Proceedings of the 2012 IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Norwegian University of Science and Technology, Trondheim, Norway, May 31 - June 1, 2012

FrPIT3.1

Recent developments in model-based optimization and control of subsurface flow in oil reservoirs [*]

Paul M.J. Van den Hof [*],[**], **Jan Dirk Jansen** [***] and **Arnold Heemink** [****]

* Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands
** Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands (email: p.m.j.vandenhof@tue.nl)
*** Department of Geoscience and Engineering, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands (email: j.d.jansen@tudelft.nl)
**** Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands.

Abstract: The past ten years have seen an increasing application of systems and control theory to porous media flow. This involves in particular the use of optimization, parameter identification, and model reduction techniques in attempts to increase the amount of oil or gas that can be recovered from subsurface hydrocarbon reservoirs. Other applications involve the control of ground water flow for drinking water or pollution control, and the subsurface storage of CO_2. The dynamic behavior of subsurface multi-phase porous media flow is typically simulated with large-scale nonlinear numerical models, containing up to millions of state variables and parameters. Moreover, a typical characteristic of these models is a very large uncertainty in the parameter values, reflecting the very large geological uncertainty of the subsurface. Traditionally they are primarily used for 'field development', i.e. the engineering of well configurations and production strategies, but an emerging use is in the 'real-time' optimization and control of oil production, known as 'closed-loop reservoir management (CLRM)'. In this paper we describe some recent contributions of our group to the use of systems and control theory for CLRM. This concerns sequential and multi-level production optimization, identifiability of model parameters, and 'control-relevant' upscaling.

Keywords: Dynamic optimization; model-based control; process control; reservoir engineering; data assimilation; model predictive control; real-time optimization; system identification; parameter estimation; identifiability; model reduction.

1. INTRODUCTION

The production of oil and gas from petroleum reservoirs is highly dynamic process of which the operation classically is driven by decisions based on operator experience and supported by scenario studies. Because of economic needs, and enabled by newly developed technology for drilling wells and instrumenting wells with actuators and sensors, there is an increasing opportunity for model-based control and optimization to develop rational and model-based decision support systems so as to optimize the economic efficiency of the process, Jansen et al. [2008], Van den Hof et al. [2009], Foss [2011]. However the size and complexity of the underlying systems, as well as the uncertainties in their models, raise additional challenges for the systems and control tools that can be applied.

Oil and gas reside inside the pores of subsurface rock layers. The production life of a petroleum reservoir generally lasts a number of decades and usually two or three production phases can be identified. In the primary production phase wells are drilled from the surface into the reservoir. The over-pressurized reservoir provides the driving force to push the oil to the surface. Although this process requires no additional effort than drilling wells, unfortunately only 5%-15% of the total amount of oil can be recovered in this manner. When the reservoir pressure decreases, the production rates drop as well. To maintain a preferred production level, the application of an external force is required. This is generally done by either injecting fluids (water or gas) into the reservoir or by installing pumps. Somewhere between 20%-70% of the oil can be recovered using secondary production methods. Tertiary production methods are aimed at changing the properties of the fluids (oil, gas or water) to improve recovery, as e.g. injecting steam, adding polymers to injected water or by injecting surfactants.

Although a *recovery factor* approaching 100% can theoretically be reached (c.f. the clean-up of hydrocarbon spills with chemical pollution recovery measures), the economical threshold lays on average between 30 to 50%. This is partly due to the relatively large costs associated with secondary and tertiary methods, but it is also the result of an inefficient deployment of these methods, due to a lack of knowledge on how to operate these phases in a structured and economically efficient way.

Reservoir simulation is traditionally applied during the development phase of an oil field to determine e.g. the optimal position of the wells, the optimal capacity of the processing facilities, or the best recovery technique (e.g. depletion drive in which case oil is produced through depletion of the natural reservoir pressure, or reservoir 'flooding' in which case oil is produced by injecting other fluids such as water, gas or chemicals in the reservoir). More recently it has been proposed to also use reservoir simulation during the operational phase of oil production. The underlying hypothesis is that it will be possible to significantly increase life-cycle value by changing reservoir management from a periodic to a near-continuous model-based controlled activity; see Jansen et al. [2009].

In this paper we will give an overview of some recent developments in model-based control and optimization of oil production, in particular for the secondary production method of *waterflooding* with the aim to improve economic performance over the producing life of the reservoir. The emphasis of the paper is on recent contributions, in particular obtained in the authors' own research group, and we do not intend to give a complete review. However, we do mention other review papers and textbooks covering the historic development of optimization and parameter identification methods related to reservoir simulation.

The paper proceeds as follows. In Section 2 the process of waterflooding is explained and a (general) reservoir model is summarized. After presenting the recovery optimization problem in Section 3, we address the principle limitations and challenges in Section 4. In four subsequent sections we then address several new developments, in terms of parameter estimation, balancing long-term and short-term optimization, time-scale separation and control-relevant upscaling.

2. WATERFLOODING

2.1 Introduction

Waterflooding is the most popular secondary recovery method. To get an idea, over 50% of USA oil production is due to waterflooding. The goal of the method is to maintain or increase reservoir pressure to boost production and to displace oil from the pores of the reservoir rock and replace it by water.

Oil is produced through *production* wells through which the oil flows to the surface. For the injection of water particular *injection* wells are drilled. Over the recent years, increasingly more so-called *smart* or *intelligent* wells are drilled. These smart wells are equipped with multiple downhole variable control valves, each for a different section of the well. Using these valves, the amount of water
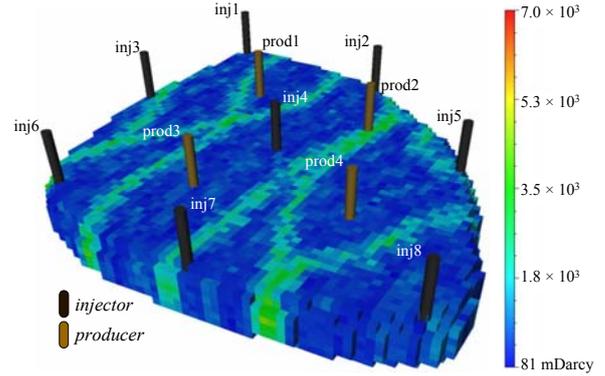
Fig. 1. *3D reservoir model with 4 production and 8 injection wells (van Essen et al. [2009]). The geological structure involves a network of meandering channels in which the fluids flows experience less resistance, due to higher permeability.*

injected into and oil produced from a specific geological layer can also be controlled. These smart wells vastly expand the opportunities for control.

An example of an -academic- 3D reservoir model is depicted in Figure 1. Its geological structure involves a network of fossilized meandering channels in which the flowing fluids experience less resistance, due to higher permeability. The reservoir model contains 8 injection wells that inject water and 4 production wells.

The injected water travels through the reservoir, away from the injection wells, but generally not in a uniform manner. This is the result of the strong heterogeneous nature of reservoir rock. The block diagram in Figure 2 shows (schematically) the multiple actuation inputs and measurement outputs available to control the water flooding process.

When in a particular region the oil-water front reaches a production well, water is produced and when the water fraction exceeds an economic limit the production well is closed (shut in). Then the production capacity of this well inlet is lost, and oil that is left in the reservoir may have been by-passed. The challenge of control here is to dynamically manipulate the spatially distributed injector and producer valves so as to achieve maximum oil production over the lifetime of the reservoir.

2.2 System equations

The waterflooding process is described by a reservoir model based on conservation of mass and momentum equations. We will limit attention to reservoirs with oil and water as the only two components involved. For a general introduction into reservoir modeling see Aziz and Settari [1979]. The mass balance is then given by:

$$\nabla \cdot (\rho_i u_i) + \frac{\partial}{\partial t} (\phi \rho_i S_i) = 0, \quad i = o, w, \tag{1}$$

where t is time, ∇ the gradient operator $(\frac{\partial}{\partial x} \ \frac{\partial}{\partial y} \ \frac{\partial}{\partial z})^T$, ϕ is the porosity (volume fraction of void space), ρ_i is the
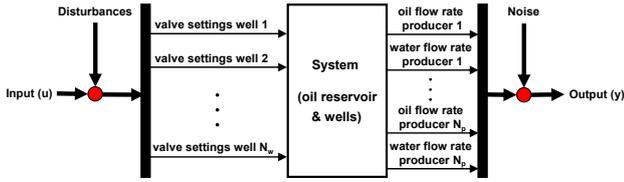
Fig. 2. *The inputs involve the pressures, or the total flow rates (sum of oil and water rates), or the valve settings of the wells. The outputs are the pressures, or the phase flow rates (individual oil and water rates) of the wells. If a pressure in a well is used as input, the rates become the output and vice versa. Note that in an injector only water rates are relevant.*

density of the phase i, u_i the superficial velocity, S_i the saturation, defined as the proportion of the pore space occupied by phase i, where o relates to the oil phase and w to the water phase. The first term in equation (1) represents the divergence of the vector field $\rho_i u_i$, while the second term represents the storage due to rock and fluid compression.

Conservation of momentum is described by the semi-empirical Darcy's equation (discarding gravity):

$$u_i = -k \frac{k_{ri}}{\mu_i} \nabla p_i, \quad i = o, w, \qquad (2)$$

where p_i is the pressure of phase i, k is the absolute permeability, k_{ri} is the relative permeability and μ_i is the viscosity of phase i. The permeability k is an inverse measure of the resistance a fluid experiences flowing through the porous medium. The relative permeability k_{ri} relates to the additional resistance phase i experiences when other phases are present. These relative permeabilities are strongly non-linear functions of the water saturation S_w, introducing nonlinear dynamics into the system.

Equations of state for oil and water relate the phase densities to the pressures, while a similar relationship relates the porosity to the pressure. In addition two closure equations complete the model, i.e.:

$$S_o + S_w = 1, \qquad (3)$$

and secondly the capillary pressure equation:

$$p_{cow} = p_o - p_w = f_{cow}(S_w). \qquad (4)$$

As a result a set of equations remains where typically the oil pressure p_o and water saturation S_w are chosen as primary state variables, , although other combinations are possible. The resulting set of coupled nonlinear partial differential equations has a dual character. The pressure behaviour is essentially diffusive, i.e. the corresponding equations are parabolic and become elliptic in the limit of zero compressibility. The saturation behaviour is mixed diffusive-convective, i.e. the corresponding equations are mixed parabolic-hyperbolic and become completely hyperbolic in the case of zero capillary pressure. The saturations can only be controlled indirectly through influencing the spatial gradients of the pressures. From a systems and control perspective the key consequence is that manipulating the inputs (i.e. the well settings) results in near

instantaneous changes in the pressures away from the wells but in (possibly very strongly) delayed changes in the saturations in those locations.

After discretization in space, leading to a system built up of a finite number of cells, referred to as *grid blocks*, and discretization in time the following state space form results:

$$g_k(u_k, x_k, x_{k-1}, \theta) = 0, \quad k = 1, ..., K, \qquad (5)$$

where subscripts refer to discrete instants k of time, with K the total number of time steps. Vector function g takes its values in $\mathbb{R}^n$, $x_k \in N \subset \mathbb{R}^n$ is a vector of reservoir state variables, $u_k \in M \subset \mathbb{R}^m$ is a vector of control variables in the wells, and $\theta \in \Theta \subset \mathbb{R}^{n_\theta}$ is a vector of model parameters. In isothermal reservoir simulation the state variables are typically pressures and phase saturations, or pressures and chemical component accumulations in each cell of the simulator. The control vector u_k can represent a combination of prescribed well flow rates, well bore pressures (either at surface or down hole) or valve settings. The parameter vector θ typically contains porosities, permeabilities (i.e. inverse resistances to flow) in each simulator cell, but may also include other parameters such as fluid properties or initial conditions). Outputs $y \in H \subset \mathbb{R}^h$ typically consist of measured pressures and flow rates in the wells, but may also be more indirect measurements such as interpreted time-lapse seismic data. For the model outputs we use a simple, explicit, relationship between inputs, states, parameters and outputs as

$$\hat{y}_k(\theta) = h_k(u_k, x_k, \theta), \quad k = 1, ..., K, \qquad (6)$$

The sets H, N, M and Θ are subsets of the set of real numbers because their elements are constrained to stay within physical limits; pressures and permeabilities, e.g., are always positive, and saturations and porosities have values between zero and one. For a more detailed introduction to reservoir simulation concepts from a process control perspective see Jansen et al. [2008]; for a classic textbook on reservoir simulation see Aziz and Settari [1979]. This text also describes more complex situations where e.g. an additional (gas) phase, or multiple chemical components, or a varying temperature is taken into account. Starting from given initial conditions x_0, the implicit recursive Eq. (5) is typically solved with the aid of a time stepping algorithm using Newton iteration to minimize the residual at each time step to a preset tolerance.

In the following we will frequently use a second, even more compact, notation to represent the simulator as

$$\mathbf{g}(\mathbf{u}, \mathbf{x}, x_0, \theta) = \mathbf{0}, \qquad (7)$$

and an explicit output equation

$$\hat{\mathbf{y}}(\theta) = \mathbf{h}(\mathbf{u}, \mathbf{x}, \theta), \qquad (8)$$

where $\mathbf{g}, \mathbf{h}, \mathbf{u}, \mathbf{x}, \mathbf{y}$ and $\hat{\mathbf{y}}(\theta)$ can be interpreted as concatenated time-sequenced vectors

$$\mathbf{g} = [g_1^T, g_2^T, ..., g_K^T]^T, \qquad (9)$$

$$\mathbf{h} = [h_1^T, h_2^T, ..., h_K^T]^T, \qquad (10)$$

$$\mathbf{u} = [u_1^T, u_2^T, ..., u_K^T]^T, \quad \mathbf{u} \in Q \subset \mathbb{R}^q, \qquad (11)$$

$$\mathbf{x} = [x_1^T, x_2^T, ..., x_K^T]^T, \quad \mathbf{x} \in P \subset \mathbb{R}^p, \qquad (12)$$

$$\mathbf{y} = [y_1^T, y_2^T, ..., y_K^T]^T, \qquad (13)$$

$$\hat{\mathbf{y}}(\theta) = [\hat{y}_1^T(\theta), \hat{y}_2^T(\theta), ..., \hat{y}_K^T(\theta)]^T. \qquad (14)$$

In these expressions $p = n \times K$ and $q = m \times K$. Note that for simplicity in notation we assume that control steps coincide with simulator time steps although this is not generally true.

3. RECOVERY OPTIMIZATION

3.1 General formulation

In recovery optimization, the typical problem to be solved is to optimize for the economic performance over the life cycle of the reservoir by dynamically manipulating the valve settings of injector and producer wells.

If a model of the reservoir dynamics is available, the recovery optimization problem can be expressed as:

$$\max_{\mathbf{u} \in Q} J_K(\mathbf{u}, x_0), \qquad (15)$$

$$\text{subject to} \quad \mathbf{g}(\mathbf{u}, \mathbf{x}) = 0, \quad x_0 = \bar{x}_0, \qquad (16)$$

$$\text{and} \quad \mathbf{c}(\mathbf{x}, \mathbf{u}) \leq 0, \quad \mathbf{d}(\mathbf{x}, \mathbf{u}) = 0 \qquad (17)$$

where $\mathbf{g}$ represents the system model as described in (7) and $\bar{x}_0$ is a vector containing the initial conditions of the reservoir. The inequality constraints $\mathbf{c}(\mathbf{x}, \mathbf{u})$ and equality constraints $\mathbf{d}(\mathbf{x}, \mathbf{u})$ represent linear or nonlinear constraints on the inputs (bound constraints), states and/or the outputs. Note that we dropped the dependence of $\mathbf{g}$ on θ assuming the model parameters to be known and fixed.

Typical objective functions are the cumulative oil produced or the net present value (NPV), the latter one representing cumulative cash flow. It is mathematically represented by:

$$J_K = \sum_{k=1}^{K}$$

$$\left[\frac{r_o \cdot q_{o,k}(y_k) - r_w \cdot q_{w,k}(y_k) - r_{inj} \cdot q_{inj,k}(u_k)}{(1+b)^{\frac{t_k}{\tau_t}}} \cdot \Delta t_k \right] (18)$$

where r_o is the oil revenue $[\frac{\$}{m^3}]$, r_w the water production costs $[\frac{\$}{m^3}]$ and r_{inj} the water injection costs $[\frac{\$}{m^3}]$, which are all assumed constant. K represents the total number of time steps k of a fixed time span and Δt_k the time interval of time step k in $[day]$. The term b represents the discount rate for a certain reference time τ_t. The terms $q_{o,k}$, $q_{w,k}$ and $q_{inj,k}$ represent the total flow rate of respectively produced oil, produced water and injected water at time step k in $[\frac{m^3}{day}]$.

A combination of reservoir simulation models and optimization algorithms has been applied to optimize the (theoretical) recovery from subsurface oil reservoirs since the 1980s; for recent review papers see Echeverría Ciaurri
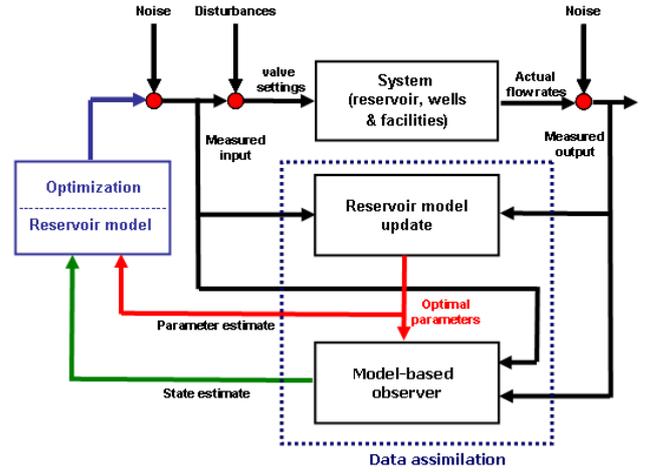
Fig. 3. *Closed-loop approach to improved reservoir management, CLRM (Closed-Loop Reservoir Management)*

et al. [2011] and Jansen [2011], and for an early textbook Ramirez [1987]. This form of model-based recovery optimization is also referred to as flooding optimization, sweep optimization, life-cycle optimization, or production optimization. We note that the latter name may be confusing because it is traditionally used for short-term optimization of well rates (without large simulation models, and on a time scale of days to months), rather than for recovery optimization over the entire life of the field (simulation-based, and on a time scale of years to decades). For these large scale dynamic optimization problems, several approaches are available, however the properties of reservoir models and the usually large number of control variables seriously limit the number of applicable dynamic optimization techniques. Sequential, gradient-based optimization using an adjoint to calculate the gradients, is currently the main candidate for solving the very large oil recovery optimization problem, Jansen [2011], and the tools for doing this are currently available in a number of corporate and commercially available reservoir simulation packages. Nevertheless also alternative methods are currently evaluated, Wang et al. [2007], Chen et al. [2009], Heirung et al. [2011].

3.2 Closed-Loop Reservoir Management (CLRM)

Similar to the framework for (Nonlinear) MPC (NMPC) Allgöwer and Zheng [2000], the model-based optimization as described in the previous section is generally implemented in a closed-loop structure using a receding (or in this case shrinking) horizon, and includes a model-based observer to estimate the current state. However in order to deal with the large uncertainty in the underlying dynamic model, it has become common practice to extend the nonlinear observer with a parameter estimation functionality as well, leading to a closed-loop set-up as indicated in Figure 3.

The combination of large scale numerical 'reservoir simulation' models and optimization algorithms for parameter estimation (using historic production data as input) has been developed in the oil industry since the 1970s. Under the names 'computer-assisted history matching',

'automatic history matching' or 'data assimilation' a large variety of techniques has been introduced; see e.g. the recent review papers of Aanonsen et al. [2009] and Oliver and Chen [2010] for overviews, and the textbooks of Datta-Gupta and King [2007], Oliver et al. [2008] and Evensen [2009] for detailed descriptions of different approaches. The most common approach is to follow a Bayesian estimation strategy where a nonlinear state observer, typically an Ensemble Kalman Filter (EnKF), Evensen [1994, 2009], is employed to estimate both the states and the parameters on the basis of past input en output data. In the EnKF the analytical propagation of the error covariance matrix is replaced by a Monte Carlo approach, in which the covariance matrix is computed from an ensemble of models. The EnKF utilizes the availability of the reservoir simulation model and does not require a linearization of this model. At the same time it remains computationally feasible for large scale models, up to millions of states, as has been shown in several applications, both in reservoir engineering problems and in different fields, see e.g. Naevdal et al. [2005], Gu and Oliver [2007], Aanonsen et al. [2009].

The popularity of EnKF in this domain is also explained by the relative ease with which not only states but also parameters can be included in the estimation problem. Following a Bayesian approach the unknown parameters are added to the states, leading to an extended state vector $\bar{x} = [x^T \ \theta^T]^T$. This extended state is then estimated with an EnKF on the basis of an ensemble of $\bar{x}$. This implies that also an ensemble of prior parameter estimates is added to the problem as a source of prior information.

In reservoir engineering problems the EnKF observer (data-assimilation) algorithm is applied to models with number of state variables up the order of 10^6. Since the permeability in each grid block is usually considered to be an unknown parameter, the dimension of θ amounts to the number of grid blocks, being of the same order as the number of states. As a result an extremely large number of variables has to be estimated from the available data. Although strong spatial correlations are present, which strongly reduce the effective number of unknown variables, the remaining number of unknowns is still very large. Because the amount of information in typical oil field production data is very limited, most of the parameter identification problems in reservoir engineering are therefore very ill-posed. We note that there remain many issues in the application of the EnKF, e.g., inconsistency between state and parameters for highly nonlinear problems, insufficient degrees of freedom to assimilate large data sets and underestimation of uncertainty. Thus, although the method has proved effective even in field cases, there are still issues remaining.

The repeated procedure of parameter/state estimation and subsequent model-based optimization is considered to significantly increase life-cycle value of the reservoir by providing the tools for a near-continuous model-based controlled activity. This approach is known as 'closed-loop reservoir management' (CLRM) or 'closed-loop production optimization'. We refer to Naevdal et al. [2006] and Sarma et al. [2006] for early 'proof of concept' applications with the aid of simple reservoir models, and to Sarma et al. [2008b], Chen et al. [2009], Jansen et al. [2009] and Peters et al. [2010] for subsequent larger-scale examples.

Overviews from a process engineering perspective were presented in Jansen et al. [2008] and Van den Hof et al. [2009], and the current paper can be considered a follow-up to these. The effect of increasing the frequency of combined data assimilation and optimization was addressed in Jansen et al. [2009], where it was shown that, for the example considered, an increased frequency leads to increased recovery. The effect of time delays in the model update was addressed in Foss and Jensen [2010]. It was concluded that delays may deteriorate the beneficial effects of CLRM, although the paper is restricted to the (restrictive) situation of single-phase (reservoir pressure) control. Specific optimization and data assimilation techniques in an CLRM setting were addressed in e.g. Chen et al. [2009], Peters et al. [2010] and Chen et al. [2010].

4. CURRENT LIMITATIONS AND CHALLENGES

The development over the last ten years has demonstrated a considerable progress in the field of reservoir optimization. CLRM in the strict sense, i.e. through combining data assimilation and flooding optimization during the operational phase of oil production, has not yet been reported to be used in practice. The individual elements, however, are being applied at increasing rates, and especially data assimilation is now regularly used during field (re)development planning. In the overall approach there are still serious limitations and challenges that are being faced. We consider the most important ones here. They will serve as a motivation for a selected set of developments that are being highlighted in the next sections.

Uncertainty

Reservoirs consist of deeply buried layers of fossilized fluvial or marine deposits, possibly folded, faulted, fractured or partially cemented. They often display very strong heterogeneities in fluid flow properties, with correlation lengths below the typical inter-well distance. Information is scarce and consists of low-resolution field-wide information on the reservoir structure, in combination with detailed but sparse information from wells. The state-of-the-art approach is therefore to use one or more ensembles of reservoir models with stochastically distributed parameter values as prior information. Data assimilation of these ensembles can give an estimate of the uncertainty in future reservoir flow. To this end various techniques, such as Markov chain Monte Carlo (McMC) in combination with proxy-models, or EnKF are increasingly applied, although it should be noted that all currently used implementations of EnKF give a relatively poor estimation of uncertainty when compared with a rigorous MCMC sampling; see for example Emerick and Reynolds [2011].

Given the limited information content of production data, the resulting models remain highly uncertain, which limits the reliability and therefore the value of long-term predictions and the subsequent optimization strategies. Explicit quantification of the uncertainty is a challenge, see also Caers [2011]. In the end the objective is to provide reliable tools for (model-based) operations and decision-making that take account of the model uncertainties present. From an optimization perspective a multi-scenario (robust) approach has been evaluated in van Essen et al. [2009].

Dynamic models and complexity.

Geological models are very complex. They typically have billions of cells, and are too complex to be used for dynamic simulations. As a result they are "upscaled" to relatively coarse dynamic reservoir flow models (containing up to millions of variables). Data assimilation and parameter estimation is then typically performed on the basis of these coarser models. However capturing and maintaining geologically realistic models in this coarser domain is not trivial. Parameter updates might make the models loose their geological realism, e.g. channel structures may be broken, or unrealistic localized features with low flow resistance may show up. Finding model representations that can focus on the essential geological properties of the reservoirs, and that allow a parametrization that can represent realistic variations in these properties is an important challenge.

At the same time it is very important to realize that not all detailed phenomena that occur in the dynamic models will have influence on the optimized trajectory of operation. Whereas geologists (and many reservoir engineers) focus on modeling of reservoir properties in the smallest possible details, from a control perspective it is sufficient to focus on those aspects that essentially contribute to the control strategy. From observability and controllability studies it has appeared that -in particular in situations of fixed well configurations - the reservoir models are typically poorly observable, poorly controllable and poorly identifiable, Zandvliet et al. [2008a], Van Doren [2010]. This implies that the essential dynamics in the process is represented by reduced-order models. On top of that the control-relevant dynamics might even be further simplified from the original models. However the relation between geological ("flow-relevant") features and the resulting reservoir response and optimal operation trajectory is not yet fully understood.

Measurement data

One could say that the current reservoir optimization problem is to a large extent a measurement problem. The use of only spatially sparse well data (periodic 'production data', i.e. pressures and flow rates in the wells) to determine the water-oil fluid front and a subsequent optimized operational strategy, clearly limits the accuracy of and induces considerable uncertainties in the models. An additional problem is the limited quality of traditional production data. The total amount of oil produced from a field is very accurately measured, for obvious fiscal and economic reasons, but the individual well production volumes, in particular the volumes of injected and produced water and gas, are usually only known with very limited accuracy, and errors in the order of tens of percents are no exceptions. An increasing use of data other than this traditional well data is warranted. In particular 'time-lapse' seismics (i.e. the repetition of seismics of over time to observe changes in reservoir pressures or fluid front positions) is an important source of field-wide information, but also other geophysical measurements such as gravity measurements, electromagnetic measurements, or subsidence measurements (satellite- or ground-based) are increasingly used. As an aside we note that the rapid development of 'smart wells', often equipped with multiple pressure and temperature sensors, and of the emerging geophysical measurement techniques, leads to new challenges related to transmission, storage, quality control, and interpretation of vastly increasing amounts of data.

Besides these new measurement opportunities, there is still room also for improving the information content of the data that is used for current estimation and data assimilation techniques, by moving from the use of normal production data, to data that results from deliberate well tests. The opportunities for adding excitation signals to well inputs, as is typically done in handling system identification problems, to make data more informative, are still not fully explored. We will further expand on this slightly in Section 7.

Nonlinearity

The reservoir model is essentially nonlinear, in particular due to the nonlinear saturation dependence of the relative permeabilities. Moreover, the nonlinear hyperbolic character of the saturation equations results in a delayed response of the saturation values in the reservoir to input changes in the wells. Phrased differently: the linearized dynamics between inputs and outputs will be essentially dependent on the location of the oil-water front. As a result of this, there is no steady state in the process, which classifies it basically as a batch process. However, and in contrast to many other batch processes, this process can only be run once, and there is no opportunity for batch-to-batch learning. Therefore, both the modeling, the model adaptation, and the optimization need to be done simultaneously. This further stresses the need for the optimal use of information that is present in the measurement data. Additionally, the dynamical properties between inputs and outputs that are observed during one phase of the reservoir depletion, will become different in another phase. This severely limits the opportunities to employ black box system identification techniques, for building models that have predictive capabilities over a large horizon. However there are good opportunities for, and a growing interest in, the estimation of relatively simple data-driven models (black box models) from well input-output measurements. They typically have a shorter predictive horizon than full-physics models and are therefore naturally suited for short-term optimization problems.

Process configuration

Model-based recovery optimization is actually not just a matter of optimizing valve settings. It increasingly involves the optimization of well locations, well trajectories or even well maintenance ('workover') strategies. This often results in mixed integer problems, and a wide variety of solution methods is being pursued to address these problems using increasingly complex models. For a simple strategy to optimize well locations on the basis of an NPV economic cost function, see Zandvliet et al. [2008b]. One of the challenges here will be to involve investment decisions (timing and location) concerning new wells to be drilled during the operational phase of the reservoir.

In the next sections we will highlight a few of the development that have been addressed recently, and that are

motivated from the challenges and limitations identified above.

5. PARAMETER ESTIMATION - WHAT CAN BE RETRIEVED FROM DATA?

When focusing now on the parameter estimation problem, we observe that generally grid block permeabilities are used to parametrize our model structure. The number of unknown parameters scales with the number of grid blocks that is used in the discretization of the underlying PDE. As a direct consequence the number of parameters in the estimation problem becomes excessively large, and a relevant question becomes whether these parameters can be reliably estimated from measurement data.

If we consider the physics-based model structure

$$\hat{\mathbf{y}}(\theta) = h(\mathbf{u}, \theta; x_0), \qquad (19)$$

the problem of assessing whether this model structure is identifiable, comes down to evaluating whether two different parameter values can lead to the same predictor. A full global analysis of this property is very cumbersome, but a local analysis is very well feasible. It is reflected in the notion of *local identifiability* assessing the injective properties of the mapping $\theta \rightarrow \hat{\mathbf{y}}(\theta)$ in a localized value $\theta = \theta_m$, see Grewal and Glover [1976].

In order to test local identifiability we consider a quadratic identification criterion in a prediction error setting (Ljung [1999]),

$$V(\theta) := \frac{1}{2}\boldsymbol{\epsilon}(\theta)^T P_v^{-1} \boldsymbol{\epsilon}(\theta), \qquad (20)$$

where the prediction error sequence $\boldsymbol{\epsilon}$ is defined as

$$\boldsymbol{\epsilon}(\theta) = \mathbf{y} - \hat{\mathbf{y}}(\theta), \qquad (21)$$

and where P_v is a weighting matrix that could represent (an estimate of) the covariance matrix of the noise sequence $\mathbf{v}$ that is supposed to act on the measured output. The Hessian of the cost function now provides a measure for the uniqueness of the parameter estimate, and thereby also as a sufficient condition for local identifiability in the parameter value $\hat{\theta}$ that is the minimizing argument of (20), under the given experimental conditions.

After approximation of $\hat{\mathbf{y}}$ with a first-order Taylor expansion around $\hat{\theta}$, the Hessian of $V(\theta)$ with respect to the parameters is

$$\frac{\partial^2 V(\theta)}{\partial \theta^2} = \frac{\partial \hat{\mathbf{y}}(\theta)^T}{\partial \theta} P_v^{-1} \left(\frac{\partial \hat{\mathbf{y}}(\theta)^T}{\partial \theta} \right)^T, \qquad (22)$$

where the derivatives are evaluated in $\theta = \hat{\theta}$, while

$$cov\ (\hat{\theta}) = J^{-1}$$

with J representing the Fisher information matrix

$$J = \mathbb{E}\left[\left. \frac{\partial^2 V(\theta)}{\partial \theta^2} \right|_{\hat{\theta}} \right], \qquad (23)$$

and $\mathbb{E}$ denoting expectation.

Local identifiability at $\hat{\theta}$ can now be investigated by evaluating the SVD of the Hessian that can be obtained from:

$$\frac{\partial \hat{\mathbf{y}}(\theta)^T}{\partial \theta} P_v^{-\frac{1}{2}} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}. \qquad (24)$$

If in this SVD a Σ_2 is found that satisfies $\Sigma_2 = 0$, then we have a lack of identifiability, and $cov(\hat{\theta})$ will become unbounded. In the situations of the reservoirs that we consider here, Σ_2 will typically contain a very large sequence of very small singular values, indicating that there is practically lack of identifiability, and a very large variance $cov(\hat{\theta})$. As a result the physical parameter estimates are actually highly unreliable. This has been shown in Van Doren et al. [2011].

In data assimilation methods, typically sequential estimation algorithms are used for joint state and parameter estimation, as the Extended Kalman Filter and the Ensemble Kalman Filter. They typically follow a Bayesian approach, which comes down to considering the (Bayesian) cost function

$$V_p(\theta) := V(\theta) + \frac{1}{2}(\theta - \theta_p)P_{\theta_p}^{-1}(\theta - \theta_p), \qquad (25)$$

where the last term represents the weighted mismatch between the parameter vector and the prior parameter vector θ_p with covariance P_{θ_p}.

When again the model output $\hat{\mathbf{y}}(\theta)$ is approximated using a first-order Taylor expansion around θ_p, the Hessian of (25) becomes:

$$\frac{\partial^2 V_p(\theta)}{\partial \theta^2} = \frac{\partial \hat{\mathbf{y}}(\theta)^T}{\partial \theta} P_v^{-1} \left(\frac{\partial \hat{\mathbf{y}}(\theta)^T}{\partial \theta} \right)^T + P_{\theta_p}^{-1}. \qquad (26)$$

Since $P_{\theta_p}^{-1}$ is positive definite by construction and the first term is positive semi-definite, the Hessian has full rank and the parameter estimate

$$\hat{\theta}_{Bayes} = \arg \min_{\theta} V_p(\theta)$$

is unique. This uniqueness is guaranteed by the prior information that has been added to the problem. Formally there can still be lack of identifiability, however it is not any more reflected in a non-unique parameter estimate. A consequence of this approach is that the obtained parameter estimate may be highly influenced by the prior information, and less by the measurement data. By analyzing the Hessian of $V(\theta)$, also in the sequential estimation case, it can however still be evaluated to which extent the data has induced (particular directions in) the parameter estimate, and to which extent the prior information.

The general observation is however, that when parametrizing the reservoir with one (permeability) parameter per grid block, the dimension of the parameter raises to levels that makes the parametrization unidentifiable, and therefore highly unreliable. If this problem is circumvented by the inclusion of prior information, it should be realized that the majority of the information in the posterior originates from the prior rather than from the data, Van Doren et al. [2011], Krymskaya et al. [2010].

In the literature several approaches have been suggested to deal with the problem of overparametrization:

- Several authors have considered to keep the grid block structure but to essentially reduce the number of grid blocks, Durlofsky et al. [1996], Zhang et al. [2008];
- Alternatively, spatial patterns are chosen to act as spatial basis functions, thereby considerably reducing the number of parameters. An example of this is the use of the discrete cosine transform, see Jafarpour and

McLaughlin [2008, 2009]. Another example is the use of (kernel) principal component analysis, see Sarma et al. [2008a].

- In Van Doren et al. [2008] it is shown that the columns of U_1 in (24) actually serve as spatial basis functions that span the identifiable space in the parameter domain. Since this is a local result, iteratively updating the basis functions is an option, as is illustrated in Van Doren et al. [2011]. In Tavakoli and Reynolds [2010] a related reparameterization is used which is claimed to be ideal in the sense that it keeps only the parameters that will lead to a reduction in uncertainty when data are actually history-matched.
- By simple capacitance resistor or channel parametrizations between injector and producer wells, simplified physical models can be represented by a limited number of parameters. This approach has been followed in Weber et al. [2009], Lin et al. [2010], Van Doren [2010]. Reducing the models even to black box linear models to be identified from production data has been addressed in Rowan and Clegg [1963], Chierici et al. [1981] and more recently in Lee et al. [2009], van Essen et al. [2010].

We note that reducing the number of parameters in reservoir parameter estimation is a topic of debate within the reservoir simulation community. In particular it is argued that reducing the number of parameters poses the risk of underestimating uncertainties and masking the presence of essential physical or geological features, see also Oliver and Chen [2010]. On the other hand, it is the limited information content in the measurement data that induces this poor identifiability (and therefore considerable uncertainty), which can only be resolved by providing additional (prior) information.

Computational issues

Due to the high computational cost of a reservoir model simulation, the number of model simulations required for the estimation of the physical parameters in the Bayesian approach must be limited. In case of a large number of parameters, the most efficient optimization method is obtained when the gradient of the objective function with respect to the model parameters is calculated by solving the adjoint problem; see e.g. Oliver et al. [2008]. The adjoint method, however, does have one very serious drawback: It requires the implementation of the adjoint of the tangent linear reservoir model. The implementation of the full adjoint equations for the parameter estimation procedure is an immense programming effort. Moreover, it also requires access to the simulation code. As a result there is a need for efficient methods that can deal with this complexity.

Clearly, the rapidly increasing popularity of the EnKF is driven to a large extent by the desire to avoid exact gradient-based optimization methods. It can be shown that the EnKF can be interpreted as a single step of a Gauss-Newton update (which is theoretically the optimal exact gradient-based optimization method), see Reynolds et al. [2006]. Indeed the EnKF can be interpreted as a sequential version of a gradient-based optimization algorithm where the gradient is approximated with the aid of an ensemble. Similarly, an emerging technique for

gradient-based but adjoint-free flooding optimization is ensemble optimization; see e.g. Chen et al. [2009]. An alternative to ensemble optimization, both for parameter estimation and flooding optimization, is to approximate the original reservoir model with a simplified linear model, for which the adjoint model is easier to implement. A generic way to obtain an approximate linear model is by the use of proper orthogonal decomposition (POD). In Vermeulen and Heemink [2006] the POD method is used to obtain an approximate low-dimensional version of the tangent linear model. As a result its adjoint can be implemented very easily and the minimization problem can be solved efficiently in reduced space. Vermeulen and Heemink [2006] applied this method to groundwater flow problems where it was shown to be computationally very efficient. In Kaleta et al. [2011] this methodology was applied to reservoir model estimation problems.

6. BALANCING LONG-TERM AND SHORT-TERM OPTIMIZATION

When performing a model-based optimization of the NPV over the life-cycle of the reservoir, as indicated in (15)-(18), an optimal input trajectory results $\hat{u}_k, k = 1, ..., K$ in terms of time-varying well pressures and/or flow rates. However from an industrial operational point of view several questions can be raised with respect to this optimized strategy:

(1) The optimized strategy is optimal provided that the underlying model is correct. Because of serious model uncertainties, both in the reservoir and in economic parameters as e.g. the oil price, this is very hard to justify. As a result the long-term predictions of the model should be considered with care. It would be hard to motivate reducing short-term production in order to gain more production at a stage far ahead in the future.
(2) In addition to this, economic objectives might be different from focusing on long-term NPV optimization only. Short term objectives may need to be taken into account (e.g. maximum instantaneous oil flow rates). Reservoir models are typically much too coarse to resolve small scale fluctuations (both in time and in space) near the wells. In particular 'coning' of water or gas near a production well is usually not modeled correctly.
(3) Operational requirements (e.g. well maintenance schedules, or surface facility constraints) may be overriding long-term strategies.

Model uncertainty can, to a certain extent, be addressed by robust optimization in which an expected value of the objective function is maximized based on one or more ensembles of reservoir models; see van Essen et al. [2009]. This requires however an accurate description and bounding of the model uncertainties. The point listed above actually suggest an extension of the optimization framework, where the optimization of the long-term NPV $J^{(1)}$ could be extended to also include short-term objective in the optimization.

In van Essen et al. [2011] a constrained optimization approach is presented, where a secondary objective function $J_K^{(2)}$ is introduced that focuses on a short-term objective.

In fact $J_K^{(2)}$ follows the same structure as J_K as presented in (18), but with a discount factor b that is substantially higher than in the long-term situation, thereby effectively stressing short-time behavior. A new optimization objective can now be formulated as follows:

- Determine $\mathbf{u}_\theta^* = \arg\max_{\mathbf{u} \in Q} J_K^{(1)}(\mathbf{u}, x_0)$ according to the setup and constraints of the original long-term nominal optimization problem (15)-(18);
- Determine

$$\max_{\mathbf{u} \in Q} J_K^{(2)}(\mathbf{u}, x_0), \qquad (27)$$

such that $J_K^{(1)}(\mathbf{u}, x_0) \geq J_K^{(1)}(\mathbf{u}_\theta^*, x_0) - \epsilon \qquad (28)$

and under the regular constraints (16)-(17), where $\epsilon > 0$ is a real-valued margin on the long-term revenues that is allowed to be compromized. This optimal input trajectory is denoted as $\tilde{\mathbf{u}}_\theta^*$.

In this way, a short-time objective is optimized under the constraint that there is a prespecified limited compromise on the long-term revenues. It has been noted that even with $\varepsilon = 0$, this problem has a feasible solution, due to the redundancy of degrees of freedom in the original optimization problem. In other words, after an optimal input $\mathbf{u}_\theta^*$ of $J_K^{(1)}$ has been established, there are still considerable degrees of freedom left to optimize a second objective. This offers the opportunity to perform hierarchical or lexicographic optimization in which the redundancy in the input variables is used to maximize short term performance (e.g. maximum instantaneous oil rate) without compromising the long-term goal (e.g. maximum NPV). The redundancy is primarily a result of the typical overparameterization of flooding optimization problems. There are usually many combinations of the control variables (i.e. the number of well control parameters times number of control time steps) that lead to (nearly) identical optimal values of the objective function.

In van Essen et al. [2011] an example of this approach is presented for a reservoir as indicated in Figure 1 with 4 production and 8 injection wells. Figure 4 shows the optimized input trajectories $\mathbf{u}_\theta^*$ and $\tilde{\mathbf{u}}_\theta^*$. The second trajectory (in red) shows substantially higher injection rates in the first phase of production. Figure 5 shows the time evolution of the NPV values of both strategies. Without any compromise on the value of $J^{(1)}$ the short term objective $J^{(2)}$ is drastically improved. The blue curve in Figure 5 shows the result if we would only optimize $J^{(2)}$ without taking account of any long-term effects. In this case a cost is incurred at the end of the life-cycle of the reservoir. See van Essen et al. [2011] for further details.

7. TIME-SCALE SEPARATION - TWO-LEVEL APPROACH

The discussion in the previous section has highlighted the importance and opportunities to include short-term production objectives into our reservoir optimization. This might also affect the models that we use as a basis for this optimization. The large scale reservoir models are typically developed for long-term prediction horizons, while for short-term predictions we could rely more and more on production measurement data. This points to a separation
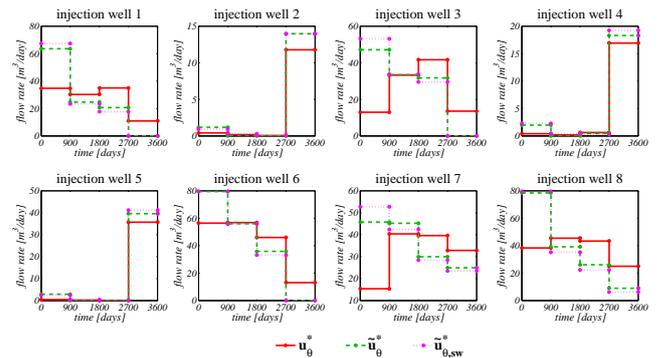
Fig. 4. *Optimized input trajectories for each of the eight injection wells over time, for the optimal solution of the long term optimization J_1 (green), and the optimal solution of the short term optimization under long term constraints J_2 (red), van Essen et al. [2011].*
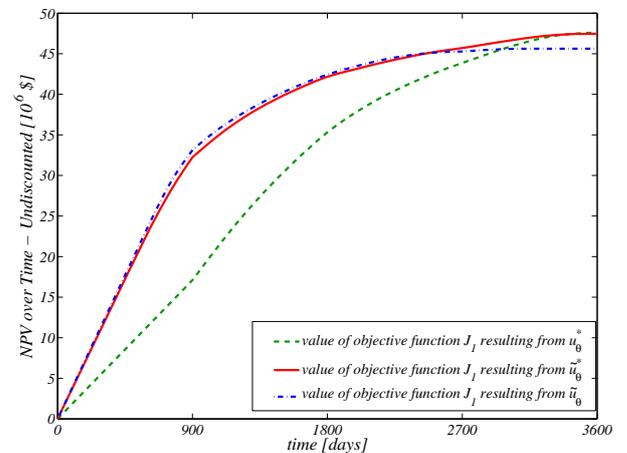
Fig. 5. *Time evolution of the NPV objective function for different optimized input strategies: long-term optimization (green), short-term optimization under long-term constraints (red), and short-term optimization (blue), van Essen et al. [2011].*

of time-scales, both in the optimization strategy and in the models used for this optimization.

The classical solution to this problem that is common in industrial process control in processing and refining industry is to introduce a multi-level control framework where long term effects (real-time optimization) are separated from short term effects (advanced process control). Saputelli et al. [2006] and Foss and Jensen [2010] have suggested to use such a multi-level approach also for upstream oil production. van Essen et al. [2010, 2012] have applied a two level strategy in which they combined long-term recovery optimization (trajectory generation), using a 'high-order' reservoir simulation model, with short-term production optimization (trajectory following) using a 'low-order' black box identified model. Fig. 6 depicts a schematic of this two-level CLRM process. The bottom half of the figure (below the dotted line) represents the traditional CLRM process which combines data assimilation with recovery optimization. The optimization is aimed at finding the optimal input $\hat{\mathbf{u}}_k$ that maximizes the NPV as expressed in equation (18). However, in this case we are primarily
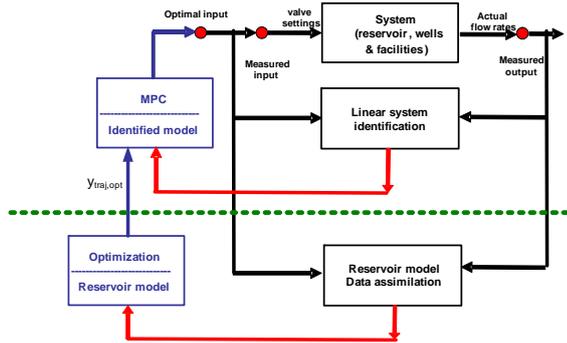
Fig. 6. Schematic representation of two-level CLRM.

interested in the optimal output $\hat{\mathbf{y}}_k$ that corresponds to $\hat{\mathbf{u}}_k$. The top half of Fig. 6 represents a state observer and an identified data-driven model that is used as the basis for a Model Predictive Control (MPC) controller aimed at tracking the optimal output $\hat{\mathbf{y}}_k$.

In van Essen et al. [2010, 2012] this strategy is applied to an the example reservoir model with 8 injectors and 4 producers. In order to assess the effect of model inaccuracies, a reference model is used to generate production data, and an (approximate) reservoir model is used as a basis for the optimization of input trajectories. In the first 75 days of the experiment excitation signals are added to the inputs, on the basis of which an 8th order black model has been identified. Next an MPC tracking controller is designed on the basis of this linear model. The results are depicted in Figure 7 showing that after the period of identification there is accurate tracking of the required reference trajectory. An open-loop implementation of the optimal trajectory on the reference model shows serious deviations, caused by the inaccuracies in the reservoir model.

8. CONTROL-RELEVANT UPSCALING

Field development is increasingly based on very detailed 'static' geological numerical models. These typically contain up to billions of cells with lithologic properties and display complex structural geologic features such as layering, faults and fractures. Dynamic simulation of subsurface flow with these large numbers of cells is computationally not feasible with current computing limitations, even with the use of massive parallel computing. Relatively coarse dynamic reservoir flow models (containing up to millions of variables) are therefore constructed by 'upscaling' very detailed 'static' geological models. However, from a system-theoretical point of view, a more fundamental argument for using upscaled models is that there is only a limited amount of information that can be observed from production data, while there is also a limited amount of control that can be exercised by adjusting the well parameters; in other words, the input-output behavior is usually of much lower dynamical order than the number of grid blocks in the model. Therefore, we developed an upscaling approach to find a coarse model that optimally describes the input-output behavior of a reservoir system. In this control-relevant method, the coarse-scale model parameters are calculated as the solution of an optimization problem that minimizes the distance between the input-output behaviors of the fine- and coarse-scale models. This distance is
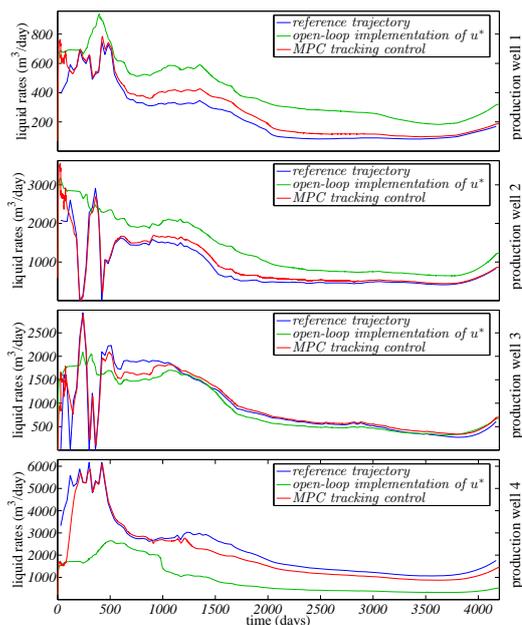
Fig. 7. *Output trajectory following for the four production wells. Reference output based on the nonlinear (approximate) reservoir model (blue), MPC tracking result on the basis of the local linear identified model (red), and open-loop implementation of optimized input trajectory on the data-generating system (green), van Essen et al. [2010].*

measured with the aid of the Hankel- or energy norms, in which we use Hankel singular values and Markov parameters as a measure of the combined controllability and observability, and response of the system, respectively. The method is particularly attractive to scale up simulation models in flooding optimization and/or history matching studies for a given configuration of wells. An advantage of our upscaling method is that it relies most heavily on those parameter values that directly influence the input-output behavior. It is a global method, in the sense that it relies on the system properties of the entire reservoir. It does not, however, require any forward simulation, neither of the full nor of the upscaled model. It also does not depend on a particular control strategy, but instead uses the dynamical system equations directly. Its dependency on well locations, however, implies that it should be (partially) repeated when those locations are changed. We tested the method on several examples and for nearly all cases obtained coarse scale models with a superior input-output behavior compared to common upscaling algorithms. For more details see Vakili-Ghahani and Jansen [2010, 2012], Zandvliet et al. [2008a].

9. CONCLUSIONS

In this paper an overview has been given of recent trends and developments as well as challenges in model-based production optimization from hydrocarbon reservoirs. Whereas a first step has been made towards using a model-based approach to optimize production, several important challenges are still ahead. The effective handling of model uncertainties is one of the key issues. The complexity of reservoir models is a serious handicap, and

at the same time it is felt that for generating optimized operational trajectories it might be possible to focus on the control/optimization relevant dynamics only, thereby limiting to relatively simple reduced complexity models. It has been illustrated that the information content in production data is too limited to validate models with too many degrees of freedom (parameters). The combination of physical reservoir models for the larger prediction horizon, and identified (black box) models for the shorter prediction horizon, seems to be a promising approach. Eventually also the objective functions have to be further extended, incorporating the possibility to include decisions on investments of to-be-installed wells, in terms of their location and timing, during the production stage of a reservoir.

10. ACKNOWLEDGMENTS

REFERENCES

S.I. Aanonsen, G. Naevdal, D.S Oliver, A.C. Reynolds, and B. Valles. The ensemble kalman filter in reservoir engineering – a review. *SPE Journal*, 14(3):393–412, 2009.

F. Allgöwer and A. Zheng. *Nonlinear Model Predictive Control*. Birkhäuser, 2000.

K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Applied Science Publishers, 1979.

J. Caers. *Modeling uncertainty in the earth sciences*. Wiley-Blackwell, 2011.

C. Chen, Y. Wang, G. Li, and A.C. Reynolds. Closed-loop reservoir management on the brugge test case. *Computational Geosciences*, 14(4):691–703, 2010.

Y. Chen, D.S. Oliver, and D. Zhang. Efficient ensemble-based closed-loop production optimization. *SPE Journal*, 14(4):634–645, 2009.

G.L. Chierici, G.A. Gottardi, and R.P. Guidorzi. Identified models for gas storage dynamics. *SPE JOurnal*, 21(2): 151–159, 1981.

A. Datta-Gupta and M.J. King. *Streamline simulation: Theory and Practice*. SPE, Richardson, 2007.

L.J. Durlofsky, R.A. Behrens, R.C. Jones, and A. Bernath. Scale up of heterogeneous three dimensional reservoir descriptions. *SPE Journal*, 1(3):313–326, September 1996.

D. Echeverría Ciaurri, T. Mukerji, and L.J. Durlofsky. Derivative-free optimization for oil field operations. In X.-S. Yang and S. Koziel, editors, *Computational Optimization and Applications in Engineering and Industry*. Springer, Berlin, 2011.

A. A. Emerick and A. C. Reynolds. Combining the ensemble kalman filter with markov chain monte carlo for improved history matching and uncertainty characterization. In *Proc. SPE Reservoir Simulation Symposium*, The Woodlands, USA, February 2011.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5): 10143–1062, 1994.

G. Evensen. *Data assimilation The ensemble Kalman filter, 2nd ed.* Springer, Berlin, 2009.

B. Foss. Process control in the upstream petroleum industries. In *Proc. of the 2011 4th International Symposium on Advanced Control of Industrial Processes*, Thousand Islands Lake, Hangzhou, P.R. China, 2011.

B. Foss and J.P. Jensen. Performance analysis for closed-loop reservoir management. *SPE Journal*, 16(1):183–190, 2010.

M. S. Grewal and K. Glover. Identifiability of linear and nonlinear dynamical systems. *IEEE Trans. Automatic Control*, 21(6):833–837, 1976.

Y. Gu and D. Oliver. An iterative ensemble kalman filter for multiphase fluid flow data assimilation. *SPE Journal*, 12(4):438–446, 2007.

T.A.N. Heirung, M.R. Wartmann, J.D. Jansen, B.E. Ydstie, and B.A. Foss. Optimization of the water-flooding process in a small 2d horizontal oil reservoir by direct transcription. In *Proc. 18th IFAC World Congress*, pages 10863–10868, Milan, Italy, 2011.

B. Jafarpour and D. B. McLaughlin. History matching with an ensemble kalman filter and discrete cosine parameterization. *Computational Geosciences*, 12(2): 227–244, 2008.

B. Jafarpour and D. B. McLaughlin. Reservoir characterization with the discrete cosine transform. *SPE Journal*, 14(1):182–201, 2009.

J.D. Jansen. Adjoint-based optimization of multiphase flow through porous media – a review. *Computers and Fluids*, 46(1):40–51, 2011.

J.D. Jansen, O.H. Bosgra, and P.M.J. Van den Hof. Model-based control of multiphase flow in subsurface oil reservoirs. *Journal of Process Control*, 18(9):846–855, 2008.

J.D. Jansen, S.G. Douma, D.R. Brouwer, P.M.J. Van den Hof, O.H. Bosgra, and Heemink. Closed-loop reservoir management. In *SPE Reservoir Simulation Symposium*, The Woodlands, USA, 2-4 February 2009.

M.P. Kaleta, R.G. Hanea, A.W. Heemink, and J. D. Jansen. Model-reduced gradient-based history matching. *Comput. Geosciences*, 15:135–153, 2011.

M. V. Krymskaya, R. G. Hanea, J. D. Jansen, and A. W. Heemink. Observation sensitivity in computer-assisted history matching. In *Proc. 72nd Europec/EAGE Annual Conference and Exhibition*, Barcelona, Spain, 14-17 June 2010.

K.H. Lee, A. Ortega, A.M. Nejad, N. Jafroodi, and I. Ershaghi. A novel method for mapping fractures and high permeability channels in waterfloods using injection and production rates. In *2009 SPE Western Regional Meeting*, page Paper SPE 121353. SPE, 2009.

Y. Lin, A. Ortega, A.M. Nejad, and I. Ershaghi. Waterflood tomography: mapping high contrast permeability structures using injection/production data. In *2010 SPE Western Regional Meeting*, 2010. Paper SPE 132699.

L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1999.

G. Naevdal, L. M. Johnsen, S. I. Aanonsen, and E. H. Vefring. Reservoir monitoring and continuous model updating using ensemble kalman filter. *SPE Journal*, 10(1):66–74, 2005.

G. Naevdal, D.R. Brouwer, and J.D. Jansen. Water flooding using closed-loop control. *Computational Geosciences*, 10(1):37–60, 2006.

D.S. Oliver and Y. Chen. Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1):185–221, 2010.

D.S. Oliver, A.C. Reynolds, and N. Liu. *Inverse theory for petroleum reservoir characterization and history matching.* Cambridge University Press, Cambridge, 2008.

L. Peters, R.J. Arts, G.K. Brouwer, C.R. Geel, S. Cullick, R.J. Lorentzen, Y. Chen, K.N.B. Dunlop, F.C. Vossepoel, R. Xu, P. Sarma, A.H. Alhuthali, and A.C. Reynolds. Results of the brugge benchmark study for flooding optimization and history matching. *SPE Reservoir Evaluation and Engineering*, 13(3):391–405, 2010.

W.F. Ramirez. *Application of optimal control theory to enhanced oil recovery.* Elsevier, 1987.

A. C. Reynolds, M. Zafari, and G. Li. Iterative forms of the ensemble kalman filter. In *Proc. 10th European Conference on the Mathematics of Oil Recovery (ECMOR X)*, Amsterdam, The Netherlands, 4 - 7 September 2006. paper A030.

G. Rowan and M.W. Clegg. The cybernetic approach to reservoir engineering. In *38th Annual Fall Meeting of the SPE of AIME*, New Orleans, USA, 1963. Paper SPE 727.

L. Saputelli, M. Nikolaou, and M.J. Economides. Real-time reservoir management: a multi-scale adaptive optimization and control approach. *Computational Geosciences*, 10(1):61–96, 2006.

P. Sarma, L.J. Durlofsky, K. Aziz, and W.H. Chen. Efficient real-time reservoir management using adjoint-based optimal control and model updating. *Computational Geosciences*, 10(1):3–36, 2006.

P. Sarma, L. J. Durlofsky, and K. Aziz. Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. *Mathematical Geosciences*, 40(1):3–32, 2008a.

P. Sarma, L.J. Durlofsky, and K. Aziz. Computational techniques for closed-loop reservoir modeling with application to a realistic reservoir. *Petroleum Science and Technology*, 26(10 & 11):1120–1140, 2008b.

R. Tavakoli and A. C. Reynolds. History matching with parameterization based on the svd of a dimensionless sensitivity matrix. *SPE Journal*, 15(12):495–508, 2010.

S.A. Vakili-Ghahani and J.D. Jansen. Control-relevant upscaling. *SPE Journal*, 15(2):471–479, 2010.

S.A. Vakili-Ghahani and J.D. Jansen. A system-theoretical approach to selective grid coarsening of reservoir models. *Computational Geosciences, in press*, 2012.

P.M.J. Van den Hof, J.D. Jansen, G.M. van Essen, and O.H. Bosgra. Model-based control and optimization of large scale physical systems – challenges in reservoir engineering. In *Proc. 21st Chinese Control & Decision Conference*, Guilin, China, 17-19 June 2009.

J. F. M. Van Doren. *Model Structure Analysis for Model-Based Operation of Petroleum Reservoirs.* Phd, Delft University of Technology, 2010.

J. F. M. Van Doren, P. M. J. Van den Hof, J. D. Jansen, and O. H. Bosgra. Determining identifiable parameterizations for large-scale physical models in reservoir engineering. In *Proc. 17th IFAC World Congress*, pages 11421–11426, Seoul, Korea, 2008.

J. F. M. Van Doren, P. M. J. Van den Hof, and J. D. Jansen. Parameter identification in large-scale models for oil and gas production. In S. Bittanti, A. Cenedese, and S. Zampieri, editors, *Proc. 18th IFAC World Congress*, pages 10857–10862, Milan, Italy, 2011.

G. M. van Essen, M. J. Zandvliet, P. M. J. Van den Hof, O. H. Bosgra, and J. D. Jansen. Robust waterflooding optimization of mutiple geological scenarios. *SPE Journal*, 14(1):202–210, 2009.

G. M. van Essen, A. Rezapour, P. M. J. Van den Hof, and J. D. Jansen. Integrated dynamic optimization and control in reservoir engineering using locally identified linear models. In *Proc. 49th IEEE Conf. Decision and Control*, pages 7643–7648, Atlanta, CA, USA, 2010.

G. M. van Essen, P. M. J. Van den Hof, and J. D. Jansen. A two-level strategy to realize life-cycle production optimization in an operational setting. In *SPE Intelligent Energy International Conference*, 2012. SPE-149736.

G.M. van Essen, P.M.J. Van den Hof, and J.D. Jansen. Hierarchical long-term and short-term production optimization. *SPE Journal*, 16(1):191–199, 2011.

P.T.M. Vermeulen and A.W. Heemink. Model-reduced variational data assimilation. *Mon. Weather Rev.*, 134 (10):2888–2899, 2006.

C. Wang, G. Li, and A.C. Reynolds. Production optimization in closed-loop reservoir management. In *SPE Annual Technical Conference and Exhibition*, Anaheim, California, U.S.A., November 2007. SPE 109805-MS.

D. Weber, T. F. Edgar, L. W. Lake, L. Lasdon, S. Kawas, and M. Sayarpour. Improvements in capacitance-resistive modeling and optimization of large scale reservoirs. In *SPE Western Regional Meeting*, San Jose, California, U.S.A., March 2009.

M. J. Zandvliet, J. F. M. Van Doren, O. H. Bosgra, J. D. Jansen, and P. M. J. Van den Hof. Controllability, observability and identifiability in single-phase porous media flow. *Computational Geosciences*, 12(4):605–622, 2008a.

M. J. Zandvliet, M. Handels, G. M. van Essen, D. R. Brouwer, and J. D. Jansen. Adjoint-based well-placement optimization under production constraints. *SPE Journal*, 13(4):392–399, 2008b.

P. Zhang, G. Pickup, and M. Christie. A new practical method for upscaling in highly heterogeneous reservoir models. *SPE Journal*, 13(1):68–76, March 2008.