UNIVERSITY OF TULSA
THE GRADUATE SCHOOL

CONDITIONING GEOSTATISTICAL MODELS TO TWO-PHASE FLOW PRODUCTION DATA

by
Zhan Wu

A dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy
in the Department of Petroleum Engineering
1999

UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

CONDITIONING GEOSTATISTICAL MODELS TO TWO-PHASE FLOW PRODUCTION DATA

by

Zhan Wu

A DISSERTATION

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Dissertation Committee

_____ Co-Chairperson

_____ Co-Chairperson

_____

_____

ABSTRACT

Zhan Wu (Doctor of Philosophy in Petroleum Engineering)

CONDITIONING GEOSTATISTICAL MODELS TO TWO-PHASE FLOW PRODUCTION DATA

(126 pp.-Chapter V)

Co-Directed by Drs. Albert C. Reynolds and Dean S. Oliver

(88 words)

A discrete adjoint method for generating sensitivity coefficients related to two-phase flow production data is derived. The procedure is applied to calculate the sensitivity of wellbore pressure and water-oil-ratio to reservoir simulator gridblock permeabilities and porosities. Using these sensitivity coefficients, an efficient form of the Gauss-Newton algorithm is applied to generate maximum a posteriori estimates and realizations of the rock property fields conditioned to a prior geostatistical model and pressure and/or water-oil ratio data obtained under two-phase (oil and water) flow conditions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would first like to express my sincere appreciation and gratitude to my thesis advisors Dr. Albert C. Reynolds, Jr., Professor of Petroleum Engineering and Mathematical Sciences of the University of Tulsa, and Dean S. Oliver, Assistant Professor of Petroleum Engineering of the University of Tulsa for their valuable assistance and guidance during this study. I also thank Dr. Leslie G. Thompson, Associate Professor of Petroleum Engineering of the University of Tulsa, and Dr. Brenton S. McLaury, Assistant Professor of Mechanical Engineering of the University of Tulsa for participating in my thesis committee and for their comments and suggestions. I would like to extend my thanks and appreciation to all other faculty members in the Petroleum Engineering and to all the staff members in the McFarlin Library, who provided me an excellent study atmosphere as a TU graduate student.

Appreciation is also extended to all my friends and my graduate student colleagues for their useful discussion and encouragement. I gratefully acknowledge the financial support from TUPREP (Tulsa University Petroleum Reservoir Exploitation Projects) during my study.

I am especially gratefully to my wife Lei Wang and my beloved son Alex Wu for their love, support, and encouragement.

CHAPTER I

INTRODUCTION

1.1 Background and Literature Review

To the best of our knowledge Jacquard and Jain [28] presented the first procedure for numerically computing sensitivity coefficients for history matching purposes. They applied their method to the estimation of permeability in a two-dimensional reservoir from pressure data. The procedure was based on an electric circuit analogue, but later, following Jacquard and Jain's basic ideas (also see Jacquard [27] and Jahns [29]), Carter et al. [9] presented an elegant derivation of a method to compute sensitivity coefficients for two-dimensional single-phase flow problems. Carter's procedure for single-phase flow has been extended to three-dimensional problems in a computationally efficient way by He et al. [23]. For nonlinear problems, e.g., multiphase flow problems, the derivations of Carter et al. [9] and He et al. [23] do not apply. Thus, we are forced to seek other alternatives. One possible choice is the adjoint or optimal control method, introduced independently for the single-phase history matching problem by Chen et al. [11] and Chavent et al. [10]. (For linear problems, Carter et al. [8] have shown that the adjoint method is equivalent to the procedure of Carter et al. [9]. Although the adjoint method has been applied to multiphase flow problems (see, for example, Lee and Seinfeld [33], Yang et al. [56] and Makhlouf et al. [34]), these implementations have proved to be computationally inefficient. Specifically, the procedure has normally been applied only to compute the gradient of an objective function based on a sum of the squares of the production data mismatch term.

Limited to this information, one is forced to minimize the objective function using a steepest descent, variable-metric, or conjugate gradient algorithm, which results in slow convergence. The sensitivity coefficients needed to form the Hessian for Newton or Gauss-Newton iteration are not available from the traditional adjoint systems, and even if they were, the Hessian would typically be ill-conditioned because the history matching problem is generally ill-posed. Because of the computational cost of repeatedly solving the adjoint system for many iterations (see, for example, Makhlouf et al. [34] in which history matching of a two-phase (water-oil) 450 cell reservoir model required 6400 CPU seconds on a CRAY X-MP/48), the complete rigorous adjoint (or optimal control) method has not been applied to history match production data for multiphase flow problems in real reservoirs. For example, when Wasserman et al. [54] applied an optimal control approach to history match a two-dimensional Saudi Arabian oil field, they actually computed adjoint variables (Lagrange multipliers) only for the overall pressure equation and used an objective function based only on pressure mismatch terms, other production data, e.g., water-oil ratios were not included in the objective function.

Another popular choice for computing sensitivity coefficients for multiphase flow problems is the method known in the petroleum engineering literature as the gradient simulator method. This method was introduced to the petroleum engineering literature by Anterion et al. [2], but was known earlier in the ground water hydrology literature as the sensitivity coefficient method; see, for example, Yeh's [57] review of parameter identification methods. In the gradient simulator method, one computes the sensitivity of pressures and saturations to model parameters at the end of a simulator time-step by solving a linear system obtained by differentiating the matrix form of the finite-difference equations with respect to a model parameter, e.g., a gridblock value of permeability or porosity. The resulting system of equations can then be solved to obtain the sensitivity of all gridblock pressures and gridblock saturations to this particular model parameter. (From these sensitivities, one can easily construct

other sensitivity coefficients, e.g., the sensitivity of wellbore pressure and water-oil-ratio to model parameters.) The advantage of the gradient simulator method is that the matrix problem solved to obtain these sensitivity coefficients involves the same coefficient matrix as the one used to solve for pressures and saturations at this time step. Moreover, the coefficient matrix does not depend on the model parameters; only the right hand side of the matrix problem depends on the model parameters. Thus, the problem reduces to solving a matrix problem with multiple right-hand side vectors, one right-hand side vector, for each model parameter. The difficulty is that if we wish to estimate (or construct realizations of) permeabilities and porosities at several thousand gridblocks, then we have several thousand right-hand sides. The number of right-hand sides is equal to the number of model parameters to be estimated. Although Killough et al. [31] developed a faster iterative solver for this problem, the effort required for each sensitivity is still of the order of 10% of a forward simulation. Of course, if the reservoir model has only a few parameters, e.g., some form of zonation ([28], [4]) is used, or the model can be reparameterized without losing significant information, the gradient simulator becomes more attractive. Effectively zonation is a reparameterization procedure in which one applies history matching to estimate permeability (or porosity) multipliers with the same multiplier applied to all grid block permeabilities within a particular zone. The history matching procedure using by Hegstad and Omre [25] is essentially a zonation procedure, albeit one embedded within a formal Bayesian estimation procedure. In their paper, history matching is done by adjusting the mean log-permeability of each layer instead of adjusting individual grid block permeabilities. It our belief that the subspace introduced into the geophysics literature Kennett and Williamson [30] (also see Oldenberg et al. [37] and Oldenberg and Li [36]) provides a reparameterization method that can significantly improve computational efficiency of simulation and estimation procedures based on inverse problem theory [52]. Reynolds et al. [47] have applied the subspace method to estimate or simulate permeability and porosity fields by history matching multi-

well single-phase flow pressure data. However, the subspace method has not been implemented to history match multiphase flow production data.

In the gradient simulator method, one actually obtains the sensitivity of all gridblock pressures (or saturations) to each model parameter. For the most part, this is useless information, as one typically only has measurements at wells. In essence, it is the computation of this useless information that detracts from the computational efficiency of the gradient simulator method. To avoid this inefficiency, Chu et al. [12] used the basic ideas of Tang et al. [51] to develop a modified generalized pulse-spectrum technique (MGPST) to estimate the sensitivity of wellbore pressure to reservoir simulator gridblock permeabilities and porosities. This method is effectively an approximation of the gradient simulator method, an approximation in which one computes only the sensitivity of well gridblock pressures to the rock property fields. The MGPST yields reasonably accurate estimates of sensitivity coefficients related to the permeability field, but, unfortunately does not yield accurate values of sensitivity coefficients related to the porosity field. Ref. [17] contains a review of parameter estimation methods including some discussion of methods for calculating sensitivity coefficients.

For the purpose of this work, history matching will refer to the determination of model parameters (reservoir simulator grid block values of permeability and porosity) which honor production data, i.e., observed wellbore pressures and water oil ratios (or water cuts). In the typical case, there are several thousand model parameters and the number of independent pressure data is far less than the number of model parameters. Thus, the data is not sufficient to determine unique gridblock values of permeability and porosity. In such cases, there exist an infinite set of reservoir descriptions which will generate predicted production data which agrees with the observed data. In mathematical terms, the inverse problem of determining the rock property fields from the production data is ill-posed, i.e., does not have a unique solution. To specify a particular solution (estimate of the rock property fields) requires

some form regularization; see, for example, Tikhonov [53] or Tang et al. [51]. Here, we also refer to the specification of additional constraints in the form of prior knowledge (see Menke [35] or Tarantola [52]), or the reduction in the number of model parameters by zonation (see Jacquard and Jain [28] or Bissel [4]), and as regularization. If the number of model parameters is greater than the number of independent data and one attempts to perform history matching by minimizing the objection function defined as the sum of production data mismatch terms squared by a Newton method, the Hessian matrix generated at each iteration would be singular. Somewhat like a Levenberg-Marquardt algorithm (see Fletcher [19] or Ref. [46]), regularization gives rise to a Hessian matrix which is nonsingular. The disadvantage of regularization is that the particular estimate of model parameters generated depends on the specific form of regularization chosen. In our work, we prefer to use a prior geostatistical model as a regularization term; see, for example, Tarantola [52] and Oliver [39] and He et al. [24]. Conceptually, this prior geostatistical model can be generated from static data (all available data and geologic knowledge or interpretation except production data).

As in Tarantola et al. [52], we use an approach based on probability theory. Specifically, we assume that probability density function (pdf) for the prior model is a multivariate Gaussian distribution. Then by applying Bayes' theorem, the a posteriori pdf can be determined and has a particularly simple form. The a posteriori pdf represents the pdf for the model parameters based on both production data and prior information so realizations (samples) of this pdf represent realizations conditioned to both prior information and data. The maximum a posteriori estimate of the model parameters can be obtained by minimizing a particular objective function (see Tarantola [52]) and an approximate sampling of the a posteriori pdf can be obtained using a method proposed in the petroleum engineering literature by Oliver et al. [42]; also see Kitanidis [32]. Here the procedure given in Oliver et al. [42] is referred to as the randomized maximum likelihood method. If an objective function

is minimized by a Newton procedure, then the generation of the Hessian matrix requires knowledge of the sensitivity coefficients. The maximum a posteriori estimate of the rock property fields is in general too smooth to be a plausible realization as it is the mean of a Gaussian distribution; see Tarantola [52]. The a posteriori pdf represents the pdf for the model parameters based on both production data and prior information. Thus, realizations (samples) of this pdf represent realizations conditioned to both prior information and production data. The uncertainty in model parameters can sometimes be estimated from the a posteriori covariance matrix (see Tarantola [52]), but our main goal is to estimate the uncertainty in predicted reservoir performance. To do so, it is important to first generate a suite of model realizations via a correct sampling of the a posteriori pdf. Then, for the proposed operating conditions, predicted performance for each realization of the rock property fields is computed using a reservoir simulator. By constructing statistics (mean, variance, histogram) for the set of outcomes for each predicted parameter (e.g., cumulative oil production, water oil ratios or break through times), one can estimate the uncertainty in the predicted parameters. Reynolds et al. [48] have applied this specific procedure to the case where geostatistical realizations from a prior model were conditioned to multiwell pressure data.

The procedure mentioned above for characterizing the uncertainty in reservoir performance predictions is reliable only if the set of realizations of the rock property fields correctly reflects the uncertainty in grid block permeabilities and porosities, i.e., represents a correct sampling of the a posteriori pdf. Thus, sampling this pdf properly is of paramount importance. Having said this, it is important to note that the randomized maximum likelihood method provides only an approximate sampling except in the case where data is linearly related to the model see, Oliver [40], Oliver et al. [42] and Reynolds et al. [48]. If the data is linearly related to the model, then the randomized maximum likelihood method provides a rigorous sampling procedure. Although a Markov chain Monte Carlo method provides a theoretically rigorous pro-

cedure to sample the a posteriori pdf correctly, current implementations of these procedures are not sufficiently computationally efficient for practical application; see, for example, Hegstad et al. [26], Oliver et al. [41] and Cunha et al. [14].

As a historical note, many of the ideas currently under investigation for regularized automatic history matching are rooted in the fundamental work of Gavalas et al. [20] and Shah et al. [50]. Although only one-dimensional single-phase flow problems were considered, the work of these authors laid the ground work for much of the research work being done today on automatic history matching.Gavalas et al. recognized that the proper incorporation of prior information or a prior model stabilizes the history-matching problem and reduces the variability in the possible values of grid block rock properties that provide an acceptable history match of pressure data. They used Gaussian type expressions for the covariance functions for permeability and porosity and their cross covariance and prior estimates of the means of porosity and permeability to incorporate prior information in the objective function. They showed that the incorporation of prior information reduced the errors in the estimates of grid block permeability and porosity and also improved the convergence properties of the minimization algorithms considered. As they considered synthetic examples, the true values of grid block permeabilities and porosity were known so the error defined as the difference between true and estimated values could be computed. They also showed that Bayesian estimation gave better approximations of the true rock property fields than were obtained by zonation. They also showed that the number of parameters to be estimated can be reduced by reparameterizing the model in terms of the eigenvectors corresponding to the largest eigenvalues of the prior covariance matrix. For a small one-dimensional single-phase flow problem, they showed that the number of parameters to be estimated could be reduced from sixty six to twenty without significantly affecting the accuracy of the estimates of the rock property fields. Later the same authors (see Shah et al. [50]) compared results obtained by Bayesian estimation, by zonation, reparameterization based on vectors of

sensitivity coefficients. Their comparisons were based on computing the trace of the a posteriori covariance matrix obtained by assuming the objective function could be linearized around the true model. As the trace is the sum of the a posteriori variances, it gives an approximate measure of the global uncertainty in the estimated values of grid block permeabilities and porosities. They found that Bayesian estimation results in the smallest uncertainty. The work of Reynolds et al. [47] contains a direct extension of Refs. [20] and [50] to two-dimensional single-phase flow problems. However, the latter reference considers the problem of conditioning to multiwell pressure data, estimates well skin factors as well as rock property fields and implements a subspace method to perform reparameterization. Finally, we note that procedures similar to those introduced by Gavalas et al. [20] and Shah et al. [50] have also been applied to estimate aquifer properties; see, for example the excellent series of papers by Carrera and Newman [5], [6] and [7].

As we have discussed above, our interest is in generating estimates or realizations of grid block porosities and permeabilities (or log-permeabilities) by conditioning a geostatistical model to dynamic data. Typically, a procedure for generating a particular estimate of the model parameters would be based on minimizing an objective function which includes both data mismatch term squared as well as a regularization term. The regularization term could be a mismatch term representing the deviation from the prior model, constraints derived from the prior model, or a smoothness constraint. It is important to choose an efficient optimization algorithm to minimize the objective function. Simulated annealing has often been applied as a minimization technique; see, for example, Ouenes et al. [43] and Hegstad et al. [25]. Simulated annealing requires many thousands of iterations to converge even for problems with a few hundred model parameters (grid block permeabilities and porosities). Thus, it is an impractical procedure if one has to run a reservoir simulator to generate the predicted data corresponding to the estimate of model parameters at each iteration. For this reason, Deutsch [16], who considered the problem of conditioning a model to

geostatistical information and pressure transient data, did not include the pressure data directly into the objective function. Instead, he formed an objective function which included the variogram mismatch and a mismatch between an "average" permeability interpreted from the pressure transient data and the corresponding predicted average permeability calculated from the permeability field obtained at each iteration of the simulated annealing algorithm. There are two major difficulties with this approach. First, the procedure lacks a rigorous basis for the relative weighting of the two component parts (variogram mismatch and average permeability mismatch) of the objective function. One is forced to construct an ad hoc weighting; see, for example, Deutsch and Cockerham [15]. Second, the choice of the objective function defines the probability density function that is being maximized, Geman and Geman [21]. This means the pdf arises from the objective function chosen, whereas, the objective function to be minimized should be derived directly from stochastic model (pdf). For example, Cunha et al. [14] have shown that if one wishes to condition a prior multivariate Gaussian distribution to pressure transient data, then minimizing the objective function used by Deutsch [16] does not yield the correct maximum a posteriori estimate (most probable model). In addition, the approach of Deutsch [16] requires the determination of an "average permeability" from the measured pressure transient data. Deutsch used a power law averaging (see, for example, Alabert [1]). In a minor modification of Deutsch's work, Sagar et al. [49] used computational techniques provided by Feitosa et al. [18] to compute average permeability. The Feitosa et al. work was based on the analytical solution of Oliver [38].

1.2 Summary of Research Objectives and Results

We consider the two-dimensional, two-phase flow of water and oil in an x-y coordinate system. Gravity and capillary effects are neglected. The reservoir may contain any number of completely-penetrating water injection wells and any number

of completely-penetrating producing wells. Water injection rates are specified but the producing well may be produced at either a specified total flow water or at a specified bottom hole pressure. A well constraint may be changed during the simulation run. Fluid properties and relative permeabilities are assumed to be known for both the forward problem and the inverse problem. Outer reservoir boundaries are no-flow boundaries. As our first task, we developed an IMPES finite-difference simulator to predict production data given all rock and fluid property data. The simulator represents our forward model.

As we believe it is preferable to include the production data mismatch in our objective function, at each iteration of the optimization algorithm, it will be necessary to run the simulator to calculate predicted data. Thus, computational efficiency requires an optimization algorithm which converges rapidly. As in Refs. [39], [12], [42] and He [23], the Gauss-Newton procedure is used as an optimization algorithm. Following these same references and Tarantola [52], we cast the estimation problem in terms of maximizing the a posteriori pdf. An approximation for the uncertainty in estimates of gridblock log-permeabilities and porosities is obtained from the a posteriori covariance matrix as in Tarantola [52]. To sample the a posteriori pdf for the purpose of characterizing the uncertainty in performance predictions, a method proposed by Oliver et al. [42] and Kitanidis [32] is used. Here, this sampling procedure is referred to as the randomized maximum likelihood method. Although this procedure gives a completely rigorous sampling only if production data is linearly related to the vector of model parameters (see Oliver et al. [42] and Reynolds et al. [48]), the results of Oliver et al. [42] suggests it generates a reasonable sampling of the a posteriori pdf.

To obtain the a posteriori pdf, it is necessary to specify production data measurement errors. Procedures for doing this have been formulated. Our modeling of these errors recognize that, on a percentage basis, the measurement error in water-oil ratio may be large when the water production rate is small.

If only long-time production data is used in the inverse procedure, we found that the Gauss-Newton method may not converge properly. This result occurs because at early iterations, the production data mismatch terms are large and the effect of the prior model as a regularization term becomes small. This results in excessively rough rock property fields at the first iteration which are difficult to correct at later iterations. We have developed a procedure to overcome this convergence problem. The method is conceptually similar to the Levenberg-Marquardt algorithm.

As the Gauss-Newton method requires the calculation of sensitivity coefficients, we have derived a completely rigorous adjoint procedure to generate the sensitivity coefficients related to model parameters (simulator gridblock permeabilities and porosities) for a two-phase flow (oil and water) problem. The procedure is derived directly from the fully discretized difference equations rather than the semidiscrete equations or original partial differential equations. With this procedure, we directly generate the sensitivity of wellbore pressures and water-oil-ratios to model parameters, i.e., we do not generate the sensitivity of gridblock pressures and saturations to the rock property fields. Unlike previous implementations of the adjoint method, we generate the sensitivity of production data to gridblock permeabilities and porosities, not just the derivative of the objective function with respect to model parameters. This allows one to generate the Hessian required for application of the Gauss-Newton procedure. Since the objective function includes information from a prior geostatistical model, as well as the sum of squares of production data mismatch terms, the Hessian matrix includes the inverse of the covariance matrix estimated from the prior model. Thus, the Hessian matrix is guaranteed to be positive definite, and hence nonsingular. This of course does not mean that there is a unique solution to the inverse problem as there exists an infinite collection of porosity and permeability fields which are consistent with the prior model and the observed data. But, it does mean that we can obtain realizations by history-matching using a computationally efficient algorithm, the Gauss-Newton method.

Given that it requires some effort to formulate the adjoint problem, one should ask how the gradient simulator procedure and our formulation of the adjoint procedure compare in terms of computational efficiency. Both procedures solve linear problems of the same size to obtain sensitivity coefficients. At least formally, both proceed by solving a set of matrix problems where each coefficient matrix is associated with multiple right-hand sides. In the gradient simulator method, the number of right hand sides is equal to the number of model parameters, but in our adjoint formulation, the number of right-hand sides is equal to the number of production data used as conditioning data. (Actually, in our formulation, the number of right hand is usually equal to about one-half the number of conditioning data.) Thus, if the number of model parameters is significantly greater than the number of production data used as conditioning data, our adjoint formulation is preferable. If not, the gradient simulator procedure would be expected to be more computationally efficient. In terms of memory required, the adjoint method does not compare favorably with the gradient simulator method. In the adjoint method, all gridblock saturations and pressures from one simulation run must be stored and all are needed when solving the adjoint equations.

The organization of the remainder of this dissertation is as follows. Chapter II discusses the reservoir simulator, the a posteriori pdf, a method for modeling production data measurement errors and the application of the Gauss-Newton method for estimation and simulation and uncertainty. Chapter III presents the derivation of the adjoint method for generating sensitivity coefficients. Chapter IV is dedicated to results and in Chapter V, we present conclusions and summarize the research contributions of this work.

CHAPTER II

ESTIMATION AND SIMULATION OF ROCK PROPERTY FIELDS

In this chapter, we discuss the reservoir simulator used for the forward problem, the a posteriori probability density function (pdf), measurement errors, the implementation of the Gauss-Newton method for estimation and simulation, and methods for estimating uncertainty.

2.1 Reservoir Simulator

Throughout, capillary pressure is neglected and oil-field units are used. Although we write some equations in general terms, all computational results presented pertain to two dimensional flow in an $x-y$ coordinate system with no gravity effects. Thus, gravity terms are not included in the flow equations even when they are written for three-dimensional flow. The constants C_1 and C_2, respectively, are defined as

$$C_1 = 1.127 \times 10^{-3} \tag{2.1}$$

$$C_2 = 5.615 \tag{2.2}$$

The flow equations are written as

$$C_1 \nabla \cdot \Big(\frac{b_m k_{rm}}{\mu_m} [k] \nabla p(x,y,z,t) \Big) = \frac{\phi}{C_2} \frac{\partial (b_m S_m)}{\partial t} + \hat{q}_m(x,y,z,t), \tag{2.3}$$

for $m = o$ (oil) and $m = w$ (water), $[k]$ is the permeability tensor, and b_m is the inverse formation volume factor in units of STB/RB. The term $\hat{q}(x,y,z,t)$ is a source/sink

term in units of STB/ft^3-day, which is positive for a producing well and negative for an injection well and is nonzero only if the point (x, y, z) is intersected by a well. We assume that the coordinate directions are aligned with the principle permeability directions so that

$$[k] = \begin{bmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & k_z \end{bmatrix} \tag{2.4}$$

We assume a rectangular parallelepiped reservoir, i.e., Eq. 2.3 applies for all $t > 0$ on

$$\Omega = \{(x, y, z) | 0 < x < L_x,\, 0 < y < L_y,\, 0 < z < L_z\}, \tag{2.5}$$

where the boundary of Ω is denoted by $\partial\Omega$. We assume no flow boundary conditions. Initial conditions are given by

$$p(x, y, z, 0) = p_0(x, y, z) \tag{2.6}$$

and

$$S_w(x, y, z, 0) = S_{wi}(x, y, z). \tag{2.7}$$

We partition Ω into gridblocks using a standard block-centered grid and let (x_i, y_j, z_k), $i = 1, 2, \ldots n_x$, $j = 1, 2, \ldots n_y$, $k = 1, 2, \ldots n_z$, denote the gridblock centers. Considering 2.3 at (x_i, y_j, z_k), we use a standard finite-difference procedure to approximate spatial derivatives in 2.3 and multiply the resulting equation by $\Delta x_i \Delta y_j \Delta z_k$ to obtain the following equation:

$$\begin{aligned} &T_{mx,i+1/2,j,k}(t)(p_{i+1,j,k}(t) - p_{i,j,k}(t)) - T_{mx,i-1/2,j,k}(t)(p_{i,j,k}(t) - p_{i-1,j,k}(t)) + \\ &\quad T_{my,i,j+1/2,k}(t)(p_{i,j+1,k}(t) - p_{i,j,k}(t)) - T_{my,i,j-1/2,k}(t)(p_{i,j,k}(t) - p_{i,j-1,k}(t)) + \\ &\quad T_{mz,i,j,k+1/2}(t)(p_{i,j,k+1}(t) - p_{i,j,k}(t)) - T_{mz,i,j,k-1/2}(t)(p_{i,j,k}(t) - p_{i,j,k-1}(t)) = \\ &\qquad \Big(\frac{\Delta x_i \Delta y_j \Delta z_k \phi_{i,j,k}}{C_2}\Big) \frac{\partial (b_m(t) S_m(t))_{i,j,k}}{\partial t} + q_{m,i,j,k}(t) \end{aligned} \tag{2.8}$$

for $m = o, w$. Note that the modified source sink terms are given by

$$q_{m,i,j,k}(t) = \Delta x_i \Delta y_j \Delta z_k \hat{q}(x_i, y_j, z_k, t) \tag{2.9}$$

and have units of STB/D. Throughout, Δx_i, Δy_j and Δz_k are the dimensions of the gridblock centered at (x_i, y_j, z_k). The x-direction boundaries of this gridblock are $x_{i-1/2}$ and $x_{i+1/2}$ so that $\Delta x_i = x_{i+1/2} - x_{i-1/2}$. Similar and obvious notation is used for the gridblock boundaries in the other directions. The T's denote transmissibilities at the gridblock boundaries. Note we have left the time variable continuous in Eq. 2.8 and written all terms which depend on pressure or saturation as functions of time. At any time t and any j and k,

$$T_{mx,i+1/2,j,k}(t) = \frac{C_1 \Delta y_j \Delta z_k k_{x,i+1/2,j,k}}{x_{i+1} - x_i} \Big(\frac{b_m(t) k_{rm}(t)}{\mu_m(t)} \Big)_{i+1/2,j,k} \tag{2.10}$$

for $m = o, w$ and all $i = 1, 2, \ldots n_x - 1$. To incorporate no flow boundaries, we set

$$T_{mx,1/2,j,k}(t) = T_{mx,n_x+1/2,j,k}(t) = 0. \tag{2.11}$$

Similarly,

$$T_{my,i,j+1/2,k}(t) = \frac{C_1 \Delta x_i \Delta z_k k_{y,i,j+1/2,k}}{y_{j+1} - y_j} \Big(\frac{b_m(t) k_{rm}(t)}{\mu_m(t)} \Big)_{i,j+1/2,k}, \tag{2.12}$$

for $m = o, w$ and $j = 1, 2, \ldots n_y - 1$,

$$T_{my,i,1/2,k}(t) = T_{mx,i,n_y+1/2,k}(t) = 0, \tag{2.13}$$

$$T_{mz,i,j,k+1/2}(t) = \frac{C_1 \Delta x_i \Delta y_j k_{z,i,j,k+1/2}}{z_{k+1} - z_k} \Big(\frac{b_m(t) k_{rm}(t)}{\mu_m(t)} \Big)_{i,j,k+1/2}, \tag{2.14}$$

for $m = o, w$ and all $k = 1, 2, \ldots n_z - 1$, and

$$T_{mz,i,j,1/2}(t) = T_{mz,i,j,n_z+1/2}(t) = 0. \tag{2.15}$$

In our simulator, time dependent terms which appear in the transmissibilities are evaluated by upstream weighting, i.e., $\Big(\big(b_m(t) k_{rm}(t) \big) / \mu_m(t) \Big)$ terms are calculated

by upstream weighting. As we consider multiple producing and injection wells, the upstream direction may be different in different regions of the reservoir. At each time step in the simulator, upstream directions are calculated using the latest approximation of the pressure field. Permeabilities at gridblock interfaces are computed as harmonic averages. Specifically, for all $1 \leq j \leq n_y$ and $1 \leq k \leq n_z$,

$$k_{x,i+1/2,j,k} = \frac{(\Delta x_i + \Delta x_{i+1}) k_{x,i,j,k} k_{x,i+1,j,k}}{\Delta x_i k_{x,i+1,j,k} + \Delta x_{i+1} k_{x,i,j,k}}, \tag{2.16}$$

for $i = 1, 2, \ldots n_x - 1$,

$$k_{x,1/2,j,k} = k_{x,1,j,k} \tag{2.17}$$

and

$$k_{x,n_x+1/2,j,k} = k_{n_x,j,k}. \tag{2.18}$$

Similarly,

$$k_{y,i,j+1/2,k} = \frac{(\Delta y_j + \Delta y_{j+1}) k_{y,i,j,k} k_{y,i,j+1,k}}{\Delta y_j k_{y,i,j+1,k} + \Delta y_{j+1} k_{y,i,j,k}}, \tag{2.19}$$

for $j = 1, 2, \ldots n_y - 1$,

$$k_{y,i,1/2,k} = k_{y,i,1,k}, \tag{2.20}$$

$$k_{y,i,n_y+1/2,k} = k_{i,n_y,k} \tag{2.21}$$

$$k_{z,i,j,k+1/2} = \frac{(\Delta z_k + \Delta z_{k+1}) k_{z,i,j,k} k_{z,i,j,k+1}}{\Delta z_k k_{z,i,j,k+1} + \Delta z_{k+1} k_{z,i,j,k}}, \tag{2.22}$$

for $k = 1, 2, \ldots n_z - 1$,

$$k_{z,i,j,1/2} = k_{z,i,j,1}, \tag{2.23}$$

and

$$k_{z,i,j,n_z+1/2} = k_{i,j,n_z}. \tag{2.24}$$

For two dimensional problems, we simply use one gridblock in the z-direction and replace Δz_1 by the reservoir thickness h and set all z-direction transmissibilities to zero, i.e., delete all terms that pertain to vertical flow in Eq. 2.8 and delete the z variable from the equations. By appropriate discretization in time and combination of the discrete equations, one can derive from Eq. 2.8 all standard seven-point finite-difference schemes ranging from IMPES to fully implicit schemes.

Our simulator is a standard IMPES simulator. At each time step, the set of finite difference equations for the overall pressure equation at each time step is solved to obtain gridblock pressures and the finite difference form of the water flow equation is solved for gridblock values of water saturation. The pressure finite-difference equations are solved using a sparse matrix routine.

2.1.1 Well Constraints

The simulator can include water injection wells and multiple producing wells. At injection wells, the water injection rate is specified. Either the flowing bottom hole pressure, the total flow rate or the oil flow rate is specified at the producing wells. The relation between a gridblock source or sink term, the gridblock pressure and flowing bottom hole pressure is specified by Peaceman's equation [45]. Specific source sink terms in a gridblock penetrated by a well must be computed from the specified well constraint. Since our actual simulator pertains only to two-dimensional $(x-y)$ two-phase (oil and water), we discuss the pertinent relations only for this special case. Assume well k is located in a gridblock centered at (x_i, y_j), which here, we also refer to as gridblock k. Applying Peaceman's equation for the two-phase flow problem at time t^n gives the phase flow rate $(q^n_{m,k})$ for well k is related to the gridblock pressure, $p^n_k = p^n_{i,j}$, and the flowing wellbore pressure, $p^n_{wf,k}$ by

$$q^n_{m,k} = WI_k \frac{k^n_{rm,k} b^n_{m,k}}{\mu^n_{m,k}} \big(p^n_k - p_{wf,k}\big), \tag{2.25}$$

where $WI_k = WI_{i,j}$ is the well index term for gridblock k which is centered at (x_i, y_j). Defining, the constant C_3 by

$$C_3 = (2\pi)1.127 \times 10^{-3}, \tag{2.26}$$

$$WI_k = \frac{C_3 h\sqrt{k_{x,k}k_{y,k}}}{\ln(r_{o,k}/r_{w,k}) + s_k}, \tag{2.27}$$

where $k_{x,k}$ and $k_{y,k}$, respectively, denote the x and y direction absolute permeabilities for the gridblock containing well k, (i.e., $k_{x,k} = k_{x,i,j}$ and $k_{y,k} = k_{y,i,j}$ and

$$r_{o,k} = \frac{0.28073\Delta x_i\sqrt{1 + \frac{k_{x,k}\Delta y_k^2}{k_{y,k}\Delta x_k^2}}}{1 + \sqrt{k_{x,k}/k_{y,k}}}. \tag{2.28}$$

Here $r_{w,k}$ is the wellbore radius of well k, s_k is the skin factor for well k, and Δx_k and Δy_k, respectively, denote the x and y dimensions of gridblock j, i.e., $\Delta x_k = \Delta x_i$ and $\Delta y_k = \Delta y_j$. Moreover the term $k^n_{rm,k}b^n_{m,k}/\mu_{m,k}$ denotes the product of the inverse formation volume factor and mobility of phase m evaluated at the estimated water saturation and pressure for the gridblock at time t^n.

Suppose a water injection well is located in gridblock k and the water injection rate is specified as $q^n_{w,k}$ at time t^n. Then for gridblock k we set $q^n_{o,k} = 0$, and set

$$q^n_{w,k} = WI_k\Big(\frac{k^n_{rw,k}b^n_{w,k}}{\mu^n_{w,k}} + \frac{k^n_{ro,k}b^n_{o,k}}{\mu^n_{o,k}}\Big)\big(p^n_k - p_{wf,k}\big), \tag{2.29}$$

where $q_{w,k} < 0$ corresponds to injection. The oil mobility term is included in Eq. 2.29 to ensure that water can be injected without an extremely large increase in the wellbore pressure. For example, if water is immobile at the initial water saturation in the gridblock, and Eq. 2.29 was applied without including the oil terms, $\big(p^n_k - p_{wf,k}\big)$ would become infinite. Of course once the gridblock fills with water so $S_w = 1 - S_{or}$, $k_{ro,k} = 0$ and the oil terms in Eq. 2.29 are automatically deleted.

Now assume k corresponds to a production well where the total production rate is specified, i.e., $q^n_{t,k} > 0$ is specified at time t^n. Then, Peaceman's equation is

$$q^n_{t,k} = WI_k\Big(\frac{k^n_{rw,k}b^n_{w,k}}{\mu^n_{w,k}} + \frac{k^n_{ro,k}b^n_{o,k}}{\mu^n_{o,k}}\Big)\big(p^n_k - p_{wf,k}\big). \tag{2.30}$$

Defining the mobility of phase m in gridblock k at time t^n by

$$\lambda^n_{m,k} = \frac{k^n_{rm,k}}{\mu^n_{m,k}}, \tag{2.31}$$

Eq. 2.30 can be expressed as

$$q^n_{t,k} = WI_k\Big(\lambda^n_{w,k}b^n_{w,k} + \lambda^n_{o,k}b^n_{o,k}\Big)\big(p^n_k - p_{wf,k}\big). \tag{2.32}$$

The oil and water difference equations at this gridblock require the sink terms $q^n_{o,k}$ and $q^n_{w,k}$ where

$$q^n_{t,k} = q_{o,k} + q_{w,k}, \tag{2.33}$$

and the unknown individual phase flow rates are still required to satisfy Peaceman's equation, Eq. 2.25. But, for $m = o$ or $m = w$,

$$q^n_{m,k} = \frac{q^n_{m,k}}{q^n_{t,k}}q^n_{t,k}. \tag{2.34}$$

Using Eqs. 2.25 and 2.32 in Eq. 2.34 and simplifying gives

$$q^n_{m,k} = \frac{\lambda^n_{m,k}b^n_{m,k}}{\lambda^n_{w,k}b^n_{w,k} + \lambda^n_{o,k}b^n_{o,k}}q^n_{t,k}, \tag{2.35}$$

for $m = 0, w$. Eq. 2.35 can be used to specify the relevant sink terms in the phase m finite-difference equation for gridblock k. After solving the IMPES equations for grid block pressures and saturations, Eq. 2.32 can solved for $p^n_{wf,k}$.

Alternatively, we can insert the right side of Eq. 2.25 directly into the finite difference equations. When this is done, the flowing wellbore pressure $p^n_{wf,k}$ adds an additional unknown to the system of difference equations. However, since $q_{t,k}$ is specified, Eq. 2.32 provides the additional equation needed so that the number

of equations is equal to the number of unknowns. Specifically, the set of IMPES pressure finite-difference equation plus Eq. 2.32 are solved simultaneously for the gridblock pressure and the wellbore pressure, $p^n_{wf,k}$.

If the flowing bottom hole pressure is specified as the rate constraint, then the right hand side of Eq. 2.25 is substituted as the sink term into the finite difference for phase m. Since $p^n_{wf,k}$ is specified, this introduces no additional unknowns into the system of finite difference equations. After solving for grid block pressures and saturations are calculated, Eq. 2.25 can be applied to calculate the phase flow rates and the producing water-oil ratio at time t^n at well k given by

$$WOR^n_k = \frac{q^n_{w,k}}{q^n_{o,k}} = \frac{\lambda^n_{o,k} b^n_{o,k}}{\lambda^n_{w,k} b^n_{w,k}}. \tag{2.36}$$

Similarly the water cut can be calculated as

$$f^n_{w,k} = \frac{q^n_{w,k}}{q^n_{t,k}} = \frac{\lambda^n_{o,k} b^n_{o,k}}{\lambda^n_{w,k} b^n_{w,k} + \lambda^n_{o,k} b^n_{o,k}}. \tag{2.37}$$

2.2 Prior and A Posteriori Probability Density Function

The notation and terminology used here is the same as used in previous related work; see for example, [12], [39] or [24]. The model parameters are gridblock log-permeabilities and gridblock permeabilities. Although the flow equations and adjoint equations (Chapter III) are formulated in terms of anisotropic permeabilities, the computational results in Chapter IV all pertain to the isotropic case. Thus, for simplicity, the formulation here assumes permeability is isotropic. However, the extension to the anisotropic case is trivial.

Each gridblock permeability is treated as a random variable which is assumed to have a log-normal distribution (log-permeability is normal) with known mean and variance σ^2_k. Each gridblock porosity is assumed to be a normal random variable with known mean and variance σ^2_ϕ. We model each of these attributes as a stationary

Gaussian random function. The correlation coefficients between the two rock property attributes is assumed to be known, but may be zero.

Assuming the simulation model is based on N gridblocks, the vector of model parameters is given by the $M = 2N$ dimensional column vector

$$m = \begin{bmatrix} m_\phi \\ m_k \end{bmatrix}, \tag{2.38}$$

where m_ϕ is the M dimensional column vector of gridblock porosities and m_k is the M-dimensional column vector of gridblock values of $\ln(k)$. Throughout, m_{prior} is the vector containing the prior estimates of these parameters.

2.2.1 The Prior Model

It is assumed that the prior stochastic model for m is a correlated stationary Gaussian random field. Thus, the prior model has a multivariate Gaussian probability density function with prior covariance matrix, C_M. For the specific problem under consideration,

$$C_M = \begin{bmatrix} C_\phi & C_{\phi k} & O \\ C_{k\phi} & C_k & O \\ O & O & C_s \end{bmatrix}. \tag{2.39}$$

In Eq. 2.39, C_ϕ is the covariance matrix for gridblock porosities (derived from the porosity variogram), C_k is the covariance matrix for gridblock $\ln(k)$'s (derived from the variogram for log-permeability), $C_{\phi k}$ is the cross covariance matrix between porosity and $\ln(k)$ at the set of gridblocks, $C_{k\phi}$ is equal to the transpose of $C_{\phi k}$ and throughout O's denote null matrices, i.e., matrices with all entries equal to zero. As in Ref. [13], the cross covariance is obtained using the screening hypothesis of Xu et al. [55]. This avoids the necessity of modeling the cross variogram between log-permeability and porosity and ensures that the covariance matrix C_M is positive definite.

2.2.2 A Posteriori Probability Density Function

The relationship between the vector d of calculated production data and the vector m of model reservoir parameters is written as

$$d = g(m). \tag{2.40}$$

The functional relationship of Eq. 2.40 represents the effect of generating d from our reservoir simulator for a given m. If the true reservoir could be described by discretization into gridblocks and the entries of m were exactly equal to the true values of gridblock rock properties, then Eq. 2.40 would predict the observed production data provided the solution of the finite difference equations was not affected by truncation or roundoff errors and data were measured exactly. However, the observed production data will be corrupted by measurement errors. We let d_{obs} denote the vector of all observed or measured production data that will be used a conditioning data in the inverse problem.

As is commonly done, we model measurement errors as independent Gaussian random variables with prescribed means and variances. If there are N_d production data, we let $\sigma^2_{d,j}$ denote the variance of the jth measurement error and define C_D to be a $N_d \times N_d$ diagonal matrix with its jth diagonal equal to $\sigma^2_{d,j}$. Throughout, C_D is referred to as the data covariance matrix. Then from basic inverse problem theory (see, for example, Tarantola [52]), an application of Bayes' theorem indicates that a posteriori probability density function is given by

$$f(m) = a \exp\Big(-\frac{1}{2}\Big[(m - m_{prior})^T C_M^{-1}(m - m_{prior}) + (g(m) - d_{obs})^T C_D^{-1}(g(m) - d_{obs})\Big]\Big), \tag{2.41}$$

where a is the normalizing constant.

2.2.3 Estimation and Simulation

Given production data, d_{obs}, the inverse problem refers to the determination of models m that are consistent with the observed data. Based on the probabilistic formulation, realizations of m can be obtained by sampling the a posteriori pdf given by Eq. 2.41. Any sample of this pdf gives a plausible set of values of gridblock log-permeabilities and porosities that are conditioned to the production data, however, based on the pdf, some realizations are more likely than others. The process of generating realizations of this pdf is referred to simulation, or to differentiate this sampling from reservoir simulation, stochastic simulation.

Estimation refers to the construction of a particular estimate of the model or model parameters that is consistent with the production data and the prior model. If a prior model were not specified, then one formulation of the estimation problem is as follows: find a model m that minimizes the data mismatch objective function $O_d(m)$ defined by

$$O_d(m) = \frac{1}{2}\Big((g(m) - d_{obs})^T C_D^{-1}(g(m) - d_{obs})\Big]. \tag{2.42}$$

Thus, we have cast the inverse problem in the form of a classical least squares problem. For the cases of interest to us, the number of independent production data is far fewer than the number of model parameters (gridblock log-permeabilities and porosities). Thus, unlike standard least squares problems, the problem is undetermined and does not have a unique solution. For example, if the production data correspond to "early times", the data may be completely insensitive to the values of permeability and porosity at grid blocks far from wells. Thus, if there exists one model m which minimizes $O(d)$, any other model obtained from m by perturbing values of permeability and porosity in these gridblocks will also minimize the objective function. In such a situation, minimization of Eq. 2.42 by a Newton method would fail because the Hessian matrix would be singular. Some form of regularization is

required to mathematically specify a particular estimate, i.e., to determine a specific model m from the set of all m which minimize Eq. 2.42. As discussed in Parker [44] and Menke [35], there exist a variety of regularization techniques. In the probabilistic approach we have taken, the prior stochastic model provides the regularization term.

2.3 The Maximum A Posteriori Estimate

Here, the most probable model refers to the estimate of m which maximizes the a posteriori pdf. More commonly, this estimate would be referred to as the maximum a posteriori (MAP) estimate. Throughout the MAP estimate of the rock property fields is denoted by m_∞. It can be obtained by minimizing the objective function

$$O(m) \;=\; \frac{1}{2}\Big[(m - m_{prior})^T C_M^{-1}(m - m_{prior}) + (g(m) - d_{obs})^T C_D^{-1}(g(m) - d_{obs})\Big]. \tag{2.43}$$

If the predicted data are linearly related to the model, i.e.,

$$d = Gm, \tag{2.44}$$

where G is an $N_d \times M$ matrix, then Eq. 2.43 has a unique global minimum which can be constructed from an analytical formula. As given in Tarantola [52], the analytical equation for the MAP is

$$m_\infty = m_{prior} - H^{-1}\Big(C_M^{-1} + G^T C_D^{-1}(Gm_{prior} - d_{obs})\Big) \tag{2.45}$$

where the $M \times M$ matrix H is given by

$$H = C_M^{-1} + G^T C_D^{-1} G. \tag{2.46}$$

Since C_M and C_D are covariance matrices and hence positive definite, H is also positive definite and hence nonsingular. Note, however, if no prior model were available,

then C_M would be deleted from the equations so that Eqs. 2.45 and 2.46 become

$$m_\infty = m_{prior} - H^{-1}\Big(G^T C_D^{-1}(Gm_{prior} - d_{obs})\Big) \tag{2.47}$$

where

$$H = G^T C_D^{-1} G. \tag{2.48}$$

In this case, there is no reason to believe that H is nonsingular. In fact, if the number of model parameters is greater than the number of data, H must be singular, Menke [35]. Thus, we see that the prior model provides regularization so that the linear inverse problem has a unique maximum a posteriori estimate. For the problem of conditioning to nonlinear production data, Eq. 2.44 does not hold and there is no guarantee that there is a unique MAP estimate.

2.4 The Gauss-Newton Method

In our work, the maximum a posteriori estimate is constructed by minimizing the objective function of Eq. 2.43 using the Gauss-Newton procedure with restricted-step. This is the same procedure used in many other papers, see, for example, Refs. [12], [47] and [24]. As will be discussed in a later section, the Gauss-Newton procedure is also applied to sample the a posteriori pdf.

The gradient of the $O(m)$ with respect to the model parameters m is given by

$$\nabla_m O(m) = G^T C_D^{-1}(g(m) - d_{obs}) + C_M^{-1}(m - m_{prior}), \tag{2.49}$$

and the Hessian for the Gauss-Newton method is given by

$$H = G^T C_D^{-1} G + C_M^{-1}, \tag{2.50}$$

where G is the matrix of sensitivity coefficients. In our work, sensitivity coefficients are generated by the adjoint method discussed in Chapter III.

In the Newton procedure, the Hessian, gradient and sensitivity coefficients change from iterate to iterate. At the lth iteration of the Gauss-Newton algorithm, we solve

$$H_l \delta m^{l+1} = -\nabla O_l \tag{2.51}$$

for the search direction δm^{l+1} and then compute the updated estimate of the model parameters from

$$m^{l+1} = m^l - \mu_l \delta m^{l+1}, \tag{2.52}$$

where μ_l is computed by the same restricted step procedure used in Refs. [12] and [24]. Fletcher [19] provides a detailed discussion of the restricted-step method. In Eqs. 2.51, l denotes the iteration index, ∇O_l denotes the gradient evaluated at the old iterate (m^l) and H_l denotes the Hessian matrix evaluated at m^l.

Similar to Tarantola [52], Chu et al. [12], show that matrix inversion lemmas can be used to rewrite Eq. 2.52 as

$$\delta m^{l+1} = \hat{\mu}_l (m_{prior} - m_l) - \mu_l C_M {G_l}^T \Big(C_D + G_l C_M {G_l}^T \Big)^{-1} \times \Big[g(m^l) - d_{obs} - G_l (m^l - m_{prior}) \Big], \tag{2.53}$$

where $\hat{\mu}_l = \mu_l$. If we set $\hat{\mu}_l = \mu_l = 1$ in Eq. 2.53, the resulting equation is equivalent to Eq. 2.51. We have used different notation for these two quantities since some authors (see, for example, Tarantola [52]) simply set $\hat{\mu}_l = 1$ in Eq. 2.53 to obtain

$$m^{l+1} = m_{prior} - \mu_l C_M {G_l}^T \Big(C_D + G_l C_M {G_l}^T \Big)^{-1} \times \Big[g(m^l) - d_{obs} - G_l (m^l - m_{prior}) \Big]. \tag{2.54}$$

As the Gauss-Newton method converges to the MAP estimate, $\mu_l \to 1$, thus setting $\hat{\mu}_l = 1$ should not cause significant difficulties, and in previous work done at the

University of Tulsa, it was found that setting $\hat{\mu}_l = 1$ did not cause any noticeable degradation in the rate of convergence or in the maximum a posteriori estimate obtained from the Gauss-Newton method. Nevertheless, in the results shown in this work, Eq. 2.53 with $\hat{\mu}_l = \mu_l = 1$ is applied to compute δm^{l+1} and then m^{l+1} is calculated from Eq. 2.52.

Although Eqs. 2.51 and 2.53 are mathematically equivalent, the computational time for the two schemes may be different. The inverse matrix on the right side of Eq. 2.53 is $N_d \times N_d$, where N_d is the number of observed production data used as conditioning data. The Hessian matrix in Eq. 2.51 is $M \times M$, where M is the number of model parameters. If $N_d << M_p$, it is most convenient to apply Eq. 2.53. In the computational results presented here production data are thinned to fewer than 25 data points per well and Eq. 2.53 is incorporated into the Gauss-Newton algorithm.

2.4.1 Convergence Criteria for the Gauss-Newton Algorithm

In this work, we assume that the Gauss-Newton procedure has converged to the maximum a posteriori estimate when the following criterion is satisfied:

$$\frac{O(m^{l+1}) - O(m^l)}{O(m^{l+1}) + 10^{-14}} \leq \epsilon_1, \tag{2.55}$$

where the convergence tolerance ϵ_1 is on the order of 10^{-3}. Another appropriate convergence tolerance is given by

$$\frac{\parallel \delta m^{l+1} \parallel}{\parallel m^l \parallel + 10^{-14}} \leq \epsilon_2, \tag{2.56}$$

where ϵ_2 is on the order of 10^{-3}.

Regardless of the tolerance used, we also check the condition

$$\sum_{i=1}^{N_d} \left(\frac{d_{obs,i} - g_i(m)}{\sigma_{d,i}} \right)^2 \leq N_d, \tag{2.57}$$

If measurement errors are Gaussian and m represents the true model, then the left side of Eq. 2.57 has a chi-squared distribution with expected value given by N_d and

variance equal to $2N_d$; see Barlow [3]. If the left hand side of Eq. 2.57 is significantly greater than N_d, we expect that something is wrong, e.g., our model of measurement errors is incorrect, the Gauss-Newton method has converged to a local minimum, or the prior model is inappropriate. In fact Eq. 2.57 can be used if conjunction with Eq. 2.55 (or Eq. 2.56). In this situation, if either equation is satisfied, we assume that the Gauss-Newton method has converged.

2.5 Modeling of Production Data Measurement Errors

As noted previously, production data measurement errors are modeled as independent Gaussian random variables. As we use two types of production data, wellbore pressure data and water-oil-ratio (or water cut) data, two distinct types of measurement errors must be considered in order to specify the data covariance matrix C_D. In our work, it is assumed that pressure measurement errors are identically distributed independent random variables with mean zero and variance σ_d^2. For water-oil-ratio data, however, it is clear that the measurement error should be related to the magnitude of the data themselves. A procedure for doing so is discussed in this section.

As both pressure and production data obtained from flow rates may be used as conditioning data, we denote the data covariance matrix for pressure measurement errors as C_{Dp}, denote the data covariance matrix for all other production data as C_{Dq} and let C_D denote the overall data covariance matrix. If only pressure data is used as conditioning data, then $C_D = C_{Dp}$. If both pressure and some other type of production data (water-oil-ratio for the computational examples presented in this work) are used as conditioning data, then

$$C_D = \begin{bmatrix} C_{Dp} & O \\ O & C_{Dq} \end{bmatrix}, \tag{2.58}$$

where a submatrix O always denotes a null matrix of appropriate order. If only

water-oil ratio data is used as production data, then $C_D = C_{Dq}$. As in our previous implementations, we assume that pressure measurements errors are independent, identically distributed random variables with mean zero and variance σ_d^2. Thus, the pressure data covariance matrix C_{Dp} is a $N_p \times N_p$ diagonal matrix, with all diagonal entries equal to σ_d^2, where N_p is the number of pressure data used as conditioning data.

2.5.1 Rate Measurement Errors

The emphasis in this section is on the model used for measurement errors in flow rate data, water-oil-ratio data, and water-cut data. Throughout, q_m, $m = o, w$ denotes the true rate (no measurement error), $q_{m,obs}$, $m = o, w$ denotes the observed or measured rate, $Var[X]$, denotes the variance of a random variable X, $E[X]$ its expectation, e_m denotes the measurement error for the rate of phase m and is considered to be a Gaussian variable of mean zero. We let $\sigma(x) = \sigma_x = \sqrt{Var[X]}$.

With the preceding notation,

$$q_{m,obs} = q_m + e_m, \tag{2.59}$$

so

$$E[q_{m,obs}] = q_m + E[e_m] = q_m, \tag{2.60}$$

and

$$Var[q_{m,obs}] = Var[q_m] + Var[e_m] = Var[e_m]. \tag{2.61}$$

Throughout, we assume that rate measurement errors are normally distributed with mean zero so that

$$E[e_m] = 0, \tag{2.62}$$

for $m = o, w$ Eqs. 2.60 and 2.61, respectively, are equivalent to

$$E[q_{m,obs} - q_m] = E[e_m] = 0, \tag{2.63}$$

and

$$Var[q_{m,obs} - q_m] = Var[e_m]. \tag{2.64}$$

The preceding equations apply to each observation of the rates. Here, we assume measurement errors are uncorrelated so the associated data covariance matrix is diagonal.

One would expect the error in a rate measurement to depend on the magnitude of the rate. For example, if the water rate is on the order of 0.5 STB/D, it does not make sense to specify $Var[q_{w,obs}] = 2.25$ (STB/D)^2. On the other hand, if the water rate is 500 STB/D, $\sigma_{q_{w,obs}} = 1.5$ STB/D probably represents an unrealistic measurement accuracy.

For simplicity, we will assume a constant relative measurement error, ϵ_m, that is, for a one percent relative measurement error in oil rate, we set $\epsilon_o = 0.01$. We approximate the relevant standard deviations by

$$\sigma(q_{m,obs} - q_m) = \sigma_{q_{m,obs}} = \sigma[e_m] \approx \epsilon_m q_{m,obs}, \tag{2.65}$$

or,

$$Var(q_{m,obs} - q_m) = Var[e_m] \approx (\epsilon_m)^2 q_{m,obs}^2. \tag{2.66}$$

2.5.2 Variance of WOR

The true water-oil ratio is given by

$$f(q_o, q_w) = WOR = \frac{q_w}{q_o}, \tag{2.67}$$

whereas, the observed water-oil ratio is given by

$$f(q_{o,obs}, q_{w,obs}) = WOR_{obs} = \frac{q_{w,obs}}{q_{o,obs}}. \tag{2.68}$$

A standard Taylor series approximation of $f(q_{o,obs}, q_{w,obs})$ about the true values of the rates is given by

$$f(q_{o,obs}, q_{w,obs}) = f(q_o, q_w) + \frac{\partial f(q_o, q_w)}{\partial q_o}(q_{o,obs} - q_o) + \frac{\partial f(q_o, q_w)}{\partial q_w}(q_{w,obs} - q_w) \tag{2.69}$$

Eq. 2.69 is equivalent to

$$\begin{aligned} e_{WOR} &= WOR - WOR_{obs} \\ &= \frac{\partial f(q_o, q_w)}{\partial q_o}(q_{o,obs} - q_o) + \frac{\partial f(q_o, q_w)}{\partial q_w}(q_{w,obs} - q_w) \\ &= \frac{\partial f(q_o, q_w)}{\partial q_o} e_o + \frac{\partial f(q_o, q_w)}{\partial q_w} e_w. \end{aligned} \tag{2.70}$$

Taking the expectation and variance, respectively, of Eq. 2.70 and applying Eq. 2.62 yields

$$E[e_{WOR}] = 0, \tag{2.71}$$

and

$$Var[e_{WOR}] = \left(\frac{\partial f(q_o, q_w)}{\partial q_o}\right)^2 Var[q_{o,obs} - q_o] + \left(\frac{\partial f(q_o, q_w)}{\partial q_w}\right)^2 Var[q_{w,obs} - q_w]. \tag{2.72}$$

The derivatives appearing in Eq. 2.72 can be formally calculated from Eq. 2.67. The true values of rates are unknown, however, so we evaluate the derivatives at $(q_{o,obs}, q_{w,obs})$. Inserting then, these partial derivatives of f into Eq. 2.72 gives

$$Var[e_{WOR}] = \frac{q_{w,obs}^2}{q_{o,obs}^4} Var[q_{o,obs} - q_o] + \frac{1}{q_{o,obs}^2} Var[q_{w,obs} - q_w]. \tag{2.73}$$

Performing trivial algebra and using Eq. 2.66, Eq. 2.73 becomes

$$Var[e_{WOR}] = \frac{q_{w,obs}^2}{q_{o,obs}^2}\left(\frac{1}{q_{o,obs}^2}Var[q_{o,obs}-q_o] + \frac{1}{q_{w,obs}^2}Var[q_{w,obs}-q_w]\right)$$
$$= WOR_{obs}^2(\epsilon_o^2+\epsilon_w^2). \quad (2.74)$$

The preceding formula results, however, in what seem to be unrealistically small values for the variance in measurements of WOR when the water rate is small. The problem is that we assumed that the same relative error in the water rate was valid for any water rate, including water rates that are approximately zero. It seems more reasonable to assume that there is some finite detection limit, below which the uncertainty in the measurement of the water rate has a fixed magnitude.

Thus, we rewrite Eq. 2.74 as

$$Var[e_{WOR}] = WOR_{obs}^2\,\epsilon_o^2 + \frac{1}{q_{o,obs}^2}\sigma_{q_{w,obs}}^2, \quad (2.75)$$

where

$$\epsilon_m^2 = \frac{Var[q_{m,obs}-q_m]}{q_{m,obs}^2} \quad (2.76)$$

and also specify

$$\sigma_{q_{w,obs}} = \max[\epsilon_w q_{w,obs}, \sigma_{q_{w,obs}}^{min}]. \quad (2.77)$$

With ϵ_o and ϵ_w specified, Eq. 2.75 is used to define the variance of the measurement error for the water-oil ratio where $\sigma_{q_{w,obs}}$ is computed with Eq. 2.77 . In the examples considered in this work, we set $\sigma_{q_{w,obs}}^{min} = 2$ STB/D, $\epsilon_o = 0.01$ and $\epsilon_w = 0.02$.

Eqs. 2.72 through 2.75 assume that the water and oil rates are measured independently and that the two rates are uncorrelated, which may not be the case. If the measurement errors of oil and water rates are correlated, then, as shown in Barlow [3], Eq. 2.72 (with partial derivatives evaluated at observed rates) should be

replaced by

$$Var[e_{WOR}] = \left(\frac{\partial f(q_{o,obs}, q_{w,obs})}{\partial q_o}\right)^2 Var[q_{o,obs}-q_o]+\left(\frac{\partial f(q_{o,obs}, q_{w,obs})}{\partial q_w}\right)^2 Var[q_{w,obs}-q_w]+$$
$$2\left(\frac{\partial f(q_{o,obs}, q_{w,obs})}{\partial q_o}\right)\left(\frac{\partial f(q_{o,obs}, q_{w,obs})}{\partial q_w}\right) cov[q_{o,obs} - q_o,\ q_{w,obs} - q_w], \quad (2.78)$$

where the covariance of the rate measurement errors is given by

$$cov[q_{o,obs} - q_o, q_{w,obs} - q_w] = \rho\sqrt{Var[q_{o,obs} - q_o]\, Var[q_{w,obs} - q_w]}. \quad (2.79)$$

Here, ρ denotes the correlation between oil and water rate measurement errors. Similarly, Eq. 2.73 should be replaced by

$$Var[e_{WOR}] = \frac{q_{w,obs}^2}{q_{o,obs}^4} Var[q_{o,obs} - q_o] + \frac{1}{q_{o,obs}^2} Var[q_{w,obs} - q_w] -$$
$$2\frac{q_{w,obs}}{q_{o,obs}^3} cov[q_{o,obs} - q_o, q_{w,obs} - q_w] \quad (2.80)$$

We can of course always apply the general formula for the variance of the WOR and then simply set $\rho = 0$ if we wish errors to be uncorrelated.

2.5.3 Variance of Water-Cut Measurement Error

The water cut is given by

$$f_w = \frac{q_w}{q_o + q_w} = \frac{q_w}{q_t} \quad (2.81)$$

Following the same procedure used to derive Eq. 2.73, we have

$$Var[e(f_w)] = \frac{q_w^2}{q_t^4} Var[q_{t,obs} - q_t] + \frac{1}{q_t^2} Var[q_{w,obs} - q_w] =$$
$$f_w^2 \left(\frac{Var[q_{t,obs} - q_t]}{q_t^2} + \frac{Var[q_{w,obs} - q_w]}{q_w^2}\right), \quad (2.82)$$

where $e(f_w)$ denotes the measurement error in water cut. Note we have assumed here that q_t and q_w are measured independently and that these two measurement errors are uncorrelated. If we actually measure q_o and q_w independently, and then construct a "measured" q_t by adding these two measurements, then

$$Var[q_{t,obs} - q_t] = Var[q_{w,obs} - q_w] + Var[q_{o,obs} - q_o]. \quad (2.83)$$

2.6 Gauss-Newton Adjustments in Data Covariance Matrices

For the examples presented in this report, all conditioning data used to construct realizations correspond to $t \geq 280$ days. In this case, the initial data mismatch terms are extremely large and the effect of the prior covariance matrix C_M which provides regularization and also promotes continuity in the simulated rock property fields (assuming that C_M corresponding to a variogram with a range that spans at least a few gridblocks) has little effect at the first iteration of the Gauss-Newton method. Because of this, a very rough rock property field is generated at the first iteration, and it is often impossible to remove the excessive roughness at subsequent iterations. To avoid this problem, we use artificially high values of the measurement error variances at early iterations, a procedure which is at least in the spirit of the Levenberg-Marquardt algorithm.

Recall that the data mismatch part of the objective function of Eq. 2.43 is 1/2 times

$$O_d = [d - d_{obs}]^T C_D^{-1} [d - d_{obs}]. \tag{2.84}$$

Since pressure measurement errors and water-oil-ratio measurement errors are uncorrelated, we can write Eq 2.84 as

$$O_d = O_p + O_q, \tag{2.85}$$

where

$$O_p = [p - p_{obs}]^T C_{Dp}^{-1} [p - p_{obs}] = \sum_{j=1}^{N_p} \frac{(p_j - p_{obs,j})^2}{Var[p_{obs,j}]} = \frac{1}{\sigma_d^2} \sum_{j=1}^{N_p} (p_j - p_{obs,j})^2, \tag{2.86}$$

and

$$O_q = [WOR - WOR_{obs}]^T C_{Dq}^{-1} [WOR - WOR_{obs}] = \sum_{j=1}^{N_q} \frac{(WOR_j - WOR_{obs,j})^2}{Var[WOR_{obs,j}]}. \tag{2.87}$$

Here $p_{obs,j}$ is the jth pressure measurement and WOR_j is the jth water-oil-ratio data point. The vector p_{obs} contains all of the observed pressure data and the vector WOR_{obs} contains all of the observed water-oil-ratio data. The last equality of Eq. 2.86 follows from the assumption that pressure data measurement errors are identically distributed Gaussian random variables with variance equal to σ_d^2. Given the assumption that measurement errors are Gaussian, Eqs. 2.86 and 2.87 both have a chi-squared probability function and the expected values of the two sub-objective functions are given by

$$E[O_p] = E\Big[[p - p_{obs}]^T C_D^{-1}[p - p_{obs}]\Big] = N_p, \tag{2.88}$$

and

$$E[O_q] = E\Big[[WOR - WOR_{obs}]^T C_{Dq}^{-1}[WOR - WOR_{obs}]\Big] = N_q. \tag{2.89}$$

Prior to performing iteration $l+1$ of the Gauss-Newton method to calculate m^{l+1}, we compute

$$O_p^l = [p^l - p_{obs}]^T C_D^{-1}[p^l - p_{obs}], \tag{2.90}$$

and

$$O_q^l = [WOR^l - WOR_{obs}]^T C_{Dq}^{-1}[WOR^l - WOR_{obs}], \tag{2.91}$$

where p^l and WOR^l denote vectors of pressure and WOR data *calculated* from m^l. If $O_p^l > 3N_p$, we replace C_{Dp} by $(O_p^l/3N_p)C_{Dp}$. If $O_q^l > 3N_q$, we replace C_{Dq} by $(O_q^l/3N_q)C_{Dq}$ when generating m^{l+1}. This somewhat ad hoc procedure results in smoother convergence and for all examples we have tried prevents convergence to a local minimum which gives an unacceptable match of the data. Typically, after a few iterations, a sufficiently low data mismatch is achieved so that the specified data covariance matrices C_{Dp} and C_{Dq} can be used.

2.7 Sampling the Posterior Probability Density Function

As discussed in the introduction, it is desirable to generate a set of realizations of the model by sampling the a posteriori pdf in order to characterize the uncertainty in reservoir performance predictions. The procedure used for sampling is the one introduced into the petroleum engineering literature by Oliver et al. [42]. The same procedure was introduced earlier into the groundwater literature without discussion of its limitations by Kitanidis [32]. The procedure is effectively the one formulated for linear problems by Oliver [40]. An alternate derivation of the method is given in Reynolds et al. [48]. Here, this sampling procedure is referred to as the randomized maximum likelihood method. Strictly speaking, the method provides a rigorous sampling of the a posteriori pdf only if the data are linearly related to the model.

To generate a single realization of the vector of model parameters by the randomized maximum likelihood method, we generate an unconditional realization of m_{uc} of the vector of model parameters by sampling the prior pdf and an unconditional realization of the data. The unconditional realization of the model can be generated using a Cholesky or square root decomposition of the prior covariance matrix or by sequential Gaussian cosimulation; see, for example, Ref. [22]. Since the data covariance matrix is a positive definite, diagonal matrix, an unconditional simulation of the data can be obtained from

$$d_{obs,uc} = d_{obs} + C_D^{1/2} Z_D, \tag{2.92}$$

where Z_D is a vector of independent standard random normal deviates. A conditional realization, m_c of the vector of model parameters is then generated by minimizing

$$O_r(m) \;=\; \frac{1}{2}\Big[(m - m_{uc})^T C_M^{-1}(m - m_{uc}) + (g(m) - d_{obs,uc})^T C_D^{-1}(g(m) - d_{obs,uc})\Big]. \tag{2.93}$$

The Gauss-Newton procedure is applied to minimize $O_d(m)$. If one wishes to generate N_r realizations, the process is repeated N_r times.

2.8 Evaluation of Uncertainty

Throughout, m_∞ denotes the approximate of the MAP estimate obtained by applying the Gauss-Newton procedure to minimize the objection function of Eq. 2.43 and G_∞ denotes the sensitivity coefficient matrix[6–8,20] evaluated at m_∞. For the problems considered here, sensitivity coefficients are calculated from the adjoint method developed in Chapter III.

As discussed in Tarantola [52], the a posteriori covariance matrix, C_{MP}, is given by

$$C_{MP} = \left(G_\infty C_D^{-1} G_\infty^T + C_M^{-1}\right)^{-1} = C_M - C_M {G_\infty}^T (G_\infty C_M {G_\infty}^T + C_D)^{-1} G_\infty C_M. \quad (2.94)$$

The diagonal elements of C_{MP} represent the a posteriori variances. Letting, $\sigma_{MP,j}$ denote the jth diagonal entry, the 95 percent confidence interval for the jth model parameter, m_j can be approximated by $[m_{\infty,j} - 2\sigma_{MP,j},\ m_{\infty,j} + 2\sigma_{MP,j}]$.

If predicted production data were linearly related to the model as in Eq. 2.44, then the a posteriori pdf of Eq. 2.41 is Gaussian with expectation m_∞ and covariance C_{MP}. In this case, one can sample the a posteriori pdf by calculating the Cholesky decomposition of C_{MP},

$$C_{MP} = LL^T, \quad (2.95)$$

and applying the equation

$$m_r = m_\infty + LZ_r, \quad (2.96)$$

to generate a realization. Here, Z_r is a vector of independent standard random normal deviates. By generating a set of N_r independent Z_r vectors, one can obtain a set of

N_r realizations. This set represents a correct sampling of the a posteriori pdf when data are linearly related to the model. However, the results of Oliver et al. [41], indicate that one can not assume that this is a valid procedure when the relationship between data and model is nonlinear. The latter reference illustrates that Eq. 2.96 may provide a very poor sampling procedure even when production data consists solely of well-test pressure data obtained under single-phase flow conditions.

As the randomized maximum likelihood method appears to be the most reliable sampling procedure available, it is the one used in this work. Note one could generate a large number of realizations from this method, and estimate means, variances and covariances from the set of realizations to evaluate the uncertainty in the rock property fields. However, characterizing the uncertainty in predicted reservoir performance is of greater interest. Thus, we simulate reservoir performance for each realization and construct statistics from the suite of predicted outcomes for each predicted parameter of interest (e.g., cumulative oil production) in order to estimate the uncertainty in each prediction parameter.

CHAPTER III

GENERATION OF SENSITIVITY COEFFICIENTS BY THE ADJOINT METHOD

Here, a formal derivation of the adjoint method used to generate sensitivity of pressures and water-oil ratios to the rock property fields is presented.

3.1 Formulation of the Adjoint Problem

In this major section, we present a derivation of the discrete adjoint equations. The adjoint variables obtained by solving this system of equations are used later to generate the sensitivity coefficients related to production data.

3.1.1 Flow Equations and Reservoir Simulator

Throughout, oil-field units are used and the constants C_1 and C_2, respectively, are defined as

$$C_1 = 1.127 \times 10^{-3} \tag{3.1}$$

$$C_2 = 5.615 \tag{3.2}$$

As derived in Chapter II (see Eq. 2.8), the semidiscrete flow equation is given by

$$\begin{aligned} &T_{mx,i+1/2,j,k}(t)(p_{i+1,j,k}(t) - p_{i,j,k}(t)) - T_{mx,i-1/2,j,k}(t)(p_{i,j,k}(t) - p_{i-1,j,k}(t)) + \\ &\quad T_{my,i,j+1/2,k}(t)(p_{i,j+1,k}(t) - p_{i,j,k}(t)) - T_{my,i,j-1/2,k}(t)(p_{i,j,k}(t) - p_{i,j-1,k}(t)) + \\ &\quad T_{mz,i,j,k+1/2}(t)(p_{i,j,k+1}(t) - p_{i,j,k}(t)) - T_{mz,i,j,k-1/2}(t)(p_{i,j,k}(t) - p_{i,j,k-1}(t)) = \\ &\qquad \Big(\frac{\Delta x_i \Delta y_j \Delta z_k \phi_{i,j,k}}{C_2}\Big)\frac{\partial (b_m(t) S_m(t))_{i,j,k}}{\partial t} + q_{m,i,j,k}(t) \end{aligned} \tag{3.3}$$

for $m = o, w$. The term $q_{m,i,j,k}(t)$ is the source or sink term for the gridblock centered at (x_i, y_j, z_k) and has units of STB/D, with $q > 0$ corresponding to production and $q < 0$ corresponding to injection.

As noted in Chapter II, the current simulator is an IMPES simulator, however, we will use a fully implicit formulation of the adjoint equations. We let Δt^l, $l = 0, 1, ...$, denote the time steps used in the reservoir simulator so that $t^{l+1} = t^l + \Delta t^l$ where $t^0 = 0$. The time steps used in the adjoint solution do not need to be identical to the time steps used in the reservoir simulator, but for now, we will assume they are identical. If they are not, one may need to interpolate between the pressure and saturation values generated in a simulation run to obtain pressure and saturation values at time steps used in the adjoint solution. However, preliminary results not discussed here indicate that one can take larger time steps in the adjoint solution than in the simulator. In this case, it is easiest to just use a subset of the simulator times (t^l's) when solving the adjoint equations. This avoids the need to do interpolation. Whether our recent observation that one can generally use far larger time steps in the adjoint solution is (a) a consequence of the fact the simulator is based on the IMPES method and the adjoint equations are solved by a fully implicit formulation, (b) a consequence of the fact that the simulator problem is nonlinear, whereas the adjoint problem is linear, (c) a consequence of the fact that the adjoint solutions are solved backward in time using a unit source at a specific points (x_i, y_j, z_k, t^l) and except near the time where the source is located, the adjoint solutions vary slowly and smoothly, (d) a combination of (a), (b) and (c), or (e) a non-general observation, is unclear at this time.

To simplify notation, we define

$$V^l_{\phi,i,j,k} = \frac{\Delta x_i \Delta y_j \Delta z_k \phi_{i,j,k}}{C_2 \Delta t^{l-1}} \tag{3.4}$$

denote the left-hand side of Eq. 3.3 by $f_{m,i,j,k}(t)$ for $m = o, w$ and let $f^l_{m,i,j,k}$ denote this term evaluated at t^l for $l = 0, 1, ...$. Using this notation, setting $t = t^{l+1}$ in Eq. 3.3

and performing a backward difference in time, it follows that

$$f^{l+1}_{m,i,j,k} = V^{l+1}_{\phi,i,j,k}\Big[\big(b_m S_m\big)^{l+1}_{i,j,k} - \big(b_m S_m\big)^{l}_{i,j,k}\Big] + q^{l+1}_{m,i,j,k}, \tag{3.5}$$

for $m = o, w$.

From this point on we will assume that there are M simulator gridblocks which have been ordered $1, 2, \ldots M$. How gridblocks are ordered is irrelevant to the discussion except, porosities, permeabilities, pressures and saturations, and the differenced flow equations must be ordered consistently. Relative to this ordering, we define relevant M-dimensional vectors. First, we define the $M \times M$ matrix, V^l_ϕ, by

$$V^l_\phi = \begin{bmatrix} V^l_{\phi,1} & 0 & 0 & \ldots & 0 \\ 0 & V^l_{\phi,2} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & V^l_{\phi,M} \end{bmatrix} \tag{3.6}$$

for all l where $V^l_{\phi,j}$ denotes the pore volume of gridblock j in ft^3 divided by $C_2\Delta t^{l_1}$. Similar notation will be used for the other variables.

Relative to the gridblock ordering, we define for $l = 0, 1, 2, \ldots$,

$$p^l = \big[p^l_1, p^l_2, \ldots, p^l_M\big]^T. \tag{3.7}$$

For $l = 0, 1, 2, \ldots$ and $m = 0, w$, we define

$$S^l_m = \big[S^l_{m1}, S^l_{m2}, \ldots, S^l_{mM}\big]^T, \tag{3.8}$$

$$(bS)^l_m = \big[b^l_{m1}S^l_{m1}, b^l_{m2}S^l_{m2}, \ldots b_{mM}S^l_{mM}\big]^T, \tag{3.9}$$

$$q^l_m = \big[q^l_{m1}, q^l_{m2}, \ldots, q^l_{mM}\big]^T, \tag{3.10}$$

and

$$f_m^l = \left[f_{m1}^l, f_{m2}^l, \ ..., \ f_{mM}^l\right]^T. \tag{3.11}$$

Throughout, the superscript T on a matrix represents the operation of taking the matrix transpose. With the preceding notation, Eq. 3.5 can be written as

$$f_m^{l+1} = V_\phi^{l+1}\left[\left(bS\right)_m^{l+1} - \left(bS\right)_m^{l}\right] + q_m^{l+1}, \tag{3.12}$$

for $m = o, w$.

3.1.2 The Adjoint Functional

Even though the finite difference equations solved in the IMPES simulator are not identical to Eq. 3.12, we formulate the adjoint solution based on this equation. In essence, we are assuming that the pressures and saturations obtained from our IMPES simulator satisfy Eq. 3.12. If this is not exactly true, some error may be introduced into our adjoint solutions, however, in all computations we have done, we have not encountered any significant inaccuracy due to this approach. However, we always use small time steps in the IMPES solutions, with a maximum time step of no more than four days. The alternative would be to form the adjoint problem directly from the IMPES finite-difference equations, but as we intend to use an implicit finite-difference procedure for three-dimensional problems, we formulate the adjoint problem assuming that Eq. 3.12 is satisfied by the IMPES finite-difference solution.

We define vectors of gridblock permeabilities and porosities as follows:

$$k_x = \left[k_{x,1}, k_{x,2}, \ ..., \ k_{x,M}\right]^T \tag{3.13}$$

$$k_y = \left[k_{y,1}, k_{y,2}, \ ..., \ k_{y,M}\right]^T \tag{3.14}$$

$$k_z = \left[k_{z,1}, k_{z,2}, \ ..., \ k_{z,M}\right]^T \tag{3.15}$$

and

$$\phi = \left[\phi_1, \phi_2, \ ..., \ \phi_M\right]^T \tag{3.16}$$

Although we use k_x, k_y, k_z and ϕ as both scalar and vector variables, which is meant should be clear from the context and the fact that the scalar forms normally have index subscripts attached to them to indicate the gridblock in which they are located. We denote the differentials of these variables by

$$dk_x = \left[dk_{x,1}, dk_{x,2}, \ ..., \ dk_{x,M}\right]^T \tag{3.17}$$

$$dk_y = \left[dk_{y,1}, dk_{y,2}, \ ..., \ dk_{y,M}\right]^T \tag{3.18}$$

$$dk_z = \left[dk_{z,1}, dk_{z,2}, \ ..., \ dk_{z,M}\right]^T \tag{3.19}$$

and

$$d\phi = \left[d\phi_1, d\phi_2, \ ..., \ d\phi_M\right]^T \tag{3.20}$$

We define a general function by

$$g = g(p^1, ..., p^L, S_{w1}, ..., S_{wL}, k_x, k_y, k_z). \tag{3.21}$$

The function is arbitrary, but the choice is dictated by the sensitivity coefficients we wish to compute. In fact, the adjoint method is formulated so that we can calculate the sensitivity coefficients of g. It is perfectly acceptable to choose g as the objective function given by the sum of squares of production data misfit terms; in fact, this is the standard choice. Our choice of Eq. 3.21 is dictated by our desire to find the sensitivity coefficients related to pressure (and/or saturation), e.g., find

$$\frac{\partial p_j^l}{\partial \alpha}, \tag{3.22}$$

where α is one of the model parameters, i.e., a gridblock permeability or porosity and p_j^l is the pressure at gridblock j at time t^l. Gridblock j represents a gridblock containing a well at which we have measured data that we wish to use as conditioning data. As shown later, flowing wellbore pressure sensitivities can be computed directly from the associated sensitivity coefficient of the associated gridblock pressure. If we wish to compute the sensitivity coefficient given by the expression of Eq. 3.22, we can simply choose

$$g(p^l, S_w^l) = p_j^l. \tag{3.23}$$

For all l, we now define two M-dimensional vectors of adjoint variables,

$$\lambda_w^l = \left[\lambda_{w,1}^l, \lambda_{w,2}^l, \ldots \lambda_{w,M}^l\right]^T, \tag{3.24}$$

and

$$\lambda_o^l = \left[\lambda_{o,1}^l, \lambda_{o,2}^l, \ldots \lambda_{o,M}^l\right]^T. \tag{3.25}$$

Here, $\lambda_{m,j}^l$ denotes the scalar adjoint variable for phase m at gridblock j at time t^l. Note at each time t^l, we have two adjoint variables per gridblock. Note this number corresponds to the reservoir simulation problem in which, at each time step, we solve for two unknowns at each gridblock, pressure and water saturation.

We now "adjoin" Eq. 3.12 to the function g (3.21) to obtain the functional J given by

$$J = g + \sum_{m=o,w} \sum_{l=0}^{L-1} \left(\lambda_m^{l+1}\right)^T \left[f_m^{l+1} - q_m^{l+1} - V_\phi^{l+1}\left((bS)_m^{l+1} - (bS)_m^l\right)\right], \tag{3.26}$$

where for now, we may think of L as the number of time steps used to generate pressure and saturation from the reservoir simulator. Since pressure and saturation obtained from the finite difference simulator are assumed to satisfy Eq. 3.12, the terms in the sum are zero for any values of the adjoint variables. This is convenient because it means that the derivative of J with respect to any model parameter α (a

gridblock permeability or porosity) is the same as the derivative of g with respect to any model parameter, i.e.,

$$\frac{\partial J}{\partial \alpha} = \frac{\partial g}{\partial \alpha}. \tag{3.27}$$

Thus the sensitivity coefficients for J and g are identical.

3.1.3 Gradients and Differentials

From the viewpoint of computing total differentials, we consider f_m^{l+1} and q_m^{l+1} as functions of permeabilities and gridblock pressures and saturations at time t^{l+1}, i.e.,

$$f_m^{l+1} = f_m(p^{l+1}, S_w^{l+1}, k_x, k_y, k_z), \tag{3.28}$$

and

$$q_m^{l+1} = q_m(p^{l+1}, S_w^{l+1}, k_x, k_y, k_z). \tag{3.29}$$

The total differential of the ith component of f_m is then given by

$$df_{mi}^{l+1} = \sum_{j=1}^{M} \Big(\frac{\partial f_{mi}^{l+1}}{\partial p_j^{l+1}} dp_j^{l+1} + \frac{\partial f_{mi}^{l+1}}{\partial S_{wj}^{l+1}} dS_{wj}^{l+1} + \frac{\partial f_{mi}^{l+1}}{\partial k_{x,j}} dk_{x,j} + \frac{\partial f_{mi}^{l+1}}{\partial k_{y,j}} dk_{y,j} + \frac{\partial f_{mi}^{l+1}}{\partial k_{z,j}} dk_{z,j} \Big). \tag{3.30}$$

If we define the gradient operator in the standard way, then

$$\nabla_{p^{l+1}} f_{mi} = \begin{bmatrix} \frac{\partial f_{mi}}{\partial p_1^{l+1}} \\ \frac{\partial f_{mi}}{\partial p_2^{l+1}} \\ \vdots \\ \frac{\partial f_{mi}}{\partial p_M^{l+1}} \end{bmatrix}, \tag{3.31}$$

$$\nabla_{k_x} f_{mi} = \begin{bmatrix} \frac{\partial f_{mi}}{\partial k_{x,1}} \\ \frac{\partial f_{mi}}{\partial k_{x,2}} \\ \vdots \\ \frac{\partial f_{mi}}{\partial k_{x,M}} \end{bmatrix}, \tag{3.32}$$

with similar notation used for gradients with respect to the M dimensional vectors S_w^{l+1}, k_x, k_y, k_z and ϕ. With this notation, Eq. 3.30 can be rewritten as

$$df_{mi}^{l+1} = \left(\nabla_{p^{l+1}} f_{mi}^{l+1}\right)^T dp^{l+1} + \left(\nabla_{S_w^{l+1}} f_{mi}^{l+1}\right)^T dS_w^{l+1} + \left(\nabla_{k_x} f_{mi}^{l+1}\right)^T dk_x + \left(\nabla_{k_y} f_{mi}^{l+1}\right)^T dk_y + \left(\nabla_{k_z} f_{mi}^{l+1}\right)^T dk_z, \quad (3.33)$$

where dk_x, dk_y and dk_z respectively, are given by Eqs. 3.17, 3.18 and 3.19,

$$dp^{l+1} = \left[dp_1^{l+1}, dp_2^{l+1}, \ldots dp_M^{l+1}\right]^T, \quad (3.34)$$

and

$$dS_w^{l+1} = \left[dS_{w1}^{l+1}, S_{w2}^{l+1}, \ldots S_{wM}^{l+1}\right]^T. \quad (3.35)$$

Using gradient notation and recalling Eq. 3.11, we also see that for any l,

$$\nabla_{k_x}\left[\left(f_m^l\right)^T\right] = \begin{bmatrix} \frac{\partial}{\partial k_{x,1}} \\ \frac{\partial}{\partial k_{x,2}} \\ \vdots \\ \frac{\partial}{\partial k_{x,M}} \end{bmatrix} [f_m^l]^T = \begin{bmatrix} \frac{\partial}{\partial k_{x,1}} \\ \frac{\partial}{\partial k_{x,2}} \\ \vdots \\ \frac{\partial}{\partial k_{x,M}} \end{bmatrix} \left[f_{m1}^l, f_{m2}^l, \ldots f_{mM}^l\right]$$

$$= \left[\nabla_{k_x} f_{m1}^l, \nabla_{k_x} f_{m2}^l, \ldots \nabla_{k_x} f_{mM}^l\right] = \begin{bmatrix} \frac{\partial f_{m1}^l}{\partial k_{x,1}} & \frac{\partial f_{m2}^l}{\partial k_{x,1}} & \cdots & \frac{\partial f_{mM}^l}{\partial k_{x,1}} \\ \frac{\partial f_{m1}^l}{\partial k_{x,2}} & \frac{\partial f_{m2}^l}{\partial k_{x,2}} & \cdots & \frac{\partial f_{mM}^l}{\partial k_{x,2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{m1}^l}{\partial k_{x,M}} & \frac{\partial f_{m2}^l}{\partial k_{x,M}} & \cdots & \frac{\partial f_{mM}^l}{\partial k_{x,M}} \end{bmatrix} \quad (3.36)$$

Expressions similar to Eq. 3.36 can be written for the gradient of f_m^l with respect to other vector variables.

3.1.4 Total Differential of Flux Terms

Applying all of the preceding notation, and using Eq. 3.33, we see that for $l = 0, 1, ...$

$$df_m^l = \begin{bmatrix} df_{m1}^l \\ df_{m2}^l \\ \vdots \\ df_{mM}^l \end{bmatrix} = \begin{bmatrix} \left(\nabla_{p^l} f_{m1}^l\right)^T dp^l + \left(\nabla_{S_w^l} f_{m1}^l\right)^T dS_w^l \\ \left(\nabla_{p^l} f_{m2}^l\right)^T dp^l + \left(\nabla_{S_w^l} f_{m2}^l\right)^T dS_w^l \\ \vdots \\ \left(\nabla_{p^l} f_{mM}^l\right)^T dp^l + \left(\nabla_{S_w^l} f_{mM}^l\right)^T dS_w^l \end{bmatrix} +$$

$$\begin{bmatrix} \left(\nabla_{k_x} f_{m1}^l\right)^T dk_x + \left(\nabla_{k_y} f_{m1}^l\right)^T dk_y + \left(\nabla_{k_z} f_{m1}^l\right)^T dk_z \\ \left(\nabla_{k_x} f_{m2}^l\right)^T dk_x + \left(\nabla_{k_y} f_{m2}^l\right)^T dk_y + \left(\nabla_{k_z} f_{m2}^l\right)^T dk_z \\ \vdots \\ \left(\nabla_{k_x} f_{mM}^l\right)^T dk_x + \left(\nabla_{k_y} f_{mM}^l\right)^T dk_y + \left(\nabla_{k_z} f_{mM}^l\right)^T dk_z \end{bmatrix} =$$

$$\begin{bmatrix} \left(\nabla_{p^l} f_{m1}^l\right)^T \\ \left(\nabla_{p^l} f_{m2}^l\right)^T \\ \vdots \\ \left(\nabla_{p^l} f_{mM}^l\right)^T \end{bmatrix} dp^l + \begin{bmatrix} \left(\nabla_{S_w^l} f_{m1}^l\right)^T \\ \left(\nabla_{S_w^l} f_{m2}^l\right)^T \\ \vdots \\ \left(\nabla_{S_w^l} f_{mM}^l\right)^T \end{bmatrix} dS_w^l +$$

$$\begin{bmatrix} \left(\nabla_{k_x} f_{m1}^l\right)^T \\ \left(\nabla_{k_x} f_{m2}^l\right)^T \\ \vdots \\ \left(\nabla_{k_x} f_{mM}^l\right)^T \end{bmatrix} dk_x + \begin{bmatrix} \left(\nabla_{k_y} f_{m1}^l\right)^T \\ \left(\nabla_{k_y} f_{m2}^l\right)^T \\ \vdots \\ \left(\nabla_{k_y} f_{mM}^l\right)^T \end{bmatrix} dk_y + \begin{bmatrix} \left(\nabla_{k_z} f_{m1}^l\right)^T \\ \left(\nabla_{k_z} f_{m2}^l\right)^T \\ \vdots \\ \left(\nabla_{k_z} f_{mM}^l\right)^T \end{bmatrix} dk_z =$$

$$\begin{bmatrix} \left(\nabla_{p^l} f_{m1}^l\right) & \left(\nabla_{p^l} f_{m2}^l\right) & \dots & \left(\nabla_{p^l} f_{mM}^l\right) \end{bmatrix}^T dp^l +$$

$$\begin{bmatrix} \left(\nabla_{S_w^l} f_{m1}^l\right) & \left(\nabla_{S_w^l} f_{m2}^l\right) & \dots & \left(\nabla_{S_w^l} f_{mM}^l\right) \end{bmatrix}^T dS_w^l +$$

$$\begin{bmatrix} \left(\nabla_{k_x} f_{m1}^l\right) & \dots & \left(\nabla_{k_x} f_{mM}^l\right) \end{bmatrix}^T dk_x + \begin{bmatrix} \left(\nabla_{k_y} f_{m1}^l\right) & \dots & \left(\nabla_{k_y} f_{mM}^l\right) \end{bmatrix}^T dk_y +$$

$$\begin{bmatrix} \left(\nabla_{k_z} f_{m1}^l\right) & \dots & \left(\nabla_{k_z} f_{mM}^l\right) \end{bmatrix}^T dk_z =$$

$$\left(\nabla_{p^l}(f_m^T)\right)^T dp^l + \left(\nabla_{S_w^l}(f_m^T)\right)^T dS_w^l +$$

$$\left(\nabla_{k_x}(f_m^T)\right)^T dk_x + \left(\nabla_{k_y}(f_m^T)\right)^T dk_y + \left(\nabla_{k_z}(f_m^T)\right)^T dk_z. \quad (3.37)$$

3.1.5 Total Differential of J

To further simplify notation, we introduce flow and accumulation vector variables, respectively, by

$$F_m^{l+1} = F_m^{l+1}(p^{l+1}, S_w^{l+1}, k_x, k_y, k_z) = f_m^{l+1} - q_m^{l+1}, \tag{3.38}$$

$$A_m^{l+1} = A_m^{l+1}(p^{l+1}, S_w^{l+1}, \phi) = V_\phi^{l+1}(bS)_m^{l+1}, \tag{3.39}$$

and

$$\tilde{A}_m^{l} = \tilde{A}_m^{l}(p^{l}, S_w^{l}, \phi) = V_\phi^{l+1}(bS)_m^{l}, \tag{3.40}$$

for $m = 0, w$ and $l = 0, 1, ..., L-1$. Note V_ϕ^{l+1} appears in both Eqs. 3.39 and 3.40, but the only reason this term is time dependent is that it involves Δt^l; see Eqs. 3.4 and 3.6. If a constant time step were used, this term would be independent of time and we would have $A_m^l = \tilde{A}_m^l$. Again, we let t^L represent the termination time of a reservoir simulation run. Definitions of relevant terms in the two preceding equations are given in Eqs. 3.6 through 3.11 . The notation used in Eqs. 3.38–3.40 indicates that F_M^{l+1} involves only the gridblock pressures and saturations at time t^{l+1} and all gridblock permeabilities, whereas A_m^{l+1} and $\tilde{A}_m^{l+1}$ explicitly involve the same pressures and saturations as well as gridblock porosities, but not permeabilities. Thus, in computing total differentials of these vector variables, we can use

$$\nabla_\phi\left[\left(F_m^{l+1}\right)^T\right] = O, \tag{3.41}$$

$$\nabla_{k_x}\left[\left(A_m^{l+1}\right)^T\right] = \nabla_{k_y}\left[\left(A_m^{l+1}\right)^T\right] = \nabla_{k_z}\left[\left(A_m^{l+1}\right)^T\right] = O, \tag{3.42}$$

and

$$\nabla_{k_x}\left[\left(\tilde{A}_m^{l}\right)^T\right] = \nabla_{k_y}\left[\left(\tilde{A}_m^{l}\right)^T\right] = \nabla_{k_z}\left[\left(\tilde{A}_m^{l}\right)^T\right] = O, \tag{3.43}$$

where O represents the $M \times M$ null matrix; i.e., all its entries are equal to zero.

Using the the notation of Eqs. 3.38–3.40,we can write Eq. 3.26 as

$$J = g + \sum_{m=o,w} \sum_{l=0}^{L-1} \left(\lambda_m^{l+1}\right)^T \left[F_m^{l+1} - A_m^{l+1} + \tilde{A}_m^l\right]. \tag{3.44}$$

Taking the total differential of Eq. 3.44, using operational formulas like Eq. 3.37, we obtain

$$\begin{aligned} dJ = dg + \sum_{m=o,w} \sum_{l=0}^{L-1} \Big(\left(\lambda_m^{l+1}\right)^T \Big[\left(\nabla_{p^{l+1}} [F_m^{l+1} - A_m^{l+1}]^T\right)^T dp^{l+1} + \\ \left(\nabla_{S_w^{l+1}} [F_m^{l+1} - A_m^{l+1}]^T\right)^T dS_w^{l+1} + \\ \left(\nabla_{k_x} [F_m^{l+1}]^T\right)^T dk_x + \left(\nabla_{k_y} [F_m^{l+1}]^T\right)^T dk_y + \left(\nabla_{k_z} [F_m^{l+1}]^T\right)^T dk_z - \\ \left(\nabla_{\phi} [A_m^{l+1}]^T\right)^T d\phi + \left(\nabla_{\phi} [\tilde{A}_m^l]^T\right)^T d\phi + \\ \left(\nabla_{p^l} [\tilde{A}_m^l]^T\right)^T dp^l + \left(\nabla_{S_w^l} [\tilde{A}_m^l]^T\right)^T dS_w^l \Big] \Big). \end{aligned} \tag{3.45}$$

By simply changing the index of summation, we have

$$\begin{aligned} \sum_{l=0}^{L-1} \left(\lambda_m^{l+1}\right)^T \left(\nabla_{p^{l+1}} [F_m^{l+1} - A_m^{l+1}]^T\right)^T dp^{l+1} = \\ \sum_{l=1}^{L-1} \left(\left(\lambda_m^l\right)^T \left(\nabla_{p^l} [F_m^l - A_m^l]^T\right)^T dp^l \right) + \\ \left(\lambda_m^L\right)^T \left(\nabla_{p^L} [F_m^L - A_m^L]^T\right)^T dp^L \end{aligned} \tag{3.46}$$

and

$$\begin{aligned} \sum_{l=0}^{L-1} \left(\lambda_m^{l+1}\right)^T \left(\nabla_{S_w^{l+1}} [F_m^{l+1} - A_m^{l+1}]^T\right)^T dS_w^{l+1} = \\ \sum_{l=1}^{L-1} \left(\left(\lambda_m^l\right)^T \left(\nabla_{S_w^l} [F_m^l - A_m^l]^T\right)^T dS_w^l \right) + \\ \left(\lambda_m^L\right)^T \left(\nabla_{S_w^L} [F_m^L - A_m^L]^T\right)^T dS_w^L. \end{aligned} \tag{3.47}$$

We define

$$BT^L = \sum_{m=o,w} \left(\lambda_m^L\right)^T \left[\left(\nabla_{p^L} [F_m^L - A_m^L]^T\right)^T dp^L + \left(\nabla_{S_w^L} [F_m^L - A_m^L]^T\right)^T dS_w^L \right]. \tag{3.48}$$

Using Eqs. 3.46, 3.47 and 3.48, Eq. 3.45 can be rewritten as

$$dJ = dg + \sum_{m=o,w}\sum_{l=1}^{L-1}\left(\lambda_m^l\right)^T\Big[\left(\nabla_{p^l}[F_m^l - A_m^l]^T\right)^T dp^l + \left(\nabla_{S_w^l}[F_m^l - A_m^l]^T\right)^T dS_w^l\Big] +$$
$$\sum_{m=o,w}\sum_{l=0}^{L-1}\Big(\left(\lambda_m^{l+1}\right)^T\Big[\left(\nabla_{k_x}[F_m^{l+1}]^T\right)^T dk_x + \left(\nabla_{k_y}[F_m^{l+1}]^T\right)^T dk_y + \left(\nabla_{k_z}[F_m^{l+1}]^T\right)^T dk_z -$$
$$\left(\nabla_{\phi}[A_m^{l+1}]^T\right)^T d\phi + \left(\nabla_{\phi}[\tilde{A}_m^{l}]^T\right)^T d\phi + \left(\nabla_{p^l}[\tilde{A}_m^{l}]^T\right)^T dp^l + \left(\nabla_{S_w^l}[\tilde{A}_m^{l}]^T\right)^T dS_w^l\Big]\Big) + BT^L. \tag{3.49}$$

Since p^0 and S_w^0 represented fixed specified initial conditions, they will not change if the rock property fields are perturbed. Thus, their differentials are zero, i.e.,

$$dp^0 = dS_w^0 = 0. \tag{3.50}$$

Thus, we can add or delete any terms that involve a matrix multiplication of these differential from any equation, e.g.,

$$\left(\nabla_{p^0}[\tilde{A}_m^0]^T\right)^T dp^0 = \left(\nabla_{S_w^0}[\tilde{A}_m^0]^T\right)^T dS_w^0 = 0. \tag{3.51}$$

In all cases of interest to us, the function g will be chosen so that it depends explicitly on one or more of the p^l and S_w^l vectors, $l = 1, \ldots L-1$ and possibly k_x, k_y and k_z. Thus, using an operational formula like the one of Eq. 3.33, the total differential of g can be written as

$$dg = \Big\{\sum_{l=1}^{L-1}\Big([\nabla_{p^l} g]^T dp^l + [\nabla_{S_w^l}]^T dS_w^l\Big)\Big\} +$$
$$[\nabla_{k_x} g]^T dk_x + [\nabla_{k_y} g]^T dk_y + [\nabla_{k_z} g]^T dk_z + [\nabla_{\phi} g]^T d\phi. \tag{3.52}$$

Using Eq. 3.52 in Eq. 3.49 and rearranging the resulting equation gives

$$
\begin{aligned}
dJ = \\
&\sum_{l=1}^{L-1}\Big\{\Big(\Big[\sum_{m=o,w}\left(\lambda_m^l\right)^T\left(\nabla_{p^l}[F_m^l - A_m^l]^T\right)^T\Big] + \left(\nabla_{p^l}g\right)^T + \Big[\sum_{m=o,w}(\lambda_m^{l+1})^T\left(\nabla_{p^l}[\tilde{A}_m^l]^T\right)^T\Big]\Big)dp^l + \\
&\Big(\Big[\sum_{m=o,w}\left(\lambda_m^l\right)^T\left(\nabla_{S_w^l}[F_m^l - A_m^l]^T\right)^T\Big] + \left(\nabla_{S_w^l}g\right)^T + \Big[\sum_{m=o,w}(\lambda_m^{l+1})^T\left(\nabla_{S_w^l}[\tilde{A}_m^l]^T\right)^T\Big]\Big)dS_w^l\Big\} + \\
&\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\left(\lambda_m^{l+1}\right)^T\left(\nabla_{k_x}[F_m^{l+1}]^T\right)^T\Big) + \left(\nabla_{k_x}g\right)^T\Big\}dk_x + \\
&\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\left(\lambda_m^{l+1}\right)^T\left(\nabla_{k_y}[F_m^{l+1}]^T\right)^T\Big) + \left(\nabla_{k_y}g\right)^T\Big\}dk_y + \\
&\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\left(\lambda_m^{l+1}\right)^T\left(\nabla_{k_z}[F_m^{l+1}]^T\right)^T\Big) + \left(\nabla_{k_z}g\right)^T\Big\}dk_z + \\
&\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\left(\lambda_m^{l+1}\right)^T\Big[\left(\nabla_{\phi}[\tilde{A}_m^l]^T\right)^T - \left(\nabla_{\phi}[A_m^{l+1}]^T\right)^T\Big]\Big) + \left(\nabla_{\phi}g\right)^T\Big\}d\phi + BT^L. \quad (3.53)
\end{aligned}
$$

3.1.6 Discrete Adjoint Equations

Next we choose the adjoint variables to insure that the coefficients multiplying dp^l and dS_w^l in Eq. 3.53 vanish, for $l = 1, 2, ..., L-1$, i.e., we require that

$$\Big\{\sum_{m=o,w}\Big[\left(\lambda_m^l\right)^T\left(\nabla_{p^l}[F_m^l - A_m^l]^T\right)^T + (\lambda_m^{l+1})^T\left(\nabla_{p^l}[\tilde{A}_m^l]^T\right)^T\Big]\Big\} + \left(\nabla_{p^l}g\right)^T = 0, \quad (3.54)$$

and

$$\Big\{\sum_{m=o,w}\Big[\left(\lambda_m^l\right)^T\left(\nabla_{S_w^l}[F_m^l - A_m^l]^T\right)^T + (\lambda_m^{l+1})^T\left(\nabla_{S_w^l}[\tilde{A}_m^l]^T\right)^T\Big]\Big\} + \left(\nabla_{S_w^l}g\right)^T = 0, \quad (3.55)$$

where the zeros on the right-side of Eqs. 3.54 and 3.55 denote M-dimensional row vectors with all entries equal to zero. Taking transposes of Eqs. 3.54 and 3.55, re-

spectively, and rearranging the resulting equations gives

$$\sum_{m=o,w} \left[\left(\nabla_{p^l} [F_m^l - A_m^l]^T \right) \lambda_m^l + \nabla_{p^l} [\tilde{A}_m^l]^T \lambda_m^{l+1} \right] = -\nabla_{p^l} g, \tag{3.56}$$

and

$$\sum_{m=o,w} \left[\nabla_{S_w^l} [F_m^l - A_m^l]^T \lambda_m^l + \nabla_{S_w^l} [\tilde{A}_m^l]^T \lambda_m^{l+1} \right] = -\nabla_{S_w^l} g, \tag{3.57}$$

for $l = 1, 2, ..., L-1$. Eqs. 3.56 and 3.57 represent the discrete system of adjoint equations. To solve for the adjoint variables, we need to specify additional constraints. As is standard, we specify these by

$$\lambda_o^L = \lambda_w^L = 0, \tag{3.58}$$

where the right side of Eq. 3.58 (as well as Eqs. 3.56 and 3.57) is the M-dimensional column vector with all entries equal to zero. Using Eq. 3.58 in Eq. 3.48, it follows that

$$BT^L = 0. \tag{3.59}$$

By applying the auxiliary condition of Eq. 3.58 we can solve the adjoint system (Eqs. 3.56 and 3.57) backward in time, i.e., solve the system for $l = L-1, L-2, \ldots 0$. We postpone discussion of the solution of this adjoint system until later. Our next task is to derive the expressions for sensitivity coefficients.

3.2 General Formulas for Computing Sensitivity Coefficients

Considering J as a function of ϕ, k_x, k_y and k_z, we can write its total differential as

$$dJ = \left(\nabla_{k_x} J \right)^T dk_x + \left(\nabla_{k_y} J \right)^T dk_y + \left(\nabla_{k_z} J \right)^T dk_z + \left(\nabla_{\phi} J \right)^T d\phi. \tag{3.60}$$

By applying Eqs. 3.56, 3.57 and 3.59, Eq. 3.53 reduces to

$$\begin{aligned} dJ = &\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\big(\nabla_{k_x}[F_m^{l+1}]^T\big)^T\Big) + (\nabla_{k_x}g)^T\Big\}dk_x + \\ &\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\big(\nabla_{k_y}[F_m^{l+1}]^T\big)^T\Big) + (\nabla_{k_y}g)^T\Big\}dk_y + \\ &\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\big(\nabla_{k_z}[F_m^{l+1}]^T\big)^T\Big) + (\nabla_{k_z}g)^T\Big\}dk_z + \\ &\Big\{\Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\Big[\big(\nabla_\phi[\tilde{A}_m^l]^T\big)^T - \big(\nabla_\phi[A_m^{l+1}]^T\big)^T\Big]\Big) + (\nabla_\phi g)^T\Big\}d\phi \end{aligned} \tag{3.61}$$

By comparing Eqs. 3.60 and 3.61, it follows that

$$(\nabla_{k_x}J)^T = \Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\big(\nabla_{k_x}[F_m^{l+1}]^T\big)^T\Big) + (\nabla_{k_x}g)^T, \tag{3.62}$$

$$(\nabla_{k_y}J)^T = \Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\big(\nabla_{k_y}[F_m^{l+1}]^T\big)^T\Big) + (\nabla_{k_y}g)^T, \tag{3.63}$$

$$(\nabla_{k_z}J)^T = \Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\big(\nabla_{k_z}[F_m^{l+1}]^T\big)^T\Big) + (\nabla_{k_z}g)^T, \tag{3.64}$$

and

$$(\nabla_\phi J)^T = \Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\lambda_m^{l+1}\big)^T\Big[\big(\nabla_\phi[\tilde{A}_m^l]^T\big)^T - \big(\nabla_\phi[A_m^{l+1}]^T\big)^T\Big]\Big) + (\nabla_\phi g)^T. \tag{3.65}$$

Eqs. 3.62 through 3.65 give the desired sensitivity coefficients. For example, Eq. 3.62 gives the sensitivity to the x-direction permeability field. Taking transposes of the equations for the sensitivity coefficients gives

$$\nabla_{k_x}J = \begin{bmatrix} \frac{\partial J}{\partial k_{x,1}} \\ \frac{\partial J}{\partial k_{x,2}} \\ \vdots \\ \frac{\partial J}{\partial k_{x,M}} \end{bmatrix} = \Big(\sum_{m=o,w}\sum_{l=0}^{L-1}\big(\nabla_{k_x}[F_m^{l+1}]^T\big)\lambda_m^{l+1}\Big) + (\nabla_{k_x}g), \tag{3.66}$$

$$\nabla_{k_y} J = \begin{bmatrix} \frac{\partial J}{\partial k_{y,1}} \\ \frac{\partial J}{\partial k_{y,2}} \\ \vdots \\ \frac{\partial J}{\partial k_{y,M}} \end{bmatrix} = \Big(\sum_{m=o,w} \sum_{l=0}^{L-1} \big(\nabla_{k_y} [F_m^{l+1}]^T\big) \lambda_m^{l+1} \Big) + (\nabla_{k_y} g), \tag{3.67}$$

$$\nabla_{k_z} J = \begin{bmatrix} \frac{\partial J}{\partial k_{z,1}} \\ \frac{\partial J}{\partial k_{z,2}} \\ \vdots \\ \frac{\partial J}{\partial k_{z,M}} \end{bmatrix} = \Big(\sum_{m=o,w} \sum_{l=0}^{L-1} \big(\nabla_{k_x} [F_m^{l+1}]^T\big) \lambda_m^{l+1} \Big) + (\nabla_{k_z} g), \tag{3.68}$$

and

$$\nabla_{\phi} J = \begin{bmatrix} \frac{\partial J}{\partial phi_1} \\ \frac{\partial J}{\partial \phi_2} \\ \vdots \\ \frac{\partial J}{\partial \phi_M} \end{bmatrix} = \Big(\sum_{m=o,w} \sum_{l=0}^{L-1} \Big[\big(\nabla_{\phi} [\tilde{A}_m^l]^T\big) - \big(\nabla_{\phi} [A_m^{l+1}]^T\big) \Big] \lambda_m^{l+1} \Big) + (\nabla_{\phi} g). \tag{3.69}$$

If the permeability field is isotropic $k_x(x, y, z) = k_y(x, y, z) = k_z(x, y, z) = k(x, y, z)$, then we do not compute sensitivity coefficients for k_x, k_y and k_z and add them together. We simply use

$$\nabla_{k} J = \begin{bmatrix} \frac{\partial J}{\partial k_1} \\ \frac{\partial J}{\partial k_2} \\ \vdots \\ \frac{\partial J}{\partial k_M} \end{bmatrix} = \Big(\sum_{m=o,w} \sum_{l=0}^{L-1} \big(\nabla_{k} [F_m^{l+1}]^T\big) \lambda_m^{l+1} \Big) + (\nabla_{k} g); \tag{3.70}$$

Eqs. 3.66–3.68 are not applied. We choose specific forms of g to generate the sensitivity coefficients we want. If g depends explicitly only on pressure and saturation, the gradients of g that appear on the right side of Eqs. 3.66–3.70 are zero.

A typical choice used by others is to set g equal to the sum of squares of the data mismatch terms, but, as discussed before, this precludes generation of

the sensitivity coefficients needed to apply a Gauss-Newton method to minimize the objective function. Instead, we wish to generate the sensitivity coefficients related to each measured data point, provided of course, that we intend to use the measured data as conditioning data. We will have to make several choices for g to get all of the sensitivity coefficients we want. As the procedure has only been implemented for two-dimensional flow problems, from this point on we consider specific implementation only for two-dimensional flow in the x-y plane.

3.3 Sensitivity of Production Data to Rock Property Fields

We assume that wellbore pressures and/or other production data are measured at times t^{l_j}, $j = 1, 2, ...r$ and we wish to compute the sensitivity of these data to the model parameters (gridblock permeabilities and porosity). For simplicity, we assume that these times corresponds to some of the t^l's used in the reservoir simulator, although this is unnecessary. We may have data measured at different wells at different times. This major section focuses on the derivation of specific formulas for calculating the sensitivity of pressure and water-oil ratios to the gridblock permeabilities and porosities.

3.3.1 Sensitivity of Gridblock Pressures or Saturations.

One procedure for generating the sensitivity coefficients related to wellbore pressures is to first calculate the sensitivity of gridblock pressures to the rock property fields and then translate these results to obtain the sensitivity of flowing wellbore pressures to the rock property fields. In our simulator, gridblock pressures are related to wellbore pressures by Peaceman's equation [45]. Thus, we assume we have a measured wellbore pressure, $p^{obs}_{wf,j}(t^r)$ at time t^r, where the superscript *obs* refers to measured (observed) data and the subscript j indicates the well is located in gridblock j. Suppose one wishes to compute the sensitivity coefficients related to the calculated

pressure, $p_{wf,j}(t^r)$, that will be predicted by the simulator for a given set of gridblock permeability and porosity values, and let p_j^r denote the corresponding calculated value of gridblock pressure at time t^r.

If we choose

$$g = p_j^r, \tag{3.71}$$

then for all l,

$$\nabla_{S_w^l} g = 0. \tag{3.72}$$

For $l \neq r$,

$$\nabla_{p^l} g = 0, \tag{3.73}$$

whereas, for $l = r$,

$$\nabla_{p^r} g = \nabla_{p^r} p_j^r = e^j, \tag{3.74}$$

where the ith component of e^j is given by

$$e_i^j = \delta_{i,j}, \tag{3.75}$$

where $\delta_{i,j}$ is the Kronecker delta function, i.e., $\delta_{j,j} = 1$ and $\delta_{i,j} = 0$ for $i \neq j$. Eqs. 3.72–3.74 give the source terms to be used in the right-side of Eqs. 3.56 and 3.57 when solving for the adjoint variables. Note Eqs. 3.73 and 3.74 indicate that there is a unit source in the j component of the right-side of Eq. 3.56 (corresponding to the jth gridblock) at time $t^l = t^r$. Thus, to obtain the sensitivity coefficients for p_j^r, we simply insert the source terms given by Eqs. 3.72–3.74 into Eqs. 3.56 and 3.57 and solve the resulting adjoint equations backward in time for $l = r, r-1, r-2, \ldots, 0$.

Similarly, if we want the sensitivity coefficients for p_m^s (gridblock pressure at time t^s in gridblock m), then we choose

$$g = p_m^s, \tag{3.76}$$

so

$$\nabla_{S_w^l} g = 0, \tag{3.77}$$

for all l

$$\nabla_{p^l} g = 0, \tag{3.78}$$

for $l \neq s$, and

$$\nabla_{p^s} g = \nabla_{p^s} p_m^s = e^m, \tag{3.79}$$

where the the ith component of e^m is given by $e_i^m = \delta_{i,j}$. Using these source terms in Eqs. 3.56 and 3.57, we solve the resulting adjoint system backward in time for $l = s, s-1, ..., 0$. Viewed in this way, we have to solve one adjoint system for each data point. However, note that as we change g, only the right-hand sides of Eqs. 3.56 and 3.57 change. The coefficient matrices on the left-hand side of these equations do not change. This results in considerable computational savings because it means we solve repeatedly the same set of linear equations with different right-hand sides.

For concreteness, suppose one only wishes to generate sensitivity coefficients related to various gridblock pressures at gridblocks j and m at at times $t^{l_1} < t^{l_2} < \cdots < t^{l_r}$, then we proceed as follows: Beginning at time t^{l_r}, we begin solving two adjoint systems (Eqs. 3.56 and 3.57) backward in time using the two different source generating functions, $g = p_j^{l_r}$ and $g = p_m^{l_r}$. These equations are solved for $l = l_r, l_r - 1, \ldots, l_{r-1}$. For each value of l, this requires solving Eqs. 3.56 and 3.57 with two different right-hand sides. In doing so, two distinct sets of adjoint variables are generated, one will be used to compute the sensitivity of $p_j^{l_r}$ to the rock property fields and one will be used to generate the sensitivity coefficient related of $p_m^{l_r}$. At $l = l_{r-1}$, we need to begin generating the adjoint variables related to the sensitivity coefficients for $p_j^{l_{r-1}}$ and $p_m^{l_{r-1}}$. Thus, at this time, we begin two additional adjoint solutions based on the two source generating functions $g = p_j^{l_{r-1}}$ and $g = p_m^{l_{r-1}}$. Note, when we begin

the solution of the adjoint systems (Eqs. 3.56 and 3.57) related to the sensitivity coefficients for these two pressures ($p_j^{l_{r-1}}$ and $p_m^{l_{r-1}}$), we set $\lambda_m^{l_{r-1}+1} = 0$. Now for $l = l_{r-1}, l_{r-1} - 1, ..., l_{r-2}$, we solve four systems backward in time, but only the right-hand side vectors differ from system to system. This basic procedure is continued backward in time until we reach $l = 0$. When time is equal to a time t^{l_j} at which we wish to generate sensitivity coefficients for the two gridblock pressures, we add two more right hand side vectors and begin generating two additional sets of adjoint variables. If, as is currently done in our code, the matrix problem of Eqs. 3.56 and 3.57 is solved using a sparse matrix technique, then only one LU decomposition needs to be done per "time step" and then the backward-forward substitution procedure is applied for each right-hand side vector.

Note the procedure is not restricted to generation of the sensitivity coefficients related to gridblock pressures. For example, we can just as easily generate the sensitivity of the water saturation in gridblock j at time t^{l_k} by choosing $g = S_w^{l_k}$. The main point about the process is we add one right hand side per sensitivity coefficient. However, if at the maximum time of interest t^{l_r}, we wish to generate only n sensitivity coefficients, we have only n right-hand side vectors to consider. If at each time, t^{l_j}, $j = 1, 2, \ldots, r$ where $t^{l_j} < t^{l_{j+1}}$, we wish to generate n sensitivity coefficients, then for $l \leq l_1$, one will have to solve a matrix problem with $n \times r$ different right-hand side vectors. If data were uniformly spaced in time, then on average, each matrix problem in the adjoint system will need to be solved with $(n \times r)/2$ right-hand side vectors.

3.3.2 Sensitivity of Wellbore Pressures

Two different methods are presented for calculating wellbore pressure sensitivity coefficients. In the first method gridblock pressure and saturation sensitivity coefficients are calculated and used to generate wellbore pressure sensitivity coefficients. In the second method, wellbore pressure sensitivity coefficients are calculated directly. Since, we have incorporated the second method into our code, the first

method is not presented in detail.

Method 1. The constant C_3 is defined by

$$C_3 = (2\pi)1.127 \times 10^{-3}. \tag{3.80}$$

Using somewhat confusing notation, we assume that well j is located in gridblock j and define the well index for well j by

$$WI_j = \frac{C_3 h\sqrt{k_{x,j}k_{y,j}}}{\ln(r_{o,j}/r_{w,j}) + s_j}, \tag{3.81}$$

where

$$r_{o,j} = \frac{0.28073\Delta x_j\sqrt{1 + \frac{k_{x,j}\Delta y_j^2}{k_{y,j}\Delta x_j^2}}}{1 + \sqrt{k_{x,j}/k_{y,j}}}. \tag{3.82}$$

Here $r_{w,j}$ is the wellbore radius of well j, s_j is the skin factor for well j, and Δx_j and Δy_j, respectively, denote the x and y dimensions of gridblock j.

Applying Peaceman's equation [45] to our two-phase flow problem gives

$$q_{t,j}^r = WI_j\Big(\frac{b_{o,j}^r k_{ro,j}^r}{\mu_{o,j}^r} + \frac{b_{w,j}^r k_{rw,j}^r}{\mu_{w,j}^r}\Big)\Big(p_j^r - p_{wf,j}^r\Big), \tag{3.83}$$

where the superscript r refers to time t^r, $q_{t,j}^r$ is the total flow rate in STB/D at well j at time t^r, p_j^r is the pressure in gridblock j at this time, and $p_{wf,j}^r$ is the wellbore pressure at the well in the jth gridblock at time t^r.

Now suppose we wish to compute the sensitivity coefficient $\partial p_{wf,j}/\partial\alpha$ where α is a gridblock permeability or porosity. If the total flow rate is specified as boundary condition (wellbore constraint), then it is fixed so

$$\frac{\partial q_{t,j}^r}{\partial\alpha} = 0, \tag{3.84}$$

for all j and r. Thus, taking the derivative of Eq. 3.83 with respect to α gives

$$0 = WI_j\Big[\frac{\partial\Big(\frac{b^r_{o,j}k_{ro,j^r}}{\mu^r_{o,j}} + \frac{b^r_{w,j}k^r_{rw,j}}{\mu^r_{w,j}}\Big)}{\partial\alpha}\big(p^r_j - p^r_{wf,j}\big) + \Big(\frac{b^r_{o,j}k^r_{ro,j}}{\mu^r_{o,j}} + \frac{b^r_{w,j}k^r_{rw,j}}{\mu^r_{w,j}}\Big)\Big(\frac{\partial p^r_j}{\partial\alpha} - \frac{\partial p^r_{wf,j}}{\partial\alpha}\Big)\Big] + \frac{\partial WI_j}{\partial\alpha}\Big(\frac{b^r_{o,j}k^r_{ro,j}}{\mu^r_{o,j}} + \frac{b^r_{w,j}k^r_{rw,j}}{\mu^r_{w,j}}\Big)\big(p^r_j - p^r_{wf,j}\big). \quad (3.85)$$

Note Eq. 3.85 involves several derivatives with respect to α and to obtain all derivatives needed, requires we first compute the sensitivity coefficients $\partial S^r_{w,j}/\partial\alpha$ and $\partial p^r_j/\partial\alpha$. For example,

$$\frac{\partial k^r_{ro,j}}{\partial\alpha} = \Big(\frac{dk_{ro}}{dS_w}\Big)_{S^r_{wj}}\Big(\frac{\partial S^r_{w,j}}{\partial\alpha}\Big). \quad (3.86)$$

However, once all of these derivatives have been generated, one can rearrange Eq. 3.85to compute $\partial p^r_{wf,j}/\partial\alpha$.

Method 2. Here, we choose the source terms in the adjoint equation so that one can directly generate the adjoint variables necessary to directly compute the sensitivity of $p^r_{wf,j}$ to the rock property fields. To do so, we define

$$\beta^r_j = \Big(\frac{b^r_{o,j}k^r_{ro,j}}{\mu^r_{o,j}} + \frac{b^r_{w,j}k^r_{rw,j}}{\mu^r_{w,j}}\Big). \quad (3.87)$$

Then the source generating term is specified by

$$g = p^r_{wf,j} = p^r_j - \Big(\frac{q^r_{t,j}}{WI_j\,\beta^r_j}\Big), \quad (3.88)$$

where the second equality of Eq. 3.88 follows from Eq. 3.83 and the definition of Eq. 3.87. To compute the source terms needed in Eqs. 3.56 and 3.57, we need to take the gradients of g with respect to p^l and $S_{w,l}$. Note since g now also explicitly involves $k_{x,j}$ and $k_{y,j}$ we need to compute the gradients of g with respect to the x and y direction permeability fields that appear on the right side of Eqs. 3.66 and 3.57 because these gradients are no longer identically zero.

We begin by considering

$$\nabla_{p^l} g = \nabla_{p^l} p^r_{wf,j}. \tag{3.89}$$

For $l \neq r$,

$$\nabla_{p^l} g = \nabla_{p^l} p^r_{wf,j} = 0. \tag{3.90}$$

For $l = r$,

$$\nabla_{p^l} g = \nabla_{p^r} p^r_{wf,j} = \begin{bmatrix} \frac{\partial p^r_{wf,j}}{\partial p^r_1} \\ \frac{\partial p^r_{wf,j}}{\partial p^r_2} \\ \vdots \\ \frac{\partial p^r_{wf,j}}{\partial p^r_M} \end{bmatrix}, \tag{3.91}$$

where from Eq. 3.88, we have

$$\frac{\partial p^r_{wf,j}}{\partial p^r_i} = 0, \tag{3.92}$$

for $i \neq j$ and

$$\frac{\partial p^r_{wf,j}}{\partial p^r_j} = 1 + \Big[\frac{q_{t,j}}{WI_j(\beta^r_j)^2}\Big]\Big(\frac{\partial \beta^r_j}{\partial p^j_r}\Big). \tag{3.93}$$

Currently, in our simulator, we evaluate the inverse formation volume factors at the gridblock pressure using the relation

$$b_m = b_m(p) = b_{m,in}\big(1 - c_m(p_{in} - p)\big), \tag{3.94}$$

for $m = o, w$ where c_m is the compressibility of phase m, p_{in} is initial pressure and $b_{m,in}$ is the inverse formation volume factor of phase m at initial pressure. Thus

$$\frac{\partial b^r_{m,j}}{\partial p^r_j} = \frac{\partial b_m(p^r_j)}{\partial p^r_j} = c_m b_{m,in}. \tag{3.95}$$

Assuming constant viscosities, and recognizing that relative permeabilities do not depend on pressure, it follows that

$$\frac{\partial \beta^r_j}{\partial p^j_r} = \frac{c_o b_{o,in} k^r_{ro,j}}{\mu_o} + \frac{c_w b_{w,in} k^r_{rw,j}}{\mu_w}, \tag{3.96}$$

and thus, Eq. 3.93 becomes

$$\frac{\partial p^r_{wf,j}}{\partial p^r_j} = 1 + \Big[\frac{q_{t,j}}{WI_j(\beta^r_j)^2}\Big]\Big(\frac{c_o b_{o,in} k^r_{ro,j}}{\mu_o} + \frac{c_w b_{w,in} k^r_{rw,j}}{\mu_w}\Big). \tag{3.97}$$

Now to obtain $\nabla_{p^r} p_{wf,j}$ for $l = r$, one simply uses Eq. 3.97 and Eq. 3.92 in Eq. 3.91. Eqs. 3.91 and 3.90, provide all results needed to compute the source term on the right-side of Eq. 3.56.

Next, the source term $\nabla S^l_w g$ that appears on the right side of Eq. 3.57 is considered. Similar to the above presentation,

$$\nabla_{S^l_w} g = \nabla_{S^l_w} p^r_{wf,j} = 0, \tag{3.98}$$

for $l \neq r$ and for $l = r$,

$$\nabla_{S^l_w} g = \nabla_{S^r_w} p^r_{wf,j} = \begin{bmatrix} \frac{\partial p^r_{wf,j}}{\partial S^r_{w,1}} \\ \frac{\partial p^r_{wf,j}}{\partial S^r_{w,2}} \\ \vdots \\ \frac{\partial p^r_{wf,j}}{\partial S^r_{w,M}} \end{bmatrix}. \tag{3.99}$$

From Eq. 3.88,

$$\frac{\partial p^r_{wf,j}}{\partial S^r_{w,i}} = 0, \tag{3.100}$$

for $i \neq j$ and

$$\frac{\partial p^r_{wf,j}}{\partial S^r_{w,j}} = \Big[\frac{q_{t,j}}{WI_j(\beta^r_j)^2}\Big]\Big(\frac{b^r_{o,j}}{\mu_o}\frac{dk_{ro}(S^r_{w,j})}{dS_w} + \frac{b^r_{w,j}}{\mu_w}\frac{dk_{rw}(S^r_{w,j})}{dS_w}\Big). \tag{3.101}$$

Having done the computations of Eqs. 3.100 and 3.101, we can form the gradient of Eq. 3.99. Then, Eqs. 3.98 and 3.99 give the source terms needed in the right side of Eq. 3.57.

By solving the adjoint system Eqs. 3.56 and 3.57, we calculate the adjoint variables needed to find the sensitivity of $p^r_{wf,j}$ to gridblock permeabilities and porosities. Again, if one wishes to compute the sensitivity of $p^r_{wf,j}$ to the rock property

fields for several values of j (i.e., we have multiple wells at which we collect pressure data) and/or several values of r (data collected at several times), then several sets of adjoint variables will be generated by solving the adjoint system with multiple right-hand sides.

One important comment is in order. After solving for the adjoint variables related to the sensitivity coefficients for $p^r_{wf,j}$, the sensitivity coefficients must be computed from Eqs. 3.66–3.69. With the choice $g = p^r_{wf,j}$, the gradients of g appearing on the right sides of these equations are not all zero vectors. For example, in Eq. 3.66,

$$\nabla_{k_x} g = \nabla_{k_x} p^r_{wf,j} = \begin{bmatrix} \frac{\partial p^r_{wf,j}}{\partial k_{x,1}} \\ \frac{\partial p^r_{wf,j}}{\partial k_{x,2}} \\ \vdots \\ \frac{\partial p^r_{wf,j}}{\partial k^r_{x,M}} \end{bmatrix}, \tag{3.102}$$

where from Eq. 3.88, it follows that

$$\frac{\partial p^r_{wf,j}}{\partial k_{x,i}} = 0, \tag{3.103}$$

for $i \neq j$, but

$$\frac{\partial p^r_{wf,j}}{\partial k_{x,j}} = \frac{q_{t,j}}{\beta^r_j \, WI^2_j} \frac{\partial WI_j}{\partial k_{x,j}}. \tag{3.104}$$

Similarly,

$$\frac{\partial p^r_{wf,j}}{\partial k_{y,i}} = 0, \tag{3.105}$$

for $i \neq j$, but

$$\frac{\partial p^r_{wf,j}}{\partial k_{y,j}} = \frac{q_{t,j}}{\beta^r_j \, WI^2_j} \frac{\partial WI_j}{\partial k_{y,j}}. \tag{3.106}$$

Formulas for the derivatives of WI_j appearing in Eqs. 3.104 and 3.106 are given by

$$\frac{\partial WI_j}{\partial k_{x,j}} = \Big(\frac{C_3 h}{2[\ln(r_{o,j}/r_{w,j}) + s_j]} \Big) \times \left[\sqrt{\frac{k_{y,j}}{k_{x,j}}} - \Big(\frac{1}{\ln(r_{o,j}/r_{w,j}) + s_j} \Big) \Big(\frac{\Delta y_j^2 \sqrt{k_{x,j} k_{y,j}}}{\Delta x_j^2 k_{y,j} + \Delta y_j^2 k_{x,j}} - \frac{\sqrt{k_{y,j}}}{\sqrt{k_{x,j}} + \sqrt{k_{y,j}}} \Big) \right], \tag{3.107}$$

and

$$\frac{\partial WI_j}{\partial k_{y,j}} = \Big(\frac{C_3 h}{2[\ln(r_{o,j}/r_{w,j}) + s_j]}\Big) \times$$
$$\left[\sqrt{\frac{k_{x,j}}{k_{y,j}}} + \Big(\frac{1}{\ln(r_{o,j}/r_{w,j}) + s_j}\Big)\Big(\frac{\Delta y_j^2 k_{x,j}\sqrt{k_{x,j}}}{\Delta x_j^2 k_{y,j}\sqrt{k_{y,j}} + \Delta y_j^2 k_{x,j}\sqrt{k_{y,j}}} - \frac{k_{x,j}}{k_{y,j} + \sqrt{k_{x,j}k_{y,j}}}\Big)\right]. \tag{3.108}$$

3.3.3 Sensitivity of Water-Oil Ratio to Rock Property Fields

For the two-dimensional problem under consideration, flow rates at a producing well at time t^r are computed from

$$q_m = WI_j\Big(\frac{b^r_{m,j}k^r_{rm,j}}{\mu_m}\Big)\big(p^r_j - p^r_{wf,j}\big), \tag{3.109}$$

for $m = 0, w$. Thus the WOR at well j at time t^r is given by

$$WOR^r_j = \Big(\frac{b^r_{w,j}k^r_{rw,j}}{\mu_w}\Big)\Big(\frac{\mu_o}{b^r_{o,j}k^r_{ro,j}}\Big). \tag{3.110}$$

To obtain the sensitivity of $g = WOR^r_j$ to the rock property fields, we solve the adjoint equations using source terms derived from the choice

$$g = WOR^r_j. \tag{3.111}$$

From Eqs. 3.56 and 3.57, it is clear that the gradients $\nabla_{p^l} WOR^r_j$ and $\nabla_{S^l_w} WOR^r_j$ are needed.

We consider the gradient with respect to the vector of gridblock pressure first. Note that for $l \neq r$,

$$\nabla_{p^l} g = \nabla_{p^l} WOR^r_j = 0, \tag{3.112}$$

and for $l = r$,

$$\nabla_{p^r} g = \nabla_{p^r} WOR^r_j = \begin{bmatrix} \frac{\partial WOR^r_j}{\partial p^r_1} \\ \frac{\partial WOR^r_j}{\partial p^r_2} \\ \vdots \\ \frac{\partial WOR^r_j}{\partial p^l_M} \end{bmatrix}. \tag{3.113}$$

However from Eq. 3.110, it follows that

$$\frac{\partial WOR_j^r}{\partial p_i^r} = 0, \tag{3.114}$$

for $i \neq j$, but

$$\frac{\partial WOR_j^r}{\partial p_j^r} = \Big(\frac{k_{rw,j}^r \mu_o}{k_{ro,j}^r \mu_w}\Big)\Big(\frac{1}{b_{o,j}^r}\frac{\partial b_{w,j}^r}{\partial p_j^r} - \frac{b_{w,j}^r}{(b_{o,j}^r)^2}\frac{\partial b_{o,j}^r}{\partial p_j^r}\Big) = \Big(\frac{k_{rw,j}^r \mu_o}{k_{ro,j}^r \mu_w}\Big)\Big(\frac{c_w b_{w,in}}{b_{o,j}^r} - \frac{c_o b_{o,in} b_{w,j}^r}{(b_{o,j}^r)^2}\Big). \tag{3.115}$$

Eqs. 3.114 and 3.115 provide the information needed to compute the gradient of Eq. 3.113. Eqs. 3.112 and 3.113 provide the source terms needed on the right-side of Eq. 3.56.

Next, we consider the computation of $\nabla_{S_w^l} WOR_j^r$. First note that

$$\nabla_{S_w^l} g = \nabla_{S_w^l} WOR_j^r = 0, \tag{3.116}$$

for $l \neq r$. For $l = r$, we need

$$\nabla_{S_w^r} g = \nabla_{S_w^r} WOR_j^r = \begin{bmatrix} \frac{\partial WOR_j^r}{\partial S_{w,1}^r} \\ \frac{\partial WOR_j^r}{\partial S_{w,2}^r} \\ \vdots \\ \frac{\partial WOR_j^r}{\partial S_{w,M}^l} \end{bmatrix}. \tag{3.117}$$

From Eq. 3.110, it follows that

$$\frac{\partial WOR_j^r}{\partial S_{w,i}^r} = 0, \tag{3.118}$$

for $i \neq j$ and

$$\frac{\partial WOR_j^r}{\partial S_{wj}^r} = \frac{b_{w,j}^r \mu_o}{b_{o,j}^r \mu_w}\Big(\frac{1}{k_{ro,j}^r}\frac{dk_{rw}(S_{w,j})}{dS_w} - \frac{k_{rw,j}^r}{(k_{ro,j}^r)^2}\frac{dk_{ro}(S_{w,j})}{dS_w}\Big). \tag{3.119}$$

Eqs. 3.118 and 3.119 provide the information needed to compute the gradient of Eq. 3.117 and then Eqs. 3.116 and 3.117 provide the source terms needed on the right-side of Eq. 3.57. Note with this choice of g, the gradients of g with respect to the rock property fields that appear on the right sides of Eqs. 3.66–3.69 are all zero.

3.4 Formulation of the Adjoint System

In this section, we provide more detailed information on the formulation of the discrete system of adjoint equations. Recall that the adjoint system of equations is given by Eqs. 3.56 and 3.57 which are repeated here as

$$\sum_{m=o,w} \left[\left(\nabla_{p^l} [F_m^l - A_m^l]^T \right) \lambda_m^l + \nabla_{p^l} [\tilde{A}_m^l]^T \lambda_m^{l+1} \right] = -\nabla_{p^l} g, \tag{3.120}$$

and

$$\sum_{m=o,w} \left[\nabla_{S_w^l} [F_m^l - A_m^l]^T \lambda_m^l + \nabla_{S_w^l} [\tilde{A}_m^l]^T \lambda_m^{l+1} \right] = -\nabla_{S_w^l} g, \tag{3.121}$$

where (see Eqs. 3.38–3.40

$$F_m^{l+1} = F_m^{l+1}(p^{l+1}, S_w^{l+1}, k_x, k_y, k_z) = f_m^{l+1} - q_m^{l+1}, \tag{3.122}$$

$$A_m^{l+1} = A_m^{l+1}(p^{l+1}, S_w^{l+1}, \phi) = V_\phi^{l+1} (bS)_m^{l+1}, \tag{3.123}$$

and

$$\tilde{A}^l = \tilde{A}_m^l(p^l, S_w^l, \phi) = V_\phi^{l+1} (bS)_m^l, \tag{3.124}$$

for $m = 0, w$ and $l = 0, 1, ..., L-1$. Also recall from Eqs. 3.4 and 3.6 that

$$V_\phi^l = \begin{bmatrix} V_{\phi,1}^l & 0 & 0 & \dots & 0 \\ 0 & V_{\phi,2}^l & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & \vdots \\ 0 & 0 & 0 & \dots & V_{\phi,M}^l \end{bmatrix}, \tag{3.125}$$

where $V_{\phi,j}^l$ denotes the pore volume of gridblock j in ft^3 divided by $C_2 \Delta t^l = 5.615 \Delta t^l$, i.e.

$$V_{\phi,j}^l = \frac{V_j \phi_j}{C_2 \Delta t^l}. \tag{3.126}$$

To write individual equations in the adjoint system, we will use basic operational formulas like the following one:

$$\nabla_{p^l}[f_m^T] = \begin{bmatrix} \frac{\partial f^l_{m,1}}{\partial p^l_1} & \frac{\partial f^l_{m,2}}{\partial p^l_1} & \cdots & \frac{\partial f^l_{m,M}}{\partial p^l_1} \\ \frac{\partial f^l_{m,1}}{\partial p^l_2} & \frac{\partial f^l_{m,2}}{\partial p^l_2} & \cdots & \frac{\partial f^l_{m,M}}{\partial p^l_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^l_{m,1}}{\partial p^l_M} & \frac{\partial f^l_{m,2}}{\partial p^l_M} & \cdots & \frac{\partial f^l_{m,M}}{\partial p^l_M} \end{bmatrix}. \tag{3.127}$$

Eq. 3.127 is general in the sense that we can replace f_m by any M-dimensional column vector and take its gradient with respect to S^l_w, k_x or any other vector variable.

We now rewrite Eq. 3.120 as

$$\begin{aligned} &\big(\nabla_{p^l}[F^l_o]^T\big)\lambda^l_o + \big(\nabla_{p^l}[F^l_w]^T\big)\lambda^l_w = \\ &\big(\nabla_{p^l}[A^l_o]^T\big)\lambda^l_o - \big(\nabla_{p^l}[\tilde{A}^l_o]^T\big)\lambda^{l+1}_o + \big(\nabla_{p^l}[A^l_w]^T\big)\lambda^l_w - \big(\nabla_{p^l}[\tilde{A}^l_w]^T\big)\lambda^{l+1}_w - \nabla_{p^l} g. \end{aligned} \tag{3.128}$$

Using the basic operational formula of Eq. 3.127 and Eq. 3.122, we can write the rth equation of Eq. 3.128 as

$$\begin{aligned} \sum_{s=1}^{M}\Big[\Big(\frac{\partial f^l_{o,s}}{\partial p^l_r} - \frac{\partial q^l_{o,s}}{\partial p^l_r}\Big)\lambda^l_{o,s}\Big] + \sum_{s=1}^{M}\Big[\Big(\frac{\partial f^l_{w,s}}{\partial p^l_r} - \frac{\partial q^l_{w,s}}{\partial p^l_r}\Big)\lambda^l_{w,s}\Big] = \\ \sum_{s=1}^{M}\Big[\Big(\frac{\partial A^l_{o,s}}{\partial p^l_r}\Big)\lambda^l_{o,s}\Big] + \sum_{s=1}^{M}\Big[\Big(\frac{\partial A^l_{w,k}}{\partial p^l_r}\Big)\lambda^l_{w,s}\Big] - \\ \sum_{s=1}^{M}\Big[\Big(\frac{\partial \tilde{A}^l_{o,s}}{\partial p^l_r}\Big)\lambda^{l+1}_{o,s}\Big] - \sum_{s=1}^{M}\Big[\Big(\frac{\partial \tilde{A}^l_{w,s}}{\partial p^l_r}\Big)\lambda^{l+1}_{w,s}\Big] - \frac{\partial g}{\partial p^l_r}. \end{aligned} \tag{3.129}$$

We now assume that r corresponds to gridblock (i,j,k) so the only $f^l_{m,s}$'s in the sums of Eq. 3.129 that depend on $p^l_r = p^l_{i,j,k}$ are the ones corresponding to $s = (i-1,j,k)$, $s = (i,j,k)$, $s = (i+1,j,k)$, $s = (i,j-1,k)$, $s = (i,j+1,k)$, $s = (i,j,k-1)$ and $s = (i,j,k+1)$. In Eq. 3.129, the derivatives of the flow rate terms, the $A_{m,s}$'s and the $\tilde{A}_{m,s}$'s are zero unless $m = r$; see, Eqs. 3.123–3.126. Thus,

in three-dimensional notation, Eq. 3.129 is equivalent to

$$\begin{aligned}
&\Big(\frac{\partial f^l_{o,i-1,j,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i-1,j,k}+\Big(\frac{\partial f^l_{o,i+1,j,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i+1,j,k}+\Big(\frac{\partial f^l_{o,i,j-1,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i,j-1,k}+\Big(\frac{\partial f^l_{o,i,j+1,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i,j+1,k}+\\
&\Big(\frac{\partial f^l_{o,i,j,k-1}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i,j,k-1}+\Big(\frac{\partial f^l_{o,i,j,k+1}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i,j,k+1}+\Big(\frac{\partial f^l_{o,i,j,k}}{\partial p_{i,j,k}}-\frac{\partial q^l_{o,i,j,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{o,i,j,k}+\\
&\Big(\frac{\partial f^l_{w,i-1,j,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i-1,j,k}+\Big(\frac{\partial f^l_{w,i+1,j,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i+1,j,k}+\\
&\Big(\frac{\partial f^l_{w,i,j-1,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i,j-1,k}+\Big(\frac{\partial f^l_{w,i,j+1,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i,j+1,k}+\\
&\Big(\frac{\partial f^l_{w,i,j,k-1}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i,j,k-1}+\Big(\frac{\partial f^l_{w,i,j,k+1}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i,j,k+1}+\Big(\frac{\partial f^l_{w,i,j,k}}{\partial p_{i,j,k}}-\frac{\partial q^l_{w,i,j,k}}{\partial p_{i,j,k}}\Big)\lambda^l_{w,i,j,k}=\\
&\Big(\frac{\partial A^l_{o,i,j,k}}{\partial p^l_{i,j,k}}\Big)\lambda^l_{o,i,j,k}-\Big(\frac{\partial \tilde{A}^l_{o,i,j,k}}{\partial p^l_{i,j,k}}\Big)\lambda^{l+1}_{o,i,j,k}+\Big(\frac{\partial A^l_{w,i,j,k}}{\partial p^l_{i,j,k}}\Big)\lambda^l_{w,i,j,k}-\Big(\frac{\partial \tilde{A}^l_{w,i,j,k}}{\partial p^l_{i,j,k}}\Big)\lambda_{w,i,j,k}-\frac{\partial g}{\partial p^l_{i,j,k}}=\\
&-\Big(\frac{V_{i,j,k}\phi_{i,j,k}}{5.615\Delta t^l}\Big)\Big[\Big(S^l_{o,i,j,k}\frac{\partial b^l_{o,i,j,k}}{\partial p^l_{i,j,k}}\Big)\Big(\lambda^{l+1}_{o,i,j,k}-\lambda^l_{o,i,j,k}\Big)+\Big(S^l_{w,i,j,k}\frac{\partial b^l_{w,i,j,k}}{\partial p^l_{i,j,k}}\Big)\Big(\lambda^{l+1}_{w,i,j,k}-\lambda^l_{w,i,j,k}\Big)\Big]\\
&-\frac{\partial g}{\partial p^l_{i,j,l}} \qquad (3.130)
\end{aligned}$$

where the last equality in Eq. 3.130 follows from Eqs. 3.123–3.126. To complete the system of equations, we simple need to compute the partial derivatives in Eq. 3.130. The derivative of the f_m terms can be easily computed from the relation

$$\begin{aligned}
f_{m,i,j,k} = {} & T^l_{mx,i+1/2,j,k}(p^l_{i+1,j,k}-p^l_{i,j,k})-T^l_{mx,i-1/2,j,k}(p^l_{i,j,k}-p^l_{i-1,j,k})+\\
& T^l_{my,i,j+1/2,k}(p^l_{i,j+1,k}-p^l_{i,j,k})-T^l_{my,i,j-1/2,k}(p^l_{i,j,k}-p^l_{i,j-1,k})+\\
& T^l_{mz,i,j,k+1/2}(p^l_{i,j,k+1}-p^l_{i,j,k})-T^l_{mz,i,j,k-1/2}(p^l_{i,j,k}-p^l_{i,j,k-1}). \qquad (3.131)
\end{aligned}$$

Computation of derivative of the inverse formation volume factor terms is straightforward; these terms are modeled using Eq. 3.94. The derivative of the flow rate terms depends on how these sink terms are incorporated in the simulator. For the two-dimensional results presented in this work, the total flow rate at production wells is specified and then

$$q^l_{m,i,j,k}=\Big(\frac{q^l_{m,i,j,k}}{q^l_{t,i,j,k}}\Big)q^l_{t,i,j,k}=\Big(\frac{(b^l_{m,i,j,k}k^l_{rm,i,j,k})/\mu_m}{(b^l_{o,i,j,k}k^l_{ro,i,j,k}/\mu_o)+(b^l_{w,i,j,k}k^l_{rw,i,j,k}/\mu_w)}\Big)q^l_{t,i,j,k}, \qquad (3.132)$$

for $m = o, w$. As discussed previously, the derivative of g that appears in Eq. 3.130 depends on our choice of the source generating function.

A typical equation of the system of Eq. 3.121 can be found using the same procedure used to obtain Eq. 3.130.

3.5 Example Sensitivities

For all problems considered in this work, a 25 by 25 grid is used as shown in Fig. 3.1, where the i and j indices, respectively, correspond to gridblocks in the x and y directions. The areal dimensions of the reservoir are $1,100$ ft by $1,100$ ft and we assume a uniform thickness with $h = 30$ ft. In all cases, the 25 by 25 simulation grid is the same as the model parameter grid meaning that we calculate the sensitivity to the permeability and porosity of every grid block.

Sensitivities are more difficult to interpret physically when the reservoir properties are nonuniform and the number of wells is large, so here we consider a reservoir model which has only two wells and uniform isotropic permeability ($k = 40.4$ md) and uniform porosity ($\phi = 0.25$). Oil and water relative permeability curves are shown in Fig. 3.2. Other relevant rock and fluid properties are given in Table 3.1.

An injector is located at gridblock (18,18) and a producer at gridblock (5,5). The producer is constrained to produce at a constant total rate of 1250 STB/day while the injector is constrained to inject at a constant rate of 1600 STB/day. Because of the differences in the formation volume factors for water and oil, the injection and producing rate are both roughly equal to 1,600 RB/D prior to breakthrough and during this time period, average reservoir pressure remains nearly constant. The producing water-oil-ratio (WOR) is shown in Fig. 3.4.

As there is some smearing of the front in the reservoir simulation results, one does not observe a perceptible increase in WOR until $t \approx 1,780$ days. However, examination of the simulation results indicates there water saturation begins to in-

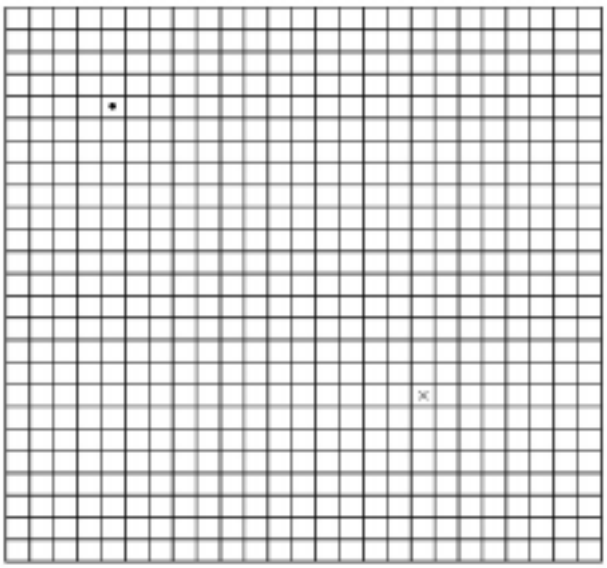

Figure 3.1: Grid System and well locations. The • denotes the location of the injection well and the × denotes the location of the producing well.

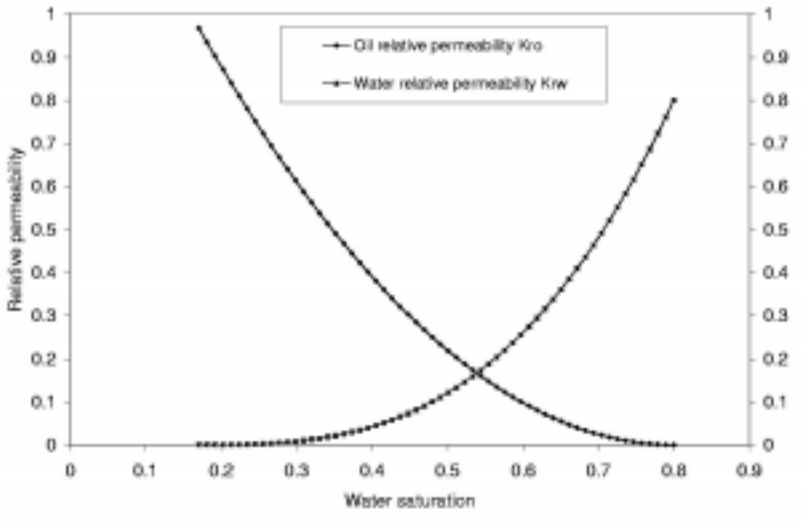

Figure 3.2: Oil and water relative permeability curves.

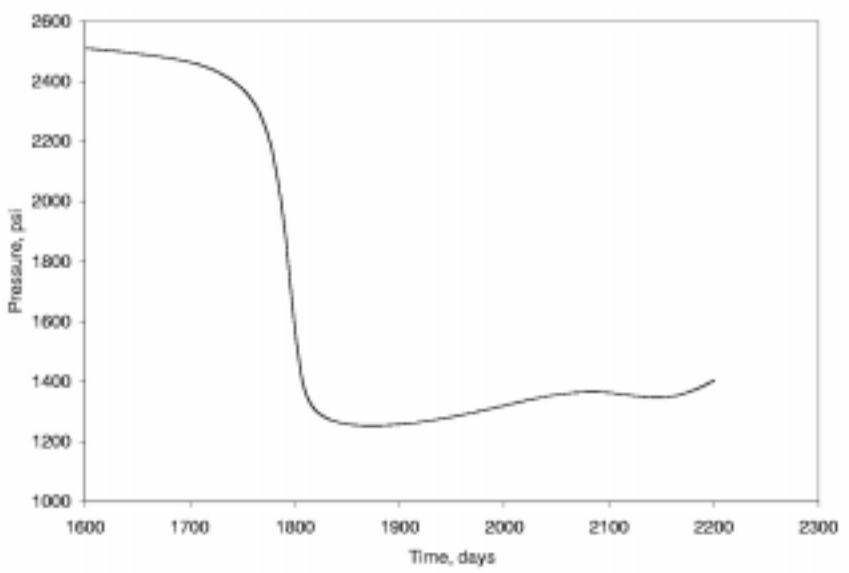

Figure 3.3: Flowing bottom-hole pressure at the producing well.

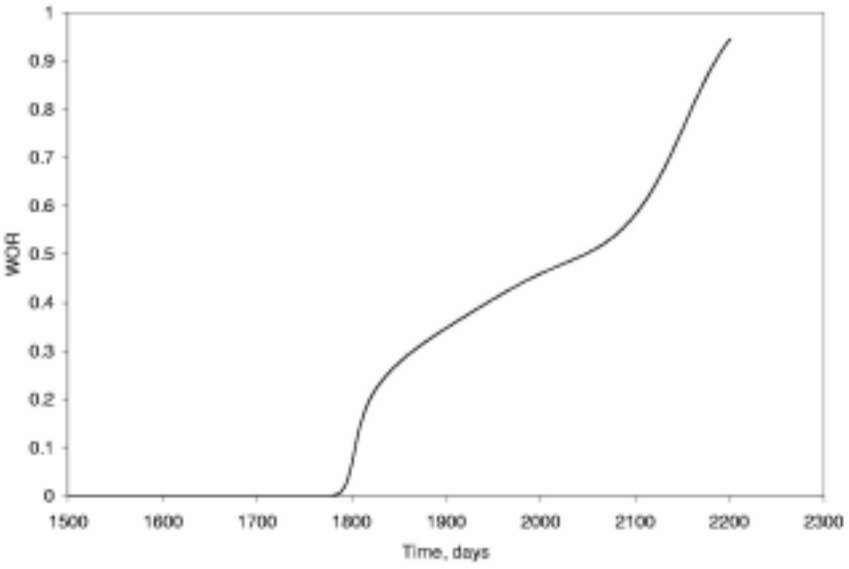

Figure 3.4: Water-oil-ratio at the production well.

Porosity	0.25
Permeability, md	40.4
Oil viscosity, cp	0.8
Water viscosity, cp	1.0
Initial water saturation	0.16
Residual oil saturation	0.20
Water compressibility, psi^{-1}	4×10^{-6}
Oil compressibility, psi^{-1}	1.0×10^{-5}
Initial pressure, psi	4,000
Oil formation volume factor at p_i	1.269
Water formation volume factor at p_i	1.004

Table 3.1: Rock and fluid properties for calculation of sensitivity coefficients.

crease very gradually in the producing block at $t \approx 1,750$ days. This small increase in water saturation causes a significant decrease in the total mobility. Thus, as the producing well is constrained to produce at a constant total rate, a rapid decrease in the total mobility in the grid block containing the producing well results in a rapid decrease in the the flowing wellbore pressure at the producer as shown in Fig. 3.3.

3.5.1 Sensitivity of p_{wf} to $\ln k$ and ϕ at 20 Days

At early times, most of the reservoir is still single phase. The pressure at the producer is highly sensitive to the permeability of the grid-block containing the producing wells and very little else. At all times, the greatest sensitivity of pressure to permeability occurs at the site of the active well as we should expect. The sensitivity to porosity at early time is somewhat less intuitive. What we see in Fig. 3.5 is largely a storage effect, that is an increase in porosity near the injector causes less pressure

support at the producer, while an increase in porosity near the producer causes a reduction in the drawdown.

3.5.2 Sensitivity of p_{wf} to $\ln k$ and ϕ at 1800 Days

At the time of break through of water to the producing well, the effects of fluid mobility are easily seen in the plot of sensitivity of p_{wf} to porosity and permeability (Figure 3.6). The total mobility has decreased in the swept zones and is dropping rapidly in the grid block containing the producing well, causing the pressure drop at the producer to increase. An increase in the permeability between the wells, would cause the water to advance faster and the pressure to decrease more rapidly. That explains the negative values of sensitivity to permeability in the interwell region. On the other hand, an increase in porosity in the interwell region would delay the advance of water, resulting in a higher total mobility and higher pressure at the producer. Thus the sensitivity of p_{wf} to porosity is positive in the interwell region.

3.5.3 Sensitivity of p_{wf} to $\ln k$ and ϕ at 1864 days

When the water-oil ratio has reached 0.3 at 1864 days, the pressure at the producer is very near to a minimum value. As a result, the effect of small changes in any time-dependent variable, such as mobility or fluid compressibility, can have only a small effect on the pressure at the producing well. Note, for example, that the magnitude of the sensitivity to porosity in Figure 3.7 is everywhere less than 50 while the maximum magnitude at t = 1800 days was over 2000.

3.5.4 Sensitivity of p_{wf} to $\ln k$ and ϕ at 2136 Days

When the water-oil ratio reaches 0.7 at 2136 days, similar effects can be seen in Fig. 3.8. The pressure is again nearly stationary although it is close to a local maximum instead of a minimum. The small increase in pressure that occurs between

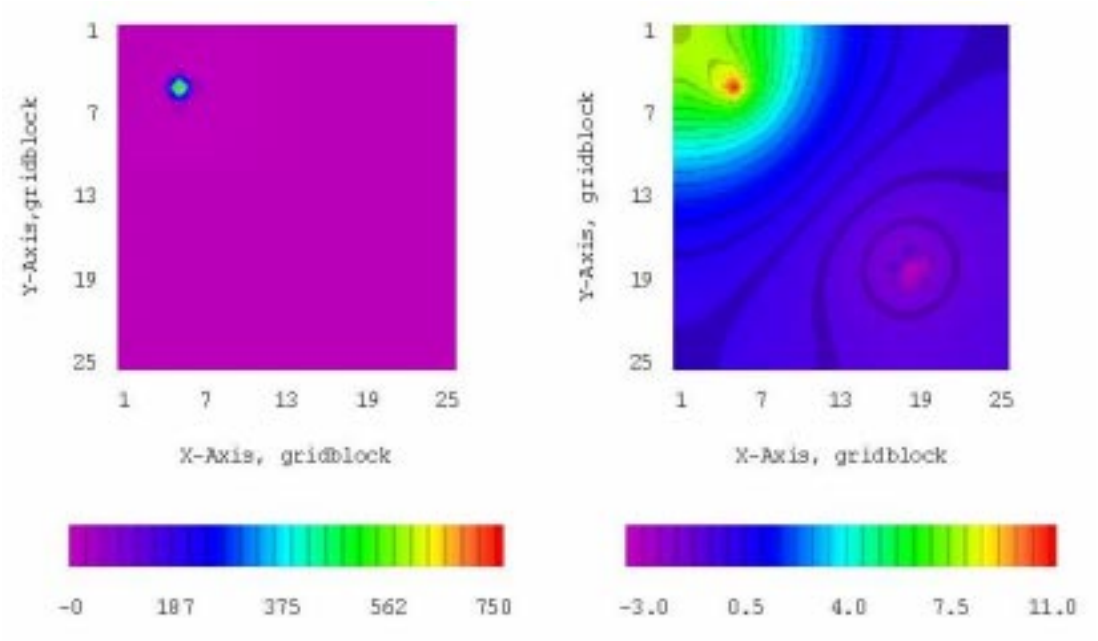

Figure 3.5: Sensitivity of bottom-hole flowing pressure to log-permeability (left) and porosity (right) at 20 days.

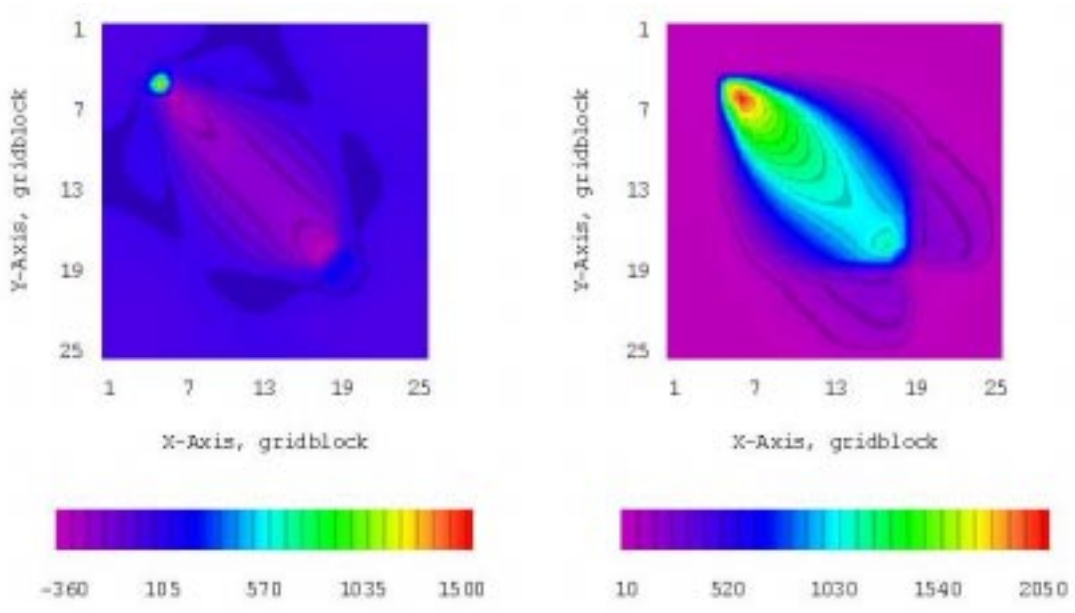

Figure 3.6: Sensitivity of bottom-hole flowing pressure to log-permeability (left) and porosity (right) at 1800 days.

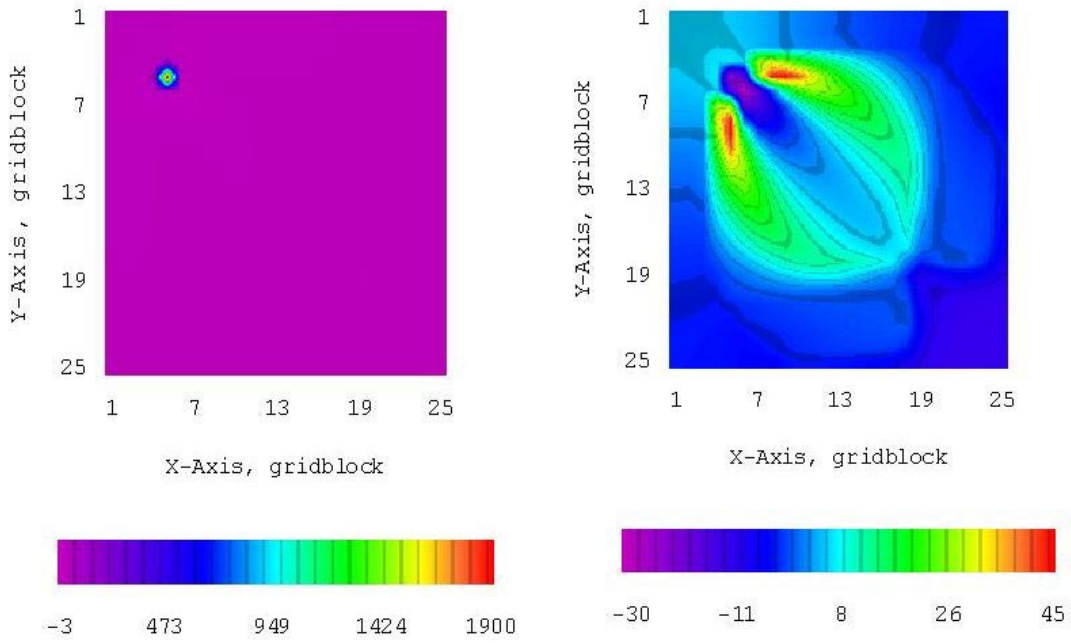

Figure 3.7: Sensitivity of bottom-hole flowing pressure to log-permeability (left) and porosity (right) at 1864 days.

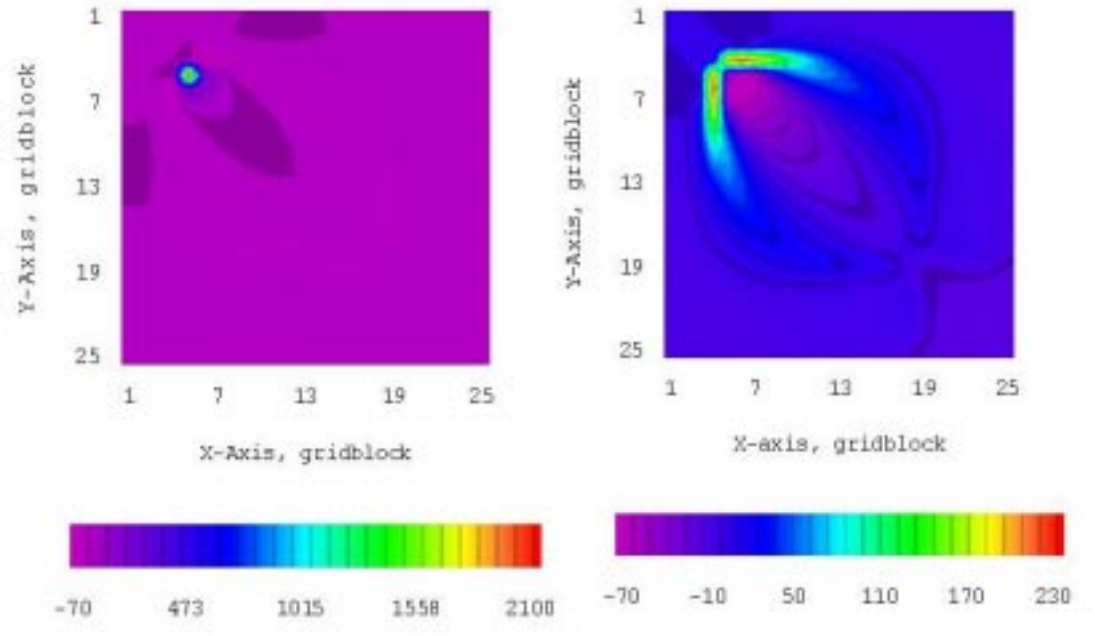

Figure 3.8: Sensitivity of bottom-hole pressure to log-permeability (left) and porosity (right) when the WOR reaches 0.7.

1864 and 2136 days is apparently due to an increase in mobility of fluids in the grid block containing the producer as the gridblock saturation approaches the residual oil saturation. Again, however, the magnitudes of sensitivities are relatively small except for the sensitivity of permeability to gridblock permeability which is not a dynamic effect. The sensitivity of p_{wf} to porosity is negative between the wells because an increase in porosity between the wells would decrease the saturation of the gridblock with the producing well. An increase in porosity at the leading edge of the saturation front causes a decrease in saturation of gridblocks "beyond" the producer. For low water saturation, a decrease in saturation results in an increase in mobility, and a positive sensitivity.

3.5.5 Sensitivity of WOR to $\ln k$ and ϕ at 1800 Days

The sensitivity of water-oil ratio with respect to porosity is relatively straight forward. The only significant effect that a change in porosity has on water-oil ratio is through a change in the velocity of propagation of saturations along streamlines. An increase in porosity will cause the water to be delayed in arriving at the producer. The effect on water-oil ratio is largest when the change in porosity occurs along a streamline which has a rapidly changing saturation at the producer. Thus, in Fig. 3.9, the largest sensitivity is directly between the injector and producer.

Qualitatively, the sensitivity of water-oil ratio with respect to permeability is similar to the sensitivity of water-oil ratio with respect to porosity, except that the sign of the sensitivity is reversed. This is because an increase in permeability causes the velocity of the flow to increase and the saturation at the well to increase.

3.5.6 Sensitivity of WOR to $\ln k$ and ϕ at 1864 Days

When the water-oil ratio reaches 0.3, more streamlines contribute water to the producing well. The boundary of non-zero sensitivities in Fig. 3.10 is a direct

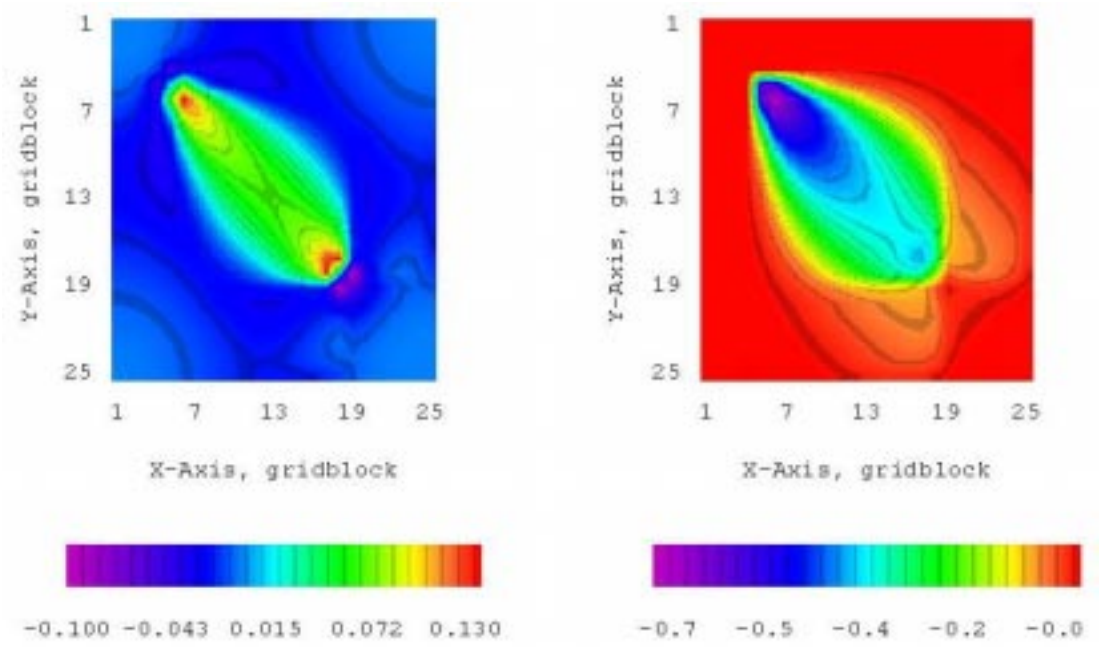

Figure 3.9: Sensitivity of producing water-oil ratio to log-permeability (left) and porosity (right) soon after breakthrough (1800 days).

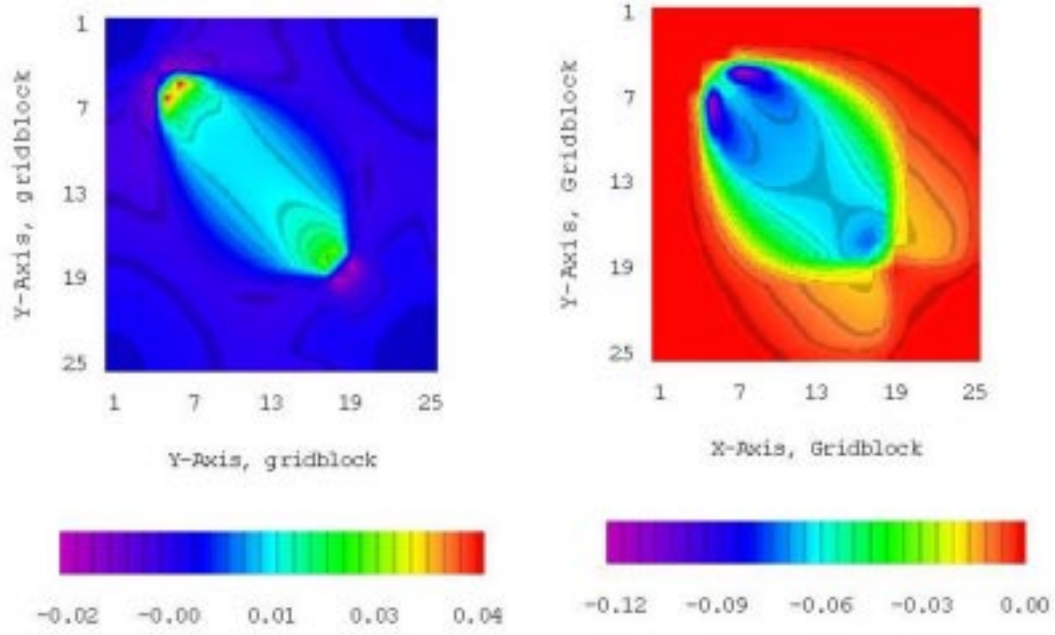

Figure 3.10: Sensitivity of producing water-oil-ratio to log-permeability (left) and porosity (right) when the WOR reaches 0.3 at 1864 days.

indication of the region that has broken through to the producer, and this region is clearly larger that it was at the time of initial water breakthrough.

3.5.7 Sensitivity of WOR to $\ln k$ and ϕ at 2136 Days

At 2136 days, the region of significant sensitivity is wider (see Fig. 3.11) than it was at 1800 days (Figure 3.10). A region of zero sensitivity of water-oil ratio to porosity occurs between the wells at gridblock (7,7). The saturation in this gridblock is no longer changing with time, so modifying the porosity cannot affect the water-oil ratio at the producer.

3.5.8 Comparison With Finite-Difference Method

To verify that the procedure developed in Chapter III yields accurate sensitivity coefficients, we have made several comparisons with sensitivity coefficients calculated with the finite-difference method (sometimes referred to as the direct method). In all cases, we obtained essentially identical results, less than one per cent difference. In the finite-difference method, the permeability (or porosity) of each grid-block is altered by a small amount and the resulting change in wellbore pressure and water-oil ratio is calculated. The change in wellbore pressure (or WOR) dividing by the change in the grid block value of permeability (or porosity) give the value of the relevant sensitivity coefficient. Although it is possible for sensitivities from the finite-difference approximation to be in error if the magnitude of the perturbation is poorly chosen, it is reasonable to suppose that agreement between sensitivities calculated using two completely different methods indicates that both methods are giving correct results. Fig. 3.12 shows a comparison of the sensitivity of the water-oil-ratio to the porosity field at $1,800$ days obtained by the finite difference method and the adjoint method. Note that the results from the two methods are indistinguishable. It is important to also note that the finite difference method required 626 runs of

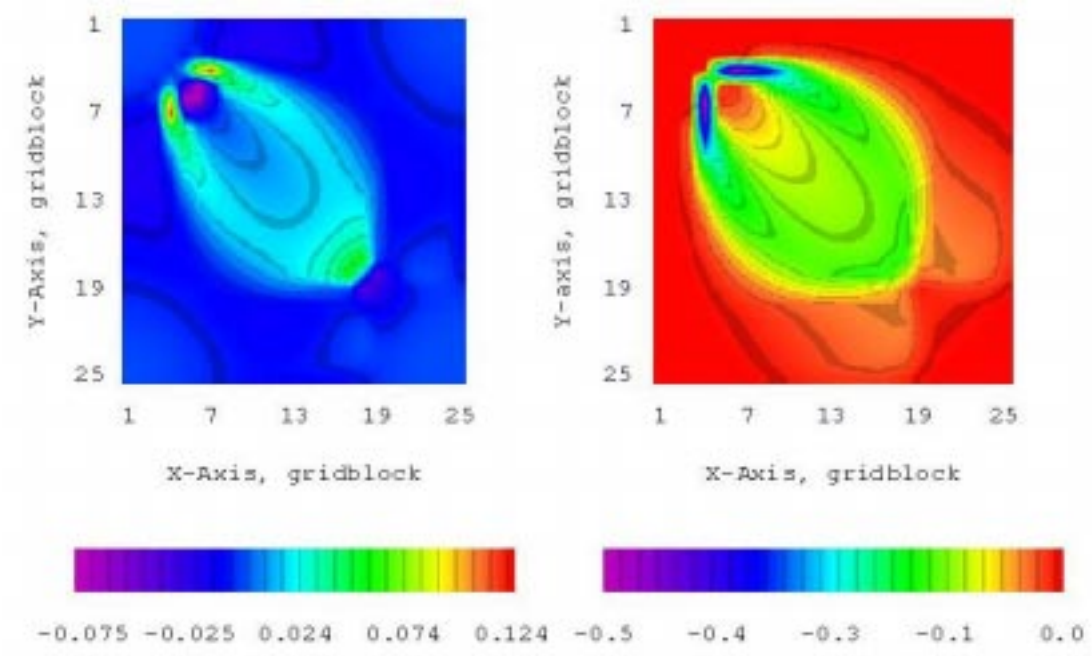

Figure 3.11: Sensitivity of producing water-oil-ratio to log-permeability (left) and porosity (right) when the WOR reaches 0.7.

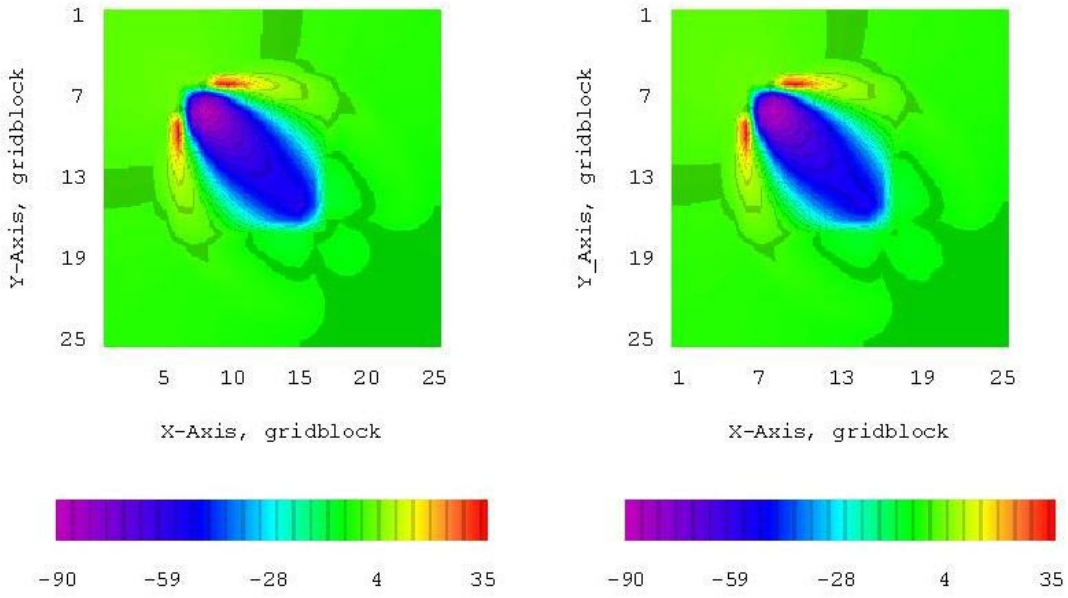

Figure 3.12: Comparison of the sensitivity of bottom-hole pressure at the producing well to porosity from the adjoint method (left) and from the "direct" method (right).

the reservoir simulator while the adjoint approach required approximately as much computer time as two simulation runs. Thus, it is clearly impractical to compute sensitivity coefficients using the the finite-difference approach.

CHAPTER IV

CONDITIONING ROCK PROPERTY FIELDS TO PRODUCTION DATA: COMPUTATIONAL RESULTS

The adjoint approach to calculation of sensitivities (Chapter III) was applied to the problem of estimating the maximum a posteriori (MAP) reservoir model consistent with measurements of water-oil ratio and with pressure data and to the problem of generating realizations that are conditional to the production data. Results for two example problems are presented in this chapter. For the second example, we also apply the procedure discussed earlier to evaluate the uncertainty in reservoir performance predictions.

The maximum a posteriori model is the model, which is the "most probable" (in the sense that it maximizes the a posteriori pdf), although it is generally too smooth to be a "plausible" model. The maximum a posteriori reservoir model should look something like a smoothed version of the true reservoir and is found by minimizing the objective function of Eq. 2.43 which is repeated here as

$$S(m) = \frac{1}{2}\left[(m - m_{prior})^T C_M^{-1}(m - m_{prior}) + (g(m) - d_{obs})^T C_D^{-1}(g(m) - d_{obs})\right] \tag{4.1}$$

where m is the vector of model parameters (gridblock porosities and log-permeabilities), d_{obs} is the vector of observed production data, and $g(m)$ is the vector of data that are calculated using m in the reservoir simulator. As discussed in Chapter II, the Gauss-Newton method is applied to minimize this objective function. For the examples considered, the number of unknowns is much less than the number of conditioning production data. Because of this, the form of the Gauss-Newton method used is given

by

$$\delta m^{l+1} = (m_{prior} - m_l) - C_M G_l{}^T \Big(C_D + G_l C_M G_l{}^T\Big)^{-1} \times \Big[g(m^l) - d_{obs} - G_l(m^l - m_{prior})\Big], \quad (4.2)$$

$$m^{l+1} = m^l - \mu_l \delta m^{l+1}, \quad (4.3)$$

In Eq. 4.2, G_l is the matrix of sensitivity coefficients calculated at the lth iteration where sensitivity coefficients are calculated by the adjoint method developed in Chapter III. Eq. 4.2 can be obtained by using Eq. 2.49 in Eq. 2.51 and then applying matrix inversion lemmas; see Tarantola [52].

As discussed in Chapter II, realizations that are conditional to production data are found by minimizing an objective function similar to the one of Eq. 4.1, in which the prior model is replaced by a unconditional realization generated from the prior model and the observed data are replaced by data to which random errors have been added.

In generating the MAP and in generating realizations, the data covariance matrix is modified by the procedure of Section 2.6. As noted previously, this results in more rapid convergence and reduces the likelihood of converging to a local minimum which gives an unacceptable match of production data.

The oil and water relative permeability curves used in both examples are given in Fig. 3.2.

4.1 Three-Zone Reservoir

The "true" model, from which the data are generated, is composed of three zones of constant permeability constructed on a 25×25 grid (Fig. 4.1). The log-permeability is 3.7 (40.44 mD) in the upper left quadrant, 4.3 (73.7 mD) in the upper

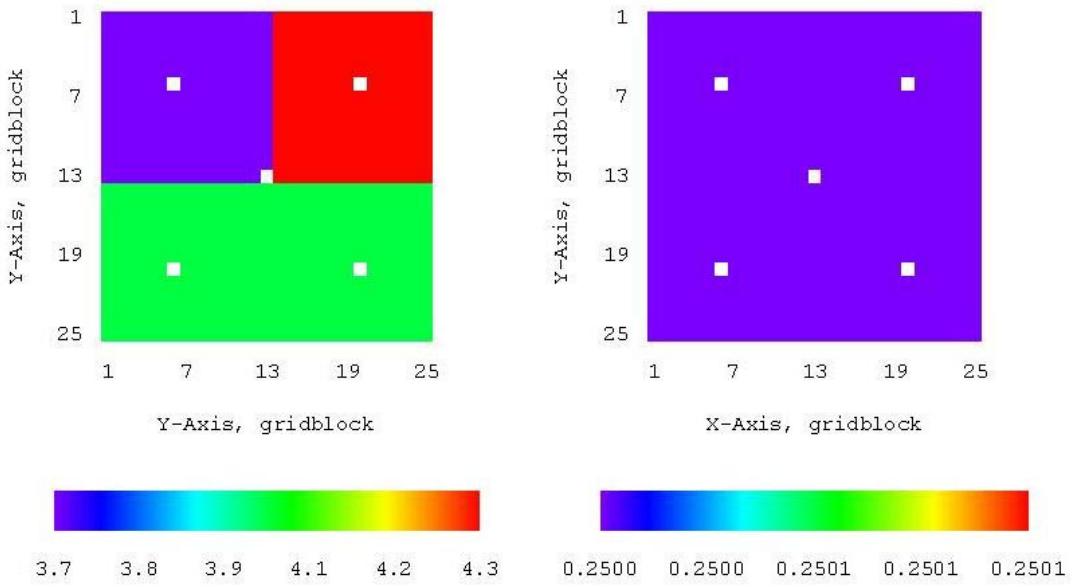

Figure 4.1: The true log-permeability (left) and porosity (right) with well locations for three-zone inverse problem.

right quadrant, and 4.0 (54.6 mD) in the lower two quadrants. The true porosity field is uniform with $\phi = 0.25$.

An injector is located in the center, and four producers are located near the corners. The square "data points" in Fig. 4.1 represent the well locations. Water breakthrough occurs at between 250 and 350 days at the various producers. The prior mean for this example is $\ln k = 4.0$ and $\phi = 0.25$ in every grid block. The prior variance for log-permeability is 0.5 and the prior variance for porosity is 0.0025. The prior spatial autocorrelation, used for calculating the maximum a posteriori estimate, is the same for both porosity and permeability: spherical with a range of 600 ft. We assumed no prior correlation between porosity and permeability. Based on the prior, the "most probable model" is uniform with all gridblock permeabilities and porosities equal to the prior means for these two parameters. It should be clear that if there were no data, or if the data are independent of the model parameters, then the maximum

a posteriori model would be the given by the prior means. Any features that appear in the maximum a posteriori estimate must be required by the data.

We explore the value of various types of data by calculating the maximum a posteriori estimate using three different sets of observations:

(i) measured water-oil ratio at the producers,

(ii) measured pressure at the producers,

(iii) both measured pressure and water-oil ratio at the producers.

All of the observed pressure and WOR data are from the period of time between 280 to 381 days. During this time period, breakthrough of water occurs at all of the producing wells.

When only the water-oil ratio data are used to calculate the best distribution of log-permeability the solution looks qualitatively correct (Fig. 4.2), i.e., the log-permeability is approximately correct at each of the producing well locations. When pressure data at the producing wells are used instead of water-oil ratio data, however, the maximum a posteriori estimate is somewhat better, in that the region of correct permeability is considerably larger than when only water-oil ratio data are used. The maximum a posteriori estimate generated from a combination of both types of data is not much different from the maximum a posteriori estimate generated from pressure data only. The primary difference is an improvement in the distribution of permeability on the left side near the center.

Figure 4.3 shows three maximum a posteriori estimates of the porosity distribution that were obtained using the three different combinations of data. All combinations give relatively uniform distributions, although the use of water-oil ratio data alone results in a maximum a posteriori solution with porosity that is somewhat high in the upper left quadrant and low in the upper right quadrant. This is because the sensitivity of water-oil ratio at a producing well to porosity has the opposite sign as the sensitivity to log-permeability; see the results of Section 3.5.

The uncertainty in the estimated porosity and log-permeability fields can be

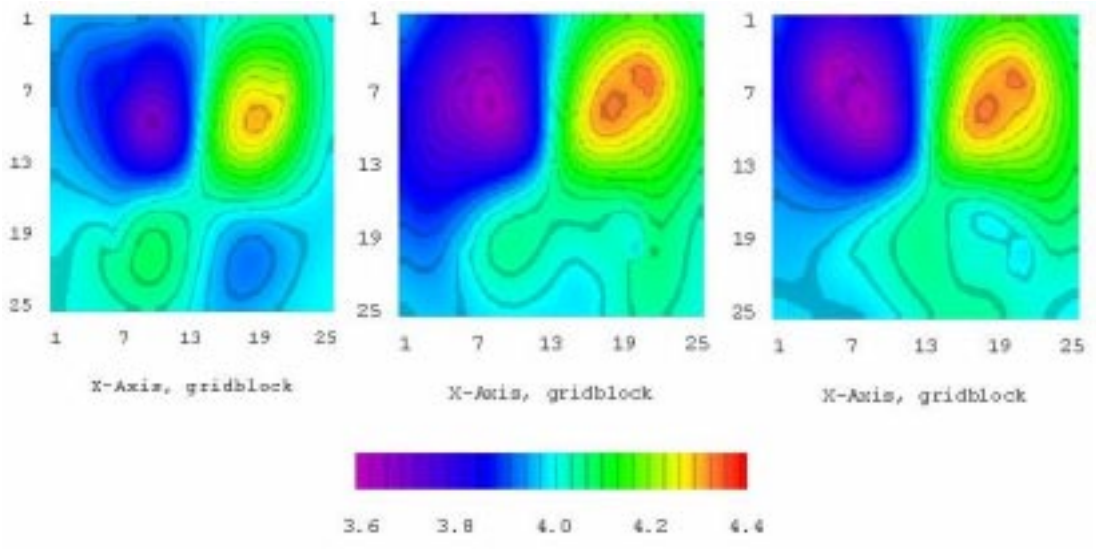

Figure 4.2: Maximum a posteriori estimate of log-permeability conditioned only to WOR (left), only to pressure (center), and to both WOR and pressure (right) for the three-zone inverse problem.

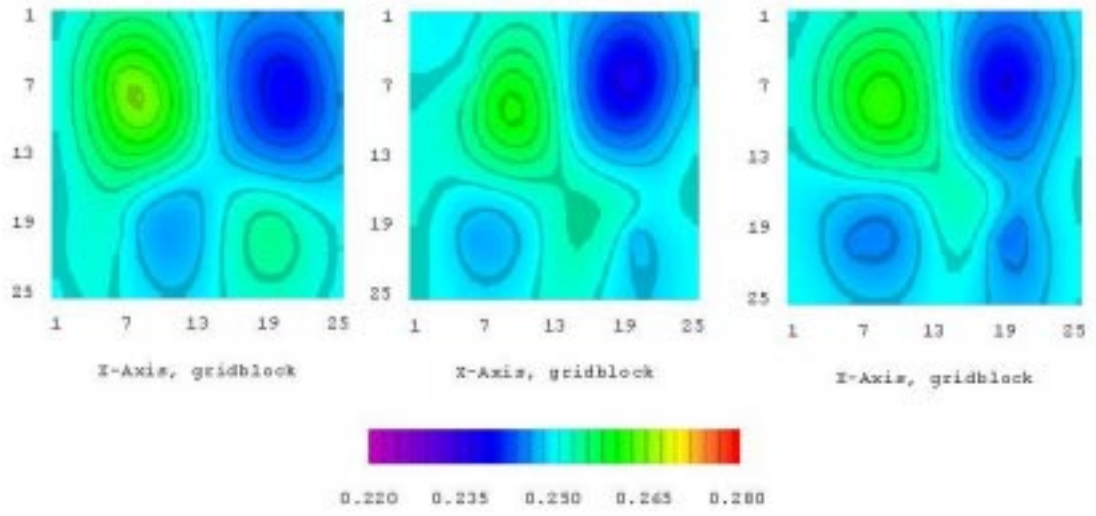

Figure 4.3: Maximum a posteriori estimate of porosity conditioned only to WOR (left), only to pressure (center), and to both WOR and pressure (right) for the three-zone inverse problem.

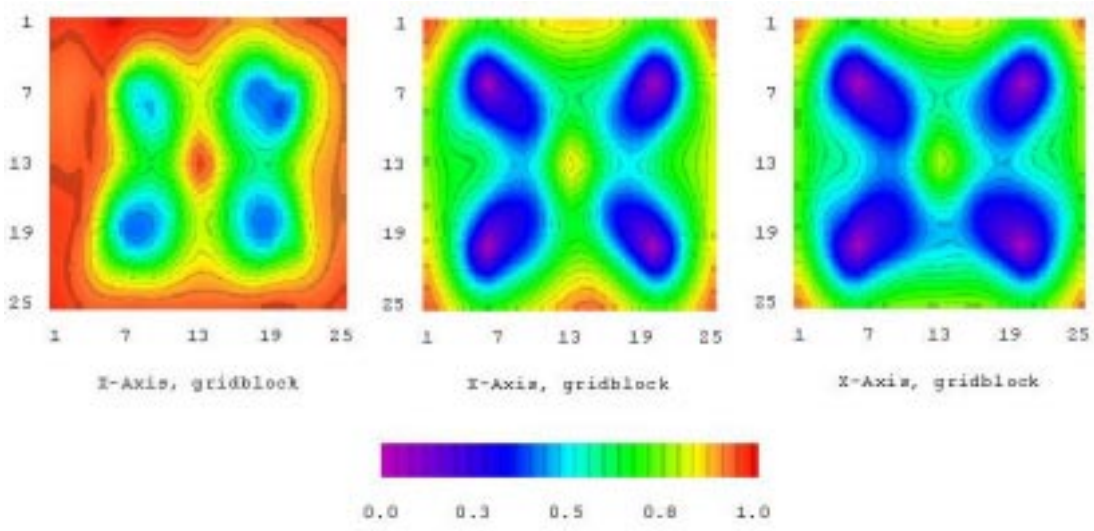

Figure 4.4: The normalized a posteriori variance of log-permeability conditioned only to WOR (left), only to pressure (center), and to both WOR and pressure (right).

characterized by the a posteriori covariance. The covariance is difficult to visualize, however, so we often examine only the variance. Figure 4.4 shows the normalized a posteriori variance for the maximum a posteriori estimates based on water-oil ratio data only, on pressure data only, and on both types of data. This variance is normalized by the prior variance so it is equal to 0.0 in areas that have no uncertainty after conditioning to data. The normalized variance is equal to 1.0 in regions at the reservoir whose variance is not reduced by the incorporation of data. Thus the best estimate of permeability in the regions near the perimeter of the reservoir are not improved by the incorporation of water-oil ratio data (left side of Fig. 4.4), the greatest reduction occurs when both water-oil ratio and pressure data are incorporated.

Figure 4.5 shows similar plots of normalized a posteriori variance of porosity after conditioning to production data. Again we see that the uncertainty in model parameters is greatest near the perimeter of the reservoir and that pressure data are better than water-oil ratio data for reducing the variance in the estimate.

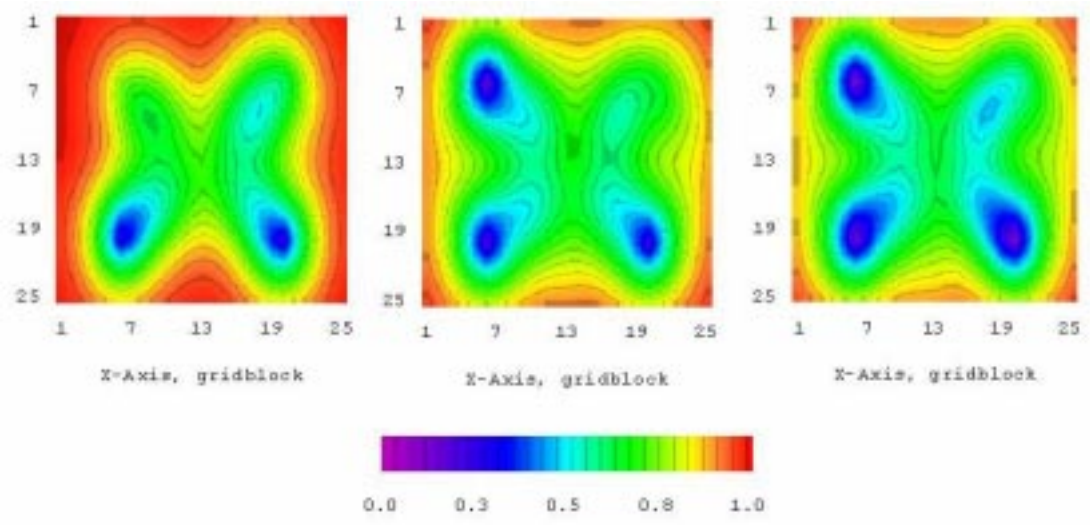

Figure 4.5: Normalized a posteriori variance of porosity conditioned only to WOR (left), only to pressure (center), and to both WOR and pressure (right).

In each case, whether conditioning to only one type of data or to both, convergence to the maximum a posteriori estimate was quite rapid. Figure 4.6 shows that only four iterations were required to match the water-oil ratio data by itself, and that only eight iterations were required to match the pressure data, or to match the combined set of water-oil ratio and pressure data. The iterations were halted when one of two conditions were met: 1) The value of the objective function reached a small fraction of the "expected value" based on the number of data, or 2) the objective function stopped decreasing.

The final matches to the data were very good in all cases. In Figs. 4.7 and 4.8, the agreement between data (shown by the discrete symbols) and the predicted pressures and water-oil ratios from the maximum a posteriori estimates can be seen to agree very well at all times and at all wells.

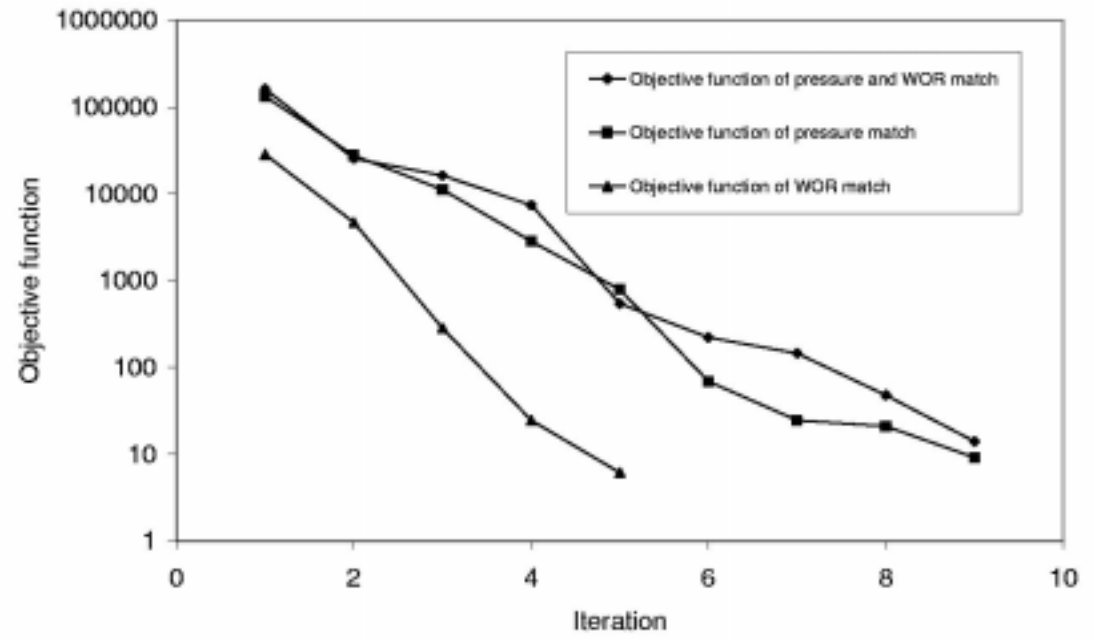

Figure 4.6: Comparison of the rate of minimization of the objective function when conditioning only to WOR, only to pressure, and to both WOR and pressure for the three-zone reservoir.

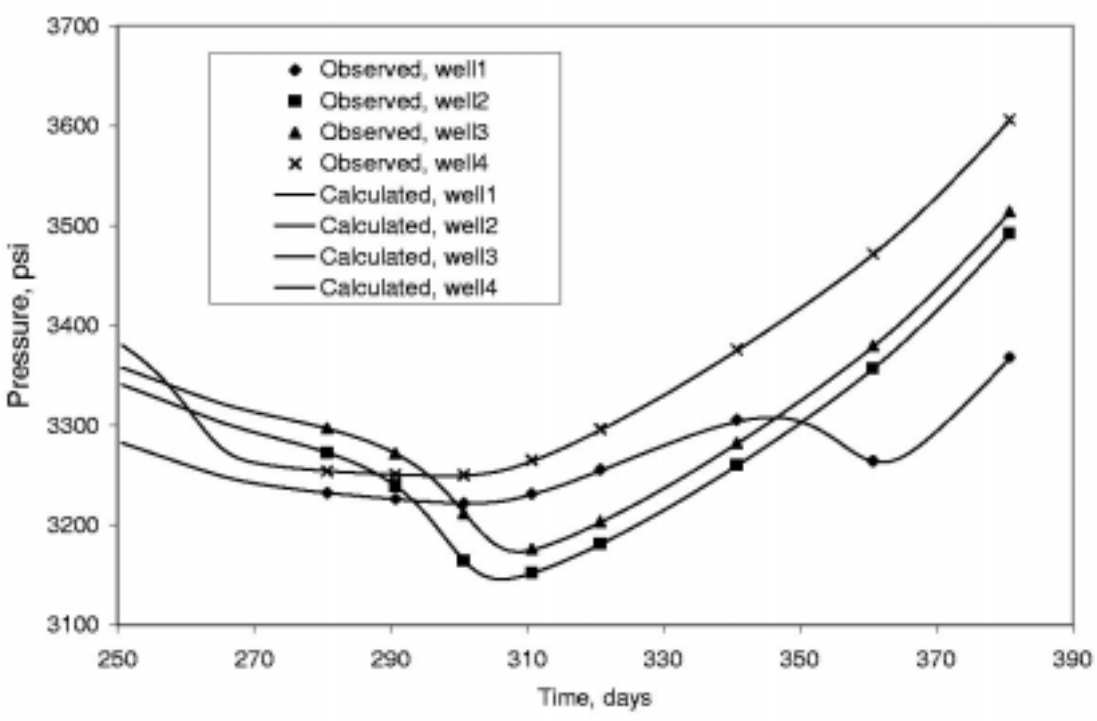

Figure 4.7: Flowing bottom-hole pressure match after conditioning to both WOR and pressure for three-zone reservoir.

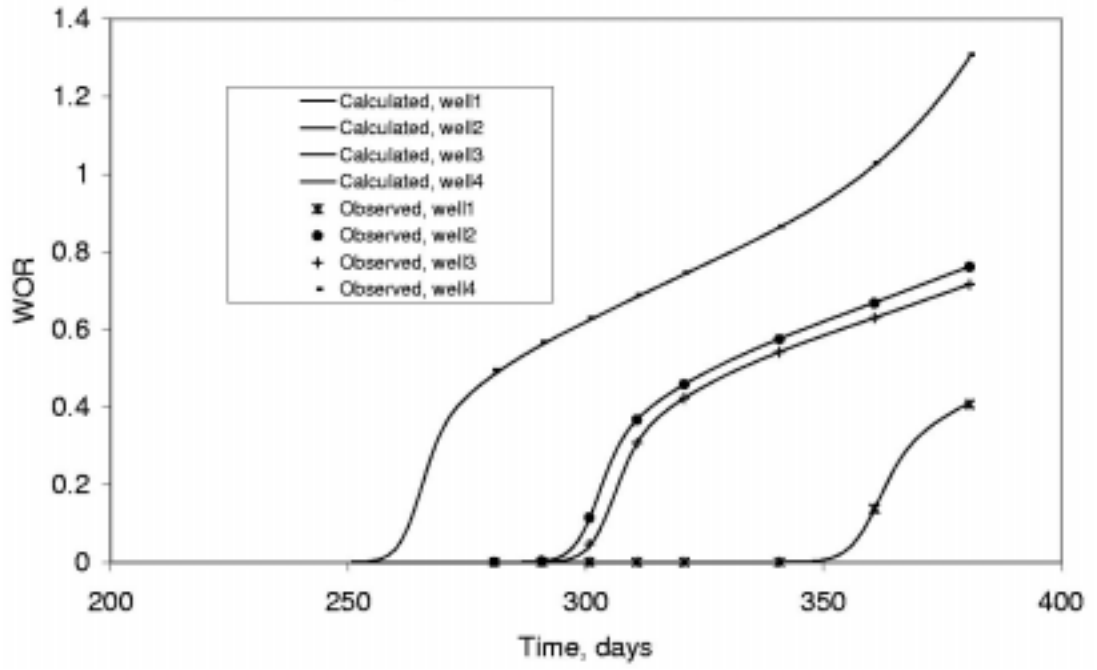

Figure 4.8: Water-oil ratio match after conditioning to both WOR and pressure for the three-zone reservoir.

4.2 MAP Estimate Conditional to Data from Heterogeneous Reservoir

The purpose of the three-zone reservoir example of the previous section was to illustrate the ability of the inversion method to recover the main features of the permeability and porosity distribution from pressure and water-oil ratio data. The composite reservoir provided a simple truth case for which comparisons could easily be made. Now we progress to the more challenging problems associated with highly heterogeneous permeability and porosity fields.

For this investigation, the reservoir itself is a square with length 1100 feet subdivided into 25 gridblocks in the x and y directions. The reservoir thickness is uniform and equal to 30 feet. Permeability is heterogeneous but isotropic. The "true" distributions for the porosity and log-permeability fields, from which the synthetic production data are generated, are correlated Gaussian random fields with anisotropic spherical variograms. The range of the variogram is 360 feet in the x-direction and 600 feet in the y direction. The correlation coefficient between log-permeability and porosity is 0.5. The Xu et al. [55] screening hypothesis is used to construct cross covariances. Production wells are located at (6,6), (6,20), (20,6) and (20,20). The injection well is located at (13,13). The total flow rate at each producer is fixed at 250 bbl/day and the rate at the injector is fixed at 1250 bbl/day. This provides approximately balanced production and injection before the breakthrough of water. Other relevant reservoir properties are listed in Table 4.1.

The true log-permeability field is shown on the left side of Fig. 4.9. The range of values is much larger than in the three-zone composite reservoir. Here the ratio of largest permeability to smallest is greater than 30. The observed water-oil ratio and pressure data were generated from the true permeability and true porosity fields. The true porosity field is shown on the left side of Fig. 4.10.

Because Well 1 (upper left quadrant) is located in a region of low porosity, it has early breakthrough (Fig. 4.11). Since the permeability in this region is also

Mean porosity	0.25
Mean log-permeability	4.0
Porosity variance	0.0025
Variance of log-permeability	0.5
Correlation coefficient between $\ln(k)$ and ϕ	0.5
Oil viscosity, cp	0.82
Variogram range in x-direction, ft	360
Variogram range in y-direction, ft	600
Water viscosity, cp	1.0
Initial water saturation	0.16
Residual oil saturation	0.20
Water compressibility, psi^{-1}	4×10^{-6}
Oil compressibility, psi^{-1}	1.0×10^{-5}
Initial pressure, psi	4,000
Oil formation volume factor at p_i	1.269
Water formation volume factor at p_i	1.004

Table 4.1: Rock and fluid properties for heterogeneous reservoir example.

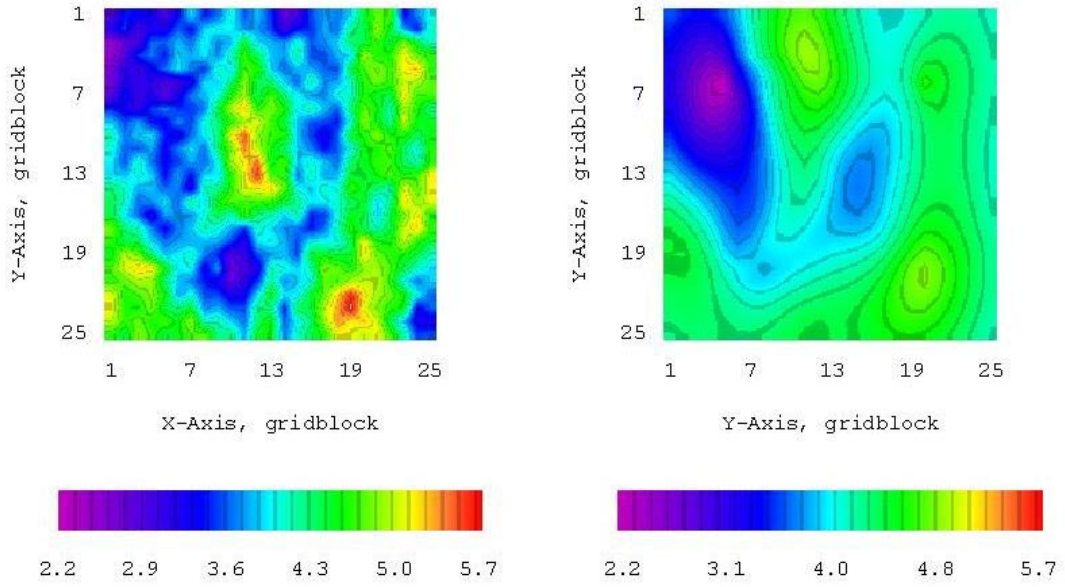

Figure 4.9: The "true" log-permeability field (left) and the MAP estimate of log-permeability conditioned to both WOR and pressure measurements (right).

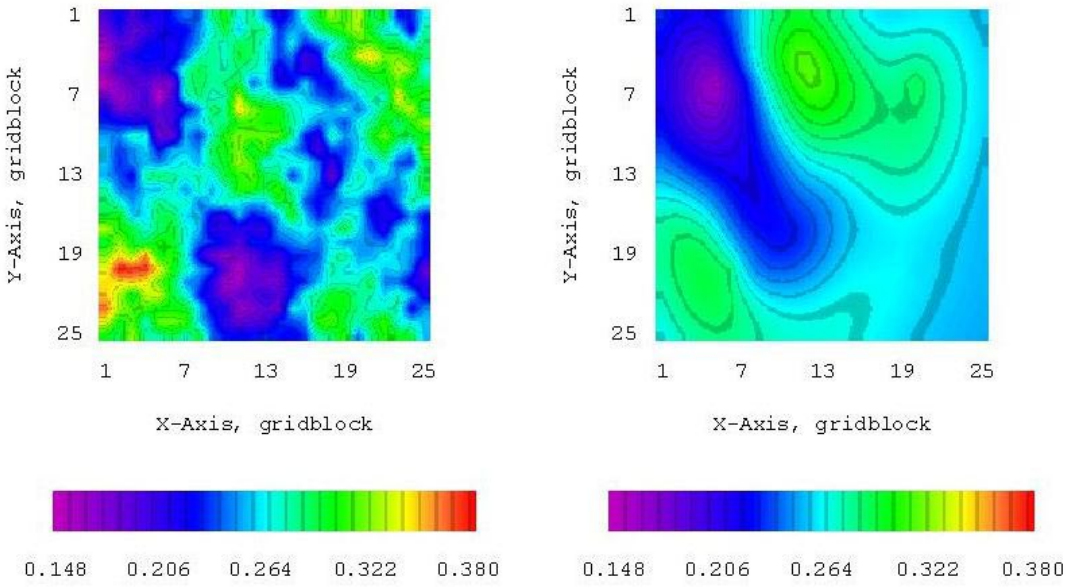

Figure 4.10: The "true" porosity field (left) and the MAP estimate of log-permeability conditioned to both WOR and pressure measurements (right).

low, the flowing bottom-hole pressure is low; see Fig. 4.12. Well 4, on the other hand, is located in a region of relatively high permeability and porosity and thus has late breakthrough and high flowing bottom-hole pressure. Wells 2 and 3 are intermediate in breakthrough times and pressure. The maximum a posteriori estimates of log-permeability and porosity, respectively, are shown on the right sides of Figs. 4.9 and 4.10. Although the estimate of permeability near the location of Well 1 (upper left quadrant) looks somewhat lower than might be expected from the true permeability field, the WOR match (Fig. 4.11) and the pressure match (Fig. 4.12) is seen to be excellent and the overall appearance of the MAP estimate is that it is a smoothed version of the true distributions.

4.3 Realization Conditional to Data from Heterogeneous Reservoir

As discussed earlier, the procedure for generating conditional realizations is very similar to the method for generating the maximum a posteriori estimate. The principal difference is that instead of finding distributions of porosity and log-permeability that are close to the prior estimates, we find distributions that are close to the unconditional realizations. On the left side of Fig. 4.13, the true log-permeability is shown. This distribution of log-permeability was used to generate synthetic observed water-oil ratio and pressure data. In the center of Fig. 4.13 is the unconditional realization that was generated using the same covariance and mean as used to generate the true realization. Note, however, that although the correlation length and range are approximately the same, the details of the distribution of log-permeability are quite different. On the right side of Fig. 4.13 is the conditional realization that corresponds to the unconditional model realization in the center of the figure and to an unconditional realization of the data. The conditional realization is quite similar to the true realization, both in correlation length and in detail.

Figure 4.14 shows the corresponding sequence of porosity realizations: the

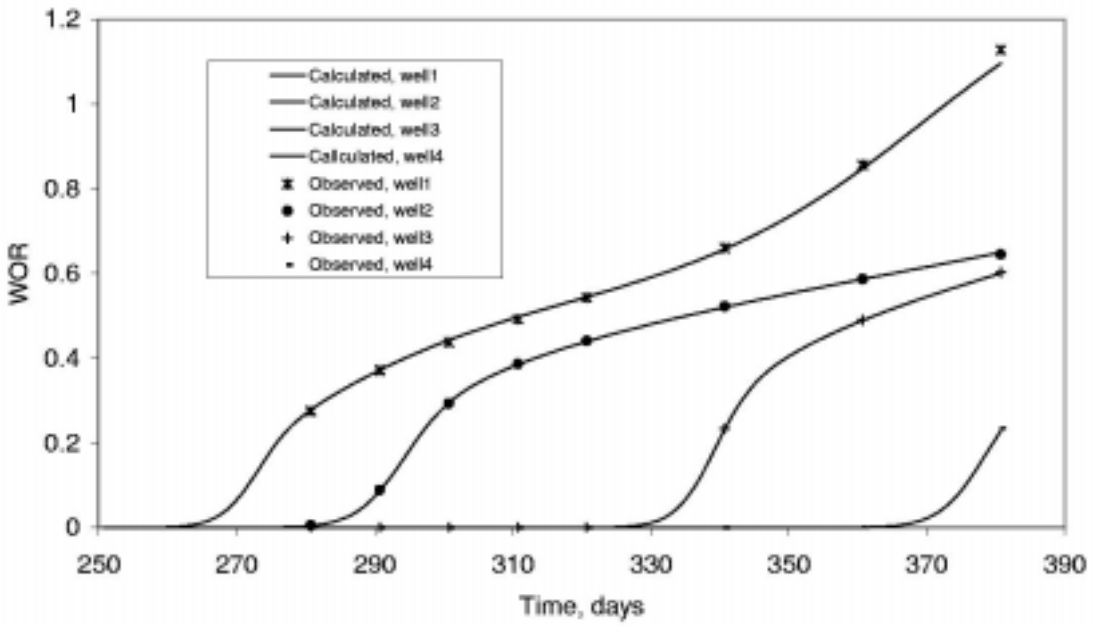

Figure 4.11: Water-oil ratio match at the four producing wells at the MAP solution with heterogeneous permeability and porosity.

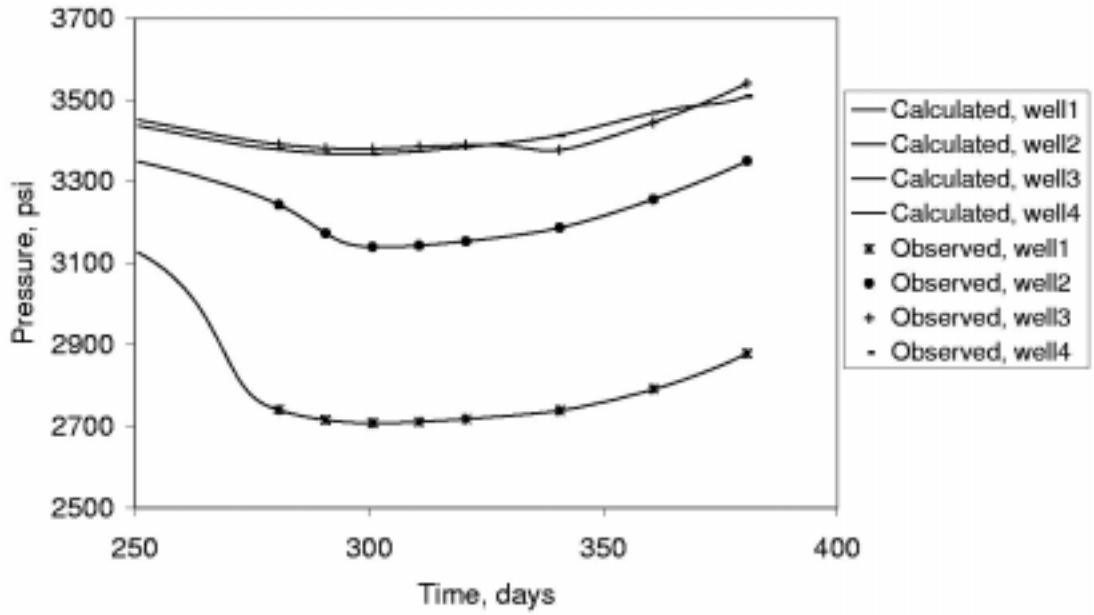

Figure 4.12: Pressure match at the four producing wells at the MAP solution with heterogeneous permeability and porosity.

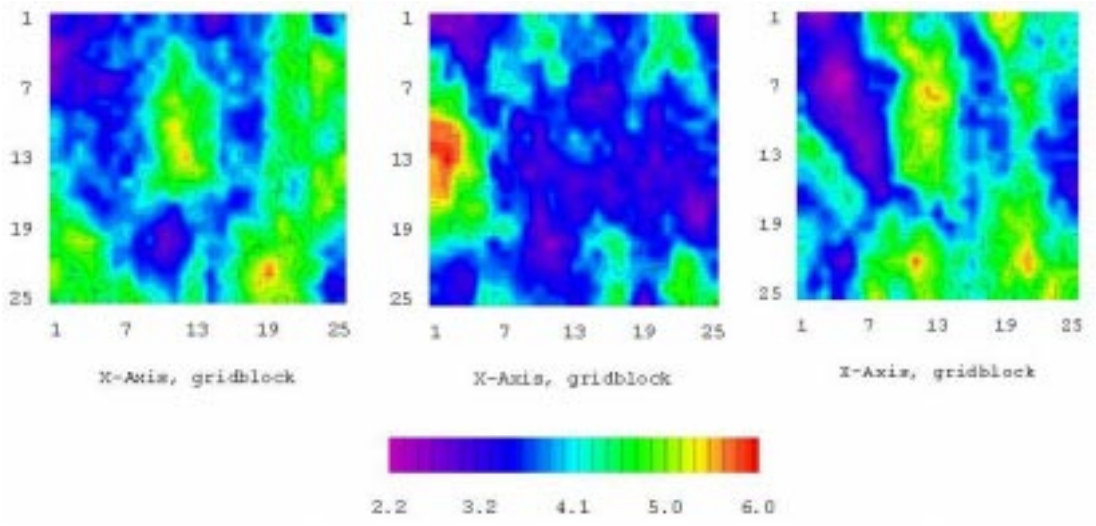

Figure 4.13: The true log-permeability field (left), an unconditional realization (center), and a conditional realization (right) that is "close" to the unconditional realization.

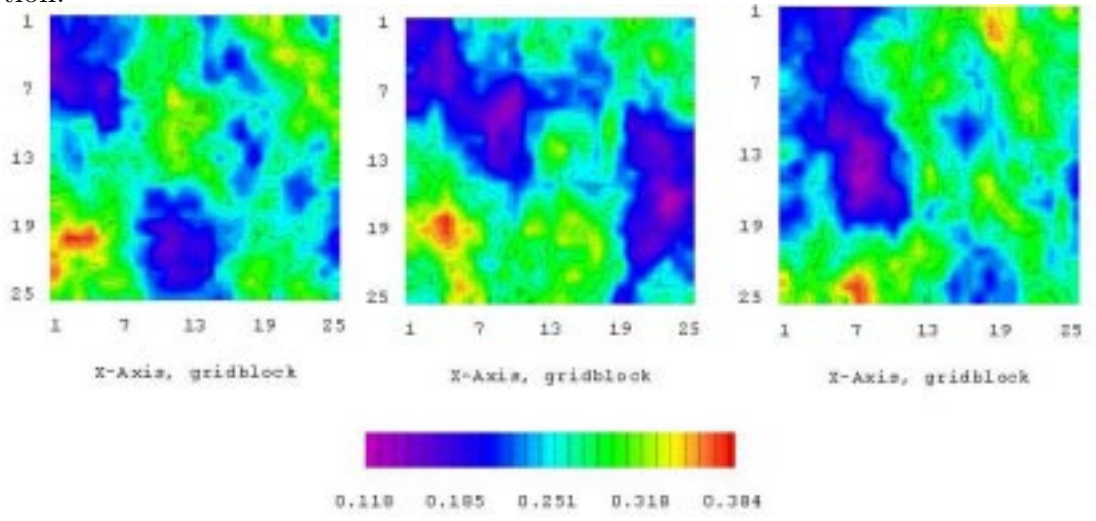

Figure 4.14: The true porosity field (left), an unconditional realization (center), and a conditional realization (right) that is "close" to the unconditional realization.

true distribution, an unconditional realization, and an conditional realization. The conditional realization of porosity is very similar to the true distribution of porosity.

The pressure and water-oil ratio data were matched closely as shown in Figs. 4.15 and 4.16, although several more iterations were required to generate a conditional realization as to generate the maximum a posteriori estimate (see Figure 4.18).

The fact that many more iterations were required to generate a realization is a source of some concern. Although it could be due to the high degree of non-linearity in the multiphase problem and the fact that we must linearize around heterogeneous permeability and porosity fields, it could simply be due to the fact that we have not optimized the damping schedule in the minimization procedure. In Fig. 4.19, we have plotted the differences between the starting (unconditional) log-permeability field and the final, conditional log-permeability field, and between the unconditional and conditional porosity fields. The fact that these differences are quite smooth causes us to believe that the convergence rate can be substantially improved.

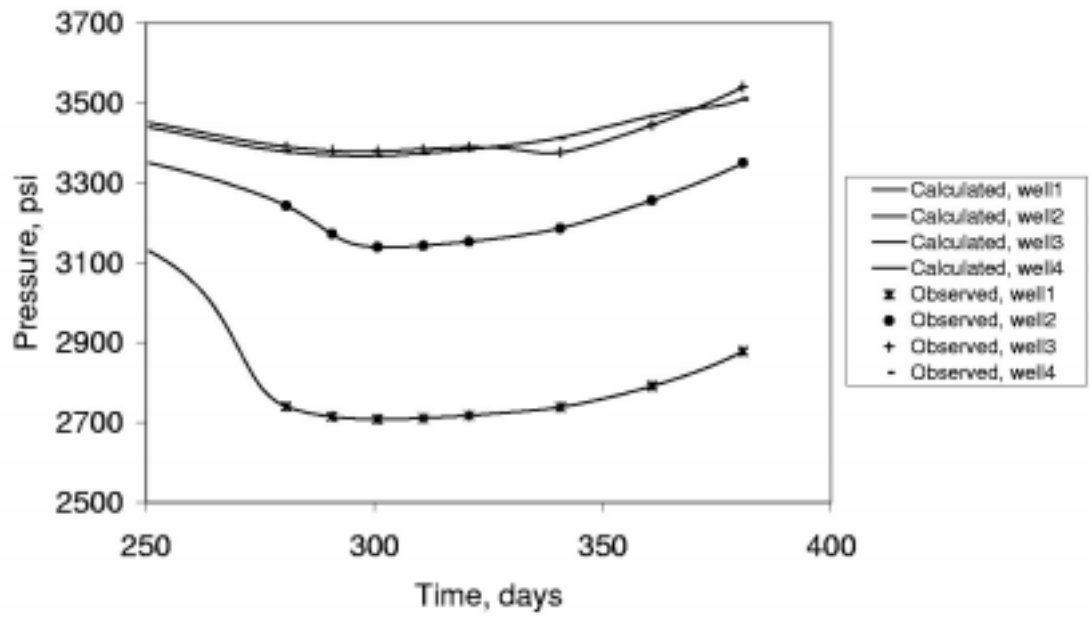

Figure 4.15: Pressure from the conditional realization compared with "observed" pressure measurements from the true model.

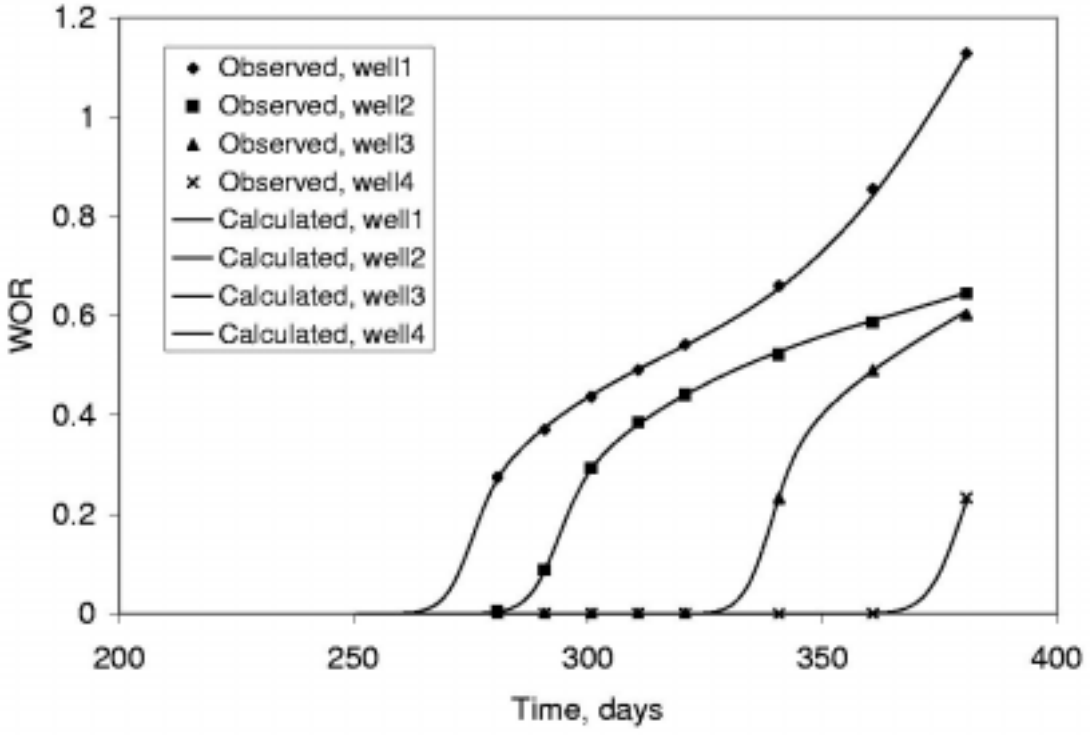

Figure 4.16: WOR from the conditional realization compared with observed WOR from the true model.

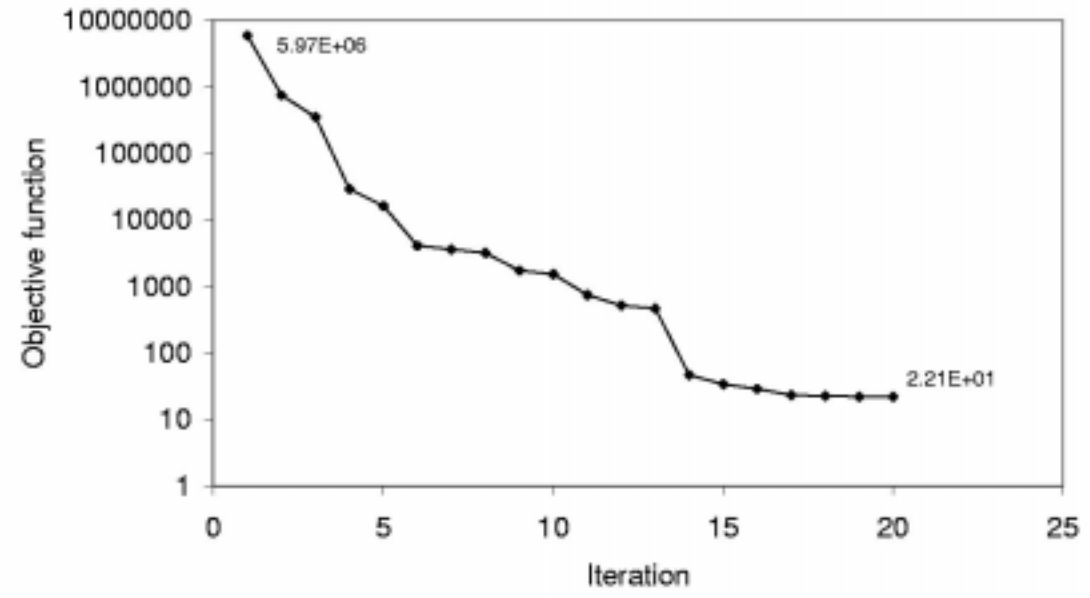

Figure 4.17: Objective function of WOR and pressure match for heterogeneous (conditional) realization.

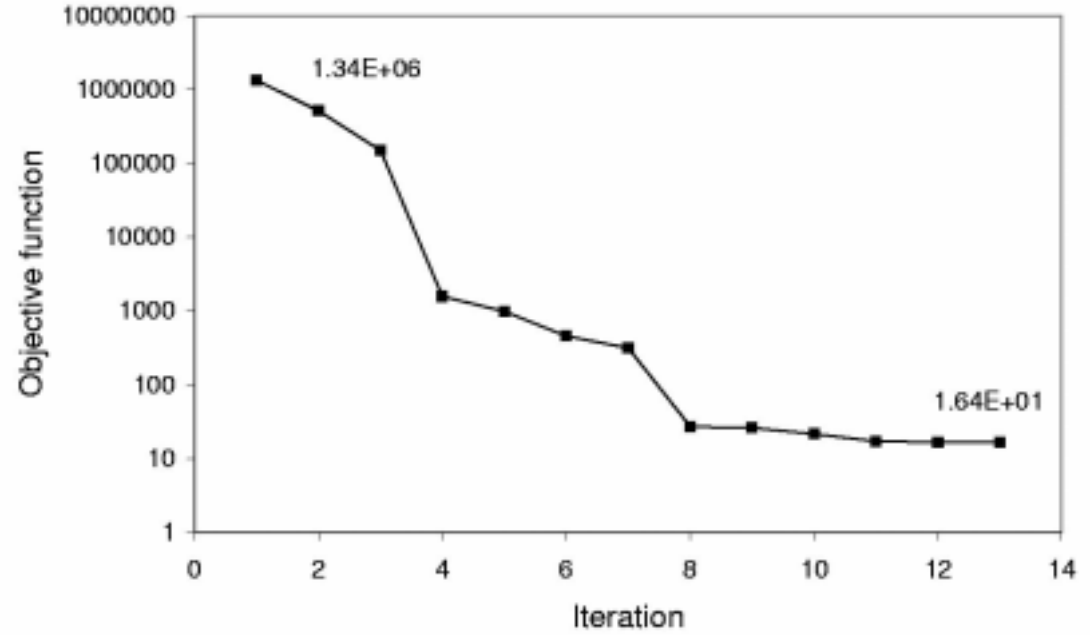

Figure 4.18: Objective function conditioned to WOR and pressure for heterogeneous reservoir. (The maximum a posteriori solution.)

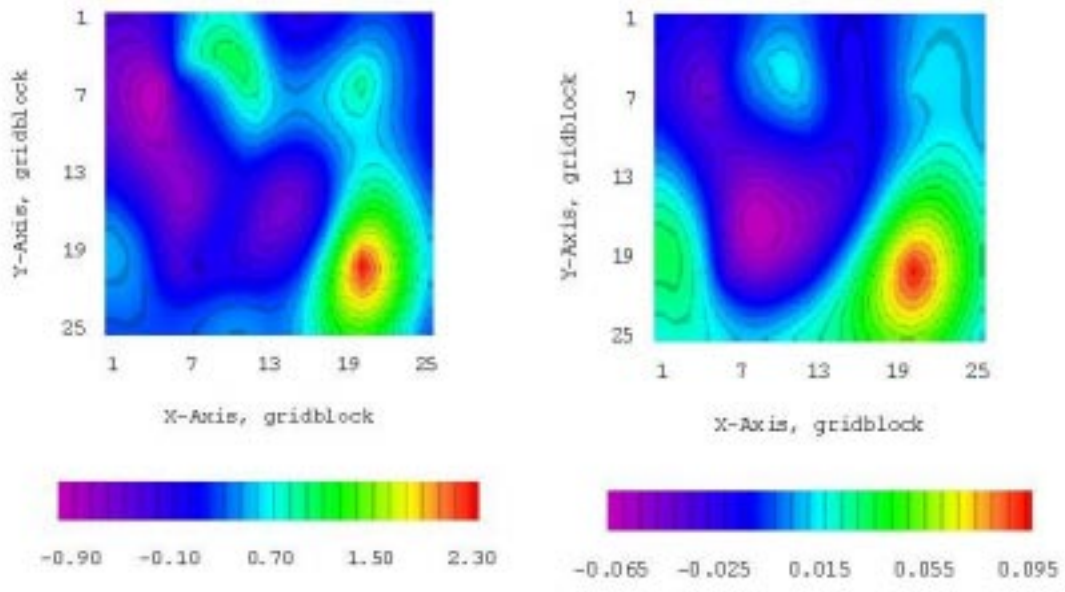

Figure 4.19: The correction to the log-permeability field (left) and the porosity field (right).

4.4 Evaluating the Uncertainty in Performance Predictions

The recommended procedure for evaluating the uncertainty in reservoir performance predictions was discussed in Section 2.8. For this purpose, fifty unconditional realizations of the rock property fields were generated for the heterogeneous reservoir example considered above. Each realization is conditioned to a realization of the production data (i.e., synthetic production data from the true reservoir with additional noise added) as discussed in Section 2.7. Both WOR and wellbore pressure data (at producing wells) were used for conditioning the realizations. The "known history" of the reservoir encompasses only 380 days. To evaluate the uncertainty in performance predictions, reservoir performance up to approximately 10 years was simulated for each conditional and unconditional realization. In each such prediction, the known rate history at each well was used to specify the wellbore constraints up to the end of the history period (380 days). At 380 days, the well constraint at all producing wells was changed to production at a fixed bottom-hole pressure (the bottom-hole pressure at the previous time step). At the injection well, the wellbore was constrained to constant rate for the entire history and prediction period.

For ten of the unconditional realizations, it was not possible to reduce the mismatch in the data to a small level. Fig. 4.20 shows that, in most cases, the value of the objective function used in minimization (Eq. 2.93) was on the order of 50. In ten cases, however, the value of the objective function remained above 150, and in some cases was greater that 10,000 (Fig. 4.21). These "realizations" were rejected so, although there are 50 unconditional realizations, there are only 40 conditional realizations.

4.4.1 Unconditional Realizations

Fig. 4.22 shows predicted oil production at Well 1 for each of the unconditional realizations and the true oil production rate. Fig. 4.23 shows similar results

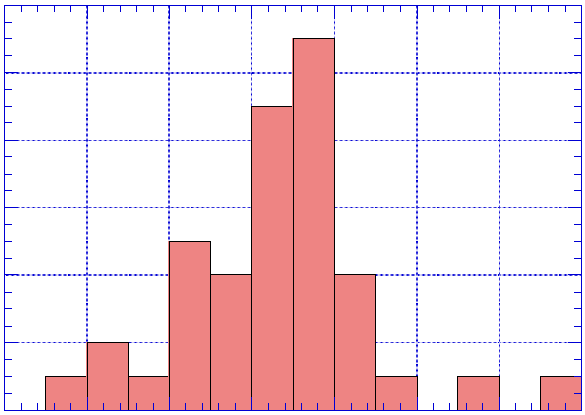

Figure 4.20: The main part of the frequency distribution of values of the objective function after conditioning.

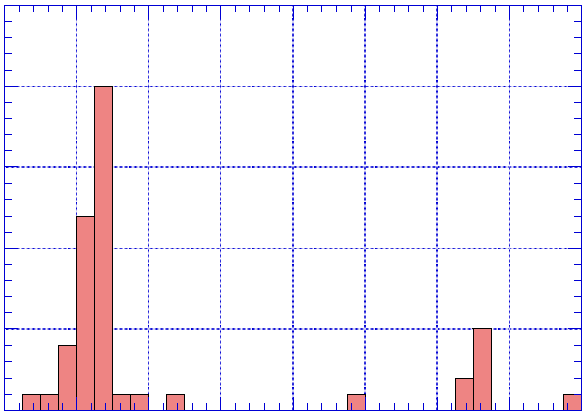

Figure 4.21: The entire frequency distribution of values of the log of the objective function after conditioning. The large values are unacceptable realizations.

for the predicted water rates at Well 1. In both figures, the prediction from the truth case is shown by a thick dashed red line. Fig. 4.24 and Fig. 4.25, respectively, show the unconditional realizations of cumulative oil production for Well 1 and for the entire field.

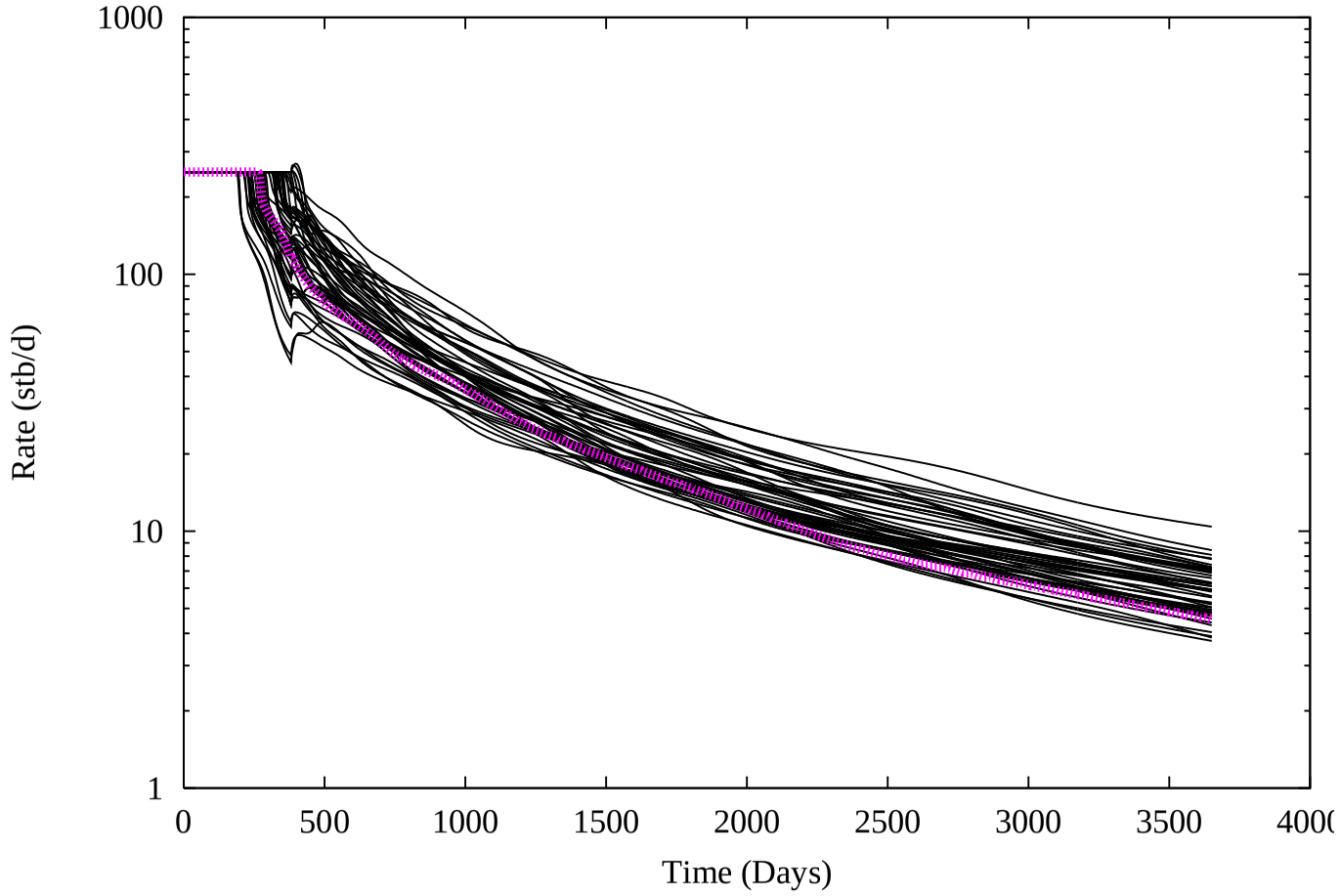

Figure 4.22: Predictions of oil production rate at Well 1 for 50 unconditional reservoir realizations. Production from the true reservoir is shown in red.

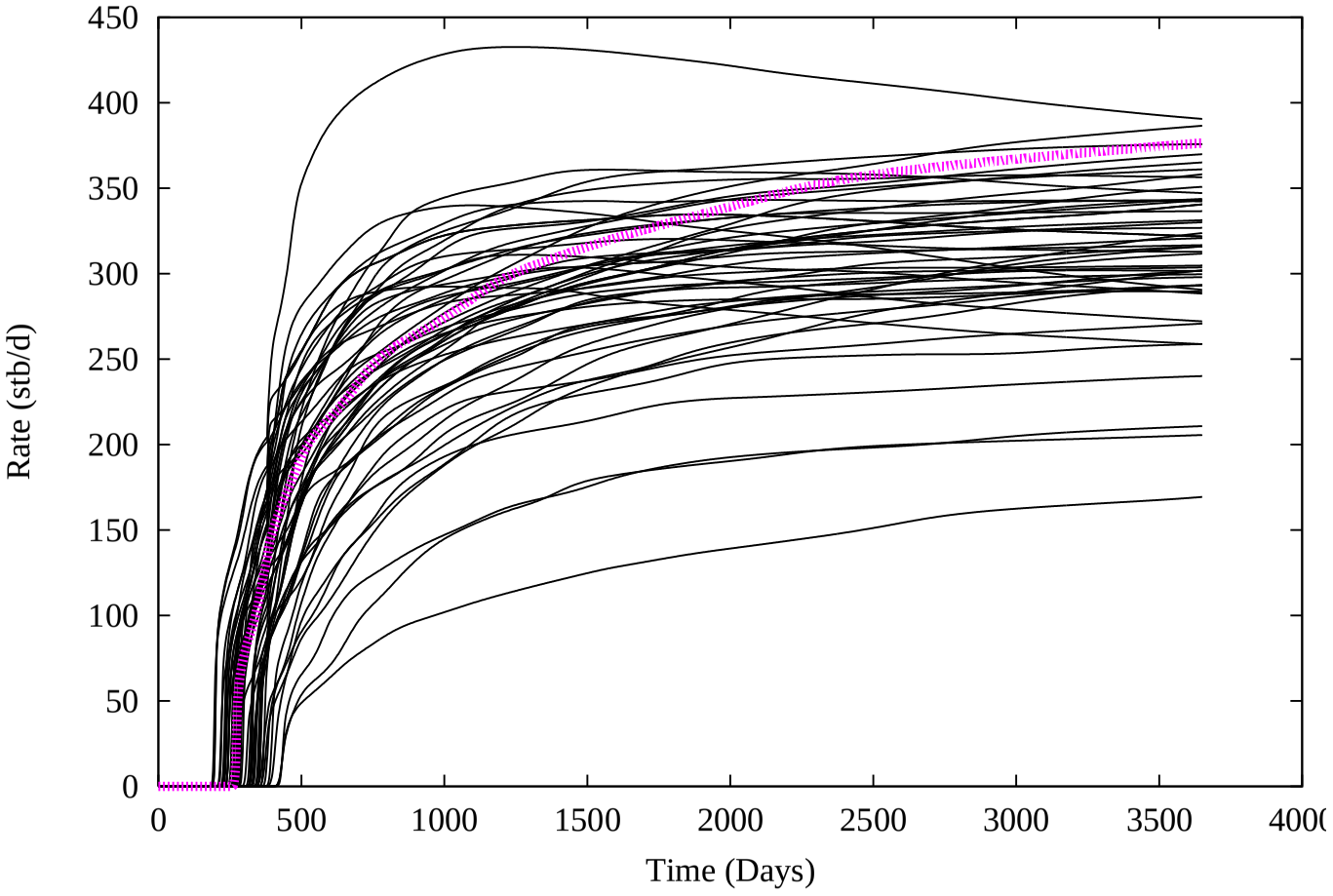

Figure 4.23: Predictions of water production rate at Well 1 for 50 unconditional reservoir realizations. Production from the true reservoir is shown in red.

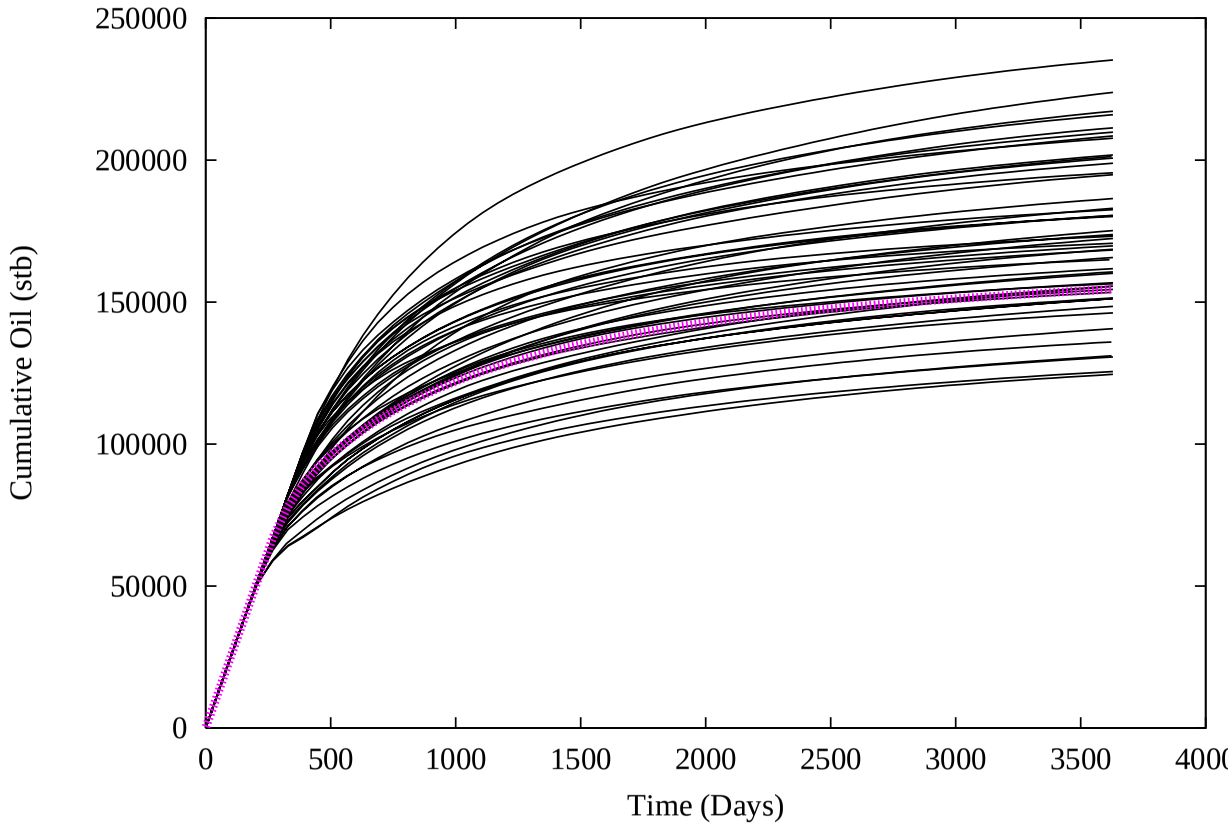

Figure 4.24: Unconditional realizations of the cumulative oil production from Well 1. The true cumulative oil is shown in red.

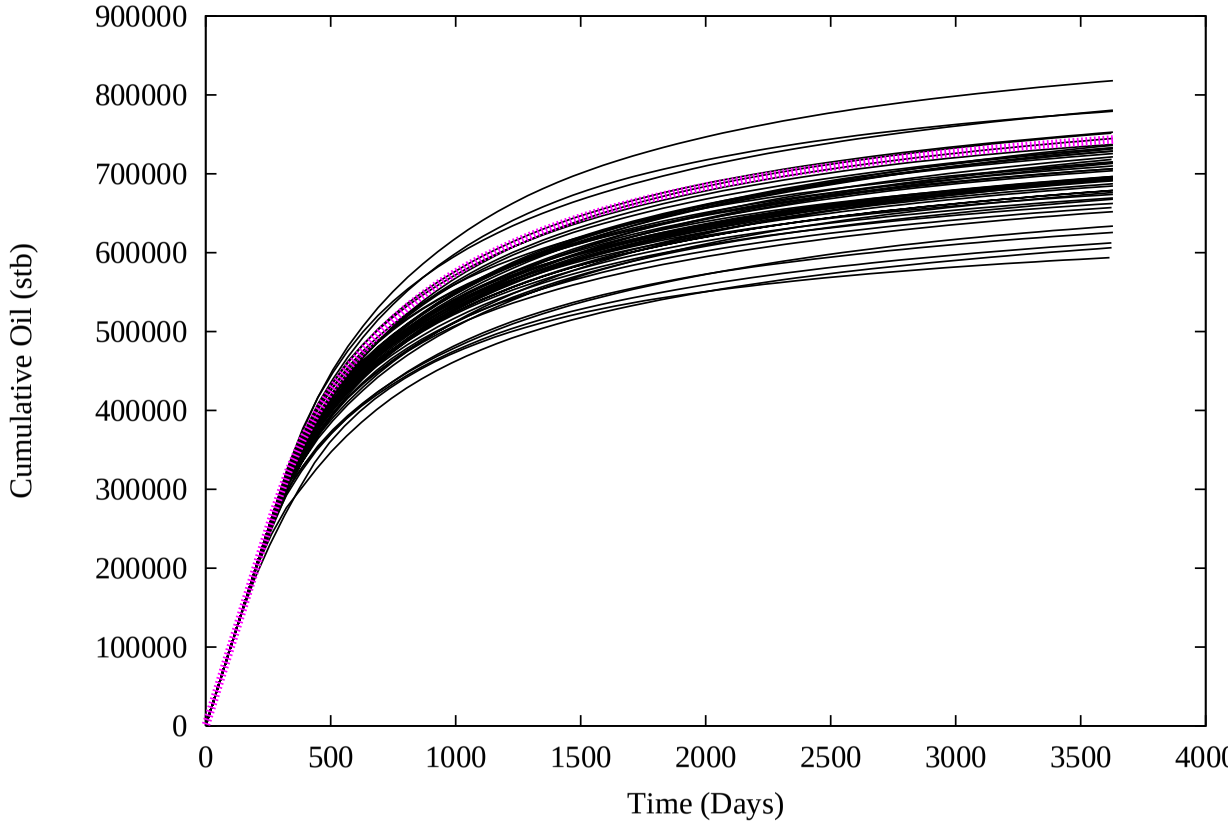

Figure 4.25: Unconditional realizations of the cumulative oil production from all wells. The true cumulative oil production from the field is shown in red.

4.4.2 Conditional Realizations

Fig. 4.26 shows predicted oil production at Well 1 for each of the unconditional realizations and the true oil production rate. Fig. 4.27 shows similar results for the predicted water rates at Well 1. In both figures, the prediction from the truth case is shown by a thick dashed red line. Fig. 4.28 and Fig. 4.29, respectively, show the conditional realizations of cumulative oil production for Well 1 (located in the upper left corner of the reservoir in gridblock (6,6)) and for the entire field. Only the accepted realizations are included in these plots.

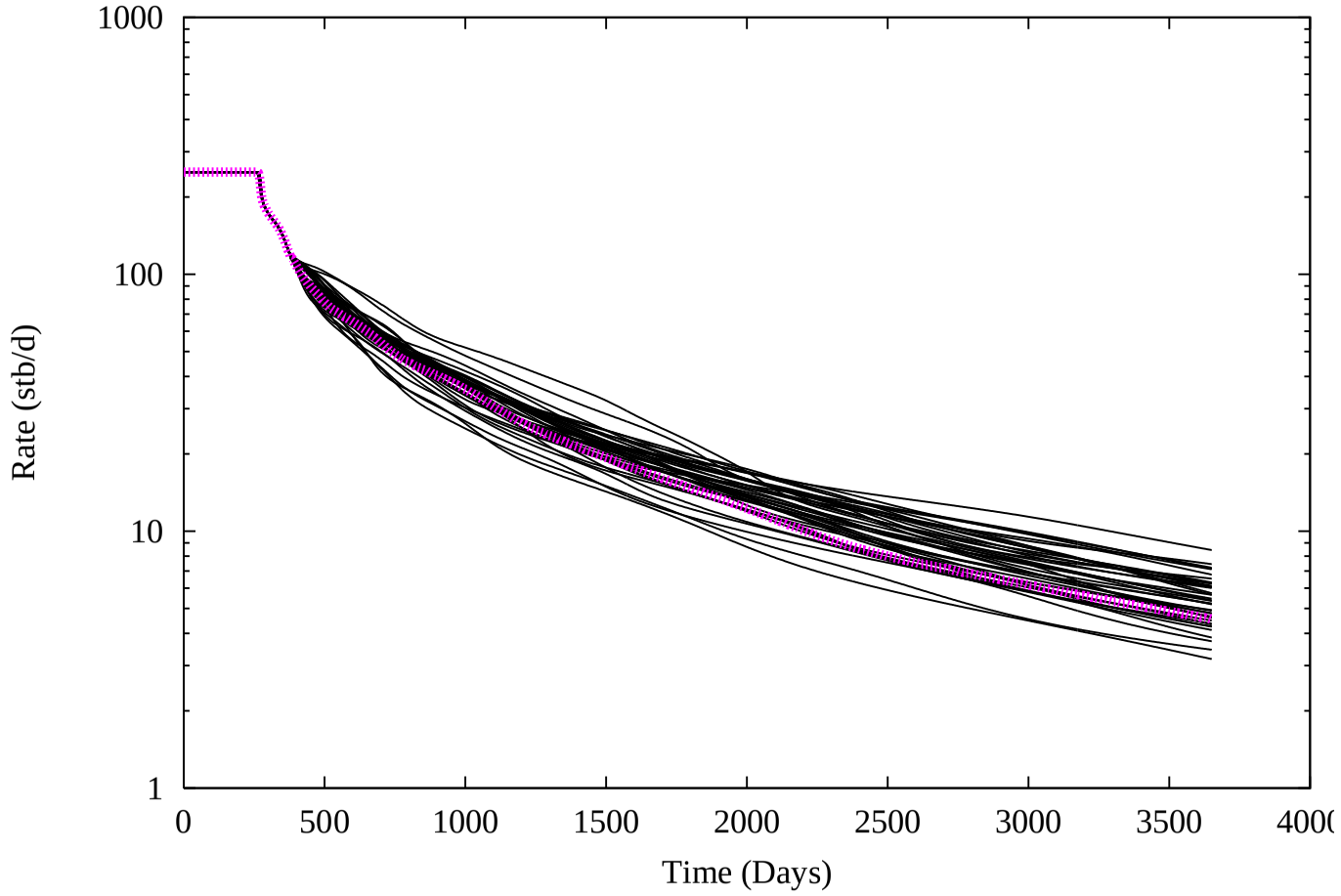

Figure 4.26: Predictions of oil production rate at Well 1 for 40 conditional reservoir realizations. Production from the true reservoir is shown in red.

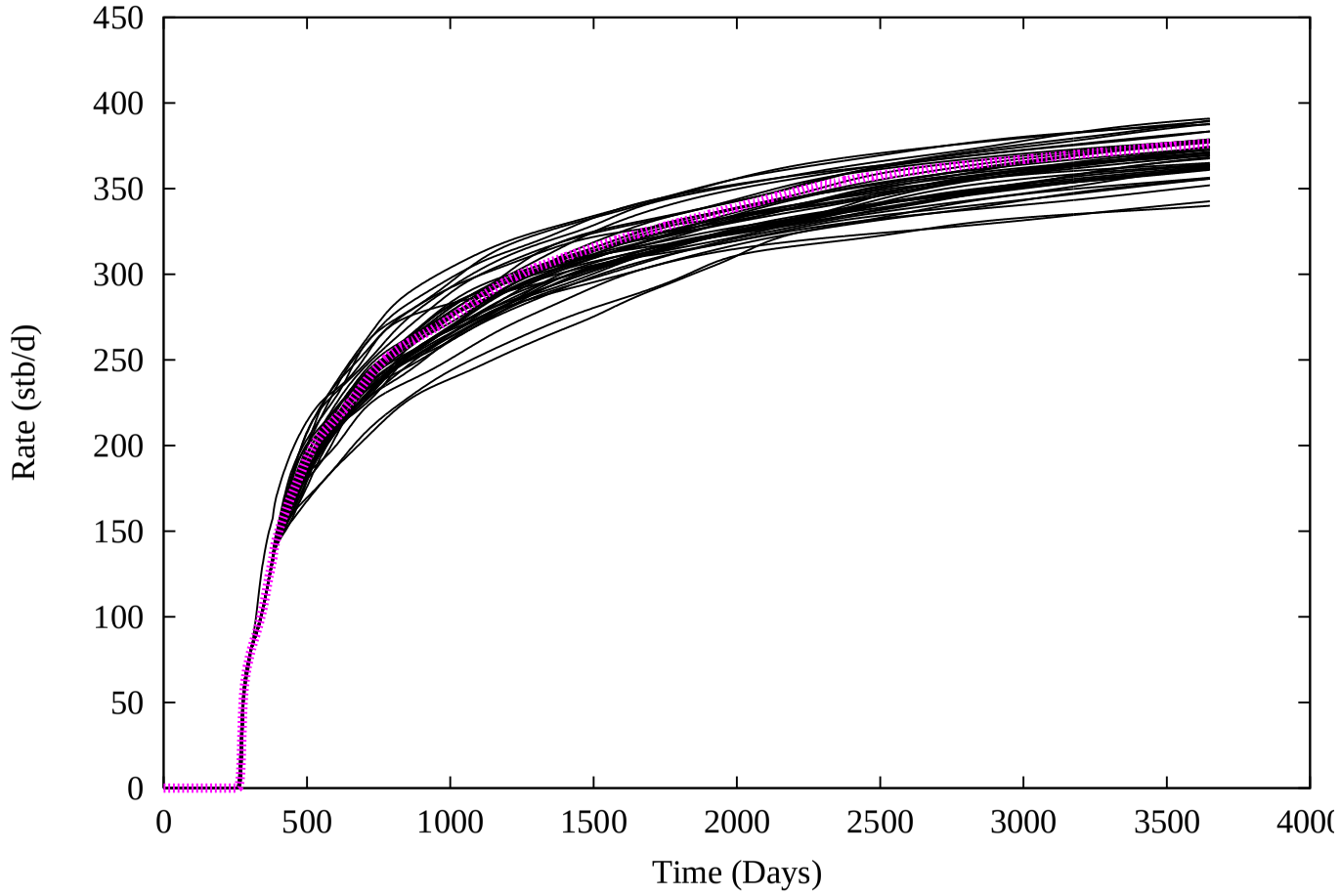

Figure 4.27: Predictions of water production rate at Well 1 for 40 conditional reservoir realizations. Production from the true reservoir is shown in red.

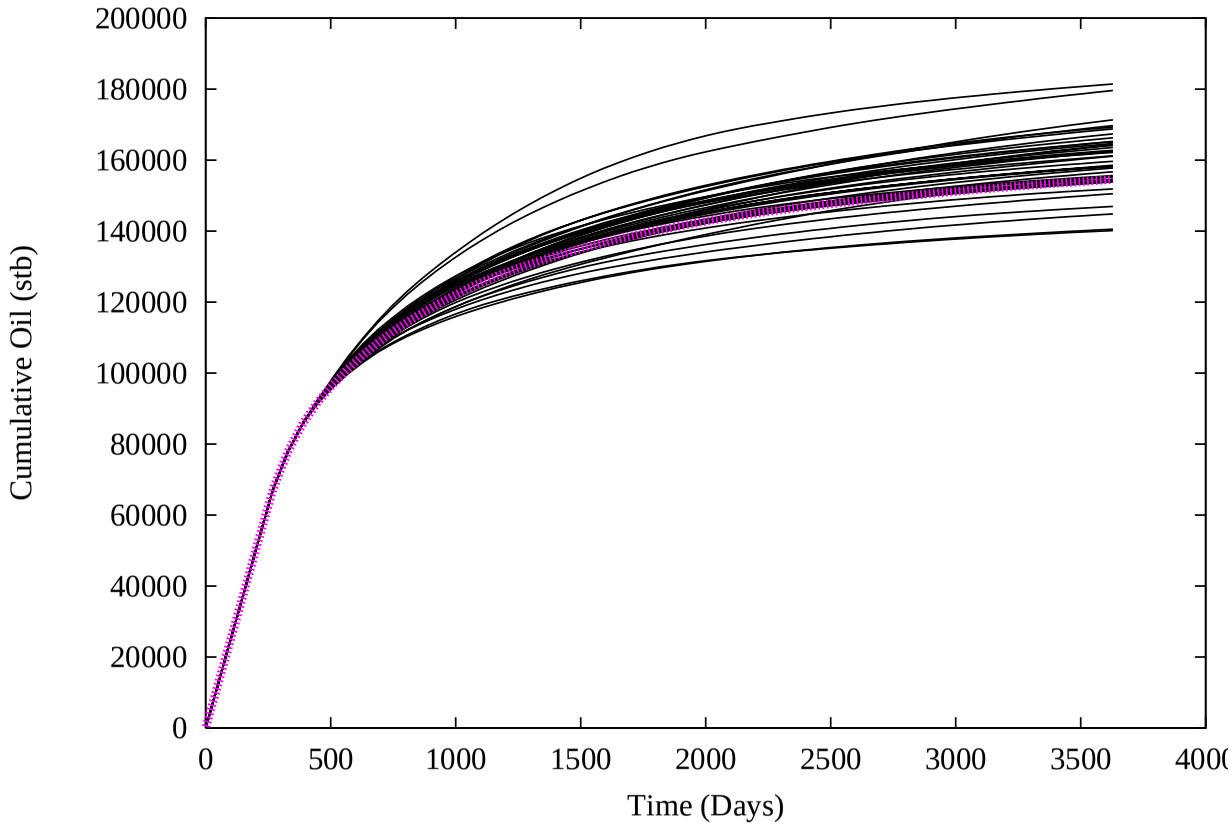

Figure 4.28: Conditional realizations of the cumulative oil production from Well 1. The true cumulative oil is shown in red.

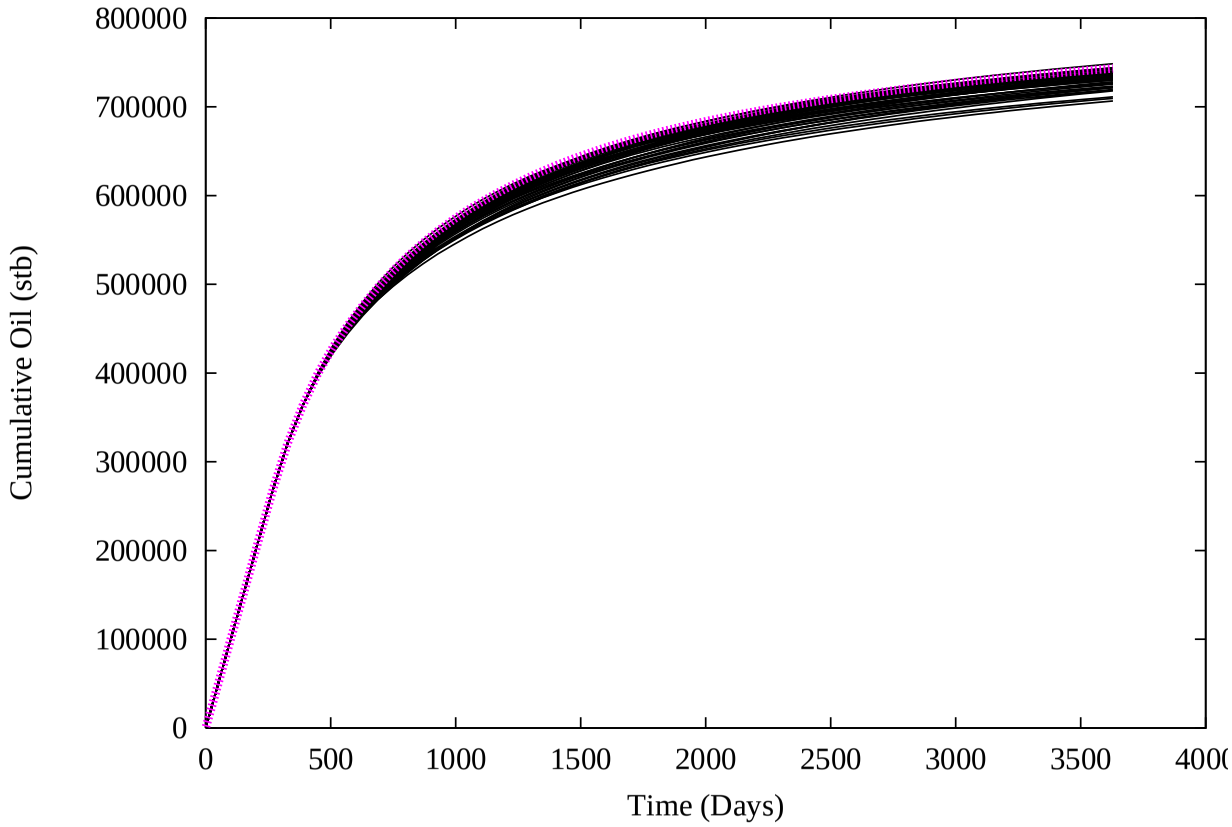

Figure 4.29: Conditional realizations of the cumulative oil production from all wells. The true cumulative oil production from the field is shown in red.

4.4.3 Comparison of Conditional and Unconditional Realizations

Although it is clear from Figs. 4.25 and 4.29, for example, that the uncertainty in the future performance of the well and the field were reduced by conditioning to the early production data, the degree of reduction is not obvious from the figures. It is far easier to quantify the uncertainty and the reduction in uncertainty, by looking at the predicted performance at a particular instant in time.

In Fig. 4.30, the distribution of unconditional realizations of cumulative production at ten years is shown. Note that the uncertainty is quite large; the standard deviation divided by the mean is approximately 0.156. After conditioning to the early production data, the distribution of realizations can be seen to be much narrower (Fig. 4.31). The standard deviation of the conditional realizations of cumulative oil production, divided by the mean is only about 0.055, so the relative uncertainty is reduced by a factor of almost three.

Figs. 4.32 and 4.33 show similar results for the distribution of realizations of cumulative oil production from the entire field. Note, however, that although the standard deviation of the total cumulative oil production is much larger than the standard deviation of the individual well cumulative production, it is only 1.6 times as large for the unconditional realizations. If the production from each well was random and independent of each of the other wells, the variance in the total should be the sum of the variances of the individual wells, so the standard deviation for the field cumulative would be twice as large as the standard deviation in the cumulative production at each well. Because it is smaller than this,we conclude that the predicted performance of the four producers is not independent, even for the unconditional realizations. Conditioning to early production data is more beneficial in reducing uncertainty in the total field production than in the production from individual wells (see Table 4.2). The relative uncertainty after conditioning is only 0.013. It is unlikely, however, that such small uncertainties are typical. In this

	Unconditional		Conditional	
	Well 1	Field	Well 1	Field
Realizations	50	50	40	40
Mean	173400.	699700.	160100.	731200.
Median	170300.	696300.	161200.	732600.
Std Deviation	27100.	43400.	8800.	9800.
Std Error	3830.	6130.	13920.	1540.
Skewness	0.250	-0.087	-0.061	-0.667
Kurtosis	-0.656	0.722	0.527	-0.019

Table 4.2: A summary of the statistics of the distribution of conditional and unconditional realizations of cumulative oil production at the end of ten years.

synthetic example, the total pore volume was determined quite well by the early data, and the original saturation and relative permeabilities were assumed to be known. If neither the initial sturation or the relative permeabilities were known, the uncertainty would be much larger.

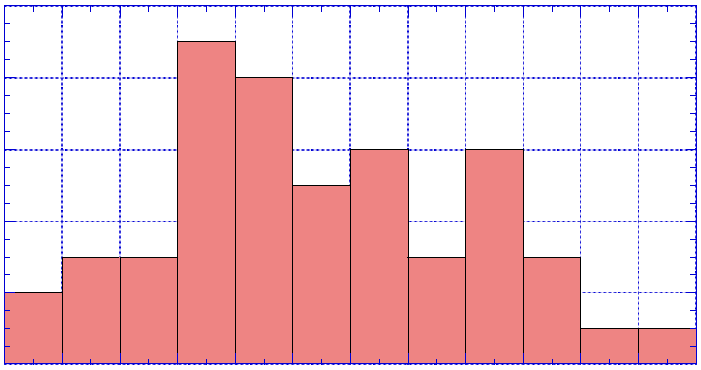

Figure 4.30: The distribution of unconditional realizations of cumulative oil production from Well 1 at the end of 10 years of production.

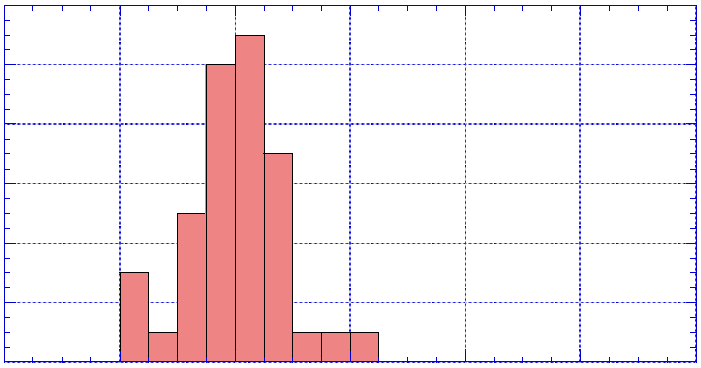

Figure 4.31: The distribution of conditional realizations of cumulative oil production from Well 1 at the end of 10 years of production.

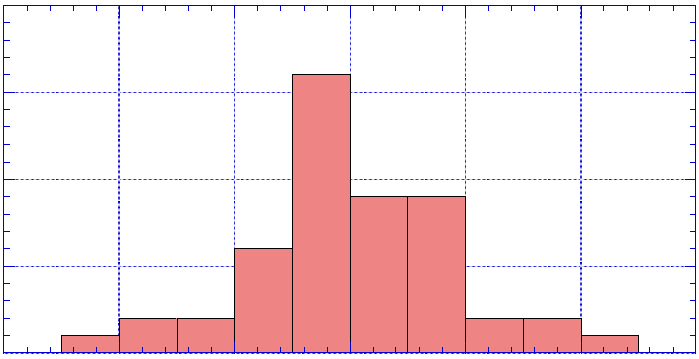

Figure 4.32: The distribution of unconditional realizations of cumulative field oil production at the end of 10 years of production.

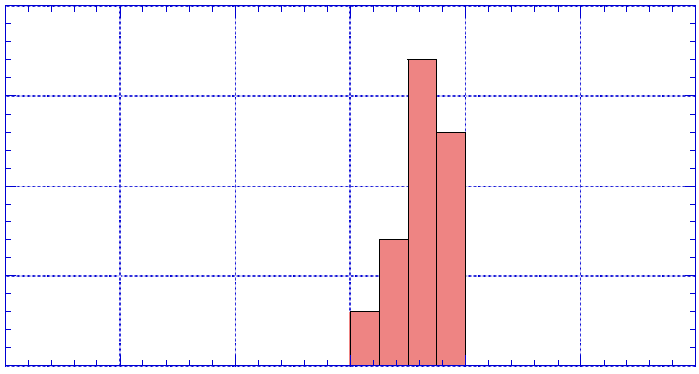

Figure 4.33: The distribution of conditional realizations of cumulative field oil production at the end of 10 years of production.

CHAPTER V

CONCLUSIONS

We have considered the problem of generating the maximum a posteriori (MAP) estimate and realizations of log-permeability and porosity fields conditioned to a prior multivariate Gaussian distribution and two-phase flow production data. Only the two-dimensional two-phase flow of oil and water has been considered. However, the basic ideas presented should extend to general three-dimensional multiphase flow problems. The rock property fields correspond to reservoir simulator gridblock permeabilities and porosities. Realizations and the MAP estimate require the minimization of an objective function related to the a posteriori probability density function (pdf) for the rock property fields. Definition of this pdf and its associated objective function requires a probabilistic specification of measurement errors in observed production data. We have presented a method to model these errors which recognizes that, on a percentage basis, the measurement error in water-oil ratio may be large when the water production rate is small. Minimization of the appropriate objective function is accomplished by applying the Gauss-Newton iteration procedure. We found, however, that when only long-time production data are used as conditioning data in the inverse procedure, the Gauss-Newton method converges slowly or may be trapped in a local minimum. This result occurs because at early iterations, the production data mismatch terms are large and dominate the search direction; the effect of the prior model, which provides regularization term is small. Because of this, excessively rough rock property fields at the are obtained at the first iteration of the Gauss-Newton algorithm and it is difficult to correct these fields at subsequent iterations. We have developed a procedure to overcome this convergence problem. The method,

which uses an artificially high variance for measurement errors at early iterations, is conceptually similar to the Levenberg-Marquardt algorithm.

In the Gauss-Newton method, a significant amount of the total computational work is expended in generating sensitivity coefficients. A rigorous discrete adjoint method for calculating the sensitivity of production data to the rock property fields has been derived. Unlike traditional adjoint approaches which generate only the gradient of an objective function based on the sum of squares of production data mismatch terms, our approach allows one to generate directly the sensitivity of calculated data to the model parameters (gridblock log-permeabilities and porosities). The gradient simulator appears to be the only available alternative to the adjoint method. Which is preferable depends directly on the number of model parameters (number of gridblock log-permeabilities and porosities for the problem considered here) and the number of production data used as conditioning data. Both methods solve linear problems of the same size to obtain sensitivity coefficients. At least formally, both proceed by solving a set of matrix problems where each coefficient matrix is associated with multiple multiple right-hand sides. In the gradient simulator method, the number of right hand sides is equal to the number of model parameters, but in our adjoint formulation, the number of right-hand sides is no greater than the number of production data used as conditioning data. Thus, if the number of model parameters is significantly greater than the number of production data used as conditioning data, our adjoint formulation is preferable. If not, the gradient simulator procedure would be expected to be more computationally efficient. In terms of memory required, the adjoint method does not compare favorably with the gradient simulator method. In the adjoint method, all gridblock saturations and pressures from one simulation run must be stored and all are needed when solving the adjoint equations.

Using sensitivity coefficients calculated with the adjoint method, we have generated the MAP estimate and realizations of the rock property fields for synthetic cases. As notes above the MAP estimated and realizations are generated by mini-

mizing the appropriate objective function using an efficient Gauss-Newton algorithm instead of relying on slowly convergent methods (e.g., steepest descent or conjugate gradient) which utilize only the gradient of the objective function. As the formulation of the Gauss-Newton method applied requires the solution of $N_d \times N_d$ matrix problems where N_d is the number of production data, the Gauss-Newton iterations consume a relatively minor portion of the overall computation time provided the number of conditioning production data is not too large. The main work is in computing the sensitivity coefficients needed at each iteration of the Gauss-Newton algorithm. This requires one basic reservoir simulation run to generate gridblock pressures and saturations and the solution of linear adjoint systems backward in time with multiple right-hand side vectors.

Finally, we have applied a procedure to evaluate the uncertainty in predicted reservoir performance. In particular, we have shown that conditioning to geostatistical models to production data reduces the uncertainty in reservoir performance predictions far below the level of uncertainty obtained if one simply makes reservoir performance predictions from realizations generated from a prior geostatistical model constructed solely from static data.

REFERENCES

[1] F. G. Alabert. Constraining description of randomly heterogeneous reservoirs to pressure test data: A Monte Carlo study, SPE-19600. *64th Annual SPE Tech. Conf.*, Formation Evaluation and Reservoir Geology:307–321, 1989.

[2] F. Anterion, B. Karcher, and R. Eymard. Use of parameter gradients for reservoir history matching, SPE-18433. In *10th SPE Reservoir Simulation Symp.*, pages 339–354, 1989.

[3] Roger Barlow. *Statistics, A Guide to the Use of Statistical Methods in the Physical Sciences.* John Wiley & Sons, Chichester, England, 1989.

[4] R. Bissel. Calculating optimal parameters for history matching. In *Proceedings of the 4th European Conference on the Mathematics of Oil Recovery, Topic E: History Match and Recovery Optimization*, 1994.

[5] J. Carrera and S. P. Neuman. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resources Research*, 22(2):199–210, 1986.

[6] J. Carrera and S. P. Neuman. Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms. *Water Resources Research*, 22(2):211–227, 1986.

[7] J. Carrera and S. P. Neuman. Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data. *Water Resources Research*, 22(2):228–242, 1986.

[8] R. D. Carter, L. F. Kemp, and A. C. Pierce. Discussion of comparison of sensitivity coefficient calculation methods in automatic history matching. *Soc. Petrol. Eng. J.*, pages 205–208, April 1982.

[9] R. D. Carter, L. F. Kemp, A. C. Pierce, and D. L. Williams. Performance matching with constraints. *Soc. Petrol. Eng. J.*, 14(4):187–196, 1974.

[10] Guy M. Chavent, M. Dupuy, and P. Lemonnier. History matching by use of optimal control theory. *Soc. Petrol. Eng. J.*, 15(1):74–86, 1975.

[11] W. H. Chen, G. R. Gavalas, John H. Seinfeld, and Mel L. Wasserman. A new algorithm for automatic history matching. *Soc. Petrol. Eng. J.*, pages 593–608, December 1974.

[12] L. Chu, A. C. Reynolds, and D. S. Oliver. Computation of sensitivity coefficients for conditioning the permeability field to well-test pressure data. *In Situ*, 19(2):179–223, 1995.

[13] L. Chu, A. C. Reynolds, and D. S. Oliver. Reservoir description from static and well-test data using efficient gradient methods (SPE-29999). In *1995 SPE International Meeting*, Beijing, 1995.

[14] Luciane Bonet Cunha, Dean S. Oliver, R. A. Redner, and A. C. Reynolds. A hybrid Markov chain Monte Carlo method for generating permeability fields conditioned to multiwell pressure data and prior information. *Soc. Pet. Eng. J.*, 3(3):261–271, Sept. 1998.

[15] C. V. Deutsch and P. W. Cockerham. Practical considerations in the application of simulated annealing to stochastic simulation. *Mathematical Geology*, 26(1):67–82, 1994.

[16] Clayton V. Deutsch. *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data*. Ph.D. thesis, Stanford University, Stanford, California, 1992.

[17] Richard E. Ewing, Michael S. Pilant, Gordon J. Wade, and A. Ted Watson. Estimating parameters in scientific computation. *IEEE Computational Science & Engineering*, 1(3):19–31, 1994.

[18] G. S. Feitosa, Lifu Chu, Leslie Thompson, and Albert C. Reynolds. Determination of permeability distribution from well-test pressure data. *Journal Petroleum Technology*, pages 607–615, July 1994.

[19] Roger Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, 1987.

[20] G. R. Gavalas, P. C. Shah, and John H. Seinfeld. Reservoir history matching by Bayesian estimation. *Soc. Petrol. Eng. J.*, 16(6):337–350, 1976.

[21] S. Geman and D. Geman. Reservoir history matching by Bayesian estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI–6(6):721–74, 1984.

[22] J. Jaime Gómez-Hernández and André G. Journel. Joint sequential simulation of multigaussian fields. In A. Soares, editor, *Geostatistic Troia 92*, pages 133–144. 1992.

[23] N. He, A. C. Reynolds, and D. S. Oliver. Three-dimensional reservoir description from multiwell pressure data and prior information. *Soc. Pet. Eng. J.*, pages 312–327, Sept. 1997.

[24] Nanqun He, Albert C. Reynolds, and Dean S. Oliver. Three-dimensional reservoir description from multiwell pressure data and prior information (SPE-36509). In *1996 SPE Annual Technical Conference and Exhibition*, 1996.

[25] Bjørn Kåre Hegstad and Henning Omre. Uncertainty assessment in history matching. In *5th International Geostatistics Congress*, 1996.

[26] Bjørn Kåre Hegstad, Henning Omre, Håkon Tjelmeland, and Kelly Tyler. Stochastic simulation and conditioning by annealing in reservoir description. In M. Armstrong and P. A. Dowd, editors, *Geostatistical Simulation*, pages 43–55. Kluwer Acad., 1993.

[27] P. Jacquard. Théorie de l'interprétation des mesures de pression. *Revue de L'Institut Français du Pétrole*, 19(3):297–334, 1964.

[28] P. Jacquard and C. Jain. Permeability distribution from field pressure data. *Soc. Petrol. Eng. J.*, 5(4):281–294, 1965.

[29] Hans O. Jahns. A rapid method for obtaining a two-dimensional reservoir description from well pressure response data. *Soc. Petrol. Eng. J.*, 6(12):315–327, 1966.

[30] B. L. N. Kennett and P. R. Williamson. Generalized subspace methods for large scale nonlinear inversion. In N. J. Vlarr et al., editor, *Mathematical Geophysics: A Survey of Recent Developments in Seismology and Geodynamics*, pages 139–154. Reidel Publishing Company, Dordrecht, 1988.

[31] J. E. Killough, Y. Sharma, A. Dupuy, and R. Bissell. A multiple right hand side iterative solver for history matching SPE 29119. In *Proceedings of the 13th SPE Symposium on Reservoir Simulation*, pages 249–255, 1995.

[32] Peter K. Kitanidis. Quasi-linear geostatistical theory for inversing. *Water Resour. Res.*, 31(10):2411–2419, 1995.

[33] T. Y. Lee and John H. Seinfeld. Estimation of two-phase petroleum reservoir properties by regularization. *J. Computational Physics*, 69:397–419, 1987.

[34] Eliana M. Makhlouf, Wen H. Chen, Mel L. Wasserman, and John H. Seinfeld. A general history matching algorithm for three-phase, three-dimensional petroleum reservoirs. *SPE Advanced Technology Series*, 1(2):83–91, 1993.

[35] William Menke. *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, San Diego, revised edition, 1989.

[36] D. W. Oldenburg and Y. Li. Subspace linear inverse method. *Inverse Problems*, 10:915–935, 1993.

[37] D. W. Oldenburg, P. R. McGillivray, and R. G. Ellis. Generalized subspace methods for large-scale inverse problems. *Geophys. J. Int.*, 114:12–20, 1993.

[38] Dean S. Oliver. The averaging process in permeability estimation from well-test data. *SPE Formation Evaluation*, 5(3):319–324, 1990.

[39] Dean S. Oliver. The incorporation of transient pressure data into reservoir characterization. *In Situ*, 18(3):243–275, 1994.

[40] Dean S. Oliver. On conditional simulation to inaccurate data. *Mathematical Geology*, 28(6):811–817, 1996.

[41] Dean S. Oliver, Luciane B. Cunha, and Albert C. Reynolds. Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology*, 29(1):61–91, 1997.

[42] Dean S. Oliver, Nanqun He, and Albert C. Reynolds. Conditioning permeability fields to pressure data. In *European Conference for the Mathematics of Oil Recovery*, pages 1–11, September 1996.

[43] A. Ouenes, B. Bréfort, G. Meunier, and S. Dupéré. A new algorithm for automatic history matching: Application of simulated annealing method (SAM) to reservoir inverse modeling. Unsolicited manuscript SPE 26297, 1993.

[44] Robert L. Parker. *Geophysical Inverse Theory*. Princeton University Press, Princeton, New Jersey, 1994.

[45] D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation with non-square grid blocks and anisotropic permeability. *Soc. Pet. Eng. J.*, pages 531–543, June, 1983.

[46] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Cambridge University Press, Cambridge, England, 1992.

[47] Albert C. Reynolds, Nanqun He, Lifu Chu, and Dean S. Oliver. Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. *Soc. Petrol. Eng. J.*, 1(4):413–426, 1996.

[48] Albert C. Reynolds, Nanqun He, and Dean S. Oliver. Reducing uncertainty in geostatistical description with well testing pressure data. In *Fourth International Reservoir Characterization Technical Conference*, 2–4 March 1997.

[49] R. K. Sagar, B. G. Kelkar, and L. G. Thompson. Reservoir description by integration of well test data and spatial statistics. In *Proceedings from the 68th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers*, pages 475–489, October 1993.

[50] P. C. Shah, G. R. Gavalas, and J. H. Seinfeld. Error analysis in history matching: The optimum level of parameterization. *Soc. Petrol. Eng. J.*, 18(6):219–228, 1978.

[51] Y. N. Tang, Y. M. Chen, W. H. Chen, and M. L. Wasserman. Generalized pulse-spectrum technique for 2-D and 2-phase history matching. *Applied Numerical Mathematics*, 5(6):529–539, 1989.

[52] Albert Tarantola. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, Amsterdam, The Netherlands, 1987.

[53] A. N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Math. Dokl.*, 4:1624–1627, 1963.

[54] M. L. Wasserman, A. S. Emanual, and J. H. Seinfeld. Practical applications of optimal-control theory to history-matching multiphase simulator models. *Soc. Pet. Eng. J.*, pages 347–354, 1975.

[55] W. Xu, T. T. Tran, R. M. Srivastava, and A. G. Journel. Integrating seismic data in reservoir modeling: the collocated cokriging approach, (SPE-24742). In *1992 SPE Annual Technical Conference and Exhibition*, 1992.

[56] P. Yang, A. T. Watson, and R. V. Armasu. Automatic history matching with variable metric methods. *SPE Res. Eng. J.*, pages 995–1001, August, 1988.

[57] William W-G. Yeh. Review of parameter identification in groundwater hydrology: The inverse problem. *Water Resources Research*, 22(2):95–108, 1986.