



SPE 99728

Multiscale Regularization of Flooding Optimization for Smart-Field Management

M. Lien, U. of Bergen; D.R. Brouwer, SPE, Shell Intl. E&P; T. Mannseth, CIPR; and J.D. Jansen, SPE, Delft U. of Technology and SIEP

Copyright 2006, Society of Petroleum Engineers

This paper was prepared for presentation at the 2006 SPE Intelligent Energy Conference and Exhibition held in Amsterdam, The Netherlands, 11–13 April 2006.

This paper was selected for presentation by an SPE Program Committee following review of information contained in an abstract submitted by the author(s). Contents of the paper, as presented, have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material, as presented, does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Papers presented at SPE meetings are subject to publication review by Editorial Committees of the Society of Petroleum Engineers. Electronic reproduction, distribution, or storage of any part of this paper for commercial purposes without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of where and by whom the paper was presented. Write Librarian, SPE, P.O. Box 833836, Richardson, TX 75083-3836, U.S.A., fax 01-972-952-9435.

Abstract

Smart fields can provide enhanced oil recovery through the combined use of optimization and data assimilation. In this paper, we focus on the dynamic optimization of injection and production rates during waterflooding. In particular, we use optimal control theory in order to find an optimal well management strategy over the life of the reservoir that maximizes an objective function (e.g. recovery or net present value). Optimal control requires the determination of a potentially large number of (groups of) well rates for a potentially large number of time periods. However, the optimal number of well groups and time steps is not known a-priori. Moreover, taking these numbers too large slows down the optimization process and increases the chance of achieving a sub-optimal solution. We investigated the use of multi-scale regularization methods to achieve grouping of the control settings of the wells in both space and time. Starting out with a very coarse grouping, the resolution is subsequently refined during the optimization. The regularization is adaptive in that the multi-scale parameterization is chosen based on the gradients of the objective function. Results for the numerical examples studied indicate that the regularization may lead to significantly different optimum strategies, but may result in a similar cumulative oil production.

Introduction

We consider the secondary recovery phase of a heterogeneous oil reservoir, where water is injected into the reservoir for pressure maintenance and sweep improvement. We consider a scenario with multiple injectors and multiple producers of which the well rates can be controlled individually. Ideally, the injected water will displace the remaining oil in the reservoir on its way from the injection wells to the production wells. Rock heterogeneities will, however, influence the path of the injected water. The water will mainly flow in the high permeability channels, which

causes only part of the oil to be produced. Recently, smart fields concepts have been proposed as a means to improve control over the water front through detailed adjustments of the injection and production rates in time using a combination of model-based flooding optimization and model updating^{1,2}. For the optimization part these “closed-loop” reservoir management strategies rely on optimal control theory, which has been proposed before as a flooding optimization method by various authors³⁻⁷. However, optimal control is computationally expensive, and detailed management of every individual well of a smart field at every moment in time is economically and technically demanding. Moreover, there may not be enough information in the system to determine the optimal production strategy uniquely. Hence, we seek to develop management strategies with a restricted number of degrees of freedom, which at the same time maintain the advantages of this new technology.

In this paper, multiscale estimation techniques are utilized to attempt to find the optimal well management level. These are hierarchical regularization methods where the number of degrees of freedom in the estimation is gradually increased as the optimization proceeds. Multi-scale methods were first applied for solving partial differential equations to speed up convergence^{8,9}. Later, through the development of wavelets, multi-scale approaches have also widely been used within inverse problems¹⁰⁻¹³.

Theory

Optimal control

We consider the problem of maximizing the net present value (NPV) of the cumulative oil production over a fixed time horizon,

$$\max \bar{J} = \sum_{n=1}^{N_T} J^n(\mathbf{x}^n, \mathbf{u}^n), \dots\dots\dots (1)$$

subject to the dynamic constraint

$$\mathbf{g}^n(\mathbf{x}^n, \mathbf{x}^{n+1}, \mathbf{u}^n) = \mathbf{0}, \quad \mathbf{x}^0 = \mathbf{x}_0. \dots\dots\dots (2)$$

Here, the dynamic system (2) describes a heterogeneous, two-phase (oil-water) reservoir with smart wells⁶. Based on a conventional, finite difference approximation, it is given as a discrete-time dynamic model, where $\mathbf{x} \in \mathbb{R}^M$ denotes the state variables with elements corresponding to oil pressures and water saturations in all grid blocks at all times, and $\mathbf{u} \in \mathbb{R}^N$ is the control variable vector, with elements denoting the settings of the different wells at all times. The parameter n represents

the discrete time. The dynamic system gives the relation between the state variables \mathbf{x} and the control variables \mathbf{u} . Hence, Eq. (2) must be fulfilled for all admissible optimal control trajectories. This constrained optimal control problem is solved by forming the *Lagrangian*,

$$\bar{L} = \sum_{n=0}^{N_T-1} L^n = \sum_{n=0}^{N_T-1} J^n(\mathbf{x}^n, \mathbf{u}^n) + (\boldsymbol{\lambda}^n)^T \mathbf{g}^n(\mathbf{x}^n, \mathbf{x}^{n+1}, \mathbf{u}^n). \dots (3)$$

Here the dynamic system (2) is formally included as a constraint by adding it to the objective function (1) with the aid of Lagrange multipliers $\boldsymbol{\lambda}$. For more details of this optimal control problem, we refer to Brouwer and Jansen⁶. The necessary conditions for an optimal solution are derived by computing the first order variation of L ; confer, for example^{14,15}. The resulting first order optimality conditions are given by:

the *state equation*,

$$\mathbf{g}^n(\mathbf{x}^n, \mathbf{x}^{n+1}, \mathbf{u}^n) = \mathbf{0}, \quad \mathbf{x}^0 = \mathbf{x}_0, \dots (4)$$

the *adjoint equation*,

$$(\boldsymbol{\lambda}^n)^T = \left[-(\boldsymbol{\lambda}^{n-1})^T \frac{\partial \mathbf{g}^n}{\partial \mathbf{x}^n} - \frac{\partial J^n}{\partial \mathbf{x}^n} - \frac{\partial J^{n-1}}{\partial \mathbf{x}^n} \right] \left(\frac{\partial \mathbf{g}^{n-1}}{\partial \mathbf{x}^n} \right)^{-1}, \dots (5)$$

and the *optimality equation*,

$$\frac{\partial L^n}{\partial \mathbf{u}^n} = \frac{\partial J^n}{\partial \mathbf{u}^n} + (\boldsymbol{\lambda}^n)^T \frac{\partial \mathbf{g}^n}{\partial \mathbf{u}^n}. \dots (6)$$

Solving the constrained optimization problem requires repetition of the following steps until the optimal \mathbf{u} is reached:

- 1) Numerical simulation of the dynamic system (2), over the time interval $[0, N_T]$, with the initial choice of \mathbf{u} .
- 2) Evaluation of the objective function (1).
- 3) Backward numerical simulation of the adjoint Eq. (5), starting from the terminal condition $\boldsymbol{\lambda}^{N_T} = \mathbf{0}$.
- 4) Computation of the gradients $\partial L / \partial \mathbf{u}$,
where $L = [L^1 \dots L^{N_T}]^T$.
- 5) Computation of an improved control vector \mathbf{u} .

In the following, the steps (1-5) will be referred to as a cycle. Convergence to the optimum is assumed if the objective function value hardly increases for a number of subsequent cycles. Because the objective functions in flooding optimization are generally non-convex, the convergence to the global optimum is not guaranteed. However, a local optimum may often represent a significant improvement over the initial control strategy³⁻⁷.

Control variable estimation

A steepest ascent method is applied for optimizing the control variable. In the steepest ascent method, \mathbf{u} is updated according to the following rule:

$$(\mathbf{u}_{k+1})^T = (\mathbf{u}_k)^T + \varepsilon_k \left(\frac{\partial L}{\partial \mathbf{u}_k} \right)^T. \dots (7)$$

Here, ε_k is the weighting factor at cycle k . The value of ε_k is determined empirically, following a procedure as described in Ray¹⁶. An additional restriction on the change in the controls is imposed by not allowing L to change by more than a certain fraction γ per cycle. The maximum allowable ΔL_k , is increased for each subsequent cycle k according to the relation

$$\Delta L_k = k\gamma, \dots (8)$$

up to some preset maximum ΔL^{max} , with an associated weighting factor, ε^{max} . Because the optimization of \mathbf{u} is gradient based, we may compute a local optimum where the result depends on the initial \mathbf{u} . Further, there may not be enough information in the optimal control problem to determine the individual well settings uniquely, so that different values of \mathbf{u} may produce the same result. Hence, we seek to regularize this problem by reducing the solution space to get a better posed problem. Decomposition of an optimization problem by scale is known to facilitate the convergence of iterative ascent techniques to the global solution, particular in the presence of multiple local minima¹³.

The dimension of \mathbf{u} is determined by the total number of wells times the number of time steps in the numerical application, as each well in principle is allowed to be adjusted at every time step. Well settings that are altered rapidly in both space and time may give oscillations in the production and injection rates, hence, a more restricted representation of \mathbf{u} may be beneficial also in order to get a more stable production profile. Multiscale techniques enable a restricted representation of \mathbf{u} , where initially one starts out with a very coarse representation, but as the optimization proceeds and one hopefully approaches the right solution, the resolution is gradually increased.

Multiscale parameterization

A restricted representation of \mathbf{u} is achieved by grouping the fine-scale elements of \mathbf{u} both in space and time. Hence, instead of adjusting each well individually, multiple wells are grouped and treated as one. Further, the well settings are held constant over time as adjustments are allowed only at a restricted number of time steps. Mathematically \mathbf{u} can be represented by the sum,

$$\mathbf{u}(k, \mathbf{y}) = \sum_{i=1}^{N(k)} v_i \boldsymbol{\varphi}_i(k, \mathbf{y}). \dots (9)$$

Here, \mathbf{y} specifies the coordinates (in space and time) of the fine-scale elements of \mathbf{u} and $N(k)$ gives the number of coarse-scale parameters v_i at cycle k . The vectors $\{\boldsymbol{\varphi}_i(k, \mathbf{y})\}_{i=1}^{N(k)}$ form a set of orthogonal basis functions with elements $\varphi_i^j(k, \mathbf{y})$ that can be represented as,

$$\varphi_i^j(k, \mathbf{y}) = \begin{cases} 1 & \text{if } \mathbf{y} \in C_i^k; \\ 0 & \text{otherwise.} \end{cases} \dots (10)$$

Here $j = 1 : N(i)$ denotes the number of fine-scale elements, and C_i^k signifies the wells and the different time steps covered by the i^{th} coarse-scale parameter at cycle k . The coarse-scale parameters v_i can be interpreted as the elements of a reduced-

order control vector \mathbf{v} . Eq. (9) may therefore also be expressed as $\mathbf{u} = \mathbf{\Phi}\mathbf{v}$ where the columns of matrix $\mathbf{\Phi}$ are the basis functions $\{\varphi_i(k, \mathbf{y})\}_{i=1}^{N(k)}$.

The parameterization of \mathbf{u} is hierarchical, where the number of parameters is increased by a fixed factor at each refinement. The dimension of \mathbf{v} will be given by the rule, $N(k) = N(k)_t + N(k)_s = \text{ceil}(2^k) + \text{ceil}(2^{k/a})$. In order to balance the number of degrees of freedom in space and time, we set the parameter $a = N_t / N_s$. Among the wells, the refinements are done successively along the x-direction and y-direction. Initially, \mathbf{u} is set to be constant over all production wells and all injection wells at all time steps. Hence, \mathbf{u} will be represented by two parameter values, one for the production wells and one for the injection wells, i.e. \mathbf{v} is a control vector with just two elements. After these values are optimized the parameterization of \mathbf{u} is refined for the next cycle. The new parameters are then again optimized, before another refinement step is conducted, giving a successively finer resolution of \mathbf{u} . The parameter refinement followed by optimization is repeated for each cycle until convergence is reached. In this application, \mathbf{u} is refined at each cycle. Hence, the control variable is not optimized to reach an optimum between each refinement. With this approach we may reduce the risk of driving the optimization into a local minimum before further refinements are conducted. Furthermore, a reduction in the CPU time is achieved, since driving the optimization into an optimum at each scale would be computationally demanding (See e.g. Sun and Yeh¹⁷ and Tsai and Yeh¹⁸). In order to update the coarse-scale parameters by the steepest ascent method, the associated coarse-scale gradients are needed. These may easily be derived from the fine-scale gradients, as is shown in the next section, see also Yoon *et al.*¹³.

Gradient calculation at a coarse scale

Let $s_j = \partial L / \partial u_j$, $j = 1, \dots, N$ represent the sensitivity of the Lagrangian with respect to a change in an element u_j of \mathbf{u} , denoted by δu_j . The change in the Lagrangian, δL , due to fine-scale perturbations, δu_j , can then be expressed in terms of the fine-scale sensitivities as follows:

$$\delta L = \sum_i^{N(k)} \sum_j^{N(i)} s_j \delta u_j \dots \dots \dots (11)$$

The inner summation term can be expressed as

$$\sum_j^{N(i)} s_j \delta u_j = s_i \delta v_i, \dots \dots \dots (12)$$

where s_i is the sensitivity of L with respect to a change in the coarse-scale parameter, δv_i . The change in the individual fine-scale elements δu_j , will equal δv_i . From Eq. (11), the coarse-scale sensitivity can therefore be expressed as

$$s_i = \sum_j^{N(i)} s_j \dots \dots \dots (13)$$

Hence, control variable gradients at various scales can be obtained simply by integrating (adding) the fine-scale gradients.

A multiscale approach results in a rapid increase in the number of parameters. The zonation is also very rigid, with a homogenous representation of \mathbf{u} where the same number of time steps and wells are covered by each parameter. A more efficient parameterization can be achieved by local refinement strategies. Here the number of parameters is increased only when it is feasible according to some predefined rule. Different methods have been developed to enable local refinement strategies, see e.g. Chavent and Bissel¹⁹ and Sun and Yeh²⁰. Adaptive multi-scale estimation as developed by Grimstad *et al.*²¹ base the parameterization of \mathbf{u} on a linearization of the model equations. With this method, the chosen parameterization will give the optimal refinement of first order. This method, however, requires the objective function to be quadratic with respect to the model equations. Ben Ameer *et al.*²² developed another method where the zonation of \mathbf{u} is determined by *refinement indicators* based on information from the gradients of the objective function. Their approach was developed for the inverse problem of transmissivity estimation for two-dimensional groundwater flow. Lately, Cominelli *et al.*²³ applied a simplified version of this method on a field-scale history-matching problem for permeability estimation. This method may be applied also to our problem and provides a tool to reach an efficient parameterization for the control variables.

Refinement indicators

In this section, the rules governing the local parameterization of \mathbf{u} are derived. Following Ben Ameer *et al.*²², the refinements at each cycle may be implemented as equality constraints in the optimal control problem. Consider the case where \mathbf{u} is represented by one parameter, i.e. such that $\mathbf{v} = v_i$. We consider dividing this parameter into potentially two new parameters, i.e. such that $\mathbf{v} = [v_i^1 \ v_i^2]^T$. This may be achieved, either by a refinement among the wells or refinement in time. This tentative partitioning may be written as $\mathbf{A}\mathbf{v} = \mathbf{b}$, where $\mathbf{A} = [1 \ -1]$. Here $\mathbf{b} = \mathbf{b}_s$ or $\mathbf{b} = \mathbf{b}_t$, determine the grouping in space (i.e. among the wells) or in time, respectively. The constraint $\mathbf{A}\mathbf{v} = \mathbf{b}$ can be extended to represent an arbitrary number of tentative parameters, $N(k)$.

In this work, \mathbf{A} has the general form

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 1 & -1 & \cdot & \cdot & \cdot \\ \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & -1 \end{bmatrix}, \dots \dots \dots (14)$$

where each row relates to one tentative cut, i.e. potential discontinuity in space or time. We start the refinement process with $v_i^1 = v_i^2$ (i.e. $\mathbf{b} = \mathbf{0}$), after which we investigate the potential gain of allowing the new parameters to take on different values. Imposing $\mathbf{A}\mathbf{v} = \mathbf{b}$ as an equality constraint to

the optimal control problem, we obtain the extended Lagrangian,

$$\tilde{L} = L + \boldsymbol{\mu}^T (\mathbf{b} - \mathbf{A}\mathbf{v}) \dots\dots\dots (15)$$

We note that the adjoint Eq. (5) is unaffected by this new constraint. However, the optimality equation (6) now reads,

$$\frac{\partial \tilde{L}}{\partial \mathbf{v}} = \frac{\partial L}{\partial \mathbf{v}} - \boldsymbol{\mu}^T \mathbf{A} = \mathbf{0}. \dots\dots\dots (16)$$

In equality-constrained optimization, the multipliers have a useful interpretation. Define $V(\mathbf{b})$ to be the value of \tilde{L} at an optimum,

$$V(\mathbf{b}) = L + \boldsymbol{\mu}^T (\mathbf{b} - \mathbf{A}\mathbf{v}). \dots\dots\dots (17)$$

Differentiating with respect to \mathbf{b} , using the chain rule, we obtain

$$V'(\mathbf{b}) = \frac{\partial L}{\partial \mathbf{x}} \frac{d\mathbf{x}}{d\mathbf{b}} + \left(\frac{\partial L}{\partial \mathbf{u}} - \boldsymbol{\mu}^T \mathbf{A} \right) \frac{d\mathbf{v}}{d\mathbf{b}} + \boldsymbol{\mu}^T. \dots\dots\dots (18)$$

At an optimum the first and second term vanish, and Eq. (18) reduces to $V'(\mathbf{b}) = \boldsymbol{\mu}^T$. Thus, the multiplier vector $\boldsymbol{\mu}$ may be interpreted as the sensitivity of the optimal Lagrangian to the perturbation \mathbf{b} . That is the marginal value of allowing new degrees of freedom in the well management. From Eq. (7), we note that $\boldsymbol{\mu}$ may be expressed by means of the gradients of L ,

$$\frac{\partial L}{\partial \mathbf{v}} = \boldsymbol{\mu}^T \mathbf{A}. \dots\dots\dots (19)$$

For the example with one tentative cut, \mathbf{b}_s or \mathbf{b}_t , the following rule applies at an optimum:

$$\frac{\partial L}{\partial v_i} = \frac{\partial L}{\partial v_i^1} + \frac{\partial L}{\partial v_i^2} = 0. \dots\dots\dots (20)$$

Hence, the partitioning of v_i into v_i^1 and v_i^2 , will have the following refinement indicator

$$\mu_i = \frac{\partial L}{\partial v_i^1} = -\frac{\partial L}{\partial v_i^2}. \dots\dots\dots (21)$$

The individual gradients of the new parameters may be interpreted as the marginal value of allowing to adjust them independently. From the solution of the adjoint Eq. (5) and the optimality condition (6), all the fine-scale gradients are available at every cycle. Hence, this method allows us to compare a high number of different cuts at virtually no cost. This gives a more flexible algorithm, with respect to reaching the optimal parameterization with a minimum number of parameters.

In this application, we compare six equally spaced cuts in space and six equally spaced cuts in time for each parameter. We note that the ordering of the tentative cuts will have an impact on the final parameterization. This has not been a topic in this application. The two cuts associated with the largest refinement indicators are selected. If the selected cuts are parallel (i.e. both cuts are conducted either in time or space, respectively) the parameter is divided into three new parameters, if they are orthogonal the parameter is divided into four new parameters. The refinement indicators are required to exceed a predefined threshold value in order for the

refinement to be conducted. Hence, each parameter may be divided into all from one to four new parameters. To maintain a restricted number of parameters, only one coarse-scale parameter is refined at each cycle. The refinement process consists of two main steps;

- 1) For each coarse-scale parameter \mathbf{v}_i select the local best refinement candidate, $\hat{\mathbf{v}}_I = \{\mathbf{v}_i^l\}_{l=1}^{N(i)}$, where $N(i) \in [0, 4]$ based on refinement indicators, $\boldsymbol{\mu}_i$.
- 2) Among the local best candidates, choose the parameterization that gives the greatest change in the Lagrangian, $\Delta \tilde{L}(\hat{\mathbf{v}}_J) > \Delta \tilde{L}(\hat{\mathbf{v}}_I) \forall I \neq J$.

Refinement away from the optimum

In this application, we consider a relatively computationally demanding forward problem, with a high number of parameters. Hence, it is important to restrict the number of forward calculations. Computing the optimum of the objective function value for all scales, would be very computationally demanding. Hence, refinement of \mathbf{v} is conducted before the optimum is reached. Outside an optimum, however, the gradient of \tilde{L} will have a constant value, different from zero. Let us consider the example of one tentative cut in time or space, where initially, \mathbf{u} is described by one parameter v_i . Starting the refinement process, the gradients will have a constant value $\partial \tilde{L} / \partial v_i = c$. In terms of the tentative new parameters v_i^1 and v_i^2 , the gradient reads

$$\frac{\partial \tilde{L}}{\partial v_i} = \frac{\partial \tilde{L}}{\partial v_i^1} + \frac{\partial \tilde{L}}{\partial v_i^2} = c. \dots\dots\dots (22)$$

Hence, we may write $\partial \tilde{L} / \partial v_i^l = c_l$, $l = 1, 2$, where the values of $\{c_l\}_{l=1}^2$ reflect how far the different parameters are from the optimum. The refinement indicator outside the optimum now reads,

$$\mu_i = \frac{\partial L}{\partial v_i^1} - c_1 = -\frac{\partial L}{\partial v_i^2} + c_2. \dots\dots\dots (23)$$

We note that for a given parameter, the refinement indicator is shifted by a constant value. In order to assess whether the gradients are valid as selection criteria away from an optimum we look at the physical interpretation of $\partial \tilde{L} / \partial \mathbf{v}$. Given the general form of \mathbf{A} (see Eq. (14)), we have the relation $\mathbf{A}\mathbf{A}^T = r\mathbf{I}$. By forming the right pseudo-inverse, we can write, $\mathbf{v} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b} = \frac{1}{r} \mathbf{A}^T \mathbf{b}$. Hence, we get the expression $d\mathbf{v} / d\mathbf{b} = \frac{1}{r} \mathbf{A}^T$. Here, $r = 2$, but will in general reflect the number of new parameters in the refinement of each \mathbf{v}_i . By inserting for $r = 2$ and $\mathbf{A}^T = [1 \ -1]^T$ into Eq. (18) and applying $\partial \tilde{L} / \partial v_i^l = c_l^T$, $l = 1, 2$, we obtain

$$V'(\mathbf{b}) = \frac{1}{2} (\mathbf{c}_1 - \mathbf{c}_2)^T + \boldsymbol{\mu}^T. \dots\dots\dots (24)$$

Here, we have assumed $\partial L / \partial \mathbf{b} \approx \mathbf{0}^T$. By inserting (23) into Eq. (24) we obtain,

$$V'(\mathbf{b}) = \frac{1}{2} \left(\frac{\partial L}{\partial \mathbf{v}_i^1} - \frac{\partial L}{\partial \mathbf{v}_i^2} \right) \dots \dots \dots (25)$$

Hence, the constant values $\{c_l\}_{l=1}^2$ vanish, and the sensitivity of the Lagrangian with respect to the perturbation \mathbf{b} will still be a function of the resulting gradients. Refinements associated with large gradients of opposite sign will represent the greatest potential for increased objective function value. We note that the gradients will always have opposite signs for refinements at an optimum, whereas this will not always be the case here.

Objective function

The objective function to be maximized is the NPV of the cumulative oil production over all times,

$$\bar{J} = \sum_{n=0}^{N_t-1} \left\{ -V_{gb} \frac{q_i^n \Delta t^n [(1-f_w^n)r_o + f_w^n r_w]}{(1+d)^t} \right\} \dots \dots \dots (26)$$

Here, V_{gb} is the grid block volume [m^3], r_o is the oil revenue [$\frac{\$}{\text{m}^3}$], r_w is the water cost [$\frac{\$}{\text{m}^3}$], d is the discount rate [$\frac{1}{\text{year}}$], t is the cumulative time [year], f_w^n is the water fraction [-] and q_i^n is the total injection rate per unit volume [$\frac{1}{\text{s}}$].

Constraints on the controls

In water flood optimization, there are additional constraints on the control variables apart from the ones that are imposed by the dynamic system. In this paper we considered one type of well operating constraints in which, the total water injection rate, q_{inj} , and the total liquid production rate, q_{prod} , are controlled directly, i.e. a rate-constrained scenario. In an earlier publication we also considered a pressure-constrained scenario in which the well flowing pressures in the injector and producer wells are held fixed²⁴.

For rate-constrained optimization, the injection and production rates are controlled directly in each well

$$\mathbf{u}^n = \left[(q_{inj})_1^n \dots (q_{inj})_{N_{inj}}^n (q_{prod})_1^n \dots (q_{prod})_{N_{prod}}^n \right]^T \dots \dots \dots (27)$$

The total injection and production rate is set to be constant throughout the considered production period. That is, the control variable must fulfill the following constraint,

$$\sum_{i=1}^{N_s} (q_s)_i = q^{max} \dots \dots \dots (28)$$

for $s \in \{inj, prod\}$. Hence, the optimal control strategy comes down to an optimal allocation of injection and production over the individual wells. The different wells are, therefore, defined to be favorable or unfavorable, depending on whether their production (injection) rates are to be increased or decreased. In this application, we apply the gradients to determine which group the individual wells belong to, where the wells associated with gradients above the mean are defined to be favorable and those with gradients below the mean to be unfavorable. The modified gradients in the rate-controlled case are, thus, given as

$$\frac{\partial L^{n,mod}}{\partial (q_s)_i^n} = \frac{\partial L^n}{\partial (q_s)_i^n} - \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{\partial L^n}{\partial (q_s)_i^n} \dots \dots \dots (29)$$

Examples

Example 1

Fig. 1 depicts a reservoir with 9 injectors and 8 producers. We considered both conventional vertical wells (single completions), each completed in all 7 layers, as well as dual completions. In the latter case the upper completion is in the upper 3 layers, and the lower completion is in the lower 4 layers. The upper completions are numbered 1, 2, 3, ..., while the lower completions are numbered 1b, 2b, 3b, The permeability field contains a number of high permeability channels, highlighted in **Fig. 2**. The porosity field was taken homogeneous at 0.2. The optimization was done for a fixed production period of 10 years, with balanced injection and production at a constant field rate of 9000 bbl/d. Further information on the field is given in **Table 1**.

As initial guess for the control function, an equal and constant rate per well was chosen. For the single completion injectors and producers this amounts to 1000 bbl/day and 1125 bbl/day respectively. For the double completion injectors and producers this amounts to 500 bbl/day and 612.5 bbl/day respectively. The objective was to maximize cumulative oil production. Because field injection and production rates are constant this is equivalent to minimizing cumulative water production. Optimization was done with an ordinary steepest ascent method (OSA), the ordinary multi-scale method (OM), and an adaptive multi-scale method by means of refinement indicators (RI).

The optimum liquid injection and production rates found in the single completion case are shown in **Fig. 3**, **Fig. 4**, and **Fig. 5**. A comparison of the cumulative production figures is shown in **Fig. 6**. All methods result in an improvement with respect to the reference case with the largest improvement being obtained with the OSA method. The second highest cumulative oil production is obtained with the RI method. **Fig. 4** shows that in the optimum case injector 2 is shut in for the entire period, indicating that even with less wells compared to the reference case an improvement in cumulative oil production is possible. Costs for wells were, however, not considered in this study. The lowest improvement was obtained with the OM method, with poor convergence toward the optimum. If undesirable, the extreme well rates achieved for this method should be avoided by setting upper limits to the rate per well.

The optimum liquid injection and production rates found with the dual completion wells are shown in **Fig. 7**, **Fig. 8**, and **Fig. 9**. Although the optimum injection and production rates vary widely with the method used, they all result in a quite similar cumulative oil production (**Fig. 10**). **Fig. 11** shows that all optimized dual completion cases gave higher cumulative oil production than the optimized single completion cases, pointing at the added value of having enhanced control functionality in the wells.

Fig. 12 shows the final saturation distribution in the reservoir. Sweep is particularly poor near the attic of the reservoir. **Fig. 13** & **Fig. 14** show that these areas are locally better swept in the optimum case, but also that a significant part of the oil near the attic of the reservoir is still not swept.

Example 2

Fig. 15 depicts a reservoir consisting of 4 layers with 36 dually completed injectors and 25 dually completed producers. For each well one completion is in the upper two layers and the second completion in the lower two layers. The upper completions are odd numbered, and the corresponding lower completions are even numbered. Main reservoir properties are shown in **Table 2**. The optimization was again done for a fixed production period of 10 years, with balanced injection and production at a constant field rate of 47000 bbl/d. As an initial control strategy all wells/completions were operated at the same rate of respectively 653 bbl/d for each injector completion, and 940 bbl/d for each producer completion.

Fig. 16 shows the water saturation distribution at 3650 days. Most of the reservoir is water saturated leaving little scope for optimization. The optimization for this example was done only with the OSA and the RI methods. The optimum liquid injection and production rates found are shown in **Fig. 17** and **Fig. 18**. Again the optimum injection production strategies vary considerably for both methods, with slightly higher ultimate recovery for the OSA method (**Fig. 19**). An important advantage of the RI derived solution is that wells rates only need to be managed at a restricted number of times during the production period. In the examples, the RI approach results in a well management strategy that provides almost the same amount of oil as the fine-scale solution.

From comparing **Fig. 3** & **Fig. 4**, **Fig. 7** & **Fig. 8**, and **Fig. 17** & **Fig. 18**, it is seen that the computed optima for the RI cases give the solutions with the lowest well rates. This is advantageous since high flow rates in the producers may lead to for instance sand production or excess gas production in a reservoir close to the bubble point. For the injectors high rates may not be desirable because it may for instance lead to reservoir fracturing. High flow rates requirements will also make a well more expensive. Based on these considerations, out of two control policies giving the same end result, the one with the lowest requirements in terms of maximum rate and ease of implementation is to be preferred.

Conclusion

Numerical examples illustrated considerable gain from applying flooding optimization. By allowing for a detailed well management (grouping of well controls in time and space) the scope of increased oil recovery is considerable. We compared different approaches to derive the optimal injection and production strategy over the life of the field. In the ordinary steepest ascent approach, each individual well is controlled near-continuously in time. In multi-scale optimization, the degree of freedom in the well management is found as a part of the solution strategy, resulting in grouping of the wells, and less frequent control. By applying refinement indicators to determine the grouping of the wells, adaptive multi-scale estimation was enabled. In the method of refinement indicators, we obtained solutions with nearly the same total oil recovery as in the ordinary steepest ascent approach. However, this was achieved with a more restricted well management strategy. We also found that this multi-scale solution gave well rates more easily attainable than those found with the ordinary steepest ascent approach.

Nomenclature

\mathbf{A}	= matrix
\mathbf{b}	= vector of grouping parameters
c	= parameter, Mt, \$
C	= set of control variables, -
d	= discount rate, 1/t, 1/a
f	= water fraction, -
\mathbf{g}	= vector function
\mathbf{I}	= unit matrix
J	= objective function (cost function), M, \$
k	= refinement cycle counter, -
L	= Lagrangian, M, \$
n	= time step, -
N_T	= final time step, -
q	= flow rate per unit volume, 1/t, 1/s
r	= number of refinements, -
r_o	= oil price per unit volume, M/L ³ , \$/m ³
r_w	= water costs per unit volume, M/L ³ , \$/m ³
\mathbf{s}	= sensitivity vector
t	= time, t, a
\mathbf{u}	= control vector, -
\mathbf{v}	= reduced order control vector, -
V	= optimal value of L , M, \$
\mathbf{x}	= state vector
\mathbf{y}	= coordinate vector
γ	= restriction on control change, -
ε	= weight factor, -
ϕ	= basis function
Φ	= transformation matrix
λ	= vector of Lagrange multipliers, -
μ	= vector of refinement indicators, -

Subscripts

gb	= grid block
w	= water
o	= oil
$prod$	= production
t	= total
w	= water

References

1. Brouwer, D.R., Naevdal, G., Jansen, J.D., Vefring, E. and van Kruijsdijk, C.P.J.W.: "Improved reservoir management through optimal control and continuous model updating," paper SPE 90149 presented at the 2004 SPE Annual Technical Conference and Exhibition, Houston, Texas, USA, 26-29 September.
2. Sarma, P., Durllofsky, L.J. and Aziz, K.: "Efficient closed-loop production optimization under uncertainty," paper 94241 presented at the SPE Europe/AEGE Annual Conference, Madrid, Spain (2005) 13-16 June.
3. Asheim H.: "Maximization of water sweep efficiency by controlling injection and production rates," paper SPE 18365 presented at the 1988 SPE European Petroleum Conference, London UK.
4. Virnovski G. A.: "Waterflooding strategy design using optimal control theory," Proc. 6th European Symposium on IOR (1991) Stavanger, Norway.

5. Sudaryanto B.: "Optimization of displacement efficiency of oil recovery in porous media using optimal control theory," PhD Thesis (1998) University of Southern California, USA.
6. Brouwer, D.R. and Jansen, J.D.: "Dynamic optimization of water flooding with smart wells using optimal control theory," *SPEJ* (December 2004) 391-402.
7. Sarma, P., Aziz, K. and Durlofsky, L.J.: "Implementation of adjoint solution for optimal control of smart wells," paper SPE 92864 presented at the SPE Reservoir Simulation Symposium, Houston, USA (2005) 31 January – 2 February.
8. Brandt A.: "Multi-level adaptive solutions to the boundary value problem," *Mathematics of Computation* (1977) **13** 333-390.
9. Briggs W.: *A Multigrid Tutorial*, Society for Industrial and Applied Mathematics (1987) Philadelphia, PA.
10. Emsellem Y. and de Marsily G.: "An automatic solution for the inverse problem," *Water Resources Research* (1971) **7** 1264—1283.
11. Chavent G. and Liu J.: "Multiscale parameterization for the estimation of a diffusion coefficient in elliptic and parabolic problems," Proc. 5th IFAC symposium on Control of Distributed Parameter Systems, Perpignan, France (1989).
12. Liu J.: "A multi resolution method for distributed parameter estimation," *SIAM J. Scientific Computing* (1993) **14** 389-405.
13. Yoon S., Malallah A. H., Datta-Gupta A., Vasco D. W. and Behrens R. A.: "A multiscale approach to production data integration using streamline methods," *SPEJ* (June 2001) 182—192.,
14. Kamien M. L. and Schwartz N. L.: *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, Elsevier North Holland Inc. (1981).
15. Stengel R. F.: *Optimal Control and estimation*, Dover Publications Inc. New York (1994).
16. Ray W. H.: *Advanced Process Control*, McGraw-Hill (1981).
17. Sun N.-Z. and Yeh W. W.-G.: "Identification of parameter structure in groundwater inverse problems," *Water Resources Research* (1985) **21** 869-883.
18. Tsai F. T.-C. and Yeh W. W.-G.: "Characterization and identification of aquifer heterogeneity with generalized parameterization and Bayesian estimation," *Water Resources Research* (2004) **40** 1-12.
19. Chavent G. and Bissell R.: "Indicators for the refinement of parameterization," *Inverse Problems in Engineering Mechanics*, Elsevier, Amsterdam (1998).
20. Sun N.-Z. and Yeh W. W.-G.: "A proposed stepwise regression method for model structure identification," *Water Resources Research* (1998) **34** 2561—2572.
21. Grimstad A.-A., Mannseth T., Nævdal G. and Urkedal H.: "Adaptive multiscale permeability estimation," *Computational Geosciences*. (2003) **7** 1-25.
22. Ben Ameer H., Chavent G. and Jaffré J.: "Refinement and coarsening indicators for adaptive parameterization: application to the estimation of hydraulic transmissivities," *Inverse Problems* (2002) **18** 775-794.
23. Cominelli A., Ferdinandi F. De Montleau P. and Rossi R.: "Using gradients to refine parameterization in field case history-match projects," paper SPE 93599 presented at the 2005 SPE Reservoir Simulation Symposium, Houston, TX, USA.
24. Lien, M., Jansen, J.D., Mannseth, T. and Brouwer, D.R.: "Multiscale regularization of dynamic waterflood

optimization," Submitted for publication to *Computational Geosciences*.

Metric Conversion Factors

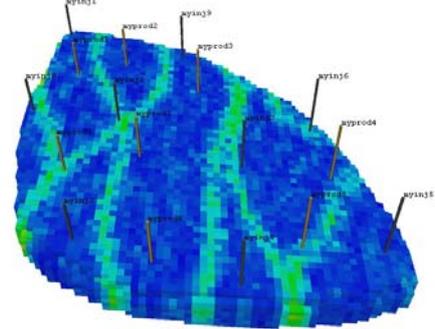
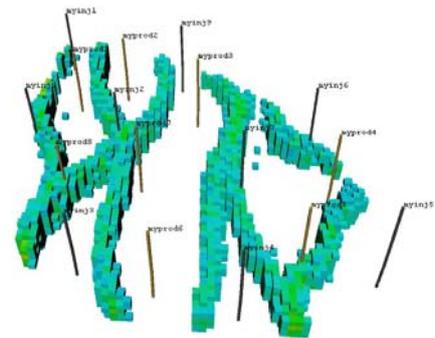
$$\begin{array}{ll} \text{bbl} \times 1.589\,873 & \text{E-01} = \text{m}^3 \\ \text{mD} \times 9.869\,233 & \text{E-04} = \mu\text{m}^2 \end{array}$$

TABLE 1 – MODEL PROPERTIES EXAMPLE 1

<u>Variable</u>	<u>Value</u>	<u>Units</u>
Nr. of grid blocks	25200 (60 x 60 x 7) -	
Grid block size (xyz)	15 x 15 x 7	m
Nr. of active grid blocks	18553	m
Initial pressure	40e6	Pa
Relative permeability model	Corey	-
Residual oil saturations	0.10	-
Endpoint relative permeability oil	1	-
Corey exponent oil	2	-
Connate water saturation	0.10	-
Endpoint relative permeability water	1	-
Corey exponent water	2	-
Oil viscosity	2e-3	Pa s
Oil compressibility	1e-10	1/Pa
Oil density at reservoir conditions	800	kg/m ³
Water viscosity	1e-3	Pa s
Water compressibility	1e-10	1/Pa
Water density at reservoir condit.	1000	kg/m ³
Capillary pressure	0	Pa

TABLE 2 – MODEL PROPERTIES EXAMPLE 2

<u>Variable</u>	<u>Value</u>	<u>Units</u>
Nr. of grid blocks	40804 (101 x 101 x 4) -	
Grid block size (xyz)	20 x 20 x 10	m
Nr. of active grid blocks	40804	m
Initial pressure	40e6	Pa
Relative permeability model	Corey	-
Residual oil saturations	0.10	-
Endpoint relative permeability oil	1	-
Corey exponent oil	2	-
Connate water saturation	0.10	-
Endpoint relative permeability water	1	-
Corey exponent water	2	-
Oil viscosity	0.75e-3	Pa s
Oil compressibility	4.35e-10	1/Pa
Oil density at reservoir conditions	846.5	kg/m ³
Water viscosity	0.31e-3	Pa s
Water compressibility	4.35e-10	1/Pa
Water density at reservoir condit.	1008.6	kg/m ³
Capillary pressure	0	Pa

**Fig. 1: Example 1. Permeability field and well locations.****Fig. 2: Example 1. Permeability range 1700-7000 mD.**

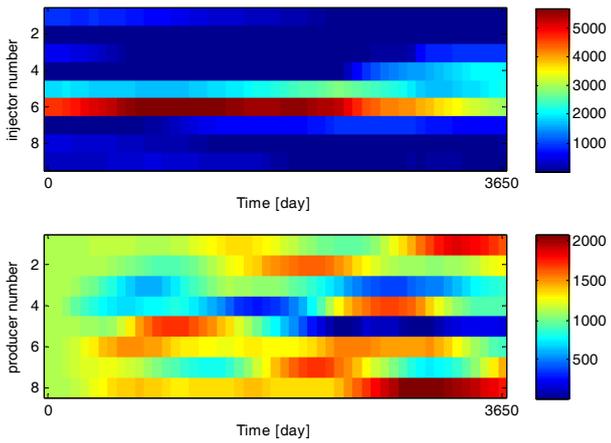


Fig. 3: Optimum well rates obtained with OSA approach and single completions wells. Optimum obtained after 58 iterations.

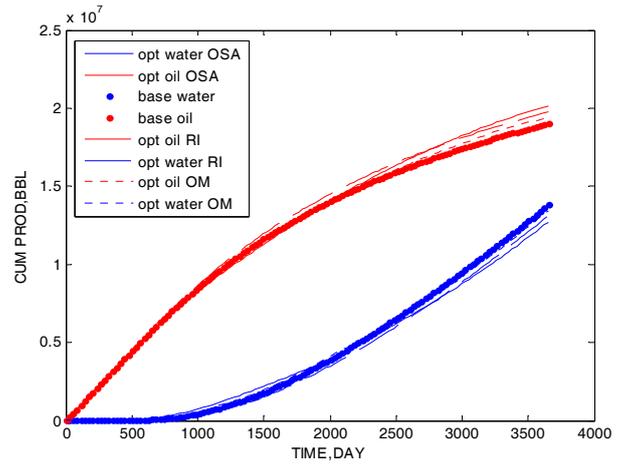


Fig. 6: Comparison of cumulative production with single completion wells for reference case, and optimized cases using OSA, OM, and RI optimization approaches.

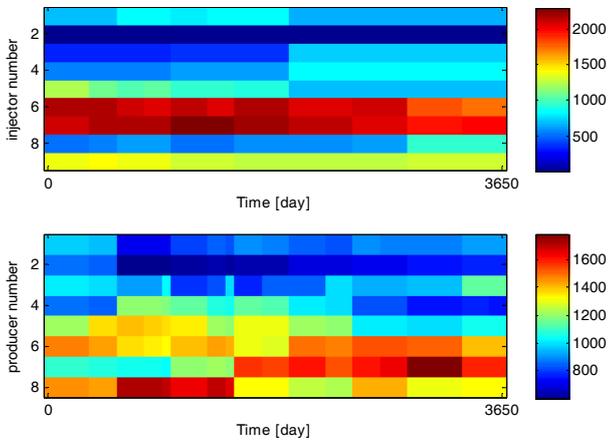


Fig. 4: Example 1. Optimum well rates obtained with RI method and single completions wells. Optimum obtained after 51 iterations.

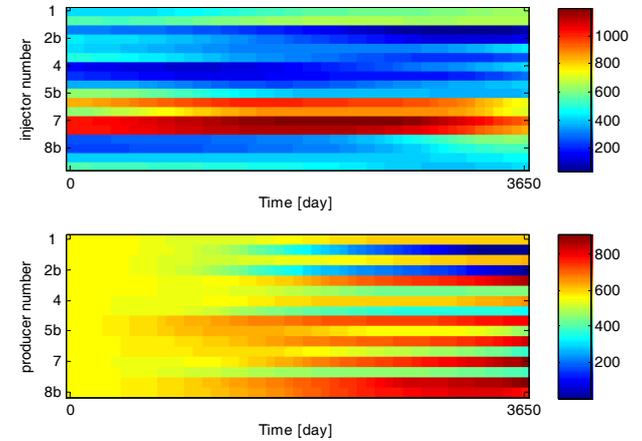


Fig. 7: Optimum well rates obtained with the OSA approach using dual completions. Optimum obtained after 43 iterations.

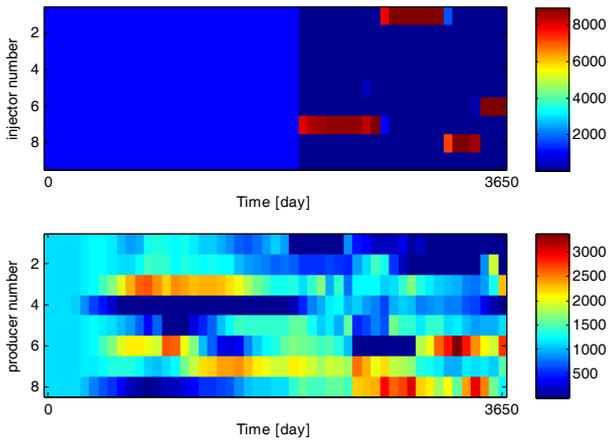


Fig. 5: Example 1. Optimum well rates obtained with OM approach and single completions wells. Optimum obtained after 32 iterations.

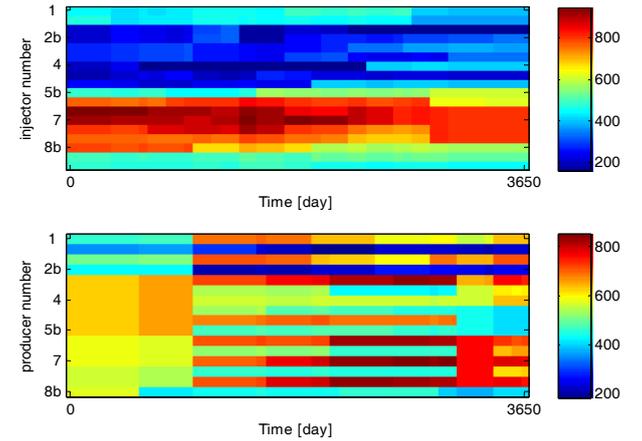


Fig. 8: Optimum well rates obtained with the RI approach using dual completions. Optimum obtained after 52 iterations.

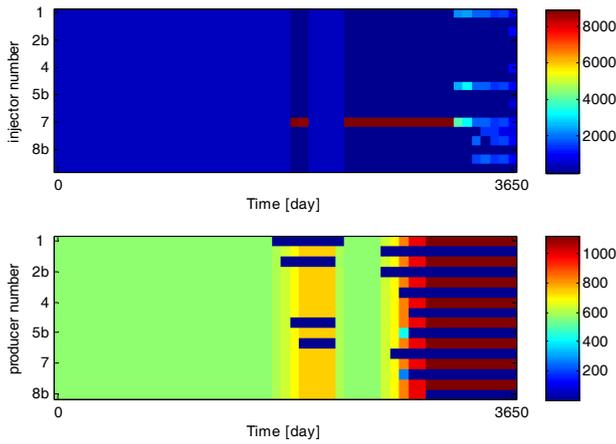


Fig. 9: Optimum well rates obtained with the OM approach using dual completions. Optimum obtained after 27 iterations.

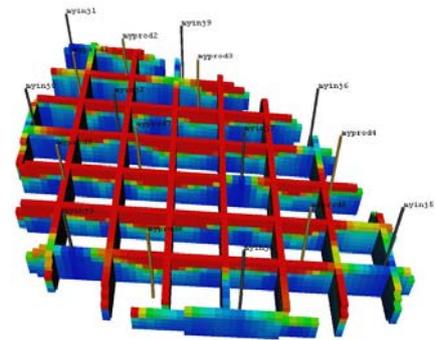


Fig. 12: Example 1. Final saturation for base case with single completions. Scale is from water saturation 0.1 (red) to 0.9 (blue).

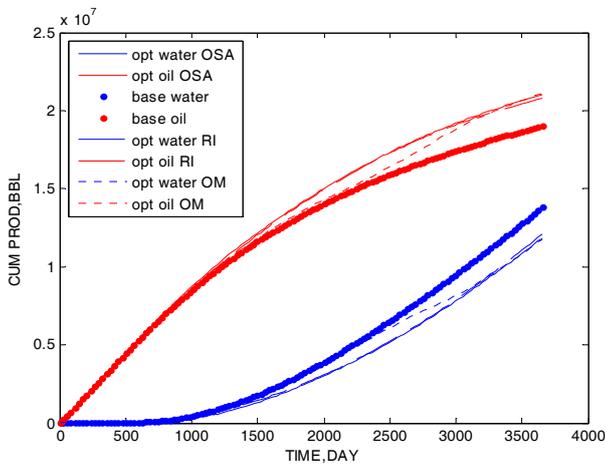


Fig. 10: Comparison of cumulative production with dual completion wells for reference case, and optimized cases using OSA, OM, and RI approaches.

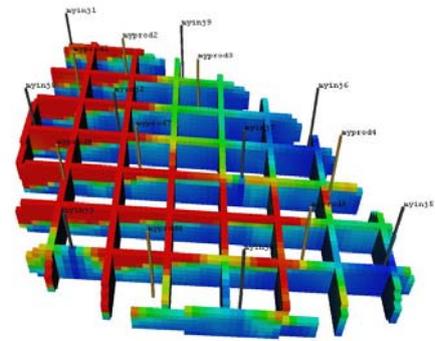


Fig. 13: Example 1. Final saturation in the single completion optimum case obtained with the OSA approach. Optimum obtained after 58 iterations.

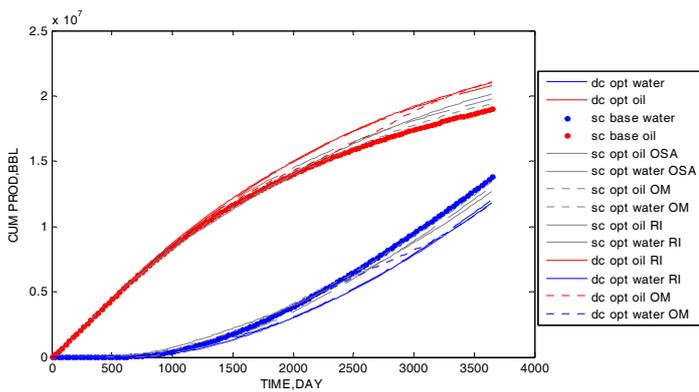


Fig. 11: Example 1. Comparison of cumulative production for all cases.

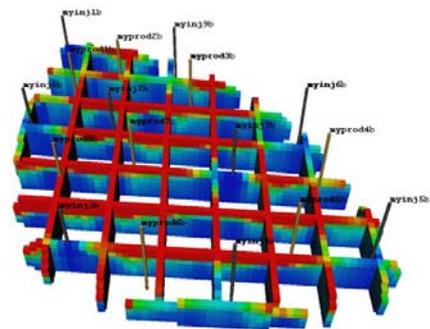


Fig. 14: Final saturation in the dual completions optimum case obtained with the RI approach. Optimum at 52 iterations.

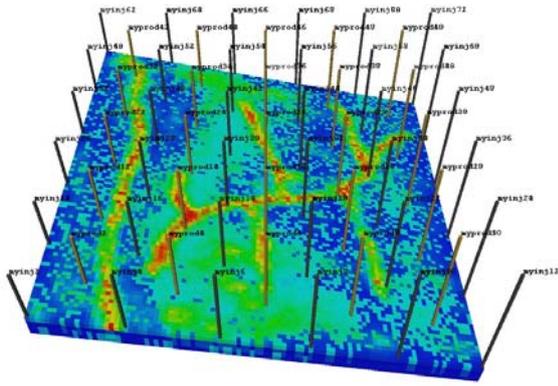


Fig. 15: Example 2. Permeability field and well locations.

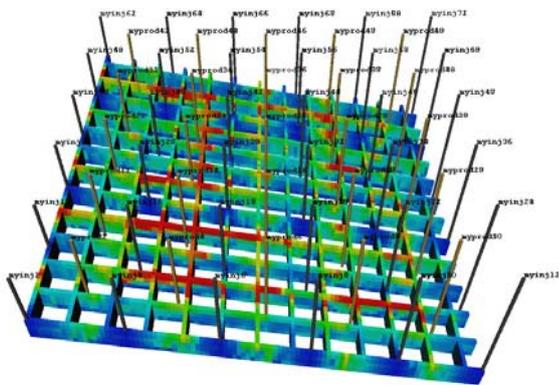


Fig. 16: Example 2. Water saturation distribution at 3650 days for base case. Scale is from water saturation 0.1 (red) to 0.9 (blue).

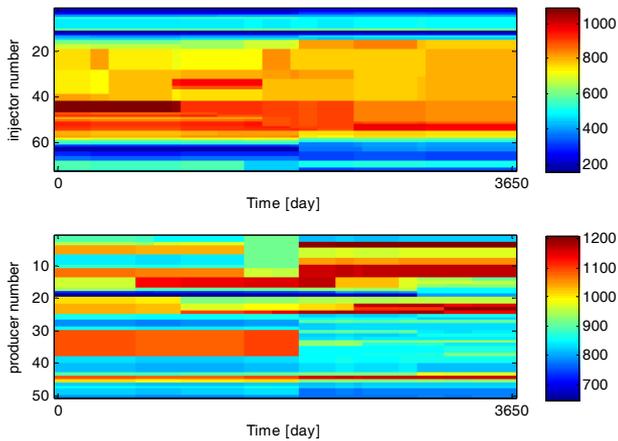


Fig. 17: Example 2. Optimum liquid injection and production rates, obtained with the RI approach.

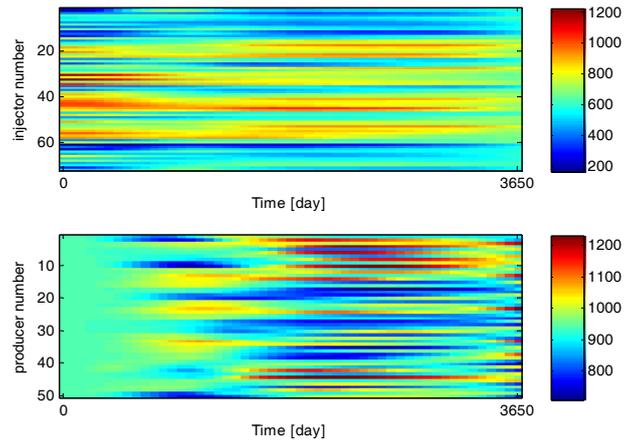


Fig. 18: Example 2. Optimum liquid injection and production rates, obtained with the OSA approach.

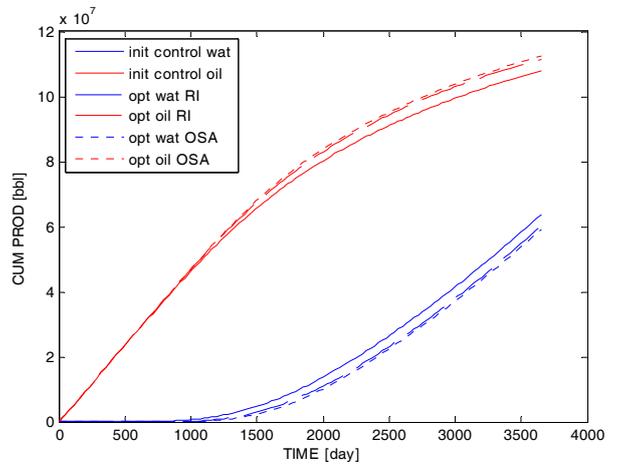


Fig. 19: Example 2. Cumulative production figures for base case, the OSA approach, and the RI approach.