History Matching using the Multi Point Approximation Approach

FRANDSEN, Poul Erik, GEUS - Geological Survey of Denmark and Greenland
REFFSTRUP, Jan, Dept. of Energy Engineering, Tech. Univ. of Denmark
ANDERSEN, Jan, GEUS - Geological Survey of Denmark and Greenland *)

Paper presented at the 5th European Conference on the Mathematics of Oil Recovery, Leoben, Austria, 3-6 Sept. 1996

ABSTRACT

In the present paper the use of the multipoint method for automated history matching purposes is demonstrated. In these types of problems, reservoir simulators are used to estimate unknown reservoir properties by matching simulated and observed production data. These types of calculations are usually characterized by a very long computation time because each function evaluation represents a full reservoir simulation. In the multi point method the optimisation process is divided into a sequence of approximate solutions. The approximate solutions are obtained by using a simplified and fast reservoir simulation model. Intermediate solutions to the optimisation problem is obtained using this simplified models. The method is implemented with a trust region technique for advancing the sequence of simplified solutions to the final solutions. The paper demonstrates the use of the technique on a simplified one-dimensional flooding problem where the flow is described analytically and on a 3D gas injection problem using a traditional black oil reservoir simulator. Further the combined use of the multi-point method and the Bayesian inversion technique is demonstrated. Bayesian approach allows for inclusion of apriori knowledge as an integrated part of the overall procedure.

INTRODUCTION

Reservoir simulators are traditionally used in an estimation procedure where the uncertain reservoir parameters are adjusted until a match between calculated and observed performance has been obtained. The matching procedure thus involves minimisation of an objective function expressing the difference between the calculated and observed

reservoir performance. Traditionally a non-linear least squares problem is formulated. The difference between characteristic performance measures such as pressure, production, and gas-oil-ratio is sampled over a set of points in time and the objective function is then the sum of squared differences. If the simulators output is differentiable with respect to the parameters to be estimated and if no numerical noise appears in the calculated response then the objective function is smooth and differentiable in this type of formulation.

The classical iterative methods for non-linear least squares have been used to solve this problem [1]. Such methods include the well-known Levenberg-Marquardt method [2,3] and later advanced derivatives by Powell [4] and Dennis, Gay and Welsh [5]. The classical methods operate in a real n-dimensional space of parameters and the solution is a real n-dimensional vector. Further, the methods are based on gradient information i.e. the objective function must be differentiable. In cases where the simulator offers derivatives with respect to the parameters, the application of general least squares solvers is a viable approach.

Derivatives are usually available when special purpose simulators are used where the simulator is integrated into the history matching environment. Only a few general purpose large scale reservoir simulators do, however, offer this option. A major problem in using large scale reservoir simulators together with traditional methods for history matching is the excessive CPU-time requirements needed to perform the estimation. Hence, if the simulator does not offer derivatives with respect to the parameters, the derivatives of the objective function must be approximated. If this is done the objective function needs to be calculated one plus the number of unknown parameters times when the

*) Author now with Kemp&Lauritzen A/S, Denmark.

gradient is needed. Since a function evaluation includes a simulation and since a single simulation can take hours, even on a fast work station, the overall CPU time can be excessive. In order to speed up the procedure the simulator may then be run with low accuracy. This, on the other hand, imposes numerical errors such that the objective function gets inflicted with a noise that stems from these errors. If the signal-to-noise ratio is large derivative approximations can still be utilized. In the final stage of the iteration, however, the combination of noise and derivative approximations will prevent the convergence of a gradient based method.

In recent formalisms including Bayesian inversion [6] stochastic simulation can be incorporated in which case the objective function gets inflicted with random noise. This makes the use of traditional least squares methods impossible. Instead methods that allow for discontinuities may be considered like the Simplex method by Nelder and Mead [7] or methods based on random generators like genetic algorithms [8]. Unfortunately all known methods that depend on function sampling alone are known to require an even larger number of evaluations of the objective function.

In this presentation we shall consider a new approach to parameter estimation which seems promising with reference to improved procedures for both CPU-excessive and noisy parameter estimation problems. The basic idea is to substitute a fast *simplified simulator* for the comprehensive simulator in the intermediate steps of the iteration. Hence the majority of the simulations should be performed with the fast simulator. If the simplified simulator response is also smooth and differentiable then at least in theory it is possible to smooth the noisy objective function by an approximate objective function which is derived from the response of the simplified simulator. In this presentation we shall consider the ability to keep the number of function evaluations and hence comprehensive simulations at a reasonable level. The effect of noise has not been studied systematically yet and we shall not consider this subject here.

Methods which use fast, simplified simulators to assist the comprehensive simulator have been proposed by Toropov [9] and Bandler et al. [10]. Toropov [9] introduces the multi point approximation principle whereas Bandler et al. [10] introduce the space mapping technique. In this presentation a trust region method which combines the multi point approximation with traditional trust region techniques is introduced. A discussion of trust region techniques can be found in Fletcher [11].

BACKGROUND

Multi point approximation methods are designed for a maximum utilisation of the comprehensive simulations performed during the parameter estimation. The simplified fast simulator, which is less accurate but much faster than the original and time-consuming comprehensive simulator, must of course simulate the same physical reality as the comprehensive one. It may be a combination of several sub-models each of which covers a subset of the overall model output domain. The difference between this response and the reservoir history defines an approximate objective function. The simplified simulator and the approximate objective function are equipped with a set of tuning coefficients. Instead of basing the calculation on the next candidate solution on the currently best set of parameters, which is common in traditional methods, basically all samples of the objective function that have been gained during the iteration are kept available. In each step the approximate objective function is fitted to the real objective function over a set of iteration points in the neighbourhood of the currently best known set of parameters. Hence, the calculation of the next candidate is based on *multiple* values of the objective function. This is obtained by solving a non-linear least squares problem using the tuning coefficients as unknowns. The fitting of the approximate objective function is denoted the tuning. The goal is a fast to evaluate approximate function that resemble the objective function in much larger regions of the parameter space than can be achieved with quadratic model functions used in the traditional methods. The next approximate solution is obtained as the minimum of the approximate objective function. Some of the best existing optimisation methods are needed for the fitting and the minimisation of the approximate objective function. Hence the multi point approach can be considered as an enhancement of classical optimisation.

The choice of tuning coefficients is denoted the tuning design and this can only be made by a *qualified user*. Hence the approach is based on the inclusion of expert knowledge expressed through the choice of the simplified simulator and how its response reacts on variations of the simulator input. The user must possess an understanding of how the simplified simulator can be manipulated to resemble the comprehensive simulator and then enable this manipulation through the tuning design. The inclusion of the user knowledge in the actual set up of the parameter estimation procedure makes it more difficult to use this new type of approach. It does, however, offer an opportunity to obtain the better approximations to the objective function mentioned above. When traditional methods are used, the only real possibility of user interference seems to be the choice of a good starting point for the

iteration and a good scaling of the problem at hand (which may not be trivial!).

OPTIMISATION PROBLEM

The objective function F depends on the formulation. Let θ denote the vector of unknown parameters to be estimated. Also, let $o_p(t)$ denote the observed reservoir history as a function of time t and let $\omega_c(t, \theta)$ denote the simulated reservoir performance obtained from the comprehensive simulator. The traditional least squares measure is defined as

$$F(\theta) = d(\theta) = \sum_{i=1}^m w_i (o_p(t_i) - \omega_c(t_i, \theta))^2 \quad \text{.....(1)}$$

where F is the objective function to be minimised. Without loss of generality we may disregard the weights w_i .

The Bayesian formalism has attracted increasing attention over the past years. Here a short summary of the formulation is given, aiming at the derivation of the objective function which is our main interest. The formalism is discussed in detail by Floris and Bos [6]. The goal is to maximise the objective function F which is derived from Bayes' rule

$$F(\theta) = p(\theta | y = o_p) = \frac{p(y = o_p | \theta) p(\theta)}{p(y = o_p)} \quad \text{.....(2)}$$

The term $p(y = o_p | \theta)$ is referred to as the likelihood and is a measure of the goodness of the fit between simulated and observed reservoir performance. It is calculated as

$$p(y = o_p | \theta) = \prod_{i=1}^m \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2} \left(\frac{o_p(t_i) - \omega_c(t_i, \theta)}{\sigma_i} \right)^2} \quad \text{.....(3)}$$

where σ_i is the standard deviation on the i 'th observation. The function $p(\theta)$ is referred to as the joint prior. It represents expert prior knowledge of the reservoir parameters to be estimated. The prior information is given as a probability density function for each of the parameters. Assuming that the inversion parameters are independent, the joint prior is obtained as

$$p(\theta) = \prod_{i=1}^n p_i(\theta_i) \quad \text{.....(4)}$$

where n is the number of parameters. The scaled product of the likelihood and the prior i.e. the objective function defined in (2) is referred to as the posterior.

The least squares objective (1) can also be modified to include the prior. If all parameters are Gaussian, $\theta_j \sim N(\mu_j, \sigma_j)$, then by taking the logarithm of the Bayesian objective in equation (2) and changing the sign, the following function to be minimised is obtained -

$$F(\theta) = c_1 \sum_{i=1}^m (o_p(t_i) - \omega_c(t_i, \theta))^2 + c_2 + \frac{1}{2} \sum_{j=1}^n \left(\frac{\theta_j - \mu_j}{\sigma_j} \right)^2 \quad \text{.....(5)}$$

where c_1 and c_2 are real constants. Hence, the prior appears as a penalty to the original objective function.

THE MULTI POINT ITERATION

Since the simulators are an integrated part of the whole approach, we have chosen to use an example to illustrate the different parts of the procedure. The Bayesian formulation is used in combination with the piston displacement model described in appendix A. The synthetic history $o_p(t)$ is generated with the piston model for a reservoir with 6 blocks of equal length and with permeabilities $(K_1, K_2, K_3, K_4, K_5, K_6) = (200, 100, 175, 175, 100, 200)$ (all in md). The pressure drop $\Delta P_w(t)$ between injection and production at 20 sampling times is used as history data. In the example let $\theta = K_2$ be the only unknown. The piston model with 6 blocks is used as the comprehensive simulator with response $\omega_c(t, \theta)$. The fast simulator with response $\omega_s(t, \theta)$ is also a piston model but only with 2 blocks which cover the original blocks $\{1, 2, 3\}$ and $\{4, 5, 6\}$ respectively. The permeabilities K_1 and K_2 of the 2 blocks are obtained as the arithmetic mean of the original blocks - $K_1 = (K_1 + K_2 + K_3)/3$ and $K_2 = (K_4 + K_5 + K_6)/3$. In order to be able to show some steps of the iteration, the arithmetic mean rather than the geometric mean is chosen. (If the geometric mean is used the approximate objective is nearly identical to the real objective function after only one single iteration). Finally, a Gaussian distribution with expectation μ and standard deviation σ_θ is chosen for the single parameter and hence for the prior -

$$p(\theta) = \frac{1}{\sigma_\theta \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\theta - \mu}{\sigma_\theta} \right)^2} \quad \text{.....(6)}$$

In figure 1 the sum of squared differences $d(\theta)$ that appears in the piston case is shown (I) together with adequately scaled versions of the prior function (II), the likelihood function (III) and the posterior that were used. Since the data were generated with $\theta = K_2 = 100$, $d(100) = 0$ whereas the likelihood attains its maximum at this point. The prior is obtained with $\mu = 120$ and $\sigma = 4$

i.e. the maximum is obtained at $\theta = 120$. The posterior attains its maximum at $\theta = K_2 = 109.52$ and the intention is to find this point.

Figure 1 Piston displacement least squares objective (I) and Bayesian objective (IV) obtained from scaled versions of likelihood (II) and prior (III).

We may now proceed to a detailed discussion of the multi point approach. The method is iterative. In the k 'th step an approximate fast to evaluate objective function, F_k , is substituted for the real objective function, F . If the simplified simulator, ω_s , is substituted for ω_c in the Bayesian objective (2) a crude posterior denoted F_S appears through (3) and (4). If the simulators react qualitatively in the same way to changes in θ , then F and F_S will be of the same nature though different in quantity. Since F_S is fixed, it cannot be used directly as approximate objective. Rather, the approximate objective must appear like F in the neighbourhood of the currently best point, θ_c , that is the problem is to find a function that can be adjusted from step to step. Such a function, denoted F_T , can be derived from F_S via the tuning design where the tuning coefficients are introduced. The tuning is case dependent. As an illustration, consider again the piston example. The first tuning coefficient to be used here, α_3 , scales θ . For different values of α_3 the function $F_S(\alpha_3\theta)$ is shifted along the θ -direction and widened as shown in figure 2.

If a least squares objective is considered, a similar tuning can be made by substituting ω_s for ω_c in equation (1) and using $\alpha_3\theta$ as input to ω_s . Also scaling and offset abilities will often be required. For the least squares case these abilities are included in the following tuneable function

$$F_T(\theta, \alpha) = \alpha_1 + \alpha_2 \sum_{i=1}^m (o_p(t_i) - \omega_s(t_i, \alpha_3\theta))^2 \quad \dots\dots\dots(7)$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3)$.

Figure 2 The crude posterior $F_S(\alpha_3\theta)$ for $\alpha_3 = 0.7, 1, 1.5$ and 2

In the Bayesian formulation the scale tuning is obtained by multiplying the exponent in the likelihood (3) with α_2 , and the offset tuning is obtained by multiplying the prior with the term $\exp(-\frac{1}{2}\alpha_1)$. In the one-dimensional piston case this yields

$$F_T(\theta; \alpha) = e^{-\frac{1}{2}\alpha_1} \prod_{i=1}^m \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\alpha_2 \left(\frac{o_p(t_i) - \omega_s(t_i, \alpha_3\theta)}{\sigma_i} \right)^2} \cdot \frac{1}{\sigma_\theta \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\alpha_3\theta - \mu}{\sigma_\theta} \right)^2} \quad \dots\dots\dots(8)$$

For any new set of α 's F_T takes a new form, but it still shares basic properties with F .

Now assume that a tuning design has been chosen. Then the calculations of a single step go as follows: Assume that F during the iteration process has been evaluated in a sequence of points θ_l , $l=1, \dots, k-1$ i.e. these are the points where information about the objective has been obtained. The best point is θ_c , $F(\theta_c) \geq F(\theta_l)$, $l=1, \dots, k-1$. Since F_k should approximate F , a vector α_k is determined such that $F_k(\theta) = F_T(\theta, \alpha_k)$ is a best fit of F in some region near θ_c . The question is then which of the available points should be used when this fit is made. Here the trust region, T_k , is introduced which is a sphere with centre θ_c and radius Δ_k i.e.

$$T_k = \{\theta / \|\theta_c - \theta\|_2 \leq \Delta_k\} \dots \dots \dots (9)$$

An extended region is also used -

$$A_k = \{\theta / \|\theta_c - \theta\|_2 \leq \gamma \Delta_k, \gamma \geq 1\} \dots \dots \dots (10)$$

Points far away from θ_c , i.e. outside A_k , are considered outdated and are not used. Points inside T_k are close to θ_c and they are included with unit weight. Points in $A_k \setminus T_k$ are also used but with a weight which decreases from unity with increasing distance to θ_c -

$$w_k = (\Delta_k / \|\theta_c - \theta_k\|_2)^2, \theta_k \in A_k \setminus T_k \dots \dots \dots (11)$$

Hence α_k is found as the solution to the problem

$$\text{Minimise } \sum_{\theta_i \in A_k} w_i (F(\theta_i) - F_T(\theta_i, \alpha))^2 \dots \dots \dots (12)$$

This problem can be solved with any good non-linear least squares solver from a subroutine library. We have used routine NL2SOL by Dennis, Gay and Welsh [5]. Note, that only new fast to calculate values of F_T are needed whereas the expensive to evaluate values of $F(\theta_i)$ are known from the former iterations.

Figure 3 The objective function F (posterior - Bayesian formulation) and the approximate objective F_1 (--) obtained by fitting the function F_T to the values of F at 125, 120 and 115.

As an illustration of the tuning process consider the example above and assume that F has been evaluated at four points - {140, 125, 120, 115} where F attains the values 0.139, 0.470, 0.605 and 0.712 respectively (3 decimal places). Since $F(115)$ is largest, $\theta_c = 115$. Also,

assume that the current radius $\Delta_k = 10$ and $\gamma = 2$. In this simple case the trust region is an interval, [105;125] and A_k is the interval [95;135]. This means that data at the points {125, 120, 115} will be used whereas $\theta = 140$ is disregarded. Minimising the sum in equation (12) over these three points with equal weights yields the coefficients $(\alpha_1, \alpha_2, \alpha_3) = (1.7164, -0.2425, 1.1593)$ and the corresponding $F_{k=1}$ shown together with F in figure 3. Note that the two functions are practically identical in the interval [114;135].

The approximate objective has now been tuned to appear like the real objective in the neighbourhood of θ_c and may then take its place for the determination of the next point θ_k . This is found as the point that optimises the tuned approximate objective $F_k(\theta)$. Since the tuning is only based on data in A_k and since there is no guarantee that F_k is bounded above, the step must be limited. Also here the trust region can be used i.e. F_k is optimised inside this region. Hence θ_k is found as the solution to the problem

$$\begin{aligned} &\text{Maximise } F_k(\theta) \\ &\text{subject to } \|\theta_c - \theta\|_2 \leq \Delta_k \dots \dots \dots (13) \end{aligned}$$

This is a standard constrained optimisation problem which can be solved with several well known methods. The Watch-dog method by Chamberlain et al. [12] has been used here. This method requires derivatives of F_k . Forward differences are sufficient for the present application. Once again note that the function in equation (13) only depends on the simplified simulator. In the piston displacement example the trust region is simply an interval. The maximum of F_1 in figure 3 is attained in the point $\theta = 107.37$ and here F is evaluated, $F(107.37) = 0.7522$.

The trust region T_k is controlled as follows: The initial radius, Δ_0 , is chosen by the user. In subsequent iterations Δ_k is updated automatically according to the agreement between changes of F and the current F_k where the number

$$\rho_k = [F(\theta_k) - F(\theta_c)] / [F_k(\theta_k) - F_k(\theta_c)] \dots \dots \dots (14)$$

and the length of the step $\delta_c = \|\theta_c - \theta_k\|_2$ is used. The step is considered a failure if $\rho_k < 0.1$ or if $\delta_c < \Delta_k$. The usual reason for a poor step is that Δ_k is too large. In this case the radius is reduced for the next iteration, $\Delta_{k+1} = \Delta_k/2$. The step is considered a success if $\rho_k > 0.75$ and if $\delta_c = \Delta_k$, i.e. if the step goes to the boundary of the trust region. In this case the calculations of the step are repeated with a trial radius $2\Delta_k$ and with the

new information $F(\theta_k)$ included. If this step is a success, the procedure is repeated with radius $3\Delta_k$ and so on until one of the conditions are not satisfied. If the last trial step is uphill, the new radius is the last trial radius. Otherwise the former trial radius is used. Note, that the trust region centre remains unchanged during these trial steps. Finally, if $0.1 < \rho_k < 0.75$ then $\Delta_{k+1} = \Delta_k$. The new trust region is always centred in the currently best point i.e. the point with largest F value. Hence θ_c either remains unchanged or is updated to θ_k or one of the trial points.

Prior to the updating of the trust region, the next iteration can be initiated and one or several new points and corresponding values of F are included, depending on how the new radius was found. The sequence of approximate objective functions obtained in the piston displacement example is shown in figure 4. After 3 iterations F_k is identical to F in the neighbourhood of the maximum of the posterior.

What remains to be discussed is what is done if the number of points in A_k is less than the number of α 's denoted L such that problem (12) does not possess a unique solution. In this case A_k is increased temporarily by an increase of γ until enough points are available. The fit may in this case be based on very distant points

until L points are in the region. Then the next iteration is initiated without changing the trust region.

Initially, a set of L points must be chosen inside the user-specified initial trust region in order that the first set of α 's can be determined and the proper iteration started. Also note that in the final phase of the iteration the radius is rapidly reduced due to the fact that the step never reaches the boundary. The iteration is terminated when the radius is below some small ϵ . Also, if F_k has its minimum at θ_c , i.e. if the approximate objective does not decrease at all the iteration must be terminated.

SYNTHETIC TEST CASE

The concept of the multi point approach was illustrated by a simple one-dimensional test example in the previous section. In the present section the objective is first of all to illustrate the approach on a larger problem, but the test problems have also been designed in such a way that both the flexibility of the multi point method and the effect of using the Bayesian inversion technique is illustrated. As mentioned previously it is essential in this type of calculations, that the number of function evaluations is kept as low as possible. In the example-calculation in the present section the iterations were stopped when the relative changes in the variables were less than 15% from one iteration step to the next. This convergence criteria was chosen to test the ability of method to approach the vicinity of the solution, since further development of the method is required to obtain a better final convergence.

The data used in the example are modified data from the SPE first comparative test study by Odeh [13]. The reservoir in this test case is a box shaped three layered reservoir with sides $10,000 \times 10,000 \text{ ft}^2$ and height 100 ft. The problem considered is gas injection in a layered oil reservoir, represented as a quarter of a five-well pattern. Gas is injected in the top layer in the corner and oil is produced from the bottom layer at a position diagonally from the injection well. In the original test example the reservoir was divided into $10 \times 10 \times 3$ simulation blocks. In the present case the same simulator was used as both the fast model and the comprehensive model. The fast model was set up as a coarse grid model with $4 \times 4 \times 3$ blocks whereas the full grid was used for the comprehensive model. All simulations were carried out with the commercial simulator Eclipse. Modified reservoir data are shown in table 1.

Synthetic production data used in the matching procedure were generated with the fine grid model.

Two test cases were run to illustrate the use of the multi point method. In the test cases two of the three permeabilities are considered unknown.

Figure 4 Piston displacement example. Sequence of approximate objectives F_k , $k = 1, 2$ and 3 near solution $\theta^* = 109.52$.

relative to θ_c , and F_k could be a poor approximation of F locally. Hence frequently new information is called for. The inclusion of extra points is made if a step is poor and if the number of points in the trust region is less than L . Extra points are chosen inside T_k along the co-ordinate directions and at a distance $\Delta_k/2$ from θ_c

Layer no.	Height (ft)	Permeability (md)
1	20	200
2	30	50
3	50	500

Well	Cell	Rate	BHP psia
Gas-inject.	(1,1,1)	65000MSCF/day	≤8500
Oil-product.	(10,10,3)	30000 STB/day	≥4100

Table 1 Reservoir data for SPE test case

First the technique was used to solve a traditional optimisation problem.

Secondly the integrated use of the multi point method and the Bayesian inversion technique is demonstrated. In this second test case it is illustrated how geological information is integrated in the parameter estimation problem by the Bayesian formulation.

In the first test case the parameters selected for inversion were the permeability of the top and bottom layers. Simulation of production from the reservoir indicated that the flow in the reservoir was primarily governed by the permeability of these layers. And hence these two parameters formed ideal parameters for a parameter estimation problem.

The objective function corresponding to this parameter choice is illustrated in figure 5. (figures at p. 10) The function was generated using the comprehensive simulator, i.e. Eclipse in a fine grid mode and defined in the usual least square sense (see equation (1)). Hence in this case the multi point method minimises the objective. The figure gives the objective function as a function of the two reservoir parameters.

Figure 6 illustrates the corresponding approximate objective function generated using the fast model. The approximate objective function was generated with the coarse grid model. The approximate objective function generated by the fast simulator was made adjustable using in all five α -coefficients, described in previous sections. α_1 and α_2 were used to scale the function along the two axis, whereas α_3 and α_4 have been used to scale absolute permeability of the top and bottom layers. α_5 scales the relative permeability of gas k_{rg} .

The initial guess of the permeability values for the top layer, K_1 and the permeability in the bottom layer, K_3 were $(K_1, K_3) = (300, 300)$. 5 initial evaluations (no. of 's) of F was needed before the real iteration could be initiated. The iteration path is illustrated in table 2. The solution found was $(K_1, K_3) = (204, 479)$ with true values of $(K_1, K_3) = (200, 500)$. In all 5 comprehensive model calls were used to initialise the model i.e. to generate α -values and 4 calls of the comprehensive model were used in the actual iteration. During this iteration process

the approximate objective function was adjusted after each comprehensive simulation. A comparison between the final simplified approximate objective function and comprehensive function is illustrated in figure 5 and 6. It became apparent, that numerical solutions of complex physical problems generate some numerical noise. The solution found in the present case was the most accurate solution, that could be obtained with the present noise level.

Comprehensive simulation no.	K_1	K_3	F
init 1	300	300	8.0917
init 2	350	300	10.567
init 3	300	350	4.9051
init 4	250	300	6.6504
init 5	300	250	13.327
1	250.85	387.08	2.0954
2	217.37	425.19	0.6602
3	192.31	467.96	0.1418
4	204.38	478.79	0.0044

Table 2

In the second test case the parameters selected for inversion were K_2 and K_3 i.e. the middle and bottom layer permeabilities.

Experience had shown, that the flow in the reservoir was very little influenced by the permeability value of the middle layer. Hence selecting these two parameters would constitute an ill-posed problem.

The problem was first solved again using the traditional optimisation.

Figure 7 illustrates the objective function for this test case generated with the comprehensive model and defined in the same manner as in above. It can be seen that the error i.e. the objective function is little sensitive to the choice of the value of K_2 . A solution was obtained after 2 iterations with values of permeabilities of $(K_2, K_3) = (150, 441)$. As can be seen from the results, the permeability of the middle layer is very poorly estimated reflecting the ill-posed nature of the problem. Again numerical noise in the solution procedure stopped the iteration, and it was not possible to obtain the original 50 md of the middle layer by inversion.

Additional information was needed to make this problem a well-posed problem. This was done by including geological data in the form of a priori distribution functions and combine a priori and the likelihood function as discussed in the previous sections.

In this way the geological information is included in the parameter estimation problem. The resulting probability density function is the posterior function the maximum of which represents the most likely parameter set.

The posterior function as in equation (5) (least squares with penalty) for this test case is illustrated in figure 8. The function has for illustrative reasons been generated using the comprehensive simulator i.e. the reservoir simulation model in a fine grid mode.

The approximate objective function for this case is illustrated in figure 9. Again five parameters were chosen to make the approximate objective function flexible enough. Two parameters were used to scale the posterior function along the parameter axes and two parameters were chosen to adjust the absolute permeability and finally one was chosen to scale the relative permeability for gas. In the a priori density distribution function chosen for the permeability for the middle layer, a mean permeability value of 100 md was chosen with a standard deviation of 80 md to reflect the uncertainty in the interpretation of geological data. As illustrated in the figure, the posterior function, which is the objective function in Bayesian terminology, is changed to a function with a well defined extremum.

The result of the optimisation process is illustrated in table 3. In all 3 iterations i.e. calls of the comprehensive simulator were used for convergence.

Comprehensive simulation no.	K_2	K_3	F
init 1	100	300	5.0574
init 2	150	300	4.1685
init 3	100	350	2.2841
init 4	50	300	6.7454
init 5	100	250	9.7604
1	124.57	396.93	0.7362
2	118.48	448.86	0.1619

Table 3

The example illustrates, that the classical optimisation problem can be integrated with geological information in the Bayesian formulation and by integrating information from flow simulation with geological information change an ill-posed optimisation problem to an well-posed problem. This also means, however, that direct geological information now plays a more important part. The final result of the simulation was permeabilities $(K_2, K_3) = (118, 448)$ md. The correct maximum of F was $(K_2, K_3) = (107, 474)$. The value of parameters with little influence on the flow conditions in the reservoir is now determined on the basis of the a priori geological information.

A sensitivity analysis of the optimisation problem should initially be carried out to identify parameters with no or little sensitivity. A priori distribution function could then be given for these problems. Such parameters should not in general be excluded from the optimisation problem unless it is certain that the parameter will not gain sensitivity at later stages in the production cause.

CONCLUSION

The multi point method with a new trust region technique has been tested as part of an automated history matching procedure. The method works without information of function derivatives. The method was tested on two reservoir engineering problems, a simplified one-dimensional water flooding problem using analytical flow models and a three-dimensional gas injection problem using a standard black oil reservoir simulator. It has been demonstrated that the method can reach the vicinity of the solution using only in the order of 7 - 9 function evaluations, representing full reservoir simulations.

The use of the Bayesian formalism as part of the procedure was further demonstrated. The introduction of a priori knowledge with the Bayesian formalism allowed for a direct inclusion of geological information as part of the optimisation procedure. This technique can effectively be used to condition ill-posed optimisation problems due to reservoir parameters with no or little effect on the reservoir flow.

Further development of the approach is required in order to improve the final convergence and to avoid the up-front comprehensive simulations which are needed to start the iteration. Finally, the complex tuning design must be studied in order to obtain rules for a systematic development and application of simplified simulators.

NOMENCLATURE

F	Objective function to me optimised
θ	Vector of parameters to be estimated
n	Number of parameters to be estimated
α	Vector of tuning coefficients
l	Number of tuning coefficients
$o_p(t)$	Observed reservoir history as function of time
t	time
$\omega_c(t, \theta)$	Simulated reservoir performance obtained from the comprehensive simulator
$\omega_s(t, \theta)$	Simulated reservoir performance obtained from the fast simplified simulator
d	Mismatch function
$p(y = o_p \theta)$	Likelihood function
$p(\theta y = o_p)$	Posterior function
$p(\theta)$	Joint prior

$p_i(\theta_i)$	Probability density function
μ	Mean value (expectation)
σ	Standard deviation
K	Absolute permeability
ΔP_w	Pressure drop
F_k	Approximate objective function in k'th iteration
F_S	Crude objective function approximation
$F_T(\theta, \alpha)$	Tuneable function
θ_c	Trust region centre, best point (set of estimation parameters) found.
θ_k	Point found in k'th iteration.
α_k	Tuning coefficients found in k'th iteration.
T_k	Trust region.
Δ_k	Trust region radius used in k'th iteration.
A_k	Extended region which includes T_k .
γ	Extended region radius is $\gamma\Delta_k$.
ρ_k	Ratio between decrease of F and decrease of F_k when moving from θ_c to θ_k
δ_k	Length of step from θ_c to θ_k .

ACKNOWLEDGEMENTS

We acknowledge the European Commission for their financial contribution to this research. We also acknowledge the fruitful contributions by Professor K.Madsen, J.V.Nørgaard and the partners in the JOULE 2 - REP5 project.

REFERENCES

- [1] K.H.Coats and J.Dempsey, "A new technique for Determining Reservoir Description from Field Performance Data", SPE Journal, 9, 1970, pp. 66-74.
- [2] K.Levenberg, "A method for the solution of certain non-linear problems in least squares", Quart. Appl. Math., 2, 1944, pp 164-168.
- [3] D.W.Marquardt, "An algorithm for least squares estimation of nonlinear parameters", J.Soc.Indust. Appl.Math., 11, 1963, pp 431-441.
- [4] M.J.D.Powell, "A hybrid method for non-linear equations", in Numerical Methods for non-linear algebraic Equations, ed. P.Rabinowitz, Gordon & Breach, 1970, p. 87.
- [5] J.E.Dennis, D.M.Gay, and R.E.Welsh, "ALGORITHM 573. NL2SOL - An adaptive nonlinear least squares algorithm", ACM.Trans.Math. Software, 7, 1981, pp. 364-383.
- [6] F.Floris and C.F.M.Bos, "Flow constrained inversion of geostatistical parameters", Proc. 4th European Conf. on the Math. of Oil Recovery, Røros, Norway, 1994.
- [7] J.A.Nelder and R.Mead, "A simplex method for Function Minimisation", Computer J., 7, 1965, p. 308.
- [8] D.E.Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, MA, USA, 1989
- [9] V.V.Toropov, "Multi Point Approximation Methods in Optimization Problems with Expensive Function values", Proc. 4th Int. Symp. on Systems Analysis and Simulation, Berlin, Germany, 1992.
- [10] J.W.Bandler, R.M.Biernacki, S.H.Chen, R.H.Hemmers and K.Madsen, "Electromagnetic Optimization Exploiting Aggressive Space Mapping", IEEE Trans. Microwave Theory and Techniques, Vol. XX, 1995.
- [11] R.Fletcher, "Practical Methods of Optimization", John Wiley, 1987.
- [12] R.M.Chamberlain, C.Lemarechal, H.C.Pedersen and M.J.D.Powell, "The Watchdog technique for forcing Convergence in algorithms for Constrained Optimization", Mathematical Programming Study, No. 16, 1982, pp. 1-17.
- [13] A.S.Odeh, "Comparison of Solutions to a Three-Dimensional Black-Oil Reservoir Simulation Problem", J. Petroleum Tech., 5, 1981, pp 13-25, 1981.
- [14] F.Floris, "GABI User's Guide", TNO, The Netherlands, September, 1994.

APPENDIX A: PISTON DISPLACEMENT MODEL

In this simple example thin oil is displaced by water in a one-dimensional reservoir. Water is injected in one end and oil is produced from the opposite end of the reservoir as shown in figure A1. To keep the flow equations simple, piston-like displacement is assumed. Figures A2 and A3 show the reservoir models used as input to the comprehensive and simplified simulator respectively. The intention is to estimate the permeability K_2 . All other reservoir parameters are considered to be known and constant. The reservoir simulator 'PISTON' described in [14] was used in the example. The simulator calculates the pressure at the producing well relative to a fixed injection well pressure in a simple reservoir with blocks of different permeability. The simulator input is: Injection rate, reservoir length, height and width, porosity, viscosity of oil and water, relative permeability factors for oil and water, fixed injection pressure, the list of absolute permeabilities of the blocks and a list of times at which the producing well pressure has to be reported.

Figure A1 Simplified reservoir model

Figure A2 Comprehensive reservoir model

Figure A3 Simplified reservoir model

Figure 5 K1-K3 case. Objective function

Figure 6 K1-K3 case: 1'st approximate objective

Figure 7 K2-K3: Least squares objective function

Figure 8 K2-K3 case. Posterior objective function (Bayesian in least squares form with penalty)

Figure 9 K2-K3 case. 1'st approximate objective