

**STOCHASTIC APPROACH
TO STEADY STATE FLOW
IN NONUNIFORM GEOLOGIC MEDIA**

by
Shlomo Orr

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

1 9 9 3

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Shlomo Orr

entitled Stochastic Approach to Steady State Flow in Nonuniform Geologic Media

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

| | | |
|------------------------|------------------------|------------------------------|
| Dr. Shlomo P. Neuman | <u>Shlomo Neuman</u> | <u>June 23, 1993</u> Date |
| Dr. Thomas Maddock III | <u>Tom Maddock III</u> | <u>6/23/1993</u> Date |
| Dr. Jim T-C Yeh | <u>Siou-Chup Yeh</u> | <u>6/23/1993</u> Date |
| Dr. Art W. Warrick | <u>Art W. Warrick</u> | <u>23 June 1993</u> Date |
| Dr. Peter J. Wierrenge | <u>Peter Wierrenge</u> | <u>6/23/1993</u> Date |

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.

Dr. Shlomo P. Neuman Shlomo Neuman June 23, 1993
Dissertation Director Date

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: Shlomo P. Neuman

APPROVAL BY DISSERTATION DIRECTOR

This dissertation has been approved on the date shown below:

Shlomo P. Neuman June 23, 1993

Shlomo P. Neuman

Date

Professor of Hydrology

ACKNOWLEDGEMENTS

This work was generously supported by the John and Margaret Harshbarger Fellowship. It was also supported jointly by the U.S. Nuclear Regulatory Commission under contract NRC-04-90-51, and the U.S. Geological Survey, under Water Resources Research Grant 14-08-0001-G2092. I am grateful to my advisor, teacher, and friend, Dr. Shlomo P. Neuman, whose ingenious ideas are in the core of this dissertation; his thorough reading of the manuscript has enriched and elevated it. I am also grateful to the members of my committee, Drs. Art Warrick, Peter Wierenga, Thomas Maddock III, Jim Yeh, Mike Sully, and Rachid Ababou, who kindly accepted the high variability in my (statistically homogeneous) research, and assisted with excellent suggestions.

I am grateful to the wonderful people at the University of Arizona Research Support Group in the Computing Center (CCIT), particularly to Lucy Carruthers, whose constant help was invaluable, and to Mark Westergaard and Mark Borgstrom, for their kind assistance, whenever needed.

Thanks are due to Randolph E. Bank (the author of PLTMG) and to Booker Bense, who helped in the implementation of PLTMG on the Cray-Y MP at the University of California, San Diego. Thanks also to Jaime Gómez-Hernández for providing and advising with GCOSIM3D.

I thank the many friends who helped me at different stages of my research: Thomas Harter, Dong-Xiao Zhang, Oscar Levin, Gordon Wittmeyer, Evangelos Paleologos, Colleen Loeven, and many more. I am grateful to Yael Neuman for her generosity and hospitality.

Finally, I am indebted to my family: Galya, Ittai, and especially Meital, who endured the harshness of student and nomad life, during all these years.

To the memory of Arie Geva and David Segev-Frankel

TABLE OF CONTENTS

| | |
|---|-----------|
| LIST OF FIGURES | 10 |
| ABSTRACT | 18 |
| 1 INTRODUCTION | 20 |
| 1.1 Problem Definition | 25 |
| 1.2 Scope | 28 |
| 1.3 Solutions of Stochastic PDE's in Terms of Ensemble Moments | 29 |
| 1.4 The concept of homogenization | 32 |
| 1.4.1 Nonlocal Effects | 37 |
| 1.4.2 Spatial Homogenization | 39 |
| 1.5 Solutions of Steady State Stochastic PDE by Perturbation Method | 61 |
| 1.5.1 Dagan (1989) | 62 |
| 1.5.2 Shvidler (1962) and Matheron (1967) | 68 |
| 1.5.3 Gutjahr et al.(1978), Bakr et al.(1978), Gelhar and Axness (1983), Gelhar (1986) | 73 |
| 1.5.4 Nonuniform Mean Flow | 79 |
| 2 OPERATOR THEORY OF STOCHASTIC PDE'S | 95 |
| 2.1 Operator Representation | 95 |
| 2.2 Expansion in Neumann Series | 101 |
| 2.2.1 Contraction, Inverse Operators, and Convergence of the Neu- mann Series | 111 |

| | | |
|----------|---|------------|
| 2.3 | Applications of Integrodifferential Equations and Neumann Series . . | 115 |
| 2.3.1 | Stochastic Representation of Microflow and Derivation of Macroequations (Beran, 1968) | 116 |
| 2.3.2 | Kraichnan's Equation | 118 |
| 2.3.3 | King (1987, 1989) | 120 |
| 2.3.4 | Kröner (1977, 1986) | 122 |
| 2.3.5 | Markov (1987) and Shetzen (1980) | 124 |
| 2.4 | Unsteady Flow and Semigroups | 128 |
| 2.4.1 | Preliminaries: Time-dependent Formulation | 128 |
| 2.4.2 | Semigroups of Operators | 129 |
| 2.4.3 | Zeitoun and Braester (1991) | 134 |
| 2.4.4 | Conclusion | 144 |
| 3 | NONLOCAL THEORY, EFFECTIVE CONDUCTIVITIES AND WEAK APPROXIMATION | 146 |
| 3.1 | Problem Definition and Nonlocal Formalism | 146 |
| 3.2 | Effective Hydraulic Conductivity | 159 |
| 3.3 | Weak Approximation of Residual Flux | 166 |
| 3.4 | Effective Hydraulic Conductivity in 2-D And 3-D Lognormal Fields Under Uniform Mean Flow | 167 |
| 3.5 | Second Conditional Moments of Prediction Errors | 171 |
| 3.6 | Geostatistical Evaluation of Hydraulic Conductivity Moments | 177 |
| 4 | MONTE CARLO SIMULATIONS AND RANDOM FIELD GEN- ERATORS | 184 |

| | | |
|----------|--|------------|
| 4.1 | The LU Decomposition Method | 188 |
| 4.2 | Spectral Methods | 190 |
| 4.3 | Turning Bands Method | 193 |
| 4.4 | Sequential Simulations | 196 |
| 4.5 | Pseudo Random Number Generators | 199 |
| 4.6 | Comparison and Conclusions | 204 |
| 5 | HIGH RESOLUTION MONTE CARLO SIMULATIONS FOR RA- | |
| | DIAL FLOW | 252 |
| 5.1 | Multigrid Methods and PLTMG | 252 |
| 5.2 | Effective Hydraulic Conductivity and Ensemble Statistics Under Ra- | |
| | dially Converging Mean Flow | 266 |
| 5.3 | Verification of the Weak Approximation | 283 |
| 6 | CONCLUSIONS | 292 |

APPENDICES:

| | | |
|----------|---|------------|
| A | RANDOM FUNCTIONS | 304 |
| B | SPACES, NORMS, AND STRONG CONTINUITY | 308 |
| C | THE EIGENFUNCTION METHOD | 312 |
| D | CONVERGENCE AND ASYMPTOTIC APPROXIMATION | 316 |
| E | SYMMETRY OF RANDOM GREEN'S FUNCTION | 320 |
| F | RELATIONSHIPS BETWEEN COVARIANCES OF K AND Y | 322 |
| G | MATHERON'S PROOF RE 2-D LOGNORMAL FIELDS | 324 |
| H | SPSSX OUTPUT: ANALYSIS OF A RANDOM FIELD | 327 |
| | LIST OF REFERENCES | 332 |

LIST OF FIGURES

| | | |
|-----|---|-----|
| 1.1 | Permeability (millidarcy) and porosity space series from laboratory analysis of core samples from a bore hole in Mt. Simon sandstone aquifer in Illinois (Bakr, 1976). | 21 |
| 1.2 | Permeability of crystalline rocks and characteristic scale of measurements: Bars mark the maximum permeability range when several values are reported; stars represent single values. The shaded region is the author's personal view of the trend (Clauser, 1992). | 22 |
| 1.3 | Naff's Fig. 2: Effective hydraulic conductivity factor, $\rho\alpha(\rho)$; the dashed curves are based on extrapolation for large ρ | 87 |
| 1.4 | A comparison between (normalized) effective (point) conductivities as a function of dimensionless distance from the well bore, as predicted by Naff (1991) and calculated by Desbarats (1992b) from (numerical) MCS. | 93 |
| 3.1 | K_e/K_g versus σ_Y^2 for isotropic 3-D domain, after Neuman et al. (1992a). | 170 |
| 4.1 | Schematic representation of a discrete random field and the turning bands lines (adapted from Zimmerman and Wilson, 1990). | 195 |
| 4.2 | Serial plots in 2-D of 2000 nonoverlapping pairs of random numbers (After Sharp and Bays, 1992). | 202 |
| 4.3 | Serial plots in 3-D of 10000 nonoverlapping triplets of random numbers (After Sharp and Bays, 1992). | 203 |

| | | |
|------|--|-----|
| 4.4 | Spatial covariances of single realizations of 64×64 fields, with $\lambda = 5$, generated by GCOSIM3D (top) and TUBA1 (bottom). | 208 |
| 4.5 | Calculated and theoretical average spatial (isotropic) covariances and variograms in the X and Y directions of 50 realizations after filtering out outliers. | 209 |
| 4.6 | Calculated and theoretical average spatial (isotropic) covariances and variograms in the X and Y directions of 100 realizations after filtering out outliers. | 210 |
| 4.7 | Averages of (multidirectional) spatial covariances of 50 fields generated by the turning bands method (TUBA1) with 16 lines. | 212 |
| 4.8 | Average spatial covariances and variograms of 500 realizations (64×64) generated by GCOSIM, with $\lambda = 8$ | 213 |
| 4.9 | Average (multidirectional) spatial covariances and variograms of 500 realizations (64×64) generated by GCOSIM, with $\lambda = 8$ | 214 |
| 4.10 | Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the FFT method and an IMSL based RNG. | 215 |
| 4.11 | Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the FFT method and an RNG from the NAG library. | 217 |

- 4.12 Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the turning bands method with 32 lines and the portable RNG DPRAND. 218
- 4.13 Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the turning bands method with 320 lines and the portable RNG DPRAND. 219
- 4.14 Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with sequential simulator GCOSIM3D and (the portable) RAN2. . . . 220
- 4.15 Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with sequential simulator GCOSIM3D and (the portable) DPRAND. 221
- 4.16 Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 realizations generated with a defected turning bands method. . 222
- 4.17 Theoretical covariances between the mid (four) points and all the other points in the field. 224

- 4.18 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM), with an IMSL based RNG. 225
- 4.19 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM), with a NAG-based RNG. 226
- 4.20 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM), with the portable RNG DPRAND. 228
- 4.21 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the turning bands method (TUBA1) with 32 lines and the portable DPRAND. 229
- 4.22 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the turning bands method (TUBA1) with 320 lines and the portable DPRAND. 230
- 4.23 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the FFT method with the IMSL-based RNG. 231
- 4.24 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the FFT method with a NAG-based RNG. 232
- 4.25 “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the sequential simulator (GCOSIM3D) with DPRAND. 233

| | | |
|------|---|-----|
| 4.26 | “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the sequential simulator (GCOSIM3D) with RAN2. | 234 |
| 4.27 | “Ensemble” covariances of 1000 realizations generated by a defected RFG, based on the turning bands method. | 236 |
| 4.28 | Realizations generated by four different random field generators (Thomas Harter, pers. com.). | 237 |
| 4.29 | “Ensemble” means of 1000 realizations of normally distributed fields generated by four different RFG’s (Thomas Harter, pers. com.). . . . | 238 |
| 4.30 | “Ensemble” variances of 1000 simulations with the four (tested) RFG’s (Thomas Harter, pers. com.). | 239 |
| 4.31 | Spatial covariances of single realizations generated by the four different methods (Thomas Harter, pers. com.). | 240 |
| 4.32 | Spatial covariances of single realizations generated by the four different methods (Thomas Harter, pers. com.). | 241 |
| 4.33 | Spatial covariances of single realizations generated by the four different methods (Thomas Harter, pers. com.). | 242 |
| 4.34 | Averages of spatial covariances over 10 realizations generated by the four different methods (Thomas Harter, pers. com.). | 244 |
| 4.35 | Averages of spatial covariances over 50 realizations generated by the four different methods (Thomas Harter, pers. com.). | 245 |
| 4.36 | Deviations (from the theoretical covariance) of average spatial covariances over 10 realizations generated by the four different methods (Thomas Harter, pers. com.). | 246 |

| | | |
|------|---|-----|
| 4.36 | Deviations (from the theoretical covariance) of average spatial covariances over 10 realizations generated by the four different methods (Thomas Harter, pers. com.). | 246 |
| 4.37 | Deviations (from the theoretical covariance) of average spatial covariances over 50 realizations generated by the four different methods (Thomas Harter, pers. com.). | 247 |
| 4.38 | Mean deviation from the theoretical covariance of field portions as functions of the number of realizations (NMC). The fields were generated by LUSIM (LU-decomposition), with five different RNG's; "r", "e", "d", "i", and "n" stand for RAN2, ESSL, DPRAND, IMSL, and NAG, respectively (Thomas Harter, pers. com.). | 248 |
| 4.39 | Mean deviation from the theoretical covariance of field portions as functions of the number of realizations (NMC). The fields were generated with the portable DPRAND RNG, used by the four different methods [where "f", "l", "g", "t" stand for FFT, LU-decomposition, sequential simulations (GCOSIM), and <i>tbm</i> , respectively] (Thomas Harter, pers. com.). | 249 |
| 4.40 | Spatial average of "ensemble" mean log-conductivities over 920 windows on fields generated with the same four methods ["f", "l", "g", "t" stand for FFT, LU-decomposition, sequential simulations (GCOSIM), and turning bands methods, respectively], with DPRAND (Thomas Harter, pers. com.). | 250 |
| 5.1 | Flow chart for the FMG method (after Fletcher, 1988). | 256 |
| 5.2 | Flow chart for the multigrid V-cycle (after Fletcher, 1988). | 259 |

| | | |
|------|---|-----|
| 5.3 | Head response to a “delta sequence” as a delta function in a circular domain. | 263 |
| 5.4 | The fundamental solution in a circular domain. | 264 |
| 5.5 | Comparison between the fundamental solution (upper left), the simulated impulse response function (upper right), and the response to a unit sink (below) with an inner radius of 0.001 units. | 265 |
| 5.6 | Illustration of adaptive grid refinement and 2-D point sink solution for $K \equiv 1$ | 267 |
| 5.7 | Comparison of analytical and numerical 2-D point sink solutions for $K \equiv 1$ | 268 |
| 5.8 | 2-D point sink solution for $\sigma_Y^2 = 1$ and $\lambda = 2$ | 270 |
| 5.9 | 2-D point sink solution for $\sigma_Y^2 = 4$ and $\lambda = 5$ | 271 |
| 5.10 | Sample variance of \mathcal{G} versus number of simulations for $\sigma_Y^2 = 1$ and $\lambda = 2, 5, 8$ | 272 |
| 5.11 | Sample variance of \mathcal{G} versus number of simulations for $\sigma_Y^2 = 4$ and $\lambda = 2, 5, 8$ | 273 |
| 5.12 | Ratio between sample mean of \mathcal{G} and G_g versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 274 |
| 5.13 | Ratio between sample mean of \mathcal{G} and G_κ versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 275 |
| 5.14 | Sample variance of \mathcal{G} versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 276 |
| 5.15 | Sample variance of specific flux q versus radial distance r for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 277 |

| | | |
|------|--|-----|
| 5.16 | Sample variance of specific flux q versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 278 |
| 5.17 | K_e/K_g versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 280 |
| 5.18 | K_e/K_h versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$ | 281 |
| 5.19 | Spatial variation of ratio between sample means $\{K'(0)K'(\chi)\mathcal{G}'(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}(\chi, 0)\}$ for $\sigma_Y^2 = 1$ and $\lambda = 2$ | 284 |
| 5.20 | Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}(\chi, 0)\}$ versus number of simulations for $\sigma_Y^2 = 1$ and $\lambda = 2, 5, 8$ | 285 |
| 5.21 | Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}(\chi, 0)\}$ versus number of simulations for $\sigma_Y^2 = 4$ and $\lambda = 2, 5, 8$ | 286 |
| 5.22 | Spatial variation of ratio between sample means $\{K'(0)K'(\chi)\mathcal{G}'_r(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}_r(\chi, 0)\}$ for $\sigma_Y^2 = 1$ and $\lambda = 2$ | 288 |
| 5.23 | Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'_r(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}_r(\chi, 0)\}$ versus number of simulations for $\sigma_Y^2 = 1$ and $\lambda = 2, 5, 8$ | 289 |
| 5.24 | Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'_r(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}_r(\chi, 0)\}$ versus number of simulations for $\sigma_Y^2 = 4$ and $\lambda = 2, 5, 8$ | 290 |

ABSTRACT

This dissertation considers the effect of measuring randomly varying local hydraulic conductivities $K(\mathbf{x})$ on one's ability to predict steady state flow within a bounded domain, driven by random source and boundary functions. That is, the work concerns the prediction of local hydraulic head $h(\mathbf{x})$ and Darcy flux $\mathbf{q}(\mathbf{x})$ by means of their unbiased ensemble moments $\langle h(\mathbf{x}) \rangle_\kappa$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ conditioned on measurements of $K(\mathbf{x})$. These predictors satisfy a deterministic flow equation in which $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa = -\kappa(\mathbf{x})\nabla\langle h(\mathbf{x}) \rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x})$ where $\kappa(\mathbf{x})$ is a relatively smooth unbiased estimate of $K(\mathbf{x})$ and $\mathbf{r}_\kappa(\mathbf{x})$ is a "residual flux." A compact integral expression is derived for $\mathbf{r}_\kappa(\mathbf{x})$ which is rigorously valid for a broad class of $K(\mathbf{x})$ fields, including fractals. It demonstrates that $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ is nonlocal and non-Darcian so that an effective hydraulic conductivity does not generally exist. It is shown analytically that under uniform mean flow the effective conductivity may be a scalar, a symmetric or a nonsymmetric tensor, or a set of directional scalars which do not form a tensor. For cases where $\mathbf{r}_\kappa(\mathbf{x})$ can neither be expressed nor approximated by a local expression, a weak (integral) approximation (closure) is proposed, which appears to work well in media with pronounced heterogeneity and improves as the quantity and quality of $K(\mathbf{x})$ measurements increase. The nonlocal deterministic flow equation can be solved numerically by standard methods; the theory here shows clearly how the scale of grid discretization should relate to the scale, quantity and quality of available data. After providing explicit approximations for the prediction error moments of head and flux, some practical methods are discussed to compute $\kappa(\mathbf{x})$ from noisy measurements of $K(\mathbf{x})$ and to calculate required second moments of the associated

estimation errors when $K(\mathbf{x})$ is log normal. Nonuniform mean flow is studied by conducting high resolution Monte Carlo simulations of two dimensional seepage to a point sink in statistically homogeneous and isotropic log normal $K(\mathbf{x})$ fields. These reveal the existence of radial effective hydraulic conductivity which increases from the harmonic mean of $K(\mathbf{x})$ near interior and boundary sources to geometric mean far from such sources for σ_Y^2 (the variance of $\ln K$) at least as large as 4. They suggest the possibility of replacing $r_\kappa(\mathbf{x})$ by a local expression at distances of few conditional integral scales from the interior and boundary sources. Special attention is paid to the “art” of random field generation, and comparisons are made between four alternative methods with five different random number generators.

CHAPTER 1

INTRODUCTION

During the last two decades hydrologists have been forced to deal with “real world” heterogeneities inherent in nature. As groundwater contamination has become a major issue, and as more information and evidences on heterogeneities on different scales have been gathered, the effect of heterogeneity on the quality of predictions of flow and solute transport in the subsurface no longer could be ignored.

As noted by Dagan (1989),

Natural porous formations are heterogeneous, i.e., they display spatial variability of their geometric and hydraulic properties... this variability is of an irregular and complex nature. It generally defies a precise quantitative description, either because of insufficiency of information or because of lack of interest in knowing the very minute details of the structure and flow field.

Of particular interest is the strong variability of the hydraulic conductivity, $K(\mathbf{x})$, in space. Analyses of core samples (e.g., Bakr, 1976, and Figure 1.1) and packer tests in fractured rocks (e.g., Neuman and Depner, 1987; Clauser, 1992, and Figure 1.2) have indicated variability of $K(\mathbf{x})$ over several orders of magnitude, while porosity values are much less variable, as illustrated in Figure 1.1.

Variances of $\ln K(\mathbf{x})$ are generally between 1 and 2, but can reach 10 in some cases, particularly in fractured rocks (de Marsily, 1986, p. 80; see also Neuman and

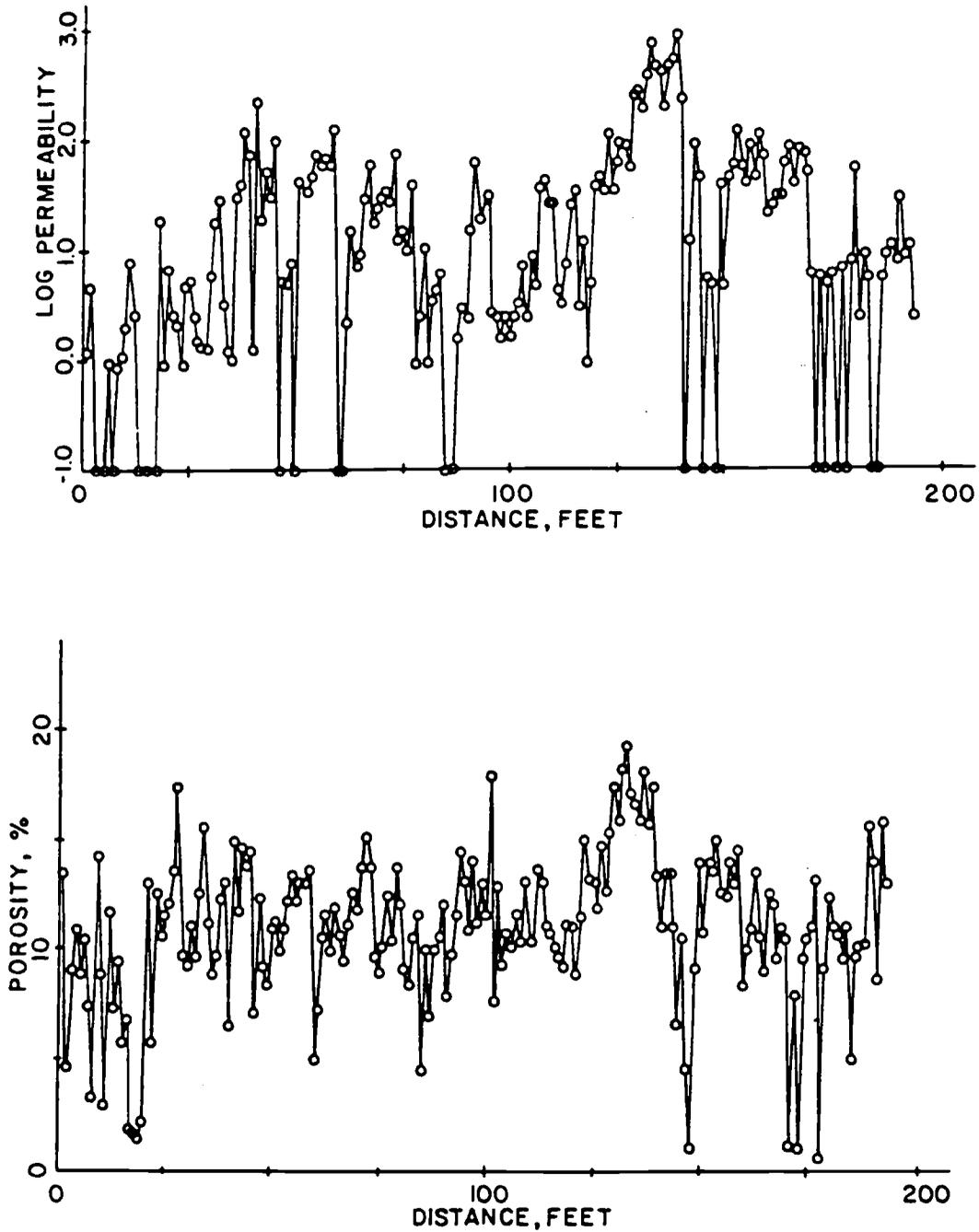


Figure 1.1: Permeability (millidarcy) and porosity space series from laboratory analysis of core samples from a bore hole in Mt. Simon sandstone aquifer in Illinois (Bakr, 1976).

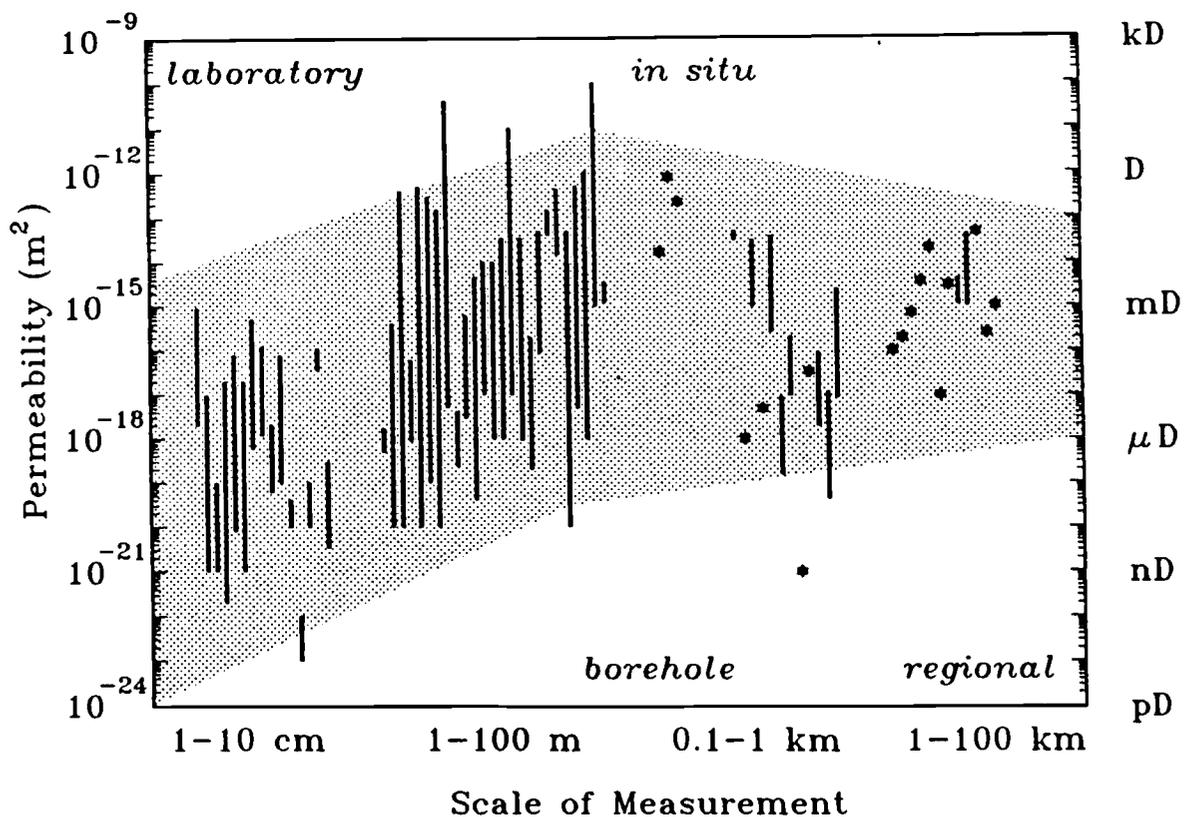


Figure 1.2: Permeability of crystalline rocks and characteristic scale of measurements: Bars mark the maximum permeability range when several values are reported; stars represent single values. The shaded region is the author's personal view of the trend (Clauser, 1992).

Depner, 1988). The distributions of hydraulic conductivities in different geologic media are found to be approximately log-normal, while the distribution of porosities is usually regarded to be normal (see Neuman, 1982, and de Marsily, 1986, for lists of references and discussions on that subject).

The variations in $K(\mathbf{x})$ cause variations in velocities, which, in turn, enhance solute dispersion in the subsurface. Further, measurements are practically limited to selected locations and are frequently done on disparate scales (as demonstrated in Figure 1.2), ranging from core samples to packer tests and pumping tests, assuming local homogeneity. The task of quantitatively relating measurements and properties on different scales is a difficult and intriguing one (see, for example, Cushman, 1986, 1990; Rubin and Gomez-Hernandez, 1990; Neuman, 1990; Follin, 1992). Although the upscaling issue is beyond the scope of this work, it will be partially discussed in the context of homogenization in the following introductory review.

The recognition that the hydraulic properties of porous media can be described in statistical terms, and the lack of information about $K(\mathbf{x})$ between measurements, has led many hydrologists to view $K(\mathbf{x})$ as a (correlated) random field. Summarizing, the “randomness” of the hydraulic conductivity field is due to (1) spatial variability of $K(\mathbf{x})$, (2) uncertainty regarding the $K(\mathbf{x})$ field, and (3) measurement and interpretive, or *model* errors, stemming from the estimation of $K(\mathbf{x})$ from local measurements of flux and heads.

The spatial variability of $K(\mathbf{x})$ and the uncertainty have also led hydrologists to search for effective (or equivalent) hydraulic conductivities, which would yield the average behavior of the system under different flow regimes, in a spatial and/or an ensemble sense. The attempts to determine these fictitious parameters are central

to both homogenization and stochastic theories (including this work), and will be discussed in the following review. In this dissertation we adopt the statistical (or, more precisely, the stochastic) viewpoint. Due to its generality and strength, the statistical approach has become a major tool for predicting flow and transport in heterogenous formations by an increasing number of hydrologists (e.g., Shvidler, 1962, Matheron, 1973, Freeze, 1975, Dagan, 1976, 1989, 1986, Neuman, 1982, 1984, 1990, de Marsily, 1986, Gelhar, 1976, Gelhar and Axness, 1983, Gutjahr et al., 1978, Hoeksema and Kitanidis, 1985, Russo and Bresler, 1980, Ababou, 1988, Ababou and Wood, 1990, Yeh et al., 1985, Naff and Veccia, 1986, Naff, 1991, Rubin and Dagan, 1988, 1989, Rubin and Gómez-Hernández, 1990, Desbarats, 1992a, 1992b, and many others). For example, in his book, Dagan (1989) explains:

... The mathematical framework used to describe heterogeneity and transition to average variables which is adopted here is the *statistical* one... (Consequently) properties of flow and transport variables are represented with the aid of *random* space functions (RSF; see Appendix A for a definition) and the actual porous formation (and processes) are regarded as a realization of the ensemble of the RSF which describe them.

The statistical framework has been adopted by scientists and engineers in virtually all fields. For example, R.E. Bellman (1963) stated that

The necessity of a theory of stochastic process for the description of physical phenomena in conceptually, analytically, and computationally tractable form is by now understood and accepted, and the steady trend in theories from linear deterministic to non linear and stochastic must be regarded as the primary goal of future research efforts.

A major disadvantage of the statistical approach, however, is the need in many data in order to obtain meaningful solutions (or estimates). A typical geostatistical analysis requires a sample of at least 50 data points (rigorously, at least 1000 points; D. Myers and S.P. Neuman, pers. com.). In practice, the ergodic hypothesis¹ has to be invoked, ensemble distributions and a (spatial) correlation structure are inferred, and a variogram model (which fits the correlation structure) is chosen. Consequently, there is a potential for errors at each stage (fortunately, some of these errors can be estimated by the statistical approach itself). As it will be demonstrated throughout the dissertation, existing deterministic “solutions” of stochastic flow and transport problems are, in fact, approximations, restricted to relatively simple forms of statistical structures and flow regimes, and suffer from some serious difficulties. The theory presented in this work overcomes some of the major limitations (above); it can (potentially) be extended to more complex flow regimes (e.g., unsteady flow), and has some direct implications to transport (as demonstrated by Neuman, 1993).

1.1 Problem Definition

We consider steady state flow of water under isothermal conditions. The governing balance equation on a local scale ω is

$$\nabla \cdot K(\mathbf{x})\nabla h(\mathbf{x}) + f(\mathbf{x}) = 0 \quad (1.1)$$

subject to the boundary conditions (BC)

$$h(\mathbf{x}) = H(\mathbf{x}) \quad \mathbf{x} \in \Gamma_D \quad (1.2)$$

¹In essence, the *ergodic* hypothesis suggests equivalence between ensemble and spatial moments.

and

$$-\mathbf{q}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = Q(\mathbf{x}) \quad \mathbf{x} \in \Gamma_N \quad (1.3)$$

where

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}) \quad (1.4)$$

is the (Darcian) flux, $f(\mathbf{x})$ is a randomly prescribed source function, $H(\mathbf{x})$ is a randomly prescribed head on Dirichlet boundary segments Γ_D , $Q(\mathbf{x})$ is a randomly prescribed flux into the flow domain, Ω , across Neumann boundary segments Γ_N , $\mathbf{n}(\mathbf{x})$ is a unit outward normal to the boundary Γ_N , and Γ is the union of Γ_D and Γ_N . h , the hydraulic head, is the “solution” (or response, or “output” of the system). All quantities (\mathbf{q} , h , f , K) are defined on a local scale ω , on which Darcy’s law is taken to apply. For simplicity, we assume that the boundary conditions $H(\mathbf{x})$ and $Q(\mathbf{x})$ are statistically independent. (Others have looked at variations of this same problem).

Because the hydraulic conductivity $K(\mathbf{x})$ is random (and/or the boundary conditions are random and/or the source is random), the response $h(\mathbf{x})$ or \mathbf{q} is also random, and the equation is classified as a *stochastic partial differential equation* (stochastic PDE), or a *stochastic system*, or simply a *random equation*, which has the form of random Poisson equation (some definitions and properties of random functions are discussed in Appendix A). Our goal, in this dissertation, is to contribute towards the solution of this stochastic PDE.

For the one dimensional steady-state case there is an exact deterministic solution for (1.1) subject to homogeneous Dirichlet and Neumann boundary condition on each end respectively, with a concentrated (point) source $f(x) = \delta(x - \xi)$ (i.e., a

unit source at $x = \xi$)², which is given by (e.g., Stakgold, 1979, p. 70):

$$h(x, \xi) = \int_0^x \frac{1}{K(y)} dy \quad 0 \leq x < \xi \quad (1.5)$$

$$h(x, \xi) = \int_x^1 \frac{1}{K(y)} dy \quad \xi < x \leq 1 \quad (1.6)$$

[More accurately, in (1.5) the boundary conditions are defined as $h(0) = 0$ (at $x = 0$) and $dh/dx = 0$ at $x = 1$, hence, $h(x, \xi) = \text{constant}$ for $\xi < x \leq 1$; in (1.6) $h(1) = 0$ (at $x = 1$) and $dh/dx = 0$ at $x = 0$, hence, $h(x, \xi) = \text{constant}$ for $0 < x \leq \xi$.] In fact, the solution $h(x, \xi)$ above is the *Green's function*, $G(x, \xi)$, of the one dimensional version of (1.1) subject to (general) Dirichlet and Neumann boundary condition on each end respectively.] Needless to say, no such exact solutions generally exist for more complex multidimensional problems associated with (1.1). Notice that the above solution relates the (deterministic) response, $h(x)$, to the parameter $K(x)$, assuming that $K(x)$ is known at each point (or, alternatively, that $K(x)$ is a known function of x); when $K(x)$ is random, (1.1) becomes a stochastic PDE, and we seek to relate the (joint) statistics of the response, $h(x)$ to that of $K(x)$ at each point x .

Notice also, that while (1.1) is a linear partial differential equation (PDE), the solution $h(\mathbf{x})$ in (1.5) – (1.6) depends *nonlinearly* on the parameters. Indeed, Frisch (1968) noted that

The principal difficulty of the problem (of stochastic partial differential equations) stems from the fact, first observed by Kraichnan (1962), that in spite of its apparent linearity, it is nonlinear in stochastic quantities, because the solution of a linear equation depends nonlinearly upon the

²The Dirac delta function is defined such that $\delta(\mathbf{x}, \mathbf{x}') = 0$ if $\mathbf{x} \neq \mathbf{x}'$, and $\int_{-\infty}^{\infty} h(\mathbf{x})\delta(\mathbf{x} - \mathbf{x}')d\mathbf{x} = h(\mathbf{x}')$.

coefficients. This gives rise to very serious difficulties, which are of a mathematical nature and are shared in common not only by the problem of turbulence and the many-body problem. Typical among these difficulties are the divergence of perturbation expansions, the appearance of secular terms³, and the fact that even the lowest order moments of the solution depend on the complete set of all the moments of the coefficients.

1.2 Scope

In this chapter we first discuss the meaning of a solution to a stochastic PDE in terms of ensemble moments, and list the major analytical and numerical methods of solutions. Next, we discuss the concept of homogenization, and contrast upscaling with the concept of effective hydraulic conductivity (in ensemble sense) as defined by existing stochastic theories. We will exemplify the latter by presenting methods to derive effective conductivities via perturbation solutions of the above PDE.

In Chapter 2 we present an operational approach to solve stochastic PDE's, which leads to integrodifferential presentations of the solutions. We discuss the properties of such solution methods and demonstrate their applications to both steady-state and transient flow problems. The transport problem will be introduced briefly.

In Chapter 3 we present a new integro-differential theory which is the heart and main contribution of this dissertation. It leads to a better definition of effective conductivity in bounded domains, accounts for information content via conditioning, and leads to a weak approximation analogous to that of field theoretical diagrammatic

³Secular terms appear in time-dependent (transient) problems. These are expressions proportional to the positive power of time, t , which increase indefinitely as $t \rightarrow \infty$; to be discussed later on.

approach in physics, but more general (in that boundaries are explicitly accounted for, and so is information). The weak approximation is invoked in order to make the solution tractable. We also provide explicit approximations for the moments of prediction errors of head and flux.

In Chapter 4 we discuss solution of the stochastic PDE via Monte Carlo simulations (MCS). Four methods for generating (unconditional, correlated) random fields for Monte Carlo Simulations are reviewed and compared. A variety of tests for random field generators are demonstrated. Special attention is paid to (pseudo) random number generators. In Chapter 5 we briefly review the basics of the multigrid methods, and use Monte Carlo simulations (with the powerful solver PLTMG) to investigate aspects of the theory introduced in Chapter 3. In particular, we check the above weak approximation by MCS for variances of $Y = \ln K$ as large as $\sigma_Y^2 = 4$. Effective conductivities under radial flow are investigated and compared with existing theories. Ensemble means and variances of heads and fluxes are also calculated and compared with known analytical results.

Chapter 6 lists conclusions derived from this work, and makes recommendations for future research.

1.3 Solutions of Stochastic PDE's in Terms of Ensemble Moments

A solution to a stochastic PDE consists of specifying the (joint) probability density function (*pdf*) of the response, $h(\mathbf{x})$, given those of $K(\mathbf{x})$, $f(\mathbf{x})$, $H(\mathbf{x})$, and $Q(\mathbf{x})$.

Unfortunately, one cannot obtain the joint cumulative distribution function (CDF) of the random response at *all* (infinite number of) points. Even for a finite set

of points, one cannot obtain closed form equations for a finite number of moments. This is the so called *closure problem*, which can be circumvented by either *closure approximations* (e.g., perturbation methods) or by numerical approximations, i.e., by Monte Carlo simulations (MCS).

Often, the most relevant information is contained in the first two ensemble moments⁴ of $h(x)$, that is, the ensemble mean (or the expected value), the variance (a measure of the uncertainty), and the covariance (or the variogram, a measure of the underlying correlation structure and scale, or, alternatively, the spectral density).

In particular, when the random function is multivariate normal (or lognormal), the second order representation is exhaustive, and may serve for deriving any other higher statistical moments. (This is one of the reasons for dealing with $Y = \ln K$ instead of (directly) with K in the stochastic hydrology literature.)

With respect to the first moment, i.e., the expected value of the response, obviously, having a deterministic coefficient which can be used in the governing equation instead of the random one (and thereby converts the stochastic PDE into a deterministic one) is of great merit. This coefficient is, then, an *effective* parameter. The existence of such an effective coefficient will become clear in Chapter 3.

The most widely used analytical method for analyzing random systems is the (small) perturbation technique, which consists of expanding random quantities around their mean values in Taylor series. The popularity of the perturbation approach stems from its relative simplicity and power. Because of its widespread use in hydrology and other fields, we give three illustrations of the perturbation method in Section 1.5.

⁴Termed *statistical measures* by Adomian (1983).

Another class of analytical methods “invert” (or transform) the stochastic PDE into random integral or integro-differential equations, mostly in the form of Neumann or Volterra series. These methods are discussed in Chapters 2 and 3.

As it will be shown in the review that follows, closure approximations are commonly limited to some form of stationarity of the hydraulic conductivity field, subject to uniform mean flow conditions. All known approximations are limited to mild heterogeneity (or small perturbations, where $\sigma_Y^2 \ll 1$). Cases with intricate geometries and boundary conditions, with high variability, and various types of excitations are usually impossible to handle with such methods. [Rubin and Dagan (1988) used small perturbations to approximate (to first order) the first two moments of heads in a two-dimensional, semi-infinite domain of stationary log-conductivity field, with a constant head boundary normal to the mean flow (steady state). Naff and Veccia (1986) approximated, to first order, the influence of impervious boundaries upon head variance in a stationary log-conductivity field under mean uniform, steady state flow in a domain of infinite horizontal extent, bounded by two (infinite) impervious horizontal planes. In both cases, the geometry, boundary conditions, and flow regimes are relatively simple; despite the existence of boundaries, in both cases the domains are partly unbounded.] These limitations are the main motivation for this dissertation. In particular, the method suggested in Chapter 3 is less restricted and more general than the other analytical methods discussed here.

As it will be seen in the following review, in almost every solution method of the stochastic flow equation, including the method proposed in Chapter 3, the determination of effective conductivity plays a major role. Therefore, we devote the following section to aspects of effective conductivity, in both the ensemble and the

spatial sense. We then review some relevant examples of the latter, and highlight differences and similarities between the ensemble and spatial concepts. A special section is devoted to solutions of the steady-state stochastic flow equation by perturbation methods.

1.4 The concept of homogenization

The problem of defining effective conductivity of random media has been subject to extensive research in different fields of physics and engineering. The (general) idea of effective hydraulic conductivity is, “on the average”, to reproduce the flow in terms of fluxes and/or energy dissipation (or, alternatively, pressure, head, temperature, or other dependent variables, depending on the system of interest).

In the previous section we defined an effective property as a deterministic coefficient which may vary in space, and which, when used in the governing equation, yields the expected value (or ensemble mean) of the response at each point. In the case of Eq. 1.1, this means that an effective hydraulic conductivity would, ideally, be a tensor, $\mathbf{K}_e(\mathbf{x})$, such that

$$\nabla \cdot \mathbf{K}_e(\mathbf{x}) \nabla \langle h(\mathbf{x}) \rangle + \langle f \rangle = 0, \quad (1.7)$$

where the response is the head at each point, $h(\mathbf{x})$. Similarly, the expected value of the flux at each point would be

$$\langle \mathbf{q}(\mathbf{x}) \rangle = -\mathbf{K}_e(\mathbf{x}) \nabla \langle h(\mathbf{x}) \rangle. \quad (1.8)$$

As we will see, such a tensor does not always exist. When it does exist, the effective hydraulic conductivity enables us to transform the random equations

into workable deterministic ones, which produce deterministic estimates of heads and fluxes. In other words, the effective parameter replaces the uncertain one, to produce the average behavior of the system at each point in the field.

When the above averaging is done spatially, rather than in the ensemble, it is termed *homogenization*, or *upscaling* or, simply, (block) averaging. In this case, by introducing an upscaled property, one replaces a heterogeneous material by an equivalent homogeneous one, to produce the average behavior of the field⁵.

Upscaling (or spatial homogenization) is used for finding the relationships between small scale “measurements” and large scale “measurements”, i.e., for “mapping” properties from one scale to another (e.g., based on their statistics), and for predicting large scale response (and perhaps, the uncertainty associated with it) from small scale measurements⁶.

While in the *ensemble* sense, the variances of the (random) property and the response are reflections of our *uncertainty*⁷, in a spatial sense, these variances reflect the extent of the heterogeneity.

\mathbf{K}_e can also be defined in terms of *energy* dissipation. The dissipation function E , defined as energy per unit weight of fluid dissipated by friction is (e.g., Dagan, 1989, p. 188):

$$E = -\mathbf{q}\nabla h \quad (1.9)$$

with residuals defined by $h' = h - \langle h \rangle$ and $\mathbf{q}' = \mathbf{q} - \langle \mathbf{q} \rangle$. For stationary \mathbf{q} and h' and

⁵Note the similarity between this and the definition in the previous paragraph.

⁶Very often, we cannot measure the properties directly; we only *infer* or *estimate* them from measurements.

⁷See, for example, Ababou and Gelhar (1990).

uniform mean flow, we have (Dagan, 1989)

$$\langle E \rangle = \langle K \nabla h \cdot \nabla h \rangle = \mathbf{J}^T \mathbf{K}_e \mathbf{J} = \langle \mathbf{q} \rangle^T \mathbf{K}_e^{-1} \langle \mathbf{q} \rangle \quad (1.10)$$

where $\mathbf{J} \equiv \nabla \langle h \rangle$, and \mathbf{J}^T and \mathbf{q}^T are the transpositions of \mathbf{J} and \mathbf{q} respectively. Based on the first law of thermodynamics, $\langle E \rangle \geq 0$; hence, \mathbf{K}_e in (1.10) is positive definite and invertible. In fact, it is also symmetric (cf. Dagan, 1989, p. 97).

According to Kröner (1977), the two definitions of effective conductivity, (1.8) and (1.10), “are mutually consistent, which means that the “Hill condition” $\langle \mathbf{q} \rangle \langle \nabla h \rangle = \langle \mathbf{q} \nabla h \rangle$ applies” [as already shown in (1.10)].

Following Batchelor (1974),

$$\langle E \rangle = \langle K \rangle \mathbf{J} \cdot \mathbf{J} - \langle K \nabla h' \cdot \nabla h' \rangle \quad (1.11)$$

which, when combined with Eq. 1.10, and examined in the directions of the principal coordinates (or eigenvectors), reveals the upper bound

$$K_{e_i} = \langle K \rangle - \langle K \left(\frac{\partial h'}{\partial x_i} \right) \rangle / J_i^2 \leq \langle K \rangle \equiv K_A \quad (1.12)$$

where K_A is the arithmetic mean, and the lower bound

$$K_{e_i}^{-1} = \langle K^{-1} \rangle - \langle \frac{\mathbf{q}' \cdot \mathbf{q}'}{K} \rangle / \langle q_i \rangle^2 \leq \langle K^{-1} \rangle \equiv K_H^{-1} \quad (1.13)$$

which means that $K_{e_i} \geq K_H$, where K_H is the harmonic mean. Thus $K_H \leq K_{e_i} \leq K_A$ for all i 's. If K is log normally distributed,

$$\exp(-\sigma_Y^2/2) \leq K_{e_i}/K_G \leq \exp(\sigma_Y^2/2) \quad (1.14)$$

where $Y = \ln K$, and $K_G = \exp(\langle Y \rangle)$ is the geometric mean.

These upper and lower bounds are readily obtained for flow parallel and normal to perfectly stratified layers, and were already found by Matheron (1967) who

concluded that for uniform mean flow (i.e., on the average, parallel flow lines), the effective permeability always ranges between the harmonic and the arithmetic mean of the local permeabilities, for any spatial correlation (of permeability) and for any number of dimensions (de-Marsily, 1986). In several other works (discussed below) narrower bounds for effective conductivities have been obtained (e.g., Hashin and Shtrikman, 1962, for multiphase medium). However, as pointed out by Dagan (1989), they do not provide an improvement over the above bounds for a continuous distribution of K between zero and infinity. The disadvantage of the arithmetic and harmonic means as upper and lower bounds is that for large variances (σ_Y^2), they are widely separated, and therefore, of limited use in such cases.

It should be noted that despite the similarity in definitions, there is a substantial difference between ensemble and spatial effective properties. In addition to the necessary conditions for the existence of (point) effective conductivities (discussed in Chapter 3), any upscaling method requires the existence of at least two distinct disparate scales of heterogeneity, such that the “large scale” contains enough “small scales” of heterogeneity (as demonstrated in the following review). However, as the (stationary) random field of interest becomes “infinitely” large (or, in a less restrictive sense, when the boundary effects diminish to zero), the ensemble and spatial averages become equivalent, through *ergodicity*; then, under mean uniform flow conditions, the effective property can be equated with the upscaled property.

Auriault correctly points out that whatever method employed, phenomenological investigation, or homogenization procedure, the basic assumption underlying the methods is that both the medium and the excitation are homogenizable, i.e., there exists an equivalent (large scale) description. When this basic assumption is not

justified, the average parameters obtained are meaningless in that they are limited to the particular flow regime (i.e., to the particular boundary values).

Similarly, Beran (1968, p. 126) noted that “for effective constants (or parameters) to be meaningful for a wide range of problems, we assume that both the joint statistics of the constants and the imposed fields are homogeneous”. In other words, the random fields have to be statistically homogeneous (or stationary), and the mean gradient has to be uniform. This is why the source term f in (1.1) and in (1.7) is usually dropped in (spatial) homogenization theories. These are rather restrictive conditions. Nevertheless, in the following sections and in Chapter 5, we will show that, with some caution, those conditions can be relaxed in practice, and effective conductivities can be defined for large portions of the flow fields under more general conditions (e.g., for non-homogeneous conductivity fields and nonuniform mean flow fields).

Because most homogenization methods assume ergodicity, there is frequently no clear distinction between ensemble and spatial effective conductivity. This ambiguity occurs in many of the works we have reviewed; some authors refer to homogenization in a spatial sense, while some other authors refer to it in an ensemble sense, and yet some other investigators interpret it in both spatial and ensemble sense. For this reason, we extend our review to include spatial homogenization methods.

For example, Auriault (1991) uses “homogenization” as a generic term, and divides it into four different schemes:

- Homogenization for fine periodic structures
- Statistical modeling

- The self consistent method
- Methods that use the (spatial) average theorems

where the “statistical modeling” of Auriault is equivalent to our “ensemble” approach. Unlike Auriault (1991) and as noted above, we discern between two major approaches to effective properties: ensemble and spatial. We, however, adopt (the remainder of) Auriault’s classification for homogenization methods which we review in the sequel. But first, we introduce a more rigorous definition of homogenization, and some important implications, which will be discussed in more detail, later, in the following chapters.

1.4.1 Nonlocal Effects

Cushman (1990) defines homogenization as the development of functional forms for constitutive variables at the scale of interest. Constitutive variables are parameters that “show our ignorance of, or lack of detailed information on the many degrees of freedom of relevance to lower scales within the system”, and usually result from the solution to an inverse problem (*ibid*). If the constitutive variable depends on what is happening at a “point” (in space-time) or a very small neighborhood of the point, the variable is said to be *local*, and derived from a *local theory* (*ibid*). Of course, such constitutive variables may still vary in space and time. On the other hand, if the constitutive variable depends on information from regions distinct from the neighborhood (of space-time) point where evaluation of the variable is to be made, then the theories and constitutive variables are said to be *nonlocal*. According to Cushman (1990), “nonlocal theories may result in constitutive variables that have two or more additional degrees of freedom (or parameters), as they usually appear in

the form of *functionals*.” (e.g., integrals). As will be shown in the next section and in Chapters 2 and 3, random integral equations have a central role in homogenization theories⁸. In particular, convolution integrals may result in a nonlocal mean flux, which refutes the existence effective conductivity (depending on mean flow conditions, as discussed in Chapter 3).

Tartar (1989, 1986) is especially aware of nonlocal effects induced by homogenization and integral equations. He explains why such effects are expected:

We were indeed looking for a convolution equation because our problem was invariant under translation in x or t , and we know that linear operators commuting with translations have to be given by convolutions.

Tartar (1989) adds that his goal is to find “what kind of memory kernel can appear by homogenization”, and that this is important “because memory effects do occur in models of continuum mechanics”.

Neuman (1993) introduces nonlocality of solute flux as a major outcome of an Eulerian-Lagrangian transport theory, and then compares his results with works done in the area of turbulent diffusion by Kraichnan (1959) and Roberts (1961). Non-locality of eddy diffusivity and of effective diffusion equations is also discussed by Avellanda and Majda (1990), and by Kraichnan (1987) in conjunction with turbulent diffusion (which is beyond the scope of this work). In the following two chapters we discuss methods that utilize random integrals, which result either in nonlocal forms or in expressions for effective conductivity. In the following section, we complete our discussion on homogenization with typical examples for each method.

⁸As well as for bounding effective conductivity of heterogeneous media, which is beyond the scope of this work.

1.4.2 Spatial Homogenization

Although we will not follow the spatial “effective” conductivity approach in our solutions to the stochastic steady-state flow equation, in the following, we review upscaling methods, for two reasons: first, as was mentioned above, and as will be seen in the sequel, in several occasions, particularly when the self consistent approach is used, there is no clear distinction between “ensemble” and “spatial” moments, which implies ergodicity of some sort; second, some of the methods are extended to nonuniform and unsteady flow conditions, and/or are compared, asymptotically, to the ensemble approaches (discussed in the next section).

Periodic and Multi-Scale Structures

A perturbation method which is somewhat different from the one used in our field (e.g., Dagan, 1989) is used in the context of periodic structures and *multiple scales* by Auriault (1991), and by Hinch (1991; pp. 126-127) respectively. Both authors emphasize two characteristic scales, macroscopic (or large scale) L and microscopic (or small scale) l , such that (Auriault, 1991)

$$\frac{l}{L} = \epsilon \ll 1 \quad (1.15)$$

Auriault (1991) starts from the macroscopic Laplace equation (i.e., $\nabla \cdot K(\mathbf{x})\nabla h(\mathbf{x}) = 0$) in a periodic porous medium, and the microscopic Stokes equation,

$$\nabla^2 \mathbf{v} = \frac{1}{\mu} \nabla p, \quad (1.16)$$

where μ and p are (fluid) viscosity and pressure respectively. Auriault then expands asymptotic series for both velocity (\mathbf{v}) and pressure, similar to Eq 1.19 (below), where

he considers p both as a macroscopic and as a microscopic variable; then he follows the derivation of Ene and Sanchez-Palencia (1975) of Darcy's law. Next, Auriault defines a dimensionless characteristic

$$Q = \frac{\nabla p}{\mu \nabla^2 \mathbf{v}} = O(pl/\mu v) \quad (1.17)$$

and concludes that the determination of an equivalent macroscopic description is possible only if $Q = O(\epsilon^{-1})^9$.

A similar methodology was used by Mei and Ariault (1989) in order to extend the theory to media with several disparate spatial scales, where Eq. 1.15 refers now to all hierarchies in the periodic medium (i.e., the medium is assumed to be periodic in both parameters and responses, on all scales). Following Ene and Sanchez-Palencia (1975), Mei and Ariault (1989) take full advantage of the perturbation technique (between zero and second orders) and succeed in defining effective permeability tensors on the different scales, which are positive definite and symmetric. Their upscaling procedure starts from the spatial (cell) average of the micro-scale conductivity as the effective conductivity of the next higher scale (their Eq. 2.21), and then continues with larger scales, such that at each scale the effective conductivity corresponds to the average zero order perturbation of the flux (or velocity).

Hinch (1991) starts from the elliptic *heat conduction equation*, which is identical to our Poisson (steady-state flow) equation

$$\nabla \cdot K \nabla T = Q \quad (1.18)$$

⁹By $f(x) = O[g(x)]$ as $x \rightarrow x_o$ we mean that $f(x)/g(x)$ is bounded as $x \rightarrow x_o$. In contrast, the *asymptotic* notation $f(x) \sim g(x)$ as $x \rightarrow x_o$ means that $[f(x)/g(x)] \rightarrow 1$ as $x \rightarrow x_o$.

where T is the temperature, and K and Q are given functions which have a fine scale structure described by the short scale variable $\xi = x/\epsilon$ (Q is equivalent to $-f$ in Eq. 1.1). The asymptotic expansion is then done on the long scale x ,

$$T(x, \epsilon) \sim T_o(\xi, x) + \epsilon T_1(\xi, x) + \epsilon^2 T_2(\xi, x) \quad (1.19)$$

and substituted into the governing equation. The main assumption is that at leading (or zero) order, the temperature does not vary on the microscale (only at the macroscale), as expressed by

$$\frac{\partial}{\partial \xi} \left(K \frac{\partial T_o}{\partial \xi} \right) = 0 \quad (1.20)$$

Consequently, at leading order (i.e., on a large scale) the temperature satisfies a standard heat conduction equation with an effective heat conductivity K^* and effective source term Q^* , such that

$$\frac{\partial}{\partial x} \left(K^* \frac{\partial T_o}{\partial x} \right) = Q^* \quad (1.21)$$

where $Q^* \equiv \langle Q(\xi, x) \rangle$, and

$$K^* = \langle K(\xi) + K(\xi) \frac{\partial A}{\partial \xi} \rangle \quad (1.22)$$

where the ensemble average is taken over the ξ -microscale. $A \equiv A(\xi)$ is a coefficient of linearity which emerges from the expansion and from the distinction between the two scales; it depends on the details of $K(\xi)$ and links between the leading order (0) and the higher order (1), i.e.,

$$T_1(\xi, x) = A(\xi) \frac{\partial T_o}{\partial x} \quad (1.23)$$

Note that $\partial A/\partial \xi$ in (1.22) must be negative in order for $K^* \leq \langle K \rangle = K_A$. On one hand, the effective conductivity defined in (1.22) is not limited to mean (large scale)

uniform flow, and is not limited to a specific distribution of K ; on the other hand, K is assumed periodic, which is generally more restrictive than a specified distribution. The dependency on the variance must be included in the second term in (1.22), and hence, it should be implicit, in part, in $A(\xi)$. However, it is not clear from the text (i.e., Hinch, 1991) how this coefficient is calculated. Interestingly, Hinch (1991) remarks that “...higher order corrections (or expansion) will find that there is a correction to the heat flux which is *non-local*, i.e., the heat flux at one point depends on the value of the temperature gradients in the neighborhood of that point”.

Keller (1987) systematically introduces theorems about effective conductivities of reciprocal media (a basic form of periodic media), and extend them to stationary random media. The most relevant theorems¹⁰ are:

- Mendelson (1975): For a two dimensional statistically stationary (or homogeneous) and isotropic random medium, $K_e = K_G$.
- Kohler and Papanicolau (1982): For a two dimensional statistically stationary (but not necessarily) isotropic random medium, $K_{e_x} = 1/\hat{K}_{e_y}$, where K_{e_x} and K_{e_y} are the effective conductivities in the x and y directions respectively, and \hat{K}_e is the effective conductivity of the reciprocal field $1/K(x)$.
- Kohler and Papanicolau (1982): For a three dimensional statistically stationary random medium, $K_{e_x} \geq 1/\hat{K}_{e_y}$.

These conclusions are very similar to the earlier work by Matheron (1967) mentioned above. While the above theorems are not limited to log-normal distribution, they are

¹⁰The theorems were proved primarily by using variational principles.

yet based on, *two component media*, i.e., with two elementary conductivities K_1 and K_2 subject to random geometrical arrangement; this implies a much more restrictive condition, i.e., two delta distributions.

Following the Taylor-Aris *method of moments*, as generalized by Brenner (1980,1982), Kitanidis (1990) derives the governing system of equations for the “effective” upscaled hydraulic conductivity of a periodic porous medium. The key assumptions made in the method of moments are: (1) the conductivity function is periodic, and in the limit, as the size of the elementary cell $l \rightarrow \infty$, it is a stationary random field¹¹, (2) the boundary conditions are periodic, and (3) the flow is *gradually* varying; particularly, the zeroth, first, and second central moments of the head field computed over the averaging volume, V , reach a steady rate of change. The averaging volume, in turn, is assumed to encompass one period of an idealized periodic porous medium.

In Kitanidis’ (1990) view, a periodic medium is only a conceptual or mathematical model of a formation whose parameters fluctuate about an average value. According to Dykaar and Kitanidis (1992), the periodicity assumption places an upper limit on the largest scale of the conductivity field. This is in contrast with Dagan’s (1986, 1989) view of periodic media as “ordered media which are seldom encountered in natural formations”.

According to the moment method, the effective conductivity is found in two steps (*ibid*). First, in an n -dimensional locally isotropic¹² flow domain, n uncoupled

¹¹The extension to stationary random function is controversial (see below).

¹²Kitanidis’ (1990) solution satisfies local anisotropy of \mathbf{K} as well.

partial differential equations

$$\sum_{p=1}^n \frac{\partial}{\partial x_p} \left(K(\mathbf{x}) \frac{\partial g^i(\mathbf{x})}{\partial x_p} \right) = \frac{\partial K(\mathbf{x})}{\partial x_i} \quad i = 1, \dots, n \quad (1.24)$$

are solved for the n unknown auxiliary functions $g^i(\mathbf{x})$ in an n dimensional parallelepiped. In 3-D, the flow domain has a volume $V = l_1 l_2 l_3$ with sides l_1, l_2, l_3 . Eq. 1.24 is subject to periodic boundary conditions, which means that g^i and all its derivatives are the same at $x_j = 0$ and at $x_j = l_j$. The functions g^i are then used in the integral

$$K_{e_{ij}} = -\frac{1}{2V} \int_V K \left(\frac{\partial g^i}{\partial x_i} + \frac{\partial g^j}{\partial x_j} \right) d\mathbf{x} + \bar{K} \delta_{ij} \quad (1.25)$$

to determine the components of the effective conductivity tensor $K_{e_{ij}}$, where δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ for $i = j$, and $\delta_{ij} = 0$ otherwise), and \bar{K} is the spatial average of K :

$$\bar{K} = \frac{1}{V} \int_V K(\mathbf{x}) d\mathbf{x} \quad (1.26)$$

Kitanidis' (1990) analysis is based on a *transient* head response to an injected impulse of water at some point in the medium; after enough time has elapsed for the mound to spread out over an area larger than the scale of conductivity fluctuations, an upscaled (or block) conductivity can be determined. Following Dagan (1982, 1989), when $t > \bar{l}^2 S / \bar{K}$, where \bar{l} is the scale of conductivity fluctuations, S is the specific storage, and \bar{K} is a typical value of conductivity, the flow is slowly or gradually varying, the scale of head fluctuations is considerably larger than the typical length scale of the hydraulic conductivity fluctuations, and the net rate of spreading converges to a constant, i.e., it reaches a quasi-steady state. Similar conclusions were reached by El-Kadi and Brutsaert (1985).

Kitanidis' approach is purely deterministic; local conductivities are specified precisely at each point, with symmetric boundary conditions for each element in

his superimposed grid (thus, emulating an infinite domain; cf. Durlofsky, 1991, below). In the asymptotic case, when the period is extended to infinity, Kitanidis' (1990) results become similar to the known small-perturbation results by Gutjahr et al. (1978) and Gelhar and Axness (1983), who assumed a normal distribution of $Y = \ln K$ and small σ_Y^2 . Further, without invoking either the Gaussian distribution assumption (in fact, without invoking the stochastic process approach at all) or uniform flow, Kitanidis' results perfectly resemble the conductivity bounds for the asymptotic cases of perfect stratification (discussed above).

Encouraged by these results, Dykaar and Kitanidis (1992) use a numerical Fourier Galerkin spectral method to discretize Eq. 1.24, and then solve the resulting system of equations with the conjugate gradient method, using FFT's for the matrix-vector multiplication. However, they find that beyond a certain variance of $\ln K$ (or, equivalently, beyond a certain condition number of the transformed matrix), the conjugate gradient method would not converge with the standard FFT deconvolution procedure. Consequently, they attenuate the high-frequency components of (the Fourier transformed) \tilde{K} by applying a low-pass filter, "which tends to round corners, and fill in the sharp peaks and valleys of $K(\mathbf{x})$ " (*ibid*). The filtering is implemented by multiplying the discrete Fourier coefficients by weighting factors. The authors do not present any criterion for judging the validity of this approach, and/or an estimate of the error induced by the discrete sampling. However they cite three possible sources of errors: (1) aliasing, caused by insufficient (discrete) sampling rate of $K(\mathbf{x})$, (2) Fourier series truncation, or low-pass filtering, and (3) numerical round-off. Defending their low-pass filtering, they (*ibid*) claim that "all numerical schemes, such as finite difference and finite elements, are subject to the same error

sources”, particularly, “the finite difference or finite element operator itself acts as a filter to varying degrees, depending on the particular differential operator”. We discuss the above argument later in this section, and in Chapter 4. In Chapter 5 we use a sophisticated multi-grid finite element method, with adaptive grid refinement, which eliminates both the smoothing and the convergence problems.

The advantage of the Fourier Galerkin method used by Dykaar and Kitanidis (1992), is its efficiency; a solution can be achieved in order $N \ln(N)$ operations and order N storage, where N is the total number of sampling points used in a flow domain. This allowed them to simulate very large fields in order to calculate upscaled conductivities of very large blocks in stationary fields of $\ln K$, such that ergodicity in the mean would be justified, an REV (representative elementary volume) could then be defined, and the block conductivity, K_b , could be considered as the corresponding homogeneous constant. Their analysis shows that in two dimensional fields having an exponential, isotropic covariance, about 80 integral scales are needed to achieve this goal, each integral scale encompassing at least 8.5 conductivity values (or elements, in our terms). These requirements were reduced by a factor of 4 for a Gaussian covariance model. In three dimensions, with a Gaussian covariance structure, about 30 integral scales were needed to consider one realization as an infinite domain. The results from the moment method show good agreement with simple “inverse” method solutions (which consider the flow domain as a flow tube, and calculate the effective conductivity as the ratio between the total flux and the imposed gradient), for both stationary and bimodally distributed hydraulic conductivity fields consisting of shale lenses embedded in sandstone. As was mentioned earlier, their result are in (striking) agreements with the Landau and Lifshitz (1960) and Matheron (1967) conjecture,

and Neuman et al. (1992) for the 3-D effective conductivity, up to variance (of $\ln K$) of 6. It seems, though, that beside the efficient (but smooth) numerical spectral method, the moment method does not have a significant advantage over the simple inverse method (where block conductivity is calculated as the ratio between the total flux and the imposed gradient), because the number of the generated conductivity values has to be large in either case; rather than solving for the auxiliary functions g^i , one could solve directly for heads and fluxes.

From our point of view, the most interesting outcome of the works of Kitani-dis (1990) and Dykaar and Kitanidis (1992) is the fact that the spreading mound is not only transient, but also resembles *radial* flow. This implies that, in the above asymptotic cases, there exists an “effective hydraulic conductivity”, as long as the investigated area is far enough from the singularity. This is in general agreement with Dagan (1989), Naff (1991), and with our Chapter 5.

Durlofsky (1991) uses a numerical procedure to determine equivalent grid block permeability tensors for heterogeneous, infinite periodic porous media under uniform flow conditions. As in the homogenization theory of Bourgeat (1984) and subsequent extensions by Saez et al. (1989) and Mei and Auriault (1989), Durlofski imposes periodic boundary conditions, such that for each cell, one pair of opposing faces is subject to equal fluxes and heads (hence, periodic in fluxes and heads), while the other pair of opposing faces is subject to simple Dirichlet boundary conditions (i.e., to an imposed uniform gradient in each cell)¹³. That is, he assumes that the pressure variation is locally linear. The calculation of the equivalent permeabilities

¹³To distinguish from Mei and Auriault (1989), whose equation is driven by a source term of the form $\nabla \cdot \mathbf{k}$, allowing their pressure-like variable (S) to remain periodic on all faces.

is based on spatial (arithmetic) averaging of the fluxes on two (adjacent, perpendicular) faces. Durlofsky (1991) distinguishes between *effective* permeability, which is “a property of the medium, and does not vary with the flow conditions”, and *equivalent* permeability, which, by his definition, is expected to vary under different flow conditions.

Due to the particular choice of (essentially periodic) boundary conditions, Durlofsky’s (1991) equivalent permeability tensors are always symmetric and positive definite.

Although the periodic media approach does not require a priori knowledge of the principal directions of the equivalent permeability tensor, it requires a priori knowledge of the two distinctive scales, in order to choose the unit cell. In fact, it can be shown that below a certain cell size (which is case specific), the method is very sensitive to cell dimensions. For a case with a simple block structure and distinctive pattern, Durlofsky (1991) demonstrates the superiority of the periodic boundary conditions over the common inverse method (which imposes no-flow BC parallel to an imposed gradient, in a “flow tube” like blocks). Durlofsky admits, however, that for the general case of heterogeneous media, “the equivalent grid block permeability of the region can no longer be specified unambiguously, as they result from the local flow field, which is not generally known a priori”. Despite the inherent limitations of the method, he presents results of numerical experiments with two different fractal fields (though at least one of them shows a distinct pattern), where calculated effective block permeabilities succeed in reproducing the average (spatial) heads and streamlines. It seems that the boundary conditions chosen for each case, are not necessarily periodic (in fact, in the second case, the common “flow-tube” is used).

In conclusion, it seems that the main limitation of periodic media analyses is the required equivalence between block size and “period”, and the consequent assignment of “periodic boundary conditions” to each block. Auriault (1991) argues that “If we limit ourselves to the discovery of the structure of the macroscopic description, periodic and random media are equivalent, if homogenization is possible”. It is not clear, however, what he means by “structure”, and when is homogenization possible. Similar arguments are given by Dykaar and Kitanidis (1992) and Keller (1987). It seems to us, that although there is some similarity, random media are (unfortunately) much more complex than periodic media; in fact, they add another dimension, the *probability space*, to the problem. Particularly, in periodic media the parameter fluctuations are *bounded* in a known manner, while in random media, the *moments*, not the fluctuations, are bounded. In periodic media averaging is always *spatial*, while in random media both spatial and ensemble averages may play a role (cf. Keller, 1987).

The Self Consistent Method

The *self consistent approach* (also termed *embedding matrix approach*, and *effective medium theory*) was originally used to estimate effective heat conductivity and dielectric constants by Budianski (1965), Hashin and Shtrikman (1962), Beran (1968), Landauer (1978), and others, and was adopted by Dagan (1979, 1981) to estimate effective hydraulic conductivity of heterogeneous *isotropic* formations; it was later extended by Poley (1988) and by Dagan (1989) to *anisotropic* formations.

According to the self consistent approach, the medium is viewed as a collection

of blocks of different conductivity, shape, size and location, set at random with no correlation between them. The self-consistent model concentrates on a representative inclusion of permeability K_j of typical dimension R_{jk} , and replaces the remaining ones by an embedding homogeneous matrix of unknown permeability K_o , which extends to infinity. The head h satisfies Laplace's equation in each block, as well as continuity of head and normal flux at the interface between blocks. At infinity (on the far boundary), $h = -\mathbf{J} \cdot \mathbf{x}$, such that for a homogeneous formation the flow is uniform (i.e., the gradient and hence, the flux are constant). The idea is to isolate one block [Poley (1988) uses ellipsoids, and Dagan (1981) uses oblate spheroids], to solve for h , and then to determine K_o of the surrounding matrix from the requirement that the field be *undisturbed*. "If such a value can be found, one can fill the entire space with composite spheroids of different sizes without affecting the flow ... consequently, one can substitute $K_e = K_o$ " (Dagan, 1979, 1989). This is reminiscent of the method of *superposition* employed in electrical circuits for the determination of effective resistance (although in this case, the shape of the resistors is, justifiably, irrelevant). Perhaps the strongest assumption is that the disturbance in the head field caused by any one inclusion is not affected by similar disturbances due to other inclusions; another limitation is the dependence on the (assumed) *shape* of the inclusion.

To allow using a known solution (Lamb, 1932) for flow around a (moving) solid ellipsoid parallel to a principal axis in a homogeneous liquid, both Poley (1988) and Dagan (1981) take all ellipsoids (or spheroids) to have parallel principal axes, and then transform the coordinates of the (regionally) anisotropic medium to an isotropic one. This way, the blocks become anisotropic and change dimensions, requiring that

flow within the inclusions be assumed uniform in the (horizontal) x direction (parallel to both regional flow and the principal axis of the inclusions). Another way to handle the global anisotropy is to equate K_o with K_h or K_v (horizontal and vertical “homogeneous matrix” conductivities, respectively), depending on the direction of the mean gradient, without stretching the coordinates, such that the inclusions remain isotropic. As was pointed out by Dagan (1989, p. 202), since the field is considered uniform in the surrounding matrix, it is only the principal component of \mathbf{K}_e in the direction of \mathbf{J} which counts. Despite the use of the same approach by the two authors, there are differences between the derivations of Dagan (1979, 1981) and Poley (1988) [Dagan’s derivation seems to be more concise, consistent, and less obscure than that of Poley]. Their derivations lead to different expressions for the effective hydraulic conductivity. However, both expressions for the effective principal hydraulic conductivities are functions of the probability density function (*pdf*) of K , the principal directions of the effective conductivity tensor (assumed known), and the (characteristic) *shape* of the assumed inclusions (see below). We concentrate, below, on Dagan’s (1979, 1981, 1989) derivation. Dagan (1979) borrows the general solutions of Lamb (1932) for flow within and outside the inclusion (spheroid/cylinder in the isotropic case, and oblate spheroids in the anisotropic case). He then uses the boundary conditions along the interface and at infinity to define the constants for the relevant solution. Once a solution for head is obtained, it is used to calculate the contribution of a particular inclusion to the total gradient and flow, by averaging this contribution over the total volume. Such averaging is justified by the assumption of complete randomness of the block centroids, \mathbf{x}_i , such that its *pdf* is uniform (i.e., the *pdf* of the block centroids, $f(\mathbf{x}_i) = 1/V$, where V is the total volume). Since the

above contributions depend on the relative distances $\mathbf{x} - \mathbf{x}_l$, the integration can be carried out over \mathbf{x} , with \mathbf{x}_l kept fixed at the origin.

The next step is to integrate all average contributions of all blocks of different conductivities over the infinite volume, provided that the *pdf* of the hydraulic conductivity, $f(K)$, is known (correlation is ignored). Now, the average flux, $\langle \mathbf{q} \rangle = K_o \mathbf{J} + \sum \langle \mathbf{q}' \rangle$, where the flux fluctuation, \mathbf{q}' , is averaged for all possible sizes of (similar) blocks, for all block conductivities (K_j) and locations (\mathbf{x}_l),

$$\langle \mathbf{q}' \rangle = \int_V \mathbf{q}' f(K) f(\mathbf{x}_l) dK d\mathbf{x}_l \quad (1.27)$$

Finally, $K_e = \langle \mathbf{q} \rangle / \mathbf{J}$, and therefore, K_e is a function of $f(K)$, K_o , and the characteristic shape of the block. The very last step is to substitute $K_e = K_o$ and to solve (iteratively) for K_e . This last step seems to be inconsistent, at first glance, because up to this point, a careful probabilistic technique combined with an accurate hydrodynamic solution were employed; consequently, one would expect that in a similar (consistent) manner, a probabilistic approach should be used for the interaction between the surfaces of the different blocks.

For isotropic media, Dagan (1979) arrives at the following effective conductivity, K_e ,

$$K_e + (n - 1)K_o = \left[\int_0^\infty \frac{f(K) dK}{(n - 1)K_o + K} \right]^{-1} \quad (1.28)$$

where n is the number of spatial dimensions, and $f(K)$ is the frequency distribution of K . [Note that the integral in (1.28) resembles the ensemble mean of $[(n - 1)K_o + K]^{-1}$.] A lower bound for the effective hydraulic conductivity is obtained if K_o in (1.28) is substituted by the harmonic mean K_H , and an upper bound is reached for $K_o = K_A$, the arithmetic mean. These bounds are also reached in the anisotropic

case (below). Finally, the self consistent argument is made, that the conductivity of the matrix surrounding each inclusion is equal to the effective conductivity, that is, $K_o = K_e$, and hence

$$K_e = \frac{1}{n} \left[\int_0^\infty \frac{f(K) dK}{(n-1)K_e + K} \right]^{-1} \quad (1.29)$$

No small variance assumption is made, but interactions between inclusions are not accounted for. Interestingly, an expression equivalent to Eq. 1.29 was obtained by Kirkpatrick (1973, Eq. 5.4) in his analysis of effective conductance of random *resistor* networks (see review by Gomez-Hernandez, 1991). It is also interesting to note that in order to solve the integral equation (1.29), one has to use successive approximations (discussed in Chapter 2).

For anisotropic media, Dagan (1989) obtains the following principal horizontal and vertical effective conductivities:

$$K_{e_h} = \frac{1}{2} \left[\int_0^\infty \frac{f(K) dK}{(K - K_{e_h})\gamma(e'') + 2K_{e_h}} \right]^{-1} \quad (1.30)$$

$$K_{e_v} = \left[\int_0^\infty \frac{f(K) dK}{(K_{e_v} - K)\gamma(e'') + K} \right]^{-1} \quad (1.31)$$

where $e'' = e(K_h/K_v)^{1/2}$ is the transformed eccentricity (i.e., the eccentricity in the transformed coordinate system) and $\gamma(e'')$ is a shape factor which depends on both the eccentricity and the principal component of the effective hydraulic conductivity. In Eqs. 1.30 and 1.31 K_o of the surrounding matrix was replaced by the principal components $K_{h_x} = K_{h_y} = K_h$, and K_v . The eccentricity can be defined as $e = \lambda_{Y_v}/\lambda_{Y_h}$, the ratio of vertical to horizontal integral scales of $\ln K$ (i.e., it expresses the magnitude of the statistical anisotropy). Note that here, there are explicit relationships between the principal directions of the effective and the statistical K tensors, and that the variances of $\ln K$ is unlimited. One way to check the validity of the resulting

effective hydraulic conductivity expressed by equations 1.30 and 1.31 is to examine some extreme situations. Dagan (1989) consider cases in which:

- $e \rightarrow \infty \quad \implies \quad \gamma = 1 \quad \rightsquigarrow \quad$ 2-D medium, i.e., $K_{e_h} \rightarrow K_G$
- $e \rightarrow 0 \quad \implies \quad \gamma = 0 \quad \rightsquigarrow \quad$ Stratified, i.e., $K_{e_h} \rightarrow K_A, \quad K_{e_v} \rightarrow K_H$
- $e \rightarrow 1 \quad \implies \quad \gamma = 2/3 \quad \rightsquigarrow \quad$ Isotropic medium, i.e., $K_{e_h} = K_{e_v} = K$

In the isotropic case, equations 1.30 and 1.31 become identical to the 3-D version of (1.29), that is $K_{e_h} = K_{e_v} = K_e$. In the 2-D case ($\gamma = 1$), Eq. 1.30 becomes identical to the 2-D version of (1.29), that is, $K_{e_h} = K_e$, but (1.31) becomes an identity (i.e., $K_{e_v} = K_{e_v}$). In the stratified case, ($e = \gamma = 0$), $K_{e_v} = K_H$ in Eq. 1.31, but Eq. 1.30 leads to an identity ($K_{e_h} = K_{e_h}$) instead of $K_{e_h} = K_A$. Perhaps for this reason, rather than substituting the above limits (as described above) in (1.30) and (1.31) explicitly, Dagan (1989) prefers to expand (1.30) and (1.31) in Taylor series for the stratified case, assuming σ_Y^2 is small; in this way he obtains $K_{e_v} \rightarrow K_H$ and $K_{e_h} \rightarrow K_A$.

Besides comparing the extreme cases with the known lower and upper bounds of the effective hydraulic conductivity, Dagan's self consistent approach shows a good agreement with isotropic effective conductivities approximated by Gutjahr et al. (1978) and Gelhar and Axness (1983) for small variances (differing by only 1% in 3-D up to $\sigma_Y^2 = 2$). Perhaps the most significant check on the method is that for the 2-D isotropic case it yields $K_e = K_G$, the geometric mean, with high degree of accuracy, for any variance of $\ln K$ (up to $\sigma_Y^2 = 7$), for a lognormal distribution of K (Dagan, 1989). Comparisons with results from numerical experiments by Desbarats (1987; Monte Carlo simulations) and by Dykaar and Kitanidis (1992; large field inverse

method) show good agreement in the cases of binary (or two-phase) permeabilities (e.g., shale lenses embedded in sandstone). However, for general three dimensional statistically homogeneous and isotropic conductivity fields, the self consistent method gives a poor approximation for the effective conductivity for $\sigma_Y^2 > 1$, as illustrated by Dykaar and Kitanidis (1992). Desbarats (1987) demonstrates a very good agreement between the self consistent method and 2-D numerical experiments with statistically isotropic and homogeneous fields which possess some correlation, but less so for uncorrelated fields; he concludes that the shortcoming of the method is that it is unable to make any distinction between media with different covariance structures (as is apparent in Eq. 1.29).

Geostatistical and Power Averaging

A considerable effort has been devoted by many researchers, to investigate and develop geostatistical models for upscaling hydraulic conductivities in two and three dimensions. Similar to the (ensemble) statistical models above, the conductivity is represented as a spatial random function, where heterogeneity is described, primarily, by the mean and the variogram of the point scale values [although, Monte Carlo simulations (MCS) do require a *pdf*, and kriging of $\ln K$ implies lognormal distribution of K]. In the following paragraphs, we will focus on works by Desbarats and co-workers, whose goal is to empirically obtain expressions for conductivities at the (large) block scale as spatial averages of point scale values, by means of numerical experiments (or MCS).

Desbarats and Dimitrakopoulos (1990) find numerically that the spatial *geometric* average is an excellent approximation of effective block transmissivity for two

dimensional statistically homogeneous and isotropic log-normal fields under mean uniform flow. [This was confirmed independently by numerical inverse MCS done by us (in an earlier work).] They also find that the mean and the variance of block-averaged values depend on the averaging scale relative to the integral scale; as the ratio L/λ increases (L being the domain's side, and λ the integral scale), mean block transmissivity decreases from the (ensemble) arithmetic average for $L/\lambda \ll 1$ towards the ensemble geometric mean, while the variance of the block (averaged) conductivity decreases to zero as L/λ becomes large, as expected, from Vanmarcke, 1983). [The dependence of block conductivity on relative block size was independently confirmed by our (early) inverse MCS with similar conductivity fields.]

The geostatistical model of block transmissivity provides relationships between the mean and variance of block-averaged quantities, the spatial area, and the spatial correlation structure of the (point) transmissivity. Desbarats and Dimitrakopoulos (1990) find that if Y_s is the spatial average log-transmissivity over a block (or area), and Y_s^* is its kriged estimate (based on point data), the block transmissivity estimate $T_s^* = \exp(Y_s^*)$ is closer to the theoretical ensemble geometric mean, than is the unbiased (corrected) block transmissivity $T_s^{**} = T_s^* \exp(\sigma_K^2/2 + \eta)$, where η is a Lagrange multiplier added to the kriging equations (in which the kriging variance is minimized) in order to satisfy the unbiasedness constraint $\sum_i \beta_i = 1$ (β_i being the weights used in the kriging estimation/interpolation). Hence, the former (i.e., the biased estimator) is superior to the latter (unbiased estimator). The additional terms, $\exp(\sigma_K^2/2 + \eta)$, compensate for the variance reduction, or “smoothing effect”, common to all least-squares methods (see also Clifton and Neuman, 1982, de-Marsily, 1986, p. 308, and Journel and Huijbregts, 1978, p. 572). Desbarats and

Dimitrakopoulos (1990) bring an example of a sandstone unit in an oil reservoir, where each “point” transmissivity is defined as the integral of permeabilities of core samples over the reservoir thickness (implying arithmetic averaging over depth). After subdividing their (rectangular) region into a regular grid of square blocks, and kriging them one by one, they conduct steady-state flow simulations, imposing no-flow boundaries along the mean flow direction, and prescribed heads on the other (transverse) boundaries. The effective transmissivity is calculated from numerical solution of the flow equation (apparently as the ratio between the average longitudinal flux and the imposed gradient) and compared with both “biased” and “unbiased” estimates of block transmissivity; this is repeated for blocks of different sizes, all of which were smaller than the integral scale. They find that the latter bias correction of kriged estimates leads to an artificial overestimation of the overall effective transmissivity of the flow field, while the original, uncorrected kriging estimator does not introduce artifacts of block size in flow simulations. However, their results indicate that the differences between the above estimates shrink as the block size increases; this suggests that for $L \gg \lambda$ the artifacts of block size may be negligible.

Desbarats (1992a) uses spatial power averaging as an empirical method for up-scaling hydraulic conductivity in three dimensional heterogeneous formations, where “point” scale conductivity field is modeled as a stationary (i.e., statistically homogeneous), ergodic, and multivariate lognormal random field. The *power average* block conductivity, K_b , is defined by him as

$$\begin{aligned} K_b &= \left(\frac{1}{V} \int_V K(\mathbf{x})^\omega dV \right)^{1/\omega} & \text{for } \omega \neq 0 \\ K_b &= \exp \left(\frac{1}{V} \int_V \ln K(\mathbf{x}) dV \right) & \text{for } \omega = 0 \end{aligned} \quad (1.32)$$

where the power ω is a function of the covariance structure, the dimensions of the flow

field, and cell dimensions (resulted from the discretization); the only way to find ω is by numerical experiments (*ibid*). Desbarats (1992a) conducts high resolution numerical experiments (in fact, MCS) of steady state flow through 3-D random, lognormal hydraulic conductivity fields with an isotropic exponential covariance structure, and $L/\lambda = 3, 2, 1$. In order to calculate block effective conductivity, Desbarats adopts the formulation of Rubin and Gomez-hernandez (1990) for the upscaled hydraulic conductivity, K_u ,

$$K_u = \frac{\int_V q_x(\mathbf{x})dV}{\int_V J_x(\mathbf{x})dV} \quad (1.33)$$

where q_x and J_x are the longitudinal flux and head, respectively, which are obtained from the numerical solution to the flow equation $\nabla \cdot \mathbf{q} = 0$. For $\sigma_Y^2 \leq 2$, Desbarats' numerical experiments (200 realizations) show an excellent agreement between K_b in (1.32) with $\omega = 1/3$ and K_u in (1.33); this agreement deteriorates for $\sigma_Y^2 = 4$ (apparently due to the limited number of simulations and/or the limited field size). The experiments (MCS) show (again) the decrease in mean block conductivity from the arithmetic mean for $L/\lambda \ll 1$ to the theoretical (and experimental) value (Aitchison and Brown, 1957)

$$K_\omega = \langle K(\mathbf{x})^\omega \rangle^{1/\omega} = \exp[\langle Y(\mathbf{x}) \rangle + \omega(\sigma_Y^2/2)] \quad (1.34)$$

for $L/\lambda \gg 1$, where K_ω is the *ensemble power mean* of $K(\mathbf{x})$ for the exponent ω (Desbarats, 1992a), which, after finding ω , becomes the effective block conductivity. Ababou (1989, 1990) notices that the above power average K_ω could be expressed as

$$K_\omega = K_G \exp[\sigma_Y^2(1/2 - g)] \quad (1.35)$$

where $K_G = e^{\langle Y(\mathbf{x}) \rangle}$ is the ensemble geometric mean, and $g = (1 - \omega)/2$ is a geometric factor. This relation is identical to an expression for effective conductivity of an

infinite 3-D anisotropic fields derived by Gelhar and Axness (1983); the factor g in their paper is a function of the anisotropy ratio of the principal integral scales, and is equal to $1/3$ in the isotropic case, which corresponds to $\omega = 1 - 2g = 1/3$. Consequently, Eq. 1.35 resembles, the relationships postulated earlier by Landau and Lifshitz (1960), Matheron (1967), and results of MCS by Dykaar and Kitanidis (1992) and by Neuman et al. (1992a; see Chapter 3). Although the latter authors use a simpler (more “crude”) definition of K_u than (1.33), i.e., they define block conductivity as the ratio of the average flux through the upstream boundary to the global imposed gradient, the results for K_u are very similar. differently than in (1.33)

Considering the infinite field case, Ababou (1988, 1990) conjectures the following expression for ω in terms of the principal integral scales:

$$\omega_i = 1 - \frac{2}{n} \frac{\lambda_H}{\lambda_i}, \quad (1.36)$$

where n is the field dimension and λ_H is defined as

$$\lambda_H = \left(\frac{1}{n} \sum_{i=1}^n \lambda_i^{-1} \right)^{-1}. \quad (1.37)$$

By combining (1.36) – (1.37) for one of the principal directions in a two dimensional domain (with K log-normal), and considering the definition of g (above), (1.35) takes the form

$$K_{\omega_1} = K_G e^{\sigma_Y^2/2} \exp \left(-\frac{\sigma_Y^2 \lambda_2}{\lambda_1 + \lambda_2} \right), \quad (1.38)$$

which (surprisingly) appears to be a particular case of a general expression obtained by Paleologos (1993; see also Neuman et al., 1993). Although (1.38) does not particularly comply with Desbarats’ experiments with anisotropy ratio of $10 : 1$, it provides correct values for ω for the special cases of perfectly stratified and isotropic fields

(Desbarats, 1992a); it would be worthwhile to test them for a broader spectrum of anisotropy ratios in two and three dimensional fields.

A fairly substantial amount of work has been devoted to power averaging of sand-shale sequences. This type of formation is characterized by a high variance, a bi-modal distribution, and a strongly anisotropic spatial covariance structure. Desbarats (1987) discusses several different approaches and formulations, including *percolation theory*, and concludes that there is no general and accurate method for estimating effective permeability in sand-shale sequences. Considerable mathematical simplification is achieved when a *binary* permeability model is assumed, where permeabilities take on one of two possible values, K_{ss} and K_{sh} , for sandstone and shale, respectively. Practical experience with operation of sand-shale oil reservoirs, as well as Desbarats's numerical experiments, seem to support this assumption (Desbarats, 1987).

A recently developed method for simulating spatially correlated *indicator* variables (termed *indicator geostatistics*) was used to investigate shale spatial covariance in sand-shale formations; the method successfully reproduced patterns of sand-shale heterogeneity observed in an outcrop of the Assakao sandstone and other similar formations (e.g., Gomez- hernandez and Journel, 1989 and Gomez-Hernandez, 1991b). The shale indicator is a binary variable that takes a value of 1 for shale locations and 0 for sand locations. The covariance of the indicator variable is directly related to the orientation and width-to-length ratio of shale segments. Journel et al. (1986), Desbarats (1987a, 1987b) and Deutsch (1987, 1989) find the block conductivity to be a function of shale proportion, block size relative to correlation scale of the indicator covariance, and conductivity contrast between sand and shale. They propose a power

average law for each principal direction, i , of the form

$$K_{V_i} = (pK_s h^{\omega_i} + (1 - p)K_s s^{\omega_i})^{1/\omega_i}, \quad (1.39)$$

where p is the proportion of shales, and ω_i are coefficients to be determined empirically (e.g., by MCS). The block is assumed to be oriented in the principal directions. The method was applied by Bachu and Cuthiell (1990) to conductivities of several core samples from a Canadian reservoir, where each core sample is considered a *block*. Values of ω_i were obtained for the different core samples.

Summarizing, the power averaging method is based on inverse Monte Carlo Simulations for defining block hydraulic conductivity with maximum generality. The most common and straightforward inverse technique is to prescribe constant heads on two sides and no-flow boundaries on all other sides of the block, emulating a (straight) stream tube with imposed (mean) gradient; the principal directions of the block conductivity tensor must be assumed. Techniques that are based on the concept of periodic-media (e.g., Durlofsky, 1990; Kasap and Lake, 1990), naturally use periodic boundary conditions, and do not require a priori knowledge of principal directions of the block conductivity tensor; however, they must assume two distinctive scales. In both methods, boundary effects are minimized and neglected; consequently, very large fields ($L \gg \lambda$) must be considered. Limitations and difficulties associated with MCS are discussed in Chapters 4 and 5.

1.5 Solutions of Steady State Stochastic PDE by Perturbation Method

We recall that the effective conductivity is not essential for the solution of the stochastic flow equation, but is rather a “desirable” property; when it exists, it

can provide the *mean* response (head or flux) at each point in the domain, directly. However, effective parameters do not allow calculations of higher moments of the response, such as variance.

1.5.1 Dagan (1989)

[Much of the following review is based on lecture notes by Neuman (1992), derived in large part from Dagan (1989).]

Dagan considers uniform mean flow in a statistically homogeneous, isotropic, infinite log-normal conductivity field subject to (uniform) mean gradient, \mathbf{J} . The perturbation method consists of expanding the head, h , in an asymptotic series (see Appendix D)

$$h = h^{(0)} + h^{(1)} + h^{(2)} + \dots \quad (1.40)$$

The flow equation $\nabla \cdot (K \nabla h) = 0$ is written in terms of $Y = \ln K$ as (Dagan, 1989, Eq. 3.3.7)

$$\nabla^2 h = -\nabla Y \cdot \nabla h. \quad (1.41)$$

The latter is solved iteratively according to (Dagan, 1989, Eq. 3.3.9)

$$\nabla^2 h^{(i+1)} = -\nabla Y \cdot \nabla h^{(i)} \quad (1.42)$$

with $h^{(-1)} \equiv 0$ and $i = -1, 0, 1, 2, \dots$. The zero order approximation, $h^{(0)}$, satisfies $\nabla^2 h^{(0)} = 0$ and prescribed *deterministic* boundary conditions. The latter are chosen such that $\nabla h^0 = -\mathbf{J} = \text{constant}$. The head perturbations, $h^{(i)}$ for $i > 0$, satisfy corresponding *homogeneous* (zero) boundary conditions (because the boundary conditions are not random). As a result, $\langle \nabla h^{(i)} \rangle \equiv 0$ for all $i > 0$.

From Darcy's law, the flux is

$$\begin{aligned}\mathbf{q}(\mathbf{x}) &= -K(\mathbf{x})\nabla h(\mathbf{x}) = -e^{\langle Y \rangle + Y'}\nabla h(\mathbf{x}) = -K_G e^{Y'}\nabla h(\mathbf{x}) \\ &= -K_G \left[1 + Y' + \frac{(Y')^2}{2!} + \dots \right] [\nabla h^{(0)} + \nabla h^{(1)} + \nabla h^{(2)} + \dots] \quad (1.43)\end{aligned}$$

where $K_G = e^{\langle Y \rangle}$ is the geometric mean of K , and $Y' = Y - \langle Y \rangle$ are zero-mean random fluctuations in $Y(\mathbf{x})$. If the flux is expanded in the asymptotic series

$$\mathbf{q} = \mathbf{q}^{(0)} + \mathbf{q}^{(1)} + \mathbf{q}^{(2)} + \dots, \quad (1.44)$$

it is easy to see that

$$\begin{aligned}\mathbf{q}^{(0)} &= -K_G \nabla h^{(0)} \\ \mathbf{q}^{(1)} &= -K_G (Y' \nabla h^{(0)} + \nabla h^{(1)}) \\ \mathbf{q}^{(2)} &= -K_G \left(\frac{(Y')^2}{2} \nabla h^{(0)} + Y' \nabla h^{(1)} + \nabla h^{(2)} \right) \\ &\vdots\end{aligned} \quad (1.45)$$

If we further assume uniform $\langle Y \rangle$ and of $\langle Y' \nabla h^{(1)} \rangle$ in an infinite domain, take the ensemble mean of heads and fluxes in the above equations, and solve the flow equations for each head perturbation subject to the corresponding boundary conditions, we obtain, to first order in σ_Y^2 (where $\sigma_Y^2 = \langle Y'^2 \rangle$, the variance of Y)

$$\langle \mathbf{q} \rangle = \langle \mathbf{q}^{(0)} \rangle + \langle \mathbf{q}^{(1)} \rangle + \langle \mathbf{q}^{(2)} \rangle = K_G [(1 + \sigma_Y^2/2)\mathbf{J} - \alpha] \quad (1.46)$$

where

$$\alpha = \langle Y' \nabla h^{(1)} \rangle \quad (1.47)$$

To determine α , Dagan makes use of the *Green's function method* (e.g., Greenberg, 1971, and Fletcher, 1988), reviewed here briefly. Consider, for example, the operator

equation¹⁴

$$Lh + f = 0 \quad (1.48)$$

subject to certain boundary conditions, where L is a (deterministic) partial differential operator, f is a source term, and h is the response (head, in our case). This operator equation can be viewed as a general form of our (governing) flow equation, (1.1). A solution can be constructed, in principle, by “inverting” the operator L . The solution is then expressed in integral form as

$$L^{-1}[-f] = h(\mathbf{x}) = - \int_{\Omega} G(\mathbf{x}, \mathbf{x}') f(\mathbf{x}') d\mathbf{x}', \quad (1.49)$$

where the kernel (the fixed part) $G(\mathbf{x}, \mathbf{x}')$ is the *Green's function* (for simplicity, boundary integrals are omitted). The Green's function may be defined as head (or response) at \mathbf{x} produced by a point source of unit strength located at \mathbf{x}' ; hence, it is also called an *impulse response function* and *influence function*. Note that $h(\mathbf{x})$ is a convolution of the Green's function with the source term. In general $G(\mathbf{x}, \mathbf{x}')$ contains information equivalent to the operator L , the boundary conditions and the domain. When $G(\mathbf{x}, \mathbf{x}')$ does not need to satisfy boundary conditions, e.g., in an “infinite domain” ($\Omega \rightarrow \infty$), it is termed a *fundamental solution* or a *principal solution* or an *elementary solution* or a *free-space Green's function*. A major difficulty is to determine what the Green's function should be to suit a particular problem. The subsequent evaluation of (1.49) is usually straightforward. Green's functions can be obtained for relatively simple linear equations. For example, the Green's function corresponding to the Laplace and the Poisson equations satisfies

$$\nabla^2 G(\mathbf{x}, \mathbf{x}') + \delta(\mathbf{x} - \mathbf{x}') = 0 \quad (1.50)$$

¹⁴Operator theory is widely discussed in Chapter 2.

subject to (generally) homogeneous boundary conditions, where $\delta(\mathbf{x} - \mathbf{x}')$ is the Dirac delta centered at \mathbf{x}' , and ∇^2 is the Laplacian operator with respect to \mathbf{x} . The two and three dimensional fundamental solutions, i.e., Green's functions for unbounded (or infinite) domains are

$$G_\infty(\mathbf{x}, \mathbf{x}') = -\frac{1}{2\pi} \ln |\mathbf{x} - \mathbf{x}'| \quad 2 - D \quad (1.51)$$

$$G_\infty(\mathbf{x}, \mathbf{x}') = \frac{1}{4\pi |\mathbf{x} - \mathbf{x}'|} \quad 3 - D \quad (1.52)$$

Consider (1.42) with $i = 0$ (a zero order approximation):

$$\nabla^2 h^{(1)} = \nabla Y \cdot \mathbf{J}, \quad (1.53)$$

where \mathbf{J} is (again) the mean uniform gradient ($\nabla h^0 = -\mathbf{J} = \text{constant}$). This equation can be viewed as a Poisson equation, with $-f$ replaced by its right hand side; it can be then be “inverted” in the manner of (1.49), to yield

$$\begin{aligned} h^{(1)}(\mathbf{x}) &= -\int_{\Omega_\infty} G_\infty(\mathbf{x}, \mathbf{x}') \nabla_{\mathbf{x}'} Y(\mathbf{x}') \cdot \mathbf{J} d\mathbf{x}' \\ &= \int_{\Omega_\infty} \nabla_{\mathbf{x}'} G_\infty(\mathbf{x}, \mathbf{x}') Y(\mathbf{x}') \cdot \mathbf{J} d\mathbf{x}' \end{aligned} \quad (1.54)$$

where Green's first identity is used. By taking the gradient of $h^{(1)}$ in (1.54), multiplying by $Y'(\mathbf{x})$, and taking the ensemble mean, we obtain, to first order in σ_Y^2 ,

$$\alpha = \int_{\Omega_\infty} \langle Y'(\mathbf{x}) Y'(\mathbf{x}') \rangle \nabla \nabla_{\mathbf{x}'}^T G_\infty(\mathbf{x}, \mathbf{x}') \mathbf{J} d\mathbf{x}' \quad (1.55)$$

Note that G_∞ is the (deterministic) fundamental solution, of the Poisson equation for a *homogeneous* medium. As was pointed out by Neuman (1992), in the case of *bounded* domains the perturbation $h^{(1)}$, and hence $\langle Y' \nabla h^{(1)} \rangle$, are no longer stationary, and $\langle \mathbf{q}^{(2)} \rangle$ includes an additional term, $-K_G \nabla \langle h^{(2)} \rangle$. We recover this term in Chapter 3.

Note that, in general, $\langle \mathbf{q}(\mathbf{x}) \rangle$ in (1.47), considering (1.55), is non-local (i.e., depends on the gradients in the entire domain, not only at \mathbf{x}). Under uniform flow, however, $\mathbf{J} = \text{constant}$, and can be taken out of the integral. Then, from (1.47) and (1.55), an effective conductivity tensor can be defined to first order as

$$\begin{aligned}\mathbf{K}_e &\approx K_G \left(1 + \frac{\sigma_Y^2}{2}\right) \mathbf{I} - \tilde{\kappa}_\infty \\ \tilde{\kappa}_\infty &= K_G \int_{\Omega_\infty} C_Y(\mathbf{x}, \mathbf{x}') \nabla \nabla_x^T G_\infty(\mathbf{x}, \mathbf{x}') d\mathbf{x}'\end{aligned}\quad (1.56)$$

where \mathbf{I} is the identity tensor, and $C_Y(\mathbf{x}, \mathbf{x}') = \langle Y'(\mathbf{x})Y'(\mathbf{x}') \rangle$ is the (two-point) covariance of $Y = \ln K$.

By using Fourier transform (FT) and Parseval's theorem, Dagan (1989) obtains (his Eq. 3.4.13)

$$\tilde{\kappa}_\infty = \frac{K_G}{(2\pi)^{3/2}} \int \frac{\mathbf{k}\mathbf{k}^T}{k^2} \hat{C}_Y(-\mathbf{k}) d\mathbf{k}\quad (1.57)$$

where \mathbf{k} is the *wave number vector*, and \hat{C}_Y is the FT of the covariance of $Y(\mathbf{x})$, i.e., the *spectrum* (or spectral density) of $Y(\mathbf{x})$. Due to the symmetry of G_∞ , $\tilde{\kappa}_\infty$ is a symmetric, semi-positive definite, second rank tensor, and consequently, \mathbf{K}_e is a dyadic. For statistically anisotropic $Y(\mathbf{x})$ with elliptical geometry, Dagan's expression implies that the principal axes of \mathbf{K}_e coincide with the principal axes of statistical anisotropy. An expression similar to Eq. 1.57 has been obtained by Gutjahr et al. (1978; see below) by direct application of the FT to head perturbations, and later used by Gelhar and Axness (1983; see below). Neuman and Depner (1988) generalized it such that $\tilde{\kappa}_\infty$, and hence \mathbf{K}_e , do not depend on C_Y (or on the spectral density of Y), but only on σ_Y^2 and on the *ratios* between the integral scales in the principal directions (not on their values). In statistically isotropic fields $\tilde{\kappa}_\infty$ is reduced

to a scalar. For three dimensional media (1.57), and consequently (1.56), reduce to

$$\begin{aligned}\tilde{\kappa}_\infty &= K_G \sigma_Y^2 / 3 & 3 - D \\ K_e &\approx K_G (1 + \sigma_Y^2 / 6),\end{aligned}\tag{1.58}$$

and for two dimensional media $\tilde{\kappa}_\infty$ and \mathbf{K}_e reduce to

$$\begin{aligned}\tilde{\kappa}_\infty &= K_G \sigma_Y^2 / 2 & 2 - D \\ K_e &\approx K_G\end{aligned}\tag{1.59}$$

(cf. Neuman and Depner, 1988; Gutjahr et al., 1978.) For perfect horizontal stratification, the principal components of the effective conductivity, parallel and transverse to the strata become

$$\begin{aligned}K_{e_h} &\approx K_G (1 + \sigma_Y^2 / 2) \\ K_{e_v} &\approx K_G (1 - \sigma_Y^2 / 2).\end{aligned}\tag{1.60}$$

The latter are first order approximations of the arithmetic and harmonic means (the upper and lower bounds) of the log-normal K field. They are free of any assumptions about the *pdf* of the conductivity field. When K is log-normal, one may view $K_e = K_G (1 + \sigma_Y^2 / 6)$ for $\sigma_Y^2 \ll 1$ as the first two terms in the expansion of $K_e = K_G \exp(\sigma_Y^2 / 6)$ for any σ_Y^2 . This (latter) extrapolation was postulated by Landau and Lifshitz (1960), Matheron (1967), and Gelhar and Axness (1983), and was shown to hold for 3-D isotropic random fields (as mentioned in the previous section and in Chapter 3). Neuman and Depner (1988) validated the above extrapolation with single-hole and cross-hole tests, for variances of up to $\sigma_Y^2 = 7$; this was also verified numerically and independently by Neuman et al. (1992), Dykaar and Kitanidis (1992), Desbarats (1992a), and Ababou et al. (1989).

1.5.2 Shvidler (1962) and Matheron (1967)

[The following review is based mainly on Dupuy (1967). Matheron's and Shvidler's works were written in French and Russian, and the review by Dupuy is the only comprehensive summary available in English. Unfortunately, some of the expressions in Dupuy's review are inconsistent with his appendices and/or with Matheron (1967); in such (and other) cases we either translated Matheron's work directly, or adopted what seemed to be the correct expressions.]

Shvidler (1962) was perhaps the first to apply a stochastic (or random field) approach in hydrogeology. By using a perturbation method, Shvidler studied steady state flow in an infinite bounded, mildly heterogeneous domain. According to Matheron (1967, p. 125), Shvidler defined a local permeability tensor in a stationary and ergodic field,

$$\mathbf{K} = \langle K \rangle (\mathbf{I} + \varepsilon \check{\mathbf{K}}') \quad (1.61)$$

where \mathbf{I} is the identity tensor, $\check{\mathbf{K}}' = \mathbf{K}' / \langle K \rangle$ (a zero mean fluctuation term), and ε is a small perturbation parameter. For simplicity, Shvidler considers only a scalar K (i.e., local isotropy). Consequently, (1.61) becomes

$$K = \langle K \rangle (1 + \varepsilon \check{K}') \quad (1.62)$$

The perturbation expansions for head gradient and flux are

$$\begin{aligned} \nabla h(\mathbf{x}) &= \nabla h^{(0)}(\mathbf{x}) + \varepsilon \nabla h^{(1)}(\mathbf{x}) + \varepsilon^2 \nabla h^{(2)}(\mathbf{x}) + \dots \\ \mathbf{q}(\mathbf{x}) &= \mathbf{q}^{(0)}(\mathbf{x}) + \varepsilon \mathbf{q}^{(1)}(\mathbf{x}) + \varepsilon^2 \mathbf{q}^{(2)}(\mathbf{x}) + \dots \end{aligned} \quad (1.63)$$

From Darcy's law, $\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x})$, hence

$$\mathbf{q}^{(0)} = -\langle K \rangle \nabla h^{(0)}$$

$$\mathbf{q}^{(i)} = -\langle K \rangle [\nabla h^{(i)} + \check{K}' \nabla h^{(i-1)}], \quad (1.64)$$

and from the continuity equation, $\nabla \cdot \mathbf{q}(\mathbf{x}) = 0$,

$$\begin{aligned} \nabla^2 h^{(0)} &= 0 \\ \nabla^2 h^{(i)} + \nabla \cdot [\check{K}' \nabla h^{(i-1)}] &= 0. \end{aligned} \quad (1.65)$$

Upon assuming that the correlation scale is much smaller than the field dimensions, Shvidler postulated that the field boundaries are “of no consequence” (Dupuy, 1967, pp. 4-5), which is like assuming an infinite domain. Like Dagan (1989; previous section), Shvidler uses the Green’s function method to invert (1.65), such that

$$h^{(i)} = - \int_{\Omega_\infty} G_\infty(\mathbf{x}, \mathbf{x}') \nabla \cdot [\check{K}' \nabla h^{(i-1)}] d\mathbf{x}', \quad (1.66)$$

and by Green’s first identity, (Matheron, 1967, Eq. 21.6)

$$\nabla h^{(i)} = - \int_{\Omega_\infty} \nabla \nabla_{\mathbf{x}'}^T G_\infty(\mathbf{x}, \mathbf{x}') \check{K}' \nabla h^{(i-1)} d\mathbf{x}', \quad (1.67)$$

where $G_\infty(\mathbf{x}, \mathbf{x}')$ is the fundamental solution of the Poisson equation. Matheron’s review implies that G_∞ is a *deterministic* function (most probably associated with the uniform permeability $\langle K \rangle = \text{constant}$). Next, Matheron (1967, p. 127) equates

$$\int \nabla \nabla_{\mathbf{x}'}^T G(\mathbf{x}, \mathbf{x}') C d\mathbf{x}' = -\frac{1}{n} \mathbf{I} C, \quad (1.68)$$

(n being the number of spatial dimensions) implying that $\nabla \nabla_{\mathbf{x}'}^T G = \mathbf{I} \nabla^2 G$. By assuming $\langle \check{K}' \nabla h^{(i-1)} \rangle = \text{constant}$, Matheron (or Shvidler) obtains

$$\langle \nabla h^{(i)} \rangle = \frac{1}{n} \mathbf{I} \langle \check{K}' \nabla h^{(i-1)} \rangle. \quad (1.69)$$

Consequently, the average flux becomes

$$\begin{aligned} \langle \mathbf{q}^{(i)} \rangle &= -\langle K \rangle [\mathbf{I} \langle \nabla h^{(i)} \rangle + \langle \check{K}' \nabla h^{(i-1)} \rangle] \\ &= -\langle K \rangle (n-1) \mathbf{I} \langle \nabla h^{(i)} \rangle. \end{aligned} \quad (1.70)$$

According Matheron (1967, p. 127), second order approximation in ε implies

$$\langle \mathbf{q}^{(1)} \rangle = \langle \nabla h^{(1)} \rangle = 0. \quad (1.71)$$

By taking the ensemble average of (1.67), and considering (1.69)–(1.71) for $i = 1, 2$, Matheron arrives at *Shvidler's tensor*,

$$\mathbf{S} = -\langle K \rangle^{-2} \int_{\Omega_\infty} \nabla \nabla_{\mathbf{x}'}^T G_\infty(\mathbf{x}, \mathbf{x}') C_K(\mathbf{x}, \mathbf{x}') d\mathbf{x}' \quad (1.72)$$

where $C_K(\mathbf{x}, \mathbf{x}')$ is the covariance of K . [Note the similarity between Shvidler's tensor and $\tilde{\kappa}_\infty$ in (1.56)]. Consequently, he obtains the following a second order approximation (in ε) for effective conductivity:

$$\mathbf{K}_e \approx \langle K \rangle (\mathbf{I} - \varepsilon^2 \mathbf{S}) \quad (1.73)$$

which is very similar to Dagan's \mathbf{K}_e in (1.56). Likewise, \mathbf{S} in (1.72) is a symmetric positive definite tensor.

Matheron (1967, p. 130) refers to a “subisotropic” case (which Dupuy, 1967 considers as *isotropic*), where $C_K \rightarrow C_K(0) = \sigma_K^2$, and $\nabla \nabla_{\mathbf{x}'}^T G = \nabla^2 G$ [cf. (1.68)]. Consequently, \mathbf{S} reduces to

$$S = \frac{1}{n} \frac{\sigma_K^2}{\langle K \rangle^2} \quad (1.74)$$

where σ_K^2 is the variance of K . Consequently, (1.73) reduces to

$$K_e \approx \langle K \rangle \left(1 - \frac{\varepsilon^2}{n} \frac{\sigma_K^2}{\langle K \rangle^2} \right) \quad (1.75)$$

From Appendix F, for log-normally distributed K ,

$$C_K(\mathbf{x}, \mathbf{x}') = K_G(\mathbf{x}) K_G(\mathbf{x}') \sigma_Y^2 [e^{C_Y(\mathbf{x}, \mathbf{x}')} - 1]. \quad (1.76)$$

If we further approximate

$$\langle K \rangle = K_G e^{\sigma_Y^2/2} \approx K_G (1 + \sigma_Y^2/2) \quad (1.77)$$

and, to first order, $e^{C_Y} \approx 1 + C_Y$, we note that Shvidler's expressions (1.72)–(1.73) approximate Dagan's expression for effective conductivity, (1.56), for $\sigma_Y^2 \ll 1$. As will be seen in the following section, under *nonuniform* flow conditions, Shvidler-Matheron's expressions for mean flow rate (Eqs. 1.102– 1.105 in the following section) are also similar to those obtained by Dagan for mean flux (i.e., Eq. 1.56 above, and Eq. 1.111 in Section 1.5.4 that follows). In Chapter 3 we use a different approach and obtain an exact formal expression for the mean flux $\langle \mathbf{q} \rangle$. After using a weak (closure) approximation, we obtain expressions for mean flux and effective conductivity, which are reminiscent of the approximations of Dagan and Shvidler-Matheron.

Matheron (1967) proves that in two dimensional statistically isotropic and homogeneous, log-normal K (and K^{-1}) fields under uniform flow, $K_e = K_G$ exactly, for any variance (cf. Eq. 1.59, and Appendix G). Following Shvidler (1962), Matheron (1967) makes the general conjecture that for lognormally distributed K in an n dimensional statistically isotropic and homogeneous medium, the effective conductivity is

$$K_e = K_G \exp \left[\sigma_Y^2 \left(\frac{1}{2} - \frac{1}{n} \right) \right] \quad (1.78)$$

where $K_A = K_G \exp(\sigma_Y^2/2)$ is the arithmetic mean, and $K_H = K_G \exp(-\sigma_Y^2/2)$ is the harmonic mean. The validity of (1.78) has been established by several independent studies and numerical experiments, including our work (see Chapter 3). The expression for $n = 3$ was conjectured earlier by Landau and Lifshitz (1960).

Another interesting result from the Shvidler–Matheron analysis is the so called *variance reduction factor* (VRF) for total flow through a (general) two dimensional

bounded domain is (Dupuy, 1967)

$$VRF_Q = \frac{(\sigma'_Q)^2}{(\sigma'_K)^2} = \frac{1}{A^2} \int_{\Omega} \rho_K(\mathbf{x} - \mathbf{x}') \mathbf{W}_2(\mathbf{x}, \mathbf{x}') d\mathbf{x}', \quad (1.79)$$

where Q is the flow rate through the entire domain, Ω , and $(\sigma'_Q)^2 = \sigma_Q^2 / \langle Q \rangle^2$, and $(\sigma'_K)^2 = \sigma_K^2 / \langle K \rangle^2$ are the normalized variances. A is the area of the (entire) domain, defined by Dupuy both as a “spatial measure”, and as $A = \int_{\Omega} d\mathbf{x}$. $\rho_K(\mathbf{x} - \mathbf{x}') = C_K(\mathbf{x} - \mathbf{x}') / \sigma_K^2$ is the autocorrelation function, or tensor. [Dupuy, inconsistently uses the same notation $K(x, y)$ for both the autocorrelation function ρ_K and the covariance C_K in the expressions for VRF_Q , particularly, his Eq. 26 and his Appendix B.] If the local permeabilities are anisotropic (second order tensors), then \mathbf{C}_K is a fourth order tensor. \mathbf{W}_2 is a “weighting factor” such that

$$\mathbf{W}_2(\mathbf{x}, \mathbf{x}') = W_1(\mathbf{x})W_1(\mathbf{x}') \quad (1.80)$$

where $W_1(\mathbf{x})$ is the *normalized* ∇h_o^2 , i.e.,

$$W_1(\mathbf{x}) = \nabla h_o^2(\mathbf{x}) \left[\int_{\Omega} \nabla h_o^2(\mathbf{x}) \right]^{-1} d\mathbf{x} \quad (1.81)$$

where h_o is the zero order (“unperturbed”) solution. Consequently, the integral in (1.79) is always smaller than one. Note that in an infinite domain, $VRF_Q \rightarrow 0$, which leaves the merit of (1.79) questionable¹⁵. Note also that W_1 , and consequently W_2 and VRF_Q are *non-local* (cf. next section). In the case of mean uniform flow, the unperturbed pressure gradient is constant everywhere, and the VRF_Q is reduced to a constant multiplied by the integral of the covariance of K . Indeed, for a rectangular domain with no-flow boundaries along the mean flow direction and prescribed

¹⁵Eq. 1.79 is taken from Dupuy’s appendix; in the text, it appears without the factor $1/A^2$ with no justifiable explanation.

heads on the transverse boundaries (i.e., upstream and downstream boundaries) the variance reduction factor becomes (Dupuy, 1967)

$$VRF_Q = \frac{1}{A^2} \int_{\Omega} \rho_K(\mathbf{x} - \mathbf{x}') d\mathbf{x} \quad (1.82)$$

i.e., VRF_Q becomes the spatial average of the correlation function, and therefore, it is directly related to the integral scale (cf. Appendix A; for a 1-D infinite domain the remaining integral is identical to the integral scale).

In the case of *radial flow* ∇h_o^2 is inversely proportional to r , and Shvidler's weighting factor in (1.79) becomes¹⁶

$$\mathbf{W}_2(\mathbf{x}, \mathbf{x}') = \frac{1}{r(\mathbf{x})r(\mathbf{x}')} \quad (1.83)$$

Dupuy (1967) shows how Shvidler's VRF_Q changes dramatically (20 fold) from uniform to radial flow. More about radial flow in Section 1.5.4. Shvidler's emphasis on total discharge rather than specific discharge apparently stems from the practical concern of uncertainty in total discharge from a well penetrating a heterogeneous formation.

1.5.3 Gutjahr et al.(1978), Bakr et al.(1978), Gelhar and Axness (1983), Gelhar (1986)

Gutjahr et al. (1978), Bakr et al. (1978), and Gelhar and Axness (1983) consider mean uniform flow in a three dimensional aquifer characterized by a log-normally distributed hydraulic conductivity field, which is locally isotropic (i.e., K is a scalar), with a known covariance function of $Y = \ln K$, which may be (statistically)

¹⁶However, comparing it to Matheron's (1967) Eq. 24.3, we suspect that Dupuy exchanged Q_o with $\langle Q \rangle$.

anisotropic. As was mentioned earlier, if Y is Gaussian (i.e., normal), then the entire statistical structure is exhausted by the first two moments, the mean $\langle Y \rangle$ and the covariance $C_Y(\mathbf{x}, \mathbf{x}') = \langle Y(\mathbf{x})Y(\mathbf{x}') \rangle - \langle Y(\mathbf{x}) \rangle \langle Y(\mathbf{x}') \rangle$. If the field is stationary (or statistically homogeneous) then $C_Y(\mathbf{x}, \mathbf{x}') = C_Y(\mathbf{d})$, where $\mathbf{d} = \mathbf{x} - \mathbf{x}'$.

In order to solve the stochastic, steady state flow equation, Gelhar and Axness (1983) start from (1.41), i.e.,

$$\nabla \cdot \mathbf{q}(\mathbf{x}) = \nabla^2 h + \nabla Y \cdot \nabla h = 0. \quad (1.84)$$

Next, they decompose both input and output into ensemble mean and zero-mean fluctuations (i.e., $Y = \langle Y \rangle + Y'$, and $h = \langle h \rangle + h'$, such that $\langle Y' \rangle = \langle h' \rangle = 0$), substitute it in (1.84), take the ensemble average, and get

$$\nabla^2 \langle h \rangle + \nabla \langle Y \rangle \cdot \nabla \langle h \rangle + \langle \nabla Y' \cdot \nabla h' \rangle = 0. \quad (1.85)$$

By subtracting this from (1.84), they obtain

$$\nabla^2 h' + \nabla \langle Y \rangle \cdot \nabla h' + \nabla Y' \cdot \nabla \langle h \rangle = -[\nabla Y' \cdot \nabla h' - \langle \nabla Y' \cdot \nabla h' \rangle]. \quad (1.86)$$

Assuming that the variance of Y is small (i.e., $\sigma_Y^2 < 0.25$), and that the fluctuations are slow, they neglect the entire right hand side of (1.86), to yield their “first order” approximations:

$$\nabla^2 h' + \nabla \langle Y \rangle \cdot \nabla h' + \nabla Y' \cdot \nabla \langle h \rangle \approx 0. \quad (1.87)$$

[This neglect is counter intuitive because derivatives of (relatively) high frequency fluctuations are expected to be larger than derivatives of slowly varying averages (which are not neglected); moreover, the neglect of $\langle \nabla Y' \cdot \nabla h' \rangle$ implies no correlation between gradients of hydraulic conductivity and head fluctuations, which, in turn,

implies no cross-correlation between input and output perturbations.] Further, assuming that Y is (second order) stationary (or statistically homogeneous), $\nabla\langle Y \rangle = 0$, and (1.87) is reduced to

$$\nabla^2 h' + \nabla Y' \cdot \nabla \langle h \rangle \approx 0, \quad (1.88)$$

which is a *linearized* form of the flow equations (because h' depends linearly on $\nabla Y'$). Note the similarity between (1.88) and the perturbation expansion (1.42), which for $i = 0$ becomes

$$\nabla^2 h^{(1)} = -\nabla Y \cdot \nabla h^{(0)}. \quad (1.89)$$

Viewing the (above) decomposition of h as $h = \langle h \rangle + h' \approx h^{(0)} + h^{(1)}$, and recalling that $\nabla\langle Y \rangle = 0$, one may conclude that by neglecting products of derivatives of fluctuations, Gelhar and his co-workers obtained a first order approximation in σ_Y , expressed by Eq. 1.88. On the other hand, as will be shown in the sequel, their resulting effective conductivity is identical to that obtained by Dagan (above) through consistent second order perturbation expansion. Neuman et al. (1987) show that the derivation of Gelhar and Axness (1983) of the macroscopic transport equation “is inconsistent in the sense that its derivation involves a first-order approximation of the mean velocity at one stage, and a second-order approximation at another stage”, and therefore “it is consistent neither with a first order, nor with a second order approach”. In a recent review by Ababou and Gelhar (1990) they reiterate (1.85) and (1.86) for stationary Y fields (instead of Eq. 1.88) and obtain

$$\nabla^2 \langle h \rangle = -\langle \nabla Y' \cdot \nabla h' \rangle \quad (1.90)$$

$$\nabla^2 h' + \nabla Y' \cdot \nabla \langle h \rangle = -\nabla Y' \cdot \nabla h' + \langle \nabla Y' \cdot \nabla h' \rangle \quad (1.91)$$

However, they notice that the above equations constitute a system of two (stochastic) equations with three unknowns ($\langle h \rangle$, h' , and $\langle \nabla Y' \cdot \nabla h' \rangle$); consequently they concluded that “a complete solution of this problem would require solving an infinite hierarchy of equations governing higher order moments...which turns out to be an impossible task in the general case” (*ibid*). This is known to be the standard closure problem (Neuman, personal communication).

Based on the assumption that h' is second order stationary, Gutjahr et al. (1978) and Bakr et al. (1978) solve (1.88) in Fourier space, by using spectral representations of $Y'(\mathbf{x})$ and $h'(\mathbf{x})$ in terms of Fourier-Stieltjes integrals (discussed in Chapter 4). Consequently, they find relationships between variances of heads and hydraulic conductivity, as functions of the correlation scale and the mean head gradient:

$$\text{1-D} \quad VRF_{h_{2D}} = \frac{\sigma_h^2}{\sigma_Y^2} \approx \mathbf{J}^2 \lambda^2 \quad (1.92)$$

$$\text{3-D} \quad VRF_{h_{3D}} = \frac{\sigma_h^2}{\sigma_Y^2} \approx \frac{\mathbf{J}^2 \lambda^2}{3} \quad (1.93)$$

where, here, $\mathbf{J} \equiv -\nabla \langle h \rangle$ (the mean head gradient). λ is the correlation length (or integral scale) of Y , and σ_h^2 is the variance of the head. Equations (1.92), (1.93) show, quantitatively, the drop of head variance compared to the variance of the log-conductivity, as a function of \mathbf{J} , λ , and the dimensionality of the model used. Note the similarity between these equations and Eq. 1.79 of Shvidler (1962). Both expressions show strong dependency of the variance reduction on the integral scale and on the average head (or pressure) gradient. Since in Shvidler's case the reduction is in the variance of the total flow rate through the domain (or block) rather than (local) head variance, the latter is diminished more strongly with respect to the conductivity variance. Similar conclusions were reached on the basis of Monte Carlo

simulations by Smith and Freeze (1979).

It is worth mentioning that Mizell et al. (1982; see also Ababou and Wood, 1990) show that unless a *hole covariance* (which is characterized, in part, by a negative covariance) is used for a two dimensional stationary *unbounded* domain, the variance of the head field becomes infinite due to artificial divergence of the spectrum at low wave numbers, known as an *infrared divergence*. Ababou and Wood (1990) explain this phenomenon by the (artificial) reduction in the degree of freedom from three to two dimensions.

In order to determine an effective hydraulic conductivity for an infinite second order stationary field under uniform mean flow, Gutjahr et al. (1978) expand $\exp(Y')$ in Taylor series, as in (1.43), and obtain a second order approximation for the flux (first order in σ_Y^2):

$$\begin{aligned} \mathbf{q}(\mathbf{x}) &= -K(\mathbf{x})\nabla h(\mathbf{x}) = -e^{(Y)+Y'}\nabla h(\mathbf{x}) = -K_G e^{Y'}\nabla h(\mathbf{x}) \\ &\approx -K_G \left[1 + Y' + \frac{(Y')^2}{2!} \right] [\nabla\langle h \rangle + \nabla h'] \end{aligned} \quad (1.94)$$

Hence,

$$\langle \mathbf{q} \rangle \approx K_G \left[\mathbf{J}(1 + \sigma_Y^2/2) - \langle Y'\nabla h' \rangle \right] \quad (1.95)$$

Rather than neglecting products of fluctuations (or their derivatives), particularly $\langle Y'\nabla h' \rangle$ in (1.95) above, the authors evaluate the latter, using the spectral representation theorem (presented in Chapter 4). They find that for a two-dimensional statistically homogeneous and isotropic Y field

$$\langle Y'\nabla h' \rangle = \mathbf{J}\sigma_Y^2/2. \quad (1.96)$$

Consequently, $\langle \mathbf{q} \rangle = K_G \mathbf{J}$, i.e., the effective conductivity in 2-D is K_G . A similar analysis for 1-D and 3-D yields the familiar approximations for effective hydraulic conductivities of n -dimensional fields under uniform mean flow [cf. (1.59), (1.58)]:

$$K_e \approx K_g \left(1 - \frac{\sigma_Y^2}{2} \right) \quad n = 1 \quad (1.97)$$

$$K_e \approx K_g \quad n = 2 \quad (1.98)$$

$$K_e \approx K_g \left(1 + \frac{\sigma_Y^2}{6} \right) \quad n = 3 \quad (1.99)$$

Gelhar and Axness (1983) later extrapolate the expressions for the (principal) effective hydraulic conductivities (particularly Eq. 1.99) to large σ_Y^2 , by viewing them as the first two terms in the expansion of Eq. 1.78 (Matheron's conjecture). They also extend the above formulae to statistically anisotropic media, and obtain an effective hydraulic conductivity tensor of the form

$$\mathbf{K}_e = K_G (1 + \sigma_Y^2/2 - \mathbf{F}) \quad (1.100)$$

where $K_G \mathbf{F} \approx \tilde{\kappa}_\infty$ in Eq. 1.57 above, such that (1.100) is identical to (1.56). Note that the derivation of \mathbf{F} by Gutjahr et al. (1978) in (1.100) leads to Eq. 1.57 without resorting to Green's function.

Perturbation methods are limited to relatively simple cases, e.g., infinite (unbounded) domains, and stationary fields. Since convergence of the series resulting from perturbation methods cannot be guaranteed, neither does the validity of the corresponding solutions (Beran, 1968; Dagan, 1989; Ghanem and Spanos, 1991). As the analysis has not been carried beyond second order, it is restricted to mildly heterogeneous media (small σ_Y^2). The accuracy of the truncation (e.g., to 2nd and 1st order approximations) as a function of the variance of the perturbations can be implied, for example, from Figure 1 of Neuman and Orr (1983; see also Chapter 3), and

from Figure 8 of Dykaar and Kitanidis (1992), where expressions for K_e obtained from numerical 3-D MCS are compared with linear and exponential expressions; the figures indicate an exponential departure of K_e from its linear approximation, beyond $\sigma_Y^2 \approx 1$.

Perturbation methods become more questionable when dealing with transient (time dependent) problems. In addition to the limited variability inherent in these methods, perturbation expansions of time-dependent problems include powers of time (called *secular terms*) which dominate higher order terms, and thereby render them non-asymptotic (this will become more clear in the following chapter). Frisch (1968) brings two simple examples with perturbation expansions that are either divergent or converge very slowly for large time values. Nevertheless (as is demonstrated throughout this chapter), perturbation methods can provide insight, and with some intuition may be extrapolated, to capture the behavior of systems with highly variable parameters and non-uniform flows (e.g., see next section).

1.5.4 Nonuniform Mean Flow

Under nonuniform mean flow $\mathbf{J}(\mathbf{x}) = -\nabla\langle h(\mathbf{x})\rangle$ is no longer constant in space. Consequently, neither upscaled nor point effective hydraulic conductivity tensors (defined above) may exist.

Shvidler (1962) and Matheron (1967)

According to Matheron (1967), Shvidler (1962) defines the total *energy* dissipated within the domain Ω under nonuniform mean flow as

$$QH_o = - \int_{\Omega} \mathbf{q} \nabla h d\mathbf{x}, \quad (1.101)$$

where Q is the (total) flow rate through the domain Ω (and consequently, through the boundaries), and $-H_o$ is a deterministic prescribed head on the boundary downstream, while the (deterministic) head prescribed on the upstream boundary is zero. Shvidler's perturbation expansion, and (involved) derivation (along the line of Section 1.5.2) leads to (Matheron, 1967, Eq. 23.4)

$$\langle QH_o \rangle = Q_o H_o - \langle K \rangle \varepsilon^2 T_2, \quad (1.102)$$

where Q_o is the total flow corresponding to a uniform medium with hydraulic conductivity $\langle K \rangle$ (i.e., a zero order approximation; see Matheron, 1967, p. 140),

$$T_2 = \int_{\Omega} \int_{\Omega} C_K(\mathbf{x}, \mathbf{x}') \nabla h^{(0)}(\mathbf{x}) \nabla h^{(0)}(\mathbf{x}') \nabla \nabla_{\mathbf{x}}^T G_{\infty}(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}', \quad (1.103)$$

and

$$Q_o H_o = \langle K \rangle \int_{\Omega} \nabla [h^{(0)}]^2 d\mathbf{x}, \quad (1.104)$$

The form of equations (1.102)–(1.103) is similar to that of equations (1.47) and (1.55), particularly after dividing the formers by the constant H_o . Likewise, T_2 in (1.103) is (also) nonlocal. Moreover, for radial flow between two concentric circles with radii R_o and R_1 , Eq. 1.102 reduces to (Matheron, 1967, p. 143; radial coordinate system)

$$\left\langle \frac{Q}{Q_o} \right\rangle = 1 - \frac{\varepsilon^2}{\log \frac{R_1}{R_o}} \int_{R_o}^{R_1} dr \int_{R_o}^{R_1} dr' \int_0^{2\pi} C_K(d) \frac{\partial^2 G(\mathbf{x}, \mathbf{x}')}{\partial r \partial r'} d\theta, \quad (1.105)$$

where $d = \sqrt{r^2 + r'^2 - 2rr' \cos \theta}$.

By accounting for additional terms in the perturbation expansion of Shvidler (1962) and by carrying it further, Matheron (1967) examined the existence of effective conductivity of a two dimensional statistically homogeneous and isotropic porous medium under radial flow conditions around a pumping well (as mentioned in Section

1.5). He assumed constant head at the well ($r = r_w$) and zero drawdown at “large” distance $r = r_1 \rightarrow \infty$. [These are rather problematic boundary conditions in two-dimensions, in view of the corresponding Green’s function ($G_\infty = -(1/2\pi)\ln r$), which approach infinity as $r \rightarrow \infty$; this, as well as the following results, will be discussed in conjunction with Naff’s analysis of radial flow.] Consequently, Matheron concluded that “if the flow is not uniform (e.g., converging radial) then there is no law of composition, constant in time, that makes it possible to define mean Darcian permeability” (de-Marsily, 1986). However, for some extreme cases, Matheron (1967) found that (Ababou and Wood, 1990)

$$K_e \rightarrow K_H \quad \text{for} \quad r/\lambda \rightarrow \infty \quad (1.106)$$

$$K_e \rightarrow K_A \quad \text{for} \quad r/\lambda \rightarrow 0 \quad (1.107)$$

for the unconditional case, while for the conditional case with $K = K_w$ at the well (where w denotes “well”), he found

$$K_e \rightarrow K_H \quad \text{for} \quad r/\lambda \rightarrow \infty \quad (1.108)$$

$$K_e \rightarrow K_w \quad \text{for} \quad r/\lambda \rightarrow 0 \quad (1.109)$$

Dagan (1989)

Following an intuitive reasoning by Saffman (1971), Dagan (1989) assumes that

$$\langle \mathbf{q}(\mathbf{x}) \rangle = \int_{\Omega} \mathbf{W}(\mathbf{x}, \mathbf{x}') \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle d\mathbf{x}' \quad (1.110)$$

where \mathbf{W} is an unspecified kernel, depending on the statistical properties of $Y = \ln K$, the choice of coordinates, and the shape of the domain Ω . Note that $\langle \mathbf{q} \rangle(\mathbf{x})$ is *non-local*, because it depends on the mean head gradient not only at \mathbf{x} but all over the

domain. Dagan further postulates that the kernel W tends to zero as the distance $|\mathbf{x} - \mathbf{x}'| > \lambda$ (i.e., beyond the integral scale of Y). Assuming further that $\langle h \rangle$ is continuous and varies slowly at the λ scale, Dagan expands it in Taylor series, and obtains the mean flux as a series of decreasing order which, though local, depends not only on the mean head gradient, but on its derivatives as well. In order to reach a more compact form; Dagan further assumes an unbounded domain, stationary Y field, and small σ_Y^2 . Recalling equations (1.47) and (1.55), which are based on the small perturbation technique, and are valid for non-uniform flows when $\mathbf{J} = \mathbf{J}(\mathbf{x})$ remains under the integral sign, Dagan reconstructs a first order approximation [$O(\sigma_Y^2)$] of the kernel \mathbf{W} :

$$\mathbf{W}(\mathbf{x}, \mathbf{x}') = -K_G[(1 + \sigma_Y^2/2)\mathbf{I}\delta(\mathbf{x} - \mathbf{x}') - C_Y(\mathbf{x}, \mathbf{x}')\nabla\nabla_x^T G_\infty(\mathbf{x}, \mathbf{x}')]. \quad (1.111)$$

In Chapter 3, we will find the exact kernel, \mathbf{W} , for both bounded and unbounded domains.

Dagan's derivation shows that even in the general case of mean nonuniform flow, an effective conductivity can be defined for statistically homogenous formations, provided that the correlation scale of the response, e.g., λ_h , is much larger than the heterogeneity scale, λ_Y (the integral scale of the log hydraulic conductivity). In his subsequent development, Dagan (1989) shows that under radial flow, the error incurred to K_e is of the order of $\sigma_Y^2(\lambda_Y/r)^2$, where r is the distance from the well; hence, the error is negligible far from the well. Close to the well, however, the effective hydraulic conductivity is expected to be diminish (compared to K_e of uniform flow) because of streamline convergence. Indeed, Gomez-Hernandez and Gorelick (1989) find that the block effective conductivity in a bounded two dimensional unconfined aquifer with pumping wells is between the harmonic and the geometric mean of K .

The reduction in K_e towards the harmonic mean (K_H) close to the well is also evident in our MCS (Chapter 5). Under spherical or radial mean flow in a statistically homogeneous and isotropic medium, far from the well ($r \rightarrow \infty$), streamlines are almost parallel, and the mean gradient is almost uniform. Consequently, $\mathbf{J} \approx \text{constant}$ [and can be taken outside of the convolution integral in (1.110)], and K_e is similar to the effective conductivity under uniform mean flow (e.g., Eq. 1.56). Near the point source, as $r \rightarrow 0$, $K(x)$ is fully autocorrelated, and is effectively constant. In 3-D,

$$q(r) = -K \frac{\partial h}{\partial r} = -\frac{Q_o}{4\pi r^2} \quad (1.112)$$

where Q_o is a given, deterministic flow (or pumping) rate. Hence,

$$\frac{d\langle h \rangle}{dr} = \langle K^{-1} \rangle \frac{Q_o}{4\pi r^2} \quad (1.113)$$

and

$$\langle q(r) \rangle = -\langle K^{-1} \rangle^{-1} \frac{d\langle h(r) \rangle}{dr} \quad (1.114)$$

where $\langle K^{-1} \rangle^{-1} \equiv K_H$, the harmonic mean. A similar analysis for 2-D shows that the effective transmissivity at the well (perimeter) is equal to the *harmonic mean* (which is smaller than the geometric mean), such that at the limit $T_e \rightarrow T_h$ when $\lambda_Y/r_w \rightarrow \infty$. This lower bound is consistent with the classical expression by Craft and Hawkins (1959; see also, Desbarats, 1992b) for effective (block) transmissivity in a circular, radially symmetric flow field, with well radius r_w and external radius r_e :

$$T_e = \ln \frac{r_e}{r_w} \left(\int_{r_w}^{r_e} \frac{1}{T(r)r} dr \right)^{-1}. \quad (1.115)$$

As was mentioned above and will be shown in detail later, the results of Dagan's first order analysis are consistent with both the theory and our numerical Monte Carlo simulations (discussed in Chapter 5), as well as the findings of Desbarats

(1992b; discussed below), but in contrast with the findings of Matheron (1967) and Naff (1991).

Naff (1991)

Naff (1991) uses a perturbation analysis, along the line of Gutjahr et al. (1978), to solve for 2-D and quasi 3-D radial flow in a statistically homogeneous and isotropic porous medium. In particular, he approximates first and second moments of head (gradient) and specific discharge in the radial direction.

Following the traditional simplified small perturbations approach of the early works of Gelhar and co-workers, as described in Section 1.5.3, Naff starts from the local continuity equation $\nabla \cdot \mathbf{q} = 0$, substitutes $K = \exp[Y]$, and, arrives at Eq. 1.88 (above), rewritten here (his Eq. 5),

$$\nabla^2 h' + \nabla Y' \cdot \nabla \langle h \rangle = 0, \quad (1.116)$$

where Y' and h' are zero mean fluctuations of h and $Y = \ln K$. To reach this equation, Naff neglects products of perturbations and perturbation gradients, particularly $\nabla h' \nabla Y'$. Consequently, there is no assurance that (1.116) is indeed a “first order approximation” in σ_Y^2 (see Neuman et al., 1987).

In his quasi 3-D analysis, Naff assigns a zero (constant) mean head at $r = \lambda$, (i.e., at $\rho \equiv r/\lambda = 1$) where λ is the horizontal correlation scale. In addition, he assumes stationarity of both conductivity fluctuations Y' and head fluctuations h' in the vertical direction, z , and in (all) angular directions θ . Next, Naff introduces a “zero order approximation” of the head gradient (which he equates with the mean head gradient)

$$\frac{\partial \langle h \rangle}{\partial r} \simeq \frac{Q_o}{2\pi r K_G} \quad (1.117)$$

(Naff's Eq. 10) which implies

$$\langle h \rangle \approx Q_o \frac{\ln(r/\lambda)}{2\pi K_G} \quad (1.118)$$

where Q_o is the volumetric discharge per unit length of well, and K_G is the geometric mean of K . The employment of a zero order approximation (where $\sigma_Y^2 = 0$) in a higher order analysis, poses substantial difficulties. For example, based on (1.117), Naff concludes that "at this level of approximation, the geometric mean becomes the effective hydraulic conductivity". However, if $\sigma_Y^2 = 0$, then $\langle K \rangle = K_G = K_e$, and (1.117) is equivalent to

$$\frac{\partial \langle h \rangle}{\partial r} \simeq \frac{Q_o}{2\pi r \langle K \rangle}. \quad (1.119)$$

Moreover, under log-normality of K , $K_G \exp(\sigma_Y^2/2) = \langle K \rangle$, and Naff's Eq. 6, $q_r = -K \partial h / \partial r$, leads to his Eq. 7, i.e.,

$$\begin{aligned} \langle q_r \rangle &= -\langle K \rangle \frac{\partial \langle h \rangle}{\partial r} - \left\langle K' \frac{\partial h'}{\partial r} \right\rangle \\ &= -K_G \left\{ \exp(\sigma_Y^2/2) \frac{\partial \langle h \rangle}{\partial r} + \left\langle e^{Y'} \frac{\partial h'}{\partial r} \right\rangle \right\}, \end{aligned} \quad (1.120)$$

which implies that $\langle K \rangle$ is the zero order approximation, not K_G . Naff combines (1.116) and (1.117) to yield a random continuity equation in cylindrical coordinates (Naff's Eq. 11),

$$\frac{\partial}{\partial r} \left[r \frac{\partial h'}{\partial r} \right] + \frac{1}{r} \frac{\partial^2 h}{\partial \theta^2} + r \frac{\partial^2 h}{\partial z^2} = -\frac{Q_o}{2\pi K_G} \frac{\partial Y'}{\partial r}, \quad (1.121)$$

where the left hand side (LHS) is identical to the first term in (1.116) multiplied by r , and the right hand side (RHS) is the second term in (1.116) multiplied by r (and taken to the RHS). While the first term of (1.116) (the LHS of Eq. 1.121) is taken in full cylindrical coordinate system, the second term in (1.116) (the RHS of Eq. 1.121) is given in radial coordinates. In other words, vertical and angular components of

conductivity fluctuations are neglected, despite analyzing *vertical* components of the *response* (i.e., heads and fluxes). This, together with the zero order approximation, renders the quasi-three dimensional analysis incomplete and inconsistent.

Naff uses the “zero order approximation”, Eq. 1.117 (his Eq. 10), (inconsistently) to calculate “first order” approximations of the first two moments of the specific discharge (his Eq. 31). For example, his Eq. 29,

$$K_e = K_G \left\{ 1 + \sigma_Y^2 \left[\frac{1}{2} + \rho\alpha(\rho) \right] \right\}, \quad (1.122)$$

[where $\rho\alpha(\rho)$ is given in Fig. 2 of Naff and Figure 1.3 below¹⁷] is obtained from his Eq. 27,

$$K_e = K_G \left\{ \exp(\sigma_Y^2/2) + \beta(\rho) \right\}, \quad (1.123)$$

where

$$\beta(\rho) \equiv \frac{\langle e^{Y'} (\partial h' / \partial r) \rangle}{\partial \langle h \rangle / \partial r}, \quad (1.124)$$

through truncation (down to the first two terms) of series expansions of both $\exp[\sigma_Y^2/2]$ (assuming $\sigma_Y^2 < 1$) and the denominator of his Eq. 28, i.e., the denominator of

$$\beta(\rho) = \frac{\exp(\sigma_Y^2) \rho\alpha(\rho)}{1 - \sigma_Y^2 \rho\alpha(\rho)}. \quad (1.125)$$

However, Eq. 1.122 can be obtained directly, without the need to truncate a series, upon dividing his Eq. 25,

$$\left\langle Y'(r) \frac{\partial h'}{\partial r} \right\rangle = \frac{Q_o}{2\pi K_G} \frac{\sigma_Y^2}{\lambda} \alpha(\rho), \quad (1.126)$$

by equation 1.117, to yield

$$\beta(\rho) = \sigma_Y^2 (r/\lambda) \alpha(\rho) \equiv \sigma_Y^2 \rho\alpha(\rho), \quad (1.127)$$

¹⁷The definition of the the “gradient factor”, $\alpha(\rho)$, is rather involved and is not essential for the following development; its role is best identified in Eqs. 1.122 and 1.129.

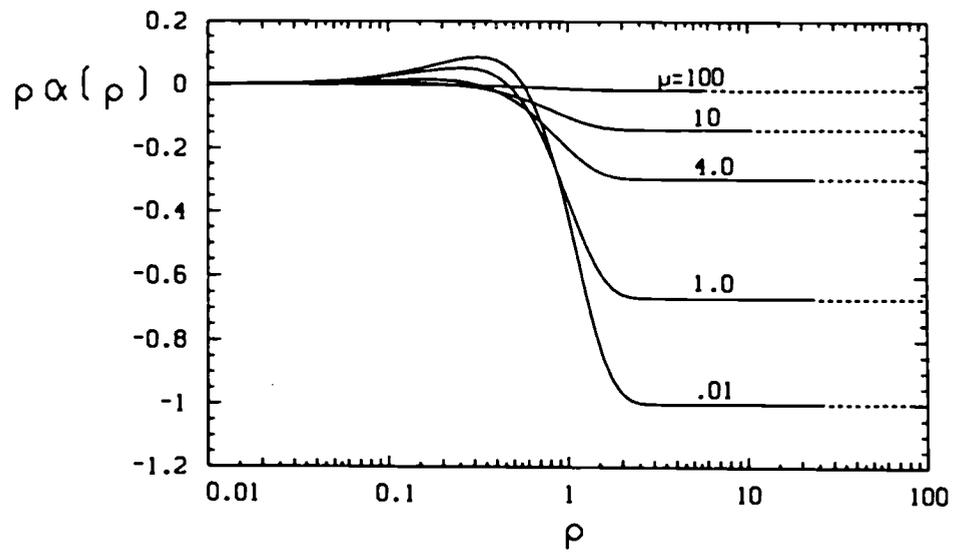


Figure 1.3: Naff's Fig. 2: Effective hydraulic conductivity factor, $\rho\alpha(\rho)$; the dashed curves are based on extrapolation for large ρ .

and thereby, avoiding the singularity in (1.125), when the denominator approaches zero. Consequently, Eq. 1.122 is a zero-order approximation. If, however, K_G in (1.117) is replaced by $\langle K \rangle$, (1.117) becomes a first order approximation, and

$$\beta(\rho) = \sigma_Y^2 \rho \alpha(\rho) \frac{\langle K \rangle}{K_G} = \sigma_Y^2 \rho \alpha(\rho) \exp(\sigma_Y^2/2). \quad (1.128)$$

Consequently,

$$\begin{aligned} K_e &= K_G \exp(\sigma_Y^2/2) [1 + \sigma_Y^2 \rho \alpha(\rho)] \\ &= \langle K \rangle [1 + \sigma_Y^2 \rho \alpha(\rho)]. \end{aligned} \quad (1.129)$$

This implies that K_e is larger than $\langle K \rangle$ for $r < \lambda$, which is theoretically impossible (as was mentioned earlier).

Eq. 1.117 is again used for the derivation of $Var[q_r]$ (the variance of the radial component of the specific discharge), subsequent to Naff's Eq. 31. It is not clear, how can q_r be random if K is constant (i.e., if $\sigma_Y^2 = 0$ then one would expect that $Var[q_r] = 0$ as well).

Naff's results show an anomalous behavior of the effective conductivity, the responses (h and q), and their variances, around $\rho \equiv r/\lambda = 1$. Naff (1991, pp. 311,313) agrees that "the somewhat anomalous behavior of these curves $0.1 < \rho < 1.0$ is not well understood and can be disconcerting, as in the case of effective hydraulic conductivity". He later postulates that it may have been caused by the Gibbs phenomenon. For the Fourier (transforms) method, the Gibbs phenomenon describes the classic oscillatory behavior of the truncated Fourier series near a point of discontinuity, and it influences the behavior not only near the point of discontinuity, but also over the entire periodic domain (Lumely and Panofski, 1964; Dykaar and Kitanidis, 1992).

“In this case, the behavior in the region $0.1 < \rho < 1.0$ is to be disregarded, as it does not relate to the physical problem”, writes Naff (1991, p. 313). This alone may discount any conclusion regarding the behavior of heads, fluxes, and particularly the “effective hydraulic conductivity” near the well.

We suspect that, besides the constraints and inconsistencies mentioned above, the boundary conditions are responsible, at least in part, for the nonphysical results. For example, Naff imposes zero mean head at $\rho = 1$ (i.e., at $r = \lambda$); at the same time, he assumes that both the Green’s function associated with head fluctuations (in his appendix B) and its radial derivative, approach zero at “infinity” (i.e., $G \rightarrow 0$ and $\partial G/\partial r \rightarrow 0$ as $r \rightarrow \infty$), and that G and its derivatives are bounded at $r \rightarrow 0$. Similar restrictions are imposed on the spectral “representation” of h' (or, more precisely, the complex Fourier amplitude associated with the head fluctuations h'), \hat{h}' , where Naff assumes that both \hat{h}' and $\partial \hat{h}'/\partial r$ are bounded at the above boundaries. This may be too restrictive. As pointed out by Jackson (1962),

It should be clear that the solution to Poisson’s equation with both h and $\partial h/\partial n$ specified on a closed boundary (Cauchy boundary condition) does not exist, since there are unique solutions for Dirichlet and Neumann conditions separately, and these will in general not be consistent. The question of whether Cauchy boundary conditions on an *open* surface define a unique problem requires more discussion than is warranted here.

See also Morse and Feshbach (1953; pp. 692-706). Greenberg (1971; pp. 82-83) brings examples of overly restrictive and overly loose boundary conditions at infinity, and warns that

Stipulation of appropriate boundary conditions at infinity is an important

part of problems in infinite domains. Quoting Friedrichs (Methods of Mathematical Physics; Lecture notes), “conditions are appropriate if they are strong enough so that at most one solution can satisfy them, and weak enough so that there exists at least one solution satisfying them”.

Naff’s boundary conditions are somewhat inconsistent with the deterministic solution for radial flow in cylindrical coordinates, as expressed by Eq. 1.117 (his Eq. 10), and by Eq. 1.118, which implies $\langle h \rangle \rightarrow \infty$ at infinity and $\langle h \rangle \rightarrow -\infty$ as $r \rightarrow 0$. This also follows from the unbounded nature of the Green’s function (or, rather, the fundamental solution) for 2-D radial flow $G(\mathbf{x}, \mathbf{x}_o) = -(1/2\pi) \ln r$, where $r = |\mathbf{x} - \mathbf{x}_o|$. If the fluctuations in h' and Y' are bounded, as implied from his appendix B, then their relative effect diminishes to zero both close to the well and very far from it. This tendency is apparent in all the figures, and may be the reason for the anomaly around $\rho = 1$.

Desbarats (1992b)

Desbarats (1992b) studies the effective block transmissivity in a 2-D square heterogeneous domain, with constant-head on all boundaries, and a well pumping at constant rate from the center. The conductivity is lognormal, statistically homogeneous and isotropic. Block transmissivities are obtained empirically by a weighted spatial averaging of point transmissivities. Desbarats’ objective is to find a (deterministic) spatial averaging law for block scale transmissivities, using geostatistics and MCS.

Desbarats defines an effective transmissivity T_{e1} through

$$Q_o = 2\pi T_{e1} \frac{(h_e - h_w)}{\ln(c_1 r_e/r_w)}, \quad (1.130)$$

where $r_e = (S/\pi)^{1/2} = L/\sqrt{\pi}$ is an apparent external radius (the area $S = L^2$), and c_1 is a shape factor such that $c_1 = 1$ for a circular field, and $c_1 = 0.956$ for a square; h_e is specified head on the boundaries (not the head at r_e), and Q_o is the (fixed) discharge rate. Desbarats introduces an alternative definition for the effective transmissivity in radial flow, similar to (1.130),

$$Q_o = 2\pi T_{e2} \frac{(\bar{h} - h_w)}{\ln(c_2 r_e/r_w)}, \quad (1.131)$$

where \bar{h} is now the “average head”, and the shape factor c_2 equal 0.6065 and 0.6015 for circular and square domains respectively. He shows later that both definitions agree with his “power averaging” (described below).

Desbarats suggests the following weighted geometric spatial average T_{S_r} for the effective transmissivity of block S,

$$\ln(T_{S_r}) \equiv Y_{S_r} = \frac{1}{W} \int_S \frac{Y(\mathbf{x})}{r^2(\mathbf{x})} d\mathbf{x}, \quad (1.132)$$

where

$$W = \int_S \frac{1}{r^2(\mathbf{x})} d\mathbf{x} = 2\pi \ln \left(c_1 \frac{r_e}{r_w} \right) \quad (1.133)$$

These expressions are then combined with the geostatistical definitions of the first two moments of Y_{S_r} , taking advantage of the Gaussianity of the (random) blocks resulting from the Gaussianity of $Y(\mathbf{x})$. Like in the uniform flow cases, Desbarats shows (theoretically, and later experimentally) that the variance of block transmissivities diminishes slowly to zero, as the block size increases with respect to the correlation scale (i.e., as $L/\lambda \rightarrow \infty$; his Figs. 2, and 3a). Similarly, the expected value of the block transmissivity, $\langle T_{S_r} \rangle \rightarrow T_A \equiv \langle T(\mathbf{x}) \rangle$ as $L/\lambda \rightarrow 0$, and $\langle T_{S_r} \rangle \rightarrow T_g \equiv \exp(\langle T(\mathbf{x}) \rangle)$ as $L/\lambda \rightarrow \infty$ (his Fig. 3).

Desbarats also investigates the (practically) important effect of a given transmissivity at the well location, T_w , (e.g., from core measurements), on the (conditional) expected value of block transmissivity, and finds the relationship

$$E[T_{S_r} | T_w] = c_o T_w^a \quad (1.134)$$

where c_o is a constant, and $a = Cov[Y_{S_r}, Y_w]$, and

$$Cov[Y_{S_r}, Y_w] = \frac{1}{W} \int_S \frac{C_Y(\mathbf{x} - \mathbf{x}_w)}{r^2(\mathbf{x})} d\mathbf{x} \quad (1.135)$$

Likewise, he finds the effect of T_w on the conditional variance of T_{S_r} , and on the conditional (first) moments of the drawdown.

The spatial averaging (1.132) is in very good agreement with corresponding effective values obtained numerically through (1.130) for low to moderate variances $\sigma_Y^2 \leq 2$, except for some outliers (particularly under high variances). Desbarats attributes these outliers to extreme values of transmissivities at the well node in some realizations, which is amplified by the harmonic averaging inherent in the finite difference scheme. We have experienced similar outliers in our radial flow simulations, despite grid refinement. [We attribute outliers to the random field generators, as is demonstrated in Chapter 4.]

Desbarats calculates *ensemble* point effective transmissivities, and plots them versus distance from the well, in comparison with Naff's (1991) results, in Figure 1.4 below (his Fig. 11). He finds that $T_e \rightarrow T_H$ as $r/\lambda \rightarrow 0$, and that T_e increases towards (but does not reach) the arithmetic mean far from the well. However, based on similarity between flow at large distance from a well and uniform mean flow (as explained above), Desbarats suggests that T_e should, in fact, approach T_G far from the well. Two possible reasons for the deviation of T_e from T_G far from the well

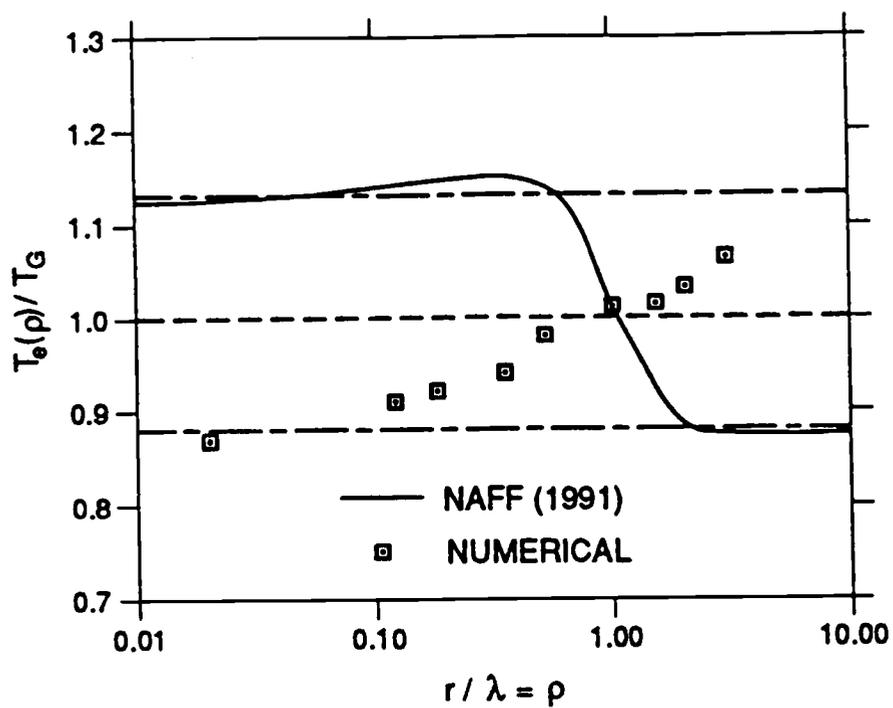


Figure 1.4: A comparison between (normalized) effective (point) conductivities as a function of dimensionless distance from the well bore, as predicted by Naff (1991) and calculated by Desbarats (1992b) from (numerical) MCS.

are: (i) the influence of the prescribed head, outer boundary condition, due to the limited domain (or block) size ($L/\lambda \leq 10$, or $r/\lambda \leq 5$, and (ii) the finite difference discretization. Desbarats' results (including his latter suggestion) are reproduced by our MCS in Chapter 5. Interestingly, our use of prescribed *flux* boundary condition on outer boundaries has caused the effective transmissivity to tend towards the *harmonic* mean, close to the boundary.

CHAPTER 2

OPERATOR THEORY OF STOCHASTIC PDE'S

2.1 Operator Representation

In the previous chapter we discussed methods for solving stochastic partial differential equation (stochastic PDE) by using closure approximations. By “solving”, we meant relating the ensemble moments of heads, $h(\mathbf{x})$, and fluxes, $\mathbf{q}(\mathbf{x})$, to the ensemble moments of the parameter, $K(\mathbf{x})$ (or its logarithm, $Y = \ln K$), particularly the first two moments. We discussed primarily perturbation techniques, in which the governing (stochastic) PDE is left in its differential form, $Y(\mathbf{x})$ is expanded in a Taylor series, while $h(\mathbf{x})$ and $\mathbf{q}(\mathbf{x})$ are either expanded in asymptotic sequences or decomposed into “mean” and (zero mean) small fluctuations; under some restrictive conditions, i.e., statistical homogeneity of Y in an unbounded domain under (mostly) uniform mean flow, first order approximations (in σ_Y^2) can be obtained for the first two moments of h and \mathbf{q} , in a closed form. Some of these methods resort, at a later stage, to an integrodifferential form, particularly by using the Green’s function approach (e.g., Dagan, 1989). For more complex situations, particularly bounded domains and nonuniform flow, the approximations either do not apply [because of unknown mixed moments that are added to the equations] or an intuitive conjecture must be invoked at some point [as in Dagan’s (1989, pp. 77-78) analysis of nonuniform flow]; even then, perturbation methods are limited to small variability (i.e., $\sigma_Y^2 \ll 1$). In the following, we discuss methods that develop, directly, explicit relationships between

the ensemble moments of $h(\mathbf{x})$ [or $\nabla h(\mathbf{x})$] and the statistics of $K(\mathbf{x})$, by using operator representations and (stochastic) integrodifferential equations. In the following, we show that this form of solutions has some definite advantages over the differential form used in the small perturbation methods; these advantages will become more obvious in Chapter 3, where we propose a new, more powerful integrodifferential formulation.

The stochastic PDE

$$\nabla \cdot [K(\mathbf{x})\nabla h(\mathbf{x})] + f(\mathbf{x}) = 0 \quad (2.1)$$

can be written in the general *operator* form:

$$\mathcal{L} h + f = 0 \quad (2.2)$$

where \mathcal{L} is a linear stochastic *differential* operator (which includes random parameters), f is an excitation (or a source term) which can be either random or deterministic, and h is the random response.

The inverse operator \mathcal{L}^{-1} is a stochastic *integral* operator such that

$$h = \mathcal{L}^{-1}(-f). \quad (2.3)$$

The stochastic operators \mathcal{L} and \mathcal{L}^{-1} in (2.2) and (2.3) transform a random variable (or function) to another random function.

The relationships between differential and integral equations can be best demonstrated by means of *Green's functions* (e.g., Greenberg, 1971; Fletcher, 1988). Consider, for example, the partial differential equation

$$Lh + f = 0, \quad (2.4)$$

subject to certain boundary conditions, where L is a deterministic differential operator. A solution can be constructed, in principle, by “inverting” the operator L , yielding the convolution integral

$$L^{-1}[-f] = h(\mathbf{x}) = - \int_{\Omega} G(\mathbf{x}, \mathbf{x}_o) f(\mathbf{x}_o) d\mathbf{x}_o, \quad (2.5)$$

where the kernel (the fixed part) $G(\mathbf{x}, \mathbf{x}_o)$ is the *Green's function*, discussed in the previous chapter (Section 1.5.1). An alternative form for the inverse operator can be obtained by the *eigenfunction method* discussed in Appendix C.

Schetzen (1980) views differential and integral equations as two possible models of the same system:

Generally, the modeling used in the study of systems can be classified as being either implicit or explicit. Implicit models are those in which the system response is expressed as an implicit operation on the system input. An example is ... a differential equation. Explicit models are those in which the system response is expressed as an explicit operation on the system input. An example is the modeling of the relation between the system response and input by a convolution integral. Neither type of model is all encompassing and each provides insights not provided by the other. The model that is best to use thus depends on the specific questions being asked and the specific understanding of the system operation being sought. For example, ...the study of the spectrum of the system response for a random input, normally is best achieved using explicit models.

While in (partial) differential equations the boundary and/or initial conditions are imposed at a final stage of the solution, integral equations incorporate boundary

conditions in two ways: (i) as limits or boundary integrals, and (ii) in the kernel itself. *The integral equation relates the unknown function not only to its values at neighboring points (derivatives), but also to its values throughout the region, including the boundary* (Arfken, 1985).

The case with a deterministic operator L and random excitation f has been studied extensively in the field of structural engineering (see Ghanem and Spanos, 1991, for references), while the hydrologic problem of steady state subsurface flow with random boundary condition and recharge was tackled by Cheng and Lafe (1991). The case where \mathcal{L} is stochastic is considerably more difficult, and only approximate solutions to the problem have been reported in the literature. Interest in this class of stochastic differential equations has its origins in quantum mechanics, wave propagation, turbulence theory, random eigenvalues, and functional¹ integration (Ghanem and Spanos, 1991).

It is common to split the stochastic operator \mathcal{L} in Eq. 2.2 into a deterministic part L , and a zero mean random part, \mathcal{R} (Frisch, 1968; Adomian, 1983; Ghanem and Spanos, 1991),

$$(L + \mathcal{R}) h + f = 0. \quad (2.6)$$

When applied to (1.1),

$$\begin{aligned} \mathcal{L} &= \nabla \cdot [K \nabla] = L + \mathcal{R} \\ \langle \mathcal{L} \rangle &= L = \nabla \cdot [\langle K \rangle \nabla] \\ \mathcal{R} &= \nabla \cdot [K' \nabla], \quad \langle \mathcal{R} \rangle = 0 \end{aligned} \quad (2.7)$$

¹A functional is a quantity which depends upon all the values a function $f(x)$ takes in some interval $a \leq x \leq b$; roughly speaking, it is a function of a function.

and

$$\begin{aligned} K(\mathbf{x}) &= \langle K(\mathbf{x}) \rangle + K'(\mathbf{x}), & \langle K'(\mathbf{x}) \rangle &= 0 \\ \mathbf{q}(\mathbf{x}) &= \langle \mathbf{q}(\mathbf{x}) \rangle + \mathbf{q}'(\mathbf{x}), & \langle \mathbf{q}'(\mathbf{x}) \rangle &= 0 \end{aligned} \quad (2.8)$$

where $\langle \cdot \rangle$ denotes ensemble mean (or mathematical expectation), $K'(\mathbf{x}) = K(\mathbf{x}) - \langle K(\mathbf{x}) \rangle$, and $\mathbf{q}'(\mathbf{x}) = \mathbf{q}(\mathbf{x}) - \langle \mathbf{q}(\mathbf{x}) \rangle$.

Solving for h results in

$$h = L^{-1}[-f] - L^{-1}\mathcal{R}h \quad (2.9)$$

which for an unbounded (infinite) domain takes the (explicit) form

$$\begin{aligned} h(\mathbf{x}) &= - \int_{\Omega_\infty} G(\mathbf{x}, \mathbf{x}_0) [f(\mathbf{x}_0) + \mathcal{R}h(\mathbf{x}_0)] d\mathbf{x}_0 \\ &= - \int_{\Omega_\infty} G(\mathbf{x}, \mathbf{x}_0) [f(\mathbf{x}_0) + \nabla \cdot K'(\mathbf{x}_0) \nabla h(\mathbf{x}_0)] d\mathbf{x}_0. \end{aligned} \quad (2.10)$$

This is an integrodifferential equation with a random forcing function f and random kernel, $\mathcal{K} = G(\mathbf{x}, \mathbf{x}_0)\mathcal{R}$. Bharucha-Reid (1972, p. 160) notes that “this form of the random kernel (is)... a multiplicative perturbation of the (deterministic) Green’s function by the random operator \mathcal{R} ”.

Integral operators are “nicer” than differential operators since integration is a smoothing process whereas differentiation has the opposite effect (Greenberg, 1978). This property is especially important when dealing with random fields. “From the probabilistic viewpoint, the integral equation formulation has the advantage that it does not make use of derivatives of random functions, which need not exist” (Frisch, 1968)². Adomian (1983) equates the integral operator \mathcal{L}^{-1} (cf. Eq. 2.2) with a stochastic *filter*. Consequently he views the resulting theory as a generalization of

²This is not the case, however, with our kernels, as seen in (2.10).

ordinary filter theory. The method presented in the third chapter of this dissertation takes full advantage of the smoothness and filtering features of the integral equations.

According to Arfken (1985), mathematical problems such as existence, uniqueness and completeness may be handled more easily and elegantly in integral form. However, this cannot be taken for granted when dealing with stochastic operators. According to Bharucha-Reid (1972, p. 113), “at the present time there are no known results which give necessary and sufficient conditions for the existence of the *inverse* of *random* operators (i.e., operators with random coefficients) of the types³

$$\mathcal{L}(\zeta)h = -f(\mathbf{x}) \quad (2.11)$$

$$\mathcal{L}(\zeta)h = -f(\mathbf{x}, \zeta) \quad (2.12)$$

where ζ stands for “elementary events” in the probability (measure) space, and h and $f(\mathbf{x})$ are continuous scalar-valued functions.” Equations 2.11 and 2.12 are equivalent to Eq. 2.2, or Eq. 1.1. While in (2.11) the source term f is deterministic, in (2.12) it is random. Fortunately, results on the inversion of random operators of the form

$$(\mathcal{L}(\zeta) - \lambda I)h = -f(\mathbf{x}) \quad (2.13)$$

$$(\mathcal{L}(\zeta) - \lambda I)h = -f(\mathbf{x}, \zeta) \quad (2.14)$$

are known [here λ is the eigenvalue (see Appendix C) and I is the identity operator], and can be used to establish the existence of the random solution of equations (2.11), (2.12) (Bharucha-Reid, 1972). According to Curtain and Pritchard (1977), a transformation between two normed linear spaces is said to be “invertible” if the algebraic inverse exists (i.e., there exists $\mathcal{L} = S^{-1}$, such that $S\mathcal{L} \equiv \mathcal{L}^{-1}\mathcal{L}$ and $\mathcal{L}S \equiv \mathcal{L}\mathcal{L}^{-1}$ are

³Translated to our operator terminology, from more general “transformation” terms.

identity transformations), *and* is continuous; moreover, “it can be shown that the properties of continuity and boundedness are equivalent” (*ibid*). In other words, in order for equations 2.9–2.10 to exist, the integrals (i.e., the inverse operators) have to be bounded, which implies boundedness of the operators under the integral (except at some singular points). The boundedness condition is particularly disturbing due to the randomness of our operators. Because existence of inverse operators is closely related to *contraction* and convergence of Neumann series (discussed in the following sections), we defer our discussion on this (important) topic to Section 2.2.1.

The random integrodifferential equation (2.9) [or (2.10)] is the first step in the solution of the original problem (1.1) or (2.2). Subsequent steps vary from one method of solution to another. In most of these methods, the second step consists of successive approximations, along the line of deterministic, one-dimensional solutions of integral equations. At some point, the equation is averaged and approximated, to yield the desired relationships between the statistical moments of the input (e.g., conductivity field) and the moments of the output (e.g., head field).

2.2 Expansion in Neumann Series

Variants of the method of *successive approximations*, or Expansion in Neumann series are commonly used for solving deterministic and stochastic integrodifferential equations. These variants include the methods of “Volterra-Wiener series” (Schetzen, 1980; Markov, 1987), “Adomian decomposition” (Adomian, 1983, 1986; Zeitoun and Braester, 1991), “Picard iterations” and to a lesser extent, “Hierarchy closure approximation” (Bharucha-Reid, 1968). Throughout this chapter, we discuss variations on and limitations of these variants. The general idea is to invert

the stochastic differential operator into a stochastic integro-differential form, then expand it by successive approximations. At some point in the substitution process, the series is averaged, and either approximated or truncated, to yield a tractable kernel, from which joint statistical moments of the system output as functions of the moments of the (input) random parameters can be obtained.

Averaging of Eq. 2.9 (repeated here for clarity)

$$h = L^{-1}[-f] - L^{-1}\mathcal{R}h \quad (2.15)$$

yields

$$\langle h \rangle = -L^{-1}\langle f \rangle - L^{-1}\langle \mathcal{R}h \rangle. \quad (2.16)$$

We note that the mean of h depends on those of f and $\mathcal{R}h$. To allow determining the latter, one applies the operator \mathcal{R} to Eq. 2.15,

$$\mathcal{R}h = \mathcal{R}L^{-1}[-f] - \mathcal{R}L^{-1}\mathcal{R}h. \quad (2.17)$$

Substituting this into Eq. 2.15 yields

$$h = L^{-1}[-f] - L^{-1}\mathcal{R}L^{-1}[-f] + L^{-1}\mathcal{R}L^{-1}\mathcal{R}h. \quad (2.18)$$

This substitution process can be continued indefinitely, which is equivalent to expansion of (2.15) in a Neumann series, according to

$$h = (1 + L^{-1}\mathcal{R})^{-1}L^{-1}[-f], \quad (2.19)$$

provided that the inverse of $(1 + L^{-1}\mathcal{R})$ exists (Arfken, 1985, p. 881), and it is legitimate to write

$$(1 + L^{-1}\mathcal{R})^{-1} = 1 - L^{-1}\mathcal{R} + (L^{-1}\mathcal{R})^2 - (L^{-1}\mathcal{R})^3 + \dots = \sum_{k=0}^{\infty} (-1)^k (L^{-1}\mathcal{R})^k. \quad (2.20)$$

This is analogous to expansion of $(1+x)^{-1}$ in a binomial power series (e.g., Gradshteyn and Ryzhik, 1980)

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots + (-x)^{n-1} + \dots = \sum_{k=0}^{\infty} (-1)^k x^k, \quad (2.21)$$

which, however, converges only for $|x| < 1$. Eq. 2.21 is an *identity* which is readily verified upon multiplying both sides by $(1+x)$. However, for $|x| \geq 1$ one must include *all* (infinite number of) terms, otherwise the finite series diverges. In other words, for practical purposes, the series diverges when $|x| \geq 1$. We therefore expect the Neumann series to diverge, unless (Ghanem and Spanos, 1991; Kröner, 1977; Adomian, 1983)

$$\|L^{-1}\mathcal{R}\| < 1 \quad (2.22)$$

where $\|\cdot\|$ represents a suitable norm (see Appendix B for definitions of different norms). According to Adomian (1983, p. 302), under the condition (2.22), the series (2.23) converges *uniformly*. More on convergence later.

The Neumann series approach consists of writing according to (2.19) and (2.20) (Adomian, 1983; Bharucha-Reid, 1972, p. 160; Ghanem and Spanos, 1991; Greenberg, 1978, p. 349)

$$h = \sum_{k=0}^{\infty} (-1)^k [L^{-1}\mathcal{R}]^k L^{-1}[-f], \quad (2.23)$$

and taking ensemble mean. Clearly, the right hand side involves only (known) joint moments of \mathcal{R} and f . In practice, the series is truncated after a final number of terms.

An alternative (so called hierarchial) approach is to continue the substitution (2.17)–(2.18) only to a finite number of terms,

$$h = \sum_{k=0}^{n-1} (-1)^k (L^{-1}\mathcal{R})^k L^{-1}[-f] + (L^{-1}\mathcal{R})^n h, \quad (2.24)$$

then take ensemble mean. Since joint moments of \mathcal{R} and h are unknown, it is common to invoke the closure approximation (Ghanem and Spanos, 1991)

$$\langle [L^{-1}\mathcal{R}]^n h \rangle \approx \langle [L^{-1}\mathcal{R}]^n \rangle \langle h \rangle. \quad (2.25)$$

Specifically, upon taking the ensemble average of (2.18), which corresponds to $n = 2$ in (2.24) (the second term on the right hand side of the equation drops if \mathcal{R} is independent of f), so

$$\langle h \rangle = -L^{-1}\langle f \rangle + L^{-1}\langle \mathcal{R}L^{-1}\mathcal{R}h \rangle. \quad (2.26)$$

The corresponding hierarchy “closure approximation” is (Frisch, 1968; Adomian, 1983)

$$\langle \mathcal{R}L^{-1}\mathcal{R}h \rangle \approx \langle \mathcal{R}L^{-1}\mathcal{R} \rangle \langle h \rangle. \quad (2.27)$$

Heuristic arguments such as “local independence” have been made to justify the decoupling behind the closure assumption in (2.25) and (2.27). The decoupling, however, has no rigorous basis; it was criticized by Keller (1962), Kraichnan (1961; see Frisch, 1968) and Adomian (1983), and was shown by the latter two authors to be equivalent to the truncation of small perturbation series (i.e., closure at a certain level of the hierarchy is equivalent to the same order of perturbation in a perturbation solution). On the other hand, upon truncating the Neumann series (2.23) after $n - 1$ terms, and comparing it with (2.24), one may wonder whether dropping the last (originally coupled) term (as implied by the Neumann series approach) is better than adding it in a decoupled form [as in (2.24)]. Adomian (1983, pp. 205-209) compared the two methods with results from numerical integration for the case of a time-dependent ordinary differential equation, and found that the resulting truncated Neumann series (after 12 terms) yielded more accurate results than

the hierarchy method; it is not clear, however, how many terms were considered in the hierarchy method.

To better understand the behavior of these solution methods, we observe on the basis of (2.10) that (2.9) [or (2.15)] has the following one-dimensional analogue,

$$h(x) = g(x) + \lambda \int \mathcal{K}(x, x_o) h(x_o) dx_o, \quad (2.28)$$

where λ is a constant, and \mathcal{K} is a *kernel* (which is frequently taken to be symmetric), such that the second term in the right hand side is analogous to $L^{-1}\mathcal{R}h$. Though confusing, a commonly used operator form of (2.28) is

$$h = g + \lambda\mathcal{K}h. \quad (2.29)$$

In the case of an infinite domain, $g(x)$ is analogous to $L^{-1}[-f]$, so

$$g(x) = - \int_{-\infty}^{\infty} G(x, x_o) f(x_o) dx_o, \quad (2.30)$$

and $\lambda\mathcal{K}(x, x_o) = -G(x, x_o)\mathcal{R}$. In the special case of an algebraic (rather than differential) operator \mathcal{R} , (2.28) is an *integral equation of the second kind* (in our case, the equation is *integrodifferential*). If the limits of integration are fixed, (2.28) is a *Fredholm equation*; if one of the limits is variable, it is a *Volterra equation*.

The Volterra integral equation of the second kind is readily solved by Picard successive approximations (e.g., Tricomi, 1957), leading to a *Volterra series*. A similar iterative process for the Fredholm equation leads to a *Neumann series* (or *Liouville-Neumann series*). Let $\Gamma(x, x_o)$ be the so called *resolvent kernel* of $\mathcal{K}(x, x_o)$

in (2.28) given by the Volterra or Neumann series

$$\Gamma(x, x_o) = \sum_{n=0}^{\infty} \lambda^n \mathcal{K}_{(n+1)}(x, x_o), \quad (2.31)$$

where the iterated kernels $\mathcal{K}_n(x, x_o)$ are defined as

$$\begin{aligned} \mathcal{K}_1(x, x_o) &= \mathcal{K}(x, x_o) \\ \mathcal{K}_2(x, x_o) &= \lambda \int \mathcal{K}(x, y) \int \mathcal{K}(y, x_o) dy \\ &\vdots \\ \mathcal{K}_n(x, x_o) &= \lambda^{(n-1)} \int \mathcal{K}_{(n-1)}(x, y) \int \mathcal{K}(y, x_o) dy. \end{aligned} \quad (2.32)$$

Then

$$h(x) = g(x) + \lambda \int \Gamma(x, x_o) f(x_o) dx_o \quad (2.33)$$

(Courant and Hilbert, 1953; Tricomi, 1957; Arfken, 1985). Adomian (1983) calls $\Gamma(x, x_o)$ a generalized Green's function.

If the Neumann series converges, then a proper truncation of (2.31) may yield an acceptable approximation for $\Gamma(x, x_o)$. The success (or accuracy) of this approximation depends, among other factors, on the quality of the initial guess (usually taken to be $h(x) \approx h_o(x) = g(x)$).

Unlike the *Neumann* series, the deterministic (one-dimensional) *Volterra* series converges almost everywhere (excluding some possible singular points or poles) for *all* λ 's, as long as the kernel is square integrable (and hence, bounded; e.g., Tricomi, 1957), i.e.,

$$\|\mathcal{K}\|^2 = \iint \mathcal{K}^2(x, y) dx dy \leq M^2 \quad (2.34)$$

where M is a finite upper bound.

The series (2.31) can be viewed as a geometric series whose convergence may be checked by the Cauchy ratio test (Tricomi, 1957, p. 51). Consequently, the deterministic (one-dimensional) Neumann series converges if

$$|\lambda| < \|\mathcal{K}\|^{-1}, \quad (2.35)$$

Similarly, Arfken (1985) and Stakgold (1979) show [based on (2.31)–(2.33) and the ratio test for geometric series] that a sufficient condition⁴ for the convergence of the (deterministic, one-dimensional) Neumann series is

$$|\lambda| < |b - a|^{-1} |\mathcal{K}|_{max}^{-1} \quad (2.36)$$

where a and b are the limits of integration. Following the operator notation of (2.29),

$$h = (1 - \lambda\mathcal{K})^{-1}g. \quad (2.37)$$

Binomial expansion leads to (2.33). Consequently, the series converges only if $\|\lambda\mathcal{K}\| < 1$ [cf. (2.22)]. At first glance, the latter condition seems, to contradict (2.36); however, note that \mathcal{K} in (2.37) is a (convolution) *integral*, not a *kernel*.

The kernel \mathcal{K} of the homogeneous (with $g(x) = 0$) Fredholm equation of the second kind, (2.28), can be expressed in terms of eigenfunctions, ϕ_n and eigenvalues, λ_n , (discussed in Appendix C) as

$$\mathcal{K}(x, x_o) = \sum_{n=1}^{\infty} \frac{\phi_n(x)\phi_n(x_o)}{\lambda_n} \quad (2.38)$$

(Tricomi, 1957; Arfken, 1985; cf. Eq. C.13 in Appendix C). A similar expression for the iterated symmetric kernel (2.33) is shown by Bharucha-Reid (1972; p. 180) to

⁴Arfken (1985, p. 881) shows an example in which the condition (2.36) does not hold, but nevertheless the series converges.

admit the (power) expansion

$$\mathcal{K}_n(x, x_o) = \sum_{k=1}^{\infty} \frac{\phi_k(x)\phi_k(x_o)}{\lambda_k^n} \quad n = 1, 2, \dots \quad (2.39)$$

all of which converge absolutely and uniformly in both x and x_o (see Appendix D for definitions). Expansion (2.39) is valid for $n \geq 2$ as long as the kernel \mathcal{K} is square integrable (an L_2 -kernel). If \mathcal{K} is continuous and positive, then (2.39) holds for $n = 1$. [A symmetric kernel $\mathcal{K}(x, y)$, $x, y \in [a, b]$ is said to be *positive* if for every function $f \in L_2[a, b]$, $\int_a^b \int_a^b \mathcal{K}^2(x, y)f(x)f(y)dxdy \geq 0$.]

In terms of eigenvalues, a necessary and sufficient condition for convergence of the Neumann series (associated with the Fredholm equation) is that $|\lambda| < |\lambda_e|$, where λ_e is the smallest eigenvalue of the corresponding homogeneous equation [when $g(x) = 0$]. According to Porter and Stirling (1990), a sufficient condition for convergence is that the largest eigenvalue of the integral operator be less than one. In other words, the Neumann series converges only for sufficiently small $|\lambda|$ (Tricomi, 1957). In contrast, it is shown by Tricomi (1957; his section 1.5 and p. 72), that a Volterra integral equation has no eigenvalues; this is why it converges uniformly almost everywhere (see Appendix D).

Although the above analyses refer mostly to one dimensional deterministic algebraic operators, the above conditions for convergence can be sometimes extended to hold for random two and three dimensional spaces, particularly for bounded, square integrable kernels. For example, Frisch (1968) shows that the (sufficient condition for) convergence of the Neumann series corresponding to the random Helmholtz equation

$$\nabla^2 u + k^2 u = f, \quad (2.40)$$

[where k is the wave number (or frequency), and f is a source term] depends inversely

on the domain size⁵ and on the upper bound of the random fluctuations, in a manner similar to (2.36).

With this in mind, we can say that when $\lambda\mathcal{K}(x, x_o) = -G(x, x_o)\mathcal{R}$, (2.35) is formally analogous to (2.22). Setting $\lambda\mathcal{K}(x, x_o) = -G(x, x_o)\mathcal{R}$ renders (2.23) analogous to (2.33). The corresponding resolvent kernel, $\Gamma(x, x_o)$, which is associated with the *deterministic Green's function* $G(x, x_o)$ is nevertheless termed *stochastic Green's function* by Adomian (1983). In this dissertation, the term stochastic Green's function will be given a different meaning. As averaged kernels of random integrodifferential equations consist of products of (derivatives of) Green's function and moments of the random coefficients, the constant λ in (2.33) can be also interpreted as (a power of) the variance of the random parameter (e.g., σ_K^2). If we consider $\lambda \approx \sigma_K^2$, then (based on Eq. 2.36) the variability of K must be small to enable convergence of the Neumann series.

While steady-state problems lead to Neumann series, time-dependent problems lead to Volterra series Adomian (1983) shows that in a *Volterra*-type series, for a bounded input (forcing function) and a stationary stochastic process on a *bounded* time interval, say, on $[0, T]$, beyond which the correlation vanishes (i.e., for $t > T$), and where the kernels of the random integral operators have a finite bound, M (hence, their fluctuations are also bounded), the series converges with an upper bound of

$$\langle h(t) \rangle \leq \sum_{n=0}^{\infty} \frac{M^{2n} t^n}{n!} \leq M e^{Mt}. \quad (2.41)$$

This upper bound, resembles the upper bound of the *deterministic* Volterra series (e.g., Tricomi, 1957; Adomian, 1983, pp. 134, 311). Adomian reasons the bounded-

⁵More precisely, on D^n , where D is the diameter (i.e., the upper bound of the distance between two points in the domain) of the n -dimensional domain.

ness of the kernels by the boundedness of the random coefficients by the finiteness of physical systems. Note that even in this case, (i) the upper bound reaches ∞ for $t \rightarrow \infty$; (ii) for a finite time interval, the number of terms that have to be included in the series to avoid error, increases with time, t , and (iii) expressions proportional to the positive power of t (possibly the first few terms in the series) increase indefinitely as $t \rightarrow \infty$; these terms are called *secular terms* (Frisch, 1968). Hence, although the limit (2.41) implies that the series converges to the random operator which solves for $\langle h \rangle$, the convergent expansion cannot be used to study asymptotic behavior of the solution (Frisch, 1968). This is a weakness of the Volterra series; because applications of successive approximations to time-dependent problems always result in expansions in Volterra series, (including semigroup expansions, as will be shown in Section 2.4.2), this weakness dominates time-dependent problems that use variants of this method (cf. Serrano and Unny, later on, in Section 2.4.3). Note that the failure of a finite order expansion, in this case, is not related to the strength of the random fluctuations (i.e., it does not depend on the *variance* of the random coefficients).

According to Bharucha-Reid (1972, p. 113), existence theorems for inversion of random operators are directly related to *contraction* operators (below), which, in turn, are directly related to the convergence (or simply success) of the successive approximations. Arfken (1985, p. 881) states that the convergence of the Neumann series is a demonstration that the inverse operator $(1 - \lambda\mathcal{K})^{-1}$ in (2.37) exists. The close relationships between contraction and inversion of random operators, and convergence of the Neumann series implies that the latter must be (either implicitly or explicitly) assumed by *any* method of solution (and/or approximation). This is quite

disturbing, particularly, when using methods that do not employ successive approximations, but are required to assume existence of inverse of the (random) operators. We explore some of these concepts briefly, below.

2.2.1 Contraction, Inverse Operators, and Convergence of the Neumann Series

Equations 2.9, 2.15, 2.28, and 2.29 can be represented by the iterative process (Hinch, 1991)

$$h = f(h) \quad (2.42)$$

According to the Newton-Raphson (iterative) method this will converge to the root h^* if

$$|f'(h^*)| < 1 \quad (2.43)$$

that is, each iteration will decrease the error if the absolute value of the derivative of the function (or functional, or operator) is less than one, provided that the iteration starts sufficiently close to the root (also called “the fixed point”). The Newton-Raphson (iterative) method is then more efficient than Picard iteration (or the successive approximation) method, due to its quadratic convergence.

The condition (2.43) for convergence is implied by the *contraction mapping theorem* (Hinch 1991, p. 14; Curtain and Pritchard, 1977, p. 289). According to Stakgold (1979), contractions (or contraction transformations) form an important class of transformations for which the method of successive approximations works. Stakgold (1979; p. 243) recasts (2.42) in the form

$$u = Tu \quad (2.44)$$

where T is a transformation of a metric space X onto itself. A contraction brings points uniformly closer together, shrinking the distance between them by a scale factor smaller than one. The contraction theorem states (Stakgold, 1979):

Let T be a contraction on a complete metric space. Eq. 2.44 then has one and only one solution, which can be obtained by iteration from any initial approximation whatever.

If \mathcal{X} is a *Banach space* (a normed metric space; see Appendix B for definitions), and $F : \mathcal{X} \rightarrow \mathcal{X}$ a *contraction*, then there is a unique $x^* \in \mathcal{X}$, such that $F(x^*) = x^*$, and x^* is called *the fixed point of F* (Curtain and Pritchard, 1977).

Stakgold (1979) uses the contraction theorem to investigate conditions under which the *nonlinear* deterministic Fredholm integral equation (where the kernel \mathcal{K} depends on the unknown function) tends *uniformly* to the one and only one continuous solution. He finds the familiar *sufficiency* condition, $M < 1/(b - a)$, where M is an upper bound of the integrand, and a, b are the integral limits (cf. Eq. 2.36). This means that either M or the size of the domain of interest have to be sufficiently small. He obtains similar results for the *linear* Fredholm integral equation. [Note that (based on Appendix D) the *uniform* convergence condition may be too restrictive.] Similarly, Stakgold shows that the Volterra iterative scheme is always a contraction, and therefore always converges successfully (for a bounded domain or limited time, as discussed earlier).

Bharucha-Reid (1972) starts by presenting Eq. 2.44 in the Banach space [which includes the probability (measure) space] as a basis for several *random fixed point theorems*. He indicates that finding a fixed point of T is equivalent to obtaining a solution to the operator equation; hence, *fixed point theorems constitute a general*

class of existence theorems. Bharucha-Reid distinguishes between *topological fixed point theorems*, which are strictly *existence* theorems [i.e., they do not provide a method for finding the fixed point (or solution) of the operator equation], and *algebraic* or *constructive fixed point theorems*, which give a method for finding the fixed point (e.g., by iterations or successive approximations). According to Bharucha-Reid, the prototype of most algebraic fixed point theorems is the contraction mapping theorem, which “can be considered as a result of geometric interpretation and abstract formulation of the classical method of successive approximations due to Picard” (*ibid*, p. 107). His subsequent (algebraic) contraction mapping theorems (in Banach space) provide *sufficient* (in contrast with necessary) conditions for existence and uniqueness of the contraction operators, as well as error estimates. The (sufficient) condition that dominates these theorems is quite restrictive:

$$\|T\| < 1. \quad (2.45)$$

Subsequently, Bharucha-Reid presents random fixed point theorems, followed by the *random contraction mapping theorem*. He stresses that random contraction mapping theorems “are of fundamental importance in probabilistic functional analysis in that they can be used to establish the existence and uniqueness of solutions of random operator equations ...and can be utilized in the study of stochastic approximation procedures”. He then shows that if there exists a unique fixed point of a random contraction operator, then if its Bochner integral exists (i.e., if $\|T\| \in L_1$; cf. *ibid*, p. 21; see also Appendix B), one can define the expectation of the random fixed point, and the fixed point of the expectation of the random contraction operator, i.e., $\langle T \rangle$.

On the other hand, existence and uniqueness of solutions to (random) equa-

tions are directly related to *inversion theorems*. According to Bharucha-Reid (1972, p. 113), in the case of deterministic operators and random input, deterministic inversion theorems can be used to obtain conditions for the existence of random solution. In the case of random operators, however, the inversion theorems for deterministic operators are not applicable, and probabilistic versions of the classical theorems are required in order to solve random operator equations of the form (2.11)–(2.14). Recalling Arfken's statement that the convergence of the deterministic (one-dimensional) Neumann series is a demonstration that the inverse operator exists, we return to its random analogue.

Conditions for almost sure convergence (with probability 1)⁶ of the Neumann series of stochastic integrals when the random parameters are *bounded* are also presented by Tsokos and Padget (1974). Benaroya and Rehak (1987) could not conclude whether the convergence of the Neumann series is restricted to small variations of the parameters.

We have seen that boundedness of the (random) operators is essential in all of the above contraction (and/or fixed point) theorems. Strictly speaking, 2-D and 3-D Green's functions relevant to steady state flow, include singularities, and random (log-normal) hydraulic conductivities are, theoretically *unbounded*; therefore, we cannot a-priori assume bounds on our kernels. However, although the particular values of the random parameters are unbounded in a strict sense, (i) *their moments are bounded* (a fact that will be crucial in the next chapter), and (ii) in reality (physically), they are, in fact, bounded. Moreover, most existence theorems for both *contraction* and *inversion* of random operators allow some singular points, or even curves to exist

⁶See Appendix D for a complete definition.

within the domain or its boundaries. Bharucha-Reid (1972) reminds us that

Although results on the existence, uniqueness and measurability of solutions are of theoretical interest, and are required for the systematic development of the theory of random equations, we are obliged to recognize that in most applied problems, it is not the solution that is of great importance, but the statistical properties of the solution.

Despite the seeming equivalence between uniform convergence and contraction (pointed out by Stakgold, 1979, above), Hinch (1991) relates contraction to *asymptotic behavior* rather than to convergence. As shown in Appendix D, the asymptotic approximations (or semiconvergent series) disregard convergence; instead, they emphasize the validity of *truncation*. Hinch (1991) demonstrates a case where a series *diverges* due to an unbounded (but invertible) differential operator, but is yet *asymptotic*; in another case, the convergence depends on the upper bound of the integral and thus, any practical truncation entails a significant error. Similarly, Arfken (1985) demonstrates the use of asymptotic but divergent series for establishing *upper and lower bounds* of the incomplete Gamma function.

We have seen that the convergence of Neumann series is closely related to invertability of operators and existence of solutions to steady-state problems. Hence, any method that attempts to solve stochastic integrodifferential equations, shares in common the restrictions that apply to Neumann series.

2.3 Applications of Integrodifferential Equations and Neumann Series

In this section, we review relevant works from different fields of engineering and science, in which the governing stochastic PDE's are transformed to random integrodif-

ferential equations, with or without operator representation. Most of these works use variants of Neumann and Volterra series in an attempt to solve the resulting random integrodifferential equations, while some works use alternative methods.

2.3.1 Stochastic Representation of Microflow and Derivation of Macroequations (Beran, 1968)

One of the “classic” examples of the derivation of Darcy’s law from the microscopic Stokes equation is given by Beran (1968). Beran starts from Stokes equation for slow viscous flow,

$$\nabla^2 \mathbf{v} = \frac{\gamma}{\mu} \nabla h, \quad (2.46)$$

where $\gamma \equiv \rho g$ is the specific weight of the fluid (ρ being fluid density, and g acceleration due to gravity), and μ is the fluid viscosity. The (deterministic) Green’s function $G_v(\mathbf{x}, \mathbf{x}_o)$ corresponding to (2.46) satisfies

$$\nabla^2 G_v(\mathbf{x}, \mathbf{x}_o) + \delta(\mathbf{x} - \mathbf{x}_o) = 0 \quad (2.47)$$

subject to the homogeneous (Dirichlet) boundary condition $G_v(\mathbf{x}, \mathbf{x}_s) = 0$ at the solid–fluid interface, where $\mathbf{v}(\mathbf{x}_s) = 0$). \mathbf{x}_s is a point on the surface of the solid phase of the porous medium, and δ is the Dirac delta function. Then

$$\mathbf{v}(\mathbf{x}) = \int_{V'_F} G(\mathbf{x}, \mathbf{x}_o) \frac{\gamma}{\mu} \nabla h(\mathbf{x}_o) d\mathbf{x}_o, \quad (2.48)$$

where V'_F indicates all points in the fluid including \mathbf{x} and \mathbf{x}_o . We see that $\mathbf{v}(\mathbf{x})$ depends upon values of $\nabla h(\mathbf{x})$ over all (fluid) space, and hence, nonlocal. By introducing a function $\beta(\mathbf{x})$ such that $\beta(\mathbf{x}) = 1$ in the fluid phase and $\beta(\mathbf{x}) = 0$ in the solid phase, we can take the integral in (2.48) over all phases, including the solid,

i.e., (2.48) is multiplied by $\beta(\mathbf{x})\beta(\mathbf{x}_0)$ and then ensemble averaged, such that

$$\langle \beta(\mathbf{x})\mathbf{v}(\mathbf{x}) \rangle = \int_V \frac{\gamma}{\mu} \langle \beta(\mathbf{x})\beta(\mathbf{x}_0)G(\mathbf{x}, \mathbf{x}_0)\nabla h \rangle d\mathbf{x}_0 \quad (2.49)$$

where, now, \mathbf{x} and \mathbf{x}_0 are any points in the fluid-solid volume V (the entire domain). The reason for introducing $\beta(\mathbf{x})$ in addition to $\beta(\mathbf{x}_0)$ is the dependence of $G(\mathbf{x}, \mathbf{x}_0)$ on two points, \mathbf{x} and \mathbf{x}_0 . Next, we note that (Beran, 1968, p. 281)

$$\langle \beta(\mathbf{x})\mathbf{v}(\mathbf{x}) \rangle = \langle \beta(\mathbf{x}) \rangle \langle \mathbf{v} \rangle \quad (2.50)$$

where $\langle \mathbf{v} \rangle$ is the average fluid velocity. This is due to the fact that \mathbf{v} may take on any value in the fluid where $\beta = 1$. Note that $\langle \beta \rangle = \phi$, the porosity; consequently, $\langle \beta\mathbf{v} \rangle = \langle \mathbf{q} \rangle$ (cf. Dagan, 1989, p. 47-48, and p. 61). Defining a *stochastic* (or random) Green's function⁷,

$$G_{v_F}(\mathbf{x}, \mathbf{x}_0) = \beta(\mathbf{x})\beta(\mathbf{x}_0)G_v(\mathbf{x}, \mathbf{x}_0), \quad (2.51)$$

(2.49) becomes

$$\langle \mathbf{q} \rangle = \int_V \frac{\gamma}{\mu} \langle \nabla h(\mathbf{x}_0)G_{v_F}(\mathbf{x}, \mathbf{x}_0) \rangle d\mathbf{x}_0. \quad (2.52)$$

In general, the local pressure gradients $\nabla h(\mathbf{x}_0)$ and the function $G_{v_F}(\mathbf{x}, \mathbf{x}_0)$ are correlated. It was felt by Beran (1968, p. 282) that since $G_v(\mathbf{x}, \mathbf{x}_0)$ and $\nabla h(\mathbf{x}_0)$ arise from the solution of different boundary value problems⁸, perhaps the correlation between them is small; moreover, "even though $G(\mathbf{x}, \mathbf{x}_0)$ and $\nabla h(\mathbf{x}_0)$ may be correlated in some instances, on the average, the overall correlation will be weak" (*ibid*). According to Beran, "if we assume this correlation is weak and may be neglected, or that fluctuations in pressure gradients are small, then Eq. 2.52 becomes"

$$\langle \mathbf{q} \rangle = - \left[- \int_V \langle G_{v_F}(\mathbf{x}, \mathbf{x}_0) \rangle d\mathbf{x} \right] \frac{\gamma}{\mu} \nabla \langle h \rangle. \quad (2.53)$$

⁷For the sake of coherence, we slightly modify (the original) derivation of Beran.

⁸Note that $G_v(\mathbf{x}, \mathbf{x}_0)$ is associated with \mathbf{v} , not with h .

This decoupling of head gradient and the stochastic Green's function, and the assumption that $\langle \nabla h \rangle = \text{constant}$, leads to a local Darcian expression, with permeability

$$k = - \int_V \langle G_{v_F}(\mathbf{x}, \mathbf{x}_o) \rangle d\mathbf{x}_o \quad (2.54)$$

(where the hydraulic conductivity is defined as $K = k\gamma/\mu$). The decoupling in (2.53) is reminiscent of Eq. 2.27. It is interesting to note that, in general, the Green's function G_{v_F} is a *vector*, associated with \mathbf{q} , which renders (2.48)–(2.54) nonphysical. After introducing Eq. 2.48, Beran considers, for the rest of his derivation, only the velocity (or flux) component parallel to the gradient direction (however, we have kept the complete spatial representation, consistently, throughout the derivation). He agrees that the reduction in dimensionality at that stage is a significant one; he quotes Scheidegger (1956) that this is “appropriate for narrow-channel flow”, but adds that “to the extent that porous media can be described by capillary models, presumably this is adequate, but the reader must keep this limitation in mind” (*ibid*, p. 280). Under this approach, G_{v_i} and k^i are scalars associated with flow in the i direction.

2.3.2 Kraichnan's Equation

Frisch (1968) presents a random operator equation corresponding to the stochastic Helmholtz equation, which has a form similar to Eq. 2.15,

$$\mathcal{G} = G_o - L^{-1}\mathcal{R}\mathcal{G}. \quad (2.55)$$

$G_o = L^{-1}\delta(\mathbf{x})$ is a deterministic (“free space”) Green's function corresponding to the deterministic governing PDE with a constant coefficient. Multiplication by \mathcal{R} ,

substitution, and averaging, yields an operator equation of the form

$$\langle \mathcal{G} \rangle = G_o + L^{-1} \langle \mathcal{R} L^{-1} \mathcal{R} \mathcal{G} \rangle, \quad (2.56)$$

similar to (2.26). Frisch, then introduces the *Kraichnan equation*, based on Kraichnan's "random coupling model" and the *diagram expansion method* (or *diagram expansion*; Kraichnan, 1961)⁹. According to Frisch (1968), diagram expansions have been used successfully in quantum-electrodynamics, the many-body problem, statistical mechanics, etc., and are equivalent to all-order perturbation series. "The diagram method introduces Feynman diagrams in a very elementary way, which requires no previous knowledge of field-theoretical concepts" (Frisch, 1968, p. 104). Frisch writes the *nonlinear* Kraichnan's equation as

$$\langle \mathcal{G} \rangle = G_o + L^{-1} \langle \mathcal{R} \langle \mathcal{G} \rangle \mathcal{R} \rangle \langle \mathcal{G} \rangle. \quad (2.57)$$

Note that this is not a simple decoupling of \mathcal{R} and \mathcal{G} , as the operator L^{-1} in (2.56) is replaced here by $\langle \mathcal{G} \rangle$. According to Frisch (1968, p. 123), "the remarkable point is that the solution of (2.57) can be considered both as the exact solution of a model equation¹⁰ and as an approximate solution of the original equation. This ensure that if Eq. 2.57 can be solved, the solution will be physically acceptable". The theory we present in the next chapter leads, in a simple and direct way, to a weak approximation of the form

$$\langle \mathcal{G} \rangle = G_o + L^{-1} \langle \mathcal{R} L^{-1} \langle \mathcal{G} \rangle \mathcal{R} \rangle \langle \mathcal{G} \rangle. \quad (2.58)$$

This differs from (2.57) only by an extra L^{-1} which, we suspect, has been accidentally dropped by Frisch from (2.57). Indeed, a similar weak approximation proposed by

⁹We leave the discussion of this (important) method to future work.

¹⁰Frisch does not clarify the term "model equation".

Neuman (1993) for solute transport, coincides exactly with Kraichnan's all-order approximation for the same problem.

2.3.3 King (1987, 1989)

King (1987) defines a (random) Green's function, $\mathcal{G}(\mathbf{x}, \mathbf{x}_0)$, for steady state flow as the solution of

$$\nabla \cdot \mathbf{K}(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}, \mathbf{x}_0) = \delta(\mathbf{x} - \mathbf{x}_0), \quad (2.59)$$

subject to deterministic Neumann (prescribed flux) boundary conditions. Though he introduces \mathbf{K} as a tensor, he treats it as a statistically homogeneous and isotropic scalar, $K(\mathbf{x})$. King also defines a deterministic Green's function, $G(\mathbf{x}, \mathbf{x}_0)$, which corresponds to a uniform domain with permeability $\langle K \rangle = \text{constant}$. After some manipulations, King obtains the following (random) integro-differential (second order Fredholm) equation for $\mathcal{G}(\mathbf{x}, \mathbf{x}_0)$,

$$\mathcal{G}(\mathbf{x}, \mathbf{x}_0) = G(\mathbf{x}, \mathbf{x}_0) - \int G(\mathbf{x}, \mathbf{x}_0') \langle K \rangle \nabla_{\mathbf{x}_0'} \cdot [K'(\mathbf{x}_0') \nabla_{\mathbf{x}_0'} \mathcal{G}(\mathbf{x}_0', \mathbf{x}_0)] d\mathbf{x}_0', \quad (2.60)$$

which is equivalent to

$$\mathcal{G} = G - L^{-1} \langle K \rangle \mathcal{R} \mathcal{G}. \quad (2.61)$$

King then takes the Fourier-transform of (2.60) and employs a diagrammatic representation of successive approximations (or Neumann series), adopted from “*field theory* which is equivalent to a zero-state Potts model”¹¹, and which he considers as a perturbation expansion, to solve for the average Green's function $\langle \mathcal{G}(\mathbf{x}, \mathbf{x}_0) \rangle$. His criterion for terminating the iterative process is “that the correlation between

¹¹No reference given.

distant points (small wave number) is small". This leads him to a term $\Sigma(\mathbf{k})$ called "self-energy, by analogy with solid state and particle physics"¹², given by

$$\Sigma(\mathbf{k}) = G^{-1}(\mathbf{k}) - \langle \mathcal{G}(\mathbf{k}) \rangle^{-1} \quad (2.62)$$

where \mathbf{k} appears to be the wave number vector (not defined in the text). The latter, in turn, yields a scalar "renormalized" effective hydraulic conductivity

$$K_e = \langle K \rangle + \frac{\Sigma(\mathbf{k})}{k^2} \quad (2.63)$$

where $k = \|\mathbf{k}\|$. Finally, King develops an expression for the effective hydraulic conductivity similar to that of Gutjahr et al. (1978),

$$K_e \approx \frac{\langle K \rangle}{n} [1 - \sigma_Y^2], \quad (2.64)$$

where n is the number of dimensions. Following Matheron's (1967) conjecture, King views this as a first order approximation to

$$K_e = \langle K \rangle \exp[-\sigma_Y^2/n] = K_G \exp\left[\sigma_Y^2 \left(\frac{1}{2} - \frac{1}{n}\right)\right] \quad (2.65)$$

for lognormal K , and large σ_Y^2 , and concludes that the mean pressure and effective hydraulic conductivity do not depend on correlation length. According to King (1987), the main purpose of his work was to show that methods used in field theory can be applied to problems of flow in heterogeneous media by recovering familiar results from perturbation methods (including Bakr et al., 1978, Gutjahr et al., 1978, Mizell et al., 1982, and Dagan, 1981, 1982). He concludes that although the method presented still assumes small fluctuations in $K(\mathbf{x})$, it does not require the assumption of linear (uniform) flow, which renders it more general. We show in the next chapter

¹²No reference given.

how a method which seems analogous to the diagram method can provide better approximations to more general problems, consider both nonuniform flows and large fluctuations in $K(\mathbf{x})$.

2.3.4 Kröner (1977, 1986)

The linear stress-strain ($\sigma - \varepsilon$) relationships (Hooke's law),

$$\sigma = \mathbf{E}\varepsilon, \quad (2.66)$$

is similar to Darcy's law, σ being analogous to \mathbf{q} , ε to $-\nabla h$, and \mathbf{E} (the elastic modulus, or elasticity), is analogous to \mathbf{K} . [It should be mentioned that in the basic stress-strain law, both the stress and the strain are second rank tensors, while the elasticity may be a fourth rank tensor; consequently, the stress-strain law may be non-local from the outset (see Kröner, 1977, 1986), and the fundamental solution $G(\mathbf{x}, \mathbf{x}_0)$ is also a tensor]. Consequently, there is much in common between methods used for defining effective moduli and methods used to define effective conductivities (and thereby solving for the first moment of h).

Kröner (1977, 1986), in a work on bounds for effective elastic moduli (or elasticity) of disordered materials, considered the Lippmann-Schwinger integral equation

$$\varepsilon + \mathbf{\Gamma} \delta \mathbf{E} \varepsilon = \varepsilon_0 \quad (2.67)$$

or equivalently,

$$\varepsilon = (I + \mathbf{\Gamma} \delta \mathbf{E})^{-1} \varepsilon_0, \quad (2.68)$$

where ε_0 is the strain in some arbitrary non-random reference medium with uniform elasticity \mathbf{E}_0 (which is not necessarily isotropic, and not necessarily identical to $\langle \mathbf{E} \rangle$),

and $\delta\mathbf{E} \equiv \mathbf{E} - \mathbf{E}_o$. [Only when $\mathbf{E}_o = \langle \mathbf{E} \rangle$, then $\delta\mathbf{E} = \mathbf{E}'$, the (usual) zero-mean fluctuations]. Here Γ is an integral operator whose kernel is the “Green tensor”¹³

$$\Gamma(\mathbf{x}, \mathbf{x}_o) = \nabla \nabla G(\mathbf{x}, \mathbf{x}_o), \quad (2.69)$$

where $G(\mathbf{x}, \mathbf{x}_o)$ is defined as the Green’s function of the uniform reference medium. Because $G(\mathbf{x}, \mathbf{x}_o)$ is symmetric, the tensor Γ is self-adjoint. According to Kröner, the integral in (2.67) is defined despite the r^{-3} singularity (resulting from $\nabla \nabla G(\mathbf{x}, \mathbf{x}_o)$) in the sense of generalized functions, as long as $\delta\mathbf{E}\varepsilon$ remains bounded.

Following Beran and MacCoy (1970), Kröner defines a *prime operator* or *deviation operator*, P , which projects the deviation part of the random function on its right side, such that $Pf \equiv f' \equiv f - \langle f \rangle$, $Pfg \equiv (fg)' \equiv fg - \langle fg \rangle$, etc, where f, g are any random functions. Applying P to (2.67), we obtain, with $P\varepsilon_o = 0$,

$$\varepsilon' + P\Gamma \delta\mathbf{E}\varepsilon \equiv \varepsilon' + P\Gamma \delta\mathbf{E}\varepsilon' + P\Gamma \delta\mathbf{E}\langle \varepsilon \rangle = 0. \quad (2.70)$$

Premultiplying (2.70) by $\delta\mathbf{E}$, solving for $\delta\mathbf{E}\varepsilon'$ and taking the ensemble average, we obtain

$$\langle \delta\mathbf{E}\varepsilon' \rangle = -\langle (\mathbf{I} + \delta\mathbf{E}P\Gamma)^{-1} \delta\mathbf{E}P\Gamma \delta\mathbf{E} \rangle \langle \varepsilon \rangle. \quad (2.71)$$

On the other hand, when effective elastic moduli, \mathbf{E}_{eff} , exist,

$$\mathbf{E}_{eff} \langle \varepsilon \rangle = \langle \mathbf{E}\varepsilon \rangle \quad (2.72)$$

(which is equivalent to our $\langle \mathbf{q} \rangle = -\langle K\nabla h \rangle = -K_e \nabla \langle h \rangle$). Consequently,

$$\mathbf{E}_{eff} \langle \varepsilon \rangle = \langle \mathbf{E} \rangle \langle \varepsilon \rangle + \langle \delta\mathbf{E}\varepsilon' \rangle. \quad (2.73)$$

¹³The operator notation here is, again, confusing. Γ “alone” is an integral operator, while $\Gamma(\mathbf{x}, \mathbf{x}_o)$ denotes the *kernel* of the operator.

Hence, expanding (2.71) in binomial series (similar to Eq. 2.23, above), we obtain the Neumann series for effective moduli,

$$\mathbf{E}_{eff} = \langle \mathbf{E} \rangle - \langle \delta \mathbf{E} P \Gamma \delta \mathbf{E} \rangle + \langle \delta \mathbf{E} P \Gamma \delta \mathbf{E} P \Gamma \delta \mathbf{E} \rangle - \dots \quad (2.74)$$

The final result can be expressed in the form

$$\mathbf{E}_{eff} = \langle \mathbf{E} \mathbf{B} \rangle, \quad (2.75)$$

where

$$\mathbf{B} \equiv (\mathbf{I} + P \Gamma \delta \mathbf{E})^{-1} \equiv \mathbf{I} - P \Gamma \delta \mathbf{E} + P \Gamma \delta \mathbf{E} P \Gamma \delta \mathbf{E} \dots \quad (2.76)$$

[Note that in contrast with (2.75), $\mathbf{E}_{eff} \neq \langle \delta \mathbf{E} \mathbf{B} \rangle$.] The solution is *formal* in that the resulting multiple integrals are not calculated. Γ is formed by means of (deterministic) Green's function belonging to the relevant boundary-value problem. Thus, the kernels of these integrals consist of contractions of "Green tensors" (Γ 's) and products of parameter fluctuations. Along the line of (our) Section 2.2, Kröner remarks that "the convergence of the series is not generally established", and that "one expects to obtain convergence if the formal inequality $|\Gamma \delta \mathbf{E}| < \mathbf{I}$, and ... it may be violated in composites with large (parameter) fluctuations" (cf. Eq. 2.22).

2.3.5 Markov (1987) and Shetzen (1980)

Along the line of Schetzen (1980; below), Markov (1987) considers an infinite heterogenous medium whose heat conductivity coefficient is a statistically homogeneous and isotropic random function. The temperature field is governed by the (random) Laplace equation for heat conduction, identical to the (random) groundwater flow equation, i.e.,

$$\nabla \cdot [\mathbf{K}(\mathbf{x}) \nabla h(\mathbf{x})] = 0. \quad (2.77)$$

The approximate solution is obtained through a “Volterra-Wiener functional series”, a variant of the Neumann series, similar to (2.31),

$$h(\mathbf{x}) = T_0(\mathbf{x}) + \int T_1(\mathbf{x} - \mathbf{y})K'(\mathbf{y})d\mathbf{y} \\ + \iint T_2(\mathbf{x} - \mathbf{y}_1, \mathbf{x} - \mathbf{y}_2)K'(\mathbf{y}_1)K'(\mathbf{y}_2)d\mathbf{y}_1d\mathbf{y}_2 + \dots \quad (2.78)$$

with deterministic kernels T_0, T_1, \dots . The difference between the Volterra-Wiener series and the Neumann series is that in the former, the kernels are completely unknown and have to be identified (a “black box” approach), while in the latter, the components of the resolvent kernel are known (are usually defined as explicit functions of the corresponding Green’s function, as in Kröner’s kernels, above), but an approximation has to be made in order to make the solution tractable.

To “center” the terms in (2.78) about $\langle h \rangle$, Markov starts from

$$h(\mathbf{x}) = \mathbf{J} \cdot \mathbf{x} + \sum_{n=1}^{\infty} \int \dots \int T_n(\mathbf{x} - \mathbf{y}_1, \dots, \mathbf{x} - \mathbf{y}_n) \mathbf{P}_n^k(\mathbf{y}_1, \dots, \mathbf{y}_n) d^3\mathbf{y}_1 \dots d\mathbf{y}_n \quad (2.79)$$

where $\mathbf{J} \equiv \nabla \langle h(\mathbf{x}) \rangle$, and

$$\mathbf{P}_n^k(\mathbf{y}_1, \dots, \mathbf{y}_n) = K'(\mathbf{y}_1) \dots K'(\mathbf{y}_n) - \mathbf{M}(\mathbf{y}_1, \dots, \mathbf{y}_n) \\ \mathbf{M}(\mathbf{y}_1, \dots, \mathbf{y}_n) = \langle K'(\mathbf{y}_1) \dots K'(\mathbf{y}_n) \rangle, \quad (2.80)$$

so that

$$\langle \mathbf{P}_n^k(\mathbf{y}_1, \dots, \mathbf{y}_n) \rangle = 0, \quad (2.81)$$

$n = 1, 2, \dots$

For slightly heterogeneous media with

$$\frac{|K'(\mathbf{x})|}{\langle K \rangle} \ll 1, \quad (2.82)$$

Markov (1987) uses a perturbation method to identify the kernels in (2.78), (2.79).

This limitation may be too restrictive; it leads to

$$\frac{|\langle K^n(\mathbf{x}) \rangle|}{\langle K \rangle^n} \ll 1, \quad (2.83)$$

which, for $n = 2$, implies $\sigma_Y^2 / \langle K \rangle^2 \ll 1$.

With

$$h_n(\mathbf{x}) = \int T_n(\mathbf{x} - \mathbf{y}_1, \dots, \mathbf{x} - \mathbf{y}_n) \mathbf{P}_n^k(\mathbf{y}_1, \dots, \mathbf{y}_n) d\mathbf{y}_1 \cdots d\mathbf{y}_n, \quad (2.84)$$

and $h_o(\mathbf{x}) = \mathbf{J} \cdot \mathbf{x}$, Markov finds

$$h_1(\mathbf{x}) = \frac{1}{4\pi \langle K \rangle} \mathbf{J} \cdot \int \nabla \frac{1}{|\mathbf{x} - \mathbf{y}|} K' d\mathbf{y}, \quad (2.85)$$

so that the first kernel in (2.78) is

$$T_1(\mathbf{x}) = \frac{1}{4\pi \langle K \rangle} \mathbf{J} \cdot \nabla \frac{1}{|\mathbf{x}|}, \quad (2.86)$$

and

$$h_2(\mathbf{x}) = \frac{1}{(4\pi \langle K \rangle)^2} \iint \nabla \frac{1}{|\mathbf{x} - \mathbf{y}_1|} \cdot \nabla \nabla \frac{1}{|\mathbf{y}_1 - \mathbf{y}_2|} \cdot \mathbf{J} K'(\mathbf{y}_1) K'(\mathbf{y}_2) d\mathbf{y}_1 d\mathbf{y}_2, \quad (2.87)$$

and so on. Terms like $|\mathbf{y}_{i-1} - \mathbf{y}_i|^{-1}$, above, resemble the familiar 3-D fundamental solution (or Green's function) for the flow (Laplace) equation (2.77), and the kernel of the perturbation series above are similar to Kröner's kernels (as well as to kernels formulated earlier by Dederichs and Zeller, 1973). Delighted by this similarity, Markov remarks that "... the (Volterra-Wiener) series have thus been repeatedly employed in this theory without being recognized as such ...". Kernels proportional to the moments of K and to respective contractions of gradients of Green's function of the form $|\mathbf{x}|^{-1}$, were also constructed by Beran (1965) with the aid of variational

principles. We have seen in the previous chapter that kernels of that form dominate the (approximate) solutions of Dagan (1989), Shvidler (1962), and Matheron (1967). Similar kernels arise also in the next chapter, as part of a more general and accurate solution to the random flow equation.

We mentioned earlier that Markov's perturbation solution is based, to a large extent, on Shetzen (1980). Rather than reviewing Shetzen's work, we merely highlight its most relevant aspect. Schetzen (1980; p. 200) recognizes that

the (deterministic Neumann) series¹⁴ representation of a physical system may converge for only limited range of the system input amplitude. This problem of convergence is the same as that encountered with the Taylor series representation of a function. This similarity should be expected, since ... the Neumann series is really a Taylor series with memory, so both the Taylor and Neumann series should have the same basic limitation.

To circumvent this problem, Wiener (according to Schetzen, 1980) formed a new set of functionals from the Volterra functionals. He called them *G-functionals* because of their special orthogonality property when their input is from a Gaussian process. The advantage of the G-functional series (or Volterra-Wiener series) is that it does not suffer from the severe convergence problems of the Neumann series. In any case, the difficulties in formulating higher order closure approximations are serious enough to limit the applicability of the method to small random fluctuations (Schetzen, 1980; p. 261).

¹⁴Termed "Volterra series" by Shetzen, despite the finite integral limits.

2.4 Unsteady Flow and Semigroups

In the following, we will briefly review two works in the field of hydrology, both of which attempt to use the (truncated) Neumann and Volterra series to predict transient flow and transport in random hydraulic conductivity fields. [In fact, these are the only known works in hydrology to use successive approximations for solving the governing integrodifferential equations.]

2.4.1 Preliminaries: Time-dependent Formulation

Time-dependent (or transient) problems of groundwater flow can be formulated as

$$\frac{\partial h(t)}{\partial t} = Lh(t) + f, \quad (2.88)$$

subject to the initial condition $h(0) = h_o$. If $L = L(t)$, this is an ordinary differential equation (ODE) having the general solution (Curtain and Pritchard, 1977, p. 119)

$$h(t) = e^{Lt}h_o + \int_0^t e^{L(t-s)}f(s)ds. \quad (2.89)$$

In the case of a stochastic ODE, (2.88) can be split into a deterministic operator L and a zero mean random operator \mathcal{R} , as in (2.6), and in the absence of source/sink it becomes (Frisch, 1968, p. 114)

$$\frac{\partial h(t)}{\partial t} = [L + \mathcal{R}]h(t). \quad (2.90)$$

Then the solution is

$$h(t) = \int_0^t \exp\{L(t - \tau)\}h(0)\delta(\tau)d\tau + \int_0^t \exp\{L(t - \tau)\}\mathcal{R}(\tau)h(\tau)d\tau, \quad (2.91)$$

which is a variant of (2.89). Eq. 2.91 can be applied to stochastic *PDE*'s (where the operators depend on both time and space) by resorting to a *semigroup* formulation (e.g., Curtain and Pritchard, 1977, p. 193).

2.4.2 Semigroups of Operators

Basic Concepts

[Based on Curtain and Pritchard (1977), Bharucha-Reid (1972), Frisch (1968), and Ahmed (1991).]

Let \mathcal{X} be a Banach space (defined in Appendix B)¹⁵, and $\{T(t), t \geq 0\}$ a family of bounded linear operators in \mathcal{X} , that is, for each $t \geq 0$, $T(t) \in \Lambda(\mathcal{X})$, where $\Lambda(\mathcal{X})$ denotes the space of bounded linear operators in \mathcal{X} . The family of operators $\{T(t), t \geq 0\}$ is said to be a *semigroup of operators* in \mathcal{X} if

1. $T(0) = I$ (identity)
2. $T(t + s) = T(t)T(s) = T(s)T(t)$ for all $t, s \geq 0$
3. $\lim_{t \downarrow 0} \|T(t)u - u\| = 0$ for each $u \in \mathcal{X}$.

Properties (1) and (3) mean that the semigroup $\{T(t), t \geq 0\}$ [defined by property (2)] is *strongly continuous*. If $\lim_{t \downarrow 0} \|T(t) - I\| = 0$, the semigroup $\{T(t), t \geq 0\}$ is said to be *uniformly continuous*. Property (3) implies that $T(t)$ is strongly continuous at $t = 0$, and hence, is known as the C_0 -semigroup. Note that by the second property, T behaves as a product of exponents, and the third property looks like a *contraction*

¹⁵A Banach space is basically an n -dimensional normed vector space.

as t decreases towards zero. Indeed, if in addition $\|T(t)\| \leq 1$ for all $t \geq 0$, then $\{T(t), t \geq 0\}$ is a *contraction semigroup of operators* (Bharucha-Reid, 1972).

If we can define a bounded operator $A_t = [T(t) - I]/t$, and if $\lim_{t \downarrow 0} A_t u$ exists for each $u \in \mathcal{X}$, then there exists an operator $A \in \Lambda(\mathcal{X})$ such that $\lim_{t \downarrow 0} A_t u = Au$ as $t \downarrow 0$, which is called the *infinitesimal generator* of a semigroup $\{T(t), t \geq 0\}$ on \mathcal{X} . The infinitesimal generator of a semigroup is a bounded linear operator if and only if the semigroup is uniformly continuous (Ahmed, 1991).

Further, it can be shown that if (and only if) $A \in \Lambda(\mathcal{X})$, then there exists one and only one semigroup, and it is given by $S(t) \equiv e^{At}$, $t \geq 0$, where the latter is a uniformly continuous semigroup of operators on \mathcal{X} , and its infinitesimal generator is A (Ahmed, 1991). Note that the exponential S is reminiscent of the (homogeneous) solution of linear ordinary differential equations of the type $u' - Au = 0$.

Many linear stochastic problems can be formulated as (Frisch, 1968; Bharucha-Reid, 1972)

$$\frac{du}{dt} = \mathcal{L}u = (L + \mathcal{R})u, \quad (2.92)$$

where $u(\zeta, t)$ is for every fixed t a random variable with values in the Banach space \mathcal{X} . Assuming that L is a deterministic linear operator on \mathcal{X} , forming the infinitesimal generator of a strongly continuous semigroup of contraction operators denoted by $S(t) = e^{At}$ (see above), and that $\mathcal{R}(\zeta)$ is a random linear operator, which satisfies almost surely $\|\mathcal{R}\| \leq M$ (i.e., it is *bounded*).

Bharucha-Reid (1972) shows (after Phillips, 1953) that under the above assumptions $T(t)u$ is strongly continuous and differentiable, that [similar to (2.92)],

$$\frac{dT(t)u}{dt} = \mathcal{L}T(t)u = (L + \mathcal{R})T(t)u, \quad (2.93)$$

and that the solution $T(t)$ admits the representation

$$T(t) = \sum_{n=0}^{\infty} T_n(t), \quad (2.94)$$

where

$$T_0(t) = S(t) \quad \text{and} \quad T_n(t) = \int_0^t S(t-\tau) \mathcal{R} S_{n-1}(\tau) d\tau. \quad (2.95)$$

It follows from the above results that for almost every event ζ , the random linear operator $L + \mathcal{R}(\zeta)$ also generates a strongly continuous semigroup of operators $T(t, \zeta)$, which is given by the following perturbation formula (Bharucha-Reid, 1972; Frisch, 1968)

$$\begin{aligned} T(t, \zeta) &= \exp\{[L + \mathcal{R}(\zeta)]t\} \\ &= \exp\{Lt\} + \int_0^t \exp\{L(t-\tau)\} \mathcal{R}(\zeta) \exp\{L\tau\} d\tau \\ &+ \int_0^t \int_0^{\tau_1} \exp\{L(t-\tau)\} \mathcal{R}(\zeta) \exp\{L(t-\tau_2)\} \mathcal{R}(\zeta) \exp\{L\tau_2\} d\tau_1 d\tau_2 \\ &+ \dots \end{aligned} \quad (2.96)$$

Note that this equation can be obtained from (2.91) [or (2.104) below] by successive substitutions (or approximations) of $h(t)$ [or $u(t)$], and assuming (2.99) below. Note also that (2.96) is considered by the above authors a *perturbation* expansion, although it resembles an ordinary Volterra series (see discussion of asymptotic stability of semigroup, below). Winter et al. (1984) and Neuman et al. (1987) use the semigroup approach in conjunction with a perturbation method to predict diffusion in random velocity fields; they employ an equation like (2.96), but with powers of ε , the perturbation parameter, multiplying each term in (2.96).

Since $S(t) = e^{Lt}$ is a *contraction* operator for every fixed t , and for any contraction semigroup $\|e^{Lt}\| \leq 1$, we have

$$\|S(t)\| = \|e^{Lt}\| \leq 1 \quad (2.97)$$

(cf. Eq. 2.45) and from (2.96). It follows from here that

$$\begin{aligned} \|T(t, \zeta)\| &= \|\exp\{[L + \mathcal{R}(\zeta)]t\}\| \\ &\leq 1 + Mt + \dots + \frac{M^n t^n}{n!} + \dots = e^{Mt} \end{aligned} \quad (2.98)$$

(cf. Eq. 2.41). Hence, the perturbation expansion (2.96) converges uniformly in any bounded interval; in particular, it converges to the random operator $T(t, \zeta)$, which is the *solution operator* of Eq. 2.92. Hence (Bharucha-Reid, 1972),

$$u(t, \zeta) = T(t, \zeta)u(0, \zeta) \quad (2.99)$$

(cf. Eqs. 2.42, 2.44).

However, as was pointed out by Frisch (1968, p. 87; see also earlier discussion on the asymptotic behavior of Volterra series) the convergent expansion cannot be used to study asymptotic behavior of $u(t, \zeta)$, because as $t \rightarrow \infty$ (e.g., when approaching steady-state flow) the number of terms which must be retained to get a good approximation increases indefinitely. This is due to the presence of *secular* terms like $M^n t^n / n!$. This is also implied by the remainder (R_n) of the series $e^t = \sum_{n=0}^{\infty} t^n / n!$, which is (Arfken, 1985)

$$R_n \leq \frac{t^n e^t}{n!}. \quad (2.100)$$

As part of their *stability theory*, Curtain and Pritchard (1977) analyze the asymptotic stability of semigroups. Roughly speaking, stability implies that small perturbations to the system do not radically alter the way in which the system is operating (*ibid*). By definition, a strongly continuous semigroup \mathcal{J}_t on a Banach space \mathcal{X} is said to be asymptotically stable if there exist constants M, β , with $\beta < 0$ such that (*ibid*),

$$\|\mathcal{J}_t\| \leq M e^{\beta t} \quad (2.101)$$

where $\mathcal{J}_t \equiv T(t)$ above.

Note that (2.101) and (2.97) imply that $\|L\| \leq 0$, and that the semigroup operator has to be a strongly decaying operator (e.g., a negative exponential function). The property of generating asymptotically stable semigroup is called *Liapunov stability (ibid)*.

Given that L is an infinitesimal generator of an asymptotically stable semigroup \mathcal{J}_t , Curtain and Pritchard (1977) investigate under what conditions $L + \mathcal{R}$ is also an infinitesimal generator of an asymptotically stable semigroup. Consider, for example, Eq. 2.92 (again):

$$\frac{du}{dt} = Lu + \mathcal{R}u, \quad (2.102)$$

for which the solution is

$$u(t) = \exp\{(L + \mathcal{R})t\}u(0), \quad (2.103)$$

or, alternatively, like in Eq. 2.91,

$$u(t) = \exp\{Lt\}u(0) + \int_0^t \exp\{L(t - \tau)\}\mathcal{R}(\tau)u(\tau)d\tau. \quad (2.104)$$

They (*ibid*) show that (2.104) can be expanded in Volterra series (like Eq. 2.96), and find that in order to generate an asymptotically stable semigroup consistent with (2.101), it is necessary to have

$$\|\mathcal{R}\| < \frac{-\beta}{M}, \quad (2.105)$$

which can theoretically be satisfied for any $\beta < 0$. Practically, this means that $\|\mathcal{R}\|$ has to be sufficiently small.

Curtain and Pritchard (1977) give an example of a simple one-dimensional parabolic equation

$$\frac{\partial z}{\partial t} = \frac{\partial^2 z}{\partial x^2}, \quad (2.106)$$

such that $x \in [0, 1]$, with $z(0, t) = z(1, t) = 0$ and $z(x, 0) = z_o(x)$; they find that the semigroup \mathcal{J}_t generated by the (spatial) operator on RHS of (2.106) has the limit

$$\|\mathcal{J}_t\| \leq \exp(-\pi^2 t). \quad (2.107)$$

By comparing (2.101) and (2.105), we find $\|\mathcal{R}\| \leq \pi^2$. However, a more thorough investigation leads Curtain and Pritchard (1977) to the less restrictive condition $\|\mathcal{R}\| < \pi^4$ (which is still low).

2.4.3 Zeitoun and Braester (1991)

Zeitoun and Braester (1991) use the Neumann expansion to analyze transient flow in heterogeneous formations with variations in both permeability and porosity (ϕ) without sources/sinks. After Laplace transforming their dimensionless governing equation, they divide it into a left hand side (LHS) that resembles that of the Helmholtz equation (defined earlier), and a right hand side (RHS) which consists of (derivatives of) their random parameters, the (transformed) random response, h_D , and the Laplace transform parameter s . They then proceed with inversion of this stochastic PDE, making use of the Green's function associated with the Helmholtz operator, and get an implicit solution for h_D in the form of a stochastic integro-differential equation with h_D residing in two domain integrals (one domain integral being associated with permeability fluctuations, the other with porosity fluctuations) and boundary integrals. They eliminate h_D from the boundary integrals without explanation, and suggest to solve the remaining integro-differential equation (their Eq. 25) by Monte Carlo simulations (why not do this directly, with their original stochastic PDE?).

As an alternative approach, Zeitoun and Braester (1991) suggest to solve their integro-differential equation by means of Neumann series, starting with the boundary integrals (which now contain only the deterministic Green's function) as initial guess. They then assume that the permeability and the porosity are uncorrelated, normally distributed, zero mean (statistically) homogeneous random fields.

In an attempt to find conditions for the convergence of their solution, Zeitoun and Braester (1991) start from the general form of the Fredholm equation. They represent the kernel by its eigenvalues and eigenfunctions as

$$K(x, t) = \sum_{n=1}^{\infty} \lambda_n \phi_n(x) \phi_n(t), \quad (2.108)$$

where λ_n are eigenvalues and ϕ_n are the corresponding eigenfunctions. [This is different from Eq. 2.38 (the common definition of a kernel by its eigenvalues and eigenfunctions), where λ_n appears in the denominator; hence, we suspect that λ_n in (2.108) should be replaced by its reciprocal.] They then use successive approximations (like in Eq. 2.33) to obtain a general expression for the iterated kernel, where both porosity and permeability are random and normally distributed. They somehow obtain a *sufficiency* condition for the convergence of the Neumann series,

$$\sigma_\nu < \frac{1}{\max |\lambda_n|}, \quad (2.109)$$

where $\nu = K'/\langle K \rangle$ is the relative permeability fluctuations, so that σ_ν (the standard deviation of ν) is equivalent to the coefficient of variation (*CV*) of the permeability. The authors, however, do not give any indication as to the (order of) magnitude of the eigenvalues associated with their kernel. From the examples in the preceding sections, we expect the eigenvalues to be the arguments of Fourier series of the kernel (or Green's function), and hence, be less than or equal to one, which implies

large σ_ν , contrary to all of the (systematic) analyses reviewed here. If, however, λ_n in (2.108) is replaced by its reciprocal, $\hat{\lambda}_n = \lambda_n^{-1}$ (as recommended above), then $\max |\hat{\lambda}_n| = [\min |\lambda_n|]^{-1}$, and consequently, the sufficiency condition (2.109) becomes $\sigma_\nu < \min |\lambda_n|$.

After setting the porosity variations to zero, and dropping the boundary conditions, they find a sufficiency condition for convergence of the Neumann series of their integro-differential equation (with random K , only),

$$\sigma_\nu < 1 + \frac{s}{\max(k_n^2)}, \quad (2.110)$$

where k_n is the wave number, and s is the Laplace transform parameter. As part of their analysis, the authors indicate that s decreases with time, and at the limit, $t \rightarrow \infty$ (steady-state), $s \rightarrow 0$. Consequently, at steady-state, $\sigma_\nu < 1$, which implies small variability of K . This is contrary to (2.109), and to the authors' belief that their Neumann series expansion method is amenable to larger variability than standard perturbation methods. Indeed, in their following example Zeitoun and Braester (1991) use $\sigma_\nu \leq 0.5$, and their results are inconclusive.

Above all, Zeitoun and Braester (1991) do not show explicitly how moments of the response (e.g., heads, fluxes) are related to moments of the random hydraulic conductivity field.

Serrano and Unny (1985, 1989, 1990,1992)

Serrano et al. (1985) apply the semigroup approach to "evolution equations" similar to (2.92),

$$\frac{\partial h}{\partial t} + A h = f, \quad (2.111)$$

with initial conditions $u(x, 0) = u_o(x)$ (without specifying boundary conditions). For the homogeneous case (i.e., $f = 0$), they write the inverse of (2.111) as

$$u(t) = \mathcal{J}_t u_o, \quad (2.112)$$

where \mathcal{J}_t , termed “evolution operator”¹⁶, is very similar to $T(t)$ in (2.99); if it is also a contraction semigroup, then $\mathcal{J}_t \equiv T(t)$. Following Curtain and Pritchard (1977; p. 156) Serrano and Unny (*ibid*) present the solution to (2.111) in the form

$$u(t) = \mathcal{J}_t u_o + \int_0^t \mathcal{J}_{t-s} f(s) ds, \quad (2.113)$$

where \mathcal{J}_t , is, again, similar to the Green’s function (hence, an “impulse response function”). For the sake of simplicity and without loss generality, the semigroup operators of Unny and Serrano can be viewed as Green’s functions¹⁷ associated with the deterministic differential operators (as discussed below).

Unny (1989) considers the one-dimensional groundwater continuity equation and the one dimensional dispersion equation, both having the general form of (2.92) and (2.93), i.e.,

$$\frac{\partial h}{\partial t} + \mathcal{L} h = f \quad (2.114)$$

with Dirichlet type boundary conditions (although the dispersion equation should not have a well-defined downstream BC), and some fixed initial conditions $h(\mathbf{x}, 0) = h_o(\mathbf{x})$. He first decomposes the random differential operator into a deterministic and a random part,

$$\mathcal{L} = L + \mathcal{R}, \quad (2.115)$$

¹⁶In their later publications, Unny (1989) and Serrano (1992) call it *semigroup*; Serrano (1990, 1992) refers to it in one place as both a *strongly continuous semigroup* and an *impulse response function*.

¹⁷Though, in a narrow sense, because of restrictions associated with semigroups.

then, writes the solution as a Volterra equation of the second kind with the integral decomposed into a deterministic and a stochastic part (f is considered nonrandom),

$$h = \mathcal{J}_t h_o + \int_0^t \mathcal{J}_{t-s} f(s) ds - \int_0^t \mathcal{J}_{t-s} \mathcal{R} h(s) ds \quad (2.116)$$

(cf. Eq. 2.113), where \mathcal{J}_t (the kernel) is a semigroup associated with the operator L (although in the text, \mathcal{J}_t is associated with \mathcal{L} , not with L , it seems to be inconsistent with Unny's subsequent derivation). Unny then expresses h in the right hand side as an infinite series $h = \sum_{i=1}^{\infty} h_i$ (cf. Eq. 2.94),

$$h = \mathcal{J}_t h_o + \int_0^t \mathcal{J}_{t-s} f(s) ds - \int_0^t \mathcal{J}_{t-s} \mathcal{R} (h_1 + h_2 + h_3 + \dots) ds. \quad (2.117)$$

By identifying h_1 as the preceding term, $\int_0^t \mathcal{J}_{t-s} f(s) ds$, and so on, such that each h_i is expressed in terms of the preceding h_{i-1} , the random integral is expanded in what seems to be the Volterra series, similar to (2.23) and (2.96),

$$\begin{aligned} h &= \mathcal{J}_t h_o + \int_0^t \mathcal{J}_{t-s} f(s) ds - \int_0^t \int_0^s \mathcal{J}_{t-s} \mathcal{R} \mathcal{J}_{s-\tau} f(\tau) d\tau ds \\ &\quad - \int_0^t \int_0^s \int_0^\tau \mathcal{J}_{t-s} \mathcal{R} \mathcal{J}_{s-\tau} \mathcal{R} \mathcal{J}_{\tau-\xi} f(\xi) d\xi d\tau ds - \dots \end{aligned} \quad (2.118)$$

According to Unny (1989) the last term in the series contains h_i , which (however) seems to be inconsistent with both the rationale of the Volterra series¹⁸ and his following derivation (Eq. 2.119 below). According to Unny,

The basic idea here is that a random semigroup operator, which may be difficult to derive in particular cases, can be determined in an easily computable series by decomposition of the differential operator into a deterministic operator whose semigroup is known or found with little effort,

¹⁸Which eliminates the necessity in the separation $\langle \mathcal{R}h \rangle \rightarrow \langle \mathcal{R} \rangle \langle h \rangle$.

and a random operator whose contribution to the total semi group can be found in series form. Basically, convergence arises because the succeeding terms are multiples of increasing number of negative exponentials . . . it is known that the series converges rapidly, and that in some circumstances considering one term is sufficiently accurate . . .

Unny (1989) and Serrano (1990; discussed next) seem to be unaware of problems associated with secular terms lack of asymptotic convergence of their series expansions, and adhere to their particular “strongly dissipating” (or decaying) semigroup operators (or Green’s functions). However, Serrano (1992) uses relatively short time intervals in his seminumeric analysis (probably, to avoid the problem with secular terms).

Based on his latter argument, Unny (1989) truncates the series after one term, then takes ensemble average, while separating between the kernel and the forcing function (or source term), based on statistical independence of the two, to obtain the expected value of the response,

$$\langle h \rangle = \mathcal{J}_t h_o + \int_0^t \mathcal{J}_{t-s} \langle f(s) \rangle ds - \int_0^t \int_0^s \mathcal{J}_{t-s} \langle \mathcal{R} \mathcal{J}_{s-\tau} \rangle \langle f(\tau) \rangle d\tau ds. \quad (2.119)$$

[As it will be shown in the sequel, we find Unny’s and Serrano’s attempts to validate the above truncation to be unconvincing.]

Serrano (1992) uses semianalytical methods in conjunction with truncated Volterra series to solve the one-dimensional perturbed (but not random) transport dispersion equation,

$$\frac{\partial c}{\partial t} - D \frac{\partial^2 c}{\partial x^2} + v \frac{\partial c}{\partial x} = -\delta v \frac{\partial c}{\partial x}, \quad (2.120)$$

subject $c(-\infty) = c(\infty) = 0$, and $c(x, 0) = c_o(x)$, where c denotes concentration, D is

the (deterministic) dispersion coefficient, v is the velocity, and δv is a (deterministic) velocity deviation (a fraction of u), which represents velocity perturbation; the perturbed part is treated as a forcing function. The semigroup operator associated with the deterministic portion of (2.120) appears as a negative exponential, and resembles the Green's function (actually the fundamental solution) associated with the differential operator on the LHS of (2.120). The resulting integro-differential equation has the form of (2.113) with c appearing on both sides of the equation,

$$c(x, t) = \mathcal{J}(x, t)c_o - v \int_0^t \mathcal{J}(x, t - \tau) \frac{\partial c_i}{\partial x} d\tau. \quad (2.121)$$

Serrano then uses successive approximations to evaluate $c(x, t)$ at some distance x , given $\delta v = v$. The strong dissipative nature of the kernel (the semigroup) brings about near-convergence to the exact solution within two iterations (i.e., with the first two terms in the series), for a given time (though, with a relatively small dispersion coefficient). Serrano shows that for such a system (i.e., 1-D Volterra series with a strongly dissipative operator), when the kernel and the velocity fluctuations are bounded (say by M), the series should converge if $Mt/2 < 1$, where M is the bound on the absolute velocity fluctuations, and t is the elapsed time. Consequently, Serrano concludes that for sufficiently small "time steps", solutions for the random one-dimensional concentration field can be reached even for relatively high velocity fluctuations. It should be mentioned, however, that so far, Serrano's analysis is of deterministic nature, and does not guarantee to work even with low variances of random parameters (or velocities, in this case).

Next, Serrano (1992) regards v in (2.120) as a correlated random field, where δv is replaced by $v' = v - \langle v \rangle$ (zero-mean fluctuations), and estimates the mean and variance of the concentration field (i.e., the response). Despite the fact that the

LHS of (2.120) is no longer deterministic, Serrano continues to use the deterministic semigroup (or impulse response function) in his derivations. In other words, he does not consider that not only v but also c is random. Serrano (incorrectly) equates the ensemble mean concentration with the zero order approximation of c (i.e., the deterministic solution with $\langle v \rangle$ as the velocity in (2.120) after dropping its RHS). He then generates one realization of v with a low variance ($CV = 0.5$) and long integral scale (about three times larger than the 1-D interval of interest), and compares it with the deterministic solution (his “mean” concentration). Of course, the simulated concentration field looks different than the mean. The single result is inconclusive.

Next, Serrano estimates the concentration variance as a function of time and space, as (his Eq. 15)

$$\sigma_{c(x,t)}^2 = \int_0^t \int_0^t \mathcal{J}(x, s, t - \tau) \mathcal{J}(x, \xi, t - \gamma) C_v(s - \xi) \frac{\partial c^o(s, \tau)}{\partial s} \frac{\partial c^o(\xi, \gamma)}{\partial \xi} d\gamma d\tau, \quad (2.122)$$

where $C_v(s - \xi)$ is the covariance of v , and c^o denotes zero order approximation of (or, the deterministic solution for) the concentration, c . Serrano does not explain how he derived (2.122); it seems to stem from (2.121) above. There is some similarity between (2.122) and a recent derivation by Neuman (1993). Serrano’s calculated standard deviation of concentration versus distance indicates a direct dependence of concentration variance on concentration gradient, as expected from (2.122).

Serrano (1992) uses a similar approach to analyze the one-dimensional transient groundwater *flow* in a random transmissivity field. He separates the governing equation to a “regular” LHS and a “random” RHS, in a manner similar to (2.120),

$$\frac{\partial h}{\partial t} - \frac{\langle T \rangle}{S} \frac{\partial^2 h}{\partial x^2} = \frac{i(t)}{S} + \mathcal{R}(x)h, \quad (2.123)$$

with $h(0, t) = h_1$, $h(l, t) = h_2$, $h(x, 0) = h_o(x)$ as boundary and initial conditions.

T and S are transmissivity and storativity respectively ($S = \text{const.}$), i is recharge, and the random operator \mathcal{R} is given by

$$\mathcal{R}(x)h = \frac{1}{S} \left(T'(x) \frac{\partial^2}{\partial x^2} + \frac{\partial T'(x)}{\partial x} \frac{\partial}{\partial x} \right) h, \quad (2.124)$$

where $T' = T - \langle T \rangle$ is zero mean fluctuations. Again, serrano consider the LHS of (2.123) as “deterministic”, ignoring the fact that h is random.

Next, Serrano eliminates the (disturbing) time dependence by discretizing the time domain in the differential equation at equal intervals Δt in the finite difference way, with x as the only independent variable, such that (2.123) reduces to

$$-\frac{d^2 h}{dx^2} + a^2 h = b i + a^2 h_o + \mathcal{R}h, \quad (2.125)$$

where $h_o(x)$ is the head at the previous time step, $a^2 = S/(\langle T \rangle \Delta t)$, and $b = 1/\langle T \rangle$. Serrano then inverts (2.125) to get $h(x)$ as a function of two (domain) integrals, the kernels of which are the Green’s function associated with the deterministic version of the LHS of (2.125), i.e.,

$$h(x) = \int_0^l G(x, \xi) [b i(\xi) + a^2 h_o(\xi)] d\xi + \int_0^l G(x, \xi) \mathcal{R}(\xi) h(\xi) d\xi. \quad (2.126)$$

Again, successive approximations (or Neumann series) must be employed in order to solve (2.126). The mean head is approximated by the zero order approximation, i.e.,

$$\langle h(x) \rangle = a^2 \int_0^l G(x, \xi) h_o(\xi) d\xi + \int_0^l G(x, \xi) i d\xi. \quad (2.127)$$

Similar derivations are repeated by Serrano (1990) in his analysis of (again, one-dimensional) infiltration in random diffusivity fields. He believes that “since (the Volterra series of semigroups) is not a perturbation approximation, arbitrarily large

variances in the stochastic terms can be included” (Serrano, 1990). Here, too, Serrano uses the deterministic impulse response function as the kernel of his random integrals, and replaces the mean of the response (water content) by its zero order approximation. In order to verify his “large variance” assumption, he simulates infiltration in random diffusivity fields generated with extremely low variances ($\sigma_D^2 \approx CV \approx 10^{-4}$). As expected (because of the highly nonlinear/sensitive $K(h)$), even then, for a particular (single) realizations, “the comparison of observed and simulated water content at corresponding times was not satisfactory” (Serrano, 1990, p. 707), however (as expected from the low variances), “the observed means were in good agreement with the simulated means” (*ibid*).

In conclusion, although the works of Serrano and Unny¹⁹ present some interesting ideas, particularly the use of semigroups and the semianalytic (or seminumeric) approach, they seem to be limited in several ways. In particular, Serrano, like Zeitoun and Braester (1991), seems to “miss the point” in that he tries to verify stochastic theories through individual realizations rather than through *moments* (as discussed in the first section). Particularly disturbing is the lack of distinction between expected value (ensemble mean) and single realization of the response in Serrano’s work. The inclusion of more than two terms in their Volterra series renders the solutions intractable, even in one dimension; their argument that it can be extended to higher dimensions in a straightforward manner has not been demonstrated.

Serrano and Unny have never verified their “large variance” claim. They do not address the problems associated with the asymptotic behavior of the Volterra series, particularly the appearance of secular terms and the bounds on perturbations.

¹⁹The emphasis here is on Serrano’s work.

The (above mentioned) asymptotic stability analysis and the contraction theorems imply that (strongly decaying) semigroup operators must be associated with *small bounded perturbations* (at least for one-dimensional models).

From the introduction to this section, it seems that several assumptions have to be made before resorting to semigroup formulations. The Green's function approach is simple and intuitive; except for linearity, it does not rely on any assumption with respect to the particular operator used, and it is not restricted to small variability of the fields. The natural question arises: why should one prefer semigroups to Green's functions? Unfortunately, Unny and Serrano do not address this question.

2.4.4 Conclusion

The above two examples demonstrate the gap between *formal* solutions of stochastic PDE's, expressed by Neumann and Volterra series, and working solutions with truncated series. None of the works described above has shown a rigorous justification for the truncation after one or two terms, nor have they addressed (directly) the question of asymptotic behavior; instead, the authors in both cases brought unrealistic (or irrelevant) examples with extremely low variability to "prove" that low order truncation of series is adequate.

In the following chapter we introduce a new integrodifferential theory which does not resort to Neumann series (or its variants); instead of truncating asymptotic series it leads to an approximation which improves as the information content (and hence, the conditioning) increases; consequently, it is not restricted to low variability of the original field (but yet, to low variability of the conditional field; see Neuman, 1993). The closest analogy to this theory is Kraichnan's equation (2.57) obtained

through the diagram method.

CHAPTER 3

NONLOCAL THEORY, EFFECTIVE CONDUCTIVITIES AND WEAK APPROXIMATION

3.1 Problem Definition and Nonlocal Formalism

Most groundwater flow models are based on the premise that Darcy's law applies on a scale (or scales) smaller than the dimensions of the flow domain being modelled. However, this scale is seldom specified in an unambiguous manner and its relationship to scales of measurement (Cushman, 1987), model discretization or quantity and quality of available data (Clifton and Neuman, 1982), are generally ill-defined. This chapter addresses formally such issues of scale and information content by starting from the premise that Darcy's law,

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}), \quad (3.1)$$

applies when the flux $\mathbf{q}(\mathbf{x})$, the hydraulic conductivity $K(\mathbf{x})$, and the hydraulic gradient $\nabla h(x)$ are representative of a bulk volume (support) ω centered about the point \mathbf{x} , such that (1) ω is small compared to the flow domain, Ω , and (2) $K(\mathbf{x})$ can be evaluated locally at points \mathbf{x} with the aid of existing measurement and interpretive methods. We refer to values of $K(\mathbf{x})$ as scalar values obtained in this manner as "local measurements" with support ω (the important case of variable support is beyond the scope of this work). Given a sufficient number of reliable measurements with such supports, it is possible in principle to test the validity of Darcy's law on the local scale ω (for example, by conducting a number of flow tests under different

pressure gradients and demonstrating that the resultant flow rate is proportional to this gradient, as has been done during the packer testing of fractured granites near Oracle, Arizona, by Hsieh et al., 1983). If it is thus found experimentally valid, Darcy's law can be regarded as a local constitutive relationship. Upon combination with the principle of mass conservation one obtains an operational theory of groundwater flow which, as advocated by Cushman (1984, 1986, 1991), has a clearly defined observational basis.

To keep matters as simple as possible, we consider the steady state continuity equation

$$-\nabla \cdot \mathbf{q}(\mathbf{x}) + f(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega \quad (3.2)$$

subject to the boundary conditions

$$h(\mathbf{x}) = H(\mathbf{x}) \quad \mathbf{x} \in \Gamma_D \quad (3.3)$$

and

$$-\mathbf{q}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = Q(\mathbf{x}) \quad \mathbf{x} \in \Gamma_N \quad (3.4)$$

Here $f(\mathbf{x})$ is a randomly prescribed source function, $H(\mathbf{x})$ is a randomly prescribed head on Dirichlet boundary segments Γ_D , $Q(\mathbf{x})$ is a randomly prescribed flux into Ω across Neumann boundary segments Γ_N , $\mathbf{n}(\mathbf{x})$ is a unit outward normal to the boundary Γ , and Γ is the union of Γ_D and Γ_N . Rather than considering cross-correlations between the source and boundary functions $f(\mathbf{x})$, $H(\mathbf{x})$, and $Q(\mathbf{x})$ as done formally by Cheng and Lafe (1991) we stipulate, for simplicity, that they be prescribed in a statistically independent manner. In complex geologic media the constitutive parameter, $K(\mathbf{x})$, often varies from point to point in an apparently random fashion. This has led many groundwater hydrologists to adopt the geostatistical

working hypothesis that $K(\mathbf{x})$ can be viewed as a random field (Delhomme, 1979; Neuman, 1982, 1984; de Marsily, 1986). Though $K(\mathbf{x})$ may exhibit directional dependence at any point \mathbf{x} , we continue in the tradition of existing stochastic subsurface flow theories (e.g., Dagan, 1989) by treating it as a scalar. A rationale for treating $K(\mathbf{x})$ values from local packer tests in fractured crystalline rocks as representative of a random field that is scalar on a local scale, but exhibits large-scale statistical and hydraulic anisotropy due to spatial nonuniformity and the effect of fracture orientations, has been offered by Neuman (1987) and applied successfully to multi-scale field data by Neuman and Depner (1988).

To develop a groundwater flow model based on (3.1) – (3.4), one must have some information about $K(\mathbf{x})$. We assume that $K(\mathbf{x})$ has been determined at selected points \mathbf{x} by means of standard methods such as local pumping or packer tests. We further assume that based on these data, one can come up with a relatively smooth unbiased estimate $\kappa(\mathbf{x})$ of the unknown random function $K(\mathbf{x})$. Clearly, $\kappa(\mathbf{x})$ is a deterministic function which, though slowly varying in space, nevertheless provides (an estimated) representation of how the local valued scalar function $K(\mathbf{x})$ varies from point to point; $\kappa(\mathbf{x})$ is a smoothed, but not upscaled, version of $K(\mathbf{x})$. The question how to obtain such an estimate is not central to our work and remains outside its scope. Nevertheless, we do recall later two known geostatistical methods which, under some practical working hypotheses, make possible the determination of $\kappa(\mathbf{x})$ on the basis of local $K(\mathbf{x})$ data in a way which minimizes (though does not completely eliminate) bias and provides information about the second spatial moments of the associated estimation errors. The latter are defined as

$$K'(\mathbf{x}) \equiv K(\mathbf{x}) - \kappa(\mathbf{x}) \quad \langle K'(\mathbf{x}) \rangle_{\kappa} \equiv 0 \quad (3.5)$$

where $\langle K'(\mathbf{x}) \rangle_\kappa$ is the conditional ensemble mean of hydraulic conductivity fluctuations about $\kappa(\mathbf{x})$, the subscript κ indicating that conditioning is done on the same data as those used to obtain $\kappa(\mathbf{x})$. Equivalently, $\langle K'(\mathbf{x}) \rangle_\kappa$ is the conditional ensemble mean of the estimation errors associated with $\kappa(\mathbf{x})$. It follows that $\kappa(\mathbf{x})$ is the conditional mean of $K(\mathbf{x})$, $\kappa(\mathbf{x}) = \langle K(\mathbf{x}) \rangle_\kappa$. One of the proposed methods treats $\kappa(\mathbf{x})$ as a global drift which does not generally pass through the measurement points and results in zero mean, wide-sense (second-order) stationary estimation errors $K'(\mathbf{x})$. The other method treats $\kappa(\mathbf{x})$ as a kriging estimate which does pass through the measurement points (unless they are corrupted by measurement noise) and results in a zero mean, nonstationary $K'(\mathbf{x})$ field. Other methods to determine $\kappa(\mathbf{x})$ may be equally acceptable or better. Whereas our theory assumes that an unbiased estimate $\kappa(\mathbf{x})$ of $K(\mathbf{x})$ is available *a priori*, there is nothing in it to prevent one from improving (or determining) $\kappa(\mathbf{x})$ and/or other relevant parameters *a posteriori* by means of suitable inverse methods. The latter implies conditioning not only on $K(\mathbf{x})$ but also on head and flux data. Though the theory does not require any prior assumptions about the probability distribution of $K'(\mathbf{x})$, the same is not necessarily true about some of its proposed applications and simplifications; this will become clear later.

Consider the conditional ensemble moments $\langle h(\mathbf{x}) \rangle_\kappa$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$, where again the subscript κ indicates that conditioning is done on those data used to obtain $\kappa(\mathbf{x})$. We define the latter so as to form conditionally unbiased predictors of $h(\mathbf{x})$ and $\mathbf{q}(\mathbf{x})$ according to

$$h(\mathbf{x}) = \langle h(\mathbf{x}) \rangle_\kappa + h'(\mathbf{x}) \quad \langle h'(\mathbf{x}) \rangle_\kappa \equiv 0 \quad (3.6)$$

$$\mathbf{q}(\mathbf{x}) = \langle \mathbf{q}(\mathbf{x}) \rangle_\kappa + \mathbf{q}'(\mathbf{x}) \quad \langle \mathbf{q}'(\mathbf{x}) \rangle_\kappa \equiv 0 \quad (3.7)$$

where $h'(\mathbf{x})$ and $\mathbf{q}'(\mathbf{x})$ are the associated prediction errors. The aim of our work is

to derive theoretical and working equations satisfied by the predictors $\langle h(\mathbf{x}) \rangle_\kappa$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$, and by second moments of the prediction errors $h'(\mathbf{x})$ and $\mathbf{q}'(\mathbf{x})$.

Taking the conditional ensemble mean of (3.2) – (3.4) gives

$$-\nabla \cdot \langle \mathbf{q}(\mathbf{x}) \rangle_\kappa + \langle f(\mathbf{x}) \rangle = 0 \quad \mathbf{x} \in \Omega \quad (3.8)$$

subject to the boundary conditions

$$\langle h(\mathbf{x}) \rangle_\kappa = \langle H(\mathbf{x}) \rangle \quad \mathbf{x} \in \Gamma_D \quad (3.9)$$

and

$$-\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa \cdot \mathbf{n}(\mathbf{x}) = \langle Q(\mathbf{x}) \rangle \quad \mathbf{x} \in \Gamma_N \quad (3.10)$$

where $\langle f(\mathbf{x}) \rangle$, $\langle H(\mathbf{x}) \rangle$, and $\langle Q(\mathbf{x}) \rangle$ are prescribed (unconditional) first moments of the statistically independent random source and boundary functions $f(\mathbf{x})$, $H(\mathbf{x})$, and $Q(\mathbf{x})$. It is thus evident that $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ satisfies a standard continuity equation driven by ensemble mean source and boundary functions. Taking the conditional ensemble mean of (3.1), considering (3.5) – (3.7), gives

$$\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa = -\kappa(\mathbf{x})\nabla \langle h(\mathbf{x}) \rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x}) \quad \mathbf{r}_\kappa(\mathbf{x}) = -\langle K'(\mathbf{x})\nabla h'(\mathbf{x}) \rangle_\kappa \quad (3.11)$$

The predictor $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ of the flux $\mathbf{q}(\mathbf{x})$ is thus seen to consist of a Darcian flux component, $-\kappa(\mathbf{x})\nabla \langle h(\mathbf{x}) \rangle_\kappa$, and a residual flux component, $\mathbf{r}_\kappa(\mathbf{x}) = -\langle K'(\mathbf{x})\nabla h'(\mathbf{x}) \rangle_\kappa$. The connotation “residual” is not meant to imply that $\mathbf{r}_\kappa(\mathbf{x})$ is necessarily small.

The residual flux is an unknown mixed moment of hydraulic conductivity and gradient fluctuations about their respective conditional means. To render the deterministic equations (3.8) – (3.11) solvable for the dependent variable $\langle h(\mathbf{x}) \rangle_\kappa$, it is necessary to express $\mathbf{r}_\kappa(\mathbf{x})$ as a function (or functional) that contains parameters (or kernels) which do not depend on $\langle h(\mathbf{x}) \rangle_\kappa$. In the stochastic groundwater literature

(cf., Dagan, 1989, and previous chapter) it is common to treat $Y(\mathbf{x}) = \ln K(\mathbf{x})$ as a stationary field with small variance, $\sigma_Y^2 \ll 1$, to expand the corresponding unconditional mixed moment in powers of σ_Y^2 , and to retain only a few leading terms of this expansion. Consequently, the results are limited to mildly heterogeneous media. Recently, Zeitoun and Braester (1991) have attempted to overcome this small variance limitation by expansion in Neumann series according to an approach proposed by Adomian (1983). Unfortunately, as seen in the first chapter, the convergence, and more importantly, the asymptotic behavior of the Neumann series expansion is limited to a relatively narrow range of conditions and we are therefore not convinced that this approach applies to large σ_Y^2 (all the examples quoted by Zeitoun and Braester concern $\sigma_Y^2 \ll 1$).

A key result of this work is the following development in of a formal expression for $\mathbf{r}_\kappa(\mathbf{x})$ which is rigorously valid for a broad class of $K(\mathbf{x})$ fields, including fractals, under arbitrary steady state flow regimes in either bounded or unbounded domains Ω (other works which consider boundary effects on homogeneous fields in a somewhat related manner have been published by Naff and Vecchia, 1986, and Rubin and Dagan, 1988, 1989)

To develop formal expressions for the residual flux $\mathbf{r}_\kappa(\mathbf{x})$ in (3.11), we restrict consideration to hydraulic conductivity fields $K(\mathbf{x})$, source functions $f(\mathbf{x})$, and boundary functions $H(\mathbf{x})$, $Q(\mathbf{x})$ for which (3.1) – (3.4) admit a solution in all but a few pathological cases. This means that there exists a random Green's function $\mathcal{G}(\mathbf{x}', \mathbf{x})$ which satisfies almost surely, or with probability 1 (i.e., for most, but not necessarily all, realizations of $K(\mathbf{x})$, $f(\mathbf{x})$, $H(\mathbf{x})$, and $Q(\mathbf{x})$), the random Poisson

equation

$$\nabla_{\mathbf{x}'} \cdot [K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x})] + \delta(\mathbf{x}' - \mathbf{x}) = 0 \quad \mathbf{x}', \mathbf{x} \in \Omega \quad (3.12)$$

subject to the homogeneous boundary conditions

$$\mathcal{G}(\mathbf{x}', \mathbf{x}) \equiv 0 \quad \mathbf{x}' \in \Gamma_D \quad (3.13)$$

and

$$\nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') \equiv 0 \quad \mathbf{x}' \in \Gamma_N \quad (3.14)$$

where the grad operator $\nabla_{\mathbf{x}'}$ is taken with respect to the space coordinates \mathbf{x}' .

Expressing (3.1) – (3.4) in terms of \mathbf{x}' , substituting (3.1) into (3.2), multiplying by $\mathcal{G}(\mathbf{x}', \mathbf{x})$, and integrating over Ω yields

$$\int_{\Omega} \{\nabla_{\mathbf{x}'} \cdot [K(\mathbf{x}') \nabla_{\mathbf{x}'} h(\mathbf{x}')] + f(\mathbf{x}')\} \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}' = 0 \quad (3.15)$$

Applying Green's identity twice, considering (3.13) and (3.14), gives

$$\begin{aligned} \int_{\Omega} \{\nabla_{\mathbf{x}'} \cdot [K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x})] h(\mathbf{x}') + f(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x})\} d\mathbf{x}' = \\ \int_{\Gamma_D} K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') h(\mathbf{x}') d\mathbf{x}' - \int_{\Gamma_N} K(\mathbf{x}') \nabla_{\mathbf{x}'} h(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}'. \end{aligned} \quad (3.16)$$

Substituting (3.3) – (3.4) and (3.12) into (3.17) while considering (3.1) yields, with probability 1,

$$\begin{aligned} h(\mathbf{x}) &= \int_{\Omega} f(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \\ &- \int_{\Gamma_D} K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') h(\mathbf{x}') d\mathbf{x}' + \int_{\Gamma_N} Q(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \end{aligned} \quad (3.17)$$

Similarly, we define the deterministic Green's function $G_{\kappa}(\mathbf{x}', \mathbf{x})$ as the solution of the deterministic Poisson equation

$$\nabla_{\mathbf{x}'} \cdot [\kappa(\mathbf{x}') \nabla_{\mathbf{x}'} G_{\kappa}(\mathbf{x}', \mathbf{x})] + \delta(\mathbf{x}' - \mathbf{x}) = 0 \quad \mathbf{x}', \mathbf{x} \in \Omega \quad (3.18)$$

subject to the homogeneous boundary conditions

$$G_\kappa(\mathbf{x}', \mathbf{x}) \equiv 0 \quad \mathbf{x}' \in \Gamma_D \quad (3.19)$$

and

$$\nabla_{\mathbf{x}'} G_\kappa(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') \equiv 0 \quad \mathbf{x}' \in \Gamma_N \quad (3.20)$$

Then, in analogy to (3.17),

$$\begin{aligned} h_\kappa(\mathbf{x}) &= \int_\Omega \langle f(\mathbf{x}') \rangle G_\kappa(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \\ &\quad - \int_{\Gamma_D} \kappa(\mathbf{x}') \nabla_{\mathbf{x}'} G_\kappa(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') \langle H(\mathbf{x}') \rangle d\mathbf{x}' + \int_{\Gamma_N} \langle Q(\mathbf{x}') \rangle G_\kappa(\mathbf{x}', \mathbf{x}) d\mathbf{x}'. \end{aligned} \quad (3.21)$$

Upon combining (3.1) and (3.2), the resulting equation can be recast as

$$\mathcal{L}h(\mathbf{x}) + f(\mathbf{x}) = 0 \quad \text{or} \quad (L + \mathcal{R})h(\mathbf{x}) + f(\mathbf{x}) = 0 \quad (3.22)$$

where \mathcal{L} and \mathcal{R} are stochastic operators, L a deterministic operator, defined such that

$$\mathcal{L} = \nabla \cdot [K(\mathbf{x})\nabla] = L + \mathcal{R} \quad \mathcal{R} = \nabla \cdot [K'(\mathbf{x})\nabla] \quad (3.23)$$

$$\langle \mathcal{L} \rangle_\kappa = L = \nabla \cdot [\kappa(\mathbf{x})\nabla] \quad \langle \mathcal{R} \rangle_\kappa \equiv 0 \quad (3.24)$$

Since \mathcal{G} exists almost surely, there also exists with probability 1 an inverse operator \mathcal{L}^{-1} such that $\mathcal{L}^{-1}\mathcal{L} = \mathcal{L}\mathcal{L}^{-1} = 1$ and

$$h(\mathbf{x}) = \mathcal{L}^{-1}[-f(\mathbf{x})] \quad \mathbf{x} \in \Omega \quad (3.25)$$

It is clear that $\mathcal{L}^{-1}[-f(\mathbf{x})]$ is given by the right-hand-side of (3.17). Similarly, there exists an inverse operator L^{-1} such that $L^{-1}L = LL^{-1} = 1$ and

$$h_\kappa(\mathbf{x}) = L^{-1}[-\langle f(\mathbf{x}) \rangle] \quad (3.26)$$

where $L^{-1}[-\langle f(\mathbf{x}) \rangle]$ is given by the right-hand-side of (3.22).

From the almost sure existence of \mathcal{L}^{-1} and the existence of L^{-1} it is clear that (3.25) can be rewritten with probability 1 as

$$h(\mathbf{x}) = \mathcal{L}^{-1}[-f(\mathbf{x})] = (1 + L^{-1}\mathcal{R})^{-1}L^{-1}[-f(\mathbf{x})] \quad (3.27)$$

One can easily verify through premultiplication by $(1 + L^{-1}\mathcal{R})$, considering the identity $(1 + L^{-1}\mathcal{R})(1 + L^{-1}\mathcal{R})^{-1} = 1$, that

$$(1 + L^{-1}\mathcal{R})^{-1} = 1 - (1 + L^{-1}\mathcal{R})^{-1}L^{-1}\mathcal{R} \quad (3.28)$$

Similarly, postmultiplication by $(1 + L^{-1}\mathcal{R})$ and consideration of the identity $(1 + L^{-1}\mathcal{R})^{-1}(1 + L^{-1}\mathcal{R}) = 1$ show that

$$(1 + L^{-1}\mathcal{R})^{-1} = 1 - L^{-1}\mathcal{R}(1 + L^{-1}\mathcal{R})^{-1} \quad (3.29)$$

Substitution of (3.28) into (3.27) gives

$$\mathcal{L}^{-1} = [1 - (1 + L^{-1}\mathcal{R})^{-1}L^{-1}\mathcal{R}]L^{-1} \quad (3.30)$$

and further substitution of (3.29) into (3.30) yields

$$\mathcal{L}^{-1} = [1 - L^{-1}\mathcal{R} + L^{-1}\mathcal{R}(1 + L^{-1}\mathcal{R})^{-1}L^{-1}\mathcal{R}]L^{-1} \quad (3.31)$$

From (3.25) and (3.27) it is obvious that $(1 + L^{-1}\mathcal{R})^{-1}L^{-1}$ operates the same way as \mathcal{L}^{-1} and so (3.31) is equivalent to

$$\mathcal{L}^{-1} = (1 - L^{-1}\mathcal{R} + L^{-1}\mathcal{R}\mathcal{L}^{-1}\mathcal{R})L^{-1} \quad (3.32)$$

This form is reminiscent of Adomian's (1983) expansion in Neumann series mentioned in the first chapter, but contrary to the latter, it contains only a few terms and is

valid for every invertible L and \mathcal{L} . Substituting (3.32) into (3.25), taking conditional ensemble mean, recalling that $f(\mathbf{x})$ is a statistically independent random function, and taking into account (3.26), yields

$$\langle h(\mathbf{x}) \rangle_\kappa = (1 + L^{-1} \langle \mathcal{R} \mathcal{L}^{-1} \mathcal{R} \rangle_\kappa) h_\kappa(\mathbf{x}) \quad (3.33)$$

Premultiplying by L and comparing with (3.8) and (3.11) reveals that

$$\langle \mathcal{R} \mathcal{L}^{-1} \mathcal{R} \rangle_\kappa h_\kappa(\mathbf{x}) = \nabla \cdot \mathbf{r}_\kappa(\mathbf{x}) \quad (3.34)$$

From the analogy between (3.25) and (3.17) it follows that $\mathcal{L}^{-1} \mathcal{R} h_\kappa(\mathbf{x})$ is given by the right-hand-side of (3.17) when $f(\mathbf{x})$ in the latter is replaced by $-\mathcal{R} h_\kappa(\mathbf{x})$. Applying Green's identity to this expression while utilizing (3.13) and the definition of \mathcal{R} in (3.23), premultiplying by \mathcal{R} , taking conditional ensemble mean, considering (3.34) and our stipulation that $H(\mathbf{x})$ and $Q(\mathbf{x})$ be statistically independent random functions, gives the desired expression for the residual flux in (3.35) and (3.42).

Note that one could add an arbitrary divergence-free vector field $\mathbf{p}(\mathbf{x})$ to the right-hand side of (3.35) without violating (3.34); however, only $\mathbf{p}(\mathbf{x}) \equiv 0$ would satisfy the requirement that $\mathbf{r}_\kappa(\mathbf{x})$, and hence $\langle \mathbf{q}(\mathbf{x}) \rangle$ in (3.11), be identically equal to zero under hydrostatic conditions.

Consequently,

$$\begin{aligned} \mathbf{r}_\kappa(\mathbf{x}) &= \int_\Omega \mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}') \nabla_{\mathbf{x}'} h_\kappa(\mathbf{x}') d\mathbf{x}' - \int_{\Gamma_D} \mathbf{b}_\kappa(\mathbf{x}, \mathbf{x}') \mathbf{n}(\mathbf{x}') \langle H(\mathbf{x}') \rangle d\mathbf{x}' \\ &\quad - \int_{\Gamma_N} \mathbf{c}_\kappa(\mathbf{x}, \mathbf{x}') \langle Q(\mathbf{x}') \rangle d\mathbf{x}' \end{aligned} \quad (3.35)$$

where $\mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}')$, $\mathbf{b}_\kappa(\mathbf{x}, \mathbf{x}')$, $\mathbf{c}_\kappa(\mathbf{x}, \mathbf{x}')$ are kernels independent of $\langle h(\mathbf{x}) \rangle_\kappa$, and $h_\kappa(\mathbf{x})$ is the solution of the deterministic equation

$$\nabla \cdot [\kappa(\mathbf{x}) \nabla h_\kappa(\mathbf{x})] + \langle f(\mathbf{x}) \rangle = 0 \quad \mathbf{x} \in \Omega \quad (3.36)$$

subject to the boundary conditions

$$h_\kappa(\mathbf{x}) = \langle H(\mathbf{x}) \rangle \quad \mathbf{x} \in \Gamma_D \quad (3.37)$$

and

$$-\mathbf{q}_\kappa(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \langle Q(\mathbf{x}) \rangle \quad \mathbf{x} \in \Gamma_N \quad (3.38)$$

where

$$\mathbf{q}_\kappa(\mathbf{x}) = -\kappa(\mathbf{x})\nabla h_\kappa(\mathbf{x}) \quad (3.39)$$

Note from (3.3) – (3.4) and (3.37) – (3.39) that

$$h_\kappa(\mathbf{x}) = \langle h(\mathbf{x}) \rangle_\kappa = \langle h(\mathbf{x}) \rangle \quad \mathbf{x} \in \Gamma_D \quad (3.40)$$

$$\mathbf{q}_\kappa(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \langle \mathbf{q}(\mathbf{x}) \rangle_\kappa \cdot \mathbf{n}(\mathbf{x}) = \langle \mathbf{q}(\mathbf{x}) \rangle \cdot \mathbf{n}(\mathbf{x}) \quad \mathbf{x} \in \Gamma_N \quad (3.41)$$

It is thus obvious that $h_\kappa(\mathbf{x})$ is simply the lowest order approximation of $\langle h(\mathbf{x}) \rangle_\kappa$ which is easily determined by standard analytical or numerical methods (more on this later). The kernels are defined precisely by the following expressions,

$$\begin{aligned} \mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}') &= \langle K'(x)K'(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \\ \mathbf{b}_\kappa(\mathbf{x}, \mathbf{x}') &= \langle K'(\mathbf{x})K(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \\ \mathbf{c}_\kappa(\mathbf{x}, \mathbf{x}') &= \langle K'(\mathbf{x})K'(\mathbf{x}')\nabla\mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \kappa(\mathbf{x}')^{-1} - \langle K'(\mathbf{x})\nabla\mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \end{aligned} \quad (3.42)$$

where $\mathcal{G}(\mathbf{x}', \mathbf{x})$ is the random solution of (3.1) – (3.4) for the case where $f(\mathbf{x})$ is a point source of unit strength at point \mathbf{x}' , given by the Dirac delta function $\delta(\mathbf{x}-\mathbf{x}')$, subject to homogeneous boundary conditions $H(\mathbf{x}) \equiv Q(\mathbf{x}) \equiv 0$. In other words, $\mathcal{G}(\mathbf{x}', \mathbf{x})$ is the random Green's function corresponding to (3.1) – (3.4). Since the statistical properties of \mathcal{G} are unknown, none of these kernels can be evaluated quantitatively without either high-resolution Monte Carlo simulations (of the kind we perform later)

or approximations (of the type we propose below). Nevertheless, the kernels are sufficiently well defined to reveal some of their most interesting fundamental properties. Since $\mathcal{G}(\mathbf{x}', \mathbf{x})$ is symmetric (Appendix E) and $\mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}')$ is a quadratic form, we see immediately that this kernel is a symmetric, positive-semidefinite, second-rank tensor (a dyadic). On the other hand, $\mathbf{b}_\kappa(\mathbf{x}, \mathbf{x}')$ is a non-symmetric second-rank tensor and $\mathbf{c}_\kappa(\mathbf{x}, \mathbf{x}')$ is a vector. Lack of symmetry is thus seen to arise from boundary sources. All three kernels are nonlocal (depending on more than one point in space) and constitute system parameters which however depend not only on the statistical properties of the hydraulic conductivity field $K(\mathbf{x})$, but also on the information that one has about this field, as embodied in its estimate $\kappa(\mathbf{x})$ and in the statistical properties of the estimation error $K'(\mathbf{x})$. The latter is also true about the local “hydraulic conductivity” $\kappa(\mathbf{x})$. Hence all the parameters entering into (3.8) – (3.39) depend on both medium properties and information content (scale, quantity and quality of data).

It is clear from (3.35) that $\mathbf{r}_\kappa(\mathbf{x})$ is generally not proportional to the local hydraulic gradient but is given by a functional consisting of three spatial convolution integrals, one of which involves $\nabla_{\mathbf{x}'} h_\kappa(\mathbf{x}')$ at points other than \mathbf{x} . As such, $\mathbf{r}_\kappa(\mathbf{x})$ is generally nonlocal and non-Darcian. This in turn implies that the flux predictor $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ is likewise nonlocal (depends on predicted head gradients at points other than \mathbf{x}) and non-Darcian (there is no effective or equivalent hydraulic conductivity valid for arbitrary directions of conditional mean flow) except in special cases. As mentioned in the previous chapter, a nonlocal expression for the unconditional mean flux in a homogeneous $K(\mathbf{x})$ field, in the absence of boundaries, has been discussed by Dagan (1989, equation 3.4.43) based on a postulate by Saffman (1971). It involves

$\nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle$ rather than $\nabla_{\mathbf{x}'} h_\kappa(\mathbf{x}')$ under the integral and does not specify the nature of $\mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}')$ though Dagan has been able to identify the latter to first order in σ_Y^2 *a posteriori*. We will demonstrate in Section 3.5 that Saffman's postulate is not rigorously valid under general conditions.

Equations (3.8) – (3.39) and (3.42) constitute an exact deterministic system of equations which can be solved for the predictors $\langle h(\mathbf{x}) \rangle_\kappa$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$, subject to arbitrary random source and boundary terms, provided the kernels $\mathbf{a}(\mathbf{x}, \mathbf{x}')$, $\mathbf{b}(\mathbf{x}, \mathbf{x}')$, and $\mathbf{c}(\mathbf{x}, \mathbf{x}')$ are known. Since these nonlocal parameters as well as the local parameter $\kappa(\mathbf{x})$ and the functions $\langle h(\mathbf{x}) \rangle_\kappa$ and $h_\kappa(\mathbf{x})$ are smooth relative to their random counterparts, they can be treated approximately as finite dimensional functions, i.e., each can be expressed as the linear combination of a finite number of basis functions in the standard manner of finite elements; such methods can then be used to solve (3.8) – (3.39) and (3.42). The dimensionality N of each approximating function, i.e., the number N of basis functions employed (which controls the number of nodes and elements in a finite element grid), should depend in a standard way on the degree of smoothness of the function being approximated. Though we defer detailed discussion of this topic to future work, we nevertheless see that our theory provides a direct answer to the important question how should the scale of grid discretization (N) relate to the scale of measurement (ω) and to the quantity and quality of available data (smoothness of conditional moments). Likewise, we propose that if one has sufficient measurements of $K(\mathbf{x})$, $h(\mathbf{x})$ and $\mathbf{q}(\mathbf{x})$ on the scale ω and/or of spatial head and flux integrals (possibly weighted in a known fashion), one may in principle be able to estimate the above parameters *a posteriori* by means of a suitable inverse method.

In the remainder of this work we use analytical and numerical approaches to

investigate conditions under which the above nonlocal deterministic equations may be recast in a more familiar local form by means of effective hydraulic conductivities under uniform and nonuniform mean flows in two and three dimensions; for cases where such localization is warranted, we elucidate the nature of the effective hydraulic conductivities with and without the common assumption that $K(\mathbf{x})$ is log normal; for cases where localization is not warranted, we propose a weak approximation (closure) which eliminates the need for moments of $K(\mathbf{x})$ beyond $\langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa$ and improves with the quality of the estimate $\kappa(\mathbf{x})$; we investigate and verify some aspects of the proposed weak approximation by high-resolution Monte Carlo simulation under mean radial flow in strongly heterogeneous two-dimensional media using an adaptive multigrid finite element solver; we develop an exact formal expression for the prediction error moment $\langle h'(\mathbf{x})h'(\mathbf{x}') \rangle_\kappa$, and explicit approximations for the latter as well as for $\langle \mathbf{q}'(\mathbf{x})\mathbf{q}'^T(\mathbf{x}') \rangle_\kappa$; we recall a two-stage geostatistical method to estimate $\kappa(\mathbf{x})$ from noisy measurements of $K(\mathbf{x})$; and we supply corresponding expressions for $\langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa$ under the assumption that $K'(\mathbf{x})$ is log normal.

3.2 Effective Hydraulic Conductivity

Thanks to the rigorous nature of our expressions (3.35) – (3.42) for $\mathbf{r}_\kappa(\mathbf{x})$, we are in a position to make some precise and fundamental statements about the existence and properties of effective hydraulic conductivities in steady state flow through randomly nonuniform media. We start by noting that for $\mathbf{r}_\kappa(\mathbf{x})$ (and hence $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$) to be Darcian, it is necessary that there exist a symmetric, positive-semidefinite tensor $\kappa''(\mathbf{x})$ of second rank (a dyadic) whose principal values (eigenvalues) do not exceed $\kappa(\mathbf{x})$, and which is additionally independent of $\langle h(\mathbf{x}) \rangle_\kappa$ and its gradient, such

that

$$\mathbf{r}_\kappa(\mathbf{x}) = \tilde{\kappa}(\mathbf{x})\nabla\langle h(\mathbf{x})\rangle_\kappa \quad \mathbf{x} \in \bar{\Omega} \quad (3.43)$$

where $\bar{\Omega}$ is the closure of Ω (the union of Ω and Γ). Then and only then is it proper to write

$$\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa = -\mathbf{K}_e(\mathbf{x})\nabla\langle h(\mathbf{x}) \rangle_\kappa \quad \mathbf{x} \in \bar{\Omega} \quad (3.44)$$

where $\mathbf{K}_e(\mathbf{x})$ is a symmetric, positive definite “effective (or equivalent) hydraulic conductivity tensor” given by

$$\mathbf{K}_e(\mathbf{x}) = \kappa(\mathbf{x})\mathbf{I} - \tilde{\kappa}(\mathbf{x}) \quad (3.45)$$

\mathbf{I} being the identity tensor.

As discussed in Chapter 1, the above definition of effective (or equivalent) hydraulic conductivity has an unambiguous operational meaning which differs from meanings ascribed to similar terms by Desbarats (1987, 1992a,b), Gómez-Hernández and Gorelick (1989), King (1989), Mei and Auriault (1989), Arbogast et al. (1990), Desbarats and Dimitrakopoulos (1990), Kitanidis (1990), Morgan and Babuska (1990), Durlofsky (1991), Gómez-Hernández (1991), and some other authors (the latter two make an especially clear differentiation between effective and equivalent hydraulic conductivities). It is more closely related to definitions based on unconditional stochastic theories of flow in infinite, statistically homogeneous hydraulic conductivity fields such as those of Shvidler (1962), Matheron (1967), Gelhar and Axness (1983), Dagan (1982, 1989), King (1987), Poley (1988), and Naff (1991). These and other authors derived approximate expressions for the unconditional equivalent of $\mathbf{r}(\mathbf{x})_\kappa$ in uniform and/or radial mean flow regimes, from which they then developed ancillary approximations for effective hydraulic conductivity by taking the log

hydraulic conductivity, $Y(\mathbf{x}) = \ln K(\mathbf{x})$, to be Gaussian.

Equation (3.11) implies that since the residual flux $\mathbf{r}_\kappa(\mathbf{x})$ is generally non-Darcian and nonlocal, so is the conditional mean flux $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$. Hence, in general, there is no effective or equivalent hydraulic conductivity, $\mathbf{K}_e(\mathbf{x})$, valid at each \mathbf{x} in $\bar{\Omega}$ under arbitrary directions of conditional mean flow, that relates $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ to the conditional mean head gradient $\nabla \langle h(\mathbf{x}) \rangle_\kappa$ in the manner of (3.44). There are, however, some special cases for which a $\mathbf{K}_e(\mathbf{x})$ can be defined. One such case is where $\bar{\Omega}$ tends formally to a point (practically to ω , the support of the hydraulic conductivity measurements which we recall must have dimensions not exceeding those of $\bar{\Omega}$) while the mean source and boundary functions $\langle f(\mathbf{x}) \rangle$, $\langle H(\mathbf{x}) \rangle$, and $\langle Q(\mathbf{x}) \rangle$ remain finite. Since uniqueness of the solution requires that at least one point on Γ act as a Dirichlet boundary, the random Green's function \mathcal{G} and its derivatives tend to zero as $\bar{\Omega}$ shrinks to a point. Hence $\mathbf{r}_\kappa(\mathbf{x})$ in (3.35) tends to zero and (3.11) reduces to

$$\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa = -\kappa(\mathbf{x}) \nabla \langle h(\mathbf{x}) \rangle_\kappa \quad (3.46)$$

implying that $\mathbf{K}_e(\mathbf{x}) = \mathbf{I}\kappa(\mathbf{x})$. This merely confirms that the hydraulic conductivity one calculates experimentally for a sample $\bar{\Omega}$ (in the laboratory or in the field) by imposing head and no flow conditions along the sample boundaries, Γ , and measuring the resultant flow rate, approaches the estimate $\kappa(\mathbf{x})$ of the scalar hydraulic conductivity $K(\mathbf{x})$ as $\bar{\Omega}$ tends to ω . This finds support in numerical experiments conducted by Desbarats and Dimitrakopoulos (1990, Fig. 7). Another case is the limit as the variance of the estimation error $K'(\mathbf{x})$ approaches zero; then clearly $\mathbf{r}_\kappa(\mathbf{x}) \rightarrow 0$ for each \mathbf{x} in $\bar{\Omega}$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ is again given by (3.46).

In the special case where $K(\mathbf{x}) \equiv K$ is a random variable independent of \mathbf{x} one

has $K'(\mathbf{x}) \equiv K'$, the estimation error is perfectly autocorrelated within $\bar{\Omega}$, and K is uncorrelated with $h(\mathbf{x})$. This is easily seen upon substituting (3.1) into (3.2), dividing throughout by the random constant K , and taking the conditional ensemble mean while recalling that $f(\mathbf{x})$ is a statistically independent function. Hence $\mathbf{r}_\kappa(\mathbf{x}) \equiv 0$, $\langle h(\mathbf{x}) \rangle_\kappa \equiv h_\kappa(\mathbf{x})$, and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ is once again given by (3.46).

The case most often discussed in the stochastic literature is that of an unbounded domain Ω_∞ in which $\kappa(\mathbf{x}) \equiv \kappa = \text{constant}$, $K'(\mathbf{x})$ is statistically homogeneous, $\langle f(\mathbf{x}) \rangle \equiv 0$, and flow is controlled by boundary conditions at infinity which result in a uniform mean hydraulic gradient, $\nabla \langle h(\mathbf{x}) \rangle_\kappa \equiv \mathbf{J} = \text{constant}$ (though we retain the subscript κ for consistency of notation, no actual conditioning is implied by it in this special case). The same is true for $\nabla h_\kappa(\mathbf{x})$ and therefore (3.35) reduces to

$$\mathbf{r}_\kappa(\mathbf{x}) = \tilde{\kappa} \mathbf{J} \quad (3.47)$$

where $\tilde{\kappa}$ is a constant, symmetric, positive-semidefinite, second-rank tensor given by

$$\tilde{\kappa} = \int_{\Omega_\infty} \langle K'(0)K'(\mathbf{x}') \nabla \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', 0) \rangle_\kappa d\mathbf{x}' \quad (3.48)$$

\mathbf{x} having been arbitrarily set equal to zero. It follows from (3.11) that the mean flux $\langle \mathbf{q} \rangle_\kappa$ is also constant and satisfies Darcy's law in the form

$$\langle \mathbf{q} \rangle_\kappa = -\mathbf{K}_e \mathbf{J} \quad \mathbf{K}_e = \kappa \mathbf{I} - \tilde{\kappa} \quad (3.49)$$

where \mathbf{K}_e is a constant, symmetric, positive-definite, second-rank effective conductivity tensor. Equation (26) is clearly valid for arbitrary (uniform) directions of mean flux.

To account for the effect of boundaries, we first consider the special case where $\langle f(\mathbf{x}) \rangle \equiv \langle Q(\mathbf{x}) \rangle \equiv 0$ and $\langle H(\mathbf{x}) \rangle \equiv \langle H \rangle = \text{constant}$ so that the flow is static in the

mean. Then $\nabla h_\kappa(\mathbf{x}) \equiv \mathbf{r}_\kappa(\mathbf{x}) \equiv 0$ and (3.35) imply that

$$\int_{\Gamma_D} \langle K'(\mathbf{x})K(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}',\mathbf{x}) \rangle_\kappa \mathbf{n}(\mathbf{x}') d\mathbf{x}' \equiv 0 \quad (3.50)$$

for every \mathbf{x} in $\bar{\Omega}$. As \mathcal{G} does not depend on the choice of source and boundary functions, only on the boundary configuration and $K(\mathbf{x})$, (3.50) must be universally valid. The same follows from the obvious fact that

$$\int_{\Gamma_D} K(\mathbf{x}')\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}',\mathbf{x})\mathbf{n}(\mathbf{x}') d\mathbf{x}' \equiv -1 \quad (3.51)$$

meaning that the total outflow rate across Γ_D due to a unit source, subject to a homogeneous Neumann boundary condition, must be equal to 1; clearly, the gradient of (3.51) is identically equal to zero.

Next, we consider a box-shaped flow domain $\bar{\Omega}$ with orthogonal plane boundaries centered about the origin of a cartesian coordinate system which is oriented parallel and normal to the box boundaries. On the boundary Γ_{D1-} traversed by the negative x_1 axis $\langle H(\mathbf{x}) \rangle \equiv \langle H \rangle_{1-} = \text{constant}$, on the opposite boundary Γ_{D1+} traversed by the positive x_1 axis $\langle H(\mathbf{x}) \rangle \equiv \langle H \rangle_{1+} = \text{constant}$, on the remaining four boundaries $\langle Q(\mathbf{x}) \rangle \equiv 0$, and in the interior $\langle f(\mathbf{x}) \rangle \equiv 0$. If in addition $\kappa(\mathbf{x}) \equiv \kappa = \text{constant}$ (but $K'(\mathbf{x})$ is not necessarily homogeneous!), then $\nabla h_\kappa(\mathbf{x}) \equiv \mathbf{J} = \text{constant}$ where \mathbf{J} has only one nonzero component, $J_1 = (\langle H \rangle_{1+} - \langle H \rangle_{1-})/L_1$, L_1 being the distance between the Dirichlet boundaries Γ_{D1-} and Γ_{D1+} . This, together with (3.35) and (3.50), imply that the residual flux can be expressed as,

$$\mathbf{r}_\kappa(\mathbf{x}) = \tilde{\kappa}(\mathbf{x})\mathbf{J} \quad (3.52)$$

where $\tilde{\kappa}(\mathbf{x})$ is a (generally) nonsymmetric second-rank tensor given by

$$\tilde{\kappa}(\mathbf{x}) = \int_{\Omega} \langle K'(\mathbf{x})K'(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}',\mathbf{x}) \rangle_\kappa d\mathbf{x}' - L_1 \int_{\Gamma_{D1}} \langle K'(\mathbf{x})K(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}',\mathbf{x}) \rangle_\kappa d\mathbf{x}' \quad (3.53)$$

and Γ_{D1} is either Γ_{D1-} or Γ_{D1+} . Here the lack of symmetry is a boundary source effect stemming from the presence of $K(\mathbf{x}')$, rather than $K'(\mathbf{x}')$, in the integral over Γ_{D1} . Hence according to (3.11), the conditional mean flux takes the form

$$\langle \mathbf{q}(\mathbf{x}) \rangle_{\kappa} = -\kappa \nabla \langle h(\mathbf{x}) \rangle_{\kappa} + \tilde{\kappa}(\mathbf{x}) \mathbf{J} \quad (3.54)$$

where, in general, $\nabla \langle h(\mathbf{x}) \rangle_{\kappa}$ differs from $\nabla h_{\kappa}(\mathbf{x}) = \mathbf{J}$. We shall refer to this form as quasi-Darcian in that $\langle \mathbf{q}(\mathbf{x}) \rangle_{\kappa}$ is linear in both $\nabla \langle h(\mathbf{x}) \rangle_{\kappa}$ and \mathbf{J} but is not proportional to either of these two gradients, the tensor $\tilde{\kappa}(\mathbf{x})$ being additionally nonsymmetric. Note that $\tilde{\kappa}(\mathbf{x})$ depends on $K(\mathbf{x})$, the quality of its estimate $\kappa(\mathbf{x})$, the location (if $K(\mathbf{x})$ is statistically nonhomogeneous) and orientation (if $K(\mathbf{x})$ is statistically anisotropic) of the box, box dimensions, and the choice of prescribed mean head and flux boundaries.

Assume now that, in addition to the above requirements, $K'(\mathbf{x})$ is statistically homogeneous and exhibits ellipsoidal (Vanmarcke, 1983) or (equivalently) geometric (Journel and Huijbregts, 1978) statistical anisotropy. One can then define a system of principal Cartesian coordinates, x^p , and corresponding principal characteristic lengths, $\lambda_1, \lambda_2, \lambda_3$, such that $K'(x^*)$ becomes isotropic when x^p is transformed according to $x^* = \lambda^{-1} x^p$ where λ is a diagonal matrix whose main diagonal consists of $\lambda_1, \lambda_2, \lambda_3$. If the box-shaped flow domain $\bar{\Omega}$ and the working coordinates \mathbf{x} are oriented parallel to the principal coordinates, then symmetry implies that $\nabla \langle h(\mathbf{x}) \rangle_{\kappa} \equiv \nabla h_{\kappa}(\mathbf{x}) \equiv J$. This, coupled with the mass balance requirement that $\mathbf{r}_{\kappa}(\mathbf{x})$ remains constant in the direction of mean flow, imply that the residual flux has only one nonzero component parallel to x_1 ,

$$\mathbf{r}_{\kappa 1}(x_2, x_3) = K_1(x_2, x_3) J_1 \quad (3.55)$$

where $K_1(x_2, x_3)$ is a directional quantity given by

$$\begin{aligned} K_1(x_2, x_3) &= \int_{\Omega} \langle K'(\mathbf{x})K'(\mathbf{x}')\partial^2\mathcal{G}(\mathbf{x}', \mathbf{x})/\partial x_1\partial x'_1 \rangle_{\kappa} d\mathbf{x}' \\ &- L_1 \int_{\Gamma_{D_1}} \langle K'(\mathbf{x})K(\mathbf{x}')\partial^2\mathcal{G}(\mathbf{x}', \mathbf{x})/\partial x_1\partial x'_1 \rangle_{\kappa} d\mathbf{x}' \end{aligned} \quad (3.56)$$

The conditional mean flux has, likewise, only one nonzero component which can be expressed by means of a scalar form of Darcy's law,

$$\langle q_1(x_2, x_3) \rangle_{\kappa} = -[\kappa - K_1(x_2, x_3)]J_1 \quad (3.57)$$

where $[\kappa - K_1(x_2, x_3)]$ is an effective directional hydraulic conductivity. Note that the latter is a function of $K(\mathbf{x})$, the quality of its estimate $\kappa(\mathbf{x})$, box dimensions, and distance from its Neumann boundaries.

If the box is at an angle to the principal axes of statistical anisotropy, symmetry is broken and $\nabla\langle h(\mathbf{x}) \rangle_{\kappa}$ is no longer uniform. Hence the latter is not equal to \mathbf{J} and Darcy's law does not hold; one must use instead the quasi-Darcian expression (3.54). Only if $K'(\mathbf{x})$ is statistically isotropic will (3.57) apply for arbitrary choices of x_1, x_2, x_3 relative to a frame of reference attached to the medium. Even then, the three quantities $[\kappa - K_1(x_2, x_3)]$, $[\kappa - K_3(x_1, x_2)]$, and $[\kappa - K_2(x_3, x_1)]$ will generally differ from each other unless $\bar{\Omega}$ is a cube. It is important to appreciate that these quantities are strictly directional and do not form the principal components of a tensor. It is further important to recall that they correspond to a closed box-shaped flow domain with two prescribed constant mean head and four zero mean flow boundaries. As heads and fluxes along interfaces between finite elements or finite difference cells in a numerical model are generally not prescribed in this fashion but are merely required to satisfy appropriate compatibility criteria, the above quantities qualify

neither as principal nor as directional equivalent hydraulic conductivities for such elements or cells.

3.3 Weak Approximation of Residual Flux

To render our nonlocal theory workable near interior sources and boundaries where localization does not appear warranted, one must either have sufficient data to evaluate the unknown kernels in (3.42) by an inverse method as was suggested earlier, or develop a working approximation for the residual flux $\mathbf{r}_\kappa(\mathbf{x})$. We propose to approximate the latter through a replacement of $\mathcal{G}(\mathbf{x}, y)$ within third moments under interior integrals, as well as within second moments under boundary integrals, in (3.35) and (3.42) by its conditional mean $\langle \mathcal{G}(\mathbf{x}, y) \rangle_\kappa$, so that

$$\begin{aligned} \mathbf{r}_\kappa(\mathbf{x}) &\simeq \int_{\Omega} \langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} h_\kappa(\mathbf{x}') d\mathbf{x}' \\ &\quad - \int_{\Gamma_D} \langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{n}(\mathbf{x}') \langle H(\mathbf{x}') \rangle d\mathbf{x}' \\ &\quad - \int_{\Gamma_N} \langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa \nabla \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \kappa^{-1}(\mathbf{x}') \langle Q(\mathbf{x}') \rangle d\mathbf{x}' \end{aligned} \quad (3.58)$$

The approximation is “weak” in that $\mathcal{G}(\mathbf{x}, y)$ is replaced by $\langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa$ only under integrals and not at individual pairs of points $(\mathbf{x}', \mathbf{x})$; the replacement is done in a weighted mean sense, not locally. It is expected to improve as the conditional ensemble moments of the estimation error $K'(\mathbf{x})$ diminish in magnitude due to an improvement in the quality of the estimate $\kappa(\mathbf{x})$.

Substituting (3.58) into (3.11) and setting $\langle f(\mathbf{x}) \rangle = \delta(\mathbf{x} - \mathbf{x}_o)$, $\langle H(\mathbf{x}) \rangle \equiv \langle Q(\mathbf{x}) \rangle \equiv 0$ in (3.8)–(3.10) shows that $\langle \mathcal{G} \rangle_\kappa$ satisfies approximately a linear integro-differential equation

$$\nabla \cdot \left[\kappa(\mathbf{x}) \nabla \langle \mathcal{G}(\mathbf{x}, \mathbf{x}_o) \rangle_\kappa - \int_{\Omega} \langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} G_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \right]$$

$$+\delta(\mathbf{x} - \mathbf{x}_o) \simeq 0 \quad (3.59)$$

subject to the homogeneous boundary conditions

$$\langle \mathcal{G}(\mathbf{x}, \mathbf{x}_o) \rangle_\kappa \equiv 0 \quad \mathbf{x} \in \Gamma_D \quad (3.60)$$

and

$$\left[\kappa(\mathbf{x}) \nabla \langle \mathcal{G}(\mathbf{x}, \mathbf{x}_o) \rangle_\kappa - \int_{\Omega} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla G_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \right] \cdot \mathbf{n}(\mathbf{x}) \equiv 0 \quad (3.61)$$

for \mathbf{x} on Γ_N , where G_κ is the solution of (3.18) – (3.20).

The approximation (3.58) also yields the following respective working equivalents for the effective hydraulic conductivity expressions in (3.48), (3.53), and (3.56),

$$\tilde{\kappa} \simeq \int_{\Omega_\infty} \langle K'(0) K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', 0) \rangle_\kappa d\mathbf{x}' \quad (3.62)$$

$$\begin{aligned} \tilde{\kappa}(\mathbf{x}) &\simeq \int_{\Omega} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' \\ &\quad - L_1 \int_{\Gamma_{D1}} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' \end{aligned} \quad (3.63)$$

$$\begin{aligned} K_1(x_2, x_3) &\simeq \int_{\Omega} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_\kappa \frac{\partial^2 \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa}{\partial x_1 \partial x_1'} d\mathbf{x}' \\ &\quad - L_1 \int_{\Gamma_{D1}} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_\kappa \frac{\partial^2 \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa}{\partial x_1 \partial x_1'} d\mathbf{x}' \end{aligned} \quad (3.64)$$

The verification of the weak approximation is described in Chapter 4.

3.4 Effective Hydraulic Conductivity in 2-D And 3-D Lognormal Fields Under Uniform Mean Flow

It is common in subsurface hydrology to treat $Y(\mathbf{x}) = \ln K(\mathbf{x})$ as a Gaussian and statistically homogeneous field. It then follows from Appendix F that $\kappa =$

$\langle K \rangle = K_g \exp(\sigma_Y^2/2)$ where $K_g = \exp\langle Y \rangle$ is the geometric mean of $K(\mathbf{x})$ and σ_Y^2 is the variance of $Y(\mathbf{x})$. Furthermore, $\langle K'(\mathbf{x})K'(\mathbf{x}') \rangle = K_g^2 \exp(\sigma_Y^2) \{ \exp[C_Y(x - \mathbf{x}')] - 1 \}$ where $C_Y(x - \mathbf{x}') = \langle Y'(\mathbf{x})Y'(\mathbf{x}') \rangle$ is the covariance of $Y(\mathbf{x})$. We start by considering an unbounded domain Ω_∞ with a uniform mean hydraulic gradient $\nabla \langle h(\mathbf{x}) \rangle_\kappa = \nabla h_\kappa(\mathbf{x}) = \mathbf{J} = \text{constant}$. We found earlier that here the effective hydraulic conductivity is a symmetric, positive-definite, second-rank tensor given by (3.48) – (3.49); like before, we retain the subscript κ for consistency of notation though no actual conditioning takes place in this case.

In statistically isotropic fields the effective hydraulic conductivity is a scalar, K_e . By recognizing that this scalar must be a linear functional of $K(\mathbf{x})$, Matheron (1967) was able to prove that, in two-dimensions, $K_e \equiv K_g$ regardless of what σ_Y^2 and C_Y may be. We restate this proof in Appendix G by basing it on the exact expressions (3.48) – (3.49) rather than on an abstract linear functional.

For the more general statistically anisotropic and three-dimensional cases we follow a less rigorous approach. Upon introducing the above relationships into (3.48) – (3.49) and disregarding terms of order σ_Y^4 , we obtain

$$\mathbf{K}_e \simeq K_g \left(1 + \frac{\sigma_Y^2}{2} \right) \left[\mathbf{I} - \int_{\Omega_\infty} C_Y(\mathbf{x}') \nabla \nabla_{\mathbf{x}'}^T G_0(\mathbf{x}', \mathbf{x}; x=0) d\mathbf{x}' \right] \quad (3.65)$$

where G_0 is the fundamental solution of the Laplace equation,

$$\nabla^2 G_0(\mathbf{x}, 0) + \delta(\mathbf{x}) = 0 \quad \mathbf{x} \in \Omega_\infty \quad (3.66)$$

This is precisely the second-order result obtained via perturbation by Dagan (1989, equations 3.4.9, 3.4.12, 3.4.16) and it corresponds exactly to a second-order result obtained earlier via spectral analysis by Gelhar and Axness (1983) (see previous chapter). Whereas the latter result involves the spectral density of $Y(\mathbf{x})$, Neuman

and Depner (1988) showed that \mathbf{K}_e depends solely on σ_Y^2 and the integral scales of $Y(\mathbf{x})$. According to them, the principal components of \mathbf{K}_e are given by

$$K_{ei} \simeq K_g \left(1 + \frac{\sigma_Y^2}{2} - F_i\right) \quad i = 1, 2, 3 \quad (3.67)$$

where

$$F_i = \frac{2\sigma_Y^2}{\pi\lambda_i^2} \int_0^{\pi/2} \int_0^{\pi/2} \frac{f_i}{f^T \lambda^{-2} f} \sin \phi \, d\phi \, d\theta \quad (3.68)$$

λ being a diagonal matrix the nonzero terms of which are the principal integral scales λ_i , and

$$f^T(\phi, \theta) = (\cos \theta \sin \phi, \sin \theta \sin \phi, \cos \phi) \quad (3.69)$$

To obtain an expression valid for large σ_Y^2 , they adopted a conjecture due to Gelhar and Axness (1983), similar to that proposed earlier for isotropic fields by Landau and Lifshitz (1960) and Matheron (1967), that the expression in parentheses constitutes the first two terms in a series expansion of $\exp(\sigma_Y^2/2 - F_i)$, yielding

$$K_{ei} \simeq \exp \left(\langle Y \rangle + \frac{\sigma_Y^2}{2} - F_i \right) \quad i = 1, 2, 3 \quad (3.70)$$

To verify (3.70) numerically, our group (Neuman et al., 1992a,b) performed high-resolution Monte Carlo simulations of uniform mean flow across a finite element grid of $35 \times 35 \times 35$ trilinear cubes in a statistically homogeneous, isotropic, log normal $K(\mathbf{x})$ field with an exponential covariance. Random realizations of $K(\mathbf{x})$ were generated with the turning band method of Tompson et al (1989) so as to obtain an integral scale of five cube lengths. Point values of $K(\mathbf{x})$, generated at the centroids of the cubes, were assigned without change to the corresponding cubes. The finite element flow equations were then solved in residual form by preconditioned conjugate gradients (with incomplete LU factorization) using the ESSL-IBM package

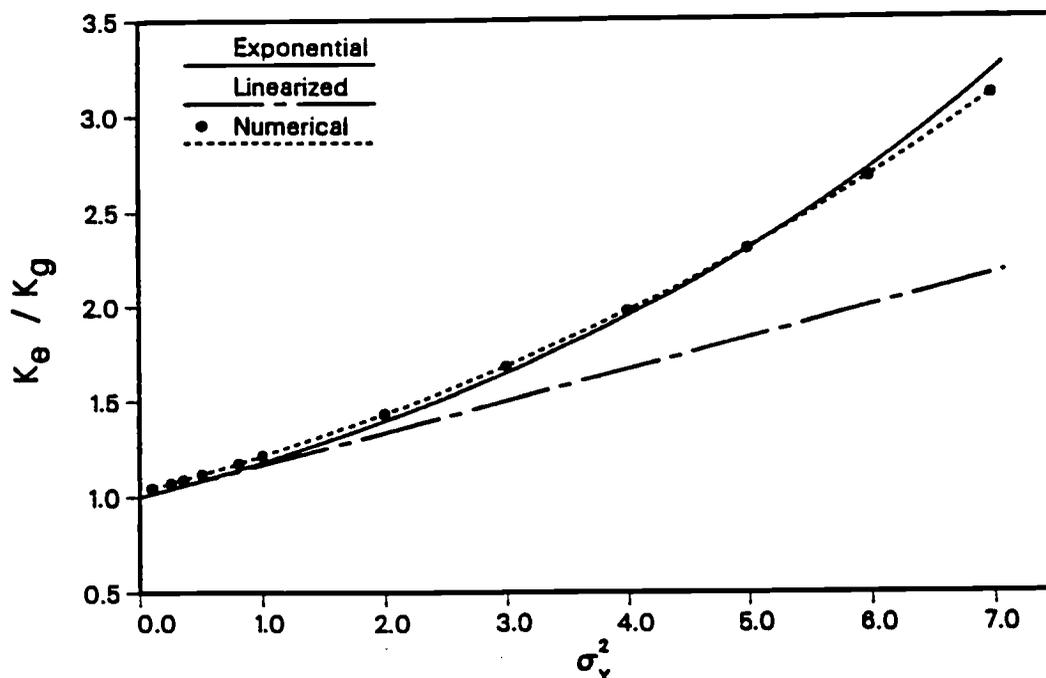


Figure 3.1: K_e/K_g versus σ_Y^2 for isotropic 3-D domain, after Neuman et al. (1992a).

DSMGCG (O. Levin, personal communication). Constant heads were imposed on two opposing sides of the cubic flow domain, no-flow conditions on the remaining sides. The global mass balance error, averaged over 500 simulations, increased from 0.002% for $\sigma_Y^2 = 0.25$ to 0.5% for $\sigma_Y^2 = 7$. The largest global mass balance error during any simulation was less than 3%.

Figure 3.1 shows the ratio between K_e and K_g for this isotropic case as a function of σ_Y^2 .

We see that the exponential expression (3.70) agrees very well with the numerical results up to at least $\sigma_Y^2 = 7$. A similar agreement with Monte Carlo results,

generated with a different numerical method, has been reported recently by Dykaar and Kitanidis (1992) for σ_Y^2 as large as 6 and by Desbarats (1992b) for $\sigma_Y^2 \leq 3$; Ababou (1988) found earlier, on the basis of a single realization in a large domain, that the agreement holds at least up to $\sigma_Y^2 = 5.3$. This may explain why, when applied by Neuman and Depner (1988) to data obtained from relatively small-scale single-hole packer tests in fractured granites, according to which $\sigma_Y^2 > 7$ and $Y(\mathbf{x})$ is statistically anisotropic, (3.70) showed consistency with the results of much larger-scale cross-hole tests conducted and interpreted independently of the single-hole tests. Note that the equivalent linearized expressions (3.65) and (3.67) underestimate K_e by an amount that grows exponentially with σ_Y^2 as the latter increases beyond 1.

Dagan (1989) derived closed-form expressions for F_i for the case where $Y(\mathbf{x})$ has an axisymmetric covariance (his equations 3.4.17 - 3.4.18). Upon noting that for uniform mean flow in a statistically homogeneous, isotropic 2-D domain, $F = \sigma_Y^2/2$, Dagan concluded on the basis of (3.67) that $K_e = K_g$ at least for small σ_Y^2 . As (3.70) also yields $K_e = K_g$ for $F = \sigma_Y^2/2$, the above numerical verification of (3.70) for $0 < \sigma_Y^2 \leq 7$ in 3-D constitutes an independent (though indirect) verification that $K_e = K_g$ in 2-D for both small and large σ_Y^2 .

3.5 Second Conditional Moments of Prediction Errors

To develop expressions for the second conditional moments of the prediction errors $h'(\mathbf{x})$ and $\mathbf{q}'(\mathbf{x})$ we start with a rigorous formalism. Combining (3.1) and (3.2), substituting (3.6) and (3.7), subtracting its conditional ensemble mean from the result, and recalling from (3.11) that $\mathbf{r}_\kappa(\mathbf{x}) = -\langle K'(\mathbf{x})\nabla h'(\mathbf{x}) \rangle_\kappa$ we obtain

$$\nabla \cdot [K(\mathbf{x})\nabla h'(\mathbf{x}) + K'(\mathbf{x})\nabla \langle h(\mathbf{x}) \rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x})] + f'(\mathbf{x}) = 0 \quad (3.71)$$

In analogy to (3.17) and (3.25) we can write

$$\begin{aligned}
h'(\mathbf{x}) &= \mathcal{L}^{-1}\{-\nabla \cdot [K'(\mathbf{x})\nabla\langle h(\mathbf{x})\rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x})] - f'(\mathbf{x})\} \\
&= \int_\Omega \{\nabla_{\mathbf{x}'} \cdot [K'(\mathbf{x}')\nabla_{\mathbf{x}'}\langle h(\mathbf{x}')\rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x}')] + f'(\mathbf{x}')\} \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \\
&\quad - \int_{\Gamma_D} K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') H'(\mathbf{x}') d\mathbf{x}' \\
&\quad + \int_{\Gamma_N} K(\mathbf{x}') \nabla_{\mathbf{x}'} h'(\mathbf{x}') \cdot \mathbf{n}(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \tag{3.72}
\end{aligned}$$

Application of Green's identity gives

$$\begin{aligned}
h'(\mathbf{x}) &= - \int_\Omega [\nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{r}_\kappa(\mathbf{x}') + K'(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \nabla_{\mathbf{x}'} \langle h(\mathbf{x}')\rangle_\kappa + f'(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x})] d\mathbf{x}' \\
&\quad - \int_{\Gamma_D} K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}') H'(\mathbf{x}') d\mathbf{x}' + \int_{\Gamma_N} Q'(\mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x}' \tag{3.73}
\end{aligned}$$

where

$$Q'(\mathbf{x}) = [K(\mathbf{x})\nabla h(\mathbf{x}) - \kappa(\mathbf{x})\nabla\langle h(\mathbf{x})\rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x})] \cdot \mathbf{n}(\mathbf{x}) \tag{3.74}$$

Since $f(\mathbf{x})$, $H(\mathbf{x})$, and $Q(\mathbf{x})$ are statistically independent and have no effect on $\mathcal{G}(\mathbf{x}', \mathbf{x})$, taking the conditional ensemble mean of (3.73) yields the interesting formal relationship

$$\int_\Omega \mathbf{r}_\kappa(\mathbf{x}') \cdot \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' = \int_\Omega \rho_\kappa(\mathbf{x}', \mathbf{x}) \cdot \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' \tag{3.75}$$

where, in analogy to the definition of $\mathbf{r}_\kappa(\mathbf{x})$, $\rho_\kappa(\mathbf{x}', \mathbf{x}) = -\langle K'(\mathbf{x}') \nabla \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_\kappa$. By virtue of (3.35) and (3.18) – (3.20),

$$\rho_\kappa(\mathbf{x}, \mathbf{x}_o) = \int_\Omega \langle K'(\mathbf{x}) K'(\mathbf{x}') \nabla \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} G_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \tag{3.76}$$

Multiplying the gradient of (3.73) by $K'(\mathbf{y})$ and taking conditional ensemble mean yields

$$\begin{aligned}
\langle K'(\mathbf{y}) \nabla h'(\mathbf{x}) \rangle_\kappa &= -\nabla \int_\Omega \langle K'(\mathbf{y}) \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{r}_\kappa(\mathbf{x}') d\mathbf{x}' \\
&\quad - \int_\Omega \langle K'(\mathbf{y}) K'(\mathbf{x}') \nabla \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' \tag{3.77}
\end{aligned}$$

Similarly,

$$\begin{aligned} \langle K'(\mathbf{y})\nabla\mathcal{G}'(\mathbf{x}, \mathbf{x}_o)\rangle_\kappa &= -\nabla \int_\Omega \langle K'(\mathbf{y})\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}', \mathbf{x})\rangle_\kappa \rho_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \\ &\quad - \int_\Omega \langle K'(\mathbf{y})K'(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}', \mathbf{x})\rangle_\kappa \nabla_{\mathbf{x}'}\langle\mathcal{G}(\mathbf{x}', \mathbf{x}_o)\rangle_\kappa d\mathbf{x}' \end{aligned} \quad (3.78)$$

These are integro-differential equations for $\langle K'(\mathbf{y})\nabla h'(\mathbf{x})\rangle_\kappa$ and $\langle K'(\mathbf{y})\nabla\mathcal{G}'(\mathbf{x}, \mathbf{x}_o)\rangle_\kappa$, respectively. When \mathbf{y} takes the value \mathbf{x} , (3.77) and (3.78) become integro-differential equations for $\mathbf{r}_\kappa(\mathbf{x})$ and $\rho_\kappa(\mathbf{x}, \mathbf{x}_o)$, constituting implicit equivalents of the respective explicit expressions (3.35) and (3.76). This shows that $\mathbf{r}_\kappa(\mathbf{x})$ is generally not an explicit functional of $\nabla\langle h(\mathbf{x})\rangle_\kappa$ as postulated for the unconditional case by Saffman (1971) and quoted by Dagan (1989, equation 3.4.43; see first chapter).

From (3.73), the second conditional moment of $h'(\mathbf{x})$ is given formally by

$$\begin{aligned} \langle h'(\mathbf{y})h'(\mathbf{x})\rangle_\kappa &= \int_\Omega \int_\Omega \mathbf{r}_\kappa(\mathbf{y}')^T \nabla_{\mathbf{y}'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{r}_\kappa(\mathbf{x}') d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_\Omega \int_\Omega \mathbf{r}_\kappa(\mathbf{y}')^T \nabla_{\mathbf{y}'} \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) K'(\mathbf{x}') \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_\Omega \int_\Omega \mathbf{r}_\kappa(\mathbf{x}')^T \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) K'(\mathbf{y}') \nabla_{\mathbf{y}'}^T \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \nabla_{\mathbf{y}'} \langle h(\mathbf{y}') \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_\Omega \int_\Omega \nabla_{\mathbf{y}'}^T \langle h(\mathbf{y}') \rangle_\kappa \langle K'(\mathbf{y}') \nabla_{\mathbf{y}'} \mathcal{G}(\mathbf{y}', \mathbf{y}) K'(\mathbf{x}') \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_\Omega \int_\Omega \langle f'(\mathbf{y}') f'(\mathbf{x}') \rangle \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_{\Gamma_D} \int_{\Gamma_D} \langle H'(\mathbf{y}') H'(\mathbf{x}') \rangle \langle K(\mathbf{y}') \nabla_{\mathbf{y}'}^T \mathcal{G}(\mathbf{y}', \mathbf{y}) \mathbf{n}(\mathbf{y}') K(\mathbf{x}') \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \mathbf{n}(\mathbf{x}') \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_{\Gamma_N} \int_{\Gamma_N} \langle Q'(\mathbf{y}') Q'(\mathbf{x}') \rangle \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \end{aligned} \quad (3.79)$$

By the same token,

$$\begin{aligned} \langle \mathcal{G}'(\mathbf{y}, \mathbf{y}_o) \mathcal{G}'(\mathbf{x}, \mathbf{x}_o) \rangle_\kappa &= \int_\Omega \int_\Omega \rho_\kappa(\mathbf{y}', \mathbf{y}_o)^T \nabla_{\mathbf{y}'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \rho_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_\Omega \int_\Omega \rho_\kappa(\mathbf{y}', \mathbf{y}_o) \nabla_{\mathbf{y}'} \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) K'(\mathbf{x}') \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}_o) \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\ &\quad + \int_\Omega \int_\Omega \rho_\kappa(\mathbf{x}', \mathbf{x}_o) \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) K'(\mathbf{y}') \nabla_{\mathbf{y}'}^T \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \nabla_{\mathbf{y}'} \langle \mathcal{G}(\mathbf{y}', \mathbf{y}_o) \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \end{aligned}$$

$$+ \int_{\Omega} \int_{\Omega} \nabla_{\mathbf{y}'}^T \langle \mathcal{G}(\mathbf{y}', \mathbf{y}_o) \rangle_{\kappa} \langle K'(\mathbf{y}') \nabla_{\mathbf{y}'} \mathcal{G}(\mathbf{y}', \mathbf{y}) K'(\mathbf{x}') \nabla_{\mathbf{x}'}^T \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}_o) \rangle_{\kappa} d\mathbf{x}' d\mathbf{y}' \quad (3.80)$$

Substituting (3.6) and (3.7) into (3.1) and subtracting (3.11) gives

$$\mathbf{q}'(\mathbf{x}) = -\mathbf{r}_{\kappa}(\mathbf{x}) - \kappa(\mathbf{x}) \nabla h'(\mathbf{x}) - K'(\mathbf{x}) \nabla \langle h(\mathbf{x}) \rangle_{\kappa} - K'(\mathbf{x}) \nabla h'(\mathbf{x}) \quad (3.81)$$

Hence the second conditional moment of $\mathbf{q}'(\mathbf{x})$ is given formally by

$$\begin{aligned} \langle \mathbf{q}'(\mathbf{y}) \mathbf{q}'(\mathbf{x})^T \rangle_{\kappa} &= -\mathbf{r}_{\kappa}(\mathbf{y}) \mathbf{r}_{\kappa}(\mathbf{x})^T \\ &+ \kappa(\mathbf{y}) \nabla_{\mathbf{y}} \nabla^T \langle h'(\mathbf{y}) h'(\mathbf{x}) \rangle_{\kappa} \kappa(\mathbf{x}) + \nabla_{\mathbf{y}} \langle h(\mathbf{y}) \rangle_{\kappa} \langle K'(\mathbf{y}) K'(\mathbf{x}) \rangle_{\kappa} \nabla^T \langle h(\mathbf{x}) \rangle_{\kappa} \\ &+ \nabla_{\mathbf{y}} \langle h(\mathbf{y}) \rangle_{\kappa} \langle K'(\mathbf{y}) \nabla^T h'(\mathbf{x}) \rangle_{\kappa} \kappa(\mathbf{x}) + \kappa(\mathbf{y}) \langle K'(\mathbf{x}) \nabla_{\mathbf{y}} h'(\mathbf{y}) \rangle_{\kappa} \nabla^T \langle h(\mathbf{x}) \rangle_{\kappa} \\ &+ \langle K'(\mathbf{y}) \nabla_{\mathbf{y}} h'(\mathbf{y}) \nabla^T h'(\mathbf{x}) \rangle_{\kappa} \kappa(\mathbf{x}) + \kappa(\mathbf{y}) \nabla_{\mathbf{y}} \langle h'(\mathbf{y}) K'(\mathbf{x}) \nabla^T h'(\mathbf{x}) \rangle_{\kappa} \\ &+ \langle K'(\mathbf{x}) K'(\mathbf{y}) \nabla_{\mathbf{y}} h'(\mathbf{y}) \rangle_{\kappa} \nabla^T \langle h(\mathbf{x}) \rangle_{\kappa} + \nabla_{\mathbf{y}} \langle h(\mathbf{y}) \rangle_{\kappa} \langle K'(\mathbf{y}) K'(\mathbf{x}) \nabla^T h'(\mathbf{x}) \rangle_{\kappa} \\ &+ \langle K'(\mathbf{y}) \nabla_{\mathbf{y}} h'(\mathbf{y}) K'(\mathbf{x}) \nabla^T h'(\mathbf{x}) \rangle_{\kappa} \end{aligned} \quad (3.82)$$

To render the above formal expressions workable, an approximation is necessary. Consistent with the weak approximation of $\mathbf{r}_{\kappa}(\mathbf{x})$ in (3.58), one can approximate (3.76) – (3.78) via

$$\rho_{\kappa}(\mathbf{x}, \mathbf{x}_o) \simeq \int_{\Omega} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_{\kappa} \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \nabla_{\mathbf{x}'} G_{\kappa}(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \quad (3.83)$$

$$\begin{aligned} \langle K'(\mathbf{y}) \nabla h'(\mathbf{x}) \rangle_{\kappa} &\simeq -\nabla \int_{\Omega} \langle K'(\mathbf{y}) \nabla_{\mathbf{x}'}^T \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \mathbf{r}_{\kappa}(\mathbf{x}') d\mathbf{x}' \\ &- \int_{\Omega} \langle K'(\mathbf{y}) K'(\mathbf{x}') \rangle_{\kappa} \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_{\kappa} d\mathbf{x}' \end{aligned} \quad (3.84)$$

$$\begin{aligned} \langle K'(\mathbf{y}) \nabla \mathcal{G}'(\mathbf{x}, \mathbf{x}_o) \rangle_{\kappa} &\simeq -\nabla \int_{\Omega} \langle K'(\mathbf{y}) \nabla_{\mathbf{x}'}^T \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \rho_{\kappa}(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \\ &- \int_{\Omega} \langle K'(\mathbf{y}) K'(\mathbf{x}') \rangle_{\kappa} \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}_o) \rangle_{\kappa} d\mathbf{x}' \end{aligned} \quad (3.85)$$

Solution of the latter integro-differential equation for $\langle K'(y)\nabla\mathcal{G}'(\mathbf{x}, \mathbf{x}_o)\rangle_\kappa$ makes possible the explicit evaluation of $\langle h'(y)h'(\mathbf{x})\rangle_\kappa$ according to (3.84).

In a similar manner, (3.79) and (3.80) simplify respectively to

$$\begin{aligned}
\langle h'(y)h'(\mathbf{x})\rangle_\kappa &\simeq \int_\Omega \int_\Omega \mathbf{r}_\kappa(y')^T \nabla_{y'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}'(y', y) \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{r}_\kappa(\mathbf{x}') d\mathbf{x}' dy' \\
&\quad - \int_\Omega \int_\Omega \mathbf{r}_\kappa(y')^T [\nabla_{y'} \langle \mathcal{G}(y', y) \rangle_\kappa \rho_\kappa^T(\mathbf{x}', \mathbf{x}) - \langle K'(\mathbf{x}') \nabla_{y'} \mathcal{G}'(y', y) \rangle_\kappa \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa] \\
&\hspace{25em} \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' dy' \\
&\quad - \int_\Omega \int_\Omega \mathbf{r}_\kappa(\mathbf{x}')^T [\nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \rho_\kappa^T(y', y) - \langle K'(y') \nabla_{\mathbf{x}'} \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{y'}^T \langle \mathcal{G}(y', y) \rangle_\kappa] \\
&\hspace{25em} \nabla_{y'} \langle h(y') \rangle_\kappa d\mathbf{x}' dy' \\
&\quad + \int_\Omega \int_\Omega \nabla_{y'}^T \langle h(y') \rangle_\kappa \langle K'(y') K'(\mathbf{x}') \rangle_\kappa \nabla_{y'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(y', y) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' dy' \\
&\quad + \int_\Omega \int_\Omega \langle f'(y') f'(\mathbf{x}') \rangle \langle \mathcal{G}(y', y) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' dy' \\
&\quad + \int_{\Gamma_D} \int_{\Gamma_D} \langle H'(y') H'(\mathbf{x}') \rangle \langle K(y') K(\mathbf{x}') \rangle_\kappa \nabla_{y'}^T \langle \mathcal{G}(y', y) \rangle_\kappa \mathbf{n}(y') \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{n}(\mathbf{x}') d\mathbf{x}' dy' \\
&\quad + \int_{\Gamma_N} \int_{\Gamma_N} \langle Q'(y') Q'(\mathbf{x}') \rangle \langle \mathcal{G}(y', y) \rangle_\kappa \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' dy' \tag{3.86}
\end{aligned}$$

$$\begin{aligned}
\langle \mathcal{G}'(y, y_o) \mathcal{G}'(\mathbf{x}, \mathbf{x}_o) \rangle_\kappa &\approx \int_\Omega \int_\Omega \rho_\kappa(y', y_o)^T \nabla_{y'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(y', y) \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \rho_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' dy' \\
&\quad - \int_\Omega \int_\Omega \rho_\kappa(y', y_o)^T [\nabla_{y'} \langle \mathcal{G}(y', y) \rangle_\kappa \rho_\kappa^T(\mathbf{x}', \mathbf{x}) \\
&\quad \quad - \langle K'(\mathbf{x}') \nabla_{y'} \mathcal{G}'(y', y) \rangle_\kappa \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa] \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}_o) \rangle_\kappa d\mathbf{x}' dy' \\
&\quad - \int_\Omega \int_\Omega \rho_\kappa(\mathbf{x}', \mathbf{x}_o)^T [\nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \rho_\kappa^T(y', y) \\
&\quad \quad - \langle K'(y') \nabla_{\mathbf{x}'} \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{y'}^T \langle \mathcal{G}(y', y) \rangle_\kappa] \nabla_{y'} \langle \mathcal{G}(y', y_o) \rangle_\kappa d\mathbf{x}' dy' \\
&\quad + \int_\Omega \int_\Omega \nabla_{y'}^T \langle \mathcal{G}(y', y_o) \rangle_\kappa \langle K'(y') K'(\mathbf{x}') \rangle_\kappa \nabla_{y'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}'(y', y) \mathcal{G}'(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}_o) \rangle_\kappa d\mathbf{x}' dy' \tag{3.87}
\end{aligned}$$

The latter is an integro-differential equation for $\langle \mathcal{G}'(y, y_o) \mathcal{G}'(\mathbf{x}, \mathbf{x}_o) \rangle_\kappa$ whose solution makes possible the explicit evaluation of $\langle h'(y)h'(\mathbf{x}) \rangle_\kappa$ according to (3.86).

To approximate (3.82) in a manner consistent with the weak approximation (3.58), it would be necessary to first develop integro-differential expressions for the third and fourth moments in a manner analogous to the derivation of (3.77) – (3.80), then replace \mathcal{G} by $\langle \mathcal{G} \rangle_\kappa$ under integrals according to the rule discussed in connection with (3.58). Rather than following this lengthy route, we present in (3.91), below, only the much stronger but simpler approximation obtained by disregarding moments of order higher than two in (3.82).

Upon consistent replacement of \mathcal{G} under integrals by its conditional mean, $\langle \mathcal{G} \rangle_\kappa$, (3.86) simplifies to the following expression for the covariance of the head prediction errors,

$$\begin{aligned}
\langle h'(\mathbf{y})h'(\mathbf{x}) \rangle_\kappa &\simeq - \int_{\Omega} \int_{\Omega} \mathbf{r}_\kappa^T(\mathbf{y}') \nabla_{\mathbf{y}'} \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{r}_\kappa(\mathbf{x}') d\mathbf{x}' d\mathbf{y}' \\
&+ \int_{\Omega} \int_{\Omega} \nabla_{\mathbf{x}'}^T \langle h(\mathbf{y}') \rangle_\kappa \langle K'(\mathbf{y}') K'(\mathbf{x}') \rangle_\kappa \nabla_{\mathbf{y}'} \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\
&+ \int_{\Omega} \int_{\Omega} \langle f'(\mathbf{y}') f'(\mathbf{x}') \rangle \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \\
&+ \int_{\Gamma_D} \int_{\Gamma_D} \langle H'(\mathbf{y}') H'(\mathbf{x}') \rangle \langle K(\mathbf{y}') K(\mathbf{x}') \rangle_\kappa \nabla_{\mathbf{y}'}^T \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \mathbf{n}(\mathbf{y}') \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \mathbf{n}(\mathbf{x}') d\mathbf{x}' d\mathbf{y}' \\
&+ \int_{\Gamma_N} \int_{\Gamma_N} \langle Q'(\mathbf{y}') Q'(\mathbf{x}') \rangle \langle \mathcal{G}(\mathbf{y}', \mathbf{y}) \rangle_\kappa \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa d\mathbf{x}' d\mathbf{y}' \tag{3.88}
\end{aligned}$$

where, according to (3.83),

$$\rho_\kappa(\mathbf{x}, \mathbf{x}_o) \simeq \int_{\Omega} \langle K'(\mathbf{x}) K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} G_\kappa(\mathbf{x}', \mathbf{x}_o) d\mathbf{x}' \tag{3.89}$$

Likewise, we obtain from (3.84) the following approximation required for (3.91) below,

$$\langle K'(\mathbf{y}) \nabla h'(\mathbf{x}) \rangle_\kappa \simeq - \int_{\Omega} \langle K'(\mathbf{y}) K'(\mathbf{x}') \rangle_\kappa \nabla \nabla_{\mathbf{x}'}^T \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_\kappa \nabla_{\mathbf{x}'} \langle h(\mathbf{x}') \rangle_\kappa d\mathbf{x}' \tag{3.90}$$

Note that upon replacing \mathbf{y} by \mathbf{x} in (3.90), one obtains an alternative approximation for $\mathbf{r}_\kappa(\mathbf{x})$ which is stronger than that in (3.58). The quality of all these approximations is expected to improve with the quality of the estimate $\kappa(\mathbf{x})$.

Upon disregarding moments of order higher than two in (3.82), we obtain a much stronger but simpler approximation for the covariance of the flux prediction errors,

$$\begin{aligned}
\langle \mathbf{q}'(\mathbf{y})\mathbf{q}'(\mathbf{x})^T \rangle_\kappa &\simeq -\mathbf{r}_\kappa(\mathbf{y})\mathbf{r}_\kappa(\mathbf{x})^T \\
&+ \kappa(\mathbf{y})\nabla_{\mathbf{y}}\nabla^T \langle h'(\mathbf{y})h'(\mathbf{x}) \rangle_\kappa \kappa(\mathbf{x}) + \nabla_{\mathbf{y}} \langle h(\mathbf{y}) \rangle_\kappa \langle K'(\mathbf{y})K'(\mathbf{x}) \rangle_\kappa \nabla^T \langle h(\mathbf{x}) \rangle_\kappa \\
&+ \nabla_{\mathbf{y}} \langle h(\mathbf{y}) \rangle_\kappa \langle K'(\mathbf{y})\nabla^T h'(\mathbf{x}) \rangle_\kappa \kappa(\mathbf{x}) + \kappa(\mathbf{y}) \langle K'(\mathbf{x})\nabla_{\mathbf{y}} h'(\mathbf{y}) \rangle_\kappa \nabla^T \langle h(\mathbf{x}) \rangle_\kappa
\end{aligned} \tag{3.91}$$

The variance of the fluxes is then given by

$$\begin{aligned}
\langle \mathbf{q}'(\mathbf{x})\mathbf{q}'(\mathbf{x})^T \rangle_\kappa &\simeq -\mathbf{r}_\kappa(\mathbf{x})\mathbf{r}_\kappa(\mathbf{x})^T \\
&+ \kappa(\mathbf{x})^2 \nabla_{\mathbf{y}}\nabla^T [\langle h'(\mathbf{y})h'(\mathbf{x}) \rangle_\kappa; \mathbf{y} = \mathbf{x}] + \langle K'(\mathbf{x})^2 \rangle_\kappa \nabla \langle h(\mathbf{x}) \rangle_\kappa \nabla^T \langle h(\mathbf{x}) \rangle_\kappa \\
&- \kappa(\mathbf{x})\nabla_{\mathbf{y}} \langle h(\mathbf{y}) \rangle_\kappa \mathbf{r}_\kappa^T(\mathbf{x}) - \kappa(\mathbf{x})\mathbf{r}_\kappa(\mathbf{x})\nabla^T \langle h(\mathbf{x}) \rangle_\kappa
\end{aligned} \tag{3.92}$$

3.6 Geostatistical Evaluation of Hydraulic Conductivity Moments

The theory presented thus far assumes that one has at his/her disposal an unbiased estimate $\kappa(\mathbf{x})$ of the hydraulic conductivity function $K(\mathbf{x})$. Our working approximations further require that one know the second conditional moment of the associated estimation errors $\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa$. In this section we recall briefly a two-step geostatistical approach to the estimation of $Y(\mathbf{x}) = \ln K(\mathbf{x})$ from noisy measurements with support ω , iterative generalized least squares (IGLS) followed by residual kriging (Neuman and Jacobson, 1984; Neuman et al., 1987), coupled with maximum likelihood cross-validation (MLCV) (Samper and Neuman, 1989a,b) to minimize bias. In the first step, the estimate $v(\mathbf{x})$ of $Y(\mathbf{x})$ is taken to be a

slowly varying “global drift” such that the “residual field” $Y'(\mathbf{x}) = Y(\mathbf{x}) - v(\mathbf{x})$ is conditionally homogeneous (wide-sense stationary in space). In the second step, the quality of the estimator $v(\mathbf{x})$ is improved via residual kriging which generally reduces the conditional variance and correlation scales of $Y'(\mathbf{x})$ but renders the latter conditionally nonhomogeneous. The functional forms and parameters of the drift and residual semivariogram models are estimated by means of IGLS and MLCV. We supply expressions for $\kappa(\mathbf{x})$ and $\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa$ corresponding to both estimation steps under the assumption that $Y(\mathbf{x}) = \ln K(\mathbf{x})$ is Gaussian. As shown in Appendix F, if $K(\mathbf{x})$ is log normal then

$$\kappa(\mathbf{x}) = K_g(\mathbf{x}) \exp\left(\frac{1}{2}\sigma_Y^2(\mathbf{x})\right) \quad K_g(\mathbf{x}) = \exp[v(\mathbf{x})] \quad (3.93)$$

where $v(\mathbf{x})$ is an unbiased estimate of $Y(\mathbf{x})$ and $\sigma_Y^2(\mathbf{x}) = \langle [Y'(\mathbf{x})]^2 \rangle_v$ is the conditional variance of the estimation error $Y'(\mathbf{x}) = Y(\mathbf{x}) - v(\mathbf{x})$, the subscript v indicating that conditioning is done on the same data as those used to obtain $v(\mathbf{x})$. By the same token,

$$\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa = \kappa(\mathbf{x})\kappa(\mathbf{y})\{\exp\langle Y'(\mathbf{x})Y'(\mathbf{y}) \rangle_v - 1\} \quad (3.94)$$

where $\langle Y'(\mathbf{x})Y'(\mathbf{y}) \rangle_v$ is a conditional covariance. Hence $\kappa(\mathbf{x})$ and $\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa$ are easily computed from $v(\mathbf{x})$ and $\langle Y'(\mathbf{x})Y'(\mathbf{y}) \rangle_v$ which in turn can be evaluated as described below. Let \mathbf{Y}^* be a vector of measured $Y(\mathbf{x})$ values on supports ω centered about J discrete points \mathbf{x}_j such that

$$Y_j^* = Y(\mathbf{x}_j) + \mu_j = v(\mathbf{x}_j) + Y'(\mathbf{x}_j) + \mu_j \quad j = 1, 2, \dots, J \quad (3.95)$$

where μ_j are unbiased measurement errors that show neither autocorrelation nor cross-correlation with $Y(\mathbf{x}_j)$. Hence

$$\langle (\mathbf{Y}' + \boldsymbol{\mu})(\mathbf{Y}' + \boldsymbol{\mu})^T \rangle_v = \mathbf{V} = \langle \mathbf{Y}'\mathbf{Y}'^T \rangle_v + \mathbf{W} \quad (3.96)$$

where \mathbf{W} is a diagonal matrix of measurement error variances which we consider known. In a manner similar to Neuman and Jacobson (1984) we designate $v(\mathbf{x})$ as drift and represent it by a finite-dimensional function

$$v(\mathbf{x}) = \sum_{m=1}^M c_m s_m(\mathbf{x}) \quad (3.97)$$

where M is a positive integer, c_m are coefficients, and $s_m(\mathbf{x})$ are prescribed coordinate functions. Neuman and Jacobson describe a stepwise regression method which, for a set of trial basis functions, seeks the smallest value of M that renders $Y'(\mathbf{x})$ conditionally homogeneous; for such M , the quantities $\langle Y'Y'^T \rangle_v$ and c_m are evaluated simultaneously by IGLS; Samper and Neuman (1989a,b) couple this process with MLCV to reduce its bias.

Let $\gamma_v(\mathbf{x}, \mathbf{s})$ be the conditional semivariogram of $Y'(\mathbf{x})$ defined as

$$\gamma_v(\mathbf{x}, \mathbf{s}) \equiv \frac{1}{2} \langle [Y'(\mathbf{x} + \mathbf{s}) - Y'(\mathbf{x})]^2 \rangle_v \quad (3.98)$$

where \mathbf{s} is a displacement (or spatial lag) vector. Viewing Y_j^* and μ_j as point values of corresponding random fields $Y^*(\mathbf{x})$ and $\mu(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_j$ allows defining a conditional semivariogram $\tilde{\gamma}_v(\mathbf{x}, \mathbf{s})$ for the residuals $Y^*(\mathbf{x}) - v(\mathbf{x}) = Y'(\mathbf{x}) + \mu(\mathbf{x})$ as

$$\begin{aligned} \tilde{\gamma}_v(\mathbf{x}, \mathbf{s}) &\equiv \frac{1}{2} \langle \{ [Y'(\mathbf{x} + \mathbf{s}) + \mu(\mathbf{x} + \mathbf{s})] - [Y'(\mathbf{x}) + \mu(\mathbf{x})] \}^2 \rangle_v \\ &= \gamma_v(\mathbf{x}, \mathbf{s}) + \frac{1}{2} [\sigma_\mu^2(\mathbf{x} + \mathbf{s}) + \sigma_\mu^2(\mathbf{x})] \end{aligned} \quad (3.99)$$

where s is the magnitude of \mathbf{s} and $\sigma_\mu^2(\mathbf{x})$ is the variance of $\mu(\mathbf{x})$. If $Y'(\mathbf{x})$ is conditionally homogeneous then $\sigma_Y^2(\mathbf{x}) \equiv \sigma_Y^2 = \text{constant}$ and the conditional covariance of $Y'(\mathbf{x})$ depends only on \mathbf{s} ,

$$\langle Y'(\mathbf{x})Y'(\mathbf{x} + \mathbf{s}) \rangle_v = C(\mathbf{s}) \quad (3.100)$$

as does γ_v , the two being related via

$$C(\mathbf{s}) = \sigma_Y^2 - \gamma_v(\mathbf{s}) \quad (3.101)$$

At $s = 0$, $\gamma_v(\mathbf{s}) = 0$; as s increases, $\gamma_v(\mathbf{s})$ tends to a constant “sill” σ_Y^2 , and $C(\mathbf{s})$ tends to zero. The components of \mathbf{V} in (3.96) are thus given by

$$V_{ij} = \sigma_Y^2 - \gamma_v(\mathbf{x}_i - \mathbf{x}_j) + \delta_{ij}\sigma_\mu^2(\mathbf{x}_i) \quad (3.102)$$

where δ_{ij} is the kronecker delta (equal to 1 if $i = j$ and to zero if $i \neq j$). The latter matrix is required by Neuman and Jacobson (1984) to estimate the drift $v(\mathbf{x})$ and to compute the covariance of the corresponding estimation error.

Neuman and Jacobson (1984) estimate $\tilde{\gamma}_v(\mathbf{s})$ by the method of moments from the so-called “experimental (or sample) semivariogram”

$$\hat{\gamma}_v(\mathbf{s}) = \frac{1}{2P(\mathbf{s})} \sum_{j=1}^J \{[Y^*(\mathbf{x}_j + \mathbf{s}) - v(\mathbf{x} + \mathbf{s})] - [Y^*(\mathbf{x}_j) - v(\mathbf{x})]\}^2 \quad (3.103)$$

where $P(\mathbf{s})$ is the number of data pairs within a given tolerance (or window) of s . As measurement error variances are considered known, one can obtain a corresponding estimate of $\gamma_v(\mathbf{s})$ from (3.99), then select M in (3.97) as the smallest integer that renders this estimate flat at high s values (implying the attainment of a constant sill, which in turn is taken as an indication that $Y'(\mathbf{x})$ is conditionally homogeneous), and finally estimate $\langle Y'(\mathbf{x})Y'(\mathbf{x} + \mathbf{s}) \rangle_v$ by means of (3.100) – (3.101). The method suffers from some bias which however is greatly reduced when one uses MLCV (Samper and Neuman, 1989a,b) rather than the method of moments to calculate $\hat{\gamma}_v(\mathbf{s})$, and cross-validation statistics rather than visual inspection to terminate the iterative process. MLCV offers additional advantages over the traditional method of moments: It provides information about the quality of the estimated semivariogram parameters and can handle data associated with variable supports.

The above geostatistical analysis allows simplifying (3.93) and (3.94) in the following manner,

$$\kappa(\mathbf{x}) = K_g(\mathbf{x}) \exp\left(\frac{1}{2}\sigma_Y^2\right) \quad K_g(\mathbf{x}) = \exp[v(\mathbf{x})] \quad (3.104)$$

$$\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa = \kappa(\mathbf{x})\kappa(\mathbf{y})\{\exp[\sigma_Y^2 - \gamma_v(\mathbf{s})] - 1\} \quad (3.105)$$

It is clear that $K'(\mathbf{x})/\kappa(\mathbf{x})$ is conditionally homogeneous because $\langle K'(\mathbf{x}) \rangle_\kappa \equiv 0$ by virtue of (F.4) in Appendix F, and $\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa/\kappa(\mathbf{x})\kappa(\mathbf{y})$ in (3.105) depends only on the displacement \mathbf{s} but not on the locations \mathbf{x} and \mathbf{y} .

It is often possible to obtain a better estimate of $\kappa(\mathbf{x})$ through a reduction in the conditional variance and correlation scales of $Y'(\mathbf{x})$ via residual kriging, but this generally renders $Y'(\mathbf{x})$ and $K'(\mathbf{x})/\kappa(\mathbf{x})$ conditionally nonhomogeneous. In a manner analogous to Neuman and Jacobson (1984) we designate the generalized least squares estimates of $v(\mathbf{x})$ and $Y'(\mathbf{x})$ by $\hat{v}(\mathbf{x})$ and $\hat{Y}'(\mathbf{x})$, respectively, such that

$$Y(\mathbf{x}) = v(\mathbf{x}) + Y'(\mathbf{x}) = \hat{v}(\mathbf{x}) + \hat{Y}'(\mathbf{x}) \quad (3.106)$$

We define the kriging estimate of $Y(\mathbf{x}_i)$ at some point \mathbf{x}_i as

$$\tilde{Y}(\mathbf{x}_i) = \hat{v}(\mathbf{x}_i) + \tilde{Y}'(\mathbf{x}_i) \quad (3.107)$$

where $\tilde{Y}'(\mathbf{x}_i)$, the kriging estimate of $Y'(\mathbf{x}_i)$, is given by

$$\tilde{Y}'(\mathbf{x}_i) = \sum_{p=1}^{P_i} \alpha_{ip} \hat{Y}'_{ip}^* \quad (3.108)$$

and $\hat{Y}'_{ip}^* = Y_{ip}^* - \hat{v}(\mathbf{x}_i)$, Y_{ip}^* being P_i measurements of $Y(\mathbf{x})$ at points \mathbf{x}_{ip} surrounding \mathbf{x}_i . Note that P_i must not exceed the total number of available measurements, P , and that a given measurement may be used to estimate $\hat{Y}'(\mathbf{x}_i)$ at more than one point

\mathbf{x}_i . The kriging coefficients α_{ip} satisfy the $(P_i + 1)$ kriging equations (e.g., Clifton and Neuman, 1982; *ibid*)

$$\sum_{r=1}^{P_i} \alpha_{ir} \gamma_v(\mathbf{x}_{ip}, \mathbf{x}_{ir}) - \alpha_{ip} \sigma_{\mu_{ip}}^2 + \beta_i = \gamma_v(\mathbf{x}_{ip}, \mathbf{x}_i) \quad p = 1, 2, \dots, P_i \quad (3.109)$$

$$\sum_{r=1}^{P_i} \alpha_{ir} = 1 \quad (3.110)$$

for every $i = 1, 2, \dots, I$, where $\sigma_{\mu_{ip}}^2$ is the variance of μ at \mathbf{x}_{ip} and β_i is a Lagrange multiplier. These equations can be solved uniquely for α_{ir} and β_i provided γ_v is positive definite.

The estimation (or kriging) errors

$$e(\mathbf{x}_i) \equiv \tilde{Y}(\mathbf{x}_i) - Y(\mathbf{x}_i) = \tilde{Y}'(\mathbf{x}_i) - \hat{Y}'(\mathbf{x}_i) \quad i = 1, 2, \dots, I \quad (3.111)$$

have zero conditional mean by virtue of (3.110). Their conditional variance and covariance are, respectively (Clifton and Neuman, 1982; Neuman and Jacobson, 1984),

$$\langle e(\mathbf{x}_i)^2 \rangle_v = \sum_{p=1}^{P_i} \alpha_{ip} \gamma_v(\mathbf{x}_{ip}, \mathbf{x}_i) + \beta_i \quad (3.112)$$

$$\begin{aligned} \langle e(\mathbf{x}_i) e(\mathbf{x}_j) \rangle_v &= -\gamma_v(\mathbf{x}_i, \mathbf{x}_j) + \sum_{p=1}^{P_i} \alpha_{ip} \gamma_v(\mathbf{x}_{ip}, \mathbf{x}_j) + \sum_{r=1}^{P_j} \alpha_{jr} \gamma_v(\mathbf{x}_{jr}, \mathbf{x}_i) \\ &+ \sum_{p=1}^{P_i} \sum_{r=1}^{P_j} \alpha_{ip} \alpha_{jr} [\gamma_v(\mathbf{x}_{ip}, \mathbf{x}_{jr}) - \delta_{ip,jr} \sigma_{ip} \sigma_{jr}] \quad i \neq j \end{aligned} \quad (3.113)$$

By virtue of (3.111), we can rewrite (F.1) in Appendix F as

$$K(\mathbf{x}) = \exp[Y(\mathbf{x})] = \exp[\tilde{Y}(\mathbf{x}) - e(\mathbf{x})] \quad (3.114)$$

Then, in analogy to (3.93) and (3.94),

$$\kappa(\mathbf{x}) = \exp[\tilde{Y}(\mathbf{x})] \exp\left(\frac{1}{2} \langle e(\mathbf{x})^2 \rangle_v\right) \quad (3.115)$$

$$\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_{\kappa} = \kappa(\mathbf{x})\kappa(\mathbf{y})\{\exp(\epsilon(\mathbf{x})\epsilon(\mathbf{y}))_{\nu} - 1\} \quad (3.116)$$

Though the conditional mean of $K'(\mathbf{x})/\kappa(\mathbf{x})$ remains zero, its conditional second moment now depends on the locations \mathbf{x} and \mathbf{y} due to the nonhomogeneity of $\epsilon(\mathbf{x})$. Note however that, since kriging is a linear estimator, $\epsilon(\mathbf{x})$ depends only on data locations \mathbf{x}_{ip} and not on their values Y_{ip}^* .

CHAPTER 4

MONTE CARLO SIMULATIONS AND RANDOM FIELD GENERATORS

The ultimate and most “straightforward” method for “solving” stochastic PDE’s is by Monte Carlo simulations (MCS). Monte Carlo simulations constitute a versatile mathematical tool capable of handling situations where all other methods fail (e.g., where high variances of hydraulic conductivities and/or intricate boundary conditions and/or non-uniform flows are involved). Therefore MCS are used to verify theories and closure approximations of the solutions of stochastic PDE’s. In this method, simulated (or generated) alternative (or equally probable) realizations of the spatial process (e.g., hydraulic conductivity), which are statistically similar (i.e., have identical probability distributions and preserve the same covariance structure), provide the inputs to flow models (see Fig. 8.3 of Christakos, 1992). These numerical models (or PDE solvers) produce an output field (e.g., of heads, velocities, etc.) for each (input) realization; the ensemble of output fields is studied statistically to yield the moments (i.e., mean, covariance, variogram) and/or the probability distributions of the output at each point in space. The generation of such realizations is the subject of this chapter, and the numerical methods and models used by us for the (actual) flow simulations are the subject of chapter 5.

When the input (generated) fields are similar statistically, but disregard actual known (or measured) data values, the simulations are *unconditional*; when “measurements”, or known values, of the input parameter at some points in space are the same

for all realizations, the simulations become *conditional*. Conditional (input) fields (or realizations) are generally non-stationary, despite possible stationarity of the original unconditional field.

MCS (i.e., simulations) of flows in random fields require large, high resolution (fine) grids, to yield meaningful results. It is generally advised to have at least 5 grid cells (preferably more than 10; Dagan and Indelman, pers. com.) per correlation scale¹ (in each spatial direction, and, when relevant, in time); to avoid obstruction by boundary effects (e.g., when an infinite domain is of interest), the domain size should exceed the correlation scale by a factor of at least 10, and preferably much more. (e.g., Ababou et al., 1989; Dykaar and Kitanidis, 1992; Follin, 1992; Desbarats, 1992a). For highly variable conductivity fields (i.e., when the variances of the log-conductivity is greater than 1), and/or steep head gradients, further subdivision of each grid cell is desirable to reduce numerical errors and smoothing (as will be discussed later). Such MCS are computationally intensive, i.e., require ample computer power, in terms of both CPU time and memory. In fact, even the present supercomputers are limited (in terms of memory and allocated CPU time) to problems of relatively small size, particularly in three dimensions (e.g., less than $100 \times 100 \times 100$). Moreover, MCS are *numerical* simulations; when applied to strongly varying fields, large numerical errors may result, and special care has to be taken to minimize the errors. Furthermore, as will be shown in the sequel, the field generation is far from being a “straightforward” and a reliable procedure. As field generation is central to MCS, we devote the rest of the chapter to this subject.

The choice of a random field generator (RFG) has a central role in Monte

¹A measure of the distance beyond which correlation ceases.

Carlo simulations. It becomes crucial when steep gradients are involved, as in the case of converging radial flow; this is due to the sensitivity of the solutions to both variability and deficiencies of the generated random fields. This sensitivity is implied from (1.92), (1.93), (1.79) in Chapter 1, where the reduction factor (i.e., the ratio of head variance to conductivity variance) is approximately proportional to the square of mean gradient, (i.e., to \mathbf{J}^2).

Knowing the random variables u_1, \dots, u_k , the conditional density of u_{k+1}, \dots, u_n is given by (Papoulis, 1984; cf. Appendix A)

$$f(u_{k+1}, \dots, u_n | u_1, \dots, u_k) = \frac{f(u_1, \dots, u_k, \dots, u_n)}{f(u_1, \dots, u_k)}. \quad (4.1)$$

Consequently, from the *chain rule*, (Papoulis, 1984; Christakos, 1992)

$$f(u_1, \dots, u_n) = f(u_1)f(u_2|u_1) \cdots f(u_n | u_1, \dots, u_{n-1}). \quad (4.2)$$

If u_i depend on the position \mathbf{x}_i in continuous space, they become spatial random functions $u(\mathbf{x}_i)$, or simply a random field.

Following Christakos (1992), based on (4.2) one can calculate the probability densities $f(u_1)$, $f(u_2|u_1)$, \dots , and $f(u_n | u_1, \dots, u_{n-1})$. The corresponding probability *distributions* can also be found, i.e.,

$$\begin{aligned} F(u_1) &= \int_0^{u_1} f(y)dy \\ F(u_2|u_1) &= \int_0^{u_2} f(y|u_1)dy, \end{aligned} \quad (4.3)$$

and so on. Next, the following system of equations of marginal and conditional distributions is developed (Christakos, 1992):

$$F(u_1) = v_1$$

$$\begin{aligned}
 F(u_2|u_1) &= v_2 \\
 &\vdots \\
 F(u_n|u_1, \dots, u_{n-1}) &= v_n,
 \end{aligned}
 \tag{4.4}$$

where v_i , $i = 1, \dots, n$ are independent random numbers.

Finally, realizations of (univariate) $u(\mathbf{x})$ can be generated by solving the system (4.4) with respect to u_i . If the generated realizations “honor” the measured values at the data locations, they become “conditioned”, and the simulations become *conditional*. Otherwise, the generated realizations are *unconditional*. According to Carr and Myers (1985), from the geostatistical point of view, unconditional simulations are “the most mathematically rigorous and fundamental aspect of conditional simulations... conditioning simply involves linear estimation and then differencing.”

In practice, several techniques are available for generating random fields consistent with the system (4.4); however, only the sequential simulation technique utilizes (4.4) directly. The methods most commonly used for generating spatially correlated random fields are:

1. Matrix (LU) decomposition.
2. Spectral methods and Fast Fourier transform (FFT).
3. Turning bands method (*tbm*).
4. Sequential simulation.

As will be shown in the sequel, the first three methods use finite linear combinations of uncorrelated random variables; consequently, the common distribution of these random variables should be such that it is preserved under finite linear combinations

(Myers, 1989); as the normal (or Gaussian) distribution is the only distribution with this property, it is most frequently used. [In addition, the normal distribution is the simplest distribution, requires only the first two moments]. While the first two methods generate only Gaussian fields, based on a single covariance function, the sequential simulator features a rich family of spatial structures, not limited by Gaussianity, including Indicator simulation². Each of the above methods can serve as a basis for generating both unconditional and conditional simulations; however, the conditioning of the sequential simulator is done by construction (Gómez-Hernández, 1991). In the following we will briefly review the basic steps used in the above four methods for generating unconditional random fields whose *pdf* is multivariate normal (or log-normal).

4.1 The LU Decomposition Method

The lower-upper triangular matrix technique was developed independently by Barr and Slezak (1972), Clifton and Neuman (1982), Elishakoff (1983), and is part of major software libraries, such as IMSL and NAG. It was recently extended by Davis (1987) and Alabert (1987) to conditional simulations, and by Myers (1989) to the use of cokriging for co-simulations of two correlated random variables (based on their cross-covariance or cross-variogram).

The steps for (simple) unconditional simulations are as follows (Christakos, 1992, Harter, 1992):

- (1) For a given covariance model, $C(s)$, develop the corresponding covariance

²Indicator formalism allows numerical or interpretive information to be commonly coded into elementary bits (0-1) that are based on the exceedence of given threshold values (Journel and Alabert, 1990).

matrix \mathbf{C} of size $n \times n$, which must be a semipositive definite symmetric matrix, and can be decomposed to a lower triangular matrix \mathbf{L} and an upper triangular matrix \mathbf{U} , say by the Cholesky algorithm:

$$\mathbf{C} = \mathbf{L}\mathbf{U}. \quad (4.5)$$

(2) For a vector \mathbf{V} of n independent standard Gaussian random variables, $N[0, 1]$ (i.e., zero mean and unit variance), define the vector

$$\mathbf{Y} = \mathbf{L}\mathbf{V}, \quad (4.6)$$

which has zero mean [this can be seen by taking the expected value of (4.6)], and covariance

$$\langle (\mathbf{L}\mathbf{V})(\mathbf{L}\mathbf{V})^T \rangle = \mathbf{L}[\mathbf{V}\mathbf{V}^T]\mathbf{U} = \mathbf{L}\mathbf{U} = \mathbf{C}, \quad (4.7)$$

as required. (\mathbf{I} is the identity matrix, and T denotes “transpose”). This means that once \mathbf{C} is decomposed as in (4.5), the correlated vector \mathbf{Y} (our random field) can be readily computed by (4.6) for any vector \mathbf{V} . The vectors \mathbf{V} are readily created for each realization by a (hopefully good) random number generator (RNG)³. This means that no matter how many fields are to be generated, the covariance matrix needs to be LU-decomposed only once. Moreover, the method is independent of the number of dimensions of the random field (except for size).

The LU decomposition method is perhaps the most “straightforward” method for generating correlated random fields, in that the covariance matrix is utilized directly. However, (computer) memory requirements become prohibitive when relatively small fields are generated, particularly in 3-dimensional domains. For example, the generation of a two-dimensional field with 60×60 (square) elements (i.e., 61×61

³Vector RNG’s are available in most (computer) libraries.

nodes), involves a correlation matrix of 3600×3600 components, which requires 68 mega-bites (MB) of memory; a 128×128 field, requires 242 MB of memory. Since Monte Carlo simulations mostly involve large fields, this elegant method has been taken over by faster methods, having lesser memory requirements.

4.2 Spectral Methods

The family of spectral methods, particularly those using Fast Fourier Transforms (FFT) (Gutjahr, 1987; Christakos, 1992; see also references for the *tbm*), take advantage of the spectral representation theorem, which expresses a spatially correlated random process in terms of uncorrelated random process in the spectral (or frequency, or wave number) domain (Harter, 1992). The methods are applied to one-, two- and three- dimensional fields, and are limited to (second order) stationary random processes. These are, perhaps, the fastest methods for generating large random fields.

The spectrum (or spectral density) of a statistically homogeneous and isotropic process (or field), Y , is the Fourier transform (FT) of its covariance function $C(s) = Cov[Y(x+s), Y(x)]$ (e.g., De Marsily, 1986; Papoulis, 1984)

$$S(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} C(s)e^{-iks} ds, \quad (4.8)$$

where k is the *wave number*. Using the inverse Fourier transform,

$$C(s) = \int_{-\infty}^{+\infty} S(k)e^{iks} dk. \quad (4.9)$$

The *representation theorem* states that if the second order stationary stochastic process $Y(x)$ is of zero mean [$\langle Y \rangle = 0$], and of covariance $C(s)$, then one can define a

complex associated process Z that satisfies

$$Y(x) = \int_{-\infty}^{+\infty} e^{iks} dZ(k), \quad (4.10)$$

such that

$$\begin{aligned} \langle dZ(k_1) dZ^*(k_2) \rangle &= 0 & \text{if } k_1 \neq k_2 \\ \langle dZ(k_1) dZ^*(k_2) \rangle &= S(k_1) & \text{if } k_1 = k_2 \end{aligned} \quad (4.11)$$

where dZ is also zero mean. Equation (4.10) is a Fourier-Stieltjes integral, and the asterisk (*) in (4.11) denotes the complex conjugate.

The $dZ(k)$ process can be constructed by setting

$$dZ(k) \equiv \sqrt{S(k)dk} [\alpha(k) + i\beta(k)], \quad (4.12)$$

where α and β are uncorrelated, zero mean, uniformly or normally distributed random variables, with variance of $\frac{1}{2}$.

Equation (4.10) can alternatively be written as (Zimmerman and Wilson, 1990)

$$Y(x) = 2Re \left\{ \int_0^{+\infty} e^{iks} dZ(k) \right\} \approx 2Re \left\{ \int_0^{\Omega} e^{iks} dZ(k) \right\}, \quad (4.13)$$

where Ω is a maximum (cut-off) frequency. In a discretised form,

$$Y(x) \approx 2Re \left\{ \sum_{m=1}^{M-1} e^{ik_m s} dZ(k) \right\}, \quad (4.14)$$

where M is the number of harmonics,

$$k_m = (m + 1/2)\Delta k, \quad (4.15)$$

and $\Delta k = \Omega/M$. Combining (4.12), (4.14), and (4.15) yields

$$Y(x) \approx 2Re \left\{ \sum_{m=0}^{M-1} e^{ik_m s} \sqrt{S(k_m)\Delta k} [\alpha_m + i\beta_m] \right\}. \quad (4.16)$$

Consequently, the FFT method consists of

- (1) Determining the cut-off frequency (Ω) and the number of harmonics, M (e.g., based on analysis and suggestions by Tompson et al., 1989), and discretizing the frequency (or wave-number) domain into intervals Δk .
- (2) For each m , (i) calculate k_m , according to (4.15), (ii) calculate $S(k_m)$, according to (4.8), and (iii) generate uncorrelated, normally (or uniformly) distributed random numbers α_m and β_m .
- (3) Use (4.16) to calculate $Y(\mathbf{x}) = \ln K(\mathbf{x})$, at each location \mathbf{x} (i.e., for each grid node).

Similar to the turning bands method (below), the discretisation and truncation used in spectral methods is crucial. For example, too low a cutoff frequency (Ω) in (4.13) would reduce the variance of the generated field (Wilson and Mantoglou, 1982); more accurately, it would cause *aliasing* (Gutjahr, 1989; Harter, 1992), which implies not only missing information contained in frequencies above the cutoff frequency (or the so called *Nyquist frequency*), but also a wrong interpretation of high frequencies as low frequencies; in other words, aliasing means that if the true spectrum has energy above the cutoff frequency, then the estimated spectrum will be in error at all values (Lumely and Panofsky, 1964, pp. 54-55).

A detailed description of the spectral method for generating both conditional and unconditional random fields is given by Harter (1992). Detailed analyses of the effects of discretisation on the characteristics of the generated fields are discussed by Mantoglou and Wilson (1982), Tompson et al. (1989), and Zimmerman and Wilson (1990). Variants of the spectral method are described by Press et al. (1986), and Christakos (1992).

4.3 Turning Bands Method

The turning bands method (*tbm*) was first suggested by Matheron (1973); it was further developed and used by Journel (1984), Journel and Huijbregts (1978), Delhomme (1979), Mantoglou and Wilson (1982), Tompson et al. (1989), and Ababou (personal communication). The method has gained popularity mainly due to its ability to rapidly generating random fields of practically any size and any dimensionality. Although initially limited to (stationary) *isotropic* fields, one could transform the generated fields into anisotropic ones, by using a suitable transformation of the coordinate system (e.g., Mantoglou and Wilson, 1982). Recently, the method has become controversial, for the reasons discussed below. The fundamental characteristic of the turning bands method is that it produces a simulation of a random function in n -D as a linear combination of independent simulations of (correlated) random functions defined in 1-D, such that the covariance function and the mean are preserved. The reduction in dimensionality, which is the most attractive property of the *tbm*, is made possible by the fact that the transformation from 3- or 2-dimensional covariance function into an equivalent 1-dimensional covariance function can be uniquely defined (Matheron, 1973; Myers, 1990).

The basic steps of the turning bands method for generating random fields are:

- (1) Determine the covariance model $C_n(s)$ or the spectral density $S_n(w)$.
- (2) Given the covariance (or spectral density) of the n -dimensional random field, find the corresponding 1-dimensional covariance, $C_1(s)$ (or spectral density, $S_1(w)$).
- (3) Generate a 1-dimensional multivariate, stationary, zero-mean process $Z_j(\xi)$ on each line, using an appropriate autoregressive–moving average technique, or a spectral method (as above). Ideally, a 2-D or a 3-D covariance is attained only as a limit

for an infinite number (N) of lines. In practice, a finite number of *evenly* spaced lines (say, $N = 100$) is used. In three dimensions, one could either take a large number of *randomly* distributed lines over a unit sphere (e.g., $N = 100$, as done by Tompson et al, 1989), or use $N = 15$ *evenly* spaced lines joining the mid-points of the opposite edges of a regular icosahedron, which is a regular polyhedron with the maximum number of faces (i.e., 20 faces; Journel and Huijbregts, 1978). In other words, from geometric considerations, 15 lines in 3-D constitute the best possible division of space to evenly spaced lines (Myers, 1990). In two dimensions, any number of *evenly* spaced lines can (theoretically) be used; Wilson and Mantoglou (1982) show that 16 evenly spaced lines yield a reasonably accurate spatial covariance function (though averaged over many realizations). However, as we will demonstrate in the sequel, at least 32 evenly spaced lines are needed in order to avoid a distortion effect associated with the appearance of line-like patterns in both the generated fields and their spatial covariances; in fact, Ababou (personal communication) recommends using 1000 lines. Each of these lines is divided into small discrete intervals of equal size (e.g., Mantoglou and Wilson, 1982, Tompson et al., 1989), each of which represents one (random) value.

(4) At each point in the field, say, \mathbf{x}_o , calculate the average contributions projected from all lines onto that point, i.e.,

$$Y(\mathbf{x}_o) = \frac{1}{\sqrt{N}} \sum_{j=1}^N Z(\xi_o, j), \quad (4.17)$$

where j is the line index, and \mathbf{x}_o is orthogonal to ξ_o (see Figure 4.1). Note that the application of the turning bands method involves discretisation at each step; moreover, theoretically, the sum in (4.17) should be an integral (Myers, 1989). When using the spectral method to generate the line processes, the Fourier transforms (FT)

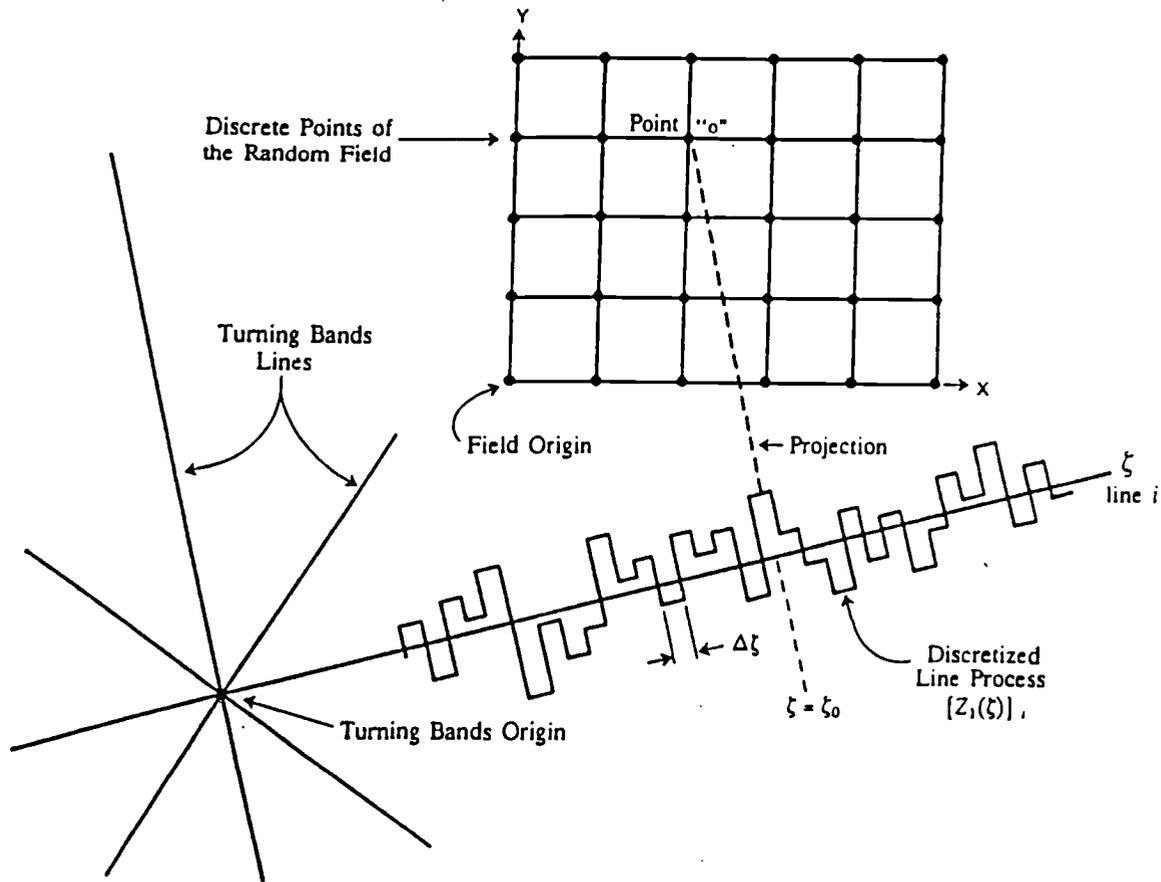


Figure 4.1: Schematic representation of a discrete random field and the turning bands lines (adapted from Zimmerman and Wilson, 1990).

that are usually employed introduce an additional discretisation problem (expressed by the number of harmonics required), as well as a truncation problem (expressed by the cut-off frequency). As will be shown, the accumulated effect of the discretisation at different stages, and the use of truncated FT, result in “stripping” pathology and/or “star” patterns seen in particular realizations (see also, Desbarats, 1987a; Gutjahr, 1989, and Tompson et. al, 1989). These patterns can be attenuated by increasing the number of lines and refining the discretisations; by doing so, however, one increases the CPU time considerably, losing some of the advantage the turning bands method.

4.4 Sequential Simulations

The method of sequential simulations (also called *the method of conditional distribution*, e.g., Johnson, 1987) was, perhaps, the first known technique for generating multivariate distributions. It was recently suggested by Journel (1984), and implemented by Gómez-Hernández (1991) and by Deutsch and Journel (1992) for generating Indicator and Gaussian random fields.

Based on equation (4.2) above, given the joint *pdf*, $f(u_1, \dots, u_n)$, one can calculate the marginal density $f(u_1)$ of a “known” variable, and then, the conditional (univariate) densities of the other (unknown) variables, $f(u_2|u_1), \dots$, and $f(u_n | u_1, \dots, u_{n-1})$. The general simulation algorithm is as follows (Harter, 1992):

Given measured (or known) data values u_1, \dots, u_k and their marginal distributions,

1. Select an unknown “data” point (value), u_{k+1} , to be generated (i.e., select the point for which you wish to generate a value of the random function u).

2. Find the conditional density of that data point, and draw u_{k+1} from it.
3. Select a second data point, u_{k+2} , to be generated.
4. Find the conditional density for that data point, given the known data u_1, \dots, u_k, u_{k+1} and their densities, and draw u_{k+2} from it.
5. Repeat the procedure until all unknown data points become “known” for that particular realization.

In the case of multivariate normal distribution, the conditional *pdf*, $f(u_{k+1}, \dots, u_n | u_1, \dots, u_k)$, is also $n - k$ variate normal. The conditional expected value of u_i ($i > k$) is given by (Dagan, 1989)

$$\langle u_i \rangle_\kappa \equiv \langle u_i | u_1, \dots, u_k \rangle = \langle u_i \rangle + \sum_{j=1}^k \lambda_{ij} (u_j - \langle u_j \rangle), \quad (4.18)$$

where $\langle u_i \rangle$ is the unconditional mean of u_i . The constant, deterministic coefficients of λ_{ij} are solutions of the linear (simple kriging) system

$$\sum_{j=1}^k \lambda_{ij} C_{jl} = C_{il} \quad l = 1, \dots, k, \quad (4.19)$$

where $C_{ij} = Cov[u_i, u_j] = \langle (u_i - \langle u_i \rangle)(u_j - \langle u_j \rangle) \rangle$. The conditional variance (also known as kriging or estimation variance) of u_i is given by (Dagan, 1989)

$$\sigma_\kappa^2 \equiv \sigma^2(u_i | u_1, \dots, u_k) = \sigma_{u_i}^2 - \sum_{l=1}^k \lambda_{il} C_{li}, \quad (4.20)$$

where σ_κ^2 and $\sigma_{u_i}^2$ are the conditional and unconditional variances of u_i respectively. As the kriging errors are (ideally) uncorrelated, the procedure for generating new values for each realization can be simplified to

1. Find (or calculate) λ_{ij} for the new data point i , based on the known covariance.

2. Calculate the estimate $\langle u_i \rangle_\kappa$ at a new point by using (4.18).
3. Generate univariate normal, zero mean, random numbers with variance $\sigma_{u_i}^2$ in (4.20), and draw a random error e_i .
4. The new value is then $u_i = \langle u_i \rangle_\kappa + e_i$.

Although it seems adequate only for *conditional* simulations, unconditional random fields can be readily generated by drawing the first value from a (zero mean) normal distribution with the unconditional variance σ_u^2 , and then, each additional point value is conditioned on the previously generated ones. As mentioned above, the sequential simulator is the most general (and hence, the most powerful) RFG. However, like the turning band simulators, some “art” is needed to construct the best input for each particular problem. While in the turning bands method, it is the discretisation and truncation that have to be carefully considered, the input considerations for the sequential simulator are similar to those of variogram estimation and kriging. For example, the search neighborhood is recommended by Gómez-Hernández (1991) to be at least as large as the “range” (which is approximately 1/3 of the integral scale of the exponential variogram we were using); we found that at least four (preferably five) integral scales should be set as the “radius” of the search neighborhood, and at least two points at each quadrant are needed for a reasonable reproduction of the (theoretical) variogram. As in kriging, at each new point, the kriging system has to be solved. Because of the large neighborhood needed, the number of additional points increases exponentially as the generation process progresses. Consequently, the simulator becomes one of the slowest RFG’s. On the other hand, in the early stage of an unconditional field generation, when only a few “data points” exist, the genera-

tion of new points suffers from insufficient points to correlate with (or weights λ_{ij} to use); hence, insufficient accuracy for the new generated points, and consequently, a deficient overall accuracy.

Perhaps the greatest drawback of the sequential simulations is the so called *random path*, which is the visiting sequence according to which new points are generated sequentially, at random locations. This requires an additional random number generator (RNG), and introduces an additional source of error. In particular, numbers generated by this RNG which are larger than the number of nodes, are discarded (Gómez-Hernández, 1991, p. 54). Nevertheless, with a new (and better) RNG, we have found the sequential simulator (GCOSIM-3D, Gómez-Hernández, 1991)⁴ to be superior, or at least not inferior, to other RFG's we have tested (see below).

4.5 Pseudo Random Number Generators

As mathematical and numerical algorithms are deterministic, they cannot generate “random numbers”, but rather pseudo random numbers. There is a vast literature devoted to uniform random number generators (RNG's) and their variants (particularly, the classic book is by Knuth, 1981). The intent of this section is not to review them, but to briefly describe the most popular family of RNG's.

Most RNG's are *linear congruential generators*, which generate a sequence of integers, I_1, I_2, I_3, \dots , each between 0 and $m - 1$ (a large number) by the recurrence relation (Press et al. 1986; Dagpunar, 1988)

$$I_{j+1} = \text{mod}(aI_j + c, m), \quad (4.21)$$

⁴The code has been modified, and improved significantly by Lucy Caruthers and her colleagues at the Computer center, University of Arizona.

where m is called the *modulus* (*ibid*), and a and c are positive integers, called the *multiplier* and the *increment* respectively. Usually, rather than an integer, I_{j+1} , a real number between 0 and 1 is returned as I_{j+1}/m (which is strictly less than 1). The recurrence (4.21) will eventually repeat itself with a period that is obviously not greater than m . If m , a and c are properly chosen, then the period will be of maximal length (i.e., of length m). I_0 is the *seed*, which, theoretically, is insignificant, because all integers between 0 and $m - 1$ occur at some point anyway. [In practice, however, the choice of seed affects the generated fields, and some RNG's include "seed generators" which ensure the randomness of very long sequences (e.g., the RNG by Maclaren, we used)]. Then I_1 becomes the "seed", and so on. Eventually, one hopes to have a uniform distribution between 0 and 1, of *uncorrelated* "random" numbers. This can be subsequently transformed into other distributions, like the Gaussian distribution, using one of several different techniques (e.g., Press et al. 1986, and IMSL).

The linear congruential method has the advantage of being very fast, requiring only a few operations per call, hence its almost universal use (Press et al. 1986). It has the disadvantage that it is not free of sequential correlation on successive calls (*ibid*), particularly if the constants a , m , and c are not very carefully chosen. In fact, the only difference between most famous RNG's, including those used by the IMSL and NAG libraries, and between "good" RNG's and "bad" RNG's is in the values of these constants (see, for example, Sharp and Bays, 1992). Another important factor that affects the performance of RNG's, particularly when generating very long sequences, is the "accuracy" of the computer (i.e., the number of *bits* available); at least a 32-bit capability is recommended (e.g., by Sharp and Bays, 1992) for generating very long

sequences (required for multiple large random fields), and some RNG's would not "work" with less than 64 bit. Our experiments (described below) were done mostly using the full 64-bit option on both the Convex C-240, and the RISC-6000 machines). In order to break up sequential correlation, a *shuffling* procedure is often used (Press et al. 1986, p. 194; Dagpunar, 1988, p. 35). This procedure is reminiscent of another common technique for generating uniform distribution of pseudo random numbers, the *generalized feedback shift register* method, recommended by Johnson (1987) and Maclaren (1988), and described briefly by Dagpunar (1988).

There are several stringent empirical tests which check uniformity, correlation, local randomness, lattice behavior (particularly 2- and 3-dimensional serial patterns), etc. However, none of them is completely exhaustive. One of the most effective and visual tests is the serial test, in which sequential pairs of random numbers are used as coordinates for plotting individual points (i.e., each two random numbers constitute one point in an $X - Y$ square plot). The same can be done in "3-D" (where each three numbers define one point in a cube). If there are correlations, they would appear as patterns in the plots. Figures 4.2 and 4.3 (from Sharp and Bays, 1992) show how different combinations of m , a , and c (appear as M , A , and C in these figures) in (4.21) affect the (undesirable) correlation structure. RNG's which exhibit good (uncorrelated) randomness on 2-D plots, may fail dramatically in 3-D plots, or when the number of triplets increases by a factor of two or more; a more dramatic change appears when rather than choosing sequential numbers continuously, one picks two (or three) consecutive numbers, skips n consecutive numbers (e.g., $n = 5$), picks the next pair (or 3 numbers), skips another n numbers, and so on (e.g., Dagpunar, 1988).

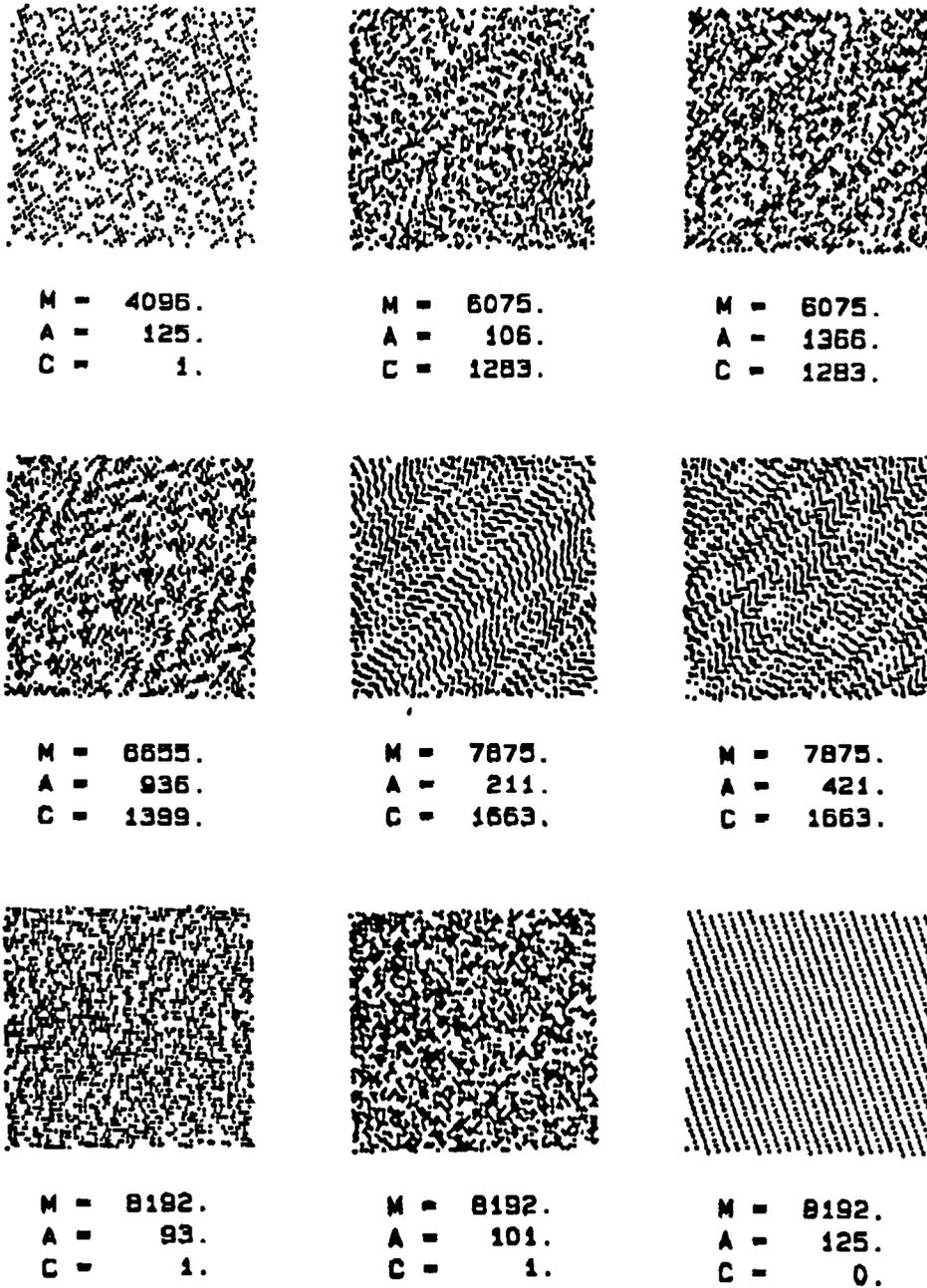
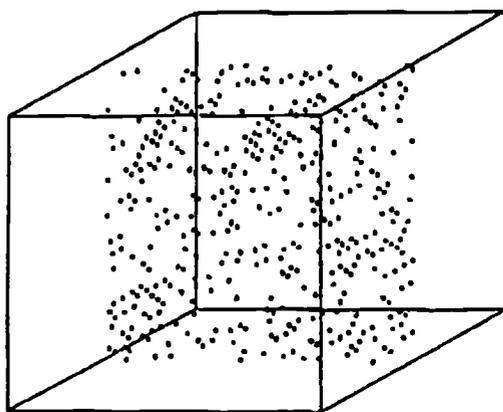
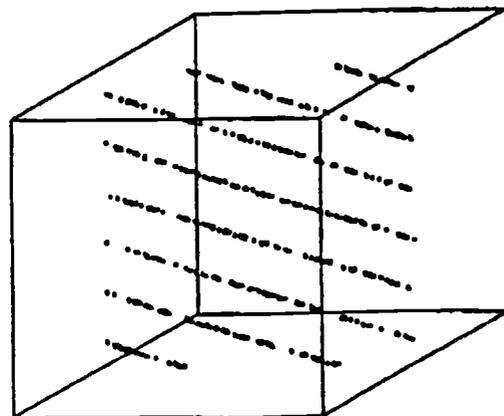


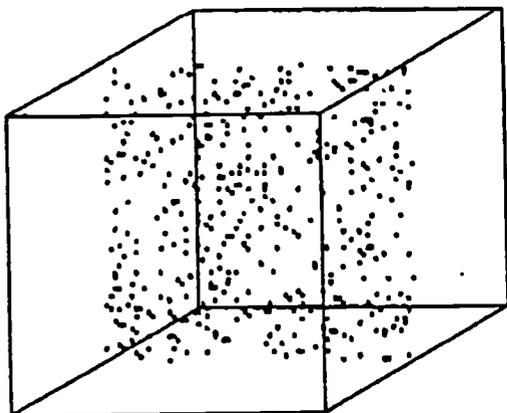
Figure 4.2: Serial plots in 2-D of 2000 nonoverlapping pairs of random numbers (After Sharp and Bays, 1992).



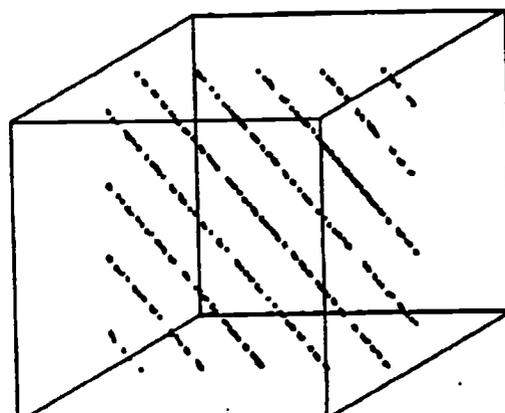
M = 30323.
A = 170.
C = 0.



M = 30259.
A = 171.
C = 0.



M = 65537.
A = 75.
C = 0.



M = 65537.
A = 254.
C = 0.

Figure 4.3: Serial plots in 3-D of 10000 nonoverlapping triplets of random numbers (After Sharp and Bays, 1992).

Obviously empirical testing of the whole cycle is not practically possible (Dagpunar, 1988); since for our Monte Carlo Simulations we need 2–4 million numbers, no test can practically assure a good performance of the RNG used. Moreover, passing all tests does not guarantee that an RNG is “good”, even for that particular (tested) portion of the cycle. The easy (and perhaps the best) way to choose an RNG is by its “reputation”, making sure it had passed all possible tests, and... remaining very very suspicious (Press et al., 1986) and unloyal (Johnson, 1987). [Indeed, we remained suspicious while using one of the portable (i.e., machine independent) random number generators, RAN2 (Press et al., 1986); after experiencing too many outliers we suspected the performance of the original RAN2, and replaced it by the new improved version (again, by Press et al., 1992), and yet, with very little loyalty, we have checked it against some other reputable RNG’s. We found later from Maclaren (personal communication) that the old version of RAN2 had failed to pass some crucial (new) tests.]

4.6 Comparison and Conclusions

Comparison between random field generators is a difficult task, and to date, there is no quick and rigorous way to define a “good” RFG. Theoretically, one needs to generate an infinite number of realizations, and verify the joint probability distribution for all points in space. A less rigorous approach would generate at least one thousand realizations and verify the (marginal) distributions at each point and the correlations between each two points in the field. Even then, the amount of work involved would render it impractical, especially when, at the same time, a comparison between several RNG’s is done in order to choose the one which will “best” fit the

particular RFG. As is shown in the sequel, the quality of the generated fields (based on the first two moments), depends on the *combined* operation of the RFG and the RNG.

[The following analysis is a product of a collaboration with Thomas Harter].

We used the five most promising and/or popular random number generators, which had passed at least a few severe tests successfully, to test the four different random field generators (RFG's) mentioned above. The initial motivation for these tests was to find how many realizations are needed to approximate *ensemble* covariances between the central four points (in fields of size 64×64) and any other point in the field. (This was essential for the exploration of the "weak approximation", discussed in Chapter 3). We found that the number of realizations (i.e., the sample size) needed to yield "ensemble" covariances⁵ to that reasonably resemble the desired (theoretical) ones, depends significantly on the choice of RNG, and varies notably from one RFG to another, exposing their flaws. Encouraged by the (modest) generality of this test, we proceeded with additional, more general RFG tests, which are based on sampling the ensemble of realizations, rather than spatial fields. As is shown in the sequel, these tests could reveal underlying defects of a random field generator.

The chosen random number generators are:

- RAN2 (Press et al. 1992; a portable RNG).
- DPRAND (Maclaren, 1992; a portable RNG).
- IMSL based RNG (1991; machine dependent).

⁵In this chapter, "ensemble" means *sample* of the ensemble.

- NAG based RNG (1990; machine dependent).
- ESSL based RNG (1991; machine dependent).

where the latter three RNG's reside in the most renowned libraries. [The RNG from the NAG library was also developed by Maclaren.] In the following figures, the first letter in the above list denotes the particular RNG (e.g., “r” for RAN2, etc.).

The most common checks of a generated random field are its spatial (marginal) statistics. In the case of log-normal fields, it is common to check the spatial statistics of their logarithm, which is normally distributed. This includes

1. Visualization of the histogram, calculation of the first (four) moments, checks for normality, outliers, and other statistics, as done routinely by the main statistical libraries, like SAS and SPSSX (see Appendix H).
2. Visualization of the spatial covariances of some sample realizations, and the average spatial covariance over many realizations.

The behavior of a spatial covariance of a particular realization depends, to a large extent, on the size of the field (because as the lag distance becomes larger, there are less sample points, and the corresponding calculated covariance function becomes “random” itself). This becomes particularly disturbing as the relative size of the correlation scale increases. By the same token, as is shown below, an “average” spatial covariance may be misleading and overly smooth. Unless stated differently, the following figures are based on square (64×64) log-normal random fields [$K = \exp(Y)$] for which the corresponding normal fields have $\lambda_Y = 5$, $\sigma_Y^2 = 1$, and $\langle Y \rangle = 0$ (the latter implies $K_g = \exp(\langle Y \rangle) = 1$).

Figure 4.4 illustrates spatial covariances of single realizations of 64×64 fields ($\lambda = 5$) generated by a sequential simulator (GCOSIM3D by Gómez-Hernández 1991; upper figure) and the turning bands method (TUBA1, based on Mantoglou and Wilson, 1982; lower figure) respectively. The calculation of the covariances in each direction is enhanced by using FFT (Gutjahr, 1989), and the correct way to view it is by “cutting” the plot along any symmetry line that passes through the mid point, and regard the three dimensional plot as a collection of directional (two-dimensional) plots of covariances (i.e., the plots are symmetric with respect to any straight line that passes through the mid-point). While the (particular) field generated by GCOSIM seems to be laterally anisotropic, the field generated by TUBA1 is both diagonally anisotropic and oscillatory (although, the oscillations are, in part, due to the limited size of the domain).

Figure 4.5 shows calculated and theoretical average spatial (isotropic) covariances and variograms in the X and Y directions of 50 realizations after filtering out outliers, which, in turn, were determined by extreme values of integral scales (calculated numerically by the trapezoidal rule). Although the average spatial functions seem to be sufficient for ensemble statistics calculations, as shown in the sequel, this is a misleading picture. In fact, we found later that the random number generator used for the generations of these fields (the old RAN2) suffers from some major flaws. Figure 4.6 shows a similar plot of averages over 100 (filtered) realizations, with GCOSIM and a much better RNG (DPRAND). Yet, the deviations from the theoretical covariance function in different directions are disturbing.

Figure 4.7 illustrates averages of (multidirectional) spatial covariances (like Figure 4.4) of 50 fields generated by TUBA1 (*tbm*) with 16 lines (or turning bands),

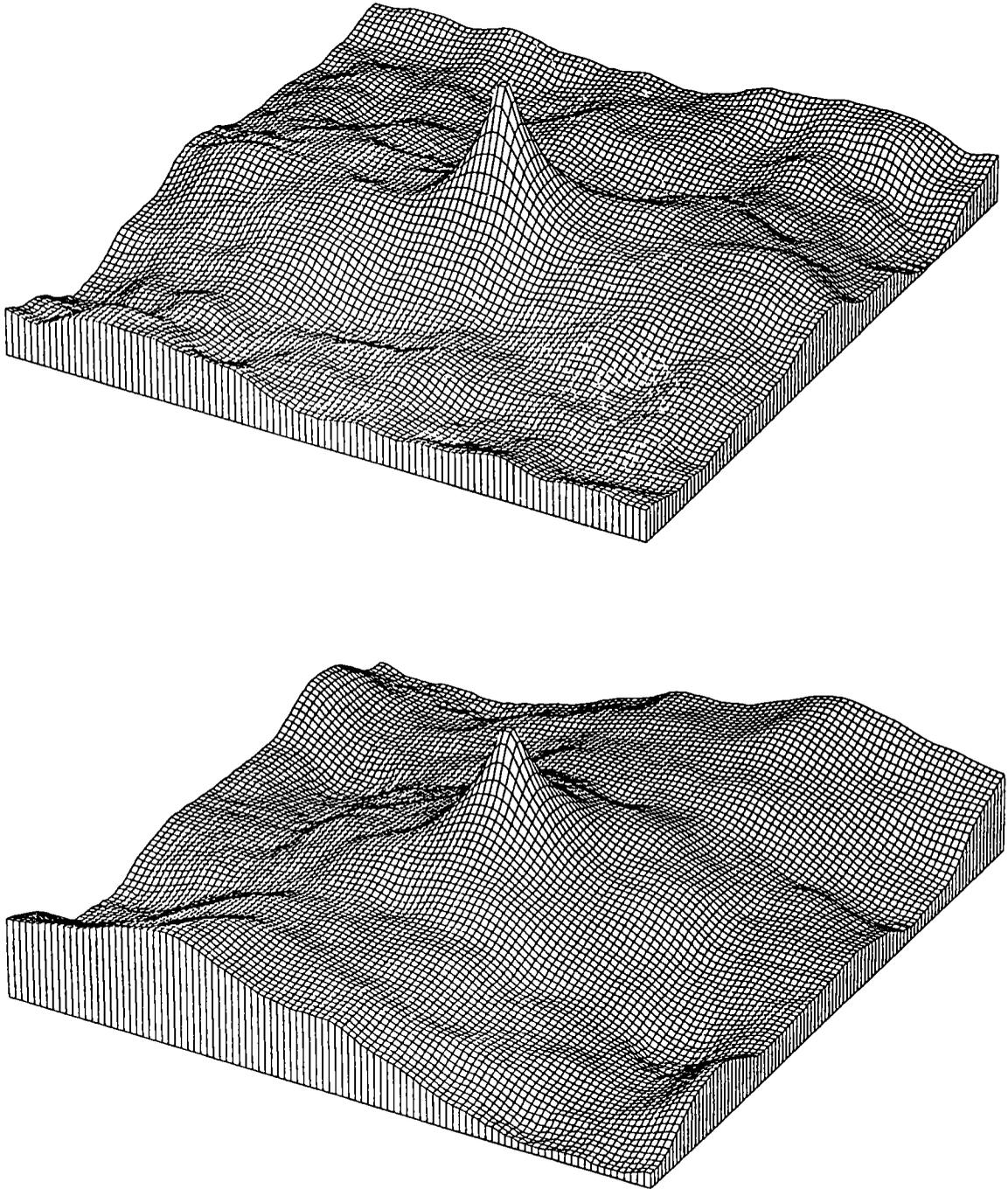


Figure 4.4: Spatial covariances of single realizations of 64×64 fields, with $\lambda = 5$, generated by GCOSIM3D (top) and TUBA1 (bottom).

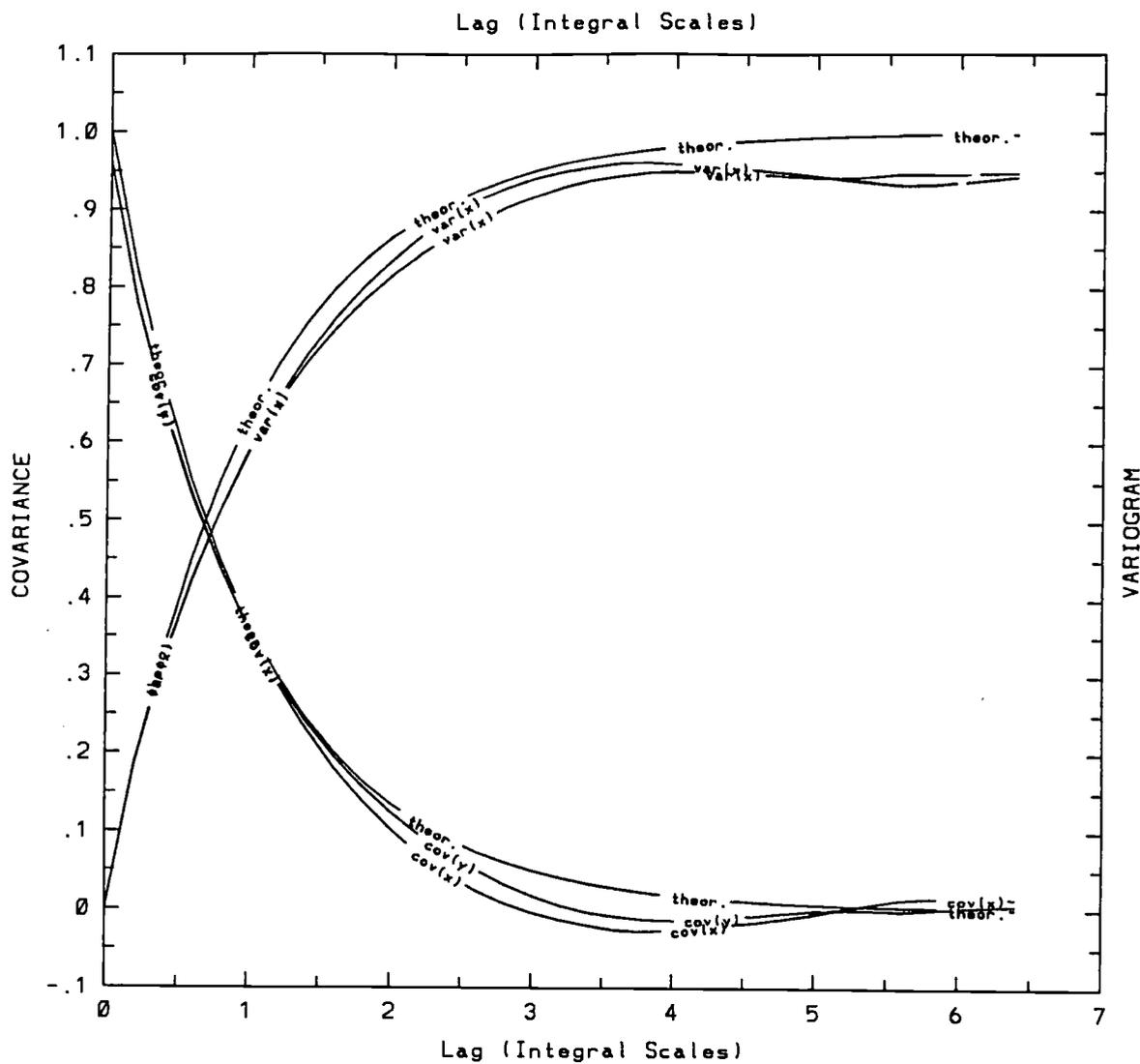


Figure 4.5: Calculated and theoretical average spatial (isotropic) covariances and variograms in the X and Y directions of 50 realizations after filtering out outliers.

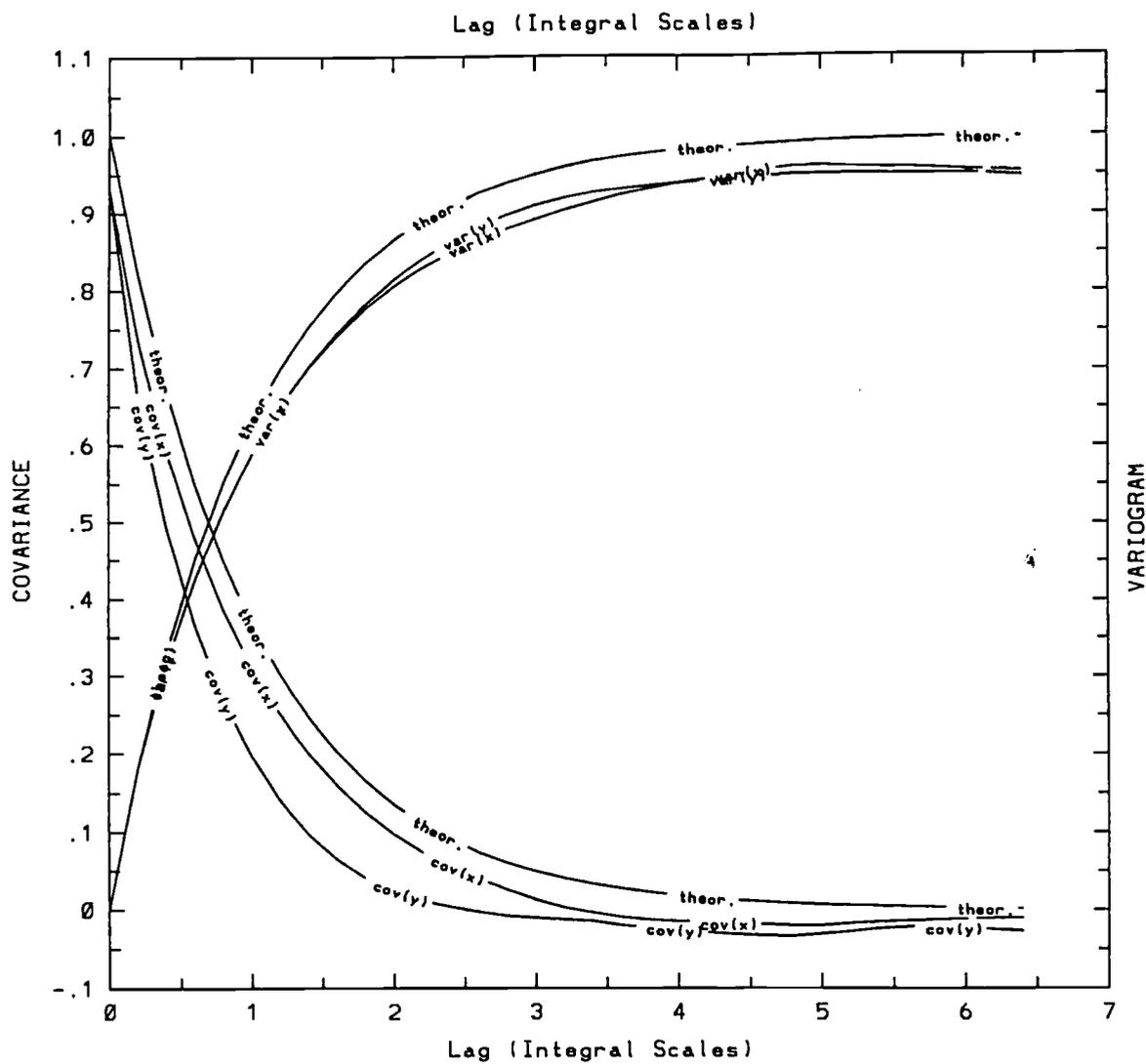


Figure 4.6: Calculated and theoretical average spatial (isotropic) covariances and variograms in the X and Y directions of 100 realizations after filtering out outliers.

as recommended by Wilson and Mantoglou (1982). The distinct pattern of (16) lines (or directions) is disturbing. In the subsequent field generation by the turning bands method we used at least 32 lines (see below).

Figure 4.8 shows average spatial covariances and variograms of 500 realizations (64×64) generated by the sequential simulator, GCOSIM, with larger correlation scale ($\lambda = 8$). The deviations from the theoretical covariance are more pronounced here (and, for some reason, not so much for the variogram). Figure 4.9 shows multi-directional “3-D” and “2-D” plots of the average covariances of the same fields. The (spatial) covariances in these figures seem to be closer to the theoretical one.

Next we investigate *ensemble statistics*, particularly by “ensemble” covariances. Beside being a direct check of input realizations to MCS, these calculations do not suffer from undersampling at large lags (low frequencies). Figure 4.10 shows averages of “ensemble” covariances in the x and y directions, as well as “minimum” and “maximum” “ensemble” covariances (in any direction), over 1000, 500, and 200 realizations generated with the FFT method with an IMSL based RNG. In this plot, the covariance functions were calculated for each pair of points and for each lag and in each direction, and then averaged to yield the “mean covariances in the x and y directions”; the extreme values of these covariances for each lag distance are plotted as points; alternatively, all the calculated two-point covariances, for all lags, could be plotted as a cloud of points distributed between those extremes. According to this figure, the average covariance is almost identical to the theoretical one, even for a sample the of (only) 200 realizations. It is clear (and expected) that as the number of realizations decreases, the dispersion of the sample covariances increases (as expressed by the gap between the minimum and the maximum calculated covari-

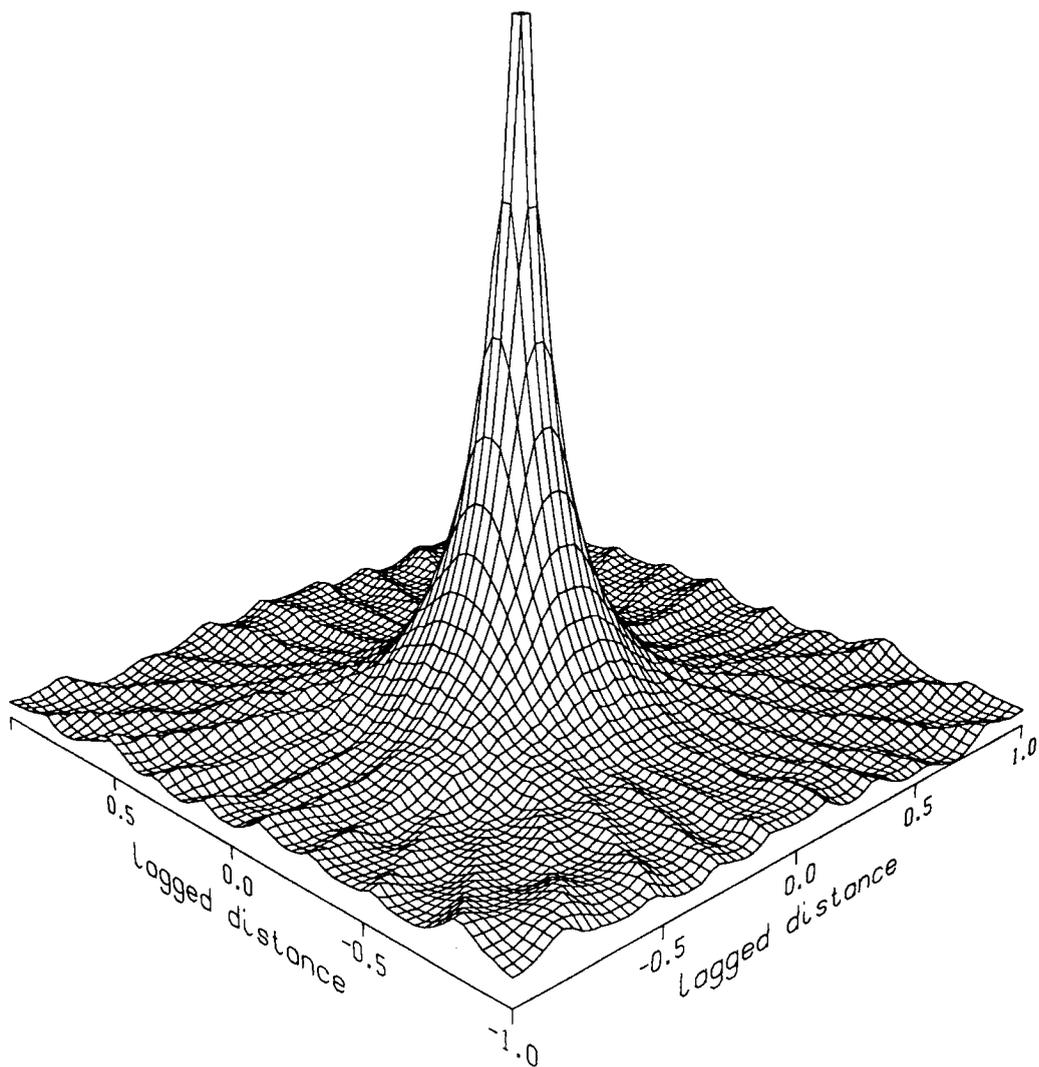


Figure 4.7: Averages of (multidirectional) spatial covariances of 50 fields generated by the turning bands method (TUBA1) with 16 lines.

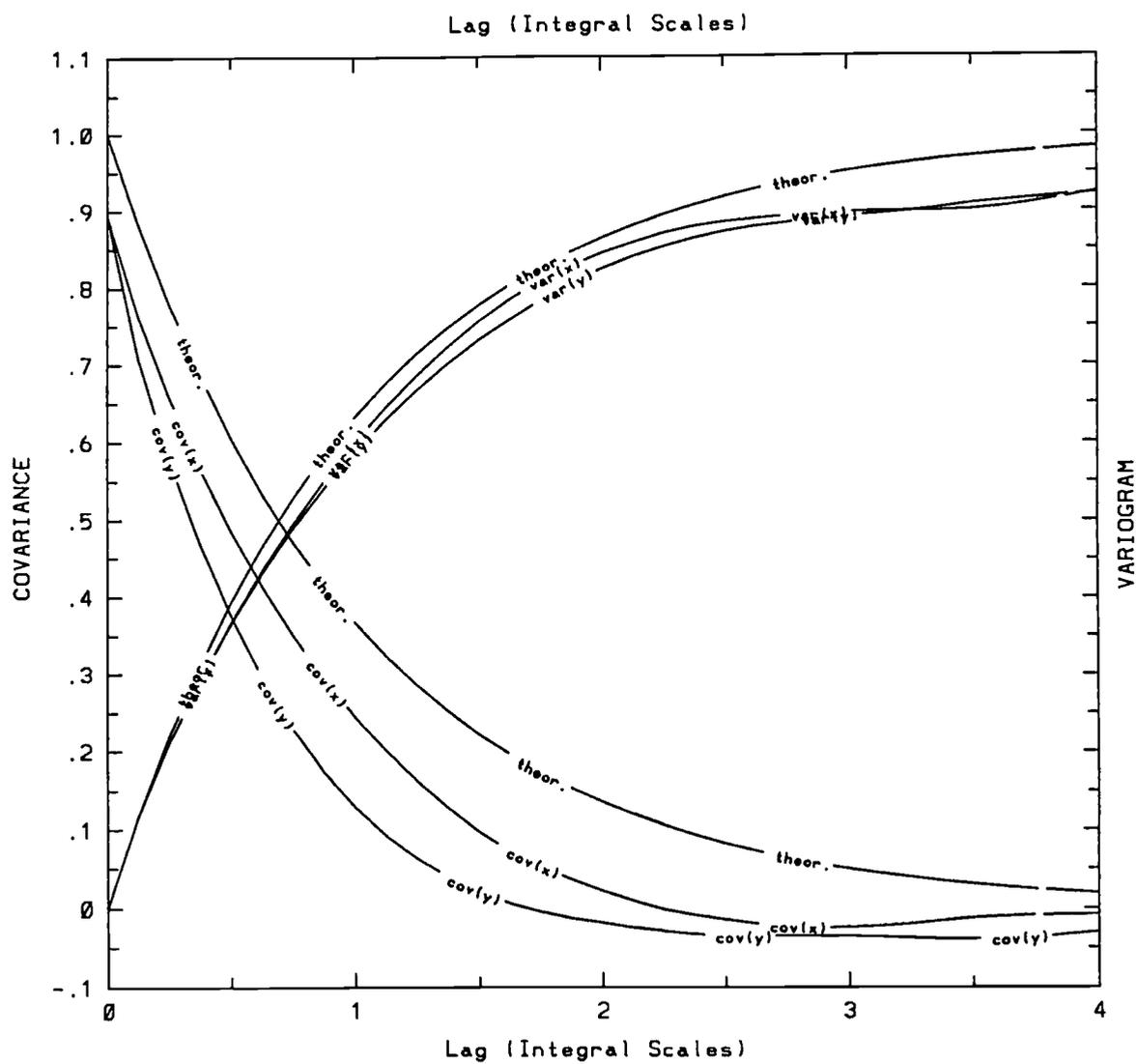


Figure 4.8: Average spatial covariances and variograms of 500 realizations (64×64) generated by GCOSIM, with $\lambda = 8$.

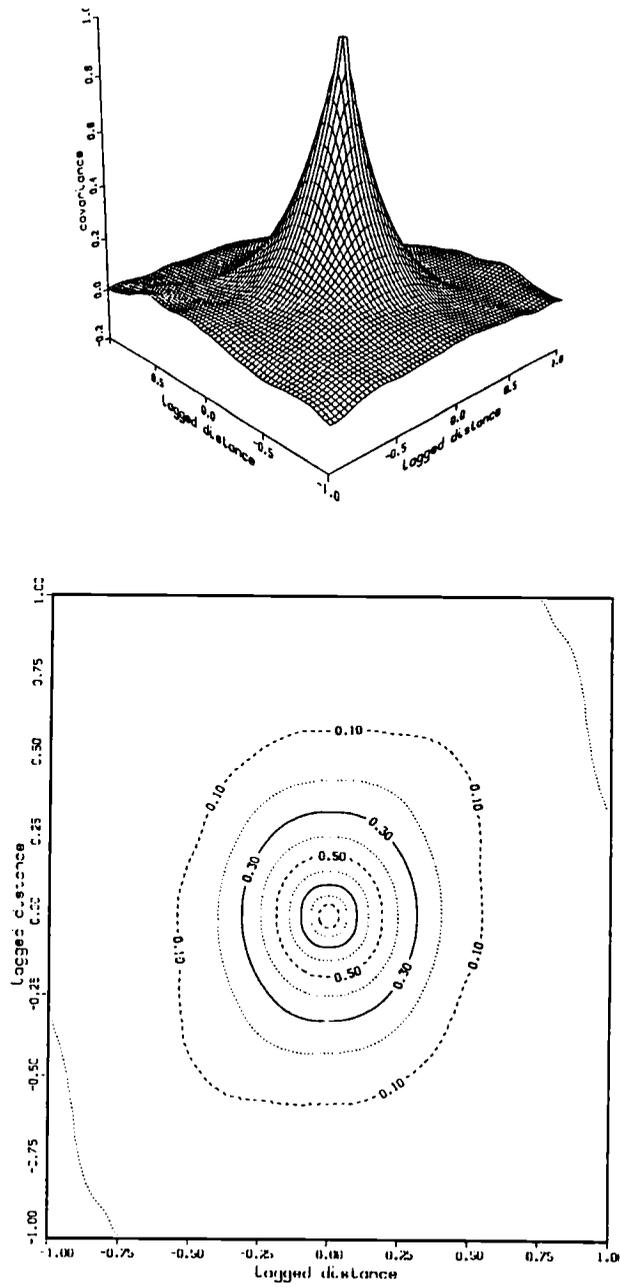


Figure 4.9: Average (multidirectional) spatial covariances and variograms of 500 realizations (64×64) generated by GCOSIM, with $\lambda = 8$.

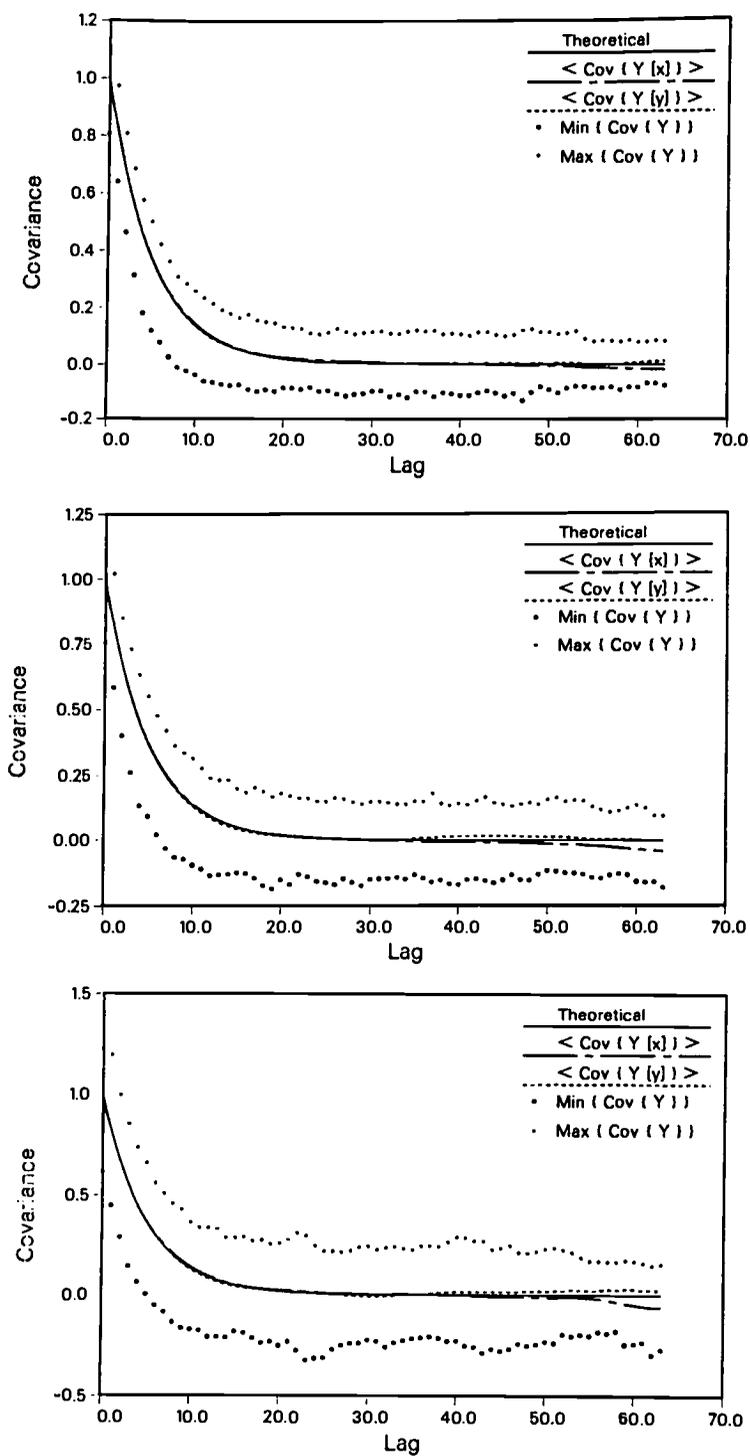


Figure 4.10: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the FFT method and an IMSL based RNG.

ances); this implies a fourth moment (a variance of the covariance), which ideally should approach zero. Note the deviations at large lags, that look like an aliasing artifact (mentioned above). Figure 4.11 shows similar sample (ensemble) covariances of realizations generated by the same RFG (based on the FFT method), but with a different RNG, this time from the NAG library. The overall behavior is as before, except that deviations at large lags do not show; this implies that what seemed to be aliasing in the previous case may be simply an artifact of undersampling (obviously, as the lag increases, there are less covariances to average).

Figure 4.12 shows similar (average) sample (ensemble) covariances, of fields generated by the *tbm* method (TUBA1), using 32 lines and the portable random number generator DPRAND. The figure exhibit a similar mean behavior and similar range of variability as before, including the deviations at large lag distances. Figure 4.13 shows mean sample (ensemble) covariances of fields generated by the turning bands method (*tbm*) and with the same RNG, except that this time the fields are generated with 320 lines (i.e., by 10 fold more lines). The covariances do not indicate a significant improvement over the previous ones. Figure 4.14 illustrates similar plots, now of fields generated by the sequential simulator GCOSIM3D and (the portable) RAN2. The slight deviations from the theoretical covariance seem to be insignificant. The same is true in Figure 4.15, where the RNG was replaced by DPRAND; the ideal behavior of the mean covariance at large lags implies some superiority of the sequential simulator over spectral methods (like FFT and TUBA1). Figure 4.16 shows the same type of sample (ensemble) covariance plots, now with another *tbm*. The deviations from the theoretical covariance clearly reveals a defect. Based on it,

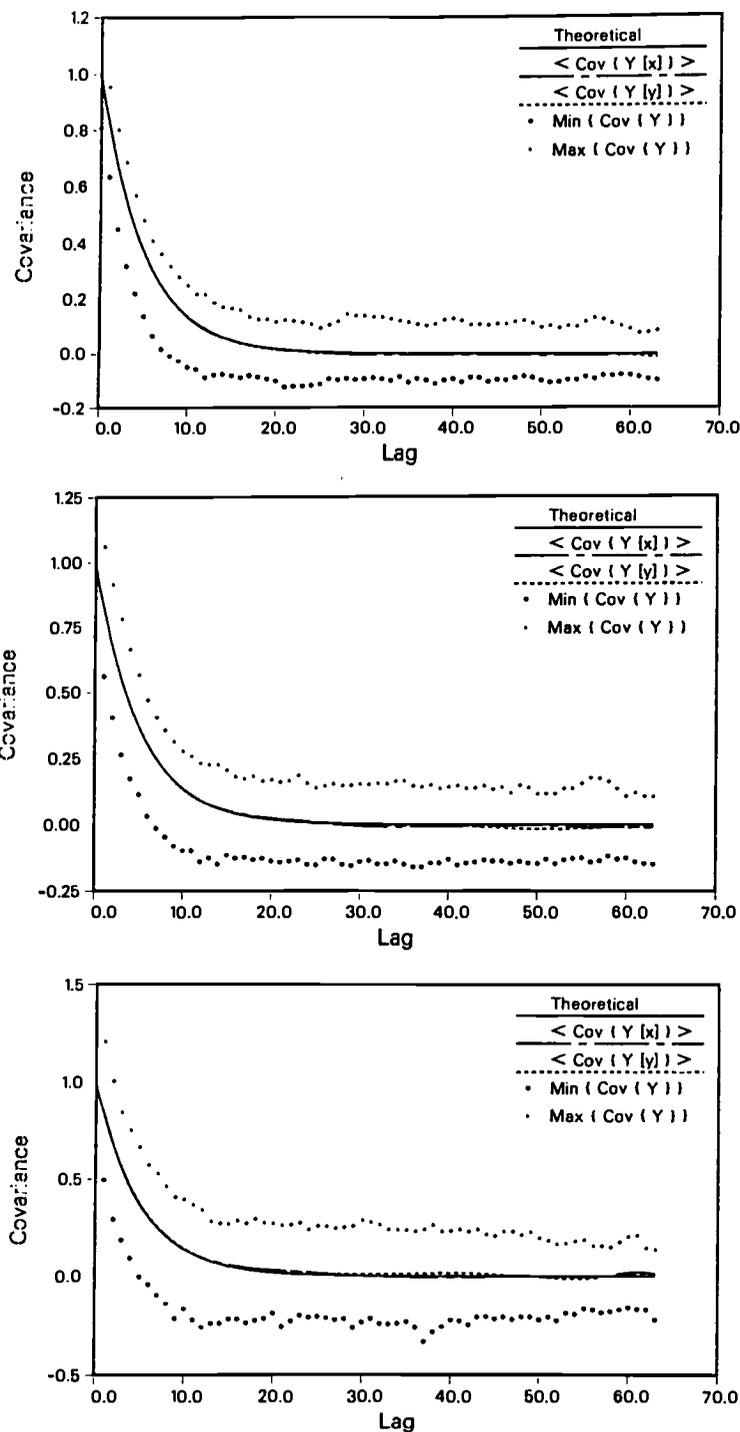


Figure 4.11: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the FFT method and an RNG from the NAG library.

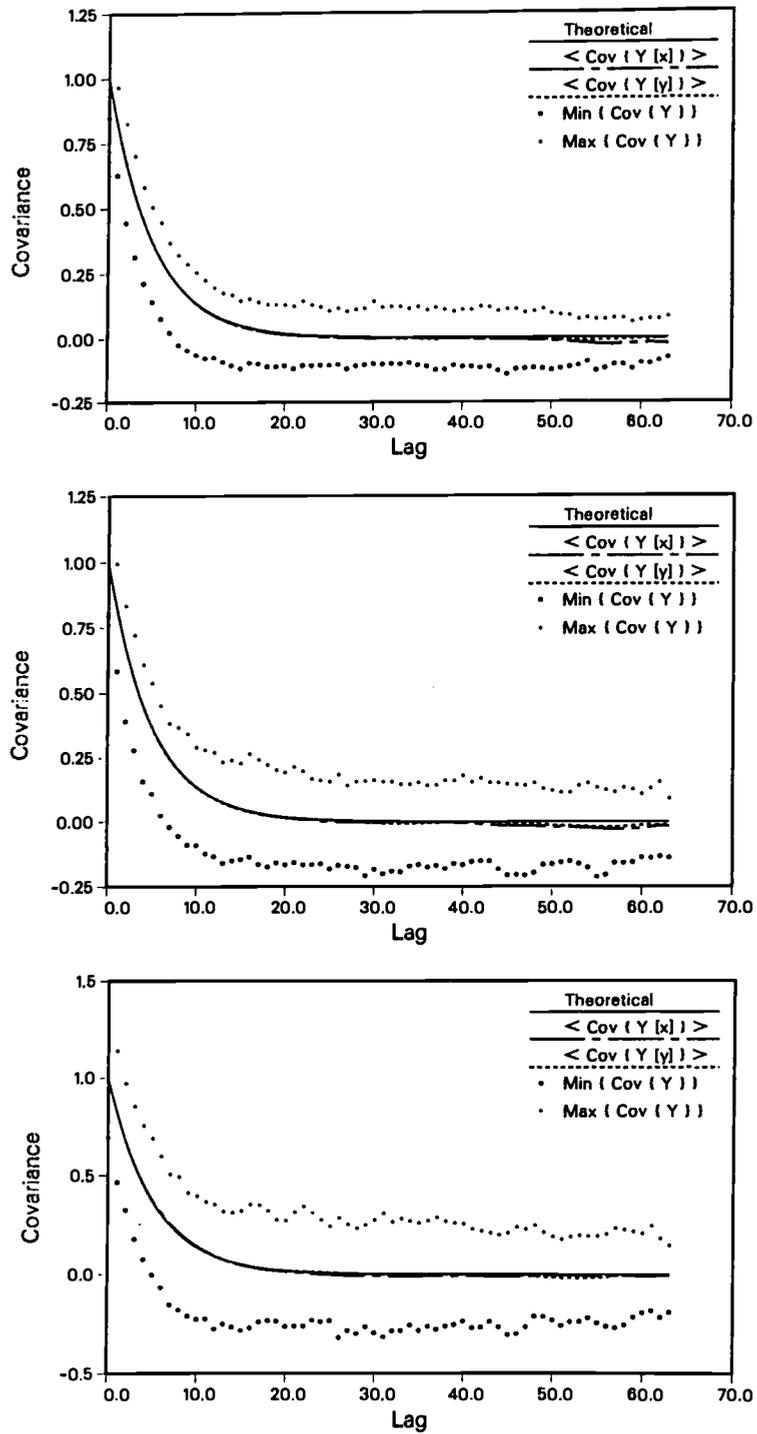


Figure 4.12: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the turning bands method with 32 lines and the portable RNG DPRAND.

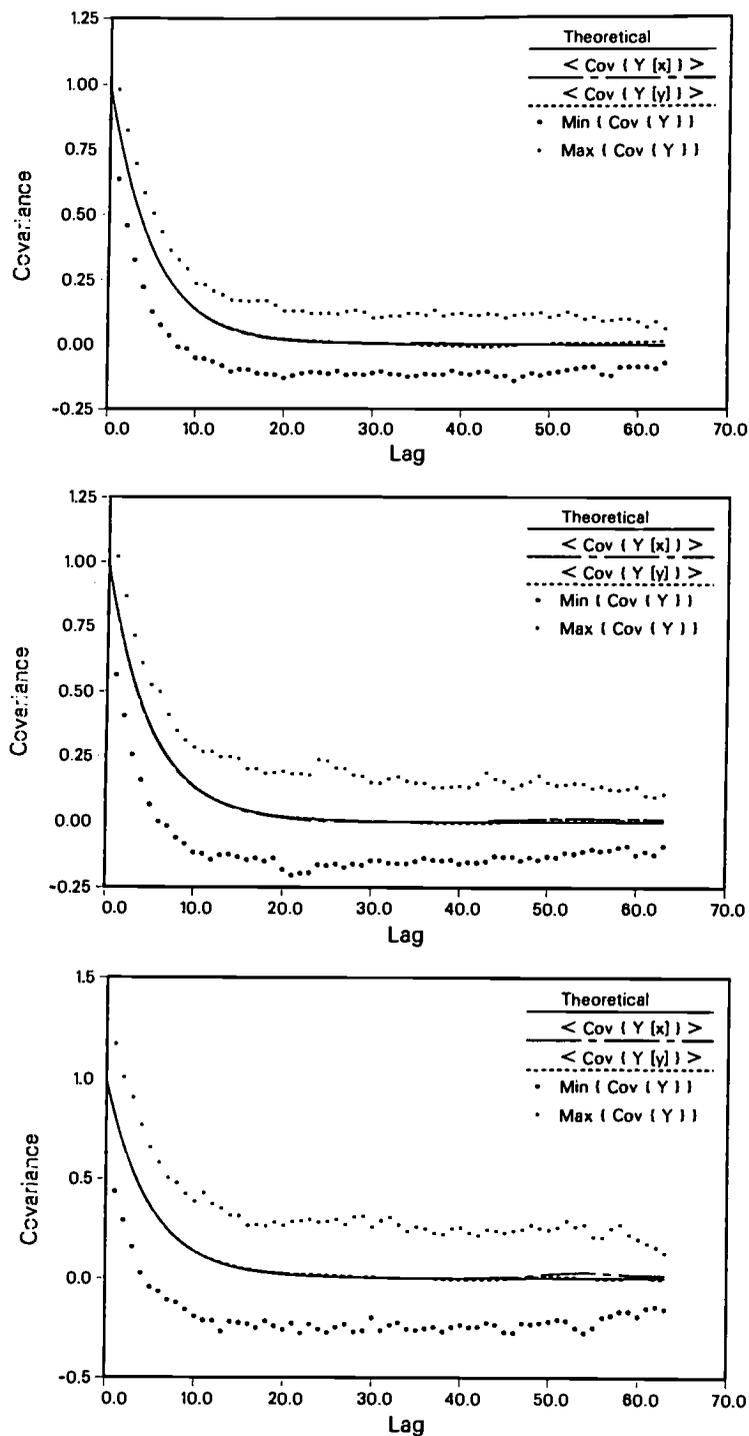


Figure 4.13: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with the turning bands method with 320 lines and the portable RNG DPRAND.

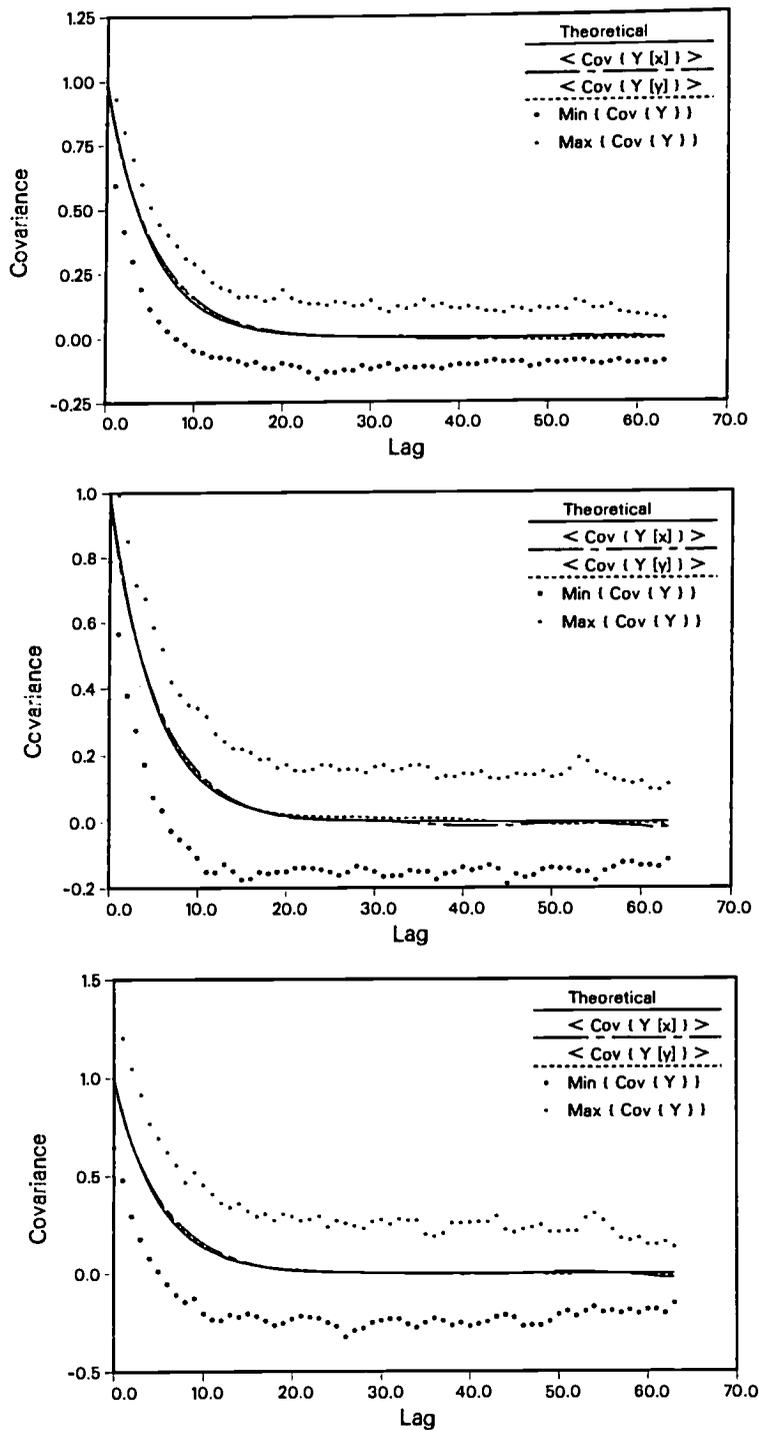


Figure 4.14: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with sequential simulator GCOSIM3D and (the portable) RAN2.

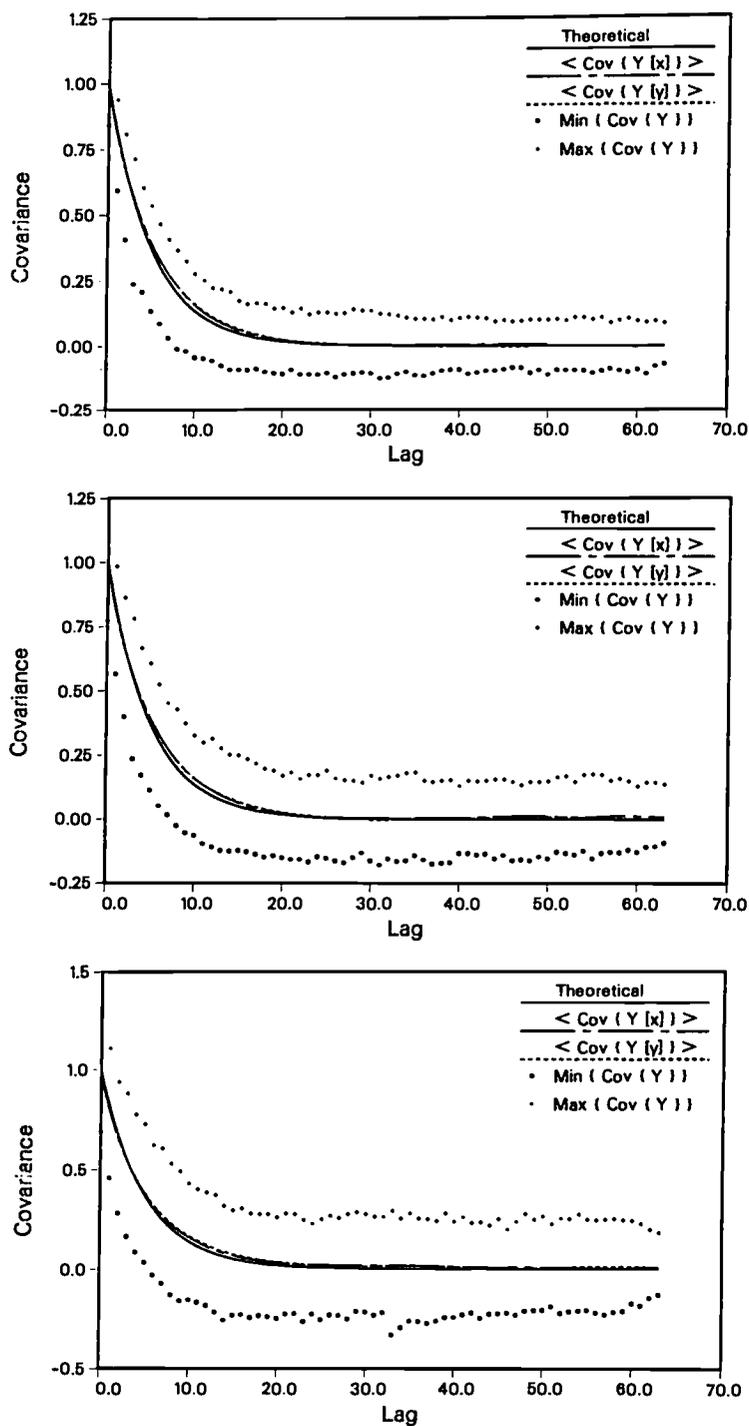


Figure 4.15: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 (top), 500 (middle), and 200 (bottom) realizations generated with sequential simulator GCOSIM3D and (the portable) DPRAND.

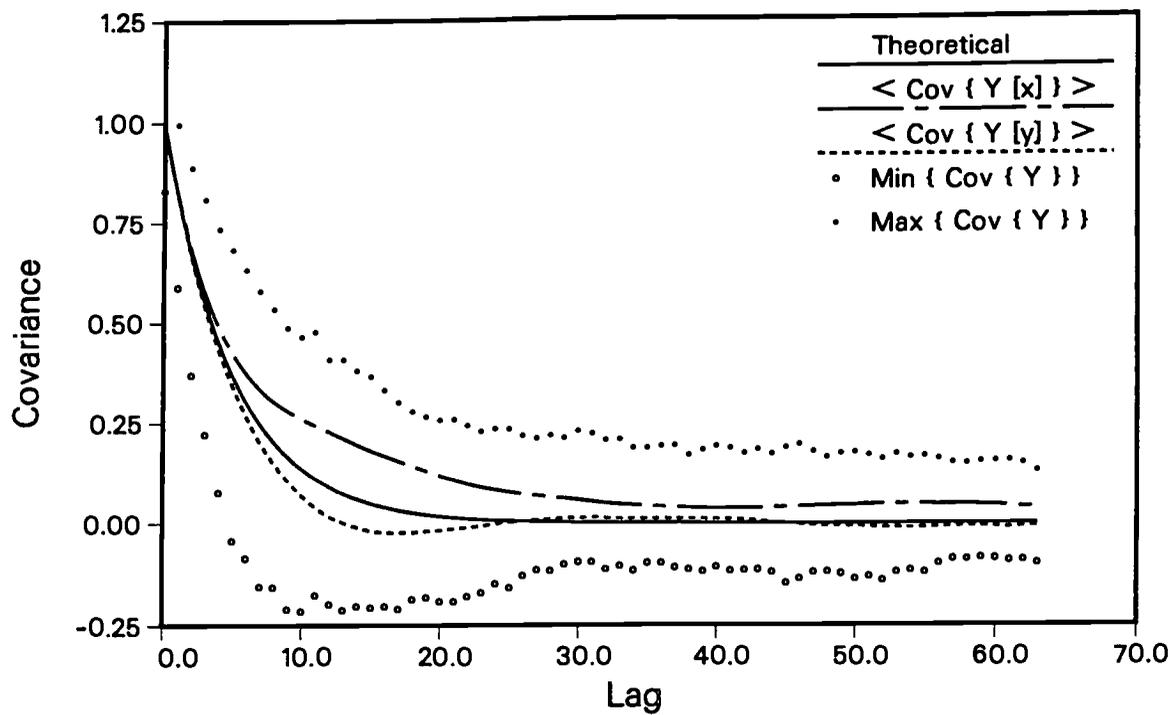


Figure 4.16: Averages of “ensemble” covariances in the x and y directions, and minimum and maximum “ensemble” covariances (in any direction), over 1000 realizations generated with a defected turning bands method.

the programmer⁶ identified a bug, and fixed it. This demonstrates that such plots constitute a test for RFG's (although they provide only "necessary", not "sufficient" conditions for the acceptance of the generated fields).

Another way to inspect the ensemble behavior of the generated fields is to plot sample (ensemble) covariances between the mid-points and all other points in the field, as described above, now on true 3-D plots, with covariances of the log-normal K fields instead of the (previous) normal Y fields. Note that according to (3.94) [or alternatively, (F.3)], for log-normal random fields with $K_g = 1$ and $\sigma_Y^2 = 1$, $\sigma_K^2 = K_g^2 \exp(\sigma_Y^2)[\exp(\sigma_Y^2) - 1] \approx 4.67$. Figure 4.17 depicts theoretical covariances between the mid (four) points and all the other points in the field. Figure 4.18 depicts sample (ensemble) covariances of 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM; Harter, personal communication), with an IMSL based RNG. In contrast with the previous set of plots, which emphasize averages of sample (ensemble) covariances of $\ln K$, the set of (multidirectional) single "ensemble" covariances of K is much more sensitive to the number of realizations; while the sample of 1000 realizations in Figure 4.18 seems reasonable, and the sample of 500 realizations is barely "acceptable", the "ensemble" covariances over 200 and 100 realizations are, in fact, unacceptable, with "peaks" of correlations between distant points, as high as the variance of K . This unfavorable behavior is somewhat less drastic when using the NAG-based RNG, with the same RFG (LUSIM), as seen in Figure 4.19. The NAG RNG is obviously advantageous in this case. An even more favorable behavior is exhibited when the same RFG is used with the RNG DPRAND⁷

⁶The programmer's name is kept confidential.

⁷Both the NAG-based RNG and DPRAND were developed by Nick Maclaren, University of Cambridge Computer Laboratory, Cambridge, England.

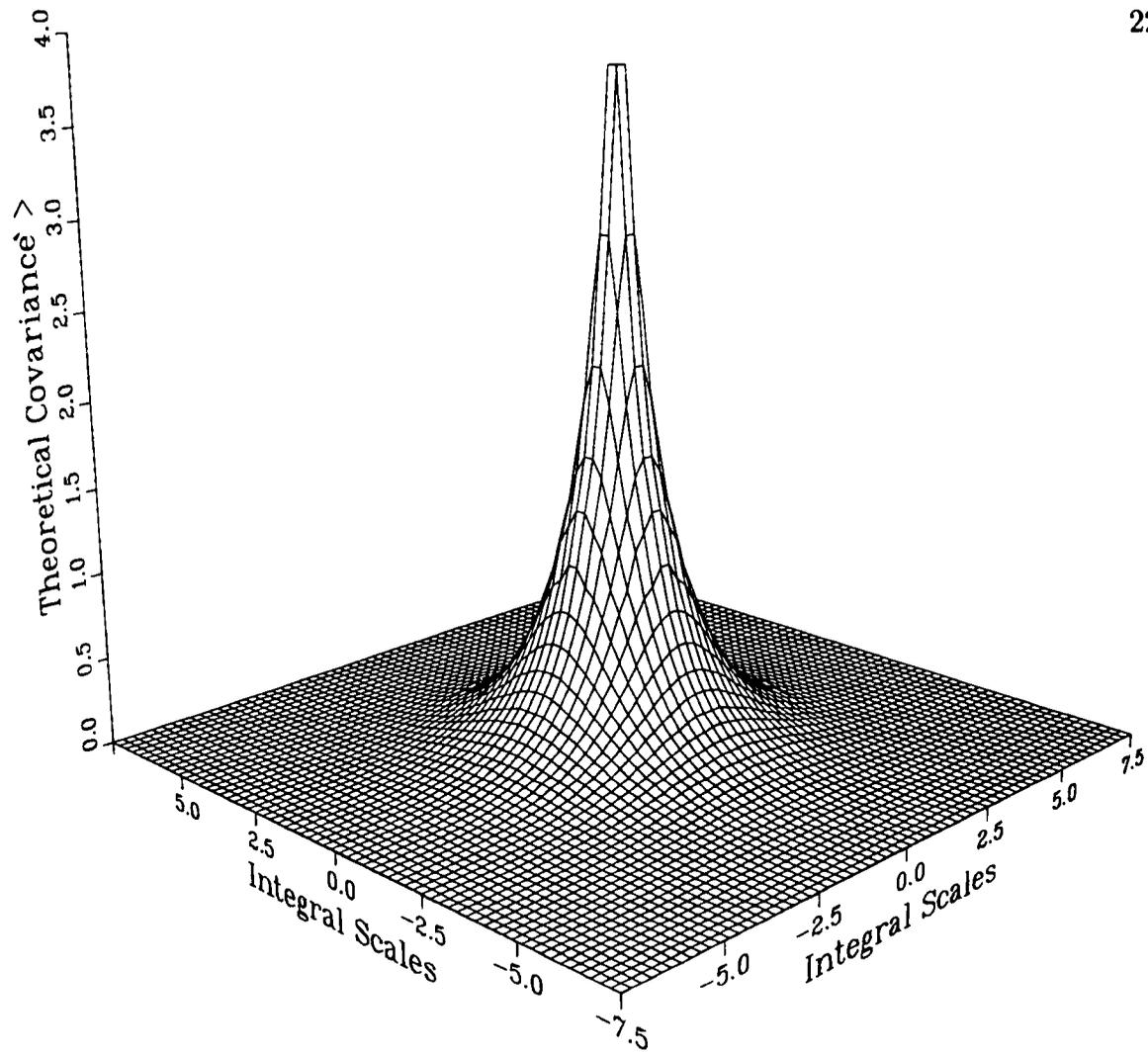


Figure 4.17: Theoretical covariances between the mid (four) points and all the other points in the field.

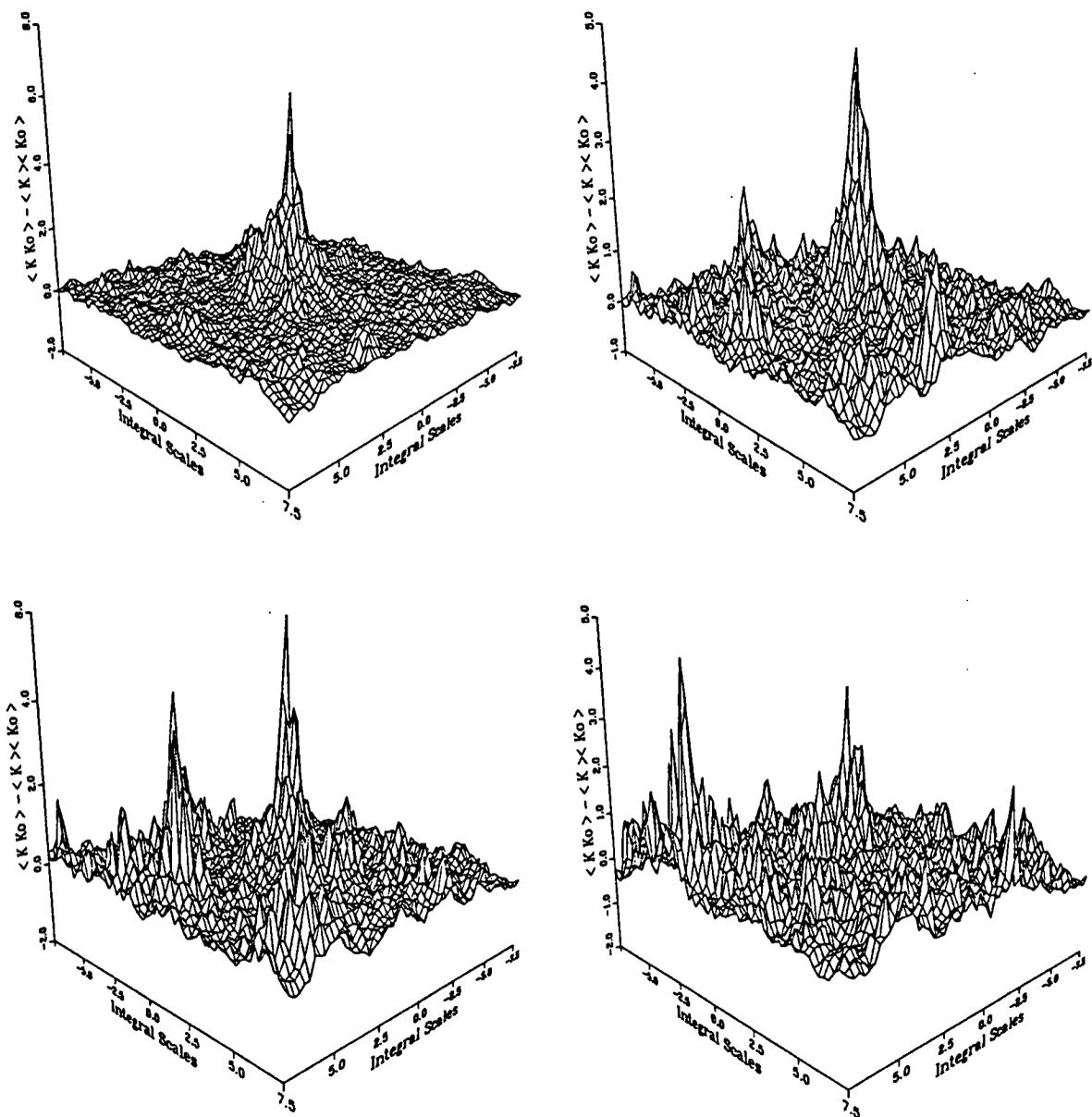


Figure 4.18: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM), with an IMSL based RNG.

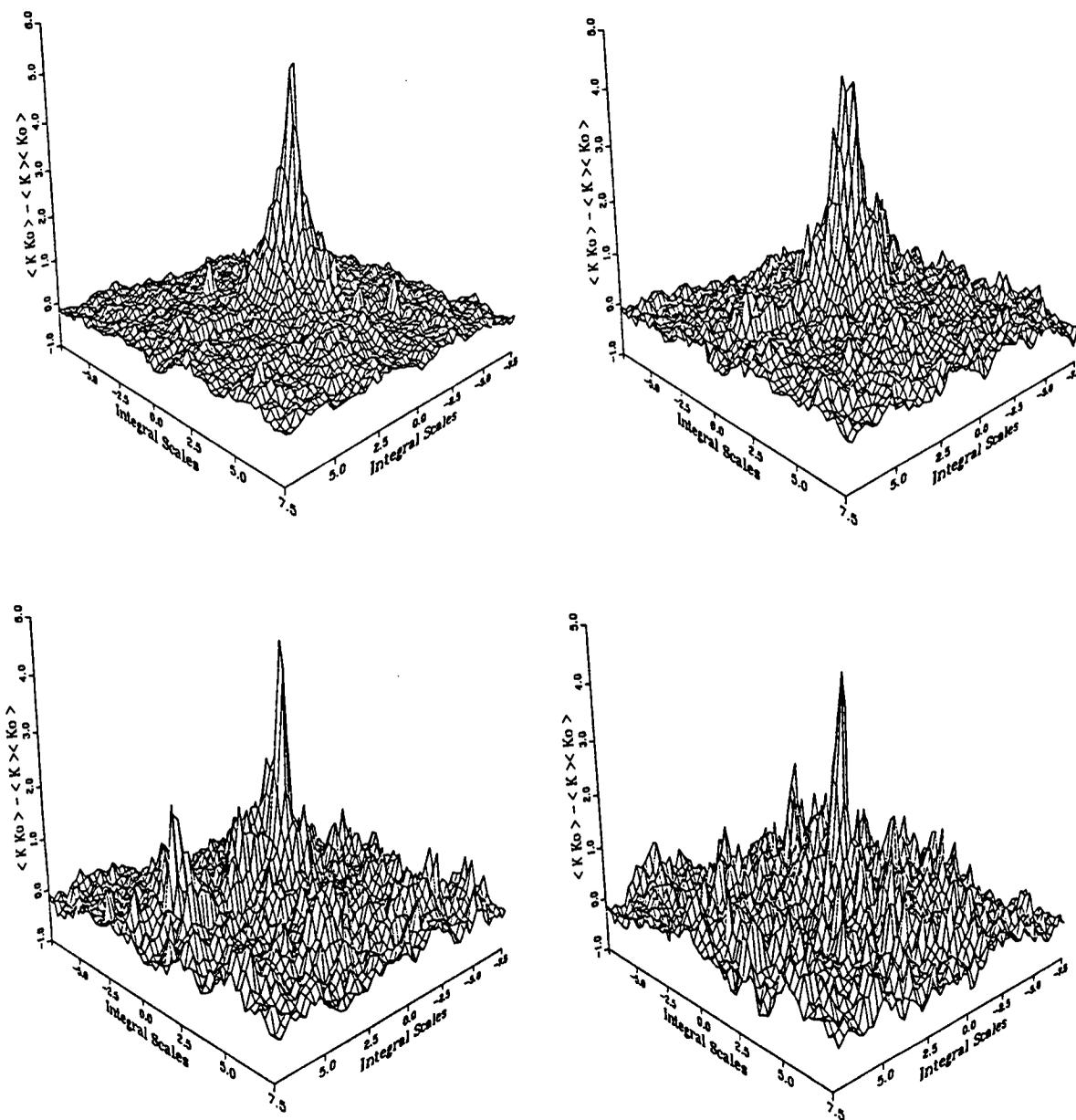


Figure 4.19: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM), with a NAG-based RNG.

as observed in Figure 4.20, below. Figure 4.21 depicts “ensemble” covariances of fields generated by the turning bands method (TUBA1, with 32 lines) for 1000, 500, 200, and 100 realizations, again with DPRAND as the random number generator. In this case, too, 200 realizations are still acceptable. Figure 4.22 shows “ensemble” covariances of fields generated by the *tbm*, (TUBA1) with 320 lines. Surprisingly, they look inferior to the (preceded) ones generated with 32 lines. Figure 4.23 shows the above number of realizations generated by the FFT method with the IMSL-based RNG. The “good” behavior of the covariances for 200 and 100 realizations is somewhat misleading because of the high variance in the mid-point (3-4 folds of the theoretical variance). Otherwise, for the 500 and 1000 realizations, the covariances are acceptable. Figure 4.24 depicts “ensemble” covariances of 1000, 500, 200 and 100 realizations, generated by the FFT method, with a NAG-based RNG. At 200 realizations, a peculiar and undesirable correlation emerges between the mid point and a few points beyond the correlation scale. At 100 realizations, this peak correlation outweighs the correlation between adjacent mid-points. This indicates that our “good impression” from the combination LUSIM and NAG-based RNG above, may not be justified, and that these type of tests are not general enough. Figures 4.25 show covariances of fields generated by the sequential simulator (GCOSIM) with DPRAND. The behavior is very good for 1000 and 500 realizations, acceptable for 200 realizations, but not for 100 realizations (again, the peculiar undesirable correlation between distant points). Figure 4.26 shows covariances of fields generated with the combination GCOSIM (sequential simulator) and RAN2. The covariances are acceptable for 1000, 500, 200 and 100 realizations. Consequently, this combination was selected as the “best” combination for our purpose (i.e., investigating the extent

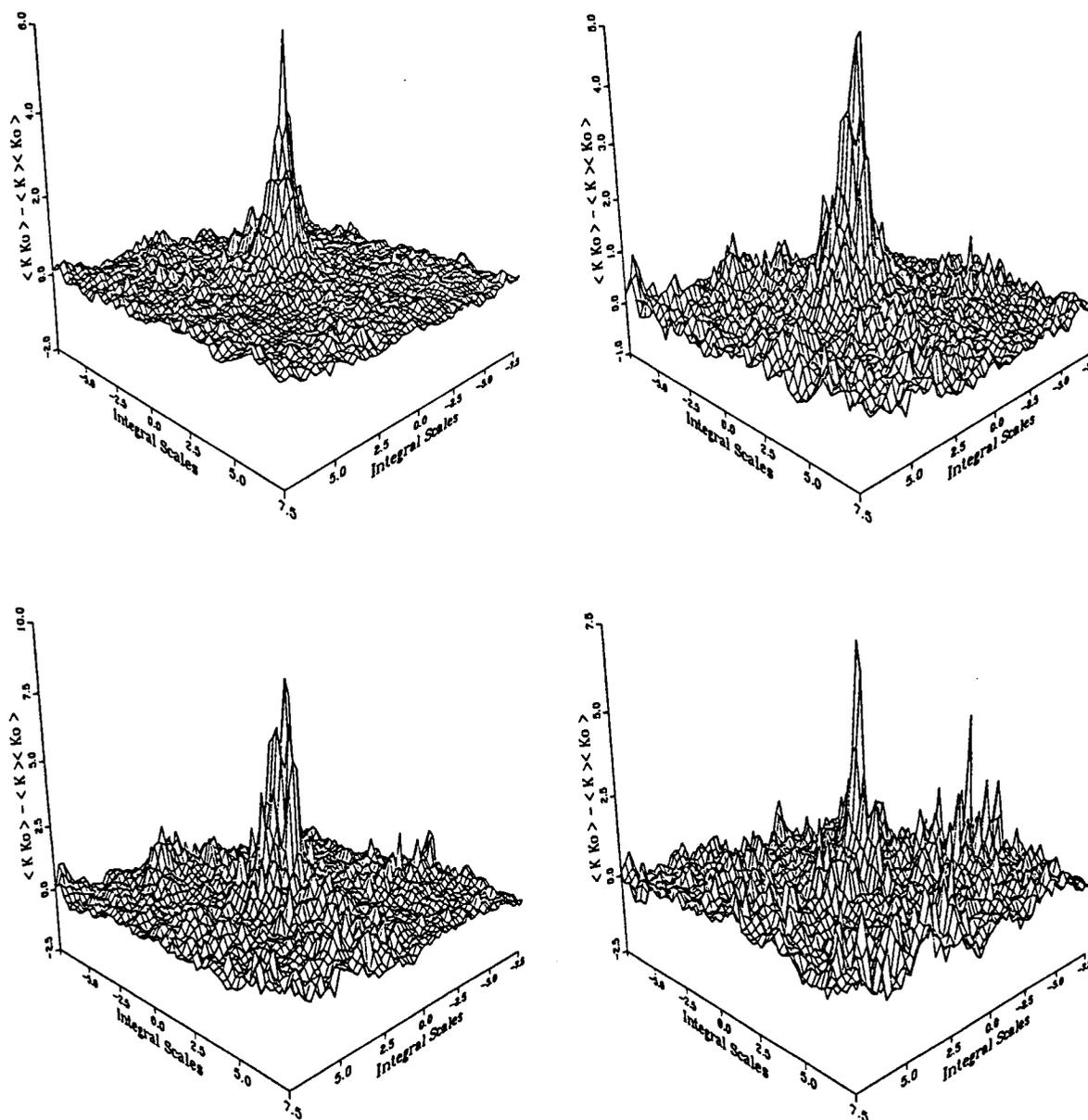


Figure 4.20: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by an LU decomposition based RFG (LUSIM), with the portable RNG DPRAND.

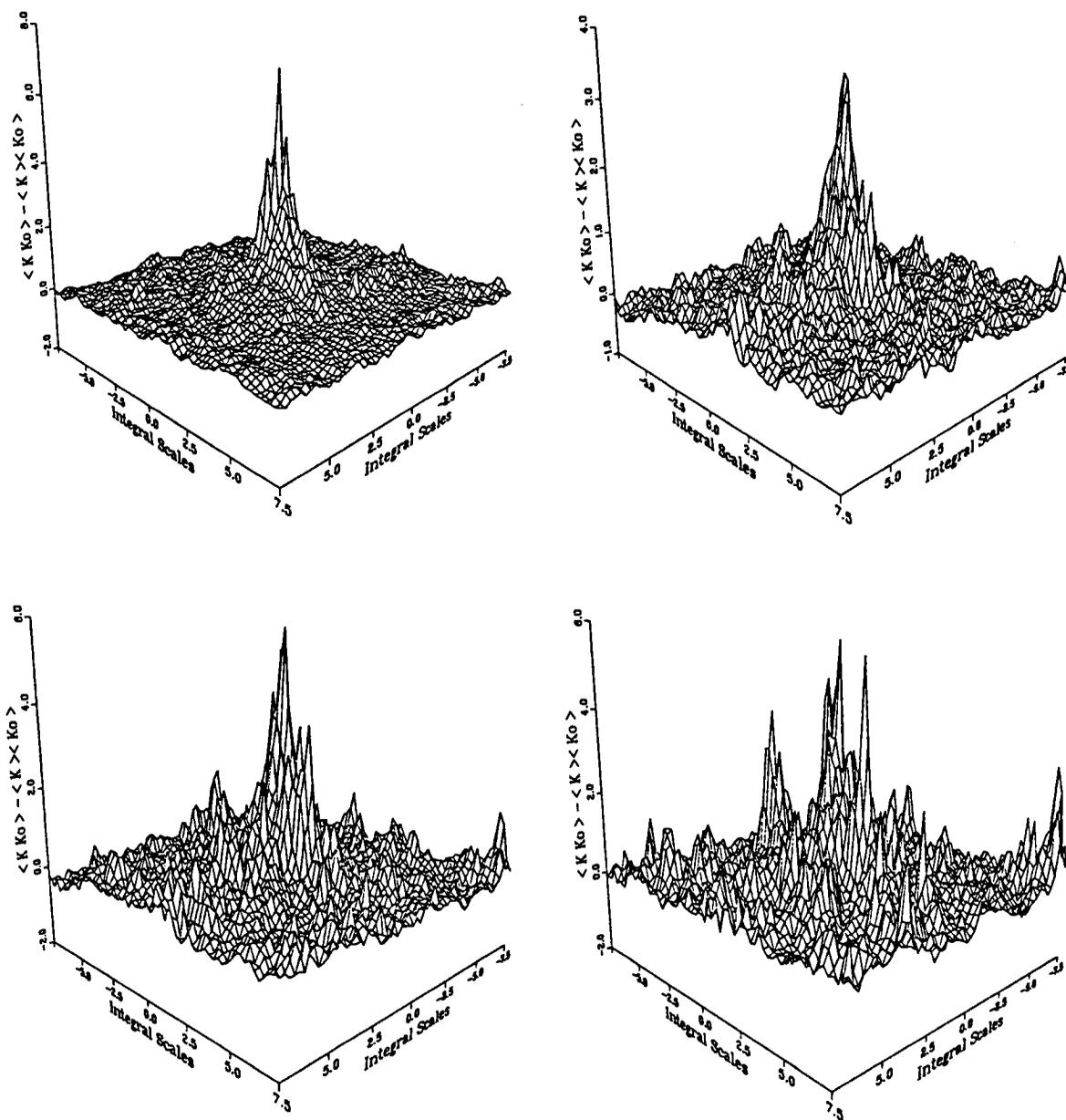


Figure 4.21: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the turning bands method (TUBA1) with 32 lines and the portable DPRAND.

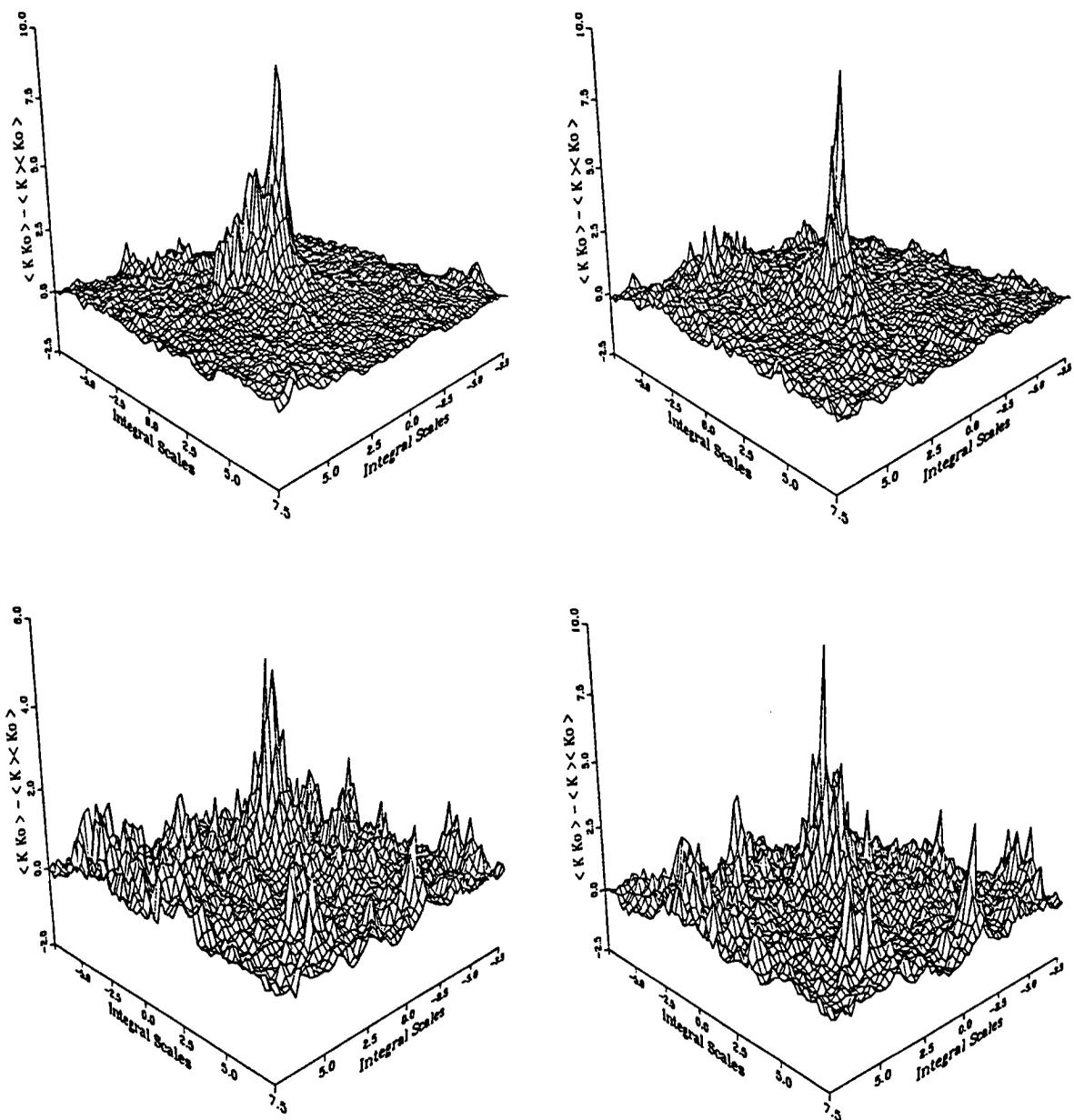


Figure 4.22: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the turning bands method (TUBA1) with 320 lines and the portable DPRAND.

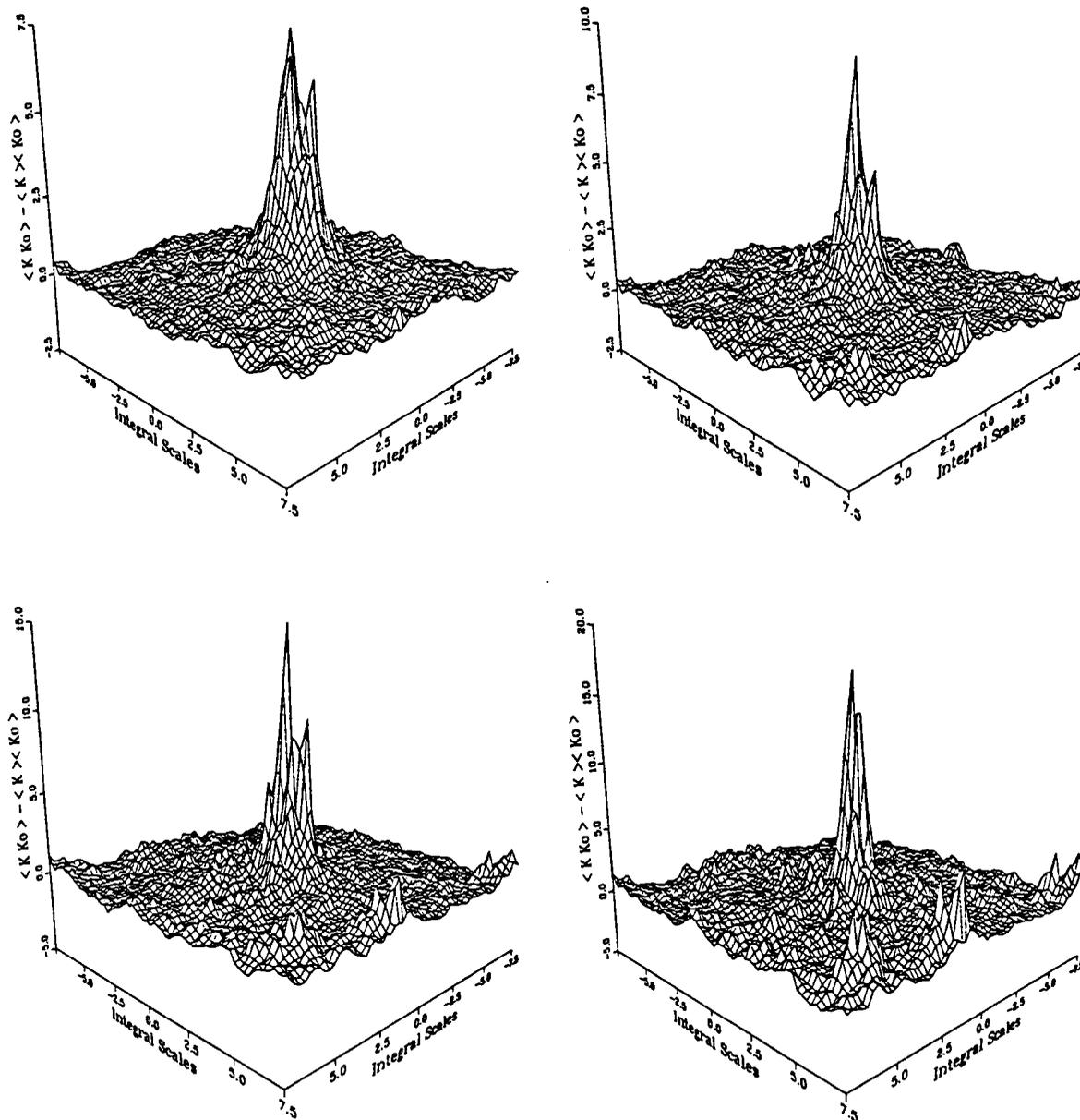


Figure 4.23: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the FFT method with the IMSL-based RNG.

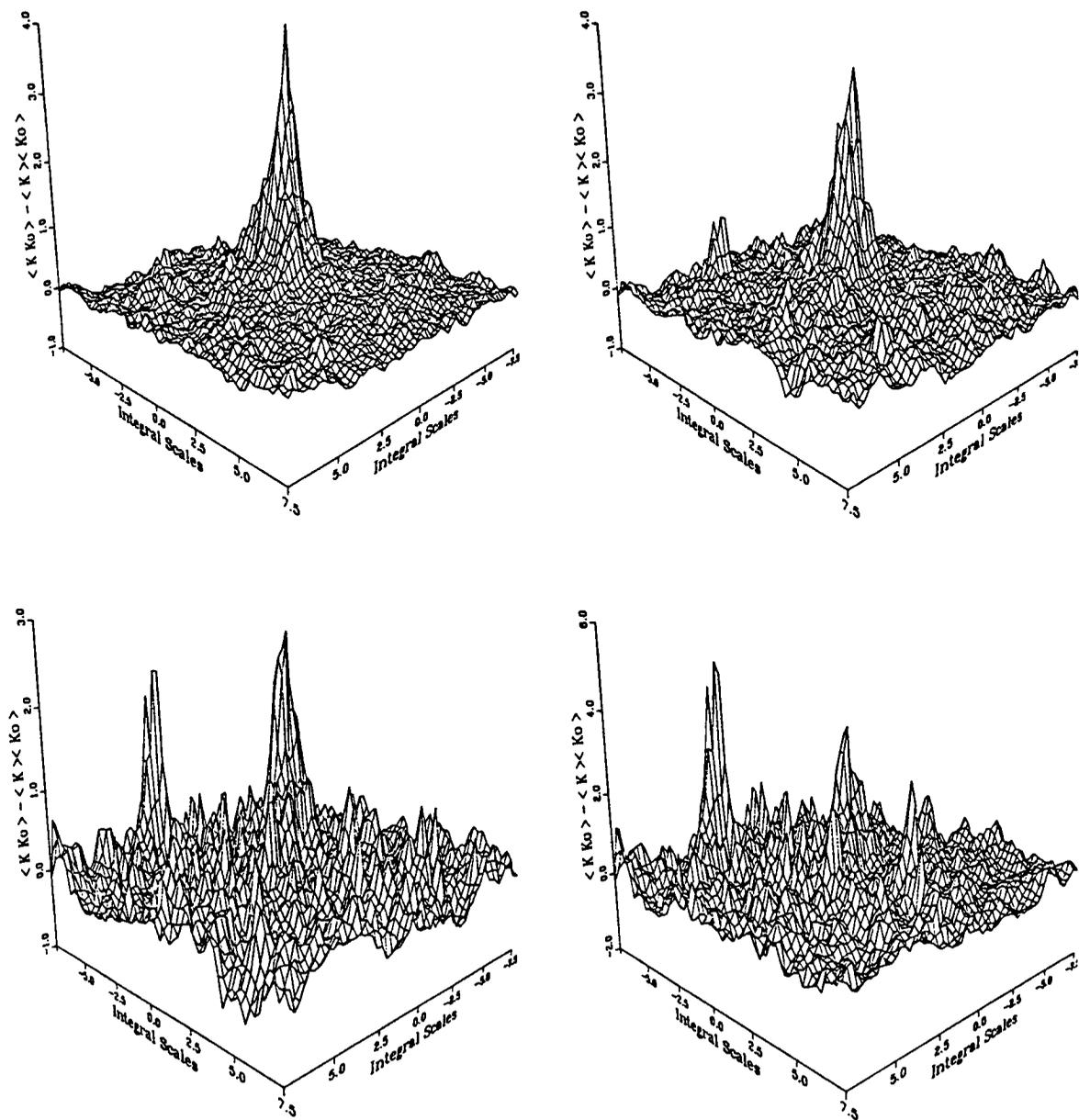


Figure 4.24: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the FFT method with a NAG-based RNG.

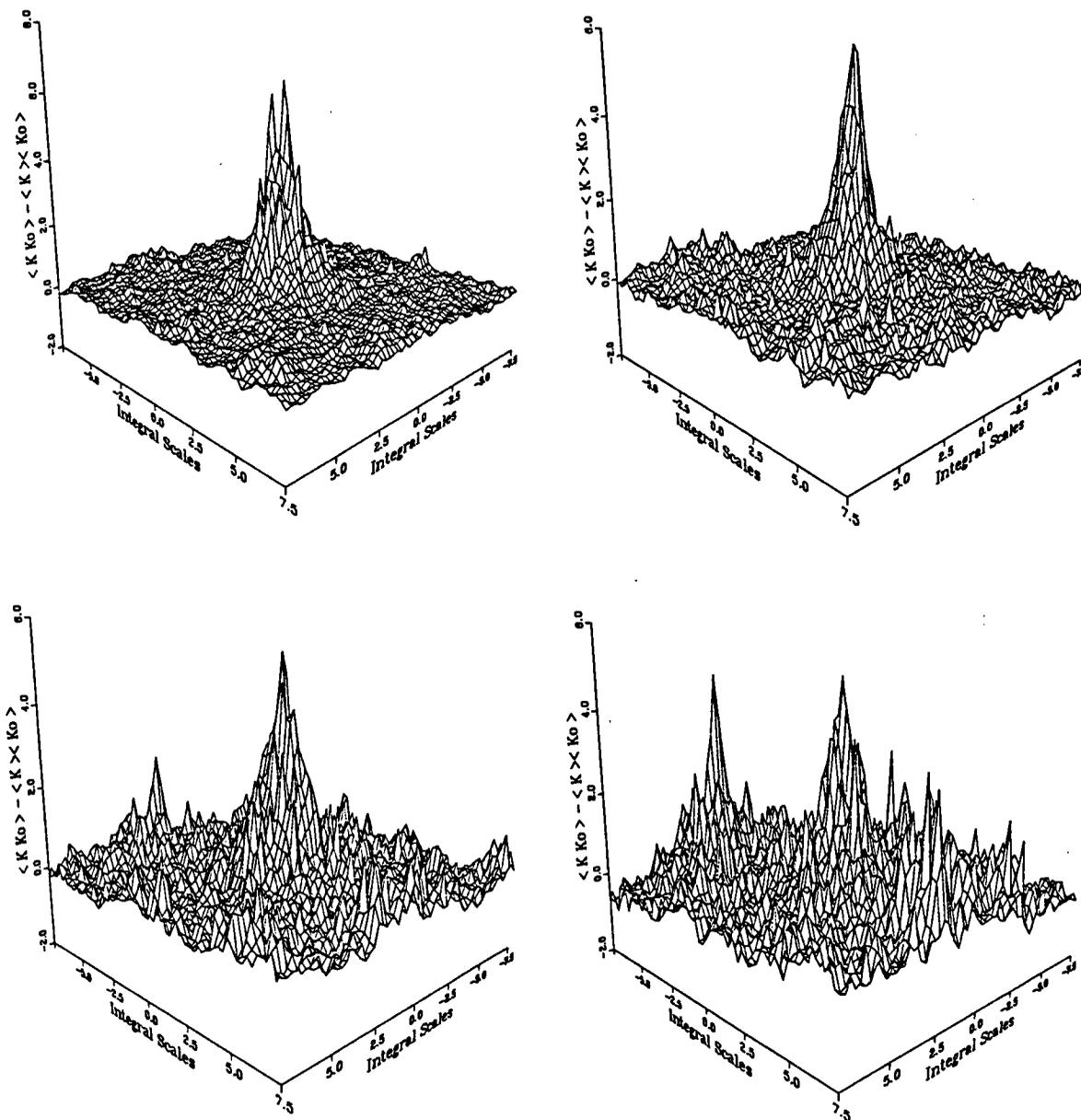


Figure 4.25: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the sequential simulator (GCOSIM3D) with DPRAND.

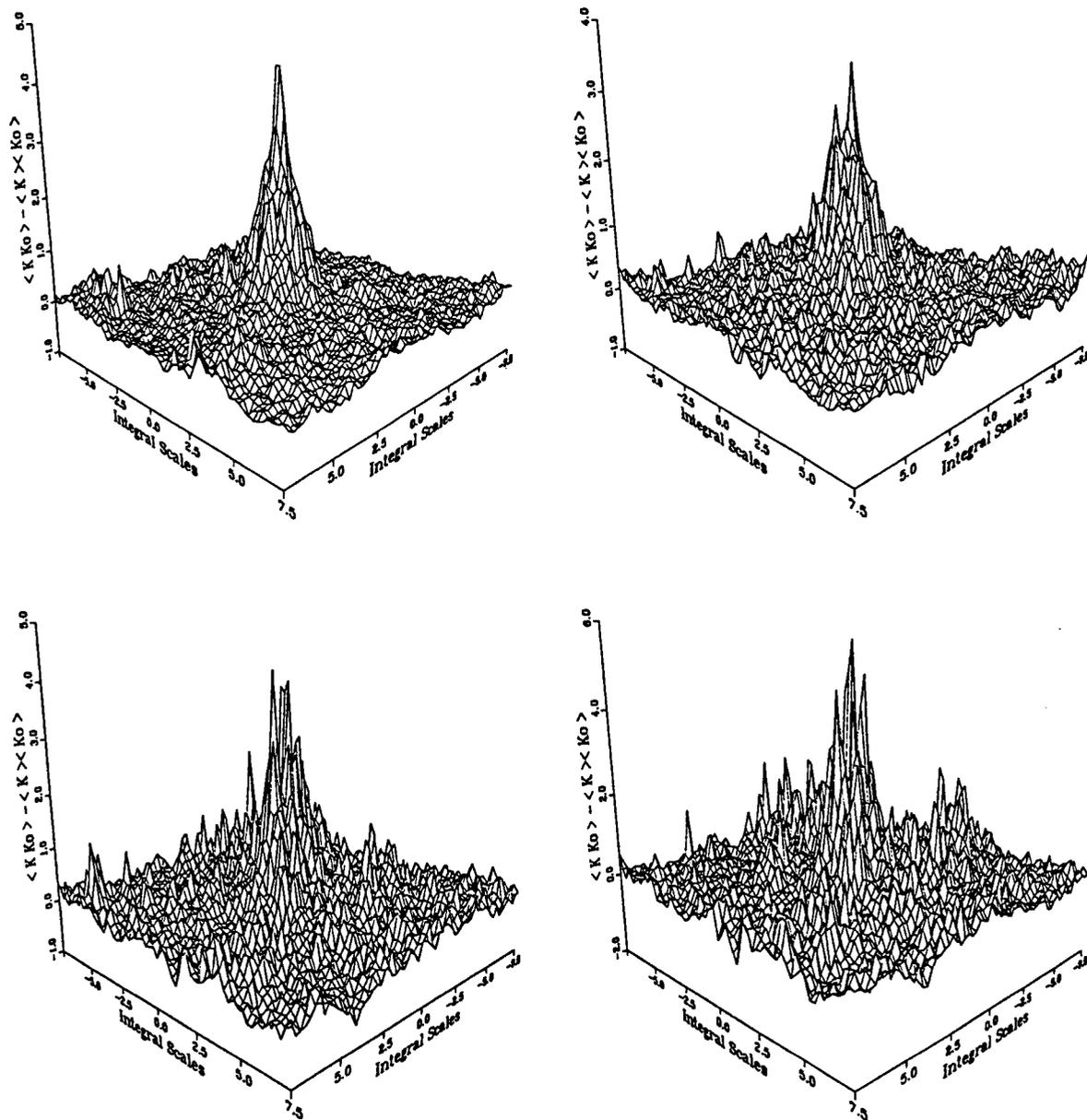


Figure 4.26: “Ensemble” covariances of (clockwise, starting from the upper left) 1000, 500, 200, and 100 realizations, generated by the sequential simulator (GCOSIM3D) with RAN2.

of validity of the weak approximation and effective conductivities). Finally, Figure 4.27 shows the “ensemble” covariances of 1000 realizations generated by the defected RFG. The “hump” of correlation (left to the major peak) is a clue to the anomaly which was clearly observed in the previous plots (Fig. 4.16).

The following figures and the approach behind them are due to Thomas Harter (personal communication). The turning bands simulator used in the following analysis has a defect (or bug, as shown in Figures 4.16, 4.27); consequently it does not represent the turning bands method, but is rather used to test the analysis itself.

Figure 4.28 illustrates (single) realizations generated by the four different RFG’s. The fields are normally distributed with the same characteristics as above (i.e., 64×64 fields with $K_g = 1$, $\sigma_Y^2 = 1$, $\lambda = 5$). There is very little one can conclude from these plots, except that the fields (indeed) seem to be random and somewhat auto-correlated. The only important conclusion from this figure is the (known) fact that the theoretically unbounded variability is practically bounded (between 3 and -3 , i.e., within a range of 3 orders of magnitude).

Figure 4.29 shows the sample (ensemble) means of 1000 realizations (normal fields); except for slight stripping by the *tbm*, it is inconclusive. Figure 4.30 shows “ensemble” variances of 1000 simulations with the four (tested) RFG’s. Strips in a form of “rays” emanating from the lower left corner of the *tbm* fields give the first clue to an anomaly. Figures 4.31, 4.32, 4.33 depict spatial covariances of single realizations generated by the four different methods. Almost each figure shows anisotropy and/or irregularity; the most irregular are the ones generated by the *tbm*, while it seems that the least irregular are the ones generated by the LU decomposition method. Figure

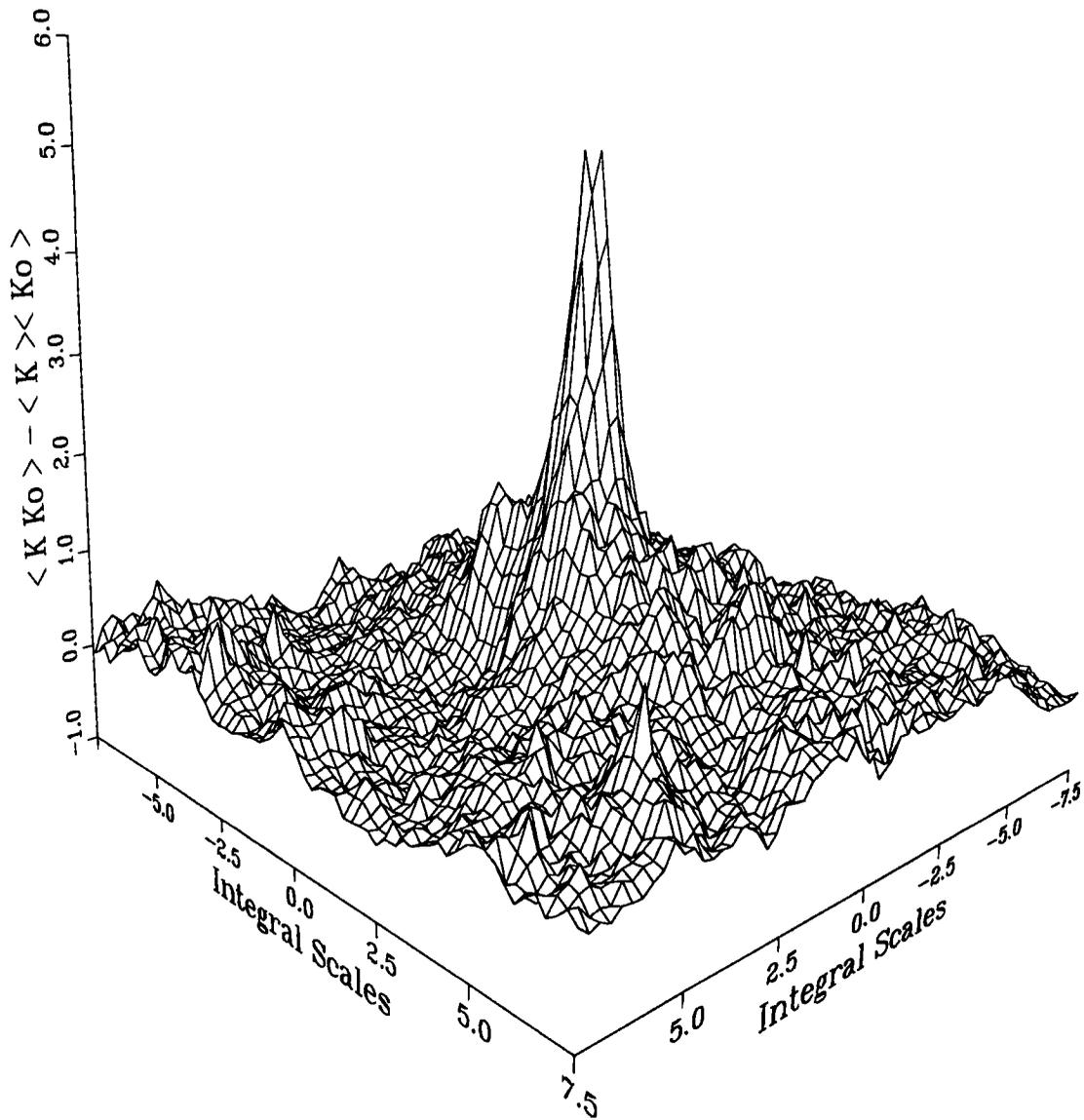


Figure 4.27: “Ensemble” covariances of 1000 realizations generated by a defected RFG, based on the turning bands method.

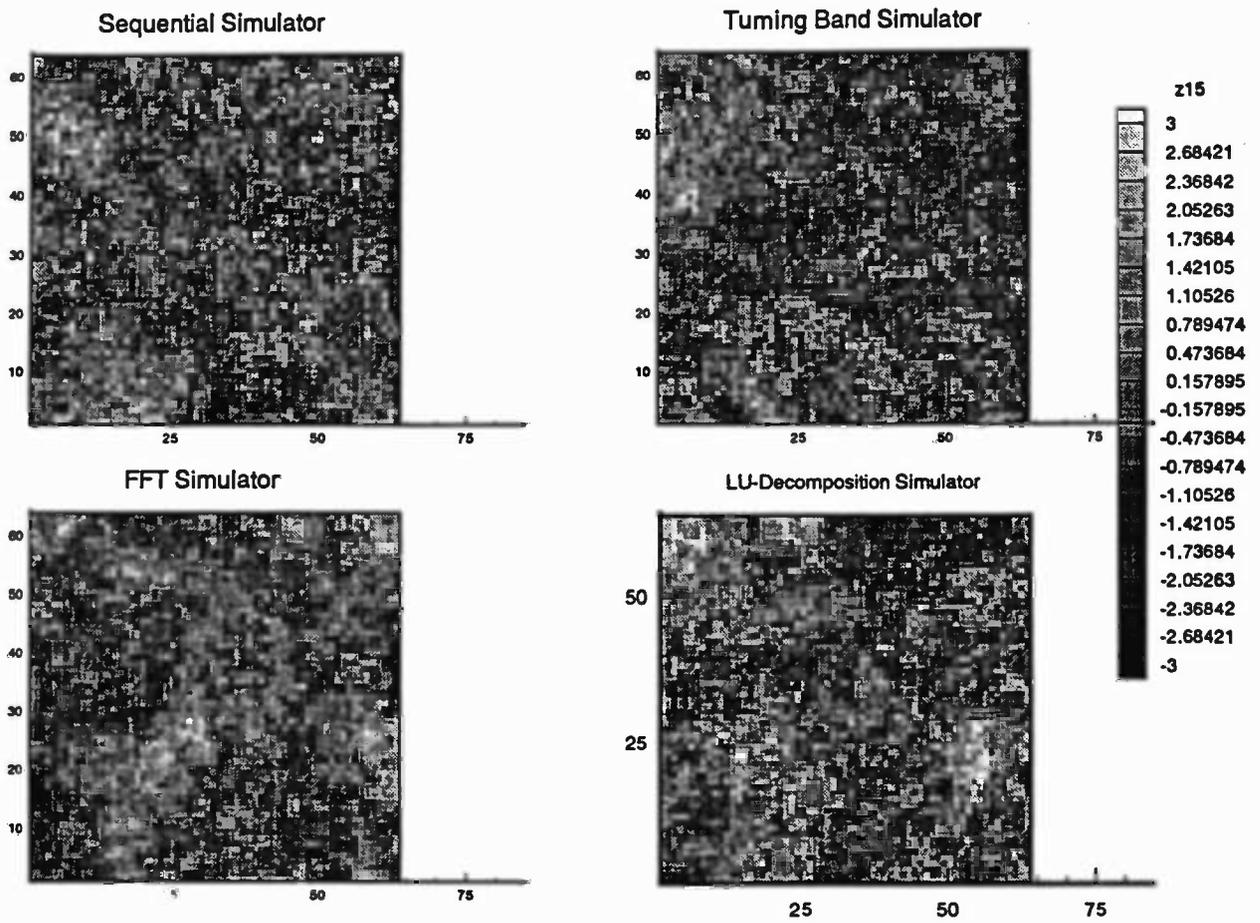


Figure 4.28: Realizations generated by four different random field generators (Thomas Harter, pers. com.).

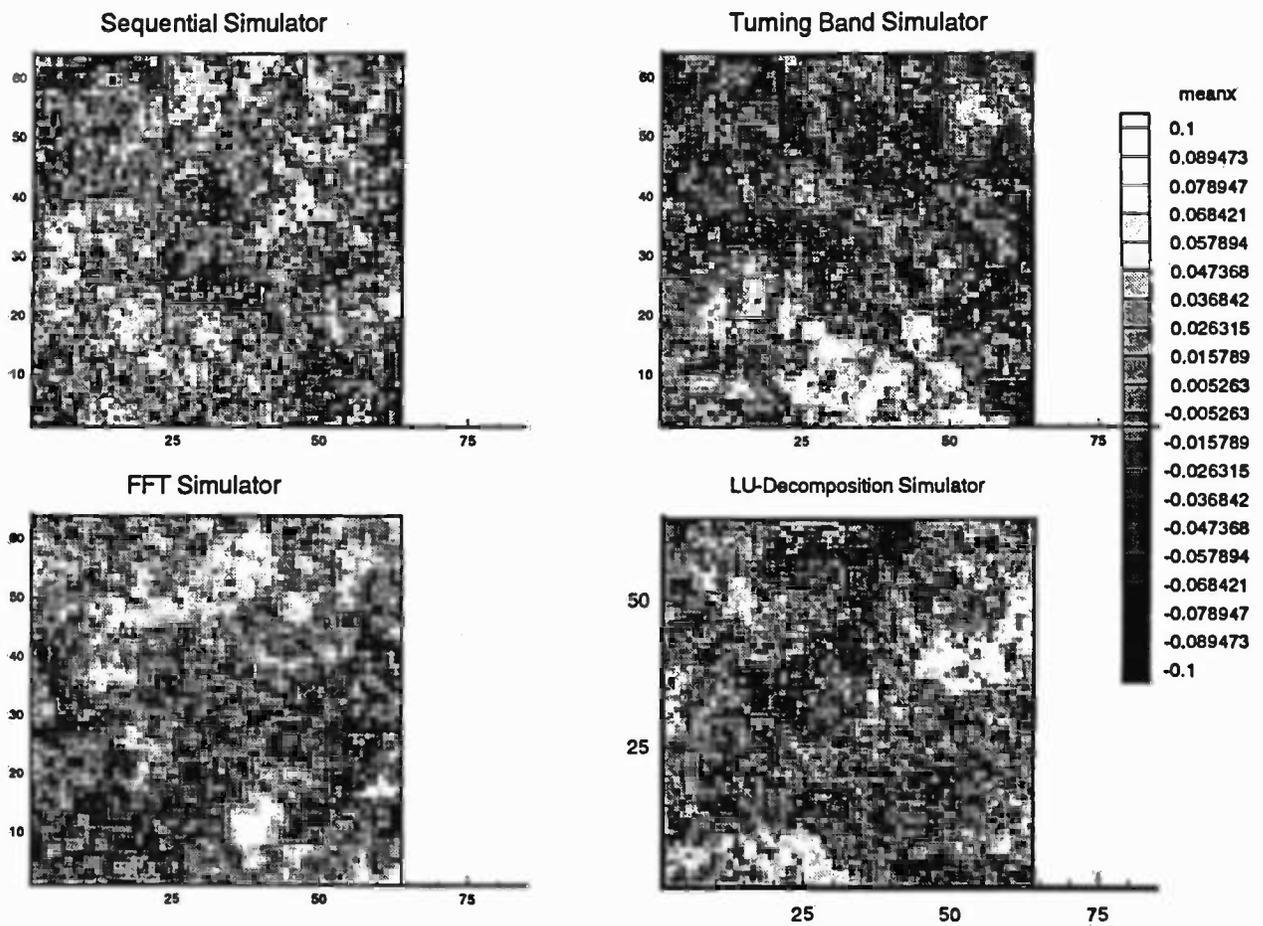


Figure 4.29: “Ensemble” means of 1000 realizations of normally distributed fields generated by four different RFG’s (Thomas Harter, pers. com.).

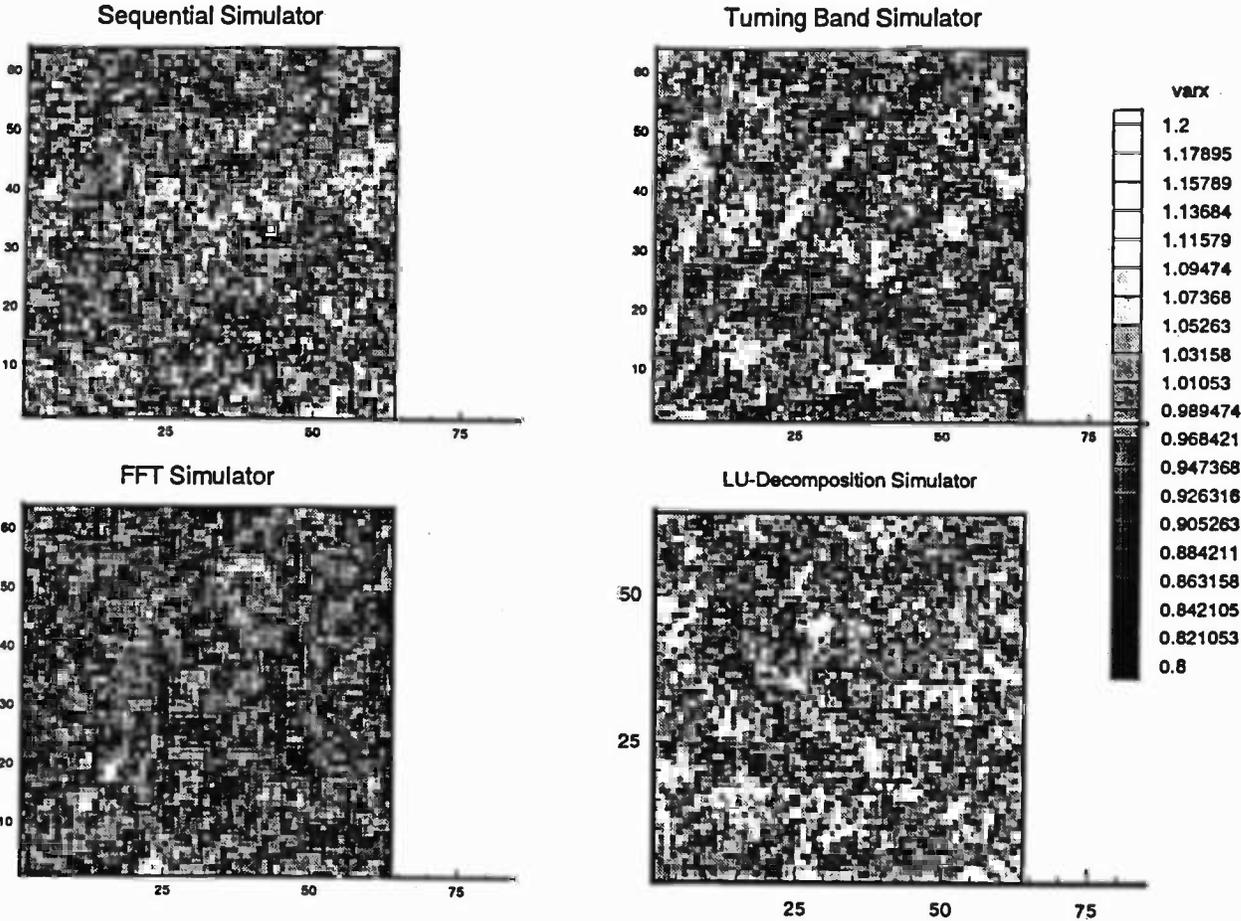


Figure 4.30: “Ensemble” variances of 1000 simulations with the four (tested) RFG’s (Thomas Harter, pers. com.).

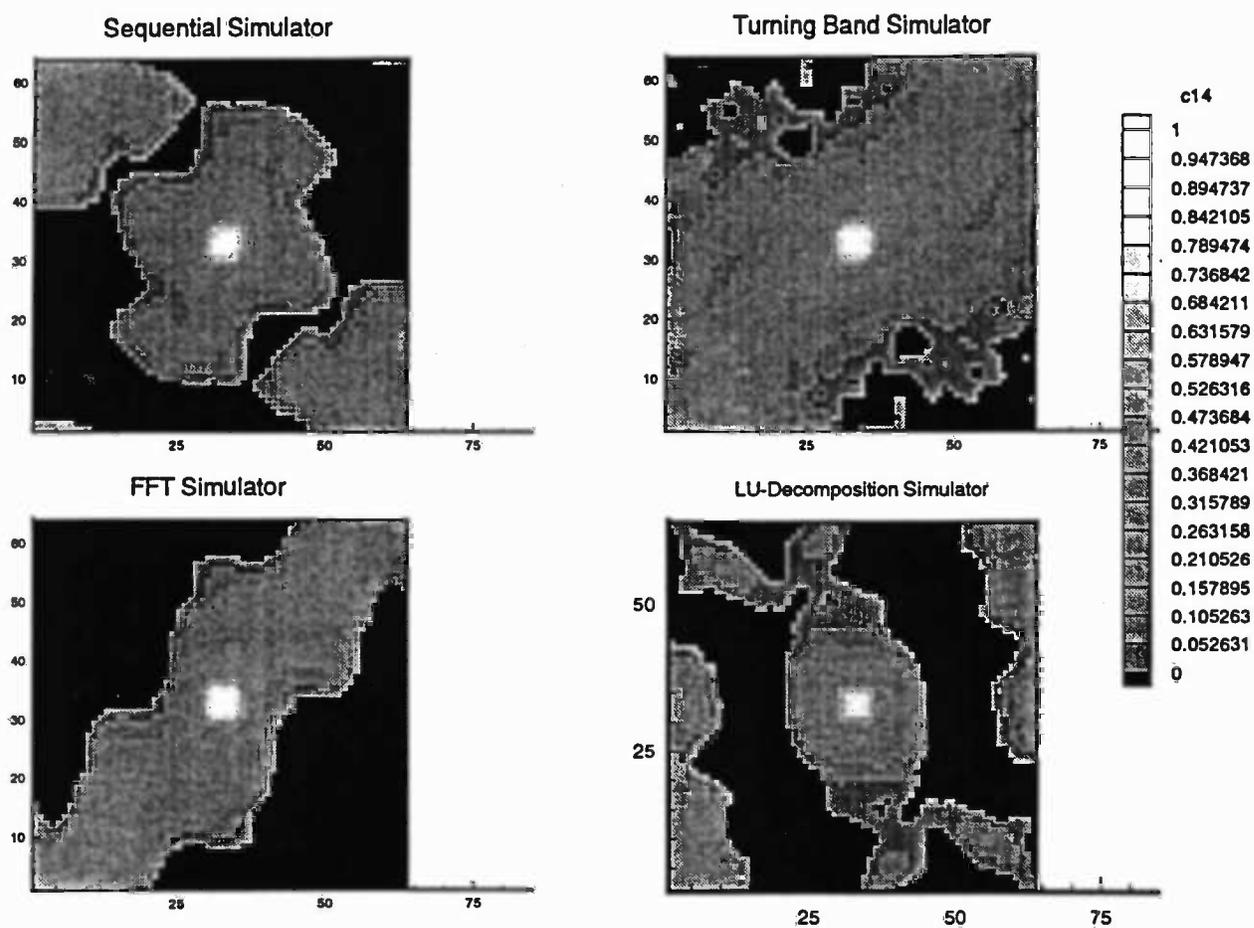


Figure 4.31: Spatial covariances of single realizations generated by the four different methods (Thomas Harter, pers. com.).

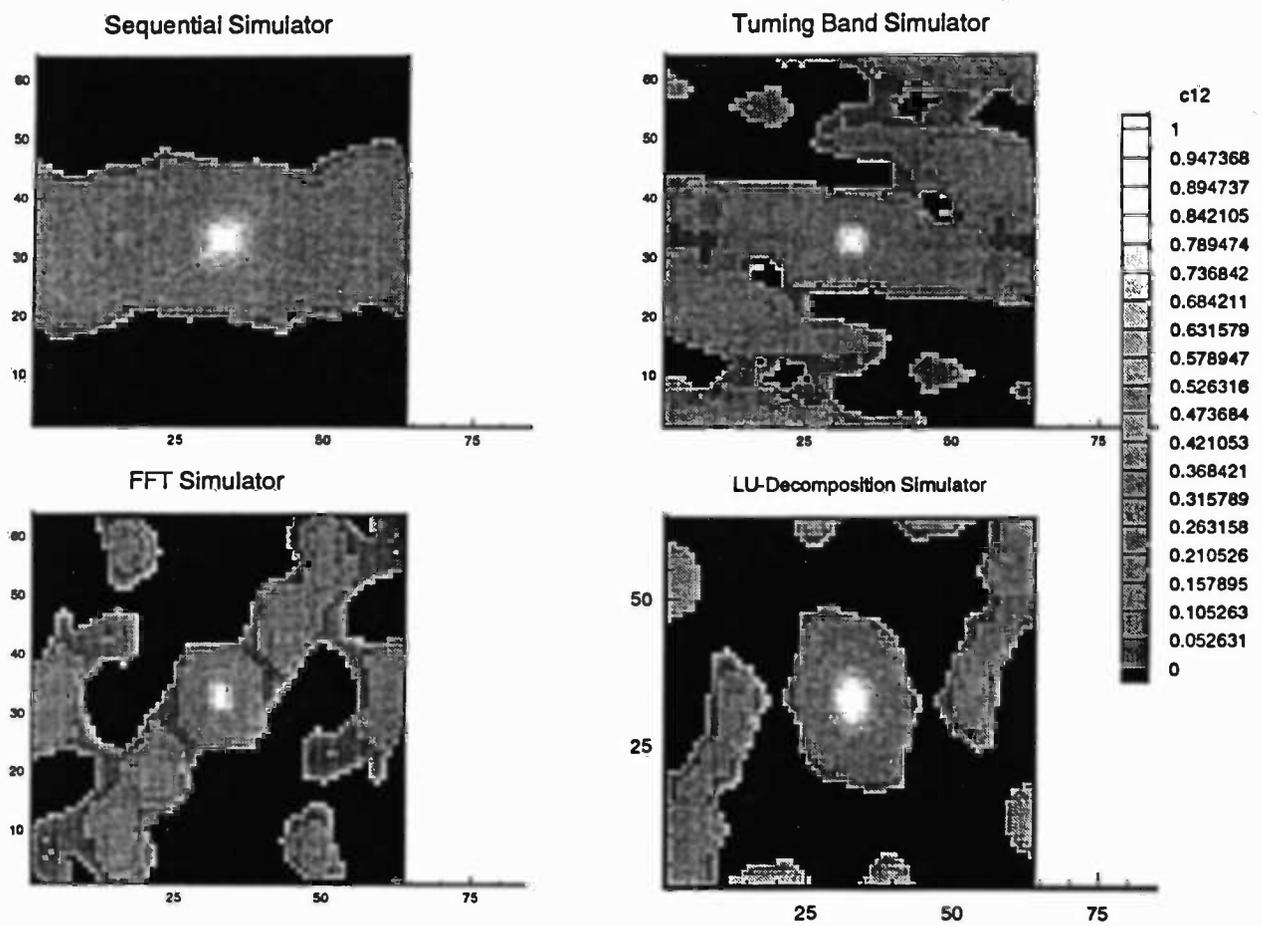


Figure 4.32: Spatial covariances of single realizations generated by the four different methods (Thomas Harter, pers. com.).

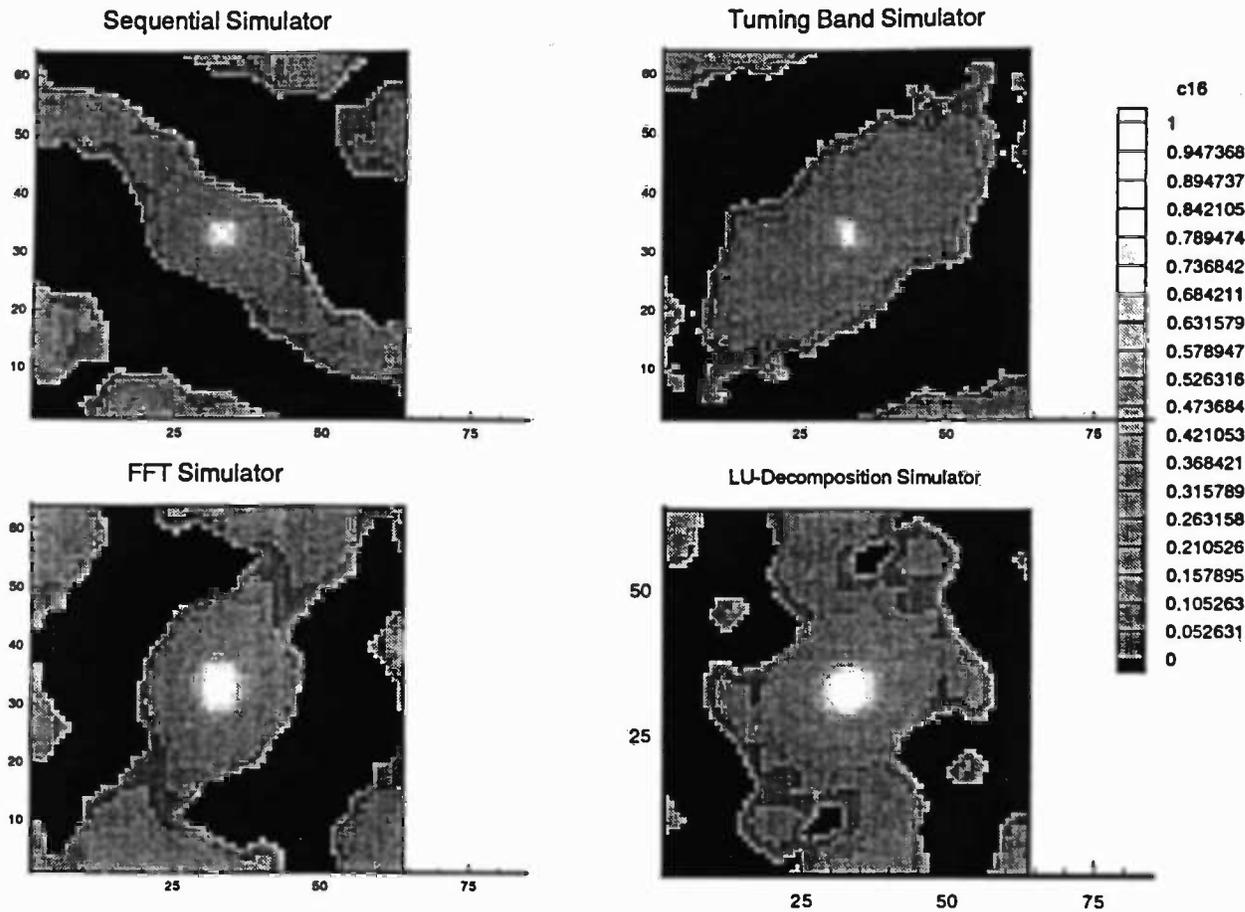


Figure 4.33: Spatial covariances of single realizations generated by the four different methods (Thomas Harter, pers. com.).

4.34 depicts averages of spatial covariances over 10 realizations, while Figure 4.35 shows the same for 50 simulations. The anomaly of the defected *tbm* starts to show up, while the other RFG's start to show the desirable (symmetric) shape of the theoretical covariance, with no particular distinction. Figures 4.36, 4.37 show the deviations from the theoretical covariances for 10 and 50 realizations respectively. The defected *tbm* exhibits striking star-like patterns which seems to stem from both the anomaly showed earlier and the 16 lines of the *tbm*. The sequential simulator (GCOSIM) shows some clear lateral anisotropy. In the following stage of the analysis, Harter (personal communication) assigns a window of size 32×32 that "moves" across the 64×64 field by one unit at a time, such that a total of 920 "ensemble" covariances could be sampled, and the statistics of these covariance "fields" could be investigated for the different RFG's with the different RNG's. His motivation was the misinformation involved with the above test of "ensemble" covariances between the mid-points and all other points. In the following tests, the mid-points move along with the "window", such that a more objective judgement can be carried out. The drawbacks, however, are (i) the fact that yet, only about 23% of the points in the field can take the roll of mid-points (instead of 4 points in the previous case), and (ii) the lose of covariance information for large lag distances, which has shown to dominate the misbehavior of the "ensemble" covariances in some cases (as seen above).

For example, the deviations from the theoretical covariance can be averaged. Figure 4.38 shows the mean deviation from the theoretical covariance of field portions as functions of the number of realizations (NMC). The field portions (32×32) were generated by LUSIM (LU-decomposition), with five different RNG's; "r", "e", "d",

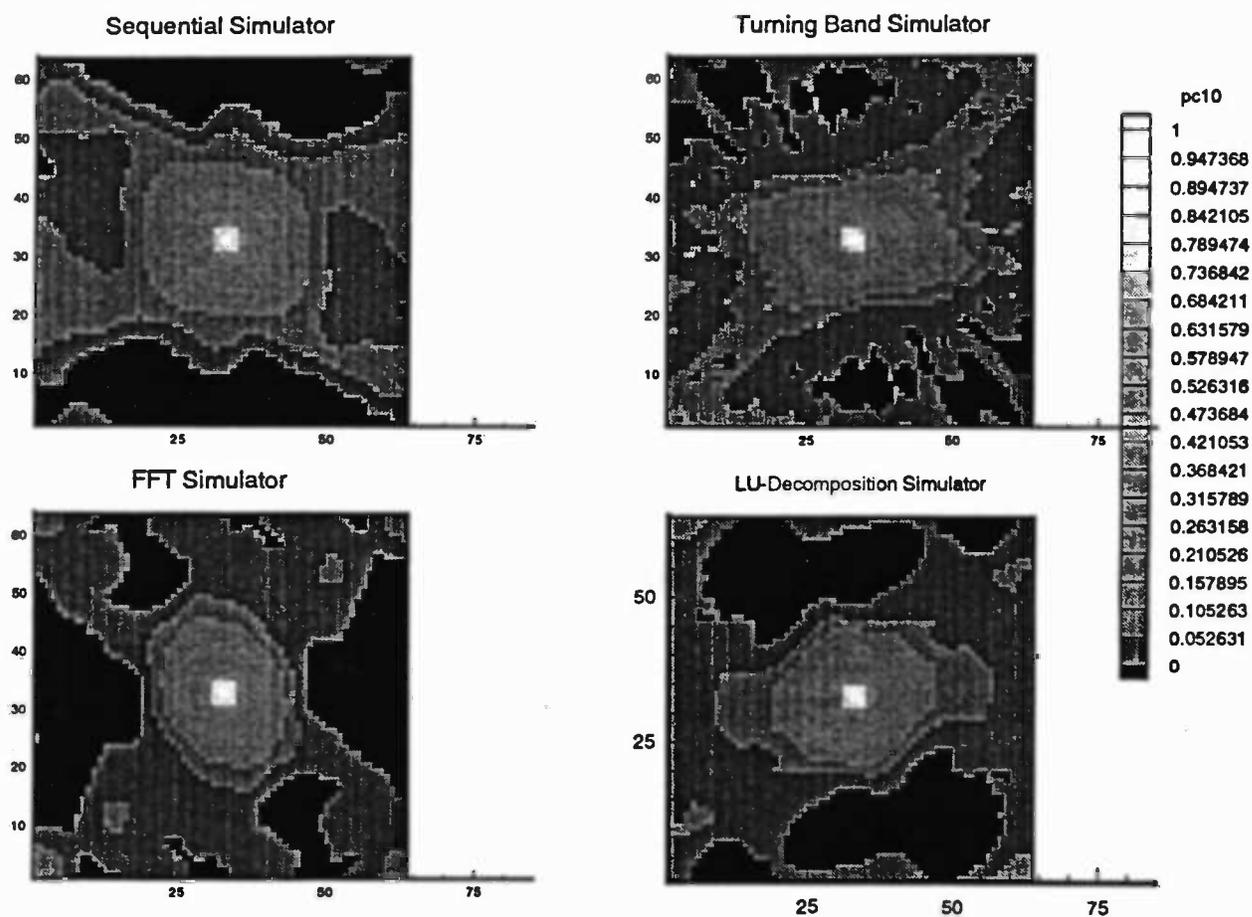


Figure 4.34: Averages of spatial covariances over 10 realizations generated by the four different methods (Thomas Harter, pers. com.).

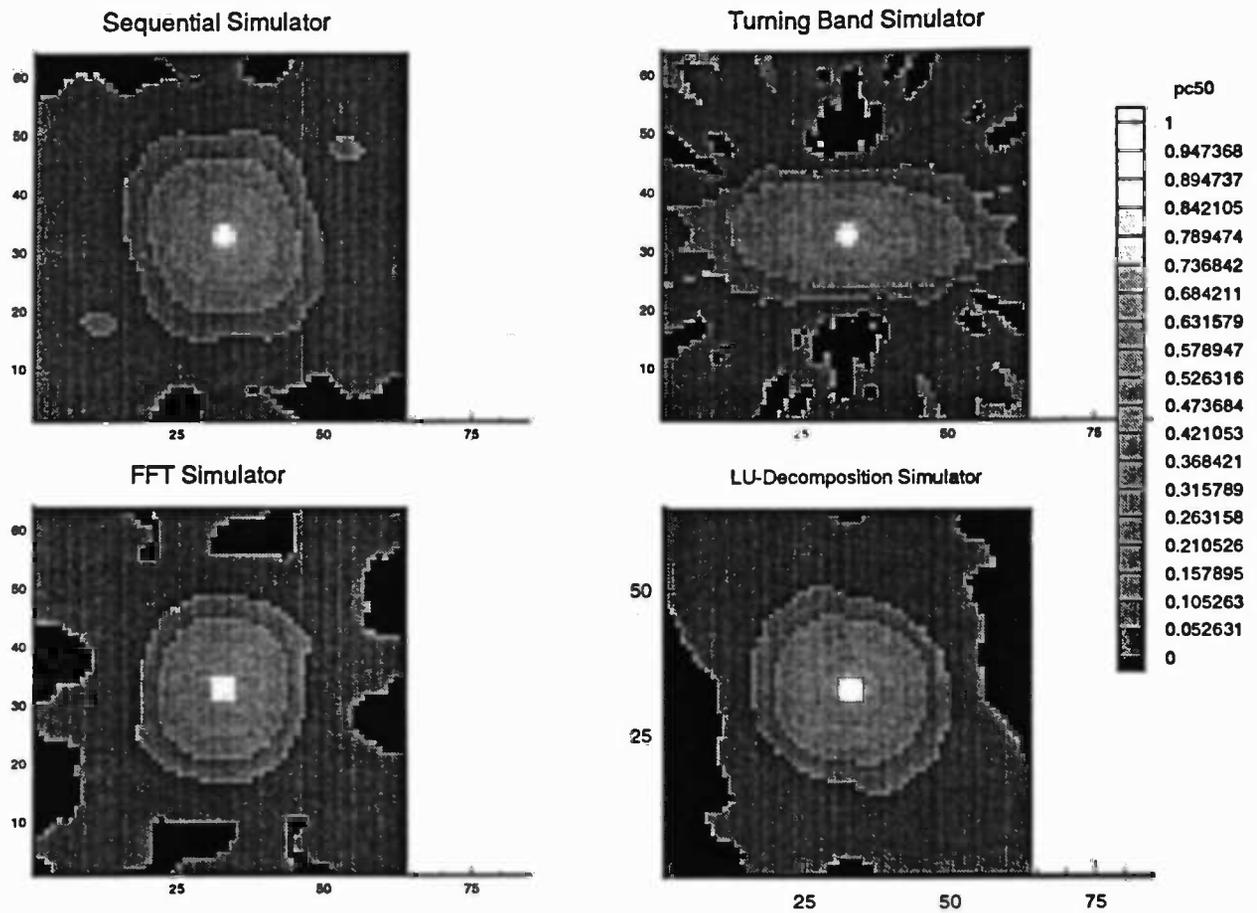


Figure 4.35: Averages of spatial covariances over 50 realizations generated by the four different methods (Thomas Harter, pers. com.).

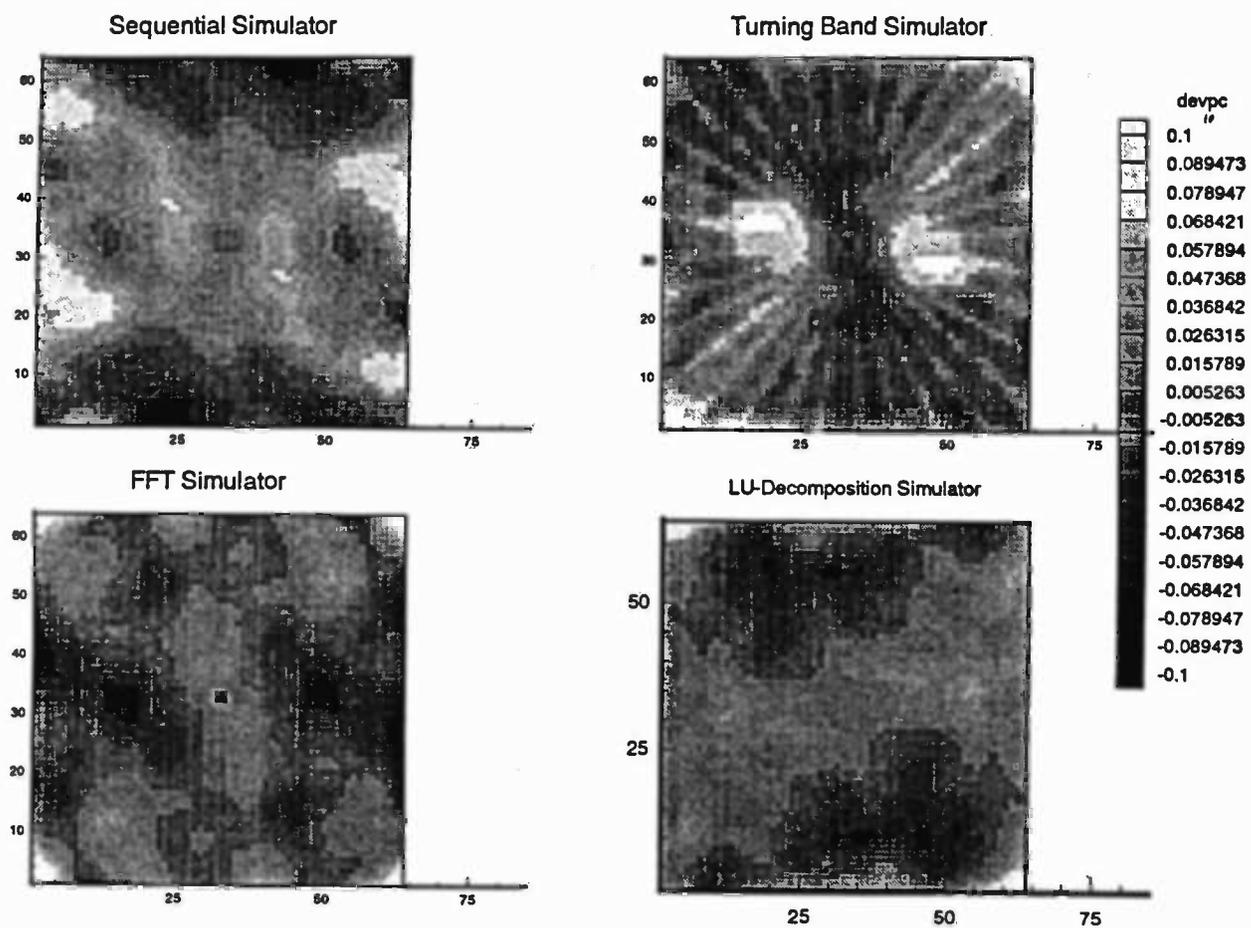


Figure 4.36: Deviations (from the theoretical covariance) of average spatial covariances over 10 realizations generated by the four different methods (Thomas Harter, pers. com.).

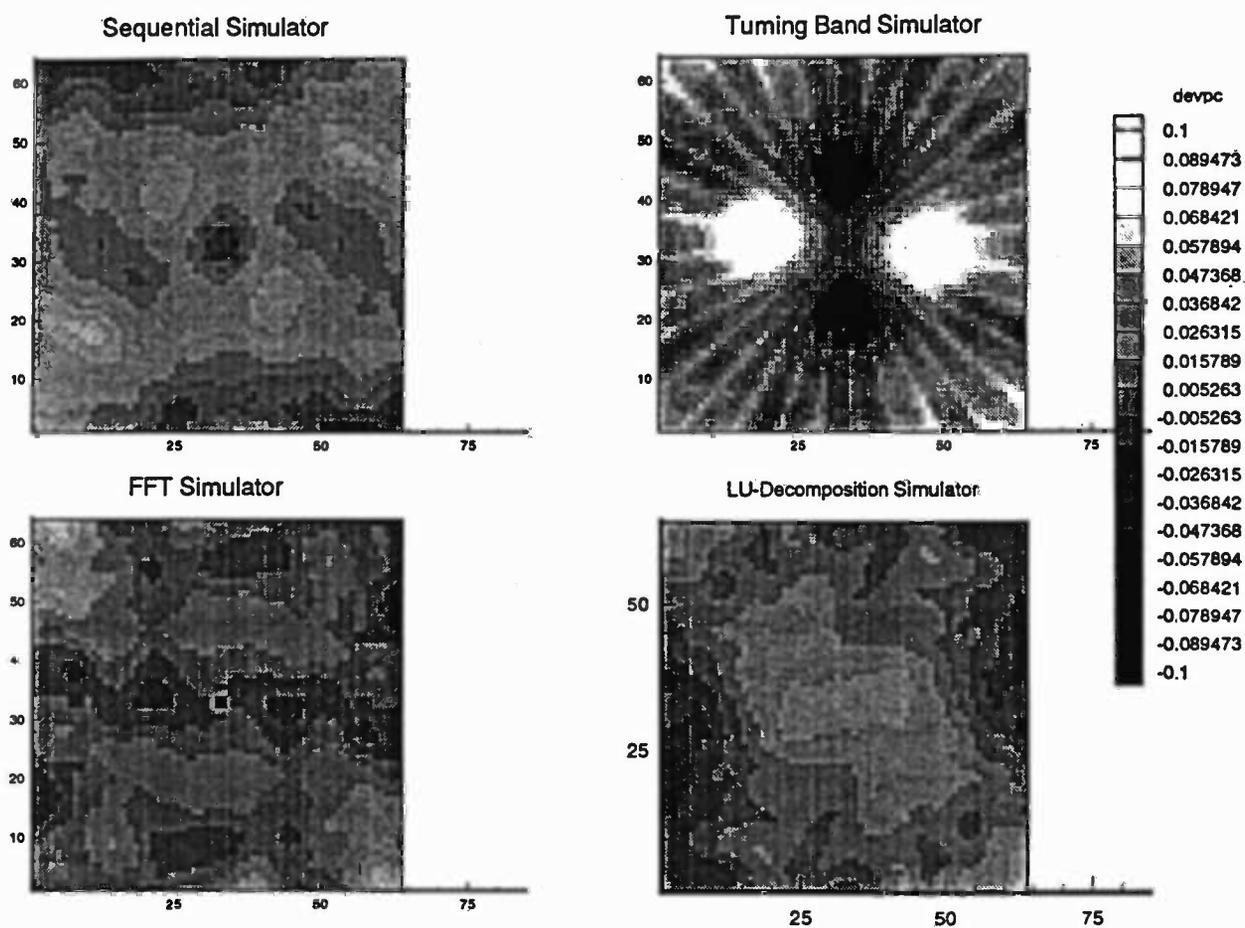


Figure 4.37: Deviations (from the theoretical covariance) of average spatial covariances over 50 realizations generated by the four different methods (Thomas Harter, pers. com.).

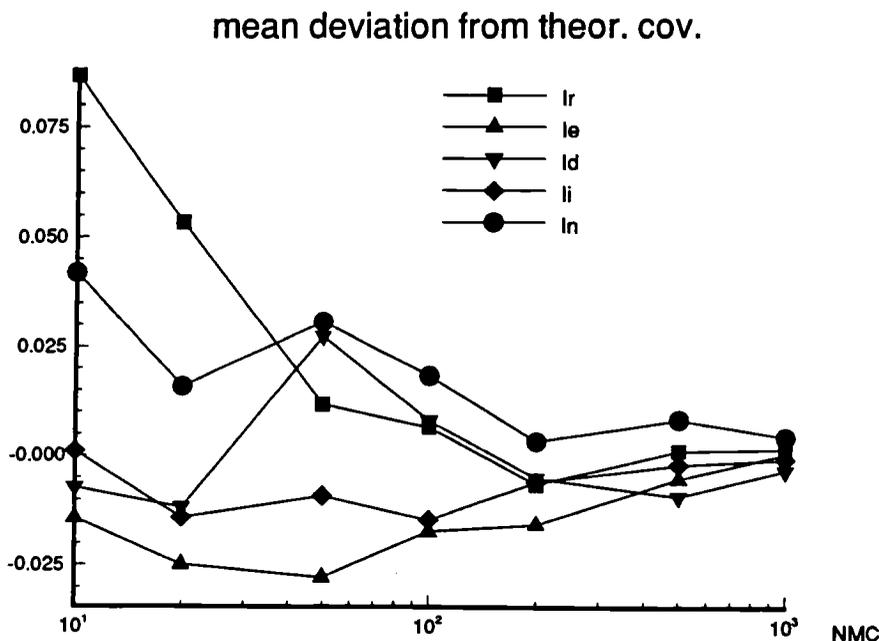


Figure 4.38: Mean deviation from the theoretical covariance of field portions as functions of the number of realizations (NMC). The fields were generated by LUSIM (LU-decomposition), with five different RNG's; "r", "e", "d", "i", and "n" stand for RAN2, ESSL, DPRAND, IMSL, and NAG, respectively (Thomas Harter, pers. com.).

"i", and "n" stand for RAN2, ESSL, DPRAND, IMSL, and NAG, respectively. As expected, the means converge to zero as soon as a few hundred realizations are averaged [or, combining the spatial and ensemble field portions, a few thousand realizations (because there are $920 \times NMC$ partial realizations)]. One may pick the 200 simulations ($NMC = 200$), as the threshold, and give the best "grade" to RAN2, DPRAND, and IMSL, for this particular test; the rest of the RNG's does not fall far behind, however. Figure 4.39 illustrates the same graph, but now, for the same RNG (DPRAND) used by the four different methods [where "f", "l", "g", "t" stand

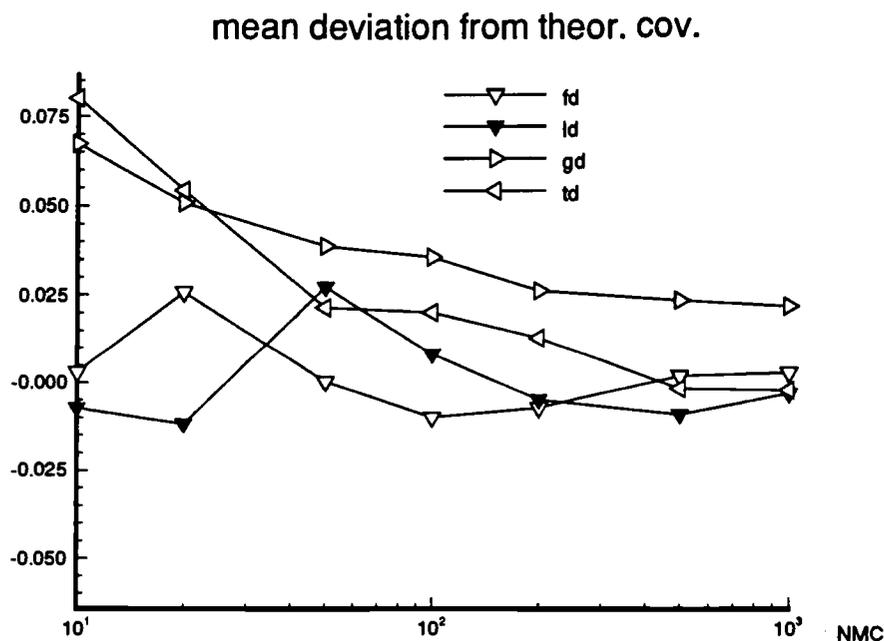


Figure 4.39: Mean deviation from the theoretical covariance of field portions as functions of the number of realizations (NMC). The fields were generated with the portable DPRAND RNG, used by the four different methods [where “f”, “l”, “g”, “t” stand for FFT, LU-decomposition, sequential simulations (GCOSIM), and *tbm*, respectively] (Thomas Harter, pers. com.).

for FFT, LU-decomposition, sequential simulations (GCOSIM), and turning bands methods, respectively]. In this case, the LU-decomposition and the FFT methods show superiority, while the sequential simulator shows some inferiority.

Figure 4.40 shows the spatial average of “ensemble” mean log-conductivities over 920 windows on fields generated by the (same) four different methods, with the same RNG (DPRAND). Beyond 100 simulations, none of the methods seems to be superior.

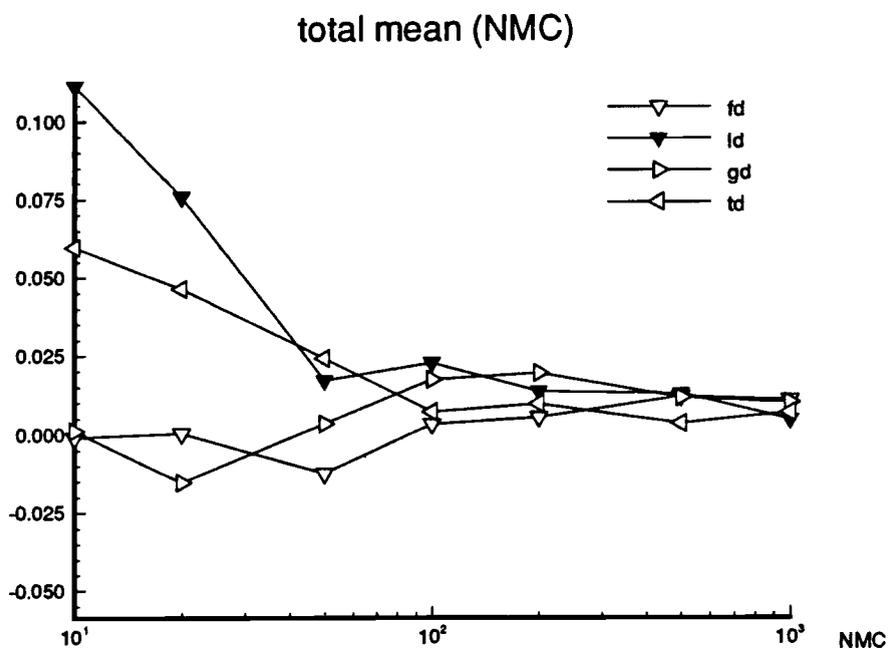


Figure 4.40: Spatial average of “ensemble” mean log-conductivities over 920 windows on fields generated with the same four methods [“f”, “l”, “g”, “t” stand for FFT, LU-decomposition, sequential simulations (GCOSIM), and turning bands methods, respectively], with DPRAND (Thomas Harter, pers. com.).

In conclusion, none of the RFG's has demonstrated superior performances for the 64×64 fields, but the defects of the particular *tbm* code could be exposed by some of the tests. It seems that the *combination* RFG & RNG has to be accounted for, in addition to the usual tests of RNG's. Based on these (combined) tests, we (hesitantly) decided upon GCOSIM (Gómez-Hernández, (1991) as our choice of RFG for our purposes. We must admit that we wonder how different (if at all) would have our MCS results been, had we used a different RFG (e.g., based on LU-decomposition).

CHAPTER 5

HIGH RESOLUTION MONTE CARLO SIMULATIONS FOR RADIAL FLOW

5.1 Multigrid Methods and PLTMG

The multigrid method¹, first introduced by Brandt (1977)², is an iterative approach for solving algebraic equations arising from boundary-value problems on discrete spatial domains, and considered by many to be the most efficient method for solving elliptic partial differential equations. A unique characteristic of multigrid methods is the use of series of coarser grids to accelerate convergence (Cole and Foote, 1990). As will be seen, the combination of fine and coarse grids provides optimal resolution of all spatial frequency components of both the solution and the errors associated with it (which we wish to minimize).

Starting with a governing equation of the form

$$L h + f = 0 \tag{5.1}$$

where, again, L is a linear differential operator, h is the solution (or response, or dependent variable), and f is the source term, we define the error e as the required *correction* to the approximate solution, v , or (Cole and Foote, 1990)

$$e = h - v \tag{5.2}$$

¹This brief introduction is based on Fletcher (1988), Briggs (1987), and Cole and Foote (1990).

²Achi Brandt had published papers on this topic several years earlier, however, the 1977 paper is the most comprehensive one.

As the solution v does not satisfy (5.1), we are left with the residual

$$\varepsilon = -f - L v \quad (5.3)$$

The error e then satisfies

$$L e \equiv L (h - v) = L h - L v = -f - L v = \varepsilon \quad (5.4)$$

which is the “residual equation” from which e can be calculated. Given e for the approximate solution v , the exact solution, h , is then estimated as $h = v + e$. This *defect correction* (or *iterative improvement*) process [which is a (known) “trick” to restore full machine precision and to overcome roundoff errors (e.g., Press et al. 1986, pp. 41-42)], is central to the multigrid method. It can be summarized by (Cole and Foote, 1990)

1. Given the approximate solution v , compute the residual ε , using (5.3).
2. Solve (5.4), the residual equation, for the correction e .
3. Add the correction e to the approximate solution v to obtain an exact solution h .

In practice, the continuous problem (5.1) is approximated on a discrete grid, having a characteristic discretization scale Δ . At the discrete level, the operator L is approximated by the (coefficient) matrix \mathbf{A} , and the vectors \mathbf{V} , \mathbf{B} , \mathbf{W} and \mathbf{R} approximate the continuous variables h , $-f$, e , and ε above.

Multigrid methods work with a sequence of grids $m = 1, \dots, M$, where the grid size ratio $\Delta_{m+1}/\Delta_m = 0.5$. The methods are applicable to both linear and

nonlinear systems. A linear system (for example) to be solved on the finest grid is written

$$\mathbf{A}^M \mathbf{V}^M = \mathbf{B}^M \quad (5.5)$$

where the matrix \mathbf{A} contains the algebraic coefficients arising from discretisation, \mathbf{V} is the vector of unknown nodal values (i.e., the solution vector), and \mathbf{B} is a vector made up of algebraic coefficients associated with discretisation and known values of \mathbf{V} (e.g., that are given by the boundary conditions and source terms). When \mathbf{A} depends on \mathbf{V} (i.e., $\mathbf{A} = \mathbf{A}(\mathbf{V})$), the equation is nonlinear. An approximation to \mathbf{V}^M in (5.5) is provided by the solution on the next coarser grid, \mathbf{V}^{M-1} . Similarly, for an intermediate grid solution, \mathbf{V}^m is a good approximation to the solution on the next finer grid \mathbf{V}^{m+1} . If an approximation to the solution of $\mathbf{A}^{m+1} \mathbf{V}^{m+1} = \mathbf{B}^{m+1}$ is denoted by $\mathbf{V}^{m+1,a}$, so that

$$\mathbf{V}^{m+1} = \mathbf{V}^{m+1,a} + \mathbf{W}^{m+1} \quad (5.6)$$

then

$$\mathbf{A}^{m+1} \mathbf{W}^{m+1} = \mathbf{B}^{m+1} - \mathbf{A}^{m+1} \mathbf{V}^{m+1,a} = \mathbf{R}^{m+1}. \quad (5.7)$$

The correction, \mathbf{W}^{m+1} , and the residual, \mathbf{R}^{m+1} , are closely approximated by the correction and the residual on the next coarser grid \mathbf{W}^m and \mathbf{R}^m , if they are smooth enough, i.e., if the amplitudes of high frequency components are small. The highest frequency (or rather, the shortest wave length) that can be represented on a discrete grid, Δ_m , is $2\Delta_m$ (e.g., Fletcher, 1988, and Briggs, 1987).

Iterative methods such as Jacobi, Gauss-Seidel, and SOR (successive over relaxation) remove high frequency components in a few iterations. It is the removal of the low frequency components of the error, and equivalently of the residual, that

causes the slow convergence of iterative (relaxation) methods on a fixed grid. However, a low frequency component on a fine grid becomes a high frequency component on a coarse grid. Multigrid methods seek to exploit the high frequency smoothing of relaxation methods.

If nothing is known about the solution, the process would start with the solution on the coarsest grid, and obtained by either a direct or a conventional iterative method. A high order interpolation is made to the next finer grid, and (say, r) multigrid V-cycles (described below) are applied to improve the solution.

A single multigrid V-cycle starts from the *finest* existing grid. For example, after solving $\mathbf{A}^1 \mathbf{V}^1 = \mathbf{B}^1$ “exactly” on the coarsest grid ($m = 1$), the solution is interpolated to the next, finer grid ($m = 2$), where the V-cycle begins to solve $\mathbf{A}^2 \mathbf{V}^2 = \mathbf{B}^2$, iteratively, as in (5.6) and (5.7); however, rather than using many iterations in an effort to reduce the residual \mathbf{R}^2 below a negligible error, ϵ , the iterative process stops after a few iterations, and the (intermediate) solution (on the fine grid, $m = 2$) is projected to the coarser grid ($m = 1$), and is solved there completely (by either a direct or an iterative method). The new (approximate) solution, or rather, the correction, \mathbf{W}^1 , is, then, interpolated to the finer grid, and the process (i.e., the V-cycle) repeats itself as many times as needed (r times), until the residual (or rather, its norm) $|\mathbf{R}^{m,r}| < \epsilon$. In a full multigrid method (FMG), the (converged) V-cycles are repeated with increasing grid refinement (and, hence, an increasing cycle length), until $m = M$. The full multigrid (FMG) method is shown schematically in Figure 5.1 (adopted from Fletcher, 1988) The function $INT(\mathbf{V}^m, m)$ represents a cubic interpolation, first in the x direction, and then, for all x^{m+1} grid points, in the y directions. The function MGI denotes one path through the multigrid V-cycle

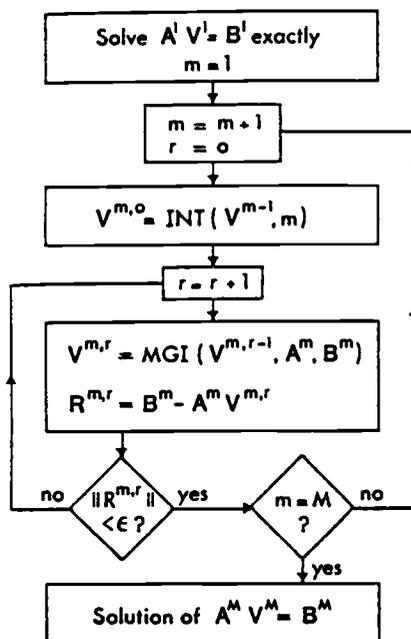


Figure 5.1: Flow chart for the FMG method (after Fletcher, 1988).

(Fletcher, 1988), discussed and shown below.

A complete (single) V-cycle starting from the *finest* grid ($m = M$), consists of the following steps:

1. Repeat a few iterations (or relaxation steps), ν_1 , to smooth the high frequency components in the correction and the residual for each grid, $m + 1$. That is, compute the residual vector, \mathbf{R}^{m+1} , and transfer (or project) it to the next coarser grid, m , using a “restriction (or projection) operator”, I_{m+1}^m , such that $\mathbf{R}^m = I_{m+1}^m \mathbf{R}^{m+1}$. Then, approximate the (new) coarse grid correction, \mathbf{W}^m , using a few iterations (ν_1), calculate an updated residual, and project it to the next coarser grid ($m - 1$), and so on, until the coarsest grid ($m = 1$) is reached.
2. Solve the exact or complete iterative solution of $\mathbf{A}^1 \mathbf{W}^1 = \mathbf{R}^1$ on the coarsest grid.
3. Repeat interpolations of the correction, \mathbf{W}^m , using a “prolongation operator”, I_m^{m+1} (usually based on bilinear interpolation in two dimensions), such that $\mathbf{W}^{m+1} I_m^{m+1} = \mathbf{W}^m$, and perform a few iterations, ν_2 , on each grid ($m + 1$). That is, on each grid, add the coarse grid correction to the finer grid solution, similar to (5.6); improve the solution \mathbf{V}^m by using a few iterations (ν_2) and get a new (smaller) correction; then interpolate it (using high order interpolation) to the next finer grid, and repeat the process until the finest grid, $m = M$, is reached.

The multigrid V-cycle is repeated until satisfactory convergence on the finest grid is achieved. According to Fletcher (1988), typically, ten cycles reduce the change in the solution to 10^{-5} , when (5.1) is a Poisson equation on a uniform grid. The flow

chart of the V-cycle is shown in Figure 5.2, (adopted from Fletcher, 1988).

The multigrid strategy can be viewed as an acceleration technique like the conjugate gradient method applied to a basic iterative scheme. Alternatively, the basic iterative scheme is interpreted as a smoother for the multigrid method, just as it was treated as a preconditioner for the conjugate gradient method (Fletcher, 1988).

PLTMG (version 6.1) due to Bank (1990), is a powerful multigrid finite element package that adaptively refines a mesh of linear triangles until the solution satisfies a number of stringent convergence criteria. PLTMG is a general solver for elliptic partial differential equations of the form (1.1) in general two dimensional domains, and handles non-linearity in both coefficients and source term. The finite element discretisations in PLTMG is based on piecewise continuous linear triangular finite elements. The package includes an initial mesh generator and several graphics packages. It was initially developed as a prototype for testing and verifying the properties of various multigrid and adaptive local mesh refinement algorithms. Since then, the software has evolved with the research interests of the author (R. Bank). The user can call for adaptive or user specified mesh refinement at any point along the solution path. There are also options for adaptive derefinement and subsequent re-refinement of an existing refined mesh. One can specify uniform grid refinement, adaptive grid refinement, or a combination consisting of uniform refinement followed by adaptive refinement. The adaptive procedure is based on the ideas of Babuska, Rheinboldt, and their co-workers, and Bank and Weiser (see references). The procedure used is commonly called a *feedback* scheme. Several algorithms, as well as some alternative criteria are used in order to achieve a required accuracy (Bank, 1986).

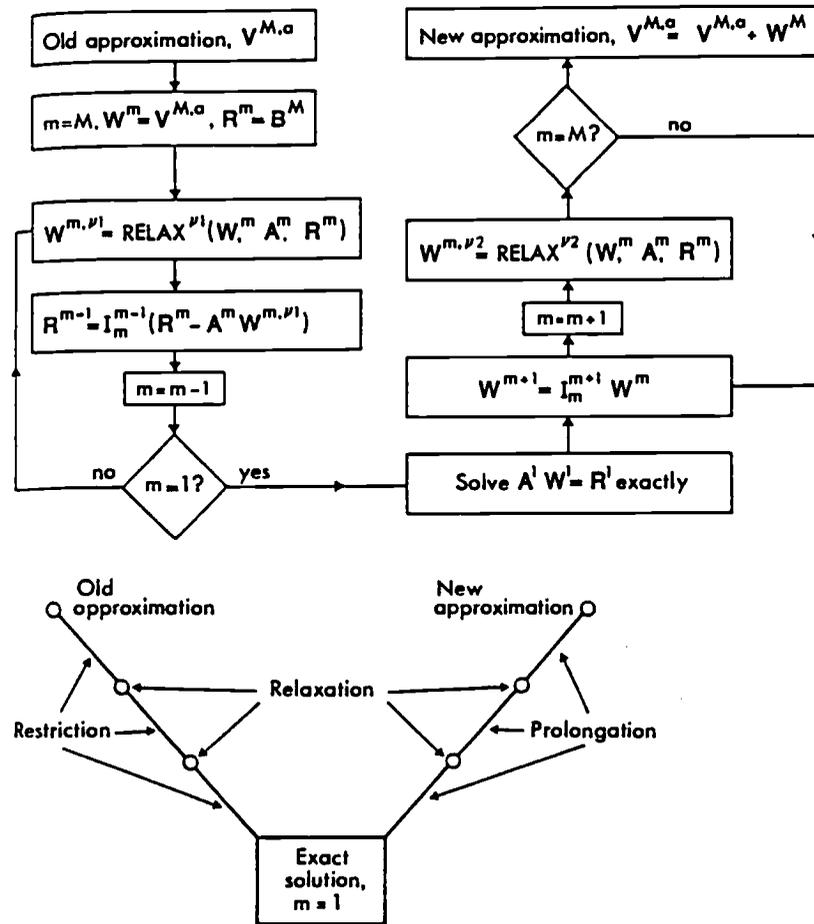


Figure 5.2: Flow chart for the multigrid V-cycle (after Fletcher, 1988).

The general procedure for solving a nonlinear problem are (Bank, 1990):

- Begin with initial guess of the solution for a given triangulation, number of vertices, and error tolerance;
- Solve a large set of linear equations, using the hierarchic basis multigrid iteration (a form of the V-cycle) to approximately solve the associated linear system (Bank and Dupont, 1980; Bank et al., 1988; Yserentant, 1985, 1986).
- For each triangle the error is estimated by solving a local Neumann problem in the triangle. The local error is approximated by a quadratic polynomial that is zero at the triangle vertices. In the simple case of linear source term, this require the assembly and solution of a 3×3 set of equations in each element, and leads to a global approximation of the error as a discontinuous piecewise quadratic polynomial that is zero at all vertices of the triangulation.
- Triangles whose error estimates are above some threshold, are refined. Often, this procedure results in a new triangulation, in which few new mesh points have been added; however, for computational efficiency, a geometric increase in the dimensions of the problem is used.

Sparse Gaussian elimination is used to solve the coarse mesh equations. Subsequently, the hierarchic basis multigrid method (the V-Cycle) is used (see Bank, 1990, p. 34 and the corresponding references for more details). This procedure is also used as a preconditioner for an overall iterative process. The solution on the refined meshes is done either by (preconditioned) conjugate gradient method (in the regular case of symmetric matrix), or by (preconditioned) biconjugate gradient iterations, for nonsymmetric matrix.

PLTMG also provides a variety of graphical displays of convergence histories, statistical data, and other interesting output. The solution includes both the dependent variable and its gradient. The input can be in the form of either coarse triangulation or as a “skeleton” of the region, which is the union of a set of non-overlapping subregions, possibly only one.

The big advantage of PLTMG is its ability to accurately simulate the response of a system (i.e., to solve the governing flow equation) near singularities like a point sink, especially in a highly variable random field. This is a major consideration when simulating random Green’s functions; i.e, when solving the governing flow equation of the form

$$\nabla_{\mathbf{x}'} \cdot [K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x})] + \delta(\mathbf{x}' - \mathbf{x}) = 0 \quad \mathbf{x}', \mathbf{x} \in \Omega \quad (5.8)$$

subject to some specified (deterministic) boundary conditions. Before solving (5.8) numerically, one has to decide

- How to simulate a delta function (or an impulse, which is a singularity).
- What boundary conditions to use (both types and values).

In order to minimize the effect of the boundary conditions (BC), we have tried to emulate an infinite domain, as much as possible. Recalling that the fundamental solution³ expresses a response to unit withdrawal at some distance r , we note that for any concentric cylinder (or circle) around the singularity (unit sink or source) in a homogeneous domain, the flux $q(r) = 1/2\pi r$, is directed toward the sink. Consequently, we prescribed the above flux at the outer circular boundary (i.e., Neumann

³That is, the Green’s function in an unbounded domain.

BC). This choice of boundary conditions was shown later to be more successful in minimizing boundary effects than Dirichlet BC (used by Desbarats, 1992a).

A straightforward way to simulate a Dirac (delta) function is to approximate the source term with the δ sequence (e.g., Greenberg, 1971, p. 61):

$$\delta \approx \frac{ke^{-kr^2}}{\pi} \quad (5.9)$$

where $k \rightarrow \infty$. We compared the resulting response (G) to the fundamental solution in 2-D

$$U = \frac{1}{2\pi} \ln(r/r_o), \quad (5.10)$$

where r_o is some (small) reference radius. Indeed, as k increases, the shape of the response becomes closer to the fundamental solution; this improvement vanishes, however, for $k > 10^3$. At first glance, the response to such a δ sequence (illustrated in Figure 5.3) resembles the fundamental solution. However, a comparison between the fundamental solution in Figure 5.4 and the “impulse” response in Figure 5.3, as well as the comparison in Figure 5.5, reveal a substantial difference between them.

Consequently, we replaced the δ function with a point sink; rather than estimating the singular δ function, we surround it with a “well” (a small circle around the center). At first, we imposed Neumann (prescribed flux) BC on the perimeter of the “well”, with the exception of a single point of prescribed head (Dirichlet) BC (in order to attain a unique solution); in practice, however, this was equivalent to assigning prescribed head ($h = 0$) on the inner circle, while maintaining the above prescribed flux on the outer boundary (thereby enforcing the same “unit” flux into the inner circle as well). Since all random field generators generate fields on square grids, we had no choice but to translate the above BC to a square region. Consequently, we prescribe $q = 0.25 L$ on each side (L being the side length). A comparison

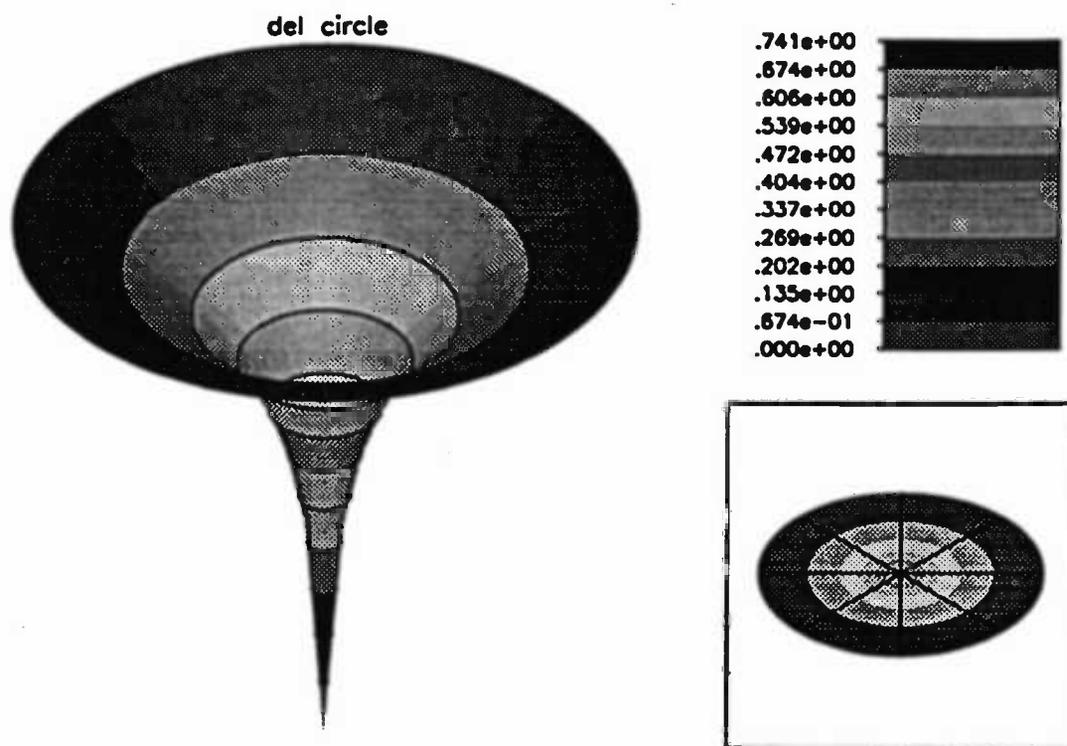


Figure 5.3: Head response to a “delta sequence” as a delta function in a circular domain.

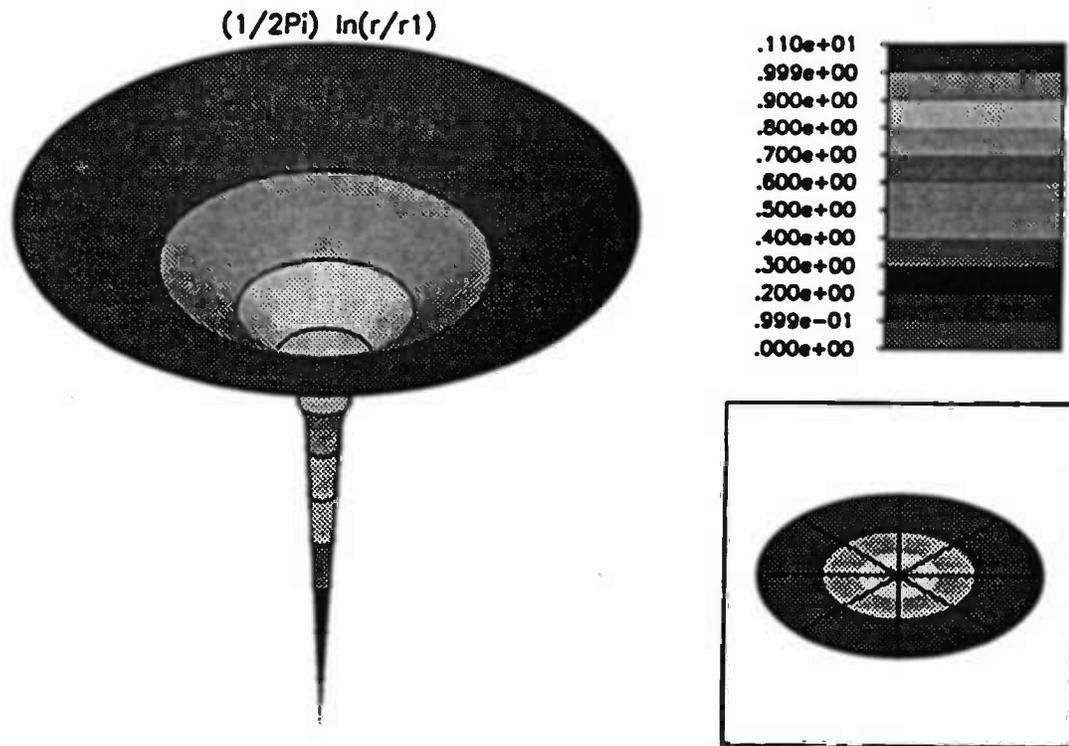


Figure 5.4: The fundamental solution in a circular domain.

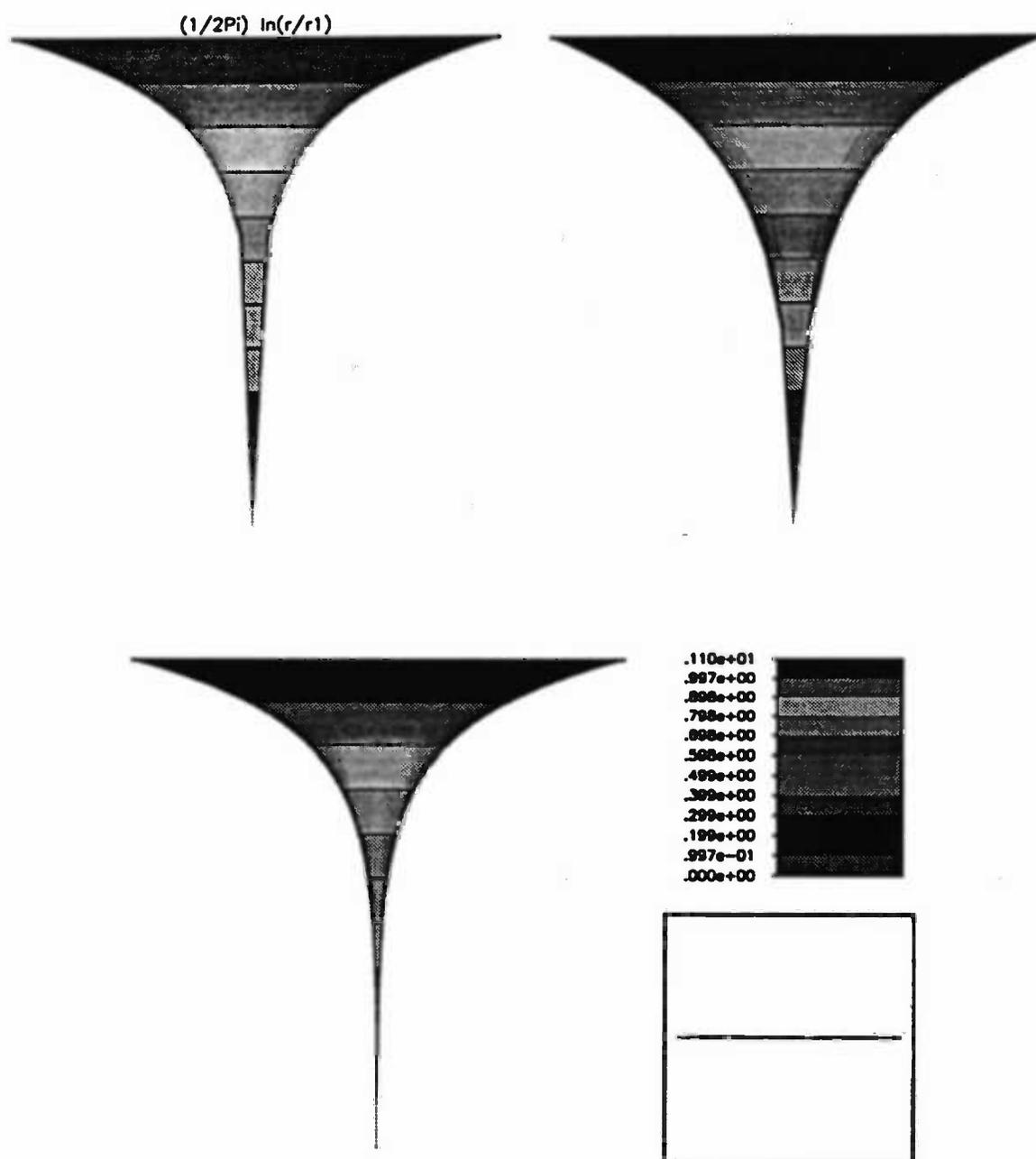


Figure 5.5: Comparison between the fundamental solution (upper left), the simulated impulse response function (upper right), and the response to a unit sink (below) with an inner radius of 0.001 units.

between the responses to two different BC on the “well” (prescribed flux with a single Dirichlet point and uniform Dirichlet BC on the inner circle) indicates that for both BC the response is close enough to the fundamental solution, as illustrated in Figure 5.7.

5.2 Effective Hydraulic Conductivity and Ensemble Statistics Under Radially Converging Mean Flow

To investigate the existence of effective hydraulic conductivities under a radially converging mean flow in a bounded domain, we performed high-resolution Monte Carlo simulations of 2-D flow to a point sink in a statistically homogeneous, isotropic, log normal $K(\mathbf{x})$ field with an exponential covariance. We generated 500 random realizations of $K(\mathbf{x})$ with a sequential simulator (GCOSIM-3D) developed by Gómez-Hernández (1991; discussed in the previous chapter). Point values of $K(\mathbf{x})$ were generated at the centroids of unit square blocks that form a 64×64 grid and assigned without change to the corresponding blocks. The flow equation was solved with PLTMG Version 6.1 (Bank 1990), described in Section 5.1. The point sink was approximated by a circle of radius 0.01 units at the center of the grid (designated as the origin of coordinates) and treated as a Dirichlet boundary with zero head. A total flow rate of unity was assigned along the outer boundary by proportioning it equally between all boundary segments. Figure 5.6 illustrates the adaptive grid refinement and solution for a uniform medium with $K \equiv 1$, starting from a coarse initial grid of 6×6 square blocks and ending with a grid of 800 triangles that is much finer near the sink than farther from it.

Figure 5.7 compares a corresponding numerical result on our standard grid

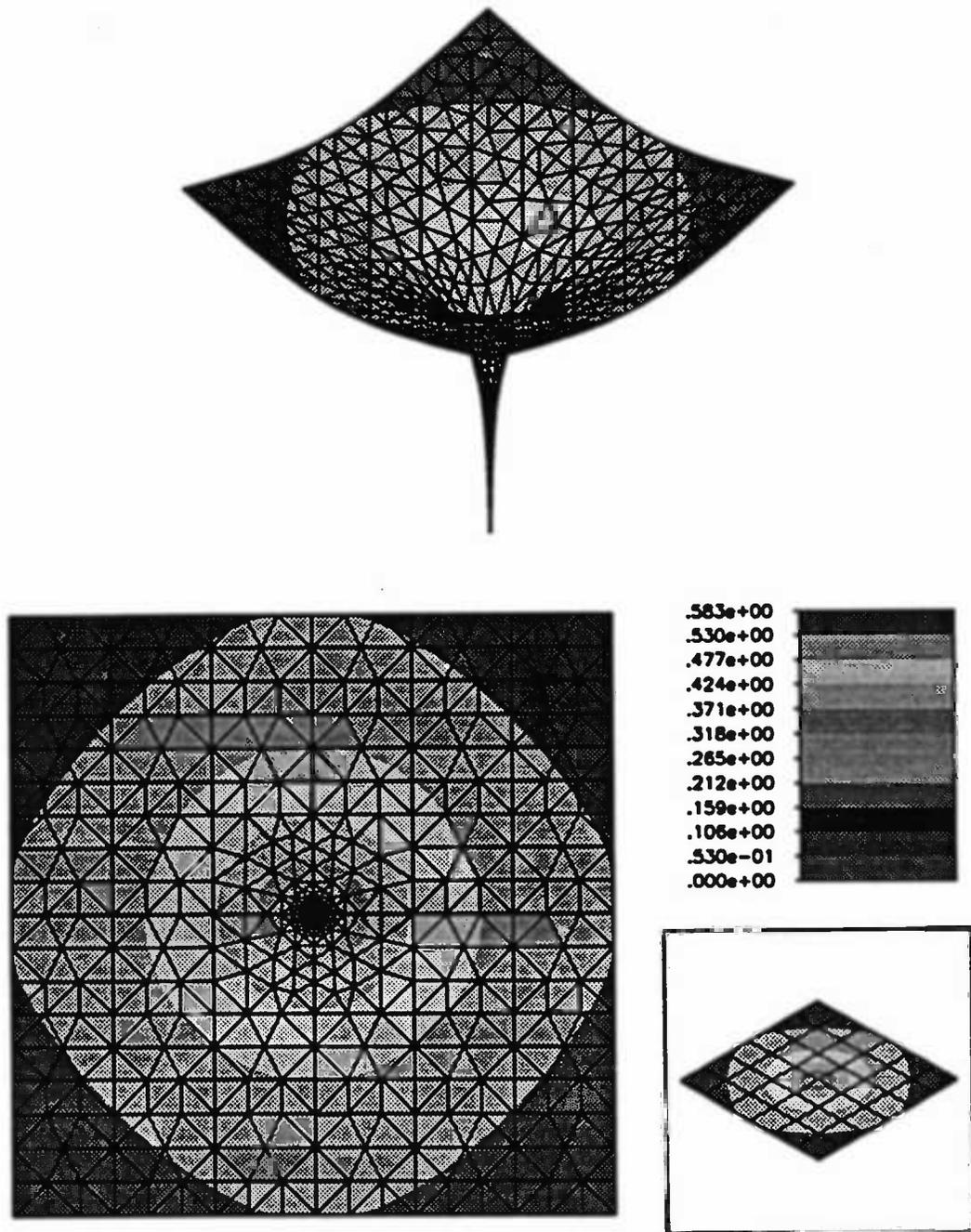


Figure 5.6: Illustration of adaptive grid refinement and 2-D point sink solution for $K \equiv 1$.

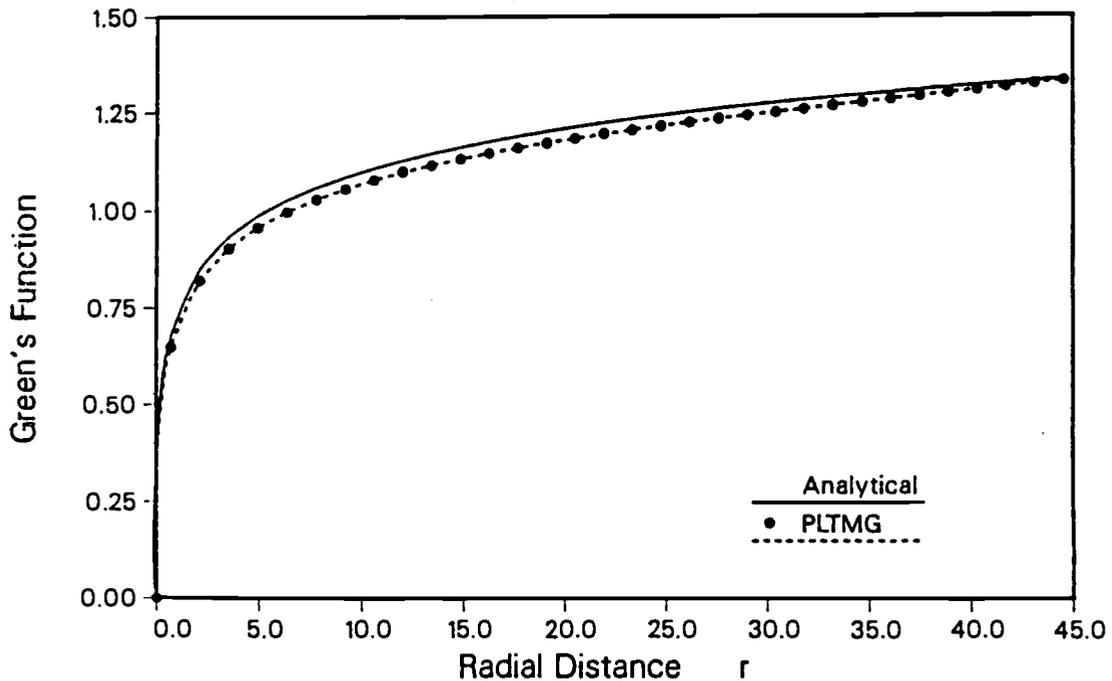


Figure 5.7: Comparison of analytical and numerical 2-D point sink solutions for $K \equiv 1$.

(consisting initially of 64×64 unit squares which, in virtually all cases, were subdivided adaptively into 65,760 triangles) with the analytical Green's function $G_0 = 1/2\pi \ln(r/0.01)$, where r is radial distance from the center. A comparison between the numerical solution and the fundamental solution is strikingly good. The slight discrepancy between the analytical and numerical solutions in the figure is an artifact of our outer boundary shape and disappears when this boundary is replaced by a circle (even though the analytical solution corresponds to an infinite rather than a finite domain). Table 1 shows the number of triangles, nt , the number of vertices,

nv , and the number of internal boundaries, nb , as well as the CPU time, and memory in mega-words (1 Mw = 8 Mb) required, on the Cray Y-MP at the San-Diego supercomputing center (SDSC), for homogeneous fields of different sizes.

| PLTMG: 2-D flow on CRAY Y-MP, $\sigma_{\ln K}^2 = 1.0$ | | | | | |
|--|-----------|-----------|-----------|------------------|-----------------|
| <i>Grid</i> | <i>nt</i> | <i>nv</i> | <i>nb</i> | <i>CPU (sec)</i> | <i>Mem (Mw)</i> |
| 20 × 20 | 6632 | 3440 | 248 | 12 | 2.2 |
| 40 × 40 | 25824 | 13156 | 488 | 54 | 4.7 |
| 60 × 60 | 57824 | 29276 | 728 | 150 | 8.5 |
| 128 × 128 | 262376 | 131960 | 1544 | 1112 | 30.3 |

This implies that

- The number of triangles (nt) needed to satisfy a modest accuracy criterion is by at least eight folds larger than the commonly used finite element discretisation (e.g., a grid of 20×20 requires 6632 triangles, although they are, naturally, not uniformly distributed).
- Due to limited memory allocation, the largest field size we can “afford” on the Cray Y-MP (at SDSC) is 128×128 . Since each simulation requires about 20 min CPU time on the Cray (and almost an hour CPU time on the Convex C-240⁴), this significantly limits both the field size and the number of MCS.

Figure 5.8 depicts a single realization of the random head in the domain, which we take to represent the random Green’s function $\mathcal{G}(\mathbf{x}', 0)$, as calculated with PLTMG for a variance $\sigma_Y^2 = 1$ and an integral scale $\lambda = 2$ (units) of $Y(\mathbf{x})$, while Figure 5.9 illustrates a single realization of $\mathcal{G}(\mathbf{x}', 0)$ for $\sigma_Y^2 = 4$ and $\lambda = 5$. The sample variance⁵ of $\mathcal{G}(\mathbf{x}', 0)$ at a point about midway between the center of the grid

⁴A minisupercomputer at the University of Arizona, with up to 1 gigabyte of memory, but limited (allowed) CPU time.

⁵Denoted by $\{\cdot\}$ in the following figures.

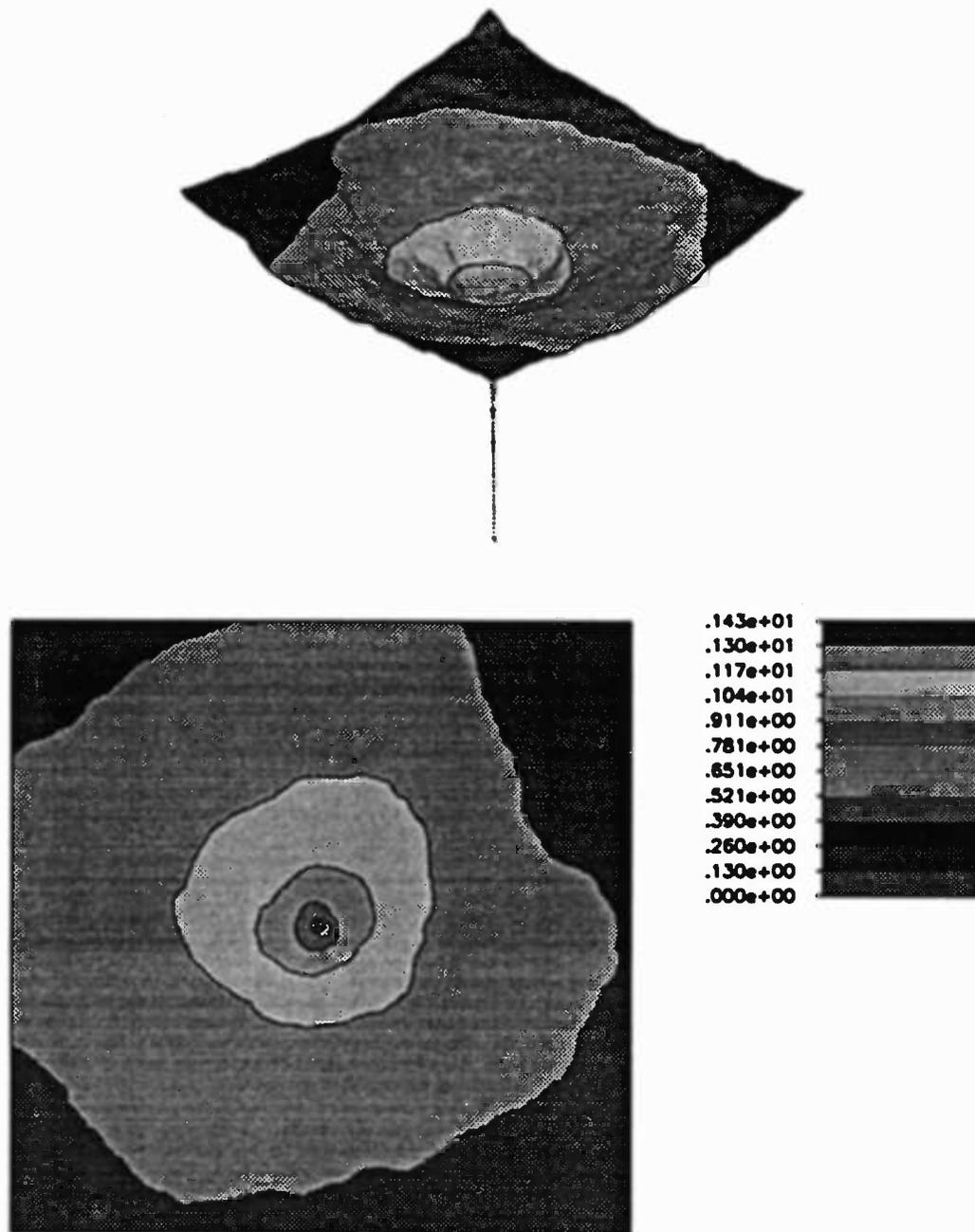


Figure 5.8: 2-D point sink solution for $\sigma_y^2 = 1$ and $\lambda = 2$.

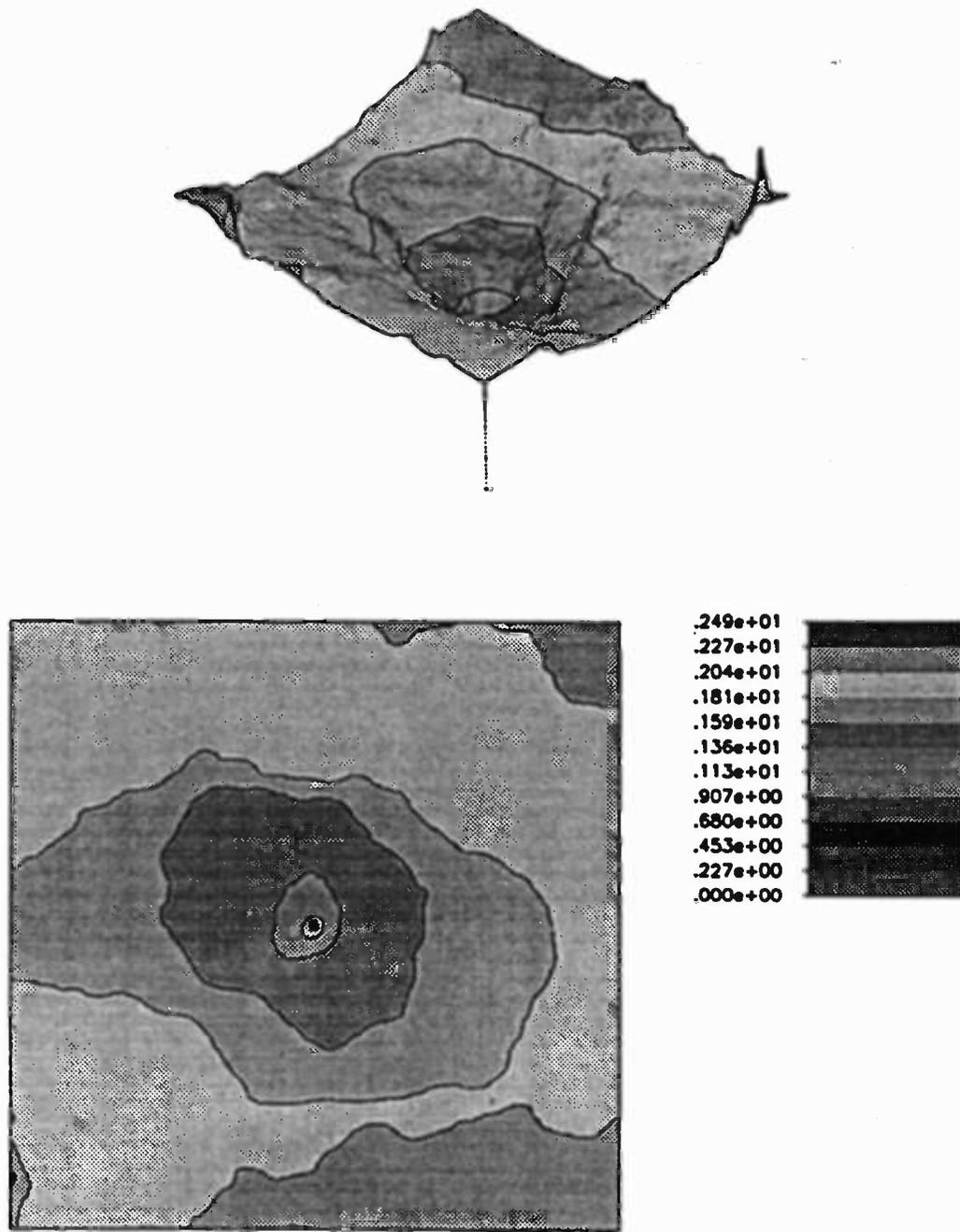


Figure 5.9: 2-D point sink solution for $\sigma_Y^2 = 4$ and $\lambda = 5$.

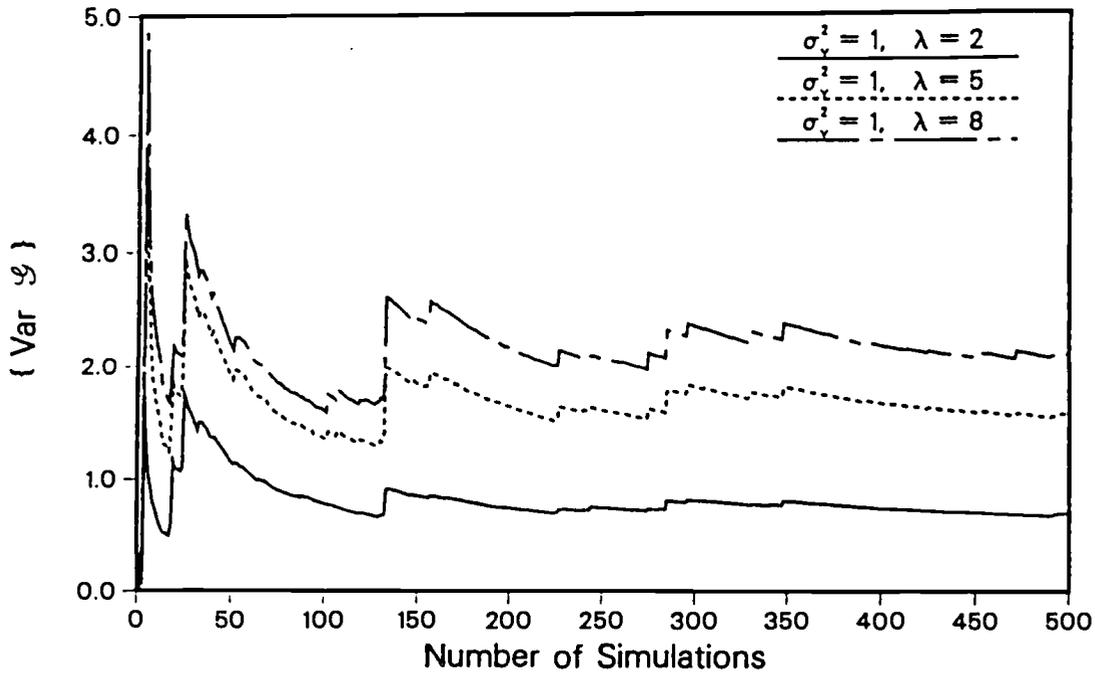


Figure 5.10: Sample variance of \mathcal{G} versus number of simulations for $\sigma_{\gamma}^2 = 1$ and $\lambda = 2, 5, 8$.

and the outer boundary is plotted as a function of the number of simulations for $\sigma_{\gamma}^2 = 1$ and $\lambda = 2, 5, 8$ in Figure 5.10, and for $\sigma_{\gamma}^2 = 4$ in Figure 5.11. It is seen in these figures that the sample variance becomes quite stable as the number of simulations approaches 500 when $\sigma_{\gamma}^2 = 1$, but not so for $\lambda = 5, 8$ when $\sigma_{\gamma}^2 = 4$. This is typical of most other points in the grid. Though it would be desirable to continue the simulation, for reasons of expediency we adopt 500 as the sample size for all statistics presented below.

Figure 5.12 illustrates how the sample mean $\{\mathcal{G}(\mathbf{x}', 0)\}$ of $\mathcal{G}(\mathbf{x}', 0)$, averaged along the four diagonals extending from the center of the grid to its corners, varies with the dimensionless distance r/λ relative to $G_g(\mathbf{x}', 0)$ for $\sigma_{\gamma}^2 = 1, 4$ and $\lambda = 1,$

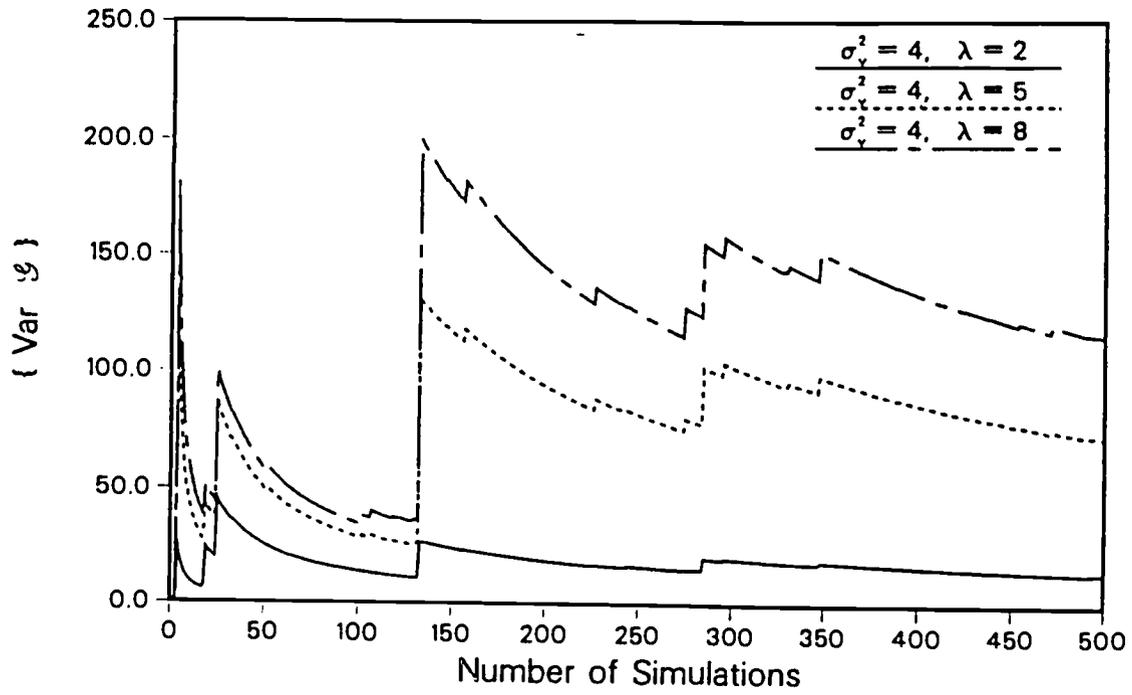


Figure 5.11: Sample variance of \mathcal{G} versus number of simulations for $\sigma_\gamma^2 = 4$ and $\lambda = 2, 5, 8$.

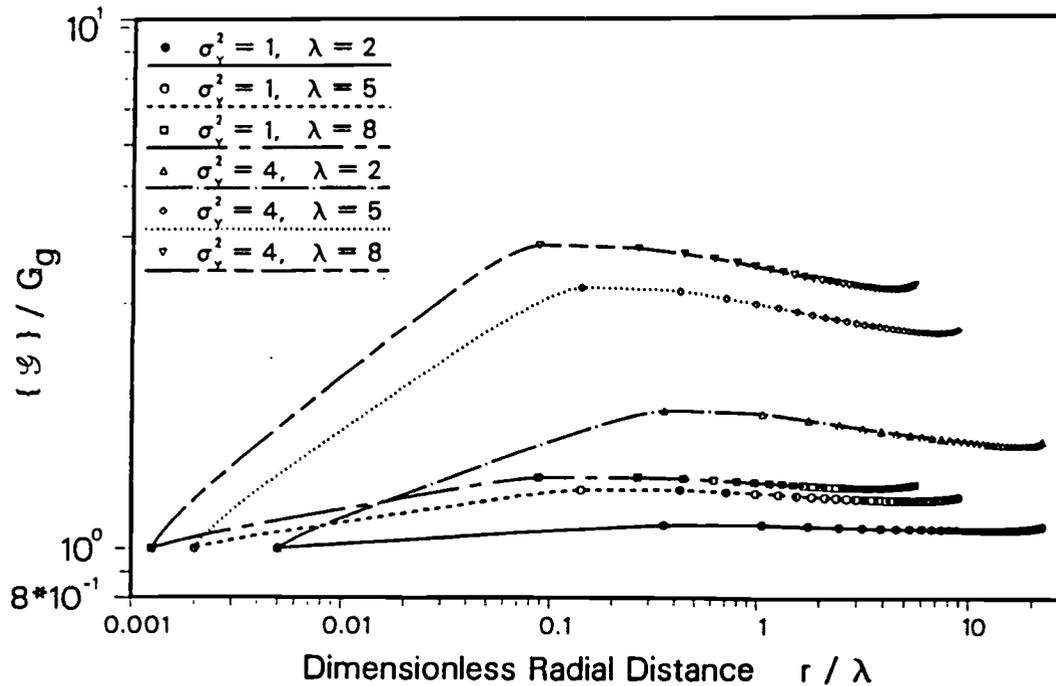


Figure 5.12: Ratio between sample mean of \mathcal{G} and G_g versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

5, 8, where $K_g G_g(\mathbf{x}', 0) = G_0(\mathbf{x}', 0)$. As we consistently set $\langle Y \rangle = 0$, our results correspond to $K_g = 1$ and hence $G_g = G_0$. We see that, in general, $\{\mathcal{G}(\mathbf{x}', 0)\}$ exceeds $G_g(\mathbf{x}', 0)$ everywhere except at the sink where both are artificially set equal to zero. The ratio between $\{\mathcal{G}\}$ and G_g climbs to a maximum within a short dimensionless distance $r/\lambda \ll 1$ and then decreases gradually to stabilize at a somewhat lower value (only to start rising again as the outer boundary is approached). The ratio generally increases as σ_Y^2 and λ become larger. When $\lambda = 8$, it exceeds 1 by a factor of at most 1.3 for $\sigma_Y^2 = 1$ and a factor of at most 3 for $\sigma_Y^2 = 4$. Figure 5.13 exhibits a similar pattern of behavior for the ratio $\{\mathcal{G}(\mathbf{x}', 0)\}/G_\kappa(\mathbf{x}', 0)$ where $\kappa G_\kappa(\mathbf{x}', 0) = G_0(\mathbf{x}', 0)$. However, $\{\mathcal{G}\}/G_\kappa$ grows more rapidly with σ_Y^2 than does $\{\mathcal{G}\}/G_g$ due to

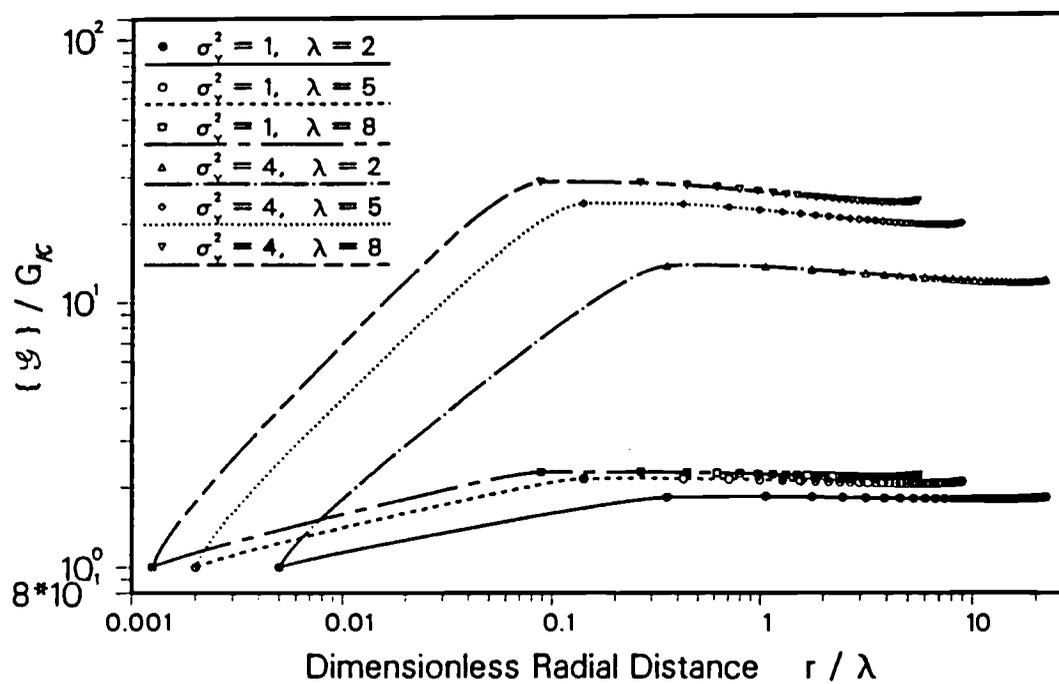


Figure 5.13: Ratio between sample mean of \mathcal{G} and G_κ versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

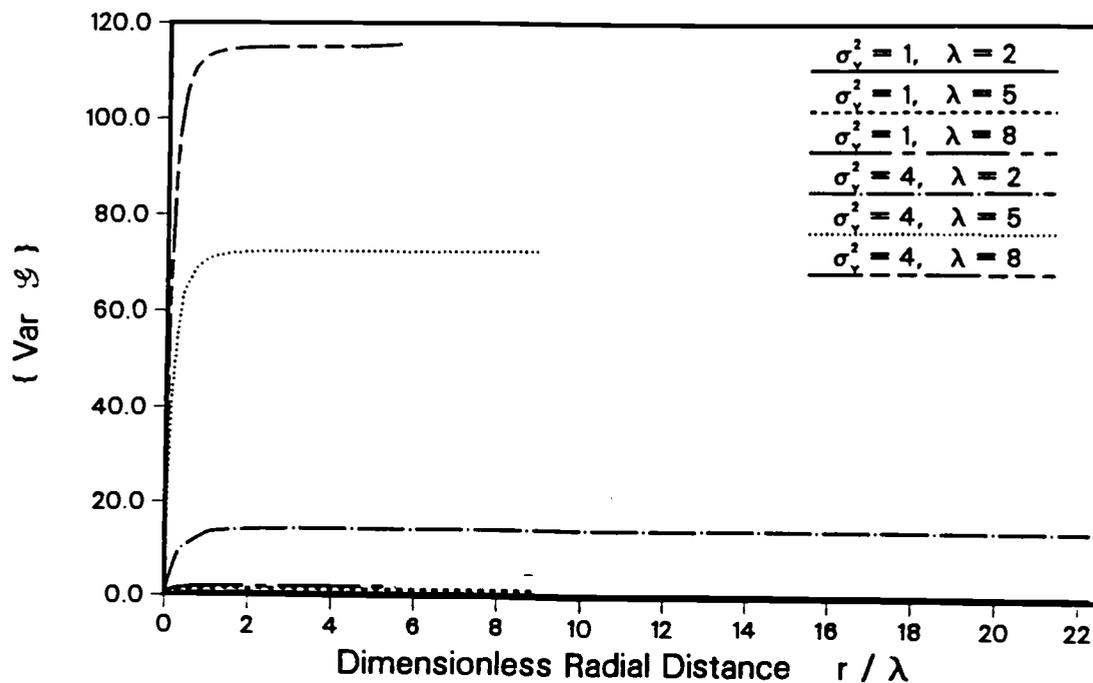


Figure 5.14: Sample variance of \mathcal{G} versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

the exponential dependence of κ on this variance. The sample variance of $\mathcal{G}(\mathbf{x}', 0)$ increases with σ_Y^2 and λ everywhere except at the sink where it is zero (Figure 5.14). It increases rapidly with dimensionless distance and tends to stabilize within one (for $\sigma_Y^2 = 1$) or two (for $\sigma_Y^2 = 4$) integral scales from the center.

Figures 5.15 and 5.16 show how the sample variance of the specific flux varies with actual and dimensionless radial distance, respectively, for the former range of σ_Y^2 and λ values. According to Figure 5.15, the variance at a given radius increases with σ_Y^2 but decreases with λ . At distances exceeding about one integral scale from the source, the variance decreases at a rate more or less proportional to r^{-2} (Figure

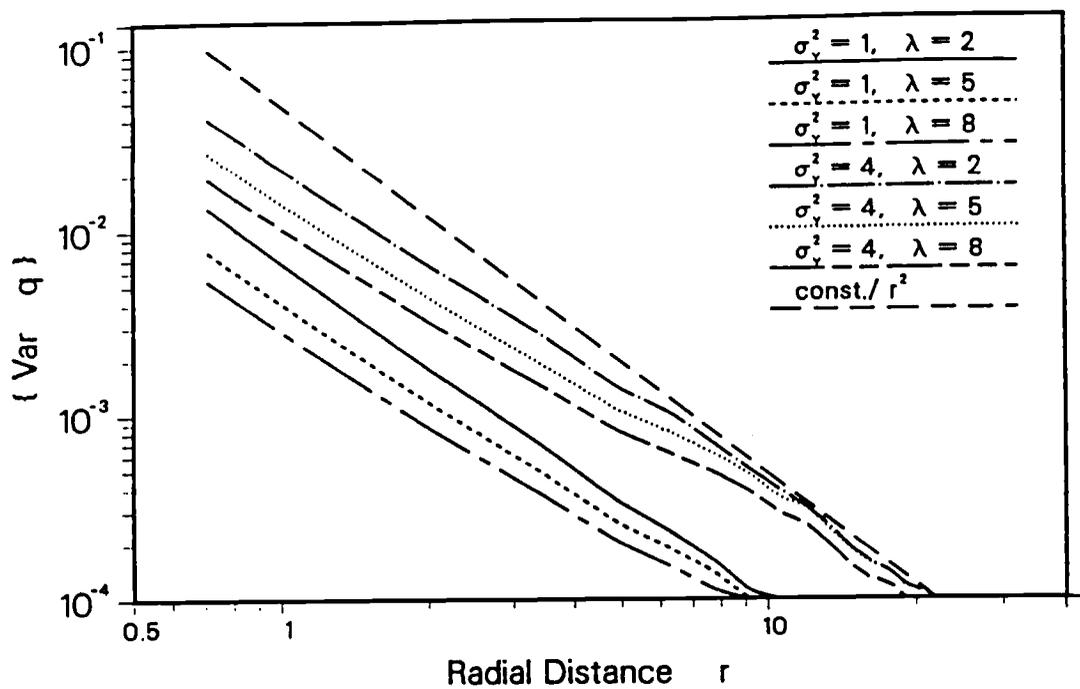


Figure 5.15: Sample variance of specific flux q versus radial distance r for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

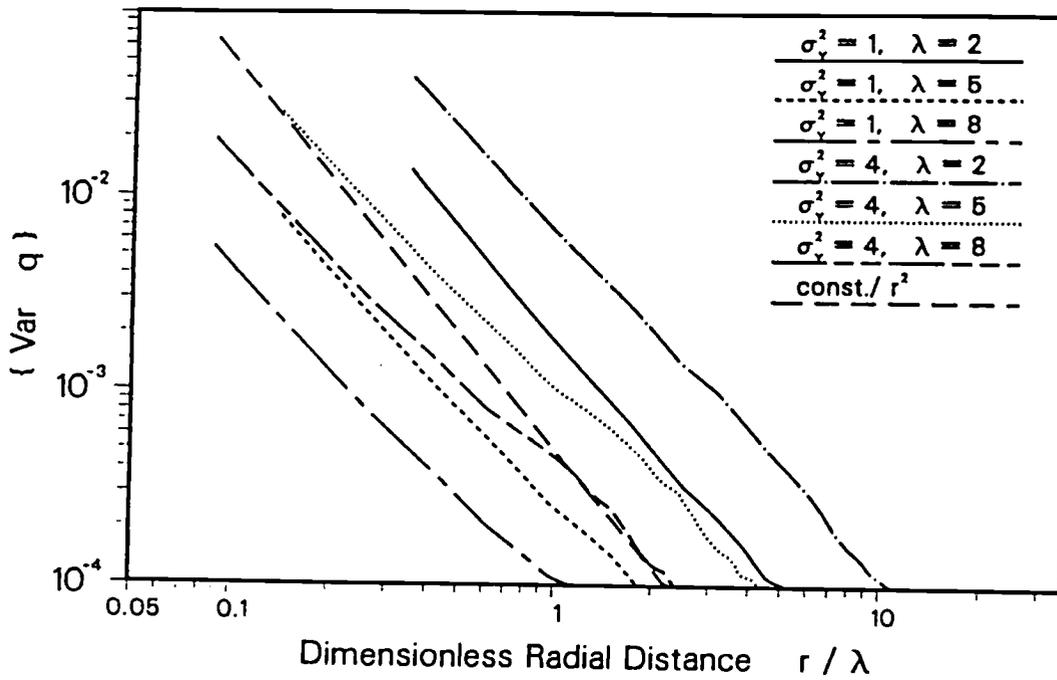


Figure 5.16: Sample variance of specific flux q versus dimensionless radial distance r/λ for $\sigma_y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

12) as proposed on the basis of theoretical considerations by Matheron (1967) and confirmed analytically by Naff (1991). Closer to the source, the rate of variation with distance is somewhat slower. This notwithstanding, our results support Naff in his conclusion that if an effective hydraulic conductivity can be defined for mean radial flow in two dimensions, then (*ibid*, p. 314) “the best estimates of the effective hydraulic conductivity will be obtained from observations at some distance from the well bore.”

Next, we examine the question whether an effective hydraulic conductivity can in fact be defined for such flow and, if so, what is its value? Dagan (1989, equation 3.4.49; see first chapter) has shown that, for small σ_Y^2 , the effect of radial flow is to reduce K_e below K_g by a quantity proportional to $\sigma_Y^2 \lambda^2 / r^2$. This implies that far from this sink, $K_e = K_g$. We can reach a similar conclusion directly, for arbitrarily large σ_Y^2 , by recognizing that sufficiently far from the point source the mean flow tends to be uniform and hence K_e must approach K_g . Dagan further suggested that close to the sink, the flow is radially symmetric, implying (*ibid*, equation 5.4.3) that K_e is equal to the harmonic mean $K_h = K_g \exp(-\sigma_Y^2/2)$ regardless of how $K(\mathbf{x})$ varies in real or probability space. All of this is confirmed and amplified by our numerical results.

In Figure 5.17 we plot the ratio K_e/K_g as obtained from our 500 simulations, and averaged further over the four diagonals extending from the origin to the four corners of the grid, versus r/λ for the previous range of σ_Y^2 and λ values. The results for $\lambda = 2$ demonstrate that K_e/K_g is virtually 1 at radial distances exceeding 2–3 integral scales from the sink and the outer boundary when $\sigma_Y^2 = 1$, but these distances increase to $r/\lambda \geq 4$ from the sink and $r/\lambda \geq 10$ from the outer boundary

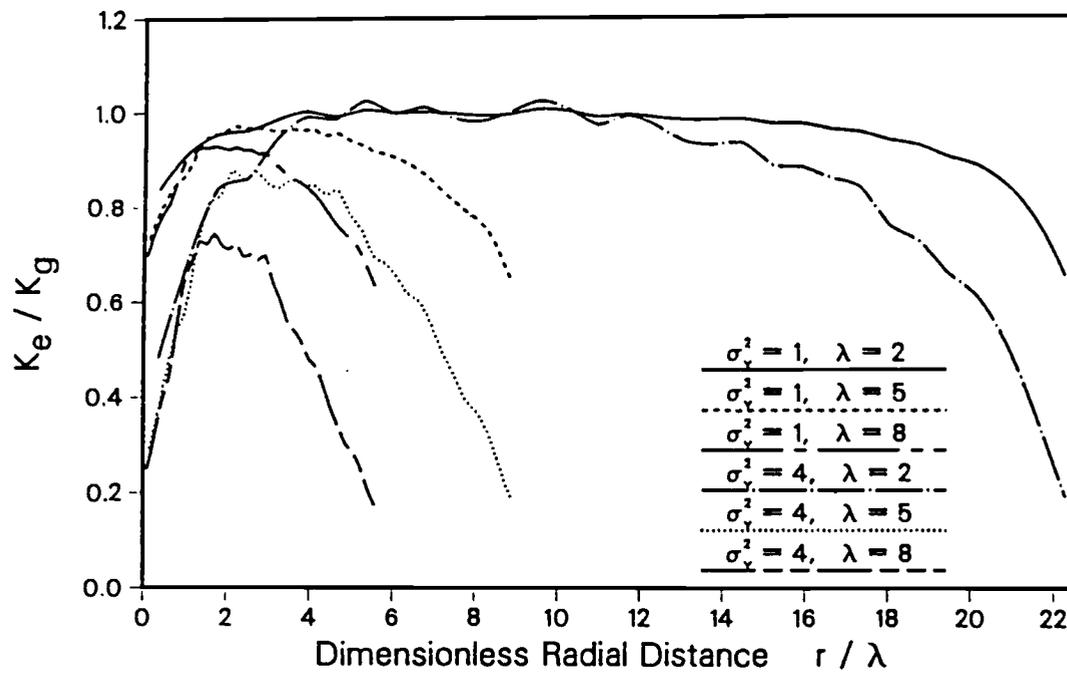


Figure 5.17: K_e / K_g versus dimensionless radial distance r / λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

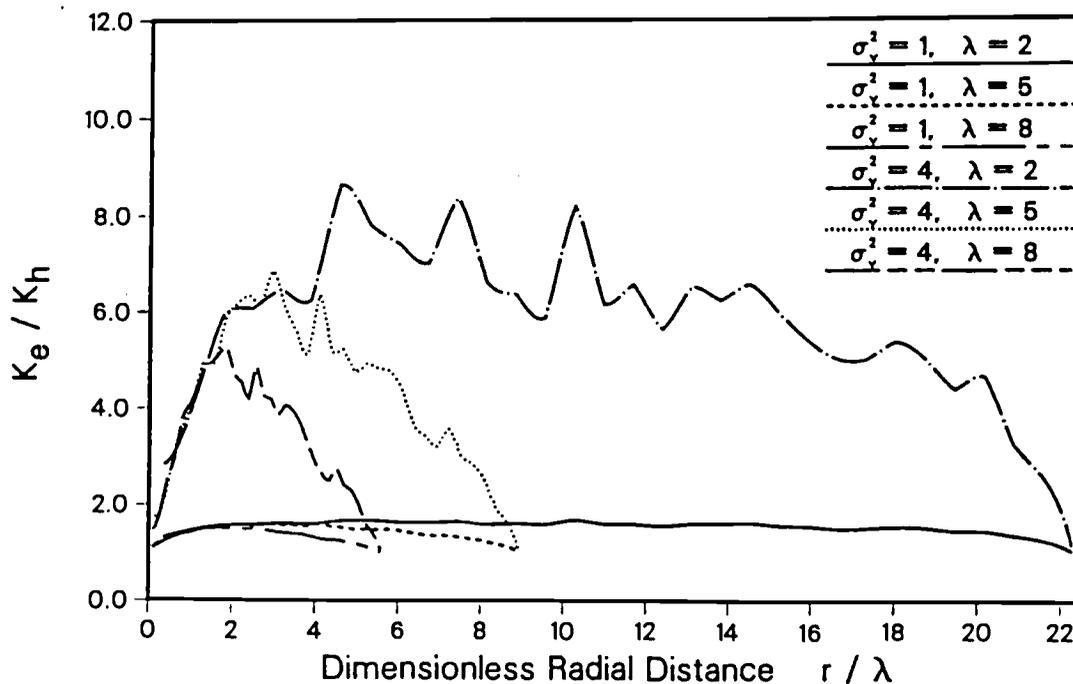


Figure 5.18: K_e/K_h versus dimensionless radial distance r/λ for $\sigma_Y^2 = 1, 4$ and $\lambda = 2, 5, 8$.

when $\sigma_Y^2 = 4$. It seems that in the cases of $\lambda = 5$ and 8 the situation is similar but masked to some extent by a more pronounced influence of the outer boundary. We take this to confirm the above theoretical prediction that, in an infinite 2-D medium, K_e approaches K_g as radial distance increases when σ_Y^2 is small, and to demonstrate further that this prediction is valid for σ_Y^2 at least as large as 4 .

Our numerical results in Figures 5.8 and 5.9 support Dagan's (1989) suggestion that, close to the sink, the flow is radially symmetric. In Figure 5.18 we plot the ratio K_e/K_h as a function of r/λ for the previous values of σ_Y^2 and λ . The plot clearly confirms Dagan's (1989) theoretical prediction (based on this assertion) that, regardless of σ_Y^2 and λ , K_e tends toward the harmonic mean of $K(\mathbf{x})$ as r/λ becomes

small; a similar result was obtained for small σ_Y^2 by Desbarats (1992b, Fig. 11; see first chapter). Figure 5.18 further demonstrates that the same is true near the outer prescribed flux boundary. Our findings are contrary to Naff's (1991, p. 313; see first chapter) conclusion that "when the flow tends to emulate a two-dimensional (infinite) system, then Matheron's (1967) asymptotic results for radial flow are generally realized: the apparent effective hydraulic conductivity becomes the arithmetic mean near the well and the harmonic mean distant from the well."

Our unconditional Monte Carlo results suggest that, in the 2-D case when the conditional variance of $Y(\mathbf{x})$ does not exceed 1, it may be appropriate to approximate $\mathbf{r}_\kappa(\mathbf{x})$ in (3.35) by

$$\mathbf{r}_\kappa(\mathbf{x}) \simeq \hat{\kappa}(\mathbf{x}) \nabla h_\kappa(\mathbf{x}') \quad (5.11)$$

for points \mathbf{x} located beyond a distance of two to three conditional integral scales (defined below) from point sources, and four such integral scales from linear boundary sources; when the conditional variance is 4, the same is true for \mathbf{x} located beyond at least four conditional integral scales from point sources and ten such scales from linear boundary sources. Here $\hat{\kappa}(\mathbf{x})$ is the "residual hydraulic conductivity tensor"

$$\hat{\kappa}(\mathbf{x}) \equiv \int_{\Omega} \mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}') d\mathbf{x}' \quad (5.12)$$

and $\mathbf{a}_\kappa(\mathbf{x}, \mathbf{x}')$ is defined in (3.42). By taking $Y'(\mathbf{x})$ to be locally homogeneous within a distance of a few conditional integral scales from \mathbf{x} , one may further approximate the residual conductivity, in a manner similar to (3.48)-(3.49) and (3.70), via

$$\hat{\kappa}(\mathbf{x}) \simeq \kappa(\mathbf{x}) \mathbf{I} - \hat{\mathbf{K}}(\mathbf{x}) \quad (5.13)$$

where $\hat{\mathbf{K}}(\mathbf{x})$ is a symmetric, second-rank tensor with principal components

$$\hat{K}_i \simeq K_g(\mathbf{x}) \exp \left[\frac{\sigma_Y^2(\mathbf{x})}{2} - F_i(\mathbf{x}) \right] \quad i = 1, 2, 3 \quad (5.14)$$

and $F_i(\mathbf{x})$ is a local equivalent of (3.68). The term “conditional integral scale” refers here to the integral scale of the log estimation errors $Y'(\mathbf{x})$ about the conditional mean of $Y(\mathbf{x})$. Since this integral scale can be reduced through the judicious acquisition of data, it follows that one has in principle control over the distance beyond which the localized expressions (5.11)-(5.14) are valid.

5.3 Verification of the Weak Approximation

To verify numerically the weak approximation (Chapter 2), we analyze further the results of our two-dimensional Monte Carlo simulations. Ideally, we would like to examine directly the conditions under which the approximation

$$\begin{aligned} \int_{\Omega_\infty} \langle K'(\mathbf{x})K'(\mathbf{x}')\nabla\nabla_{\mathbf{x}'}^T\mathcal{G}(\mathbf{x}',\mathbf{x})\rangle_\kappa\nabla_{\mathbf{x}'}h_\kappa(\mathbf{x}')d\mathbf{x}' \\ \simeq \int_{\Omega_\infty} \langle K'(\mathbf{x})K'(\mathbf{x}')\rangle_\kappa\nabla\nabla_{\mathbf{x}'}^T\langle\mathcal{G}(\mathbf{x}',\mathbf{x})\rangle_\kappa\nabla_{\mathbf{x}'}h_\kappa(\mathbf{x}')d\mathbf{x}' \end{aligned} \quad (5.15)$$

may or may not hold. As this has so far been beyond our computational means, we consider instead the related weak approximations

$$\int_{\Omega_\infty} \langle K'(\mathbf{x})K'(\mathbf{x}')\mathcal{G}(\mathbf{x}',\mathbf{x})\rangle_\kappa d\mathbf{x}' \simeq \int_{\Omega_\infty} \langle K'(\mathbf{x})K'(\mathbf{x}')\rangle_\kappa\langle\mathcal{G}(\mathbf{x}',\mathbf{x})\rangle_\kappa d\mathbf{x}' \quad (5.16)$$

and

$$\int_{\Omega_\infty} \langle K'(\mathbf{x})K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',\mathbf{x})\rangle_\kappa d\mathbf{x}' \simeq \int_{\Omega_\infty} \langle K'(\mathbf{x})K'(\mathbf{x}')\rangle_\kappa\langle\mathcal{G}_r(\mathbf{x}',\mathbf{x})\rangle_\kappa d\mathbf{x}' \quad (5.17)$$

evaluated at $\mathbf{x} = 0$, where $\mathcal{G}_r(\mathbf{x}',\mathbf{x})$ is the partial derivative of $\mathcal{G}(\mathbf{x}',\mathbf{x})$ with respect to the radius r in the \mathbf{x}' system of coordinates.

We start with (5.16) by investigating the manner in which the ratio between $\langle K'(0)K'(\mathbf{x}')\mathcal{G}'(\mathbf{x}',0)\rangle = \langle K'(0)K'(\mathbf{x}')\mathcal{G}(\mathbf{x}',0)\rangle - \langle K'(0)K'(\mathbf{x}')\rangle\langle\mathcal{G}(\mathbf{x}',0)\rangle$ and

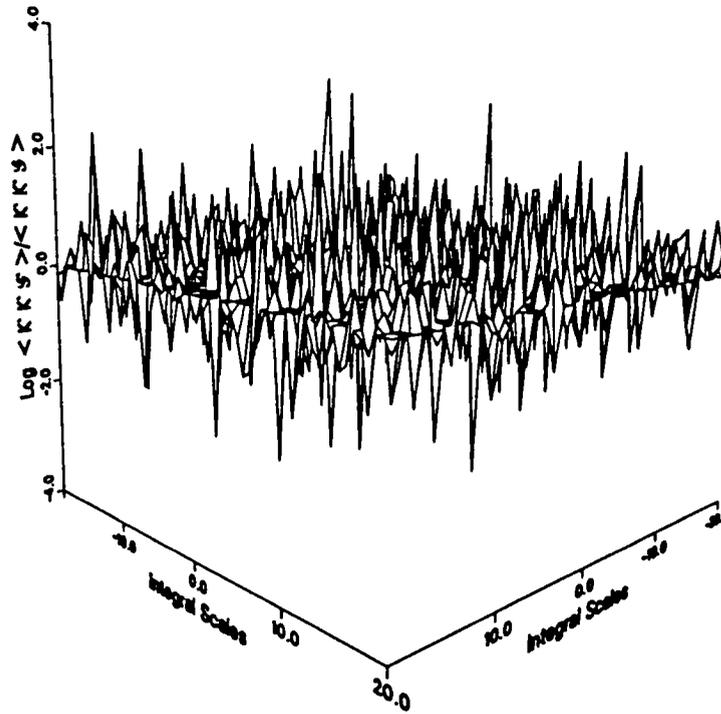


Figure 5.19: Spatial variation of ratio between sample means $\{K'(0)K'(\chi)\mathcal{G}'(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}(\chi, 0)\}$ for $\sigma_Y^2 = 1$ and $\lambda = 2$.

$\langle K'(0)K'(\mathbf{x}')\mathcal{G}(\mathbf{x}', 0) \rangle$ behaves at individual points \mathbf{x}' . This should enable us to comment on the extent to which the strong approximation (analogous to one often made in the stochastic groundwater literature)

$$\langle K'(\mathbf{x})K'(\mathbf{x}')\mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \simeq \langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_{\kappa} \langle \mathcal{G}(\mathbf{x}', \mathbf{x}) \rangle_{\kappa} \quad (5.18)$$

is or is not valid for various σ_Y^2 and λ . Let $\{K'(0)K'(\mathbf{x}')\mathcal{G}'(\mathbf{x}', 0)\}$ be the sample mean of $K'(0)K'(\mathbf{x}')\mathcal{G}'(\mathbf{x}', 0)$, and let $\{K'(0)K'(\mathbf{x}')\mathcal{G}(\mathbf{x}', 0)\}$ be the sample mean of $K'(0)K'(\mathbf{x}')\mathcal{G}(\mathbf{x}', 0)$. Figure 5.19 shows how the ratio between $\{K'(0)K'(\mathbf{x}')\mathcal{G}'(\mathbf{x}', 0)\}$ and $\{K'(0)K'(\mathbf{x}')\mathcal{G}(\mathbf{x}', 0)\}$ varies spatially after 500 simulations when $\sigma_Y^2 = 1$ and $\lambda = 2$. We see that this ratio oscillates with a very high frequency and amplitude (note the logarithmic scale). Though these oscillations may be partly due to numerical errors,

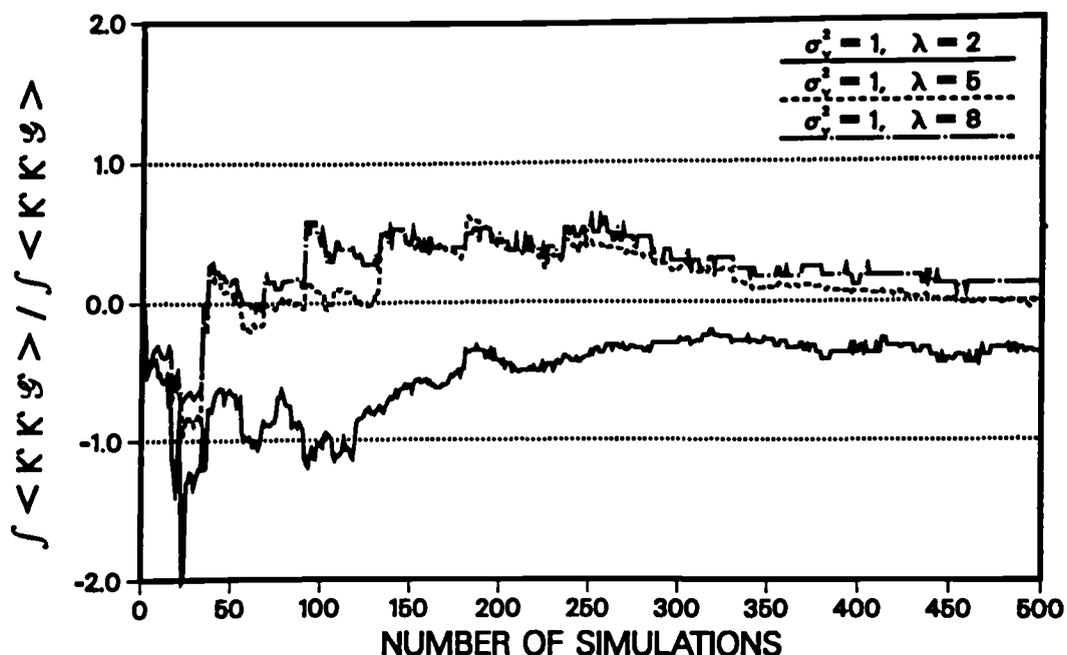


Figure 5.20: Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}(\chi, 0)\}$ versus number of simulations for $\sigma_Y^2 = 1$ and $\lambda = 2, 5, 8$.

they nevertheless raise a serious question about the validity of the local (strong) approximation (5.18).

In contrast, we show in Figure 5.20 how the ratio between the spatial integrals of $\{K'(0)K'(\mathbf{x}')\mathcal{G}'(\mathbf{x}', 0)\}$ and $\{K'(0)K'(\mathbf{x}')\mathcal{G}(\mathbf{x}', 0)\}$, calculated by the trapezoidal rule over the 64×64 grid of unit squares, varies with the number of simulations and λ for $\sigma_Y^2 = 1$, and in Figure 5.21 for $\sigma_Y^2 = 4$. We see that, in all cases, the magnitude of this ratio becomes less than 1 after at most 350 simulations and appears to be approaching very small values (zero?) asymptotically. At 500 simulations, the maximum value is less for $\sigma_Y^2 = 1$ than for $\sigma_Y^2 = 4$. We therefore feel much more comfortable about the weak approximation (5.16) than we do about the strong

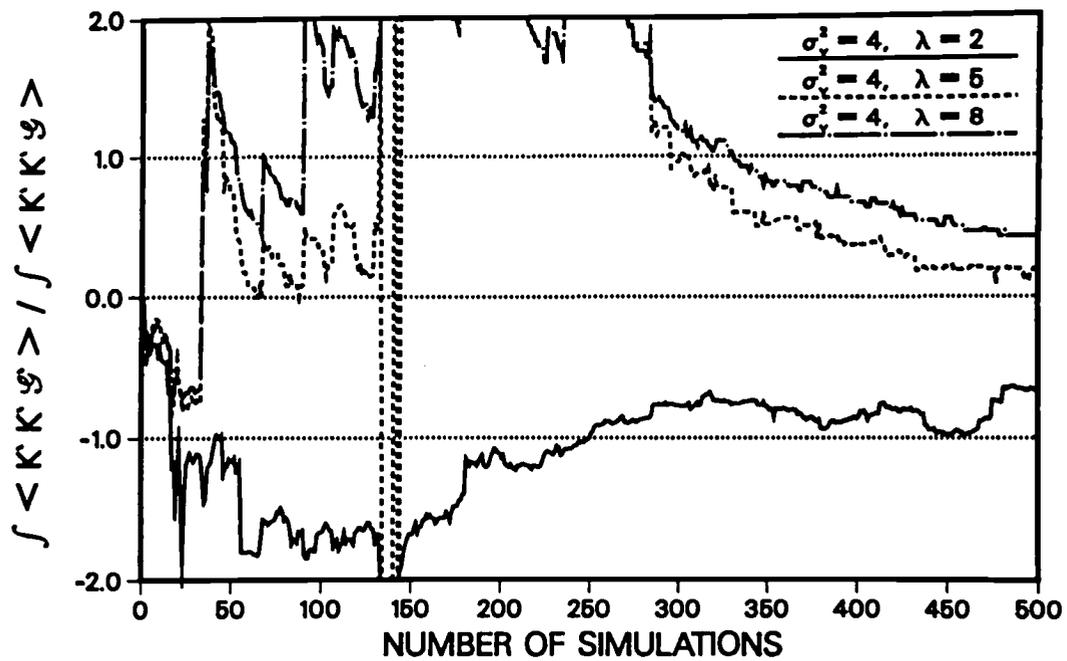


Figure 5.21: Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}(\chi, 0)\}$ versus number of simulations for $\sigma_y^2 = 4$ and $\lambda = 2, 5, 8$.

approximation (5.18), and note that the former improves further as σ_Y^2 decreases.

We continue with (5.17) by investigating the manner in which the ratio between $\langle K'(0)K'(\mathbf{x}')\mathcal{G}'_r(\mathbf{x}',0) \rangle$ and $\langle K'(0)K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',0) \rangle$ behaves at individual points \mathbf{x}' . This should enable us to comment on the extent to which the strong approximation (analogous, again, to one often made in the stochastic groundwater literature)

$$\langle K'(\mathbf{x})K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',\mathbf{x}) \rangle_\kappa \simeq \langle K'(\mathbf{x})K'(\mathbf{x}') \rangle_\kappa \langle \mathcal{G}_r(\mathbf{x}',\mathbf{x}) \rangle_\kappa \quad (5.19)$$

is or is not valid for various σ_Y^2 and λ . Let $\{K'(0)K'(\mathbf{x}')\mathcal{G}'_r(\mathbf{x}',0)\}$ be the sample mean of $K'(0)K'(\mathbf{x}')\mathcal{G}'_r(\mathbf{x}',0)$ and let $\{K'(0)K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',0)\}$ be the sample mean of $K'(0)K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',0)$. Figure 5.22 shows how the ratio between $\{K'(0)K'(\mathbf{x}')\mathcal{G}'_r(\mathbf{x}',0)\}$ and $\{K'(0)K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',0)\}$ varies spatially after 500 simulations when $\sigma_Y^2 = 1$ and $\lambda = 2$. We see again that this ratio oscillates with a very high frequency and amplitude (note the logarithmic scale). Though these oscillations may, like before, be partly due to numerical errors, they nevertheless raise a serious question about the validity of the local (strong) approximation (5.19).

In contrast, we show in Figure 5.23 how the ratio between the spatial integrals of $\{K'(0)K'(\mathbf{x}')\mathcal{G}'_r(\mathbf{x}',0)\}$ and $\{K'(0)K'(\mathbf{x}')\mathcal{G}_r(\mathbf{x}',0)\}$, calculated by the trapezoidal rule over the 64×64 grid of unit squares, varies with the number of simulations and λ for $\sigma_Y^2 = 1$, and in Figure 5.24 for $\sigma_Y^2 = 4$. We see that, in all cases, the magnitude of this ratio becomes less than 1 after at most 450 simulations and appears to be approaching very small values (zero?) asymptotically. At 500 simulations, the maximum value is less for $\sigma_Y^2 = 1$ than for $\sigma_Y^2 = 4$. We therefore again feel much more comfortable about the weak approximation (5.17) than we do about the strong approximation (5.19), and note once more that the former improves further as σ_Y^2 decreases.

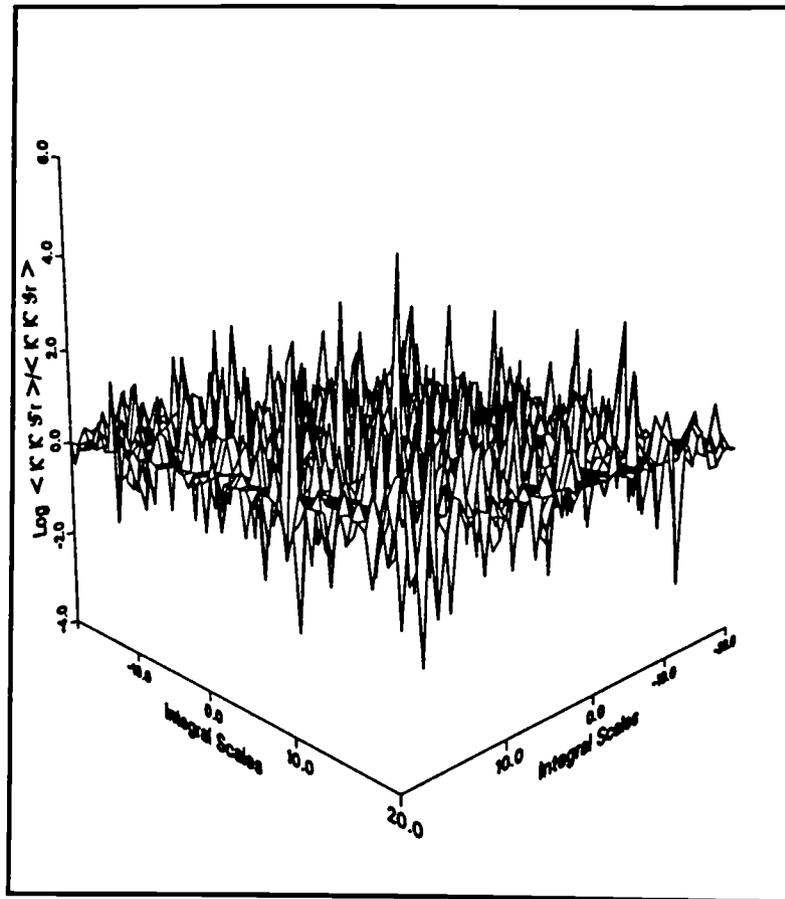


Figure 5.22: Spatial variation of ratio between sample means $\{K'(0)K'(\chi)\mathcal{G}'_r(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}'_r(0)\}$ for $\sigma_Y^2 = 1$ and $\lambda = 2$.

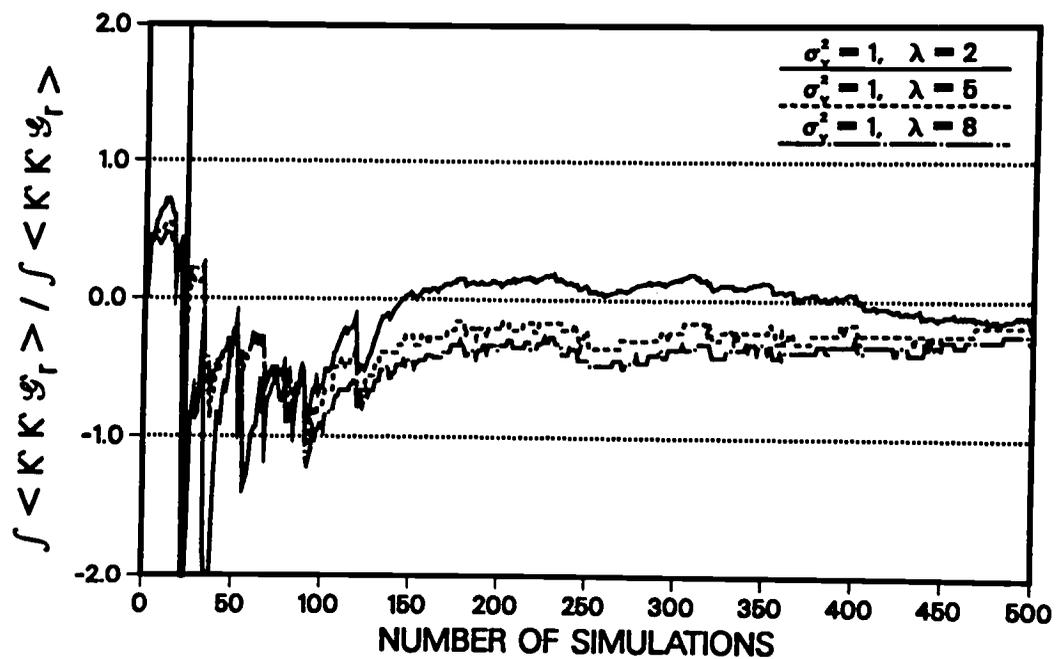


Figure 5.23: Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'_r(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}_r(\chi, 0)\}$ versus number of simulations for $\sigma_y^2 = 1$ and $\lambda = 2, 5, 8$.

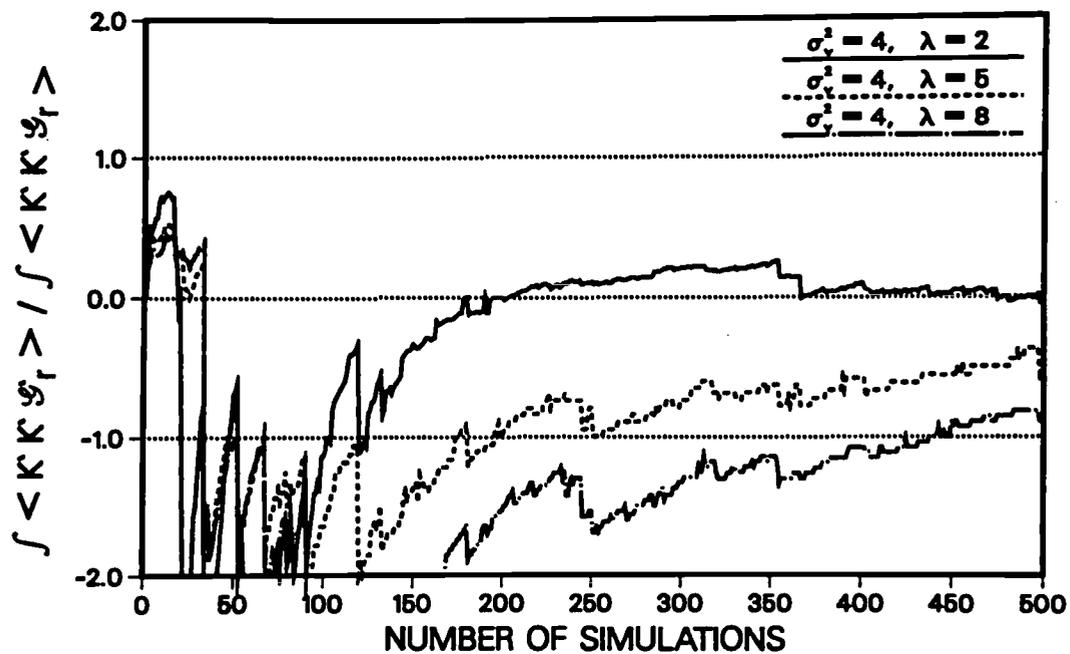


Figure 5.24: Ratio between spatial integrals of sample means $\{K'(0)K'(\chi)\mathcal{G}'_r(\chi, 0)\}$ and $\{K'(0)K'(\chi)\mathcal{G}_r(\chi, 0)\}$ versus number of simulations for $\sigma_y^2 = 4$ and $\lambda = 2, 5, 8$.

Based on the above results, we consider our weak approximation (3.58) to be more readily justifiable than the equivalent strong approximation (or analogues of the latter commonly found in the stochastic groundwater literature), and to improve with the quality of the estimate $\kappa(\mathbf{x})$ as the latter is conditioned on measurements of $K(\mathbf{x})$.

CHAPTER 6

CONCLUSIONS

Our analysis leads to the following major conclusions:

1. Starting from the premise that Darcy's law applies locally with a random hydraulic conductivity field $K(\mathbf{x})$ defined on a support ω , we have considered the effect of measuring $K(\mathbf{x})$ on one's ability to predict steady state flow within a bounded domain Ω , driven by statistically independent random source and boundary functions. Upon postulating the availability of a relatively smooth unbiased estimate $\kappa(\mathbf{x})$ of $K(\mathbf{x})$ based on local measurements, we proposed to predict the local hydraulic head $h(\mathbf{x})$ and Darcy flux $\mathbf{q}(\mathbf{x})$ on the scale ω by means of their unbiased conditional ensemble moments, $\langle h(\mathbf{x}) \rangle_\kappa$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$, where the subscript κ indicates that conditioning is done on the same data as those used to obtain $\kappa(\mathbf{x})$. These predictors satisfy a standard deterministic flow equation driven by ensemble mean source and boundary functions in which $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa = -\kappa(\mathbf{x})\nabla\langle h(\mathbf{x}) \rangle_\kappa + \mathbf{r}_\kappa(\mathbf{x})$ where $\mathbf{r}_\kappa(\mathbf{x})$ is a "residual flux." We have developed a compact integral expression for $\mathbf{r}_\kappa(\mathbf{x})$ which is rigorously valid for a broad class of $K(\mathbf{x})$ fields, including fractals, and which demonstrates that $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$ is generally nonlocal (depends on deterministic head gradients at points other than \mathbf{x}) and non-Darcian (there is no effective or equivalent conductivity valid for arbitrary directions of conditional mean flow) except in special cases.

2. Our expression for $\mathbf{r}_\kappa(\mathbf{x})$ consists of three spatial convolution integrals with kernels that cannot be evaluated quantitatively without either high-resolution Monte Carlo simulation or approximation. Nevertheless, these kernels are sufficiently well defined to reveal some of their most interesting fundamental properties. The kernel under the domain integral is a symmetric, positive-semidefinite, second-rank tensor (a dyadic); that under the Dirichlet boundary integral is a non-symmetric second-rank tensor; and that under the Neumann boundary integral is a vector. These three nonlocal kernels constitute system parameters which however depend not only on the statistical properties of the hydraulic conductivity field $K(\mathbf{x})$, but also on the information that one has about this field, as embodied in its estimate $\kappa(\mathbf{x})$ and in the statistical properties of the associated zero-mean estimation error $K'(\mathbf{x})$. The same is true about the local “hydraulic conductivity” $\kappa(\mathbf{x})$. Hence all the parameter in the deterministic flow equations depend on medium properties and on information content (scale, quantity and quality of data). As the quality of the estimator $\kappa(\mathbf{x})$ improves, information is transferred from the nonlocal/non-Darcian residual flux term $\mathbf{r}_\kappa(\mathbf{x})$ to the Darcian term $\kappa(\mathbf{x})\nabla\langle h(\mathbf{x})\rangle_\kappa$, and vice versa.
3. Our exact deterministic equations can be solved for the predictors $\langle h(\mathbf{x})\rangle_\kappa$ and $\langle \mathbf{q}(\mathbf{x})\rangle_\kappa$, subject to arbitrary random source and boundary terms, provided the kernels are known. Since these nonlocal parameters as well as the local parameter $\kappa(\mathbf{x})$ and the functions $\langle h(\mathbf{x})\rangle_\kappa$ and $h_\kappa(\mathbf{x})$ are smooth relative to their random counterparts, they can be treated approximately as finite dimensional functions. In other words, each can be expressed as the linear combination of a finite number of basis functions in the standard manner of finite elements; such

methods can then be used to solve the governing equations. The dimensionality N of each approximating function, i.e, the number N of basis functions employed (which controls the number of nodes and elements in a finite element grid), should depend in a standard way on the degree of smoothness of the function being approximated. Though we defer detailed discussion of this topic to future papers, we nevertheless conclude that our theory provides a direct answer to the important question how should the scale of grid discretisation (N) relate to the scale of measurement (ω) and to the quantity and quality of available data (smoothness of conditional moments). Likewise, we propose that if one has sufficient measurements of $K(\mathbf{x})$, $h(\mathbf{x})$ and $\mathbf{q}(\mathbf{x})$ on the scale ω and/or of spatial head and flux integrals (possibly weighted in a known fashion), one may in principle be able to estimate the above parameters a posteriori by means of a suitable inverse method. The latter implies conditioning not only on $K(\mathbf{x})$ but also on head and flux data.

4. Thanks to the rigorous nature of our integral expression for $\mathbf{r}_\kappa(\mathbf{x})$, we are in a position to make some precise and fundamental statements about the existence and properties of effective hydraulic conductivities in steady state flow through randomly nonuniform media. For $\mathbf{r}_\kappa(\mathbf{x})$ (and hence $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$) to be Darcian, it is necessary that there exist a symmetric, positive-semidefinite tensor $\tilde{\kappa}(\mathbf{x})$ of second rank (a dyadic) whose principal values (eigenvalues) do not exceed $\kappa(\mathbf{x})$, and which is additionally independent of $\langle h(\mathbf{x}) \rangle_\kappa$ and its gradient, such that $\mathbf{r}_\kappa(\mathbf{x}) = \tilde{\kappa}(\mathbf{x}) \nabla \langle h(\mathbf{x}) \rangle_\kappa$. Then and only then is it proper to write $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa = -\mathbf{K}_e(\mathbf{x}) \nabla \langle h(\mathbf{x}) \rangle_\kappa$ where $\mathbf{K}_e(\mathbf{x})$ is a symmetric, positive definite “effective (or equivalent) hydraulic conductivity tensor” given by $\mathbf{K}_e(\mathbf{x}) = \kappa(\mathbf{x})\mathbf{I} - \tilde{\kappa}(\mathbf{x})$, \mathbf{I}

being the identity tensor. This definition of effective (or equivalent) hydraulic conductivity is nontraditional in that it does not represent relations between upscaled fluxes and gradients, but between relatively smooth estimates of fluxes and gradients defined on the local scale ω . As such, our definition has an unambiguous operational meaning. In the unconditional case, it reduces to the definition commonly employed by stochastic subsurface hydrologists. We list below a number of special cases for which a $\mathbf{K}_e(\mathbf{x})$ can be defined rigorously.

5. When the flow domain $\bar{\Omega}$ tends formally to a point (practically to ω) while the mean source and boundary terms remain finite, $K_e(\mathbf{x})$ tends to the scalar $\kappa(\mathbf{x})$. This merely confirms that the hydraulic conductivity one calculates experimentally for a sample $\bar{\Omega}$ (in the laboratory or in the field) by imposing head and no flow conditions along the sample boundaries, Γ , and measuring the resultant flow rate, approaches the estimate $\kappa(\mathbf{x})$ of the scalar hydraulic conductivity $K(\mathbf{x})$ as $\bar{\Omega}$ tends to ω . The same happens for any flow domain in the limit as the variance of the estimation error $K'(\mathbf{x})$ approaches zero, or in the special case where $K(\mathbf{x}) \equiv K$ is a random variable independent of \mathbf{x} .
6. The case most often discussed in the stochastic literature is that of an unbounded domain Ω_∞ in which $\kappa(\mathbf{x}) \equiv \kappa = \text{constant}$, $K'(\mathbf{x})$ is statistically homogeneous, and flow is controlled entirely by boundary conditions at infinity which result in a uniform mean hydraulic gradient, $\nabla\langle h(\mathbf{x}) \rangle_\kappa \equiv \mathbf{J} = \text{constant}$ (we retain the subscript κ for consistency of notation even though no conditioning is implied by it in this case). Here the mean flux $\langle \mathbf{q} \rangle_\kappa$ is also constant and satisfies Darcy's law in the form $\langle \mathbf{q} \rangle_\kappa = -\mathbf{K}_e \mathbf{J}$ where $\mathbf{K}_e = \kappa \mathbf{I} - \tilde{\kappa}$, $\tilde{\kappa}$ being a constant, symmetric, positive-semidefinite, second-rank tensor and \mathbf{K}_e

an effective hydraulic conductivity tensor with similar properties but positive-definite.

7. Consider a box-shaped flow domain $\bar{\Omega}$ with prescribed mean head on two opposing faces and prescribed zero mean flux normal to the remaining four faces. If internal sources are zero in the mean while $\kappa(\mathbf{x}) \equiv \kappa = \text{constant}$ (but $K'(\mathbf{x})$ is not necessarily homogeneous !), then the flux predictor is given by $\langle \mathbf{q}(\mathbf{x}) \rangle_{\kappa} = -\kappa \nabla \langle h(\mathbf{x}) \rangle_{\kappa} + \tilde{\kappa}(\mathbf{x}) \mathbf{J}$ where \mathbf{J} is the externally imposed hydraulic gradient across the box and $\tilde{\kappa}(\mathbf{x})$ is a (generally) nonsymmetric, nonconstant, second-rank tensor. Here the lack of symmetry is strictly a boundary source effect which disappears when the influence of such boundaries dies out. This and the fact that $\langle \mathbf{q}(\mathbf{x}) \rangle_{\kappa}$ is linear in both $\nabla \langle h(\mathbf{x}) \rangle_{\kappa}$ and \mathbf{J} but is not proportional to either of these two gradients justifies referring to the above expression as quasi-Darcian. The tensor $\tilde{\kappa}(\mathbf{x})$ depends on $K(\mathbf{x})$, the quality of its estimate $\kappa(\mathbf{x})$, the location (if $K(\mathbf{x})$ is statistically nonhomogeneous) and orientation (if $K(\mathbf{x})$ is statistically anisotropic) of the box, box dimensions, and the choice of prescribed mean head and flux boundaries.

8. If $K'(\mathbf{x})$ is statistically homogeneous, exhibits ellipsoidal (geometric) statistical anisotropy, and the above box-shaped flow domain $\bar{\Omega}$ as well as the working coordinates \mathbf{x} are oriented parallel to the principal coordinates of statistical anisotropy, then the conditional mean flux is given by Darcy's law in the form $\langle q_1(x_2, x_3) \rangle_{\kappa} = -[\kappa - K_1(x_2, x_3)] J_1$ where J_1 is the externally imposed hydraulic gradient (taken to act parallel to x_1) and $[\kappa - K_1(x_2, x_3)]$ is an effective directional hydraulic conductivity. The latter is a function of $K(\mathbf{x})$, the quality of its estimate $\kappa(\mathbf{x})$, box dimensions, and location in the (x_2, x_3) plane normal to

the mean direction of flow. If the box is at an angle to the principal axes of statistical anisotropy then flow symmetry is broken, Darcy's law no longer holds, and one must use instead the quasi-Darcian expression introduced earlier. Only if $K'(\mathbf{x})$ is statistically isotropic will $\langle \mathbf{q}_1(x_2, x_3) \rangle_\kappa = -[\kappa - K_1(x_2, x_3)]J_1$ apply for arbitrary choices of x_1, x_2, x_3 relative to a frame of reference attached to the medium. Even then, the three quantities $[\kappa - K_1(x_2, x_3)]$, $[\kappa - K_3(x_1, x_2)]$, and $[\kappa - K_2(x_3, x_1)]$ will generally differ from each other unless $\bar{\Omega}$ is a cube. It is important to appreciate that these quantities are strictly directional and do not form the principal components of a tensor. It is further important to recall that they correspond to a closed box-shaped flow domain with two prescribed constant mean head and four zero mean flow boundaries. As heads and fluxes along interfaces between finite elements or finite difference cells in a numerical model are generally not prescribed but merely required to satisfy appropriate compatibility criteria, the above quantities qualify (in a strict sense) neither as principal nor as directional equivalent hydraulic conductivities for such elements or cells.

9. It is common in subsurface hydrology to treat $Y(\mathbf{x}) = \ln K(\mathbf{x})$ as a Gaussian, statistically homogeneous field. Matheron (1967) considered uniform mean flow in an unbounded, two-dimensional, statistically isotropic $Y(\mathbf{x})$ field in which the effective hydraulic conductivity is a scalar, K_e . By recognizing that this scalar must be a linear functional of $K(\mathbf{x})$, Matheron was able to prove that K_e is equal to the geometric mean, K_g , of $K(\mathbf{x})$ regardless of how strongly the latter fluctuates. We have restated Matheron's proof by basing it on our new integral expression for $\mathbf{r}_\kappa(\mathbf{x})$ rather than on an abstract linear functional.

10. In the case of uniform mean flow within an unbounded, three-dimensional, statistically anisotropic $Y(\mathbf{x})$ field the effective hydraulic conductivity tensor \mathbf{K}_e is described very well by a modified version (Neuman and Depner, 1988) of an expression originally proposed by Gelhar and Axness (1983). This expression is supported by high-resolution 3-D finite element Monte Carlo simulations (Neuman et al., 1992) at least for (a) the statistically isotropic case where (b) the variance σ_Y^2 of $Y(\mathbf{x}) = \ln K(\mathbf{x})$ is as large as 7 and (c) the integral scale measures one seventh the characteristic length of the flow domain. A similar agreement with Monte Carlo results, generated with a different numerical method, has been reported by Dykaar and Kitanidis (1992) for σ_Y^2 as large as 6; Ababou (1988) found on the basis of a single realization in a large domain that the agreement holds at least up to $\sigma_Y^2 = 5.3$. The same expression is also supported by field evidence (Neuman and Depner, 1988) from a fractured granitic rock terrain for $\sigma_Y^2 \geq 7$. Its verification confirms indirectly that, in statistically isotropic 2-D hydraulic conductivity fields, the scalar K_e is equal to the geometric mean, K_g , for any variance or integral scale. In the statistically isotropic case, the linearized version of the above expression for \mathbf{K}_e underestimates the scalar K_e by an amount that grows exponentially with σ_Y^2 as the latter increases beyond 1.
11. To investigate the existence of effective hydraulic conductivities under a radially converging mean flow in a bounded domain, we performed high-resolution Monte Carlo simulations of 2-D flow to a point sink in a statistically homogeneous, isotropic, log normal $K(\mathbf{x})$ field with an exponential covariance and an integral scale λ . Our simulations confirm a theoretical prediction by Da-

gan (1989), simplified considerably in this paper, that in the absence of an outer boundary the corresponding scalar effective hydraulic conductivity, K_e , approaches the geometric mean, K_g , as the dimensionless radial distance r/λ from the sink increases when σ_Y^2 is small. They demonstrate further that K_e is virtually equal to K_g at radial distances exceeding 2 – 3 integral scales from the sink when $\sigma_Y^2 = 1$, and $r/\lambda \geq 4$ when $\sigma_Y^2 = 4$, provided the outer boundary is far enough. The latter boundary influences K_e within 2 – 4 dimensionless distances when $\sigma_Y^2 = 1$, but well within 10 such distances when $\sigma_Y^2 = 4$. Our simulations also confirm Dagan's (1989) theoretical prediction that, regardless of σ_Y^2 and λ , K_e reduces toward the harmonic mean $K_h = K_g \exp(-\sigma_Y^2/2)$ as r/λ decreases. They demonstrate further that the same is true near the outer boundary source. Our findings contradict Naff's (1991, p. 313) conclusion that "when the flow tends to emulate a two-dimensional system, then Matheron's (1967) asymptotic results for radial flow are generally realized: the apparent effective hydraulic conductivity becomes the arithmetic mean near the well and the harmonic mean distant from the well."

12. Our unconditional Monte Carlo results suggest that, in the 2-D case when the conditional variance of $Y(\mathbf{x})$ does not exceed 1, it may be appropriate to approximate $\mathbf{r}_\kappa(\mathbf{x})$ by the local expression $\mathbf{r}_\kappa(\mathbf{x}) \simeq \hat{\kappa}(\mathbf{x})\nabla h_\kappa(\mathbf{x}')$ for points \mathbf{x} located a distance of at least two to three conditional integral scales from point sources, and four such integral scales from linear boundary sources; when the conditional variance is 4, the same is true for \mathbf{x} located at least four conditional integral scales from point sources and ten such scales from linear boundary sources. Here $\hat{\kappa}(\mathbf{x})$ is a "residual hydraulic conductivity tensor," and "conditional in-

tegral scale” is that of the log estimation errors $Y'(\mathbf{x})$ about the conditional mean of $Y(\mathbf{x})$. Since this integral scale can be reduced through the judicious acquisition of data, it follows that one has in principle control over the distance beyond which such a local approximation is valid.

13. Our 2-D simulations show that the sample variance of the specific flux at a given radial distance from the point sink increases with σ_Y^2 but decreases with λ . At distances that exceed about one integral scale from the sink, this variance decreases at a rate more or less proportional to r^{-2} as proposed on the basis of theoretical considerations by Matheron (1967) and confirmed analytically by Naff (1991). Closer to the source, the rate of variation with distance is somewhat slower. Our results thus support Naff in his conclusion that if an effective hydraulic conductivity can be defined for mean radial flow in two dimensions, then (*ibid*, p. 314) “the best estimates of the effective hydraulic conductivity will be obtained from observations at some distance from the well bore.”
14. To render our nonlocal theory workable near interior sources and boundaries where localization may not be warranted, one must either have sufficient data to evaluate the unknown kernels by an inverse method as was suggested earlier, or develop a working approximation for the residual flux $\mathbf{r}_\kappa(\mathbf{x})$. We proposed a weak approximation (closure) which eliminates the need for moments of $K(\mathbf{x})$ beyond $\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa$. Due to its conditional nature, the approximation is expected to improve with the quality of the estimate $\kappa(\mathbf{x})$ as the conditional ensemble moments of the estimation error $K'(\mathbf{x})$ diminish. The same approximation also yields working expressions for the terms $\tilde{\kappa}$, $\tilde{\kappa}(\mathbf{x})$, and $K_1(x_2, x_3)$

which enter into our definitions of effective hydraulic conductivity.

15. Analogues of the proposed weak approximation have been investigated directly by means of high-resolution finite element Monte Carlo analysis of 2-D flow to a point sink. Our results suggest that weak (integral) approximations are much more robust than their strong (local) counterparts, that they are acceptable for σ_Y^2 at least as large as 4 and λ at least as large as 8 (in a 64×64 domain), and that they improve as the variance of the estimation error $K'(\mathbf{x})$ goes down due to an improvement in the quality of the estimate $\kappa(\mathbf{x})$, as a result of conditioning on data.
16. Our weak approximation requires the solution of an integro-differential equation for the conditional ensemble mean Green's function $\langle \mathcal{G}(\mathbf{x}, \mathbf{y}) \rangle_\kappa$. This might be done by means of a quasi-analytical finite element scheme, followed by finite element computation of the head predictor $\langle h(\mathbf{x}) \rangle_\kappa$ and the flux predictor $\langle \mathbf{q}(\mathbf{x}) \rangle_\kappa$. As all three functions vary much more slowly in space than their random counterparts, one should feel comfortable approximating them by quasi-analytical and/or finite-dimensional functions as required by our proposed approach.
17. Though it is possible to express the head prediction error moment $\langle h'(\mathbf{x})h'(\mathbf{y}) \rangle_\kappa$ and the flux prediction error moment $\langle \mathbf{q}'(\mathbf{x})\mathbf{q}'^T(\mathbf{y}) \rangle_\kappa$ at a level of approximation compatible with our weak scheme, we consider the resulting expressions to be unduly complex. We therefore supplemented (in the case of $\langle h'(\mathbf{x})h'(\mathbf{y}) \rangle_\kappa$) or replaced (in the case of $\langle \mathbf{q}'(\mathbf{x})\mathbf{q}'^T(\mathbf{y}) \rangle_\kappa$) these expressions by stronger but much simpler explicit approximations which we hope to evaluate by finite elements in the future.

18. Our theory assumes that one has at his/her disposal an unbiased estimate $\kappa(\mathbf{x})$ of the hydraulic conductivity function $K(\mathbf{x})$. Our weak approximation requires further that one know the second conditional moment of the associated estimation errors $\langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_{\kappa}$. The question how to obtain these quantities is not central to our paper and remains outside its scope. Nevertheless, we have recalled two known geostatistical methods which, under some practical working hypotheses, make possible the determination of $\kappa(\mathbf{x})$ on the basis of local $K(\mathbf{x})$ data in a way which minimizes, though does not completely eliminate, bias. The first method treats $\kappa(\mathbf{x})$ as a global drift which does not generally pass through the measurement points and results in zero mean, wide-sense (second-order) stationary estimation errors $K'(\mathbf{x})$. The second method treats $\kappa(\mathbf{x})$ as a kriging estimate which does pass through the measurement points (unless they are corrupted by measurement noise) and results in a zero mean, nonstationary $K'(\mathbf{x})$ field. Other methods to evaluate $\kappa(\mathbf{x})$ may be equally acceptable or better, including inverse methods as mentioned earlier.
19. The theory in this dissertation has dealt with the prediction of heads and fluxes on scales compatible with the support ω of $K(\mathbf{x})$. Once the spatial distribution of the corresponding predictors $\langle h(\mathbf{x}) \rangle_{\kappa}$ and $\langle \mathbf{q}(\mathbf{x}) \rangle_{\kappa}$ has been determined, one can integrate them spatially (in a weighted manner if necessary) to represent averages such as those “seen,” for example, by wells completed over depth intervals which exceed the vertical dimension of ω (as when this dimension is the length of a core sample or the injection interval in a packer test). The question whether such spatially averaged heads and fluxes can be related functionally by means of “upscaled” parameters defined on supports larger than ω can then

be addressed directly.

20. The investigation of four different methods for generating random fields (abbreviated RFG's) with five different random number generators (RNG's) suggests that the choice of a combination of RFG & RNG depends on the particular purpose and needs; everything being equal, the sequential simulator is the most powerful RFG, while the RFG based on LU-decomposition (of the covariance matrix) is potentially most reliable. In our case, i.e., 64×64 square fields ($\lambda = 5$) with steep gradients and variances of $\ln K$ up to 4 (i.e., $\sigma_Y^2 \leq 4$), at least 500 realizations are needed to provide reliable ensemble moments of both the (input) generated fields and the (output) simulated heads and fluxes. As the number of simulations increases, the results (particularly the "two-point" covariances) become less sensitive to the choice of RFG.

APPENDIX A

RANDOM FUNCTIONS:

A random function (RF) is an infinite dimensional function, $u(\mathbf{x})$, of the spatial coordinate \mathbf{x} (or of time, t). Let u_i be the values of u at a point $\mathbf{x} = \mathbf{x}_i$ (Dagan, 1989). The *joint probability distribution function* (or *joint cumulative distribution function*, or joint CDF) $F(u_1, u_2, \dots, u_N)$, with $N \rightarrow \infty$, contains all the probability information about u . The passage to an infinite number of components poses difficult problems, and the appropriate mathematical tool is that of functional analysis; however, a less rigorous definition, but sufficient for most applications, is commonly used (e.g., Dagan, 1989): it assumes that the joint CDF of u is known for any set of arbitrary, but finite, number of points. Further, if the joint CDF of u is invariant under translation of the points \mathbf{x}_i , and depends only on their *relative* positions, then u is a *stationary* (or homogeneous) random function (or field). In this case, all moments are independent of space (or time); particularly, the mean¹ $\langle u \rangle = \text{const}$, and the covariance C_u (or, alternatively, the autocorrelation coefficient) depend only on distances (or time-lags) between the points, not on their actual locations, i.e.,

$$C_u(\mathbf{x}_1, \mathbf{x}_2) = \langle u(\mathbf{x}_1)u(\mathbf{x}_2) \rangle - \langle u(\mathbf{x}_1) \rangle \langle u(\mathbf{x}_2) \rangle = C_u(\mathbf{x}_1 - \mathbf{x}_2) = C_u(\mathbf{r}), \quad (\text{A.1})$$

¹ $\langle \cdot \rangle$ denotes *ensemble* mean or mathematical expectation.

where \mathbf{r} is a separation vector. If only the first two moments satisfy these requirements (i.e., they are independent of space/time), the function is *second order* or *weakly* or *wide sense* stationary.

The variance σ_u^2 and the autocorrelation coefficient ρ_u are defined as

$$\sigma_u^2(\mathbf{x}) = C_u(\mathbf{x}, \mathbf{x}) = \langle u(\mathbf{x})^2 \rangle - \langle u(\mathbf{x}) \rangle^2, \quad (\text{A.2})$$

$$\rho_u(\mathbf{x}_1, \mathbf{x}_2) = \frac{C_u(\mathbf{x}_1, \mathbf{x}_2)}{\sigma_u(\mathbf{x}_1)\sigma_u(\mathbf{x}_2)}. \quad (\text{A.3})$$

A stationary random function is *isotropic* if the joint CDF is also invariant under rotation. In this case, the autocovariance and the autocorrelation coefficients become functions of $|\mathbf{r}|$, the modulus of the separation vector \mathbf{r} , and $C_u(\mathbf{r}) = \sigma_u^2 \rho_u(\mathbf{r})$.

Random functions $u(\mathbf{x})$ whose joint CDF for any set of points \mathbf{x}_i are multivariate *normal* (or Gaussian) are of special interest; here, the entire statistical structure of u is exhausted by the first two moments [i.e., the mean $\langle u(\mathbf{x}) \rangle$ and the covariance $C_u(\mathbf{r})$]. If a Gaussian RF is weakly stationary, then the residual $u'(\mathbf{x}) = u(\mathbf{x}) - \langle u(\mathbf{x}) \rangle$ is stationary in a strict sense, and all the statistical moments of $u(\mathbf{x}_i)$ can be easily calculated (Dagan, 1989).

The *conditional mean* of a random variable u , given the random variable $v = v_o$, is (Papoulis, 1984)

$$\langle u | v_o \rangle \equiv \langle u \rangle_{v_o} = \int_{-\infty}^{\infty} u f(u | v) du, \quad (\text{A.4})$$

where $f(u | v)$ is the conditional probability density function (conditional *pdf*)

$$f(u | v) = \frac{f(u, v)}{f(v)}, \quad (\text{A.5})$$

$f(u, v)$ is the joint *pdf* of u and v , and the marginal *pdf* of v ,

$$f(v) = \int_{-\infty}^{\infty} f(u, v) du, \quad (\text{A.6})$$

is also the *total probability* (Papoulis, 1984, p. 161). If the joint CDF for two variables, $F(u, v)$, is twice differentiable, then the joint *pdf* of u and v is, by definition

$$f(u, v) = \frac{\partial^2 F(u, v)}{\partial u \partial v}, \quad (\text{A.7})$$

For a sequence of N variables, the conditional mean of u_1 , given u_2, \dots, u_N is

$$\langle u_1 | u_2, \dots, u_N \rangle = \int_{-\infty}^{\infty} u_1 f(u_1 | u_2, \dots, u_N) du_1, \quad (\text{A.8})$$

where $f(u_1 | u_2, \dots, u_N)$ is the conditional *pdf*, and $\langle u_1 | u_2, \dots, u_N \rangle$ is, by itself, a random variable, such that (Papoulis, 1984, p. 183)

$$\langle \langle u_1 | u_2, \dots, u_N \rangle \rangle = \langle u_1 \rangle. \quad (\text{A.9})$$

Differentiation and integration of random functions:

A random function $u(\mathbf{x})$ is called continuous in the mean square (m.s.) at a point \mathbf{x} if (Dagan, 1989)

$$\lim_{|\mathbf{r}| \rightarrow 0} \langle [u(\mathbf{x} + \mathbf{r}) - u(\mathbf{x})]^2 \rangle = 0. \quad (\text{A.10})$$

A random function $u(\mathbf{x})$ is differentiable in the m.s. at \mathbf{x} and has a random function $du/d\mathbf{x}$ as derivative at \mathbf{x} if

$$\lim_{\mathbf{r} \rightarrow 0} \left[\frac{u(\mathbf{x} + \mathbf{r}) - u(\mathbf{x})}{r} - \frac{du(\mathbf{x})}{d\mathbf{x}} \right]^2 = 0. \quad (\text{A.11})$$

This definition is easily extended to partial differentiation. A necessary and sufficient condition for $du/d\mathbf{x}$ to exist is that the covariance function $C_u(\mathbf{x}, \mathbf{x} + \mathbf{r})$ of the random parameter be twice differentiable at the origin, i.e., as $\mathbf{r} \rightarrow 0$ (Dagan, 1989; p. 15).

A stochastic integral $U = \int_V a(\mathbf{x})u(\mathbf{x})d\mathbf{x}$ is defined as the mean square (m.s.) limit of the infinite sum $U_n = \sum_{i=1}^n a(\mathbf{x}_i)u(\mathbf{x}_i)\Delta\mathbf{x}_i$, $a(\mathbf{x})$ being a deterministic function. A necessary and sufficient condition for the existence of the stochastic integral of $a(\mathbf{x})u(\mathbf{x})$ over V (in the ordinary, Riemann sense) is that its variance $\sigma_U^2 \equiv \langle U'^2 \rangle$ satisfy (Dagan, 1989)

$$\langle U'^2 \rangle = \int_V \int_V a(\mathbf{x}')a(\mathbf{x}'')C_u(\mathbf{x}', \mathbf{x}'')d\mathbf{x}'d\mathbf{x}'', \quad (\text{A.12})$$

where the expected value of U is

$$\langle U \rangle = \int a(\mathbf{x})\langle u(\mathbf{x}) \rangle d\mathbf{x}. \quad (\text{A.13})$$

In the stationary and isotropic case, the *integral scale* of $u(\mathbf{x})$ is defined as (Dagan, 1989, p. 19)

$$\lambda_u = \left[n \int_0^\infty \rho_u(r)r^{n-1}dr \right]^{1/n}, \quad (\text{A.14})$$

where n is the dimensionality of \mathbf{x} . The integral scale is a measure of the distance between two points beyond which $u(\mathbf{x})$ and $u(\mathbf{x} + r)$ cease to be correlated.

APPENDIX B

SPACES, NORMS, AND STRONG CONTINUITY:

Spaces and Norms: (After Collatz, 1966; Curtain and Pritchard, 1977; Bharucha-Reid, 1972)

A space R is called a *metric space* if for any two elements $f, g \in R$ there is a real number $d = d(f, g)$ called *the distance*, which is symmetric and positive semidefinite,

$$d(f, g) = d(g, f), \quad d(f, g) \geq 0 \quad \forall f, g \in R \quad (\text{B.1})$$

(where \forall means “for all”) with $d = 0$ if and only if $f = g$, and

$$d(f, h) \leq d(f, g) + d(h, g) \quad f, g, h \in R. \quad (\text{B.2})$$

A space is called *linear* (with respect to a field F) if there is defined an addition and a multiplication of elements with numbers of the field F . These operations satisfy the rules of vector algebra, particularly the associative and distributive laws.

A *linear transformation*, T , from a linear space X to a linear space Y over the same field F is a map $T : X \rightarrow Y$, such that

$$T(\alpha x + \beta y) = \alpha T x + \beta T y \quad (\text{B.3})$$

for all $x, y \in X$ and all $\alpha, \beta \in F$.

A space R is *normed* if it is linear, and if each element $f \in R$ is associated with a real number $\|f\|$, the norm of f , which satisfies the three relations:

1. $\|f\| = 0$ if and only if $f = \Theta$ (definiteness).
2. $\|f + g\| \leq \|f\| + \|g\|$ (triangle inequality).
3. $\|cf\| = |c| \|f\|$ (homogeneity for $f \in R, c \in F$)

A norm may be thought of as the distance d of a function (or variable) f from its null value Θ (or from the origin), that is $\|f\| = d(f, \Theta)$. The norm of elements (points) $\mathbf{x} = (x_1, \dots, x_n)$ can be defined either by

$$\|\mathbf{x}\| = \max_j (p_j |x_j|^q), \quad p_j \text{ real } > 0 \quad (\text{B.4})$$

or

$$\|\mathbf{x}\| = \left[\sum_{j=1}^n p_j |x_j|^q \right]^{1/q}, \quad 1 \leq q < \infty, \quad (p_j > 0), \quad (\text{B.5})$$

where both the coefficients, p_j , and the exponent, q , define the type of norm. In particular, if $p_j \equiv 1$, and $q = 2$, this becomes an *Euclidean norm*.

Similarly the norm of a function f can also be defined as

$$\|f(\mathbf{x})\| = \max [p(\mathbf{x}) |f(\mathbf{x})|], \quad (\text{B.6})$$

where $p(\mathbf{x})$ is a given positive function. For $p(\mathbf{x}) \equiv 1$

$$\|f(\mathbf{x})\| = \max |f(\mathbf{x})|. \quad (\text{B.7})$$

In the case of functions with continuous (partial) derivatives on a closed interval (or a closed n -dimensional region), the norm can be defined as

$$\|f(\mathbf{x})\| = \left[\int \sum_{j=0}^n p_j(\mathbf{x}) |f^{(j)}(\mathbf{x})|^q d\mathbf{x} \right]^{1/q}, \quad (\text{B.8})$$

or as

$$\|f(\mathbf{x})\| = \sup \sum_{j=0}^n p_j(\mathbf{x}) |f^{(j)}(\mathbf{x})|, \quad (\text{B.9})$$

where $f^{(0)} = f$, and superscripts indicate partial derivatives of some order (j); p_j are given continuous nonnegative functions.

The following are alternative definitions for a *Banach space*:

- A complete¹ normed linear space is called *Banach space*.
- A *Banach space* is an n -dimensional vector space with a norm.
- A nonempty set \mathcal{X} is a *Banach space* if (1) \mathcal{X} is a linear space, (2) \mathcal{X} has a norm, and (3) \mathcal{X} is a complete metric space with metric $d(x, y) = \|x - y\|$.

In particular:

1. The norm in (B.5), with \mathbf{x} representing n -tuples of real numbers, is a real Banach space with respect to coordinatewise addition and scalar multiplication, and with the *Euclidean norm* where $p_j \equiv 1$, and $q = 2$.
2. The space $C[B]$ of *continuous functions* on a closed and bounded domain B with the norm (B.6) or (B.7) is a Banach space.
3. The *Lebesgue spaces* L_p , $1 \leq p < \infty$, with the function space $L_p(S)$ defined for any real number p , and any positive measure space which consists of those scalar-valued measurable functions $f(s)$ on S for which $\int_S |f(s)| ds$ is finite. The L_p spaces are Banach spaces with norm defined by $\|x\|_p = [\int_S |f(s)|^p ds]^{1/p}$

Bounded transformations:

$T : X \rightarrow Y$, where X and Y are normed linear spaces is said to be *bounded* if (Curtain

¹A subset N of a metric space R is “complete” if there is a limit element $f \in N$ to each Cauchy sequence f_m with the distance $d(f_m, f) \rightarrow \Theta$ (Collatz, 1966).

and Pritchard, 1977)²

$$\|Tx\|_Y \leq M\|x\|_X \quad (\text{B.10})$$

for some constant $M > 0$ and all $x \in X$.

Norm of a linear bounded transformation: (Curtain and Pritchard, 1977)

If T is a linear bounded transformation between two normed linear spaces X, Y , we define its norm $\|T\|$ as

$$\|T\| = \sup \left\{ \frac{\|Tx\|_Y}{\|x\|_X} \right\}, \quad (\text{B.11})$$

where $x \in X$ and $x \neq 0$. Consequently,

$$\|Tx\|_Y \leq \|T\|\|x\|_X, \quad (\text{B.12})$$

and we can define the normed linear space $\Lambda(X, Y)$ to be the space of bounded transformations $T : X \rightarrow Y$ with the above norm.

Strong continuity: (Curtain and Pritchard, 1977)

Let $T(t) \in \Lambda(X, Y)$ for every $t \in [a, b]$, then $T(t)$ is *strongly continuous* at t_0 if

$$\|T(t)x - T(t_0)x\|_Y \rightarrow 0 \quad \forall x \in X \quad \text{as } t \rightarrow t_0 \quad (\text{B.13})$$

Curtain and Pritchard (1977) bring, as example, the operator e^{At} , where $A \in \Lambda(X)$, as a strongly (or uniformly) continuous operator. In fact, since e^{At} can be expressed as a convergent series (see Appendix 1-D for definitions),

$$e^{At} = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \cdots \quad (\text{B.14})$$

where I is the identity tensor, then $e^{At} \in \Lambda(X)$. [The operator e^{At} is fundamental to the semigroup approach.]

² $\|\cdot\|_X$ denotes the norm of \cdot on X .

APPENDIX C

THE EIGENFUNCTION METHOD

An alternative approach to the inverse operator, or Green's function method, is the *eigenfunction method* (Greenberg, 1971). Since this method requires a good deal of theoretical development, we will only highlight some of its major properties and its relation to Green's functions and Fourier series expansions.

In order to solve (2.4) by the eigenfunction method, we consider the associated Sturm-Liouville eigenvalue problem¹

$$Lh + \lambda h = 0, \tag{C.1}$$

subject to the same *homogeneous*² boundary conditions as in (2.4). λ is a parameter (associated with the eigenvalues). The procedure is to expand the various quantities in (2.4) in terms of the eigenfunctions ϕ_n of (C.1):

$$\begin{aligned} h(\mathbf{x}) &= \sum_{n=1}^{\infty} a_n \phi_n(\mathbf{x}) \\ f(\mathbf{x}) &= \sum_{n=1}^{\infty} c_n \phi_n(\mathbf{x}), \end{aligned} \tag{C.2}$$

where the a_n 's are unknown, and c_n 's are the coefficients³ of $f(\mathbf{x})$, and thus can be computed from ϕ_n and f :

$$c_n = \frac{(f, \phi_n)}{(\phi_n, \phi_n)}, \tag{C.3}$$

¹Although Greenberg (1971) discusses the original ordinary 1-D problem, we generalize it here to higher dimensions.

²The homogeneity of the boundary conditions is basic to the eigenfunction method.

³If $f(\mathbf{x})$ in (C.2) is a Fourier series, then c_n are Fourier coefficients.

where the *inner product* $(f(\mathbf{x}), g(\mathbf{x}))$ is defined as

$$(f(\mathbf{x}), g(\mathbf{x})) \equiv \int_{\Omega} f(\mathbf{x})g(\mathbf{x})d\mathbf{x}, \quad (\text{C.4})$$

particularly

$$(\phi_n, \phi_n) \equiv \int_{\Omega} \phi_n(\mathbf{x})\phi_n(\mathbf{x})d\mathbf{x}, \quad (\text{C.5})$$

such that $(\phi_n, \phi_m) = 0$ for $m \neq n$, i.e., the eigenfunctions are mutually orthogonal. Greenberg (1971) shows that in (C.2), $a_n = -c_n/\lambda_n$, where λ_n are the *eigenvalues* (see below).

For example, the nontrivial solution to the one dimensional boundary value problem

$$\frac{d^2h}{dx^2} = f(x); \quad h(0) = h(l) = 0 \quad (\text{C.6})$$

with the associated eigenvalue problem

$$\frac{d^2h}{dx^2} + \lambda h = 0; \quad h(0) = h(l) = 0 \quad (\text{C.7})$$

is

$$h(x) = A \sin \sqrt{\lambda}x + B \cos \sqrt{\lambda}x. \quad (\text{C.8})$$

where A and B are constants. Now, the boundary condition (BC) $h(0) = 0$ implies that $B = 0$, while the BC $h(l) = 0$ implies that either $A = 0$ or $\sqrt{\lambda}l = n\pi$, where $n = 1, 2, 3, \dots$. Consequently,

$$\lambda = \frac{n^2\pi^2}{l^2} \equiv \lambda_n, \quad (\text{C.9})$$

and

$$h(x) = A \sin \left(\frac{n\pi x}{l} \right) \equiv A\phi_n(x), \quad (\text{C.10})$$

where A is an arbitrary constant, λ_n are the *eigenvalues* (or characteristic values), and the corresponding *eigenfunctions* (or characteristic functions) are $\phi_n = \sin(n\pi x/l)$,

which happen to be precisely the quantities needed for Fourier series representation of functions defined over $0 \leq x \leq l$. In a sense, this is where the terms of a Fourier sine series “come from” (Greenberg, 1971).

The general solution of (2.4) consists of first determining the eigenvalues λ_n and eigenfunctions $\phi_n(\mathbf{x})$ of the associated Sturm-Liouville problem (C.1), then computing Fourier coefficients c_n of $f(\mathbf{x})$ with respect to these eigenfunctions, and finally inserting these various quantities into

$$h(\mathbf{x}) = - \sum_{n=1}^{\infty} \left(\frac{c_n}{\lambda_n} \right) \phi_n(\mathbf{x}). \quad (\text{C.11})$$

Note that the solution fails to exist if, say $\lambda_j = 0$ and $c_j \neq 0$. If, however, $c_j = 0$ then there is an infinite number of solutions (Greenberg, 1971, p. 48).

A substitution of (C.3) in (C.11) yields the following solution, h (Greenberg, 1971) :

$$h(\mathbf{x}) = \int_{\Omega} \left\{ - \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x}') \phi_n(\mathbf{x})}{\lambda_n(\phi_n, \phi_n)} \right\} f(\mathbf{x}') d\mathbf{x}'. \quad (\text{C.12})$$

In the case of a bounded domain, boundary (or surface) integrals should be added to (C.12). A close look at the kernel of the integral (i.e., the fixed part, as discussed below) in (C.12) identifies it as the Green's function (Greenberg, 1971)

$$G(\mathbf{x}', \mathbf{x}) = - \sum_{n=1}^{\infty} \frac{\phi_n(\mathbf{x}') \phi_n(\mathbf{x})}{\lambda_n(\phi_n, \phi_n)}. \quad (\text{C.13})$$

Thus, the Fourier series expansion of any Green's function yields the eigenfunctions and eigenvalues of the particular differential equation, and vice versa. Note that from (C.13), the Green's function fails to exist if one of the eigenvalues of the associated Sturm-Liouville system, say λ_j is zero. Summarizing, if zero is not an eigenvalue of (C.1), then both the Green's function G and the solution h exists, and are given by (C.13) and (C.11) respectively. If, on the other hand, some $\lambda_j = 0$, then the Green's

function fails to exist, and either the solution h fails to exist, or it has an infinity of solutions.

APPENDIX D

CONVERGENCE AND ASYMPTOTIC APPROXIMATION

Deterministic functions:

Convergence:

A series u_N converges to its limit u if for any $\epsilon > 0$ there exists a number $M > 0$, such that $|u_N - u| < \epsilon$ for all $N > M$, and in such case we write $\lim_{N \rightarrow \infty} u_N = u$ (e.g., Spiegel, 1971). Thus, an expansion converges if its terms eventually decay sufficiently rapidly (Hinch, 1991). Note that convergence of a series does not guarantee *accuracy* of its approximation by a finite number of terms (to be demonstrated in the sequel).

Uniform convergence:

A sum of a series of functions $u_1(x) + u_2(x) + \dots + u_N(x) = S_N(x)$ converges to its limit $S(x)$ in a region R if for any $\epsilon > 0$ there exists a number N , which in general depends on both ϵ and x , such that $|S_N(x) - S(x)| < \epsilon$ for all $N > M$. Note that a necessary (but not sufficient) condition for this convergence to a limit is that $\lim_{N \rightarrow \infty} u_N(x) = 0$. If N depends only on ϵ , not on x , the series converges *uniformly* to $S(x)$ in R (Spiegel, 1971). In other words, if for sufficiently large N , the remainder (or tail) of the series is less than an arbitrarily small ϵ in a given interval, the series converges uniformly (Arfken, 1985).

Random functions:

Convergence almost everywhere (a.e.):

If the set of outcomes ζ such that

$$\lim_{N \rightarrow \infty} u_N(\zeta) = u(\zeta) \quad (\text{D.1})$$

exists, and its probability equals 1, then the random sequence u_N converges a.e (or with probability 1), or equivalently it converges *almost surely* (a.s.). This is written in the form

$$P\{u_N \rightarrow u\} = 1 \quad \text{as} \quad N \rightarrow \infty \quad (\text{D.2})$$

where P denotes probability.

Convergence in the Mean Square (or m.s.) sense:

A sequence of random variables $u_N(\zeta)$ converges to $u(\zeta)$ in the *mean square* sense if $|u_N - u|$ converges in the mean to zero as $N \rightarrow \infty$, i.e.,

$$\lim_{N \rightarrow \infty} \langle |u_N - u|^2 \rangle = 0. \quad (\text{D.3})$$

Similarly, we define (Beran, 1968) *Convergence in the r -th Mean:*

A sequence of random variables $u_N(\zeta)$ converges to $u(\zeta)$ in the r -th mean sense if $|u_N - u|^r$ converges in the r -th mean to zero as $N \rightarrow \infty$, i.e.,

$$\lim_{N \rightarrow \infty} \langle |u_N - u|^r \rangle = 0. \quad (\text{D.4})$$

Convergence in Probability (p):

The probability $P\{|u - u_N| > \varepsilon\}$ of the event $\{|u - u_N| > \varepsilon\}$ is a sequence of

numbers depending on ε . If this sequence tends to zero

$$\lim_{n \rightarrow \infty} P\{|u - u_N| > \varepsilon\} = 0 \quad (\text{D.5})$$

for any $\varepsilon > 0$, then we say that the sequence u_N tends to u in probability (or in measure). This is called also *stochastic convergence* (Papoulis, 1984).

Convergence in Distribution (d).

If $F_n(u)$ and $F(u)$ denote the (probability) distributions of u_N and u , such that

$$\lim_{n \rightarrow \infty} F(u_N) = F(u) \quad (\text{D.6})$$

for every point u of continuity of $F(u)$, then the sequence u_N tends to u in distribution. In this case, $u_N(\zeta)$ need not converge for any ζ .

Asymptotic approximation:

The partial sum $\sum_{i=1}^N u_i(x)$ is said to be an asymptotic approximation to (or asymptotic representation of) a function $u(x)$ as $\varepsilon \rightarrow 0$ if for each $M \leq N$

$$\frac{u(\varepsilon) - \sum_{i=1}^M u_i(\varepsilon)}{u_M(\varepsilon)} \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \quad (\text{D.7})$$

i.e., the remainder is smaller than the last term included, once ε is sufficiently small (Hinch, 1991). Often the terms $u_N(\varepsilon)$ are powers of ε multiplied by some coefficient, i.e., $f \sim \sum_{i=1}^N a_N \varepsilon^N$ which implies that $\varepsilon < 1$. Therefore, asymptotic series are also called *asymptotic power series*, or *semi convergent series* (Arfken, 1985). Both Arfken (1985) and Hinch (1991) show examples in which partial sums of a series are very good approximations to the (original) functions, while the complete series expansions diverge.

[Note the similarity between uniform convergence (above) and asymptotic approximation; on one hand, uniform convergence implies asymptotic approximation, but not vice versa (i.e., uniform convergence is more restrictive); on the other hand, uniform convergence does not depend on a small parameter, ϵ , as does asymptotic approximation. This distinction is essential for the truncation of Neumann and Volterra series, where practically, the asymptotic representation is more crucial than the convergence of the series, and where ϵ is analogous to σ_Y^2 (the variance of log-conductivity).]

APPENDIX E

SYMMETRY OF THE RANDOM GREEN'S FUNCTION

According to (3.12) – (3.14), $\mathcal{G}(\mathbf{x}', \mathbf{x})$ is the solution of

$$\nabla_{\mathbf{x}'} \cdot [K(\mathbf{x}') \nabla_{\mathbf{x}'} \mathcal{G}(\mathbf{x}', \mathbf{x})] + \delta(\mathbf{x}' - \mathbf{x}) = 0 \quad \mathbf{x}', \mathbf{x} \in \Omega \quad (\text{E.1})$$

subject to the homogeneous boundary conditions at \mathbf{x}' on Γ . Analogously, $\mathcal{G}(\mathbf{x}, \mathbf{x}')$ is the solution of

$$\nabla \cdot [K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}, \mathbf{x}')] + \delta(\mathbf{x} - \mathbf{x}') = 0 \quad \mathbf{x}, \mathbf{x}' \in \Omega \quad (\text{E.2})$$

subject to the homogeneous boundary conditions at \mathbf{x} on Γ . Multiplying (E.2) by $\mathcal{G}(\mathbf{x}', \mathbf{x})$, integrating over Ω , applying Green's identity twice, and rearranging gives

$$\begin{aligned} & \int_{\Omega} \mathcal{G}(\mathbf{x}, \mathbf{x}') \nabla \cdot [K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}', \mathbf{x})] d\mathbf{x} + \int_{\Omega} \delta(\mathbf{x} - \mathbf{x}') \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x} \\ &= \int_{\Gamma} K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \mathcal{G}(\mathbf{x}, \mathbf{x}') d\mathbf{x} - \int_{\Gamma} K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}, \mathbf{x}') \cdot \mathbf{n}(\mathbf{x}) \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x} \end{aligned} \quad (\text{E.3})$$

$$(\text{E.4})$$

Substituting (E.2), together with the corresponding homogeneous boundary conditions, into (E.4) yields

$$\begin{aligned} & \int_{\Omega} \{ \mathcal{G}(\mathbf{x}, \mathbf{x}') \nabla \cdot [K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}', \mathbf{x})] - \nabla \cdot [K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}, \mathbf{x}')] \mathcal{G}(\mathbf{x}', \mathbf{x}) \} d\mathbf{x} \\ &= \int_{\Gamma_N} K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}', \mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \mathcal{G}(\mathbf{x}, \mathbf{x}') d\mathbf{x} - \int_{\Gamma_D} K(\mathbf{x}) \nabla \mathcal{G}(\mathbf{x}, \mathbf{x}') \cdot \mathbf{n}(\mathbf{x}) \mathcal{G}(\mathbf{x}', \mathbf{x}) d\mathbf{x} \end{aligned} \quad (\text{E.5})$$

$$(\text{E.6})$$

Since this must hold for arbitrary $K(\mathbf{x})$, Ω , Γ_N , and Γ_D , each of the integrands in (E.6) must vanish independently of the rest. This clearly happens if and only if

$$\mathcal{G}(\mathbf{x}, \mathbf{x}') = \mathcal{G}(\mathbf{x}', \mathbf{x}) \quad \mathbf{x}, \mathbf{x}' \in \bar{\Omega} \quad (\text{E.7})$$

The symmetry of \mathcal{G} is thus seen to follow from the self-adjoint nature of the operator $\nabla \cdot [K(\mathbf{x})\nabla]$ under homogeneous Dirichlet and Neumann boundary conditions.

APPENDIX F

RELATIONSHIPS BETWEEN COVARIANCES OF K AND Y

Let $v(\mathbf{x})$ be an unbiased estimate of $Y(\mathbf{x}) \equiv \ln K(\mathbf{x})$ and let $\sigma_Y^2(\mathbf{x}) = \langle [Y'(\mathbf{x})]^2 \rangle_v$ be the conditional variance of the estimation error $Y'(\mathbf{x}) = Y(\mathbf{x}) - v(\mathbf{x})$.

Then

$$K(\mathbf{x}) = \exp[Y(\mathbf{x})] = \exp[v(\mathbf{x}) + Y'(\mathbf{x})] = K_g(\mathbf{x}) \exp[Y'(\mathbf{x})], \quad (\text{F.1})$$

where

$$K_g(\mathbf{x}) = \exp[v(\mathbf{x})]. \quad (\text{F.2})$$

If $Y'(\mathbf{x})$ is Gaussian then (Mood and Graybill, 1963)

$$\kappa(\mathbf{x}) \equiv \langle K(\mathbf{x}) \rangle_v = K_g(\mathbf{x}) \langle \exp[Y'(\mathbf{x})] \rangle_v = K_g(\mathbf{x}) \exp\left(\frac{1}{2}\sigma_Y^2(\mathbf{x})\right), \quad (\text{F.3})$$

and hence

$$K'(\mathbf{x}) \equiv K(\mathbf{x}) - \kappa(\mathbf{x}) = -K_g(\mathbf{x}) \left(\exp\left(\frac{1}{2}\sigma_Y^2(\mathbf{x})\right) - \exp[Y'(\mathbf{x})] \right). \quad (\text{F.4})$$

It follows that

$$\begin{aligned} \langle K'(\mathbf{x})K'(\mathbf{y}) \rangle_\kappa &= \\ & K_g(\mathbf{x})K_g(\mathbf{y}) \langle \{ \exp[\sigma_Y^2(\mathbf{x})/2] - \exp[Y'(\mathbf{x})] \} \{ \exp[\sigma_Y^2(\mathbf{y})/2] - \exp[Y'(\mathbf{y})] \} \rangle_v \\ &= K_g(\mathbf{x})K_g(\mathbf{y}) \{ \langle \exp[Y'(\mathbf{x}) + Y'(\mathbf{y})] \rangle_v - \langle \exp[Y'(\mathbf{x})] \rangle_v \exp[\sigma_Y^2(\mathbf{y})/2] \\ & \quad - \langle \exp[Y'(\mathbf{y})] \rangle_v \exp[\sigma_Y^2(\mathbf{x})/2] + \exp[\sigma_Y^2(\mathbf{x})/2 + \sigma_Y^2(\mathbf{y})/2] \} \\ &= K_g(\mathbf{x})K_g(\mathbf{y}) \exp\left(\frac{1}{2}[\sigma_Y^2(\mathbf{x}) + \sigma_Y^2(\mathbf{y})]\right) \{ \exp\langle Y'(\mathbf{x})Y'(\mathbf{y}) \rangle_v - 1 \} \\ &= \kappa(\mathbf{x})\kappa(\mathbf{y}) \{ \exp\langle Y'(\mathbf{x})Y'(\mathbf{y}) \rangle_v - 1 \}, \end{aligned} \quad (\text{F.5})$$

where $\langle \cdot \rangle_v$ designates moments conditioned on $v(\mathbf{x})$.

APPENDIX G

MATHERON'S PROOF Re 2-D LOGNORMAL FIELDS

We consider the case of two-dimensional flow governed locally by

$$\nabla \cdot [K(\mathbf{x})\nabla h(\mathbf{x})] = 0. \quad (\text{G.1})$$

The stream function $\psi(\mathbf{x})$ is known to satisfy (cf., Frind and Matanga, 1985) an analogous equation

$$\nabla \cdot [\hat{K}(\mathbf{x})\nabla\psi(\mathbf{x})] = 0, \quad (\text{G.2})$$

where $\hat{K}(\mathbf{x}) \equiv K(\mathbf{x})^{-1}$. These two equations can be rewritten, respectively, as

$$\nabla^2 h(\mathbf{x}) + \nabla Y(\mathbf{x}) \cdot \nabla h(\mathbf{x}) = 0, \quad (\text{G.3})$$

$$\nabla^2 \psi(\mathbf{x}) + \nabla \hat{Y}(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}) = 0, \quad (\text{G.4})$$

where $\hat{Y}(\mathbf{x}) \equiv \ln[K(\mathbf{x})^{-1}]$. Since the latter is equal to $-Y(\mathbf{x})$, (G.4) can also be written as

$$\nabla^2 \psi(\mathbf{x}) - \nabla Y(\mathbf{x}) \cdot \nabla \psi(\mathbf{x}) = 0. \quad (\text{G.5})$$

In coordinates \mathbf{y} that are rotated by 90° relative to \mathbf{x} , (G.5) takes the form

$$\nabla^2 \psi(\mathbf{y}) + \nabla Y(\mathbf{y}) \cdot \nabla \psi(\mathbf{y}) = 0. \quad (\text{G.6})$$

Since $Y(\mathbf{x})$ is Gaussian, $Y'(\mathbf{x}) = Y(\mathbf{x}) - \langle Y \rangle$ and $\hat{Y}'(\mathbf{x}) = \hat{Y}(\mathbf{x}) - \langle \hat{Y} \rangle$ differ only by sign, otherwise they are distributed symmetrically about zero. Hence $Y'(\mathbf{x})$

and $\hat{Y}'(\mathbf{x})$ have identical moments of all order and are thus identical in probability. Hence random fluctuations in $\psi(\mathbf{y})$ about its mean, caused solely by variations in $Y(\mathbf{y})$, must be statistically identical to fluctuations in $h(\mathbf{x})$ about its mean. Taking ensemble moments of (G.3) and (G.6) thus reveals that when $\langle h(\mathbf{x}) \rangle = (J, 0, 0)^T$ then $\langle \nabla \psi(\mathbf{y}) \rangle = (C, 0, 0)^T$ where C is a constant.

The local form of Darcy's law,

$$\mathbf{q}(\mathbf{x}) = -K(\mathbf{x})\nabla h(\mathbf{x}) \quad (\text{G.7})$$

has an equivalent (dual) form

$$\hat{\mathbf{q}}(\mathbf{x}) = -\hat{K}(\mathbf{x})\nabla \psi(\mathbf{x}) \quad (\text{G.8})$$

The corresponding effective hydraulic conductivities are defined, respectively, as

$$K_e \equiv -\frac{\langle q_1(\mathbf{x}) \rangle}{\partial \langle h(\mathbf{x}) \rangle / \partial x_1} = \frac{\langle K(\mathbf{x}) \partial h(\mathbf{x}) / \partial x_1 \rangle}{\partial \langle h(\mathbf{x}) \rangle / \partial x_1}, \quad (\text{G.9})$$

$$\hat{K}_e \equiv -\frac{\langle \hat{q}_1(\mathbf{x}) \rangle}{\partial \langle \psi(\mathbf{x}) \rangle / \partial x_1} = \frac{\langle \hat{K}(\mathbf{x}) \partial \psi(\mathbf{x}) / \partial x_1 \rangle}{\partial \langle \psi(\mathbf{x}) \rangle / \partial x_1}. \quad (\text{G.10})$$

These definitions imply, in analogy to $\hat{K}(\mathbf{x}) = K(\mathbf{x})^{-1}$, that

$$\hat{K}_e = \frac{1}{K_e}. \quad (\text{G.11})$$

From (3.48) – (3.49) it follows that

$$K_e = \langle K \rangle - \int_{\Omega_\infty} \langle K'(0)K'(x') \partial^2 \mathcal{G}(x', 0) / \partial (x'_1)^2 \rangle dx', \quad (\text{G.12})$$

$$\hat{K}_e = \langle \hat{K} \rangle - \int_{\Omega_\infty} \langle \hat{K}'(0)\hat{K}'(x') \partial^2 \hat{\mathcal{G}}(x', 0) / \partial (x'_1)^2 \rangle dx', \quad (\text{G.13})$$

where $\hat{\mathcal{G}}(\mathbf{x}, 0)$ is the random Green's function associated with (G.2), i.e., the solution of (G.2) under the sole action of a unit point source at the origin. Let $k \equiv K / \langle K \rangle$, $\hat{k} \equiv \hat{K} / \langle \hat{K} \rangle$, $\eta \equiv \langle K \rangle \mathcal{G}$, $\hat{\eta} \equiv \langle \hat{K} \rangle \hat{\mathcal{G}}$ so that η and $\hat{\eta}$ satisfy

$$\nabla \cdot [k(\mathbf{x})\nabla \eta] + \delta = 0 \quad \nabla \cdot [\hat{k}(\mathbf{x})\nabla \hat{\eta}] + \delta = 0, \quad (\text{G.14})$$

where δ is the Dirac delta. It follows that (G.12) – (G.13) can be rewritten as

$$K_e = \langle K \rangle \left(1 - \int_{\Omega_\infty} \langle k'(0)k'(\mathbf{x}') \partial^2 \eta(\mathbf{x}', 0) / \partial(x'_1)^2 \rangle d\mathbf{x}' \right), \quad (\text{G.15})$$

$$\hat{K}_e = \langle \hat{K} \rangle \left(1 - \int_{\Omega_\infty} \langle \hat{k}'(0)\hat{k}'(\mathbf{x}') \partial^2 \hat{\eta}(\mathbf{x}', 0) / \partial(x'_1)^2 \rangle d\mathbf{x}' \right). \quad (\text{G.16})$$

Since $K(\mathbf{x})$ and $\hat{K}(\mathbf{x})$ are log normal, $k(\mathbf{x})$ and $\hat{k}(\mathbf{x})$ have identical distributions and, by virtue of (G.14), so do η and $\hat{\eta}$. Hence the integrals in (G.15) and (G.16) are identical, and we have

$$\hat{K}_e = \frac{\langle \hat{K} \rangle}{\langle K \rangle} K_e. \quad (\text{G.17})$$

Combining this with (G.11) yields

$$K_e = \sqrt{\langle K \rangle / \langle K^{-1} \rangle}. \quad (\text{G.18})$$

Equation (F.3) implies that $\langle K \rangle = K_g \exp(\sigma_Y^2/2)$ and $\langle K^{-1} \rangle = K_g^{-1} \exp(\sigma_Y^2/2)$. It therefore follows immediately that

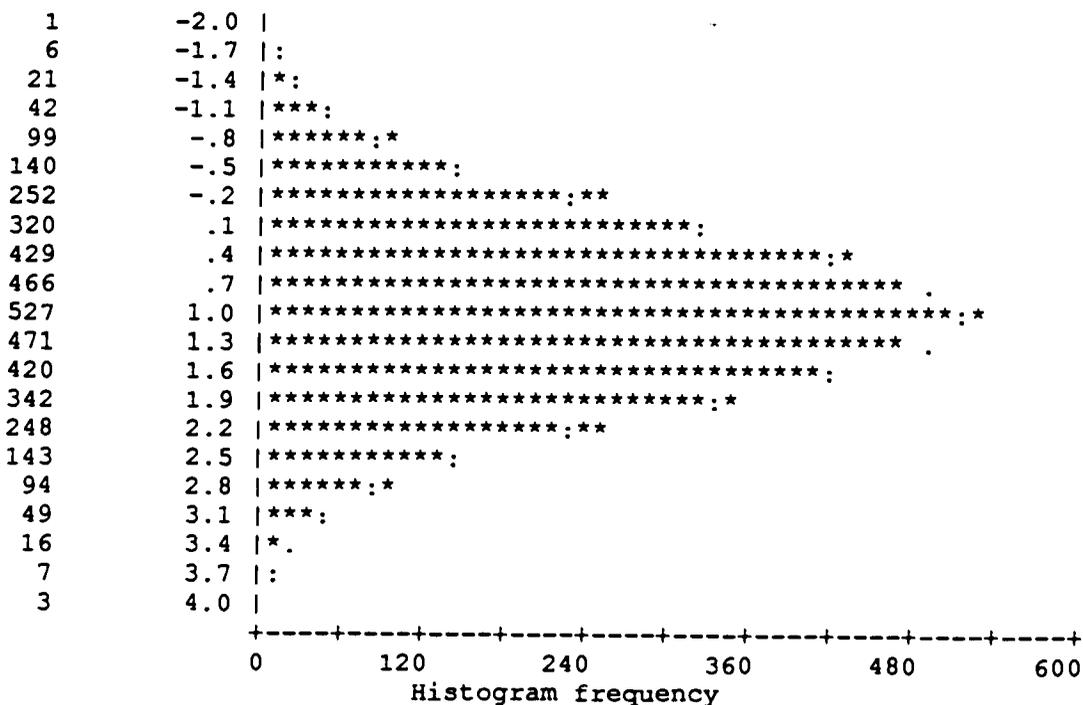
$$K_e = K_g. \quad (\text{G.19})$$

APPENDIX H

SPSSX OUTPUT: ANALYSIS OF A RANDOM FIELD

K

Count Midpoint One symbol equals approximately 12.00 occurrences



| | | | | | |
|----------|-------|----------|----------|----------|--------|
| Mean | 1.003 | Std err | .015 | Median | 1.002 |
| Mode | 1.046 | Std dev | .942 | Variance | .887 |
| Kurtosis | -.208 | S E Kurt | .076 | Skewness | .013 |
| S E Skew | .038 | Range | 5.860 | Minimum | -1.889 |
| Maximum | 3.971 | Sum | 4106.298 | | |

Valid cases 4096 Missing cases 0

(continues)

K

Valid cases: 4096.0 Missing cases: .0 Percent missing: .0

| | | | | | | | |
|---------|--------|----------|-------|-------|---------|----------|--------|
| Mean | 1.0025 | Std Err | .0147 | Min | -1.8890 | Skewness | .0125 |
| Median | 1.0019 | Variance | .8866 | Max | 3.9712 | S E Skew | .0383 |
| 5% Trim | 1.0019 | Std Dev | .9416 | Range | 5.8602 | Kurtosis | -.2083 |
| | | | | IQR | 1.2987 | S E Kurt | .0765 |

M-Estimators

| | | | |
|------------------------------|--------|---------------------|--------|
| Huber (1.339) | 1.0013 | Tukey (4.685) | 1.0010 |
| Hampel (1.700, 3.400, 8.500) | 1.0012 | Andrew (1.340 * pi) | 1.0009 |

(continues)

K

| Frequency | Bin Center | | Frequency | Stem & | Leaf |
|-------------|-------------|-------|-----------|----------|----------------|
| 5.00 | Extremes | | 5.00 | Extremes | (-1.9), (-1.8) |
| 58.00 | -1.50000 | | | Extremes | (-1.7) |
| 549.00 | -.50000 | ***** | 11.00 | -1 f | & |
| 1434.00 | .50000 | ***** | 17.00 | -1 t | & |
| 1445.00 | 1.50000 | ***** | 29.00 | -1 * | 0& |
| 537.00 | 2.50000 | **** | 47.00 | -0 . | 89 |
| 58.00 | 3.50000 | | 83.00 | -0 s | 677 |
| 10.00 | Extremes | | 92.00 | -0 f | 445 |
| | | | 132.00 | -0 t | 22233 |
| Bin width : | 1.00000 | | 196.00 | -0 * | 0000111 |
| Each star: | 121 case(s) | | 214.00 | 0 * | 0001111 |
| | | | 278.00 | 0 t | 2222233333 |
| | | | 295.00 | 0 f | 44444555555 |
| | | | 293.00 | 0 s | 6666677777 |
| | | | 354.00 | 0 . | 8888889999999 |
| | | | 335.00 | 1 * | 00000011111 |
| | | | 325.00 | 1 t | 222222333333 |
| | | | 307.00 | 1 f | 44444555555 |
| | | | 250.00 | 1 s | 666667777 |
| | | | 228.00 | 1 . | 88889999 |
| | | | 172.00 | 2 * | 000011 |
| | | | 149.00 | 2 t | 22233 |
| | | | 84.00 | 2 f | 455 |
| | | | 86.00 | 2 s | 667 |
| | | | 46.00 | 2 . | 89 |
| | | | 34.00 | 3 * | 0& |
| | | | 17.00 | 3 t | & |
| | | | 7.00 | 3 f | & |
| | | | 7.00 | Extremes | (3.6), (3.8) |
| | | | 3.00 | Extremes | (3.8), (3.9) |

Stem width: 1.00000

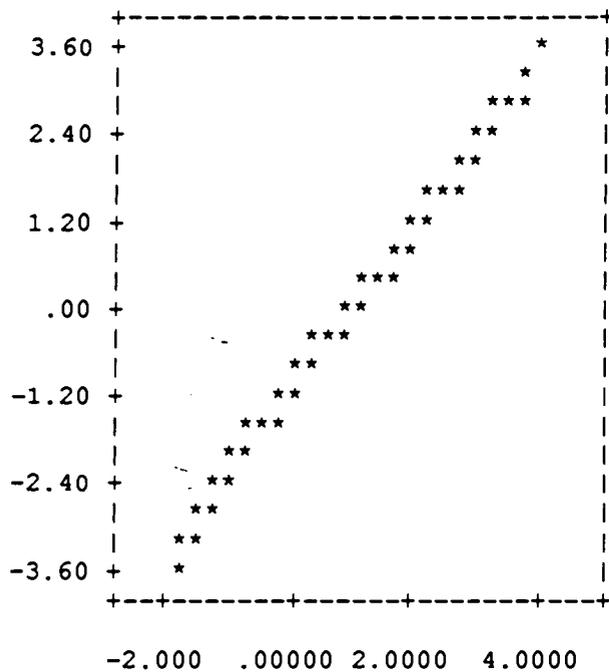
Each leaf: 28 case(s)

& denotes fractional leaves.

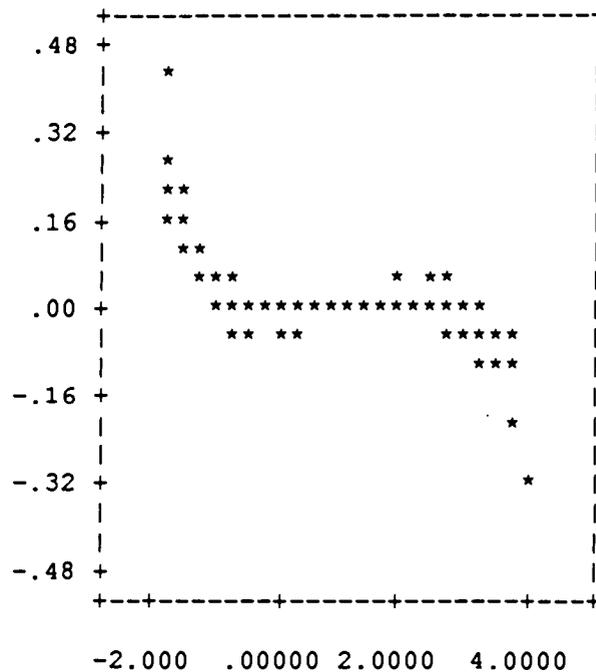
(continues)

Extreme Values

| 5 | Highest | Case # | 5 | Lowest | Case # |
|---|---------|----------|---|----------|----------|
| | 3.97120 | CASE3414 | | -1.88900 | CASE2131 |
| | 3.90020 | CASE681 | | -1.82780 | CASE2265 |
| | 3.89950 | CASE550 | | -1.79050 | CASE1408 |
| | 3.80770 | CASE421 | | -1.78660 | CASE1466 |
| | 3.78430 | CASE549 | | -1.65760 | CASE63 |



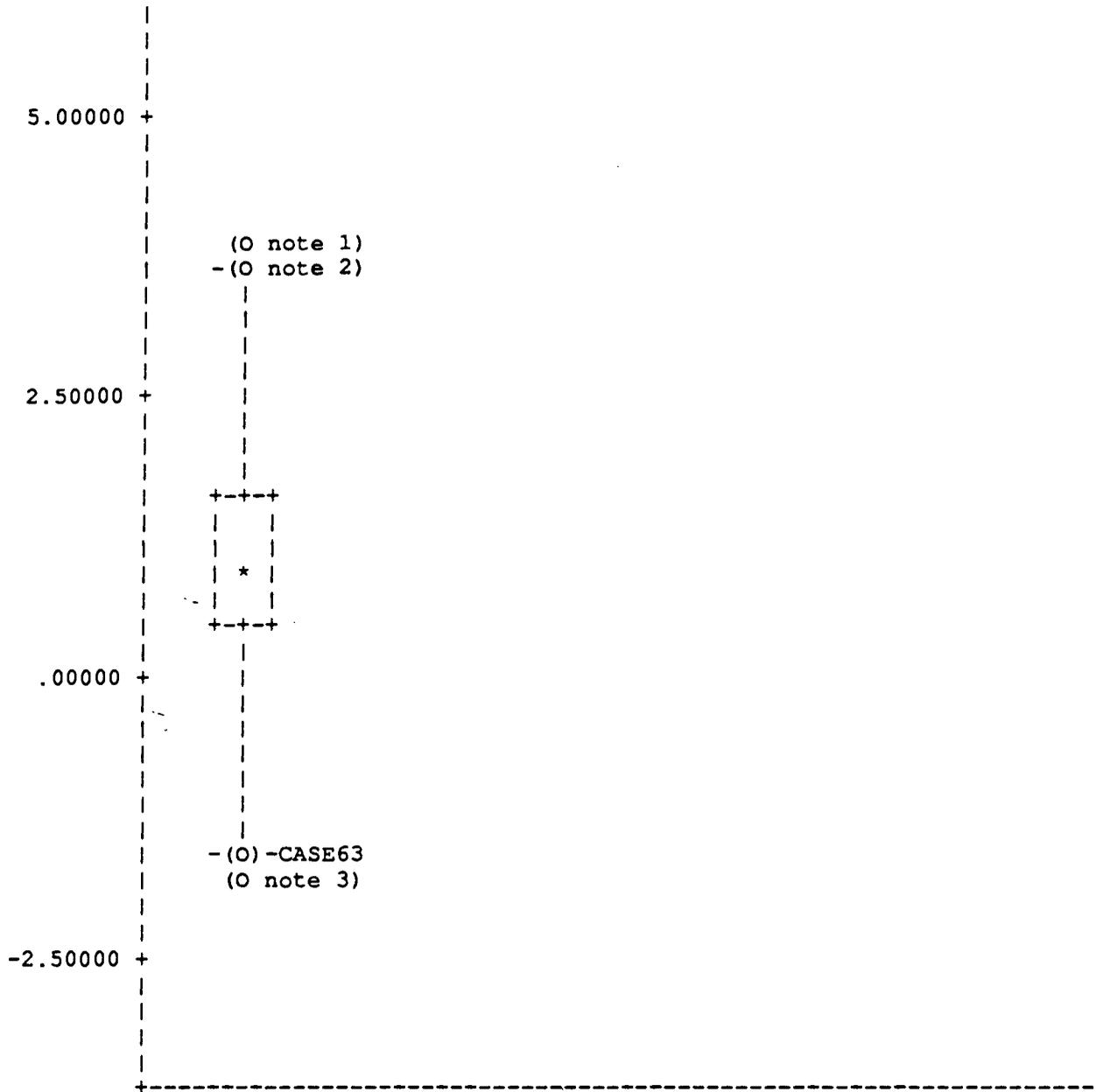
Normal Plot



Detrended Normal Plot

| | Statistic | df | Significance |
|------------------|-----------|------|--------------|
| K-S (Lilliefors) | .0108 | 4096 | > .2000 |

(continues)



Variables K
 N of Cases 4096.00

Symbol Key: * - Median (O) - Outlier (E) - Extreme

LIST OF REFERENCES

- Ababou, R., Three-Dimensional Flow in Random Porous Media, Ph.D. dissertation, Mass. Inst. of Technol., Cambridge, 1988.
- Ababou, R., Effective conductivity in randomly stratified aquifers: some recent developments, *EOS Trans.*, 70, p. 1112, 1989.
- Ababou, R., D. McLaughlin, L.W. Gelhar, and A.F.B. Tompson, Numerical simulation of three-dimensional saturated flow in randomly heterogeneous porous media, *Transport in Porous Media*, 4, 549-565, 1989.
- Ababou, R., Identification of effective conductivity tensor in randomly heterogeneous and stratified aquifers, *5th Canadian-American Conference on Hydrogeology*, Sept. 18-20, Calgary, Alberta, Canada, 1990.
- Ababou, R. and E.F. Wood, Comment on "Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge" by J.J. Gomez-Hernandez and S.M. Gorelick, *Water Resour. Res.*, 26(8), 1843-1846, 1990a.
- Ababou, R. and E.F. Wood, Correction to "Comment on 'Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge' by J.J. Gomez-Hernandez and S.M. Gorelick", *Water Resour. Res.*, 26(12), 2945, 1990b.
- Ababou, R. and L.W. Gelhar, Self-similar randomness and spectral conditioning: analysis of scale effects in subsurface hydrology, *Dynamics of fluids in hierarchical porous media*, edited by J.H. Cushman, Academic Press, 393-428, 1990.

- Abramowitz, M. and I.A. Stegun , *Handbook of Mathematical Functions*, Dover Publications, 1972.
- Adomian, G., *Stochastic Systems*, Academic Press, New York, 1983.
- Arbogast, T., J. Douglas, Jr., and U. Hornung, Derivation of the double porosity model of single phase flow via homogenization theory, *SIAM Jour. Math. Anal.*, 21(4), 823-836, 1990.
- Ahmed, N.U., *Semigroup Theory with Applications to Systems and Control*, Longman Scientific & Technical with John Wiley & Sons, New York, 1991.
- Alabert, F., The practice of fast conditional simulations through the LU decomposition of the covariance matrix, *Math. Geology*, 19(5), 369-386, 1987.
- Arfken, *Mathematical Methods for Physicists*, 3rd. ed. Academic Press, 1985
- Auriault, J.L., Heterogeneous medium. Is an equivalent macroscopic description possible?, *Int. J. Engng. sci.* 29(7), 785-795, 1991.
- Avellanda M., and A.J. Majda, Mathematical models with exact renormalization for turbulent transport, *Comm. Math. Phys.*, 131, 381- 429, 1990.
- Babuška, I., Feedback, adaptivity, and a-posteriori estimates in finite elements: aims, theory, and experience, in *Accuracy Estimates and Adaptivity for Finite Elements*, J. Wiley and Sons, New York, 1986.
- Babuška I. and A. K. Noor, *Quality assessment and control of finite element solutions*, Tech. Report BN-1049, Institute for Physical Science and Technology, University of Maryland, 1986.
- Babuška I. and W. C. Rheinboldt, Error estimates for adaptive finite element computations, *SIAM J. Numer. Anal.*, 15 (1978), pp. 736-754.

- Bachu, S., and D. Cuthiell, Effects of core-scale heterogeneity on dsteady-state and transient fluid flow in porous media: Numerical analysis, *Water Resour. Res.* 26(5), 863-874, 1990.
- Bank R. E., Analysis of a local a-posteriori error estimate for elliptic equations, in *Accuracy Estimates and Adaptivity for Finite Elements*, 119-128, J. Wiley and Sons, New York, 1986.
- Bank R. E., A posteriori error estimates, adaptive local mesh refinement, and multigrid iteration, in *Multigrid Methods II: Proceedings*, Cologne 1985 (Lecture Notes in Mathematics 1228), 7-23, Springer-Verlag, Heidelberg, 1986.
- Bank R. E., J. F. Bürgler, W. Fichtner, and R. K. Smith, Some upwinding techniques for finite element approximations of convection-diffusion equations, *Numer. Math.*
- Bank R. E., and T. F. Chan, PLTMGC: A multi grid continuation program for parameterized nonlinear elliptic systems, *SIAM J. Sci. Stat. Comput.*, 7, 540-559, 1986.
- Bank R. E., and T. F. Dupont, *Analysis of a two level scheme for solving finite element equations*, Tech. Report CNA-159, Center for Numerical Analysis, University of Texas at Austin, 1980.
- Bank R. E., T. F. Dupont, and H. Yserentant, The hierarchical basis multigrid method, *Numer. Math.*, 52, 427-458, 1988.
- Bank R. E., and H. D. Mittelmann, Continuation and multigrid for nonlinear elliptic systems, in *Multigrid Methods II: Proceedings*, Cologne 1985 (Lecture Notes in Mathematics 1228), Springer-Verlag, Heidelberg, 24-38, 1986.

- Bank R. E., and H. D. Mittelmann, Stepsize selection in *Continuation procedures and damped Newton's method*, *J. Comp. and Appl. Math.*, *26*, 67-78, 1989.
- Bank R. E. and D. J. Rose, *Global approximate Newton methods*, *Numer. Math.*, *37*, pp. 279-295, 1981.
- Bank R. E. and D. J. Rose, *A multi-level Newton method for nonlinear finite element equations*, *Math. Comp.*, *39*, 453-465, 1982.
- Bank R. E. and D. J. Rose, *On the complexity of sparse Gaussian elimination via bordering*, *SIAM J. Sci. Stat. Comput.*, *11*, 145-160, 1990.
- Bank R. E., A. H. Sherman, and A. Weiser, *Refinement algorithms and data structures for regular local mesh refinement*, in *Scientific Computing*, Imacs/North Holland, 3-17, 1983.
- Bank R. E., and R. K. Smith, *General sparse elimination requires no permanent integer storage*, *SIAM J. Sci. Stat. Comput.*, *8*, 574-584, 1987.
- Bank R. E., and A. Weiser, *Some a-posteriori error estimators for elliptic partial differential equations*, *Math. Comp.*, *44*, 283-301, 1985.
- Bank R. E., and H. Yserentant, *Some remarks on the hierarchical basis multi-grid method*, in *Domain Decomposition Methods*, 140-146, SIAM, Philadelphia, 1988.
- Bank, R.E., *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, User's Guide 6.0*, SIAM, Philadelphia, PA, 1990.
- Bakr, A.A., *Effect of Spatial Variations of Hydraulic Conductivity on Groundwater Flow*, Ph.D. dissertation, N. M. Inst. of Min. and Technol., Socorro, 1976.

- Bakr, A.A., L.W. Gelhar, A.L. Gutjahr, and J.R. MacMillan, Stochastic analysis of spatial variability in subsurface flows 1. Comparison of one- and three dimensional flows, *Water Resour. Res.*, 14(2), 263-271, 1978.
- Barr, D.R, and N.L. Slezak A comparison of multivariate normal generators, *Commun. Assoc. Comput. Mach.*, 15, 1048-1049, 1972.
- Batchelor, G.K., Transport properties of two phase materials with random structure, *Annual Reviews of Fluid Mechanics*, 6, 227-254, 1974.
- Baveye, P. and G. Sposito, The operational significance of the continuum hypothesis in the theory of water movement through soils and aquifers, *Water Resour. Res.*, 20(5), 521-530, 1984.
- Baveye, P. and G. Sposito, Macroscopic balance equations in soils and aquifers: The case of space- and time-dependent instrumental response, *Water Resour. Res.*, 21(8), 1116-1120, 1985.
- Bear, J., *Dynamics of Fluids in Porous Media*, American Elsevier, 1972.
- Bear, J., *Hydraulics of Groundwater*, McGraw-Hill, 1979.
- Bècus, G.A., Successive approximations solutio of a class of random equations, in *Approximate Solutions of Random Equations*, edited by A.T. Bharucha-Reid, North Holland, New York, 1978.
- Bellman, R.E., Stochastic transformations and functional equations. In *Stochastic Processes in Mathematical Physics and Engineering*, edited by R.E. Bellman, Vol.XVI, 171-177, Amer. Math. Soc., 1963.
- Benaroya, H. and Rehak, M., The Neumann series/born approximation applied to

- parametrically excited stochastic systems. *Probabilistic Engineering Mechanics*, 2(2), 1987.
- Bensoussan, A., J.L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures in Mathematics and its Applications*, 5, North-Holland, Amsterdam, 1978.
- Beran, M.J., Use of variational approach to determine bounds for the effective permittivity of a random medium, *Nuovo Cimento*, 38, 771-782, 1965.
- Beran, M.J., *Statistical Continuum Theories*, Interscience, New York, 1968.
- Beran, M.J. and J.J. McCoy, Mean field variation in random media, *Quart. Appl. Math.*, 28, 245, 1970.
- Bharucha-Reid, A.T., *Random Integral Equations*, Mathematics in Science and Engineering, Vol.96, Academic Press, 1972.
- Bleistein, N. and R.A. Handelsman, *Asymptotic Expansions of Integrals*, Dover Publications, 1986.
- Bourgeat, A., Homogenized behavior of two-phase flows in naturally fractured reservoirs with uniform fracture distribution, *Comput. Methods. Appl. Mech. Eng.*, 47, 205-216, 1984.
- Brenner, H., A general theory of Taylor dispersion phenomena, *PCH, Physicochem. Hydrodyn.*, 1, 91-123, 1980.
- Brenner, H., A general theory of Taylor dispersion phenomena, II, An extension. *PCH, Physicochem. Hydrodyn.*, 3(2), 139-157, 1982.
- Briggs, W.L., *A Multigrid Tutorial*, SIAM, Philadelphia, 1987.

- Carr, J, and D.E. Myers, COSIM: A Fortran IV program for co-conditional simulation, *Computers & Geosciences*, 11(6), 675-705, 1985.
- Cheng, A.H.-D. and O.E. Lafe, Boundary element solution for stochastic groundwater flow: Random boundary condition and recharge, *Water Resour. Res.*, 27(2), 231-242, 1991.
- Christakos, G., *Random Field Models in Earth Sciences*, Academic Press, 1992.
- Clauser, C., Permeability of crystalline rocks, *EOS, Trans. AGU*, 73 (21), 1992.
- Clifton, P.M., Statistical Inverse Modeling and Geostatistical Analysis of the Avra Valley Aquifer, M.S. thesis, University of Arizona, 1981.
- Clifton, P.M. and S.P. Neuman, Effects of Kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in Southern Arizona, *Water Resour. Res.*, 18(4), 1215-1234, 1982.
- Cole, C.R., and H.P. Foote, Multigrid methods for solving multiscale transport problems. In *Dynamics of fluids in hierarchical porous media*, edited by J.H. Cushman, Academic Press, 1-5, 1990.
- Collatz, L., *Functional Analysis and Numerical mathematics*, Academic Press, New York, 1966.
- Craft B.C., and M.F. Hawkins, *Applied Petroleum Reservoir Engineering*, 437 pp., Prentice-Hall, Englewood Cliffs, N. J., 1959.
- Curtain, R., and Pritchard, A.J., *Functional Analysis in Modern Applied Mathematics*, Academic Press, 1977.
- Cushman, J. H., Volume averaging, probabilistic averaging, and ergodicity, *Adv. Water Resour.*, 6, 182-184, 1983.

- Cushman, J.H., On unifying the concepts of scale, instrumentation, and stochastics in the development of multiphase transport theory, *Water Resour. Res.*, 20(11), 1668-1676, 1984.
- Cushman, J.H., On measurement, scale, and scaling, *Water Resour. Res.*, 22(2), 129-134, 1986.
- Cushman, J.H., Development of stochastic partial differential equations for subsurface hydrology, *Stochastic Hydrology and Hydraulics*, 1(4), 241-262, 1987.
- Cushman, J.H., An introduction to hierarchial porous media. *Dynamics of fluids in hierarchical porous media*, edited by J.H. Cushman, Academic Press, 1-5, 1990.
- Cushman, J.H., On the inseparability of the measurement process from theories of transport in heterogeneous materials, in *Concise Encyclopedia of Environmental Systems*, edited by P. Young, in press, 1991.
- Dagan, G., Comment on "A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media" by R.A. Freeze, *Water Resour. Res.*, 12, 567, 1976.
- Dagan, G., Models of groundwater flow in statistically homogeneous porous formations, *Water Resour. Res.*, 15(1), 47-63, 1979.
- Dagan, G., Analysis of flow through heterogeneous random aquifers by the method of embedding matrix 1. Steady flow, *Water Resour. Res.*, 17(1), 107-121, 1981.
- Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities 1. Conditional simulation and the direct problem, *Water Resour. Res.*, 18(4), 813-833, 1982a.

- Dagan, G., Analysis of flow through heterogeneous random aquifers, 1. Steady flow in confined formations, *Water Resour. Res.*, 18(5), 1541-1555, 1982.
- Dagan, G., Analysis of flow through heterogeneous random aquifers 2. Unsteady flow in confined formations, *Water Resour. Res.*, 18(5), 1571-1585, 1982b.
- Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities: The inverse problem, *Water Resour. Res.*, 21(1), 65-72, 1985.
- Dagan, G., Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale, *Water Resour. Res.*, 22(9), 120S-134S, 1986.
- Dagan, G., *Flow and Transport in Porous Formations*, Springer-Verlag, New York, 1989.
- Dagpunar J., *Principles of Random Variate generation*, Clarendon Press – Oxford Science publications, 1988.
- Davis, M.W., Production of conditional simulations via the LU decomposition of the covariance matrix, *Math. Geology*, 19(2), 91-98, 1987.
- Dederichs, P.H., and R. Zeller, Variational treatment of elastic constants of disordered materials, *Z. Phys.*, 259, 103-116, 1973.
- Delhomme, J.P., Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach, *Water Resour. Res.*, 15(2), 269-280, 1979.
- Desbarats, A.J., Numerical estimation of effective permeability in sand-shale formations, *Water Resour. Res.*, 23(2), 273-286, 1987a.

- Desbarats, A.J., Stochastic Modeling of Flow in Sand-Shale Sequences, Ph.D. dissertation, Stanford University, 1987b.
- Desbarats, A.J., Spatial averaging of hydraulic conductivity in three-dimensional heterogeneous porous media, *Math. Geol.*, Vol. 24(3), 249-267, 1992a.
- Desbarats, A.J., Spatial averaging of transmissivity in heterogeneous fields with flow toward a well, *Water Resour. Res.*, 28(3), 757-767, 1992b.
- Desbarats, A.J. and R. Dimitrakopoulos, Geostatistical modeling of transmissibility for 2D reservoir studies, *SPE Formation Evaluation*, 5(4), 437-443, 1990.
- Deutsch, C., Calculating effective absolute permeability in sandstone/shale sequences, *SPE Form. Eval.*, Sept. 343-348, 1989.
- Deutsch, C.V., and A.G. Journel, *GSLIB: Geostatistical Software Library and User Guide*, Oxford University Press, 1992).
- Dupuy, M.P., Some new mathematical approaches for heterogeneous porous medium flow studies, *SPE paper NO. 1840*, presented at the 42nd annual fall meeting, *Soc. Pet. Eng.*, Houston, Tex., Oct. 1-4, 1967.
- Durlofsky, L.J., Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media, *Water Resour. Res.*, 27(5), 699-708, 1991.
- Dykaar, B.B. and P.K. Kitanidis, Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach 1. Method, *Water Resour. Res.*, 28(4), 1155-1166, 1992a.
- Dykaar, B.B. and P.K. Kitanidis, Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach 1. Results, *Water Resour. Res.*, 28(4), 1167-1178, 1992b.

- Elishakoff, L., *Probabilistic Methods in the Theory of Structures*, John Wiley & Sons, New York, 1983.
- El-Kadi, A.I. and W. Brutsaert, Applicability of effective parameters for unsteady flow in nonuniform aquifers, *Water Resour. Res.*, 21(2), 183-198, 1985.
- Follin, S., Numerical Calculations on Heterogeneity of Groundwater Flow, Ph.D. dissertation, The Royal Inst. of Technol., Stockholm, 1992.
- Freeze, R.A., A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res.*, 11(5), 725-741, 1975.
- Fletcher, C.A.J., *Computational Techniques for Fluid Dynamics*, Springer-Verlag, New York, 1988.
- Frind, E.O. and G.B. Matanga, The dual formulation of flow for contaminant transport modeling. 1. Review of theory and accuracy aspects, *Water Resour. Res.*, 21(2), 159-169, 1985.
- Frisch, U., Wave propagation in random media, in *Probabilistic Methods in Applied Mathematics*, edited by A.T. Bharucha-Reid, 1, 75-198, Academic Press, New York, 1968.
- Gelhar, L.W., Stochastic analysis of flow in aquifers, Advances in groundwater hydrology, *Proceedings, Symposium of the Amer. Water Resources Assoc.*, Minneapolis, Minnesota, 1976.
- Gelhar, L.W. and C.L. Axness, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, 19(1), 161-180, 1983.
- Gelhar, L.W., Stochastic subsurface hydrology. From theory to applications, *Water Resour. Res.*, 22(9), 135S-145S, 1986.

- Gelhar, L. W., Stochastic analysis of solute transport in saturated and unsaturated porous media, in *Advances in Transport Phenomena in Porous Media*, NATO ASI Ser., edited by J. Bear and M. Y. Corapcioglu, 657-700, Martinus Nijhoff, Dordrecht, Netherlands, 1987.
- Ghanem, R.G., and P.D. Spanos, *Stochastic Finite Elements: A spectral Approach*. Springer-Verlag, New York, 1991.
- Gomez-Hernandez, J.J., A Stochastic Approach to the Simulation of Block Conductivity Fields Conditioned upon Data Measured at a Smaller Scale, Ph.D. dissertation, Stanford University, 1991.
- Gomez-Hernandez, J.J. and S.M. Gorelick, Effective groundwater model parameter values: Influence of spatial variability of hydraulic conductivity, leakance, and recharge, *Water Resour. Res.*, 25(3), 405-419, 1989.
- Gomez-Hernandez, J.J. and S.M. Gorelick, Reply, *Water Resour. Res.*, 26(8), 1847-1848, 1990.
- Gradsteyn, I.S. and I.M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 1980.
- Greenberg, M.D., *Application of Green's Functions in Science and Engineering*, Prentice-Hall, 1971.
- Greenberg, M.D., *Foundations of Applied Mathematics*, Prentice-Hall, 1978.
- Gutjahr, A.L., L.W. Gelhar, A.A. Bakr, and J.R. MacMillan, Stochastic analysis of spatial variability in subsurface flows 2. Evaluation and application, *Water Resour. Res.*, 14(5), 953-959, 1978.

- Gutjahr, A.L. and L.W. Gelhar, Stochastic models of subsurface flow: Infinite versus finite domains and stationarity, *Water Resour. Res.*, 17(2), 337-350, 1981.
- Gutjahr, A.L., Fast Fourier transforms for random field generation, project report for Los Alamos Grant to New Mexico Tech., Contract no. 4-R58-2690R, Dept. of Mathematics, New Mexico Tech., Socoro, NM, 1989.
- Haberman, R., *Elementary Applied Partial Differential Equations*, Prentice-Hall, 1987.
- Harter, T., Conditional simulation: A comprehensive review of some important numerical techniques, unpublished manuscript, 1992.
- Hashin, Z., and S. Shtrikman, A variational approach to the theory of effective magnetic permeability of multiphase materials, *Jour. of Appl. Physics*, 33, 3125-3131, 1962.
- Hoeksema, R.J. and P.K. Kitanidis, Analysis of the spatial structure of properties of selected aquifers, *Water Resour. Res.*, 21(4), 563-572, 1985.
- Hinch, E.J., *Perturbation Methods*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 1991.
- Hoeksema, R.J. and P.K. Kitanidis, Analysis of spatial structure of properties of selected aquifers, *Water Resour. Res.*, 21, 563-572, 1985.
- Hsieh, P.A., S.P. Neuman, and E.S. Simpson, Pressure Testing of Fractured Rocks — A Methodology Employing Three-Dimensional Cross-Hole Tests, *Top. Rep. NUREG/CR-3213*, U.S. Nuclear Regulatory Commission, Washington, D.C., 1983.

- Hille, E. and R. Phillips, *Functional Analysis and Semi-Groups*, 2nd ed., Amer. Math. Soc. Providence, Rhode Island, 1957.
- IMSL, *International Mathematical and Statistical Library*, Edition 2.1, Houston, Texas, IMSL Ltd., 1989.
- Isaaks, E.H. and R.M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, 1989.
- Jackson, J.D., *Classical Electrodynamics*, John Wiley & Sons, New York, 1962.
- Johnson, M.E., *Multivariate Statistical Simulation*, John Wiley & Sons, New York, 1987.
- Journel, A.G., Geostatistics for conditional simulation of orebodies, *Econ. Geol.*, 69(5), 673-687, 1974.
- Journel, A.G. and Ch.J. Huijbregts, *Mining Geostatistics*, Academic Press, 1978.
- Journel, A.G., *Indicator Approach to Toxic Chemical site*, Report of project n. CR-811235-02-0, EPA, EMSL, Las-Vegas, 1984.
- Kadanoff, L.P., Scaling and Universality in statistical physics, *Physica A*, 163, 1-14, 1990.
- Kasap, E., and L.W. Lake, Calculating the effective permeability tensor of a grid block, *SPE Form. Eval.*, June, 192-200, 1990.
- Keller, J.B., Wave propagation in random media, *Proc. Symp. Appl. Math.*, 13, 227-246, Amer. Math. Soc., Providence, Rhode Island, 1962.
- Keller, J.B., Stochastic equations and wave propagation in random media, *Proc. Symp. Appl. Math.*, 16, 145-170, Amer. Math. Soc., Providence, Rhode

- Island, 1964.
- Keller, J.B., Effective conductivity of reciprocal media, in *Random Media*, edited by G. Papanicolaou, The IMA Vol. in Math. and its Appl., Vol.7, Springer-Verlag, New York, 1987.
- King, P.R., The use of field theoretic methods for the study of flow in a heterogeneous porous medium, *J. Phys. A: Math. Gen.*, 20, 3935-3947, 1987
- King, P.R., The use of renormalization for calculating effective permeability, *Transport in Porous Media*, 4, 37-58, 1989.
- Kirkpatrick, S., Percolation and Conduction, *Reviews of Modern Physics*, 45(4), 574-588, 1973.
- Kitanidis, P.K., Effective hydraulic conductivity for gradually varying flow, *Water Resour. Res.*, 26(6), 1197-1208, 1990.
- Knopp, K., *Theory and Applications of Infinite Series*, Hafner Pub. Co., 1947.
- Knuth, D.E., Seminumerical Algorithms, in *The Art of Computer Programming*, 2nd edn., Vol. 2, Addison-Wesley, Reading, Mass., 1981.
- Koch, D. L., and J. F. Brady, A non-local dispersion of advection- diffusion with application to dispersion in porous media, *J. Fluid Mech.*, 180, 387-403, 1987.
- Koch, D. L., and J. F. Brady, Anomalous diffusion in heterogeneous porous media, *Phys. Fluids*, 31, 965-973, 1988.
- Kohler, W.E., and G. Papanicolaou, Bounds for the effective conductivity of random media, in *Macroscopic properties of Disordered Media*, edited by R. Burridge, S. Childress, and G. Papanicolaou, *Lecture Notes in Physics*, 154, Springer-Verlag, New York, 1982. 111-130

- Kossack, C.A., and S.T. Opdal, Scaling up heterogeneities with pseudofunctions, *SPE, Form. Eval.*, Sept. 1990.
- Kraichnan, R.H., The structure of isotropic turbulence at very high Reynolds numbers, *Jour. Fluid. Mech.*, 5, 497-543, 1959.
- Kraichnan, R.H., Dynamics of nonlinear stochastic systems, *J. Mathematical Phys.*, 2, 124-148, 1961.
- Kraichnan, R.H., The closure problem of turbulence theory, *Proc. Symp. Appl. Math.*, 13, 199-225, Amer. Math. Soc., Providence, Rhode Island, 1962.
- Kraichnan, R.H., Eddy viscosity and diffusivity: exact formulas and approximations, *Complex Systems*, 1, 805-820, 1987.
- Kröner, E., Bounds on effective elastic moduli of disordered materials, *J. Mech. Phys. Solids*, 25, 137-155, 1977.
- Kröner, E., Statistical modelling, in *Modelling Small Deformations of Polycrystals*, edited by J. Dittus and J. Zarka, 229-291, Elsevier, London, 1986.
- Lamb, H., *Hydrodynamics*, Cambridge Univ. Press, New York, 1932.
- Landau, L.D. and E.M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press, Oxford, England, 1960.
- Landauer, R., Electrical conductivity in inhomogeneous media, in *Electrical Transport and Optical Properties of Inhomogeneous Media*, edited by J.C. Garland and D.B. Tanner, AIP Conference Proceedings, No. 40, American Institute of Physics, 1978.
- Lumely, J., and A. Panofski, *The Structure of Atmospheric Turbulence*. John Wiley & Sons, New York, 1964.

- Maclaren, N.M., The generation of multiple independent sequences of pseudorandom numbers, *Appl. Statist., Royal Statist. Soc.*, 38(2), 351-359, 1989.
- Matheron, G., *Elements Pour une Theorie des Milieux Poreux*, 166 pp., Masson et Cie, Paris, 1967.
- Matheron, G., The intrinsic random functions and their applications, *Advan. Appl. Prob.*, 5, 439-468, 1973.
- Mantoglou, A. and J.L. Wilson, The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resour. Res.*, 18(5), 1379-1394, 1982.
- Markov, K.Z., Application of Volterra-Wiener series for bounding the overall conductivity of heterogeneous media. I. General procedure, *SIAM J. of Appl. Math.*, 4, 831-849.
- de Marsily, G., *Quantitative Hydrogeology*, Academic Press, 1986.
- Matheron, G., *Elements pour une Theorie des Milieux Poreux*, Masson et Cie, Paris, 1967.
- Mei, C.C. and J.-L. Auriault, Mechanics of heterogeneous porous media with several spatial scales, *Proc. R. Soc. London, Ser. A*, 426, 391-423, 1989.
- Mizell, S.A., A.L. Gutjahr, and L.W. Gelhar, Stochastic analysis of spatial variability in two-dimensional steady groundwater flow assuming stationary and nonstationary heads, *Water Resour. Res.*, 18, 1053-1067, 1982.
- Mood, A.M. and F.A. Graybill, *Introduction to the Theory of Statistics*, McGraw-Hill, New York, 443 p., 1963.

- Morgan, R.C. and I. Babuska, An approach for constructing families of homogenized equations for periodic media. I: An integral representation and its consequences, *SIAM Jour. Math. Anal.*, 22(1), 1-15, 1990.
- Morse, P.M., and H. Feshbach, *Methods of Theoretical Physics*, MacGraw-Hill, New York, 1953.
- Myers, D.E., A series of lecture notes on conditional simulations, in *Advanced Geostatistics* course, University of Arizona, 1990.
- Myers, D.E., Vector conditional simulation, in *Geostatistics, Vol. I*, edited by A. Armstrong, 283-293, 1989.
- Naff, R.L., Radial flow in heterogeneous porous media: An analysis of specific discharge, *Water Resour. Res.*, 27(3), 307-316, 1991.
- Naff, R.L., and A.V. Vecchia, Stochastic analysis of three-dimensional flow flow in a bounded domain, *Water Resour. Res.*, 22(5), 695-704, 1986.
- NAG *Numerical Algorithms Group*, Mark 14, routine G05AAF, Oxford: Numerical Algorithm Group, 1974-1992.
- Nashed, M.Z., and H.W. Engle, Random generalized inverse and approximate solutions of random operator equations, in *Approximate Solutions of Random Equations*, edited by A.T. Bharucha-Reid, North Holland, New York, 1978.
- Neuman, S.P., Statistical characterization of aquifer heterogeneities: an overview. In *Recent trends in hydrology*, edited by T.N. Narasimhan, Special Paper 189, 81-102, Geol. Soc. Am. Spec. Pap. 1982.
- Neuman, S.P., Role of geostatistics in subsurface hydrology, in *NATO ASI Series, Geostatistics for Natural Resources Characterization*, edited by G. Verly, M.

David, A.G. Journel, and A. Marechal, D. Reidel,

Neuman, S.P. and E.A. Jacobson, Analysis of nonintrinsic spatial variability by residual kriging with application to regional groundwater levels, *Math. Geol.*, 16(5), 499-521, 1984. Hingham, Mass., 2, 127-140, 1984.

Neuman, S.P., Stochastic continuum representation of fractured rock permeability as an alternative to the REV and fracture network concepts, *28th US Symposium on Rock Mechanics*, edited by I.W. Farmer, J.J.K. Daemen, C.S. Desai, C.E. Glass, and S.P. Neuman, Balkema, Rotterdam, The Netherlands, 533-561, 1987.

Neuman, S.P., C.L. Winter, and C.M. Newman, Stochastic theory of field-scale Fickian dispersion in anisotropic porous media, *Water Resour. Res.* 23(3), 453-466, 1987.

Neuman, S.P. and J.S. Depner, Use of variable-scale pressure test data to estimate the log hydraulic conductivity covariance and dispersivity of fractured granites near Oracle, Arizona, *Journal of Hydr.*, 102, 475-501, 1988.

Neuman, S. P., Universal scaling of hydraulic conductivities and dispersivities in geologic media, *Water Resour. Res.*, 26(8), 1749-1758, 1990.

Neuman, S. P., Reply to Comments by M. Anderson on "Universal scaling of hydraulic conductivities and dispersivities in geologic media", *Water Resour. Res.*, 27(6), 1383-1384, 1991.

Neuman, S.P., F.J. Samper, and M. Hernandez, Reply to comment by G. Gambolati and G. Galeati on "Analysis of nonintrinsic variability by residual kriging with application to regional groundwater levels, *Math. Geol.*, 19(3), 259-266, 1987.

- Neuman, S.P., S. Orr, O. Levin, and E. Paleologos, Theory and high-resolution finite element analysis of 2-D and 3-D effective permeabilities in strongly heterogeneous porous media, in *Computational Methods in Water Resources IX*, edited by T.F. Russell, R.E. Ewing, C.A. Brebbia, W.G. Gray, and G.F. Pinder, Elsevier Applied Science, New York, 1992.
- Neuman, S.P. and S. Orr, Prediction of steady state flow in nonuniform geologic media by conditional moments: exact nonlocal formalism and weak approximation, *Water Resour. Res.*, 29(2), 341-364, 1993.
- Neuman, S.P., O. Levin, S. Orr, E. Paleologos, D. Zhang, and Y. -K. Zhang Nonlocal representations of subsurface flow and transport by conditional moments, *Computational Stochastic Mechanics*, edited by A.H.D. Cheng, and C.Y. Young, Computational Mechanics Publications, 1992.
- Neuman, S.P., Eulerian-Lagrangian theory of transport in space-time nonstationary velocity fields: 1. Exact nonlocal formalism by conditional moments and weak approximation, *Water Resour. Res.*, 29(3), 619-631, 1993.
- Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.
- Paleologos, E., Effective Hydraulic Conductivity of Bounded, Strongly Heterogeneous Porous Media, Ph.D. dissertation, University of Arizona, 1993.
- Poley, A.D., Effective permeability and dispersion in locally heterogeneous aquifers, *Water Resour. Res.*, 24, 1921-1926, 1988.
- Porter, D, and D.S.C. Stirling, *Integral Equations, A Practical Treatment from Spectral Theory to Applications*, Cambridge Texts in Applied Mathematics, Cam-

- bridge University Press, 1990.
- Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge, 1986.
- Priestley, M.B., *Spectral Analysis and Time Series*, Academic Press, 1989.
- Rheinboldt, W. C., Error estimates and adaptive techniques for nonlinear parameterized equations, in *Accuracy Estimates and Adaptivity for Finite Elements*, J. Wiley and Sons, New York, 163-180, 1986.
- Rheinboldt, W. C., *Error questions in the computation of solution manifolds of parameterized equations*, Tech. Report ICMA-87-112, Institute for Computational Mathematics and Applications, University of Pittsburgh, 1987.
- Roberts, P.H., Analytical theory of turbulent diffusion, *Jour. Fluid. Mech.*, 11, 257-283, 1961.
- Rubin, Y. and G. Dagan, Stochastic analysis of boundaries effects on head spatial variability in heterogeneous aquifers 1. Constant head boundary, *Water Resour. Res.*, 24(10), 1689-1697, 1988.
- Rubin, Y. and G. Dagan, Stochastic analysis of boundaries effects on head spatial variability in heterogeneous aquifers 2. Impervious boundary, *Water Resour. Res.*, 25(4), 707-712, 1989.
- Rubin, Y., and Gómez-Hernández, J.J., A stochastic approach to the problem of upscaling of conductivity in disordered media: theory and unconditional numerical simulations, *Water Resour. Res.*, 26(4), 691-701, 1990.

- Russo, D., and Bresler, E., Scaling soil hydraulic properties of a heterogeneous field. *Soil Sci. Soc. Amer. J.*, 44, 691-684, 1980.
- Saez, A.E., C.J. Otero, and I. Rusinek, The effective homogeneous behavior of heterogeneous porous media, *Transport in Porous Media*, 4, 213-238, 1989.
- Saffman, P.G., On the boundary condition at the surface of a porous medium, *Studies in Applied Math.*, L(2), 93-101, 1971.
- Samper, F.J. and S.P. Neuman, Estimation of spatial covariance structures by adjoint state maximum likelihood cross validation, 1. Theory, *Water Resour. Res.*, 25(3), 351-362, 1989a.
- Samper, F.J. and S.P. Neuman, Estimation of spatial covariance structures by adjoint state maximum likelihood cross validation, 3. Application to hydrochemical and isotopic data, *Water Resour. Res.*, 25(3), 373-384, 1989b.
- Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley & Sons, New York, 1980.
- Serrano, S.E., T.E. Unny, and W.C. Lennox, Analysis of stochastic groundwater flow problems, 1, Deterministic partial differential equations in groundwater flow, a functional-analytic approach, *Journal of Hydr.*, 82(3-4), 247-263, 1985a.
- Serrano, S.E., T.E. Unny, and W.C. Lennox, Analysis of stochastic groundwater flow problems, 2, Stochastic partial differential equations in groundwater flow, a functional-analytic approach, *Journal of Hydr.*, 82(3-4), 265-284, 1985b.
- Serrano, S.E., T.E. Unny, and W.C. Lennox, Analysis of stochastic groundwater flow problems, 3, Approximate solutions of stochastic partial differential equations, *Journal of Hydr.*, 82(3-4), 285-306, 1985c.

- Serrano, S.E., Stochastic differential equations models of erratic infiltration, *Water Resour. Res.*, 26(4), 703-711, 1990.
- Serrano, S.E., Semianalytical methods in stochastic groundwater transport, *Appl. Math. Modelling*, 16, 1992.
- Sharp, W.E., and C. Bays, 1992, A review of portable random number generators, *Computers & Geosciences*, 18(1), 79-87, 1992.
- Smith, J.L., spatial variability of flow parameters in a stratified sand, *Math. Geol.*, 13(1), 1-21, 1981.
- Smith, L. and R.A. Freeze, Stochastic analysis of steady state groundwater flow in a bounded domain 1. One- dimensional simulations, *Water Resour. Res.*, 15(3), 521-528, 1979a.
- Smith, L. and R.A. Freeze, Stochastic analysis of steady state groundwater flow in a bounded domain 1. dimensional simulations, *Water Resour. Res.*, 15(6), 1543-1559, 1979b.
- Spiegel, M.R., *Advanced Mathematics*, Schaum's Outline Series, 1971.
- Stakgold, I., *Green's Functions and Boundary Value Problems*, John Wiley & Sons, 1979.
- Shvidler, M.I., Filtration Flows in Heterogeneous Media (in Russian), *Izv. Akad. Nauk SSSR Mech.*, 3, 185-190, 1962.
- Tartar, L., Remarks on homogenization, *Homogenization and Effective Moduli in Materials and Media*, edited by J.L. Ericksen, D. Kinderlehrer, R. Kohn, and J.L. Lions, The IMA Vol. in Math. and its Appl., Vol.7, 228-246, Springer-Verlag, New York, 1986.

- Tartar, L., Nonlocal effects induced by homogenization, in *Partial differential equations and the calculus of variations, Vol. 2, Essays in honor of Ennio De Giorgi*, edited by F. Colombini, A. Marino, L. Modica, and S. Spagnolo, *Progress in nonlinear differential equations and their applications*, 925-938, Birkhäuser, Boston, 1989.
- Tompson, A.F., R. Ababou and L.W. Gelhar, Implementation of the three-dimensional turning bands random field generator, *Water Resour. Res.*, 25(10), 2227-2243, 1989.
- Tricomi, F.G., *Integral Equations*, Dover Publications, Inc., New York, 1957.
- Tsokos C.P., and W.J. Padgett, Random Integral Equations with Applications to Life Sciences and Engineering, *Mathematics in Science and Engineering*, 108, Academic Press, New York, 1974.
- Winter C.L., C.M. Newman, and S.P. Neuman, A perturbation expansion for diffusion in a random velocity field, *SIAM J. Appl. Math.*, 44(2), 411-424, 1984.
- Weiser, A., *Local-Mesh, Local-Order Adaptive Finite Element Methods with A-Posteriori Error Estimators for Elliptic Partial Differential Equations*, PhD thesis, Yale University, 1981.
- Unny, T.E., Stochastic partial differential equations in groundwater hydrology, *Stochastic Hydrol. and Hydraul.*, 3, 135-153, 1989.
- Van Kampen, N. G., Stochastic differential equations, *Phys. Rep.*, 24C, 171-228, 1976.
- Van Kampen, N. G., *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1981.

- Vanmarcke, E.H., *Random Fields: Analysis and Synthesis*, MIT Press, Cambridge, Mass., 1983.
- Yeh, T.C.J, L.W. Gelhar, and A.L. Gutjahr, Stochastic analysis of unsaturated flow in heterogeneous soils, 1. statistically isotropic media. *Water Resour. Res.*, 21:447-456, 1985.
- Yserentant, H., Hierarchical bases of finite element spaces in the discretization of nonsymmetric elliptic boundary value problems, *Computing*, 35, 39-49, 1985.
- Yserentant, H., On the multi-level splitting of finite element spaces for indefinite elliptic boundary value problems, *SIAM J. Numer. Anal.*, 23, 581-595, 1986.
- Zeitoun, D.G. and C. Braester, A Neumann expansion approach to flow through heterogeneous formations, *Stochastic Hydrology and Hydraulics*, 5, 207-226, 1991.
- Zimmerman, D.A., and J.L. Wilson, *TUBA: A computer code for generating two-dimensional random fields via the turning bands method*, A user guide, SeaSoft, Albuquerque, NM, 1990.