Abstract: An emerging method to increase the recovery of oil from subsurface reservoirs is the application of measurement and control techniques to better control the multiphase flow in the reservoir over the entire production period. In particular the use of sensors and remotely controllable valves in wells and at surface, in combination with large-scale reservoir flow models is promising. Various elements from process control may play a role in such ‘closed-loop’ reservoir management, in particular optimization, parameter estimation and model reduction techniques.

Keywords: petroleum, oil, reservoir, multiphase, flow, model-based, control, optimization, data, assimilation, parameter, estimation, adjoint, robust.

1. INTRODUCTION

Most oil reservoirs consist of relatively thin slabs of porous rock buried at depths of hundreds to thousands of meters. They are typically in the order of tens of meters thick and cover several square kilometers. After the drilling of wells the oil usually flows to the surface naturally, but after some years this primary recovery phase ends, and it will be necessary to inject water or gas into the reservoir to maintain the reservoir pressure and to displace the oil from the injection wells towards the production wells. However, even when using such secondary recovery techniques, most of the oil remains trapped in the pores of the rock, and often the oil recovery factor stays somewhere between 10 and 50%. A further increase of the recovery factor of existing oil fields is therefore a good alternative to finding new ones. Sometimes this is possible during a tertiary recovery phase through the use of ‘enhanced’ oil recovery techniques such as the injection of surfactants, polymers or steam. These techniques are relatively expensive, and, depending on the type of oil and the subsurface conditions, only economically feasible at oil prices even above the current high level. An alternative, emerging, method to increase the recovery factor is the application of measurement and control techniques to improve the control of subsurface flow. In particular the use of sensors and remotely controllable valves in wells and at surface, in combination with large-scale subsurface flow models is an increasing area of research, which is known in the oil industry under various names, such as ‘smart fields’, ‘intelligent fields’, ‘real-time reservoir management’, or ‘closed-loop reservoir management’. We will discuss the use of such ‘smart’ technology to maximize reservoir performance in terms of oil recovery or another economic objective, over the life of the reservoir which is in the order of years to decades. Sources of inspiration for our research are at one hand model-based control concepts as used in the process industry, which offer a wide variety of solutions to cope with uncertainties, nonlinearities and multi-scale optimization. At the other hand we draw inspiration from disciplines like meteorology and oceanography where advanced data assimilation techniques have been developed to condition large-scale flow models (with more than $10^6$ state variables) to measured data while honoring measurement and model errors.

2. RESERVOIR MANAGEMENT

Fig. 1 depicts reservoir management as a model-based controlled process (Jansen et al., 2005). The system, at the top of the figure, comprises of one or more reservoirs, wells, and facilities for the separation and treatment of oil, gas and water. Generally, the system boundaries can be specified accurately for the wells and the surface facilities, but are much more uncertain for the reservoir of which the geometry is usually deduced from seismics with a limited resolution. Also the parameters of the system are known to varying degrees: the fluid properties can usually be determined quite well, but the reservoir properties are only really known at the wells. The subsurface is very heterogeneous, and the
parameters relevant to flow are correlated at different length scales, but often over distances smaller than the inter-well spacing. As a consequence, the uncertainties in the model parameters of the subsurface part of the system are very large, and during the design phase of an oil field development it is therefore customary to construct multiple subsurface models to simulate the flow of fluids for different geological ‘realizations’. The typical number of state variables in these system models is in the order of $10^4$ to $10^6$, with similar numbers for the model parameters. Based on these large-scale system models it is possible to make a design for the oil recovery process (known as the field development plan). This concerns, for example, determining the number and position of wells, or the optimal water injection and oil production flow rates over the life of the reservoir. During the past decades the possibilities to control subsurface flow have increased considerably. This concerns complex well configurations, e.g. ‘meandering’ horizontal wells, or multi-lateral wells with multiple branches, and the installation of control valves in ‘smart’ wells or at surface. Model-based optimization is therefore a rapidly growing activity within the reservoir simulation community. During the oil production process, more or less regular measurements are performed at the top of the wells and in the facilities, which give an indication of the pressures and phase rates (i.e. oil, gas and water flow rates) in the surface part of the production system. Traditionally these measurements are performed monthly or quarterly and with a limited accuracy. During the past years, however, an increasing amount of sensors is being installed that give near-continuous information about the system pressures and phase rates, not only at surface but more and more also downhole. In addition, other measurement techniques have emerged that give an impression of the changes in reservoir pressure and fluid saturations in between the wells. This concerns in particular ‘time-lapse’ seismic measurements, which allow for monitoring the displacement of oil-water or oil-gas fronts between injection and production wells at regular (say quarterly to yearly) intervals. By combining the measured response of sensors and the simulated response of the system models it is possible to judge to what extent the models represent reality. With the aid of systematic algorithms for data assimilation it is then, to some extent, possible to adjust the parameters of the individual grid blocks of the numerical models such that the simulated response better matches with the measured data, and, hopefully, such that the models give better predictions of the future system response. The data assimilation process (‘automatic history matching’ in oil industry vernacular) is indicated at the bottom of Fig. 1.

3. SYSTEM EQUATIONS

3.1. Reservoir models

As an example of a reservoir model we consider the flow of oil and water through a heterogeneous porous medium. We make the strongly simplifying assumptions that the reservoir is horizontal and of constant height, and that gravity and capillary forces can be neglected. As derived in Appendix A, the flow behavior is governed by a set of partial differential equations in terms of the oil pressures and the water saturations. (The water saturation $S_w$ is the fraction of the pore space filled with water. By definition, the
water and the oil saturations add up to one). After discretization of the equations in space into a number of grid blocks e.g. with a finite difference or finite element procedure, we obtain the following system of nonlinear first-order differential equations,

\[
\begin{bmatrix}
V_x p(s) & V_{xw} p \\
V_{wx} s & V_w s
\end{bmatrix} + \begin{bmatrix}
T_o(s) & 0 \\
0 & T_w(s)
\end{bmatrix} \begin{bmatrix}
p \\
s
\end{bmatrix} = \begin{bmatrix}
q_o \\
q_w
\end{bmatrix},
\]

(1)

where \( p \) and \( s \) are vectors of pressures \( p_{xw} \) and water saturations \( S_w \) in the grid block centers, \( V \) is an accumulation matrix (containing the porosity \( \phi \), and the oil, water and rock compressibilities \( c_o, c_w \), and \( c_r \)), \( T \) is a transmissibility matrix (containing the rock permeabilities \( k \), the oil and water relative permeabilities \( k_{ro} \) and \( k_{rw} \), and the oil and water viscosities \( \mu_o \) and \( \mu_w \)), and \( q_o \) and \( q_w \) are vectors of oil and water flow rates with non-zero elements corresponding to grid blocks penetrated by a well. Permeability is a measure of the ability of the rock to transmit fluids, and the relative permeabilities are saturation-dependent multipliers that represent the effect of the presence of one fluid on the flow of the other one. Both \( V \) and \( T \) are functions of \( s \), either directly or through the parameters. In injection wells we can prescribe \( q_o \), while \( q_w \) is equal to zero. In production wells \( q_o \) and \( q_w \) depend on the saturation-dependent relative permeabilities in the near-well area. Therefore we can not directly control them but we can control the total flow rates \( q = q_o + q_w \) according to

\[
\begin{bmatrix}
q_o \\
q_w
\end{bmatrix} = \begin{bmatrix}
F_o(s) \\
F_w(s)
\end{bmatrix} \begin{bmatrix}
q_o \\
q_w
\end{bmatrix},
\]

(2)

where \( F_o \) and \( F_w \) are diagonal matrices with saturation-dependent terms (known as ‘fractional flow functions’) which add more nonlinearities to the system. More complex expressions result if we take into account the effects of capillary pressures and gravity. Moreover, reservoir simulators usually also model a third phase, gas, which has a pressure-dependent oil-solubility. A next step in complexity is obtained when individual hydrocarbon components are modeled, rather than just oil and gas, or when chemical interactions or thermal effects are taken into account; see e.g. Aziz and Settari (1979). However in all these cases it is possible to obtain a set of system equations that can be expressed in a form similar to equation (1), or more compactly, as

\[
V(x)x + T(x)x = F(x)q,
\]

(3)

where \( x = [x^T, s^T]^T \). Equation (3) can be recasted in a generalized nonlinear state space form

\[
g(u, x, \theta) = 0,
\]

(4)

where \( g \) is a nonlinear vector-valued function and \( \theta \) is a parameter vector. The input vector \( u \) is related to the vector of well flow rates \( q \) as \( u = L_w q \), with a selection matrix \( L_w \) that selects the non-zero elements of \( q \), i.e. those elements that correspond to grid blocks penetrated by a well. Often the actual controlled input variables are not the well flow rates but rather the well bore pressures, either in the reservoir or at surface, or the wellhead valve settings, in which case we use additional model equations that can be simply incorporated in equation (4). In general, the input vector \( u \) therefore contains well flow rates, well pressures, and/or valve settings in those grid blocks that are penetrated by wells. Time discretization of the space-discretized system equations is usually performed fully implicitly, which implies that each time step the nonlinear equations are solved iteratively using a Newton-Raphson scheme. Using some form of time discretization, the continuous-time equation (4) can be rewritten as

\[
g(u_s, x_{s+1}, \theta) = 0, \quad x_0 = \bar{x}_0
\]

(5, 6)

where \( k \) is discrete time and where \( \bar{x}_0 \) are the initial conditions. Output variables \( y \), combined in an output vector \( y \), are a function of the input and state variables according to

\[
y_k = h(u_s, \bar{x}_s)
\]

(7)

where \( h \) is a vector-valued function. Typical outputs are wellbore pressures and phase flow rates, either measured at surface or downhole. Typical grid block sizes in a reservoir simulation model are in the order of tens to hundreds of meters in directions aligned with the geological layers, and meters to tens of meters in the direction perpendicular to the layers, and reservoir models may contain from tens of thousands up
to a few millions of grid blocks. Typical simulation time steps are in the order of weeks to months, and a single reservoir simulation of the producing life of a field requires hours to days of computing time.

3.2. Nature of the equations

The governing equations for multi-phase flow through porous media are a set of mildly nonlinear parabolic (diffusion) equations, describing the rate of change of pressure, coupled to a set of strongly nonlinear parabolic-hyperbolic (diffusion-convection) equations, describing the rate of change of phase saturations or component concentrations. The nature of the equations can be recognized in the space-discretized representation in equation (1) where the zeros in the second block-column of the transmissibility matrix $T$ indicate the absence of a diffusive term for the saturations. In reality some small diffusive/dispersive effects will always be present due to capillary forces and sub-grid geological heterogeneities. The very low fluid velocities in reservoir flow imply that inertia effects may usually be neglected. Moreover, the flow is strongly dissipative, such that the response to disturbances is typically over-critically damped and instability of the flow in time is not an issue. (Numerical instabilities during simulation of the discretized equations may of course still occur, notably when the solution methods used are not fully implicit.) The time constants of the pressure equation are typically in the order of hours to months, whereas the diffusive parts of the saturation equations may have time constants up to thousands of years. Also the convective terms (i.e. the fluid velocities) are usually so small that the propagation speeds of oil-water or oil-gas fronts are typically much lower than those of the pressure waves. Under some mild assumptions the system of equations (1) can therefore be rewritten as two separate systems, one for the pressures and one for the saturations, that are only coupled through the coefficients. The pressure equations can then often be approximated as linear with slowly time-varying coefficients. The saturation equations, however, are inherently nonlinear, and in the limit of zero diffusion (i.e. no capillary pressures or geologic dispersion, as in our example) may exhibit typical properties of hyperbolic equations such as shocks and rarefaction waves. The coefficients of the equations are generally very poorly known, and moreover often vary spatially up to four orders of magnitude.

4. OPTIMIZATION

4.1. Optimization methods

For a given configuration of wells, and in particular for a flooding scenario involving multiple injectors and producers, we can use the well rates or pressures to optimize the flooding process over the life of the reservoir. First we will address open-loop optimization, i.e. optimization without data assimilation to update the reservoir model. The objective function can be expressed as:

$$ J(u_{ik}, y_{ik}(u_{ik})) = \sum_{k=1}^{K} J_k(u_i, y_i), $$

where $K$ is the total number of time steps, and where $J_k$ represents the contribution to $J$ in each time step. Note that actually all inputs up to time $k$ may play a role in $J_k$ as follows from recursive application of equations (5) and (7). We could therefore formally write $J_k(u_{i}, y_{i}(u_{i}, x_{i}(u_{i}))),$ but to keep the notation tractable we use $J_k(u_{i}, y_{i})$ instead. An example of $J_k$ in a typical objective function is given by

$$ J_k = \left[ \sum_{i=1}^{N_W} r_{wi} \left(u_{wi} \right)_k + \sum_{j=1}^{N_o} r_{op} \left(y_{op,j} \right)_k + r_{re} \left(y_{re,j} \right)_k \right] \Delta t, $$

where the control variables $u_{wi,i}$ are the water injection rates in wells $i=1,...,N_w$, the output variables $y_{op,j}$ and $y_{re,j}$ are the water and oil production rates in wells $j=1,...,N_o$, $r_{wi}$ and $r_{op}$ are the (negative valued) unit costs for water injection and water production, $r_{re}$ is the unit income for oil production, and $t_k$ and $\Delta t = t_{k+1} - t_k$ are the time and the time interval corresponding to time step $k$. The term in the denominator is a discount factor that represents the time-value of money, where $b$ is the discount rate (cost of capital) for a reference time $\tau$. The objective function (8) with $J_k$ as expressed in equation (9) represents the present value of the oil production minus the present value of the water injection and water

production costs over the life of the field. Constraints can be expressed in terms of the state variables or
the input variables and may be equality or inequality constraints, which we represent in a general form as
\[ c(u_k, x_k) \leq 0. \] (10)

Typical input constraints are limits on the total water injection capacity, and typical state constraints are
maximum and minimum pressures in the injection and production wells respectively. The optimization
problem can now be formulated as finding the input vector \( u_k \) that maximizes \( J \) over the time interval
\[ k = 1, \ldots, K, \] subject to system equations (5), initial conditions (6), output equations (7) and constraints
(10). Sometimes the problem is nonlinear in the inputs and the constraints, and it is nearly always
nonconvex. Many numerical techniques are available to solve this type of optimization problems. In our
work we have been using a gradient-based optimization technique where the derivative information is
obtained through the use of an adjoint equation; see Brouwer and Jansen (2004), Van Essen et al. (2006)
and Zandvliet et al. (2007). Much earlier, adjoint-based techniques were introduced in reservoir
engineering for the optimization of tertiary recovery processes such as polymer or CO\(_2\) flooding; see
by, among others, Virmovsky (1991), Zakirov et al. (1996) and Sudaryanto and Yortsos (2000). However,
industry uptake of these methods was almost absent until quite recent, when the advent of ‘smart well’
and ‘smart fields’ technology caused a revival of interest. Alternative methods to perform field-life
optimization, and in particular those addressing well placement optimization, use ‘non-classical’ methods
such as genetic algorithms; see e.g. Yeten et al. (2003) and Güyagüler and Horne (2004). Moreover,
applications to optimize more complex reservoir flow processes, such as alternating-water-gas (WAG)
injection, are beginning to receive more attention, sometimes in combination with reduced-physics
models such as streamline models or response surfaces generated with experimental design (Esmaiel et
al., 2005).

4.2. Optimal control

A very efficient way to obtain gradients of the objective function \( J \) with respect to the inputs \( u_k \) is given
by ‘optimal control theory’ which makes use of an adjoint formulation; see e.g. Stengel (1994). Once the
gradients have been obtained, a wide variety of gradient-based techniques is available to iterate to a
(locally) optimal solution; see e.g. Gill et al. (1981). Appendix B gives a brief overview of adjoint-based
optimization. Implementation of the adjoint formulation in a numerical reservoir simulator is
conceptually simple if the simulator is fully implicit, because in that case the Jacobian matrix \( \partial g_k / \partial x_k \),
which is required in the adjoint formulation, is already available; see Sarma et al. (2005). In practice the
programming effort is still considerable because of the complexity of modern reservoir simulation
programs, which may contain up to millions of lines of code. Another, more theoretical, problem is the
systematic incorporation of the constraints \( c \) as specified in equation (10). Some possible solutions are
given in Sarma et al. (2006a), Montleau et al. (2006), and Kraaijevanger et al. (2007). One of the
disadvantages of gradient-based techniques is their tendency to arrive at a local optimum rather than a
global one. This is particularly the case if we have a large number of controls (wells) and a large number
of points in time at which we may change the control setting, resulting in a very large number of possible
control trajectories. Several regularization techniques can be applied to ‘smooth’ the control trajectories
and to limit the freedom in choosing control settings. Although this may result, theoretically, in a sub-
optimal global optimum, it will hopefully result in less local optima, and in an increased speed of the
iterative optimization process; see e.g. Stengel (1994). An adaptive multi-scale regularization technique
for water flooding optimization was implemented by Lien et al. (2008). Recently we also started to
investigate the use of adjoint-based techniques to optimize well locations, and the first results are
promising; see Zandvliet et al. (2008).

4.3. ‘Smart well’ optimization

The implementation of a dynamic water flooding optimization strategy, as e.g. obtained with the aid of
optimal control theory, requires the availability of adjustable valves. Mostly, wells are controlled at the
‘well head’, i.e. at the point where the well reaches the surface. A recent development are so-called ‘smart
wells’, equipped with down hole ‘inflow control valves’ (ICVs) which allow for control of the inflow
from the reservoir into individual well segments; see Fig. 2. Initially, the use of smart well technology
was strongly focused on short-term production optimization; see e.g. Naus et al. (2006). However, as
shown in Brouwer and Jansen (2004) and in several publications thereafter, there may be considerable scope to achieve increased ultimate recovery using optimal control over the entire life of the reservoir. If the ICVs are located in a single reservoir the scope to control the flow far away from the well is limited, due to the diffusive nature of the pressure propagation process. Especially if the reservoir is homogeneous, the ‘radius of influence’ around the well is small (Ramakrishnan, 2007). Although in a strongly heterogeneous reservoir there is more scope to control the flow away from a single well equipped with multiple ICVs, in general flow control at a reservoir scale requires wells that are spatially distributed over the reservoir. This could be achieved with multiple horizontal or multi-lateral wells, or with conventional vertical wells, as long as they are equipped with remotely controllable valves (surface chokes and/or ICVs), and some form of pressure and flow rate measurement. In the operational practice of controlling wells it is often more convenient to simply switch off a well rather than to try to keep its production at a predefined rate or pressure. Moreover, on/off valves are also cheaper than continuously variable valves, especially in case of downhole valves which may cost tens of thousands of dollars each. Fortunately, some water flooding control problems appear to have an optimal solution that is close to or sometimes equal to ‘bang-bang’, i.e. it is an optimal strategy to just open or close valves rather than to gradually adjust them; see Zandvliet et al. (2007).

4.4. Robust optimization

One of the major challenges in reservoir engineering is taking decisions in the presence of very large uncertainties about the subsurface structure and the parameters that influence fluid flow. One of the ways to cope with this uncertainty during the field development phase of a reservoir is to use multiple subsurface models, also known as geological realizations. Recently we implemented a robust ensemble-based optimization strategy to maximize the expected value of the objective function $J$ according to

$$
\max_{u_{ik}} E_u \left[ J \left( \mathbf{u}_{ik}, y_{i,k}^N \left( \mathbf{u}_{ik} \right), \theta_{i,k} \right) \right] \approx \max_{u_{ik}} \frac{1}{N_R} \sum_{i=1}^{N_R} J \left( \mathbf{u}_{ik}, y_{i,k}^N \left( \mathbf{u}_{ik} \right), \theta' \right)
$$

where $\theta'$ and $y'$ are the parameter and output vectors of realizations $i = 1, \ldots, N_R$; see van Essen et al. (2006). As an example, consider Fig. 3 which displays a single realization, out of an ensemble of 100, of a reservoir that has a fluvial structure with high-permeability sandstone channels (light gray) in a background of low permeability claystone (dark gray). Fig. 4 displays the results, expressed as cumulative distribution functions of $J$, for three optimization methods as applied to the hundred realizations. The thick dotted black curve corresponds to an often used reactive water flooding strategy, where the production wells are shut-in once the water/oil ratio exceeds a preset maximum. The thin gray curves correspond to ‘nominal’ optimization strategies based on the individual realizations. The thick solid black line curve corresponds to the robust optimization strategy based on hundred realizations. We also applied the same robust strategy to a different set of 100 realizations drawn from the same population of reservoir models, and obtained nearly identical results which confirms the robustness of the strategy (Van Essen et al., 2006). The curves clearly show the value of optimization compared to reactive control, and the additional benefit of a robust optimization strategy: not only is the mean highest for the robust
results, also the standard deviation is lowest. The price to pay is the need to perform forward and adjoint simulations for each realization during every iterations step in the optimization procedure.

Fig. 3. Realization of a channelized reservoir with 4 production wells and 8 water injection wells. (After Van Essen et al., 2006).

Fig. 4. Probability density functions for different control strategies as applied to all 100 ensemble members. Thick dotted black line: conventional, reactive strategy. Thin gray lines: nominal optimizations. Thick solid black line: robust optimization. (After Van Essen et al., 2006)

5. DATA ASSIMILATION

5.1. Formulation as optimization problem

We use the term data assimilation in a broad sense to indicate combined state and parameter updating. In our case this implies updating states $x$ and parameters $\theta$ using measured output data $d$. In practice, the data assimilation problem is more often formulated as an optimization problem for the parameters only, with an objective function defined in terms of the mismatch between measured and simulated output data plus a regularization term:

$$J(y^\times(\theta)) = \sum_i \left[ (d^i - y^i(\theta))^T P^{-1}_y d^i - y^i(\theta) + (\theta - \theta^0)_T P^{-1}_\theta (\theta - \theta^0) \right].$$  \hspace{1cm} (12)

Here $P^{-1}_y$ and $P^{-1}_\theta$ are time-invariant weight matrices which are often chosen as the inverse of the spatial covariance matrices of the measurement errors and the parameters respectively. The counter $i = 1, \ldots, \kappa$ indicates the different points in time at which the measurements are made. The regularization term penalizes large deviations between the updated parameter values $\theta$ and the prior values $\theta^0$. In this way, the prior information more or less fixes the parameter values in those areas where the measurements do not provide sufficient information. The optimization problem is usually solved with the aid of an adjoint-based method; see e.g. Chavent et al. (1975) and Li et al. (2003). If the states are also updated, the
uncertainty is taken into account with an additional time-varying covariance matrix \( P_i \). Joint updating of uncertain states and parameters, including the time-varying error and covariance terms, can be done using adapted versions of the ‘representer method’ which was first introduced for state estimation in very large ocean models; see Bennett (2002). For early applications to reservoir engineering, including an extension to parameter estimation, see Rommelse et al. (2006), and Przybysz-Jarnut et al. (2007). An alternative way to estimate uncertain state variables in very large nonlinear models is through the use of the ensemble Kalman filtering method, as developed in oceanography based on the much older conventional Kalman filter. An application to reservoir engineering, including parameter estimation, will be briefly discussed below. Specialized model updating methods have been developed in the reservoir engineering community using, e.g., streamline simulation to rapidly derive sensitivities of saturation changes along streamlines (Vasco et al., 1999). Other specialized methods perform parameter updating under geostatistical constraints, such as the probability perturbation method (Caers, 2003) and the gradual deformation method (Roggero and Hu, 1998), or emphasize the quantification of uncertainty (Erbaş and Christie, 2007).

5.2. Ensemble Kalman filtering

As is well known in the process control community, state estimation can be effectively performed sequentially, i.e. whenever data becomes available, with the aid of a Kalman filter. It can be shown that for linear systems, and assuming Gaussian measurement and process noise, the representer method results in exactly the same answers as the, much older, Kalman filter approach (Bennett, 2002). For nonlinear problems, the ordinary Kalman filter breaks down because the nonlinearity results in non-Gaussian noise when propagated through the system. In the ensemble Kalman filter (EnKF) the analytical error propagation is replaced by a Monte Carlo approach, in which the model error covariance is computed from an ensemble of model realizations that are all propagated in time. The EnKF method has proved to be very successful in oceanographic applications where very large models, containing millions of state variables, are frequently updated using a variety of data sources; see Evensen (2006). During the forecast step a simulation is run for each of the model realizations up to the time where new measurements become available. All realizations in the model are updated by combining the new real measurements with forecasted measurements from the ensemble. Recently a large number of publications have appeared that apply the EnKF to reservoir engineering problems; see e.g. Nævdal et al. (2005) for an early application and Evensen (2006) for an overview. These implementations of the EnKF also treat parameters as unknowns, which leads to the use of an extended state vector \( \mathbf{x} = [\mathbf{x}^T \theta^T]^T \). Model updating using the EnKF relies on the cross-covariance between the measurements and the (extended) state. However, because the EnKF uses a low-order representation of the cross-covariance matrix, based on a relatively small ensemble, spurious updates may occur. Various ‘localization’ schemes are therefore being investigated to restrict updates to regions close to the measurements; see e.g. Devegowda et al (2007). Other outstanding system-theoretical aspects of the EnKF include understanding the conditions that sometimes lead to filter divergence and the performance in case of non-Gaussian statistics.

5.3. Parameterization

In our parameter and state estimation problems we are dealing with a very large number of ‘inputs’ (parameters and states) that need to be adjusted to obtain a best match between modeled and real data. A typical reservoir model may contain millions of unknown parameters, such as grid block permeabilities and porosities, fault transmissibilities and initial conditions. Fortunately most of these parameters display spatial correlations that can be used to reduce the dimension of the parameter space, and various techniques to regularize the parameter estimation problem have been proposed using, e.g., zonation, wavelets, Karhunen-Loève decomposition or discrete cosine transforms; see Jafarpour and McLaughlin (2007) for a recent overview. It has been shown that it is also possible to make use of spatial correlations in the states (pressures, saturations) to reduce the order of reservoir models using system theoretical techniques, but application of these possibilities either in optimization or in data assimilation has hardly yet been pursued. For some early attempts, see Heijn et al. (2004), Van Doren et al. (2006), Markovinović and Jansen (2006), and Gildin et al. (2006). In general the amount of information that can be obtained from well data is rather limited, especially because the pressure propagation through a reservoir is a diffusive process. Sometimes it is possible to obtain areal information through the repetition of seismics in time, which may give an indication of those reservoir areas where pressures or saturations have...
changed; see e.g. Skjervheim et al. (2007). However, the data obtained from production measurements and time-lapse seismics are never sufficient to fully characterize the states and parameters in a traditional reservoir flow model, and data assimilation in reservoir engineering is therefore an inherently ill-posed problem. Especially if reservoir models are used for field re-development planning, involving e.g. the drilling of new wells, geological models are essential to constrain the solution space of the data assimilation problem.

### 6. CLOSED-LOOP RESERVOIR MANAGEMENT

Finally, we consider an example of full closed-loop reservoir management, as indicated in Fig. 1, by combining optimization with data assimilation. The results of this example were taken from Overbeek et al. (2004), and are comparable to other early results reported in Brouwer et al. (2004), Nævdal et al. (2006) and Sarma et al. (2006b). Fig. 5 depicts the top view of a two-dimensional reservoir model with 110 x 110 grid blocks, which was used as ‘truth’ to generate synthetic ‘measured’ data. We will refer to this model as the synthetic reservoir model (SRM). Its only function is to test the closed-loop reservoir management concept, and in a real implementation it would be replaced by a real reservoir that would produce real oil and real measured data. The 10 crosses at the left of the SRM represent a row of vertical water injection wells, and the 10 circles at the right a row of producers. Just as in the previous example, the reservoir contains high-permeable channels (light) amidst a low-permeable background (dark). We assumed that noisy measurements of pressures and total liquid (oil plus water) rates were available in all wells. We used optimal control theory with a steepest descent method for the flooding optimization and the EnKF method for updating the unknown grid block states (pressures and saturations) and parameters (permeabilities) in an ensemble of 100 coarse reservoir models of 10 x 10 grid blocks each. The optimization objective was oil revenue minus water production costs over the total field life. The controlled variables (i.e. the input variables for the optimization) were the total rates in the injectors and the producers. The total water injection rate in the 10 injectors was constrained to remain constant and equal to the total liquid production rate in the 10 producers. Therefore, the optimization problem can, for this simple example, be interpreted as the question how to optimally distribute fixed total rates of injected water and produced liquids over the 10 injectors and the 10 producers respectively such that the oil production becomes maximum over the reservoir life.

The injection and production rates form a control vector with 20 components, and we refer to the evolution of this vector in time as the control trajectory. A nominal control trajectory is defined as one followed during ‘conventional’ water flooding where the rates are such that the pressures in all injection wells are equal, and the pressures in all production wells are equal (but lower than in the injection wells). The first row in Fig. 6 depicts three snapshots in time corresponding to such a nominal control trajectory. As expected, the injected water rapidly flows through the highly permeable channels, which results in early water production in some of the producers. After one pore volume of water has been injected all oil could, in the ideal case, have been produced, but it is clear that because of the heterogeneous reservoir structure a lot of oil has been left behind; see the top-right figure. The second row of Fig. 6 depicts the results corresponding to an optimal control trajectory obtained using the following procedure:

1. **Start water injection and liquid production in the SRM using a nominal control trajectory up to the first measurement time.**
2. **Simulate the flow in all 100 ensemble members up to the first measurement time.**
3. **Obtain ‘measured’ production data from the SRM.**
4. **Update the states and model parameters in all 100 ensemble members.**
5. **Compute an optimal control trajectory for the remaining field life.**
6. **Continue water injection and liquid production in the SRM using the optimal control trajectory up to the next measurement time.**
7. **Simulate the flow in all 100 ensemble members up to the next measurement time.**
8. **Return to step 3.**

End do loop.
Fig. 5. Top view of the ‘true’ reservoir used to generate synthetic data. (After Overbeek et al., 2004)


Note that the optimization is performed over the entire remaining field life, but that the resulting optimal control trajectory is only implemented up to the next measurement time step. The figures at the third row show that the initial average permeability estimate is nearly uniform ($t = 0$ days), but that after a while a heterogeneous pattern has emerged ($t = 116$ days) that does not really change very much any more until the end of the flooding period ($t = 750$ days). As follows from comparison of the final (rightmost) figures in the first two rows, the optimized water flooding strategy results in a significantly improved oil recovery. In this example the initial ensembles did not have a marked heterogeneity, but just a Gaussian random spatial structure, and therefore we based the optimization on the ensemble average, rather than
using a robust strategy as in the previous example. The fourth row in Fig. 6 shows the relative distribution of the flow rates over the 10 injection wells and illustrates that the optimization results in a dynamic strategy of changing the flow rates over time. Although extremely simplified, this example illustrates the basic concept of closed-loop reservoir management. In our current research we apply the concept to models and data of real reservoirs, modeled with multiple ensembles, where each ensemble is based on a different geological scenario, and we aim at extending it to more complex flooding mechanisms, geological settings and data sources.

7. DISCUSSION

7.1. System-theoretical aspects

The idea of closed-loop reservoir management and production optimization has been described in different forms before; see e.g. Chierici (1992), or Nyhavn (2000), with further references given in Jansen et al. (2005). However none of these earlier papers makes use of systematic techniques for both optimization and data assimilation. It could be argued that the closed-loop procedure as described here is a form of non-linear model predictive control (NMPC) with a gradually ‘shrinking horizon’, although our approach is not aimed at following a predefined optimal trajectory, as is NMPC used in the process industry, but rather at finding that optimal trajectory to begin with. Closed-loop reservoir management could therefore also be classified as a form of real time optimization (RTO). However, due to the slow dynamics of the oil production process, both the data assimilation and the optimization may be performed off-line, unlike in true RTO. Still, a key aspect of our approach is that the data assimilation and model-based optimization are performed at a much higher frequency than in conventional reservoir management, e.g. once every few weeks instead of once every few years. This makes it arguably a ‘real-time’ procedure at the timescale governed by the front displacement between wells, which could be anywhere between months to decades. The examples shown in this paper are very simplistic, and in a realistic field, with realistic well constraints, the scope for optimization will be smaller. In addition, many more model parameters than just the permeabilities will be to a large extent unknown, e.g. porosities, reservoir boundaries and initial conditions, such that the parameter estimation problem becomes even more formidable. However, the examples illustrate some essential aspects of our approach:

- Systematic optimization of well rates over the producing life of a reservoir produced with water flooding offers scope for increased oil recovery and reduced water production.
- The effect of uncertain reservoir parameters can be reduced through 1) robust optimization over an ensemble of reservoir models, and 2) regular updating of the models using production data.
- A relatively simple reservoir model may still give acceptable results when used for optimization of a fixed configuration of injection and production wells.

Especially the last point raises some interesting system-theoretical questions which are topic of our current research. The observability of reservoir pressures (which are required to estimate the permeabilities) from the wells is probably very small because of the diffuse nature of pressure propagation in porous media. However the controllability of the pressure field (which drives the saturation changes) is equally limited, which may explain why a relatively simple model works so well to optimize flooding in a fixed configuration of wells. We note that this would to a much lesser extent be true if we were to drill new wells, in which case additional geological information would be required. Related to the question to what extent the state variables are observable is the question to what extent the model parameters are identifiable. Some early attempts to analyze these questions systematically have recently been reported in Van Doren et al. (2008) and Zandvliet (2008).

7.2. Operational aspects

Another, more practical, aspect that may benefit from input from the process control community involves the combination of short-term production optimization and long-term reservoir management. An open question is how to best implement the optimal trajectory, which has been computed off-line, in the daily operation of an oil field. Traditionally, day-to-day valve settings are chosen such as to maximize instantaneous oil production, limited by constraints on the processing capacities of gas and water that are co-produced with the oil. These settings are usually determined with heuristic operating protocols, with

the occasional use of off-line model-based optimization, such as sequential linear or quadratic programming, to optimize instantaneous production; see Bieker et al. (2007) for a recent review. In addition, simple on-line feedback control is used to stabilize the flow rates and pressures in the processing facilities that separate the oil, gas and water streams from the wells. The combination of optimization and control at different time scales requires a layered control structure where longer-term optimization results provide set points and constraints for the short-term optimization, which in turn provides set points for field controllers. Such a systematic control structure, which is well known in the process industry, has been proposed for use in oil field management by Saputelli et al (2006), but heuristic approaches still dominate the industry.

8. CONCLUSION

There is ample scope to use a variety of results from process control theory and practice to further develop the techniques for model-based control of multiphase flow in subsurface oil reservoirs. This concerns techniques for off-line optimization, parameter estimation and model reduction, and in particular systematic methods to combine off-line, long-term reservoir management with short-term production optimization and on-line process control.

REFERENCES


Conference on the Mathematics of Oil Recovery (ECMOR X), Amsterdam, Netherlands, 4-7 September.


**ACKNOWLEDGEMENTS**

Financial support for our research is provided by the following organizations: ISAPP (sponsored by TNO and Shell), VALUE (sponsored by Shell and Senter/Novem), DELPHI (multiple industry sponsors), and the Kuwait Institute for Scientific Research.
APPENDIX A: MULTIPHASE FLOW THROUGH POROUS MEDIA

As in all branches of fluid mechanics, the physics of flow through a porous medium can be described with the aid of partial differential equations that represent conservation of mass, momentum and energy, and equations of state that describe the fluid properties as a function of pressure and temperature; see e.g. Aziz and Settari (1979). Except in cases of steam flooding or in-situ combustion we can assume that reservoir flow is isothermal, which implies that we may disregard the energy balance equation. Moreover, the movement of fluids is usually so slow that we can disregard inertial effects, and that instead of the momentum balance equation we may use an empirical relationship between pressure drop and flow velocity known as Darcy’s law. For the simple case of two-phase (oil-water) flow we can write the mass balance equation for each phase in vector notation as

\[ \nabla \cdot (\rho_i \mathbf{v}_i) + \frac{\partial (\rho_i \phi S_i)}{\partial t} - \rho_i q_i^\prime = 0, \]  

(A.1)

where \( \rho \) is fluid density, \( \mathbf{v} \) is (superficial) fluid velocity, \( \phi \) is porosity, \( S \) is fluid saturation of the pore space (0 \( \leq \) S \( \leq \) 1), \( t \) is time, \( q'' \) is flow rate per unit volume, and the subscript \( i \in \{o, w\} \) indicates the oil and water phases respectively. Darcy’s law can be expressed as

\[ v_i = -\frac{k_{ri}}{\mu_i} (\nabla p_i - \rho_i g \nabla d), \]  

(A.2)

where \( K \) is the permeability tensor, \( \mu \) fluid viscosity, \( k \) relative permeability, \( p \) pressure, \( g \) acceleration of gravity and \( d \) depth. The permeability tensor \( K \), whose elements have units of surface area, represents how easily the fluids flow through the rock in different directions. Usually the orientation of the coordinate system can be aligned with the geological layering in the reservoir such that \( K \) is a diagonal matrix:

\[ K = \text{diag}(k_x, k_y, k_z), \]  

(A.3)

where \( k_x, k_y \) and \( k_z \) are directional permeabilities in the \( x, y \) and \( z \) coordinate directions. The dimensionless relative permeabilities \( k_{ri} \) are functions of \( S_i \), and are reduction factors that represent the increase in flow resistance caused by multi-phase effects. The resistance to concurrent flow of oil and water is generally much higher than the sum of the resistances to flow of the individual phases, and the relative permeabilities are therefore a major source of nonlinearity in the multi-phase equations. Combining equations (A.1) and (A.2) results in

\[ -\nabla \left[ \frac{\rho k_{wn}}{\mu_w} K \left( \nabla p_w - \rho g \nabla d \right) \right] + \frac{\partial (\rho_S \phi)}{\partial t} - \rho q_w^\prime = 0. \]  

(A.4)

Equations (A.4) (one for each phase) contain four unknowns, \( p_o, p_w, S_w \) and \( S_o \), two of which can be eliminated with aid of the relationships

\[ S_o + S_w = 1, \quad p_o - p_w = p_c(S_w) \]  

(A.5, A.6)

where \( p_c(S_w) \) is the oil-water capillary pressure which is another source of nonlinearity in the flow equations. Substituting equations (A.5) and (A.6) in equations (A.4), expanding the right-hand sides, applying chain-rule differentiation, and substituting oil, water and rock compressibilities

\[ c_o = \frac{1}{\rho_o} \frac{\partial \rho_o}{\partial p_o} \bigg|_{p_o}, \quad c_w = \frac{1}{\rho_w} \frac{\partial \rho_w}{\partial p_w} \bigg|_{p_w}, \quad c_r = \frac{1}{\phi} \frac{\partial \phi}{\partial p_w} \]  

(A.7, A.8, A.9)

allows us to express equations (A.4) in terms of \( p_o \) and \( S_w \) as follows:

\[ -\nabla \left[ \frac{\rho k_{wn}}{\mu_w} K \left( \nabla p_w - \frac{\partial \rho_o}{\partial p_o} \nabla S_o - \rho_w g \nabla d \right) \right] + \rho_o \phi \left[ S_o(c_o + c_r) \frac{\partial p_o}{\partial t} + \frac{\partial S_o}{\partial t} \right] - \rho_w q_w^\prime = 0. \]  

(A.10)
\[-\nabla \cdot \left( \frac{\rho k}{\mu_s} \mathbf{K} (\nabla p_o - \rho g \nabla h) \right) + \rho_o \phi \left( 1 - S_w \right) (c_o + c_e) \frac{\partial p_o}{\partial t} - \frac{\partial S_w}{\partial t} \right] = 0. \tag{A.11}
\]

The two-phase flow equations as formulated in expressions (A.10) and (A.11) contain two state variables: the oil pressure \( p_o \) and the water saturation \( S_w \). The equations are nonlinear because of the saturation dependence of the capillary pressure \( p_c \) and the relative permeabilities \( k_r \). In the more general case there may also be a pressure dependency of the densities \( \rho \), the porosity \( \phi \), and the compressibilities \( c \), in particular if the formulation is extended to include gas as a third phase. Boundary conditions are usually chosen as no-flow at the entire boundary, with occasionally a constant-pressure condition for part of the boundary if only a segment of a reservoir is modeled. Initial conditions are specified for pressure and saturation.

**APPENDIX B: ADJOINT-BASED OPTIMIZATION**

Consider the optimization problem

\[
\max_{u_{ik}} J(u_{ik}, y_{ik}, \lambda_{ik}, \mu_{ik} \mid u_{ik}), \tag{B.1}
\]

with objective function (8). We aim to compute the optimal control \( u_{ik} \) with the aid of a gradient-based algorithm, which requires the derivatives of \( J \) with respect to \( u_{ik} \). The problem in determining the derivatives is the indirect dependence of the variation \( \delta J \) in the objective function on a variation \( \delta u_{ik} \) of the input. A variation \( \delta u_{ik} \), of element \( i \) of vector \( u \) at time \( k \) does not only directly influence \( J \) at time \( k \), but also, as follows from recursive application of equation (5), the states, which in turn, through equation (7), influence the outputs and thus \( J \) at later times. The effect of a single variation \( \delta u_{ik} \) should therefore be computed, using the chain rule for differentiation, as

\[
\delta J = \sum_k \left( \frac{\partial J_k}{\partial u_k} + \sum_i \sum_x \left( \frac{\partial y_x}{\partial u_k} + \frac{\partial x_k}{\partial u_k} \right) \frac{\partial x_k}{\partial u_{ik}} \right) \delta u_{ik} \tag{B.2}
\]

where \( \kappa \) is a dummy variable. The term \( \delta x_k / \delta u_k \) gives problems because we need to solve the recursive system of equations (5) to connect the state vectors \( x_k \), \( k = k + 1, \ldots, K \) to the input \( u_k \) at time \( k \). This dependence can be taken into account by considering equation (5) as a set of additional constraints, and applying the technique of Lagrange multipliers to solve the constrained optimization problem; see e.g. Stengel (1994). Moreover, we also consider the initial condition (6) and the output equation (7) as constraints, and, setting aside the ‘ordinary constraints’ \( c \), we can therefore define a modified objective function

\[
\tilde{J}(u_{ik}, x_{ik}, y_{ik}, \lambda_{0k}, \mu_{0k} \mid u_{ik}) = \sum_k \left( J_1 + \lambda_1 (x_0 - \bar{x}_0)^T T_k \right) + \sum \lambda_k \left( g_{k1} (u_{ik}, x_{ik}) + \mu_k (y_{ik} - h_{ik} (u_{ik}, x_{ik})) \right), \tag{B.3}
\]

where the constraints have been ‘adjoined’ to \( J \) with the aid of vectors of Lagrange multipliers \( \lambda \) and \( \mu \). The Kronecker delta \( \delta \) ensures that the initial condition constraint is included in the summation. A necessary condition for an optimum is stationarity of \( \delta \tilde{J} \) for all variations, which leads to:

\[
\begin{align*}
\frac{\partial J_1 + \lambda^T R_1 \lambda}{\partial u_{ik}} + \frac{\partial g_{k1}}{\partial u_{ik}} + \mu_k \frac{\partial h_{k1}}{\partial u_{ik}} &= 0^T \\
\lambda^T R_1 \lambda + \lambda^T R_1 \lambda &= 0^T \\
\lambda^T R_k \lambda - \mu_k \frac{\partial h_k}{\partial x_k} &= 0^T \\
\lambda^T \frac{\partial h_k}{\partial x_k} + \mu_k \frac{\partial h_k}{\partial x_k} &= 0^T
\end{align*}
\]

\[ \frac{\partial J_k}{\partial y_k} + \mu_k = 0^r \]  
(B.8)

\[ (x_0 - \bar{x}_0)^T = 0^r \]  
(B.9)

\[ g^T (u_{k+1}, x_{k+1}) = 0^r \]  
(B.10)

\[ [y_{k+1} - h(u_{k+1}, x_{k+1})]^T = 0^r \]  
(B.11)

Equations (B.11), (B.10) and (B.9) are identical to output equation (7), system equation (5) and initial condition (6) respectively, and are therefore automatically satisfied. Equation (B.8) allows us to compute the Lagrange multipliers \( \mu_k \). Next we can use equation (B.7) to compute multiplier \( \lambda_k \) for the final discrete time \( K \), and thereafter the discrete-time differential equation (B.6) to recursively compute the multipliers \( \lambda_k \) for \( k = K-1, \ldots, 0 \), i.e. backward in time. Equation (B.5) represents the effect of changing the initial condition \( x_0 \) on the value of the objective function, while keeping all other variables fixed. However, because we prescribed the initial condition through equation (6) this term is in our case only of theoretical relevance. Finally, equation (B.4) represents the effect of changing the control on the value of the objective function, while keeping all other variables fixed. For a non-optimal control this term is not equal to zero, but then its residual is just the derivative (i.e. the transposed gradient) that we require to iteratively obtain the optimal control using a gradient-based algorithm. Solution of the optimization problem now consists of choosing an initial control vector \( u_{1:K} \) and repeating the following steps:

1. Compute the states \( x_{1:K} \) and outputs \( y_{1:K} \) using equations (5), (6) and (7).
2. Compute the value of the objective function \( J \) using equation (8). If converged stop, else continue.
3. Compute the Lagrange multipliers \( \mu_{1:K} \) and \( \lambda_{1:K} \) using equations (B.8), (B.7) and (B.6).
4. Compute the derivatives of the objective to the controls from the residuals of equation (B.4):

\[ \frac{\partial J_k}{\partial u_k} = \frac{\partial J_k}{\partial u_{k+1}} + \lambda_k \frac{\partial g_k}{\partial u_{k+1}} \]  
(B.12)

5. Compute an improved estimate of the control vector \( u_{1:K} \), using the derivatives as obtained from equation (B.12), and a gradient-based minimization routine of choice.