Comput Geosci (2006) 10: 321–342
DOI 10.1007/s10596-006-9026-6

Reservoir description using dynamic parameterisation selection with a combined stochastic and gradient search

Lars Kristian Nielsen · Sam Subbey · Mike Christie · Trond Mannseth

Received: 18 April 2004 / Revised: 15 May 2006 / Accepted: 23 May 2006 / Published online: 26 August 2006

Abstract There is a correspondence between flow in a reservoir and large scale permeability trends. This correspondence can be derived by constraining reservoir models using observed production data. One of the challenges in deriving the permeability distribution of a field using production data involves determination of the scale of resolution of the permeability. The Adaptive Multiscale Estimation (AME) seeks to overcome the problems related to choosing the resolution of the permeability field by a dynamic parameterisation selection. The standard AME uses a gradient algorithm in solving several optimisation problems with increasing permeability resolution. This paper presents a hybrid algorithm which combines a gradient search and a stochastic algorithm to improve the robustness of the dynamic parameterisation selection. At low dimension, we use the stochastic algorithm to generate several optimised models. We use information from all these produced models to find new optimal refinements, and start out new optimisations with several unequally suggested parameterisations. At higher dimensions we change to a gradient-type optimiser, where the initial solution is chosen from the ensemble of models suggested by the stochastic algorithm. The selection is based on a predefined criterion. We demonstrate the robustness of the hybrid algorithm on sample synthetic cases, which most of them were considered insolvable using the standard AME algorithm.

Keywords Adaptive Multiscale Estimation · gradient optimiser · inverse problem · Neighbourhood Approximation algorithm · permeability estimation · reservoir simulation · stochastic search algorithm · two-phase flow

L. K. Nielsen (✉)
Department of Mathematics, University of Bergen,
Bergen, Norway
e-mail: larskn@mi.uib.no

T. Mannseth
CIPR—Centre for Integrated Petroleum Research,
University of Bergen,
Bergen, Norway
e-mail: Trond.Mannseth@cipr.uib.no

S. Subbey
Institute of Marine Research,
Bergen, Norway
e-mail: samuels@imr.no

M. Christie
Institute of Petroleum Engineering, Heriot Watt University,
Edinburgh, UK
e-mail: Mike.Christie@pet.hw.ac.uk

1. Introduction

Predictions of the reservoir behaviour require estimates of the reservoir property values, such as permeability and porosity, on a grid block scale. Even if all available data sources are utilised this can be a very difficult task. Large scale permeability trends, like barriers and channels, have large impact on how the fluids flow in the porous medium. If a channel or barrier has not been detected by a seismic survey, and there is no well through it, it is hardly accounted for in a model simulating the fluid flow.

Available data types for estimation of the permeability are static and dynamic well data and geological

data. The static well data can be obtained from core samples in the wells, and the geological data are usually coarsened geological permeability maps. The dynamic well data are time series of pressures and flow rates in the wells, but they give no direct information of the permeability field. Using the equations of fluid flow we can use this indirect information to estimate the permeability trends. The issue in this paper is the inverse problem of estimating the permeability field using pressure data from the wells.

Inverse problems are often ill-posed, and that means that even if the uncertainties in the measurements are small, the errors in the solution may become large. To reduce the possibility of introducing large errors we have to regularise the problem in a proper way. This can be done by restricting the parameter space to exclude non-physical solutions.

Different strategies for regularisation of the inverse problem described above have been suggested. A common method is to compensate for over-parameterisation by penalising deviations from a priori knowledge of the permeability given by a geological model. An example of this is Bayesian estimation, see e.g. [1]. Another regularisation approach is to fix the resolution a priori by using a zonation. This means that the reservoir is divided into clusters of grid blocks, where each cluster has a constant permeability value. One problem with this approach is that we may choose a too low resolution, which will not reconcile the observed data. In the opposite case, with a too high resolution, the computational work will be unnecessarily high and the produced permeability values are often clearly out of range.

A multiscale estimation method overcomes the problem related to manually choosing the resolution a priori by solving a sequence of parameter estimation problems with increasingly higher resolution. The optimisation is performed sequentially until a satisfactory match between model predictions and observed data is obtained and the inverse problem is solved.

Grimstad et al. [2] presented the inversion technique called Adaptive Multiscale Estimation (AME). The method starts out with a clearly too low resolution, and seeks to increase the resolution gradually only in the regions where it is warranted by the data. In this way the refinements will not necessarily be the same all over the reservoir. The goal of the process is to produce *one* coarse scale solution solving the inverse problem. If required, the solution can later be downscaled by adding fine scale information, see e.g. [3, 4].

One of the drawbacks with the AME strategy is the risk of getting no acceptable solution or a solution not satisfying the coarse scale property. A way to reduce this problem is to apply more than one parameterisation for each stage in the optimisation sequence. Another related approach is to use information from several points in the parameter space when selecting the refinements to apply in the next stage. Because we are dealing with an ill-conditioned, inverse problem with no unique solution these approaches are reasonable. Especially in the early stages it might be important to have a spread in the way of approaching the final parameterisation, and hence the solution. Later we are hopefully closer to the global minima of the original inverse problem, and at this point the suggested refinements will be more reliable.

The AME algorithm is based on a dynamic parameterisation selection, which can be used together with any suitable optimisation routine. In earlier studies the Levenberg–Marquardt algorithm, which is a deterministic gradient method, has been used to solve the optimisation problem. To derive an ensemble of promising parameterisations in each stage of the algorithm, we can apply a stochastic search algorithm to derive several solutions of the optimisation problem. Generating a larger number of candidate solutions than the required number of refinements will further stabilise the selection of refinements.

The use of several parameterisations at each stage and the incorporation of a stochastic optimiser will increase the computational cost of the algorithm. To get faster convergence we can change from the stochastic method to a gradient optimiser when the number of parameters gets higher. For this we need a way to combine the stochastic and the gradient approach.

In this work we have chosen to start with a stochastic method, the Neighbourhood Approximation (NA) algorithm, and change to a gradient optimiser, Levenberg–Marquardt, when the dimension of the parameter space is above a certain level. We also apply the gradient method if we continue with an unchanged parameterisation in the next stage. The starting point for the gradient optimiser will be chosen as the most promising set of parameters from last parameterisation, in order to decrease the objective function value.

The applied stochastic search method is in this paper the Neighbourhood Approximation algorithm (NA) [5]. This method starts out by randomly generating points in the parameter space by a Uniform Monte Carlo method [6]. It then uses the concept of Voronoi Cells [7] to determine regions of good fit. These regions are then selectively sampled to generate good data-fit models. Thus, instead of searching for only a single model giving the supposed global minimum, this method is able to find an ensemble of models that fit the data to some degree of accuracy.

In earlier work [5, 8, 9], the method of using several candidate solutions have been used to do uncertainty assessments of a solution with a fixed parameterisation. In the current work the ensemble of candidate solutions is used to stabilise the parameterisation selection. An uncertainty study could also in this case have been carried out by the methodology described in this paper, but this is beyond the scope of this work.

The aim of this paper is to strengthen the regularisation of the AME approach in order to increase the possibility to achieve an acceptable coarse scale solution.

This paper is organised in the following way. The model equations for the inverse problem are defined in Section 2 and the concept of the parameter estimation problem is given in Section 3. Section 4 presents the dynamic parameterisation selection from the AME algorithm, and in Section 5 the issue of the parameter search, and especially a description of the NA-algorithm, is treated. Further the coupling of the NA-algorithm with the dynamic parameterisation selection is discussed in Section 6, while numerical results are presented in Section 7. Conclusions and remarks are given in Sections 8 and 9, respectively.

2. Model equations

Assuming only oil and water present in a porous medium with isotropic permeability, the conservation equations for two-phase incompressible, immiscible, horizontal flow are

$$\phi(\mathbf{x})\frac{\partial s_o}{\partial t} - \nabla\cdot\left(\frac{k(\mathbf{x})k_{ro}(s_o)}{\mu_o}\nabla p_o\right) = q_o(\mathbf{x}), \tag{1}$$

$$\phi(\mathbf{x})\frac{\partial s_w}{\partial t} - \nabla\cdot\left(\frac{k(\mathbf{x})k_{ro}(s_w)}{\mu_w}\nabla p_w\right) = q_w(\mathbf{x}), \tag{2}$$

where the subscripts o and w refer to the phases, water and oil, respectively. s_i denotes the saturation, μ_i the viscosity, p_i the pressure, q_i the external volumetric flow rate and k_{ri} is the relative permeability, where i is the fluid phase. The porosity and the absolute permeability are given by $\phi(\mathbf{x})$ and $k(\mathbf{x})$, respectively, where $\mathbf{x}$ is the spatial position in the porous medium. In addition we assume a completely saturated medium,

$$s_o + s_w = 1, \tag{3}$$

and suppose we have a function P_c defining the capillary pressure,

$$p_o - p_w = P_c. \tag{4}$$

The quantities ϕ, k, k_{ri} and P_c are all dependent of the porous medium and are not accessible through direct measurements.

The problem treated in this paper is to find an estimate of the absolute permeability, k, when ϕ, k_{ri} and P_c are assumed to be known. Equations (1)–(4) define this task as an inverse problem.

3. Parameter estimation problem

Let the permeability $k(\mathbf{x})$ be given by the parameterisation

$$k_N(\mathbf{x}) = k(\mathbf{c}_N, \{\psi_i(\mathbf{x})\}_{i=1}^N) = \sum_{i=1}^N c_i\psi_i(\mathbf{x}), \tag{5}$$

where $\mathbf{c}_N \in \mathbb{R}^N$ is a vector of parameter values, and $\{\psi_i\}_{i=1}^N$ is a set of real valued piecewise constant basis functions, spanning the space defined by k, which is to be estimated. Further we let $\mathbf{d} \in \mathbb{R}^M$ denote the measurements, in this case time series of production data, and $\mathbf{m}(\mathbf{c}_N) \in \mathbb{R}^M$ denote the corresponding simulated values calculated using the model equations.

To measure the misfit between the simulated and the measured data we define an objective function (misfit function)

$$J(\mathbf{c}_N) = (\mathbf{d} - \mathbf{m}(\mathbf{c}_N))^T D^{-1}(\mathbf{d} - \mathbf{m}(\mathbf{c}_N)), \tag{6}$$

where D is the covariance matrix which weights the data according to the measurement errors.

The inverse problem is solved when the objective function is minimised to a value which can be explained by the measurement errors. Assuming normally distributed measurement errors with zero mean and variance given by the diagonal elements of D, and further that the objective function is close to linear in a region around a minimum, J will be χ^2-distributed with $M - N$ degrees of freedom [10]. The statistical expectation value of J at solution of the inverse problem is therefore given by $M - N$, and the standard deviation of J is equal to $\sqrt{2(M-N)}$ at the same point. Based on these assumptions we have decided to terminate the sequence of estimation problems when

$$J(\mathbf{c}_N) < (M - N) + \sqrt{2(M-N)}. \tag{7}$$

4. Dynamic parameterisation selection

This paper uses the dynamic parameterisation selection from the Adaptive Multiscale Estimation (AME) in the process of reconstructing the true permeabilities in a reservoir. This approach seeks a coarse scale solution to

the inverse problem by applying a regularisation which restricts the number of required parameters. We will here present the motivation behind this regularisation approach and describe the main ideas of the method. For a complete description, see [2]. Related work can also be found in [11–14].

For equations corresponding to single phase flow a number of studies [15–18] indicate low parameter sensitivities (high uncertainties) and large model non-linearities related to small scale variations in the solutions. The relation between low sensitivities and small scale oscillations means there should exist a coarse scale solution of the problem.

A correspondence between sensitivities, non-linearities and scales has, to our knowledge, not been verified for the more complicated two phase flow equations. Numerical results, see e.g. [2, 4, 19, 20], though indicate that, with sparsely distributed data, there seems to be a similar relation for these equations as for the simpler studied equations. The AME method seeks a coarse scale solution of the inverse problem on the assumption that the mentioned correspondence is true for the two-phase flow equations.

In addition to coarse scale solutions there will exist finer scale solutions. The fine scale solutions will normally be more expensive to calculate because they require a higher number of parameters. More parameters give possibilities for an increase in the small scale variations, and hence the parameter uncertainties will normally be higher. The associated high non-linearities will in addition make these solutions more difficult to find.

The basic idea in multiscale estimation is to start out with a clearly too low resolution and increase the resolution gradually. Instead of increasing the resolution by the same amount in the entire grid, which is the case in an ordinary multiscale estimation, see e.g. [4], the AME approach only increases the resolution in regions where it is expected to be productive. The way of doing refinements is in this work to split each region into two equally sized subregions by a vertical or horizontal line (see figure 1). The AME method has in other studies also been used with other possible ways of doing refinements, see e.g. [13].

For each chosen parameterisation, P_N, we solve the optimisation problem, finding

$$\mathbf{c}_N^* = \arg\min_{\mathbf{c}_N} J(P_N, \mathbf{c}_N). \tag{8}$$

If the optimal values do not solve the inverse problem according to the solution criterion in Eq. (7), these values are used in the process of finding a new refinement and eventually as a starting value for the optimiser in the next step if the optimiser requires an initial value.

When selecting the next parameterisation we calculate a predicted minimum value, $\tilde{J}$, of the objective function for each new candidate refinement. To calculate the value of $\tilde{J}$ we assume the mathematical model is linear around the actual point in the parameter space, and then use a linear approximation of $\mathbf{m}(\mathbf{c})$ inserted in J. If N' is the new dimension of the parameter space, $\tilde{J}$ can be expressed as (see Appendix or [2] for derivation)

$$\tilde{J}(P_{N'}) = \Delta\mathbf{d}^T(D^{-1} - [D^{-1}A_{N'}(A_{N'}^T D^{-1} A_{N'})^{-1} \times A_{N'}^T D^{-1}])\Delta\mathbf{d}, \tag{9}$$

where $\Delta\mathbf{d} = (\mathbf{d} - \mathbf{m}(\mathbf{c}_{N'}^I))$, $P_{N'}$ is a refinement of P_N containing $N' \geq N$ parameters and $A_{N'} = \mathbf{m}'(\mathbf{c}_{N'}^I)$ is the sensitivity matrix. As a starting point for the linearisation we use the parameters $\mathbf{c}_{N'}^I$ producing the same permeability field as the solution, $\mathbf{c}_N^F$, of the previous optimised problem.

To select the new parameterisation we go through two steps. First we find the lowest $\tilde{J}$ for all possible $P_{N'}$ for each $N' = N, N+1, N+2, \ldots$. In the next step we determine which dimension to use by deciding whether the reduction of $\tilde{J}$ will be significant in comparison with the increase in N'. We consider an reduction in $\tilde{J}$ significant if

$$\tilde{J}(P_{N'}) + \tilde{\sigma}(\tilde{J}(P_{N'})) \leq \tilde{J}(P_{N'-1}), \tag{10}$$

where $N' > N$ and $\tilde{\sigma}$ is an approximation of the standard deviation. This approximation is based on the assumption that the expectation value of $\tilde{J}$ is equal to $\tilde{J}$, and $\tilde{\sigma}$ is given by (see Appendix or [2])

$$\tilde{\sigma}(\tilde{J}(P_{N'})) = \sqrt{4\tilde{J}(P_{N'}) - 2(M - N')}. \tag{11}$$

The purpose of this methodology is to keep the number of parameters as low as possible, but introduce new parameters if it seems to make a considerable decrease in the misfit function. This will at the end of the sequence hopefully produce a coarse scale solution. In practice, undesirable refinements are sometimes chosen. If an undesirable refinement is introduced in one stage, this results in a higher total number of parameters required to produce a low misfit value.

It turns out that the refinement with the lowest $\tilde{J}$ is not necessarily the optimal refinement for producing a coarse scale solution of the inverse problem. Figure 2 shows an example where the AME process is capable

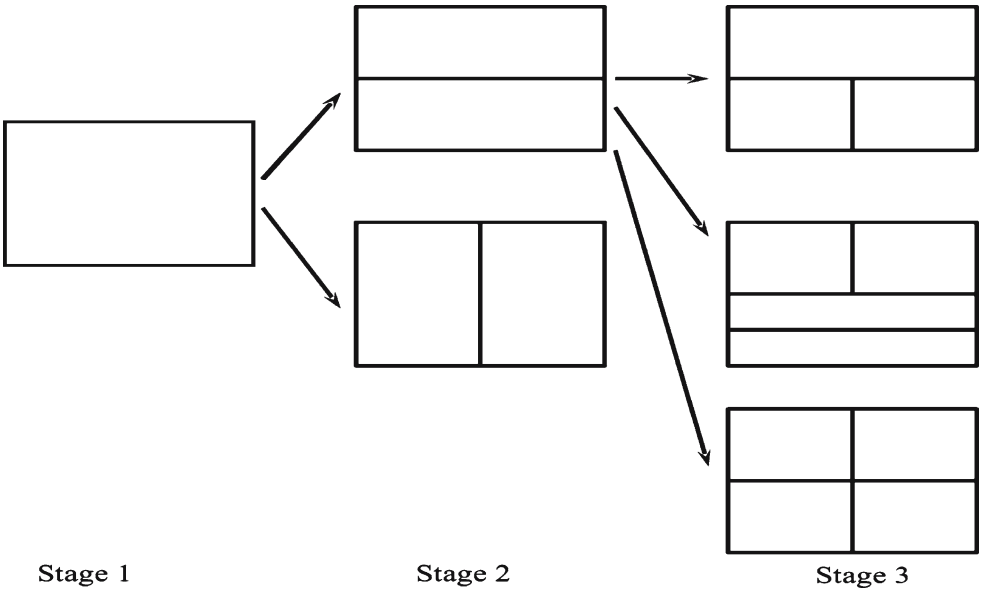

Figure 1 Allowed subdivisions. For each step it is allowed to divide a region into two equally sized subregions by either *a horizontal or a vertical line*. From the initial parameterisation with one parameter there are two possible refinements. In the next step there is a large number of allowed refinements. Three of them are presented here.

of introducing a number of reasonable refinements in order to reproduce the true permeability field. The sequence also introduces a few extra unnecessary parameters, but the most important refinements are performed at an early stage. Figure 3 presents another example where the same true permeability field is applied, but the positions of two of the wells (marked with white dots in the figures) are slightly changed. In this particular example the AME sequence chooses an undesirable parameterisation for the second estimation (figure 3c). This results in problems finding the necessary refinements later on, and in this example the final result is not reproducing the structures of the true field. A large number of parameters, most unnecessary compared with the true field, is introduced in order to decrease J below the solution criterion. The resulting permeability field also contains more small scale variations than the solution in the first example.

A contribution to the AME method in this work is to strengthen the regularisation such that unnecessary parameters, judged on the ability to produce a coarse scale field, are less often introduced.

5. Parameter search

The optimisation problem, finding

$$\mathbf{c}_N^* = \arg\min_{\mathbf{c}_N} J(P_N, \mathbf{c}_N), \tag{12}$$

where J is defined in Eq. (6), can be solved by either a gradient method or by a stochastic optimisation approach. The gradient methods use information from the gradients of the mathematical model with respect to the parameters to find the optimal set of parameters. These methods have fast convergence rates. The drawback is that gradient algorithms in some cases can fail to converge or converge to a local minimum, depending on the shape of the misfit function and the initial input value. The stochastic methods do not require any gradients to derive the search of the minimum, but have generally much slower convergence rates than the gradient methods. Many of the stochastic methods, for instance simulated annealing and genetic algorithms, are theoretically classified as global methods, though

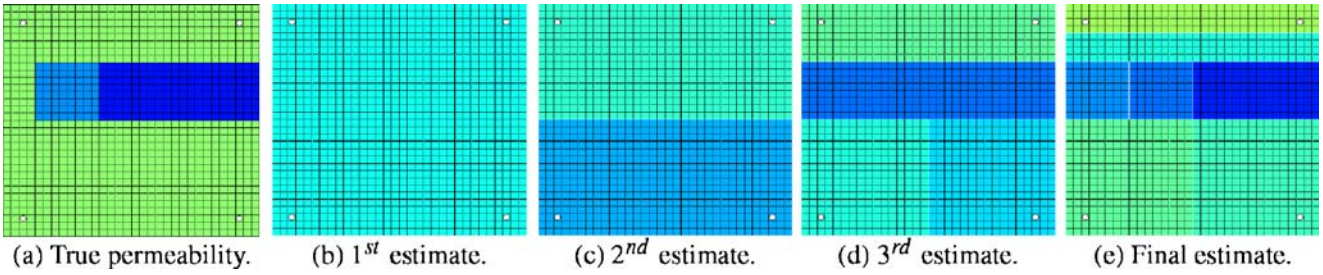

Figure 2 The *left figure* shows the true permeability field we want to reproduce. The three first stages and the final estimate of the AME sequence are shown.

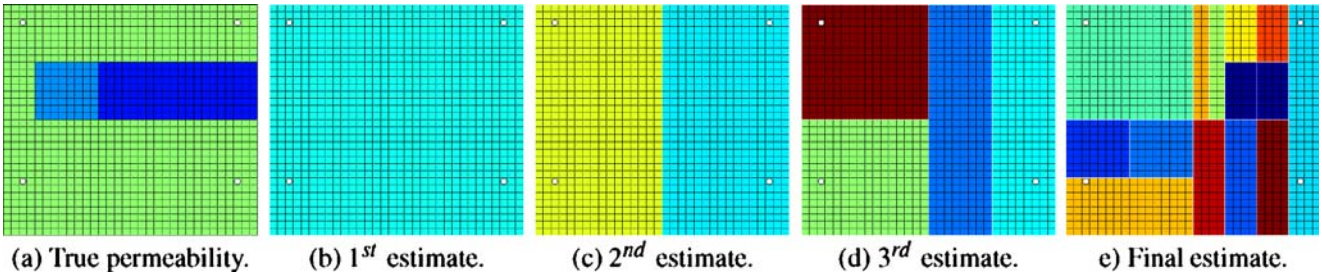

Figure 3 In this example we have an identical permeability field (**a**) as in the previous example, but the positions of two of the wells are slightly changed. The estimates between the third and the final estimate are not shown here. The parameterisation applied in the second estimate (**c**) is not the preferred refinement according to the true permeability field. This results in a final estimate which contains large small scale oscillations and is far from the desired solution.

in practice also these methods may converge to a local minimum of the objective function [5].

One of the advantages with the AME method is that the initial point, given to the gradient optimiser, is gradually improved through the stages of increasing dimension of the parameter space. In this way the risk of converging to a minimum which is not the global one decrease. As discussed earlier, the process though may fail in some cases.

The main purpose of applying a stochastic method in this work is its possibility to derive several candidate solutions to each optimisation problem. This gives more information of which new parameterisation to apply and larger possibilities to keep on with several refinements for each stage.

In this paper we use the stochastic search algorithm *Neighbourhood Approximation* (NA) for the parameter optimisation when the number of parameters is low. On the other hand, with higher dimension of the parameter space or if we want to continue with an unchanged parameterisation, we apply the *Levenberg–Marquardt* (LM) algorithm, which is a gradient type optimiser. Details about theory and implementation of the LM optimiser can be found in [21].

5.1. Neighbourhood Approximation (NA)

The Neighbourhood Approximation (NA) algorithm was originally developed by Sambridge for solving nonlinear inverse problems in seismology [22–24]. In petroleum engineering the same algorithm has successfully been applied by Subbey and co-workers [5, 8, 9].

The algorithm is a stochastic search algorithm which searches the parameter space in order to find good fitting models. It uses *Voronoi cells* to discretise the parameter space. In the following it is important not to confuse the division of the parameter space, achieved by the Voronoi Cells, and the refinements of the reservoir, which is performed by divisions of rectangular regions and used in the multiscale parameterisation. We will now define the Voronoi cells and describe how the NA algorithm works in practice.

For N parameters we specify a region $\Omega \subset \mathbb{R}^N$ within which to sample. For the defined space $\Omega \subset \mathbb{R}^N$, let ϕ denote a set of n_s points in Ω. A Voronoi cell $V(\mathbf{c}_\alpha)$ is defined as the nearest neighbour region around a point $\mathbf{c}_\alpha \in \phi$ compared to the other points in ϕ. Mathematically this can be expressed as

$$V(\mathbf{c}_\alpha) = \{\mathbf{c} \in \Omega : \| \mathbf{c} - \mathbf{c}_\alpha \| \leq \| \mathbf{c} - \mathbf{c}_\beta \| \quad \text{for} \quad \beta \neq \alpha \quad (\alpha, \beta = 1, 2, ..., n_s)\},$$

where the norm $\| \cdot \|$ is the Euclidean distance.

The Voronoi cells give us a way of dividing a real space of any dimension into a number of unique N-dimensional convex polygons. The points on the edges of the polygons will be equidistant from exactly two sampling points, and the points on the vertices are equidistant from at least three. The regions fill the space and each region will have size inversely proportional to the sampling density of the points in ϕ.

A *Voronoi diagram* can be used to approximate the values of a scalar function $J(\mathbf{c})$, where $\mathbf{c} \in \mathbb{R}^N$. Given a set of points $\{\mathbf{c}_\alpha\} \in \phi$ dividing the space $\mathbb{R}^N$ into Voronoi cells $V(\mathbf{c}_\alpha)$, we can construct a Voronoi diagram by setting the value in cell $V(\mathbf{c}_\alpha)$ constant and equal to the function value $J(\mathbf{c}_\alpha)$ for each sampling point $\mathbf{c}_\alpha \in \phi$. The resulting diagram will then be a non-smooth, piecewise constant interpolation of the function J. The Neighbourhood Approximation algorithm uses the information from the Voronoi diagram and refines the sampling density in the cells with lowest J-values. For our problem J is the misfit function defined in Eq. (6).

The search starts out by generating n_s' points in the parameter space by a uniform Monte Carlo method. The next step is to select the n_r models producing

the lowest misfits. In the Voronoi cells of these n_r parameter values we generate totally new n_s models by a uniform random walk, which is accomplished by a Gibbs sampler, see [9] for details. Thereafter all the generated models are collected and we return to the step where we select the new n_r best fitting models. Schematically this will be:

Generate initial n_s' reservoir models uniformly in the parameter space and evaluate the misfit.

1. Select the n_r models giving the lowest misfit.
2. Generate new n_s/n_r models randomly in each of the n_r chosen cells, and evaluate the misfit.
3. Go to 1.

The main input-parameters controlling the algorithm are n_r an n_s. These numbers will determine how explorative the algorithm will be [23]. The larger the ratio n_s/n_r is, the less explorative the sampling of the parameter space will be. If this ratio is large the algorithm will give less weight to the previous models and the convergence will be fast. In the opposite case, using a small ratio, the sampling will be spread over more cells (less local) and the possibility of finding a local minimum is smaller. A general increase of both numbers without changing the ratio, will also give slower convergence and a more explorative algorithm.

Both the search direction used in the algorithm, and the size and shape of the neighbourhoods are determined without any external influence. The search is guided by the introduction of models in the good fitting areas, and this requires only the relative fit to the data. In regions which have the lowest misfit the algorithm will increase the resolution and in the regions with larger misfit, the resolution will stay coarse.

6. Coupling of the dynamic parameterisation selection with the neighbourhood approximation search

In this work we try to take advantage of the spread in candidate solutions from the Neighbourhood Algorithm in order to stabilise the AME process of reproducing a coarse scale permeability field. The main difference between the approach in this paper and the standard AME approach is that the NA algorithm gives us more information of how to choose the next refinements and extended possibilities to select several candidate refinements at each stage of the dynamic parameterisation selection procedure. In cases where we find it advantageous to start new optimisations with more than one of the suggested parameterisations we get a ramification of the algorithm and several ways to approach the solution. Especially early in the AME sequence, when we may not be very close to the solution of the inverse problem, this can give large improvements in order to make the approach more robust. The ramifications will of course make the method computationally more expensive, but restricting this option to the first stages, and using a low number of branches, limits the expenses.

The strategy with producing several candidate models is also a common way to do uncertainty assessments of a produced solution [5, 8, 9]. In that case, which is not carried out here, we have to keep on with the suggested parameterisations for the chosen ensemble of solutions until the end of the estimation.

We will in the following denote the new combination algorithm as AME–NA to distinguish this from the standard AME method.

6.1. Selection of candidate models

When selecting the candidate solutions from the NA algorithm we want to select an ensemble of reservoir models which fits the observed data to some degree of accuracy. In addition it is important that the corresponding points in the parameter space are not too close. If this is the case, we will probably not get the desired improvement in the reliability of the $\tilde{J}$-values because we linearise the mathematical model around approximately the same point. The result is that the suggested new refinements probably will be equal or similar. This will not be in accordance with the ramification idea behind this work, which is to generate several candidate solutions which hopefully will include at least one reasonable new parameterisation.

We will here explain the method we have applied to select the candidate solutions of the optimisation problems. There are several ways to do this selection. We have chosen one which seems to be working satisfactorily for our problem.

The selection is based on the misfit value for each generated reservoir model. Two typical plots of the models and their misfit values are shown in figure 4. In the left figure the models are not sorted (the last generated models have the highest index number on the x-axis), and in the right one they are sorted by decreasing misfit. As we can see from these plots, we typically generate a large number of models with approximately the same misfit (relative to the complete range of misfit values). All these models have a relatively low misfit and experiments have shown that they often will be very close in parameter space.

To pick out a number of models which differs in parameter space and still have a relatively low misfit value, we therefore first check where the absolute value

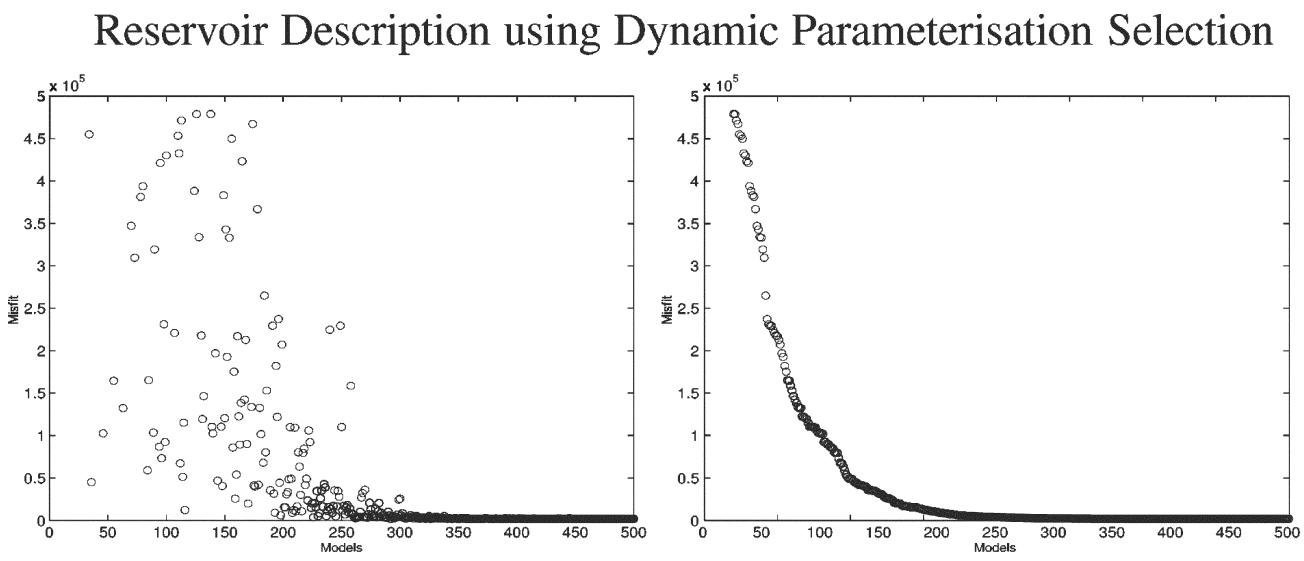

Figure 4 The *left plot* shows the generated reservoir models from the NA algorithm on the x-axis with the objective function value on the y-axis. The *right plot* contains the same models sorted by decreasing misfit. In both cases we have omitted the models giving objective function values greater than $5 \cdot 10^5$ (in this case there were misfit values as high as $1 \cdot 10^{10}$).

of the gradient of the right curve in figure 4 is below some threshold. We then only consider models among those who have lower misfit than the model at this point. To prevent getting several models from the area with approximately equal misfits, we divide the range of the *misfit* in the selected area into equal intervals and select one model from each interval. In this way we will keep the model with the lowest misfit, and still get a reasonable spread of the other models in the ensemble.

In the standard AME approach the optimisation is stopped at each stage when J either reaches the actual value of $\tilde{J}$ or if the gradient optimiser converges before $J \leq \tilde{J}$. In the NA algorithm there is no stopping criterion based on the value of the minimum misfit value. The algorithm has to be run to a fixed number of iterations, and then the candidate models are selected. In the way we are selecting models we will always continue using models which have misfit higher than $\tilde{J}$, and if the NA search produces models with misfit lower than $\tilde{J}$, one or more of these values will be selected as well.

6.2. Global selection of refinements

For each optimisation stage in the AME–NA approach we use one or several candidate solutions produced by the optimiser and calculate $\tilde{J}$ for all possible new refinements. Using the suggested refinements from all candidate solutions will rapidly include a large number of branches and the number of required optimisations for each iteration will increase very quickly. To keep the number of applied refinements low, we have introduced a global selection of new candidate refinements such that only the most promising solutions are kept.

The NA optimiser produces several candidate solutions of the parameter estimation problem (Eq. (12)). For a given iteration in the dynamic parameter selection of AME–NA, we may have a number of solutions from several optimisation problems with unequal parameterisations $\{P_r\}$. The dimension of these spaces may not always be equal. In the following we denote the candidate solutions from one iteration as $\mathbf{c}_i$ where $i = 1, 2, ..., k$ and k is the total number of candidate solutions for all the optimisations at this iteration.

For each produced candidate solution we apply the same selection criterion as described in Section 4 to choose a new refinement. That is, given a candidate $\mathbf{c}_i$ for the optimal parameters of a parameterisation P_r we define the selected new parameterisation as

$$P^i = \text{optimal}_j \tilde{J}(P_{rj}, \mathbf{c}_i^I), \tag{13}$$

where P_{rj} are possible new refinements of P_r, and $\mathbf{c}_i^I$ are the initial parameter values in the new parameterisation (the values we linearises around) corresponding to the solution for the previous optimisation using P_r. The notation *optimal* means that we select the parameterisation which produces the lowest $\tilde{J}$, but fulfilling the restriction that the dimension of the parameter space is increased only if $\tilde{J}$ will be significantly lower. That

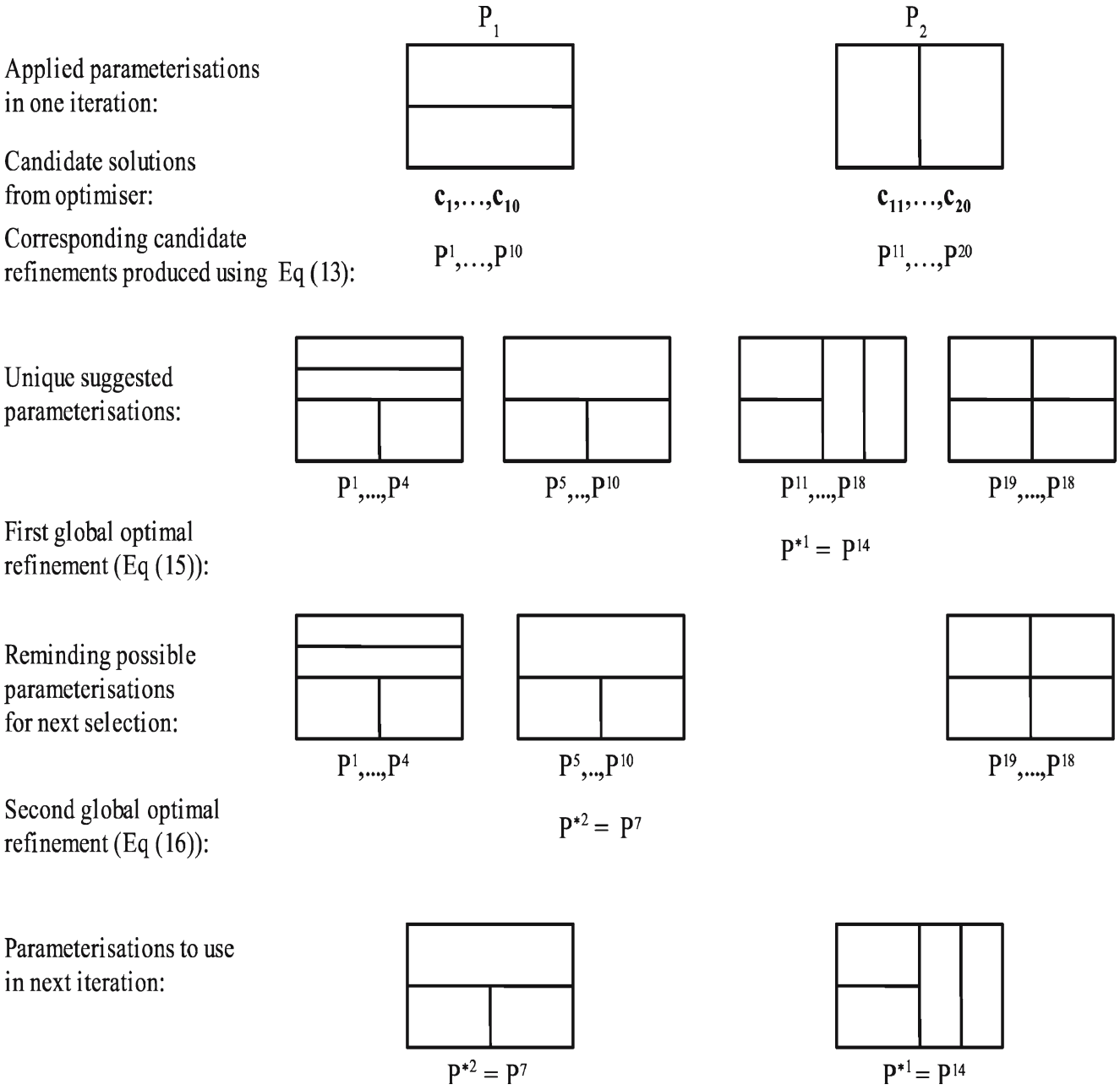

Figure 5 An example of the selection of refinements in one iteration of the dynamic parameterisation process. The optimisation problem is solved for two unequal parameterisations P_1 and P_2, and the optimiser produces ten candidate solutions for each parameterisation. The global selection produces two new parameterisations, P^7 and P^{14}, which is passed to the next iteration.

is, the dimension should only be increased from N_1' to N_2' if

$$\tilde{J}(P_{N_2'}) + \tilde{\sigma}(\tilde{J}(P_{N_2'})) \leq \tilde{J}(P_{N_1'}), \tag{14}$$

where $N_2' > N_1' \geq N$, and $P_{N_i'}$, $i = 1, 2$, are parameterisations with N_i' parameters. This is a slightly more general statement than in Eq. (10). The modification of Eq. (10) is done because we in the following may not have all dimensions of parameter spaces present in the actual ensemble of parameterisations, that is, we may have $N_2' - N_1' > 1$. For all $\mathbf{c}_i$ the selection in Eq. (13) gives us a set of parameterisations $\{P^i\}_{i=1}^k$.

In figure 5 we show an example where the optimisation problem is solved for two unequal parameterisations P_1 and P_2. The stochastic optimiser is used and we have selected ten candidate solutions, $\mathbf{c}_i$, for each case. Totally, we then have a set of 20 candidate parameter vectors and correspondingly 20 new suggested

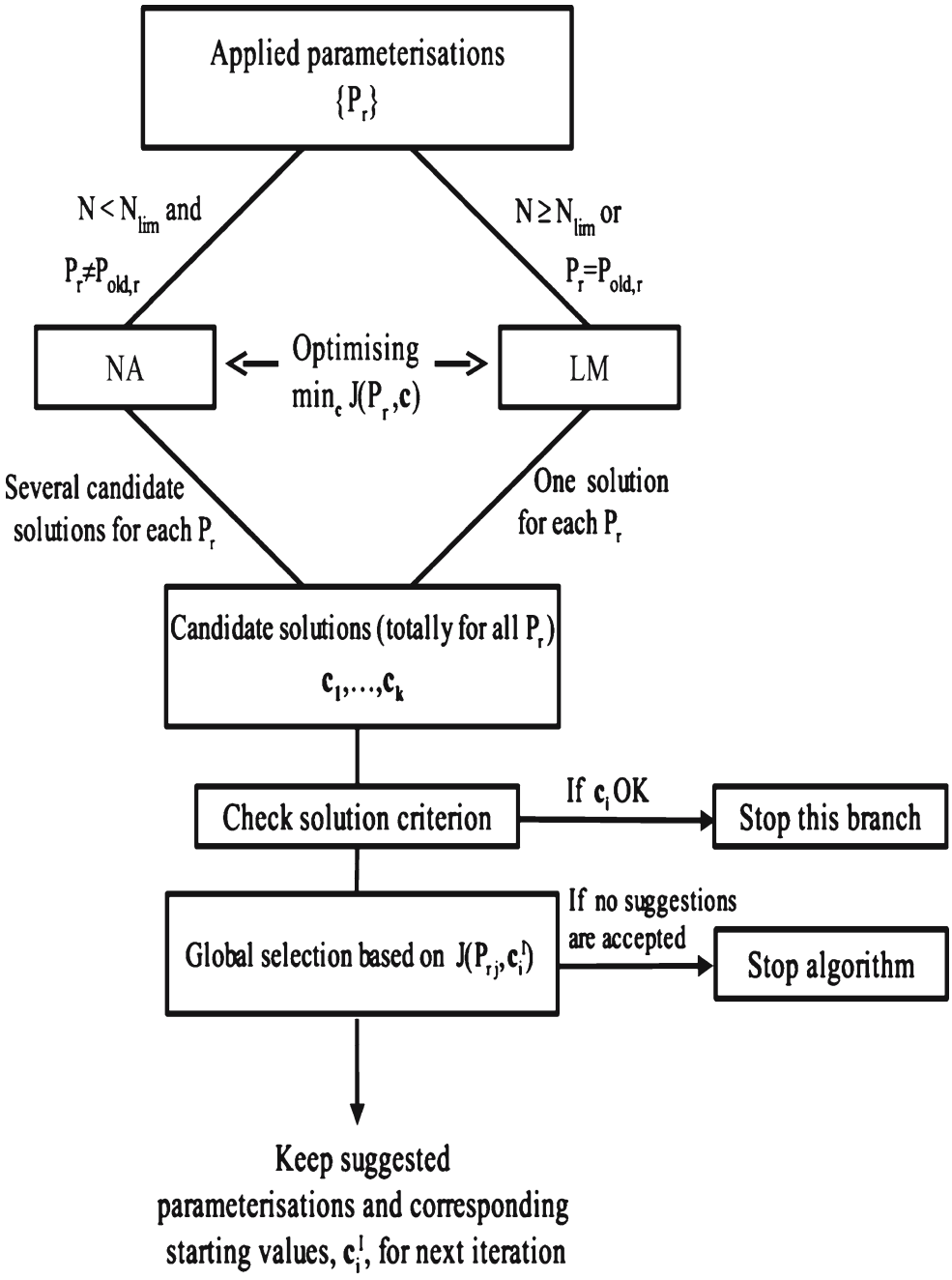

Figure 6 One iteration of the AME–NA algorithm. When N is greater or equal to a fixed dimension, N_{lim}, we change to the gradient optimiser (Levenberg–Marquardt). $P_{old,r}$ is the parameterisation from last iteration producing the new parameterisation P_r.

parameterisations $\{P^i\}_{i=1}^{20}$ (not all unequal) produced by Eq. (13).

In order to find the most promising parameterisation of the previously suggested refinements $\{P^i\}_{i=1}^{k}$ we define a new selection based on the same principles as the first one. The new *global optimal refinement* for this iteration is denoted by P^{*1} and is given by

$$P^{*1} = \text{optimal}_{P \in U_1} \tilde{J}(P) \text{ where } U_1 = \{P^i\}_{i=1}^{k}. \quad (15)$$

In the example in figure 5 there are four different new parameterisations among the set $\{P^i\}_{i=1}^{20}$, and it is assumed that Eq. (15) gives $P^{*1} = P^{14}$.

Often we want to start a new iteration with several new refinements in order to make the dynamic parameterisation selection more robust. This is achieved by eliminating all the parameterisations which are equal to one of the earlier produced P^*s in the ensemble of refinements $\{P^i\}_{i=1}^{k}$, and go through a new selection as in Eq. (15). In the special case where this resulting set of available refinement should be empty, we

choose new refinements among all the possible new parameterisations, $\{P_{rj}\}$, provided they are different from the already selected ones. For the next selection that will be

$$P^{*_2} = \text{optimal}_{P\in U_2}\tilde{J}(P), \tag{16}$$

$$\text{where } U_2 = \begin{cases} U_2{}' = \{\text{all } P^i \text{ unequal } P^{*_1}\} \text{ if } U_2{}' \neq \Phi \\ U_2{}'' = \{\text{all } P_{rj} \text{ unequal } P^{*_1}\} \text{ else.} \end{cases}$$

In the example in figure 5 this means that we eliminate $P^{11}, .., P^{18}$ when searching for the second most promising refinement. In this case we still have more unique suggested parameterisations produced by Eq. (13) and $U_2 = U_2{}'$.

The process can then, if wanted, be continued until there are no more unique parameterisations left in the set $\{P_{rj}\}$. A more compact notation for the global selection, indexing the global optimal refinements by q, is

$$P^{*_q} = \text{optimal}_{P\in U_q}\tilde{J}(P), \tag{17}$$

$$\text{where } U_q = \begin{cases} U_q{}' = \{P^i \ : \ P^i \neq P^{*_g} \\ \qquad \forall\, i, \forall\, g{=}1,...,q{-}1\} \qquad \text{if } U_q{}' \neq \Phi \\ U_q{}'' = \{P_{r,j} \ : \ P_{rj} \neq P^{*_g} \\ \qquad \forall\, r, \ \forall\, j, \ \forall\, g{=}1, ..., q{-}1\} \ \text{ else.} \end{cases}$$

The set of chosen new parameterisations, $\{P^{*_q}\}$, will be applied in the next stage, and new candidate solutions will be found by the optimiser. In the example the parameterisations passed to the next iteration are P^7 and P^{14}.

Another way of doing the selection of the refinements, which is notationally easier and in practice slightly different from this one, is to collect all P_{rj} for all $\mathbf{c}_i$, and apply Eq. (17) directly with $U_q = U_q{}''$. This approach will not, in contrast to the approach presented in this paper, force the algorithm to choose parameterisations based on different points in the parameter space if this is possible.

6.3. Ramification

The number of applied parameterisations for each stage does not have to be constant throughout the AME–NA sequence. In the beginning it can be advantageous to have several candidate refinements, but later, when the algorithm is approaching the solution, it may be sufficient to apply less parameterisations. This is due to the fact that in cases where the standard dynamic parameterisation selection fails, it is most likely that it has chosen an undesirable refinement at an early stage. This may lead to a higher number of parameters to obtain a parameterisation which will decrease the value of J. In some cases we will not be able to reduce J below the solution criterion, but for other cases the solution criterion will be reached with a (much) higher number of parameters than strictly required. Numerical experiments have shown that for both these cases the permeability values can sometimes be completely out of the expected range. See for example figures 14 and 15, where the solution criterion is fulfilled for the AME method, and figure 16, where it is not.

These circumstances make it advantageous to use more resources to avoid wrong selections of refinements in the first stages where it is most critical to make the right choices, and then constrict the possibilities of ramification when the number of parameters is higher.

Table 1 Properties for the simulations.

Reservoir dimensions	$1,600 \times 1,600 \times 20$ m
Simulation grid	$32 \times 32 \times 1$ cells
Porosity	0.3
Viscosity	$\mu_w = 1.0 \cdot 10^{-3}$ Pa · s
	$\mu_o = 1.3 \cdot 10^{-3}$ Pa · s
Relative permeability functions	$k_{rw} = S_w$
	$k_{ro} = 1 - S_w$
Initial saturation	$S_w = 0$
	$S_o = 1$
Capillary pressure function	$P_c(S_w) \equiv 0$ kPa
Total injection rate per well	$1,000$ m^3/day
Production rate	Constant BHP = 100.0 barsa
Number of observations per well	$n_{pw} = 400$
Observation times	$t = i\Delta t \qquad \Delta t = 20$ days,
	i=1, .., n_{pw}
Noise level	1 barsa

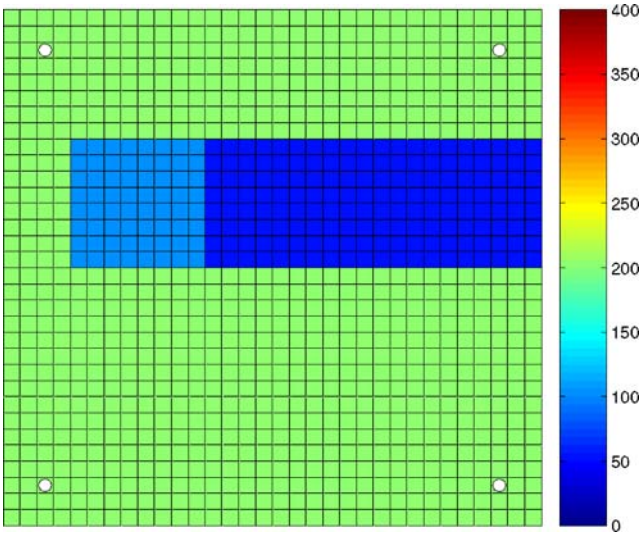

Figure 7 Base case (*upper right horizontal barrier*). The wells are marked with *white dots*.

Table 2 Input parameters for the NA algorithm.

N	1	2	4	5	6
K	3	4	6	8	10

n_s is chosen equal to $K \cdot N$ where K is an integer value, dependent of N, given in the table.
$n_r = n_s/2$ for all dimensions except for $N = 1$. In this case $n_r = 1$. Parameterisations where N is equal to 3 has, because the parameterisation selection never chose this dimension, not been required to optimise in this study.

Also, the optimisations becomes heavier and more time consuming the more parameters we use.

6.4. Change to gradient optimiser

The reasons given in Section 6.3 also make it reasonable to change to a gradient optimiser when the number of parameters is above some level. This is because the convergence is much faster for a gradient optimiser. If the selection of new refinements described in Section 6.2 suggests to retain a parameterisation we have also in this case, independent of the number of parameters, the gradient method is applied for (only) the next optimisation independent of the number of parameters. Figure 6 shows a schematic diagram of the stages in one iteration of the AME–NA algorithm. When we change to the gradient optimiser we apply the selection criterion from Section 6.2 to decide the starting, initial value, $\mathbf{c}_i^I$, given to the optimiser. This means we will not necessarily apply the parameters $\mathbf{c}_i$ giving the lowest value of $J(P_r)$, but those giving the lowest value of $\tilde{J}(P^{*_q})$. The $\tilde{J}$-values are based on a linear approximation and it therefore makes sense to apply the parameters predicting the lowest misfit as the initial point in a gradient optimiser.

7. Numerical results

In this section we will present some examples where we study the performance of the AME–NA approach, and compare with results achieved by the original AME method. We present results based on a synthetic reservoir field. The test reservoir is square and horizontal with constant thickness and no-flow outer boundaries. Except for the permeability, the fluid and rock properties are held fixed throughout the simulations. The properties for the simulations are listed in table 1.

To obtain pressure data observations we have run the simulator with a known true permeability field, and then added uncorrelated normally distributed noise to the logged pressure values. The noise level is represented by the standard deviation, σ, of the noise distribution, given in table 1.

For the studied examples we have either printed the resulting permeability field from all (or most of) the iterations or in some cases only the final result.

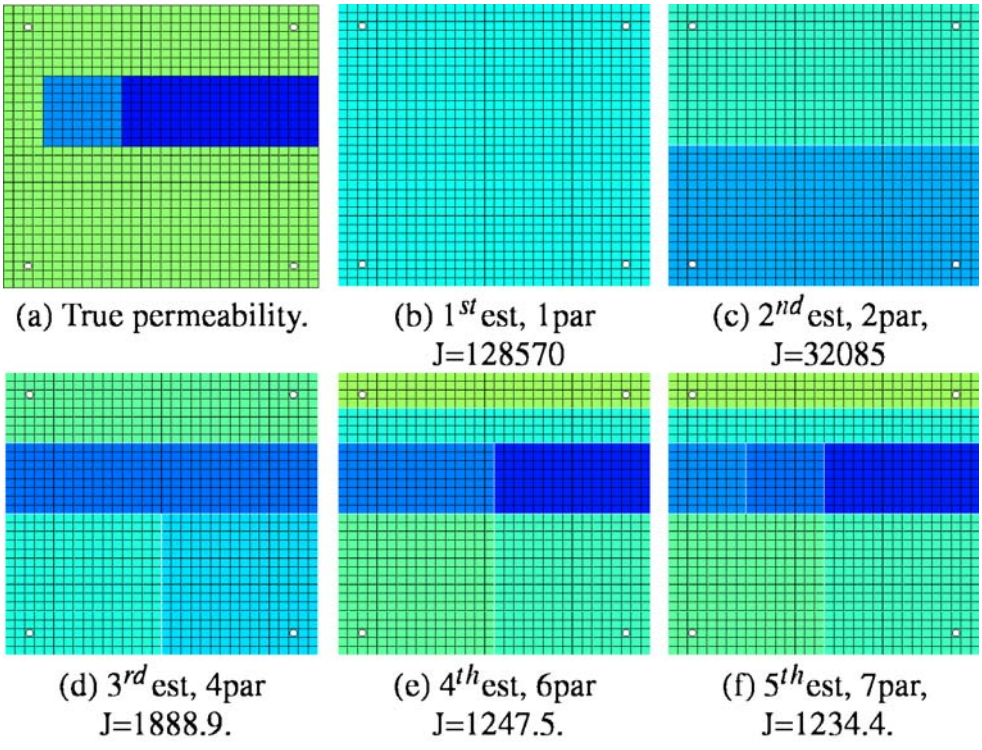

Figure 8 Example 1: *Upper right horizontal barrier* (base case). Results using the AME method. Seven parameters are required to solve the inverse problem.

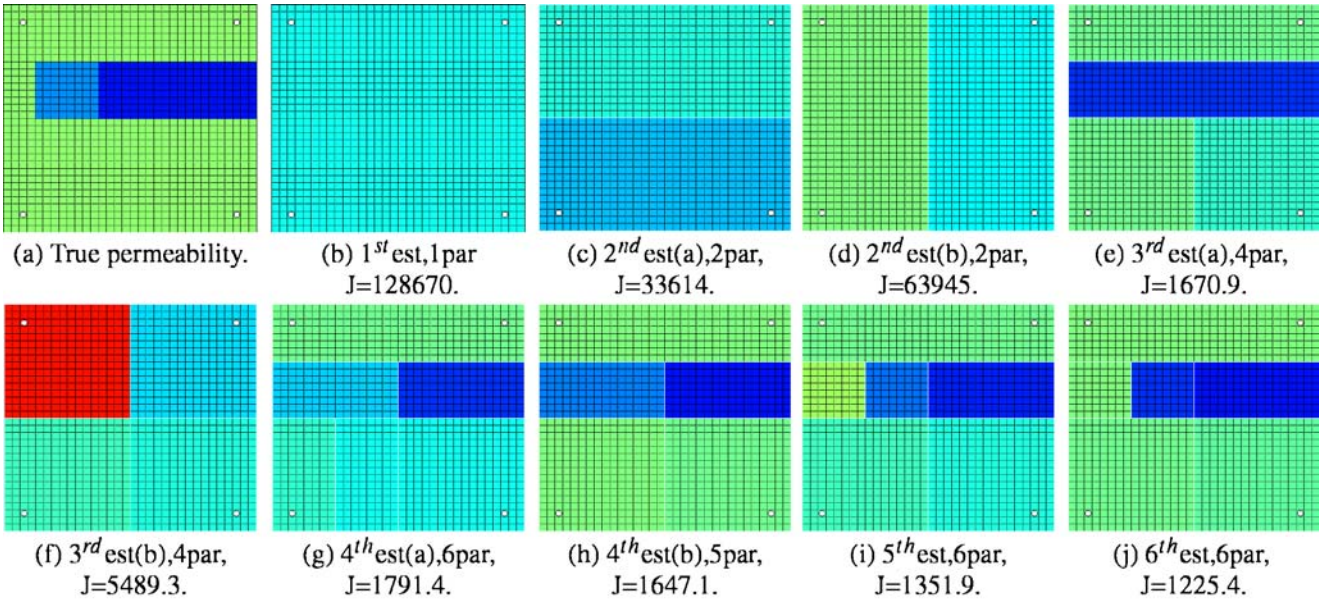

Figure 9 Example 1: *Upper right horizontal barrier* (base case) using AME–NA. From the first estimate both the allowed refinements are suggested as new parameterisations. These two cases are optimised independently of each other, and referred to as 2^{nd}est(a) and 2^{nd}est(b). In the next step, estimate 3, one refinement from each of the parameterisations in estimate 2 are selected as the most promising ones. From estimate 3 till estimate 4 we get a case where both new parameterisations are refinements of 3^{rd}est(a). This means that no refinement of 3^{rd}est(b) is applied, and the branch giving this parameterisation is eliminated in the further process. When starting step 5 we restrict the ramification such that we only select one optimal new parameterisation, which in this case is a refinement of estimate 4^{th}est(b). In the 6^{th} estimate no refinement of the last parameterisation is done, and that means we apply the gradient optimiser in this case. With six parameters the value of J is below the solution criterion and the inverse problem is solved.

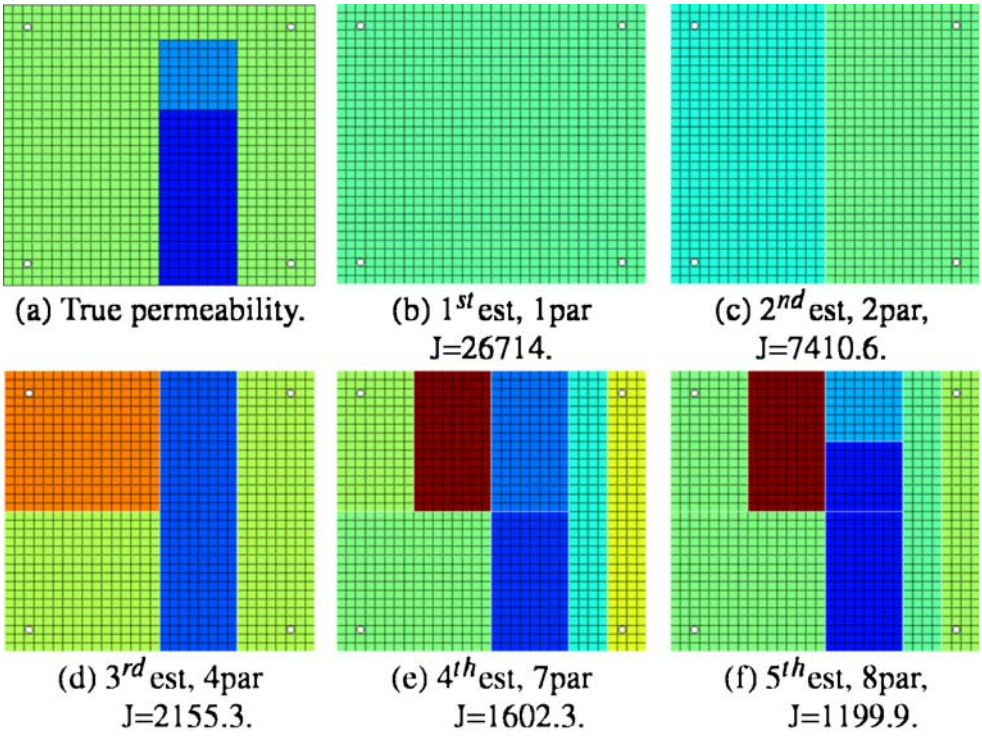

Figure 10 Example 2: *Right lower vertical barrier* (otherwise equal properties as in the base case). Results using AME.

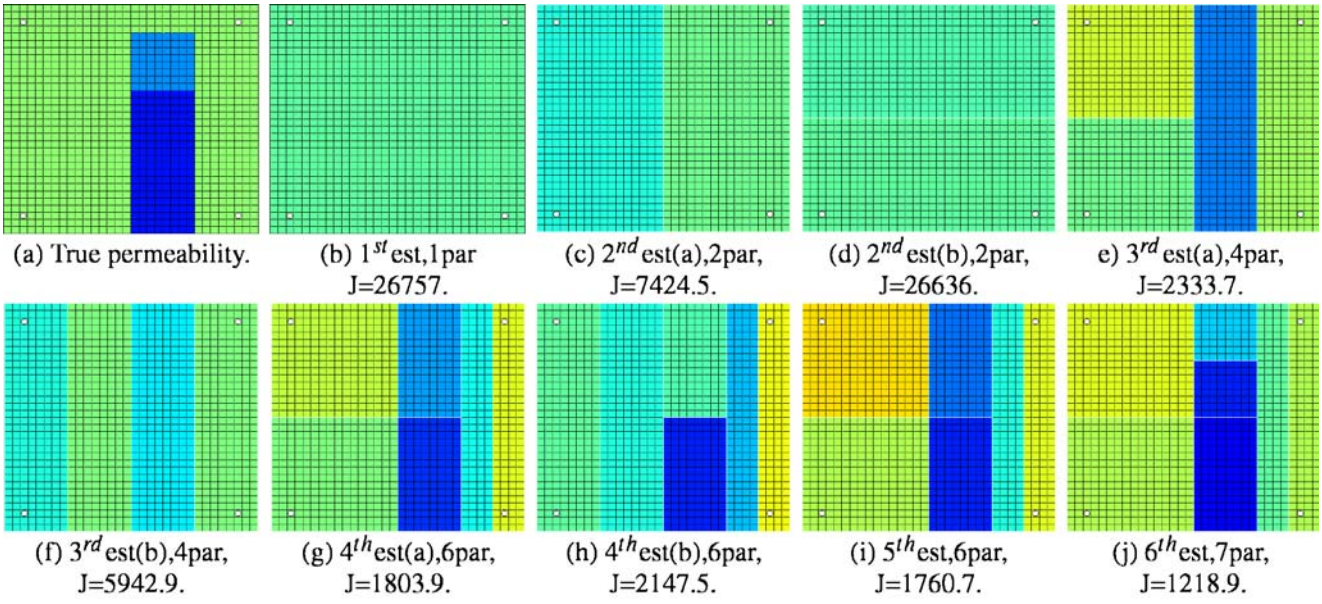

Figure 11 Example 2: *Right lower vertical barrier* (otherwise equal properties as in the base case). Results using AME–NA.

Under each plot the iteration number in the sequential parameterisation selection (*est*), number of parameters (*par*) and the actual value of the objective function (J) is given.

The chosen examples have a focus on cases where the AME method is clearly unsuccessful, but will also include successful AME sequences. We start out with a field referred to as the *base case* (figure 7). A wide range of the other studied examples are perturbations of this field. The base case has four wells, marked with white dots, positioned in the corners of the reservoir. The well in the upper right corner is a production well, and the remaining three are used for injection. In the field there are three discrete values of the permeability, which are equal to 50 mD (dark blue), 100 mD (light blue) and 200 mD (green). Unless otherwise stated, the properties for the studied fields will be identical to the properties for the base case.

In the studies we select ten candidate solutions from each NA optimisation and have chosen to change to the gradient optimiser when the dimension of the parameter space, N, is greater or equal to 7 (that is, $N_{lim} = 7$ in figure 6). The ramification is limited to two possible parameterisations up till the fourth iteration, and from the fifth one we optimise for only one parameterisation in each AME–NA stage (see figure 9 for an example).

The input values for the Neighbourhood Approximation search, n_s', n_s and n_r, are chosen dependently of the dimension of the parameter space. An explanation of how n_s and n_r are chosen is given under table 2. To

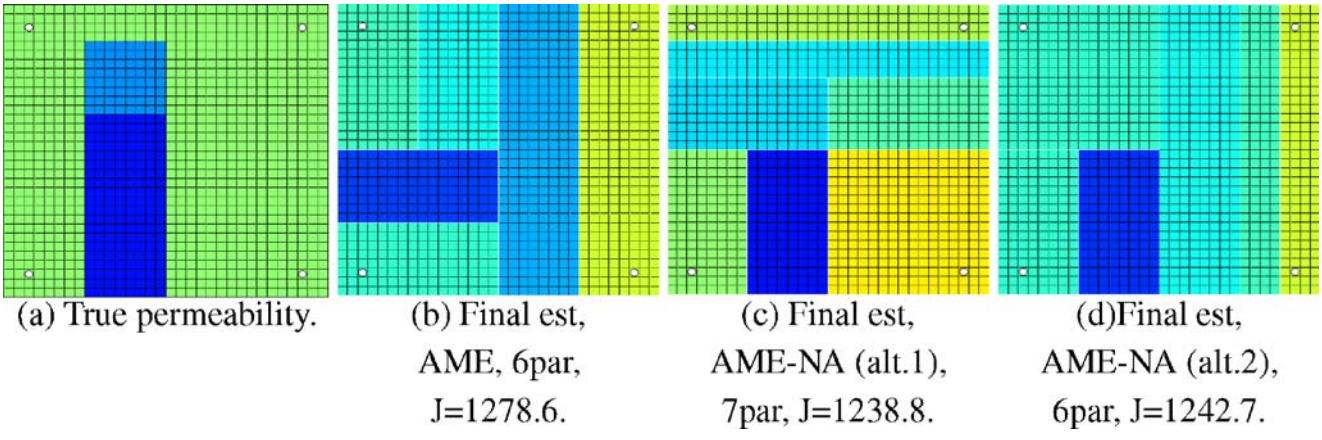

Figure 12 Example 3: *Left lower vertical barrier* (otherwise equal properties as in the base case). Results for both methods. The solution criterion is not reached when applying the AME method. AME–NA produces two different solutions (because of the ramification), both solving the inverse problem.

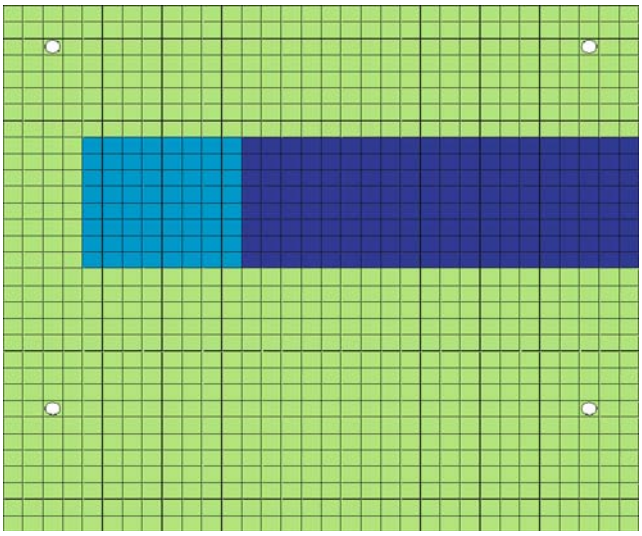

Figure 13 In this example the two lowest wells (marked with *white dots*) are moved up compared to the base case.

make the search more robust the algorithm is run twice with a slightly different number of initial samples for the two cases. The initial sampling, n'_s, is equal to or slightly lower than n_s for all dimensions.

The widest range of the applied true permeability fields is values from 10 to 200 mD. In the Neighbourhood Approximation the sampling is done for parameters corresponding to permeability values between 0 and 400 mD, which are also the range used in the coloured plots. In the plots dark blue colour corresponds to low values while dark red is high values of the permeability.

7.1. Location of barriers

In the first three examples we will study the performance of recovering a field with a simple barrier. We will test the base case and two similar fields where the barrier is rotated relative to the wells.

7.1.1. *Example 1: Upper right horizontal barrier (base case)*

This is an example where the AME method approximately identifies the true permeability (figure 8), and the AME–NA algorithm is giving an almost equal result using one parameter less (figure 9). Be aware that the optimisations are stopped when reaching the solution criterion in Eq. (7), and that the optimiser is not necessarily run till convergence. Because of this, it may be incidental that the result from the AME–NA method is slightly better looking than the result from the AME approach.

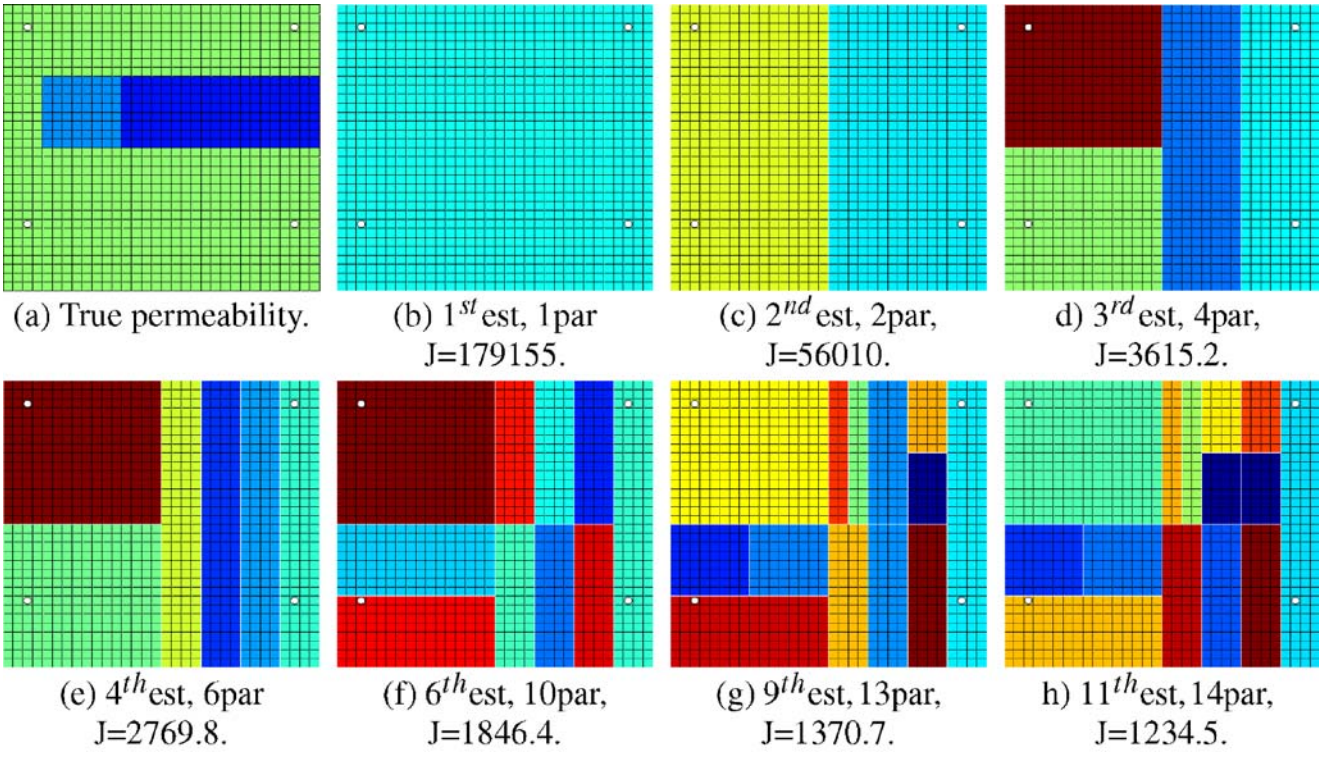

Figure 14 Example 4: *Lower wells* moved up (compared to the base case). Results from the AME approach.

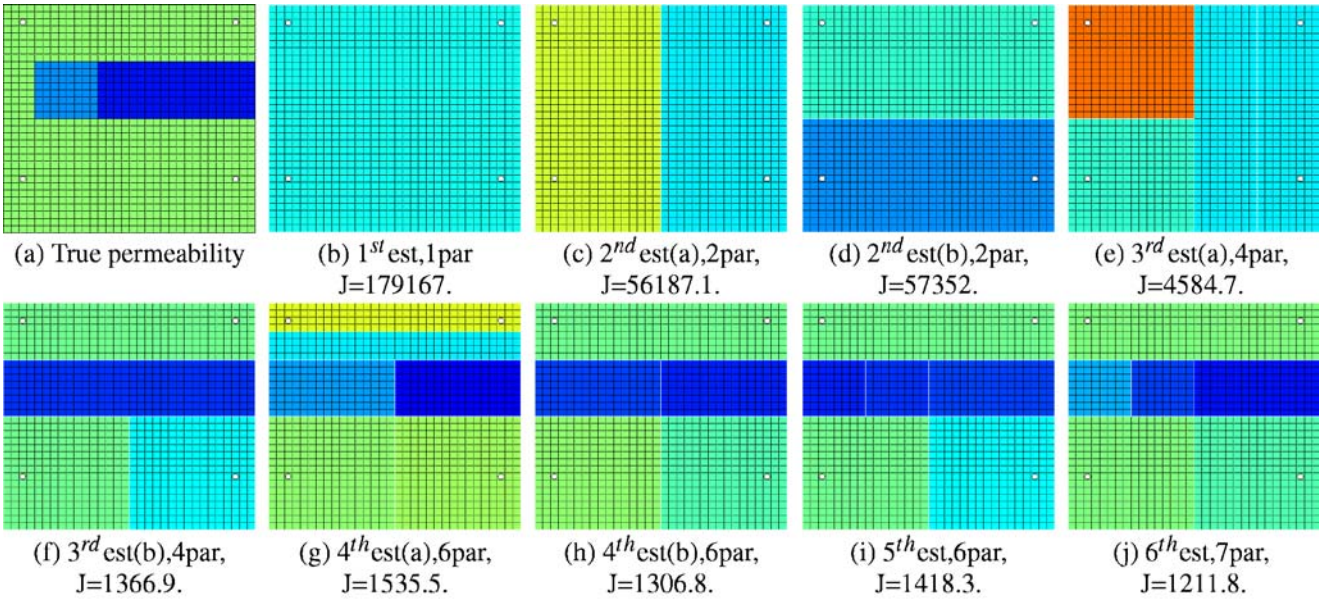

Figure 15 Example 4: *Lower wells* moved up (compared to the base case). Results from the AME–NA approach.

7.1.2. *Example 2: Right lower vertical barrier*

Also in this case the inverse problem is solved with both methods and with one less parameter for the AME–NA algorithm. The resulting permeability field produced by AME (figure 10) gives one parameter completely out of range compared to the true field applied to produce the measurements. In the real case, when the true field is unknown, it had been non-trivial to determine which solution is the best one (figure 11).

7.1.3. *Example 3: Left lower vertical barrier*

In this example the solution criterion (which is equal to 1,243 for six parameters) is not completely fulfilled using the AME method, but for the new approach we have two different solutions solving the inverse problem. The reason for *two* suggested solutions from the AME–NA method is that one of them is produced at an early stage before we restrict the amount of ramifications to allow only one branch (figure 12).

7.2. Displacement of wells

We will present one example where we have applied an equal permeability field as in the base case, but changed the position of two of the wells.

7.2.1. *Example 4: Lower wells moved up*

In this example the two lowest wells in the field are moved up such that they are not that close to the corners any more, see figures 13 and 14. The AME–NA algorithm (figure 15) produces a solution close to the true field, while the result from the AME approach (figure 14) gives a completely different estimate. Even

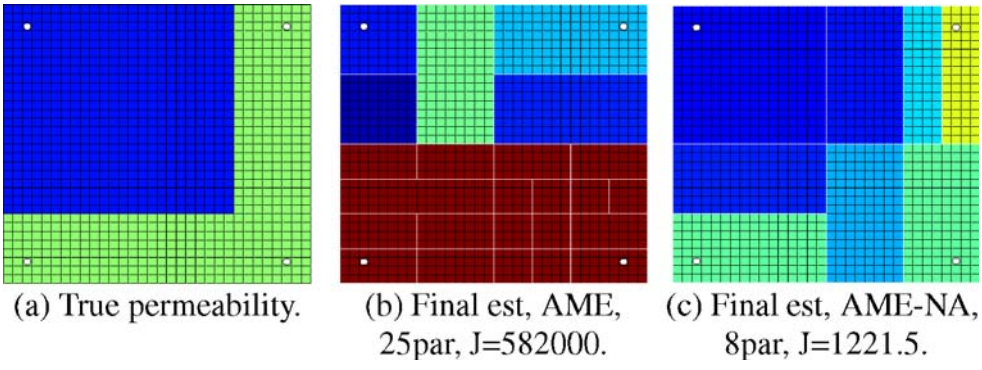

Figure 16 Example 5: *Lower right channel* (wells as in the base case). Results for both methods. Only the AME–NA method solves the inverse problem.

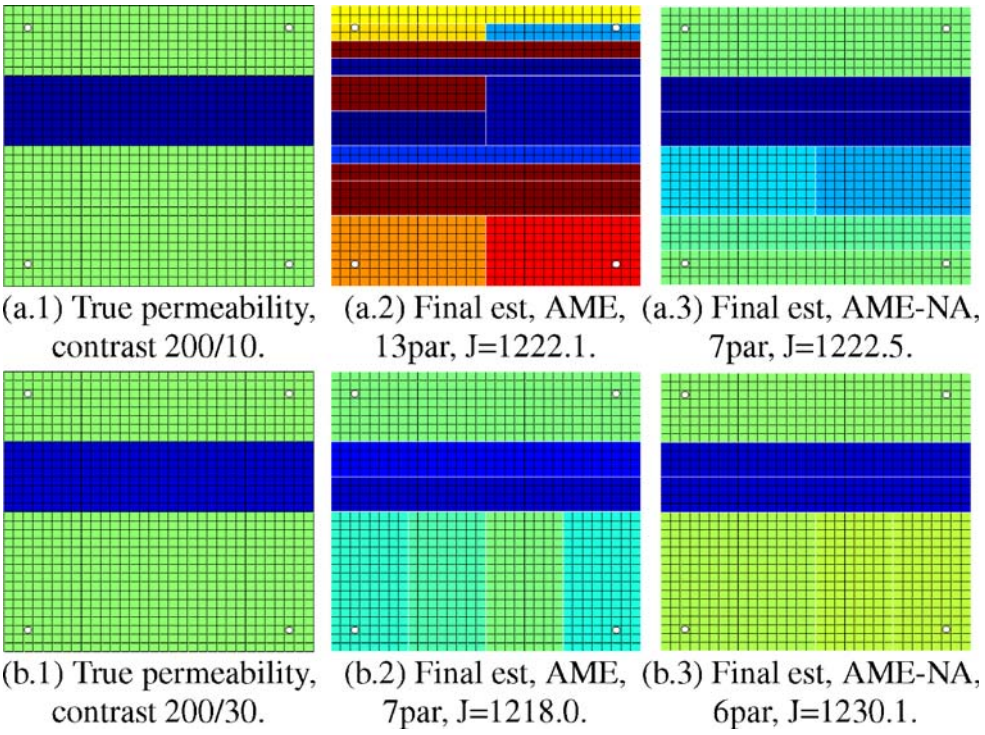

Figure 17 Example 6a and 6b: *Long barrier* with different permeability contrasts (wells as in the base case).

if the AME result is far from the expected solution it is able to solve the inverse problem, but it requires a larger number of parameters and there are large oscillations on small scales for this solution.

In Section 4 we discussed the desired properties of the solution using the dynamic parameterisation technique. Because the information in the data may notbe suitable for reproducing small scale variations, the applied methods seek a solution which reproduces the coarse structures of the field. Results like this substantiate the hypothesis that the uncertainties are high when there is a large number of small scale parameters in the estimate.

7.3. High permeable channel

We have tested one example where we have applied a simple channel with higher permeability than the rest of the reservoir. The permeability values in the true field are 200 mD for the channel and 50 mD for the rest of the field.

7.3.1. *Example 5: Lower right channel*

In this case we got completely different results using the two methods (figure 16). The AME estimation introduces a large number of parameters and the optimisation problems become difficult to solve with the current starting values. This results in producing a local minimum of J which is much higher than the solution criterion. The AME–NA algorithm is able to reduce the misfit value, and gets a permeability estimate which is close to a high permeable channel in the same area as the true field.

7.4. Permeability contrasts

In the next examples we have tested the performance of the two algorithms on two different fields where we vary the permeability contrasts between the barrier and the rest of the field.

7.4.1. *Example 6a and 6b: Different contrast on fields with a longer barrier*

In Example 6a (figure 17a.1–a.3) and 6b (figure 17b.1–b.3) we have a long horizontal barrier across the complete field, and have used different contrast between the two discrete permeability values. Both methods solve the inverse problem for the two test fields, but the result from the AME method in example 6a (figure 17a.2) is not a coarse scale estimate and contains small scale variations. In this case a parameterisation which is clearly able to reproduce the true field is obtained, but the found minimum of J is not corresponding to a field close to the true permeability.

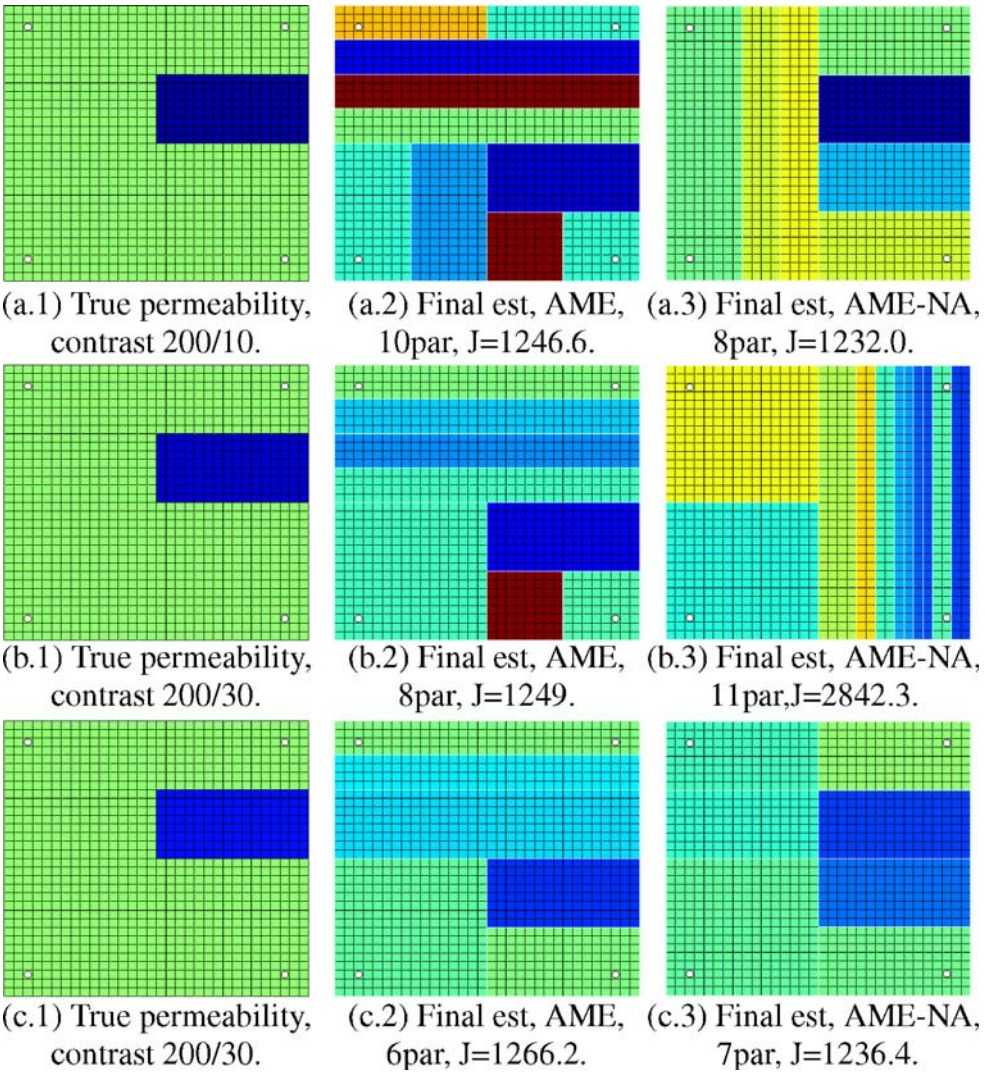

Figure 18 Example 7a–7c: *Short barrier* with different permeability contrasts (wells as in the base case).

7.4.2. *Example 7a–7c: Different contrasts on fields with a short barrier*

The next three examples are also a study where we have varied the contrast between two permeability values. In this case a shorter barrier is used. None of the results from the AME method solves the inverse problem according to our solution criterion, even though they are close. The AME–NA algorithm also fails in the middle case (figure 18b.3), and has in this case introduced a large number of unnecessary parameters.

7.5. Continuous fields

The last two examples are fields with more continuously varying permeability. Neither AME nor AME–NA should be expected to reproduce a field like this exactly, but if the algorithm is successful a coarse scale approximation could be obtained.

7.5.1. *Example 8a and 8b: Continuous fields*

For these examples it is discussable which method is giving the preferable result. In the first example (figure 19a.1–a.3) both methods solve the inverse problem with approximately equal number of parameters. The AME result is though slightly better looking. The last example (figure 19b.1–b.3) gives permeability plots which are approximately equally close to the true field, but in this case the AME method is not completely reducing J to the limit for solving the problem.

8. Conclusion

We have demonstrated a methodology for improving the robustness in deriving coarse scale permeability fields using the Adaptive Multiscale Estimation algorithm. The methodology uses the Neighbourhood

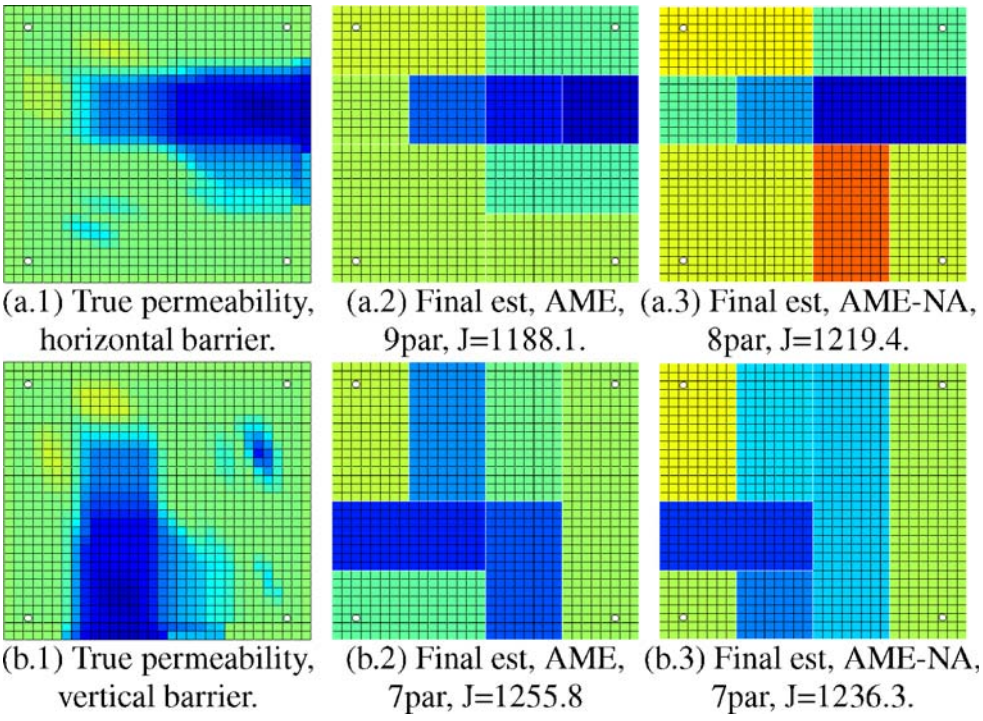

Figure 19 Example 8a and 8b: *Continuous field* with barriers (wells as in the base case).

Approximation algorithm to generate several candidate solutions. The spread in the candidate solutions from the NA algorithm results in:

- more reliable suggestions of the next refinement.
- larger possibilities for ramifications, which give several ways to approach the solution.

Our results show that while it is possible to solve the optimisation problem at low dimensions, this becomes CPU-intensive for higher dimensions. For higher dimensions the gradient algorithm ensures faster convergence.

The same selection criterion which is applied to find optimal refinements, is also used to select starting values for the gradient optimiser. This gives us a way of combining the stochastic and gradient optimiser in order to take advantage of strong parts from both methods.

The numerical results presented in this paper involve mostly pathological cases for the standard AME method. The AME–NA algorithm considerably reduced the number of such cases.

The new algorithm usually introduces fewer parameters to solve the inverse problem than the case for the AME method. In this way the risk of over-parameterisation is reduced and the problem is better conditioned. Totally we have a more robust way of solving the inverse problem.

9. Remarks and future work

The ramification part of this work could have been achieved by using a gradient optimiser producing *one* solution for each optimisation through the complete process. The same selection of refinement as described in this paper could then have been applied to get a limited ramification. In some cases this might have improved the final estimates compared to standard AME. An investigation of this approach, and comparison of the robustness with AME and AME–NA have not been performed.

The methodology from this paper can be applied to do uncertainty assessments of a solution by keeping several solutions from each stage and optimise for each suggested parameterisation. To get reliable uncertainty estimates we need a wider ramification than used in this work.

A more explicit study of the uncertainty related to fine scale variations of the permeability is also desirable to improve our knowledge of different kinds of solutions of the inverse problem studied in this paper.

In this work we have focused on producing a coarse scale estimate of the permeability. There is a demand

for combining this type of estimate with geostatistical information. In the literature there have been proposed different approaches for including geostatistical information in the optimisation process related to this problem, see, for example, [25–28]. It could be interesting to combine these types of methodologies with the presented method.

Acknowledgments This work is done with support from European Community and access to research infrastructure at Heriot Watt University, Institute of Petroleum Engineering. The work is also a part of a RF-Rogaland Research program within inverse problems. This program is financed by ENI SpA, Norsk Hydro ASA and the Norwegian Research Council. We gratefully acknowledge the support.

Appendix

In this appendix we will for completeness reproduce the derivation of formulas (9) and (10), as was performed in [2].

Assume that we want to evaluate refinements to the parameterisation after terminating the estimation with P_N at some parameter value $\mathbf{c}_N^F$. An approximation of the minimum of J (see Eq. (6)) for an extended parameterisation $P_{N'}$ can be found by linearisation of $\mathbf{m}(\mathbf{c}_{N'}^I)$,

$$\mathbf{m}(\mathbf{c}_{N'}) \approx \mathbf{m}(\mathbf{c}_{N'}^I) + A_{N'}\Delta\mathbf{c}_{N'}, \tag{18}$$

where $A_{N'} = \mathbf{m}'(\mathbf{c}_{N'}^I)$ and $\mathbf{c}_{N'} = \mathbf{c}_{N'}^I + \Delta\mathbf{c}_{N'}$, for some $\Delta\mathbf{c}_{N'}$.

Recall that $\Delta\mathbf{d}=\mathbf{d}-\mathbf{m}(\mathbf{c}_{N'}^I)$. Substituting (18) into (6) gives

$$J(\mathbf{c}_{N'}) \approx \left[\Delta\mathbf{d} - A_{N'}\Delta\mathbf{c}_{N'}\right]^T D^{-1}\left[\Delta\mathbf{d} - A_{N'}\Delta\mathbf{c}_{N'}\right]. \tag{19}$$

Using this approximation of $J(\mathbf{c}_{N'})$ it follows (by computing its minimum) that Eq. (9),

$$\tilde{J}(P_{N'})=\Delta\mathbf{d}^T\left(D^{-1}-\left[D^{-1}A_{N'}(A_{N'}^TD^{-1}A_{N'})^{-1} \times A_{N'}^TD^{-1}\right]\right)\Delta\mathbf{d}, \tag{20}$$

may be used as a measure for the objective function value attainable with the parameterisation $P_{N'}$.

In the following part we will derive an uncertainty estimate for $\tilde{J}$ which motivates Eq. (10). We assume that the model is correct for the unknown true set of parameters $\mathbf{c}^*$, i.e., that we can write

$$\mathbf{d} = \mathbf{m}(\mathbf{c}^*) + \mathbf{e}, \tag{21}$$

with normally distributed measurement errors $\mathbf{e}$ with expectation zero and covariance D, $\mathbf{e} \sim N_M(0, D)$. Then, at a stage in the estimation sequence, with parameters $\mathbf{c}_{N'}$ we have a model error (also unknown)

$$\mathbf{z}_{N'} = \mathbf{m}(\mathbf{c}^*) - \mathbf{m}(\mathbf{c}_{N'}). \tag{22}$$

Using this notation, we can write the data misfit vector as a sum of the two unknowns:

$$\Delta\mathbf{d} = \mathbf{d} - \mathbf{m}(\mathbf{c}_{N'}) = \mathbf{m}(\mathbf{c}^*) + \mathbf{e} - \mathbf{m}(\mathbf{c}_{N'}) = \mathbf{z}_{N'} + \mathbf{e}. \tag{23}$$

Since a covariance matrix is symmetric, we can find a matrix V so that $D^{-1} = V^TV$. This enables us to write the expression for $\tilde{J}$ as

$$\begin{aligned}\tilde{J}(P_{N'}) &= \Delta\mathbf{d}^TV^T\Big(I_M - VA_{N'}\left((VA_{N'})^T(VA_{N'})\right)^{-1} \\ &\quad\times(VA_{N'})^T\Big)V\Delta\mathbf{d} \\ &= (V\Delta\mathbf{d})^T(I_M - P_{VA})(V\Delta\mathbf{d}),\end{aligned} \tag{24}$$

where I_M is the $M \times M$ identity matrix, and P_{VA} is a projection matrix into the subspace spanned by the columns of $VA_{N'}$. Because of the measurement errors $\mathbf{e}$, the predicted attainable objective function value is a random variable. Since $\mathbf{e} \sim N_M(0, D)$, we have, using the definition (23):

$$\begin{aligned}V\Delta\mathbf{d} &= V\mathbf{z}_{N'} + V\mathbf{e} \sim N_M(V\mathbf{z}_{N'}, VD^{-1}V^T) \\ &= N_M(V\mathbf{z}_{N'}, I_M).\end{aligned} \tag{25}$$

It follows (using Theorem B.4 in [10]) that $\tilde{J}$ will have a non-central χ^2 distribution with $\mathrm{tr}(I_M - P_{VA})$ degrees of freedom and non-centrality parameter $\mathbf{z}_{N'}^TV^T(I_M - P_{VA})V\mathbf{z}_{N'}$:

$$\tilde{J}(P_{N'}) \sim \chi^2_{\mathrm{tr}(I-P_{VA})}(\mathbf{z}_{N'}^TV^T(I_M - P_{VA})V\mathbf{z}_{N'}). \tag{26}$$

If there are no redundant parameters, $\mathrm{tr}(P_{VA}) = N'$, and the expectation of $\tilde{J}$ is simply

$$E(\tilde{J}(P_{N'})) = \mathbf{z}_{N'}^TV^T(I_M - P_{VA})V\mathbf{z}_{N'} + (M - N'), \tag{27}$$

although this expression cannot be used to calculate $E(\tilde{J}(P_{N'}))$ since $\mathbf{z}_{N'}$ is unknown.

To calculate the variance of $\tilde{J}$, we use the result $\mathrm{Var}(\boldsymbol{\epsilon}^T B\boldsymbol{\epsilon}) = 2\mathrm{tr}(B^2)$, valid when $\boldsymbol{\epsilon} \sim N_n(0, I_n)$ (Appendix A12 in [29]), and $(I_M - P_{VA})^2 = (I_M - P_{VA})$, valid for all projection matrices. We have

$$\begin{aligned}
\mathrm{Var}(\tilde{J}(\mathbf{c}_{N'}^I)) &= \mathrm{Var}\big[(V\mathbf{z}_{N'}+V\mathbf{e})^T\{I_M-P_{VA}\}(V\mathbf{z}_{N'}+V\mathbf{e})\big] \\
&= \mathrm{Var}\big[(V\mathbf{z}_{N'})^T(I_M-P_{VA})(V\mathbf{z}_{N'})+2(V\mathbf{z}_{N'})^T \\
&\quad \times (I_M-P_{VA})(V\mathbf{e})+(V\mathbf{e})^T(I_M-P_{VA})(V\mathbf{e})\big] \\
&= \mathrm{Var}[2(V\mathbf{z}_{N'})^T(I_M-P_{VA})(V\mathbf{e})]+\mathrm{Var}[(V\mathbf{e})^T \\
&\quad \times (I_M-P_{VA})(V\mathbf{e})] = 4(V\mathbf{z}_{N'})^T(I_M-P_{VA})^T \\
&\quad \times \mathrm{Cov}(V\mathbf{e})(I_M-P_{VA})(V\mathbf{z}_{N'}) \\
&\quad + 2\mathrm{tr}((I_M-P_{VA})^2) = 4(V\mathbf{z}_{N'})^T(I_M-P_{VA}) \\
&\quad \times (V\mathbf{z}_{N'})+2\mathrm{tr}(I_M-P_{VA}) = 4(E(\tilde{J}(P_{N'})) \\
&\quad - (M-N'))+2(M-N') \\
&= 4E(\tilde{J}(P_{N'}))-2(M-N'). \qquad (28)
\end{aligned}$$

To obtain the third equality, we have utilised that $\mathrm{Cov}(\mathbf{v}^T B\boldsymbol{\epsilon}, \boldsymbol{\epsilon}^T B\boldsymbol{\epsilon}) = 0$ when $\boldsymbol{\epsilon} \sim N(0, I_N)$.

As in (27), the expression for $E(\tilde{J})$ contains the unknown model error, so the variance may not be calculated directly. Instead, by assuming that the actual predicted attainable objective function value does not differ too much from its expectation, we perform the substitution $E(\tilde{J}) \approx \tilde{J}$ in (28), and get the approximate expression for the variance:

$$\mathrm{Var}(\tilde{J}(P_{N'})) \approx 4\tilde{J}(P_{N'}) - 2(M-N'). \qquad (29)$$

From this we get our estimate for the uncertainty in $\tilde{J}$:

$$\tilde{\sigma}(\tilde{J}(P_{N'})) = \sqrt{4\tilde{J}(P_{N'}) - 2(M-N')}, \qquad (30)$$

which is given in Eq. (11).

References

1. Duijndam, A.J.W.: Bayesian estimation in seismic inversion. Part I. Principles. Geophys. Prospect. **36**, 878–898 (1998)
2. Grimstad, A.-A., Mannseth, T., Nævdal, G., Urkedal, H.: Adaptive multiscale permeability estimation. Comput. Geosci. **7**(1), 1–25 (2003)
3. Grimstad, A.-A., Mannseth, T., Nævdal, G., Urkedal, H.: Scale splitting approach to reservoir characterization. In: SPE (Society of Petroleum Engineers) 66394, presented at the SPE Reservoir Simulation Symposium, Huston, Texas, February 2001
4. Yoon, S., Malallah, A., Datta-Gupta, A., Vasco, D.W., Behrens, R.A.: A multiscale approach to production data integration using streamline models. In: Proc. of 1999 SPE Annual Technical Conference and Exhibition, Houston, Texas, October 1999. (SPE 56653)
5. Subbey, S., Christie, M., Sambridge, M.: Uncertainty reduction in reservoir modeling. SIAM Con. Math., pp. 457–467 (2002)
6. Sambridge, M., Mosegaard, K.: Monte Carlo methods in geophysical inversion problems. Rev. Geophys. **40**, 3–29 (2002)
7. Aurenhammer, F.: Voronoi diagrams: a survey of a fundamental geometric data structure. ACM Comput. Surv. **23**, 345–405 (1991)
8. Christie, M., Subbey, S., Sambridge, M.: Prediction under uncertainty in reservoir modeling. In: Proc. of 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany, 3–6 September 2002
9. Subbey, S., Christie, M., Sambridge, M.: A strategy for rapid quantification of uncertainty in reservoir performance prediction. In: Proc of SPE Reservoir Simulation symposium, pp. 1–12, Houston, Texas, 3–5 February 2003. (SPE 79678)
10. Sen, A., Srivastava, M.: Regression Analysis; Theory, Methods, and Applications. Springer, Berlin Heidelberg New York (1990)
11. Grimstad, A.-A., Mannseth, T., Aanonsen, S.I., Aavatsmark, I., Cominelli, A., Mantica, S.: Identification of unknown permeability trends from history matching of production data. Soc. Pet. Eng. J. **9**(4), 419–428 (Dec 2004)
12. Berre, I., Lien, M., Mannseth, T.: A level set corrector to an adaptive multiscale permeability prediction. Submitted 2004
13. Krüger, H., Mannseth, T.: Extension of the parameterization choices in adaptive multiscale permeability estimation. Inverse Probl. Sci. Eng. **13**, 469–484 (2005)
14. Ben Ameur, H., Chavent, G., Jaffre, J.: Refinement and coarsening indicators for adaptive parameterization: application to the estimation of hydraulic transmissivities. Inverse Probl. **18**, 775–794 (2002)
15. Liu, J.: A multiresolution method for distributed parameter estimation. SIAM J. Sci. Comput. **14**(2), 398–405 (March 1993)
16. Chavent, G., Liu, J.: Multiscale parameterization for the estimation of a diffusion coefficient in elliptic and parabolic problems. In: Fifth IFAC Symposium on Control of Distributed Parameter Systems, Perpignian, France, June 1987
17. Chavent, G., Kunisch, K.: New results on the non-linearity and the sensitivity of the estimation of the diffusion coefficient in a 2D elliptic equation. In: Proc. 3rd International Conference on Inverse Problems in Engineering: Theory and Practice, Port Ludlow, Washington, USA, June 1999
18. Grimstad, A.-A., Mannseth, T.: Nonlinearity, scale, and sensitivity for parameter estimation problems. SIAM J. Sci. Comput. **21**(6), 2096–2113 (2000)
19. Grimstad, A.-A., Kolltveit, K., Mannseth, T., Nordtvedt, J.-E.: Assessing the validity of a linearized accuracy measure for a nonlinear parameter estimation problem. Inverse Probl. **17**, 1373–1390 (2002)
20. Cheng, H., Datta-Gupta, A., He, Z.: A comparison of travel time and amplitude matching for field-scale production data integration: Sensitivity, non-linearity and practical implications. In: Proc. of SPE Annual Technical Conference and Exhibition, Denver, Colorado, October 2003
21. Morè, J.J.: The Levenberg–Marquardt algorithm: Implementation and theory. In: Watson, G.A. (ed.) Numerical Analysis; Lecture Notes in Mathematics, vol. 630, pp. 105–116. Springer, Berlin Heidelberg New York (1997)

22. Sambridge, M.: Exploring multidimensional landscape without a map. Inverse Probl. **14**, 427–440 (1998)
23. Sambridge, M.: Geophysical inversion with a neighbourhood algorithm. I. Searching the parameter space. Geophys. J. Int. **138**, 479–494 (1999)
24. Sambridge, M.: Geophysical inversion with a neighbourhood algorithm. II. Appraising the ensemble. Geophys. J. Int. **138**, 727–746 (1999)
25. Grimstad, A.-A., Mannseth, T.: Comparison of methods for downscaling of coarse scale permeability estimates. In: Proc. of 9th European Conf. on the Mathematics of Oil Recovery, Cannes, France, 30 August–2 September 2004
26. Sahni, I., Horne, R.N.: Generating multiple history-matched reservoir model realizations using wavelets. In: Proc. of SPE Annual Technical Conference and Exhibition, Houston, Texas, 26–29 September 2004
27. Aanonsen, S.I.: Efficient history matching using a multiscale technique. In: Proc. of SPE Reservoir Simulation Symposium, Houston, Texas, 31 January–2 February 2005
28. Aanonsen, S.I., Eydinov, D.: A multiscale method for distributed parameter estimation with application to reservoir history matching. Comput. Geosci. (March 2006)
29. Seber, G.A.F., Wild, C.J.: Nonlinear Regression. Wiley, New York (1989)