Report Title:	**Mapping of Reservoir Properties and Facies Through Integration of Static and Dynamic Data**
Report Type:	Annual Technical Report
Reporting Period Start Date:	October 1, 2002
Reporting Period End Date:	September 30, 2003
Principal Authors:	Albert C. Reynolds, Dean S. Oliver, Yannong Dong, Ning Liu, Guohua Gao & Ruijian Li
Date Report Issued:	March 2004
DOE Award Number:	DE-FC26-00BC15309

Petroleum Engineering Department
The University of Tulsa
600 South College Avenue
Tulsa, Oklahoma 74104

Contents

List of Figures

ABSTRACT

Knowledge of the distribution of permeability and porosity in a reservoir is necessary for the prediction of future oil production, estimation of the location of bypassed oil, and optimization of reservoir management. But while the volume of data that can potentially provide information on reservoir architecture and fluid distributions has increased enormously in the past decade, it is not yet possible to make use of all the available data in an integrated fashion. While it is relatively easy to generate plausible reservoir models that honor static data such as core, log, and seismic data, it is far more difficult to generate plausible reservoir models that honor dynamic data such as transient pressures, saturations, and flow rates. As a result, the uncertainty in reservoir properties is higher than it could be and reservoir management can not be optimized. The goal of this project is to develop computationally efficient automatic history matching techniques for generating geologically plausible reservoir models which honor both static and dynamic data. Solution of this problem is necessary for the quantification of uncertainty in future reservoir performance predictions and for the optimization of reservoir management.

Facies (defined here as regions of relatively uniform petrophysical properties) are common features of all reservoirs. Because the flow properties of the various facies can vary greatly, knowledge of the location of facies boundaries is of utmost importance for the prediction of reservoir performance and for the optimization of reservoir management. When the boundaries between facies are fairly well known, but flow properties are poorly known, the average properties for all facies can be determined using traditional techniques. Traditional history matching honors dynamic data by adjusting petrophysical properties in large areas, but in the process of adjusting the reservoir model ignores the static data and often results in implausible reservoir models. In general, boundary locations, average permeability and porosity, relative permeability curves, and local flow properties may all need to be adjusted to achieve a plausible reservoir model that honors all data. In this project, we will characterize the distribution of geologic facies as an indicator random field, making use of the tools of geostatistics as well as the tools of inverse and probability theory for data integration.

EXECUTIVE SUMMARY

Bayesian statistics provides a framework for the automatic history matching of multiphase flow production data to construct estimates or realizations of reservoir properties that are consistent with time-lapse seismic data, production data and static data obtained from logs, cores and geologic and geophysical interpretation. The automatic history matching procedure developed and implemented in our work requires the minimization of an objective function which consists of the sum of a regularization term and production data mismatch terms. The regularization term represents a geostatistical model constructed from static data. If the number of production data is small or the number of reservoir variables to be estimated is small, we showed in our previous DOE annual report on this project that a modified Levenberg-Marquardt algorithm or a Gauss-Newton method can be applied to minimize the appropriate objective function. The Levenberg-Marquardt algorithm and Gauss-Newton method, however, require the computation of individual sensitivity coefficients and this is not computationally feasible for large scale problems where the number of production data to be matched is greater than a few hundred and the number of reservoir variables is on the order of a few thousand to tens of thousands. In previous reports on this project, we discussed an implementation of a scaled limited memory Broyden-Fletcher-Goldfard-Shanno (LBFGS) algorithm and showed that it is a robust and computationally efficient algorithm for large scale history matching problems.

Since our last report, we have modified our history matching code so that we can condition a reservoir model to both production and time lapse seismic data. The history matching procedure enables us to construct estimates of gridblock porosities and log-permeabilities. Gridblock log-permeabilities are converted to permeabilities when running the simulator. The reason for constructing estimates of log-permeabilities is that the theoretical justification of our approach to automatic history matching is based on Bayesian statistics and in this setting the permeability fields are assumed to be log-normal. The rock properties estimated are referred to throughout as model parameters.

Because history matching is an inherently ill-conditioned problems, overshooting and

undershooting often occurs. Overshooting refers to obtaining an estimate of a model parameter that is unreasonably high and undershooting refers to obtaining an estimate of a model parameter that is unreasonably low. To control undershooting and overshooting, we have added the option of applying constraints in our LBFGS algorithm. We apply a constrained LBFGS algorithm to to the Tengiz field example and show that the method controls the overshooting we experienced previously. The Tengiz reservoir is an undersaturated reservoir produced by forty four wells. In this example, we estimated permeability fields on a $59 \times 49 \times 9$ upscaled reservoir simulation grid by conditioning the permeability field generated from a geostatistical model to pressure buildup data by automatic history matching of buildup data. By application of the constrained optimization algorithm we found that the reported pressure and rate data were inconsistent and are currently analyzing the data to try to properly account for these inconsistencies.

In the chapter entitled "Time Lapse Seismic Data", we discuss the status of our work on incorporating time-lapse seismic data. As time-lapse seismic data covers the areal extent of the reservoir and is related to vertically averaged pressure and fluid saturations, it is clear that matching these seismic data, simultaneously with production data, will reduce the uncertainty in our estimates of the rock property fields. Applying a gradient based optimization algorithm to condition rock property fields to seismic data requires the calculation of sensitivity of the seismic data to the rock property fields. In this report, two synthetic reservoir examples are considered to investigate our proposed approach for integrating seismic data; one example pertains to waterflooding an oil reservoir and the second is a large three-dimensional model based on a solution-gas drive reservoir in the middle east. In both cases, seismic impedance data is available at two or more times where the distribution of phases saturations in the reservoir are quite different so useful time-lapse impedance data can be constructed. We have implemented the adjoint procedure to calculate the sensitivity of seismic impedance to the permeability and porosity fields. The gradient is used in the limited memory BFGS algorithm to estimate rock property fields by history matching production data and seismic data simultaneously.

Because we will always use a reservoir simulator to calculate production data for a given reservoir description, it is convenient to include reservoir simulator gridblock log-permeabilities and porosities in the set of reservoir variables to be estimated by history-matching. In many reservoirs, however, the primary control on the distribution of permeability and porosity is the facies; typically the variation of rock properties between facies is much larger than the variation within a facies. Because of this, it is critical to estimate facies boundaries in the history matching process. In this work, we have developed a truncated

plurigaussian model for the generation of facies maps. Unlike previous implementations of this method, we have used intersecting lines as thresholds. With this approach, we show that it is still possible to (1) generate a rich variety of textures and shapes, (2) estimate the locations of the threshold lines, (3) generate approximations of the sensitivity coefficients needed to condition reservoir models to facies distributions, and (4) implement the new method into the existing history-matching code.

We have also introduced uncertainty into the relative permeability curves under the assumption that relative permeability curves can be described by power law models. For some two-dimensional flow cases, we have shown that reasonable estimates of relative permeability curves can be constructed from three-phase flow production data if a prior model for the parameters defining the power law curves is available from laboratory derived relative permeability curves.

INTRODUCTION

Automatic history matching is based on minimizing an objective function which includes a sum of squared production data mismatch terms. Typically, minimization is done based on a derivative based optimization routine, such as the Gauss-Newton and Levenberg-Marquardt algorithms, because algorithms which do not use derivative information converge too slowly for practical applications. We refer to the reservoir variables to be estimated as model parameters. The model parameters typically represent reservoir simulator gridblock porosities and permeabilities (or log-permeabilities), may also represent well skin factors, transmissibility multipliers, parameters describing relative permeability curves or parameters describing facies boundaries. Standard implementations of the Gauss-Newton method or Levenberg-Marquardt algorithm require calculation of sensitivity coefficients, which formally represent the derivative of predicted production data with respect to the model parameters.

For automatic history matching problems of interest to us, the number of model parameters is greater than the number of independent production data and thus the history matching problem does not have a unique solution. If the Gauss-Newton procedure is applied to minimize an objective function consisting of only the sum of squared production data misfit terms, the Hessian matrix will be singular and the optimization algorithm will be unstable. This instability problem can be avoided by adding a regularization term to the objective function to be minimized; see Tikhonov (1963) and Parker (1994). With a proper regularization, the Hessian matrix in the Gauss-Newton method will be real symmetric positive definite and hence nonsingular. In this work, we use a prior geostatistical model to provide regularization. With this approach, the history matching problem is equivalent to a Bayesian estimation problem (Gavalas et al., 1976; Tarantola, 1987; He et al., 1997; Wu et al., 1999).

The Gauss-Newton method is popular because it converges quadratically in the neighborhood of a minimum; see, for example, Fletcher (1987). Sometimes, however, if the initial guess in the Gauss-Newton method results in a large initial data mismatch, the Gauss-Newton will converge to a reservoir model which represents a local minimum and does not

give an acceptable match of production data; see, Wu et al. (1999)and Li et al. (2001a). For this reason, we often apply a Levenberg-Marquardt algorithm instead of the Gauss-Newton method if the production data misfit based on the initial reservoir model is very large.

Automatic history matching traces its roots to research conducted in the 1960's by Jacquard (1964), Jacquard and Jain (1965) and Jahns (1966). To the best of our knowledge, Jacquard and Jain (1965) 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 obtained under single-phase flow conditions. They used a combination of zonation (less than twenty distinct values) and an algorithm conceptually similar to the Levenberg-Marquardt algorithm to provide regularization. Jahns estimated transmissibility (kh/μ) values and storativity ($\phi c_t h$) simultaneously by history matching single-phase flow pressure data. He used the finite difference method to compute sensitivity coefficients and applied the Gauss-Newton method with an exact line search to estimate the rock property fields by minimizing an objective function consisting only of the sum of squared pressure mismatch terms. Zonation was used to provide regularization. Jahns actually used a sequence of minimization steps where the number of zones, and hence the number of parameters was increased at each minimization step. The maximum number of parameters estimated was nine, eight zonal transmissibilities (or permeabilities) and total storativity. The finite difference method used to compute sensitivities requires $N_m + 1$ runs of the simulator where N_m is the number of model parameters estimated. This procedure would not be feasible when thousands of model parameters are estimated.

Jacquard and Jain (1965) based their procedure for computing sensitivity coefficients on an electric-circuit analogue. Later, motivated by Jacquard and Jain's ideas, Carter et al. (1974) presented an elegant derivation of a method to compute sensitivity coefficients for two-dimensional single-phase flow problems. As originally presented, the procedure of Carter et al. can be applied to compute the sensitivity of simulator gridblock pressures to all gridblock permeabilities and porosities. If each well penetrates only a single gridblock, one can compute the sensitivity of the wellbore pressure to model parameters from the well's gridblock pressure sensitivities, provided the simulator uses a formula like the one of Peaceman (1978) to relate wellbore pressure and gridblock pressure. For two-dimensional single-phase flow problems with pressure measurements at N_w wells, this procedure requires $N_w + 1$ reservoir simulation runs to compute all sensitivity coefficients regardless of the number of model parameters and regardless of the number of pressure data. For three-dimensional problems, the number of simulation runs required would be equal to one plus the number of gridblocks penetrated by wells. If the number of such gridblocks is large, the Carter et al. procedure becomes

less computationally attractive. However, He et al. (1996) have developed an approximate three-dimensional version of the Carter et al. method which is computationally efficient. Regardless of the number of gridblocks penetrated by wells, the method of He et al. requires only $N_w + 1$ reservoir simulation runs to compute the sensitivity of all well pressure data to all gridblock permeabilities and porosities. The method is only approximate and does not always yield accurate results if vertical flow is significant in gridblocks penetrated by wells. It is not clear, however, whether the fact that the sensitivity coefficients are approximate leads to significant errors in the history matching process.

For nonlinear problems, e.g., multiphase flow problems, the derivations of Carter et al. (1974) and He et al. (1996) 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. (1974) and Chavent et al. (1975). (For single-phase flow problems, Carter et al. (1982) have shown that their method is equivalent to the adjoint method.)

Unlike the Carter et al. (1982) method, however, the adjoint method can be applied to compute sensitivity coefficients in multiphase flow problems. Unfortunately, the procedure requires N_d adjoint solutions where N_d is the number of production data to be history matched. The sensitivities can be calculated easily once the adjoint variables have been computed. Solving an adjoint problem is similar to solving the simulation finite difference equations with two distinct differences: (i) to find the adjoint variables needed to computed the sensitivity of a particular production data at the time t_l, the appropriate adjoint problem is solved backward in time, from time t_l to time zero; (ii) unlike the forward problem (simulator problem), the adjoint problem is linear. At each time step in the adjoint solution, a matrix problem is solved. The coefficient matrix is independent of the production data but the right hand side of the matrix problem is determined directly from the specific production data. If N_d production data are uniformly spaced in time and the final time at which we have measured production data is t_L, then computing all adjoint solutions needed to compute the sensitivities of all production data effectively requires solving a sequence of matrix problems related to solving the adjoint problem backward in time from t_L. At each time step, the matrix problem is solved with an average of $N_d/2$ right-hand side vectors; see Wu et al. (1999) for additional discussion. Even if one uses a procedure based on solving a matrix problem with multiple right hand sides, the solution of the adjoint systems needed to compute sensitivities for N_d production data will not be feasible when N_d is large. If one assumes that solving the adjoint matrix systems with an average of $N_d/2$ right hand side vectors is equivalent to at least $(0.1N_d)/2$ simulation runs, the number of equivalent

simulation runs required is prohibitive if there are several hundred production data to be matched. Because of this, the adjoint method traditionally has been used only in conjunction with optimization methods which require only the gradient of the objective function, e.g., conjugate gradient or variable metric methods; see, for example, Wasserman et al. (1975), yong Lee and Seinfeld (1987); Lee and Seinfeld (1987), Yang and Watson (1988), Makhlouf et al. (1993). Computation of the gradient of the objective function requires only the solution of a single adjoint system and thus requires no more computational time than one reservoir simulation run. Unfortunately, the implementations of these methods have resulted in slow convergence. For example, Makhlouf et al. (1993) reported that history matching a two-phase flow 450 cell reservoir model required 6400 CPU seconds on a CRAY X-MP/48. In their work, a conjugate gradient method was used as the optimization algorithm. For one three-phase flow problem with 450 grid blocks, 222 iterations of the conjugate gradient algorithm were required to obtain convergence.

Largely because of the results of Makhlouf et al. (1993), until recently, our work on automatic history matching has focused on using the Gauss-Newton and Levenberg-Marquardt algorithms instead of conjugate gradient or variable metric algorithms. Wu et al. (1999) were the first to use the adjoint method in conjunction with the Gauss-Newton method to perform history matching. They implemented the adjoint method to compute the sensitivity of all production data to gridblock permeabilities and porosities. In their work, they constructed estimates and realizations of permeability and porosity fields by conditioning a prior geostatistical model to pressure and water-oil ratio data. They considered only two-dimensional, two-phase flow (water-oil) systems. As noted in our previous DOE annual report on this process, we have extended the Gauss-Newton and Levenberg-Marquardt algorithms in conjunction with the adjoint method for sensitivity calculation to the problem of history matching production data for three-dimensional, three-phase flow problems. The resulting history matching process, however, is not practically feasible for problems of interest to us where both the number of data and the number of model parameters exceed a few hundred.

Perhaps because it is simple to implement, the so-called gradient method is frequently used to compute sensitivity coefficients needed for automatic history matching. This method was introduced into the petroleum engineering literature by Anterion et al. (1989), but was known earlier in the ground water hydrology literature as the sensitivity coefficient method; see, for example, the review of parameter identification methods by Yeh (1986). In this procedure, the sensitivity of pressures and saturations to model parameters at the end of a simulator time-step can be obtained by solving a matrix problem 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. From the pressure and saturation sensitivities, one can easily construct other sensitivity coefficients, e.g., the sensitivity of gas-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. With the fast iterative solver developed by Killough et al. (1995), it appears that the computational time to compute a single sensitivity coefficient is on the order of 10% of a forward simulation. For the gradient simulator to be practical, the number of model parameters must be small. This means, if the underlying reservoir simulation problem involves tens of thousands of gridblocks, one must reduce the number of parameters estimated directly in the optimization algorithm by some form of reparameterization, e.g., zonation (Jacquard and Jain, 1965) or gradzones (Bissell et al., 1994; Bissell, 1994; Tan, 1995), pilot points (de Marsily et al., 1984; RamaRao et al., 1995; Bissell et al., 1997) or subspace methods (Kennett and Williamson, 1988; Oldenburg et al., 1993; Reynolds et al., 1996; Abacioglu et al., 2000).

When the number of model parameters and number of production data to be matched are both large and can not be reduced by some reparameterization technique without incurring a significant loss of information, one must seek an alternative to computing and storing the full sensitivity coefficient matrix, G. One can write the Gauss-Newton method such that each iteration requires the solution of an $N_d \times N_d$ matrix problem where N_d represents the number of production data to be matched. If this matrix problem is solved by a conjugate gradient method (Hestenes and Stiefel, 1952; Fletcher and Reeves, 1964; Shanno, 1978a,b), the explicit direct computation of G is not required. Each iteration requires only the product of G times a vector and the product of the transpose of G times a vector. A procedure for computing these matrix vector products without first computing G was introduced into the petroleum engineering literature by Chu et al. (2000) although the basic idea appeared earlier in a somewhat simpler context in the geophysics literature; see Mackie and Madden (1993). Although computation of the matrix products is relatively efficient, the conjugate gradient

method may require up to N_d iterations to obtain convergence if the matrix is poorly conditioned and no good preconditioning matrix is available (see Axelsson (1994)). If N_d is large, this would render the algorithm impractical. Although the convergence of the conjugate gradient method can be considerably accelerated by the choice of a good preconditioner, it is not clear that one can construct a good preconditioner since the coefficient matrix for the matrix problem that is solved is not explicitly computed. To compute this coefficient matrix would require the explicit computation of the full sensitivity coefficient matrix G.

One can also avoid explicit computation of all sensitivity coefficients if history matching is done using a nonlinear optimization method, that requires only the gradient of the objective function. As mentioned previously, Makhlouf et al. (1993) found that a nonlinear conjugate gradient algorithm could require over two hundred iterations to converge even for a small three-phase flow history matching problem. As each conjugate gradient iteration requires roughly the equivalent of three reservoir simulation runs, history matching a large problem using a nonlinear conjugate gradient method does not appear to be feasible based on the results of Makhlouf et al. (1993). However, Makhlouf et al. (1993) did not apply preconditioning. If a good preconditioning matrix can be found for nonlinear conjugate gradients, it is conceivable that convergence could be considerably accelerated.

Quasi-Newton or variable metric methods, which are based on generating an approximation to the inverse of the Hessian matrix, require only the gradient of the objective function. The methods differ in how they correct or update the inverse Hessian approximation at each iteration. The rank one correction formula was first suggested by Broyden (1967). Another formula, now called the DFP algorithm, was first suggested by Davidon in 1959 and later presented by Fletcher and Powell (1963). The BFGS correction formula, suggested independently by Broyden (1970), Fletcher (1970), Goldfarb (1970) and Shanno (1970), and several variants of the BFGS formula (like the self-scaling variable metric (SSVM) by Oren (1973), limited memory BFGS by Nocedal (1980) and Liu and Nocedal (1989)) have also been advanced as useful variable metric methods.

The conjugate gradient method was originally proposed by Hestenes and Stiefel (1952) for solving linear systems and extended to nonlinear optimization by Fletcher and Reeves (1964) to obtain the Fletcher-Reeves algorithm. Later Polak (1971) proposed a different formula to calculate the coefficient involved in the search direction update equation. Powell (1977) presented some numerical results and theoretical reasons which indicate that the Polak-Ribière algorithm is superior to the Fletcher-Reeves algorithm. The efficiency of the conjugate gradient method depends primarily on the preconditioner used.

The limited memory BFGS (LBFGS) was designed for the purpose of solving large scale

problems which involve thousands of variables. Limited memory methods originated with the work of Shanno (1978a), and were subsequently developed and analyzed by Buckley (1978), Nazareth (1979), Nocedal (1980), Shanno (1978b), and Buckley and Lenir (1983). Liu and Nocedal (1989), and Nash and Nocedal (1991) tested LBFGS method with a set of problems. They concluded that LBFGS performs better than conjugate gradient in terms of computational efficiency, except in cases where the function evaluation is inexpensive. Nash and Nocedal (1991) also tested a truncated-Newton method in their work. From their comparison, none of the algorithms is clearly superior to the other.

The self-scaling variable metric (SSVM) method was used by Yang and Watson (1988) on hypothetical water floods of both 1D and 2D reservoir models. The 1D reservoir model consisted of 10 gridblocks with an injection well at one end and a producing well at the other end. Sixty data from each well were used for history matching. Four cases based on this 1D reservoir model were tested. The reservoir was characterized by different parameters in different cases. The number of model parameters varied from 9 to 19. Two other cases were based on a quarter of a five-spot 2D model which consisted of a 10×10 grid. Again sixty data from each well were history matched. The number of model parameters for these two cases were 4 and 11 respectively. In this paper, the authors tested four different algorithms, BFGS, SSVM, conjugate gradient and steepest descent. They concluded that (i) the self-scaling variable metric method is significantly more efficient than the BFGS method; (ii) the SSVM and BFGS methods are more efficient and robust than the conjugate gradient method, except in the case where the objective function is nearly quadratic; and (iii) both SSVM and BFGS methods perform significantly better than the steepest descent method.

Masumoto (2000) applied the SSVM method to a water-oil two phase fluid flow problem. The author considered a 1D reservoir model with 20 gridblocks. With a fixed porosity field, the author estimated the gridblock permeabilities. The objective function he minimized included a pressure mismatch part and the pressure derivative mismatch part. The author did not give any information about how many data he history matched or any assessment of the minimization algorithm. Savioli and Grattoni (1992) compared four different minimization algorithms: Davidon-Fletcher-Powell (DFP), Fletcher-Reeves (FR), BFGS and Levenberg-Marquardt (LM). The authors presented two examples. In the first example, they estimated one permeability value and one porosity value by applying these four algorithms. The second example they considered was an oil-water two phase water flooding problem. They estimated the exponent used to define the relative permeability and capillary curves with a power law function (only one adjustable parameter for each curve). They concluded that among these four algorithms, BFGS performed best in terms of computational efficiency and stability.

Given the small number of parameters estimated, it is difficult to know whether these results will extrapolate to large scale problems.

As noted in our previous annual report, all computations we have done suggest that the limited memory BFGS (LBFGS) procedure is the most viable optimization algorithm for automatic history matching of multiphase flow production data for problems where the number of reservoir variables to be estimated and the number of data are both large. Thus, unless specifically noted otherwise, the reader should assume that the history matching examples presented in this report were done with the LBFGS algorithm. In particular, the Tengiz true field example as well as the synthetic Tengiz examples were history matched with the LBFGS algorithm.

Time-lapse seismic is the process of repeating 3D seismic surveys over a producing reservoir to monitor changes in saturation and pressure. The potential impact on reservoir engineering and reservoir management is large because time-lapse seismic may allow direct imaging of rock properties that are closely related to vertically averaged fluid saturations and pressure. This is much different from the current limitation of measurements of these quantities at well locations. In general, seismic images are sensitive to the spatial variation of two distinct types of reservoir properties (Arenas et al., 2001):

- Non-time-varying *static* geologic properties such as lithology, porosity, cementation, and shale content.
- Time-varying *dynamic* fluid-flow properties such as fluid saturation and pore pressure.

If data were available from only one 3D seismic survey, it would not be possible to differentiate between the effects of static features and those due to changes in saturation and pressure. By comparing the data from 3D surveys acquired at different times in the same location, however, it is possible to eliminate the effects of unknown static properties to focus on the dynamic changes in production related properties.

The simplest, most direct method of using time-lapse seismic data is to qualitatively monitor reservoir changes due to production. In this approach, one simply identifies regions in which the amplitude or impedance has changed with time and attributes these changes to changes in saturation, pressure, or temperature. The first tests of this concept were carried out by Arco in the Holt Sand fireflood from 1981 to 1983 (Hughes, 1998). Similar studies have been reported by Cooper et al. (1999) at the Foinhaven Field and Lumley et al. (1999) at the Meren Field in Nigeria. The primary objectives at Foinhaven were simply to map fluid movements and to identify by-passed oil. The authors of the study concluded that the time-lapse signal qualitatively agreed with the expected reservoir performance. At Meren,

the goal was to identify pathways of injected water, sealing faults, and compartments that may have by-passed oil. The authors concluded that the data was successful in achieving these objectives.

The other, more difficult, approach is to use the time-lapse data to estimate the reservoir flow parameters, such as permeability and porosity. Advances in automatic or computer-assisted history matching have allowed researchers to consider the integration of time-lapse seismic data with production data. All quantitative approaches for doing this involve the minimization of an objective function that includes the mismatch between the synthetic changes in seismic data and the observed changes. Using optimization methods, a distribution of parameters that minimize the objective function is sought. The type of seismic data used in the objective function has varied among the researchers. Huang et al. (1997) used amplitude difference or other seismic attributes difference while Arenas et al. (2001) used velocity difference. Landa and Horne (1997) assumed that saturation changes could be obtained directly from time-lapse surveys.

While a number of geophysicists (Tura and D.Lumley, 1999; Landro, 2001; Meadows, 2001) have proposed estimation of saturation and pressure directly from time-lapse seismic data (including amplitude variation with offset data), it is clearly less restrictive to use all data (including production data) in the estimation of saturation and pressure to ensure that the saturations and pressures are consistent with material balance. Thus we use the seismic data in the objective function—not saturations and pressures.

In studies to date, the sensitivity of time-lapse seismic data to changes in model parameters has either been computed by the finite-difference method (Huang et al., 1997, 1998, 2001; van Ditzhuijzen et al., 2001) or the gradient simulator method (Landa and Horne, 1997). It is not feasible to compute sensitivity coefficients using either of these methods when the number of model parameters is large, however. The only reasonable approach is to use the adjoint method to integrate seismic impedance data into our objective function and to compute the sensitivity of data to model parameters. It will also be necessary to use more efficient optimization methods in the history matching than those used previously to get optimum model parameters. A practical method for doing this will be outlined in this document.

Researchers have been building tools for history matching of permeability and porosity distributions to honor production data for several years. The assumption is almost always made that the rock properties are distributed randomly and that the randomness can be adequately described by the mean and the spatial covariance of the property fields. If there is more than one type of rock, region or facies, the assumption is usually made that the

location of the boundaries of these regions is known. Bi et al. (2000) and Zhang et al. (2002) relaxed this restriction by allowing the boundaries of a three dimensional channel to be adjusted interactively during the history matching process. While the method worked quite well for a single channel in a background low-permeability facies, it became apparent that the extension to a reservoir with large numbers of channels would be impractically difficult.

As a result, we began consideration of the truncated plurigaussian model for the description of facies boundaries. The truncated plurigaussian is attractive for modeling facies for several reasons.

1. The model is capable of generating a wide variety of facies shapes and neighbor relations.

2. The model is based on Gaussian random fields, which are well-suited to current in history matching codes.

3. The truncation, or threshold maps, can be described by relatively few parameters.

In this report, we describe progress on two aspects of the history matching problem. The first problem has to do with the specification of a prior geostatistical model, the purpose of which is to ensure plausibility of realizations. This is considerably more complex for the truncated plurigaussian model than for many geostatistical models because it is necessary to specify at least two covariance models (types, ranges, variances, and orientations), as well as the threshold parameters for the truncation. The second problem is adjustment of the facies boundaries for a fixed set of geostatistical model parameters. This requires efficient minimization of an objective function that is not differentiable.

In this report, we also consider the simultaneous estimation of the absolute permeability field and relative permeability curves from three-phase flow production data. Irreducible water saturation, critical gas saturation and residual oil saturations are assumed to be known. The two-phase relative permeability curves for an oil-gas system and the two-phase relative permeability curves for an oil-water system are represented by power law models. The three-phase oil relative permeability curve is calculated from the two sets of two-phase curves using Stone's Model II. The adjoint method is applied to three-phase flow problems to calculate the sensitivity of production data to the absolute permeability field and the parameters defining the relative permeability functions. Using the calculated sensitivity coefficients, absolute permeability and relative permeability fields are estimated by automatic history matching of production data. As the example problems considered here are relatively small, we apply the Levenberg-Marquardt algorithm to do history matching. As the detailed history given

below indicates, this problem is not new, although to the best of our knowledge, this is the first attempt to three-phase relative permeabilities by history matching.

It appears that Archer and Wong (1973) were the first authors to consider the estimation of relative permeability curves by applying a reservoir simulator to history match laboratory core flood data. They estimated only parameters that define the shape of relative permeability curves for simple empirical relative permeability models and adjusted relative permeabilities by a trial and error method during the history matching.

Sigmund and McCaffery (1979) were the first to apply nonlinear regression to the problem of history matching laboratory core flood data. They used power law expressions to model relative permeability curves and estimated only the two exponential parameters in these formulas. Kerig and Watson (1986) considered a similar problem. They calculated predicted data from the Buckley-Leverett model, used cubic splines to parameterize relative permeability curves and compared relative permeability estimates obtained with such a representation to those obtained using a power law functional form. They showed that, in general, power law models do not contain enough degrees of freedom to represent the truth well, whereas cubic splines with a small number of knots appear to be sufficiently flexible to yield more accurate estimates of true relative permeability curves. In their results, they assume absolute permeability is known. In a later paper, Kerig and Watson (1987) showed how to impose constraints to ensure that the estimated relative permeabilities are concave up (convex downward), nonnegative and monotonic. The Levenberg-Marquardt modification of the Gauss-Newton method was applied for optimization. With cubic spline representations, not all coefficients are independent. Kerig and Watson (1986) presented a procedure to determine the parameters that should be adjusted when matching core flood data. Later, Watson et al. (1988) avoided this difficulty by modeling each relative permeability function as a sum of B-splines. With B-splines, the independent adjustable parameters are simply the coefficients of the B-splines in the sum. They showed that with parabolic splines, it is easy to constrain the derivatives at the knots and endpoints to be nonnegative which automatically constrains the relative permeability curves to be monotonic, but does not constrain them to be concave up.

Lee and Seinfeld (1987) considered the simultaneous estimation of the absolute permeability field and relative permeabilities for a two-dimensional, two-phase flow oil-water system. They assumed power law relative permeability curves and assumed that the end point values of relative permeabilities were known. Thus only the two exponents in the power law relative permeability functions were estimated. They modeled the two-dimensional isotropic heterogeneous permeability field using bi-cubic B-splines. In the specific examples considered,

they matched pressure and water cut data at wells producing from an oil reservoir under waterflood. Tikhonov (1963) regularization was used to stabilize the nonlinear least squares problem. Matching of data was accomplished by a three-step process with the steepest descent algorithm applied for minimization of the objective function which includes the sum of squared data mismatch terms. In the application of steepest descent, the gradient of the objective function was calculated using the adjoint method (Chen et al., 1974; Chavent et al., 1975). In the first minimization step, regularization is not applied, but the permeability field is assumed homogeneous so only three parameters are estimated, absolute permeability and the two relative permeability parameters. In the second step, the maximum likelihood estimate of gridblock permeabilities and the two relative permeability parameters are estimated using the results of the first step as the initial guess. In the last step, the absolute permeability field and relative permeabilities (the two power law exponents) are estimated simultaneously with the degree of regularization estimated from results of the previous two steps. For the synthetic examples presented, the procedure presented was robust to the initial guesses and good estimates of absolute and relative permeability were obtained. For the examples considered, however, the permeability fields were smooth and the level of heterogeneity was low with the ratio of the maximum permeability to the minimum permeability equal to two.

Yang and Watson (1991) considered the estimation of relative permeability curves using a Bayesian approach with relative permeability functions modeled as a linear combination of B-splines. In this approach, the objective function to be minimized is the sum of two terms, a production data mismatch term and a term which measures the deviation from a prior relative permeability model. (A prior model would typically be developed from laboratory core floods or by analogy with similar reservoirs.) The authors presented a procedure to estimate the relative weighting of the two terms of the objective function so that estimated relative permeabilities will remain as close as possible to the prior model but still yield calculated data that gives an acceptable match of observed production data. They considered only homogeneous reservoirs and assumed that all physical properties except relative permeabilities were known. Minimization of the objective function was accomplished with a Broyden-Fletcher-Goldfarb-Shanno optimization algorithm. They illustrated their methodology by applying it to a synthetic two-dimensional, two-phase flow waterflooding problem with a single injection well and a single producing well. They matched pressure data at both wells and WOR data at the producing well. Relative permeabilities were constrained to be monotonic and bounded above by unity. For the example considered, water saturations at the producer corresponded to relatively low values. When conventional

history matching with no prior model was used, the estimated relative permeabilities were accurate only for relatively low water saturation values. For the Bayesian approach, the high saturation portion of the estimated curves was close to the prior relative permeability curves and the low saturation portion was resolved accurately by the production data.

Watson et al. (1980) considered the simultaneous estimation of porosity, absolute permeability and relative permeability by automatic history matching of production data, but restricted their application to two-dimensional oil-water systems where permeability and porosity are homogeneous. In these examples, they use power law relative permeabilities and estimated only the two exponents in these two power law functions, i.e., they assume end-point relative permeabilities are known.

Kulkarni and Datta-Gupta (2000) also consider the estimation of relative permeabilities from water-cut and pressure data for a two-phase flow oil-water system. Like earlier work, they modeled relative permeabilities with both B-splines and power law functions, but used a streamline simulator instead of a finite-difference simulator to perform automatic history matching. In one example, they showed that using only a prior model (zeroth order derivative) to provide regularization resulted in estimated relative permeability curves which exhibited changes in concavity, but if both zeroth order derivatives and second order derivatives were used in the regularization term, then the estimated relative permeability functions were non-oscillatory. Similar to the classical analytical results of Watson et al. (1984) on linear flow problems, they found that there is considerable uncertainty in estimated relative permeability curves when only water-cut data is matched but that the uncertainty is less when both pressure data and water cut data are used to constructs estimate. Adding regularization further reduced the uncertainty. They found that the uncertainty was greater for heterogeneous permeability fields than for a homogeneous permeability field. For heterogeneous permeability fields, they found that applying regularization and matching both water cut and pressure data are necessary to obtain reasonable estimates of both absolute permeability and the relative permeability parameters.

All of the papers discussed above considered only oil-water system. Labban and Horne (1991) considered the estimation of relative permeabilities for three-phase black-oil problems. They used stabilized inflow performance relations for multiphase flow problems to postulate linear relations between relative permeabilities and fluid properties and production data. The coefficients in these linear relations were determined by two simulation runs with two different relative permeability curves. They also assumed a power-law relation between phase saturation and phase relative permeability for two different simulation runs at equivalent times. With these relations, they estimate relative permeability at different times from

historical data and then use the saturation/relative permeability relation mentioned above to otain updated relative permeabilities as functions of phase saturations for the next simulation runs. They claim the procedure gives reasonable estimates of relative permeability. The procedure assumes that absolute permeability and porosity are known.

Li et al. (2001a) implemented a procedure to estimate absolute permeability fields by automatic history matching of three-phase flow production data. They assumed that relative permeabilities were known. To the best of our knowledge, the current paper is the first one that discusses the simultaneous estimation of absolute permeability fields and relative permeability curves under three-phase flow conditions. A power law form of relative permeability curves is used, but the procedure is general and could be applied using a B-spline representation of relative permeability curves. A prior model for absolute permeability and relative permeability parameters is assumed to provide regularization, i.e., Bayesian estimation is applied to generate estimates. Model parameters, which are estimated by automatic history matching of production data, consist of gridblock absolute log-permeabilities and the parameters defining the relative permeability curves. All results are presented in terms of oil field units.

EXPERIMENTAL

Experimental work is not applicable to the research tasks and goals of this project. Consequently, no experimental work has or will be done.

HISTORY MATCHING OF DATA

Model Estimation and Simulation

Here, we define the reservoir model parameters and the a posteriori probability density function (pdf) for these parameters. This pdf, which is conditional to observed data, determines the set of plausible reservoir descriptions. We discuss the computation of the maximum a posteriori (MAP) estimate of reservoir variables. The MAP estimate is the model which maximizes the a posteriori pdf and is thus conveniently referred to as the most probable model. A method for sampling this pdf to generate multiple realizations of reservoir variables is discussed briefly.

The Prior Model.

For simplicity, the reservoir is assumed to be a rectangular parallelepiped which occupies the region

$$\Omega = \{(x, y, z) \mid 0 < x < L_x, 0 < y < L_y, 0 < z < L_z\}. \tag{1}$$

The forward model is a fully-implicit finite-difference simulator based on a block centered grid. The principle permeability directions are assumed to be aligned with the coordinate directions so that the permeability tensor is diagonal. Fluid properties are assumed to be known. Given two-phase oil-water and two-phase oil-gas relative permeabilities, the three-phase oil relative permeability is constructed from Stone's Model II; see Aziz and Settari (1979). Wellbore constraints are handled using the equation of Peaceman (1983).

Here, model parameters or reservoir variables include only gridblock porosities, horizontal log-permeabilities, vertical log-permeabilities and the skin factor at each well. In a later chapter, we discuss our initial work on estimating the location of the boundaries between facies to match production data. We assume the permeability is areally isotropic so that $k_x = k_y$. Thus we present our equations for the case where the model parameters are simulator gridblock porosities, horizontal log-permeabilities, vertical log-permeabilities and

well skin factors. Thus, if there are N simulator gridblocks and N_w wells, the total number of model parameters is equal to $N_m = 3N + N_w$. Specifically, the vector of model parameters is given by

$$m = [m_\phi^T, m_k^T, m_{k_z}^T, m_s^T]^T, \tag{2}$$

where m_ϕ is an N-dimensional column with its jth entry equal to the porosity of gridblock j, where m_k is an N-dimensional column with its jth entry equal to the horizontal log-permeability for gridblock j, m_{k_z} is an N-dimensional column with its jth entry equal to the vertical log-permeability for gridblock j, and m_s is an N_w dimensional column vector with its jth entry given by the skin factor at the jth well. These reservoir parameters are modeled as random variables, so m is a random vector. From a purely history matching point of view, we wish to construct an estimate of m from production and time lapse seismic data (dynamic data) and static data. However, there are an infinite number of models which will give equally reasonable matches of the data, and it is desirable to define a procedure for generating a particular estimate or to characterize the uncertainty in reservoir descriptions. From both the philosophical and practical points of view (see Tarantola (1987) and Omre et al. (1993)), the most challenging part of the inverse problem is the determination of a representative pdf for reservoir parameters. Similar to the recent work on automatic history matching by He et al. (1997) and Wu et al. (1999), we follow ideas that can be found in Tarantola (1987) and simply assume that a prior geostatistical model for $m_r = [m_\phi, m_k^T, m_{k_z}^T]^T$ can be constructed from static data, e.g. log, core and other geologic data. In our work, we assume this prior geostatistical model can be represented by a multivariate Gaussian distribution for m_r with a given mean and covariance matrix. In practice, the prior covariance matrix for the rock property fields can be generated from semivariograms by assuming that horizontal, vertical permeability can be modeled as stationary random functions. In our implementation, we make this assumption and then apply the Xu et al. (1992) screening hypothesis to generate the prior covariance matrix for m_r; see, Chu et al. (1995b). In the prior model, each well skin factor is treated as an independent Gaussian variable with specified mean and variance. If the skin factor was estimated by fitting pressure data with a classical well testing model solution using nonlinear regression, then the estimate of the skin factor would be its prior mean and its variance can be constructed directly from the same information used to construct confidence intervals.

The vector of prior means is given by

$$m_{\text{prior}} = \begin{bmatrix} m_{\phi,\text{prior}} \\ m_{k,\text{prior}} \\ m_{k_z,\text{prior}} \\ m_{s,\text{prior}} \end{bmatrix}. \tag{3}$$

We let C_ϕ denote the prior covariance matrix for m_ϕ, C_k denote the prior covariance matrix for m_k, C_{k_z} denote the prior covariance for m_{k_z}, C_{k,k_z} denote the cross covariance matrix between m_k and m_{k_z} and let C_s denote the $N_w \times N_w$ model covariance matrix for the vector of well skin factors. Then the prior model covariance matrix is given by the $N_m \times N_m$ matrix

$$C_M = \begin{bmatrix} C_\phi & C_{\phi,k} & C_{\phi,k_z} & O \\ C_{\phi,k} & C_k & C_{k,k_z} & O \\ C_{\phi,k_z} & C_{k,k_z} & C_{k_z} & O \\ O & O & O & C_s \end{bmatrix}, \tag{4}$$

where the O's denote null submatrices of the appropriate size. If porosity is uncorrelated with horizontal and vertical permeability, then $C_{\phi,k}$ and C_{ϕ,k_z} are null matrices. If horizontal and vertical permeability are not correlated, then C_{k,k_z} is also a null matrix.

The prior pdf for m is then given by

$$\pi_p(m) = a \exp\Big\{ -\frac{1}{2}(m - m_{\text{prior}})^T C_M^{-1}(m - m_{\text{prior}}) \Big\}, \tag{5}$$

where a is the normalizing constant. Note the model which has the highest probability based on Eq. 5 is $m = m_{\text{prior}}$, thus it is convenient to think of m_{prior} as the best estimate of the model based on static data.

The a Posteriori Probability Density Function.

We wish to determine the conditional pdf for m given observed production data and/or time-lapse seismic data. Here, we consider only three types of production data, wellbore pressure (p_{wf}), producing water-oil ratio (WOR) and producing gas-oil ratio (GOR). The WOR and GOR data are not actually measured directly but are constructed from rate measurements. Nevertheless, we will refer to the values of WOR and GOR as measured or observed data. The column vector $d_{\text{obs,w}}$ contains all observed WOR data that will be used as conditioning data. The column vector $d_{\text{obs,g}}$ contains the set of GOR conditioning data and $d_{\text{obs,p}}$ contains all conditioning pressure data. Throughout, the N_d dimensional column vector d_{obs} includes all data that will be used to condition the model m. This may include one type of data, e.g.,

only GOR data or multiple types of data, e.g., pressure, WOR and GOR data and time-lapse seismic data. Although, we do not consider it here, d_{obs} may even include hard data, e.g., observed porosities at wells.

Pressure measurements errors are modeled as independent identically distributed Gaussian random variables with mean zero and variance σ_p^2. GOR measurement errors are modeled as independent identically distributed Gaussian random variables with mean zero and variance σ_g^2. WOR ratio measurement errors are modeled by the procedure introduced by Wu et al. (1999). In this model, the WOR measurement error depends on the magnitude of the measurement. Specifically, the variance of a particular measurement error is defined as

$$\text{Var}(e_{WOR}) = WOR_{\text{obs}}^2 \epsilon_o + \frac{1}{q_{o,\text{obs}}^2} \max\left[\epsilon_w^2 q_{\text{w,obs}}^2,\, \sigma_{\text{w,min}}^2\right], \tag{6}$$

where e_{WOR} denotes the error in the "measurement" of WOR constructed from the observed oil and water rates, $q_{o,\text{obs}}$ and $q_{\text{w,obs}}$. Here, ϵ_m denotes the relative measurement error for the flow rate of phase m. For example, if the relative measurement error in the oil flow rate is two per cent, then $\epsilon_o = 0.02$. The term $\sigma_{\text{w,min}}$ is used so that we do not prescribe unrealistically small measurement errors for the WOR when the WOR is small. To use this model, one must specify values of ϵ_w, ϵ_o, and $\sigma_{\text{w,min}}$. The three diagonal matrices, $C_{D,p}$, $C_{D,w}$ and $C_{D,g}$, respectively, denote the covariance matrices for pressure data measurement errors, WOR measurement errors and GOR measurement errors. In addition, we let $C_{D,s}$ denote the data covariance matrix for time lapse seismic "measurement" errors. It is generally incorrect to model $C_{D,s}$ as a diagonal matrix, but construction of $C_{D,s}$ is difficult; see Aannonsen et al. (2002).

If the total number of conditioning data is N_d, i.e., the dimension of d_{obs} is N_d, then the overall data covariance matrix is given by the following $N_d \times N_d$ diagonal matrix:

$$C_D = \begin{bmatrix} C_{D,p} & O & O & O \\ O & C_{D,w} & O & O \\ O & O & C_{D,g} & O \\ O & O & O & C_{D,s} \end{bmatrix}. \tag{7}$$

We of course do not need to use all types of data as conditioning data. For example, if we wish to history match only GOR data, then $d_{\text{obs}} = d_{\text{obs,g}}$ and $C_D = C_{D,g}$.

For a given model m, d denotes the predicted, true or calculated data corresponding to d_{obs}. If m is the true reservoir from which d_{obs} was obtained and there are no measurement errors, then $d = d_{\text{obs}}$. As d depends on the model, we write

$$d = g(m), \tag{8}$$

to represent the operation of calculating d given m. In our work, Eq. 8 represents the operation of running the reservoir simulator to calculate d.

Bayes' theorem (see Tarantola (1987)) implies that the a posteriori pdf for the model m conditional to the observed data is proportional to the product of the prior pdf and the likelihood function for the model, and is thus given by

$$f(m|d_{\text{obs}}) = a \exp\{-O(m)\}, \tag{9}$$

where a is the normalizing constant and

$$O(m) = \frac{1}{2}\Big[\big(m - m_{\text{prior}}\big)^T C_M^{-1}\big(m - m_{\text{prior}}\big) + \big(g(m) - d_{\text{obs}}\big)^T C_D^{-1}\big(g(m) - d_{\text{obs}}\big)\Big]. \tag{10}$$

Construction of the MAP Estimate and Realizations.

The maximum a posteriori (MAP) estimate is denoted by m_∞ and is defined to be the model that maximizes the pdf of Eq. 9, or equivalently minimizes the objective function of Eq. 10. Although gradient based methods appear to be the only feasible way to construct a minimum of $O(m)$, there is no guarantee that Eq. 10 has a unique global minimum, or that a gradient-based optimization procedure will converge to a global minimum. In fact, if a gradient method is applied to minimize $O(m)$, it is important to check the results to ensure that the method did not converge to a local minimum which yields an unacceptable match of production data, or unreasonable reservoir properties.

If one wishes to generate multiple realizations of the model, it is necessary to sample the conditional pdf of Eq. 9. The most common way to do this is to apply the method proposed by Oliver et al. (1996) and Kitanidis (1995). In our work this method is referred to as the randomized maximum likelihood method. To generate a realization with this procedure, we calculate an unconditional realization m_{uc} from

$$m_{\text{uc}} = m_{\text{prior}} + C_M^{1/2} z_M, \tag{11}$$

where z_M is N_m-dimensional column vector of independent standard random normal deviates. The matrix $C_M^{1/2}$ is a square root of C_M and is normally chosen as $C_M^{1/2} = L$ where

$$C_M = LL^T, \tag{12}$$

is the Cholesky decomposition of C_M. For large problems, generation of the Cholesky decomposition is not feasible, and we apply sequential Gaussian co-simulation to generate an

unconditional realization of the model; see Gómez-Hernández and Journel (1992). Similarly a realization of the data is generated from

$$d_{\text{uc}} = d_{\text{obs}} + C_D^{1/2} z_D, \tag{13}$$

where z_D is an N_d-dimensional column vector of standard random normal deviates. The conditional realization of m is then obtained by minimizing

$$O_r(m) = \frac{1}{2}(m - m_{\text{uc}})^T C_M^{-1}(m - m_{\text{uc}}) + \frac{1}{2}(d - d_{\text{uc}})^T C_D^{-1}(d - d_{\text{uc}}). \tag{14}$$

It can be argued (see Zhang et al. (2001)) that $O(m)$ can be approximated as a chi-squared distribution with expectation given by $E(O(m)) = N_d$ and standard deviation given approximately by $\sigma(O(m)) \approx \sqrt{2N_d}$. Virtually all samples should be within five standard deviations of the mean. Thus, if applying an optimization algorithm to minimize Eq. 14 gives a result m_c, we accept m_c as a legitimate realization if and only if

$$N_d - 5\sqrt{2N_d} \leq O(m_c) \leq N_d + 5\sqrt{2N_d}. \tag{15}$$

Although Eq. 15 has proved to be very reliable for single-phase flow history matching synthetic problems and for multiphase flow problems, we occasionally encounter situations where we are unable to decrease the objective function to a value consistent with Eq. 15. This aspect needs further investigation.

We wish to be able to history match several hundred production data to generate realizations of tens of thousands of model parameters. Thus computational efficiency is an extremely important consideration. For such problems, it is not feasible to compute the individual sensitivity coefficients required by standard implementation of the Gauss-Newton and Levenberg-Marquardt algorithms. Thus, we are forced to focus on algorithms which require only the gradient of the objective function. As steepest descent often exhibits poor convergence properties (see, for example, Fletcher (1987)). The only viable algorithms in this category appear to be variable metric (quasi-Newton) methods and preconditioned conjugate gradient (PCCG) methods. The efficiency of PCCG methods largely rest on finding a good preconditioner. This, however, is not an easy task and despite significant effort we have not been able to find a preconditioner that yields a PCCG method that is as reliable as the quasi-Newton methods we have implemented.

The Gauss-Newton method with restricted step has often been used to minimize $O(m)$; see Chu et al. (1995a). However, if the initial guess for the model yields a very poor match of the observed production data, a straightforward application of the method may converge

extremely slowly or may converge to a model which yields an unacceptable match of production data; see Wu et al. (1999). Wu et al. (1999) overcame this problem by using an artificially high value for the variance of data measurement errors at early iterations. Here, we avoid this difficulty by using a form of the Levenberg-Marquardt algorithm introduced by Bi (1999). This algorithm can be written in two different forms. The first comes from a modification of the standard Gauss-Newton method and is given by

$$\left[(1+\lambda_l)C_M^{-1} + G_l^T C_D^{-1} G_l\right]^{-1} \delta m^{l+1} = -\left[C_M^{-1}(m^l - m_{\text{prior}}) + G_l^T C_D^{-1}(g(m^l) - d_{\text{obs}})\right], \quad (16)$$

$$m^{l+1} = m^l + \alpha_l \delta m^{l+1} \quad (17)$$

where $\alpha_l = 1$. Here l, as either a subscript or superscript, refers to the iteration index. The matrix G_l denotes the $N_d \times M$ sensitivity coefficient matrix evaluated at m^l. The entry in the ith row and jth column of G_l represents the sensitivity of the ith calculated data g_i to the jth model parameter evaluated at m^l, i.e., this entry is $\partial g_i(m^l)/\partial m_j$, where m_j is the jth entry of m. If $O(m^{l+1}) < O(m^l)$, we set $\lambda_{l+1} = \lambda_l/10$, and if the objective function does not decrease, we increase the Levenberg-Marquardt parameter by a factor of 10. We start with an initial value of $\lambda = 10,000$. For the multiphase flow problems we have considered to date, this simple procedure works well.

Calculation of the Gradient of the Objective Function

The equations that must be solved to compute sensitivity coefficient with the adjoint method were presented in the first annual report on this project. These results can also be found in Li et al. (2001a). Implementation of methods for calculating the gradient of the objective function for use in the limited memory Broyden-Flection-Goldfarb-Shanno (LBFGS) algorithm were given in the second annual report on this project and can also be found in Zhang and Reynolds (2002a) and Zhang (2002). Thus, we do not present here the specific equations for the calculations of sensitivities of the gradient of an objective function. As our algorithm of choice for large problems is LBFGS, we do include a description of the algorithm; for additional details see the last annual report on this project, Zhang and Reynolds (2002a) and Zhang (2002).

LBFGS Algorithm

The search direction in the Newton's method can be written as

$$d_{k+1} = -H_k^{-1} g_k, \quad (18)$$

where H_k and g_k, respectively, denote the second derivative (Hessian matrix) and the first derivative (gradient) of the objective function evaluated at m_k and k is the iteration index. With $O(m)$ given by either Eq. 10 or Eq. 14, the Gauss-Newton Hessian matrix is estimated by

$$H_k = C_M^{-1} + G_k^T C_D^{-1} G_k, \tag{19}$$

where G_k is the sensitivity matrix evaluated at m_k. As noted before, if both the number of model parameters and the number of data are large, the evaluation of G_k is computationally expensive. In quasi-Newton methods, H_k^{-1} is approximated by a symmetric positive definite matrix $\tilde{H}_k^{-1}$ which is corrected or updated from iteration to iteration. With this Hessian inverse approximation matrix, the search direction can be written as

$$d_{k+1} = -\tilde{H}_k^{-1} g_k. \tag{20}$$

Because the matrix $\tilde{H}_k^{-1}$ takes the place of H_k^{-1} in Eq. 18, the method with search direction given by Eq. 20 is called a quasi-Newton method. This method is also called a variable metric method.

In a quasi-Newton method, the key issue is how to generate the approximation to the inverse Hessian matrix. Different quasi-Newton methods use different formulas to calculate $\tilde{H}_{k+1}^{-1}$ from $\tilde{H}_k^{-1}$. All updating formulas satisfy the quasi-Newton condition given by

$$\tilde{H}_{k+1}^{-1} y_k = s_k, \tag{21}$$

where

$$y_k = g_{k+1} - g_k, \tag{22}$$

and

$$s_k = m_{k+1} - m_k. \tag{23}$$

Various possible updating formulas honor this quasi-Newton condition. The Broyden family equation is given by

$$\tilde{H}_{k+1}^{-1} = \tilde{H}_k^{-1} + \frac{s_k s_k^T}{s_k^T y_k} - \frac{\tilde{H}_k^{-1} y_k y_k^T \tilde{H}_k^{-1}}{y_k^T \tilde{H}_k^{-1} y_k} + \theta_k v_k v_k^T, \tag{24}$$

where $\theta_k \in [0, 1]$ and

$$v_k = (y_k^T \tilde{H}_k^{-1} y_k)^{1/2} \Big(\frac{s_k}{s_k^T y_k} - \frac{\tilde{H}_k^{-1} y_k}{y_k^T \tilde{H}_k^{-1} y_k} \Big). \tag{25}$$

Given that the line search is exact and the initial Hessian inverse approximation is real symmetric positive definite, the Hessian inverse approximation generated by Eq. 24 is guaranteed to be symmetric positive definite; see details in Zhang (2002). In our procedure, we

use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) correction equation proposed by Broyden (1970), Fletcher (1970), Goldfarb (1970) and Shanno (1970) independently, which is a special case of Broyden family obtained by setting $\theta_k = 1$ in Eq. 24. The BFGS update equation is given by

$$\tilde{H}_{k+1}^{-1} = \tilde{H}_k^{-1} + \frac{s_k s_k^T}{s_k^T y_k} - \frac{\tilde{H}_k^{-1} y_k y_k^T \tilde{H}_k^{-1}}{y_k^T \tilde{H}_k^{-1} y_k} + v_k v_k^T. \tag{26}$$

The limited memory BFGS (LBFGS), which uses a limited number of previous vectors (y_k's and s_k's) to construct the inverse Hessian approximation at each iteration, is an appropriate method for large scale problems where it is not feasible to explicitly store and compute the full matrix $\tilde{H}_k^{-1}$. In our work, the algorithm proposed by Nocedal (1980) was implemented and applied. In order to derive the limited memory BFGS, the normal BFGS formula Eq. 26 can be written as

$$\tilde{H}_{k+1}^{-1} = V_k^T \tilde{H}_k^{-1} V_k + \rho_k s_k s_k^T, \tag{27}$$

where $\rho_k = 1/y_k^T s_k$ and $V_k = I - \rho_k y_k s_k^T$. Nocedal (1980) suggested a procedure where only the L previous vectors are used when constructing the new $\tilde{H}_{k+1}^{-1}$. When $k < L$, the update equation is still given by Eq. 27 which can be rewritten as

$$\begin{aligned} \tilde{H}_{k+1}^{-1} =& V_k^T V_{k-1}^T \cdots V_0^T \tilde{H}_0^{-1} V_0 \cdots V_{k-1} V_k \\ &+ V_k^T \cdots V_1^T \rho_0 s_0 s_0^T V_1 \cdots V_k \\ &\vdots \\ &+ V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\ &+ \rho_k s_k s_k^T. \end{aligned} \tag{28}$$

For $k + 1 > L$ the update equation is

$$\begin{aligned} \tilde{H}_{k+1}^{-1} =& V_k^T V_{k-1}^T \cdots V_{k-L+1}^T \tilde{H}_0^{-1} V_{k-L+1} \cdots V_{k-1} V_k \\ &+ V_k^T \cdots V_{k-L+2}^T \rho_{k-L+1} s_{k-L+1} s_{k-L+1}^T V_1 \cdots V_k \\ &\vdots \\ &+ V_k^T \rho_{k-1} s_{k-1} s_{k-1}^T V_k \\ &+ \rho_k s_k s_k^T. \end{aligned} \tag{29}$$

Unless the dimension of $\tilde{H}_k^{-1}$ is small, direct application of Eqs. 28 and 29, which involve matrix products, is inefficient. Instead, we form the product $\tilde{H}_k^{-1} g_k$, which is used to construct the search direction, directly by using the algorithm proposed by Nocedal (1980).

The calculation of $\tilde{H}_k^{-1} g_k$ only involves vector products instead of matrix products. Because only the L most recent vectors from the set of s_k and y_k are used to construct $\tilde{H}_{k+1}^{-1}$, this algorithm is called the limited memory BFGS method.

The basic LBFGS algorithm we use to minimize an objective function (Eq. 10 or Eq. 14) is given below.

Step 1 Initialization

(a) Provide an initial guess, m_0, of the model, calculate the objective function corresponding to m_0 and evaluate the gradient of the objective function at m_0, i.e., compute g_0;

(b) provide an initial Hessian inverse approximation $\tilde{H}_0^{-1}$ (e.g., C_M in our examples), set the initial iteration index k=0.

Step 2 Calculate the search direction $d_k = -\tilde{H}_k^{-1} g_k$ and check whether it is a downhill direction, i.e., check to see if $d_k^T g_k < 0$. If d_k is not a downhill search direction, set $d_k = -\tilde{H}_0^{-1} g_k$.

Step 3 Calculate the step size α_k by a line search procedure as discussed later.

Step 4 Update the model to $m_c = m_k + \alpha_k d_k$.

Step 5 Calculate the objective function based on m_c.

Step 6 Determine if the Wolfe conditions are satisfied; if they are satisfied, then set $m_{k+1} = m_c$ and go to step 7, otherwise do

(a) fit a quadratic and find a step size by minimizing this quadratic, then go to step 4;

(b) if a quadratic fit has already been done, cut the step size by a specified factor (in our examples we cut the step size by a factor of 10) and go to step 4. All computations we have done suggest this case does not occur very often.

Step 7 Determine if the stopping criteria are satisfied. If satisfied, then stop; otherwise go to step 8.

Step 8 Calculate $s_k = m_{k+1} - m_k = \alpha_k d_k$ and $y_k = g_{k+1} - g_k$. Apply Eq. 28 or 29 to update the inverse Hessian approximation $\tilde{H}_{k+1}^{-1}$. Set $k = k + 1$ and then go to step 2.

Although the procedure presented above is convenient for discussion, we never explicitly update the inverse Hessian approximation as part of Step 8 after the first iteration. Instead we use the algorithm presented by Nocedal (1980) to compute $d_k = -\tilde{H}_k^{-1} g_k$ in Step 2. This eliminates the need to explicitly compute or store H_k^{-1} for $k > 0$.

For BFGS and LBFGS, scaling can have a significant effect on the rate of convergence; Shanno (1970), Oren and Luenberger (1974), Oren (1974), Oren and Spedicato (1976). If scaling is used, Step 8 needs to be modified. As scaling was considered in detail in last year's annual report, here we only give the scaling algorithm we use. Although this scaling algorithm was shown to be robust for several examples, there is no theoretical result that guarantees that it will always work well.

For the LBFGS algorithm, we have found a variant of the optimal switching rule given by Oren and Spedicato (1976) which works well. Specifically, we compute

$$\tilde{\tau}_k = \frac{s_k^T \tilde{H}_0 s_k}{s_k^T y_k}, \tag{30}$$

$$\tilde{\sigma}_k = \frac{s_k^T y_k}{y_k^T \tilde{H}_0^{-1} y_k}, \tag{31}$$

and then determine the scaling factor γ_k by the following rule:

$$\gamma_k = \begin{cases} \tilde{\tau}_k & \text{if} \quad \tilde{\tau}_k < 1.0 \\ \tilde{\sigma}_k. & \text{otherwise.} \end{cases} \tag{32}$$

For the LBFGS algorithm with initial scaling, we just replace $\tilde{H}_0^{-1}$ in Eq. 28 at the first iteration and and then use Eqs. 28 and 29 without modification at all subsequent iterations. For the LBFGS with scaling at all iterations, we replace $\tilde{H}_0^{-1}$ in Eqs. 28 and 29 by $\gamma_k \tilde{H}_0^{-1}$ in computing $\tilde{H}_{k+1}^{-1}$. In our examples, we use scaling at all iterations. The efficient LBFGS method given by Nocedal (1980) avoids formation of $\tilde{H}_k^{-1}$ for $k \geq 1$, only $\tilde{H}_k^{-1} g_k$ is calculated at each iteration. However, $\tilde{H}_0^{-1}$ must be provided as the initial approximation to the inverse Hessian. In our implementation, we use $\tilde{H}_0^{-1} = C_M$, where C_M is the prior covariance matrix.

Convergence Criteria

In our results, the following stopping criteria are used to terminate the algorithm:

1. $$\frac{\mid O_{k+1} - O_k \mid}{O_k + 10^{-14}} < \varepsilon_1 \tag{33}$$

   and

   $$\frac{\parallel m_{k+1} - m_k \parallel_2}{\parallel m_k \parallel_2 + 10^{-14}} < \varepsilon_2 \tag{34}$$

   where k denotes the iteration index and $\| \cdot \|_2$ denotes the l_2 norm of a vector. Both conditions must be satisfied to terminate the iteration. If we use only Eq. 33 as the

convergence criterion, the algorithm may converge prematurely especially when the objective function decreases very slowly at the early iterations. Because at the early iteration, the objective function is relatively big such that Eq. 33 becomes easier to be satisfied.

2. Specify a maximum allowable iteration number. If the number of iterations exceeds the specified number, we force the iteration to stop. In our examples, we usually specify the maximum number of iterations as 300.

Line Search

In our implementation of conjugate gradient and quasi-Newton methods, the line search is performed using one iteration of the Newton-Raphson method followed by a quadratic fit if necessary. We do not do an exact line search, but terminate the line search when the Wolfe conditions are satisfied; see, for example, Fletcher (1987). The Wolfe conditions are used to ensure that step sizes are not too small and that the reduction in the objective function is not negligible. In addition, the Wolfe conditions are side conditions for the exact line search; see Kolda et al. (1998). At each iteration, we perform one Newton-Raphson iteration to find a step size. Then we check whether this step satisfies the Wolfe conditions. If it does, we accept this step size. Otherwise we find an optimum step size by fitting a quadratic, as discussed in the last annual report, and then check whether the new step satisfies the Wolfe conditions. If it does, we accept this new step. Otherwise, we check whether the objective function increases or decreases. If it increases, we cut the step size by a factor of 10. If it decreases, we accept this new step size no matter whether the Wolfe conditions are satisfied or not. Our experience shows that for most of the iterations, the step size generated by one Newton-Raphson iteration satisfies the Wolfe conditions and two-thirds or more of the step sizes satisfy the Wolfe conditions after the quadratic fit. However, technically the Wolfe conditions should be satisfied at each iteration to guarantee convergence. Because of this, we have implemented a new line search procedure which may prove to be both more computationally efficient and robust. We are currently testing the new line search algorithm.

HISTORY MATCHING, TENGIZ FIELD EXAMPLE

Tengiz Reservoir

In this chapter, we consider matching static pressures from pressure buildup surveys for the Tengiz reservoir, which is an undersaturated oil reservoir. We consider history matching buildup data from the Tengiz reservoir which is located in the Pri-Caspian Basin. Tengiz is a carbonate reservoir which was formed during Devonian and Carboniferous time. Figure 1 shows that the central or "platform" portion of the reservoir is relativly flat with localized structural highs on the south and eastern edges. As shown in Fig. 1, the platform is bounded by faults or lithologic breaks and surrounded by gently sloping "flanks" of carbonate debris; see Chambers (1997) and He and Chambers (1999) for additional details on the geology. The index on the x and y axes is the number of gridblocks, i.e., the areal grid for the original reservoir model was 90×100.

Tengiz is an undersaturated oil reservoir produced by 44 wells. With very rare exception, all flowing bottomhole pressures have been maintained above bubble point pressure, which is equal to 3586 psi. Initial reservoir pressure is 11950 psi at a datum of 14765 ft subsea. Current average reservoir pressure remains more than twice the bubble point pressure, and consequently, the oil flows as a single phase in the reservoir.

Our history match of static pressure data from pressure buildup surveys is based on an upscaled reservoir model of Tengiz. The upscaled reservoir model of Tengiz was created by removing most of the sloping flanks near the outer edges of the reservoir and upscaling the remainder of the reservoir to a $59 \times 49 \times 9$ grid. In the upscaled model, the gridblock sizes in the x and y directions are almost uniform with values between 815 and 825 ft. Gridblock sizes in the z direction are non-uniform with values varying between 15 and 150 ft. Figure 2 shows a contour map of the top of the reservoir with well locations.

The initial permeability and porosity field are from a geostatistical model, generated by

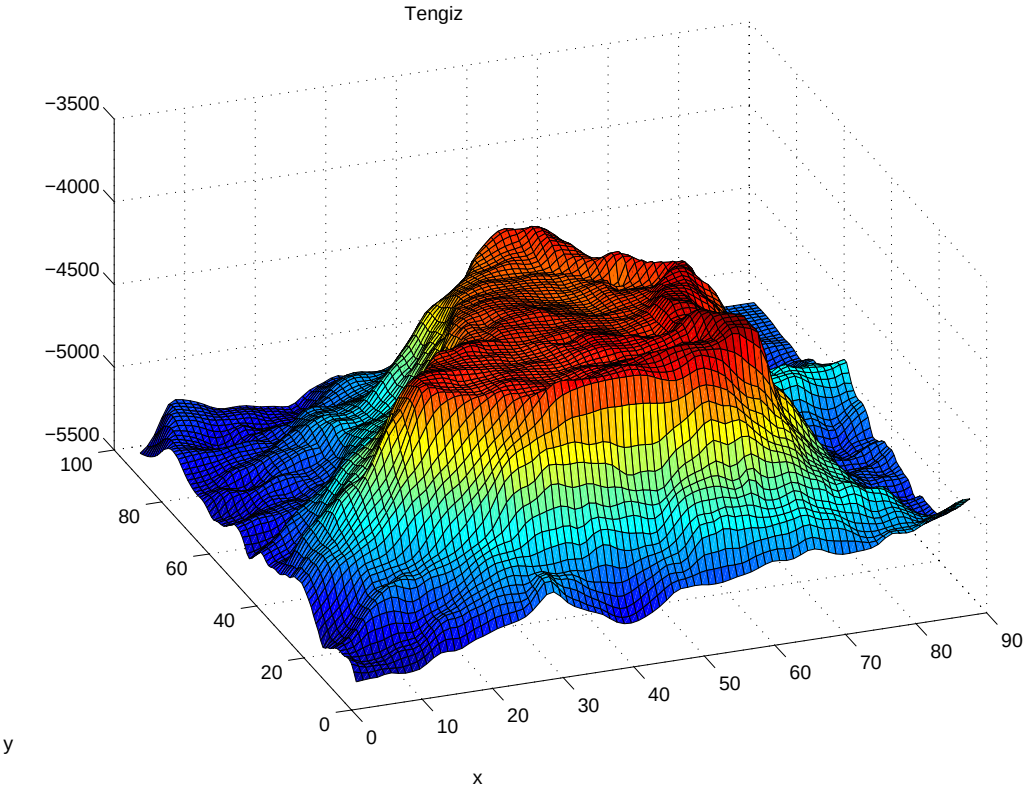


Figure 1: A 3D plot of the Tengiz field.

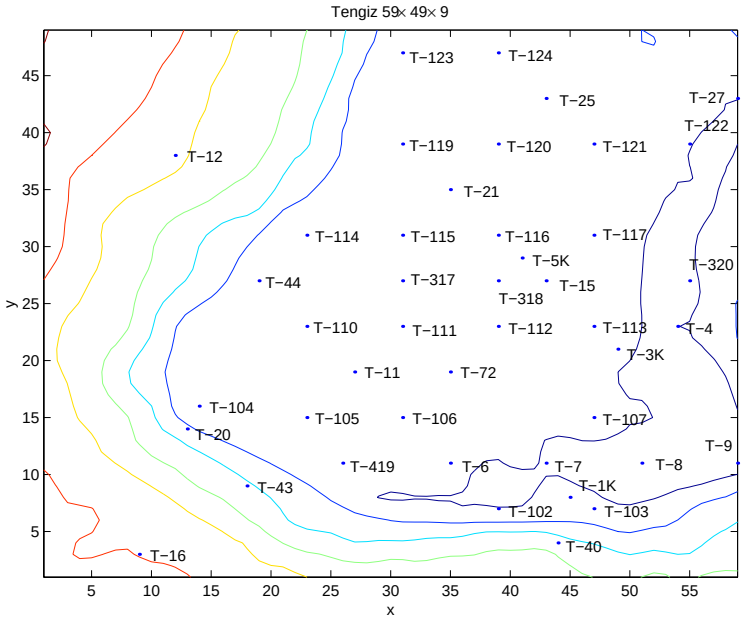


Figure 2: A contour plot of the upscaled model with the well locations.

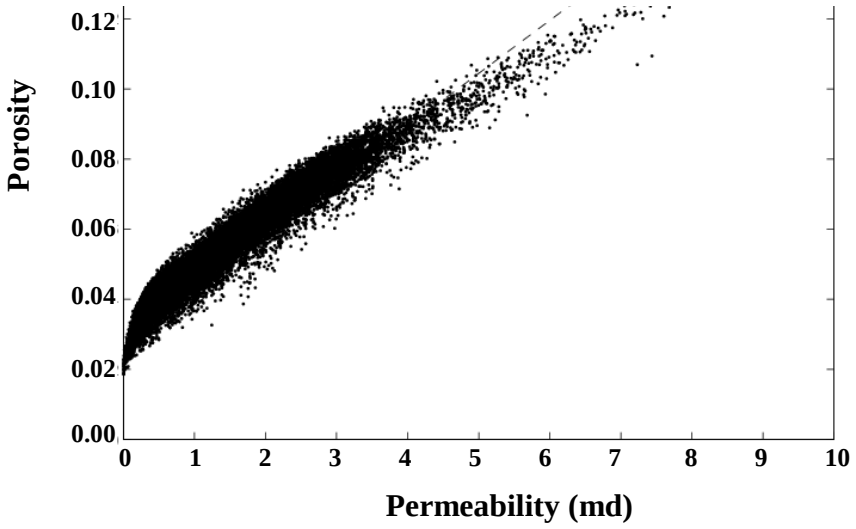


Figure 3: Cross plot of porosity to permeability.

Chevron geoscientists. The cross plot of the permeability versus porosity in Figure 3 shows an overall least square regression trend of increasing porosity with increasing permeability. The porosity varies from less than 2 percent to over 14 percent, while permeability varies from less than 0.01 md to over 9 md.

Overshooting/Undershooting

For large scale history matching problems where the number of reservoir variables estimated or simulated is too large to be resolved by the number of independent data, the history matching problem is ill-conditioned. Ill-conditioning can lead to undershooting/overshooting problems where the estimates of some variables are unreasonably small or unreasonably large. Regularization by using a prior geostatistical model does not usually alleviate this problem if the initial data mismatch is large. As mentioned previously, Li et al. (2003) used a modified Levenberg-Marquardt algorithm to damp changes in model variables at early iterations to reduce undershooting/overshooting. However, the modified Levenberg-Marquardt algo-

rithm is impractical if we wish to estimate or simulate thousands of reservoir variables. Thus, we consider imposing constraints on the model parameters in order to control the undershooting/overshooting problem. We investigate two methods for transforming the resulting constrained optimization problem to a unconstrained optimization problem. The first method simply adds a penalty function to the objective function. The second method uses a logarithmic transformation.

As noted in the last annual DOE report, history matching pressure data from the Tengiz reservoir resulted in abnormally large changes in permeability in some grid blocks. The right figure in Fig. 4 shows the horizontal permeability field for the second layer of the Tengiz reservoir that was obtained by history matching pressure data. The left figure shows the corresponding layer from an initial model provided by ChevronTexaco. This represents an unconditional realization of the log-permeability field. This unconditional realization was used as the initial guess in the LBFGS algorithm which at convergence gave the log-permeability field shown by the right figure in Fig. 4. In the unconditional realization the maximum value of horizonatl log-permeability is about 2.2 (9 md). After history matching production data, the values of $ln(k)$ increased to 6 (403 md) in some reservoir simulator grid blocks, but decreased to around -5 (0.007 md) in some other grid blocks.

Motivated by the results obtained for the Tengiz example, we investigate algorithms to overcome under and overshooting. Damping the data mismatch part is one way to control this problem that we have used previously. Applying constraints is another option.

Regularized History Matching

As noted in Chapter 3, the the a posteriori pdf of the model parameters given the vector of observed data, d_{obs}, is estimated by Eqs. 9 and 10 which are repeated here as

$$f(m|d_{\rm obs}) = a \exp\{-O(m)\}, \tag{35}$$

where a is the normalizing constant and

$$O(m) = \frac{1}{2}\Big[\big(m - m_{\rm prior}\big)^T C_M^{-1}\big(m - m_{\rm prior}\big) + \big(g(m) - d_{\rm obs}\big)^T C_D^{-1}\big(g(m) - d_{\rm obs}\big)\Big]. \tag{36}$$

As also noted in Chapter 3, to generate a realization of the model by the randomized maximum likelihood method, we minimize

$$O(m) = \frac{1}{2}\Big[\big(m - m_{\rm uc}\big)^T C_M^{-1}\big(m - m_{\rm uc}\big) + \big(g(m) - d_{\rm uc}\big)^T C_D^{-1}\big(g(m) - d_{\rm uc}\big)\Big], \tag{37}$$

where m_{uc} is an unconditional realization of the prior model.

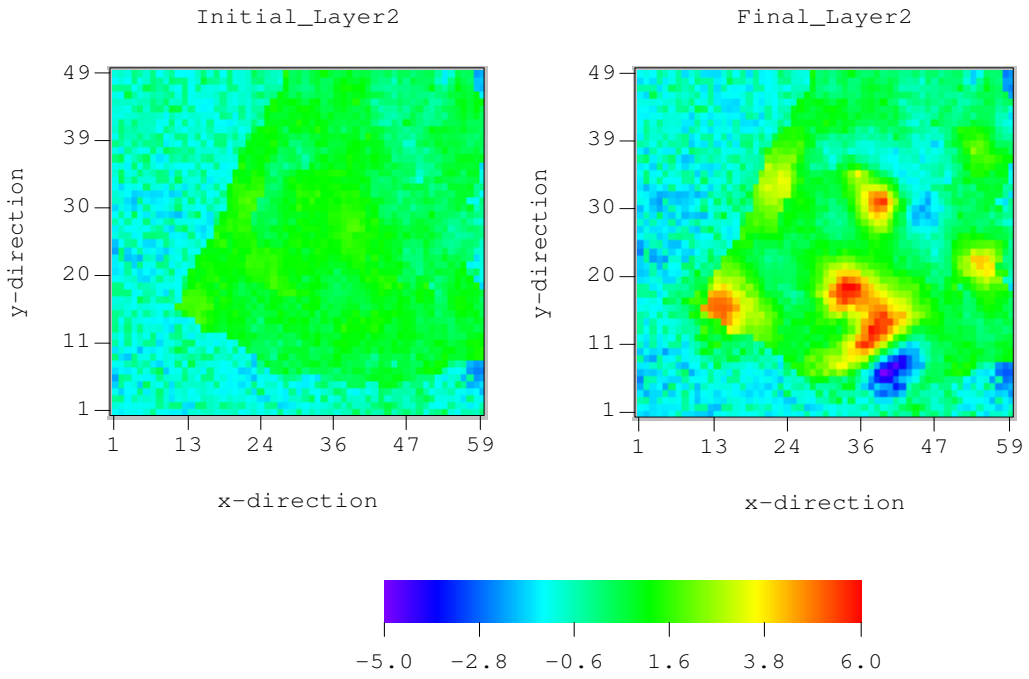


Figure 4: Example of Overshooting.

In practice, we can apply a damping factor η to avoid the undershooting/overshooting problem. When the damping function is introduced, the objective function is modifed to the following:

$$O(m) = \frac{\eta}{2}\Big[d_{obs} - g(m)\Big]^T C_D^{-1}\Big[d_{obs} - g(m)\Big] \qquad (38)$$
$$+ \frac{1}{2}(m - m_{\text{prior}})^T C_M^{-1}(m - m_{\text{prior}}).$$

It may be best to look at this approach as a modification of C_D. If η is small, we have effectively reduced C_D^{-1}, or equivalently, increased C_D. Increasing C_D weights the data less in the overall objective function and results in damping of the changes in the model over an iteration. This tends to help control undershooting/overshooting. There are many variations of this approach, for example, we modify the C_D matrix directly and use different modifications for different types of data (pressure, GOR and WOR). In fact, we can estimate a modified C_D at each iteration as some fraction of the variances estimated from the current data mismatch term. We can also rescale the measurement errors to a relative basis before modifying C_D. Although we have experimented with a variety of procedures, we have no clear cut conclusions at this time. Moreover, the constrained algorithms presented below

are more efficient when the LBFGS algorithm is used for minimization so we focused on the constrained algorithms.

Constrained Optimization

Let m_l and m_u, respectively, denote vectors with entries representing the lower and upper bounds of the model parameters; these bounds would be specified from physical knowledge of the reservoir model. We present two methods for constraining the variables based on these specified bounds, the penalty function method and a log-transformation method.

Penalty Function method

Penalty functions represent "soft" constraints. In this approach a penalty function term is added to the original objective function and then the modified objective function is minimized. If model parameter exceeds its upper or lower bounds, the penalty function will be very large, and hence such a values of model parameters can not represent a minimum of the modified objective function and the objective function will not converge to such a model.

Before constructing the penalty function, we first apply a linear transformation to rescale the model parameters by their upper and lower bounds:

$$\begin{aligned} m_{m,j} &= \frac{1}{2}(m_{u,j} + m_{l,j}) \\ m_{r,j} &= \frac{1}{2}(m_{u,j} - m_{l,j}) \\ s_j &= \frac{m_j - m_{m,j}}{m_{r,j}} \end{aligned} \tag{39}$$

Here, s_j represents the rescaled variable corresponding to the jth entry of the vector m of model parameters. When $m_j \to m_{u,j}$, $s_j \to 1$, and when $m_j \to m_{l,j}$, $s_j \to -1$.
The penalty function is defined by

$$O_b(m) = \mu \sum_{j=1}^{N_m} \{\exp[a(s_j - 1)] + \exp[-a(s_j + 1)]\} \tag{40}$$

where $\mu > 0$ and $a > 0$ are constants which determine the weight and shape of the penalty function. The partial derivative of the penalty function with respect to s_j is:

$$\frac{\partial O_b(m)}{\partial s_j} = \mu a \sum_{j=1}^{N_m} \{\exp[a(s_j - 1)] - \exp[-a(s_j + 1)]\} \tag{41}$$

By adding the penalty function to the original objective function (Eq. 37 or 36), we transfer the constrained optimization problem into an unconstrained optimization problem. The LBFGS optimization algorithm discussed above can be applied to the new objective function.

Logarithmic Transformation Method

Another procedure for transforming the constrained optimization problem into an unconstrained optimization problem is to find a suitable transformation that can map the upper bound to ∞ and the lower bound to $-\infty$. Through this transformation, the boundaries are removed. In general, the transformation needs to be invertible (one-to-one) which is normally achieved by using a monotonic transformation. One choice for such a transformation is given by the following logarithmic transformation:

$$\begin{aligned} m_j &= m_{m,j} + m_{r,j}\frac{\exp(s_j)-1}{\exp(s_j)+1} \\ &= m_{m,j} + m_{r,j}\frac{1-\exp(-s_j)}{1+\exp(-s_j)} \end{aligned} \tag{42}$$

$$s_j = \ln\Big(\frac{m_j - m_{l,j}}{m_{u,j} - m_j}\Big) \tag{43}$$

The relationship between the derivatives of any function $f(m)$ with respect to m_j and the derivative with respect to s_j can be obtained from the chain rule as follows:

$$\begin{aligned} \frac{dm_j}{ds_j} &= 2m_{r,j}\frac{\exp(s_j)}{\{\exp(s_j)+1\}^2} \\ &= 2m_{r,j}\frac{\exp(-s_j)}{\{\exp(-s_j)+1\}^2} \end{aligned} \tag{44}$$

$$\begin{aligned} \frac{\partial f}{\partial s_j} = \frac{\partial f}{\partial m_j}\frac{dm_j}{ds_j} &= 2m_{r,j}\frac{\exp(s_j)}{\{\exp(s_j)+1\}^2}\frac{\partial f}{\partial m_j} \\ &= 2m_{r,j}\frac{\exp(-s_j)}{\{\exp(-s_j)+1\}^2}\frac{\partial f}{\partial m_j} \end{aligned} \tag{45}$$

If $m_j \to m_{u,j}$, then $s_j \to \infty$; and if $m_j \to m_{l,j}$, then $s_j \to -\infty$. Thus, the boundary constraints are removed. Note, however, that for very large values of $|s_j|$, we may encounter overflow when evaluating the exponentials. To avoid this problem, we apply the equation that involves $\exp(-s_j)$ when $s_j > 0$; otherwise, we apply the equation with expression that involves $\exp(s_j)$.

	$\ln k$	$\ln k_z$	ϕ
Var	0.8073	0.9376	0.000261
	$(\ln k, \ln k_z)$	$(\ln k, \phi)$	$(\ln k_z, \phi)$
ρ	0.5	0.9096	0.662

Table 1: Statistical Parameters for Tengiz.

History Matching Pressure Data

As noted previously, the unconditional realization of the permeability and porosity fields are from a geostatistical model generated by ChevronTexaco geoscientists. As examples, Figs. 5 through 10 show the initial porosity, horizontal log-permeability and vertical log-permeability fields for the third and fourth layers of the reservoir. The geostatistical model parameters used to generate an unconditional realization are given in Table 1. We used a spherical covariance structure with ranges in the x, y and directions, respectively, equal to 6560 ft, 4920 ft and 165 feet.

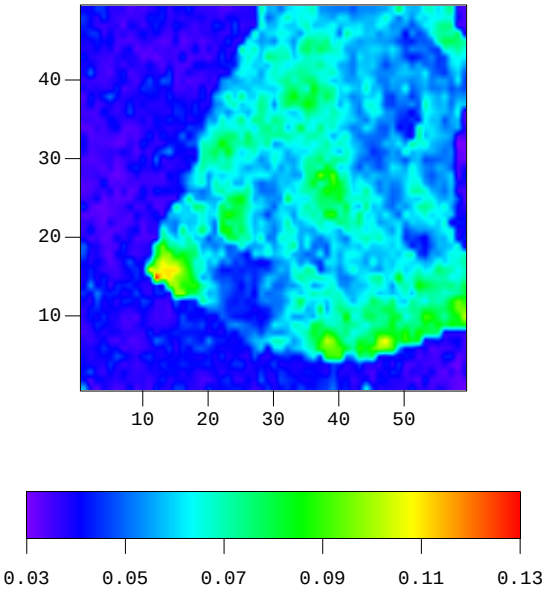


Figure 5: Initial Porosity for Layer 3 of Tengiz

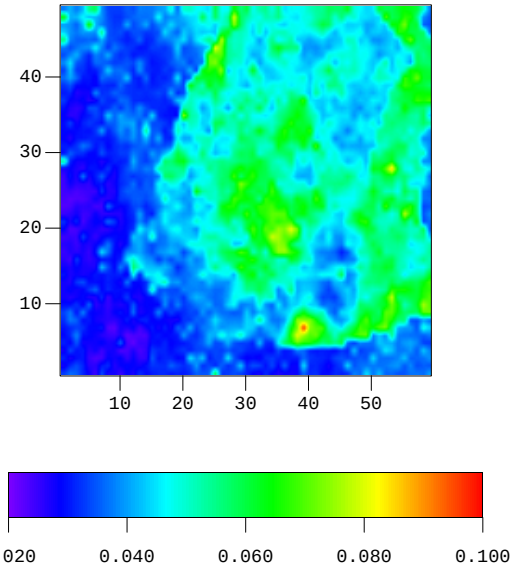


Figure 6: Initial Porosity of Layer 4 of Tengiz

A total of 104 observed pressure data from 40 wells were history matched. These pressure data span the historical period from April 1991 to January 1998. We specified the

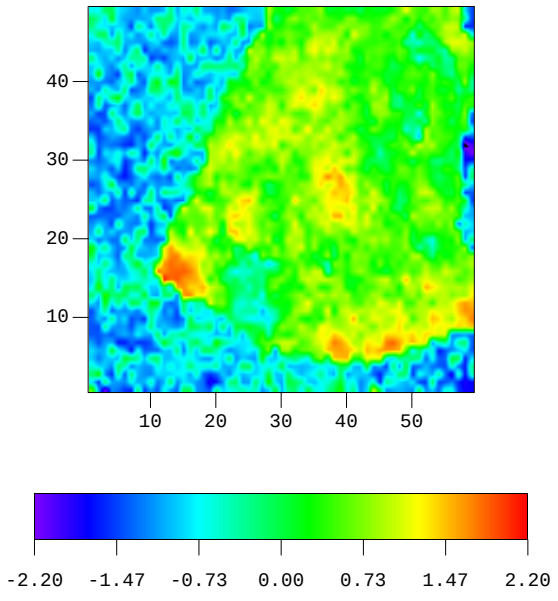


Figure 7: Initial Horizontal Log-Permeability of Layer 3 of Tengiz

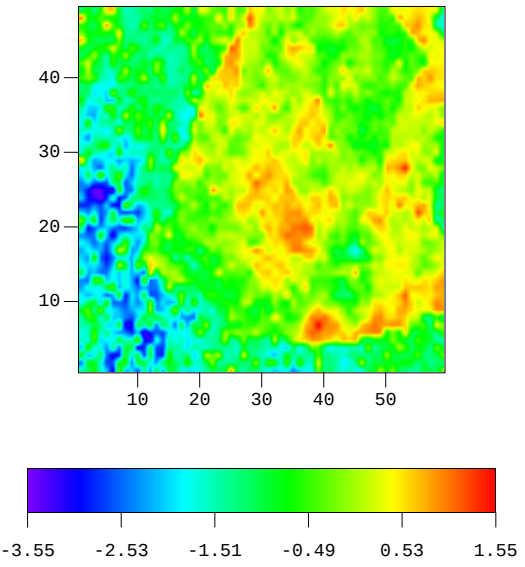


Figure 8: Initial Horizontal Log-Permeability for Layer 4 of Tengiz

oil production rate for each well based on monthly historical production data, and also set a minimum bottom hole pressure of 3000 psi for all field simulation runs. In the history matching process, we treated the initial model provided by ChevronTexaco as an unconditional realization, m_{uc}. For the unconstrained case, we applied the LBFGS to minimize the objective function

$$O(m) = \frac{1}{2}\Big[d_{obs} - g(m)\Big]^T C_D^{-1}\Big[d_{obs} - g(m)\Big] + \frac{1}{2}(m - m_{\rm uc})^T C_M^{-1}(m - m_{\rm uc}), \tag{46}$$

with the variance of pressure measurement errors set equal to 1 psi^2. For the constrained case, we modified the objective function as described previously. The model m_{uc} was used as the initial guess in all cases.

Results of History Matching with Different Algorithms

Three different optimization algorithms were applied for this case, the basic unconstrained LBFGS optimization algorithm, constrained LBFGS with the log-transformation algorithm and constrained LBFGS with a penalty function. The upper bounds and lower bounds for

Upper($\ln k$)	Lower($\ln k$)	Upper($\ln k_z$)	Lower($\ln k_z$)	Upper ϕ	Lower ϕ
4.21	-12.01	1.52	-12.05	0.2	0.01

Table 2: Upper and Lower Bounds for Constrained Optimization, Tengiz Reservoir.

constrained optimization are listed in Table 2.

The behavior of the objective function during the history matching process for the different algorithms are shown in Fig. 11. The objective function stopped at a rather high level for all algorithms but lower values were achieved with the constrained algorithms. Figs. 12 through Fig. 14 show the pressure matches obtained with the three algorithms, in these figures, all pressure data at all wells are shown, the value on the abscissa simple represents the data index. A comparison of pressure mismatches obtained with the different algorithms is shown in Fig. 15. We can see that the maximum pressure mismatch is about 700 psi when log-transformation is made before applying LBFGS, about 900 psi for the penalty function method and 1100 psi for unconstrained LBFGS algorithm. Overall, the log-transformation procedure results in the the lowest pressure mismatches with most pressure mismatch terms less that 300 psi. Although these values are still higher than desired it is important to note that the pressure data provided correspond to buildup data, but rates were provided only on a monthly basis. Thus most buildup pressures actually correspond to periods when the well was still flowing and we treated the pressure data as data for flowing wellbore pressure during the history matching process.

The horizontal log-permeability, vertical log-permeability and porosity fields of layers 3 and 4 obtained with the unconstrained optimization algorithm are shown in Figs. 16 through 21. For the unconstrained case, the maximum of $\ln k$ reaches 7.0 in the third layer which is unreasonable based on the geostatistical model. Figs. 22 through 27 pertain to conditional realizations of the horizontal log-permeability, vertical log-permeability and porosity fields obtained by history matching pressure data using the constrained optimization algorithm based on the log-transformation. For constrained case, the minimum $\ln k_z$ reaches its lower bound, but nevertheless, a more reasonable log-permeability field was obtained than was generated using the unconstrained algorithm. From careful examination of the data, we found that if a pressure buildup datum fell in a time period where the average monthly rate was quite high, then it was difficult to match the datum. This suggests that the conversion of such a buildup pressure to an equivalent flowing bottomhole pressure resulted in some pressures that were inconsistent with the flow rate data.

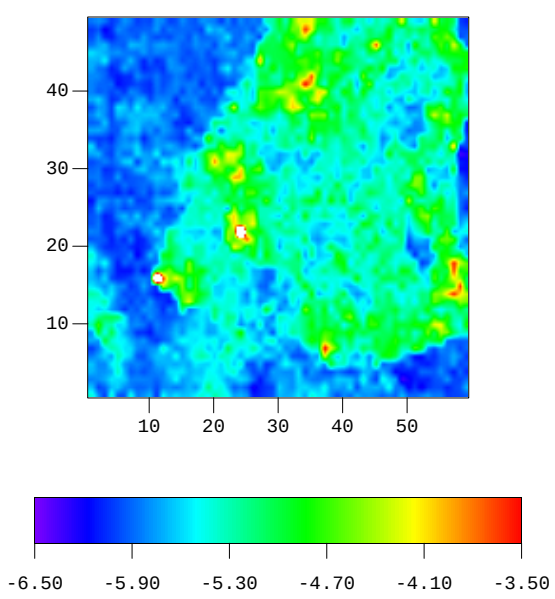


Figure 9: Initial Vertical Log-Permeability of Layer 3 of Tengiz

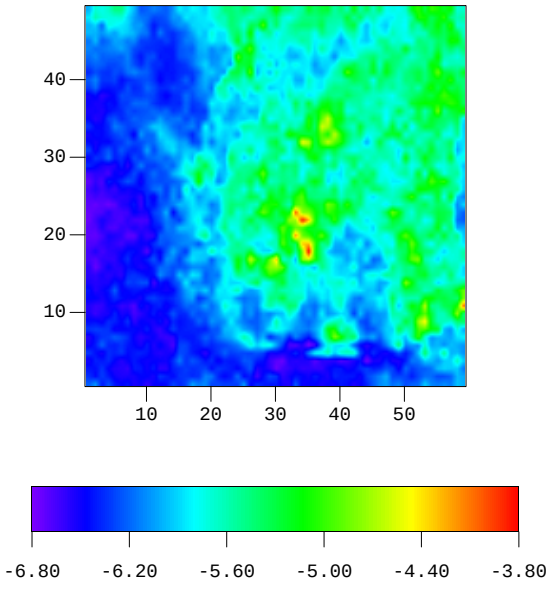


Figure 10: Initial Vertical Log-Permeability for Layer 4 of Tengiz

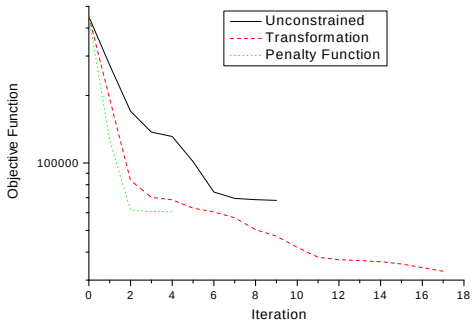


Figure 11: Objective Function Plots for Tengiz

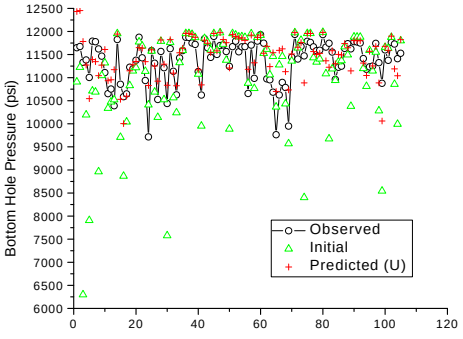


Figure 12: Pressure Match for Tengiz, Unconstrained Algorithm.

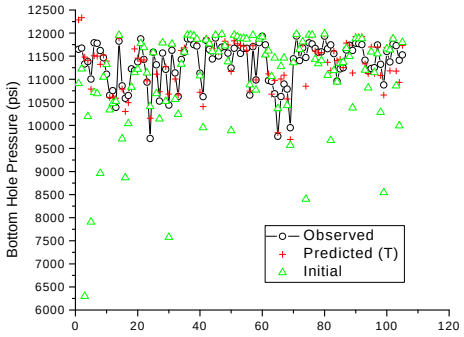


Figure 13: Pressure Match for Tengiz, Transformation Method

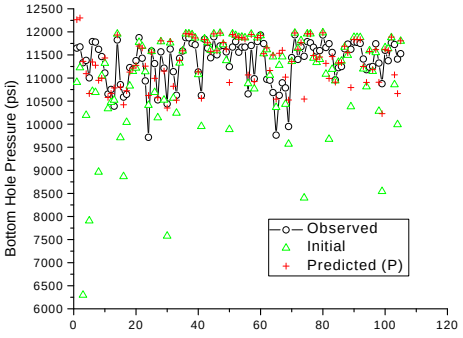


Figure 14: Pressure Match for Tengiz, Penalty Function Method.

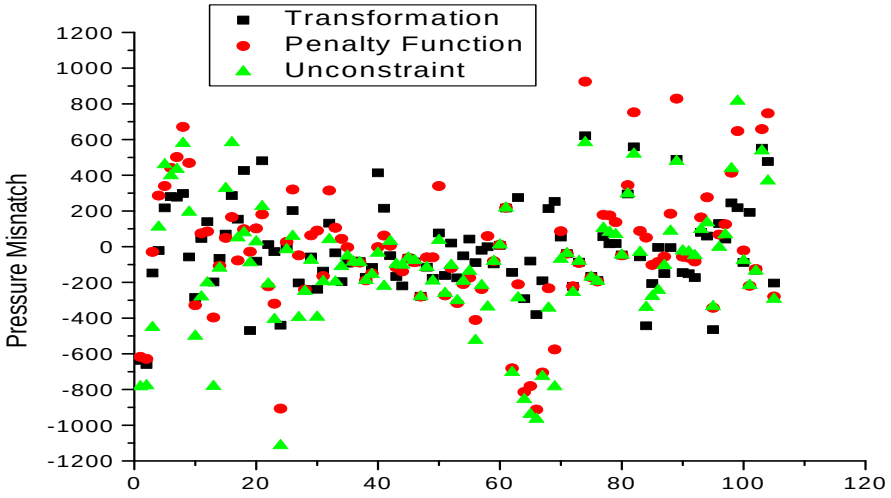


Figure 15: Comparison of Pressure Mismatches for the Tengiz Field.

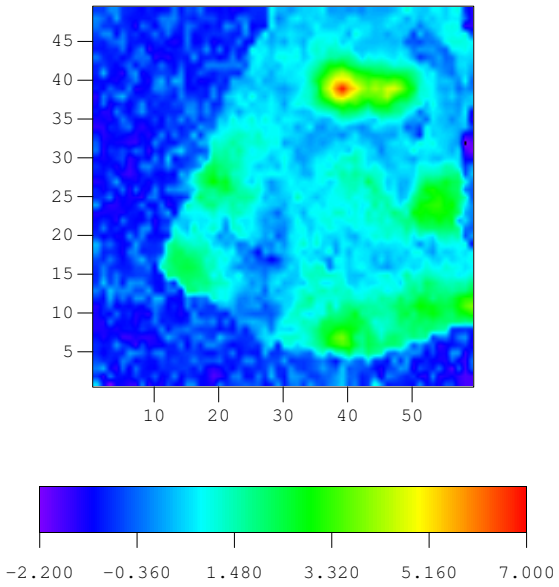


Figure 16: Log-Permeability of Layer 3 for Tengiz, Unconstrained Algorithm.

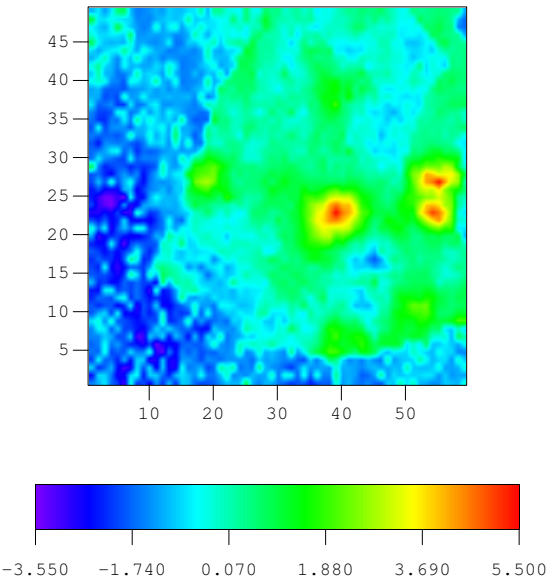


Figure 17: Log-Permeability of Layer 4 for Tengiz, Unconstrained Algorithm.

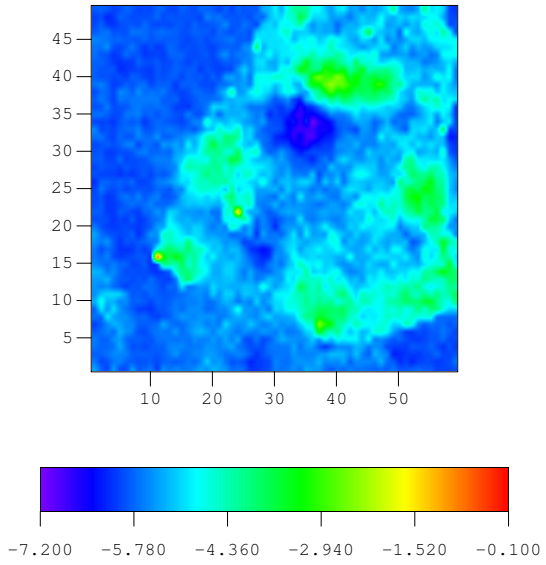


Figure 18: Vertical Log-Permeability of Layer 3 for Tengiz, Unconstrained Algorithm.

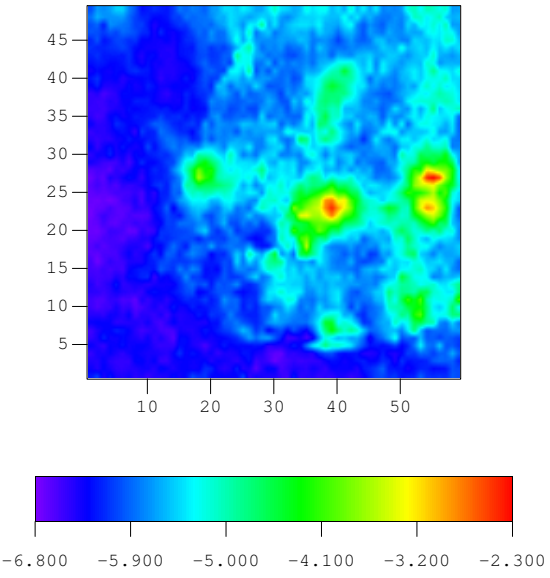


Figure 19: Vertical Log-Permeability of Layer 4 for Tengiz, Unconstrained Algorithm.

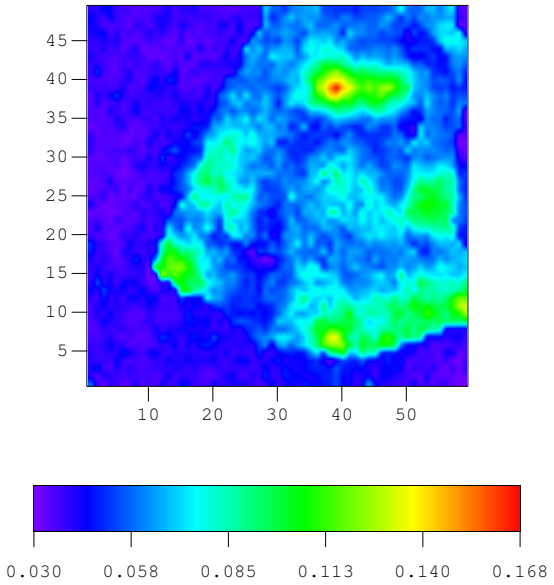


Figure 20: Porosity of Layer 3 for Tengiz, Unconstrained Algorithm

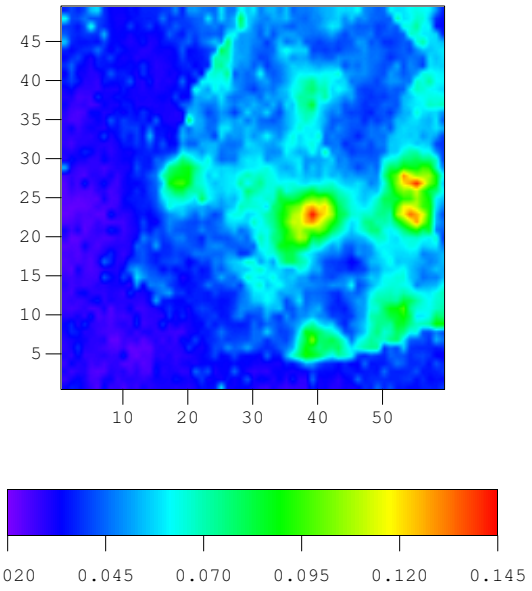


Figure 21: Porosity of Layer 4 for Tengiz, Unconstrained Algorithm.

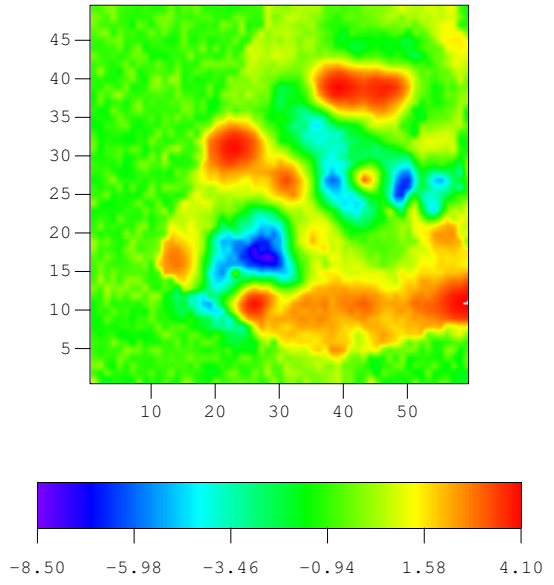


Figure 22: Log-Permeability of Layer 3 for Tengiz, Log-Transformation Algorithm.

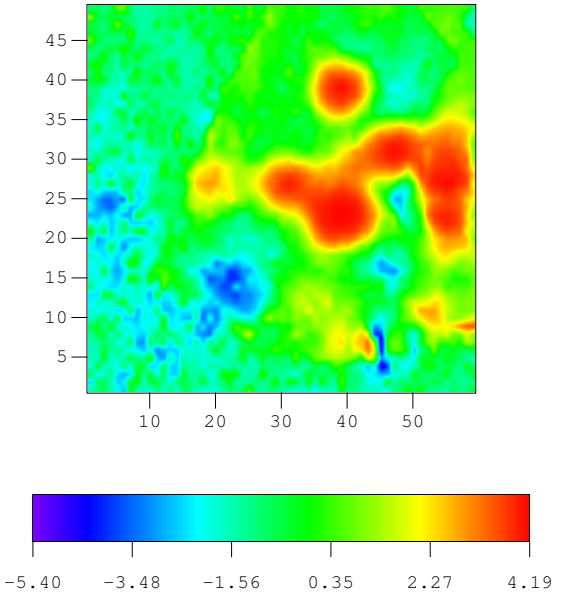


Figure 23: Log-Permeability of Layer 4 for Tengiz, Log-Transformation Algorithm.

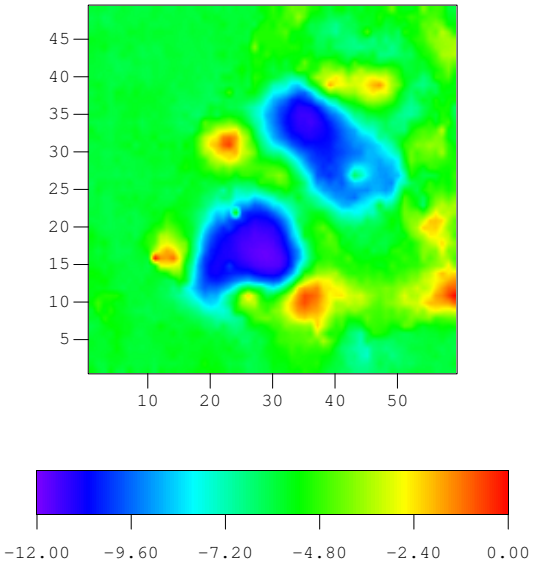


Figure 24: Vertical Log-Permeability of Layer 3 for Tengiz, Log-Transformation Algorithm.

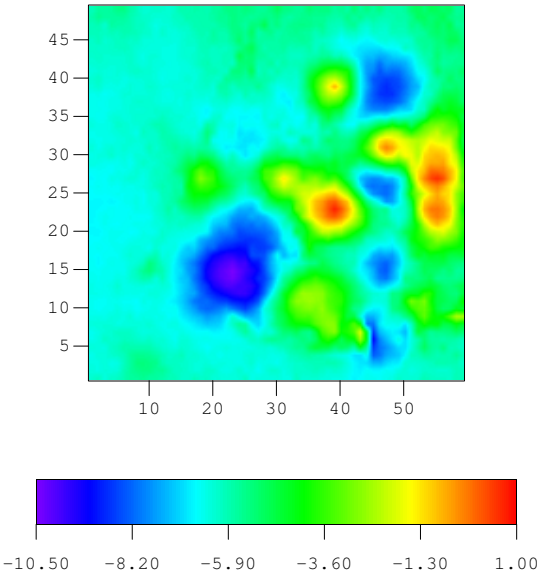


Figure 25: Vertical Log-Permeability of Layer 4 for Tengiz, Log-Transformation Algorithm.

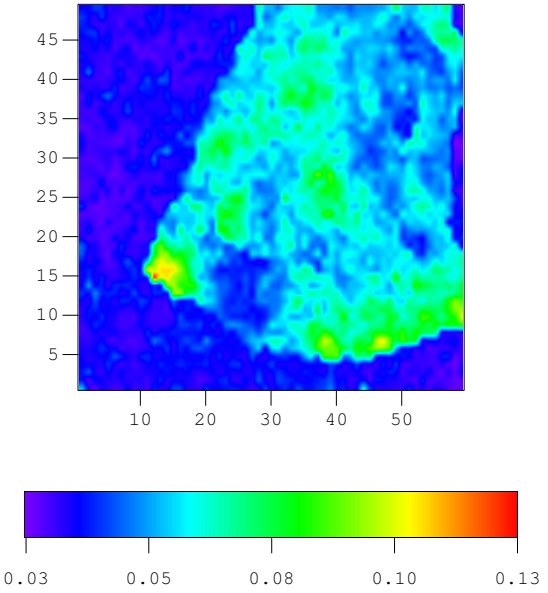


Figure 26: Porosity of Layer 3 for Tengiz, Log-Transformation Algorithm.

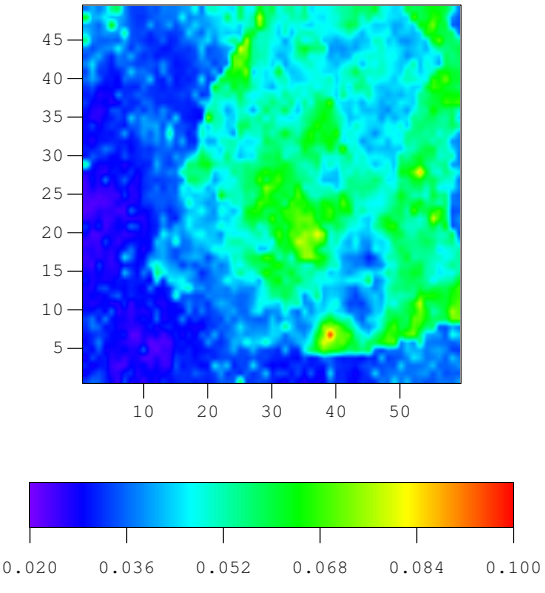


Figure 27: Porosity of Layer 4 for Tengiz, Log-Transformation Algorithm.

TIME LAPSE SEISMIC DATA

Although inverse problems in general, and automatic history matching problems in particular, are underdetermined, results derived from small amounts of data still provide estimates that are less than satisfying. It is clearly beneficial to make use of some type of "space-dense" data to improve the resolution of the estimate in the gridblocks far away from well locations. Of all usual data related to petroleum engineering, seismic data is the most promising candidate for improved spatial coverage. In a preliminary study (Dong and Oliver, 2002), we have showed that seismic impedance change data directly reflects the elastic properties of reservoir and are sensitive to permeability and porosity. We believe that seismic impedance change data will often provide useful constraints to our history matching procedure.

Since automatic history matching can ultimately be reduced to a minimization problem whose objective function includes both model mismatch and data mismatch parts, the choice of an efficient optimization method is very important, especially when processing large-scale problems. For our applications, we only consider gradient based optimization methods because non-gradient based optimization methods require far too many iterations to be of practical usefulness. Among those gradient based methods, some require the formation of the Hessian matrix as well as computation of the gradient of objective function. Formation of the Hessian is very expensive to compute in large-scale problems, so we have eliminated them from consideration. Instead, we use methods that only require the computation of the gradient of the objective function. According to Zhang and Reynolds (2002b), of the methods they evaluated, the limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method is likely to be the most successful method for large-scale history matching problems. Thus, for our history-matching problem with both production data and seismic impedance change data, we have only used LBFGS method for the minimization.

It is usually possible to assume that the errors in production data are uncorrelated. Because of the way the seismic impedance data are derived, it is not possible to make the same assumption of measurement error independence for the seismic impedance change data, which means that the part of the C_D matrix attributed to seismic impedance data will almost

certainly not be diagonal. When we apply our method to field problems, such correlation among seismic impedance change data will have to be carefully evaluated (Aannonsen et al., 2002).

Automatic History Matching

The easiest way to include seismic impedance change data is to assume that one set of data is available before production begins, and the other set is available after a period of production. Then, for these two sets of seismic impedance data, it is necessary to compute their derivatives with respect to primary variables and model parameters, which depends on the relationships of impedance to saturations and pressure. In this work, we use Gassmann (1951) and Han (1986) equations. We will summarize the Gassmann and Han equations and their derivatives with respect to primary variables and model parameters in this report. The impedance Z is given by

$$Z = \rho V_p = \sqrt{\rho K + \frac{4}{3}\rho^2 V_s^2}\,, \tag{47}$$

where, K is the bulk modulus, ρ is bulk density and V_s is shear wave velocity. All of them can be computed by Gassman and Han equations and poro-elastic properties as

$$\rho = (\rho_o S_o + \rho_w S_w + \rho_g S_g)\phi + (1-\phi)\rho_{\text{solid}}\,, \tag{48}$$

$$K = K_{\text{grain}} \frac{K_{\text{frame}} + Q}{K_{\text{grain}} + Q}\,, \tag{49}$$

$$Q = \frac{K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{frame}}\right)}{\phi\left(K_{\text{grain}} - K_{\text{fluid}}\right)}\,, \tag{50}$$

$$\log_{10} K_{\text{frame}} = \log_{10} K_{\text{grain}} - 4.25\phi\,, \tag{51}$$

$$K_{\text{grain}} = \frac{1}{2}\left[\gamma K_c + (1-\gamma)K_s + \frac{K_s K_c}{K_s\gamma + K_c(1-\gamma)}\right]\,, \tag{52}$$

$$\frac{1}{K_{\text{fluid}}} = \frac{S_w}{K_w} + \frac{S_g}{K_g} + \frac{S_o}{K_o}\,, \tag{53}$$

$$V_s = 3520.0 - 4910.0\phi - 1890.0\gamma\,. \tag{54}$$

Using the chain rule, the derivatives of impedance with respect to pressure and saturation are

$$\frac{\partial Z}{\partial P} = \frac{1}{2}\left(\rho K + \frac{4}{3}\rho^2 V_s^2\right)^{-1/2} \times \left(K\frac{\partial \rho}{\partial P} + \frac{8}{3}V_s^2\rho\frac{\partial \rho}{\partial P}\right)\,, \tag{55}$$

where

$$\frac{\partial \rho}{\partial P} = \phi\left(S_o\frac{\partial \rho_o}{\partial P} + S_w\frac{\partial \rho_w}{\partial P} + S_g\frac{\partial \rho_g}{\partial P}\right)\,, \tag{56}$$

and

$$\frac{\partial Z}{\partial S_w} = \frac{1}{2}\left(\rho K + \frac{4}{3}\rho^2 V_s^2\right)^{-1/2} \times \left(K\frac{\partial \rho}{\partial S_w} + \rho\frac{\partial K}{\partial S_w} + \frac{8}{3}V_s^2\rho\frac{\partial \rho}{\partial S_w}\right), \tag{57}$$

where

$$\frac{\partial \rho}{\partial S_w} = \phi\left(\rho_w - \rho_o\right), \tag{58}$$

$$\begin{aligned}\frac{\partial K}{\partial S_w} = &\frac{\left(\left(K_{\text{grain}}^2 - K_{\text{grain}}K_{\text{frame}}\phi - K_{\text{grain}}K_{\text{frame}}\right)\frac{\partial K_{\text{fluid}}}{\partial S_w}\right)}{\left(K_{\text{grain}}^2\phi + K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\right)} \\ &- \frac{\left(\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\frac{\partial K_{\text{fluid}}}{\partial S_w}\right)}{\left(K_{\text{grain}}^2\phi + K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\right)} \\ &\times \frac{\left(K_{\text{grain}}^2K_{\text{frame}}\phi + K_{\text{fluid}}\left(K_{\text{grain}}^2 - K_{\text{grain}}K_{\text{frame}}\phi - K_{\text{grain}}K_{\text{frame}}\right)\right)}{\left(K_{\text{grain}}^2\phi + K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\right)},\end{aligned} \tag{59}$$

and

$$\frac{\partial K_{\text{fluid}}}{\partial S_w} = \frac{\frac{1}{K_o} - \frac{1}{K_w}}{\left(\frac{S_w}{K_w} + \frac{1-S_w-S_g}{K_o} + \frac{S_g}{K_g}\right)^2}, \tag{60}$$

$$\frac{\partial Z}{\partial S_g} = \frac{1}{2}\left(\rho K + \frac{4}{3}\rho^2 V_s^2\right)^{-1/2} \times \left(K\frac{\partial \rho}{\partial S_g} + \rho\frac{\partial K}{\partial S_g} + \frac{8}{3}V_s^2\rho\frac{\partial \rho}{\partial S_g}\right), \tag{61}$$

where

$$\frac{\partial \rho}{\partial S_g} = \phi\left(\rho_g - \rho_o\right), \tag{62}$$

$$\begin{aligned}\frac{\partial K}{\partial S_g} = &\frac{\left(\left(K_{\text{grain}}^2 - K_{\text{grain}}K_{\text{frame}}\phi - K_{\text{grain}}K_{\text{frame}}\right)\frac{\partial K_{\text{fluid}}}{\partial S_g}\right)}{\left(K_{\text{grain}}^2\phi + K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\right)} \\ &- \frac{\left(\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\frac{\partial K_{\text{fluid}}}{\partial S_g}\right)}{\left(K_{\text{grain}}^2\phi + K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\right)} \\ &\times \frac{\left(K_{\text{grain}}^2K_{\text{frame}}\phi + K_{\text{fluid}}\left(K_{\text{grain}}^2 - K_{\text{grain}}K_{\text{frame}}\phi - K_{\text{grain}}K_{\text{frame}}\right)\right)}{\left(K_{\text{grain}}^2\phi + K_{\text{fluid}}\left(K_{\text{grain}} - K_{\text{grain}}\phi - K_{\text{frame}}\right)\right)},\end{aligned} \tag{63}$$

and

$$\frac{\partial K_{\text{fluid}}}{\partial S_g} = \frac{\frac{1}{K_o} - \frac{1}{K_g}}{\left(\frac{S_w}{K_w} + \frac{1-S_w-S_g}{K_o} + \frac{S_g}{K_g}\right)^2}, \tag{64}$$

$$\frac{\partial Z}{\partial \phi} = \frac{1}{2}\left(\rho K + \frac{4}{3}\rho^2 V_s^2\right)^{-1/2} \times \left(\left(K + \frac{8}{3}\rho V_s^2\right)\frac{\partial \rho}{\partial \phi} + \rho\frac{\partial K}{\partial \phi} + \frac{8}{3}\rho^2 V_s\frac{\partial V_s}{\partial \phi}\right), \tag{65}$$

where

$$\frac{\partial V_s}{\partial \phi} = -4910.0\,, \tag{66}$$

$$\frac{\partial \rho}{\partial \phi} = \rho_o\left(1 - S_w - S_g\right) + \rho_w S_w + \rho_g S_g - \rho_{\text{solid}}\,, \tag{67}$$

$$\begin{aligned}\frac{\partial K}{\partial \phi} = & \frac{\left(\left(K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}}\right)\left(\phi\frac{\partial K_{\text{frame}}}{\partial \phi} + K_{\text{frame}}\right) - K_{\text{fluid}}K_{\text{grain}}\frac{\partial K_{\text{frame}}}{\partial \phi}\right)}{\phi\left(K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}}\right) + K_{\text{fluid}}K_{\text{grain}} - K_{\text{fluid}}K_{\text{frame}}} \\ & - \frac{\left(K_{\text{frame}}\phi\left(K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}}\right) + K_{\text{fluid}}K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}}K_{\text{frame}}\right)}{\left(\phi\left(K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}}\right) + K_{\text{fluid}}K_{\text{grain}} - K_{\text{fluid}}K_{\text{frame}}\right)} \\ & \times \frac{\left(K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}} - K_{\text{fluid}}\frac{\partial K_{\text{frame}}}{\partial \phi}\right)}{\left(\phi\left(K_{\text{grain}}^2 - K_{\text{fluid}}K_{\text{grain}}\right) + K_{\text{fluid}}K_{\text{grain}} - K_{\text{fluid}}K_{\text{frame}}\right)}\,,\end{aligned} \tag{68}$$

and

$$\frac{\partial K_{\text{frame}}}{\partial \phi} = -4.25\ln(10)K_{\text{grain}}10^{-4.25}\,. \tag{69}$$

There is no direct dependance between seismic impedance and permeability k, so the sensitivity of seismic impedance with respect to permeability depends entirely on the indirect effect of permeability on pressure and saturation.

LBFGS Method

Previous studiess in TUPREP have showed that BFGS method is the most successful quasi-Newton method for history matching. However, its drawback is that it needs to store the Hessian matrix approximation, which will be impractical when large scale models are considered. The alternative is the limited memory BFGS (LBFGS) method of Nocedal (1980). LBFGS method only requires storage of a few vectors and uses these vectors to implicitly construct Hessian matrix approximation. A detailed discussion of the application of the LBFGS method to large scale history-matching problems can be found in Zhang and Reynolds (2002b).

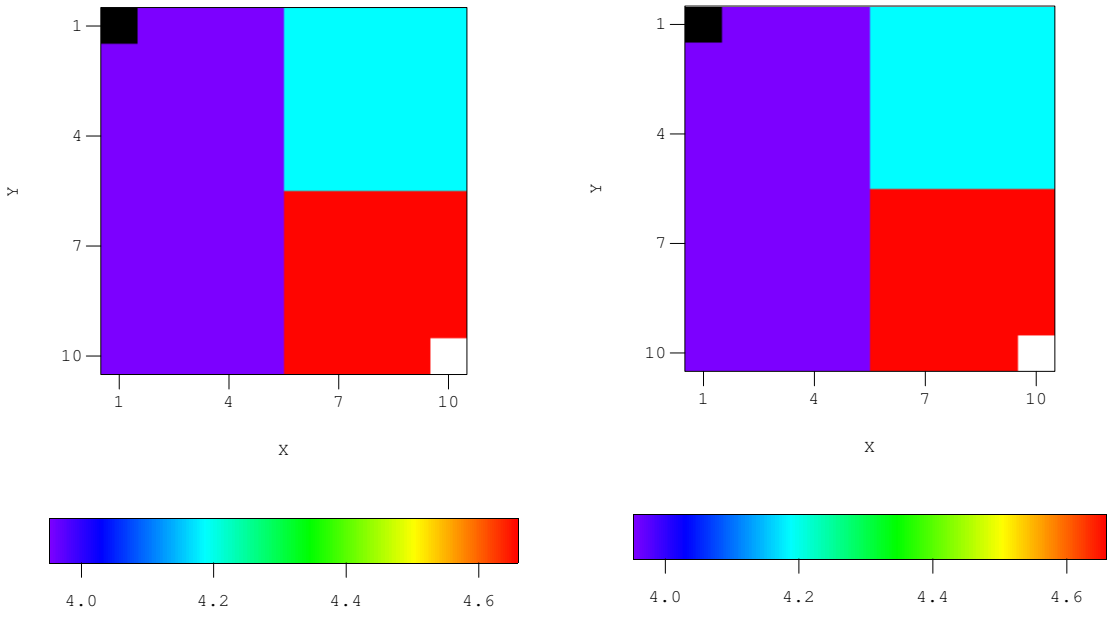
Results

To test the effect of integration of both seismic impedance change data and production data, we used two models. One is a small synthetic model and the other one is a semi-synthetic model created from a middle east oil field. In the following sections, we will discuss them in detail.

Synthetic Model

This small synthetic model has two layers. Each layer has 10×10 gridblocks with size equal to 40 ft. The vertical size of each gridblock is 30 ft. The only model parameter we adjust is log of horizontal permeability. For vertical permeability, we use a multiplier, here equal to 0.1, and for porosity, we fix it at the true value 0.2 in every gridblock. To compare matching results before integration of seismic impedance change data and after such integration, we divide each layer into three different zones. In each zone, horizontal permeability is homogeneous. However, in each zone, the horizontal permeability has a different value. Since there are very obvious edges among these three zones, it is easy to decide if one matching result is better than another. If a method is performing well, such edges would presumably be clearer. Here, we only use bottom hole pressure P_{wf} as production data.

The true values of $\ln(k)$ in three zones are 4.0, 4.2 and 4.6, which can be seen from Fig. 28. The white point in lower-right corner denotes an injection well and the black one is



(a) Top Layer (b) Bottom Layer

Figure 28: True log horizontal permeability ($\ln(k)$) field

a production well. Both of them perforate thoroughly in two layers. This is a three-phase problem.

Matching Production Data Only

In each well, we use 10 pressure data to do the history matching so that totally we have 20 pressure data to adjust 200 model parameters. We use a homogeneous value, 4.0, as the initial guess and the prior model. Final matching results are in Fig. 29 and Fig. 30. From

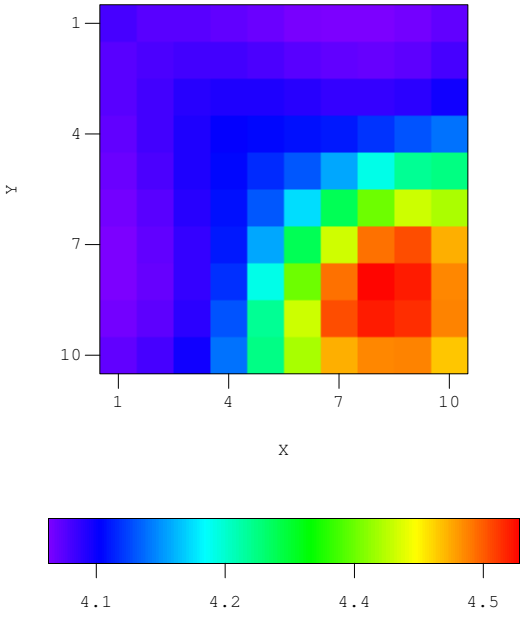


Figure 29: Estimate of log horizontal permeability field in top layer by integration of production data only

Fig. 29 and Fig. 30, we can see that from production data alone, it is not possible to obtain a good estimate of the permeability field. The boundary between the blue area and the red area is more like arc than straight line, which simply reflects the symmetry of the problem and the area of water saturation change from the injection well in that corner.

Combination of Seismic and Production Data

In this section, we evaluate the improvement in the estimate after the integration of seismic impedance change data. Results for the same problem can be seen in Fig. 31 and Fig. 32. Compared to Fig. 29, Fig. 30 and Fig. 28, the permeability estimate from integration of both seismic impedance change data and production data is better than the estimate obtained from only using production data. Moreover, it is much closer to the true field, especially its boundary of the homogeneous regions is much clearer. The reason for the improvement

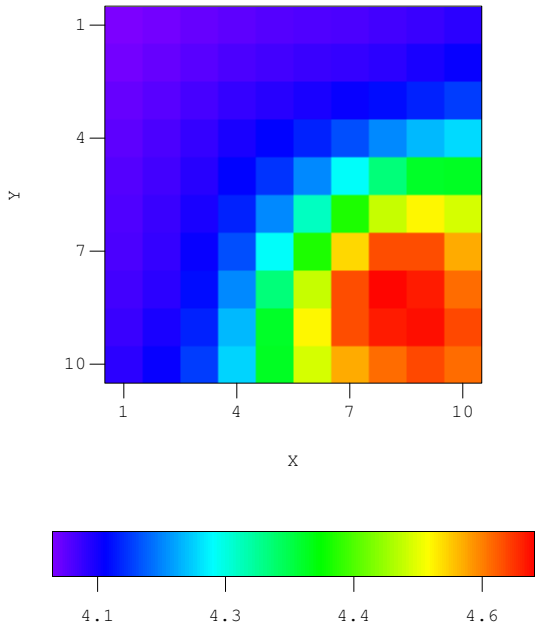


Figure 30: Estimate of log horizontal permeability field in bottom layer by integration of production data only

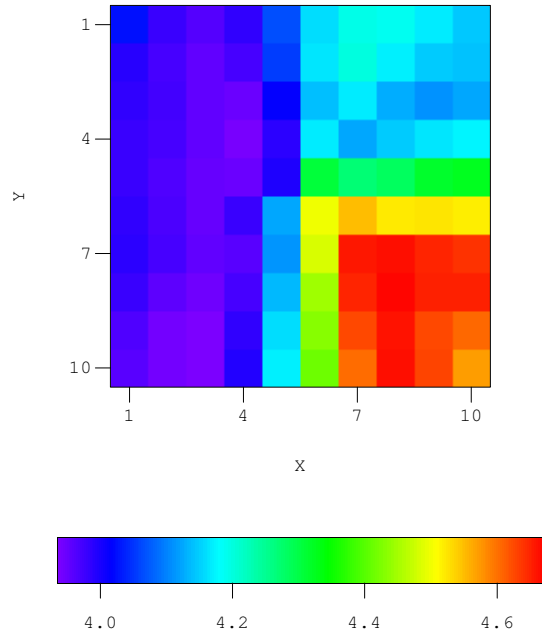


Figure 31: Estimate of log horizontal permeability field in top layer by integration of both production data and seismic impedance change data

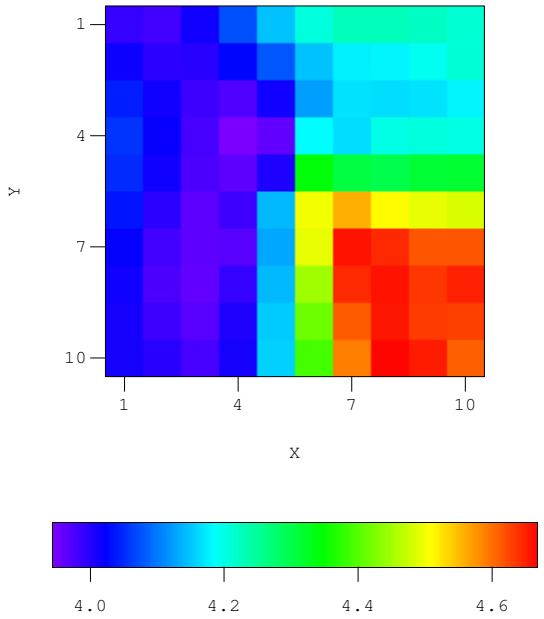


Figure 32: Estimate of log horizontal permeability field in bottom layer by integration of both production data and seismic impedance change data

is that the spatial density of seismic impedance change data provides more constraints in gridblocks that are far from well locations. Obviously, seismic impedance change data makes the possible solution space narrower and can provide much better estimates of properties, which is what we predicted when we performed the preliminary investigation work. In summary, adding seismic impedance change data into automatic history matching is feasible and provides improved estimates.

Semi-Synthetic Model

Because our final goal is to apply the method to a field problem, it is important to test the method on real field data. As an intermediate step, before application to real field data, we are applying the method to a large "field-scale" problem rather than just using it on small synthetic model. One possible candidate is the simulation model from a field provided by an industry member to TUPREP. When TUPREP obtained that data, the initial aim was to test automatic history matching method on a large scale single-phase real field problem. At this time, we intend to investigate if seismic impedance change data can give a reasonable estimates of properties in history matching problems. This goal requires comparison with a true field. We used the first five layers of the reservoir created by the company geoscientists

as the true geological model. It would presumably be unknown, except for observations at well locations. Using a covariance estimated from the model, we created a new synthetic model by Sequential Gaussian Simulation (SGS). A comparison of the supplied model with our model can be seen from Table 3.

Parameters	**True Field**	**Semi-Synthetic Model**
True Model	N/A	Synthetic model
Prior Model	Synthetic model with facies change	Generated by SGS
Well Completion	Partially perforated	Fully perforated
Initial Reservoir Pressure	11950 Psi	4000 Psi
Bubble Point Pressure	3586 Psi	3586 Psi
Number of Layers	9	5
Gridblocks in Each Layer	59×49	59×49

Table 3: Comparison between real field and semi-synthetic model

The real field has a very high initial reservoir pressure and relatively low bubble point pressure. Under such conditions, it remains single-phase even after a long production period. A deep, single-phase reservoir, would not be a good candidate for 4D seismic. Thus, in order to create a more realistic example, we changed the initial reservoir pressure to be slightly above the bubble point pressure, to ensure that free gas will evolve soon after production begins.

Creation of Prior Model

Sequential Gaussian Simulation was used to create the prior porosity field, which also served as the initial guess. The prior horizontal permeability field, was generated directly from the porosity field using a functional relationship. That correlation equation is generated from cross plot of porosity and horizontal permeability in well locations, which can be seen in Fig. 33. Using regression, we have the relationship as

$$\ln k_h = 2.41073 - 7.3652 \times \exp\left(-\frac{\phi}{0.04419}\right) . \tag{70}$$

The prior horizontal permeability field was computed directly from Eq. 70 once we had simulated the porosity field. The relationship between vertical and horizontal permeability was also estimated from a crossplot. A satisfactory relationship is provided by

$$k_v = a \times k_h \, , \tag{71}$$

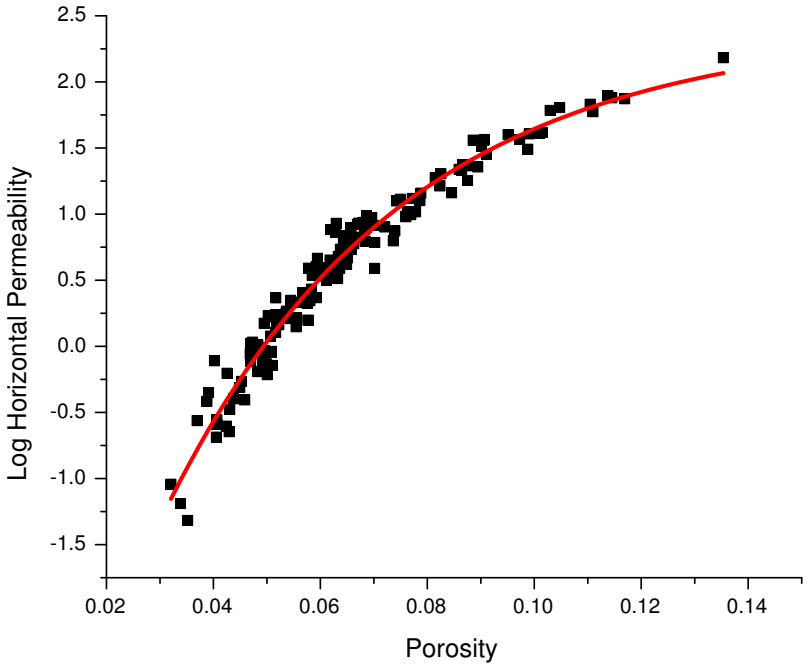


Figure 33: Correlation between porosity and log horizontal permeability in well locations

where, a is a constant multiplier. In our semi-synthetic model, we assume that $a = 0.002$ to create the vertical permeability field.

To summarize, we use the true model to create observation data, here seismic impedance change data. Then, we adjust both the prior porosity field and horizontal permeability field to match the data from the true field. We intend to match our observation data as well as possible. In this example, we do not use any production data, in order to assess the constraint the seismic impedance change data provides in such large scale problems.

The true horizontal permeability field in the first four layers can be seen in Fig. 34. The prior permeability field for the first four layers is in Fig. 35. The fifth layer of both true permeability field and prior field can be seen in Fig. 36 and Fig. 37. Similarly, true and prior porosity field of each layer are in Figs. 38, 39, 40 and 41. From true permeability and porosity field, we can see that there is an obvious discontinuity in properties between the right lower part and the left upper part of the reservoir in each layer. This is a depth effect, higher porosities and permeabilities occur at shallower depths. Most of the wells are also located in this area, which gives more gas after production than in the low permeability and porosity area. This difference makes the seismic impedance change quite different in these

two areas, which can be seen in Fig. 42 and Fig. 43. In the left region, because of lower gas saturation, the seismic impedance change value is low. The region on the right side has a higher value because of higher gas saturation. Moreover, with increase of depth, reservoir pressure becomes higher, which makes it more difficult for gas to come out, then seismic impedance change values become smaller with depth.

History Matching

In this report, we used only the seismic impedance change data—not the change in amplitude—and assumed that each gridblock had a seismic impedance change data so that the number of data was as same as the number of gridblocks. We adjusted both horizontal permeability and porosity, so the number of model parameters was twice the number of gridblocks. The minimization required 100 iteration loops, which is the maximum number of iterations allowed in our code. The square summation of data mismatch at the end of the iterations was lower than the total number of data, which implies that the resulting model was acceptable. The objective function behavior and data mismatch part decrease can be seen in Fig. 44(a) and Fig. 44(b). The matching results of horizontal permeability field and porosity field can be seen in Fig. 45, Fig. 46, Fig. 47 and Fig. 48.

From these maps, we make the following observations,

1. Seismic impedance change data provides useful constraints in history matching problem. Especially in large scale models, the use of seismic impedance change data can decrease uncertainty. The results for both permeability and porosity include features which are similar to the true model.

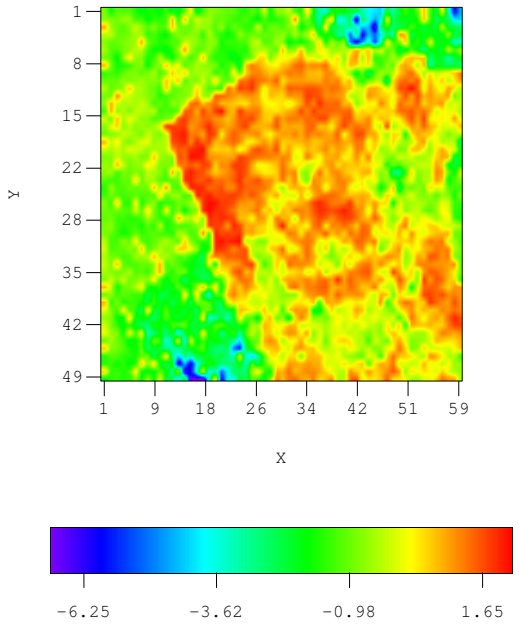
2. Estimates of the porosity field are better than the estimates for permeability. The reason is almost certainly that seismic impedance change data is more sensitive to porosity than to permeability.

3. From the top layer to the bottom layer, reservoir pressure increases, which means that there less free gas evolves in the deeper layers. This decrease in gas saturation results in smaller changes in seismic impedance in deeper layers. Thus, the results in deeper layers were not as good as the results in top layers. This can be observed clearly from Fig. 45 and Fig. 46.

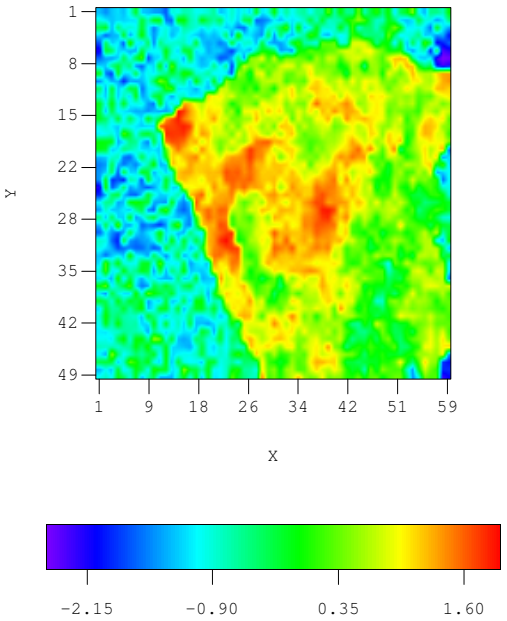
4. The properties in the upper left region of the simulation model do not change very much after integration of time-lapse seismic because there are almost no wells and the

depth is greater, which makes it more difficult to have gas accumulated there. Thus, we do not have significant seismic impedance change in that region.

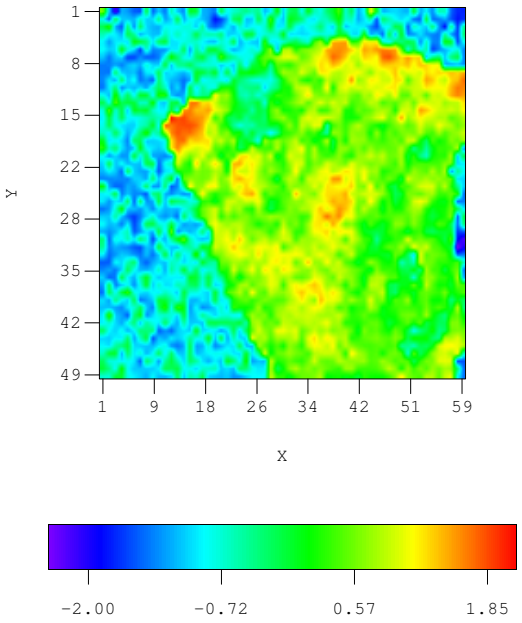
Finally, we note that integration of seismic impedance change data into automatic history matching seems to provide dramatically improved reservoir models, even when the data are noisy. The spatial density of the data appears to compensate for the sparsity of production data, especially in large scale models.



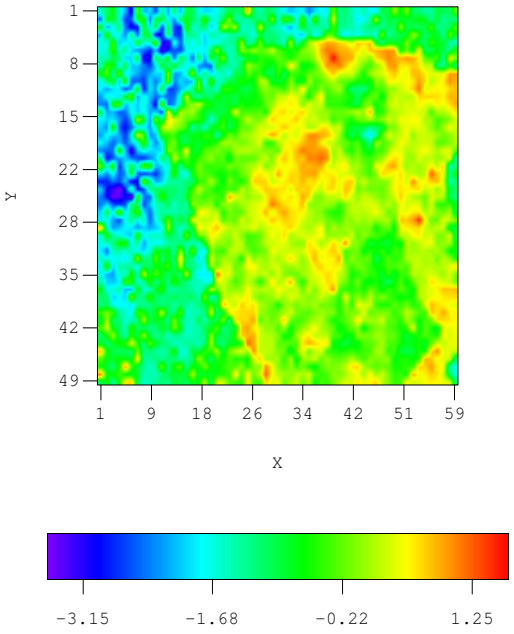
(a) Layer One



(b) Layer Two

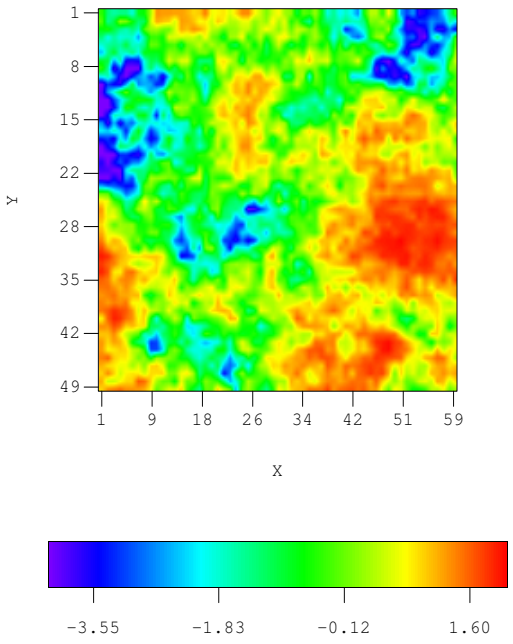


(c) Layer Three

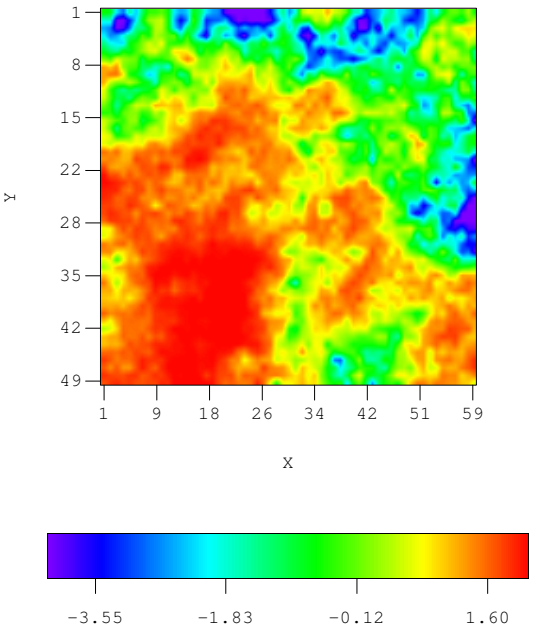


(d) Layer Four

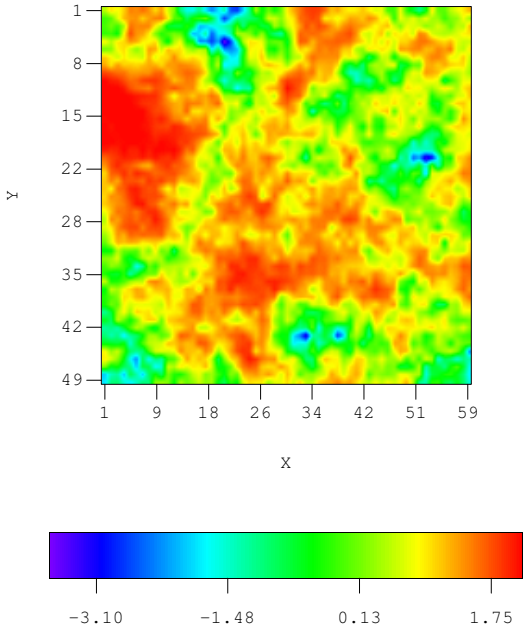
Figure 34: True log horizontal permeability field in first four layers



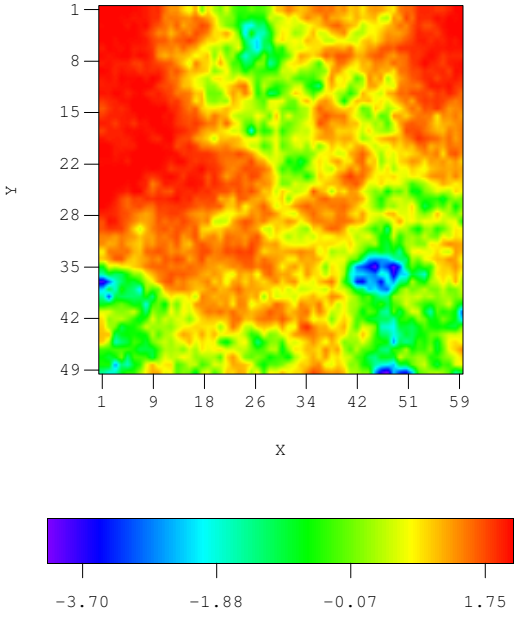
(a) Layer One



(b) Layer Two



(c) Layer Three



(d) Layer Four

Figure 35: Prior log horizontal permeability field in first four layers

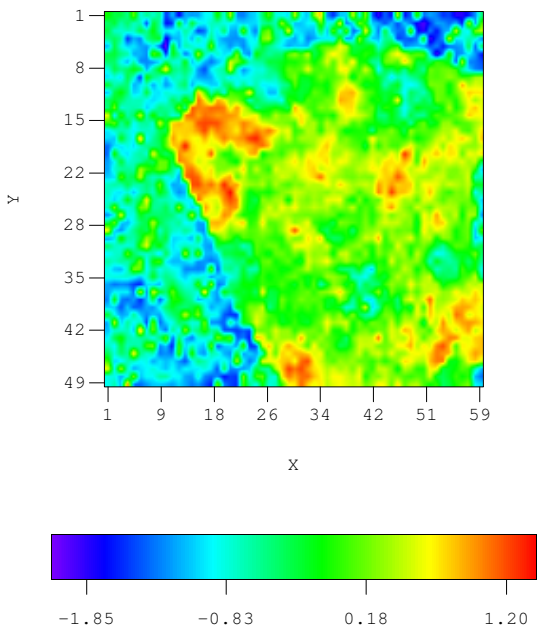


Figure 36: True log horizontal permeability field in the fifth layer

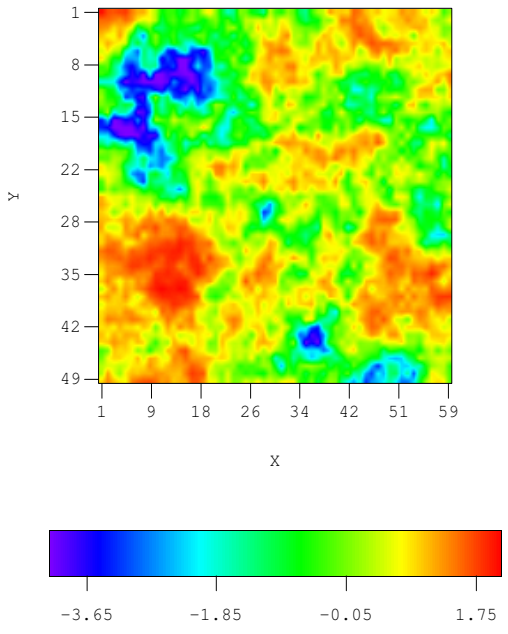
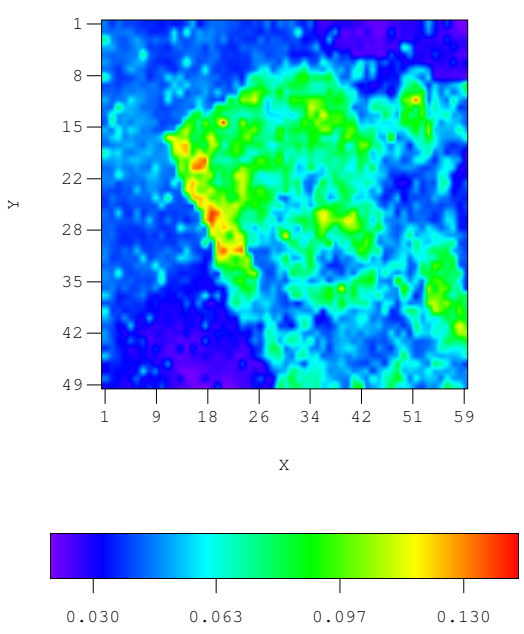
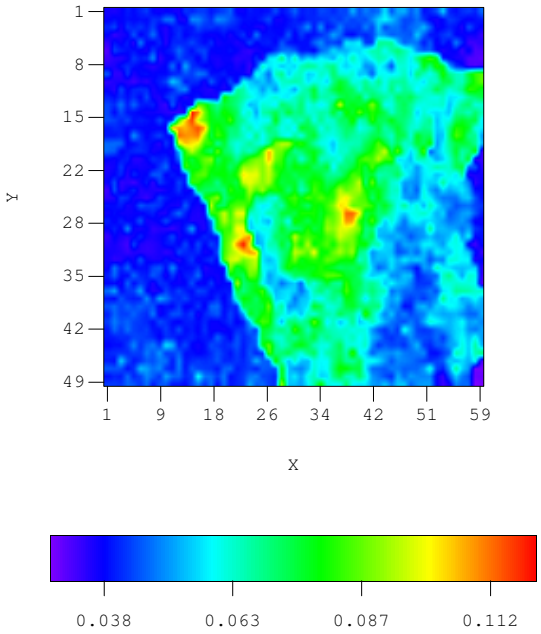


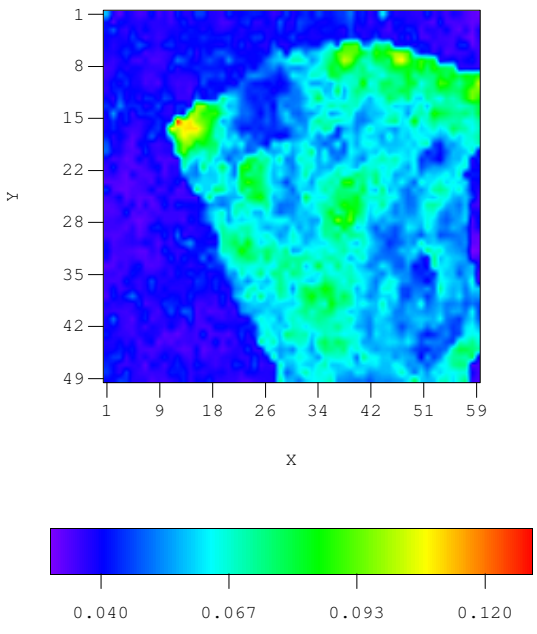
Figure 37: Prior log horizontal permeability field in the fifth layer



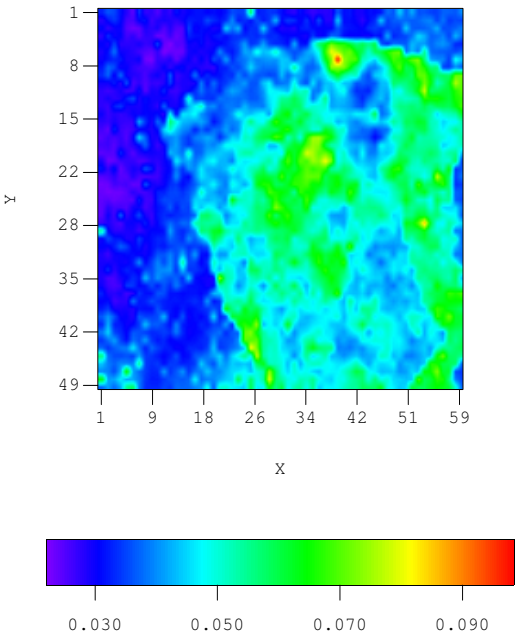
(a) Layer One



(b) Layer Two

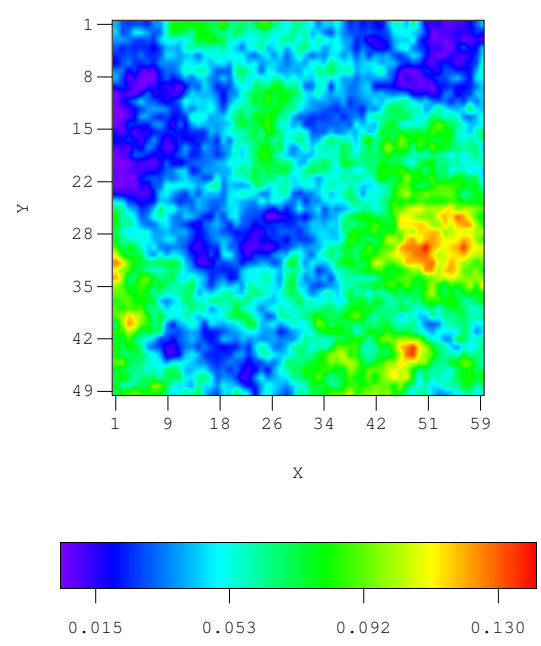


(c) Layer Three

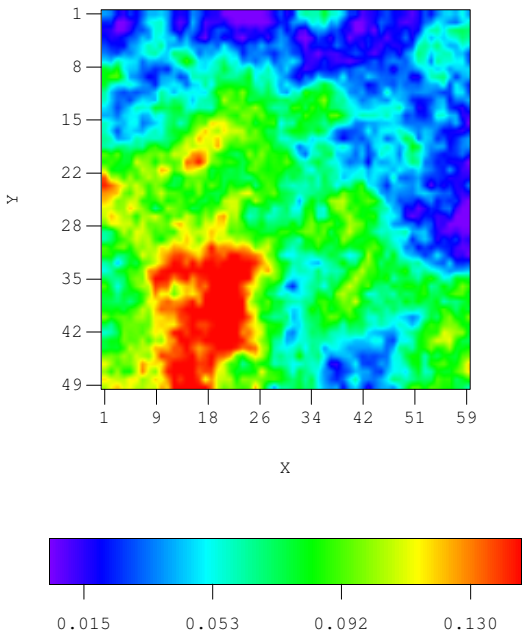


(d) Layer Four

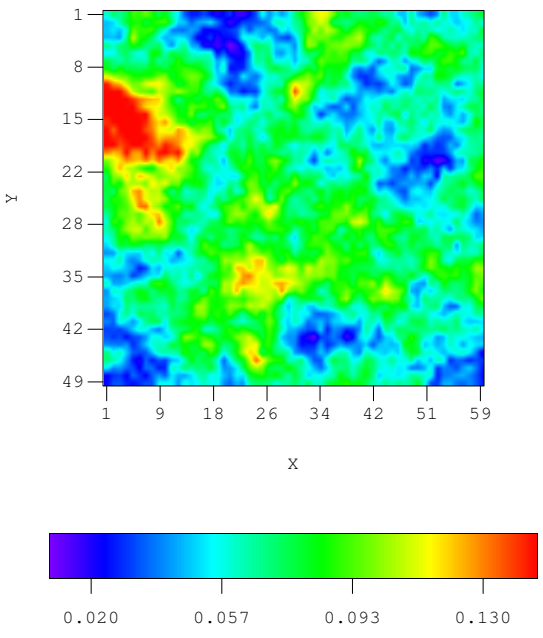
Figure 38: True porosity field in first four layers



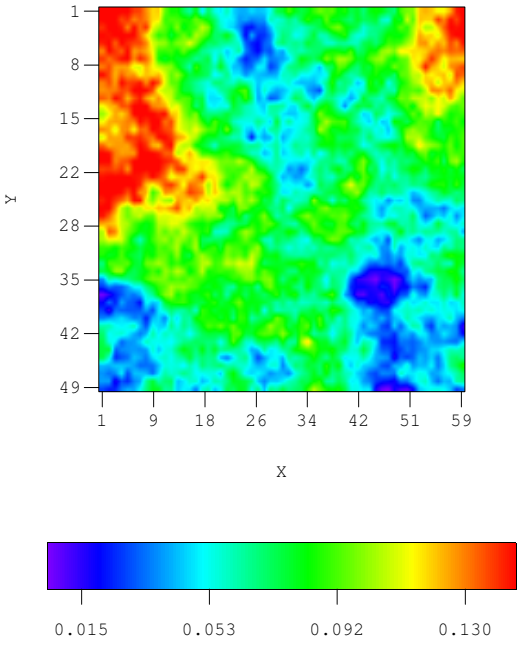
(a) Layer One



(b) Layer Two



(c) Layer Three



(d) Layer Four

Figure 39: Prior porosity field in first four layers

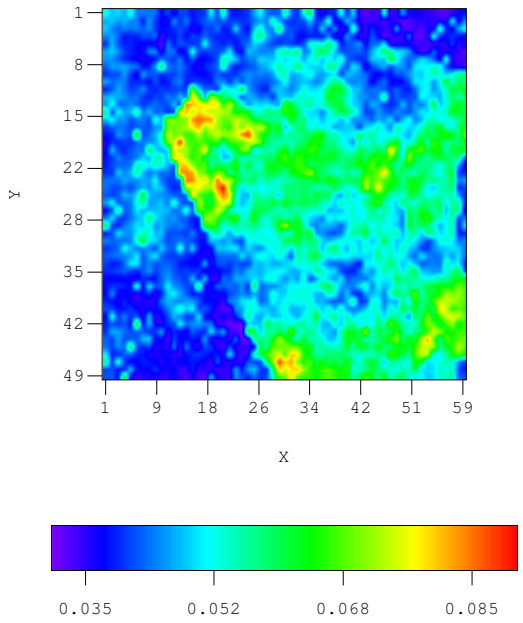


Figure 40: True porosity field in the fifth layer

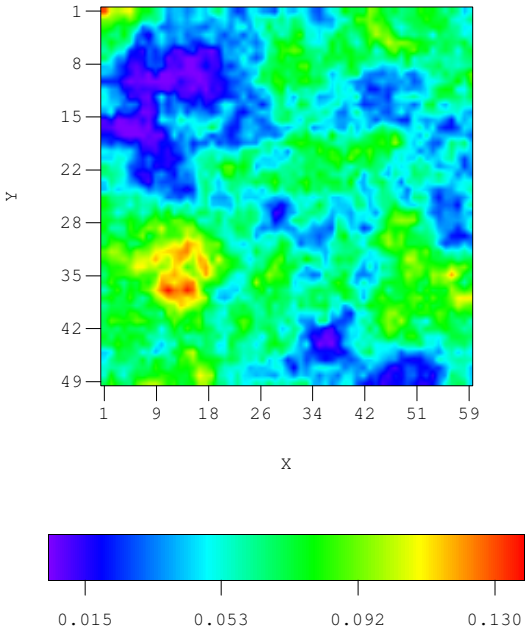
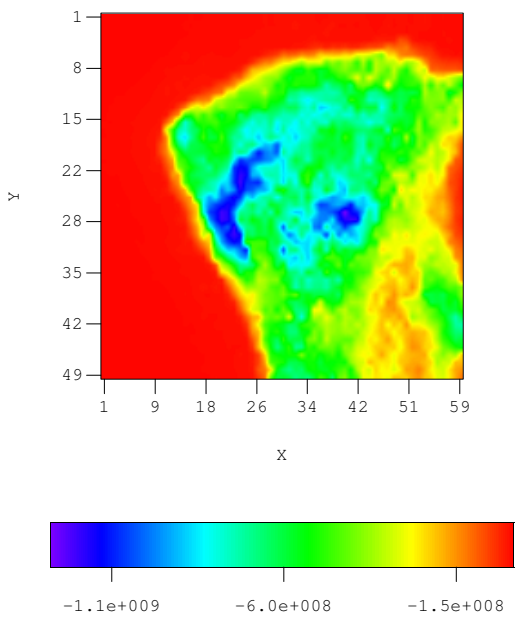
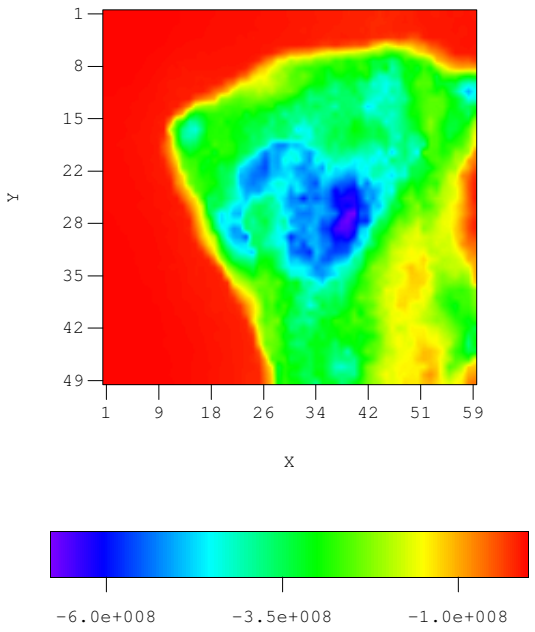


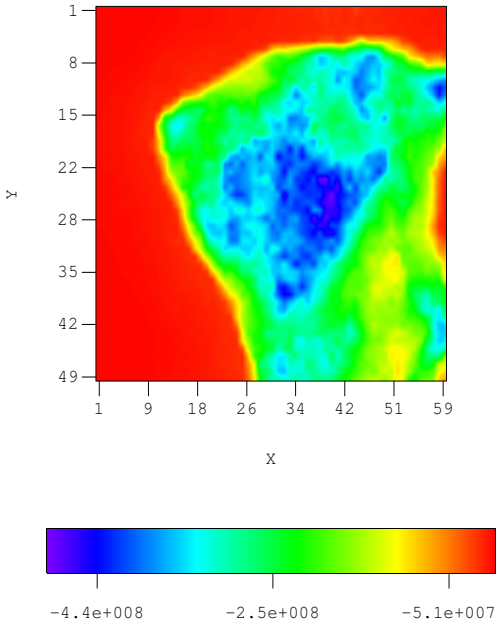
Figure 41: Prior porosity field in the fifth layer



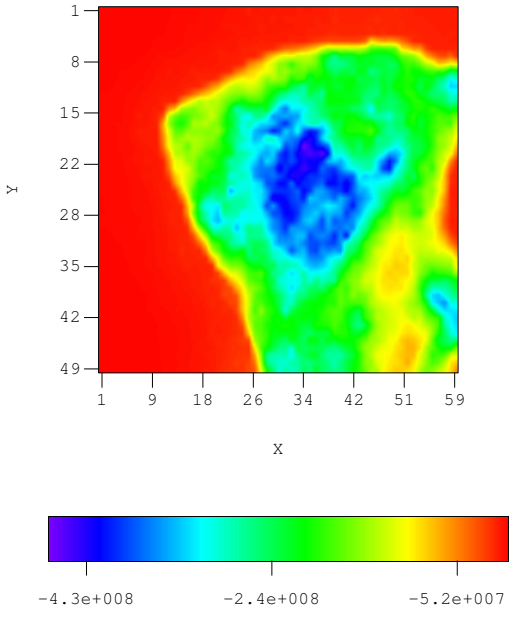
(a) Layer One



(b) Layer Two



(c) Layer Three



(d) Layer Four

Figure 42: Seismic impedance change in first four layers

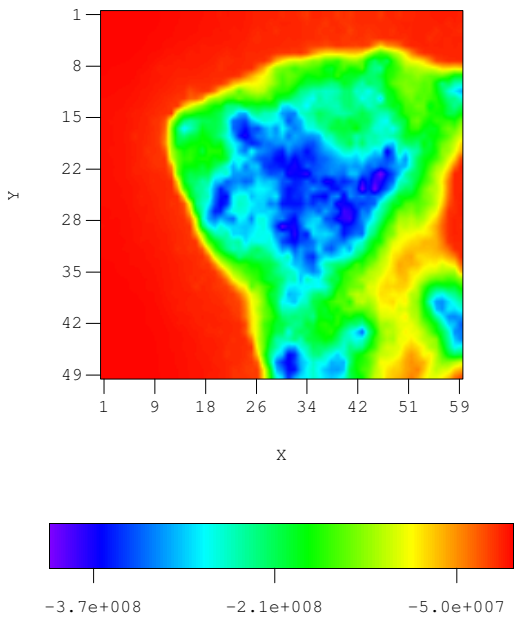
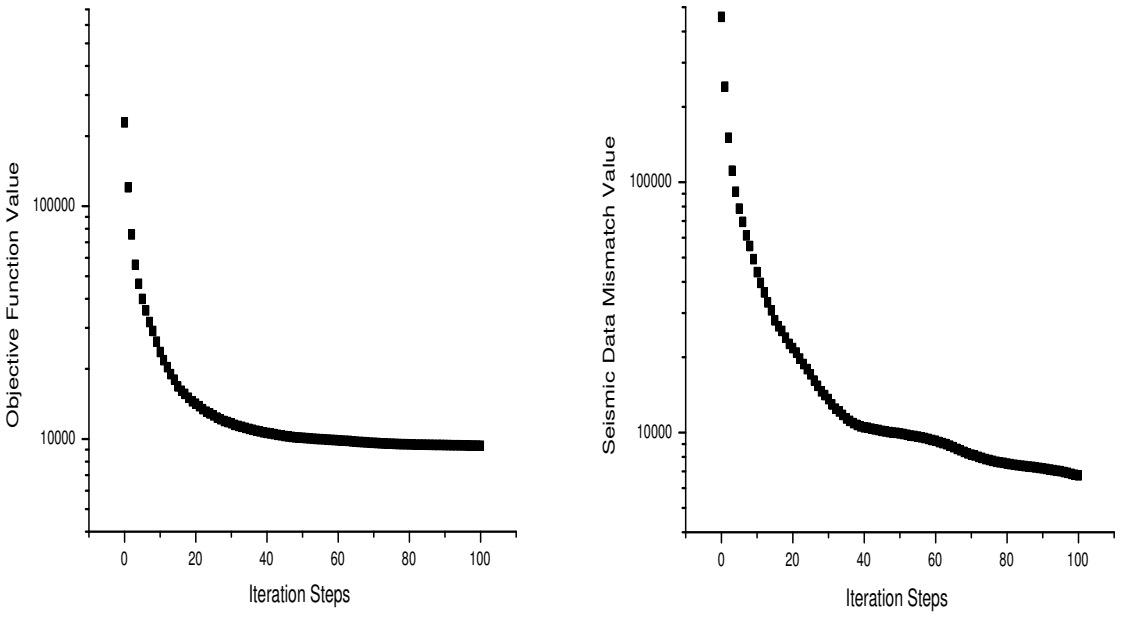


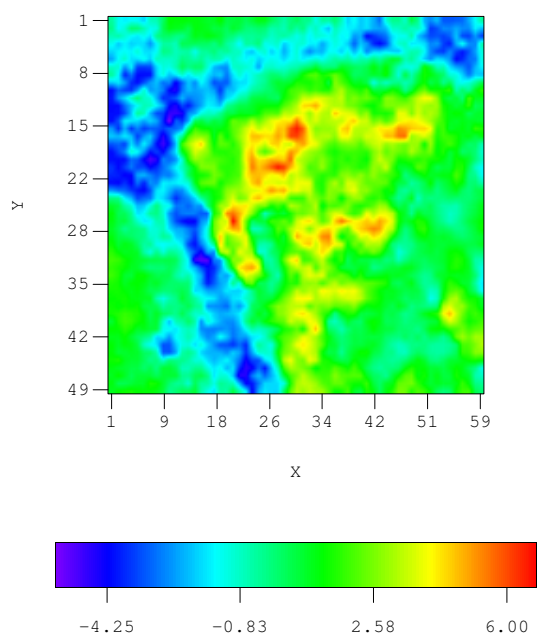
Figure 43: Seismic impedance change in the fifth layer



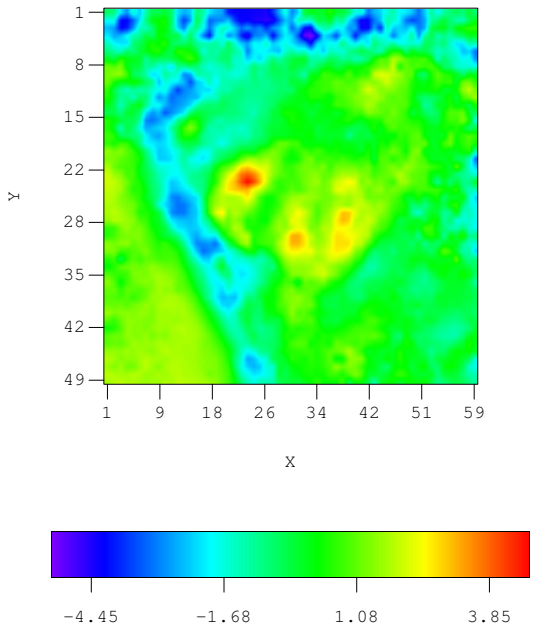
(a) Objective Function

(b) Seismic Data Mismatch

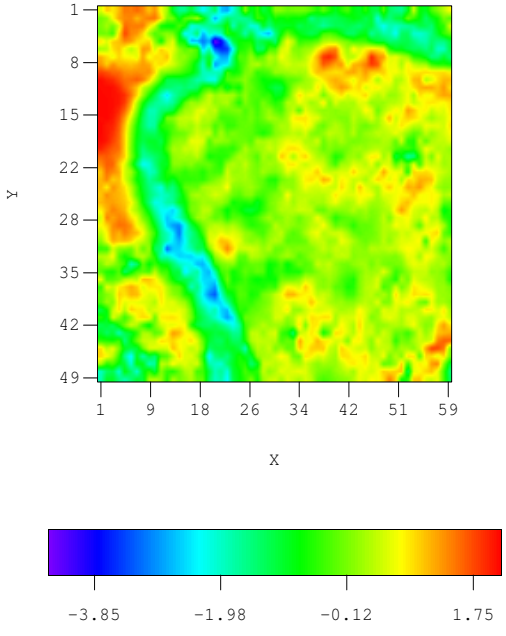
Figure 44: Objective function and seismic data mismatch decrease



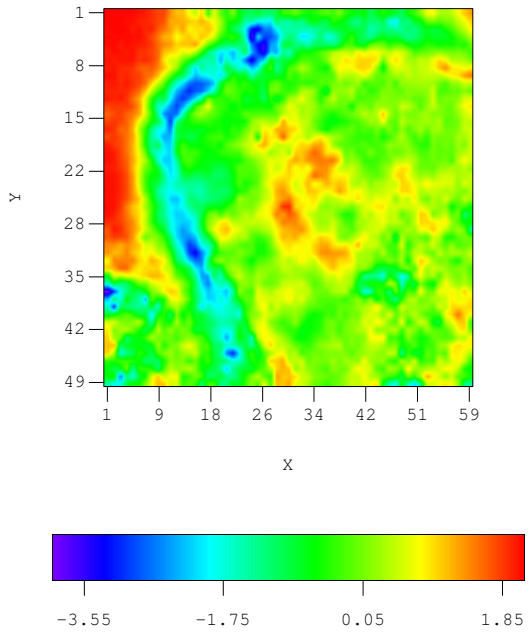
(a) Layer One



(b) Layer Two

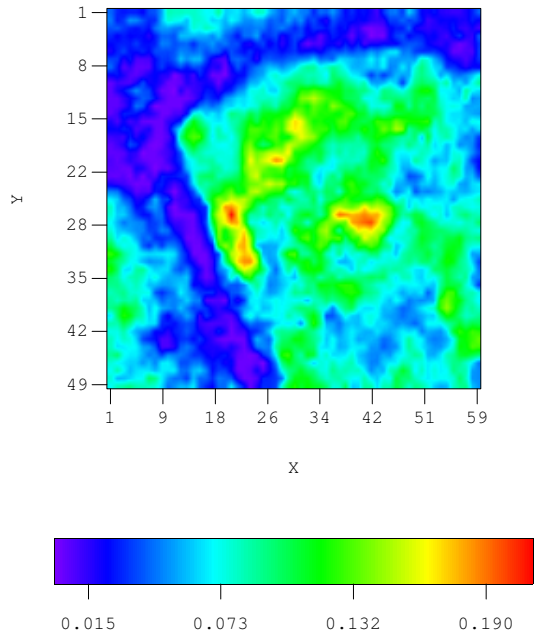


(c) Layer Three

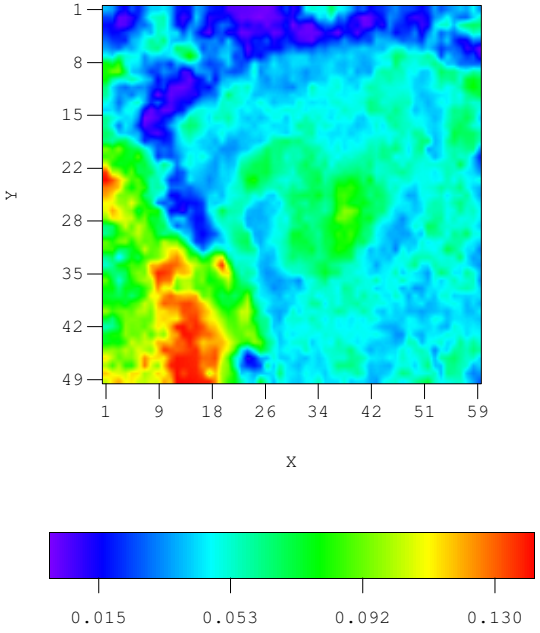


(d) Layer Four

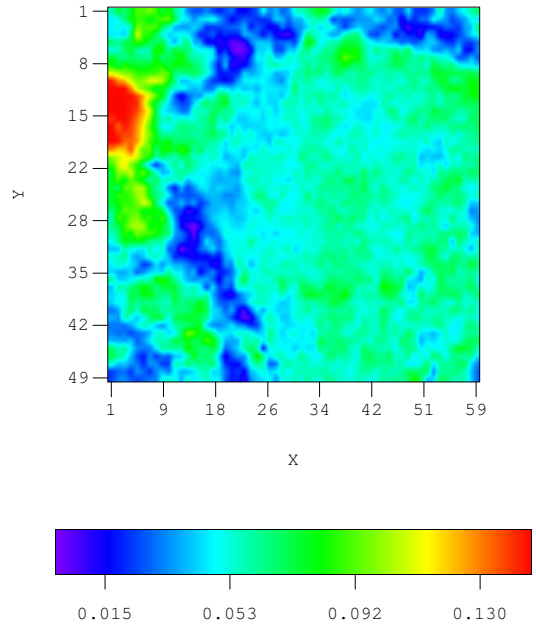
Figure 45: Estimate of log horizontal permeability field in first four layers



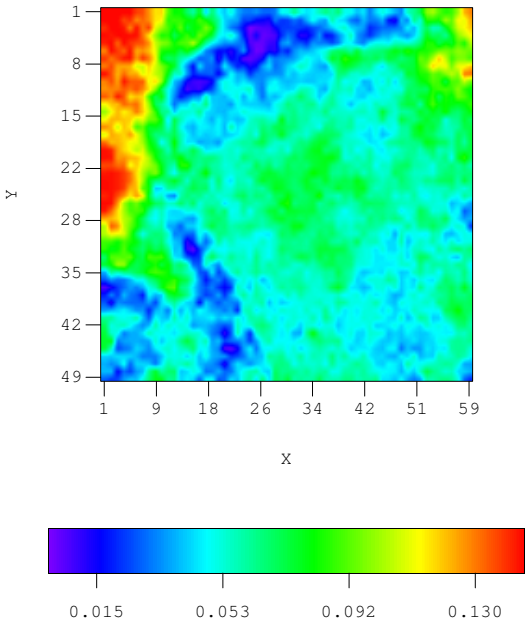
(a) Layer One



(b) Layer Two



(c) Layer Three



(d) Layer Four

Figure 46: Estimate of porosity field in first four layers

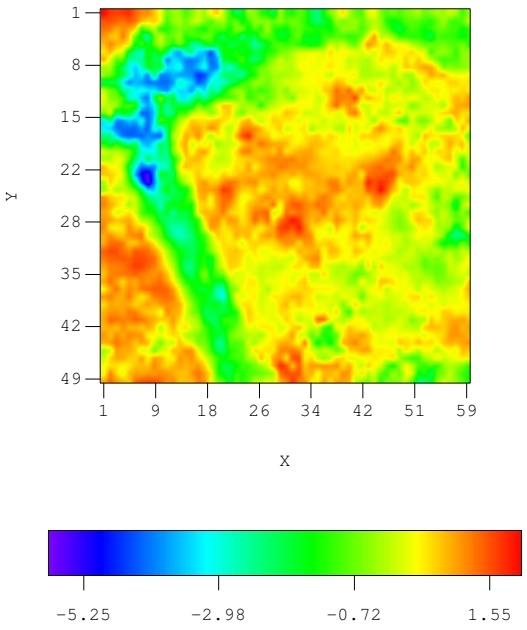


Figure 47: Estimate of log horizontal permeability field in the fifth layer

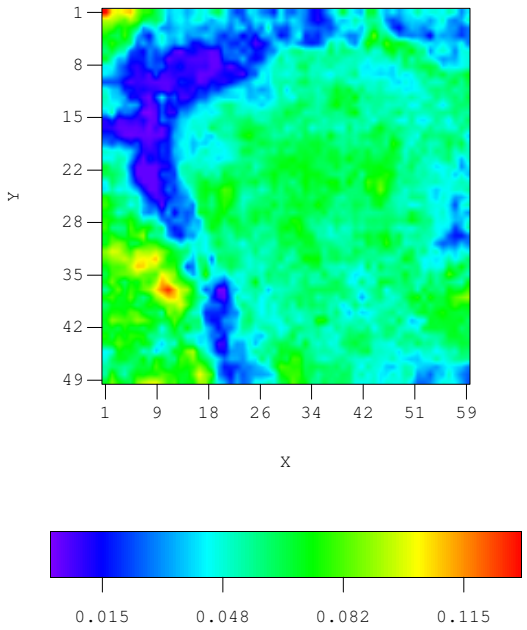


Figure 48: Estimate of porosity field in the fifth layer

HISTORY MATCHING OF FACIES DISTRIBUTIONS

Background

Major improvements in the application of the truncated Gaussian method for lithofacies simulations based on indicators were developed mostly by French scholars. By analyzing the limitations and the potential of the truncated Gaussian method, Galli et al. (1994) found a way to apply this method to a 3-D problem with vertical non-stationarity in the proportions of lithofacies. They showed that this method preserved the consistency of the indicator variograms and cross variograms. The major achievement of this paper is the introduction of the truncated plurigaussian method, which allowed more complex neighbor relations than the standard truncated Gaussian model. In the same period, Le Loc'h et al. (1994) showed the flexibility of the truncated plurigaussian method by truncating two Gaussian functions. They pointed out that even if the two underlying Gaussian functions are uncorrelated, the resulting facies sets obtained by truncating are not independent. The correlation depends on the construction of thresholds of lithotypes. Using uncorrelated Gaussian functions they found that complex theoretical indicator variograms can be produced in combining various anisotropies by choosing different Gaussian functions. They suggested that the choice of a truncation method to the Gaussian functions should be as simple as possible to have easier control over the problem.

Later, Le Loc'h and Galli (1997) presented an insight to implementing the algorithm both for practical structural analysis and conditional simulations. In demonstrating the influence of the thresholds chosen for truncation, the partition of facies was accomplished using rectangles. But even with this relatively simple thresholding method, it is not at all straightforward to choose appropriate thresholds. The difficulty in estimating model parameters that will result in the desired facies distributions has restricted the practical application of this method. An example of a truncated plurigaussian simulation conditional

to facies data at well locations was presented with a very slow convergence. This problem was attributed to the instability of the Gaussian covariance matrix. Lantuéjoul (2002) discusses the problem of conditioning truncated plurigaussian models to facies observations extensively. Assuming known threshold parameters, the truncated plurigaussian simulation scheme was able to simulate the Gaussian random fields to match given lithofacies observations. As his simulation problem was small, the Markov chain Monte Carlo sampling method was applied to evolve Gaussian random fields. While, once again, the great potential of the truncated plurigaussian method in simulating lithofacies distribution was revealed, two major problems were left unsolved and seem to be limiting the application of this method beyond France. First is the difficulty in estimation of geostatistical parameters, i.e. the geostatiscal quantities such as the range, the variance, the covariance type (Gaussian, Exponential, Spherical, etc.) and the thresholds for discrimination of facies. Second, the application of the truncated plurigausian method in practical conditional simulation problems requires more efficient methods of sampling to deal with reservoir history matching problems.

Conditional simulation of reservoir facies distributions is of great interest of reservoir engineers. Bi et al. (2000) and Zhang et al. (2002) approached the problem of simulating a channel sand by simulating the location of the centerline, the width and thickness of the channel all along the channel length. In both cases, the Levenberg-Merquardt or Gauss-Newton methods were used for the history matching and the chain rule was used to compute the derivative of the production data mismatch to the values of channel width for example. They were able to do this because many of the intermediate matrices in the computation of the sensitivities were sparse, and because the number of sensitivity coefficients to be computed was relatively small.

In the article by Rahon et al. (1996), they considered two problems in simulating locations of lithofacies conditional to well pressure data. In the first, they attempt to estimate the permeability of each facies whose locations have been fixed. In the second problem, they altered the size of facies whose permeabilities have been fixed. The gradient calculation relating to lithofacies has been successfully implemented in an implicit single-phase fluid flow model. Rahon et al. (1997) applied similar idea in the problem of simulating channel sand locations. This paper parameterizes a channel by triangularization of surface with the nodes of the triangles representing the parameters. The centerline of the channel is assumed known and fixed and the permeability and porosity in both the channel and non-channel facies are assumed to be known. Sensitivities of the well pressure observations with respect to the parameters of the nodes were computed to adjust the size of the channel. Although the idea of using gradient method to adjust parameters deciding the size of lithofacies was valuable, their

work was limited to the kind of problems with known locations of lithofacies. Landa et al. (2000) integrated well test, logging, and geological data to obtain a reservoir description using gradient method. They calculate the sensitivity matrix for permeability by solving the system n times (where n is the number of parameters or gridblocks) and assume that permeability and porosity are perfectly correlated. Their method is computationally unaffordable for problems with large number of model variables. The method of truncated plurigaussian simulation is very flexible in simulating distribution of lithofacies, for instance, the location, width and sinuosity of multi-channels. However, to obtain a satisfactory resolution in the lithofacies map, the Gaussian random fields used need to be large. Thus the way of computing gradients has to be more efficient.

The Geostatistical Model

We consider a truncated bi-Gaussian field for which two independent Gaussian random fields, y_1 and y_2 are used to generate a facies map.

Generating Thresholds

The choice of the truncation method for the Gaussian variables is important in applying truncated Gaussian simulation in automatic history matching to generate reservoir models satisfying geological requirements. Our intention is to use three or more intersecting lines as thresholds. In this report, I will focus on introducing truncated Gaussian simulation using three threshold lines. Three randomly generated lines intersecting each other without all passing through the same point divide the plane into 7 regions. A facies type can be attributed to each region, so up to 7 different kinds of facies can be included in the same plane with appropriate relative percentage. This number of facies is generally enough for geology maps in petroleum reservoir study, but if not, another line could be added. The three lines are thresholds for different rock properties. Given an angle θ and a distance r, a threshold line could be described by the following equation:

$$y = \tan(\theta - \frac{\pi}{2})(x - \frac{r}{\cos\theta}), \tag{72}$$

i.e., the threshold line is perpendicular to the line passing through the origin with the slope θ and intersects the line at a distance r.

Fig. 49 is an example illustrating the truncation scheme of intersecting threshold lines. The Gaussian random field y_1 has Gaussian type covariance and y_2 has exponential type

covariance. The coordinates of the threshold map (Fig. 49(c)) are y_1 and y_2 respectively. Three kinds of lithotypes, A, B, and C are assigned to the seven regions in the threshold map. Facies type at any gridblock in the field is decided by taking the y_1 and y_2 value of that gridblock to the threshold map. For instance, the gridblock $(20, 40)$ has low values for both its y_1 and y_2. (They both are in areas with dark shade.) So it corresponds to the area in threshold map assigned facies A. We can tell from the facies map (Fig. 49(d)) that the gridblock $(20, 40)$ was assigned facies A. Calculation of the Gaussian fields y_1 and y_2 will be discussed in the later section.

Continuous Variables

We begin by considering the continuous variables. Facies are defined by the truncation of two continuous random fields. At the basic level, the variables on the grid are independent normal deviates with mean 0 and variance 1. These independent deviates must be converted to correlated random fields for truncation. The parameters of the two covariance functions (such as the ranges of the covariances for the two fields, the principal directions) are variables in this problem.

If we partition the truncation map based on truncation lines, then the locations of the lines are also continuous variables. We let the number of lines be denoted by N_l.

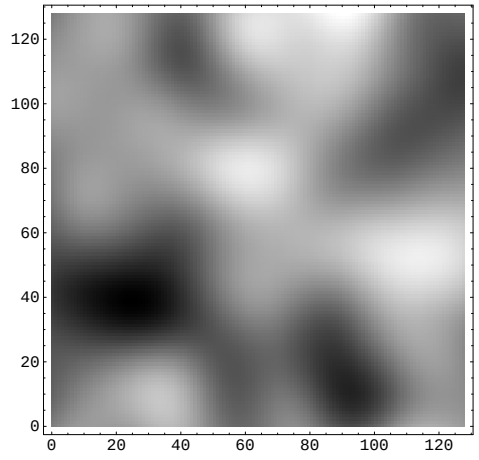
$\{Z_{11}, Z_{12}, \ldots, Z_{1N_g}\}$	Independent normal deviates with mean 0 and variance 1.
$\{Z_{21}, Z_{22}, \ldots, Z_{2N_g}\}$	Independent normal deviates with mean 0 and variance 1.
$\{a_{11}, a_{12}, \theta_{c1}\}$	Ranges of covariance and principal direction of anisotropy.
$\{a_{21}, a_{22}, \theta_{c2}\}$	Ranges of covariance and principal direction of anisotropy.
$\{(r_1, \theta_1), \ldots, (r_{N_l}, \theta_{N_l})\}$	Locations of truncation lines.

The variables Z_1 and Z_2 are assumed to be multivariate normal, so

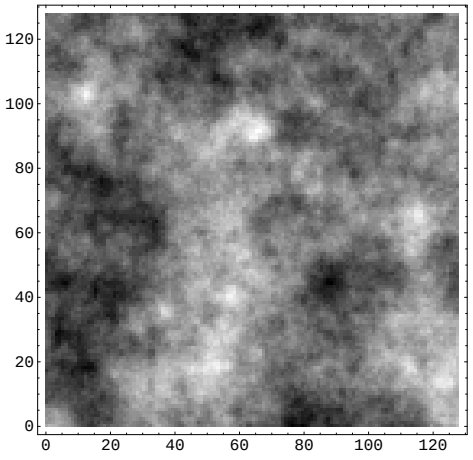
$$P(Z_1) \propto \exp\big(-\frac{1}{2} Z_1^T Z_1\big) \tag{73}$$

$$P(Z_2) \propto \exp\big(-\frac{1}{2} Z_2^T Z_2\big). \tag{74}$$

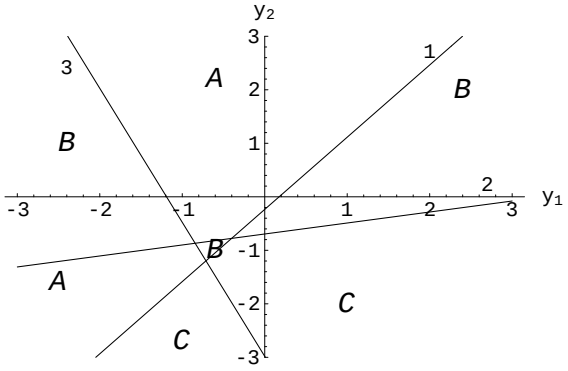
The prior pdfs for the ranges of the variograms $(a_{11}, a_{12}, a_{21}, a_{22})$ can probably also be assumed to be normal distributions with 0 means and fairly large variance. We may want to truncate the distributions so that only positive ranges are allowed. This same assumption seems reasonable for the distance of the truncation lines from the origin, except that here we use a variance of 1 so that the lines are intersected close to the origin.



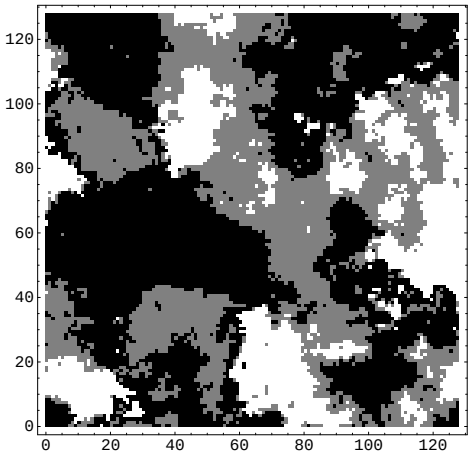
(a) The Gaussian type random field (y1).



(b) The Exponential type random field (y2).



(c) Threshold map.



(d) Calculated facies field map.

Figure 49: Simulation of lithofacies distribution in the field by truncation of random Gaussian fields y_1 and y_2 using intersecting line thresholds.

It seems quite reasonable to assume that the prior distributions for the orientation of the partitioning lines $\theta_1, \ldots, \theta_{N_l}$ are uniform on the interval $(0, \pi)$. In this case, the probability density is a constant and can be ignored (or, more accurately, absorbed into the overall constant). The same is true of θ_{c1} and θ_{c2}.

We can then define m_{cG} to be the vector of continuous variables whose prior distribution

is Gaussian and m_{cU} to be the vector of continuous variables whose prior distribution is uniform or

$$m_{cG} = \begin{bmatrix} Z_1 \\ Z_2 \\ a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ r_1 \\ \vdots \\ r_{N_l} \end{bmatrix} \quad \text{and} \quad m_{cU} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{N_l} \end{bmatrix}. \tag{75}$$

Discrete Variables

There will also be a few parameters or variables that are uncertain but not continuous (and hence not differentiable). The number of partitioning lines is clearly discrete. The covariance model (that is, Gaussian, exponential, spherical, Whittles, etc.) to be used for each of the random fields is not continuous or even numerical. Finally, each region of the truncation map must be assigned a facies. Like the covariance models, the facies do not take continuous values so this is also not continuous or numerical.

N_l	The number of partitioning lines.
$\{m_1, m_2\}$	Covariance models for fields Y_1 and Y_2.
$\{F_1, F_2, \ldots, F_{N_r}\}$	Facies assignment for each partitioning region.

We begin by assuming that $N_l = 3$ and that the covariance models are known. As for the probability for assigning a particular facies type to one of the partitioning regions: it seems reasonable to assign equal probability prior to the incorporation of any information. As soon as any information on relative abundance of facies is available, the probabilities will not be equal.

Prior Probability Density

The prior probability for facies distribution F on a grid is denoted as $P(F)$. We can write

$$\begin{aligned} P(F) &= P(F|m_{cG}, m_{cU}, m_d)P(m_{cG}, m_{cU}, m_d) \\ &\propto P(m_{cG}, m_{cU}, m_d) \\ &= P(m_{cG})P(m_{cU})P(m_d) \\ &\propto P(m_{cG})P(m_d) \end{aligned} \tag{76}$$

The first term on the right, the conditional probability for F given values of all the model variables, can be ignored as the relationship is deterministic once the variables are given. We have also assumed independence of the variables in the prior distributions, which explains the third line of Eq. 76. The fourth line is a result of the uniform distribution for some of the variables.

The prior joint probability density for the continuous variables can be written in a compact form as

$$P(m_{cG}) \propto \exp\big(-\frac{1}{2} m_{cG}^T C_M^{-1} m_{cG}\big) \tag{77}$$

where C_M is the diagonal matrix of variances. For Z_1 and Z_2 the variances are all equal to 1.

The Posterior Probability Density

Our goal is to generate samples from the posterior distribution, i.e., the distribution of F conditioned to observations, d_{obs}. Bayes' theorem tells us that

$$P(m_{cG}, m_{cU}, m_d|d_{\text{obs}}) \propto P(d_{\text{obs}}|m_{cG}, m_{cU}, m_d)P(m_{cG}, m_{cU}, m_d). \tag{78}$$

The first term on the right, the likelihood of the model, can be approximated by the following Gaussian expression,

$$P(d_{\text{obs}}|m_{cG}, m_{cU}, m_d) \approx A \exp\Big[-\frac{1}{2}(F - F_{\text{obs}})^T C_D^{-1}(F - F_{\text{obs}})\Big] \tag{79}$$

where the data or measurement error covariance matrix C_D simply reflects the possibility of error in the identification or modeling of the observed facies. Of course, the vector $F - F_{\text{obs}}$ must be defined in a reasonable way. Facies have no intrinsic numerical value and even if they were assigned numerical values for computation, it might not be reasonable to assume that the difference between Facies 1 and Facies 3 is larger than the difference between Facies 1 and Facies 2. It seems reasonable, for the purpose of conditional simulation and estimation, to assume that the facies are either the same (in which case $F - F_{\text{obs}} = 0$) or they are different (in which case $F - F_{\text{obs}} = 1$).

Minimization

At this stage, we assume homogeneity within facies, i.e. both the permeability and the porosity are constant for the same kind of facies type. However, the property fields are discontinuous at the facies boundaries. Define the difference between the facies type from optimization and the "true" facies type as:

$$f_i = \begin{cases} 0 & \text{if } F_i = F_{\text{obs,i}} \\ 1 & \text{if } F_i \neq F_{\text{obs,i}} \end{cases}$$

where i indicates the ith gridblock. The general objective function for minimization is

$$O(m) = \frac{1}{2}(g(m) - d_{\text{obs}})^T C_D^{-1}(g(m) - d_{\text{obs}}) + \frac{1}{2}(m - m_{\text{pr}})^T C_M^{-1}(m - m_{\text{pr}}). \tag{80}$$

When the objective is to minimize the difference between a facies realization and a training image, the objective function takes the form:

$$\begin{aligned} O_F(m) &= \frac{1}{2}(F(m) - F_{\text{obs}})^T C_D^{-1}(F(m) - F_{\text{obs}}) + \frac{1}{2}(m - m_{\text{pr}})^T C_M^{-1}(m - m_{\text{pr}}) \\ &= \frac{1}{2}f(m)^T C_{DF}^{-1} f(m) + \frac{1}{2}(m - m_{\text{pr}})^T C_M^{-1}(m - m_{\text{pr}}), \end{aligned} \tag{81}$$

where f is a vector of f_i for $i = 1, \ldots, N_g$, if there were facies observation in each gridblock. Otherwise, when the number of facies observations N_f is less than N_g, the dimensions of f is $N_f \times 1$. C_{DF} is the covariance matrix of the facies observation. Assume the facies in each gridblock are independently observed with an unbiased Gaussian error $N(0, \sigma_F^2)$, the C_{DF} matrix is equivalent to an identity matrix multiplied by σ_F^2. The variance of the error in facies observation approaches to zero, therefore the weighting of the model mismatch is much smaller than the data mismatch, and can be ignored from the objective function. We would like the objective function to be as small as possible since in that case we should have a match between the observations and the model realization. The problem is that this function is not differentiable so we cannot use gradient-based methods to find a minimum. One solution is to redefine the function f so that it is differentiable (only for the purpose of computing the minimum). We do this by introducing an artificial transition region between facies.

Exploration on Optimization of Threshold Lines

The gradient of the facies mismatch to model parameters has shown that the parameters deciding threshold lines have much greater sensitivity terms than those deciding random

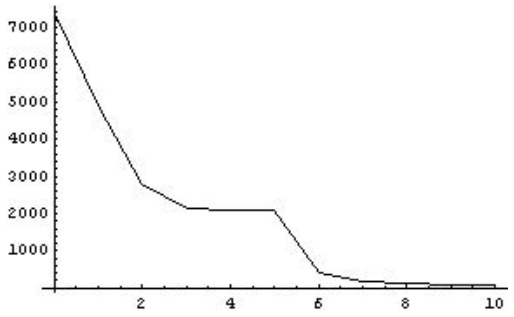
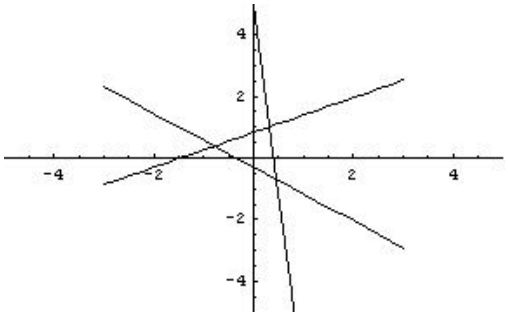


Figure 50: The reduction of the objective function with LM iterations.

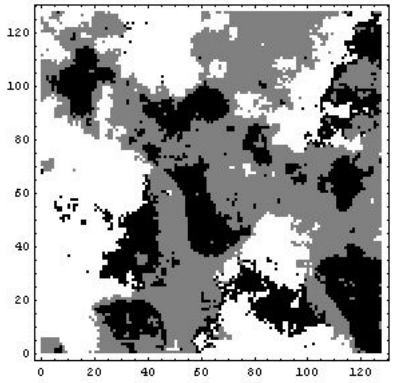
Gaussian fields, which indicates that adjustment on θ's and r's mostly controls the optimization process. So here I take the first step towards a complete optimization of all parameters – fix the two Gaussian random fields as the "true" and leave the threshold parameters as the only set of variables to be optimized. Another purpose of leaving Gaussian random field fixed is to be able to test the validity of the gradient of the objective function about threshold parameters. By fixing the Gaussian random fields, the optimization problem becomes fairly small (with only 6 variable parameters) and it is affordable to calculate inverse of Hessian directly, such that the Levenberg-Marquardt(LM) algorithm could be used instead of LBFGS.

The optimization problem was applied on a fine grid field of 128×128. Prior experience with this research has proven that coarse grid such as 10×10 would make the optimization rather tough. The objective function is the squared difference between the facies realization map and a given training image with facies observations at each grid. A set of threshold lines was generated randomly as a prior model, which was a pure guess and contains no prior knowledge about the truth. By truncating the "true" Gaussian random fields with this prior threshold model, the objective function and a search direction were calculated. After 10 steps of LM iterations, the objective function was reduced from the prior 7383 to 68. The total number of gridblocks is $128 \times 128 = 16384$, and it is reasonable to accept the convergence when less than 1% of the total gridblocks have different facies type to the truth.

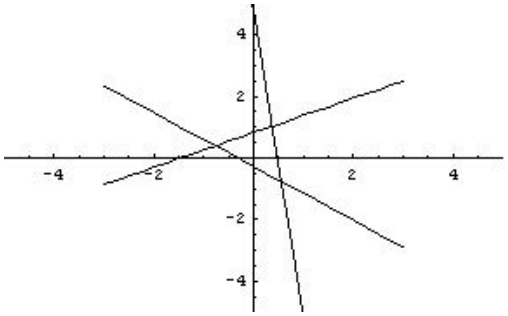
The reduction of the objective function with LM iterations is shown in Fig. 50. In Fig. 51, the threshold model after the tenth step and its corresponding facies distribution were compared with the truth case. The data mismatch of 68 is small enough that the difference between the calculated and the truth could be hardly recognized. Three set of threshold lines, the random prior, the posterior and the true are put together in Fig. 52. A slight mismatch of the posterior threshold and the truth could be observed and the lines in the posterior model have been shifted far from their prior location.



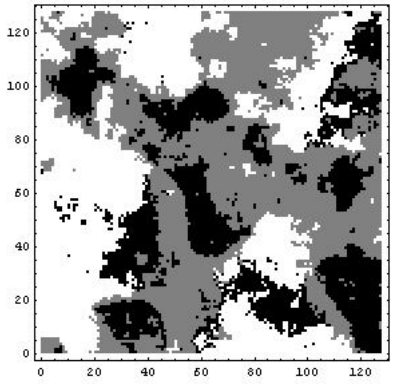
(a) The true threshold map.



(b) The true facies field.

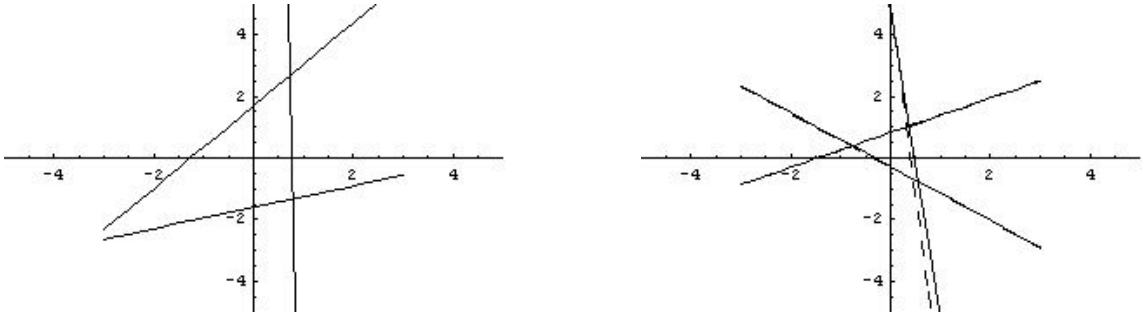


(c) Calculated threshold map.



(d) Calculated facies field.

Figure 51: The comparison of an optimized estimation after 10 LM iterations with the "true" threshold map and facies field.



(a) The prior threshold model.

(b) Comparison between calculated and the truth.

Figure 52: Comparison of the prior, the posterior and the true threshold map. Solid lines are stochastic model estimations and dashed lines are the truth.

A more realistic test problem, in which neither the parameters of the threshold lines or the variables on the grid are known, was used to test the ability to create conditional realizations on large grids, and to explore the pdf of threshold line parameters. A 2-D training image of 128×128 points with known facies was used as data for conditioning. The model variables were defined on an augmented grid whose dimensions were 192×192. Because two of these grids are need for truncation there are approximately 73,000 variables and 16,000 data in the problem.

Three threshold lines were used for truncation and three lithotypes were assigned to the threshold map. Both Y_1 and Y_2 have isotropic Gaussian type covariance with range of 90 ft. Fig. 53 shows the facies distribution from the initial guess of model parameters. The initial facies mismatch is 12525, i.e. 76% of the gridblocks have wrong facies types compared to the training image [Fig. 54 (left)].

This problem is far too large for Levenberg-Marquardt, so the limited memory version of the BFGS algorithm (Nocedal, 1980) was used for the minimization of the objective function because it requires storage of only a set of vectors instead of the whole inverse Hessian matrix. The inverse Hessian is iteratively updated based on computations of the gradient, for which we used an adjoint method with automatic code generation (Corliss et al., 2001).

Fig. 53 (right) shows the gradient of the objective function with respect to Y_1. The gradient of the objective function to Y_1 clearly reflects the regions of high sensitivity. Comparing the right and left sub-figures in Fig. 53, we may conclude that the mismatch function is most sensitive to changes in the values of the random variables that are near facies boundaries. The width of the region of sensitivity depends on the width of the transition region and its

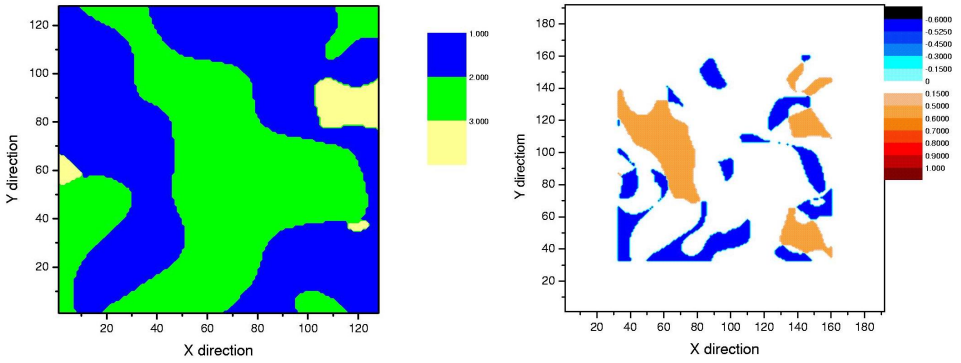


Figure 53: The initial facies map (left) and the gradient of the objective function with respect to the field evaluated at the initial map (right).

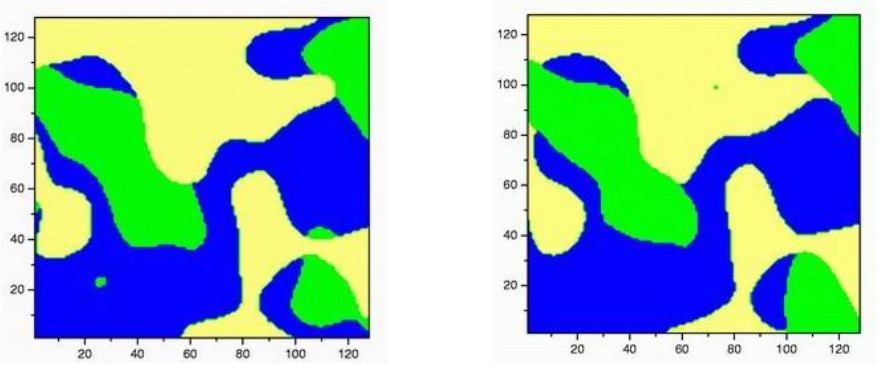


Figure 54: Comparison of the facies map generated from the minimization of the objective function (right) with the training image (left).

choice affects the rate of convergence. Finally the training image is shown together with the final facies map in Fig. 54. The facies in 95.6% of the gridblocks in the final result matched the facies in the training image.

We repeated the minimization procedure 200 times; each time we started with unconditional realizations of the parameters describing locations of the threshold lines and unconditional realizations of the random variables on the grid. The Randomized Maximum Likelihood (RML) method was used for the sampling, as it seems to do a relatively good job of sampling for this type of problem (see Liu and Oliver (2003)). In approximately half of the cases the resulting value of the objective function seemed satisfactory (less than 2000). Fig. 55 summarizes the distribution of realizations of the orientations of the first two threshold lines.

Although the orientation of the threshold lines were sampled randomly from a uniform

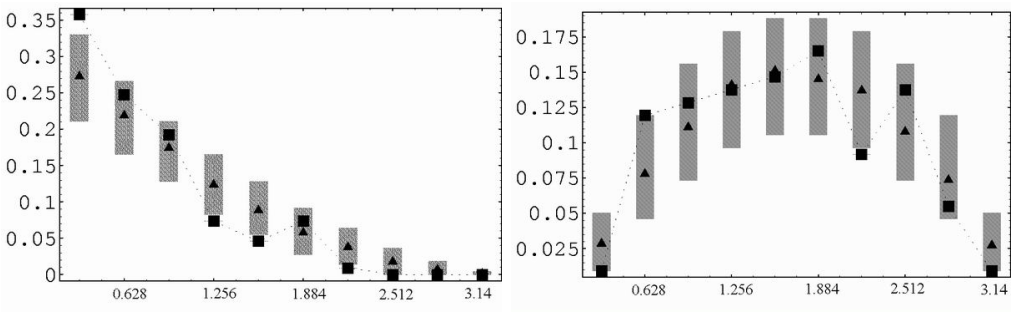


Figure 55: Comparison of the estimated pdf (squares) for θ_1 (left) and θ_2 (right) to the prior pdf (triangles). The width of the gray bars indicates the variability due to limited sample size of 109 ordered sets of 3 orientations from a uniform distribution. 10% are higher and 10% are lower.

distribution, the facies were assigned in a non-random and non-uniform manner. As a result, it was necessary to apply and ordering to the threshold lines with $\theta_1 < \theta_2 < \theta_3$. Once this is done, the distributions of orientations are no longer uniform. It would be unlikely, for example, for the smallest of the three angles (θ_1) to be close to π, and this is seen in Fig. 55. From the two plots, it is not apparent that the estimate of the posterior distribution for θ_1 or θ_2 (indicated by the black squares) is significantly different from the prior distribution(indicated by the black triangles). If there had been rotational symmetry to the threshold plots, we would have expected the orientations to cluster along lines in crossplots of variables. Crossplots of θ_1 and θ_2 failed, however, to show any obvious patterns. In order to honor the proportions of facies correctly, there must be some fairly strong constraints on relationships among the threshold line parameters, but they are not obvious from the conditional realizations. It would be easy, however, to use the realizations generated in this procedure in a Monte Carlo method.

History Matching to Production Data

Practically the hard data of facies observations in a formation layer are only available at well bores with coring operations. The training image from geologists provides expected facies distribution pattern and features. PDF of the geostatistical parameters, such as the ones deciding truncation lines and the ones deciding the covariance of the Gaussian fields. These geostatistical parameters are then used in a Bayesian scheme for simulation of facies conditional to logs or production data.

When the PDF of the geostatistical parameters are known, realizations of facies map can

be generated by minimizing the objective function with respect to the random field Z_1 and Z_2. The objective function in this case is:

$$\begin{aligned} O(m) &= \frac{1}{2}[P(m)-P_{\text{obs}}]^T C_P^{-1}[P(m)-P_{\text{obs}}] + \frac{1}{2}[F(m)-F_{\text{obs}}]^T C_F^{-1}[F(m)-F_{\text{obs}}] \\ &\quad + \frac{1}{2}(m-m_{\text{pr}})^T C_M^{-1}(m-m_{\text{pr}}) \\ &= \frac{1}{2}[d(m)-d_{\text{obs}}]^T C_D^{-1}[d(m)-d_{\text{obs}}] + \frac{1}{2}(m-m_{\text{pr}})^T C_M^{-1}(m-m_{\text{pr}}), \end{aligned} \tag{82}$$

where P_{obs} is the observed production data, such as the bottom whole pressure, the WOR and GOR, and the production rate. $P(m)$ is the simulated production data. F_{obs} is the facies observation at a few cored locations. Both C_P and C_F are diagonal covariance matrix of the data observations, the facies observation is very accurate so its variance is much smaller than that of production data. The mismatch to the hard data dominates the matching before it becomes small enough, so that the facies type at well locations are forced strongly to honor the facies observations. The production data and the hard data can be combined as d_{obs}.

The gradient of the objective function $O(m)$ with respect to the model parameters is:

$$\begin{aligned} g &= \nabla_m O(m) \\ &= G_P(m)^T C_P^{-1}[P(m)-P_{\text{obs}}] + G_F(m)^T C_F^{-1}[F(m)-F_{\text{obs}}] + C_M^{-1}(m-m_{\text{pr}}) \\ &= g_P + g_F + g_m \\ &= \nabla_m O_d(m) + \nabla_m O_m(m). \end{aligned} \tag{83}$$

When the geostatistical model has been decided, generating a facies realization that matches the production data and logging data is to optimize the random field Z_1 and Z_2. The gradient of the data mismatch with respect to Z_1 can be derived by chain rule as:

$$\begin{aligned} \nabla_{Z_1} O_d(m) &= \nabla_{Z_1}[O_P(m) + O_F(m)] \\ &= L_1 \cdot [\nabla_{Y_1} K(m) \cdot \nabla_K O_P(m) + \nabla_{Y_1}\phi(m) \cdot \nabla_\phi O_P(m) + \nabla_{Y_1} O_F(m)]. \end{aligned} \tag{84}$$

The terms $\nabla_K O_P(m)$ and $\nabla_\phi O_P(m)$ are the production data mismatch with respect to the property fields, which are obtained by adjoint method in the history matching program. $\nabla_{Y_1} K(m)$ and $\nabla_{Y_1}\phi(m)$ are diagonal matrices with the diagonal elements:

$$\frac{\partial K_i(m)}{\partial Y_{1,i}} = \frac{dK_i}{dd_l}\frac{\partial d_l}{\partial Y_{1,i}}, \tag{85}$$

$$\frac{\partial \phi_i(m)}{\partial Y_{1,i}} = \frac{d\phi_i}{dd_l}\frac{\partial d_l}{\partial Y_{1,i}}. \tag{86}$$

d_l is the distance of $(Y_{1,i}, Y_{2,i})$ to the closest threshold line, which is computed in Eq. 87.

$$\begin{aligned} d(\vec{v},\vec{u}) &= |\vec{u}| - \frac{\vec{u}\cdot\vec{v}}{|\vec{u}|} \\ &= \frac{\tan(\theta-\frac{\pi}{2})x - y - \tan(\theta-\frac{\pi}{2})\frac{r}{\cos\theta}}{\sqrt{1+\tan^2(\theta-\frac{\pi}{2})}} \\ &= \sin(\theta-\frac{\pi}{2})x - \cos(\theta-\frac{\pi}{2})y + r. \end{aligned} \tag{87}$$

The property fields K and ϕ are regarded homogenous within a facies region, but the values are discontinuous across the facies boundaries. To compute the gradient of the property fields with respect to the Gaussian fields, the property fields have to be differentiable. As the facies observations are only available at a few cored locations, it is not possible to tell whether the facies type corresponds to $(Y_{1,i}, Y_{2,i})$ is correct. Therefore the transition zone of the permeability and porosity values is made in both sides along the closest threshold line to point $(Y_{1,i}, Y_{2,i})$. If the property fields in the transition zone were defined by the square root of the distance to the threshold line as:

$$K(d_l) = \begin{cases} \frac{K_1+K_2}{2} + Sign(d_l)\frac{K_2-K_1}{2}\sqrt{\frac{|d_l|}{\epsilon}} & \text{for } |d_l| \le \epsilon \\ 0 & \text{for } |d_l| > \epsilon, \end{cases} \tag{88}$$

and K_1 and K_2 are permeability values assigned to each facies, then the gradient $\frac{dK_i}{dd_l}$ goes to infinite when the distance to the threshold line approaches to zero. Therefore linear interpolation was chosen for the transition zone of the property fields. Let the width of the transition zone on each side of the threshold line be ϵ, the permeability along the direction perpendicular to a threshold line is:

$$K(d_l) = \begin{cases} K_1 - \frac{1}{2}(K_1-K_2)(1-\frac{|d_l|}{\epsilon}) & \text{for } |d_l| \le \epsilon \\ 0 & \text{for } |d_l| > \epsilon, \end{cases} \tag{89}$$

where K_1 is the assigned facies permeability at the same side of the threshold line with $(Y_{1,i}, Y_{2,i})$, and K_2 is the assigned permeability on the other side. ϵ is the absolute distance to the threshold line. In Fig. 56, the permeability at point B can be computed by Eq. 89, and K_A should be computed by:

$$K_A = K_2 - \frac{1}{2}(K_2-K_1)(1-\frac{|d_{l,a}|}{\epsilon}), \tag{90}$$

as A is on the same side with K_2.

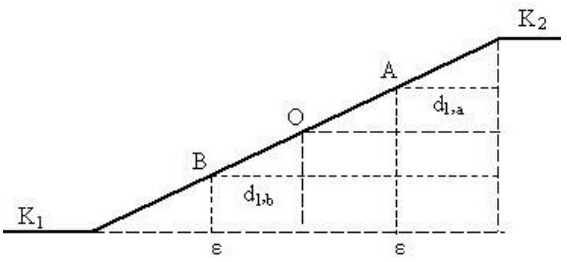


Figure 56: The gradient of the permeability is derived from this linear interpolation model. K_1 and K_2 are permeability values assigned to two adjacent regions in the threshold map. Point O is the cross-section with the threshold line, which is also the middle point of the transition zone in this 1-D plot.

The gradient $\frac{dK_i}{d|d_l|}$ in the linear interpolation case is:

$$\frac{dK_i}{d|d_l|} = \begin{cases} \frac{K_1 - K_2}{2\epsilon} & \text{for } |d_l| \leq \epsilon \\ 0 & \text{for } |d_l| > \epsilon. \end{cases} \tag{91}$$

Substituting Eq. 91 into Eq. 85, the gradient of the permeability at the ith grid with respect to Y_1 at the ith grid is:

$$\begin{aligned} \frac{\partial K_i(m)}{\partial Y_{1,i}} &= \frac{dK_i}{d|d_l|}\frac{d|d_l|}{dd_l}\frac{\partial d_l}{\partial Y_{1,i}} \\ &= Sign(d_l)\frac{K_1 - K_2}{2\epsilon}\sin(\theta_l - \frac{\pi}{2}). \end{aligned} \tag{92}$$

Similarly,

$$\begin{aligned} \frac{\partial \phi_i(m)}{\partial Y_{1,i}} &= \frac{d\phi_i}{d|d_l|}\frac{d|d_l|}{dd_l}\frac{\partial d_l}{\partial Y_{1,i}} \\ &= Sign(d_l)\frac{\phi_1 - \phi_2}{2\epsilon}\sin(\theta_l - \frac{\pi}{2}). \end{aligned} \tag{93}$$

The third term of $\nabla_{Z_1} O_d(m)$ is the gradient of facies mismatch about Z_1. The computation of $\nabla_{Y_1} O_F(m)$ is similar with the case where a training image is given. O_F is the squared difference between the simulated facies and facies observations. When the simulated and the observed facies are the same type, $f_i = F_i(m) - F_{\text{obs,i}} = 0$, otherwise $f_i = 1$. As $\nabla_{Y_1} O_F(m) = G_{F,Y_1}(m)^T C_F^{-1} f$, the key is computing the sensitivity matrix G_{F,Y_1}, which is an $N_F \times N_g$ sparse matrix with maximum one non-zero element in each row. The non-zero elements are the sensitivities of the facies difference f at the facies observation locations with respect to Y_1 at the corresponding grid. Because the logical process of deciding the values of $\frac{df_i}{dY_{1,i}}$ is beyond the description capacity of equations, the following pseudo code is provided.

DO $l = 1,\ N_l$

$$\begin{aligned} d(l) &= \frac{\tan(\theta_l - \frac{\pi}{2})Y_{1,i} - Y_{2,i} - \tan(\theta_l - \frac{\pi}{2})\frac{r_l}{\cos\theta_l}}{\sqrt{1 + \tan^2(\theta_l - \frac{\pi}{2})}} \\ &= \sin(\theta_l - \frac{\pi}{2})Y_{1,i} - \cos(\theta_l - \frac{\pi}{2})Y_{2,i} + r_l \end{aligned}$$

END DO.

$d_l = \min(d)$

IF $F_i = F_{\text{obs,i}}$ THEN

$\frac{df_i}{dY_{1,i}} = 0$

ELSE

IF $|d_l| > \epsilon$ THEN

$\frac{df_i}{dY_{1,i}} = 0$

ELSE

IF $(F_{ab,i} = F_{\text{obs,i}})$ THEN

$$\frac{df_i}{dY_{1,i}} = \frac{d_l}{|d_l|}\frac{1}{2\sqrt{\epsilon|d_l|}}\sin(\theta_l - \frac{\pi}{2})$$

ELSE

$\frac{df_i}{dY_{1,i}} = 0$

END IF

END IF

END IF

$F_{ab,i}$ is the facies type on the other side of the closest threshold line. When the facies at grid i decided by $(Y_{1,i}, Y_{2,i})$ does not match the facies observation at that location, but the other side of the threshold line has the correct facies type, i.e. $F_{ab,i} = F_{\text{obs,i}}$, there exists a transition zone on the side of the threshold line that is closer to $(Y_{1,i}, Y_{2,i})$.

The non-zero terms in $\nabla_{Y_1} O_F(m) = G_{F,Y_1}(m)^T C_F^{-1} f$ is:

$$\frac{\partial O_F(m)}{\partial Y_{1,i}} = \frac{d_l}{|d_l|}\frac{1}{2\sigma_F^2\epsilon}\sin(\theta_l - \frac{\pi}{2}). \tag{94}$$

Similarly, the gradient of the squared data mismatch with respect to Z_2 is:

$$\begin{aligned} \nabla_{Z_2} O_d(m) &= \nabla_{Z_2}[O_P(m) + O_F(m)] \\ &= L_2 \cdot [\nabla_{Y_2} K(m) \cdot \nabla_K O_P(m) + \nabla_{Y_2}\phi(m) \cdot \nabla_\phi O_P(m) + \nabla_{Y_2} O_F(m)]. \end{aligned} \tag{95}$$

$\nabla_{Y_2} K(m)$ and $\nabla_{Y_2} \phi(m)$ are $N_g \times N_g$ diagonal matrices with diagonal elements:

$$\begin{aligned} \frac{\partial K_i(m)}{\partial Y_{2,i}} &= \frac{dK_i}{d|d_l|}\frac{d|d_l|}{dd_l}\frac{\partial d_l}{\partial Y_{2,i}} \\ &= Sign(d_l)\frac{K_1 - K_2}{2\epsilon}\sin(\theta_l - \frac{\pi}{2}). \end{aligned} \tag{96}$$

$$\begin{aligned} \frac{\partial \phi_i(m)}{\partial Y_{2,i}} &= \frac{d\phi_i}{d|d_l|}\frac{d|d_l|}{dd_l}\frac{\partial d_l}{\partial Y_{2,i}} \\ &= Sign(d_l)\frac{\phi_1 - \phi_2}{2\epsilon}\sin(\theta_l - \frac{\pi}{2}). \end{aligned} \tag{97}$$

$\nabla_{Y_2} O_F(m)$ is a very sparse vector with a maximum of N_F non-zero elements. The non-zero elements are decided very similarly with those in $\nabla_{Y_1} O_F(m)$.

$$\frac{\partial O_F(m)}{\partial Y_{2,i}} = \frac{d_l}{|d_l|}\frac{1}{2\sigma_F^2\epsilon}\sin(\theta_l - \frac{\pi}{2}). \tag{98}$$

Generate Initial Model

The initial model for history matching of production data need to honor the facies observations. Knowing the geological and geostatistical model, the initial model Z_1 and Z_2 can be generated in the following procedure.

1. Generate two multivariate Gaussian deviates Z_1 and Z_2 from $N(\mathbf{0}, I)$.
2. Simulate the facies map from Z_1 and Z_2 and check whether the simulated facies at observation locations match the facies observation.
3. For simulated facies that do not honor the facies observation, generate new random variables in patches of appropriate size (30×30 in this application) and replace the patches to the variables centered at the observation locations that do not match.
4. Run the simulation again and check the matching of the facies observation. If there is mismatch, repeat step 3. If not, output the random fields Z_1, Z_2 as a set of initial model.

Twelve initial models are generated and the initial facies maps are shown in Fig. 57.

As the assignment of facies to the seven regions in the threshold map is ad hoc and non-differentiable at this stage, the true facies arrangement is assumed to be known as a part of the geostatistical model. The history matching process then uses the true threshold map

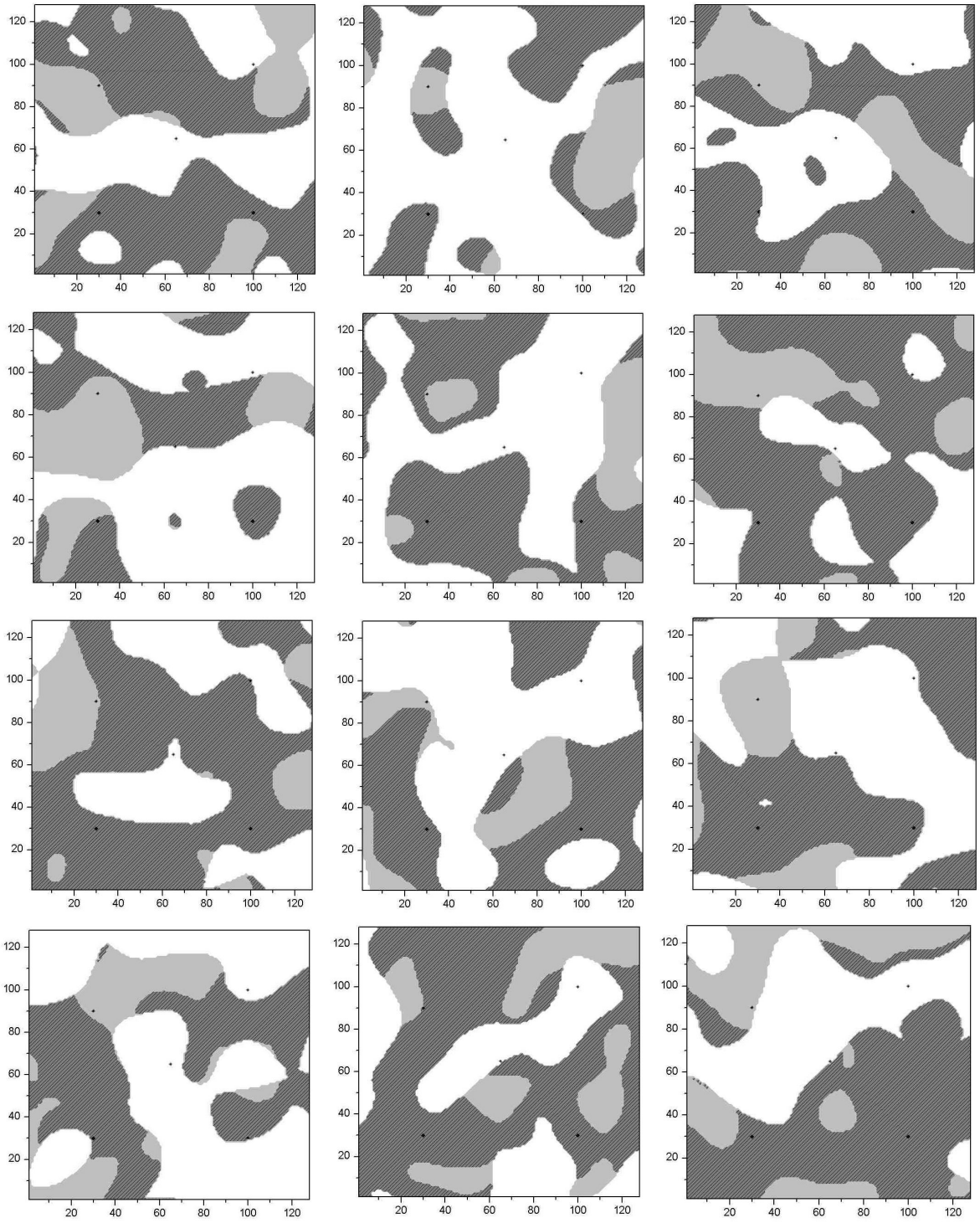


Figure 57: Initial facies maps that honors the facies observations.

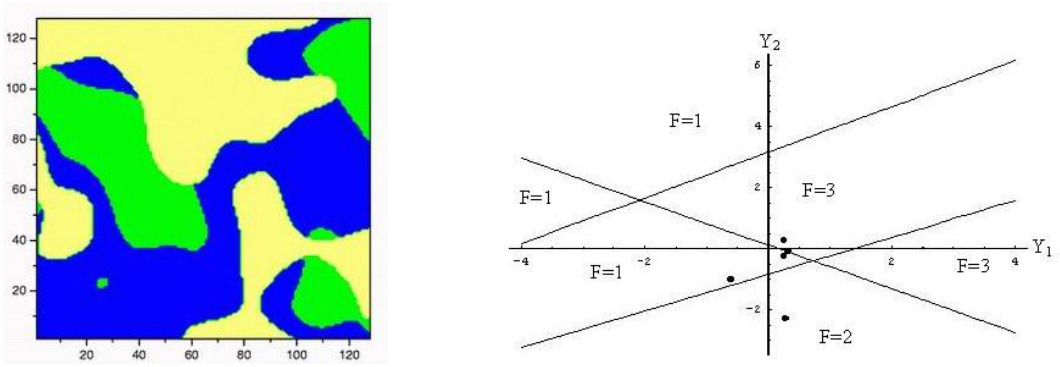


Figure 58: The true facies map and the true threshold map with the Gaussian variables (Y_1, Y_2) at each facies observation of the true facies map.

to do the truncation. One potential problem with fixing the threshold map and the facies assignment is that the grids with Gaussian random variables far from threshold lines may be difficult to be modified to the correct facies. To make this point clear, the true threshold map is shown in Fig. 58 with the pairs of Gaussian variables (Y_1, Y_2) at the observation locations of the true facies map. The regions in the threshold map are assigned three types of facies: $F = 1, 2,$ or 3. Both the facies observations with facies $F = 1$ can be close to facies 2 and 3 in the true facies map. An initial map can be very possibly generated matching the facies observations, but with the pairs of Gaussian variables in a different region of the threshold map. For instance, if either of the two pairs of (Y_1, Y_2) giving facies 1 is in the top region where Y_2 has large positive value, the observation location will be hardly adjacent to facies 2, as it is not likely for that point to move around the corner with the facies 3 region and get down to the lower facies 1 region. When there are production data at that facies observation location, it may never lead to converge because the facies around the well is incorrect.

Therefore the initial model should honor the regions at facies observation locations to make the convergence easier. The procedures for generating the initial models that not only match the facies observations, but match the regions of the facies observations are shown as following:

1. Generate two multivariate Gaussian deviates Z_1 and Z_2 from $N(\mathbf{0}, I)$.

2. Simulate the facies map from Z_1 and Z_2, then check whether the pairs of the Gaussian variables (Y_1, Y_2) at observation locations are in the correct region.

3. For Gaussian pairs at observation locations that are not in the correct region in the threshold map, generate new random variables in patches of appropriate size (30×30

in this application) and replace the patches to the variables centered at the observation locations that do not match.

4. Run the simulation again and check the matching of the Gaussian variable region. If there is mismatch, repeat step 3. If not, output the random fields Z_1, Z_2 as a set of initial model.

Investigation on Convergence

The convergence of the objective function is largely dependent on the transition zone width. Each grid has two Gaussian variables (Y_1, Y_2), and it has non-zero gradients of the objective function with respect to both Gaussian variables only if it is in the transition zone. When the transition zone is wide, grids that are far within facies regions respond to the perturbation to model parameters Z_1 and Z_2. The advantage of wide transition zone might be that the modifications to more grids can be made in each iteration. However, as the transition is just made to compute the gradient and the objective function does not consider the transition zone, gradients from a narrower width should provide a more accurate adjustment direction. Fig. 59 shows the objective function along the search direction in the first optimization iteration. The top row is with the transition width of 0.2, and the bottom row is 1. The two figures in the left column include points with non-zero facies mismatch at facies observations. As all the five facies observations are at well locations, the facies mismatch caused great production data mismatch at well locations. Therefore the objective function jumped by a factor of 20.

The reduction of the objective function before the facies alternation is small relative to its scale when the facies at well locations are different with observation. Therefore the data points look flat. To have a close look at the trend of the objective function with respect to the step size along the search direction, the flat parts are plotted out as shown in the right column of Fig. 59. Comparing the two figures with the transition width of 0.2 and 1 respectively, the reduction of the objective function is greater for the case with wider transition width.

When the proposed model does not satisfy both Wolfe conditions at the same time, a quadratic fit will be made to reduce the objective function by optimizing the step size. The quadratic function is in the form of:

$$q(\alpha) = a\alpha^2 + b\alpha + c, \tag{99}$$

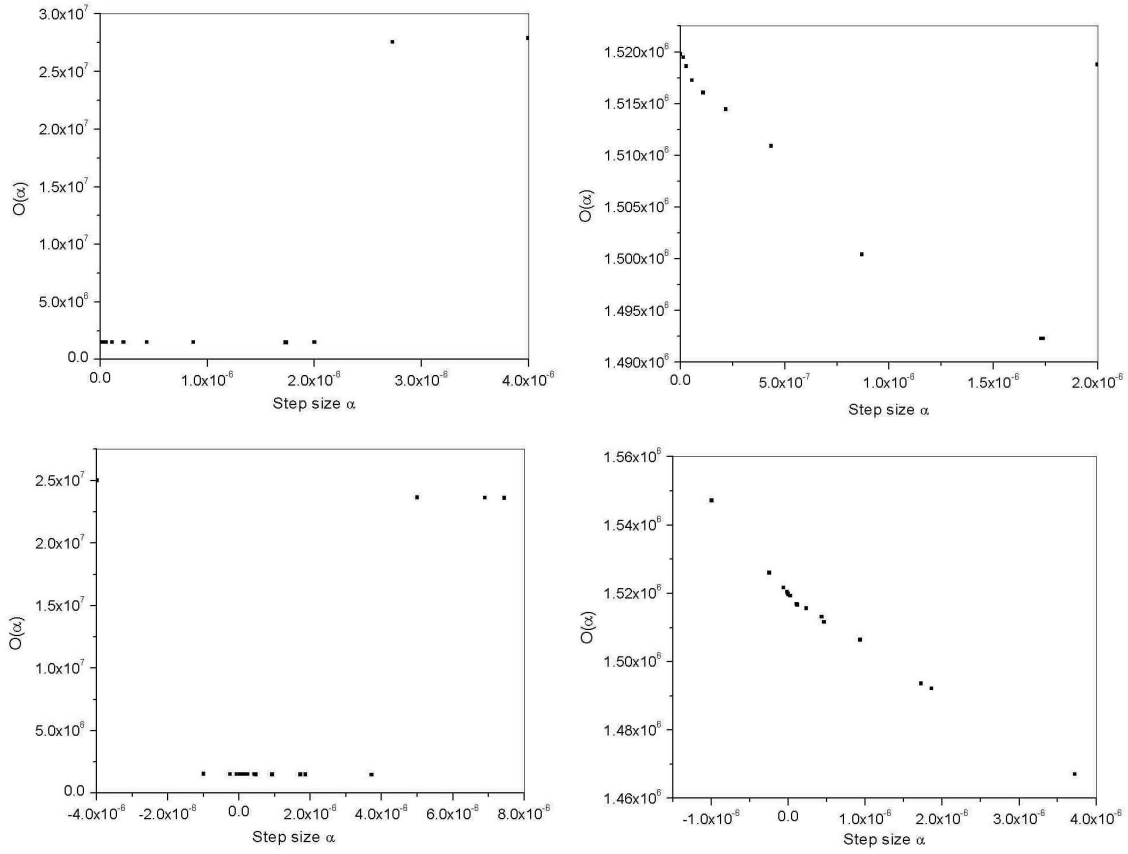


Figure 59: All the four figures are the objective function along the first search direction. The transition zone width in the first row is 0.2, in the second row is 1.0. The figures on the right column are amplifications of the flat region in the figures on the left column.

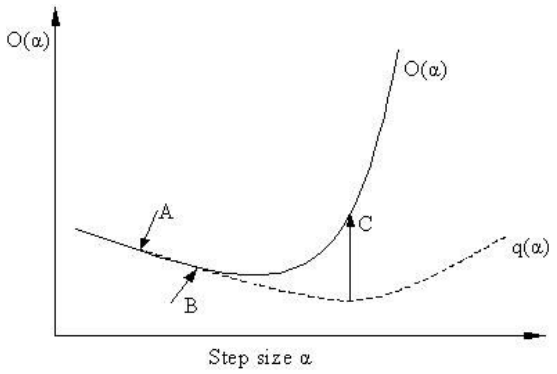


Figure 60: A schematic plot of the quadratic fit to an objective function curve with the typical shape for this minimization problem. The quadratic fit gives a higher objective function value than that from the Newton-Raphson iteration.

where the coefficients a, b, and c are computed as:

$$\begin{aligned} a &= \frac{f(m_k + \hat{\alpha}_k d_k) - f'(m_k)\hat{\alpha}_k - f(m_k)}{\hat{\alpha}_k^2}, \\ b &= f'(m_k), \\ c &= f(m_k). \end{aligned} \tag{100}$$

The step size $\hat{\alpha}_k$ is computed from the Newton-Raphson iteration right before the quadratic fit. d_k is the search direction along which we try to find a minimum of the objective function. m_k is the current model from which the search direction is computed. $f(m_k)$ is the objective function at the current model m_k. $f'(m_k)$ is the gradient of the objective function about model parameters at m_k.

Both curves of the objective function with different transition zone width are nearly linear before the abrupt jump. Fig. 60 is a schematic plot showing the typical shape of the objective function along the search direction for this type of minimization problem. Point A represents the starting model m_k, and B is the temporary model by Newton-Raphson search. Although the objective function has been reduced from point A to point B, the change in the model parameters might not be large enough to satisfy the second Wolfe condition. Therefore a quadratic fit is made through point A and B. The step size corresponding to point C is at the minimum point in the quadratic fit function $q(\alpha)$, but gives a higher objective function than both point A and B. In this case, the new model at point C should be discarded.

The process of optimization is fairly complicated. A flow chart is provided to give a better illustration of the structure of the code.

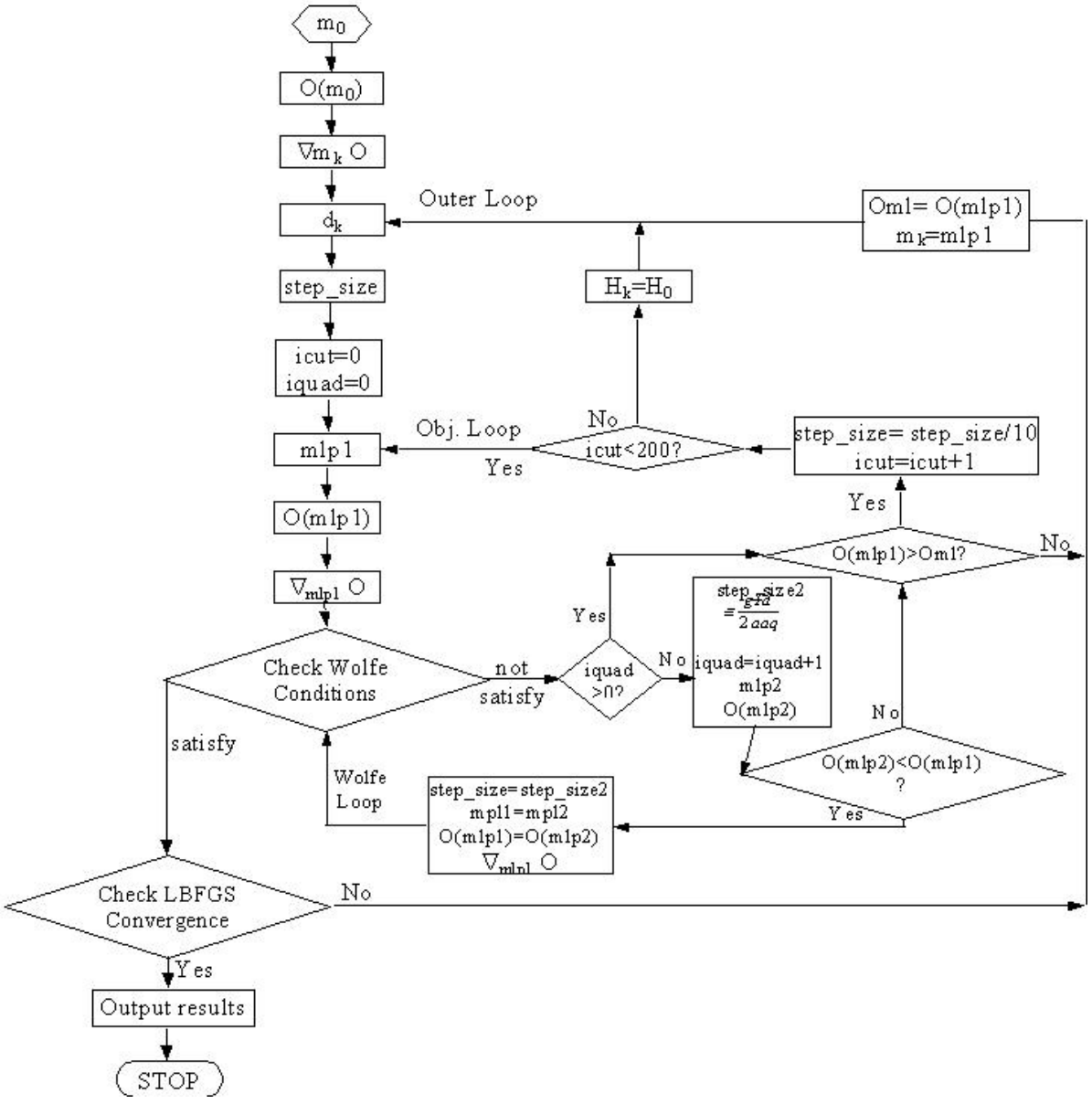


Figure 61: Flow chart for the automatic history matching process.

Constrained Optimization

There are two types of data available in this history matching problem. One type is the production data, the other type is the facies observation data, which is regarded as hard data. The process of optimizing the model to match the production data is pretty much similar with common automatic history matching problems. The hard data in our problem, however, is not the same type as the model parameters. The hard data are the facies observations, and the underlying model parameters Z_1 and Z_2 have to be constrained to the facies observations while matching the production data.

There are two aspects to consider on matching the facies observations. One is in generating initial models, which have been discussed in one of the previous sections. The other is in maintaining the facies type at observation locations in the process of optimization.

In line search for the optimized step size, we take one Newton-Raphson iteration as:

$$\alpha_1 = -\frac{(\nabla O(m_k))^T d_k}{d_k^T H(m_k) d_k}, \tag{101}$$

where the computation of the denominator $d_k^T H(m_k) d_k$ requires the term $G_k d_k$. $G_k d_k$ is computed in this project by finite difference method:

$$\begin{aligned} Gd_k &= \|d_k\| \frac{dO(m_k)}{d\alpha} \\ &\approx \|d_k\| \frac{O(m_k + \epsilon d_k) - O(m_k)}{\epsilon \|d_k\|} \\ &= \frac{O(m_k + \epsilon d_k) - O(m_k)}{\epsilon}, \end{aligned} \tag{102}$$

where ϵ is a small constant chosen based on the infinite norm of d_k:

$$\epsilon = \frac{10^{-3}}{\|d_k\|_\infty}. \tag{103}$$

The model $m_k + \epsilon d_k$ may not match the facies observation, which makes it meaningless and impractical to compute the objective function knowing it will be discarded. Therefore we first check the facies mismatch for the model $m_k + \epsilon d_k$, if it is non-zero, then the ϵ is cut back: $\epsilon = \epsilon/10$, otherwise, the model is put into the simulator to compute the objective function $O(m_k + \epsilon d_k)$.

After the step size is computed from the line search, again we check the facies mismatch for the model $m_k + \alpha_1 d_k$. If it is non-zero, then $\alpha_1 = \alpha_1/10$, otherwise, the new model $m_k + \alpha_1 d_k$ is put into the simulator to compute the objective function $O(m_k + \alpha_1 d_k)$.

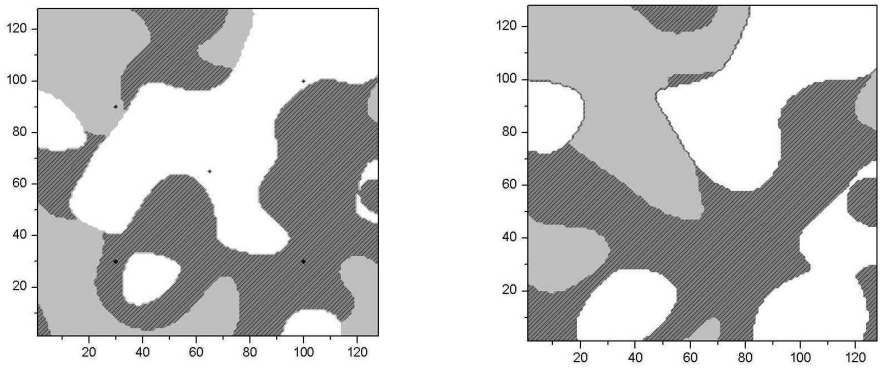


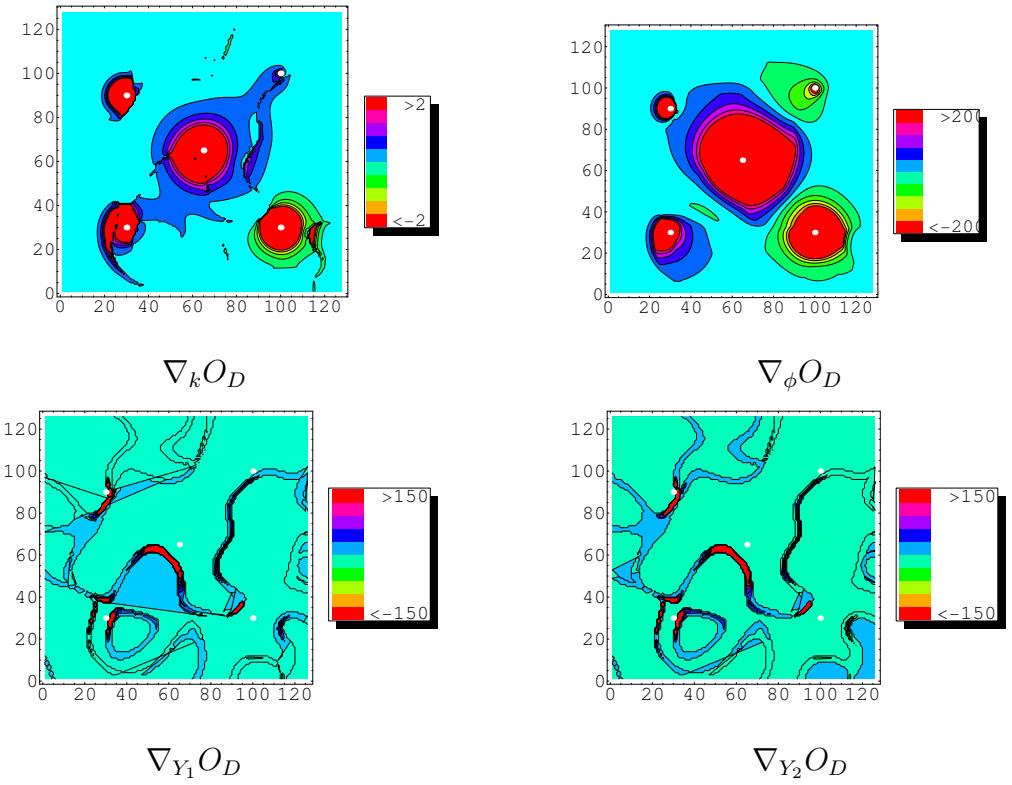
Figure 62: The initial facies map and the final facies map after convergence. The objective function reduced to 1% of the original objective function at the final model.

A Case Study

An initial model has been generated matching the regions of facies observations. The initial facies map is shown on the left of Fig. 62. After 13 LBFGS iterations, the objective function reduced to 1% of the initial value. The facies map for the converged model is shown on the right of Fig. 62. The production data is from the true case shown in Fig. 54.

The intermediate results for computing the gradient of the squared data mismatch O_D with respect to model parameters Z_1 and Z_2 are shown in Fig. 63. The gradient of squared data mismatch with respect to the permeability and the porosity fields are computed with the adjoint method. Then the chain rule is applied to compute the gradient of the squared data mismatch with respect to each of the Gaussian fields. As well4 at the upper right corner is surrounded by facies 3, which has very high permeability, the sensitivity of rock properties to production data is relatively small.

The final matching of production data for each well is shown in Fig. 64. The production data for well 2 has the best matching, mostly because the gradient is large at that low permeability region. The final simulated data from well 4 is further away from the observation data, in comparison with the simulated data from the initial model. The reason might be that the gradient around well 4 is relatively small than those around other wells, thus the facies modification close to well 4 is dominated by the gradients from other wells. If the objective function can be further reduced, the gradients from the data mismatch at well 4 will finally dominate and the facies can be improved around well 4 towards reductions of data mismatch of well 4.



$\nabla_k O_D$ $\nabla_\phi O_D$

$\nabla_{Y_1} O_D$ $\nabla_{Y_2} O_D$

Figure 63: The gradient of the objective function with respect to intermediate parameters. The first row are output from the normal adjoint computation. The second row is the gradient of the objective function with respect to each of the two Gaussian fields Y_1 and Y_2.

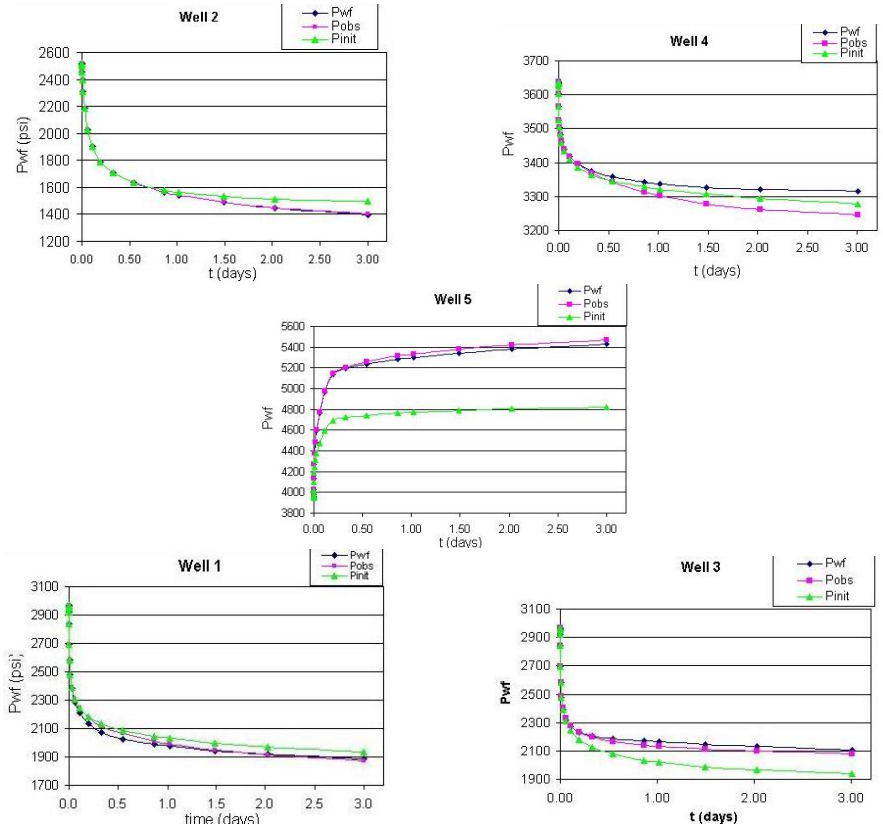


Figure 64: Comparison of the production data from the initial and the final model with the observation data.

RESULTS ON ESTIMATION OF RELATIVE PERMEABILITY CURVES

Generation of Estimates

For the synthetic examples presented here, the reservoir is assumed to be a rectangular parallelepiped and the permeability field is assumed to be isotropic. The forward model used to calculate production data is a fully-implicit finite-difference simulator based on a block centered grid. Additional information on the simulator can be found in Li et al. (2001a) The model parameters to be estimated include the gridblock log-permeabilities and the parameters involved in the power law expressions for the relative permeability curves. All other rock and fluid properties are assumed to be known. Log-permeabilities, instead of absolute permeabilities, are used as parameters because it is assumed that absolute permeability is log-normal.

Throughout, m denotes the vector of model parameters (gridblock log-permeabilities and parameters defining relative permeability curves). In the Bayesian approach used here, m is considered to be random vector. We assume that m has a prior multivariate Gaussian distribution with covariance matrix C_M and prior mean m_{prior}. It is assumed that the model parameters defining the relative permeability functions are not correlated with each other or with the gridblock log-permeabilities. In practice, the prior means and covariance function for log-permeability will be obtained from a geostatistical model constructed from static data, e.g. well log, core and seismic data but not production data. The prior covariance matrix and prior means for relative permeability parameters could be constructed from laboratory measurements of relative permeability curves, from lab data for analogue reservoirs, or relative permeability correlations.

All observed production data that will be history matched are stored in the vector d_{obs}.

The associated calculated or predicted production data corresponding to a given model m is represented by the relation

$$d = g(m). \tag{104}$$

In this work, Eq. 104 represents the operation of running the simulator with given values of the model parameters to compute the production data d. In the history matching process, we wish to determine m such that d is in agreement with d_{obs}. In general, measurement and modeling errors preclude the possibility of finding a model m so that d and d_{obs} are in exact agreement.

Production data measurement errors are assumed to be independent, random variables with mean zero and prescribed variance so that the data covariance matrix, denoted by C_D, is diagonal. The variances of all pressure measurement errors are assumed to be identical, i.e., the pressure measurement error is independent of time and independent of the magnitude of pressure. Similarly, the variance of the gas-oil ratio measurement errors is assumed to be constant.

The "measured" WOR is not actually measured directly, but is constructed from measured oil and water rates. Using basic statistical principles for constructing the variance of the ratio of two random variables, Wu et al.Wu et al. (1999) showed that the variance of the WOR measurement error can be approximated by

$$\mathrm{Var}(e_{WOR}) = WOR_{obs}^2\Big(\epsilon_o^2 + \epsilon_w^2\Big), \tag{105}$$

where ϵ_m^2 denotes the variance of the measurement error of q_m (flow rate of phase m) divided by the observed value of q_m. (For example, $\epsilon_o = 0.01$ corresponds to setting the standard deviation of the measured oil rate equal to one per cent of the measured oil rate.) As the observed water-oil ratio (WOR_{obs}) approaches zero, however, Eq. 105 indicates that the variance of the measurement error also approaches zero, and this is unrealistic because very small values of the water rate, and hence very small values of the WOR, will not be measured accurately. Thus, it makes practical sense to specify a minimum possible value of $\mathrm{Var}(e_{WOR})$; we use the following formula:

$$\mathrm{Var}(e_{WOR}) = WOR_{obs}^2\epsilon_o^2 + \frac{1}{q_{o,obs}^2}\max\Big[\epsilon_w^2 q_{w,obs}^2,\, \sigma_{w,\min}^2\Big], \tag{106}$$

where for the examples considered in this work, we set $\sigma_{w,\min} = 0.32$ std m^3/d (2 STB/D) . Note that if $\epsilon_w^2 q_{w,\mathrm{obs}} \geq 2$, then Eq. 106 reduces to Eq. 105.

Under the assumptions delineated above, the maximum a posteriori estimate of model parameters can be obtained by minimizing the following objective function Tarantola (1987);

Wu et al. (1999); Li et al. (2001a)

$$O(m) = \frac{1}{2}\Big[\big(m - m_{\text{prior}}\big)^T C_M^{-1}\big(m - m_{\text{prior}}\big) + \big(g(m) - d_{\text{obs}}\big)^T C_D^{-1}\big(g(m) - d_{\text{obs}}\big)\Big]. \quad (107)$$

In the examples, considered here, the modified Levenberg-Marquardt algorithm originally presented in Bi (1999) is applied to minimize $O(m)$ with the Li et al. (2001b) implementation of the adjoint method used to calculate sensitivity coefficients.

Relative Permeability Models

Throughout, S_{wc} denotes irreducible water saturation, S_{gc} denotes critical gas saturation, S_{org} denotes the residual oil saturation for a two-phase, gas-oil system and S_{orw} denotes residual oil saturation for a two-phase, water-oil system. As is common, we define oil relative permeability under three-phase flow conditions as a combination of the two sets of two-phase relative permeability functions. In the simulator, and hence in the examples presented, the historically popular Stone'sStone (1973) Model II is applied. The form of this model applied here is the modification that is described in Aziz and SettariAziz and Settari (1979). Our results, however, indicate this model has a peculiar characteristic that is troublesome.

As the examples presented are synthetic, simple analytical formulas are used for the relative permeability functions. Letting

$$S_{w,\max} = 1 - S_{orw}, \quad (108)$$

the water relative permeability function is given by

$$k_{rw} = \begin{cases} 0 & \text{if } S_w \le S_{wc}, \\ k_{rwcw}\left(\frac{S_w - S_{wc}}{1 - S_{orw} - S_{wc}}\right)^{n_{rw}} & \text{if } S_{wc} \le S_w \le S_{w,\max} \\ k_{rwcw} & \text{if } S_w \ge S_{w,\max}. \end{cases} \quad (109)$$

The endpoint relative permeability, k_{rwcw}, represents the maximum value of water relative permeability that can be obtained. The corresponding oil relative permeability function for a two-phase oil-water system is given by

$$k_{row} = \begin{cases} k_{rocw} & \text{if } S_w \le S_{wc}, \\ k_{rocw}\left(\frac{1 - S_{orw} - S_w}{1 - S_{orw} - S_{wc}}\right)^{n_{row}} & \text{if } S_{wc} \le S_w \le S_{w,\max} \\ 0 & \text{if } S_w \ge S_{w,\max}. \end{cases} \quad (110)$$

We let k_{rog} denote the oil relative permeability for a two-phase flow oil-gas system. In the Stone Model II, it is assumed that the two-phase oil-gas relative permeabilities depend only on gas saturation and are measured in the presence of irreducible water saturation. Thus the endpoint values of oil relative permeability for a two-phase oil-water system and a two-phase oil-gas system are identical, i.e.,

$$k_{rog}(S_g = 0) = k_{row}(S_w = S_{wc}) = k_{rocw}. \tag{111}$$

The gas relative permeability function used in this work is given by

$$k_{rg} = \begin{cases} 0 & \text{if } S_g \leq S_{gc}, \\ k_{rgcw}\left(\frac{S_g - S_{gc}}{1 - S_{org} - S_{wc} - S_{gc}}\right)^{n_{rg}} & \text{if } S_{gc} \leq S_g \leq S_{g,\max} \\ k_{rgcw} & \text{if } S_g \geq S_{g,\max}, \end{cases} \tag{112}$$

where

$$S_{g,max} = 1 - S_{org} - S_{wc}. \tag{113}$$

Note k_{rgcw} denotes the maximum value of gas relative permeability and occurs when $S_w = S_{wc}$ and $S_o = S_{org}$. The oil relative permeability function for a two-phase oil-gas system is given by

$$k_{rog} = \begin{cases} k_{rocw} & \text{if } S_g \leq S_{gc}, \\ k_{rocw}\left(\frac{1 - S_{org} - S_{wc} - S_g}{1 - S_{org} - S_{wc} - S_{gc}}\right)^{n_{rog}} & \text{if } S_{gc} \leq S_g \leq S_{g,\max} \\ 0 & \text{if } S_g \geq S_{g,\max}. \end{cases} \tag{114}$$

Equations 112 and 114 indicate that oil and gas relative permeabilities are a function of only S_g for a two-phase oil-gas system, whereas, the two-phase oil and water relative permeabilities depend only on water saturation.

The water and gas relative permeabilities of Eqs. 109 and 112 apply regardless of the number of phases present. If all three phases are mobile, the oil relative permeability function is given by

$$k_{ro} = k_{rocw}\Big(\Big[\frac{k_{row}}{k_{rocw}} + k_{rw}\Big]\Big[\frac{k_{rog}}{k_{rocw}} + k_{rg}\Big] - \Big[k_{rw} + k_{rg}\Big]\Big). \tag{115}$$

Equation 115 represents the modification of Stone's Model II as presented in Aziz and SettariAziz and Settari (1979). This equation gives oil relative permeability as a function of S_w and S_g and applies for any values of saturations, subject to the proviso that k_{ro} is set equal to zero whenever the formula predicts a negative value. If $S_g = 0$, it is easy to show that Eq. 115 reduces to the oil relative permeability function of the two-phase oil-water system and if $S_w = S_{wc}$, then Eq. 115 reduces to the oil relative permeability function for the two-phase oil-gas system.

We also assume residual oil saturations, critical gas saturation and irreducible water saturation are known. Thus, in the general three-phase flow case, estimation of relative permeability curves requires estimation of the components of the vector

$$m_{\rm rel} = \left[k_{rwcw}, k_{rocw}, n_{rw}, n_{row}, k_{rgcw}, n_{rg}, n_{rog}\right]^T. \tag{116}$$

In general the components of the model m include gridblock log-pemeabilities and the components of $m_{\rm rel}$. If absolute permeability is assumed to be known, then $m = m_{\rm rel}$.

Comments on Estimation of Relative Permeabilities

The standard black-oil equations for three-phase flow involve only the product of absolute and relative permeability, i.e., kk_{rm} for $m = o, w, g$. Suppose we are given the true permeability field and true relative permeabilities as one plausible set of model parameters which match production data. If we generate a second set of model parameters by multiplying the absolute permeability field by any positive constant a and dividing the true relative permeabilities by a, then effective permeabilities as functions of saturations for the resulting set of model parameters will be identical to the true model. Moreover, the two models will predict identical production data when input into a reservoir simulator. Thus, it is clear that there is an infinite set of model parameters that will result in the same predictions of reservoir performance. In terms of the inverse problem of interest here, this means that we can not expect to generate accurate estimates of absolute and relative permeability by history matching observed production data. Any such estimate will have an infinite uncertainty attached to them unless the uncertainty is constrained by a prior model or by imposing a constraints on the permissible values of absolute permeability and relative permeabilities.

For the two sets of two-phase relative permeability curves, the effective permeabilities as functions of saturations do not change as long as kk_{rocw}, kk_{rgcw} and kk_{rwcw} are held constant. Thus, under two-phase flow conditions, production data will not allow one to construct a reliable estimate of k and the endpoint relative permeabilities. Quite interestingly, the modified Stone's Model II three-phase oil relative permeability function has the strange characteristic that it is not possible to keep all three effective permeabilities constant if k is varied. To see this, we multiply Eq. 115 by absolute permeability and rearrange the resulting equation to obtain

$$kk_{ro} = kk_{rocw}\left(\left[\frac{kk_{row}}{kk_{rocw}} + \frac{kk_{rw}}{k}\right]\left[\frac{kk_{rog}}{kk_{rocw}} + \frac{kk_{rg}}{k}\right] - \left[\frac{kk_{rw}}{k} + \frac{kk_{rg}}{k}\right]\right). \tag{117}$$

If kk_{rocw}, kk_{rwcw} and kk_{rgcw} are held constant as k is varied, then all two-phase effective permeability functions remain unchanged, but Eq. 117 indicates that the three phase effective oil permeability function kk_{ro} varies with k. Thus, changes in the absolute permeability field will result in changes in predicted production data if three phases are mobile. Also note that if kk_{rocw} is held constant as k is varied, then both two phase flow effective oil permeabilities, kk_{row} and kk_{rog} will be unchanged, but the three-phase flow effective permeability function kk_{ro} will change unless effective gas and/or effective water permeability are altered. Thus, under three-phase flow conditioned it is not possible to alter absolute permeability while keeping all effective permeabilities constant. This means that changes in absolute permeability will result in changes in predicted reservoir performance. Thus, unlike the two-phase flow case, it may be possible to obtain reasonable estimates of both absolute permeability and relative permeability functions from three-phase flow production data. This result, however, is simply a peculiarity of Stone's Model II. For example, if one used the Dietrich and Bonder (1976) modification of Stone's model, then it is possible to obtain identical effective permeabilities under three-phase flow conditions for two different absolute permeability fields. The Dietrich-Bonder model, however, has other disadvantages; see Aziz and Settari (1979).

History Matching Examples

We consider two synthetic, two-dimensional three-phase flow problems. In the first example, the reservoir consists of three zones of uniform absolute permeability. In this case, we consider the estimation of relative permeability model parameters when absolute permeability is known and then consider the simultaneous estimation of absolute and relative permeabilities. In the second example, the true permeability field is an unconditional realization generated from a prior geostatistical model. In this case we generate a realization, rather than an estimate, of absolute and relative permeabilities. The information given in the following three paragraphs applies to all examples presented.

The reservoir simulation grid is $15 \times 15 \times 1$. The grid is uniform with areal dimensions given by $\Delta x = \Delta y = 12.2$ m (40 ft) and $\Delta z = 9.1$ m (30 ft). Throughout, gridblock (i, j) refers to the gridblock centered at (x_i, y_j, z_1). A water injection well (referred to as Well 5) is completed in gridblock $(8, 8)$. Water is injected through this well at a constant rate of 87.4 std m^3/d (550 STB/D) for all times. Producing wells are located in gridblocks $(3, 3)$ (Well 1), $(13, 3)$ (Well 2), $(13, 13)$ (Well 3) and $(3, 13)$ (Well 4). Each of these wells is produced at a constant total fluid rate of 34.98 m^3/d (220 RB/D). The synthetic production data used

to generate estimates is generated by running the reservoir simulator with the true model. Observed production data corresponds to the pressure, water-oil ratio and producing gas-oil ratio output from the simulator at thirty days intervals, i.e., at $t = 30, 60, \ldots, 300$ days at each producing well. At the water injection well, the flowing wellbore pressure at these times is also recorded. Thus, 130 data points are available for estimating the model parameters by automatic history matching. In the examples, we consider the effect of using only one type of data. Initial reservoir pressure is $31,026$ kPa (4500 psi) and initial bubble-point pressure is $30,454$ kPa (4417 psi) so pressure in gridblocks containing producing wells drop below bubble-point pressure very shortly after the beginning of production. For all examples considered water breaks through at all the producing wells prior to the end of the 300 day production period. After water breaks through at a well, all three phases are mobile in the well's gridblock.

Throughout, the natural logarithm of absolute permeability is simply referred to as log-permeability and denoted by $\ln(k)$. To generate the prior model, we assumed that log-permeability is a stationary Gaussian random function with the associated covariance function defined from an isotropic spherical variogram (Journel and Huijbregts, 1978). The range of the variogram is equal to 73.2 m (240 feet) and its sill is equal to 0.5. Thus the correlation range of $\ln(k)$ is 73.2 m and the variance is 0.5. In all examples, the prior mean for $\ln(k)$ is set equal to 4.0, so all elements of the vector m_{prior} (see Eq. 2) which correspond to gridblock log-permeabilities are set equal to 4.0. Relative permeabilities are represented by power law expressions. In all cases $S_{wc} = 0.2$, $S_{gc} = 0.05$, $S_{org} = 0.1$ and $S_{orw} = 0.2$. Throughout, porosity is assumed to be homogeneous with $\phi = 0.22$.

For the examples considered here, the variance of the pressure measurement error is set equal to 47.5 kPa^2 (1 psi^2), the variance of the GOR measurement error is set equal to 0.79 (std m^3)/(std m^3) (25 (scf/STB)^2). The WOR measurement error is defined by Eq. 106 with $\epsilon_o = 0.01$, $\epsilon_w = 0.02$ and $\sigma_{w,\min} = 0.32$ std m^3/d. These values were used to construct the data covariance matrix C_D. In the examples presented, however, no noise was added to the data generated from the simulation runs, that is d_{obs} was set equal to the vector of production data output from the simulator.

Example 1a.

In this case, the true reservoir model is not generated from an unconditional realization of the prior model, but is prescribed as a three zone reservoir as shown in Fig. 65. In all gridblocks contained within the lower left zone, $\ln(k) = 3.7$, at all gridblocks in the lower right zone, $\ln(k) = 4.3$ and in all gridblocks in the upper half $\ln(k) = 3.9$. The prior variances and prior

means for the parameters which characterize the power law relative permeability curves are defined in Table 4.

First we consider the case where the true log-permeability field is fixed at its true value and attempt only to estimate the model parameters defining the power law relative permeability curves. We consider the three cases where we construct the estimate of model parameters by history matching only one type of data (pressure, gas-oil-ratio, or water-oil ratio) and the case where we history match all 130 data to construct the estimate. The results of these four history matches are summarized in Table 5. In this and similar tables presented later, a column labeled true contains the true values of the parameters, a column labeled p_{wf} contains estimates of parameters obtained by history matching only pressure data, a column labeled GOR contains estimates of parameters obtained by history matching only producing gas-oil ratio data, a column labeled WOR contains estimates obtained by history matching only water-oil ratio data, and a column labeled "All" contains estimates obtained by history matching all observed data, pressure, GOR and WOR.

The results of Table 5 indicate that the estimates of the relative permeability parameters are in extremely close agreement with the true values when we history match all available data, only pressure data, or only GOR data. The estimates obtained by history matching only the WOR data are significantly less accurate. Fig. 66 shows the estimated two-phase gas and oil relative curves obtained by history matching only pressure data compared with the true relative permeability curves and the initial guesses for the relative permeability curves. In this and similar figures, the true relative permeability curves are shown as solid curves with no data points on them and the initial guesses for relative permeability curves used in the automatic history matching procedure are shown as curves through triangular data points. The estimated relative permeability curves are always represented in our results as open circular data points and open square data points. We did not draw the complete relative permeability curves through these open data points as it makes the figures too cluttered. The initial guesses for relative permeability curves are generated using the prior values of m_r, the model parameters that characterize the relative permeability curves. Note the initial guesses for relative permeability curves are far from the truth. As can be deduced from the results of Table 5, the relative permeability curves estimated by history matching only GOR data are also very close to the true case. Factors that contribute to the excellent estimates obtained for this example include the following: (i) the observed and predicted data are generated from the same simulator with identical spatial grids and time steps, so no modeling error (Tarantola, 1987) occurs; (ii) the observed production data that are history matched are not corrupted by measurement error, i.e. the observed production data are identical to data

obtained from the simulator run based on the true model; (iii) all rock and fluid properties except relative permeability parameters are known, so only the seven relative permeability model parameters are adjusted during the matching process (fewer model parameters tend to yield less uncertain estimates); (iv) the saturation ranges that occur during the 300 day time frame spanned by the production data vary significantly; gas saturation varies from 0.0 to almost 0.25 at producing well gridblocks and water saturation varies from about 0.2 to over 0.5 at producing well gridblocks and increases up to 0.8 at the injection well gridblock. Thus, production data is influenced by the values of relative permeabilities over large saturation ranges.

The results of Table 5 indicate that history matching either wellbore pressure data or producing GOR data yields good estimates of the parameters k_{rwcw} and n_{rw} which define the water relative permeability function. As pressure data at producing wells are quite sensitive to changes in total mobility (Thompson and Reynolds, 1997), it is not completely surprising that one can obtain reasonable estimates of water relative permeability model parameters even if we history match only pressure data corresponding to times prior to water breakthrough. It is not clear why the information content of the GOR data is sufficient to resolve the parameters defining the oil and water relative permeability curves for the two-phase water-oil system. But, as shown next, an improved understanding of these results can be obtained by considering dimensionless sensitivity coefficients.

As discussed in Zhang et al. (2002), dimensionless sensitivity coefficients provide a relative measure of how different data affect estimates of model parameters, and the uncertainty in these estimates. Here, the dimensional sensitivity of data d_i to model parameter m_j is given by

$$s_{i,j} = \frac{\partial d_i}{\partial m_j} \frac{\sigma_{m_j}}{\sigma_{d_i}}, \tag{118}$$

where $\sigma^2_{m_j}$ denotes the prior variance for model parameter m_j and $\sigma^2_{d_i}$ denotes the variance of the measurement error for the ith observed data. In essence, data corresponding to higher values of dimensionless sensitivity are expected to result in better estimates of model parameters, i.e., if $s_{i,j} > s_{k,j}$, then history matching data $d_{obs,i}$ is expected to give a better estimate of model parameter m_j than is obtained by history matching data $d_{obs,j}$; see Zhang et al. (2002) for additional discussion.

Fig. 67 shows the dimensionless sensitivity of the flowing bottom hole pressure at well 1 to the seven parameters defining the two sets of relative permeability curves as a function of time, and Fig. 68 presents a similar plot for the GOR dimensionless sensitivity coefficients. The results of Fig. 67 show that pressure is slightly sensitive to the water relative permeability parameters (n_{rw}, k_{rwcw}), even at times prior to breakthrough which occurs at about 200 days.

In fact up to 180 days, the dimensionless sensitivity of pressure to k_{rwcw} is roughly equal to the dimensionless sensitivity of pressure to the corresponding gas relative permeability parameter, k_{rgcw}. On the other hand, the results of Fig. 68 indicates that prior to 180 days, the GOR is almost insensitive to the water relative permeability parameters, but becomes very sensitive to these parameters after breakthrough. This indicates that the GOR data obtained after water breakthrough is primarily responsible for the good estimate of the water relative permeability curve (see Table 5) obtained by history matching GOR data.

To check the preceding supposition, we repeated the history match using only data up to 150 days. The results are shown in Table 6. Note in this case, history matching only GOR data gives relatively poor estimates of the parameters n_{rw} and k_{nrwc}, which define the water relative permeability curve. In fact, as expected, the results indicate that the estimates of these two parameters are essentially equal to their prior means. Matching only pressure data, however, gives good estimates of all relative permeability model parameters, although the estimates are not as accurate as those obtained by matching pressure data up to 300 days; see Table 5. Also note that the results of Table 6 indicate that the best estimates are obtained by matching both GOR and pressure data. In particular, the estimate of k_{rgcw} obtained by matching both pressure and GOR data is superior to that obtained by matching only one type of data. The dimensionless sensitivity of both GOR and pressure to this endpoint gas relative permeability are both fairly small, but the information content of the two sets of data is different because pressure is strongly influenced by the total mobility, but the GOR involves gas mobility divided by oil mobility. Because the information content for the two data types is different, history matching both pressure and GOR data yields an improved estimate of k_{rwcw}.

From the results of Figs. 67 and 68, one should note that both the dimensionless sensitivity of pressure to n_{rg} and the dimensionless sensitivity of GOR to n_{rg} are relatively large in magnitude. Thus, these dimensionless sensitivities suggest that n_{rg} should be accurately resolved by matching either pressure or GOR data. The results of Table 6 confirm this expectation. Note, however, that the GOR is much less sensitive to the gas relative permeability endpoint value k_{rgcw}, and prior to water breakthrough, the sensitivity of GOR to k_{rgcw} is very roughly equal to the negative of the sensitivity of GOR to the k_{rocw}, the end-point oil relative permeability in the two-phase gas-oil system. (The difference in sign of these sensitivities is expected because the GOR involves k_{rg}/k_{rog} at times prior to water breakthrough.) Thus, even though the GOR data resolves n_{rg} accurately, the estimates of the other parameters in the two-phase oil relatively permeability curves are not very accurate even though the corresponding estimates of k_{rg}/k_{rog} are quite close to the true ratio at values

of gas saturations that occur during the first 150 days of production. On the other hand, the results of Fig. 67 indicate that pressure is highly sensitive to k_{rocw}, n_{rg} and n_{rog}, three out of the four parameters that define the two- phase oil-gas relative permeability curves. Thus, history matching pressure should resolve these three parameters well. Pressure is somewhat less sensitive to the gas relative permeability endpoint k_{rgcw}, and the results of Table 6 indicate that a poorer estimate of this parameter is obtained by history matching pressure.

Example 1b.

Next, the assumption that the true $\ln(k)$ field was known was removed. In this case, $\ln(k)$ at each gridblock and the relative permeability parameters were estimated simultaneously by history matching various combinations of production data. The estimated log-permeability field obtained by history matching only pressure data is shown in Fig. 69. The corresponding results obtained by history matching wellbore pressure, gas-oil ratio and water-oil ratio are shown in Fig. 70. Matching all the data results in a more accurate estimate of the absolute permeability field than is obtained by matching only pressure data. In particular, matching only pressure results in several values of gridblock $\ln(k)$'s in the lower right quadrant and the upper center which are higher than the true values. Matching all data yields values somewhat closer to the truth. The over estimates of log-permeabilities are balanced by under estimates of the endpoint relative permeabilities, which are shown in Table 7. In this case, it looks like pressure data is effectively resolving effective permeabilities but not absolute permeability and relative permeabilities individually, which might cause one to question our discussion of Stone's model for three-phase relative permeabilities. Note, however, that when we condition to all data, we obtain good estimates of both log-permeability and the relative permeability parameters; see Fig. 70, Table 7, and Figs. 71 and 72. Fig. 71 shows the two-phase oil-gas relative permeability curves obtained by matching only pressure data, and Fig. 72 shows the improved results obtained by matching all data (wellbore pressures, producing WOR and producing GOR). As can be deduced from the results of Table 7, results of similar accuracy were obtained for the two-phase water-oil relative permeability curves. Although we did not show data matches for the examples considered, in all cases, we obtained excellent data matches. In fact, the matches were of similar quality to the ones that will be shown for the next example.

Example 2

In this example, an unconditional realization of log-permeability generated from the prior model is used as the true log-permeability field. The prior mean for $\ln(k)$ is still 4.0, and the prior means for relative permeability model parameters are shown in Table 8. In this case, we generate a realization of the model instead of the maximum a posteriori estimate. The randomized maximum likelihood method (Kitanidis, 1995; Oliver et al., 1996) is applied to generate a realization. To generate a realization with this procedure, we replace m_{prior} by m_{uc} (an unconditional realization of m generated from the prior model) and replace d_{obs} by a realization of the data d_{uc} (i.e., add noise) in Eq. 2 and then minimize the modified objective function. The unconditional realization m_{uc} is used as the initial guess when history matching d_{uc}.

Table 8 shows the realization of relative permeability model parameters obtained by history matching pressure, GOR and WOR. Figs. 73 and 74 present plots of the resulting realization of the relative permeability curves compared with the true curves and the initial guesses for the relative permeability curves. Fig. 75 shows the unconditional realization of the log-permeability field which was used as the initial guess in the randomized maximum likelihood method and Fig. 76 shows the realization of the log-permeability field estimated by history matching production data. Note the history matching process has radically altered the unconditional realization in the interwell region. This alteration is necessary to obtain a good match to the production data. Fig. 77 shows the producing gas-oil ratio predicted at well 1 based on the initial guess for model parameters (solid diamond data points) and the GOR data (solid triangular data points). The continuous curve represents the GOR predicted from the model obtained by history-matching. Note the initial model results in a GOR much lower than the observed GOR at late times. GOR history matches of the same quality were obtained at other wells. Fig. 78 shows the water-oil-ratio data and the data predicted from the model obtained by history matching. Note that good history matches of WOR were obtained at all four producing wells.

Remarks

It is important to note that data can directly provide information only on the part of the relative curves that correspond to phase saturations that exist within the reservoir. Since a power law functional form is used to represent all relative permeability curves, with each curve described by only two parameters, resolving relative permeabilities corresponding to low saturations accurately is essentially equivalent to resolving the complete curve accurately.

This explains why accurate estimates of relative permeability curves were obtained even though maximum gas saturation was less than 0.25. If the gas relative permeability function were modeled as a B-spline, the portion of the gas relative permeability curve corresponding to high gas saturations would not be resolved accurately by the available production data. For example, if the reservoir pressure remains above bubble-point pressure, the estimate of gas relative permeability model parameters would be determined by the prior model and would not be improved by matching production data.

Even though, arguments have been presented in favor of the Stone model (see, for example, Aziz and SettariAziz and Settari (1979)), Stone's model II leads to unsettling results. Namely, under two-phase flow conditions, one can not estimate accurately both absolute permeability and the coefficients in the power law relative permeabilities. An infinite set of values of absolute permeabilities and relative permeability coefficients will produce exactly the same production data when input to the simulator. We have shown, however, that this does not occur under three-phase flow conditions when Stone's Model II is used. In this case, reasonable estimates of both absolute and relative permeability functions can sometimes be obtained. Although a mathematical argument supporting this conclusion has been presented, we can provide no cogent physical explanation of why this should be the case, and believe a critical re-examination of Stone's Model II may be warranted.

When the objective is to estimate absolute permeability fields by history matching of production data, Li et al.Li et al. (2001a) have shown that the information content of pressure data is higher than the information content of GOR and both of these data types generally have more information content than WOR data. The results presented here suggest that this is also true when estimating absolute and relative permeabilities simultaneously.

	True	Mean	Var
n_{rw}	1.90	2.17	1.0
k_{rwcw}	0.40	0.58	0.04
n_{rg}	2.40	2.14	1.0
k_{rgcw}	0.90	0.49	0.04
n_{row}	2.60	2.05	1.0
n_{rog}	1.70	1.74	1.0
k_{rocw}	0.80	0.49	0.04

Table 4: Prior means and variances of relative permeability model parameters.

	True	p_{wf}	GOR	WOR	All
n_{rw}	1.90	1.91	1.91	1.95	1.90
k_{rwcw}	0.40	0.40	0.39	0.34	0.40
n_{rg}	2.40	2.38	2.40	2.32	2.40
k_{rgcw}	0.90	0.86	0.87	0.54	0.90
n_{row}	2.60	2.61	2.61	2.66	2.60
n_{rog}	1.70	1.69	1.73	1.75	1.70
k_{rocw}	0.80	0.80	0.79	0.68	0.80

Table 5: True and estimated relative permeability parameters; absolute permeability known.

	True	p_{wf}	GOR	p_{wf} + GOR
n_{rw}	1.90	1.96	2.17	1.91
k_{rwcw}	0.40	0.41	0.55	0.40
n_{rg}	2.40	2.34	2.34	2.39
k_{rgcw}	0.90	0.79	0.60	0.89
n_{row}	2.60	2.63	2.78	2.61
n_{rog}	1.70	1.66	1.93	1.69
k_{rocw}	0.80	0.80	0.65	0.80

Table 6: True and estimated relative permeability parameters; absolute permeability known; match of data up to 150 days.

	True	p_{wf}	All
n_{rw}	1.90	1.91	1.91
k_{rwcw}	0.40	0.35	0.39
n_{rg}	2.40	2.36	2.40
k_{rgcw}	0.90	0.72	0.87
n_{row}	2.60	2.68	2.62
n_{rog}	1.70	1.64	1.70
k_{rocw}	0.80	0.71	0.78

Table 7: True and estimated relative permeability parameters; $\ln(k)$ estimated simultaneously; Example 1b.

	True	Mean	All
n_{rw}	1.90	2.3	1.84
k_{rwcw}	0.40	0.5	0.40
n_{rg}	2.40	2.1	2.40
k_{rgcw}	0.90	0.8	0.96
n_{row}	2.60	2.3	2.59
n_{rog}	1.70	2.1	1.69
k_{rocw}	0.80	0.6	0.84

Table 8: True and estimated relative permeability parameters; $\ln(k)$ estimated simultaneously; heterogeneous reservoir example 2.

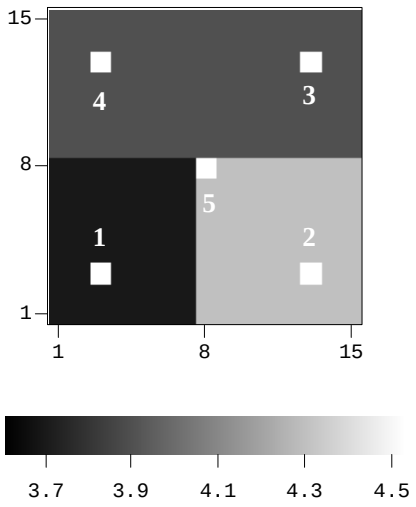


Figure 65: Three-zone reservoir, true model; Examples 1a and 1b.

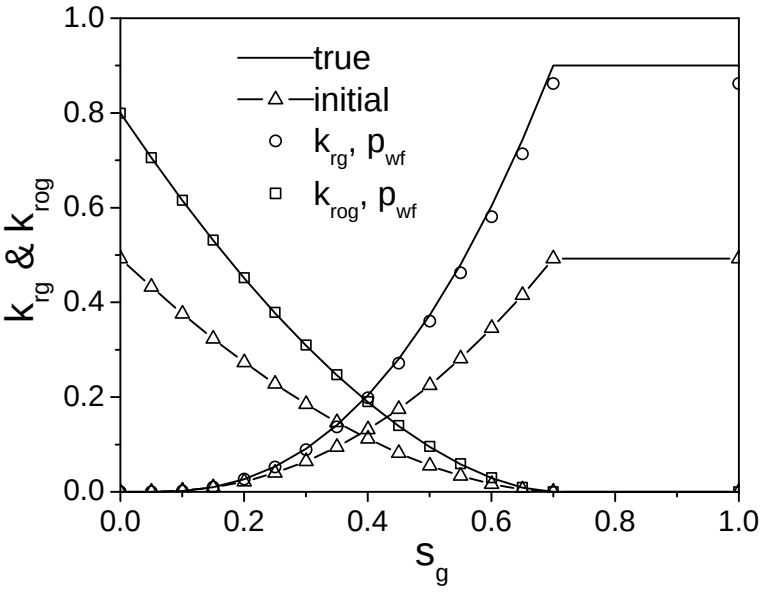


Figure 66: Prior, true and estimated gas-oil relative permeabilities, $\ln(k)$ known; Ex. 1a.

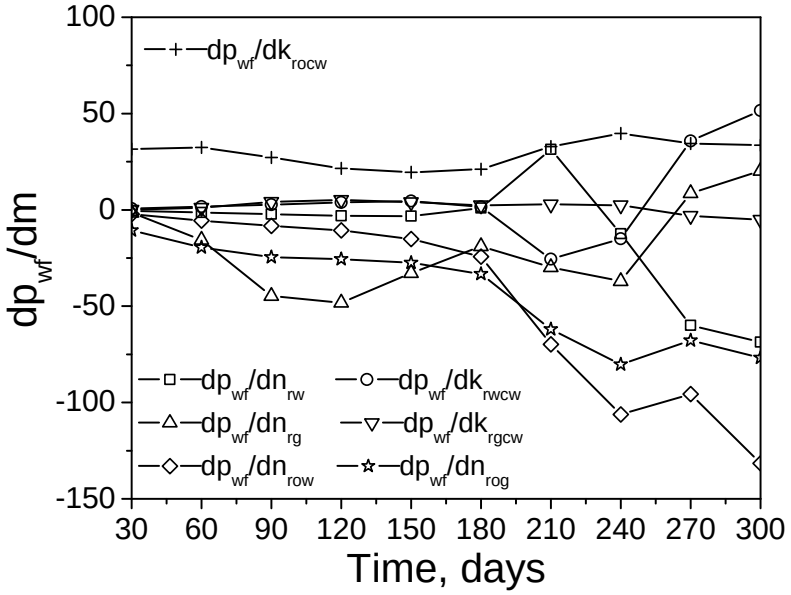


Figure 67: Dimensionless sensitivity of well 1 pressure to relative permeability model parameters; three-zone reservoir.

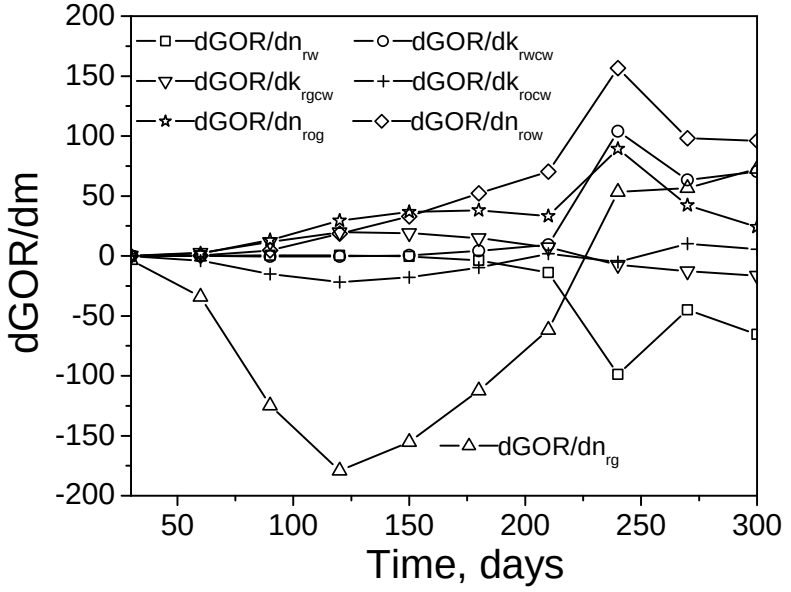


Figure 68: Dimensionless sensitivity of well 1 GOR to relative permeability model parameters; three-zone reservoir.

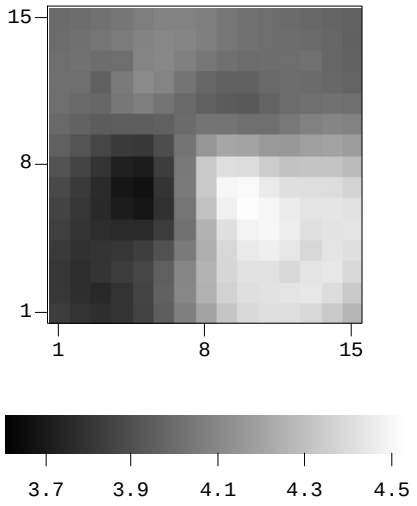


Figure 69: Log-permeability estimated from history match of pressure; Ex. 1b.

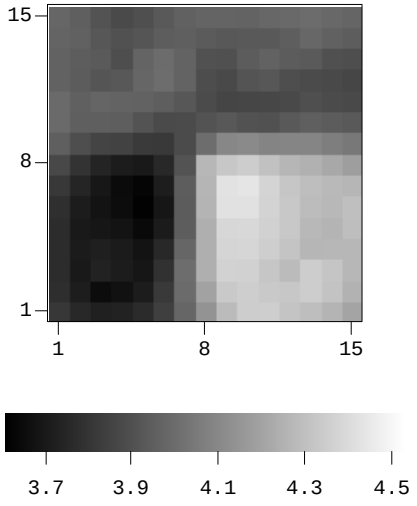


Figure 70: Log-permeability estimated from history match of pressure, GOR and WOR; Ex. 1b.

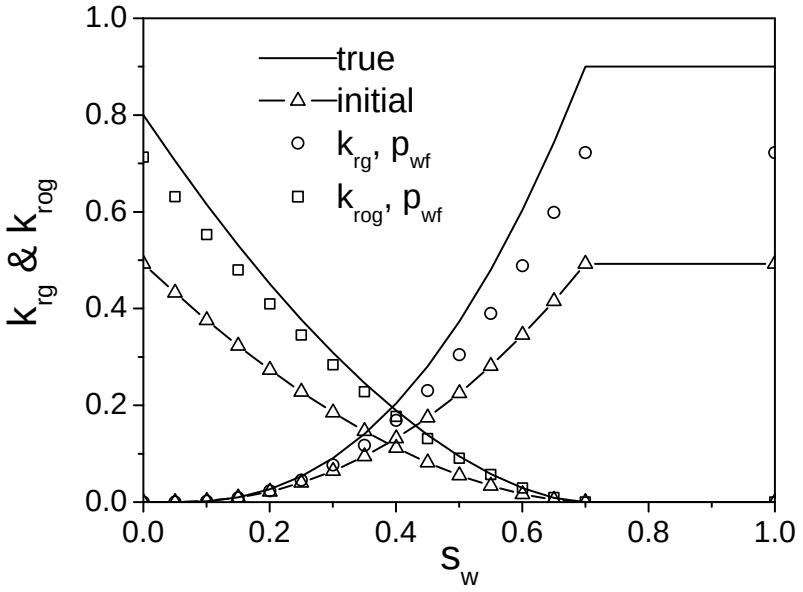


Figure 71: Prior, true and estimated gas-oil relative permeabilities, history match p_{wf}; Ex. 1b.

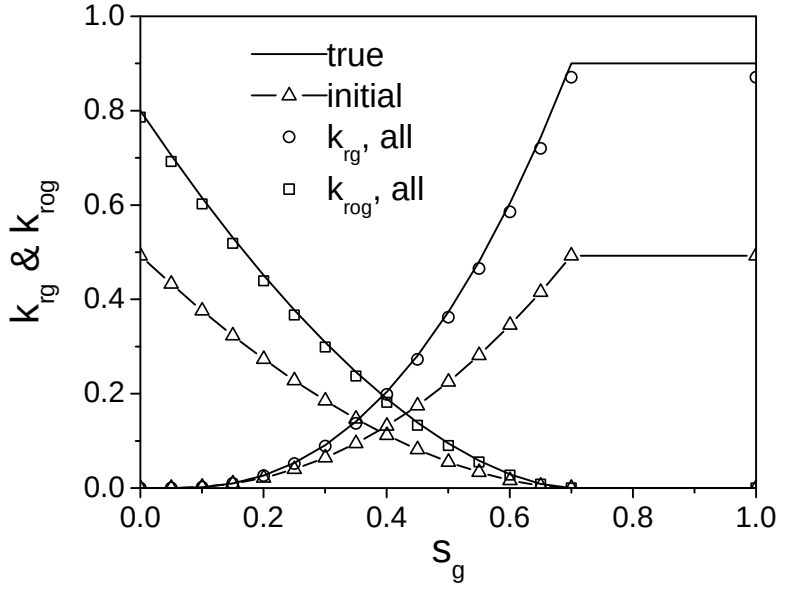


Figure 72: Prior, true and estimated gas-oil relative permeabilities, history match p_{wf}, GOR and WOR; Ex. 1b.

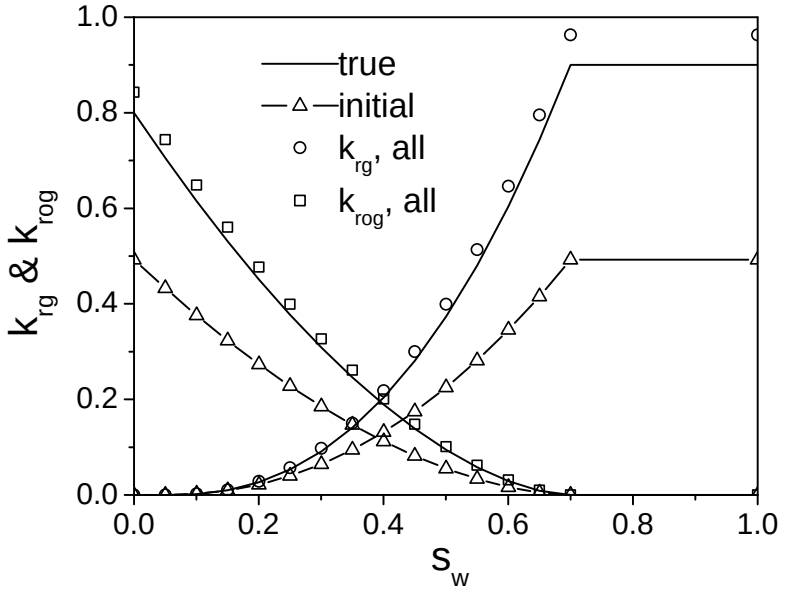


Figure 73: Prior, true and estimated gas-oil relative permeabilities, history match p_{wf}, GOR and WOR; Ex. 2.

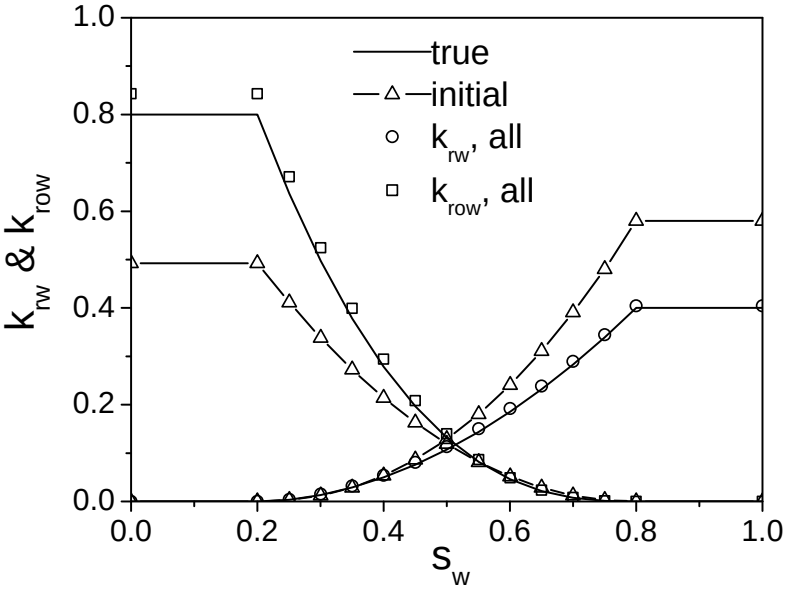


Figure 74: Prior, true and estimated water-oil relative permeabilities, history match p_{wf}, GOR and WOR; Ex. 2.

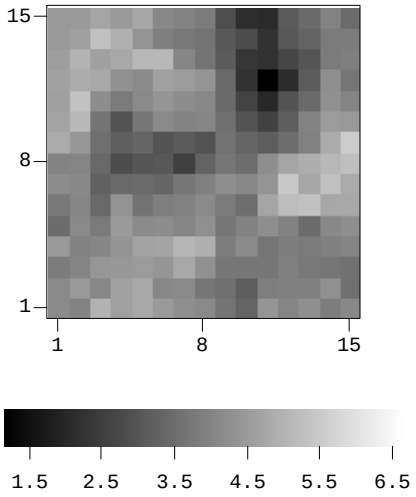


Figure 75: Unconditional realization of log-permeability.

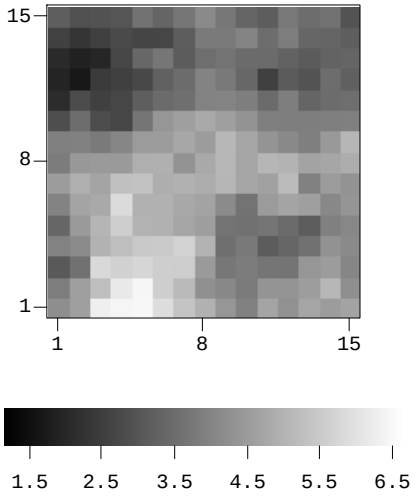


Figure 76: Conditional Realization of log-permeability.

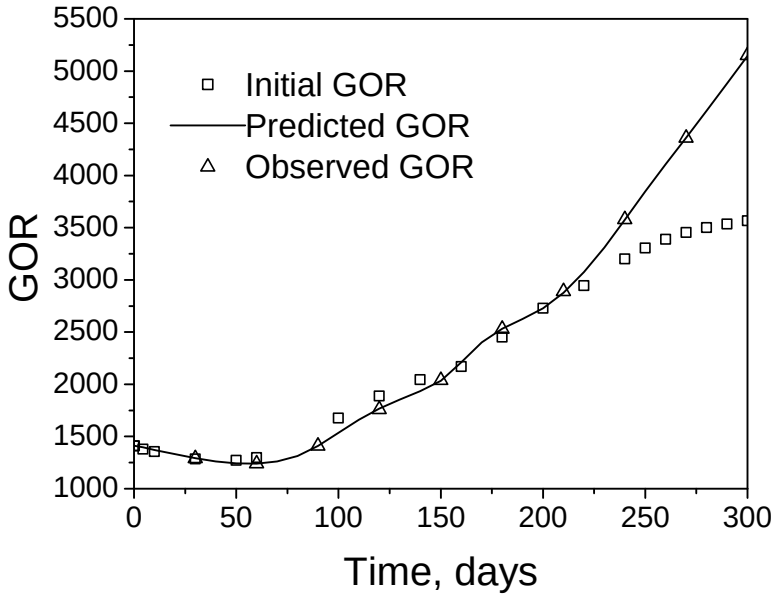


Figure 77: GOR data, GOR predicted with initial model and GOR predicted with history matched model.

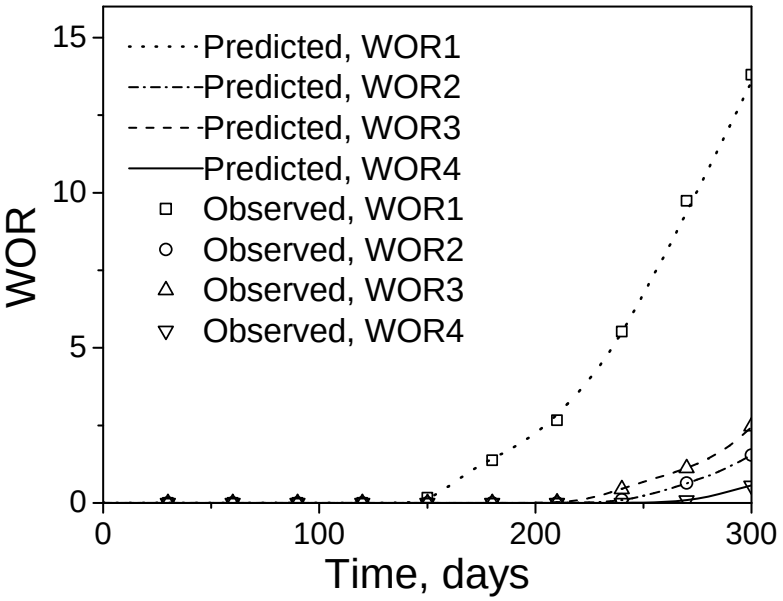


Figure 78: WOR data and history matched WOR.

CONCLUSIONS

The constrained limited memory BFGS algorithm based on log-transformation of model parameters can be used to alleviate undershooting and overshooting problems when estimating rock property fields by history matching production data.

As illustrated by our consideration of the Tengiz field example, the limited memory BFGS provides a feasible algorithm for realistic history matching problems. This example, illustrates, however, that when data are inconsistent, even a robust optimization algorithm can yield unrealistic estimates of rock property fields.

The limited memory BFGS algorithm was effective for history matching changes in acoustic impedance data for a synthetic solution-gas drive reservoir based on an example from the middle east. Results obtained using a combination of seismic and production data were more accurate than those obtained from either seismic or production data alone.

In many reservoirs knowledge of the location of geologic facies and their boundaries are critical to the problem of prediction of production. We have completed the early investigation of the problem of estimating geostatistical parameters for a truncated plurigaussian model, that will make it honor a training image. We have also demonstrated the feasibility of adjusting the locations of boundaries automatically to honor production data.

We have shown that it is feasible to construct reasonable estimates or realizations of the relative permeability curves and log-permeability fields by history matching production data obtained under three-phase flow conditions. The results assume a prior model is available for absolute permeability and the parameters that characterize the relative permeability functions. Although the results indicate that reasonably good estimates of model parameters may be obtainable by history matching only pressure data, history matching pressure, gas-oil ratio and water-oil ratio data together gives better results.

FUTURE WORK

In the future, we intend to consider the following aspects.

- Develop and implement code so that different zones or layers of the reservoir can have a different geostatistical model. In particular, this will allow the means for permeabilities and porosity to vary with depth.
- Test and an improved line search algorithm to improve the convergence performance of the limited memory BFGS algorithm.
- Redo the Tengiz field example using a variable weighting on data mismatch terms to account for the inconsistency between monthly rate data and pressure data.
- Apply our history matching procedure to the well known PUNQS3 example.
- Quantify the measurement and modelling error of time-lapse impedance data. This is important because an estimate of these errors is needed to time the objective function O_s which includes the seismic data misfit terms.
- Test the history-matching code for estimating rock property fields from time-lapse seismic data from a field in the Gulf of Mexico.
- Develop a method to estimate the probability density function (pdf) for the model parameters which define the plurigaussian model for facies so that this pdf is consistent with cross section maps provided by geologists.
- Refine the procedure for incorporating production data to estimate the location of boundaries between geologic facies as part of the history matching process.

Technical References

S. I. Aannonsen, A. Cominelli, O. Gosselin, I. Aavatsmark, and T. Barkve. Integration of 4d data in the history match loop by investigating scale dependent correlations in the acoustic impedance cube. In *8th European Conference on the Mathmatics of Oil Recovery*, 2002.

Y. Abacioglu, D. S. Oliver, and A. C. Reynolds. Efficient history-matching using subspace vectors. In *TUPREP Research Report 17 (May 22, 2000)*, pages 69–90, 2000.

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.

J. S. Archer and S. W. Wong. Use of a reservoir simulator to interpret laboratory waterflood data. *Soc. Petrol. Eng. J*, 12(6):343–347, 1973.

Eliana Arenas, Cor van Kruijsdijk, and Tanja Oldenziel. Semi-automatic history matching using the pilot point method including time-lapse seismic data. *SPE* 71634, 2001.

O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, New York, 1994.

Khalid Aziz and A. Settari. *Petroleum Reservoir Simulation*. Elsevier Applied Science Publishers, London, 1979.

Zhuoxin Bi. *Conditioning 3D Stochastic Channels to Well-Test Pressure Data*. Ph.D. thesis, University of Tulsa, Tulsa, Oklahoma, 1999.

Zhuoxin Bi, Dean S. Oliver, and Albert C. Reynolds. Conditioning 3D stochastic channels to pressure data. *SPE Journal*, 5(4):474–484, 2000.

Robert Bissell. Calculating optimal parameters for history matching. In *4th European Conference on the Mathematics of Oil Recovery*, 1994.

Robert Bissell, O. Dubrule, P. Lamy, P. Swaby, and O. Lepine. Combining goestatistical modelling with gradient information for history matching: The pilot point method, SPE 38730. In *Proceedings of the 1997 SPE Annual SPE Technical Conference and Exhibition*, pages 139–154, 1997.

Robert Bissell, Yogeshwar Sharma, and J. E. Killough. History matching using the method of gradients: Two case studies, SPE 28590. *SPE 69th Annual Technical Conference and Exhibition*, pages 275–289, 1994.

C. G. Broyden. Quasi-Newton methods and their application of function minimization. *Maths. Comp.*, 21:368–381, 1967.

C. G. Broyden. The convergence of a class of double rank minimization algorithm parts i and ii. *J. of Institute of Mathematics and its Applications*, 6:76–90 and 222–231, 1970.

A.G. Buckley. A combined conjugate-gradient quasi-Newton minimization algorithm. *Mathematical Programming*, 15:200–210, 1978.

A.G. Buckley and A. Lenir. QN-like variable storage conjugate gradients. *Mathematical Programming*, 27:155–175, 1983.

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, 1982.

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.

Kevin T. Chambers. Characterization of a carbonate reservoir using pressure-transient test and production logs: Tengiz field, kazakstan SPE 38657. In *Proceedings of the 1997 SPE Annual Technical Conference and Exhibition*, pages 155–170, 1997.

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.

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, 1974.

Lifu Chu, M. Komara, and R. A. Schatzinger. An efficient technique for inversion of reservoir properties using iteration method. *SPE Journal*, 5(1):71–81, 2000.

Lifu Chu, Albert C. Reynolds, and Dean S. Oliver. Computation of sensitivity coefficients for conditioning the permeability field to well-test data. *In Situ*, 19(2):179–223, 1995a.

Lifu Chu, Albert C. Reynolds, and Dean S. Oliver. Reservoir description from static and well-test data using efficient gradient methods. *SPE 29999 (1995 SPE International Meeting in Beijing)*, page 16 pages, 1995b.

M. Cooper, E.Thorogood, A.O'Donovan, P.Kristiansen, and P.Christie. Foinaven active reservoir management: The time-lapse signal. *SEG 1999 Expanded Abstracts*, 1999.

George Corliss, Christele Faure, Andreas Griewank, Laurent Hascoet, and Uwe Naumann, editors. *Automatic Differentiation of Algorithms, from Simulation to Optimization*. LNCSE. Springer, 2001.

G. de Marsily, G. Lavedan, M. Boucher, and G. Fasanino. Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model. In *Geostatistics for Natural Resources Characterization, Part 2*, pages 831–849. D. Reidel, 1984.

J. K. Dietrich and P. L. Bonder. Three-phase oil relative permeability models, SPE 6044. In *Proceedings of the 1976 SPE Annual Fall Meeting*, 1976.

Yannong Dong and Dean S. Oliver. Sensitivity of seismic impedance change to permeability and porosity. Technical Report TUPREP 2001, University of Tulsa, 2002.

R. Fletcher. A new approach to variable metric algorithms. *Computer Journal*, 13:317–322, 1970.

R. Fletcher and M. J. D. Powell. A rapidly convergent descent method for minimization. *Computer Journal*, 6:163–168, 1963.

R. Fletcher and C. M. Reeves. Function minimization by conjugate gradient. *Computer Journal*, 7:149–154, 1964.

Roger Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, New York, second edition, 1987.

A. Galli, H. Beucher, G. Le Loc'h, B. Doligez, and Heresim Group. The pros and cons of the truncated Gaussian method. In *Geostatistical Simulations*, pages 217–233. Kluwer Academic, Dordrecht, 1994.

F. Gassmann. Elastic waves through a packing of spheres. *Geophysics*, 16:673–685, 1951.

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.

D. Goldfarb. A family of variable-metric methods derived by variational means. *Maths. Comp.*, 24:23–26, 1970.

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.

D. Han. *Effects of porosity and clay content on acoustic properties of sandstones and unconsolidated sediments*. PhD thesis, Stanford University, 1986.

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, 1997.

Nanqun He and Kevin T. Chambers. Calibrate flow models with well-test data to improve history matching, SPE 56681. In *Proceedings of the 1999 SPE Annual Technical Conference and Exhibition*, pages 1–12, 1999.

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.

M. R. Hestenes and E. S. Stiefel. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand*, 46:409–536, 1952.

Xuri Huang, Laurence R. Bentley, and Claude Laflamme. Integration of production history and time-lapse seismic data guided by seismic attribute zonation. *SPE* 68819, 2001.

Xuri Huang, Laurent Meister, and Rick Workman. Improvement and sensitivity of reservoir characterization derived from time-lapse seismic data. *SPE* 49146, 1998.

Xuri Huang, Robert Will, Mashiur Khan, and Larry Stanley. Reservoir characterization by integration of time-lapse seismic and production data. *SPE* 38695, 1997.

John K. Hughes. Examination of seismic repeatability as a key element of time-lapse seismic monitoring. *SPE* 50627, 1998.

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

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

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.

André Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, New York, 1978. 600 p.

B. L. N. Kennett and P. R. Williamson. Subspace methods for large-scale nonlinear inversion. In *Mathematical Geophysics*, pages 139–154. D. Reidel, 1988.

P. D. Kerig and A. T. Watson. Relative-permeability estimation from displacement experiments: An error analysis. *SPE Res. Eng.*, 1(1):175–182, 1986.

P. D. Kerig and A. T. Watson. A new algorithm for estimating relative permeabilities from displacement experiments. *SPE Res. Eng.*, 1(1):103–112, 1987.

J. E. Killough, Yogeshwar Sharma, Alain Dupuy, Robert Bissell, and John Wallis. 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.

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

Tamara K. Kolda, Dianne P. O'Leary, and Larry Nazareth. BFGS with update skipping and varying memory. *SIAM J. Optim*, 8(4):1060–1083, 1998.

K. N. Kulkarni and Akhil Datta-Gupta. Estimating relative permeability from production data: A streamline approach. *SPE Journal*, 5(4):402–411, 2000.

Z. S. Labban and Roland N. Horne. Estimating relative permeabilities by automated history matching using linear modeling techniques (SPE-21206). In *11TH SPE Reservoir Simulation Symp.*, pages 47–55, 1991.

Jorge L. Landa and Roland N. Horne. A procedure to integrate well test data, reservoir performance history and 4-D seismic information into a reservoir description (SPE-38653). In *1997 SPE Annual Technical Conference and Exhibition*, 1997.

Jorge L. Landa, Roland N. Horne, Medhat M. Kamal, and C. D. Jenkins. Reservoir characterization constrained to well test data: A field example. *SPE Reservoir Evaluation & Engineering*, 3(4):74–79, 2000.

M. Landro. Discrimination between pressure and fluid saturation changes from time-lapse seismic data. *Geophysics*, 66(3):836–844, 2001.

Christian Lantuéjoul. *Geostatistical Simulation: Models and Algorithms*. Springer, Berlin, 2002.

G. Le Loc'h, H. Beucher, A. Galli, B. Doligez, and Heresim Group. Improvement in the truncated Gaussian method: Combining several Gaussian Functions. In *Proceedings of ECMOR IV, Fourth European Conference on the Mathematics of Oil Recovery*, 1994.

G. Le Loc'h and A. Galli. Truncated plurigaussian method: Theoretical and practical points of view. In *Geostatistical Simulations*, pages 211–222. Kluwer Academic, 1997.

Tai-Yong Lee and John H. Seinfeld. Estimation of absolute and relative permeabilities in petroleum reservoirs. *Inverse Problems*, 3(4):711–728, 1987.

Ruijian Li, A. C. Reynolds, and D. S. Oliver. History matching of three-phase flow production data, SPE 66351. In *Proceedings of the 2001 SPE Reservoir Simulation Symposium*, 2001a.

Ruijian Li, A. C. Reynolds, and D. S. Oliver. History matching of three-phase flow production data. *SPE J.*, 8(4):328–340, 2003.

Ruijian Li, Albert C. Reynolds, and Dean S. Oliver. Sensitivity coefficients for three-phase flow history matching. In *2001 Canadian International Petroleum Meeting*, 2001b.

D. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Math. Programming*, 45:503–528, 1989.

Ning Liu and Dean S. Oliver. Automatic history matching of geologic facies, SPE 84594. *Proceedings of the 2003 SPE Annual Technical Conference and Exhibition*, (History matching of facies boundaries using truncated pluri-Gaussian method.):1–15, 2003.

D.E. Lumley, A.G.Nunns, G.Delorme, and M.F.Bee. Meren field, nigeria: A 4d seismic case study. *SEG 1999 Expanded Abstracts*, 1999.

Randall L. Mackie and Theodore R. Madden. Three-dimensional magnetotelluric inversion using conjugate gradients. *Geophys. J. Int.*, 115:215–229, 1993.

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.

Kiyoshi Masumoto. Pressure derivative matching method for two phase fluid flow in heterogeneous reservoir. *SPE-59462*, 2000.

M.A. Meadows. Enhancements to landro's method for separating time-lapse pressure and saturation changes. *San Antonio 2001 Expanded Abstracts*, 2001.

Stephen G. Nash and Jorge Nocedal. A numerical study of the limited memory BFGS method and the truncated-Newton method for large scale optimization. *SIAM. J. OPTM.*, 1(3): 358–372, 1991.

L. Nazareth. A relationship between the BFGS and conjugate gradient algorithms and its implications for new algorithms. *SIAM. J. NUMER. ANAL.*, 16(5):794–800, 1979.

Jorge Nocedal. Updating quasi-Newton matrices with limited storage. *Math. Comp.*, 35: 773–782, 1980.

D. W. Oldenburg, P. R. McGillivray, and R. G. Ellis. Generalized subspace methods for large-scale inverse problems. *Geophys. J. Int.*, 114(1):12–20, 1993.

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

Henning Omre, Håkon Tjelmeland, Yuanchang Qi, and Leif Hinderaker. Assessment of uncertainty in the production characteristics of a sand stone reservoir. In *Reservoir Characterization III*, pages 556–603. PennWell Books, Tulsa, OK, 1993.

S. S. Oren. Self-scaling variable metric algorithm without line-search for unconstrained minimization. *Mathematics of Computation*, 27:873–885, 1973.

S. S. Oren. Self-scaling variable metric (SSVM) algorithms II: Implementation and experiments. *Management Science*, 20:863–874, 1974.

S. S. Oren and D.G. Luenberger. Self-scaling variable metric (SSVM) algorithms I: Criteria and sufficient conditions for scaling a class of algorithms. *Management Science*, 20:845–862, 1974.

S. S. Oren and E. Spedicato. Optimal conditioning of self-scaling variable metric algorithms. *Mathematical Programming*, 10:70–90, 1976.

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

D. W. Peaceman. Interpretation of well block pressures in numerical reservoir simulation. *Soc. Pet. Eng. J.*, 18(6):183–194, 1978.

D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation with non-square grid blocks and anisotropic permeability. *Soc. Pet. Eng. J.*, 23(6):531–543, 1983.

E. Polak. *Computational methods in optimization: a unified approach.* Academic Press, London, 1971.

M.J.D. Powell. Restart procedures for the conjugate gradient method. *Mathematical Programming*, 12:241–254, 1977.

Daniel Rahon, Georges Blanc, and Dominique Guérillot. Gradients method constrained by geological bodies for history matching. In *Proceedings of the 5th European Conference on the Mathematics of Oil Recovery*, pages 283–293, 1996.

Daniel Rahon, P. F. Edoa, and M. Masmoudi. Inversion of geological shapes in reservoir engineering using well-tests and history matching of production data. *1997 SPE Annual Technical Conference and Exhibition SPE 38656*, pages 141–150, 1997.

Banda S. RamaRao, A. Marsh LaVenue, Ghislain de Marsily, and Melvin G. Marietta. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1. Theory and computational experiments. *Water Resour. Res.*, 31 (3):475–493, 1995.

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.

G. B. Savioli and C. A. Grattoni. On the inverse problem application to reservoir characterization. *SPE-025522*, 1992.

D. F. Shanno. Conditioning of quasi-Newton method for function minimization. *Maths. Comp.*, 24:647–656, 1970.

D. F. Shanno. Conjugate gradient methods with inexact searches. *Mathematics of Operation Research*, 3:244–256, 1978a.

D. F. Shanno. On the convergence of a new conjugate gradient algorithm. *SIAM Journal on Numerical Analysis*, 15:1247–1257, 1978b.

P. M. Sigmund and F. G. McCaffery. An improved unsteady-state procedure for determining the relative-permeability characteristics of heterogeneous porous media. *Soc. Petrol. Eng. J.*, 19(1):15–28, 1979.

H. L. Stone. Estimation of three-phase relative permeability and residual oil data. *J. Can. Pet. Tech.*, 12(4):53–61, 1973.

Thomas B. Tan. A computationally efficient gauss-newton method for automatic history matching, SPE-29100. *Proceedings of the 13th SPE Symposium on Reservoir Simulation*, pages 61–70, 1995.

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

L. G. Thompson and A. C. Reynolds. Well testing for radially heterogeneous reservoirs under single and multiphase flow conditions. *SPE Formation Evaluation*, 12(1):57–64, 1997.

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

A. Tura and D.Lumley. Estimating pressure and saturation changes from time-lapse avo data. *Extended Abstracts, SEG Annual Meeting*, 1999.

R. van Ditzhuijzen, T. Oldenziel, and van Kruijsdijk. Geological parameterization of a reservoir model for histroy matching incorporating time-lapse seismic based on a case study of the statfjord field. *SPE* 71318, 2001.

M. L. Wasserman, A. S. Emanuel, and J. H. Seinfeld. Practical applications of optimal-control theory to history-matching multiphase simulator models. *Soc. Petrol. Eng. J.*, 15 (4):347–355, 1975.

A. Ted Watson, G. R. Gavalas, and John H. Seinfeld. Identifiability of estimates of two-phase reservoir properties in history matching. *Soc. Petrol. Eng. J.*, 24(6):697–706, 1984.

A. Ted Watson, J. H. Seinfeld, G. R. Gavalas, and P. T. Woo. History matching in two-phase petroleum reservoirs. *Soc. Petrol. Eng. J.*, 20(12):521–532, 1980.

W. T. Watson, P. C. Richmond, P. D. Kerig, and T. M. Tao. A regression-based method for estimating relative relative permeabilities from displacement experiments. *SPE Res. Eng.*, 43(5):953–958, 1988.

Zhan Wu, A. C. Reynolds, and D. S. Oliver. Conditioning geostatistical models to two-phase production data. *Soc. Petrol. Eng. J.*, 3(2):142–155, 1999.

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.

Pin-Huel Yang and A. Ted Watson. Automatic history matching with variable-metric methods. *SPE Reservoir Engineering*, 3(3):995–1001, 1988.

Pin-Huel Yang and A. Ted Watson. A Bayesian methodology for estimating relative permeability curves. *SPE Reservoir Engineering*, 6(2):259–265, 1991.

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

Tai yong Lee and John H. Seinfeld. Estimation of two-phase petroleum reservoir properties by regularization. *J. Computational Physics*, 69:397–419, 1987.

F. Zhang and A. C. Reynolds. Optimization algorithms for automatic history matching of production data. *Proceedings of 8th European Conference on the Mathematics of Oil Recovery*, page 10 pages, 2002a.

F. Zhang and A. C. Reynolds. Optimization algorithms for history matching. In *TUPREP Research Report 19*, pages 14–105. The University of Tulsa, 2002b.

F. Zhang, A. C. Reynolds, and D. S. Oliver. Evaluation of the reduction in uncertainty obtained by conditioning a 3D channel to multiwell pressure data. TUPREP report, The University of Tulsa, 2001.

F. Zhang, A. C. Reynolds, and D. S. Oliver. Evaluation of the reduction in uncertainty obtained by conditioning a 3d stochastic channel to multiwell pressure data. *Mathematical Geology*, 34(6):713–740, 2002.

Fengjun Zhang. *Automatic History Matching of Production Data for Large Scale Problems*. Ph.D. thesis, University of Tulsa, Tulsa, Oklahoma, 2002.