20). Let

$$\pi = [x_1, x_2, \ldots, x_P]^T$$  \hspace{1cm} (4.3)

be a vector defining the pilot point model utilized in an inverse problem, where $x_1, \ldots, x_P$ are pilot points and $P$ is their number. In general, $P$ should be small compared to the total number of head and flux measurements to avoid over-parameterization. One may arbitrarily postulate different $\pi$-models, solve the inverse problem for each, and select the solution that attains best results under given criteria. In Chapter 6, I illustrate such a type of selection by means of a sensitivity analysis. Alternatively, one may locate pilot points using the adjoint-state sensitivity procedure
of LaVenue and Pickens [1992], which places pilot points at the locations (on a grid laid out for that purpose) with largest sensitivity to changes in $\ln K$. This technique is ad hoc when the inverse solution is obtained using an adjoint state method; I do not consider such a type of solution in this dissertation.

My numerical examples in Chapter 5 are conditioned on flux only through a known, error-free source term (a well) entering the nonlocal mean head equations (2.25)–(2.34). Hereafter, I consider only head measurements in my inverse development but substitution by, or addition of, flux data is straightforward.

### 4.2 Statistics of Errors

Consider $J$ head measurements $h^*_j = h^*(x_j)$ located at points $x_j$ with associated measurement errors $\varepsilon^*_h = \varepsilon^*_h(x_j)$. The latter are defined through

$$h^*_j = h_j + \varepsilon^*_h = h(x_j) + \varepsilon^*_h(x_j) \quad (j = 1, 2, \ldots, J)$$

(4.4)

where $h_j = h(x_j)$ is (unknown) head at point $x_j$. Thus, I can define vectors of head measurements and measurement errors,

$$h^* = [h^*(x_1), h^*(x_2), \ldots, h^*(x_J)]^T$$

(4.5)

$$\varepsilon^*_h = [\varepsilon^*_h(x_1), \varepsilon^*_h(x_2), \ldots, \varepsilon^*_h(x_J)]^T$$

(4.6)

respectively.
Let \( \mathbf{Y}_{PE}^{*} \) be a prior hydraulic-parameter vector of dimension \( N_{Y}^{*} \) \((0 \leq N_{Y}^{*} \leq N_{Y})\) that is a subvector of vector \( \mathbf{Y}_{E}^{*} \) (Equation 3.20),

\[
\mathbf{Y}_{PE}^{*} = \left[ Y_{PE}^{*}(x_{1}), Y_{PE}^{*}(x_{2}), \ldots, Y_{PE}^{*}(x_{N_{Y}^{*}}) \right]^T
\]  
(4.7)

Its entries are initial estimates of corresponding entries of \( \mathbf{Y}_{PE} \) (some entries of \( \mathbf{Y}_{PE} \) may be regarded as having no priors if the latter are of poor or of uncertain quality). \( \mathbf{\varepsilon}_{Y_{PE}}^{*} \) is vector of prior errors and defined as

\[
\mathbf{\varepsilon}_{Y_{PE}}^{*} = \left[ \varepsilon_{Y_{PE}}^{*}(x_{1}), \varepsilon_{Y_{PE}}^{*}(x_{2}), \ldots, \varepsilon_{Y_{PE}}^{*}(x_{N_{Y}^{*}}) \right]^T
\]  
(4.8)

where errors \( \varepsilon_{Y_{PE}}^{*} \) form a subvector of errors \( \varepsilon_{Y_{E}}^{*} \) defined in Equation (3.22).

Let \( \mathbf{h}_{m} \) be a vector of interpolated (model) head predictions at head-measurement points,

\[
\mathbf{h}_{m} = \mathbf{M} (\mathbf{h}, \mathbf{h}_{D})
\]  
(4.9)

where \( \mathbf{M} \) is a \( J \times (N + N_{D}) \) interpolation matrix constructed with the bilinear Lagrange shape functions \( \psi \) in Section 2.4; \( N \) and \( N_{D} \) are numbers of non-Dirichlet and Dirichlet nodes on the finite element grid, respectively; \( \mathbf{h} \) (Equation (2.38)) is vector of second-order-approximation mean heads at non-Dirichlet nodes; and \( \mathbf{h}_{D} \) is vector of prescribed (deterministic) heads at Dirichlet nodes.

I adopt the following statistical assumptions made by Carrera and Neuman [1986a, b]:

- Head measurement \( \varepsilon_{h}^{*} \) and prior \( \varepsilon_{Y_{PE}}^{*} \) errors have Gaussian distributions with zero mean and covariance matrices \( \mathbf{S}_{h} \) and \( \mathbf{S}_{Y} \), respectively.
Both $S_h$ and $S_Y$ are symmetric and positive definite. $S_h = \sigma^2 z S_h^*$, where $S_h^*$ is covariance matrix of head measurement errors specified up to an unknown constant of multiplication $\sigma^2 z$; $S_Y = S_Y(\Theta)$, where $\Theta$ may be unknown.

Errors $\varepsilon_h^*$ are further assumed to be uncorrelated with themselves, $\varepsilon_{Y^*PE}, Y$, and $h$:

\begin{equation}
\langle \varepsilon_h^*(x) \varepsilon_h^*(y) \rangle = 0 \quad x \neq y \quad x, y \in \Omega \tag{4.10}
\end{equation}

\begin{equation}
\langle \varepsilon_h^*(x) \varepsilon_{Y^*PE}(y) \rangle = 0 \quad x, y \in \Omega \tag{4.11}
\end{equation}

\begin{equation}
\langle \varepsilon_h^*(x) Y(y) \rangle = 0 \quad x, y \in \Omega \tag{4.12}
\end{equation}

\begin{equation}
\langle \varepsilon_h^*(x) h(y) \rangle = 0 \quad x, y \in \Omega. \tag{4.13}
\end{equation}

• Head measurement errors are approximated by head residuals: $\varepsilon_h^* \approx h^* - h_m$.

• Prior errors are approximated by $Y$ residuals: $\varepsilon_{Y^*PE}^* \approx Y_{PE}^* - Y_{PE}$.

• Carrera [1984] presented a detailed discussion of head errors entering parameter estimation. He assumed that the combined head-error distribution is Gaussian with zero mean, which renders $\langle h^* - h_m \rangle = 0$ if the flow system is truly simulated by the model using optimum parameters. This is an ad hoc hypothesis for maximum likelihood theory I follow herein.
4.3 Nonlinear Parameter Estimation

4.3.1 Maximum Likelihood Formulation

The maximum likelihood estimate of $\beta$ given data $z^* = (h^*, Y^*_{PE})^T$ is the value of $\beta$ that maximizes the probability density of $z^*$ given $\beta$ [Casella and Berger, 2001; Carrera et al., 1997]. According to Carrera and Neuman [1986a] as well as Carrera et al. [1997], the likelihood function $L$ is given by

$$L(\beta | z^*) = f(z^* | \beta) = \frac{1}{\sqrt{(2\pi)^M |S_z|}} \exp \left[ \frac{- (z^* - z)^T S_z^{-1} (z^* - z)}{2} \right]$$

(4.14)

where $z = (h_m, Y_{PE})^T$ is a vector of model-computed counterparts of priors $z^*$; $(z^* - z)^T \approx (\varepsilon^*_h, \varepsilon^*_Y)^T$ is a vector of all residuals; $S_z$ is covariance of all priors given by a block diagonal matrix,

$$S_z = \begin{bmatrix} S_h & 0 \\ 0 & S_Y \end{bmatrix}$$

(4.15)

of dimension $M = J + N_Y^*$; $S_h$ is a diagonal $J \times J$ matrix of entries given by the actual covariances between measurement errors, $S_{hjk} = \langle \varepsilon_{hj} \varepsilon_{hk} \rangle$; and $S_Y = S_Y(\theta)$ is an $N_Y^* \times N_Y^*$ matrix whose components are given by the actual error covariances between $Y$ priors, $S_{Ykl} = \langle \varepsilon_{YPEk} \varepsilon_{YPEl} \rangle$. The block-diagonal property of the matrix $S_z$ follows from assumption (4.11); however, $S_Y$ is generally nondiagonal.

For simplicity, only head measurements and priors are considered in my inverse formulation, but statistically complying boundary-condition and forcing terms may be readily incorporated [Carrera and Neuman, 1986a].
When reliable prior information about the hydraulic parameters is unavailable, it may be better to ignore such information and rely exclusively on calibration against head and/or flux data. In such a case, $I = N_{Y*} = 0$, $z^* = h^*$, $z = h_m$, and $S_z = S_h$.

### 4.3.2 Simplification of Likelihood Function

Maximizing $L$ is equivalent to minimizing the negative log likelihood criterion $S$,

$$S = -2 \ln [L (\beta|z^*)]$$

(4.16)

which is easier to work with. $S$ can be written as [Carrera and Neuman, 1986a]

$$S = M \ln (2\pi) + \ln |S_z| + (z^* - z)^T S_z^{-1} (z^* - z)$$

(4.17)

While Carrera and Neuman [1986a, b, c] as well as Carrera et al. [1997] treated $S_Y$ as unknown up to a constant of multiplication, I allow $\theta$ and, thus, $S_Y$ to be totally unknown. I estimate $\theta$ a posteriori (Section 4.3.5) by performing an inversion for each $\theta$ belonging to a set of postulated $\theta$-vectors; in each inversion, $\theta$ is known. Thus, in minimizing $S$ the first two terms in Equation (4.17) are constant and may be dropped.

One then gets the criterion of “generalized sum of squared residuals”

$$F = (z^* - z)^T S_z^{-1} (z^* - z)$$

(4.18)

akin to nonlinear regression. If $\theta$ is unknown and no reliable $Y$ measurements are available one may set $S_z = S_h$ and still be able to estimate $\theta$ as described in Section 4.3.5.
Without losing generality, I further separate $S_z$ into unknown ($\sigma^2_z$) and partially known ($S_z^*$) parts,

$$S_z = \sigma^2_z S_z^* = \sigma^2_z \begin{bmatrix} S^*_h & 0 \\ 0 & S^*_Y \end{bmatrix}$$  \hspace{1cm} (4.19)$$

where $S^*_Y$ is potentially unknown (due to possibly unknown $\theta$) but provided via prior-kriging error variances-covariances and $Y$ measurement errors. A simplified criterion can be written as

$$F = (z^* - z)^T (S_z^*)^{-1} (z^* - z)$$  \hspace{1cm} (4.20)$$

Instead of minimizing $F$, it is numerically easier to optimize the equivalent criterion [Doherty, 2002a]

$$\tilde{F} = (\tilde{z}^* - \tilde{z})^T (\tilde{S_z^*})^{-1} (\tilde{z}^* - \tilde{z})$$  \hspace{1cm} (4.21)$$

in which $\tilde{S_z^*}$ is the diagonalized matrix of $S_z^*$ given by

$$\tilde{S_z^*} = \Theta^T S_z \Theta$$  \hspace{1cm} (4.22)$$

where $\Theta$ is an orthogonal matrix of dimension $M \times M$ whose columns are formed by the normalized eigenvectors of $S_z$; $\tilde{z}^*$ and $\tilde{z}$ are rotated vectors of $z^*$ and $z$,

$$\tilde{z}^* = \Theta^T z^* \hspace{1cm} \tilde{z} = \Theta^T z$$  \hspace{1cm} (4.23)$$

respectively. Note that Equation (4.21) is a "weighted least squares residual" criterion with both independent weights given by square roots of the diagonal entries of $(\tilde{S_z^*})^{-1}$ and rotated residuals ($\tilde{z}^* - \tilde{z}$). Since $S_h$ is diagonal, only $Y$ components in $z$ and $z^*$ may undergo rotation; head components and residuals are unchanged.
4.3.3 Minimization Algorithm

To estimate hydraulic parameters \( \beta \), the values of \( \Theta, S^*_x \), and \( z^* \) must be computed prior to starting the minimization of \( \tilde{F} \). Various numerical methods are available in the optimization literature to minimize \( \tilde{F} \). I use PEST \([Doherty, 2002a]\) which is based on a variant of the Levenberg-Marquardt algorithm that has proven efficient \([Cooley, 1985b; Carrera et al., 1997]\) and popular \([Hill, 1998; Poeter and Hill, 1998; Finsterle, 1999; Hill et al., 2000; Doherty, 2002a]\). The algorithm is based on a first-order approximation of the mathematical relationship between head and \( Y \). PEST iteratively updates the hydraulic parameter vector \( Y^i_{PE} \) by a vector \( d^i \),

\[
Y^{i+1}_{PE} = Y^i_{PE} + d^i
\]

where superscript \( i \) denotes iteration number and \( d^i \) is determined through the following set of calculations:

\[
\tilde{z}^i = \Theta^T z^i
\]

\[
\tilde{F}^i = (\tilde{z}^* - \tilde{z})^T \left( S^*_z \right)^{-1} (\tilde{z}^* - \tilde{z})
\]

\[
\sigma^2 z^i = \frac{\tilde{F}^i}{M - N_Y} \quad [Doherty, 2002a]
\]

\[
J^i = \frac{\partial h^i_n}{\partial Y^i_{PE}} = (J \times N_Y) \text{ Jacobian or sensitivity matrix (determined as described in Section 4.3.4)},
\]

\[
J_z^i = \begin{bmatrix} J^i \\ M_1 \end{bmatrix} = (M \times N_Y) \text{ extended Jacobian matrix},
\]

\[
(4.24)
\]

\[
(4.25)
\]

\[
(4.26)
\]

\[
(4.27)
\]

\[
(4.28)
\]

\[
(4.29)
\]
\( \tilde{J}_z^i = \Theta^T J_z^i \) = rotated extended Jacobian matrix, \( (4.30) \)

\[ D^i = \left( \tilde{J}_z^i \right)^T \left( \tilde{S}_z^i \right)^{-1} \tilde{J}_z^i \] \( (4.31) \)

\[ g^i = -2 \left( \tilde{J}_z^i \right)^T \left( \tilde{S}_z^i \right)^{-1} (\tilde{z}^i - \tilde{z}^i) = \text{gradient of } F, \] \( (4.32) \)

\[ Q^i = -2\sigma_e^2 (g^i)^{-1} \] \( (4.33) \)

= estimate of parameter-estimation error covariance matrix,

\( \mu^i = \) Marquardt parameter (see below),

\[ \delta^i = - (D^i + \mu^i I)^{-1} \frac{g^i}{2} = \text{updating vector}, \] \( (4.34) \)

\[ \gamma^i = \tilde{J}_z^i \delta^i \] \( (4.35) \)

\[ c^i = \frac{(\tilde{z}^i - \tilde{z}^i)^T \left( \tilde{S}_z^i \right)^{-1} \gamma^i}{\gamma^i \left( \tilde{S}_z^i \right)^{-1} \gamma^i} = \text{adjustment factor of updating vector}, \] \( (4.36) \)

\[ d^i = c^i \delta^i = \text{adjusted updating vector}, \] \( (4.37) \)

where \( M_1 \) is an \( N_y^* \times N_y \) submatrix filled with ones and \( I \) is the identity matrix.

At every iteration \( i \), the values of \( J^i, J^i_z, \) and \( \tilde{J}_z^i \) are computed once but the other quantities are calculated generally several times. The iterative process is started with \( Y_{PE}^0 \) set equal to \( Y_{PE} \) (initial guesses) if priors (no priors) are available. The process is repeated until convergence is attained and estimates of parameters, \( \hat{Y}_{PE} \), and common error variance, \( \hat{\sigma}_e^2 \), are obtained. Estimates of parameters \( Y_{PE} = Y_{PE}(\theta) \) are optimum only when \( \theta \) is (see Section 4.3.5). To start, \( \mu^1 \) is set to a relatively large value which is made to decrease as optimization progresses. At iteration \( i \), \( \mu^i \) is repeatedly decreased and \( \hat{F} \) evaluated until no further lowering of \( \hat{F} \) is achieved;
then, $\mathbf{Y}_{PF}$ is updated and a new iteration is started. I impose lower and upper limits on hydraulic parameters to prevent them from attaining physically unreasonable values and to avoid numerical errors. Suitable procedures for adjusting the Marquardt parameter can be found in the works of Cooley and Naff [1990], Hill [1992; 1998], and Doherty [2002a]. PEST’s procedure is described in the latter work.

As mentioned in Section 3.3.2, the estimate of the parameter estimation covariance matrix, $\mathbf{Q}$ (Equation (4.33)), enters posterior kriging equations. If $\mathbf{Q}$ is ill-conditioned (possibly at first iterations), it is better to include it in kriging computations only until its condition number has been lowered. Otherwise, the system of Equations (4.25)–(4.37) in the minimization algorithm may become ill-conditioned and the problem may turn unsolvable.

Localized Flow Simulator. The localized solution [Guadagnini and Neuman, 1999a] of the nonlocal mean flow equations was described in Section 2.6. It is simpler to compute than the full nonlocal solution but it provides no measures of predictive uncertainty. However, one may use a localized flow simulator to get approximate parameter estimates by inversion; these may then be used as initial guesses in a full nonlocal inversion. For example, by doing so I was able to reduce computational execution time by about 30% in performing the inverse Case 3 in Chapter 5 based on head and $\ln K$ measurements without priors. Nevertheless, in replicating Case 2 which considers head information alone, employing the localized forward model resulted in
an execution time increase of about 55% as the full nonlocal solution converged in
five iterations while the localized approach converged in 11 "localized" plus three
"nonlocal" iterations. In each case example, the inverse results were comparable
between the two approaches.

4.3.4 Calculation of Jacobian

The Jacobian or sensitivity matrix $J$ is the most computationally intensive quantity
in the iterative inverse process. Three known approaches to estimate a Jacobian
include 1) approximating its partial derivatives through finite differences, 2) using an
adjoint-state or variational approach, or 3) solving a system of sensitivity equations.

The first approach (e.g., Poeter and Hill [1998], Doherty [2002a]) is the most
common; it is model-independent and requires no extra programming. It is well suited
for parallel computing where forward-model runs require similar computational efforts
and are conducted in parallel. Nevertheless, it becomes computationally demanding
as $N_Y$ increases. $N_Y + 1$ (or $2N_Y + 1$, if using a central finite-difference scheme)
forward-model runs need to be performed to evaluate finite differences in heads $h_m$
due to changes in parameters $Y_{PE}(x_k) (k = 1, 2, \ldots N_Y)$.

The adjoint state method [Chavent, 1975; Neuman, 1980; Carrera et al., 1997]
requires additional development and programming of adjoint state equations to com-
pute $J$. In particular, both the derivation and implementation of adjoint state equations
for the finite-element nonlocal model in Chapter 2 are complex. This method per-
forms $J+1$ forward-model runs and, therefore, may be faster than the finite difference approach when $N_y > J$ (or $N_y > J/2$ if using a central finite-difference scheme). This may occur in geostatistical parameterizations.

The sensitivity-equation method [Yeh, 1986; Sun, 1994] requires solving an equation system similar to the governing equations $N_y + 1$ times and is amenable to parallel computing. However, the implementation of this method is demanding in my case due to the complexity of the nonlocal mean head equations in Chapter 2.

To estimate $J$, I have opted for the finite difference approach because it does not require further computational code development and inversions may be run in parallel. However, since the reprogramming work I performed on the forward model code of Guadagnini and Neuman (as described in Section 2.5.3) rendered the code’s performance much faster, I did no longer need parallelization. Nevertheless, the latter might be required if one considers 1) a grid finer than mine (Section 2.4), 2) a larger domain, 3) a larger database, 4) transient or 5) three-dimensional flow.

In my inversions I utilized a forward finite-difference scheme at initial stages with large rates of reduction in $\tilde{F}$; later, I switched to a central scheme when more refinement was needed.

4.3.5 Inverse Estimation of Statistical Parameters

If statistical parameters $\theta$ are partially or totally unknown, one may get inverse estimates of them based on head and flux data with or without $K$ measurements. It
would be more rigorous to estimate them jointly with parameters $Y_{PE}(\theta)$ by minimizing the log likelihood criterion (Equation (4.17)) but it is more convenient to obtain them after getting $\bar{Y}_{PE}$ [Carrera and Neuman, 1986a; Vesselinov et al., 2001a, b]. To do so, one can postulate a set of likely $\theta$-vectors and perform an inversion for each. For example, one may set a sequence of potential values for each of the unknown parameters in $\theta = (n_0, S_0, \rho_1, \rho_2, \ldots)$ and form competing $\theta$-vectors out of possible combinations. A "suboptimal" vector $\tilde{\theta}$ may then be identified by comparing the inversions using suitable model-fit criteria such as $\tilde{F}$ (Section 4.4).

The inverse estimate $\tilde{\theta}$ obtained in this manner is conditional on head and/or flux as well as possibly conductivity data. In general, there is no guarantee that it is optimum because only a few points in the vector space of $\theta$ are usually searched in an arbitrary fashion. Although using a very large set of $\theta$ values may render $\tilde{\theta}$ optimum.

If prior estimates of parameters $\theta$ are available one may incorporate them in the estimation process. This was done recently by Vesselinov et al. [2001a, b] to estimate a statistical parameter in unsaturated fractured tuff by inversion of pneumatic cross-hole test data. They proposed a series of values for a variogram power exponent and selected a suboptimum exponent as the value that attained the best trade-off between a pressure residual criterion and a previously inferred statistical-parameter prior.

If the model (or form) of $C_Y$ or $\gamma_Y$ is unknown one may estimate it in the same manner I propose to estimate a suboptimal parameterization of $\ln K(x)$ (Section 4.4).
That is, by postulating different models and selecting the best among them according to suitable criteria.

In my numerical examples in Chapter 5, I consider only known $C_Y$ (both its form and $\theta$) and, thus, cannot utilize the procedure described above.

### 4.3.6 Linearized Error Analysis of Parameter Estimates

Parameter estimation errors are characterized by their covariance matrix $Q \equiv \langle (Y - Y_{PE})(Y - Y_{PE})^T \rangle$, where $Y$ is the vector of actual (unknown) $Y$ values, $Y_k$, at parameter locations, $x_k$ ($k = 1, 2, \ldots, N_Y$). According to the Cramer-Rao principle, a lower bound on $Q$ is given by the inverse of the Fisher information matrix $F$, whose components are given by [Seber and Wild, 1989]

$$F_{kl} = \frac{1}{2} \left\langle \frac{\partial^2 \tilde{F}}{\partial Y_{PEk} \partial Y_{PEl}} \right\rangle_{Y_{PE} = \tilde{y}_{PE}(\tilde{\theta})} \quad (4.38)$$

Note that $F = \frac{1}{2}H$, where $H$ is the Hessian matrix of $\tilde{F}$. I take

$$Q \approx F^{-1} \approx \left( \frac{H_1}{2} \right)^{-1} = \sigma_v^2 \left[ (J_z^i)^T S_z^{-1} (z^i - \bar{z}^i) \right]^{-1} \quad (4.39)$$

where $H_1$ is a linear approximation of $H$. $Q$ would be exact if head depended linearly on the hydraulic parameters. Deviation from such linearity implies an underestimation of $Q$. The lower bound $F^{-1}$ becomes closer to the exact $Q$ as the head residuals approach zero.

After computing $Q$ (Section 4.3.3), I input it into a posterior kriging calculation at each iteration of the optimization process, as explained in Section 3.3. Thus, $Q$
conveys information on both parameter uncertainty and correlation into the geostatistical conditioning process.

\( F \) and its inverse \( F^{-1} \) represent measures of the information about parameters \( Y_{PE} \) contained in \( z^* \) and of parameter uncertainty, respectively. Since \( Q \approx F^{-1} \), maximum (minimum) information content corresponds to minimum (maximum) parameter estimation uncertainty.

Information concerning parameter estimation uncertainty and correlation is encapsulated in the eigenvalues and eigenvectors of \( Q \) [Carrera and Neuman, 1986a]. The eigenvalues define the axes of an \( N_Y \)-dimensional hyperellipsoid. The ratio between the largest and smallest eigenvalues is termed “condition number” (\( c_Q \)) and represents a measure of the range of parameter estimation errors [Carrera and Neuman, 1986a]. If all parameters have identical variances and are mutually uncorrelated, then the hyperellipsoid becomes a hyperspheroid. Incorporating unbiased prior information about parameters entails homogenizing the eigenvalues, reducing their \( c_Q \) toward one, and ultimately improving the inverse problem’s well-posedness. Each eigenvector has \( N_Y \) components associated with corresponding parameter estimates. If an eigenvector associated with a relatively large eigenvalue is associated with two or more parameters to more-or-less equal degrees, the corresponding parameter estimates are cross-correlated.

If parameter estimation errors are multivariate normal and the model behaves linearly in the neighborhood of \( \hat{Y}_{PE} \), parameter estimation uncertainty may also be
assessed through separate \((1 - \alpha) \times 100\%\) linear confidence intervals, which are defined as

\[
|Y_k - Y_{PEk}| \leq t(N_{df}, 1 - \alpha/2) Q_{kk}^{1/2} \quad (k = 1, 2, \ldots, N_Y)
\]  

(4.40)

where \(t(N_{df}, 1 - \alpha/2)\) is a quantile of Student’s \(t\)-distribution with \(\alpha\) usually taken to be 0.05 (corresponding to a 95\% significance level), \(N_{df} = M - N_Y\) denotes degrees of freedom, and \(Q_{kk}\) is the \(k\)th diagonal entry of matrix \(Q\) representing variance of the \(k\)th parameter estimate. Small intervals imply greater accuracy if the above premises are fulfilled. In each of my numerical examples (Chapter 5), I regard computed linear confidence intervals as good indicators of parameter estimation uncertainty only if an analysis of residuals justifies the premises.

More accurate (nonlinear) evaluations of parameter estimation uncertainty may be found in the works of Vecchia and Cooley [1987], Carrera et al. [1990], Hill [1994], Cooley [1997, 1999], and Christensen and Cooley [1999a, b], but their computations are much more costly than those of linear confidence intervals.

4.3.7 Assessment of Statistical Assumptions and Parameter Model

The geostatistical inverse method given here provides parameter estimates that are in general limited in accuracy due to the 1) assumptions and (linear) approximations made in its derivation and 2) suitability of the \(\pi\)-model parameterizing the \(\ln K(x)\) field. For each potential model the assumptions of unbiasedness, optimality, and
normality may be validated a posteriori by means of analysis of residuals.

As mentioned in page 82, various $\pi$-models may be postulated to represent potential pilot point parameterizations. The most suitable $\pi$-model is selected based upon model discrimination and quality criteria such as those given in the next section. The arbitrariness with which alternative models are postulated renders the selected $\pi$-model suboptimal but best among the set of competing models.

4.4 Identification of Parameter Model

Here I provide criteria to choose the most suitable pilot point parameterization among a set of likely alternatives.

4.4.1 Discrimination Criteria incorporating Likelihood

Discrimination criteria based on statistical likelihood are provided below [Carrera and Neuman, 1986a, c]. These criteria are based on minimizing a measure of closeness between the true (unknown) model and the proposed models using information theory:

\[
\text{AIC} = S + 2N_Y = \text{Akaike identification criterion} \tag{4.41}
\]

\[
\text{MAIC} = S + N_Y \ln M = \text{modified Akaike identification criterion} \tag{4.42}
\]

\[
\text{HIC} = S + 2N_Y \ln (\ln M) = \text{Hannan identification criterion} \tag{4.43}
\]

\[
\text{KIC} = S + N_Y \ln \frac{M}{2\pi} - \ln |Q| = \text{Kashyap identification criterion} \tag{4.44}
\]
where $S$ is log likelihood criterion (Equation 4.17), $N_Y$ is number of (unknown) parameters, $M$ is total number of data, and $Q$ is parameter estimation covariance matrix. These criteria were originally derived by Akaike [1974], Akaike [1977], Hannan [1980], Kashyap [1982], respectively. Each criterion makes a trade-off among model complexity or parsimony, model fit, and database size in a different manner. One can choose, among the postulated models, the one with consistently smaller criterion values.

Previous works having applied likelihood model discrimination in hydrogeology are those of Hoeksema and Kitanidis [1985], Carrera and Neuman [1986a,c], Samper [1986], Samper and Neuman [1989a, b], Chen et al. [2000], and Vesselinov et al. [2001a, b], among others. In particular, Carrera and Neuman [1986a] as well as Samper and Neuman [1989a, b] have utilized these criteria and shown that the KIC criterion performs best as it considers data quality too.

### 4.4.2 Quality Criteria using Statistical and Graphical Analyses

Analysis of residuals and statistics of fit are useful means to discriminate among parameter models. A comprehensive description of these criteria is provided in the guidelines and methods of Hill [1998]. Parameter estimation (or calibration) criteria in her work include the aforementioned and following criteria:

1. The common error variance is defined in Equation (4.27). It is clearly sensitive
to over-parameterization.

2. The correlation coefficient $R_c$ is defined as

$$R_c \equiv \frac{(w^* - m_w^*)^T (w - m_w 1)}{[\{(w^* - m_w^*)^T (w^* - m_w^*)\}(w - m_w^*)^T (w - m_w 1)]^{1/2}}$$  \hspace{1cm} (4.45)

where

$$w^* = S_{z^*}^{-1/2} z^* \quad w = S_z^{-1/2} z$$  \hspace{1cm} (4.46)

$m_w^*$ and $m_w$ are averages of the components of vectors $w^*$ and $w$, respectively; and $1$ is an $M$-dimension vector with unit entries. $R_c$ is a measure of calibration fit that is independent of the number of data in $z^*$. This allows comparing model calibrations corresponding to different database sizes. Hill suggests discarding inversions having $R_c \leq 0.9$.

3. Graphs of weighted head or flux residuals against their corresponding weighted estimates should display a random behavior about 0. If not, these graphs can help visualize value range(s) at which correlation occurs. Also, analysis of the maximum, minimum, and mean of residuals may reveal data outliers or inconsistent data.

4. One- or two-dimensional plots of weighted head residuals should be random. To check for randomness numerically one can use the nonparametric "runs" test [Draper and Smith, 1998]. A run is a sequence of values (residuals) having the same sign. $u$ is number of runs in a list of residuals; $n_1$ and $n_2$ are number of
positive and negative residuals, respectively. If \( n_1, n_2 \geq 10 \) it is reasonable to assume \( u \) is normally distributed with mean

\[
\mu_u = \frac{2n_1n_2}{n_1 + n_2} + 1
\]

and variance

\[
\sigma_u^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}
\]

\( u \) should lie between critical values,

\[
z_{u,l} = \frac{u - \mu \pm 0.5}{\sigma_u}
\]

### 4.4.3 Usage of Discrimination and Quality Criteria

Comparing among proposed models so as to select the best among them may lead to different results depending on the criteria being used. My intention is not to favor a particular criterion but rather to use as many criteria as possible so as to better support my model selection. If no model seems to be superior while comparing all criteria's results then (a) it is reasonable to assume that all or various models may form equally likely representations of the true hydrogeological system, and (b) some criteria might suggest discarding poor or physically unlikely models.

### 4.5 Sensitivity Analysis of Inverse Model

To investigate the performance of the nonlocal inverse model under different simulated field conditions or calibration parameter values I conduct a sensitivity analysis of
it. The inverse model behavior is tested as a function of the following variables or conditions:

1. Type(s) of conditioning data,

2. availability of \( Y \) priors,

3. \( Y \) measurements being treated as parameters,

4. number of pilot points (\( P \)),

5. spatial distribution or configuration of pilot points (\( x_1, x_2, \ldots, x_P \)),

6. number of head measurements (\( J \)),

7. number of \( Y \) measurements (\( I \)),

8. unconditional \( Y \) variance (\( \sigma^2_\text{Y} \)),

9. integral scale components (\( \rho_1, \rho_2 \)), and

10. presence/absence of a conditioning pumping well (e.g., \( Q \)).

The sensitivity analysis is presented in Chapters 5 and 6. It is based on comparing minimum, maximum, mean, and root mean square (RMS)

\[
\text{RMS of } Z = \sqrt{\frac{1}{n} \sum_{i=1}^{n} Z_i^2} \quad (4.50)
\]

values of residuals of \( \langle Y(x^*_k) \rangle \), \( \langle h^{[\theta]}(x_i) \rangle \), \( \langle q_{21}^{[\theta]}(x^*_k) \rangle \), and \( \langle q_{22}^{[\theta]}(x^*_k) \rangle \) (where \( k = 1, 2, \ldots, N_E \); \( i = 1, 2, \ldots, N \); \( N_E \) = number of grid elements; \( N \) = number of grid
nodes) at grid element-centers or nodes of synthetic case examples. Residuals are computed with respect to synthetic reference or “true” values “known” everywhere. The sensitivity analysis also includes comparing minimum, maximum, mean, and RMS values of variances $\sigma^2_Y(x_i^e), \sigma^2_h(x_i) \approx C^{[2]}_h(x, x), \text{ and } C^{[2]}_{\varphi ij}(x_i^e, x_j^e), (i, j = 1, 2)$ at grid element-centers or nodes of the same case examples.

In Chapter 5, I also show a more thorough comparison of calibration error and residual analyses as well as of discrimination criterion values for the first six conditions listed above. The comparison additionally involves sections and maps of: (a) inversely estimated $\langle Y(x) \rangle$ and $\sigma^2_Y(x)$ fields; (b) up-to second-order conditional predictions of $h(x), q_{21}(x),$ and $q_{22}(x)$; and (c) associated second-order approximations of variances-covariances $\sigma^2_h(x), C^{[2]}_h(x, y), \text{ and } C^{[2]}_{\varphi ij}(x, y)(i, j = 1, 2).$
Chapter 5
NUMERICAL EXAMPLES OF GEOSTATISTICAL INVERSION

In this chapter, I present five numerical case examples that illustrate the performance of the geostatistical inverse methodology of conditioning the nonlocal moment flow equations in Section 2.4 on hydraulic head and conductivity parameters.

All my numerical inverse examples were performed on the rectangular numerical domain in Section 2.4 having prescribed (deterministic) head of 10 and 0 (measured in arbitrary consistent length units) along the left and right boundaries, respectively and no-flow boundaries at the top and bottom. A well pumping at unit rate is located at the center \((x_1 = 9, x_2 = 4)\). It allows obtaining inverse solutions based solely on head data.

The numerical forward model (i.e., the finite-element nonlocal mean head equations (2.25)-(2.34)) was coupled with both a (posterior) kriging program I wrote for this purpose based on the posterior kriging equations (3.32)-(3.39) and the inverse parameter estimation code PEST [Doherty, 2002a] to carry out geostatistical inversions. The forward model and geostatistical (FORTRAN) subroutines were optimized at the programming and compiling stages as described in Sections 2.5.3 and 3.3.4. The stopping criteria and calibration parameter values used in all inversions in this dissertation are the default values provided by PEST [Doherty, 2002a], unless
otherwise noted. All numerical inversions were performed on an SGI Origin 2000 supercomputer at The University of Arizona using a single processor.

5.1 Reference Fields of \( Y, \) Head, and Flux

I consider \( Y(x) \) to be multivariate Gaussian, statistically homogenous, isotropic, and having an unconditional exponential covariance with zero nugget

\[
C_Y(s) = S_0 e^{-s/\rho}
\]

where \( S_0 = \sigma_{Y,u}^2 \) is the sill, \( s \) is separation distance, and \( \rho \) \((\rho_1 = \rho_2)\) is (isotropic) integral or autocorrelation scale.

Figure 5.1 shows an unconditional \( \ln K(x) \) realization with \( \sigma_{Y,u}^2 = 1 \) that was generated using the Gaussian sequential simulator SGSIM [Deutsch and Journel, 1997]. This realization served as the reference ("true") \( Y(x) \) field from which \( Y \) "measurements" were taken to condition my numerical inversions on \( \ln K \).

Then, a reference ("true") head field (Figure 5.2) was computed using the unconditional \( Y(x) \) realization and a standard finite element algorithm so as to obtain head "measurements" required to perform inversions.

For the sake of comparison, a reference flux field (Figure 5.3) was determined from the reference head distribution by means of a standard finite element algorithm and bilinear Lagrange interpolation.
Figure 5.1. Reference (unconditional) realization of $Y(x)$ with $\sigma_{Y_u}^2 = 1$.

Figure 5.2. Reference head field corresponding to the unconditional $Y(x)$ realization with $\sigma_{Y_u}^2 = 1$. 
FIGURE 5.3. Reference field of flux corresponding to the unconditional $Y(x)$ realization with $\sigma^2_Y = 1$: (a) longitudinal and (b) transverse components.
FIGURE 5.4. Hydrogeological domain with layout of $Y$ measurement points ($\times$) for Case 1. A pumping well (•) with $Q = 1$ is located at the center.

5.2 Case 1: Conditioning only on $Y$

The first step was to solve the nonlocal moment equations in a forward mode by conditioning them solely on $Y$ measurements. This is similar to the conditional examples presented by Guadagnini and Neuman [1999a,b]. They estimated $\langle Y(x) \rangle_c$ and $\langle Y'(x)Y'(y) \rangle_c$ throughout the domain ($\Omega$) by averaging values obtained from conditional realizations of $Y$. I do the same using a universal kriging algorithm that considers correlated measurements (Section 3.3).

To perform forward modeling, I sampled 16 “measurements” from the reference $Y(x)$ field (Figure 5.1) at regularly spaced points shown in Figure 5.4, which also displays the boundary conditions and pumping well. This layout of measurement points is the same as that used by Guadagnini and Neuman [1999a,b]. The “mea-
Table 5.1. \( Y \) "measurements" extracted from the reference field.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Y</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.3399</td>
<td>0.262</td>
</tr>
<tr>
<td>2</td>
<td>-0.6829</td>
<td>0.505</td>
</tr>
<tr>
<td>3</td>
<td>-2.2117</td>
<td>0.110</td>
</tr>
<tr>
<td>4</td>
<td>-1.3581</td>
<td>0.257</td>
</tr>
<tr>
<td>5</td>
<td>-1.0701</td>
<td>0.343</td>
</tr>
<tr>
<td>6</td>
<td>0.9919</td>
<td>2.696</td>
</tr>
<tr>
<td>7</td>
<td>1.3276</td>
<td>3.772</td>
</tr>
<tr>
<td>8</td>
<td>-0.8931</td>
<td>0.409</td>
</tr>
<tr>
<td>9</td>
<td>-0.8101</td>
<td>0.445</td>
</tr>
<tr>
<td>10</td>
<td>1.2718</td>
<td>3.567</td>
</tr>
<tr>
<td>11</td>
<td>-2.0580</td>
<td>0.128</td>
</tr>
<tr>
<td>12</td>
<td>1.1871</td>
<td>3.278</td>
</tr>
<tr>
<td>13</td>
<td>0.6951</td>
<td>2.004</td>
</tr>
<tr>
<td>14</td>
<td>-0.7621</td>
<td>0.467</td>
</tr>
<tr>
<td>15</td>
<td>-1.2087</td>
<td>0.299</td>
</tr>
<tr>
<td>16</td>
<td>0.4591</td>
<td>1.583</td>
</tr>
</tbody>
</table>

measured" \( Y \) values are listed in Table 5.1 along with their corresponding \( K \) values. Their arithmetic \( Y \) mean is equal to -0.4039. They were utilized to estimate \( \langle Y(x) \rangle_c \) and \( \langle Y'(x)Y'(y) \rangle_c \) in \( \Omega \) as described above, using the "true" covariance function 
\[
C_Y(s) = \sigma_Y^2 \exp(-s/\rho) \text{ with } \sigma_Y^2 = 1 \text{ and } \rho = 1.
\]

The histograms and (normal) probability plots of head, \( Y \), and longitudinal and transverse flux residuals on the grid are shown in Figures 5.5, 5.6, 5.7, and 5.8, respectively.

Figure 5.9 depicts estimated fields of \( \langle Y(x) \rangle_c \) and \( \sigma_Y^2(x) \) as well as longitudinal and transverse sections of \( \sigma_Y^2(x) \). \( \langle Y(x) \rangle_c \) shows a smooth variability that is characteristic of kriged fields. Effect of conditioning is evident in the computed field of \( \sigma_Y^2(x) \) (Figures 5.9(b)-5.9(d)) at or near measurement points, but it becomes unno-
Figure 5.5. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1.

Figure 5.6. (a) Histogram of all $Y$ residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1.
Figure 5.7. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1.

Figure 5.8. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 1.
Figure 5.9. Fields of conditional (a) $\langle Y(x) \rangle_c$ and (b) $\sigma^2_Y(x)$ for Case 1. (c) Longitudinal and (d) transverse sections of $\sigma^2_Y(x)$. 
noticeable along the transverse section passing through the well (Figure 5.9(d)) due to the relatively short integral scale of 1.

Figure 5.10 displays contours and sections of computed head $\langle h(x) \rangle_c \approx \langle h^{[2]}(x) \rangle_c$ conditioned on $Y$; the reference head is included for comparison. $\langle h^{[2]}(x) \rangle_c$ is in relatively good agreement with the reference head, especially when viewed in sections. One should keep in mind that the conditional mean head is, in an average sense, close to each of all possible conditional realizations (of which the reference head is only one) and does not exhibit the random characteristic these realizations display.

Second-order approximation of head variance $\sigma_h^2(x) \approx C_h^{[2]}(x)$ is shown in Figure 5.11. It is zero at (deterministic) Dirichlet boundaries, with higher values upstream of the well than downstream of it, and it exhibits a sharp peak at the well. These results are more or less similar to those obtained by Guadagnini and Neuman [1999b] for a similar case with $\sigma^2_{\gamma u} = 4$; they compared favourably their nonlocal moment flow solution with a Monte Carlo simulation.

Figure 5.12 depicts head covariance $C_h^{[2]}(x, y)$ for two points of reference: one at the well (Figures 5.12(a) and 5.12(b)) and the other near a conditioning (measurement) point (Figures 5.12(c) and 5.12(d)). The longitudinal and transverse sections of $C_h^{[2]}(x, y)$ are normalized in their horizontal coordinates by $L_1$ and $L_2$, respectively. Note that the head covariance in the longitudinal sections peaks at or near the reference points and decays to zero at the Dirichlet boundaries in a nonsymmetric, smooth manner (more pronounced downstream than upstream of the well). In the transverse
FIGURE 5.10. (a) Contours of (forward) conditional mean head $\langle h^{[d]}(x) \rangle_c$ for Case 1. Sections of forward mean and reference heads at (b) $x_2 = 4$, (c) $x_2 = 5$, (d) $x_1 = 9$, and (e) $x_1 = 11$. 
FIGURE 5.11. (a) Contours of up-to second-order approximation of conditional head variance $\sigma_h^2$ for Case 1; (b) longitudinal and (c) transverse sections.
Figure 5.12. (a) Contours and (b) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (9, 4)$. (c) Contours and (d) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (6.8, 4.8)$. Case 1.
sections, $C_h^{[2]}(x,y)$ reaches its largest value at or near the reference point $y$ and decreases to nonzero values at the top and bottom boundaries in a gradual manner. As in the work of Guadagnini and Neuman [1999b], I find that head covariance is greater in the transverse than in the longitudinal direction.

Contours and sections of longitudinal $(q_{xz1}(x))_c \approx (q_{xz1}^{[2]}(x))_c$ and transverse $(q_{xz2}(x))_c \approx (q_{xz2}^{[2]}(x))_c$ fluxes are shown in Figure 5.13; reference components of flux are included in the sections for comparison. Nonlocal flux components $(q_{xz1}^{[2]}(x))_c$ and $(q_{xz2}^{[2]}(x))_c$ clearly capture the mean trend of reference counterparts even at zones of high gradient near the well. The effect of conditioning is more evident in estimated $q_{xz1}(x)$ (Figures 5.13(c) and 5.13(e)) than in estimated $q_{xz2}(x)$ (Figures 5.13(h) and 5.13(j)).

Figure 5.14 displays contours and sections of flux variances $C_{q1x}^{[2]}(0)$, $C_{q2x}^{[2]}(0) = C_{q21}(0)$, and $C_{q22}(0)$. $C_{q11}(0)$ and $C_{q22}(0)$ exhibit positive peaks at the well, while $C_{q12}(0)$ exhibits both positive and negative peaks at the well. The variance of longitudinal flux shows is generally larger than that of transverse flux. The zero-lag cross covariance between estimates of $q_{xz1}$ and $q_{xz2}$ is consistently smaller and less variable than the flux variances. Values of longitudinal and transverse flux variance are generally larger upstream than downstream of the well. The effect of conditioning in reducing uncertainty can be seen in the contours and sections of $C_{q11}(0)$ and $C_{q22}(0)$, although it is less discernable in the contours and sections of $C_{q12}(0)$. However, some conditioning points ($(6.9,4.9)$, $(10.9,2.9)$, and $(14.9,4.9)$) are associated with large Y-
FIGURE 5.13. (a) Contours of (forward) conditional mean longitudinal flux \( \langle q_{z1}(x) \rangle_c \); sections of forward and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), and (e) \( x_1 = 10.9 \).  (f) Contours of (forward) conditional mean transverse flux \( \langle q_{z2}(x) \rangle_c \); sections of forward and reference transverse fluxes at (g) \( x_2 = 3.9 \), (h) \( x_2 = 4.9 \), (i) \( x_1 = 8.9 \), and (j) \( x_1 = 10.9 \). Case 1.
Figure 5.13. (continued)
Figure 5.14. (a) Contours of conditional variance of longitudinal flux $C_{q_1}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q_2}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q_1q_2}(0) = C_{q_2q_1}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 1.
measurement values (see Figure 5.9(a) and Table 5.1), creating peaks in $C_{q22}^{(2)}(0)$ and, to a lesser degree, $C_{q11}^{(2)}(0)$ at those points (see Equation 2.52).

Depicted in Figure 5.15 are longitudinal and transverse sections of the tensor of flux covariance, $C_{qij}(x, y) \approx C_{qij}^{(2)}(x, y)$ ($i, j = 1, 2$), at two selected points of reference: near the well at (8.9,3.9) and at a measurement point (6.9,4.9). The abscissas represent separation distance normalized by $L_1$ and $L_2$ for the longitudinal and transverse sections, respectively. For the first reference point, all com-
FIGURE 5.14. (continued)
Figure 5.15. (a)—(d) Sections of conditional flux covariance $C^{(2)}_{qij}(x, y)$ $(i, j = 1, 2)$ with reference to point $y = (8.9, 3.9)$. (e)—(h) Sections of conditional flux covariance $C^{(2)}_{qij}(x, y)$ $(i, j = 1, 2)$ with reference to point $y = (6.9, 4.9)$. Case 1.
ponents of the flux covariance tensor display a positive peak at zero lag, with $C^{(2)}_{q12}(x, y_1 = 8.9, y_2 = 3.9)$ showing also a negative peak (Figures 5.15(a)-5.15(d)). Except for $C^{(2)}_{q12}(x, y_1 = 8.9, y_2 = 3.9)$, all components of the flux covariance exhibit a smooth, more-or-less symmetric decay with separation distance, attaining relatively small negative values at small separation distances. For the second reference point (Figures 5.15(e)-5.15(h)), $C^{(2)}_{q11}(x, y_1 = 6.9, y_2 = 4.9)$ and $C^{(2)}_{q22}(x, y_1 = 6.9, y_2 = 4.9)$ exhibit a behavior similar to that described above, with a less pronounced decay. $C^{(2)}_{q12}(x, y_1 = 6.9, y_2 = 4.9)$ reveals zero values at zero lag and higher correlation at normalized separation distances less than 0.2. $C^{(2)}_{q21}(x, y_1 = 6.9, y_2 = 4.9)$ has zero values at zero lag and a nonsymmetric pattern, with fluctuations about zero at large separation distance along the transverse section. Note that the the tensor of flux covariance is generally nonsymmetric (i.e., $C^{(2)}_{q12}(x, y) \neq C^{(2)}_{q21}(x, y)$) (see Figures 5.15(b)-5.15(c) and 5.15(f)-5.15(f)). Only at zero lag this tensor is symmetric; that is, $C^{(2)}_{q12}(0) = C^{(2)}_{q21}(0)$.

To help assess the overall uncertainty of the estimates, Table 5.2 presents a summary of residual and variance statistics (minimum, maximum, mean, and RMS) computed from all grid nodes or elements in the domain. $Y(x)$, $h(x)$, $q_{x1}(x)$, and $q_{x2}(x)$ represent “true” (reference) values; $Y_{K}(x)$ is (posterior) kriging estimate of $Y(x)$ (Section 3.3). All residual mean values are relatively small, as expected for unbiased estimates. The minimum (-3.49) and maximum (2.67) values of $Y$ residuals illustrate the large difference that may exist between true and mean $Y$ values; the same can be
TABLE 5.2. Grid summary statistics of computed residuals and variances for Case 1.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>-3.5</td>
<td>2.7</td>
<td>-0.15</td>
<td>0.90</td>
</tr>
<tr>
<td>$\sigma_Y^2(x)$</td>
<td>0.00000</td>
<td>1.1</td>
<td>0.88</td>
<td>0.89</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>-0.70</td>
<td>0.59</td>
<td>0.079</td>
<td>0.26</td>
</tr>
<tr>
<td>$\sigma_h^2(x)$</td>
<td>0.0028</td>
<td>0.93</td>
<td>0.34</td>
<td>0.38</td>
</tr>
<tr>
<td>$\langle q_{ax}^{[2]}(x) \rangle_c - q_{ax}(x)$</td>
<td>-1.9</td>
<td>0.48</td>
<td>-0.047</td>
<td>0.26</td>
</tr>
<tr>
<td>$\langle q_{az}^{[2]}(x) \rangle_c - q_{az}(x)$</td>
<td>-0.68</td>
<td>0.53</td>
<td>-0.018</td>
<td>0.13</td>
</tr>
<tr>
<td>$C_{q11}(0)$</td>
<td>0.013</td>
<td>0.32</td>
<td>0.056</td>
<td>0.065</td>
</tr>
<tr>
<td>$C_{q12}(0) = C_{q21}(0)$</td>
<td>-0.17</td>
<td>0.14</td>
<td>-0.00061</td>
<td>0.0081</td>
</tr>
<tr>
<td>$C_{q22}(0)$</td>
<td>0.00022</td>
<td>0.16</td>
<td>0.013</td>
<td>0.018</td>
</tr>
</tbody>
</table>

said for the residual of longitudinal flux and, to a lesser degree, the residuals of head and transverse flux. The RMS values of mean $Y$ and head residuals are relatively large, 0.90 and 0.26, respectively. The computed mean of $\sigma_Y^2(x)$ (0.875) is smaller than the unconditional value of 1 as a consequence of conditioning. $\sigma_h^2(x)$ attains a maximum of 0.931 at the well (Figure 5.11) and has a rather large mean value (0.343). Maximum values of $C_{q11}(0)$, $C_{q12}(0) = C_{q21}(0)$, and $C_{q22}(0)$ are observed at or near the well (Figure 5.14). Whereas the longitudinal and transverse fluxes display small means of variance (0.0561 and 0.0134, respectively) due to conditioning, the mean of zero-lag cross covariance between estimates of $q_{ax}$ and $q_{az}$ exhibits a value (-0.0006) close to zero.

### 5.3 Case 2: Conditioning only on Head

My first inverse case example considers only head data. I sampled the head reference field (Figure 5.2) at 36 “measurement” points, located randomly in the interiors of
Figure 5.16. Hydrogeological domain with layout of head measurement (+) and (numbered) pilot (P) points for Case 2. A pumping well (•) with $Q = 1$ is located at the center.

evenly spaced subdomains of $2 \times 2$ dimensions each. These points are denoted by “+” signs in Figure 5.16, where pilot points as well as previously stated boundary and flux conditions in the domain are shown. Pilot points are numbered in the figure and are located where $Y$ measurements were positioned in Case 1.

Geostatistical inversion was performed based on the given head data, assuming a known covariance function $C_Y(s) = \sigma^2_Y \exp(-s/\rho)$ with $\sigma^2_Y = 1$ and $\rho = 1$ for estimating $\langle Y(x) \rangle_c$ and $\langle Y'(x)Y'(y) \rangle_c$ in $\Omega$. (In practice, it is often the case that $C_Y$ or $\gamma_Y$ is partially or totally unknown. One may then estimate it using an inverse procedure as that described in Section 4.3.5.) The correlation coefficient $R_c$ (Equation 4.45) of the calibration attained a value of 0.9989 and the standard residual variance (Equation 4.27) was equal to $4.0 \times 10^{-2}$. 
Table 5.3 lists the resulting pilot point estimates with associated 95% confidence intervals and corresponding hydraulic conductivity values. Note that pilot point estimates 13 and 16 reached the upper limit value of 3.5 that I set on PEST for all inversions in this chapter. A redistribution of pilot points might have improved the sensitivity to these parameters, but I kept the distribution fixed to allow direct comparison with Case 1. Parameter estimates in Table 5.3 do not generally correspond to “measured” reference values in Table 5.1. Note also that all pilot point estimates are associated with wide confidence intervals (i.e., large uncertainty).

Figure 5.17 shows an image of the covariance matrix of parameter estimation errors, Q.
Figure 5.17. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 2.

Figure 5.18 shows an image of computed eigenvalues and eigenvectors of the covariance matrix of parameter estimation errors, $\mathbf{Q}$. The condition number ($c_Q$) of $\mathbf{Q}$ (ratio between the largest and smallest eigenvalues) was equal to $3.8 \times 10^3$. Components of the eigenvector associated with the largest eigenvalue show that Parameters 7 and 8 were significantly correlated and, thus, more uncertain than the other parameters.

The histograms and (normal) probability plots of head residuals at match or calibration points and on the grid shown in Figures 5.20 and 5.21, respectively. The hypothesis that head residuals are normally distributed is not rejected at a 95% confidence level.

The histograms and (normal) probability plots of $Y$, and longitudinal and transverse flux residuals on the grid are shown in Figures 5.22, 5.23, and 5.24, respectively.
Figure 5.18. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 2.

Figure 5.19. (a) Measured ($h^*$) and second-order approximation nonlocal mean ($h_m$) heads at measurement points; (b) head residuals ($h^* - h_m$) at the same points. Case 2.
Figure 5.20. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2.

Figure 5.21. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2.
Figure 5.22. (a) Histogram of all Y residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2.

Depicted in Figure 5.25 are estimated fields of $\langle Y(x) \rangle_c$ and $\sigma_Y^2(x)$ conditioned on head data (through inversion), as well as longitudinal and transverse sections of $\sigma_Y^2(x)$. From Figure 5.25(b) one can see that $\sigma_Y^2(x)$ achieved minimum values of only about 0.4. This is due to the absence of Y measurements. In general, variances are larger on the right half of $\Omega$ than on its left half (Figure 5.25(b)). Away from the influence of pilot points, $\sigma_Y^2(x)$ displays values higher than the unconditional $\sigma_Y^2 = 1$. The effect of conditioning is unnoticed in the vicinity of head measurement points because conditioning is achieved through kriging Y values only at pilot and (possibly) measurement points.

Contours and sections of computed head $\langle h^{[2]}(x) \rangle_c$ are presented in Figure 5.26, where the reference head is included for comparison. $\langle h^{[2]}(x) \rangle_c$ is, in the mean, close
Figure 5.23. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2.

Figure 5.24. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 2.
Figure 5.25. Fields of conditional (a) \( \langle Y(x) \rangle_c \) and (b) \( \sigma_Y^2(x) \) for Case 2. (c) Longitudinal and (d) transverse sections of \( \sigma_Y^2(x) \).
FIGURE 5.26. (a) Contours of (inverse) conditional mean head \( \langle h^2(x) \rangle_c \) for Case 2. Sections of inverse mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), and (e) \( x_1 = 11 \).
Up-to second-order approximation of head variance $\sigma_h^2(x)$ is depicted in Figure 5.27. The behavior of $\sigma_h^2(x)$ is similar to that in Case 1, but now $\sigma_h^2(x)$ is consistently smaller throughout $\Omega$, especially at the well, as a result of conditioning on head data. This is evident when Figure 5.27 is compared with Figure 5.11 and Table 5.4 is compared with Table 5.2.

Head covariance $C_h(x, y) \approx C_h^{[2]}(x, y)$ is depicted in Figure 5.28 for two reference
Figure 5.28. (a) Contours and (b) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (9, 4)$. (c) Contours and (d) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (6.8, 4.8)$. Case 2.

points located at the well (Figures 5.28(a) and 5.28(b)) and near a pilot point (Figures 5.28(c) and 5.28(d)). $C_h^{[2]}(x, y)$ behaves in a manner similar to that in Case 1 (Figure 5.28). However, it is now smaller than it was there due to its dependence on $\sigma^2_h(x)$ and $\sigma^2_h(y)$ [Guadagnini and Neuman, 1999b], both of which are now (Table 5.4) smaller than are there (Table 5.2).

Figure 5.29 displays contours and sections of mean longitudinal ($\langle q_{z1}(x) \rangle_c$) and transverse ($\langle q_{z2}(x) \rangle_c$) flux; the sections include reference $q_{z1}(x)$ and $q_{z2}(x)$. Mean longitudinal flux in Figures 5.29(a)–5.29(e) provides a poorer reflection of the reference flux than was the case in Figures 5.13(a)–5.13(e). In other words, conditioning
FIGURE 5.29. (a) Contours of (inverse) conditional mean longitudinal flux \( \langle q_{x1}(x) \rangle_c \); sections of inverse and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), and (e) \( x_1 = 10.9 \). (f) Contours of (inverse) conditional mean transverse flux \( \langle q_{x2}(x) \rangle_c \); sections of inverse and reference transverse fluxes at (g) \( x_2 = 3.9 \), (h) \( x_2 = 4.9 \), (i) \( x_1 = 8.9 \), and (j) \( x_1 = 10.9 \). Case 2.
on $Y$ data alone has yielded more accurate estimates of $q_{x1}(x)$ than did conditioning only on head data. This is not the case for transverse flux which appears to be equally accurate in both cases: compare Figures 5.13(f)–5.13(j) with Figures 5.29(f)–5.29(j).

Figure 5.30 displays contours and sections of flux variances $C^{[2]}_{q11}(0)$ and $C^{[2]}_{q22}(0)$ as well as cross covariance $C^{[2]}_{q12}(0) = C^{[2]}_{q21}(0)$ at zero lag. In general, $C^{[2]}_{q11}(0)$, $C^{[2]}_{q22}(0)$, and $C^{[2]}_{q12}(0)$ behave in the same manner as they did in Case 1 (Figure 5.14) throughout $\Omega$ (including at the well). However, longitudinal and transverse flux variances are now generally larger than they were in Case 1 (compare Table 5.4 with Table 5.2). They also display larger numbers in the upper half of $\Omega$ than in the lower half. Influence of conditioning on head data through pilot points can be seen at most of these points, where $C^{[2]}_{q11}(0)$ and $C^{[2]}_{q22}(0)$ were reduced (Figures 5.30(a) and 5.30(d)). However, these variances show relative highs at Pilot Points 10, 12, 13, and 16 (see Table 5.3 and Figure 5.16) as these points hold the largest $Y$ estimates. Zero-lag cross covariance between estimates of $q_{x1}$ and $q_{x2}$ shows similar values (Figures 5.30(g) and 5.30(h)) as it did in Case 1 (Figures 5.14(g) and 5.14(h)). Though it displays now fluctuations about zero in the left part of its longitudinal sections near pilot points.

Figure 5.31 presents longitudinal and transverse sections of the tensor of flux covariance, $C^{[2]}_{qij}(x, y)$ $(i, j = 1, 2)$, at reference points $y_1 = (8.9, 3.9)$ (near the well) and $y_2 = (6.9, 4.9)$ (at Pilot Point 10). At $y_1$ every component of the flux tensor (Figures 5.31(a)–5.31(d)) is similar to its counterpart in Case 1 (Figures 5.15(a)–
Figure 5.29. (continued)
FIGURE 5.30. (a) Contours of conditional variance of longitudinal flux $C_{q11}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q22}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q12}(0) = C_{q21}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 2.
Figure 5.30. (continued)
FIGURE 5.30. (continued)
Figure 5.31. (a)−(d) Sections of conditional flux covariance $C_{q_{ij}}^{[2]}(x, y) \ (i, j = 1, 2)$ with reference to point $y = (8.9, 3.9)$. (e)−(h) Sections of conditional flux covariance $C_{q_{ij}}^{[2]}(x, y) \ (i, j = 1, 2)$ with reference to point $y = (6.9, 4.9)$. Case 2.
Table 5.4. Grid summary statistics of computed residuals and variances for Case 2.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_R(x) - Y(x)$</td>
<td>-3.1</td>
<td>3.1</td>
<td>0.25</td>
<td>0.95</td>
</tr>
<tr>
<td>$\sigma_Y^2(x)$</td>
<td>0.44</td>
<td>1.4</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>-0.62</td>
<td>0.49</td>
<td>-0.029</td>
<td>0.21</td>
</tr>
<tr>
<td>$\sigma_1^2(x)$</td>
<td>0.0035</td>
<td>0.85</td>
<td>0.33</td>
<td>0.36</td>
</tr>
<tr>
<td>$\langle q_{x1}^{[2]}(x) \rangle_c - q_{x1}(x)$</td>
<td>-1.7</td>
<td>0.81</td>
<td>0.22</td>
<td>0.36</td>
</tr>
<tr>
<td>$\langle q_{x2}^{[2]}(x) \rangle_c - q_{x2}(x)$</td>
<td>-0.63</td>
<td>0.58</td>
<td>-0.028</td>
<td>0.13</td>
</tr>
<tr>
<td>$C_{q11}^{[2]}(0)$</td>
<td>0.060</td>
<td>0.55</td>
<td>0.22</td>
<td>0.23</td>
</tr>
<tr>
<td>$C_{q12}^{[2]}(0) = C_{q21}^{[2]}(0)$</td>
<td>-0.15</td>
<td>0.15</td>
<td>-0.0035</td>
<td>0.014</td>
</tr>
<tr>
<td>$C_{q22}^{[2]}(0)$</td>
<td>0.00079</td>
<td>0.17</td>
<td>0.027</td>
<td>0.031</td>
</tr>
</tbody>
</table>

5.15(d)), except that it exhibits now generally larger values. At $y_2$ all components
$C_{q_{yj}}^{[2]}(x, y_2)(i, j = 1, 2)$ (Figures 5.31(e)-5.31(h)) are more-or-less similar to their equivalents in Case 1 (Figures 5.15(e)-5.15(h)) but differ as follows. The longitudinal and transverse flux covariances are now much larger and relatively smaller than they were in Case 1, respectively; and $C_{q11}^{[2]}(x, y_2)$ is now more strongly affected by conditioning than it was in Case 1.

Table 5.4 lists a summary of residual and variance statistics computed on the grid. $Y$ residuals have a positive bias (0.25) and a RMS value (0.95) that are now relatively larger in magnitude than they were (-0.15 and 0.90, respectively) in Case 1. The minimum (0.44) and mean (1.2) of $\sigma_Y^2(x)$ on the grid are now significantly larger than they were (0.0 and 0.88, respectively) in Case 1; these results are affected by the large uncertainty of parameter estimates. Head residuals have now a reasonably unbiased mean (-0.029) and a RMS value (0.21) smaller than they had (0.26) in Case 1 as a consequence of conditioning on head data. The mean of $\sigma_1^2(x)$ is similar to that
Figure 5.32. Hydrogeological domain with layout of head (+) and \( Y \) (\( \times \)) measurement points as well as (numbered) pilot points (\( \mathbb{P} \)) for Case 3. A pumping well (\( \bullet \)) with \( Q = 1 \) is located at the center.

in Case 1 (Table 5.2). Longitudinal-flux residuals have now larger mean magnitude (0.22) and RMS value (0.36) than they had (0.047 and 0.26, respectively) in Case 1. The statistics of transverse-flux residuals are similar to those in Case 1. Mean values for \( C_{qy}^{(i)}(x, y)(i, j = 1, 2) \) are now significantly larger than they were in Case 1.

5.4 Case 3: Conditioning on both Head and \( Y \) without Priors

The second inverse case considers head and log conductivity data, with no priors. Head data are the same considered in Case 2. \( Y \) measurements were sampled from the reference field (Figure 5.1 at 16 “measurement” points located regularly as shown in Figure 5.32. Pilot points are also shown in this figure.

Geostatistical inversion was performed based on the given data, assuming a the
TABLE 5.5. Pilot point estimates obtained for Case 3. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates.

<table>
<thead>
<tr>
<th>Pilot point</th>
<th>Y</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.922 ± 1.2502</td>
<td>18.57</td>
</tr>
<tr>
<td>2</td>
<td>-2.554 ± 1.8816</td>
<td>0.078</td>
</tr>
<tr>
<td>3</td>
<td>1.533 ± 1.6081</td>
<td>4.632</td>
</tr>
<tr>
<td>4</td>
<td>0.829 ± 4.1493</td>
<td>2.290</td>
</tr>
<tr>
<td>5</td>
<td>-3.500 ± 5.7157</td>
<td>0.030</td>
</tr>
<tr>
<td>6</td>
<td>-0.007 ± 1.9983</td>
<td>0.993</td>
</tr>
<tr>
<td>7</td>
<td>1.877 ± 3.2366</td>
<td>6.531</td>
</tr>
<tr>
<td>8</td>
<td>-0.384 ± 3.4447</td>
<td>0.681</td>
</tr>
<tr>
<td>9</td>
<td>0.060 ± 2.7949</td>
<td>1.062</td>
</tr>
<tr>
<td>10</td>
<td>-2.993 ± 3.1920</td>
<td>0.050</td>
</tr>
<tr>
<td>11</td>
<td>0.775 ± 2.7804</td>
<td>2.170</td>
</tr>
<tr>
<td>12</td>
<td>1.076 ± 3.1722</td>
<td>2.932</td>
</tr>
<tr>
<td>13</td>
<td>1.235 ± 4.1125</td>
<td>3.439</td>
</tr>
<tr>
<td>14</td>
<td>-1.116 ± 3.4003</td>
<td>0.328</td>
</tr>
<tr>
<td>15</td>
<td>3.500 ± 1.8978</td>
<td>33.12</td>
</tr>
<tr>
<td>16</td>
<td>-2.258 ± 2.2709</td>
<td>0.105</td>
</tr>
</tbody>
</table>

The same (known) covariance function in Case 2. The correlation coefficient $R_c$ (Equation 4.45) of the calibration attained a value of 0.9989 and the standard residual variance (Equation 4.27) was equal to $4.0 \times 10^{-2}$. The results of this case can be seen in the figures and tables in this section and in Table 5.12 at the end of this chapter. Results can be described similarly as it was done in Case 2. I rather compare their summary statistics with those in the other cases.
FIGURE 5.33. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 3.

FIGURE 5.34. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 3.
Figure 5.35. (a) Measured ($h^*$) and second-order approximation nonlocal mean ($h_m$) heads at measurement points; (b) head residuals ($h^* - h_m$) at the same points. Case 3.

Figure 5.36. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3.
FIGURE 5.37. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3.

FIGURE 5.38. (a) Histogram of all Y residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3.
FIGURE 5.39. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3.

FIGURE 5.40. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 3.
FIGURE 5.41. Fields of conditional (a) $\langle Y(x) \rangle_c$ and (b) $\sigma_Y^2(x)$ for Case 3. (c) Longitudinal and (d) transverse sections of $\sigma_Y^2(x)$. 
FIGURE 5.42. (a) Contours of (inverse) conditional mean head $\langle h^{[2]}(x) \rangle_c$ for Case 3. Sections of inverse mean and reference heads at (b) $x_2 = 4$, (c) $x_2 = 5$, (d) $x_1 = 9$, (e) $x_1 = 11$, and (f) $x_1 = 10$. 
FIGURE 5.42. (continued)

FIGURE 5.43. (a) Contours of up-to second-order approximation of conditional head variance $\sigma_h^2$ for Case 3; (b) longitudinal and (c) transverse sections.
FIGURE 5.44. (a) Contours and (b) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (9, 4)$. (c) Contours and (d) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (6.8, 4.8)$. Case 3.
FIGURE 5.45. (a) Contours of (inverse) conditional mean longitudinal flux \( \langle q_{x1}(x) \rangle_c \); sections of inverse and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \),
(d) \( x_1 = 8.9 \), (e) \( x_1 = 10.9 \), and (f) \( x_1 = 9.9 \). (g) Contours of (inverse) conditional mean transverse flux \( \langle q_{x2}(x) \rangle_c \); sections of inverse and reference transverse fluxes at (h) \( x_2 = 3.9 \), (i) \( x_2 = 4.9 \), (j) \( x_1 = 8.9 \), (k) \( x_1 = 10.9 \), and (l) \( x_1 = 9.9 \). Case 3.
Transverse component of flux

Longitudinal component of flux

FIGURE 5.45. (Continued)
Transverse component of flux
Figure 5.46. (a) Contours of conditional variance of longitudinal flux $C_{q11}^{[2]}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q22}^{[2]}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q12}^{[2]}(0) = C_{q21}^{[2]}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 3.
FIGURE 5.46. (Continued)
Figure 5.46. (Continued)
TABLE 5.6. Grid summary statistics of computed residuals and variances for Case 3.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>-2.7</td>
<td>2.5</td>
<td>0.11</td>
<td>0.79</td>
</tr>
<tr>
<td>$\sigma^2_x(x)$</td>
<td>0.00000</td>
<td>1.0</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>-0.35</td>
<td>0.59</td>
<td>0.032</td>
<td>0.15</td>
</tr>
<tr>
<td>$\sigma^2_h(x)$</td>
<td>0.0014</td>
<td>0.50</td>
<td>0.19</td>
<td>0.22</td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{g21}(x) \rangle_c - q</em>{g21}(x)$</td>
<td>-1.8</td>
<td>0.51</td>
<td>0.052</td>
<td>0.23</td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{g22}(x) \rangle_c - q</em>{g22}(x)$</td>
<td>-0.67</td>
<td>0.62</td>
<td>-0.014</td>
<td>0.12</td>
</tr>
<tr>
<td>$C_{\theta l}^{[2]}(\mathbf{0})$</td>
<td>0.0017</td>
<td>0.38</td>
<td>0.078</td>
<td>0.088</td>
</tr>
<tr>
<td>$C_{\theta 1}(\mathbf{0}) = C_{\theta 2}^{[2]}(\mathbf{0})$</td>
<td>-0.17</td>
<td>0.14</td>
<td>-0.000016</td>
<td>0.011</td>
</tr>
<tr>
<td>$C_{\theta 22}^{[2]}(\mathbf{0})$</td>
<td>0.00030</td>
<td>0.16</td>
<td>0.019</td>
<td>0.024</td>
</tr>
</tbody>
</table>
Figure 5.47. (a)–(d) Sections of conditional flux covariance $C_{ij}^{(2)}(x, y)$ ($i, j = 1, 2$) with reference to point $y = (8.9, 3.9)$. (e)–(h) Sections of conditional flux covariance $C_{ij}^{(2)}(x, y)$ ($i, j = 1, 2$) with reference to point $y = (6.9, 4.9)$. Case 3.
TABLE 5.7. Pilot point estimates obtained for Case 4. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates.

<table>
<thead>
<tr>
<th>Pilot point</th>
<th>$Y$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.419 ± 0.4445</td>
<td>0.657</td>
</tr>
<tr>
<td>2</td>
<td>-0.390 ± 0.4061</td>
<td>0.677</td>
</tr>
<tr>
<td>3</td>
<td>-0.645 ± 0.4741</td>
<td>0.525</td>
</tr>
<tr>
<td>4</td>
<td>-0.560 ± 0.4698</td>
<td>0.571</td>
</tr>
<tr>
<td>5</td>
<td>-0.466 ± 0.4551</td>
<td>0.628</td>
</tr>
<tr>
<td>6</td>
<td>-0.408 ± 0.4613</td>
<td>0.665</td>
</tr>
<tr>
<td>7</td>
<td>-0.663 ± 0.4927</td>
<td>0.515</td>
</tr>
<tr>
<td>8</td>
<td>-0.308 ± 0.4675</td>
<td>0.735</td>
</tr>
<tr>
<td>9</td>
<td>-0.508 ± 0.4721</td>
<td>0.602</td>
</tr>
<tr>
<td>10</td>
<td>-0.660 ± 0.4913</td>
<td>0.517</td>
</tr>
<tr>
<td>11</td>
<td>-0.669 ± 0.4955</td>
<td>0.512</td>
</tr>
<tr>
<td>12</td>
<td>-0.230 ± 0.4801</td>
<td>0.794</td>
</tr>
<tr>
<td>13</td>
<td>-0.193 ± 0.4843</td>
<td>0.824</td>
</tr>
<tr>
<td>14</td>
<td>-0.265 ± 0.4910</td>
<td>0.767</td>
</tr>
<tr>
<td>15</td>
<td>-0.357 ± 0.4918</td>
<td>0.700</td>
</tr>
<tr>
<td>16</td>
<td>-0.077 ± 0.4660</td>
<td>0.926</td>
</tr>
</tbody>
</table>

5.5 Case 4: Conditioning on both Head and $Y$ with Priors

This case considers the same database and pilot point layout as Case 3 did, except that now priors, supplied by (prior) kriging, are considered in the calibration criterion $\tilde{F}$ (Equation (4.21)).

Geostatistical inversion was performed based on the given data, assuming a the same (known) covariance function in Case 2. The results of this case can be seen in the figures and tables in this section and in Table 5.12 at the end of this chapter. Results can be described similarly as it was done in Case 2. I rather compare their summary statistics with those in the other cases.
FIGURE 5.48. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 4.

FIGURE 5.49. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 4.
FIGURE 5.50. (a) Measured ($h^*$) and second-order approximation nonlocal mean ($h_m$) heads at measurement points; (b) head residuals ($h^* - h_m$) at the same points. Case 4.

FIGURE 5.51. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4.
Figure 5.52. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4.

Figure 5.53. (a) Rotated prior and estimated ln hydraulic conductivity for each eigenvalue of covariance matrix $S_Y$; (b) ln $K$ residuals at the same points. Case 4.
FIGURE 5.54. (a) Histogram of rotated $Y$ residuals corresponding to the eigenvalues of matrix $S_Y$ and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4.

FIGURE 5.55. (a) Histogram of all $Y$ residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4.
FIGURE 5.56. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4.

FIGURE 5.57. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 4.
FIGURE 5.58. Fields of conditional (a) $\langle Y(x) \rangle_c$ and (b) $\sigma_Y^2(x)$ for Case 4. (c) Longitudinal and (d) transverse sections of $\sigma_Y^2(x)$. 

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1.1
1.0
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

section at $x_1 = 3.9$

section at $x_1 = 4.9$

section at $x_1 = 8.9$

section at $x_1 = 10.9$

section at $x_1 = 9.9$

section at $x_1 = 10.9$
FIGURE 5.59. (a) Contours of (inverse) conditional mean head \( h^{[2]}(x) \) for Case 4. Sections of inverse mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), (e) \( x_1 = 11 \), and (f) \( x_1 = 10 \).
FIGURE 5.59. (continued)

FIGURE 5.60. (a) Contours of up-to second-order approximation of conditional head variance $\sigma^2_h$ for Case 4; (b) longitudinal and (c) transverse sections.
Figure 5.61. (a) Contours and (b) sections of conditional head covariance $C_h^{[2]}(\mathbf{x}, \mathbf{y})$ for reference point (●) $\mathbf{y} = (9, 4)$. (c) Contours and (d) sections of conditional head covariance $C_h^{[2]}(\mathbf{x}, \mathbf{y})$ for reference point (●) $\mathbf{y} = (6.8, 4.8)$. Case 4.
Figure 5.62. (a) Contours of (inverse) conditional mean longitudinal flux $\langle q_{x_1}(x) \rangle_c$; sections of inverse and reference longitudinal fluxes at (b) $x_2 = 3.9$, (c) $x_2 = 4.9$, (d) $x_1 = 8.9$, (e) $x_1 = 10.9$, and (f) $x_1 = 9.9$. (g) Contours of (inverse) conditional mean transverse flux $\langle q_{x_2}(x) \rangle_c$; sections of inverse and reference transverse fluxes at (h) $x_2 = 3.9$, (i) $x_2 = 4.9$, (j) $x_1 = 8.9$, (k) $x_1 = 10.9$, and (l) $x_1 = 9.9$. Case 4.
FIGURE 5.62 (Continued)

(i) 

(j) 

Longitudinal component of flux

Transverse component of flux
Figure 5.62. (Continued)
FIGURE 5.63. (a) Contours of conditional variance of longitudinal flux $C_{q11}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q22}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q12}(0) = C_{q21}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 4.
FIGURE 5.63. (Continued)
FIGURE 5.63. (Continued)
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>-3.4</td>
<td>2.5</td>
<td>-0.15</td>
<td>0.88</td>
</tr>
<tr>
<td>$\sigma^2_{\gamma}(x)$</td>
<td>0.00000</td>
<td>1.0</td>
<td>0.74</td>
<td>0.76</td>
</tr>
<tr>
<td>$\langle h^{(2)}(x) \rangle_c - h(x)$</td>
<td>-0.64</td>
<td>0.44</td>
<td>0.014</td>
<td>0.20</td>
</tr>
<tr>
<td>$\sigma^2_{\delta}(x)$</td>
<td>0.0023</td>
<td>0.61</td>
<td>0.20</td>
<td>0.22</td>
</tr>
<tr>
<td>$\langle q_{x1}^{(2)}(x) \rangle_c - q_{x1}(x)$</td>
<td>-1.9</td>
<td>0.45</td>
<td>-0.050</td>
<td>0.26</td>
</tr>
<tr>
<td>$\langle q_{x2}^{(2)}(x) \rangle_c - q_{x2}(x)$</td>
<td>-0.67</td>
<td>0.53</td>
<td>-0.018</td>
<td>0.12</td>
</tr>
<tr>
<td>$C_{111}^{(2)}(0)$</td>
<td>0.0010</td>
<td>0.31</td>
<td>0.046</td>
<td>0.054</td>
</tr>
<tr>
<td>$C_{112}^{(2)}(0) = C_{q21}^{(2)}(0)$</td>
<td>-0.15</td>
<td>0.13</td>
<td>-0.00043</td>
<td>0.0066</td>
</tr>
<tr>
<td>$C_{222}^{(2)}(0)$</td>
<td>0.00020</td>
<td>0.14</td>
<td>0.011</td>
<td>0.015</td>
</tr>
</tbody>
</table>
FIGURE 5.64. (a)–(d) Sections of conditional flux covariance $C_{ij}^{(2)}(x,y) \ (i, j = 1, 2)$ with reference to point $y = (8.9, 3.9)$. (e)–(h) Sections of conditional flux covariance $C_{ijkl}^{(2)}(x,y) \ (i, j = 1, 2)$ with reference to point $y = (6.9, 4.9)$. Case 4.
Table 5.9. Pilot point estimates obtained for Case 5. Upper and lower limit values of 3.5 and -3.5, respectively were imposed on parameter estimates.

<table>
<thead>
<tr>
<th>Pilot point</th>
<th>Y</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.460 ± 0.4584</td>
<td>0.232</td>
</tr>
<tr>
<td>2</td>
<td>-0.765 ± 0.4911</td>
<td>0.465</td>
</tr>
<tr>
<td>3</td>
<td>-2.276 ± 0.5013</td>
<td>0.103</td>
</tr>
<tr>
<td>4</td>
<td>-1.226 ± 0.4973</td>
<td>0.293</td>
</tr>
<tr>
<td>5</td>
<td>-1.223 ± 0.4603</td>
<td>0.294</td>
</tr>
<tr>
<td>6</td>
<td>0.817 ± 0.4783</td>
<td>2.265</td>
</tr>
<tr>
<td>7</td>
<td>1.251 ± 0.4893</td>
<td>3.495</td>
</tr>
<tr>
<td>8</td>
<td>-0.740 ± 0.4922</td>
<td>0.477</td>
</tr>
<tr>
<td>9</td>
<td>-0.864 ± 0.4551</td>
<td>0.421</td>
</tr>
<tr>
<td>10</td>
<td>1.078 ± 0.4745</td>
<td>2.937</td>
</tr>
<tr>
<td>11</td>
<td>-2.062 ± 0.4995</td>
<td>0.127</td>
</tr>
<tr>
<td>12</td>
<td>1.341 ± 0.4908</td>
<td>3.822</td>
</tr>
<tr>
<td>13</td>
<td>0.792 ± 0.4470</td>
<td>2.207</td>
</tr>
<tr>
<td>14</td>
<td>-0.900 ± 0.4859</td>
<td>0.406</td>
</tr>
<tr>
<td>15</td>
<td>-1.197 ± 0.4931</td>
<td>0.302</td>
</tr>
<tr>
<td>16</td>
<td>0.580 ± 0.4962</td>
<td>1.786</td>
</tr>
</tbody>
</table>

5.6 Case 5: Conditioning on both Head and Y, considering Y Measurements as Parameters

The last inverse case considers the same database in Cases 3 and 4. However, pilot points are now located on top of Y measurements (not shown). That is, Y measurements are now priors of parameters and considered imprecisely known.

Geostatistical inversion was performed based on the given data, assuming a the same (known) covariance function in Case 2. The results of this case can be seen in the figures and tables in this section and in Table 5.12 at the end of this chapter. Results can be described similarly as it was done in Case 2. I rather compare their summary statistics with those in the other cases.
Figure 5.65. Images of symmetric (a) covariance and (b) correlation matrices of parameter-estimation errors. Case 5.
Figure 5.66. (a) Eigenvalues and (b) image of eigenvectors corresponding to the covariance matrix of parameter estimation errors obtained for Case 5.

Figure 5.67. (a) Measured ($h^*$) and second-order approximation nonlocal mean ($h_m$) heads at measurement points; (b) head residuals ($h^* - h_m$) at the same points. Case 5.
FIGURE 5.68. (a) Histogram of calibrated-head residuals at measurement points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

FIGURE 5.69. (a) Histogram of all head residuals at non-Dirichlet nodes and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.
Figure 5.70. (a) Prior ($Y_{PE}$) and estimated ($\tilde{Y}_{PE}$) ln hydraulic conductivity at pilot points; (b) ln $K$ residuals ($Y_{PE} - \tilde{Y}_{PE}$) at the same points. Case 5.

Figure 5.71. (a) Histogram of $Y$ residuals at pilot points and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.
FIGURE 5.72. (a) Histogram of all $Y$ residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.

FIGURE 5.73. (a) Histogram of all longitudinal-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.
Figure 5.74. (a) Histogram of all transverse-flux residuals at element centers and a fitted normal probability density function; (b) probability plot of the same residuals. Case 5.
FIGURE 5.75. Fields of conditional (a) $\langle Y(x) \rangle_c$ and (b) $\sigma^2_Y(x)$ for Case 5. (c) Longitudinal and (d) transverse sections of $\sigma^2_Y(x)$. 
FIGURE 5.76. (a) Contours of (inverse) conditional mean head \( \langle h^{[2]}(x) \rangle_c \) for Case 5. Sections of inverse mean and reference heads at (b) \( x_2 = 4 \), (c) \( x_2 = 5 \), (d) \( x_1 = 9 \), and (e) \( x_1 = 11 \).
Figure 5.77. (a) Contours of up-to second-order approximation of conditional head variance $\sigma_h^2$ for Case 5; (b) longitudinal and (c) transverse sections.
FIGURE 5.78. (a) Contours and (b) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (9, 4)$. (c) Contours and (d) sections of conditional head covariance $C_h^{[2]}(x, y)$ for reference point (●) $y = (6.8, 4.8)$. Case 5.
FIGURE 5.79. (a) Contours of (inverse) conditional mean longitudinal flux \( q_{x1}(x) \); sections of inverse and reference longitudinal fluxes at (b) \( x_2 = 3.9 \), (c) \( x_2 = 4.9 \), (d) \( x_1 = 8.9 \), and (e) \( x_1 = 10.9 \). (f) Contours of (inverse) conditional mean transverse flux \( q_{x2}(x) \); sections of inverse and reference transverse fluxes at (g) \( x_2 = 3.9 \), (h) \( x_2 = 4.9 \), (i) \( x_1 = 8.9 \), and (j) \( x_1 = 10.9 \). Case 5.
Transverse component of flux

Figure 5.78. (continued)
FIGURE 5.80. (a) Contours of conditional variance of longitudinal flux $C_{q11}(0)$; (b) longitudinal and (c) transverse sections. (d) Contours of conditional variance of transverse flux $C_{q22}(0)$; (e) longitudinal and (f) transverse sections. (g) Contours of conditional cross covariance $C_{q12}(0) = C_{q21}(0)$ between longitudinal and transverse fluxes; (h) longitudinal and (i) transverse sections. Case 5.
TABLE 5.10. Grid summary statistics of computed residuals and variances for Case 5.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>-3.5</td>
<td>2.6</td>
<td>-0.17</td>
<td>0.90</td>
</tr>
<tr>
<td>$\sigma^2_t(x)$</td>
<td>0.046</td>
<td>1.1</td>
<td>0.88</td>
<td>0.90</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>-0.70</td>
<td>0.43</td>
<td>-0.0038</td>
<td>0.21</td>
</tr>
<tr>
<td>$\sigma^2_h(x)$</td>
<td>0.0027</td>
<td>0.99</td>
<td>0.34</td>
<td>0.38</td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{c1}(x) \rangle_c - q</em>{x1}(x)$</td>
<td>-1.9</td>
<td>0.46</td>
<td>-0.057</td>
<td>0.26</td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{c2}(x) \rangle_c - q</em>{x2}(x)$</td>
<td>-0.67</td>
<td>0.53</td>
<td>-0.019</td>
<td>0.13</td>
</tr>
<tr>
<td>$C^{[2]}_{c11}(0)$</td>
<td>0.0017</td>
<td>0.31</td>
<td>0.054</td>
<td>0.062</td>
</tr>
<tr>
<td>$C^{[2]}<em>{c12}(0) = C^{[2]}</em>{c21}(0)$</td>
<td>-0.16</td>
<td>0.14</td>
<td>-0.00069</td>
<td>0.0076</td>
</tr>
<tr>
<td>$C^{[2]}_{c22}(0)$</td>
<td>0.00021</td>
<td>0.16</td>
<td>0.013</td>
<td>0.016</td>
</tr>
</tbody>
</table>
FIGURE 5.80. (continued)
**Figure 5.81.** (a)–(d) Sections of conditional flux covariance $C_{ij}^{[2]}(x, y) \ (i, j = 1, 2)$ with reference to point $y = (8.9, 3.9)$. (e)–(h) Sections of conditional flux covariance $C_{ij}^{[2]}(x, y) \ (i, j = 1, 2)$ with reference to point $y = (6.9, 4.9)$. Case 5.
5.7 Identification of Model

In this section, I compare statistics of residuals, variances, database size, calibration, and likelihood criteria of the five cases presented previously.

Table 5.11 lists statistics of residuals and variances on the grid with an added significant figure, compared to results in previous sections, to allow for comparison. Mean residuals and variances as well as RMS residuals for each case are ranked based on magnitude, and a mean ranking is computed for each case by assigning equal weights to considered statistics. Residuals reveal accuracy of estimation or prediction, and variances are measures of uncertainty. Note that conditioning the moment equations on both log conductivity and head data is generally better than conditioning them on only one of these data sets. Conditioning on $Y$ data alone leads to better estimates of $Y$ and fluxes than conditioning only on head, while conditioning only on head data results in better estimates of $h(x)$ than in the opposite case.

Table 5.11 indicates also that ignoring prior information about log conductivities leads to closer fits between computed and true log conductivities, heads, and fluxes. This is so because absence of constraining prior information makes it possible to fit the model more closely to the available head data. Yet, taking prior information into account leads to a significant reduction in the estimation variance of log conductivity and the predictive uncertainty of flux, while resulting in only an insignificant increase in the predictive uncertainty of head. This confirms that a good model fit does not
necessarily insure superior predictive capabilities.

Considering $Y$ measurements as priors of parameters yields results that are poorer than those when this is not done. This is so because the inverse solution is not fully (i.e., geostatistically) conditioned on the $Y$ data, only through regression.

The rankings in Table 5.11 do not account for differences in database size or model complexity or fit; results from model discrimination criteria below do consider these differences.

Table 5.12 lists statistics of calibration, database size, and of model discrimination criteria, as well as their individual and mean rankings for inverse Cases 2–5. This table considers residuals only at match or calibration points. Based on all rankings, Case 3 offers the best model among all, despite its few degrees of freedom and slow convergence. The remaining cases perform in the mean more-or-less similarly.

In terms of only model discrimination criteria, Cases 2 and 3 provide the best models because of their better model fit resulting from not considering priors. Models in Cases 4 and 5 are comparable between each other, considering these criteria, due to their similar measures of model fit. These two cases further show that the hypothesis of their head residuals being normally distributed was rejected at a 95% confidence level.
Table 5.11. Summary statistics of residuals and variances computed on the grid for one forward (Case 1) and four inverse cases. Ranking of each case is shown in parentheses for proper statistics.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>Minimum</td>
<td>-3.49</td>
<td>-3.09</td>
<td>-2.68</td>
<td>-3.43</td>
<td>-3.51</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.67</td>
<td>3.09</td>
<td>2.53</td>
<td>2.51</td>
<td>2.61</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.149 (2)</td>
<td>0.254 (5)</td>
<td>0.107 (1)</td>
<td>-0.153 (3)</td>
<td>-0.173 (4)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.895 (3)</td>
<td>0.946 (5)</td>
<td>0.789 (1)</td>
<td>0.815 (2)</td>
<td>0.896 (4)</td>
</tr>
<tr>
<td>$\sigma^2_P(x)$</td>
<td>Minimum</td>
<td>0.00000</td>
<td>0.443</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.0459</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.07</td>
<td>1.41</td>
<td>1.04</td>
<td>1.03</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.875 (3)</td>
<td>1.26 (5)</td>
<td>0.781 (2)</td>
<td>0.739 (1)</td>
<td>0.885 (4)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.891</td>
<td>1.26</td>
<td>0.794</td>
<td>0.755</td>
<td>0.900</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>Minimum</td>
<td>-0.701</td>
<td>-0.618</td>
<td>-0.347</td>
<td>-0.645</td>
<td>-0.701</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.586</td>
<td>0.491</td>
<td>0.592</td>
<td>0.441</td>
<td>0.426</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.0788 (5)</td>
<td>-0.0288 (3)</td>
<td>0.0318 (4)</td>
<td>0.0141 (2)</td>
<td>-0.00377 (1)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.256 (5)</td>
<td>0.206 (3)</td>
<td>0.147 (1)</td>
<td>0.200 (2)</td>
<td>0.210 (4)</td>
</tr>
<tr>
<td>$\sigma^2_H(x)$</td>
<td>Minimum</td>
<td>0.00278</td>
<td>0.00352</td>
<td>0.00142</td>
<td>0.00227</td>
<td>0.00271</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.931</td>
<td>0.846</td>
<td>0.498</td>
<td>0.609</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.343 (5)</td>
<td>0.326 (3)</td>
<td>0.194 (1)</td>
<td>0.201 (2)</td>
<td>0.340 (4)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.381</td>
<td>0.364</td>
<td>0.219</td>
<td>0.223</td>
<td>0.381</td>
</tr>
</tbody>
</table>
### Table 5.11. (continued)

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle q_{x1}(x) \rangle_c - q_{x2}(x)$</td>
<td>Minimum</td>
<td>-1.91</td>
<td>-1.74</td>
<td>-1.79</td>
<td>-1.91</td>
<td>-1.93</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.485</td>
<td>0.813</td>
<td>0.514</td>
<td>0.454</td>
<td>0.460</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.0469 (1)</td>
<td>0.216 (5)</td>
<td>0.0523 (3)</td>
<td>-0.0503 (2)</td>
<td>-0.0574 (4)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.257 (3)</td>
<td>0.355 (5)</td>
<td>0.233 (1)</td>
<td>0.256 (2)</td>
<td>0.259 (4)</td>
</tr>
<tr>
<td>$\langle q_{x2}(x) \rangle_c - q_{x2}(x)$</td>
<td>Minimum</td>
<td>-0.677</td>
<td>-0.632</td>
<td>-0.673</td>
<td>-0.673</td>
<td>-0.667</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.531</td>
<td>0.575</td>
<td>0.621</td>
<td>0.533</td>
<td>0.531</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.0183 (3)</td>
<td>-0.0285 (5)</td>
<td>-0.0140 (1)</td>
<td>-0.0176 (2)</td>
<td>-0.0193 (4)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.125 (3)</td>
<td>0.132 (5)</td>
<td>0.121 (1)</td>
<td>0.124 (2)</td>
<td>0.126 (4)</td>
</tr>
<tr>
<td>$C_{q_{11}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.00131</td>
<td>0.0601</td>
<td>0.00169</td>
<td>0.00103</td>
<td>0.00172</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.321</td>
<td>0.549</td>
<td>0.383</td>
<td>0.310</td>
<td>0.307</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.0561 (3)</td>
<td>0.217 (5)</td>
<td>0.0776 (4)</td>
<td>0.0460 (1)</td>
<td>0.0536 (2)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.0650</td>
<td>0.230</td>
<td>0.0879</td>
<td>0.0537</td>
<td>0.0618</td>
</tr>
<tr>
<td>$C_{q_{12}}^{[2]}(0) = C_{q_{21}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>-0.165</td>
<td>-0.149</td>
<td>-0.169</td>
<td>-0.150</td>
<td>-0.161</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.135</td>
<td>0.151</td>
<td>0.139</td>
<td>0.134</td>
<td>0.135</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.00061</td>
<td>-0.00350</td>
<td>-0.0002</td>
<td>-0.00043</td>
<td>-0.00069</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.00812 (3)</td>
<td>0.0136 (5)</td>
<td>0.0114 (4)</td>
<td>0.00662 (1)</td>
<td>0.00765 (2)</td>
</tr>
<tr>
<td>$C_{q_{22}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.00022</td>
<td>0.00079</td>
<td>0.00030</td>
<td>0.00020</td>
<td>0.00021</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.159</td>
<td>0.165</td>
<td>0.160</td>
<td>0.141</td>
<td>0.156</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.0134 (3)</td>
<td>0.0270 (5)</td>
<td>0.0192 (4)</td>
<td>0.0114 (1)</td>
<td>0.0126 (2)</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.0176</td>
<td>0.0314</td>
<td>0.0237</td>
<td>0.0147</td>
<td>0.0165</td>
</tr>
</tbody>
</table>

Mean ranking: (3.2) (4.5) (2.2) (1.8) (3.3)
Table 5.12. Calibration, residual, and likelihood statistics. Ranking of each case is shown in parentheses for proper statistics. All cases consider $J = 36$ and $P = 16$. N.A. stands for “Not Applicable”.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of $Y$ data ($I$)</td>
<td>0</td>
<td>(4)</td>
<td>16</td>
<td>(1)</td>
</tr>
<tr>
<td>Number of data ($I + J$)</td>
<td>36</td>
<td>(4)</td>
<td>52</td>
<td>(1)</td>
</tr>
<tr>
<td>Number of priors ($N_Y$)</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>$M = J + N_Y$</td>
<td>36</td>
<td>36</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>Prior kriging?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Degrees of freedom ($N_{df}$)</td>
<td>20</td>
<td>(3)</td>
<td>20</td>
<td>(3)</td>
</tr>
<tr>
<td>Number of iterations</td>
<td>5</td>
<td>(3)</td>
<td>10</td>
<td>(4)</td>
</tr>
<tr>
<td>Correlation coefficient ($R_c$)</td>
<td>0.9989</td>
<td>(2)</td>
<td>0.9995</td>
<td>(1)</td>
</tr>
<tr>
<td>Condition number ($c_Q$)</td>
<td>$3.8 \times 10^3$</td>
<td>(4)</td>
<td>$1.4 \times 10^3$</td>
<td>(3)</td>
</tr>
<tr>
<td>Common error variance ($\sigma_e^2$)</td>
<td>0.04</td>
<td>(2)</td>
<td>0.016</td>
<td>(1)</td>
</tr>
<tr>
<td>Runs test</td>
<td>not rejected</td>
<td>not rejected</td>
<td>not rejected</td>
<td>not rejected</td>
</tr>
<tr>
<td>Normality of head residuals</td>
<td>not rejected</td>
<td>(1)</td>
<td>not rejected</td>
<td>rejected (3)</td>
</tr>
<tr>
<td>Normality of $Y$ residuals</td>
<td>N.A.</td>
<td>N.A.</td>
<td>not rejected</td>
<td>not rejected</td>
</tr>
<tr>
<td>Mean weighted head-residual</td>
<td>0.014</td>
<td>(4)</td>
<td>-0.0011</td>
<td>(1)</td>
</tr>
<tr>
<td>RMS weighted head-residual</td>
<td>0.15</td>
<td>(2)</td>
<td>0.095</td>
<td>(1)</td>
</tr>
<tr>
<td>Mean weighted $Y$-residual</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.0037</td>
<td></td>
</tr>
<tr>
<td>RMS weighted $Y$-residual</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Simplified log likelihood criterion ($F$)</td>
<td>0.80</td>
<td>(2)</td>
<td>0.32</td>
<td>(1)</td>
</tr>
<tr>
<td>Log likelihood criterion ($S$)</td>
<td>18.6</td>
<td>(2)</td>
<td>17.7</td>
<td>(1)</td>
</tr>
<tr>
<td>ln[$Q$]</td>
<td>9.83</td>
<td>(4)</td>
<td>-10.2</td>
<td>(3)</td>
</tr>
<tr>
<td>AIC</td>
<td>50.6</td>
<td>(2)</td>
<td>49.7</td>
<td>(1)</td>
</tr>
<tr>
<td>MAIC</td>
<td>75.9</td>
<td>(2)</td>
<td>75.0</td>
<td>(1)</td>
</tr>
<tr>
<td>HIC</td>
<td>59.4</td>
<td>(2)</td>
<td>58.5</td>
<td>(1)</td>
</tr>
<tr>
<td>KIC</td>
<td>36.7</td>
<td>(1)</td>
<td>55.8</td>
<td>(2)</td>
</tr>
<tr>
<td>Mean ranking</td>
<td>(2.6)</td>
<td>(1.6)</td>
<td>(2.4)</td>
<td>(2.8)</td>
</tr>
</tbody>
</table>
Chapter 6
SENSITIVITY ANALYSIS

This chapter presents most of the results of the sensitivity analysis performed on the nonlocal inverse model as outlined in Section 4.5; the other results of the sensitivity analysis are in Chapter 5.

Each of the following sections offers sensitivity results per calibration parameter or field condition. Due to space limitations I do not provide here a complete analysis of results as I do in the previous chapter. I rely on model calibration and grid summary statistics to carry out the comparative analysis. The inversions in this chapter were performed numerically in the same manner as those in Chapter 5. In all cases pilot points were placed at the centers of grid elements.

6.1 Number of Pilot Points \((P)\)

Three inverse case layouts with \(P = 8, 16, 32\) were designed as shown in Figures 6.1, 5.16 (pilot-point layout of Case 2 in Chapter 5), and 6.2, respectively. They were conditioned on the same error-free data utilized in Chapter 5 that consist of 36 head measurements and a pumping-well flux rate of \(Q = 1\).

Table 6.1 shows a brief summary of the calibration performance for each inverse case. One can notice that model fit (as measured by both the regression coefficient
**Figure 6.1.** Hydrogeological domain with layout of eight pilot points (P). A pumping well (●) with \( Q = 1 \) is located at the center.

**Figure 6.2.** Hydrogeological domain with layout of 32 pilot points (P). A pumping well (●) with \( Q = 1 \) is located at the center.
Table 6.1. Calibration statistics for three cases with different number of pilot points.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$P = 8$</th>
<th>$P = 16$</th>
<th>$P = 32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ($R_c$)</td>
<td>0.9974</td>
<td>0.9989</td>
<td>1.000</td>
</tr>
<tr>
<td>Common error variance ($\sigma^2$)</td>
<td>0.063</td>
<td>0.040</td>
<td>0.0079</td>
</tr>
<tr>
<td>Condition number ($c_Q$) of $Q$</td>
<td>$5.8 \times 10^2$</td>
<td>$3.8 \times 10^3$</td>
<td>$4.1 \times 10^5$</td>
</tr>
</tbody>
</table>

$R_c$ and common error variance $\sigma^2$ improves with increasing number of pilot points. However, parameter estimates become more uncertain with increasing $P$ as reflected by the condition number $c_Q$ of the covariance matrix of parameter estimation errors. This may be due to over-parameterization and associated parameter correlation.

A summary of residual and variance statistics of $\ln K$, head, and longitudinal and transverse fluxes as well as (zero-lag) cross-covariance statistics of estimates of the latter fluxes on the grid is presented in Table 6.2. $Y$-residual ($Y_K(x) - Y(x)$) statistics are comparable for all values of $P$; however, mean and root-mean-square (RMS) values of variance of $\ln K$ ($\sigma^2_Y(x)$) are largest for the case with smallest number (8) of pilot points. Head residuals ($\langle h^{[2]}(x) \rangle_c - h(x)$) get less biased and smaller in the mean with increasing $P$. Augmenting $P$ also reduces head uncertainty as indicated by head variance ($\sigma^2_h$) statistics. Whereas transverse-flux residual ($\langle q_{z2}^{[2]}(x) \rangle_c - q_{z2}(x)$) statistics are comparable for all cases, longitudinal flux residuals ($\langle q_{z1}^{[2]}(x) \rangle_c - q_{z1}(x)$) become relatively larger as more pilot-point parameters are used. In general, employing more pilot-point parameters results in increasing uncertainty in both longitudinal and transverse fluxes as well as in the zero-lag cross covariance $C_{q_{z1}^{[2]}(x)}^{[2]}(0) = C_{q_{z2}^{[2]}(x)}^{[2]}(0)$ between nonlocal predictions of $q_{z1}(x)$ and $q_{z2}(x)$. This is likely an effect of over-
Table 6.2. Grid summary statistics of computed residuals and variances for three cases having different number of pilot points.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>P = 8</th>
<th>P = 16</th>
<th>P = 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_k(\mathbf{x}) - Y(\mathbf{x})$</td>
<td>Minimum</td>
<td>-3.3</td>
<td>-3.1</td>
<td>-2.6</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.8</td>
<td>3.1</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.057</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.92</td>
<td>0.95</td>
<td>0.84</td>
</tr>
<tr>
<td>$\sigma^2_Y(\mathbf{x})$</td>
<td>Minimum</td>
<td>0.44</td>
<td>0.44</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.9</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.7</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>1.7</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>$\langle h^{[2]}(\mathbf{x}) \rangle_c - h(\mathbf{x})$</td>
<td>Minimum</td>
<td>-1.0</td>
<td>-0.62</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.54</td>
<td>0.49</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.16</td>
<td>-0.029</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.29</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>$\sigma^2_h(\mathbf{x})$</td>
<td>Minimum</td>
<td>0.0037</td>
<td>0.0035</td>
<td>0.0017</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.0</td>
<td>0.85</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.43</td>
<td>0.33</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.49</td>
<td>0.36</td>
<td>0.24</td>
</tr>
<tr>
<td>$\langle q_{x1}^{[2]}(\mathbf{x}) \rangle_c - q_{x1}(\mathbf{x})$</td>
<td>Minimum</td>
<td>-1.6</td>
<td>-1.7</td>
<td>-1.7</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.69</td>
<td>0.81</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.19</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.32</td>
<td>0.36</td>
<td>0.37</td>
</tr>
<tr>
<td>$\langle q_{x2}^{[2]}(\mathbf{x}) \rangle_c - q_{x2}(\mathbf{x})$</td>
<td>Minimum</td>
<td>-0.66</td>
<td>-0.63</td>
<td>-0.70</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.57</td>
<td>0.58</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.022</td>
<td>-0.028</td>
<td>-0.030</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>$C_{q_{11}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.048</td>
<td>0.060</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.44</td>
<td>0.55</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.23</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.23</td>
<td>0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>$C_{q_{12}}^{[2]}(0) = C_{q_{21}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>-0.14</td>
<td>-0.15</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.0021</td>
<td>-0.0035</td>
<td>-0.0059</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.010</td>
<td>0.014</td>
<td>0.034</td>
</tr>
<tr>
<td>$C_{q_{22}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.00051</td>
<td>0.00079</td>
<td>0.00056</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.16</td>
<td>0.17</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.019</td>
<td>0.027</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.022</td>
<td>0.031</td>
<td>0.034</td>
</tr>
</tbody>
</table>
parameterization (as suggested by results in Table 6.1) because increasing the number of pilot points on the grid leads to the creation of zones (not shown) of higher mean log conductivity and, thus, of higher flux uncertainty (see Equation 2.52). This conclusion can be readily drawn only by means of conditional second moment expressions for head and flux.

Thus, pilot points should be as numerous as possible while maintaining parameters uncorrelated.

6.2 Spatial Distribution of Pilot Points

Five different configurations or \( \pi \)-models of 32 pilot points each were used to investigate the sensitivity of the nonlocal inverse model to them. The first model, \( \pi_1 \), consists of the regular pilot-point layout shown in Figure 6.2. Figure 6.3 displays Model \( \pi_2 \) where pilot points are placed at locations of either (known) relative maxima, minima, or large gradients of \( \ln K \). Model \( \pi_3 \) in Figure 6.4 is formed by pilot points positioned as closely as possible to most of the 36 head-measurement points (recall that pilot points are consistently located at grid element centers). Figure 6.5 shows Model \( \pi_4 \) which is composed of pilot points placed randomly inside 32 \((2 \times 2)\) adjacent subdomain squares. A random layout of pilot points throughout the domain forms Model \( \pi_5 \) as displayed in Figure 6.5. Each \( \pi \)-model inversion was conditioned on the same data used in the previous section.

The calibration performance for each \( \pi \)-model inversion may be seen in Table
Figure 6.3. Hydrogeological domain with layout (π2) of 32 pilot points ([P]) positioned in zones of either relative maxima, minima, or large gradients of ln K. A pumping well (●) with Q = 1 is located at the center.

Figure 6.4. Hydrogeological domain with layout (π3) of pilot points ([P]) located at head-measurement points. A pumping well (●) with Q = 1 is located at the center.
FIGURE 6.5. Hydrogeological domain with layout (π₄) of 32 pilot points (P) placed randomly in the interiors of evenly spaced subdomains of 2 × 2 dimensions each. A pumping well (●) with Q = 1 is located at the center.

FIGURE 6.6. Hydrogeological domain with layout (π₅) of 32 pilot points (P) located randomly. A pumping well (●) with Q = 1 is located at the center.
Table 6.3. Calibration statistics for five cases having different distributions of 32 pilot points.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$\pi_1$</th>
<th>$\pi_2$</th>
<th>$\pi_3$</th>
<th>$\pi_4$</th>
<th>$\pi_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ($R_c$)</td>
<td>1.000</td>
<td>0.9999</td>
<td>0.9996</td>
<td>0.9996</td>
<td>0.9993</td>
</tr>
<tr>
<td>Common error variance ($\sigma^2_e$)</td>
<td>0.0079</td>
<td>0.0082</td>
<td>0.069</td>
<td>0.063</td>
<td>0.12</td>
</tr>
<tr>
<td>Condition number ($c_Q$) of $\mathbf{Q}$</td>
<td>$4.1 \times 10^5$</td>
<td>$9.6 \times 10^5$</td>
<td>$3.0 \times 10^5$</td>
<td>$2.6 \times 10^6$</td>
<td>$7.0 \times 10^6$</td>
</tr>
</tbody>
</table>

6.3. Model fit was better for Models $\pi_1$ and $\pi_2$, and it was worst for Model $\pi_5$. Computed condition numbers of $\mathbf{Q}$ yield Models $\pi_1$, $\pi_2$, and $\pi_3$ as probably having more certain parameter estimates and Model $\pi_5$ as likely holding the most uncertain and correlated parameter estimates.

Table 6.4 lists a grid summary of residual and variance statistics of $\ln K$, head, and longitudinal and transverse fluxes, as well as of $C_{q_1q_2}(0)$. All $\pi$-models yield similar $Y$-residual statistics except for Models $\pi_2$ and $\pi_5$ that exhibit the least biased residuals and worst statistics, respectively. Comparable $Y$-variance statistics are shown for all models except for $\pi_5$ which evidence the poorest results. While Models $\pi_1$ and $\pi_2$ indicate better fit of true heads, Model $\pi_5$ displays the least accurate predictions of head. Head prediction uncertainty is lowest for results associated with Model $\pi_1$ and highest for those corresponding to Model $\pi_5$. Table 6.4 shows (clearly) the worst and (relatively) the best performances by Models $\pi_5$ and $\pi_2$, respectively in terms of residual statistics of longitudinal and transverse fluxes. The same observation applies to longitudinal and transverse flux variances. Nevertheless, the zero-lag cross covariance between estimates of $q_{x1}(x)$ and $q_{x2}(x)$ shows better results for Models $\pi_2$, $\pi_3$, and $\pi_4$ and worst values for Model $\pi_5$. 
Table 6.4. Grid summary statistics of computed residuals and variances for five cases having different distributions of 32 pilot points.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>( \pi_1 )</th>
<th>( \pi_2 )</th>
<th>( \pi_3 )</th>
<th>( \pi_4 )</th>
<th>( \pi_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_K(x) - Y(x) )</td>
<td>Minimum</td>
<td>-2.6</td>
<td>-2.7</td>
<td>-2.9</td>
<td>-3.0</td>
<td>-1.9</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.8</td>
<td>2.4</td>
<td>2.9</td>
<td>3.1</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.23</td>
<td>0.020</td>
<td>0.37</td>
<td>0.33</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.84</td>
<td>0.76</td>
<td>1.0</td>
<td>1.0</td>
<td>1.8</td>
</tr>
<tr>
<td>( \sigma_Y^2(x) )</td>
<td>Minimum</td>
<td>0.41</td>
<td>0.23</td>
<td>0.58</td>
<td>0.47</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.7</td>
<td>1.5</td>
<td>1.7</td>
<td>1.6</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.3</td>
<td>1.3</td>
<td>1.4</td>
<td>1.4</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>1.3</td>
<td>1.3</td>
<td>1.4</td>
<td>1.4</td>
<td>2.9</td>
</tr>
<tr>
<td>( \langle h^{[2]}(x) \rangle_c - h(x) )</td>
<td>Minimum</td>
<td>-0.40</td>
<td>-0.53</td>
<td>-0.53</td>
<td>-0.69</td>
<td>-1.3</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.67</td>
<td>0.38</td>
<td>0.56</td>
<td>0.48</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.10</td>
<td>-0.10</td>
<td>-0.091</td>
<td>-0.070</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.18</td>
<td>0.18</td>
<td>0.23</td>
<td>0.23</td>
<td>0.43</td>
</tr>
<tr>
<td>( \sigma_h^2(x) )</td>
<td>Minimum</td>
<td>0.0017</td>
<td>0.00027</td>
<td>0.0021</td>
<td>0.0034</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.89</td>
<td>0.75</td>
<td>0.52</td>
<td>0.53</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.22</td>
<td>0.25</td>
<td>0.23</td>
<td>0.27</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.24</td>
<td>0.30</td>
<td>0.27</td>
<td>0.31</td>
<td>0.38</td>
</tr>
<tr>
<td>( \langle q_{x1}^{[2]}(x) \rangle_c - q_{a1}(x) )</td>
<td>Minimum</td>
<td>-1.7</td>
<td>-1.6</td>
<td>-1.5</td>
<td>-1.8</td>
<td>-0.67</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.5</td>
<td>0.61</td>
<td>1.0</td>
<td>0.87</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.25</td>
<td>0.10</td>
<td>0.33</td>
<td>0.30</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.37</td>
<td>0.25</td>
<td>0.41</td>
<td>0.40</td>
<td>3.5</td>
</tr>
<tr>
<td>( \langle q_{x2}^{[2]}(x) \rangle_c - q_{a2}(x) )</td>
<td>Minimum</td>
<td>-0.70</td>
<td>-0.63</td>
<td>-0.66</td>
<td>-0.60</td>
<td>-2.1</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.67</td>
<td>0.62</td>
<td>0.58</td>
<td>0.60</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.030</td>
<td>-0.020</td>
<td>-0.028</td>
<td>-0.022</td>
<td>-0.16</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.13</td>
<td>0.11</td>
<td>0.13</td>
<td>0.13</td>
<td>0.51</td>
</tr>
<tr>
<td>( C_{q11}^{[2]}(0) )</td>
<td>Minimum</td>
<td>0.063</td>
<td>0.021</td>
<td>0.085</td>
<td>0.080</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.8</td>
<td>0.57</td>
<td>0.76</td>
<td>0.58</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.25</td>
<td>0.15</td>
<td>0.31</td>
<td>0.28</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.28</td>
<td>0.17</td>
<td>0.32</td>
<td>0.29</td>
<td>9.5</td>
</tr>
<tr>
<td>( C_{q12}^{[2]}(0) = C_{q21}^{[2]}(0) )</td>
<td>Minimum</td>
<td>-0.16</td>
<td>-0.15</td>
<td>-0.17</td>
<td>-0.13</td>
<td>-3.9</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.14</td>
<td>0.16</td>
<td>0.13</td>
<td>0.17</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.0059</td>
<td>-0.00023</td>
<td>-0.0042</td>
<td>-0.0022</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.034</td>
<td>0.021</td>
<td>0.019</td>
<td>0.016</td>
<td>1.0</td>
</tr>
<tr>
<td>( C_{q22}^{[2]}(0) )</td>
<td>Minimum</td>
<td>0.00056</td>
<td>0.00024</td>
<td>0.00089</td>
<td>0.00090</td>
<td>0.0055</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.24</td>
<td>0.16</td>
<td>0.17</td>
<td>0.17</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.027</td>
<td>0.018</td>
<td>0.032</td>
<td>0.029</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.034</td>
<td>0.022</td>
<td>0.036</td>
<td>0.033</td>
<td>0.58</td>
</tr>
</tbody>
</table>
Overall, Models 1 and 2 yield the best performances among all models. In practice, zones of high or low $K$ value, or of high $K$ gradient, can be seldom delineated a priori. Thus, a regular layout (e.g., Model 2) of pilot points is suggested by these results. An alternative would be to (automatically) locate pilot points optimally (based on some statistical criterion) on the grid in my numerical inverse algorithm. This can be accomplished through an optimization procedure such as those related to observation network design (e.g., Carrera and Neuman [1986c], Chilès and Delfiner [1999]) or the one in the work of LaVenue and Pickens [1992].

6.3 Number of Head Measurements ($J$)

To investigate the sensitivity of the nonlocal inverse model to number of head measurements three different measurement layouts with $J = 18, 36, \text{ and } 72$ were designed as depicted in Figures 6.7, 5.16, and 6.8, respectively. 72 “exact” head measurements were extracted from the “true” head field described in Section 5.1 to be partially or totally included in each of the layouts shown in the figures. An inversion was performed using 16 pilot-point parameters for each case, which was conditioned on its corresponding $J$ head-measurements and the pumping-well flux rate of $Q = 1$ used in the previous sections.

Table 6.5 lists a summary of calibration results for each inverse case. All cases display comparable correlation-coefficient values but both $\sigma^2_\varepsilon$ and $c_Q$ become smaller with increasing number of head measurements. The value of $\sigma^2_\varepsilon = 0.17$ for the case
FIGURE 6.7. Hydrogeological domain with layout of 18 head-measurement (+) and 16 pilot (P) points. A pumping well (●) with \( Q = 1 \) is located at the center.

FIGURE 6.8. Hydrogeological domain with layout of 72 head-measurement (+) and 16 pilot (P) points. A pumping well (●) with \( Q = 1 \) is located at the center.
TABLE 6.5. Calibration statistics for three cases with different numbers of head measurements.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$J = 18$</th>
<th>$J = 36$</th>
<th>$J = 72$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ($R_c$)</td>
<td>0.9992</td>
<td>0.9989</td>
<td>0.9989</td>
</tr>
<tr>
<td>Common error variance ($\sigma^2_c$)</td>
<td>0.17</td>
<td>0.040</td>
<td>0.028</td>
</tr>
<tr>
<td>Condition number ($c_Q$) of $Q$</td>
<td>$2.3 \times 10^4$</td>
<td>$3.8 \times 10^3$</td>
<td>$3.3 \times 10^3$</td>
</tr>
</tbody>
</table>

with $J = 18$ is relatively large due to its small number ($N_{df} = M - N_Y = J - P = 2$) of degrees of freedom (see Equation (4.27)). This led to over-parameterization, which caused a relatively large condition number $c_Q$ of $2.3 \times 10^4$.

In Table 6.6 a grid summary of residual and variance statistics of $\ln K$, head, and longitudinal and transverse fluxes as well as of $C_{q_{12}}^{[2]}(0)$ is shown. Whereas statistics of $Y$ residuals are generally comparable for the three cases, uncertainty in $Y$ estimates decreases when $J$ rises, as indicated by statistics of $\sigma^2_Y(x)$. The ill-posedness of the case with $J = 18$ is clearly reflected in its large values of head-residual statistics as well as of variances of $\ln K$, head, and longitudinal flux. However, over-parameterization seems to exert little or no influence on its other computed quantities. Increasing the number of head measurements renders more accurate prediction of $q_{x1}(x)$ and more certain estimates of $\ln K(x)$ and $q_{x1}(x)$ as shown by the statistics of longitudinal-flux residuals and of variances $\sigma^2_Y(x)$ and $C_{q_{11}}^{[2]}(0)$. No definite dependence of the rest of the quantities on $J$ is apparent in Table 6.6.

Thus, it seems that making as many as possible head measurements available to nonlocal inverse modeling improves parameter estimation and state variable prediction. Then, $P$ should be small (compared to $J$) enough to avoid over-
Table 6.6. Grid summary statistics of computed residuals and variances for three cases with different numbers of head measurements.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>$J = 18$</th>
<th>$J = 36$</th>
<th>$J = 72$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>Minimum</td>
<td>-3.4</td>
<td>-3.1</td>
<td>-3.3</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.8</td>
<td>3.1</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.089</td>
<td>0.25</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.92</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>$\sigma^2_Y(x)$</td>
<td>Minimum</td>
<td>1.7</td>
<td>0.44</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>3.8</td>
<td>1.4</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>3.0</td>
<td>1.3</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>3.0</td>
<td>1.3</td>
<td>1.1</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>Minimum</td>
<td>-2.4</td>
<td>-0.62</td>
<td>-0.62</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.39</td>
<td>0.49</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.36</td>
<td>-0.029</td>
<td>-0.083</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.49</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td>$\sigma^2_h(x)$</td>
<td>Minimum</td>
<td>0.0042</td>
<td>0.0035</td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>5.1</td>
<td>0.85</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.51</td>
<td>0.33</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.59</td>
<td>0.36</td>
<td>0.37</td>
</tr>
<tr>
<td>$\langle q_{x1}^{[2]}(x) \rangle_c - q_{x1}(x)$</td>
<td>Minimum</td>
<td>-1.5</td>
<td>-1.7</td>
<td>-1.9</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.83</td>
<td>0.81</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.37</td>
<td>0.22</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.45</td>
<td>0.36</td>
<td>0.32</td>
</tr>
<tr>
<td>$\langle q_{x2}^{[2]}(x) \rangle_c - q_{x2}(x)$</td>
<td>Minimum</td>
<td>-0.65</td>
<td>-0.63</td>
<td>-0.64</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.57</td>
<td>0.58</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.022</td>
<td>-0.028</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>$C_{q_{11}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.22</td>
<td>0.060</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.54</td>
<td>0.55</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.37</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.37</td>
<td>0.23</td>
<td>0.17</td>
</tr>
<tr>
<td>$C_{q_{12}}^{[2]}(0) = C_{q_{21}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>-0.14</td>
<td>-0.15</td>
<td>-0.15</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.0028</td>
<td>-0.0035</td>
<td>-0.0029</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.011</td>
<td>0.014</td>
<td>0.016</td>
</tr>
<tr>
<td>$C_{q_{22}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.00043</td>
<td>0.00079</td>
<td>0.00042</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.15</td>
<td>0.17</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.014</td>
<td>0.027</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.017</td>
<td>0.031</td>
<td>0.028</td>
</tr>
</tbody>
</table>
parameterization, especially in cases (such as these) where conductivity data are unavailable.

### 6.3.1 Small Number of Head Measurements

Three inversions with a small number of head measurements were performed. The first case had $J = 5$ and $P = 4$ as shown in Figure 6.9; the second case used $J = 9$ and $P = 4$ as depicted in Figure 6.10; and the third case considered $J = 9$ and $P = 8$ as displayed in Figure 6.11. A different version of the latter case was performed with pilot points evenly placed along the line $x_2 = 4$, but its results are not shown because they were of poor quality. All head measurements were extracted from the computed head field in Section 5.1. Unlike all other cases in this chapter, each of the three
FIGURE 6.10. Hydrogeological domain with layout of nine head-measurement (+) and four pilot (P) points. A pumping well (•) with $Q = 1$ and a (hidden) head-measurement point are located at the center.

FIGURE 6.11. Hydrogeological domain with layout of nine head-measurement (+) and eight pilot (P) points. A pumping well (•) with $Q = 1$ and a (hidden) head-measurement point are located at the center.
Table 6.7. Calibration statistics for three cases having a small number of head measurements.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$J = 5$</th>
<th>$P = 4$</th>
<th>$J = 9$</th>
<th>$P = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ($R_e$)</td>
<td>0.9999</td>
<td>0.9961</td>
<td>0.9992</td>
<td></td>
</tr>
<tr>
<td>Common error variance ($\sigma^2$)</td>
<td>0.046</td>
<td>0.13</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Condition number ($c_Q$) of $Q$</td>
<td>$4.2 \times 10^2$</td>
<td>$4.5 \times 10^2$</td>
<td>$3.6 \times 10^3$</td>
<td></td>
</tr>
</tbody>
</table>

cases considered head to be known at the center of $\Omega$. Including this measurement is expected to improve hydraulic conductivity and state variable characterization around the well. Each case was conditioned on head and on the pumping well flux of $Q = 1$.

Table 6.7 lists performance measures of calibration for each case and indicates that model fit was best in the first inverse case. The first and third cases have only one degree of freedom, but over-parameterization is more evident in the latter as its $c_Q$ was one order of magnitude larger than those in the other inversions.

Residual and variance statistics of $Y$, head, and longitudinal and transverse fluxes as well as of $C_{q12}(0)$ on the grid are summarized in Table 6.8. In general, the first case shows the best results of statistics of both $\ln K$ and head, which is in agreement with its superior model-fit performance. However, its statistics of both fluxes are the worst. The second case consistently attains the worst results of statistics of both $Y$ and head, maybe because its number of pilot-point parameters is too small ($P = 4$) to fit reasonably its nine head measurements. Nevertheless, its statistic results for both fluxes are comparable to those of the third case. The latter case displays the best results for statistics of longitudinal flux. The second and third cases present large
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>J = 5</th>
<th>J = 9</th>
<th>J = 9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P = 4</td>
<td></td>
<td></td>
<td>P = 8</td>
</tr>
<tr>
<td>Y_K(x) - Y(x)</td>
<td>Minimum</td>
<td>-2.9</td>
<td>-3.7</td>
<td>-3.6</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>3.2</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.33</td>
<td>-0.46</td>
<td>-0.38</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.98</td>
<td>1.0</td>
<td>0.99</td>
</tr>
<tr>
<td>σ₉²_Y(x)</td>
<td>Minimum</td>
<td>0.35</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.5</td>
<td>3.2</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.4</td>
<td>2.9</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>1.4</td>
<td>2.9</td>
<td>2.3</td>
</tr>
<tr>
<td>⟨h<a href="x">2</a>⟩c - h(x)</td>
<td>Minimum</td>
<td>-0.54</td>
<td>-4.1</td>
<td>-3.1</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.87</td>
<td>0.36</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.21</td>
<td>-0.71</td>
<td>-0.47</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.35</td>
<td>0.90</td>
<td>0.63</td>
</tr>
<tr>
<td>σ₉²_h(x)</td>
<td>Minimum</td>
<td>0.0044</td>
<td>0.0044</td>
<td>0.0040</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.3</td>
<td>13</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.47</td>
<td>1.1</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.52</td>
<td>1.4</td>
<td>0.91</td>
</tr>
<tr>
<td>⟨q_x<a href="x">2</a>⟩c - q_x1(x)</td>
<td>Minimum</td>
<td>-1.5</td>
<td>-1.7</td>
<td>-1.8</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.74</td>
<td>0.54</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.29</td>
<td>0.10</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.38</td>
<td>0.27</td>
<td>0.25</td>
</tr>
<tr>
<td>⟨q_x<a href="x">2</a>⟩c - q_x2(x)</td>
<td>Minimum</td>
<td>-0.60</td>
<td>-0.59</td>
<td>-0.60</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.57</td>
<td>0.58</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.015</td>
<td>-0.014</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>C_q[2]_{11}(0)</td>
<td>Minimum</td>
<td>0.17</td>
<td>0.086</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.52</td>
<td>0.34</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.29</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.29</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>C_q[2]<em>{12}(0) = C_q[2]</em>{21}(0)</td>
<td>Minimum</td>
<td>-0.15</td>
<td>-0.13</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.15</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.00000</td>
<td>0.00046</td>
<td>0.00001</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.0068</td>
<td>0.0052</td>
<td>0.0061</td>
</tr>
<tr>
<td>C_q[2]_{22}(0)</td>
<td>Minimum</td>
<td>0.0012</td>
<td>0.00017</td>
<td>0.00020</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.17</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.032</td>
<td>0.0071</td>
<td>0.0082</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.035</td>
<td>0.0094</td>
<td>0.010</td>
</tr>
</tbody>
</table>
FIGURE 6.12. Hydrogeological domain with layout of twelve pilot points (P). A pumping well (●) with Q = 1 is located at the center.

maxima of both $\sigma^2_T(x)$ and $\sigma^2_R(x)$. Thus, it is apparent that the first case obtains the best results overall despite its small numbers of head data and pilot points as well as of degree of freedom. However, due to the small numbers of both $J$ and $P$ it is adventurous to draw any general conclusion from these cases as the relative configuration among the few measurement/pilot points might have a significant role in the performance of each case.

6.4 Number of Log Conductivity Measurements ($I$)

The sensitivity of the nonlocal inverse model to $I$ is explored with three inverse cases having 0, 16, and 32 $Y$ measurements as shown in Figures 6.12, 6.13, and 6.14, respectively. Each case shared the same pilot-point model with $P = 12$ as
FIGURE 6.13. Hydrogeological domain with layout of twelve pilot ([P]) and 16 Y-measurement (×) points. A pumping well (●) with Q = 1 is located at the center.

TABLE 6.9. Calibration statistics for three cases with different numbers of ln K measurements.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>I = 0</th>
<th>I = 16</th>
<th>I = 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ($\rho_c$)</td>
<td>0.9986</td>
<td>0.9985</td>
<td>0.9983</td>
</tr>
<tr>
<td>Common error variance ($\sigma^2$)</td>
<td>0.042</td>
<td>0.047</td>
<td>0.048</td>
</tr>
<tr>
<td>Condition number ($c_Q$) of Q</td>
<td>$2.5 \times 10^3$</td>
<td>$1.8 \times 10^3$</td>
<td>$1.5 \times 10^3$</td>
</tr>
</tbody>
</table>

displayed in the figures. Each inversion was conditioned on the 36 head measurements and pumping-well flux rate of $Q = 1$ previously considered in Section 5.1. In the latter two inverse cases conditioning on log conductivity data was accomplished only geostatistically; that is, no priors were considered in their objective functions.

Statistics of calibration are presented in Table 6.9 and show that all cases performed numerically alike in the parameter estimation process.

Computed residuals and variances of head and fluxes as well as $C_{412}(0)$ on the grid are listed in Table 6.10. From this table it is clear that, in general, there is a
### Table 6.10. Grid summary statistics of computed residuals and variances for three cases with different numbers of \(ln K\) measurements.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>(I = 0)</th>
<th>(I = 16)</th>
<th>(I = 32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Y_K(x) - Y(x))</td>
<td>Minimum</td>
<td>-3.2</td>
<td>-3.5</td>
<td>-3.0</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.9</td>
<td>2.6</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.13</td>
<td>-0.17</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.93</td>
<td>0.90</td>
<td>0.79</td>
</tr>
<tr>
<td>(\sigma^2_y(x))</td>
<td>Minimum</td>
<td>0.30</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.4</td>
<td>1.1</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.3</td>
<td>0.86</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>1.3</td>
<td>0.88</td>
<td>0.74</td>
</tr>
<tr>
<td>(\langle h^{[2]}(x) \rangle_c - h(x))</td>
<td>Minimum</td>
<td>-0.60</td>
<td>-0.67</td>
<td>-1.1</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.52</td>
<td>0.42</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.013</td>
<td>0.0033</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.21</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>(\sigma^2_h(x))</td>
<td>Minimum</td>
<td>0.0037</td>
<td>0.0026</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.0</td>
<td>0.91</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.36</td>
<td>0.30</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.40</td>
<td>0.34</td>
<td>0.18</td>
</tr>
<tr>
<td>(\langle q_{x1}^{[2]}(x) \rangle_c - q_{x1}(x))</td>
<td>Minimum</td>
<td>-1.8</td>
<td>-1.9</td>
<td>-1.9</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.74</td>
<td>0.48</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.13</td>
<td>-0.056</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.30</td>
<td>0.26</td>
<td>0.23</td>
</tr>
<tr>
<td>(\langle q_{x2}^{[2]}(x) \rangle_c - q_{x2}(x))</td>
<td>Minimum</td>
<td>-0.61</td>
<td>-0.66</td>
<td>-0.63</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.57</td>
<td>0.53</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.025</td>
<td>-0.020</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.13</td>
<td>0.13</td>
<td>0.12</td>
</tr>
<tr>
<td>(C_{q11}^{[2]}(0))</td>
<td>Minimum</td>
<td>0.036</td>
<td>0.0013</td>
<td>0.0011</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.37</td>
<td>0.30</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.16</td>
<td>0.053</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.17</td>
<td>0.061</td>
<td>0.056</td>
</tr>
<tr>
<td>(C_{q12}^{[2]}(0) = C_{q21}^{[2]}(0))</td>
<td>Minimum</td>
<td>-0.14</td>
<td>-0.16</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.15</td>
<td>0.14</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.0021</td>
<td>-0.00068</td>
<td>-0.00073</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.010</td>
<td>0.0078</td>
<td>0.0077</td>
</tr>
<tr>
<td>(C_{q22}^{[2]}(0))</td>
<td>Minimum</td>
<td>0.00062</td>
<td>0.00020</td>
<td>0.00024</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.16</td>
<td>0.15</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.021</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.025</td>
<td>0.017</td>
<td>0.016</td>
</tr>
</tbody>
</table>
FIGURE 6.14. Hydrogeological domain with layout of twelve pilot (P) and 32 Y-measurement (×) points. A pumping well (●) with Q = 1 is located at the center.

trend of improving the matching between estimates of the above quantities with their corresponding true values when increasing or making available hydraulic conductivity measurements. Their corresponding measures of uncertainty are also lowered when I is augmented. Although head prediction (as manifested by head residuals) seems to be the only quantity not affected by variation in I, its measure of uncertainty ($\sigma_h^2(x)$) and that of longitudinal flux prediction ($C_{q11}(0)$) exhibit the greatest reduction among all statistics when Y data are included or I is increased.

6.5 Unconditional Variance of Log Hydraulic Conductivity ($\sigma_{Y_{uv}}^2$)

To investigate the dependency of the nonlocal model on different degrees of heterogeneity I ran two inversions for each of the following values of $\sigma_{Y_{uv}}^2$: 1, 2, and 4. For
TABLE 6.11. Calibration statistics for three cases having different variance of \( Y \); each case performed twice as in Cases 3 and 4 in Chapter 5.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>As in Case 3</th>
<th>( \sigma_Y^2 = 1 )</th>
<th>( \sigma_Y^2 = 2 )</th>
<th>( \sigma_Y^2 = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ( (R_c) )</td>
<td>0.9995</td>
<td>0.9983</td>
<td>0.9987</td>
<td></td>
</tr>
<tr>
<td>Common error variance ( (\sigma_Y^2) )</td>
<td>0.016</td>
<td>0.059</td>
<td>0.048</td>
<td></td>
</tr>
<tr>
<td>Condition number ( (c_Q) ) of ( Q )</td>
<td>( 1.4 \times 10^3 )</td>
<td>( 3.0 \times 10^3 )</td>
<td>( 1.5 \times 10^3 )</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistic</th>
<th>As in Case 4</th>
<th>( \sigma_Y^2 = 1 )</th>
<th>( \sigma_Y^2 = 2 )</th>
<th>( \sigma_Y^2 = 4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ( (R_c) )</td>
<td>0.9981</td>
<td>0.9898</td>
<td>0.9917</td>
<td></td>
</tr>
<tr>
<td>Common error variance ( (\sigma_Y^2) )</td>
<td>0.062</td>
<td>0.53</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Condition number ( (c_Q) ) of ( Q )</td>
<td>5.6</td>
<td>6.8</td>
<td>14.</td>
<td></td>
</tr>
</tbody>
</table>

For this purpose I had previously generated corresponding reference ("true") unconditional realizations of \( \ln K(x) \) (Figures 5.1, 6.15(a), and 6.16(a), respectively), using the Gaussian sequential simulator SGSIM [Deutsch and Journel, 1997], and computed associated reference head fields (Figures 5.2, 6.15(b), and 6.16(b), respectively). From each set of reference \( Y(x) \) and \( h(x) \) realizations I sampled 16 log-conductivity and 36 head measurements, respectively, according to the layout shown in Figure 5.32. Then, for each sampled dataset I performed two inversions separately in the manners of Cases 3 and 4 in Chapter 5, using the same layout of 16 pilot points in each and conditioning also on a pumping well flux of \( Q = 1 \) at the center of \( \Omega \). Whereas Case 3 considered no prior information on pilot-point parameters, Case 4 did include it via prior kriging.

Calibration results in Table 6.11 indicate that all inversions performed similarly, although the cases with \( \sigma_Y^2 = 1 \) and \( \sigma_Y^2 = 2 \) were slightly the best and worst, respectively, regardless of priors being considered. Note that including priors in the
FIGURE 6.15. Reference (a) unconditional realization of $Y$ with $\sigma^2_{Y_u} = 2$ and (isotropic) $\rho = 1$. (b) Associated reference head field considers boundary conditions and pumping well with $Q = 1$. 
Figure 6.16. Reference (a) unconditional realization of $Y$ with $\sigma_{Y_u}^2 = 4$ and (isotropic) $\rho = 1$. (b) Associated reference head field considers boundary conditions and pumping well with $Q = 1$. 
calibration process yields a relatively poorer model fit (as indicated by the values of $R_c$ and $\sigma^2_z$), but a better-posed problem’s solution (as shown by the values of $c_Q$).

Table 6.12 lists residual and variance results for $\ln K$, head, and flux. All results consistently indicate that model performance decreases with increasing level of heterogeneity ($\sigma^2_{Y,u}$). However, all residuals stay relatively unbiased with varying $\sigma^2_{Y,u}$, as shown by their mean values. Conditional variance of $Y$ varies from 0 to $\sigma^2_{Y,u}$ on the grid of each case, but its mean (on the grid) is roughly 80% of its unconditional value in all inversions with or without priors. This percent value rather depends on $I$, $J$, and $\pi$-model, as it can be noted in previous sections of this chapter. The same type of dependence is found therein for $\sigma^2_h$, which is seen here to vary from approximately zero to 80% of $\sigma^2_{Y,u}$ on the grid, but whose mean value is between 22% and 24% of $\sigma^2_{Y,u}$ in all inversions regardless of whether or not priors are included. Maximum values for $C_{q11}(0)$ and $C_{q22}(0)$ are significantly larger for the two cases with $\sigma^2_{Y,u} = 4$. However, mean values for these variances and the RMS value of $C_{q12}(0)$ decrease when priors are considered.

6.6 (Isotropic) Integral Scale ($\rho$)

Two unconditional realizations of $\ln K(x)$ having (isotropic) $\rho$ values of 1 and 4 (and $\sigma^2_{Y,u} = 1$) were generated by means of the simulator SGSIM to study the behavior of the nonlocal model under variation of this parameter. While the first realization is the same one presented in Section 5.1, the second one is shown in Figure 6.17(a).
Table 6.12. Grid summary statistics of computed residuals and variances for three cases having different variance of $Y$; each case performed twice as in Cases 3 and 4 in Chapter 5.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>As in Case 3</th>
<th>$\sigma_Y^2 = 1$</th>
<th>$\sigma_Y^2 = 2$</th>
<th>$\sigma_Y^2 = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>Minimum</td>
<td>-2.7</td>
<td>-4.3</td>
<td>-5.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.5</td>
<td>4.3</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.11</td>
<td>0.22</td>
<td>0.081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.79</td>
<td>1.2</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_Y(x)$</td>
<td>Minimum</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.0</td>
<td>2.1</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.78</td>
<td>1.6</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.79</td>
<td>1.6</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>Minimum</td>
<td>-0.35</td>
<td>-0.88</td>
<td>-1.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.59</td>
<td>0.63</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.032</td>
<td>-0.093</td>
<td>0.022</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.15</td>
<td>0.26</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_h(x)$</td>
<td>Minimum</td>
<td>0.0014</td>
<td>0.0012</td>
<td>0.0014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.50</td>
<td>0.89</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.19</td>
<td>0.40</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.22</td>
<td>0.47</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{z1}(x) \rangle_c - q</em>{z1}(x)$</td>
<td>Minimum</td>
<td>-1.8</td>
<td>-3.1</td>
<td>-5.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.51</td>
<td>1.1</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.052</td>
<td>0.089</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.23</td>
<td>0.37</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{z2}(x) \rangle_c - q</em>{z2}(x)$</td>
<td>Minimum</td>
<td>-0.67</td>
<td>-2.4</td>
<td>-3.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.62</td>
<td>2.1</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.014</td>
<td>0.012</td>
<td>-0.043</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.12</td>
<td>0.21</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>$C^{[2]}_{q11}(0)$</td>
<td>Minimum</td>
<td>0.0017</td>
<td>0.0030</td>
<td>0.00055</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.38</td>
<td>0.75</td>
<td>3.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.078</td>
<td>0.18</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.088</td>
<td>0.20</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>$C^{[2]}<em>{q12}(0) = C^{[2]}</em>{q21}(0)$</td>
<td>Minimum</td>
<td>-0.17</td>
<td>-0.28</td>
<td>-0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.14</td>
<td>0.26</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.0000</td>
<td>0.0035</td>
<td>0.0068</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.011</td>
<td>0.021</td>
<td>0.063</td>
<td></td>
</tr>
<tr>
<td>$C^{[2]}_{q22}(0)$</td>
<td>Minimum</td>
<td>0.00030</td>
<td>0.00083</td>
<td>0.00095</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.16</td>
<td>0.36</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.019</td>
<td>0.040</td>
<td>0.090</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.024</td>
<td>0.049</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>Quantity</td>
<td>Statistic</td>
<td>$\sigma_Y^2 = 1$</td>
<td>$\sigma_Y^2 = 2$</td>
<td>$\sigma_Y^2 = 4$</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-----------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>Minimum</td>
<td>-3.4</td>
<td>-4.4</td>
<td>-6.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.5</td>
<td>4.5</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.15</td>
<td>0.11</td>
<td>-0.0061</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.88</td>
<td>1.2</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>$\sigma_Y^2(x)$</td>
<td>Minimum</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.0</td>
<td>2.1</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.74</td>
<td>1.6</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.76</td>
<td>1.6</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
<td>Minimum</td>
<td>-0.64</td>
<td>-1.3</td>
<td>-2.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.44</td>
<td>0.65</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.014</td>
<td>-0.33</td>
<td>-0.090</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.20</td>
<td>0.45</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>$\sigma_h^2(x)$</td>
<td>Minimum</td>
<td>0.0023</td>
<td>0.0031</td>
<td>0.0042</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.61</td>
<td>0.92</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.20</td>
<td>0.42</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.22</td>
<td>0.47</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>$\langle q_{z1}^{[2]}(x) \rangle_c - q_{x1}(x)$</td>
<td>Minimum</td>
<td>-1.9</td>
<td>-3.2</td>
<td>-5.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.45</td>
<td>0.85</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.050</td>
<td>0.047</td>
<td>-0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.26</td>
<td>0.37</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>$\langle q_{z2}^{[3]}(x) \rangle_c - q_{x2}(x)$</td>
<td>Minimum</td>
<td>-0.67</td>
<td>-2.5</td>
<td>-3.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.53</td>
<td>2.2</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.018</td>
<td>0.013</td>
<td>-0.063</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.12</td>
<td>0.21</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>$C_{q_{11}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.0010</td>
<td>0.0023</td>
<td>0.00036</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.31</td>
<td>0.48</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.046</td>
<td>0.15</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.054</td>
<td>0.16</td>
<td>0.36</td>
<td></td>
</tr>
<tr>
<td>$C_{q_{12}}^{[2]}(0) = C_{q_{21}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>-0.15</td>
<td>-0.25</td>
<td>-0.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.13</td>
<td>0.28</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.00043</td>
<td>0.0029</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.0066</td>
<td>0.016</td>
<td>0.044</td>
<td></td>
</tr>
<tr>
<td>$C_{q_{22}}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.00020</td>
<td>0.0012</td>
<td>0.00063</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.14</td>
<td>0.32</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.011</td>
<td>0.034</td>
<td>0.071</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.015</td>
<td>0.041</td>
<td>0.096</td>
<td></td>
</tr>
</tbody>
</table>
Table 6.13. Calibration statistics for two cases with different $Y$ integral scales.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>As in Case 3</th>
<th></th>
<th>As in Case 4</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho = 1$</td>
<td>$\rho = 4$</td>
<td>$\rho = 1$</td>
<td>$\rho = 4$</td>
</tr>
<tr>
<td>Correlation coefficient ($R_c$)</td>
<td>0.9995</td>
<td>0.9992</td>
<td>0.9981</td>
<td>0.9971</td>
</tr>
<tr>
<td>Common error variance ($\sigma^2_e$)</td>
<td>0.016</td>
<td>0.040</td>
<td>0.062</td>
<td>0.12</td>
</tr>
<tr>
<td>Condition number ($c_Q$) of $Q$</td>
<td>$1.4 \times 10^3$</td>
<td>$2.1 \times 10^3$</td>
<td>6.8</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Their associated realizations of $h(x)$ were computed on a standard finite element grid and are presented in Figures 5.2 and 6.17(b), respectively. Each pair of $Y(x)$ and $h(x)$ realizations were measured at locations shown in Figure 5.32. Then, two inversions, as in Cases 3 and 4 in Section 5.1, were performed for each set of measurements using parameters located at the 16 pilot points in Figure 5.32 and additionally conditioning on a pumping flux rate of one positioned at the center of the domain.

Table 6.13 presents performance measures of calibration for all inversions. One may note that the two inversions for $\rho = 1$ had a slightly better model fit than their corresponding counterparts for $\rho = 4$ had.

Residual and variance results of log conductivity, head, and flux components computed on the grid are summarized in Table 6.14. Residual statistics of log conductivity are comparable between cases having different $\rho$ but equal availability of priors; although, the cases with $\rho = 4$ yield significantly less biased mean values of $Y$ than their counterparts with $\rho = 1$ do. These observations are also applicable for longitudinal and transverse flux statistics. Head residual is the only quantity that presents better results for the cases with $\rho = 1$ than it does for those with $\rho = 4$. Its mean and RMS values are more biased and larger, respectively, for the latter cases than their
Figure 6.17. Reference (a) unconditional realization of $Y$ with (isotropic) $\rho = 4$ and $\sigma_{\gamma_u}^2 = 1$. (b) Associated reference head field considers boundary conditions and pumping well with $Q = 1$. 
Table 6.14. Grid summary statistics of computed residuals and variances for two cases with different $Y$ integral scales.

<table>
<thead>
<tr>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
</tr>
<tr>
<td>Statistic</td>
</tr>
<tr>
<td>$\rho = 1$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$\sigma^2_{Y}(x)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x) \rangle_c - h(x)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$\sigma^2_{K}(x)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{21}(x) \rangle_c - q</em>{21}(x)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$\langle q^{[2]}<em>{22}(x) \rangle_c - q</em>{22}(x)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$C^{[2]}_{q11}(0)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$C^{[2]}<em>{q12}(0) = C^{[2]}</em>{q21}(0)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>$C^{[2]}_{q22}(0)$</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>RMS</td>
</tr>
</tbody>
</table>
related values in the former cases. This is mainly due to two zones in $\Omega$ with large residuals ($> 0.5$); one zone is adjacent to the upper boundary between $x_1 = 5.2$ and $x_1 = 10.2$, and the other zone is adjacent to the lower boundary between $x_1 = 5.0$ and $x_1 = 7.2$. These zones are associated with a steep head gradient in the reference $h(x)$ field oriented along $x_1 \approx 7$ (Figure 6.17(b)) that is difficult to characterize. Note that comparing (hypothetical) cases with different $\rho$ entails using different reference $Y(x)$ and $h(x)$ fields; thus, comparison results should be taken with caution. The mean values for variances $\sigma_Y^2(x)$, $\sigma_h^2(x)$, $C_{q_{11}}^{[2]}(0)$, and $C_{q_{22}}^{[2]}(0)$ as well as the RMS value for zero-lag cross-covariance $C_{q_{12}}^{[2]}(0)$ for the two cases having $\rho = 4$ are between 36% and 61% of their counterparts in the cases with $\rho = 1$. This may be explained by larger number of pilot-point parameters per integral scale when $\rho = 4$, which should render a relatively better characterization of the $Y(x)$ field.

In general, it seems that increasing $\rho$ decreases uncertainty of log conductivity and state variables significantly.

### 6.7 Presence of a conditioning Pumping Well

To investigate the effect that the conditioning pumping well flux rate ($Q = 1$) used in all previous inversions has on parameter estimation I performed two inversions without it following the inversions of Cases 3 and 4 in Chapter 5, separately. Both inversions included log conductivity data, which, in principle, insures unique inverse solutions. The data considered here are the same described in those cases.
Table 6.15. Calibration statistics for two cases with and without the presence of a pumping well with $Q = 1$ at (9, 4).

<table>
<thead>
<tr>
<th>Statistic</th>
<th>As in Case-3 with well</th>
<th>As in Case-4 with well</th>
<th>As in Case-3 without well</th>
<th>As in Case-4 without well</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation coefficient ($R_c$)</td>
<td>0.9995</td>
<td>0.9995</td>
<td>0.9981</td>
<td>0.9981</td>
</tr>
<tr>
<td>Common error variance ($\sigma^2_e$)</td>
<td>0.016</td>
<td>0.015</td>
<td>0.062</td>
<td>0.066</td>
</tr>
<tr>
<td>Condition number ($c_Q$) of $Q$</td>
<td>$1.4 \times 10^3$</td>
<td>$1.8 \times 10^3$</td>
<td>6.8</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Listed in Table 6.15 are the calibration statistics corresponding to the two inversions without the well and those corresponding to the Cases 3 and 4 in Chapter 5. These results show that inverse model performance stays practically unchanged if the pumping well is “removed” from the field. The sets of parameter estimates (not shown), obtained separately, are comparable, which suggests that the two inversions are stable (i.e., they undergo small variation under small changes in data).

Table 6.16 presents residual and variance statistics for log conductivity, head, and components of flux computed on the grid. The values for log conductivity statistics are almost identical between the cases undergoing the same type of inversion with and without the well. Superimposing an exactly-measured convergent flow on the mean-uniform flow regime yields statistic values for head and flux that are a little more accurate and certain than when not doing so.

### 6.8 Comparative Analysis across Cases in Previous Sections

A few conclusions can be drawn by comparing results of the inverse cases in the sections of this chapter:
Table 6.16. Grid summary statistics of computed residuals and variances for two cases with and without the presence of a pumping well with $Q = 1$ at (9, 4).

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Statistic</th>
<th>As in Case-3 with well</th>
<th>Without well</th>
<th>As in Case-4 with well</th>
<th>Without well</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_K(x) - Y(x)$</td>
<td>Minimum</td>
<td>-2.7</td>
<td>-2.7</td>
<td>-3.4</td>
<td>-3.4</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.11</td>
<td>0.11</td>
<td>-0.15</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.79</td>
<td>0.79</td>
<td>0.88</td>
<td>0.88</td>
</tr>
<tr>
<td>$\sigma^2_Y(x)$</td>
<td>Minimum</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
<td>0.00000</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.78</td>
<td>0.78</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.79</td>
<td>0.79</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>$\langle h^{[2]}(x)_c - h(x) \rangle$</td>
<td>Minimum</td>
<td>-0.35</td>
<td>-0.32</td>
<td>-0.64</td>
<td>-0.57</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.59</td>
<td>0.51</td>
<td>0.44</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.032</td>
<td>0.028</td>
<td>0.014</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.15</td>
<td>0.14</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>$\sigma^2_h(x)$</td>
<td>Minimum</td>
<td>0.0014</td>
<td>0.0018</td>
<td>0.0023</td>
<td>0.0031</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.50</td>
<td>0.33</td>
<td>0.61</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.19</td>
<td>0.19</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.22</td>
<td>0.21</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>$\langle q^{[2]}_{x1}(x)<em>c - q</em>{x1}(x) \rangle$</td>
<td>Minimum</td>
<td>-1.8</td>
<td>-1.6</td>
<td>-1.9</td>
<td>-1.7</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.51</td>
<td>0.57</td>
<td>0.45</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.052</td>
<td>0.053</td>
<td>-0.050</td>
<td>-0.047</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.23</td>
<td>0.23</td>
<td>0.26</td>
<td>0.25</td>
</tr>
<tr>
<td>$\langle q^{[2]}_{x2}(x)<em>c - q</em>{x2}(x) \rangle$</td>
<td>Minimum</td>
<td>-0.67</td>
<td>-0.64</td>
<td>-0.67</td>
<td>-0.55</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.62</td>
<td>0.67</td>
<td>0.53</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.014</td>
<td>-0.015</td>
<td>-0.018</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>$C_{q11}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.0017</td>
<td>0.0020</td>
<td>0.0010</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.38</td>
<td>0.27</td>
<td>0.31</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.078</td>
<td>0.078</td>
<td>0.046</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.088</td>
<td>0.087</td>
<td>0.054</td>
<td>0.052</td>
</tr>
<tr>
<td>$C_{q12}^{[2]}(0) = C_{q21}^{[2]}(0)$</td>
<td>Minimum</td>
<td>-0.17</td>
<td>-0.037</td>
<td>-0.15</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.14</td>
<td>0.047</td>
<td>0.13</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>-0.00002</td>
<td>-0.00035</td>
<td>-0.00043</td>
<td>-0.00004</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.011</td>
<td>0.011</td>
<td>0.0066</td>
<td>0.0054</td>
</tr>
<tr>
<td>$C_{q22}^{[2]}(0)$</td>
<td>Minimum</td>
<td>0.00030</td>
<td>0.00041</td>
<td>0.00020</td>
<td>0.00029</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>0.16</td>
<td>0.13</td>
<td>0.14</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.019</td>
<td>0.019</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>RMS</td>
<td>0.024</td>
<td>0.023</td>
<td>0.015</td>
<td>0.014</td>
</tr>
</tbody>
</table>
1. Goodness of fit (as measured by $R_c$ since $\sigma^2_e$ values may not be compared between cases having different $N_{df}$ values) improves when: (a) conditioning on $Y$ data is done only geostatistically (e.g., Case 3 in Chapter 5), (b) $N_{df}$ gets small, or (c) $\sigma^2_{Y,u}$ or $\rho$ decreases.

2. Quality of parameter estimation (as measured by $c_Q$) decreases as $N_{df}$ gets smaller or (to a lesser degree) conductivity data become fewer or unavailable.

3. Accuracy of $Y$ residuals (as measured by their RMS value) improves as the number of conductivity data or pilot points increases, or as the level of heterogeneity decreases. $Y$ residuals become more biased as $P$ increases or head measurements are available in smaller numbers. Estimation uncertainty of $Y(x)$ is reduced as integral scale, number of head or log-conductivity measurements, or number of pilot points is increased.

4. Head residuals become biased as the number of measured heads gets small. Accuracy of head residuals increases, in the mean, when the number of pilot points or head measurements gets larger or when $\sigma^2_{Y,u}$ or $\rho$ decreases. Predictive uncertainty of head is reduced as the number of pilot points, number of head or $Y$ data, or integral scale is increased, or as $\sigma^2_{Y,u}$ is decreased.

5. The accuracy of longitudinal and transverse flux prediction decreases with increasing level of heterogeneity. Yet, the accuracy and unbiasedness of $q_{x1}(x)$ prediction seem to improve as the number of head or $Y$ data increases, or (to a
lesser degree) as \( P \) gets smaller.

6. Measures of predictive uncertainty of flux, as given by RMS values of \( C_{q11}^{[2]}(0) \), \( C_{q22}^{[2]}(0) \), and \( C_{q12}^{[2]}(0) \), increase with augmenting values of \( P \) or \( \sigma_{Y_i}^2 \), with decreasing values of \( \rho \), or with lack of \( Y \) data. Yet, whereas \( C_{q11}^{[2]}(0) \) gets larger when the number of head measurements increases, the opposite happens to \( C_{q22}^{[2]}(0) \) and \( C_{q12}^{[2]}(0) \).

7. The inverse case with only five head measurements, four pilot points, and no conductivity data compares more-or-less well, in the mean, with cases having more head data. The main disadvantage of the former is that its head prediction is relatively more biased, less accurate, and more uncertain than those of the latter.
Chapter 7
CONCLUSIONS

The main conclusions I can draw from my dissertation are the following:

1. It is possible and computationally feasible to condition nonlocal ensemble moment equations of steady state flow in bounded, randomly heterogeneous media on measurements of hydraulic head and conductivity through geostatistical inversion, utilizing pilot point parameters and universal kriging.

2. The deterministic form of the recursive conditional mean flow equations allows for the use of traditional inverse methods to condition such equations on measurements of hydraulic head and flux in addition to or instead of conductivity data. Using the methodology of Carrera and Neuman [1986a] coupled with universal kriging and pilot point parameterization has proven successful for this purpose. The approach yields estimates of both and an assessment of their estimation errors.

3. Second order moments of head and flux are subsequently conditioned by using the latter estimated quantities in their calculation. While first moments are computed iteratively, second moments are obtained in a single run.

4. Inversion yields the first two conditional moments of hydraulic conductivity, head, and flux. First moments constitute optimum unbiased predictors of these
quantities, and second moments are measures of their predictive uncertainty. Traditional nonstochastic inverse methods yield only parameter estimation errors, not predictive errors.

5. It is possible to condition individual Monte Carlo realizations of hydraulic conductivity on head and/or flux data separately. To obtain accurate conditional ensemble statistics one needs to compute a large number of such conditional realizations, which is very time consuming. My approach conditions these statistics (mean and variance-covariance) directly, without resorting to invert individual realizations.

6. A direct linear solver using symbolic multifrontal Choleski factorization for sparse symmetric matrices is at least one order of magnitude faster in computing the Green’s functions than is possible with iterative incomplete Choleski conjugate gradients. This is so because the direct method needs only one (common) factorization to obtain all Green’s functions during each forward model run.

7. When reliable conductivity measurements are available one may kriging them to obtain priors of unknown parameters at pilot points. In such a case the kriging covariance matrix becomes the matrix of prior errors in the calibration criterion. This takes into account the correlation among priors, due to kriging, in the parameter estimation process, something that is often neglected in standard
8. My parameter estimates and predictions of head and flux are given in terms of statistical mean values and are thus smoother than their random counterparts. However, using the estimation statistics of hydraulic parameters, it is possible to generate random conditional realizations of these parameters and then to obtain associated conditional realizations of head and flux, through solution of the stochastic flow equations.

9. Forward solution of localized equations is much faster than that of nonlocal equations. There is, however, no computational advantage in running localized inversions as a precursor of nonlocal inversion. Additional cases need to be run to further verify this conclusion.

10. Previous inverse methods using the pilot point approach have not included prior information in the objective function (e.g., RamaRao [1995], Gómez-Hernández [1997]). I show that considering reliable priors may lead to better posed problems and smaller parameter estimation errors. The priors enter twice: first into kriging, and next into the calibration criterion. The resulting parameter estimates are physically more reasonable and accurate; the predictive flux uncertainty is significantly reduced; but the mean residuals of log conductivity, head, and flux are relatively increased.

11. My examples show that whereas conditioning on conductivity or head data alone
may lead to a closer correspondence between these quantities and the model, conditioning on both yields improved parameter estimates and predictions of head and flux.

12. The examples show that $\sigma_Y^2(x)$, $\sigma_n^2(x)$, and flux variances are reduced in the neighborhood of $Y$ measurement points. The same effect is unnoticed in the vicinity of head measurement points. This is so because conditioning is achieved through kriging $Y$ values only at pilot and (possibly) measurement points.

13. Treating $Y$ measurements as priors of unknown parameters yields poorer results in terms of mean residuals of $Y$, head, and flux than not doing so.

14. Increasing the number of pilot points improves the quality of head prediction and parameter estimation at the cost of increasing uncertainty in flux significantly. This is caused by over-parameterization through the creation of zones of high mean log-hydraulic-conductivity value. The number of degrees of freedom should be then as small as possible while maintaining parameters reasonably uncorrelated.

15. Distributing pilot points regularly in the numerical grid works better than doing it randomly or placing them at head-measurement points.

16. Augmenting the number of either head or log-hydraulic-conductivity measurements improves the quality of both parameter estimates and prediction of state
variables.

17. My examples included domains with $\sigma^2$ ranging from 1 to 4. The performance of the inverse nonlocal model is found to be very satisfactory for these cases, though it is observed to decrease with increasing degree of heterogeneity.

18. The quality of head and flux predictions improves with increasing integral scale or correlation length.

19. The nonlocal inverse method may perform satisfactorily even in cases with few head (such as five) and no conductivity data. In such cases, head prediction may be relatively more biased, less accurate, and more uncertain than those of cases with more head data.

20. The quality of parameter estimation is seen to decrease when the number of degrees of freedom of the calibration gets smaller or when conductivity data become fewer or unavailable.

Future work related to my dissertation should include:

A. In my numerical examples I consider that all statistical parameters are known. Further examples assuming these parameters are (partially or totally) unknown need to be carried out to obtain (inverse) estimates of them. By comparing these estimates with corresponding true values or reliable priors one could evaluate the performance of my methodology in assessing such parameters. The estimates
would be conditioned on state variable data in addition to possibly conductivity measurements.

B. It would be worthy to automatically locate pilot points optimally (based on some statistical criterion) in my numerical inverse algorithm. This can be accomplished through an optimization procedure such as those related to observation network design (e.g., Carrera and Neuman [1986c], Chilès and Delfiner [1999]) or the one in the work of LaVenue and Pickens [1992].

C. My inverse method takes into account both hydraulic conductivity and state-variable measurement errors. However, in my examples I consider only exact data and am unable to investigate how these errors propagate both to parameter estimation and to the calculation of first and second moments of head and flux. It is desirable to consider such errors in future examples.

D. A logical step for continuing the work in this dissertation is to condition numerical transient groundwater-flow moment equations, such as those in the recent work by Ye [2002], jointly on state-variable and conductivity data using a transient inverse method such as the one of Carrera and Neuman [1986a].
Appendix A

DERIVATION OF UNIVERSAL POINT KRIGING EXPRESSIONS WITH CORRELATED MEASUREMENTS

The posterior universal kriging expressions in Chapter 3 are derived herein. Prior kriging expressions are not derived because they are a particular case of the posterior ones, in which $Q_{ij} = \delta_{ij} \cdot \text{Var}[\varepsilon_Y(x_i)]$.

In this appendix, expected or mean values are given in angle brackets and summations are taken from 1 to $I+P$ (number of real plus fictitious measurements) unless otherwise indicated.

A.1 Kriging Estimator

Recalling our statistical premises in Chapter 3, the $Y$ field is decomposed into deterministic and random components:

$$Y(x) \equiv m(x) + Y'(x)$$

(A.1)

and its (real and fictitious) measurements $Y_i$ are allowed to contain (correlated) errors:

$$Y_i \equiv Y(x_i) + \varepsilon_{Y_i}; \quad \langle \varepsilon_{Y_i} \rangle = 0; \quad (i = 1, 2, \ldots, I+P)$$

(A.2)

where $\varepsilon_{Y_i} \equiv \varepsilon_Y(x_i)$; drift is given by

$$m(x) = \sum_{r=1}^{R} a_r f_r(x)$$

(A.3)
Let the kriging estimator $Y_K(x)$ be linear in measurements (Equation (3.32)),

$$Y_K(x) \equiv \sum_i \omega_i(x)Y_i$$  \hspace{1cm} (A.4)

and unbiased,

$$\langle Y_K(x) \rangle \equiv \langle Y(x) \rangle$$  \hspace{1cm} (A.5)

Introducing Equations (A.1), (A.3), and (A.4) into Equation (A.5) yields

$$\sum_i \omega_i(x)m_i = \sum_i \omega_i(x) \left[ \sum_{r=1}^{R} a_rf_r(x_i) \right] = m(x)$$  \hspace{1cm} (A.6)

where $m_i \equiv m(x_i)$. By interchanging summations and using Equation (A.3) we obtain Equation (3.35),

$$\sum_i \omega_i(x)f_r(x_i) = f_r(x) \quad (r = 1, 2, \ldots, R)$$  \hspace{1cm} (A.7)

### A.2 Kriging Variance and Covariance

The (kriging) estimation error covariance is defined as:

$$\langle Y'_K(x)Y'_K(y) \rangle \equiv \langle [Y(x) - Y_K(x)][Y(y) - Y_K(y)] \rangle$$  \hspace{1cm} (A.8)

Upon equating Equations (A.1) and (A.2) at measurement points, we obtain

$$Y_i = m_i + Y'_i + \epsilon_{Y_i} \quad (i = 1, 2, \ldots, I+P)$$  \hspace{1cm} (A.9)
where \( Y_i' = Y'(x_i) \). Using Equations (A.1), (A.4), (A.6), (A.9) and rearranging transforms Equation (A.8) into

\[
\langle Y_K'(x)Y_K'(y) \rangle = \left\langle \left[ Y(x) - \sum_i \omega_i(x)Y_i \right] \left[ Y(y) - \sum_j \omega_j(y)Y_j \right] \right\rangle 
\]

\[
= \left\langle \left[ m(x) + Y'(x) - \sum_i \omega_i(x) (m_i + Y_i' + \varepsilon_{Y_i}) \right] \left[ m(y) + Y'(y) - \sum_j \omega_j(y) (m_j + Y_j' + \varepsilon_{Y_j}) \right] \right\rangle 
\]

\[
= \sum_i \sum_j \omega_i(x)\omega_j(y) \langle Y_i'Y_j' \rangle + \sum_i \sum_j \omega_i(x)\omega_j(y) \langle \varepsilon_{Y_i}\varepsilon_{Y_j} \rangle 
\]

\[
+ \langle Y'(x)Y'(y) \rangle + \sum_i \sum_j \omega_i(x)\omega_j(y) \left[ \langle Y_i'\varepsilon_{Y_j} \rangle + \langle Y_j'\varepsilon_{Y_i} \rangle \right] 
\]

\[
- \sum_i \omega_i(x) \langle Y_i'Y'(y) \rangle - \sum_j \omega_j(y) \langle Y_j'Y'(x) \rangle 
\]

\[
- \sum_i \omega_i(x) \langle \varepsilon_{Y_i}Y'(y) \rangle - \sum_j \omega_j(y) \langle \varepsilon_{Y_j}Y'(x) \rangle \quad \text{(A.10)}
\]

Considering Equations (3.24)–(3.26) allows simplifying the estimation error covariance in terms of covariances of \( Y \),

\[
\langle Y_K'(x)Y_K'(y) \rangle = \sum_i \sum_j \omega_i(x)\omega_j(y) \left[ C_Y(x_i - x_j) + Q_{ij} \right] + C_Y(x - y) 
\]

\[
- \sum_i \omega_i(x)C_Y(y - x_i) - \sum_j \omega_j(y)C_Y(x - x_j) \quad \text{(A.11)}
\]

or in terms of variograms (by substituting \( \sigma_{Y_{ij}}^2 - \gamma_Y(x) \) for \( C_Y(x) \)),

\[
\langle Y_K'(x)Y_K'(y) \rangle = - \sum_i \sum_j \omega_i(x)\omega_j(y) \left[ \gamma_Y(x_i - x_j) - Q_{ij} \right] - \gamma_Y(x - y) 
\]

\[
+ \sum_i \omega_i(x)\gamma_Y(y - x_i) + \sum_j \omega_j(y)\gamma_Y(x - x_j) \quad \text{(A.12)}
\]

Setting \( y = x \) in Equations (A.11) and (A.12) yields the estimation error variance.
in terms of covariances of $Y$

$$\left\langle [Y'_K(x)]^2 \right\rangle = \sum_i \sum_j \omega_i(x) \omega_j(x) [C_Y(x_i - x_j) + Q_{ij}]$$
$$+ \sigma^2_{yu} - 2 \sum_i \omega_i(x) C_Y(x - x_i)$$  \hspace{1cm} (A.13)

or variograms

$$\left\langle [Y'_K(x)]^2 \right\rangle = -\sum_i \sum_j \omega_i(x) \omega_j(x) [\gamma_Y(x_i - x_j) - Q_{ij}]$$
$$+ 2 \sum_i \omega_i(x) \gamma_Y(x - x_i)$$  \hspace{1cm} (A.14)

### A.3 Kriging Equations

To obtain the best (minimum error variance) estimator one minimizes the constrained

(via Equation A.7) criterion

$$\mathcal{L} \equiv \left\langle [Y'_K(x)]^2 \right\rangle - \frac{1}{2} \sum_{r=1}^{R} \mu_r(x) \left[ \sum_i \omega_i(x) f_r(x_i) - f_r(x) \right]$$  \hspace{1cm} (A.15)

using Lagrange multipliers $\mu_r$ and Equation (A.13). This yields the kriging Equations (3.33) and (3.34). The latter allow rewriting the variances and covariances in Equations (A.11)–(A.14) as Equations (3.36)–(3.39). Whereas the former are explicitly symmetric, the latter are simpler to evaluate numerically.


Carrera, J., and S. P. Neuman, Estimation of aquifer parameters under transient and


