SPE 92867

Critical Evaluation of the Ensemble Kalman Filter on History Matching of Geologic Facies

N. Liu, SPE, and D.S. Oliver, SPE, U. of Oklahoma

This paper was prepared for presentation at the SPE 2005 Reservoir Simulation and Symposium held in Houston, Texas,U.S.A., 31 January 2005 - 2 February 2005.

Abstract

The objective of this paper is to compare the performance of the ensemble Kalman filter (EnKF) to the performance of a gradient-based minimization method for the problem of estimation of facies boundaries in history matching. EnKF is a Monte Carlo method for data assimilation that uses an ensemble of reservoir models to represent and update the covariance of variables. In several published studies, it outperforms traditional history matching algorithms in adaptability and efficiency.

Because of the approximate nature of the EnKF, the realizations from one ensemble tend to underestimate the uncertainty especially for problems that are highly nonlinear. In this paper, the distributions of reservoir model realizations from 20 independent ensembles are compared with the distributions from 20 randomized maximum likelihood (RML) realizations for a 2D water-flood model with one injector and four producers. RML is a gradient based sampling method that generates one reservoir realization in each minimization of the objective function. It is an approximate sampling method, but its sampling properties are similar to Markov chain Monte Carlo method (McMC) on highly nonlinear problems and relatively more efficient than the McMC.

Despite the nonlinear relationship between data such as production rates and facies observations, and the model variables, the EnKF was effective at history matching the production data. We find that the computational effort to generate 20 independent realizations was similar for the two methods, although the complexity of the code is substantially less for the EnKF.

Introduction

Several questions regarding the use of the ensemble Kalman filter for history matching are addressed in this paper. The most important is a comparison of the efficiency with a gradient-based method for a history matching problem with known facies properties but unknown boundary locations. Secondly, the ensemble Kalman filter and a gradient-based method are unlikely to give identical estimates of model variables, so it is also important to know if one method generates better realizations. Finally, since there is often a desire to use the history matched realizations to quantify uncertainty, it is important to determine if one of the methods is more efficient at generating independent realizations.

Gradient-based history matching can be performed several ways (e.g. assimilating data in batch or sequentially); a variety of minimization algorithms can be used (e.g. conjugate gradient or quasi-Newton), and several different methods for computing the gradient are available (e.g. adjoint or sensitivity equations). In this paper, we use what we believe is the most efficient of the traditional gradient-based methods: an adjoint method to compute the gradient of the squared data mismatch [1;2;3;4] and the limited memory Broyden-Fletcher-Goldfarb-Shanno method [5] (LBFGS) to compute the direction of the change [6;7;8]. The remaining choice is whether to incorporate all data at once or to incorporate the data sequentially.

Simultaneous, or batch, inversion of all data is clearly a well-established history matching procedure. Although data from wells or sensors may arrive nearly continuously, the practice of updating reservoir models as the data arrive is not common. There are several reasons that make sequential assimilation of data difficult for large, nonlinear models: (1) the covariance for all model variables must be updated as new data are assimilated, but the covariance matrix is very large, (2) the covariance may not be a good measure of uncertainty for nonlinear problems, and (3) the sensitivity of a datum to changes in values of model variables is expensive to compute. Bayesian updating in general is described by Woodbury [9]. Modifying a method described by Tarantola [10], Oliver [11] evaluated the possibility of using a sequential assimilation approach for transient flow in porous media. He found that the results from sequential assimilation could be almost as good as from batch assimilation if the order of the data was carefully selected. The problem was quite small, however, and an extension to large models was impractical.

Although a sequential method has the advantage that it generates a sequence of history matched models that may all be useful at the time they are generated, our comparisons of efficiency will primarily be based on the effort required to assimilate all of the data. If the intermediate predictions are needed (as they would be for control of a reservoir) the comparison provided here will underestimate the value of the sequential assimilation.

A secondary objective of history matching is often to assess the uncertainty in the predictions of future reservoir performance or in the estimates of reservoir properties such as permeability, porosity or saturation. In general, uncertainty is estimated from an examination of a moderate number of conditional simulations of the prediction or properties. Unless the realizations are generated fairly carefully and the sample is sufficiently large, however, the estimate of uncertainty could be quite poor. Two large comparative studies of the ability of Monte Carlo methods to quantify uncertainty in history matching have been carried out, one in groundwater[12] and one in petroleum [13]. Neither was conclusive, partly because of the small sample size. Liu and Oliver[14] used a smaller reservoir model (fewer variables), but much larger sample size. They found that the method that minimizes an objective function containing a model mismatch part and a data mismatch part, with noise added to observations, created realizations that were distributed nearly the same as realizations from Markov chain Monte Carlo.

The ensemble Kalman filter is a Monte Carlo method for updating reservoir models. It solves several problems with the application of the Kalman filter to large nonlinear problems[15;16]. It has been applied to reservoir flow problems with generally good results[17;18;19]. There has been no examination, however, of the distribution of the members of a single ensemble. The adequacy of the uncertainty estimate is completely unknown.

In the first paper on the ensemble Kalman filter, Evensen[15] described how the evolution of the probability density function for the model variables can be approximated by the motion of "particles" or ensemble members in phase space. Any desired statistical quantities can be estimated from the ensemble of points. When the size of the ensemble is relatively small, however, the approximation of the covariance from the ensemble almost certainly contains substantial errors. Houtekamer and Mitchell[20] noted the tendency for reduction in variance due to "inbreeding". When the ensemble estimate is used in a Kalman filter, van Leeuwen[21] explained how nonlinearity in the covariance update relation causes growth in the error as additional data are assimilated.

In this paper, the comparison is made using history matching on a truncated plurigaussian model[22;23] for geologic facies. It provides a difficult history matching problem with significant nonlinearities[24] that make both the ensemble Kalman filter and the limited memory BFGS method difficult to apply.

Methodologies

The problem of estimating the location of facies boundaries is difficult in history matching, partly because of the geological complexity. The truncated plurigaussian method for modeling geologic facies is useful in this aspect not only for the wide variety of textures and shapes that can be generated, but also because of the internal consistency of the stochastic model[23;25]. For both gradient and EnKF approaches, the truncated plurigaussian method is used to simulate the facies distributions. In the first step for simulating a facies field, two random Gaussian fields Y_1 and Y_2 are generated based on the geological spatial features, so that each gridblock is assigned two Gaussian variables. In the second step, a set of truncation thresholds is applied to the Gaussian variables, and the facies are assigned to gridblocks depending on the location of Y_1 and Y_2 within the truncation map. We have elected to use three intersecting lines as truncation thresholds. Three non-parallel randomly generated lines that do not all intersect at 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 modelled with appropriate relative percentage.

Randomized Maximum Likelihood. A standard method for quantify the uncertainty in reservoir simulation predictions is to generate multiple conditional reservoir model realizations, and predict future performance of each. The Randomized Maximum Likelihood (RML) method generates realizations conditional to nonlinear data from unconditional realizations in a Gaussian random field by a process of minimization. It has been shown to have good sampling properties for history matching problems with highly nonlinear relationship between the data and the model parameters[14]. If the prior covariance of the model parameters and the variance of the observed data are known, matched models can be generated as follows:

1. Generate an unconditional realization of the model parameters, $m_u \leftarrow N[m_{\rm pr}, C_M]$.
2. Generate a realization of the data, $d_u \leftarrow N[d_{\rm obs}, C_D]$.
3. Compute the set of model variables, m, that minimizes the function:

$$O(m) = \frac{1}{2}(m-m_u)^{\rm T}C_M^{-1}(m-m_u) + \frac{1}{2}(g(m)-d_u)^{\rm T}C_D^{-1}(g(m)-d_u)$$

Solving a minimization problem is required to generate each matched model. In this study, the limited memory BFGS method is used in computing the search direction, and the gradients of the objective function with respect to the permeability and the porosity fields, $\nabla_k O$ and $\nabla_\phi O$, were obtained from the adjoint method for general automatic history matching of reservoir property fields[26].

Ensemble Kalman Filter. The idea of Ensemble Kalman filter for continuous model updating is propagating an ensemble of initial reservoir models along time to assimilate data, and the statistical information carried among the models at each observation time is used to update the model covariance. The ensemble of state vectors is denoted by Ψ:

$$\Psi = \{y_1, y_2, \ldots, y_{N_e}\},$$

where N_e is the number of ensemble members; y_i for $i = 1, N_e$ are state vectors. Each of the state vectors in the ensemble Kalman filter contains all the uncertain and dynamic variables that define the state of the system. At a certain time step t_k for $k = 1, N_t$, the state vector for the reservoir model is expressed as:

$$y_k = [(m_k)^{\mathrm{T}}, d(m_k)^{\mathrm{T}}]^{\mathrm{T}},$$

where m_k consists of variables for rock properties and flow system in every gridblock, $d(m_k)$ is the simulated data from the model state m_k. The number of simulated data in the vector $d(m_k)$ does not have to be constant since it depends on the number of observation data at time step t_k.

The methodology of ensemble Kalman filter for data assimilation only consists of two sequential steps. One is the forecast forward in time based on solution of the dynamical equations for flow and transport in the reservoir:

$$y^p_{k,j} = f(y^u_{k-1,j}), \text{ for } j = 1, N_e,$$

where $f(\cdot)$ is the reservoir simulator. The superscript p indicates the "predicted" state. This step does not modify the rock properties, but replaces the pressure, saturation, and simulated data in the predicted state vector. The initial ensemble for $k = 1$ refers to the collection of initial state vectors, which are sampled from the prior probability density function of the state vector before any data assimilation.

The second step is model updating by correcting the variables describing the state of the system to honor the observations. The update to each ensemble member is made using the Kalman update formula:

$$y^u_j = y^p_j + K_e(d_j - Hy^p_j), \text{ for } j = 1, \ldots, N_e$$

where the superscript u denotes "updated", and K_e is the ensemble Kalman gain. H is the measurement operator that extracts the simulated data from the state vector y^p:

$$H_k = \begin{bmatrix} 0 & I \end{bmatrix}.$$

d_j is the observation data at current time plus random error that has the same distribution with the measurement error:

$$d_j = d_{\mathrm{obs}} + \epsilon_j, \text{ for } j = 1, \ldots, N_e.$$

The ensemble Kalman gain is computed as:

$$K_e = C_{\Psi,e}H^T(HC_{\Psi,e}H^{\mathrm{T}} + C_D)^{-1},$$

where the covariance matrix of the state vectors $C_{\Psi,e}$ at any time can be estimated from the ensemble members by the standard statistical definition:

$$C_{\Psi,e} = \frac{1}{N_e - 1}\sum_{i,j=1}^{N_e}(y^p_i - \bar{y}^p)(y^p_j - \bar{y}^p)^{\mathrm{T}},$$

where the indices i and j are for numbering of the ensemble members. $\bar{y}^p$ is the mean of the N_e ensemble members at the current data assimilation step. If the size of each state vector is N_y, the covariance matrix $C_{\Psi,e}$ is $N_y \times N_y$. It is not possible to compute or store $C_{\Psi,e}$ for problems that are not too small. Fortunately, the formulation of the ensemble Kalman gain allows us to compute $HC_{\Psi,e}$ instead of $C_{\Psi,e}$ itself. $HC_{\Psi,e}$ is the last N_d rows of $C_{\Psi,e}$.

In this study, the geostatistical models for generating the two Gaussian fields are assumed to be known, and the static variables to be modified in history matching are the random Gaussian fields Y_1 and Y_2. As the hard data measurements do not depend on the dynamic states of the reservoir fluid flow, the state vector for cases with only facies measurements is $y_j = \{Y_1, Y_2, d_{\mathrm{sim}}\}$. The facies measurements can be assimilated one at a time to simulate the process of sequential well placement, in which case d_{sim} is the facies type of the simulated facies field at the current observation location. When there are production data in d_{sim}, the state vector includes the pressure and the saturation in every gridblock, $y_j = \{Y_1, Y_2, P, S, d_{\mathrm{sim}}\}$. Both Gaussian fields have the same size as the reservoir grid, therefore the size of the state vector is $N_y = 4 \times n_{\mathrm{grid}} + n_d$, where n_d is the number of data obtained at each observation time.

Matching Hard Data and Production Data

The test case is a reservoir model on a $50 \times 50 \times 1$ grid. The dimensions of each gridblock is 30 ft$\times$30 ft$\times$20 ft. The covariance of the random Gaussian field Y_1 is Gaussian type with the principle direction 60° east of north. The range in the principle direction is one third of the field width, and twice the range in the perpendicular direction. The covariance of the random Gaussian field Y_2 is same with that of Y_1 except that the anisotropy angle is 45° east of north. Figs. 1(a) and 1(b) show a pair of unconditional Gaussian fields (Y_1, Y_2) with the specified covariances. Three facies are present in the field, which are denoted as facies 1, facies 2, and facies 3. An unconditional facies map as shown in Fig. 1(d) is generated by truncating the two unconditional Gaussian fields Y_1 and Y_2 using a truncation scheme in Fig. 1(c). The facies in dark grey is facies 1, in light grey is facies 2, and in white is facies 3. The covariances of the two Gaussian fields and the truncation scheme are assumed to be known during history matching, and used to simulate all the facies realizations.

The true facies field is shown in Fig. 2. There is one injector near the center and four producers at the corners.

The facies observations are listed in Table 1 with the well number and locations. The rock properties are constant within a facies type, but distinct among the facies. Table 2 presents the permeability and the porosity for each facies type. The injection rate is fixed at 4600 rb/day for well 1, and the production rates are fixed at 1300 reservoir barrel of total fluid per day for wells 2, 3, and 4, respectively. Well 5 is in a low productivity region, and the production rate is fixed at 600 rb/day. The initial reservoir pressure is 3800 psia, which is far above the bubble point of 500 psia. The field is produced for 80 days and the bottom-hole-pressure at all wells are recorded at day 2, and every 10 days beginning at day 10. There are 45 bottom-hole-pressure measurement and 5 facies observations from all wells. The measurement error for pressure data is assumed to be distributed normally with mean 0 and standard deviation of 3 psi. The facies observations are assumed to be exact. The same reservoir model was used to evaluate both the RML and the EnKF methods. Because of small differences in the way date are entered for the two reservoir simulators CLASS and ECLIPSE, the "observed" data from the two simulators are only compared with the "true" data from the same simulator.

Gradient approach. The traditional approaches to history matching assimilate all the data at the same time. Because the simulation of the process is so expensive, an efficient method of modifying the model parameters to match the observations is required. In our approach, the adjoint method is used to compute the gradient of the objective function with respect to the model variables Y_1 and Y_2. The adjoint system is complicated in development, and dependent on the specific reservoir simulator for which it was developed. Our adjoint system was built for the Chevron Limited Application Simulation System (CLASS) simulator[26;8].

We start the history matching with the initial models having the correct facies types at well locations. The matching of facies data is done efficiently using the ensemble Kalman filter method. During the process of history matching to production data, the step-size of the modifications to the model parameters is restricted to ensure that the facies at well locations are maintained to be correct.

Fig. 3 uses box plots to compare the distributions of the simulated BHP data from the 20 accepted RML realizations to the observed data. The "boxes" includes the range from P_{25} to P_{75}, while the lines include the range from P_{10} to P_{90}. The distributions of the simulated bottom-hole pressure from the 20 RML models are much wider than expected, based on the assumption that the magnitude of the noise is about 3 psi. As the facies type is an indicator variable, in order to compute the gradient of the objective function with respect to the facies, we added a transition zone at facies boundaries. This transition zone is only for the purpose of gradient approximation, and does not exist in the simulation process. Unlike history matching of permeability and porosity, the minimization of the objective function for history matching of the discontinuous facies often stops at a relatively high objective function value, because the gradient is only approximately correct. A typical minimization required approximately 11 iterations. We generated 100 initial models, but only used the 20 models with the lowest final objective function value.

EnKF for history matching. The states variables in the Kalman filter consist of two Gaussian fields for facies description, pressure and saturation fields, simulated bottom-hole pressure data from each well, and simulated facies at each well location: $\{Y_1, Y_2, P, S, P_{\mathrm{wf,sim}}, F_{\mathrm{sim}}\}$. 40 state vectors are included in an ensemble. The initial ensemble of state vectors are conditional to the facies observations using EnKF. Unless care is taken, updating of the Gaussian fields from matching production data may change the facies type at well locations. Once the facies type at a well location is wrong, the Kalman correction to the Gaussian fields can become large, and may cause over-shoot of the Gaussian variables. An EnKF iteration for facies observations is made after each model update to ensure the updated rock properties at well locations are always correct.

The first row of Fig. 4 shows 4 out of the 40 initial facies maps, and the second row shows the corresponding facies maps simulated from the final ensemble members after assimilating all the production data. The initial facies maps all have unique local structures, and the final facies maps have developed common features among the ensemble members. Some of the common features do not exist in the true facies map.

Box plots are used to represent the distributions of the simulated production data from all the 40 ensemble members over the 80 days of production. The reservoir simulator ECLIPSE is used in the EnKF approach for adaption to supercomputers. The simulated bottom-hole pressure and the observed data are plotted together in Fig. 5. The box plots in the first column show the bottom-hole pressure from the initial ensemble conditional only to facies observations. The box plots in the second column show bottom-hole pressure after assimilation to pressure data. The observed injection rate is plotted in the thick line. In the second column of Fig. 5, the distributions of the bottom-hole pressure from the 40 final reservoir models are much narrower than the initial distributions and are centered at the observed data.

A single ensemble of 800 state members was also generated and assimilated to the observed data. Column 3 of Fig. 5 shows the distributions of the simulated BHP from all the 800 matched final models. The matching quality is clearly at least as good as that using 40 ensemble members, as the width of the boxes is very narrow.

Discussion

This test problem is somewhat unusual in that the objective function for the problem, which includes facies data mismatch, is not differentiable, so it was necessary to intro-

duce an approximate objective function with a transition between facies. The quasi-Newton minimization method should have converged fairly quickly near the minimum, but did not do so in this case. And the final values of the objective function were larger than expected. We simply chose the 20 best realizations from 100 minimizations. A typical history matching minimization required 11 iterations, each iteration requiring CPU time equivalent to approximately 5 forward simulation runs. The effort required to generate the 20 independent realizations using LBFGS was approximately 100 starting models $\times$ 11 iterations per model $\times$ 5 simulations per iterations giving 5500 simulations, or 275 simulations per history matched model.

The ensemble Kalman filter method with 40 members in an ensemble required 40 simulation runs to generate 40 history matched models, but it is clear from Fig. 4 that the variance within an ensemble is too small, so the 40 models are not independent realizations. Realizations from the EnKF method, as judged by the median value of the objective function (6220), are much better than those from the LBFGS method which gave a larger median value for the objective function (21,300).

The tendency for the variance within an ensemble to reduce to a level that is smaller than required by the data has been previously noted for the EnKF method. Although the quality of the matching to the observations is satisfactory for all the 40 ensemble members, the significant reduction of the variations among the ensemble of matched facies models lead us to investigate the capability of EnKF in uncertainty quantification.

In practice, the number of state members in an ensemble is empirically chosen between 40 and 100. When the number is too small, the variation carried from the initial ensemble is not sufficient in approximation of the state covariances for each model updating. Consequently, the matched models may become highly correlated, and clustered at a small subspace of the real model probability distribution. In that case, the 40 final models from the same ensemble are equivalent to one well matched RML realization for uncertainty quantification. On the other hand, each ensemble member requires one reservoir simulation run plus some computational overhead for data assimilation. The EnKF approach becomes more computationally expensive the more ensemble members are included.

We also generated 20 ensembles with 40 state members in each, such that the total number of states is 800. Each of the matched models in the 20 small ensembles and the large ensemble of 800 members was simulated in ECLIPSE for water cut prediction from each of the four producers at day 140. The distributions of the predicted water cut from each ensemble are plotted in box-plots as shown in Fig. 6. The boxes numbered from 1 to 20 are distributions from the ensembles with 40 members. The last box in each plot at number 21 is the distribution from the 800 member ensemble. The thick lines are predicted water cut from the true model, where the 80 days production data were generated. None of the distributions of the 20 small ensembles covers the predicted water cut for all four producers. Some of the ensembles predicted very small uncertainties, such as number 3 and 18. And none of the small ensembles yields distributions that are large enough to represent the distributions from all other small ensembles.

The water cut predictions from the large ensemble of 800 state members cover the true water cut from each well. They also largely cover the distributions from the small ensembles. However, except for Well 4, the truth is far from the center of the distributions of the predictions are heavily biased. Well 5 does not have water breakthrough in the true model, and only 2 out of the 800 predictions are correct.

Not all of the history matched realizations are of equivalent quality — with quality in this case measured by the magnitude of the data mismatches after data assimilation. The higher quality realizations have smaller sums of squared mismatch between predicted and observed data. We used the limited memory BFGS method to history match pressure data for 100 reservoir models, but only retained the 20 with the lowest value of the objective function after minimization. The median value of the objective function from the best 20 was 21,300. The median value of the final objective function from the ensemble Kalman filter method with 40 members in an ensemble was 6220. The median value from EnKF using a single ensemble of 800 was 4780. It is not surprising that better matches to the data are obtained from the larger ensemble, simply because there are more degrees of freedom. It is somewhat surprising that the minimization method that used the adjoint to compute the gradient achieved the poorest matches to data. This has not been our experience with other history matching problems, and we attribute it in this case to the necessity of introducing a transition region between facies so that the gradient is computed from an approximation to the actual objective function.

Conclusions

The ensemble Kalman filter method outperformed the gradient-based minimization method in both computation efficiency and applicability for the problem of estimation of facies boundaries in history matching. It took approximately 5500 equivalent simulation runs to obtain the 20 accepted RML model realizations, and only 800 simulation runs for the 20 independent EnKF model realizations. The quality of the matching from the EnKF model realizations are better than those from the gradient approach. The number of 40 state vectors in an ensemble is insufficient for uncertainty quantification, as obvious correlations have been developed among the state vectors during data assimilation.

Better quality matches were obtained from the larger ensemble, where 800 simulation runs were made, than the 20 independent ensembles with 40 state vectors in

each. We have not yet evaluated the distribution carefully enough to know how many could be considered to be independent realizations. Based on the experience in this problem, however, it appears that it might be more efficient to use one large ensemble to access uncertainty, than to use many small ensembles.

Nomenclature

$C =$ covariance matrix
$d =$ vector of data (units depend on data type)
$H =$ measurement operator
$m =$ vector of model parameters
$N =$ number of data or model parameters
$O =$ objective function
$P =$ pressure
$S =$ saturation
$Y =$ Gaussian field
$y =$ state vector
$\Psi =$ ensemble of state vectors
$\sigma =$ standard deviation

Subscripts

$D =$ data
$d =$ data
$e =$ ensemble
$k =$ iteration index, or permeability
$M =$ model
obs = observed
pr = prior model
sim = simulated
u = unconditional
wf = well flow
$\phi =$ porosity

Superscripts

$p =$ predicted
$T =$ transpose
$u =$ updated
$-1 =$ inverse

Acknowledgments

Support for Ning Liu was provided by the U.S. Department of Energy, under Award No. DE-FC26-00BC15309. However, any opinions, findings, conclusions or recommendations herein are those of the authors and do not necessarily reflect the views of the DOE. Licences for the ECLIPSE^{TM} Black Oil Reservoir Simulator were provided by Schlumberger. The assistance of Henry Neeman, Director of the OU Supercomputing Center for Education & Research, is gratefully acknowledged.

References

[1] Chen, W. H. et al.: "A new algorithm for automatic history matching," *Soc. Petrol. Eng. J.*, pages 593–608, 1974.

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

[3] Makhlouf, E. M. et al.: "A general history matching algorithm for three-phase, three-dimensional petroleum reservoirs," *SPE Advanced Technology Series*, 1(2):83–91, 1993.

[4] Li, R., Reynolds, A. C., and Oliver, D. S.: "Sensitivity coefficients for three-phase flow history matching," *Journal of Canadian Petroleum Technology*, 42(4):70–77, 2003.

[5] Nocedal, J.: "Updating quasi-Newton matrices with limited storage," *Math. Comp.*, 35:773–782, 1980.

[6] Fletcher, R.: *Practical Methods of Optimization*, John Wiley & Sons, New York, second edition, 1987.

[7] Deschamps, T. et al.: "The results of testing six different gradient optimisers on two history matching problems," In *Proceedings of the 6th European Conference on the Mathematics of Oil Recovery*, pages B–24, 1998.

[8] Zhang, F. and Reynolds, A. C.: "Optimization algorithms for automatic history matching of production data," *Proceedings of 8th European Conference on the Mathematics of Oil Recovery*, page to appear, 2002.

[9] Woodbury, A. D.: "Bayesian updating revisited," *Mathematical Geology*, 21:285–308, 1989.

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

[11] Oliver, D. S.: "Incorporation of transient pressure data into reservoir characterization," *In Situ*, 18(3):243–275, 1994.

[12] Zimmerman, D. A. et al.: "A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow," *WRR*, 34(6):1373–1413, 1998.

[13] Floris, F. J. T. et al.: "Methods for quantifying the uncertainty of production forecasts: A comparative study," *Petroleum Geoscience*, 7(SUPP):87–96, 2001.

[14] Liu, N. and Oliver, D. S.: "Evaluation of Monte Carlo methods for assessing uncertainty," *SPE Journal*, 8(2):188–195, 2003.

[15] Evensen, G.: "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, 99(C5):10143–10162, 1994.

[16] Anderson, J. L. and Anderson, S. L.: "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Monthly Weather Review*, 127(12):2741–2758, 1999.

[17] Nævdal, G., Mannseth, T., and Vefring, E. H.: "Near-well reservoir monitoring through ensemble Kalman filter: SPE 75235," In *Proceedings of SPE/DOE Improved Oil Recovery Symposium*, April 13–17 2002.

[18] Nævdal, G. et al.: "Reservoir monitoring and continuous model updating using ensemble Kalman filter," *SPE 84372*, 2003.

[19] Gu, Y. and Oliver, D. S.: "The ensemble Kalman filter for continuous updating of reservoir simulation models," *Computational Geosciences*, submitted in 2004.

[20] Houtekamer, P. L. and Mitchell, H. L.: "Data assimilation using an ensemble Kalman filter technique," *Monthly Weather Review*, 126(3):796–811, 1998.

[21] van Leeuwen, P. Jan: "Comment on 'Data assimilation using an ensemble Kalman filter technique'," *Monthly Weather Review*, 127(6):1374–1377, 1999.

[22] Galli, A. et al.: "The pros and cons of the truncated Gaussian method," In *Geostatistical Simulations*, pages 217–233. Kluwer Academic, Dordrecht, 1994.

[23] Lantuéjoul, C.: *Geostatistical Simulation: Models and Algorithms*. Springer, Berlin, 2002.

[24] Liu, N. and Oliver, D. S.: "Automatic history matching of geologic facies," SPE 84594, *SPE Journal*, 9(4):1–15, 2004.

[25] Liu, N. and Oliver, D. S.: "Conditional simulation of truncated random fields using gradient methods," In *Proceedings of the IAMG 2003 Portsmouth Meeting*, 2003.

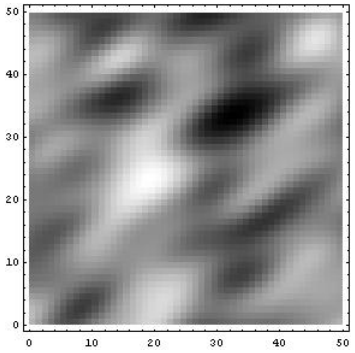
[26] Li, R., Reynolds, A. C., and Oliver, D. S.: "History matching of three-phase flow production data," *SPE Journal*, 8 (4):328–340, 2003.

Well	1	2	3	4	5
x	25	9	9	41	41
y	25	5	45	45	5
facies	3	1	1	3	2

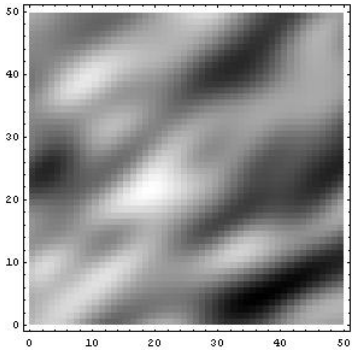
Table 1: Facies observations from each of the five wells.

index	Facies 1	Facies 2	Facies 3
Permeability (md)	174.0	80.0	372.0
Porosity	0.18	0.146	0.25

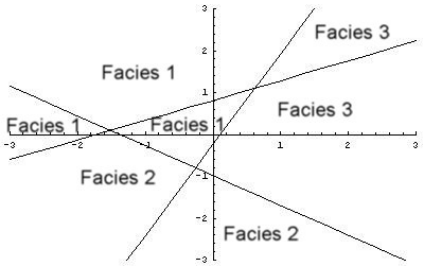
Table 2: Properties of each the litho-facies in the synthetic problem.



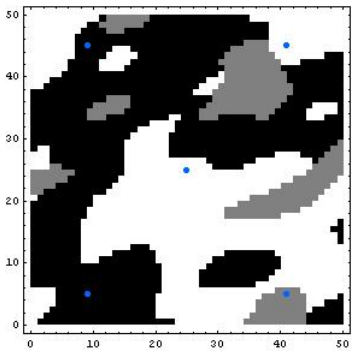
(a) Gaussian field (Y_1).



(b) Gaussian field (Y_2).



(c) Threshold map.



(d) Calculated facies field map.

Figure 1: Simulation of lithofacies distribution in the field by truncation of random Gaussian fields Y_1 and Y_2 using intersecting line thresholds.

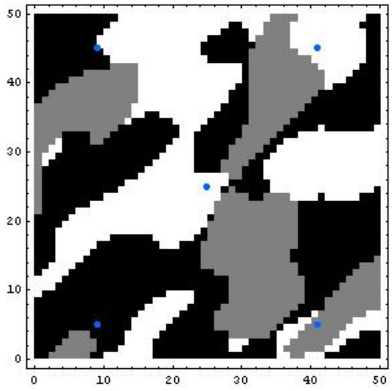


Figure 2: The true facies field of the synthetic model contains three facies types. The five dots are well locations.

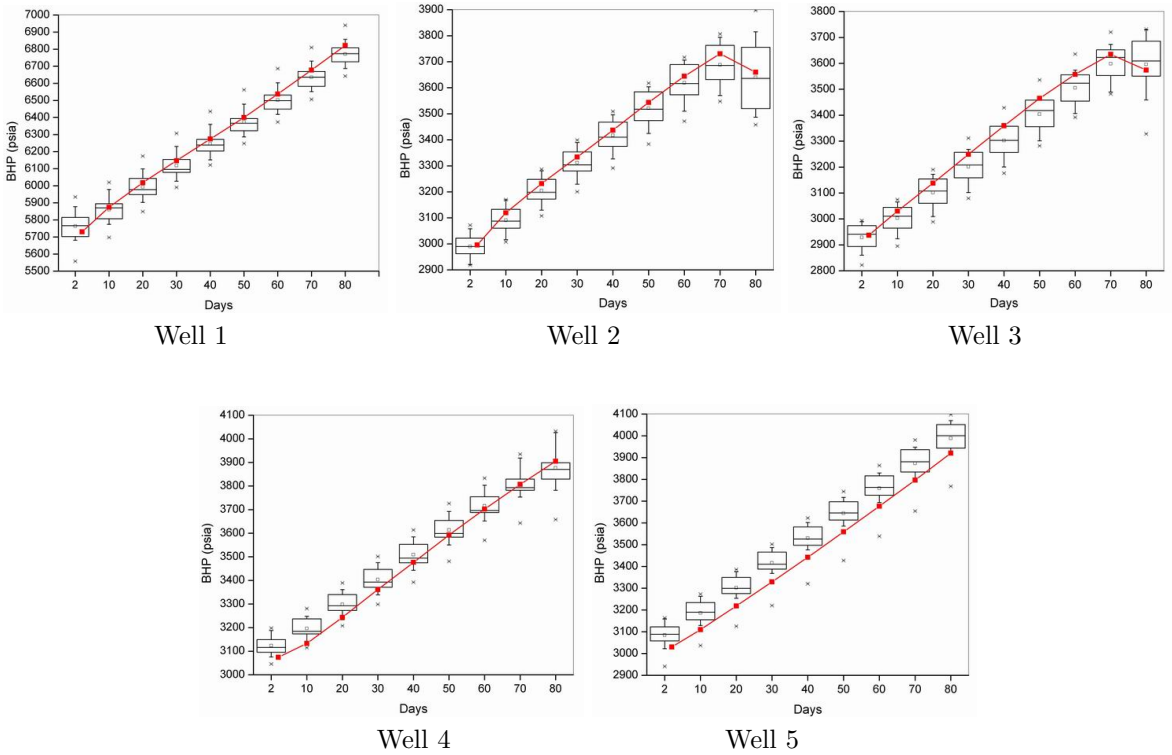


Figure 3: The box plots present the distributions of the simulated bottom hole pressure from 20 history matched models. The thick lines are the observed bottom pressure data from each well.

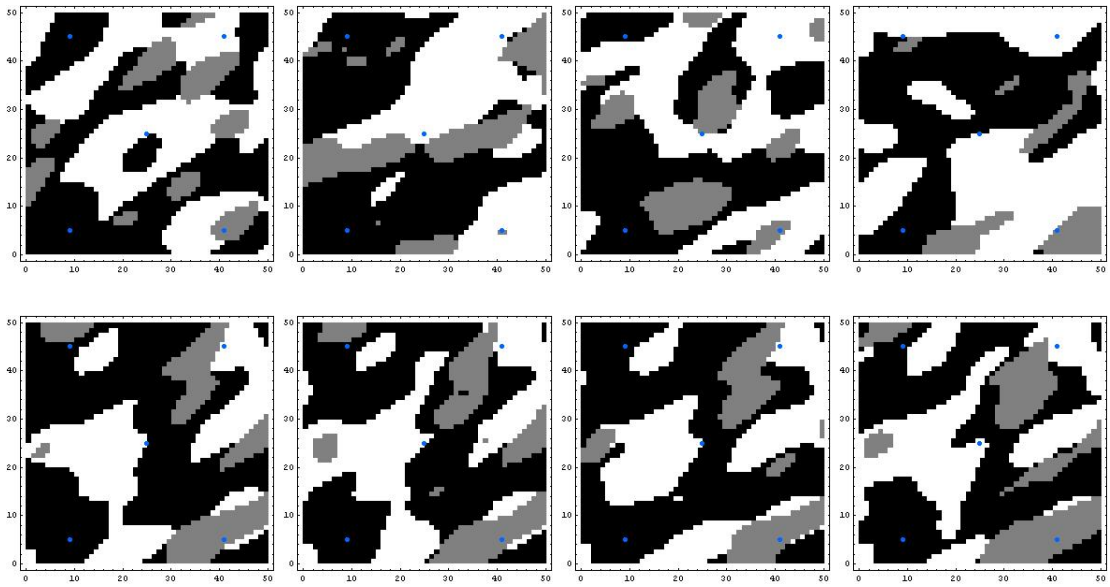


Figure 4: The first four initial facies map in an ensemble of 40 members (top row), and the corresponding final facies map after history matched to production data and hard data (bottom row). The dots in each facies map are well locations.

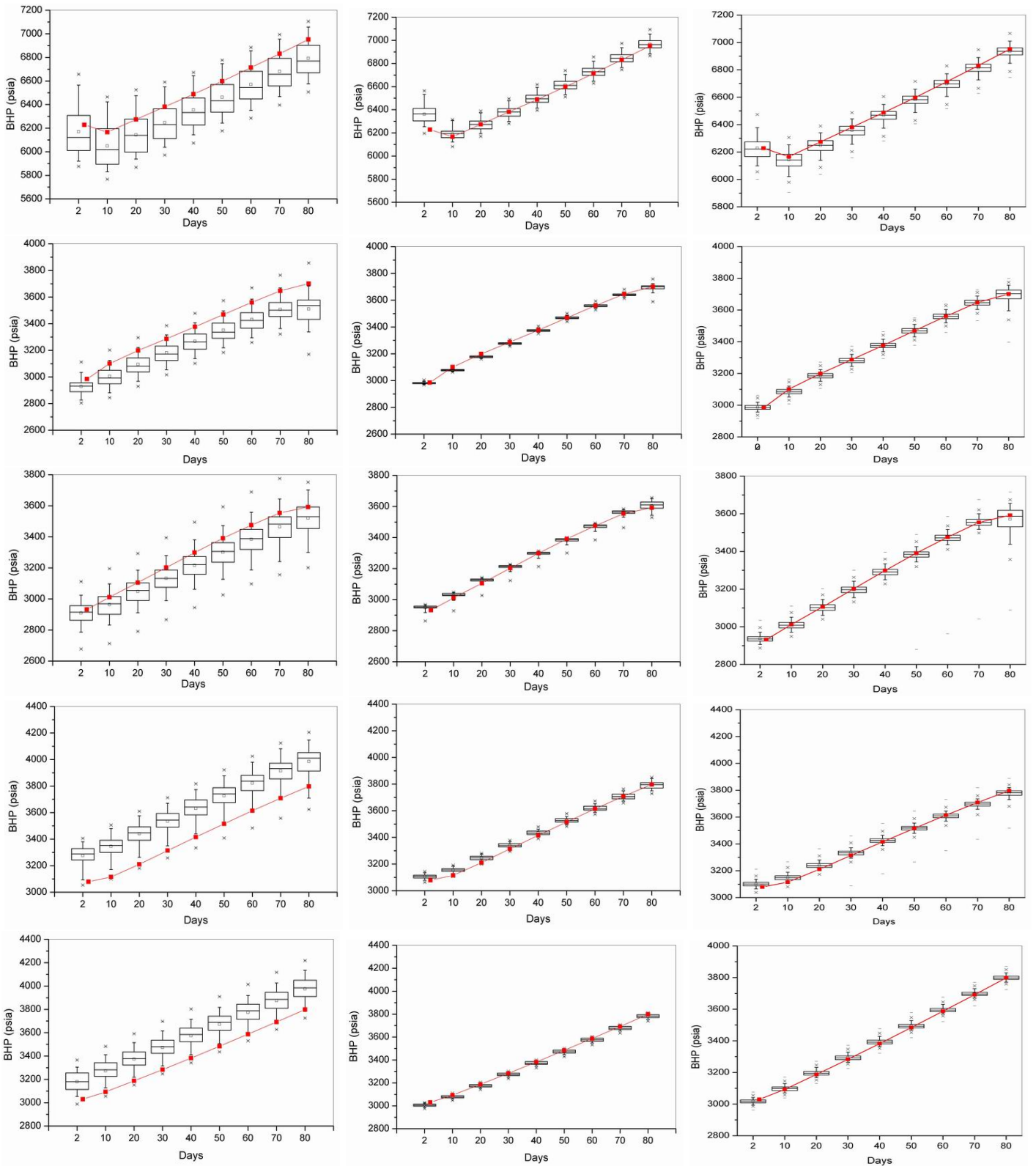


Figure 5: Distributions of the simulated bottom hole pressure from the initial ensemble of 40 (first row), and from the history matched ensemble (second row). The observed data from each well is plotted in a thick line for comparison. The box plots shown are for Wells 1 through 4.

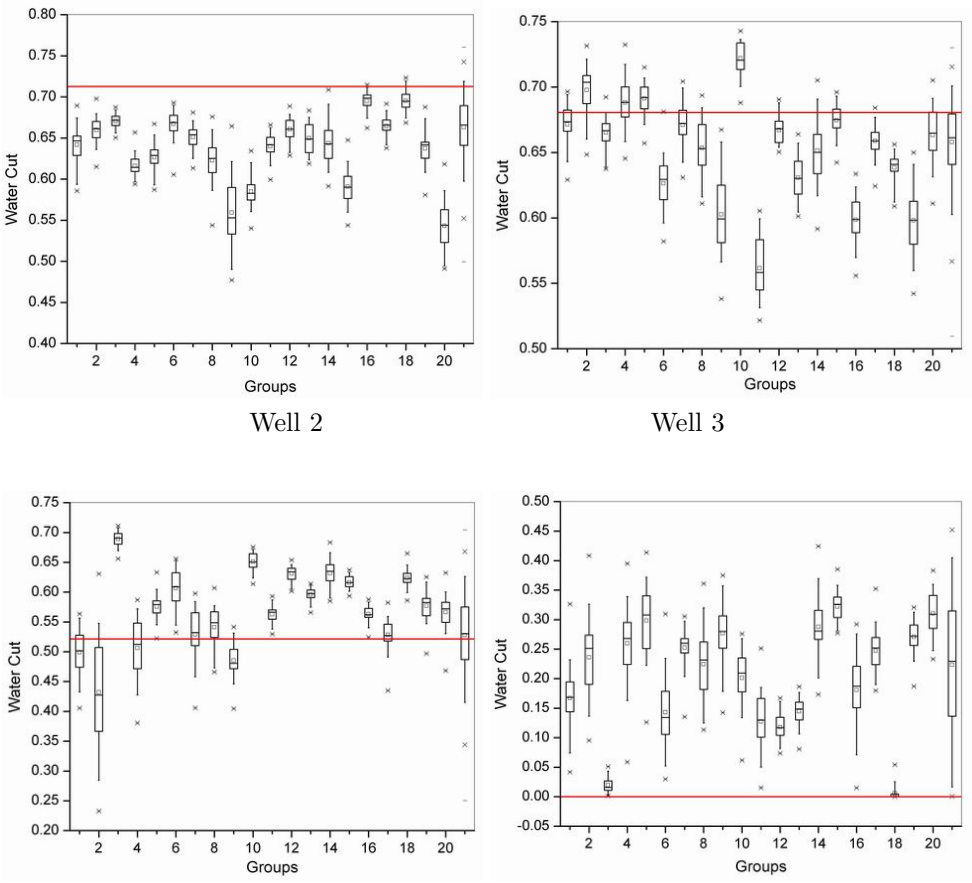


Figure 6: Distributions of water cut prediction on day 140 from 20 ensemble groups. The straight lines are water cut predicted from the true model.