Abstract

We consider estimation of the absolute permeability, and enforce regularization by seeking to select the parameter space such that over-parameterization is avoided. We perform the parameterization selection during the history match, by solving a sequence of parameter estimation problems with increasing resolution for the permeability. The number and spatial location of new degrees of freedom to be introduced in the permeability representation at a stage in the estimation sequence are selected according to predetermined criteria. Alternative criteria are discussed and tested.

Introduction

Predictive simulations of reservoir behavior require that reservoir property values, such as permeability and porosity, can be specified in every simulator grid block. Obtaining a reliable reservoir description at the grid-block scale is a challenge, even if all available data sources are utilized.

The data types available for permeability identification represents information about permeability variation on different length scales, from inter-well scales (production data) through a few grid blocks (prior model) to sub-grid block scales (hard data at wells).

Coarse-scale permeability trends, like channels and barriers, have great impact on reservoir fluid flow. If a channel/barrier has not been revealed by seismic surveys, and if it is not penetrated by any wells, it will most likely remain unknown to the reservoir engineer when simulating fluid flow. The spatial extent, and the impact on fluid flow, of such structures imply, however, that they should be identifiable from production data.

Figure 1 illustrates what is – and what is not – the aim of the paper. The left plot shows a permeability field with only small-scale variation. Production data from a small number of wells can not be expected to give sufficient information to recover this permeability field, so we do not aim at this. The plot in the middle shows a permeability field with a single coarse-scale feature. It is our aim to recover such structures from production data without any prior knowledge of their existence. The right plot shows a permeability field where small-scale variation has been put on top of a coarse-scale trend. For such cases we aim to recover the coarse-scale trend.

We perform the history match as a Multiscale Estimation (ME). That is, we solve a sequence of parameter estimation problems with increasing resolution for the permeability until a sufficiently good match is obtained.

In some previous works on ME (see, e.g.,{)} the parameterization resolution at a stage in the estimation sequence is the same all over the reservoir. This is generally not optimal with respect to identification of unknown permeability trends. A uniform parameterization will generally be too coarse in some reservoir regions and too fine in other regions. With Ordinary Multiscale Estimation (OME), a new refinement of the parameterization in the entire reservoir is then needed - even though the current parameterization is already too fine in some regions. Hence, while OME reduce the risk of over-parameterization, it does not eliminate it.
Inversion Methodology

General Methodology

Permeability estimation requires the permeability to be parameterized. In this paper this is obtained by setting \( k_N(x) = c_N^T \psi_N(x) \), where \( c_N \in \mathbb{R}^N \) is the parameter vector and \( \psi_N \) is an \( N \)-vector of real-valued piecewise constant basis functions spanning the space within which \( k \) is to be estimated. Let \( d \in \mathbb{R}^M \) denote the measured time series of the production data, and let \( m(c) \in \mathbb{R}^M \) denote the corresponding production data calculated using the model equations.

The estimation problem is solved by finding the parameter vector \( c_N^M \in \mathbb{R}^N \) which minimizes the data misfit (output least squares)

\[
J(c_N) = (d - m(c_N))^T D^{-1} (d - m(c_N)),
\]

subject to \( k > 0 \). Here, \( D \) is the covariance of the measurement errors. In order to solve this minimization problem the Levenberg-Marquardt algorithm, as implemented in \( ^9 \), is used.

Regularization

The issue of how to select the number of parameters, \( N \), and the set of basis functions, \( \psi_N \), for the parameter estimation problem remains. An improper selection could lead to either under-parameterization (too lean parameterization) or over-parameterization (too rich parameterization). With under-parameterization the match of the data will be poor, while over-parameterization can lead to huge parameter uncertainties, parameter values clearly out of range, and an unnecessarily high computational work load.

One regularization strategy is then to compensate for a deliberate over-parameterization by penalizing deviations from \( a \ priori \) knowledge about the permeability, like in Bayesian estimation (see, e.g., \(^4\)) or in Tikhonov regularization (see, e.g., \(^7\)). Another strategy is to seek to select the dimension of the parameter space such that under- or over-parameterization is avoided. Multiscale estimation, and hence, the approach in this paper, belongs to the latter class of regularization strategies.

For permeability estimation, restricted resolution of the parameter space has been enforced, e.g., by
zonation, i.e., a division of the reservoir into clusters of grid blocks with a constant permeability value within each cluster. Hence, the resolution – i.e., the regularization – is selected a priori. There are virtually no guidelines for doing this properly. The pilot point method (see, e.g.,3) represents an alternative approach, but also here the parameterization is selected a priori.

**Multiscale Estimation**

The purpose of multiscale estimation is to avoid guessing at a fixed resolution and subsequently estimate all the corresponding parameters simultaneously. Multiscale estimation solves a sequence of parameter estimation problems with initially low, but increasing resolution.

Several sequential estimations are performed, each (from the second one) starting from the permeability estimate from the previous estimation, but with a more detailed representation of the permeability. The sequence of estimations is terminated upon reconciliation of the production data.

**Ordinary multiscale estimation**

With ordinary multiscale estimation (OME) the resolution is gradually increased by the same amount in the entire grid. Each parameter sub-region (i.e., spatial region associated with a parameter) will be divided into $2^l$ (where $l$ is the spatial dimension of the model reservoir) new sub-regions at the next stage in the estimation sequence. A risk of over-parameterization is, thus, still present due to the rapid increase in the number of new parameters.

**Adaptive multiscale estimation**

With adaptive multiscale estimation5 (AME) the resolution is gradually increased only in regions of the grid where it is expected to be productive. The risk of over-parameterization is therefore reduced. Figure 2 illustrates the difference between OME and AME.

To perform AME, a way to select parameterization refinement is needed. A description of such a method follows next. More details on the description of this method as well as derivation of mathematical expressions is given in5. An alternative parameterization selection methodology is presented

---

Figure 2: Illustration of excessive parameterization with OME. From left to right: True permeability, OM 4-parameter representation, OM 16-parameter representation, Sufficiently detailed AM representation.

in$^2$.

**Selection of parameterization.** Introduction of new parameters will facilitate reduction of the objective function value, but also increase computational cost and, in general, increase parameter uncertainty. Hence, the parameterization must be selected in a proper way in order to obtain a successful result.

**Objective function value reduction.** In order to select the most productive parameterization it would be helpful to know the minimum objective function values, $\min_c J$, achievable with each candidate refinement. This is, however, clearly unattainable due to the high computational costs associated with performing the corresponding minimizations numerically.

To avoid such costly computations, the predicted attainable objective function value, $\tilde{J}$, is introduced. $\tilde{J}$ is an approximation of $\min_c J$ based on a linearization of $m(c)$ and is therefore available by analytical calculation when using an output least squares formulation. The resulting expression for $\tilde{J}$ is

$$\tilde{J}(\mathcal{P}_Q) = \Delta d^T (D^{-1}G) \Delta d, \quad (2)$$

$$G = D^{-1} A_Q (A_Q^T D^{-1} A_Q)^{-1} A_Q^T D^{-1}, \quad (3)$$

where $\mathcal{P}_Q$ is a parameterization containing $Q$ parameters. The sensitivity matrix for the parameterization $\mathcal{P}_Q$ is denoted $A_Q$, and $\Delta d = d - m(c)$ is the current residual.

In order to find the best parameterization $\mathcal{P}_Q$ with $Q$ parameters, $\tilde{J}$ for different parameterizations containing $Q$ parameters are compared. The parameterization with the lowest $\tilde{J}$-value is the best choice with $Q$ parameters (local winner).
Figure 3: Example of ranking of predictions from refinement analysis. $P_L$ is the parameterization before refinement. In this case $Q = L + 2$, since $J_{L+3} + \sigma_{L+3} \geq J_{L+2}$ and $J_{L+2} + \sigma_{L+2} < J_{L+1}$.

The local winners for different numbers of parameters are then compared in order to find the parameterization to be used in the next step of the estimation sequence.

**Restriction of parameterization dimension.**

Generally, an increase in the number of added parameters will decrease the value of $J$. Hence, selecting the parameterization with the lowest $J$ will introduce as many new parameters as is allowed for. In order to determine the number of new parameters to be introduced it is therefore not optimal to have only $\min P J$ as selection criterion. Using a measure of the uncertainty in $J$, we can assess whether the reduction in $J$ when introducing an increased number of parameters is significant.

The measurement errors are random and therefore $J$ is also random. The standard deviation, $\sigma(J)$, of $J$ can be used as a measure of the uncertainty in $J$, but $\sigma(J)$ is unknown. Instead we use an approximation $\tilde{\sigma}(J)$, $\tilde{\sigma}(J)$, to $\sigma(J)$.

$$\tilde{\sigma}(J(P_Q)) = \sqrt{4J(P_Q) - 2(M - Q)} \quad (4)$$

(Recall that $M$ is the number of production data measurements).

An increase in the number of parameters will reduce $J$ but if $J(P_{Q+1}) + \tilde{\sigma}(J(P_{Q+1})) \geq J(P_Q)$ the reduction is not considered to be significant and $P_Q$ is selected as the new parameterization. Figure 3 shows an example of selecting the parameterization.

**Alternative selection of parameterization.**

The parameterization selection methodology has two ingredients; Prediction of attainable objective function value, and; Restriction of parameterization dimension. For comparison when applied to numerical examples, we also present two alternatives to the approach outlined above, for the latter task.

**Restriction of parameterization dimension – Alternative 1: Assessing parameter uncertainties.** The first alternative is to also take into consideration the predicted parameter uncertainty for the candidate parameterizations, and avoid increasing the parameterization dimension beyond the point where the uncertainty in the estimated parameters becomes too large, see also.$^6$ This may be done by ranking the parameterizations according to $J(P_Q)U(P_Q)$, where $U(P_Q)$ predicts the uncertainty of the estimated parameters with the parameterization $P_Q$.

The influence of the measurement errors on parameter uncertainties is given by the covariance matrix for the parameters, $R_Q$. Linearization of $m(c)$ leads to

$$R_Q \approx (AQD^{-1}AQ) = HQ.$$ 

Two simple measures for parameter uncertainty derived from $H_Q$ are $\det(H_Q)$ and $\lambda_{\text{max}}(H_Q)$. In this paper we will assess parameter uncertainty using a determinant measure, $U(P_Q) = (\det(H_Q))^{1/Q}$, where the $Q$-th root is added to the expression to get a meaningful comparison between parameterizations with different dimensions.

**Restriction of parameterization dimension – Alternative 2: A coarser measure for the uncertainty in $J$.** Let all observations have identical standard deviations, denoted by $\sigma$, and let $z$ be an $M$-vector where $|z_i| = 1; i \in [1, M]$. A measure for the error level in the observations, $d$, can then be written $e_d = \sigma z$.

The error in the objective function can then be found, using standard formulas for error propagation,

$$e_J = (\nabla_d J)^T e_d = 2((d^T m(c))^{T} D^{-1})e_d.$$ \quad (5)

This leads to the following upper bound for $|e_J|$,

$$|e_J| \leq 2\|D^{1/2}(d^T m(c))\|\|z\| = 2\sqrt{JM}. \quad (6)$$

A basic assumption for the parameterization selection methodology is that $J \approx \min_{\mathcal{C}} J$ for any given parameterization. The effect of using $|e_J| = 2\sqrt{JM}$ as an alternative to $\tilde{\sigma}(J)$ for restriction of
<table>
<thead>
<tr>
<th>Reservoir dimensions</th>
<th>150m×150m×10m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porosity</td>
<td>( \phi = 0.3 )</td>
</tr>
<tr>
<td>Water viscosity</td>
<td>( \mu_w = 1.0 \times 10^{-3} \text{Pa}\cdot\text{s} )</td>
</tr>
<tr>
<td>Oil viscosity</td>
<td>( \mu_o = 1.3 \times 10^{-3} \text{Pa}\cdot\text{s} )</td>
</tr>
<tr>
<td>Water rel. perm.</td>
<td>( k_{rw} = s_w^2 )</td>
</tr>
<tr>
<td>Oil rel. perm.</td>
<td>( k_{ro} = s_o^2 )</td>
</tr>
<tr>
<td>Capillary pressure</td>
<td>( P_c(s_w) \equiv 0 \text{kPa} )</td>
</tr>
</tbody>
</table>

The parameterization dimension will be illustrated at the end of the paper.

**Termination of estimation sequence.** The sequence of estimations is terminated when the value of the objective function can be explained by random measurement errors, i.e. the sequence is terminated when

\[
J(P_Q) < (M \quad Q) + \sqrt{2(M \quad Q)},
\]

(see, e.g.,\(^{10}\)). Termination is also effectuated whenever the criterion for introduction of new parameters advises against introducing more parameters.

**Numerical examples**

The numerical examples are performed on a square horizontal reservoir with constant thickness and no-flow outer boundaries. The model equations describing two-phase, incompressible, immiscible flow, are discretized spatially on a uniform Cartesian 16 x 16 grid, using standard methods for multi-phase porous-media flow (see, e.g.,\(^{1}\)). Discretization in the time domain is fully implicit – i.e., the model equations are solved simultaneously – with adaptive selection of the time step lengths.

With the exception of the permeability, the physical characteristics of the examples are held fixed. Table 1 lists the relevant properties.

In all simulations, the reservoir contains initially 100% oil. Water is injected into the reservoir at a constant total rate, divided equally among three (two in one example) injection wells. (Wells are shown as white dots on the figures.) The pressure in the producer is fixed at 3.5 MPa. Oil and (after water breakthrough) water are produced from a single production well, located in the upper right corner on the figures.

**Base cases.** Four permeability distributions are selected for the base cases. Figure 4 shows these permeabilities.

This section presents the results using AME on the four base cases, in Figures 5 to 8. The leftmost plot in each figure shows the true permeability for that case, while the rightmost plot shows the final estimate. Intermediate plots show the parameterization at the corresponding stage in the estimation sequence. (The single-parameter estimate for each case is not shown, however, since this plot is not of any interest for the parameterization development.)

For Examples b-1 and b-4 we recover the true permeability to high precision, while for Examples b-2 and b-3 visible differences on fine scales exist.

**Perturbed base cases.** Next, the base cases are perturbed with respect to the true permeability. True permeabilities and final estimates are shown in Fig-

---

Figure 4: True permeability distributions used for the base cases, b-1, b-2, b-3, b-4. Permeability values in mD.
Figure 5: Example b-1. From left to right: True permeability, sequence of (four) parameterization plots, final estimate.

Figure 6: Example b-2

Figure 7: Example b-3

Figure 8: Example b-4

Figure 9: Examples with perturbed coarse-scale permeability distribution. From left to right: Example p-1; True permeability, Example p-1; Final estimate, Example p-2; True permeability, Example p-2; Final estimate.

Figure 10: Comparison of AME and OME. From left to right: Example b-1-4; AM estimate, Example b-1-4; OM estimate, Example b-1-3; AM estimate, Example b-1-3; OM estimate.

Here, the termination criterion, Equation 7, is fulfilled only for Example p-2, while Example p-1 comes quite close. Coarse-scale trends in the true permeabilities are recovered for both examples.

**Comparison of OME and AME.** This section compares results with AME and OME on Example b-1-4 (identical to Example b-1) and on Example b-1-3, which is the same as Example b-1-4, except that the injector in the upper left corner has been removed. Figure 10 shows the results from Example b-1-4 and Example b-1-3.

In Example b-1-4, the termination criterion is fulfilled for the OM estimate. However, the true permeability is clearly not recovered, and some characteristics of over-parameterization are visible.

In Example b-1-3, the AME fulfills the termination criterion and identifies the true permeability. For OME the termination criterion is not fulfilled. The true permeability is not recovered, and the signs of over-parameterization are more pronounced than for Example b-1-4. Several values are outside the [0, 300] mD-range used for the shading on all examples in this paper.
Utilization of fluid rate data. So far we have utilized only pressure data in the injectors for the AME. Utilizing also fluid rate data in the producer could potentially allow AME to recover more complex permeability patterns. Figure 11 shows results obtained using pressure data only (P-estimate), results obtained using pressure and fluid rate data simultaneously (PR-estimate), and results obtained using pressure and fluid rate data sequentially (sequential PR-estimate). (With the latter estimation strategy, a PR-estimation is started at the point where the P-estimation finished.) It is seen from the P-estimate and the PR-estimate that the expectation of an improved estimate when including fluid rates in the data is not always met. Potential reasons for this is discussed in8, and potential improvements are also suggested. The sequential PR-estimate seen on the right plot is the result of such an improvement.

Comparison of standard and alternative parameterization selection strategies. Finally, we illustrate use of the different alternatives for parameterization selection (see the sections following Figure 3) with AME (again using only pressure data). We concentrate on the selection of parameterization following the four-parameter step of Example b-1 (Figure 5, third plot).

With Alternative 1, we rank the parameterization candidates according to the value of the product $S = J(\mathcal{P}_Q)\bar{U}(\mathcal{P}_Q)$. This is illustrated graphically in Figure 12, where we have plotted $J(\mathcal{P}_Q)$ vs. $\bar{U}(\mathcal{P}_Q)$ for some of the parameterization candidates. Using log-log scaling of the axes, the contour lines of the function $J\bar{U}$ are straight lines, and the contour line passing through the point corresponding to the parameterization with the lowest value of $S$ is shown. The points belonging to the “best” parameterizations with one, two and three new subdivisions are indicated by the corresponding numbers. Figure 13 shows the division of the reservoir with the “best” parameterizations. We note that the “global best” parameterization, with two new parameters, seems to be better than the choice made by the standard selection method when the goal is to be able to represent the true permeability with as few parameters as possible. We also note that the one-parameter extension would be even better in this respect. With a little more weight on the uncertainty measure, the one-parameter extension would be the “global best” parameterization. The relative importance of $J$ with respect to $\bar{U}$ can be varied if we consider ranking according to $J(\mathcal{P}_Q)(\bar{U}(\mathcal{P}_Q))^\alpha$.

While Figure 13 indicates that using the alternative method for parameterization selection could be beneficial, it has not produced consistently better results than the standard method.
Figure 14: Left: The parameterization selection using the coarser measure on the error in $\hat{J}$. The magnitude of $\bar{\sigma}$ is exaggerated. Right: The selected one-parameter extension to the parameterization.

Figure 14 shows the results when using Alternative 2 for the parameterization selection. This method has a coarser measure for the uncertainty in $\hat{J}$, with $|\epsilon J| = 2\sqrt{J M}$. In this example this coarser measure leads to a selection of the one-parameter extension, since $J_6 + e_6 > J_5$. We note that with the standard measure, we would go on to compare $J_6$ and $J_7$, since $J_6 + \bar{\sigma}_6 < J_5$. The division of the reservoir corresponding to $P_5$ is shown to the right in Figure 14. In this case this is not a productive refinement of the parameterization.

Summary

We have presented parameterization selection in adaptive multiscale estimation (AME). The ability of the method to identify various permeability trends has been illustrated with several examples, and compared to ordinary multiscale estimation (OME). Different criteria for the selection of parameterization refinement has been discussed, and implications for the choice of parameterization in the examples has been illustrated.

Acknowledgements

The research presented in this paper has been financially supported by ENI-Agip SpA, Norsk Hydro ASA, and The Research Council of Norway.

References