Model-Based Control and Optimization of Large Scale Physical Systems
Challenges in reservoir engineering

Paul M.J. Van den Hof

co-authors:
Jorn Van Doren, Jan Dirk Jansen, Sippe Douma, Gijs van Essen, Okko Bosgra, Maarten Zandvliet

ERNSI Workshop, Cambridge, UK, 29 September 2010
Contents

• Introduction – problems in reservoir engineering
• The models
• Model-based optimization – OL & CL
• Parameter estimation, identifiability, model structure approximation
• 2D Example
• Prospects and conclusions
Introduction: reservoir engineering

Water flooding

- Involves the injection of water through the use of injection wells
- Goal is to displace oil by water
- Production is terminated when (too much) water is being produced
Oil production from a reservoir can last for periods of 15-25 years.
Introduction: reservoir engineering

The system involves the reservoir, wells and sometimes surface facilities.

- **Inputs:** control valve settings of the wells (injectors and producers)
  - Smart wells: multiple (subsurface) valves
- **Outputs:** (fractional) flow rates and/or bottomhole pressures
  - Smart wells: multiple (subsurface) measurement devices
Introduction: reservoir engineering

Objective

Optimize Net Present Value (NPV):

• Optimize economic revenues related to oil recovery, as a function of dynamic valve settings

\[
J = \sum_{k=1}^{N} \frac{\Delta(t_k)[r_o q_{o,k} - r_w q_{w,k} - r_i q_{i,k}]}{(1 + b)^{\frac{t_k}{\tau}}}
\]

Under constraints: \( c(x_k, u_k) \leq 0 \)

typically limits on water injection capacity, and max/min pressures in injection/production wells

• Rationalise the reservoir operational decisions
The Model

isothermal two-phase (oil-water) flow

Mass balance:
\[ \nabla (\rho_i u_i) + \frac{\partial}{\partial t} (\phi \rho_i S_i) = 0 \quad i = \{o, w\} \]

Momentum (Darcy's law):
\[ u_i = -k \frac{k_{ri}}{\mu_i} \nabla p_i \quad i = \{o, w\} \]

Variables: \( p_o, p_w, S_o, S_w \)

Saturations satisfy: \( S_o + S_w = 1 \)

Simplifying assumptions, a.o.: \( p_o = p_w \)
Discretization in space and time

State space model:

\[
V(x_t)\dot{x}_t = T(x_t)x_t + q_t; \quad x_0
\]
\[
y_t = h(x_t)
\]
\[
y^T = \begin{bmatrix} p_w^T \quad q_w^T \quad q_w^T \end{bmatrix}
\]
\[
x^T = \begin{bmatrix} p_o^T \quad S_w^T \end{bmatrix}
\]

After discretization in space (and time):

\[
g(x_{k+1}, x_k, u_k, \theta) = 0 \quad \text{dim}(x) \approx 10^4 - 10^6
\]
\[
y_k = h(x_k)
\]

and \( \theta \) typically the permeabilities in each grid block
Characteristic of process

- Nonlinear batch process (one-go), of which the dynamics is essentially dependent on the location of the -moving- oil/water front
Model-based Optimization

Optimization problem:

$$\max_q J(q) = \max_q \sum_{k=1}^{N} L(x_k, q_k)$$

such that:  
$$g(x_{k+1}, x_k, q_i, k) = 0, \quad x_0 = x(0)$$

$$q_{min} \leq q_k \leq q_{max}$$

$$q_o, k + q_w, k = q_i, k$$

Non-convex optimization, solved by gradient-based method:  
Adjoint-variables calculation through backward integration of the related (Hamiltonian based adjoint) equation.  
(feasible for systems of this size)
12-well example

- 3D reservoir
- 8 injection / 4 production wells
- Period of 10 years
- High-permeability channels
- 18,553 grid blocks
- Minimum rate of 0.1 stb/d
- Maximum rate of 400 stb/d
- No discount factor
- $r_o = 20 \$/stb, r_w = 3 \$/stb and r_i = 1 \$/stb
- Optimization of economic benefit

(Gijs van Essen et al., CAA 2006)
Why this wouldn’t work

• Model has some simplifying assumptions
• Optimization over life-time reservoir (changing economic circumstances)
• ...........

• Open-loop strategy
• We do not know the reservoir model!
Closed-loop Reservoir Management

• Moving from (batch-wise) open-loop optimization to on-line closed-loop control

• However we need a model as a basis for e.g. a receding/shrinking horizon strategy

Obtaining a model

• First-principle models (geology) are very much uncertain

• Opportunities for identification are limited
  (nonlinear behaviour dependent on front-location, single batch process, experimental limitations)

• Option: estimate physical parameters (permeabilities) in first principles model; starting with initial guess
Closed-loop Reservoir Management

Receding/shrinking horizon control strategy:

- Use a state-estimator to reconstruct the current state
- Run the optimization algorithm to evaluate future scenario’s
- Implement the optimized valve settings until the next state update
- This is actually a NMPC in a shrinking horizon implementation
- However no trajectory following but trajectory finding, i.e. real-time dynamic optimization (RTO)
Closed-loop Reservoir Management

- **Optimization**
  - Reservoir model

- **System**
  - (reservoir, wells & facilities)

- **Reservoir model update**

- **Model-based observer**
  - State estimate
  - Parameter estimate

- **Actual flow rates**
  - Measured output

- **Noise**
  - Disturbances
  - Valve settings
  - Optimal input

- **Noise**
Closed-loop Reservoir Management

Several options for nonlinear state and parameter estimation:

Available from oceanographic domain:

Ensemble Kalman filter (EnKF) (Evensen, 2006)

- Kalman type estimator, with analytical error propagation replaced by Monte Carlo approach (error cov. matrix determined by processing ensemble of model realizations)
- Ability to handle model uncertainty (in some sense)
- In reservoir engineering used for estimation of states and parameters (history matching)
Ensemble Kalman Filter

- As prior information an ensemble of initial states \( \{\tilde{x}_k|k\} \) is generated from a given distribution.

- By simulating every ensemble member, corresponding ensembles \( \{\tilde{x}_{k+1}|k\} \) and \( \{\tilde{y}_{k+1}|k\} \) are generated, and stored as columns of matrices \( \tilde{X} \) and \( \tilde{Y} \) respectively.

- The measurement update of a EKF is applied to every element of the ensemble, where the covariance matrices are replaced by sampled estimates on the basis of \( \tilde{X} \) and \( \tilde{Y} \).

- The update becomes: \( \tilde{x}_{k+1|k+1} = \tilde{x}_{k+1|k} + K_{k+1} [y_{k+1} - \tilde{y}_{k+1|k}] \), where \( K_{k+1} \) is given by:

\[
K_{k+1} = \tilde{X} \tilde{Y}^T \cdot [\tilde{Y} \tilde{Y}^T + R]^{-1} \quad \text{(BLUE)}
\]

- The result is a new ensemble \( \{\tilde{x}_{k+1|k+1}\} \).
Closed-loop simulation example

- Model with high-perm channels assumed to be ‘reality’
- Permeabilities are unknown in closed-loop control
- Period of 8 years
- Objective function: NPV
  - $r_o = 10 \$/stb, r_w = 1 \$/stb$ and $r_i = 0 \$/stb$
  - Annual discount factor: 15%
- Measurements
  - Fractional flow rates (oil/water)
  - Bottom-hole pressures
- Yearly updates of parameters and control strategy
Closed-loop simulation example

Initial ensemble
Closed-loop simulation example

Ensemble updates at different times
Closed-loop simulation example

Results

- 3 study cases: reactive control, optimal open-loop control based on perfect (‘reality’) model, optimal closed-loop control

![Graph showing the results of different control strategies]

- Reactive control: +8.3%
- Optimal open-loop control: +8.8%
Closed-loop reservoir management

Questions:

• Why are such poor models working so well?
• Does this mean that we don’t need geology?
Reservoir dynamics live in low-order space

- **Observation and control in the wells**
  - Models will typically be poorly observable and/or poorly controllable
  - Real (local) input-output dynamics is of limited order

- **Parameter estimation:**
  - Physical parameters (permeabilities) determine predictive quality but one parameter per grid block leads to excessive over-parametrization (not to be validated)
• How to estimate the parameters in this model?
  • local dynamics versus global physics
  • non-linear batch process
  • extrapolate beyond what you have seen so far
Identifiability

- Consider nonlinear model structure \( \hat{y} = h(\theta, u; x_0) \)
  with \( \hat{y} \) being a prediction of \( y := [y_1^T \cdots y_N^T]^T \)

- Locally identifiable in \( \theta_m \) for given \( u \) and \( x_0 \) if in neighbourhood of \( \theta_m \):
  \[
  \{ h(u, \theta_1; x_0) = h(u, \theta_2; x_0) \} \Rightarrow \theta_1 = \theta_2
  \]
  [Grewal and Glover 1976]

- Global properties are generally hard to analyze
Identifiability

- Notion of *identifiability* is instrumental in analyzing model structure properties

- It determines whether it is feasible at all to relate unique values to the physical parameter variables, on the basis of measured data
Testing local identifiability in identification

- In Prediction Error framework, identification criterion
  \[ V(\theta) := \frac{1}{2} \epsilon(\theta)^T P_v^{-1} \epsilon(\theta), \quad \epsilon(\theta) = y - \hat{y} = y - h(\theta, u; x_0), \]

- Hessian given by
  \[ \frac{\partial^2 V(\theta)}{\partial \theta^2} = \frac{\partial \hat{y}^T}{\partial \theta} P_v^{-1} \left( \frac{\partial \hat{y}^T}{\partial \theta} \right)^T + S \]

- Local identifiability test in \( \hat{\theta} = \text{arg min} \ V(\theta) \) : Hessian > 0

- With quadratic approximation of cost function around \( \hat{\theta} \): Hessian given by
  \[ \frac{\partial \hat{y}^T}{\partial \theta} P_v^{-1} \left( \frac{\partial \hat{y}^T}{\partial \theta} \right)^T \]
Testing local identifiability in identification

• Rank test on Hessian through SVD

\[
\left. \frac{\partial \hat{y}^T}{\partial \theta} P_v^{-\frac{1}{2}} \right|_{\theta = \hat{\theta}} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}
\]

• If \( \Sigma_2 = 0 \) then lack of local identifiability

• SVD can be used to reparameterize the model structure through

\[ \theta = U_1 \rho, \quad \text{dim}(\rho) \ll \text{dim}(\theta) \]

in order to achieve local identifiability in \( \rho \)

• Columns of \( U_1 \) are basis functions of the identifiable parameter space
Testing local identifiability in identification

\[ \frac{\partial \hat{y}^T}{\partial \theta} P_v^{-\frac{1}{2}} \bigg|_{\theta = \hat{\theta}} = \left[ \begin{array}{cc} U_1 & U_2 \end{array} \right] \left[ \begin{array}{cc} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{array} \right] \left[ \begin{array}{c} V_1^T \\ V_2^T \end{array} \right] \]

• What if \( \Sigma_2 \neq 0 \) but contains (many) small singular values?

No lack of identifiability, but possibly very poor variance properties

• Identifiability mostly considered in a yes/no setting: qualitative rather than quantitative [Bellman and Åström (1970), Grewal and Glover (1976)]

• Approach: \textit{quantitative} analysis of appropriate parameter space, maintaining physical parameter interpretation
Model structure approximation

• How to reduce the model structure in terms of its parameter space? (different from “classical” model reduction, in which the model dynamics of a single model is reduced)

• Objective: obtain a physical parametrization (model structure) in which the parameters can be reliably estimated from data.
Approximating the identifiable parameter space

Asymptotic variance analysis: \( \text{cov}(\hat{\theta}) = J^{-1} = \left( \mathbb{E} \left[ \frac{\partial^2 V(\theta)}{\partial \theta^2} \bigg| \hat{\theta} \right] \right)^{-1} \)

with \( J = \) Fisher Information Matrix.

- Sample estimate of parameter variance, on the basis of \( V(\theta) \):

\[
\text{cov}(\hat{\theta}) = \begin{cases} 
\left[ \begin{array}{cc}
U_1 & U_2 \\
0 & \Sigma_2^{-2} \\
\infty & 0
\end{array} \right] 
\left[ \begin{array}{cc}
\Sigma_1^{-2} & 0 \\
0 & \Sigma_2^{-2}
\end{array} \right] 
\left[ \begin{array}{c}
V_1^T \\
V_2^T
\end{array} \right] 
& \text{for } \Sigma_2 > 0 \\
\infty & \text{for } \Sigma_2 = 0
\end{cases}
\]

\[
\text{cov}(U_1 \hat{\rho}) = U_1 \Sigma_1^{-2} U_1^T
\]

\( \text{cov}(\hat{\theta}) > \text{cov}(U_1 \hat{\rho}) \) if \( \Sigma_2 > 0 \)
Approximating the identifiable parameter space

\[ \text{cov}(\hat{\theta}) > \text{cov}(U_1 \hat{\rho}) \quad \text{if } \Sigma_2 > 0 \]

- Discarding singular values that are small, reduces the variance of the resulting parameter estimate.
- Particularly important in situations of (very) large numbers of small s.v.’s.
- Model structure approximation (local).
- Quantified notion of identifiability – related to parameter variance.
Approximating the identifiable parameter space

- Interpretation:
  Remove the parameter directions that are poorly identifiable (have large variance)

- This is different from removing the (separate) parameters for which the value 0 lies within the confidence bound [Hjalmarsson, 2005]
A Bayesian approach

- Often applied method for dealing with overdetermination in parameter space:

- Incorporate prior knowledge term (regularization) in cost function

\[ V_p(\theta) := V(\theta) + \frac{1}{2}(\theta - \theta_p)P_{\theta_p}^{-1}(\theta - \theta_p) \]

where \( \theta_p \) is the prior parameter vector (with covariance \( P_{\theta_p} \)).

- Model output approximated with first-order Taylor expansion. Hessian is

\[ \frac{\partial^2 V_p(\theta)}{\partial \theta^2} = \frac{\partial h(\theta)^T}{\partial \theta} P_v^{-1} \left( \frac{\partial h(\theta)^T}{\partial \theta} \right)^T + P_{\theta_p}^{-1} \]

- “Always” identifiable, since \( P_{\theta_p} \) full rank by construction!!
A Bayesian approach

Implications

• Bayesian methods seem not to suffer from identifiability problems......

• This includes all (extended) Kalman filter type algorithms. Where parameters are recursively estimated by augmenting the states

• Unique parameter estimates usually result, but

• In the parameter subspace that is poorly identifiable, estimated parameters will be heavily dominated by the prior information.

• Analysis of $V(\theta)$ can show identifiable directions (locally)
Simple reservoir example

(top view)

21 x 21 grid block permeabilities
1 injection well (center); 4 production wells (corners);
3 permeability strokes
Simple reservoir example

Singular vectors can be projected on the grid:

First 12 singular vectors; identifiability case
Simple reservoir example

Using the reduced parameter space – iteratively – in identification:

Observation:
Only grid block permeabilities around well are identifiable.
Simple reservoir example

- Grid block properties far away from wells are poorly identifiable
- There are indications that they might not be very important for the optimal control strategy……
Discussion

- Estimating physical parameters in reservoirs is challenging and highly relevant
- Nonlinear (one-go) batch process complicates this
- Model structure approximation is required in order to guarantee identifiability
- Analysis can only be done locally linearized
- Identification of local linear models can serve a shorter time optimization (Van Essen, CDC 2010)
- Many model-based optimization challenges in this field

Cooperative project, together with:

Jorn Van Doren, Gijs van Essen, PhD students

Jan Dirk Jansen
Dept. of Geotechnology TUD and Shell Intern. Exploration and Production

Sippe Douma, Shell, and former PhD student

Okko Bosgra, DCSC, TUD

Consortia: VALUE (Shell, TU Delft, MIT), ISAPP (Shell, TNO, TU Delft)