Adaptive Multiscale Streamline Simulation and Inversion for High-Resolution Geomodels
V. R. Stenerud, SPE, Norwegian U. of Science and Technology; V. Kippe, SINTEF ICT; K.-A. Lie, SINTEF ICT; and A. Datta-Gupta, SPE, Texas A&M U.

Abstract
A particularly efficient reservoir simulator can be obtained by combining a recent multiscale mixed finite-element flow solver with a streamline method for computing fluid transport. This multiscale-streamline method has shown to be a promising approach for fast flow simulations on high-resolution geologic models with multimillion grid cells. The multiscale method solves the pressure equation on a coarse grid while preserving important fine-scale details in the velocity field. Fine-scale heterogeneity is accounted for through a set of generalized, heterogeneous basis functions that are computed numerically by solving local flow problems. When included in the coarse-grid equations, the basis functions ensure that the global equations are consistent with the local properties of the underlying differential operators. The multiscale method offers a substantial gain in computation speed, without significant loss of accuracy, when basis functions are updated infrequently throughout a dynamic simulation.

In this paper we propose to combine the multiscale-streamline method with a recent ‘generalized travel-time inversion’ method to derive a fast and robust method for history matching high-resolution geo-cellular models. A key point in the new method is the use of sensitivities that are calculated analytically along streamlines with little computational overhead. The sensitivities are used in the travel-time inversion formulation to give a robust quasilinear method that typically converges in a few iterations and generally avoids much of the time-consuming trial-and-errors seen in manual history matching. Moreover, the sensitivities are used to enforce basis functions to be adaptively updated only in areas with relatively large sensitivity to the production response. The sensitivity-based adaptive approach allows us to selectively update only a fraction of the total number of basis functions, which gives substantial savings in computation time for the forward flow simulations.

We demonstrate the power and utility of our approach using a simple 2D model and a highly detailed 3D geomodel. The 3D simulation model consists of more than one million cells with 69 producing wells. Using our proposed approach, history matching over a period of seven years is accomplished in less than twenty minutes on an ordinary workstation PC.

Introduction
It is well known that geomodels derived from static data only – such as geological, seismic, well-log and core data – often fail to reproduce the production history. Reconciling geomodels to the dynamic response of the reservoir is critical for building reliable reservoir models. In the past few years, there have been significant developments in the area of dynamic data integration through the use of inverse modeling. Streamline methods have shown great promise in this regard (Vasco et al. 1999; Wang and Kovscek 2000; Milliken et al. 2001; He et al. 2002; Al-Harbi et al. 2005; Cheng et al. 2006). Streamline-based methods have the advantages that they are highly efficient “forward” simulators and allow production-response sensitivities to be computed analytically using a single flow simulation (Vasco et al. 1999; He et al. 2002; Al-Harbi et al. 2005; Cheng et al. 2006). Sensitivities describe the change in production responses due to small perturbations in reservoir properties such as porosity and permeability and are a vital part of many methods for integrating dynamic data.

Even though streamline simulators provide fast forward simulation compared with a full finite-difference simulation in 3D, the forward simulation is still the most time-consuming part of the history-matching process. A streamline simulation consists of two steps that are repeated: (i) solution of a 3D pressure equation to compute flow velocities; and (ii) solution of 1D transport equations for evolving fluid compositions along representative sets of streamlines, followed by a mapping back to the underlying pressure grid. The first step is referred to as the pressure step and is often the most time-consuming. Consequently, history matching and flow simulation are usually performed on upscaled simulation models, which imposes the need for a subsequent downscaling if the dynamic data are to be integrated in the geomodel. Upscaling and downscaling may result in loss of important fine-scale information.
Recently, so-called multiscale methods have proven to be a promising alternative to standard upscaling, both with respect to accuracy and efficiency (Gautier et al. 1999; Arboagast and Bryant 2002; Jenny et al. 2004; Aarnes et al. 2005). These methods are specially designed to perform well when the underlying parameters exhibit a multiscale structure; that is, when the parameter values span several orders of magnitude or the correlation lengths of the heterogeneity structures vary over several orders. Like standard upscaling methods, the multiscale methods compute pressure and/or velocities by solving the global flow problem on a coarsened grid. However, whereas upscaling methods use local decoupled flow problems to derive upscaled permeabilities or transmissivities, and thus only preserve the local flow in an averaged sense, multiscale methods use the solutions of these localized flow problems as building blocks to form a global flow solution that is correct in an averaged sense on the coarse grid and at the same time contains representative subscale variations on the original fine grid.

Multiscale methods are primarily targeted at dynamic flows where the pressure needs to be computed repeatedly. Since temporal changes in the coefficients of the pressure equation are typically moderate compared to the spatial variability, it is seldom necessary to recompute the local flow problems each time the pressure is updated. Instead, local flow problems are computed initially as part of a preprocessing step (that is embarrassingly simple to parallelize) and typically only updated if the local domain is swept by a strong front in the fluid compositions or the global flow pattern changes significantly due to shut-in of wells, infill drilling, well conversion, etc. Hence, a pressure update typically consists of recomputing a few local flow problems and then solving a global flow problem on the coarse grid. This means that one can obtain an approximate solution on the original grid at the cost of solving the same problem on a much coarser grid.

In this paper, we combine multiscale-streamline simulation and streamline-based history matching in one efficient approach. As a flow solver we apply the multiscale mixed finite-element method (MsMFEM) (Chen and Hou 2002; Aarnes 2004). MsMFEM produces mass-conservative solutions both on the coarse grid and on the underlying fine grid, is flexible with respect to grid representation (geometry/topology), and has a rigorous mathematical framework.

For the history matching we use the generalized travel-time inversion method (Vasco et al. 1999; He et al. 2002), which has previously been successfully applied to many field cases. There are several advantages associated with travel-time inversion of production data. First, it is robust and computationally efficient. Unlike conventional ‘amplitude’ matching, which can be highly nonlinear, it has been shown that the travel-time inversion has quasilinear properties (Vasco et al. 1999; Wu and Datta-Gupta 2002). As a result; the minimization proceeds rapidly even if the initial model is not close to the solution. Second, travel-time sensitivities are typically distributed more uniform between wells compared to ‘amplitude’ sensitivities that tend to be localized near the wells. This prevents over-correction in the near-well regions (Wu and Datta-Gupta 2002). Finally, in practical field applications, production data are often characterized by multiple peaks. Under such conditions, the travel-time inversion can prevent the solution from converging to secondary peaks in the production response (Vasco et al. 1999).

Central in the inversion method is the computation of analytic streamline sensitivities in terms of simple 1-D integrals along streamlines. The sensitivities can be computed using a single streamline simulation. The second novel idea in this paper is a strategy based on sensitivity thresholding for reducing the workload for the forward simulation and for the inversion process. Altogether, the analytic sensitivities are used for three purposes: (i) in the inversion method, (ii) to reduce the computational complexity of the forward simulations by reducing the number of local flow solves, and (iii) to reduce computational complexity of the inversion process.

The outline of our paper is as follows. First, we discuss the basic steps in our proposed approach and illustrate the history-matching procedure using a simple synthetic example. Second, we describe the multiscale-streamline flow simulation and the history-matching procedure. Third, we discuss and demonstrate the impact of selective sensitivity-based workload reduction. Finally, we present a high-resolution history-matching example to demonstrate the efficiency and the practical applicability of our method.

**Background and Illustration of the Procedure**

Streamline-based history matching utilizes streamline-derived sensitivities to calibrate geomodels to dynamic data. The major steps involved in the proposed process are: (i) Multiscale-streamline flow simulation to compute production responses at the wells. (ii) Quantification of the mismatch between observed and computed production responses via a generalized travel time. An optimal ‘travel-time shift’ is computed by systematically shifting the computed production responses towards the observed data until the cross-correlation between the two is maximized (He et al. 2002). (iii) Computation of streamline-based analytic sensitivities of the production responses (water-cuts) to reservoir parameters, specifically permeability. (iv) Updating of reservoir properties to match the production history via inverse modeling. We propose a sensitivity-based thresholding strategy to reduce the computational work for this step.

This four-step process is repeated until a satisfactory match in production data is obtained. To reduce the computational workload for the forward simulation, we propose to reuse basis functions in regions with low sensitivity to the production responses.

In the next sections we will discuss the details of the mathematical formulation behind the multiscale mixed finite-element formulation and the inversion method, and propose a sensitivity-based strategy for selective work reduction. However, for clarity of exposition, we first illustrate the history-matching procedure using a synthetic 2-D example.

**A Synthetic Example.** This synthetic case (Case 1) involves reconstruction of a reference permeability distribution on a uniform 21×21 grid, based on the observed water-cut production history from a 9-spot pattern. For the forward simulation we apply the multiscale-streamline
Aarnes (2004) to ensure local mass conservation also for element discretization. The method was later modified by Chen and Hou (2002) to ensure local mass conservation on the coarse and fine grid, which is solved to obtain a flux for each face in the coarse grid. Fine-scale flow velocities are then obtained by multiplying the coarse-grid fluxes with the corresponding multiscale basis function and summing over all faces in the coarse grid.

The flow is described using quadratic relative permeability curves with zero residual oil and water saturations and endpoint mobility ratio $M_{\text{end}} = \mu_o/\mu_w$. We consider three flow cases: two with favorable mobility ratios ($M_{\text{end}}=0.2$ and $M_{\text{end}}=0.5$) and one with unfavorable mobility ratio ($M_{\text{end}}=0.1$). Synthetic dynamic data were generated by adding 5% white noise to the water-cut responses obtained from the reference permeability field using the streamline method with a standard two-point pressure solver. We treat these as the observed data. Next, starting from a homogenous initial permeability field, we match the water-cut data via the generalized travel-time inversion. To demonstrate the robustness of this method, we match the observed data for $M_{\text{end}}=0.2, 0.5, 10$ starting from a homogeneous initial permeability field. Here the permeability in each cell is treated as an adjustable parameter, giving a total of 441 unknown parameters to be estimated.

Plots of the time-shift and amplitude residuals in Fig. 3 show that the iteration converges very fast (after 3-4 iterations). Figure 4 shows a comparison of the initial and final match of the water-cut curves for $M_{\text{end}}=0.5$ for the three wells with lowest initial, highest initial, and highest final mismatch after six iterations. Overall, the match to the production data is quite satisfactory for all three mobility ratios. The corresponding permeability fields after six iterations are shown in Fig. 5. The three permeability models clearly capture the large-scale trends of the reference permeability field; a unique solution is not obtained since the data integration is ill-posed.

Mathematical Formulation

Multiscale Flow Simulation. An important aspect of the proposed history-matching algorithm is the use of a multiscale mixed finite-element method (MsMFEM) for the pressure equation. This method belongs to a family of multiscale finite-element methods, first introduced by Hou and Wu (1997). The basic idea of the methods is to construct special finite-element basis functions that are adaptive to the local properties of the elliptic differential operator. To ensure local mass conservation on the coarse and fine grid, Chen and Hou (2002) introduced a multiscale method based on a mixed finite-element discretization. The method was later modified by Aarnes (2004) to ensure local mass conservation also for
blocks containing source terms. In the current paper, we use a slightly different formulation due to Aarnes and Lie (2004).

**Governing Equations.** We consider incompressible two-phase flow of oil and water in a non-deformable permeable medium. For simplicity, we neglect the effects of gravity, compressibility and capillary forces. For simplicity, we assume no-flow boundary conditions. The flow equations can be formulated as an elliptic equation for the pressure $p$ and the total velocity $u$:

\[
    u = -\lambda k \nabla p, \quad \nabla \cdot u = q. \tag{1}
\]

Here $q$ is a source term representing injection and production wells, $k$ is the absolute permeability, and $\lambda_t = \lambda(S_w)$ is the total mobility. We will solve Eq. 1 for the fine-scale velocity field $u$ using MsMFEM, for which the details will be described in the next subsections.

The velocity field is used to obtain a streamline distribution. Along each streamline the 3D transport equation reduces to a 1D transport equation with the time-of-flight $\tau$ as the spatial coordinate:

\[
    \frac{\partial S_w}{\partial t} + \frac{\partial f_w(S_w)}{\partial \tau} = 0. \tag{2}
\]

The time-of-flight is defined as

\[
    \tau(r) = \int_0^{\phi(\xi)} \frac{1}{u(\xi)} d\xi. \tag{3}
\]

and expresses the time it takes a passive particle to travel a distance $r$ along the streamline. Equation 2 is solved forward in time along each streamline using front tracking (Holden and Risebro 2002). This method is unconditionally stable and therefore avoids the usual CFL-constraint that would otherwise have put a severe limitation on the size of the time step.

**Mixed Finite Elements.** The mixed finite-element formulation of the flow equation (Eq. 1) in a domain $\Omega$ seeks a pair $(u, p)$ in $U \times V$, such that

\[
    \int_{\Omega} u \cdot (\lambda k)^{-1} v \, dx - \int_{\Omega} p \nabla \cdot v \, dx = 0, \quad \forall v \in U, \quad \cdots (4)
\]

\[
    \int_{\Omega} \nabla \cdot u \, dx = \int_{\Omega} q \, dx, \quad \forall l \in V. \quad \cdots (5)
\]

Here $U$ and $V$ are (finite-dimensional) function spaces for pressure and velocity, respectively. Now, letting $\{\Psi\}$ and $\{\Phi\}$ be bases for $U$ and $V$, respectively, we obtain approximations $u = \sum u_i \Psi_i$ and $p = \sum p_k \Phi_k$, where the coefficients $u = \{u_i\}$ and $p = \{p_k\}$ solve a linear system of the form

\[
    \begin{bmatrix}
        B & C \\
        C^T & 0
    \end{bmatrix}
    \begin{bmatrix}
        u \\
        p
    \end{bmatrix}
    =
    \begin{bmatrix}
        0 \\
        q
    \end{bmatrix}. \tag{6}
\]

where $B=\{b_j\}$, $C=\{c_{ik}\}$ and $q=\{q_k\}$ are defined by

\[
    b_j = \int_{\Omega} \Psi^T_j \cdot (\lambda k)^{-1} \Psi_j \, dx, \quad \cdots (7)
\]

\[
    c_{ik} = \int_{\Omega} \Phi_i \nabla \cdot \Psi_j \, dx, \quad \cdots (8)
\]

\[
    q_k = \int_{\Omega} \Phi_k q \, dx. \quad \cdots (9)
\]

**Multiscale Basis Functions.** In a standard discretization, the spaces $U$ and $V$ typically consist of low-order piecewise polynomials. In multiscale methods, $U$ and $V$ are given by the solution of local flow problems. For incompressible flows, the actual pressure solution is immaterial for the flow simulation, and so only the velocity field is needed. We will therefore only construct an accurate multiscale approximation space $U^{ms}$ for the velocity and use a standard approximation space $V$ for pressure consisting of piecewise constant functions.

Let $\{K_m\}$ be a partitioning of $\Omega$ into mutually disjoint (fine) grid cells. Furthermore, let $\{T_i\}$ be a coarse partitioning of $\Omega$, in such a way that whenever $K_m \cap T_i \neq \emptyset$, then $K_m \subset T_i$ (see Fig. 1). Let $\Gamma_{ij}$ denote the non-degenerate interfaces $\Gamma_{ij} = \partial T_i \cap \partial T_j$. For each $\Gamma_{ij}$, we assign a basis function $\Psi_{ij}$ in $U^{ms}$, and for each $T_i$ we assign a basis function $\Psi_i$ in $V$. The basis function $\Psi_{ij}$ is obtained by forcing a unit flow from block $T_i$ to $T_j$; that is, by solving a local flow problem in $\Omega_{ij} = T_i \cup T_j$

\[
    \Psi_{ij} = -\lambda_j k \nabla \Phi_j, \quad \nabla \cdot \Psi_{ij} = \begin{cases} w_i(x), & x \in T_i, \\ -w_i(x), & x \in T_j, \end{cases} \tag{10}
\]

with $\Psi \cdot n = 0$ on the boundary of $\Omega$. Here the total mobility $\lambda_j = \lambda(S_{ij})$ is given on the underlying fine grid $K_m$. To give a unit flow from $T_i$ to $T_j$, the source terms $w_i(x)$ are normalized

\[
    w_i(x) = W_i(x) \left( \int_{T_j} W_i(\xi) d\xi \right)^{-1}. \tag{11}
\]

To ensure a conservative approximation of $\nu$ on the fine grid, we choose $W_i=q$ for coarse blocks containing a well (Aarnes 2004). For coarse blocks where $q=0$, we scale $W_i$ according to the trace of the permeability tensor (Aarnes et al. 2006); i.e., we use

**Fig. 5 – Case 1: The plots show, from upper left to lower right, reference permeability field and matches obtained after six iterations for mobility ratios $M_{nl}=0.2, 0.5$ and 10.**
\[ W_r(x) = \begin{cases} \text{trace}(k(x)), & \text{if } q(x)_0 = 0, \\ q(x), & \text{otherwise} \end{cases} \quad (12) \]

The local flow problems in Eq. 10 can be solved numerically by any consistent and conservative method; here we use the standard two-point flux-approximation (TPFA) scheme. The corresponding basis functions can be seen as generalizations of the lowest-order Raviart-Thomas basis functions in a standard mixed method (Raviart and Thomas 1975). Figure 2 illustrates the \( x \)-velocity basis functions in two different cases.

**Implementation of MSMFEM.** We will briefly describe some implementation aspects related to the efficiency and generality of MSMFEM. The mixed formulation leads to an indefinite global system (Eq. 6), which may be more difficult to solve efficiently than the symmetric positive-definite (SPD) systems that typically arise from standard discretization methods. However, it is possible to obtain an SPD system also for MSMFEM by reformulating Eq. 4 and Eq. 5 to an equivalent (so-called) hybrid system. Like the indefinite system in Eq. 6, the hybrid system will be sparse because the basis functions have local support, and the solution can be obtained using one of the efficient linear solvers specialized for sparse SPD systems. The hybrid formulation is described in more detail by Aarnes et al. (to appear). We note that in our current implementation, we solve the global system in Eq. 6 using a direct sparse solver, since we only deal with moderately sized coarse systems.

Most of the computational work in MSMFEM is associated with solving the local flow problems defined by Eqs. 10 to 12, and the choice of solution strategy for these equations is crucial to the overall performance of the method. The local problems are usually small to moderately sized, and the resulting systems can be solved using iterative or direct sparse linear solvers. The optimal choice of linear solver typically depends on the problem size, and we recommend having available a range of solvers tuned to different system sizes. Alternatively, if one has access to a highly efficient solver for large sparse systems, it may be beneficial to lump together several local problems to form a larger system. Solving larger systems may be advantageous because the most efficient linear solvers typically require an initial setup phase. Regardless of the choice of solution strategy, efficient parallelization is easy, since the local flow problems are completely decoupled.

In the examples presented in this paper, we only use Cartesian grids. However, MSMFEM is flexible with respect to the choice of both fine and coarse grids. Given a fine-grid solver, basis functions can be defined for almost any collection of connected fine-grid cells (Aarnes et al. 2006). Recently, the method has been implemented for (matching) corner-point and tetrahedral grids in 3D (Aarnes et al., to appear), and based on this experience we are confident that the methodology presented in the current paper is easily extended to corner-point grid models.

**Integration of Production Data.** In our approach, integration of production data is carried out using a ‘generalized travel-time inversion’ as described by He et al. (2002). First, the production-data mismatch is determined by computing a ‘generalized travel-time misfit’ for the water-cut at each producing well. This is accomplished by shifting the computed water-cuts towards the observed data until the correlation between the two is maximized. The inversion algorithm simultaneously minimizes the travel-time misfit for all the wells using an iterative least-square minimization algorithm (LSQR) (Vasco et al. 1999; He et al. 2002). The basic underlying principles behind the history-matching algorithm are briefly as follows:

- Match the field-production history within a specified tolerance. This is accomplished by minimizing the travel-time misfit for water-cut.
- Preserve geological realism by keeping changes to the prior geological model minimal, if possible. The prior model already incorporates static data (well and seismic data) and available geological information.
- Only allow for smooth and large-scale changes; the production data has low resolution and cannot be used to infer small-scale variations in reservoir properties.

**Formulation of Inverse Problem.** Mathematically, this algorithm leads to the minimization of a penalized misfit function consisting of the following three terms (Vasco et al. 1999; He et al. 2002):

\[ \| \Delta \tilde{t} - G \delta m \| + \beta_1 \| \delta m \| + \beta_2 \| L \delta m \| \quad \text{subject to} \quad \Delta \tilde{t} = \text{observed water-cuts} \quad (13) \]

Here \( \Delta \tilde{t} \) is the vector of generalized travel-time shifts at the wells, \( G \) is the sensitivity matrix containing the sensitivities of the generalized travel time with respect to changes \( \delta m \) in the reservoir properties, and \( L \) is a second-order spatial difference operator. The first term ensures that the difference between the observed and calculated production response is minimized. The two remaining terms are standard regularization terms. The second term is a norm constraint that penalizes deviations from the initial (prior) geological model and as such helps to preserve the geological realism in the history match. The third term, which is a ‘roughness’ constraint that measures the regularity of the changes, is introduced to stabilize the inversion. Physically, it only allows for large-scale changes that are consistent with the low resolution of the production data. The weights \( \beta_1 \) and \( \beta_2 \) determine the relative strengths of the prior model and the roughness term.

The minimum in Eq. 13 can be obtained by an iterative least-squares solution to the augmented linear system

\[
\begin{pmatrix}
G \\
\beta_1 I \\
\beta_2 L
\end{pmatrix}
\begin{pmatrix}
\Delta \tilde{t} \\
\delta m \\
0
\end{pmatrix} =
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix}
\quad (14)
\]

This system is solved with the iterative least-square minimization algorithm, LSQR (Paige and Saunders 1982), for which the computational cost scales linearly with respect to the number of degrees-of-freedom (Vega et al. 2004). Fine-grid sensitivities close to zero are eliminated, which makes the system more sparse and reduces the number of arithmetic operations for the LSQR-iterations. In the next section we will discuss an approach to further reduce the number of nonzero sensitivities based on thresholding of coarse-grid sensitivities.

In our implementations we focus on inverting water-cut data. However, the generalized travel-time inversion method has earlier been extended to compressible three-phase flow, so
that water-cut and gas-oil ratios are incorporated jointly (Cheng et al. 2006).

**Water-Cut Sensitivities.** A unique feature of streamline methods is that the parameter sensitivities can be computed using a single flow simulation, leading to very fast history-matching or inverse-modeling algorithms. Moreover, because the sensitivities are simple integrals along streamlines, the runtime scales very favorable with respect to the number of grid cells, thus making streamlines the preferred approach for history matching highly-detailed geological models.

To derive the desired streamline-based sensitivities, we consider the velocity of propagation for a given saturation contour $S_0$ along a streamline,

$$\frac{\partial \tau}{\partial t} \bigg|_{S_0} = \frac{df_w}{dS_w}, \quad \text{(15)}$$

from which it follows that the arrival time $t_a$ of the saturation contour will be,

$$t_a = \int \frac{df_w}{dS_w}, \quad \text{(16)}$$

We use the above relationship to compute the sensitivity of the arrival time of the saturation contour based on the sensitivity of the time-of-flight (Vasco et al. 1999; He et al. 2002). Specifically, the sensitivity of the arrival time of the saturation front with respect to reservoir parameter $m$ is computed as,

$$\frac{\partial t_a}{\partial m} = \frac{\partial \tau}{\partial m} \bigg|_{S_0}, \quad \text{(17)}$$

Here the sensitivity of the time-of-flight is computed analytically from a single streamline simulation under the assumption that the streamlines do not shift because of small perturbations in reservoir properties. For example, the time-of-flight sensitivity with respect to permeability in grid cell $i$, under the assumption of the same permeability for the whole grid cell, will be given by (Vasco et al. 1999)

$$\frac{\partial \tau}{\partial k_i} = \int \frac{s(\xi)}{\lambda_i k_i} d\xi = \frac{\Delta \tau_i}{\Delta k_i}, \quad \text{(18)}$$

where the integral is along the streamline trajectory $\Sigma$ and $s(x)$ is the ‘slowness’ defined as the reciprocal of the total interstitial velocity

$$s(x) = \frac{\phi(x)}{\mu(x)} = \frac{\phi(x)}{\lambda_i k_i |\nabla p|}. \quad \text{(19)}$$

Similarly, the time-of-flight sensitivities can be calculated with respect to mobility or to the product of mobility and permeability.

Finally, the sensitivity of the shift $\Delta \vec{T}$ in the generalized travel time with respect to reservoir parameters is given by

$$\frac{\partial \Delta \vec{T}}{\partial m} = -\frac{1}{N_d} \sum_{a=1}^{N_d} \frac{\partial \Delta \vec{T}}{\partial m}, \quad \text{(20)}$$

where $N_d$ represents the number of observed data for a well. Worth mentioning here is an important practical aspect. Our experience indicates that the selective work reduction and the data integration are more robust if the sensitivities are made dimensionless by calculating the sensitivities of the logarithm of the time shifts as described by He et al. (2002).

Finally, we remark that the streamline-based sensitivity computation has also been addressed for cases including gravity, changing field conditions, and fractured reservoirs (He et al. 2002; Al-Harbi et al. 2005).

**Sensitivity-Based Selective Work Reduction**

In this section, we discuss how the sensitivities introduced above can be used to reduce the computational complexity of the history matching with negligible loss in quality of the derived match. To this end, we compute a sensitivity coefficient for each coarse block by summing the corresponding fine-grid sensitivities (Yoon et al. 2001). Based on these sensitivity coefficients, we decide when to update and when to not update the corresponding basis functions. Similarly, we will reduce the inverse system by only including fine-scale sensitivities from coarse blocks having a sufficiently high sensitivity coefficient as described in more detail below.

**Selective Updating of Basis Functions.** We propose to reduce the computational work for MsMFEM by only updating basis functions in areas with large production-response sensitivities. To determine which basis functions to update, one can either: (i) use a predefined threshold for the sensitivity values, or (ii) update a predefined fraction of the basis functions. The first approach is fully adaptive in the sense that the number of updated basis functions may change from iteration to iteration. However, this approach requires (general) guidelines for setting the threshold. The second approach requires us to sort the sensitivities. This is a minor concern since the number of operations for sorting $N$ numbers scales like $N \log N$ and the number of basis functions scales with the number of coarse blocks. For our implementations we therefore stick to the second approach.

By inspecting Eq. 10, we notice that there are three factors that may require the basis functions to be updated before a new pressure solve. First of all, we notice that if the absolute permeability $k(x)$ changes, the basis functions will change, too. In history matching, the absolute permeability will typically change in certain regions from one forward simulation to the next. Secondly, if the well rate $q$ changes, the source terms $w_i(x)$ will change and hence basis functions with support in well-blocks will change. Finally, if the total mobility $\lambda_i$ changes, due to changes in saturation (or viscosities), the basis functions will change.

In the first flow simulation of the history-matching procedure, we update all basis functions in every pressure step, because no sensitivities are yet available. (In a more sophisticated implementation, one would typically have used another kind of indicator to ensure that basis functions are only updated near the saturation front (Aarons 2004; Jenny et al. 2004)). After the first simulation, the permeability field is updated by the inversion method. Since the permeability field has changed, we should, at least in principle, recalculate all basis functions for the first pressure step of the next flow simulation. For the subsequent pressure steps of the simulation, we apply the proposed selective updating strategy.
For the subsequent simulations we repeat the strategy of the second simulation. The approach described in this paragraph, when \( x \)% of the basis functions are updated dynamically each time step, is referred to as \( x \)% DU (dynamical update). Finally, we remark that for 0% DU, sensitivities are not really needed. Therefore, if we for this strategy choose to not update basis functions during the first flow simulation, we will denote the strategy 0% DU*. This special case deviates from what we specified above.

An extended approach would be to reuse basis functions from the previous forward simulation in the coarse blocks where the absolute permeability has undergone small or no changes in the last inversion step, or generally in coarse blocks that have little effect on the overall production characteristics. We will refer to this strategy, where \( x \)% of the basis functions are updated initially and the remaining \((100-x)\)% are kept from the previous flow simulation, as \( x \)% IU (initial update).

**Selective Reduction of the Inversion System.** Since the water-cut data contain limited information about fine-scale variations, it can be advantageous to avoid involving areas of low sensitivity in the inversion, and instead focus on resolving large-scale structures in areas with higher sensitivities. We therefore propose to eliminate fine-scale sensitivities from the LSQR-system (Eq. 14) in areas of low sensitivity to reduce the computational work in the inversion process. To determine the areas of low and high sensitivity, we use the sensitivity coefficients of the coarse grid-blocks. That is, we introduce a threshold and only include fine-scale sensitivities associated with cells inside coarse blocks having a summed sensitivity above the given threshold. The coarse blocks that are eliminated in this process will usually mainly contain cells with zero or low sensitivity.

The constraints involved in Eq. 13 are important for the elimination of coarse blocks to work. As for the thresholding of basis functions, we can either use a predefined threshold for the sensitivity values or a predefined fraction of coarse blocks; here we use the second approach. Henceforth, keeping \( y \)% of the coarse blocks is referred to as \( y \)% CB. It should be noted that eliminating fine cells for a fraction of the coarse blocks having low sensitivity will not necessarily decrease the number of fine-grid sensitivities in the inverse system by the same fraction. The reason is cells with zero or small sensitivity are already eliminated, and such fine-grid sensitivities are more likely represented in coarse blocks with low sensitivity.

**Impact of Selective Work Reduction** To investigate the accuracy of the proposed selective work reduction, we apply it to the synthetic 9-spot case presented earlier in this paper (Case 1). We will still refer to this case as Case 1 even though we will vary some parameters and strategies for selective work reduction. To further assess the quality of the data integration, we will in the following also report the average discrepancy between the reference and matched permeability field measured by

\[
\Delta \log k = \frac{1}{N} \sum_{i=1}^{N} \log(k_{i}^{\text{reference}}) - \log(k_{i}^{\text{derived}})
\]

Further, we also report time-shift and amplitude residuals measured by \( \| \Delta f \|_{2} \) and

\[
\left( \sum_{j=1}^{M} \left( f_{j}^{\text{obs}}(t_{j}) - f_{j}^{\text{cal}}(t_{j}) \right)^{2} \right)^{1/2}
\]

respectively. Here \( f_{j}^{\text{obs}} \) and \( f_{j}^{\text{cal}} \) are the observed and calculated water-cut data, respectively, in well \( k \) at time \( j \).

To test our work reduction strategy we use a \( 5 \times 5 \) test matrix with \( x \)% dynamical update and \( y \)% initial update for \( x,y=0.25,...,100 \). Figure 6 shows the reduction of time-shift and amplitude residuals after six iterations, as well as the discrepancy between matched and reference permeability fields, compared with similar results obtained by using the full method with a TPFA pressure solver. Judging from the amplitude residual and the permeability discrepancy, the data are well matched for all parameters \( x \) and \( y \), and the quality of the history match does not seem to decline dramatically compared with the TPFA solver. For the time-shift residual, the use of the multiscale pressure solver gives better match than for the TPFA solver. On the other hand, this measure is also more sensitive to the choice of parameters for the selective work reduction. The quality of the match generally decays with decreasing percentage of cells being dynamically updated, except for the time-residual for \( M_{\text{end}}=0.5 \), which somewhat surprisingly shows the opposite trend.

The derived permeability fields for the unfavorable mobility ratio do not seem to change much when reducing the number of dynamically updated basis functions. Following Aarnes (2004), one can argue that it is in general quite safe to reduce the number of dynamically updated basis functions for unfavorable flow cases, since these are characterized by weak shocks and mostly smooth variations in the total mobility. For the favorable mobility ratio \( (M_{\text{end}}=0.2) \), the derived permeability fields seem to change more by reducing the fraction of basis functions updated; see Fig. 7. In this case, the flow will generally have strong saturation fronts, which induce major changes in the basis functions as the leading water fronts move through the corresponding grid blocks.

Finally, we investigate the effect of the proposed strategy for selective reduction of the inverse system. To this end, we keep the fine-grid sensitivities corresponding to 100%, 75%, and 50% of the coarse blocks, selected by thresholding the summed sensitivities in the coarse grid. This strategy was tested in combination with \( x \)% dynamical update for \( x=0.25,...,100 \). Reducing the number of parameters in the inversion had little influence on the convergence of the inversion and small effect on the quality of the final match (see Fig. 8 and Table 1). Some derived permeability fields for \( M_{\text{end}}=0.5 \) are shown in Fig. 9. On the other hand, the inverse system could not be reduced much further than 50%. For lower values, we typically observed lack of convergence and/or (highly) non-realistic final matches. We also note that for some cases, the selective reduction of the inverse system resulted in a slightly slower convergence for the inversion. The method converged to the same residual level as without selective work reduction, but the inversion required one or two additional iterations, thus resulting in increased total computation time. Even though the selective reduction of the inversion system can result in a slightly slower convergence, our experiments demonstrate robustness for the generalized travel-time inversion.
Fig. 6 – Case 1: Robustness of selective work reduction for basis function updates.

Fig. 7 – Case 1: Derived permeability field updating (from upper left to lower right) 100%, 75%, 50%, and 25% of the basis functions dynamically for mobility ratio $M_{end}=0.2$. The reference permeability field is shown in Fig. 5.
History Matching a Full 3D Geomodel

In this section we demonstrate the feasibility of the approach for field studies by application to a high-resolution 3-D example (Case 2). As mentioned before, streamlines and the time-of-flight are used to compute the sensitivity of the production data with respect to reservoir parameters as described above. In this synthetic field-scale example, water-cuts were matched to update the reservoir permeability distribution using the multiscale-streamline simulator for the forward simulation.

Model Description. The geomodel consists of a fine grid with 256×128×32 cells, which gives a total of 1,048,576 grid cells, each of size 10×10×2 m. The fine-grid cells are collected into a uniform 32×16×8 coarse grid, so that each coarse block consists of 8×8×4 cells in the fine grid. All the cells are treated as active.

The permeability is log-normally distributed in the range 0.017 mD to 79.5 mD with mean 2.2 mD (see Fig. 10b). The correlation length is about 270 meters in the horizontal direction and about 90 meters in the vertical direction. For our purposes, this permeability field was used as a true model, from which we generated our synthetic production data using the standard TPFA method directly on the fine grid.

A total number of 32 injectors and 69 producers were included in the simulation model (see Fig. 11). All the wells are vertical and intersect all layers. The production history consists of 2475 days of water-cut data from the 69 producers (Fig. 12). The water injectors were injecting at constant total reservoir volume rate of 1609 bbl/day, and each producer was producing with constant reservoir volume rate fulfilling the total voidage rate. For each simulation, we used 15 pressure steps of length 165 days, quadratic relative permeability curves, and end-point mobility ratio of $M_{\text{end}}=5$. 

### Table 1 – Case 1: REDUCTION OF MISFIT FOR SELECTIVE WORK REDUCTION OF INVERSE SYSTEM

<table>
<thead>
<tr>
<th>DU</th>
<th>IU</th>
<th>CB</th>
<th>$M_{\text{end}}=0.2$</th>
<th>$M_{\text{end}}=0.5$</th>
<th>$M_{\text{end}}=10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>100.0</td>
<td>100.0</td>
<td>1.045</td>
<td>1.045</td>
<td>1.045</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
<td>75%</td>
<td>5.8</td>
<td>5.8</td>
<td>5.8</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
<td>50%</td>
<td>8.1</td>
<td>8.1</td>
<td>8.1</td>
</tr>
<tr>
<td>50%</td>
<td>100%</td>
<td>75%</td>
<td>4.8</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>50%</td>
<td>100%</td>
<td>50%</td>
<td>3.8</td>
<td>3.8</td>
<td>3.8</td>
</tr>
</tbody>
</table>
Integration of Production Data. To generate an initial permeability model, we treat the permeability values in the well-blocks of the reference model as known data. By conditioning on the well blocks, sequential Gaussian simulation was used to generate multiple realizations of the permeability model (Deutsch and Journel 1998).

In the following we will mainly consider three approaches: MsMFEM [full], MsMFEM [reduced], and TPFA. The two first approaches are multiscale approaches, while the last one simulates directly on the fine grid. Further, for the first and the last approach no selective work reduction occurs. For MsMFEM [reduced], the extended approach for reducing the computations of basis functions is applied. For each new forward simulation, the basis functions are sorted according to summed sensitivities and the basis functions having the lowest 50% sensitivities are kept from the previous flow simulation. The remaining 50% of the basis functions are updated once before the first pressure solve. Moreover, selective reduction of the inverse system is used to keep fine-grid sensitivities only for 50% of the coarse blocks. In other words:

- MsMFEM [full] = 100% DU + 100% IU + 100% CB,
- MsMFEM [reduced] = 0% DU* + 50% IU + 50% CB.

Figure 13 and Table 2 show the convergence of the inversion algorithm (residuals given by Eq. 22). In six iterations, all misfits in time-shift and amplitude for the water-cut dropped appreciably for all three approaches. Reference, initial, and matched water-cut curves are shown in Fig. 12 for a few selected producers. Some of the wells had a quite good match initially, and at the end of the history matching all wells had a quite satisfactory match.

Figure 10 compares the initial and reference permeability models with the updated (derived) models. The scale is logarithmic and the minimum permeability is 0.017 mD and the maximum is 110.3 mD. The three approaches gave almost identical derived permeability fields. Therefore, just one of the derived permeability fields (for MsMFEM [full]) is picked for closer inspections. From a casual look, it is hard to discern the changes made to the initial model. This is because the inversion algorithm is designed to preserve the geologic continuity and the initial geologic features to the maximum possible extent. However, a careful comparison reveals many differences between the initial and the updated geologic models.
Next, we examine if the changes made to the initial model are consistent with the ‘reference’ permeability model. **Figure 14** shows the differences between the updated and initial permeability model. These differences represent ‘changes made’. This is to be compared with the ‘changes needed’, which is the difference between the reference and the initial permeability model. We see that there is close agreement, particularly in regions where the permeability needs to be reduced (negative changes). As might be expected, there are also some discrepancies. Many of the wells had a good match initially even though the permeability fields differ. Because the water-cut data curves are a result of the total flow pattern between a producer and one or more injectors, this data source may have limited spatial information. Some of the changes occur in correct horizontal position, but incorrect vertical position. This can occur because the water-cut data has no vertical spatial resolution. Finally, it is worth pointing out that this inversion problem is highly ill-posed, and therefore a variety of possible solutions exist. Table 2 shows average discrepancies between the reference and the derived permeability fields (see Eq. 21) for TPFA, MsMFEM [full], MsMFEM [reduced]. The average discrepancies indicate that the history-matching procedure is stable with respect to the selective work-reduction strategies. We have also investigated some other selective work-reduction strategies, and the results with respect to both misfit and average discrepancies turned out to be as stable as for Case 1.

To sum up, the changes made to the permeability field preserved the geologic realism, were mostly in accordance with the ‘changes needed’ (see Fig. 14), and resulted in satisfactory match of the water-cut data. Further, the different strategies for selective work-reduction turned out to give stable results with respect to ‘changes made’ and misfit (see Table 2).

**Fig. 12** – Case 2: Water-cut match for 12 of the 69 production wells included in the history match of the geologic model (MsMFEM [full]). For each plot the solid line, the dashed and the dash-dotted line represents the reference, the initial and the updated water-cut curve, respectively.

**Fig. 13** – Case 2: Reduction of residuals for all producers. Forward simulation: MsMFEM [full] (solid curve), MsMFEM [reduced] (dashed curve) and TPFA (dash-dotted curve).
Fig. 14 – Case 2: Comparison of the “derived-initial” permeability difference and the “true-initial” permeability (MsMFEM [full]).

| TABLE 2 – Case 2: REDUCTION OF MISFIT AND EFFICIENCY FOR HISTORY-MATCHING PROCEDURE |
|----------------------------------|--------|-------------|-----------|----------------|----------|----------|---------------|----------|
|                                   | T (res. %) | A (res. %) | Δlog(k)   | Total simulation time (Wall clock) | Total CPU-time: |
|                                   |         |           |          | PC1 | PC2 | PC1 | PC2 | PC1 | PC2 | PC1 | PC2 |
| Initial                           | 100.0   | 100.0     | 0.821    | -   | -   | -   | -   | -   | -   | -   | -   |
| TPFA                              | 8.9     | 53.5      | 0.806    | 2h 12min | 1h 04min | 1h 02min | 33min | 54min | 28min |
| MsMFEM [full]                    | 10.3    | 53.1      | 0.796    | 2h 42min | 2h 29min | 1h 17min | 1h 54min | 1h 06min | 32min |
| MsMFEM [reduced]                 | 7.8     | 53.7      | 0.823    | 1h 34min | 43min   | 9min    | 7min   | 1h 07min | 32min |
| MsMFEM [reduced - SL]            | 7.6     | 48.7      | 0.808    | 36min  | 17min   | 9min    | 7min   | 12min  | 6min  |
Computational Efficiency. Finally, we will assess the efficiency of our multiscale method compared to a standard streamline method using a TPFA pressure solver. To this end, we consider two different computers running Linux: PC 1 is a laptop PC with a 1.7 GHz Intel Dothan Pentium M processor, 2Mb cache and 1.5 Gb memory. PC 2 is a workstation with a 2.4 GHz Intel Core 2 Duo, 4Mb cache, and 3 Gb memory.

Table 2 reports simulation times observed on the two computers. Here the total simulation time includes time for inversion, IO, and seven forward simulations, each with fifteen pressure steps. Similarly, we report the total time for the pressure solves and the transport solves (including mappings and tracing of streamlines).

When all basis functions are updated in all steps, the multiscale solver is, as expected, about 25% slower than TPFA with an optimal algebraic multigrid (AMG) solver on the laptop (PC1). On the other hand, the memory requirements for MsMFEM are quite low and this solver could easily have been run on larger models, as opposed to the TPFA solver, for which AMG almost exceeded the available memory. Moreover, on highly skewed, non-Cartesian grids (e.g., corner-point grids), MsMFEM uses a much better spatial discretization (Aarnes et al., to appear) and will therefore give more accurate predictions of flow.

The comparison of TPFA and MsMFEM [full] is not very interesting on the workstation (PC2). Due to an immature compiler for the particular hardware, we were not able to optimize the direct solver used to compute basis functions, while AMG could be (almost) fully optimized by using a vendor-specific compiler. The runtimes for the pressure solves (and the total runtime) on PC2 are therefore somewhat higher than expected, and will probably improve significantly when a more mature compiler becomes available in a few months.

By MsMFEM [reduced], we were able to reduce time for pressure solves by about 80% on both computers. In MsMFEM [reduced] the basis functions to be reused were read from file. Slow disc access on the laptop therefore prevented a further reduction in runtime. The workstation, on the other hand, had a faster disc, but further reductions in runtime were prevented by the unoptimized linear solver (as discussed above).

Reduction of the inverse system was expected to have a very small effect on the runtime, since a fully optimized compilation on a GHz processor gives a floating-point performance that would make the reduced number of arithmetic operations insignificant compared to other kinds of operations, which indeed is consistent with what we observe in Table 2. However, the results from the reduction of the inverse system indicate robustness for the generalized travel-time inversion method.

Finally, to speed the method further up, and to make our simulations comparable to state-of-the-art commercial streamline solvers, we apply a method for improved mass conservation for streamline simulation proposed by Kippe et al. (2007). Using this method, the total number of streamlines could be reduced from 500 000 to 50 000, thereby reducing the time for the transport solves by 80%. Altogether, this meant that the full history match could be performed in an impressive runtime of 17 minutes on the workstation (PC2) and 36 minutes on the laptop (PC1). For the workstation there is an obvious potential for further improvements by using a better compiler. Moreover, on the Core 2 Duo processor one should also exploit the natural parallelism in updating basis functions and in the streamline computations.

Summary and Conclusions
A novel approach to history matching using multiscale-streamline simulation and analytic sensitivities is presented. The power and utility of our proposed approach is demonstrated using both a synthetic and a field-scale example. The synthetic case includes matching of water-cut from a 9-spot pattern and is used to validate the method. The field-scale example consists of more than a million grid cells. Starting with a prior geomodel, production data were integrated using a generalized travel time inversion. The entire history matching process took less than 40 minutes using a laptop PC and about 17 minutes using an ordinary workstation PC. The permeability changes were found to be reasonable and geologically realistic.

Some specific conclusions from this paper can be summarized as follows.

1. A multiscale-streamline flow simulator was used for history matching by generalized travel-time inversion.
2. By utilizing the production-response sensitivities provided by the generalized travel-time inversion, we were able to reduce the total workload for the multiscale simulator considerably and still preserve the accuracy of the flow simulation.
3. By utilizing the production-response sensitivities, we were able to selectively reduce the number of non-zero sensitivities in the inverse system considerably without reducing the accuracy of the production data integration. This demonstrates robustness for the generalized travel-time inversion.
4. The approach proved applicable and efficient for a high-resolution reservoir model.

Acknowledgements
The research of Stenerud was funded by the Uncertainty in Reservoir Evaluation (URE) program at the Norwegian University of Science and Technology. The research of Kippe and Lie was funded by the Research Council of Norway under grant number 152731/S30.

Nomenclature

- \( u = \) total Darcy velocity
- \( p = \) pressure
- \( l, v = \) test functions
- \( V, U = \) function spaces
- \( K = \) fine grid cells/elements
- \( T = \) coarse grid blocks/elements
- \( \Omega = \) domain
- \( \Gamma = \) coarse block interface
- \( n = \) unit normal vector
- \( \Psi = \) basis function velocity

\[ \text{[SPE 106228]} \]
\( \Phi \) = basis function pressure  
\( q \) = total rate (source/sink)  
\( f_w \) = fractional flow function (water)  
\( S_w \) = saturation of water  
\( k \) = absolute permeability  
\( \lambda_t \) = total mobility  
\( M_{end} \) = end-point mobility ratio  
\( m \) = reservoir parameter  
\( N_q \) = number of data points  
\( N_r \) = number of grid cells  
\( t_a \) = arrival time  
\( r \) = time-of-flight

**Subscripts**

\( ms \) = multiscale

**References**


