Model reduced variational data assimilation: An ensemble approach to model calibration

Arnold Heemink
Delft University of Technology
Joint work with Gosia Kaleta, Remus Hanea, Jan Dirk Jansen, Martin Verlaan and Umer Altaf
Outline

• Background of parameter estimation or history matching using 4DVar
• Motivation for an efficient and adjoint-free history matching procedure
• Model-reduced gradient-based history matching:
  Proper Orthogonal Decomposition (POD)
  Balanced Proper Orthogonal Decomposition (BPOD)
• Model reduced Variational Data assimilation
• Results:
  Reservoir models
  Tidal model of the North Sea
• Conclusions
History matching

Parameter values are identified by minimizing an objective function that represents the mismatch between modeled and observed production data

\[ \min_{\theta} J(\theta) = \frac{1}{2} (\theta - \theta^\text{prior})^T P^{-1}_\theta (\theta - \theta^\text{prior}) + \frac{1}{2} \sum_{i=1}^{N_D} (y^\text{obs}_i - y_i(\theta))^T P^{-1}_i (y^\text{obs}_i - y_i(\theta)) \]

where

\[ y(t_i) = h_i(x(t_i), \theta) \in \mathbb{R}^m, \quad m \sim O(10) \]
\[ x(t_i) = f_i(x(t_{i-1}), \theta) \in \mathbb{R}^h, \quad h \in \mathbb{R}^h, \quad h \sim O(10^6) \]

\( x_i \) represents system variables at time \( t_i \)
\( f_i \) represents the system evolution at time
\( h_i \) represents the analytical relation between the system variable and data
Iterative gradient-based optimization scheme (often BFGS) where the gradients are computed by using the adjoint model.

To compute the gradient an adjoint model is implemented

\[ \lambda(t_i) = \left( \frac{\partial f_{i+1}[x(t_i), \theta]}{\partial x(t_i)} \right)^T \lambda(t_{i+1}) + \left( \frac{\partial J(\theta, x(t_1), \ldots, x(t_N))}{\partial x(t_i)} \right)^T \]

where \( f \) represents the reservoir model, \( \lambda \) represents the adjoint variable and the gradient is given by:

\[ \frac{dJ(\theta)}{d\theta} = \sum_{i=1}^{n} -[\lambda(t_i)]^T \left[ \frac{\partial f_i[x(t_{i-1}), \theta]}{\partial \theta} \right] \]
Motivation

4DVar or the adjoint method
- **Numerically efficient** way to calculate the gradient (one gradient calculation requires only one forward solution and one adjoint solution regardless of the number of model parameters)
- Very **difficult to implement** the adjoint of the tangent linear approximation of the forward model
- Requires **access to the simulation code**

The reservoir system
Large-order reservoir models → the intrinsic order of the system is (much) lower than the number of grid blocks in the model (small “input space”)
Very sparse data in space → gives information only around the wells (small “output space”)

Sketch of the method
Model-reduced gradient-based history matching

**Projection based method**
Suppose the dynamics of a system are described by

\[ x(t_i) = f_i(x(t_{i-1}), \theta) \in \mathbb{R}^h, \quad h \sim O(10^6) \]
\[ y(t_i) = h_i(x(t_{i-1}), \theta) \in \mathbb{R}^m, \quad m \sim O(10) \]

Petrov-Galerkin projection specifies the dynamics of a variable \( z(t_i) \in \text{span}\{\varphi_1, \ldots, \varphi_k\} \)
by

\[ z(t_i) = \Psi^T f_i(\Phi z(t_{i-1}), \theta) \in \mathbb{R}^k, \quad k \sim O(10^2) \]
\[ y(t_i) = h_i(\Phi z(t_i), \theta) \in \mathbb{R}^m, \quad m \sim O(10) \]

**Proper Orthogonal Decomposition**
Suppose we have a set of data \( \mathbf{X} = \{x_1(t_1), \ldots, x_i(t_N), \ldots, x_p(t_1), \ldots, x_p(t_N)\} \), \( x_j(t_i) \in \mathbb{R}^n \)
We seek a projection \( \Pi = \Phi \Psi^T \) of a fixed rank \( k \) such that minimizes the total error

\[ \sum_{j=1}^p \sum_{i=1}^N \left\| x_j(t_i) - \Pi x_j(t_i) \right\|_2 \]

The optimal subspace of dimension \( k \) is given by the first \( k \) eigenvectors of the covariance matrix of state variables generated by the data \( \mathbf{X} \) and the state can be approximated as

\[ x(t_i) \approx \Phi r(t_i) \]

where \( \Phi \) consists of \( k \) first eigenvectors of \( \mathbf{X} \mathbf{X}^T \) and \( \Psi = \Phi \)
Model-reduced gradient-based history matching

The tangent linear approximation of the reservoir model is given by

$$\Delta x(t_i) = F_x \Delta x(t_{i-1}) + F_\theta \Delta \theta$$

Then assuming that $\Delta x(t_i) \approx \Phi r(t_i)$ we obtain

$$r(t_i) = \Psi^T F_x \Phi r(t_{i-1}) + \Psi^T F_\theta \Delta \theta$$

It is low-order approximation of the original model and has easily available adjoint model

$$\zeta(t_i) = \Phi^T F_x^T \Psi \zeta(t_{i-1}) + \left( \nabla_{r(t_i)} J \right)^T$$

The $F_x \Phi$ and $F_\theta \Psi$ are approximated by finite differences:

$$F_x \Phi_j = \frac{\partial f_i(x(t_{i-1}), \theta)}{\partial x(t_{i-1})} \Phi_j \approx \frac{f_i(x(t_{i-1}), \theta) - f_i(x(t_{i-1}) + \varepsilon \Phi_j, \theta)}{\varepsilon}$$

$$F_\theta \Psi^j = \frac{\partial f_i(x(t_{i-1}), \theta)}{\partial \theta} \Psi^j \approx \frac{f_i(x(t_{i-1}), \theta) - f_i(x(t_{i-1}), \theta + \varepsilon \Phi^j, \theta)}{\varepsilon}$$
Model-reduced gradient-based history matching

The linear reduced model can now be given in state space form by:

\[
\begin{bmatrix}
    r(t_i) \\
    \Delta \theta(t_i)
\end{bmatrix} = \begin{bmatrix}
    \Psi^T F_x \Phi \\
    0
\end{bmatrix} \begin{bmatrix}
    r(t_{i-1}) \\
    \Delta \theta(t_{i-1})
\end{bmatrix}
\]

So the (variation of the) original parameter vector is still part of the reduced model. The reduction of the number of parameters is done separately.

Remark:
The reduced model should only be able to reproduce the input-output behavior of the original model.
Model-reduced gradient-based history matching

**Balanced Proper Orthogonal Decomposition**

*Controllability Gramian* measures to what extent the state of the model can be influenced by manipulating the input; Can be approximated by snapshots of the forward model.

*Observability Gramian* measures to what extent the state influences the outputs; Can be approximated by snapshots of the adjoint model.

Solve the SVD of the matrix

\[
Y^T X = U \Sigma V^T = [U_1 \quad U_2] \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}
\]

where \( X \) is the set of snapshots the reservoir model and \( Y \) is the set of snapshots from the adjoint model.

The balancing transformation is given as

\[
\Phi = XV_1 \Sigma^{-1/2}
\]

And:

\[
\Psi^T = \Sigma^{-1/2} U_1^T Y^T
\]

This is a **balanced POD method**. We need the adjoint for the state now!
Model-reduced gradient-based history matching

1. Initial Parameters
2. High-order Reservoir Model Simulation
3. Objective Function Calculation

Converged?

- NO
  - Runs of the simulator to get snapshots
  - Building of the Low-order Model
  - Low-order Model Simulation
  - Low-order Adjoint Model Simulation
  - Gradient Calculation
  - Parameters Update
  - Reduced Objective Function Calculation
  - Converged?
    - NO
      - Improve Sub Optimal Parameters
      - Improve Parameters in reduced space
    - YES
      - Sup Optimal Parameters
  - YES
    - Done
Remarks

• The computational effort is dominated by the generation of the reduced model (generating snapshots, computing the sensitivity matrices). The number of model simulations is roughly the dimension of the reduced model.

• The approach is very efficient in case the simulation period of the ensemble of model simulations can be chosen very small compared to the calibration period (unfortunately this is not the case for reservoir modelling problems).

• The number of outer loop iterations is usually very small: 2-5. For most iterations the reduced model can be the same and only the residuals have to be updated.

• Model-reduced history matching is very well suited for parallel processing since the ensemble of model simulations can be created completely independent of each other.

• If the complete tangent linear model is available the reduced model can be obtained easily and the approach is very efficient: the number of model simulations required is a little more the number of parameters.
Results: **Synthetic example 1 and 2**

**Reservoir (2D)**
21x21x1 grid blocks

**Phases**
Oil and water; relative permeability curves are known

**Reservoir simulator**
Simplifications: absence of gravity forces and capillary pressures; isotropic permeability; parameter independence on pressure

**Setup for experiments**
- Five spot injection - production pattern
- Reservoir is operated on rate constraint in the injection well and on bottom hole pressure constraints in production wells

**Measurements**
- Measurements taken each 30 days during 250 days, before water breakthrough
- Bottom hole pressure measurements from injection well with 10% error
- Flow rate measurements from production wells with 5% error
Results: Parameters reduction

The prior knowledge is given by an ensemble
Results: Synthetic example 1

Gradients comparison

POD
POD-based gradient

ADJ
ADJ gradient

BPOD
BPOD-based gradient

States reconstruction

Original model at timestep 30

POD
Reduced-order model at timestep 30

BPOD
Reduced-order model at timestep 30

TU Delft
Results: Synthetic example 1

Log permeability fields

True Prior POD BPOD ADJ

Numerical efficiency data

<table>
<thead>
<tr>
<th>Method</th>
<th>Objective function</th>
<th>Reduction</th>
<th>Time in simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial (Prior)</td>
<td>1657</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Adjoint-based approach (30 iter)</td>
<td>20.05</td>
<td>441 (sat) + 441 (prf) + 20 (perm)</td>
<td>61 (=1+30*2)</td>
</tr>
<tr>
<td>Adjoint-based approach</td>
<td>18.65</td>
<td>441 (sat) + 441 (prf) + 20 (perm)</td>
<td>135 (= 67*2+1)</td>
</tr>
<tr>
<td>POD-based approach</td>
<td>20.27</td>
<td>41(99%) + 31(99.9%) + 20</td>
<td>115 (=1+20+72+20+1+2)</td>
</tr>
<tr>
<td>Balanced POD-based approach</td>
<td>18.73</td>
<td>42(99.9%) + 6 (99.9%) + 20</td>
<td>114 (=1+2*20+48+20+5)</td>
</tr>
</tbody>
</table>
Results: Synthetic example 1

Prediction of the water production rate

![Graphs showing water production rates for Producer South West and Producer South East with different log types: true logθ, prior logθ, ADJ logθ, FD logθ, POD logθ, and BPOD logθ.](image)
**Results: Synthetic example 2**

### Log permeability fields

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>Prior</th>
<th>BPOD</th>
<th>POD</th>
<th>ADJ</th>
</tr>
</thead>
</table>

### Numerical efficiency data

<table>
<thead>
<tr>
<th>Method</th>
<th>Objective function</th>
<th>Reduction</th>
<th>Time in simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial (Prior)</td>
<td>226.49</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>Adjoint-based approach (30 iter)</td>
<td>21.21</td>
<td>441 (sat) + 441 (prf) + 20 (perm)</td>
<td>61 (=1+30*2)</td>
</tr>
<tr>
<td>Adjoint-based approach</td>
<td>20.33</td>
<td>441 (sat) + 441 (prf) + 20 (perm)</td>
<td>113 (=1+56*2)</td>
</tr>
<tr>
<td>POD-based approach</td>
<td>22.04</td>
<td>42 (99%) + 30 (99.9%) + 20</td>
<td>114 (=1+20+72+20)</td>
</tr>
<tr>
<td>BPOD-based approach</td>
<td>20.44</td>
<td>49 (99.9%) + 9 (99.9%) + 20</td>
<td>121 (=1+2*20+58+20)</td>
</tr>
</tbody>
</table>
Results: **Synthetic example 2**

The prediction of water breakthrough time and produced water flow rates

![Graphs showing water rate versus time for different logs.

Producer North East:
- true logθ
- prior logθ
- ADJ logθ
- FD logθ
- POD logθ
- BPOD logθ

Producer North West:
- true logθ
- prior logθ
- ADJ logθ
- FD logθ
- POD logθ
- BPOD logθ

TU Delft logo located at the bottom right corner of the page.
More realistic study case

- **Reservoir model assumption**
  - 3 dimensional (60x60x7 with 18553 active grid blocks)
  - Two-phase (oil-water)
  - No-flow boundaries at all sides

- **Measurements**
  - Bottom hole pressures from injectors each 60 days during 3 years
  - Flow rates from producers each 60 days during 3 years

Gijs van Essen [2006]
## Results

<table>
<thead>
<tr>
<th>Methods</th>
<th>Nr of model simulations</th>
<th>Objective function</th>
<th>Permeability patterns</th>
<th>State patterns</th>
<th>Number of snapshots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial (Prior)</td>
<td>-</td>
<td>346</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Adjoint-based approach</td>
<td>~ 15*2 + 45 (6)</td>
<td>98</td>
<td>22</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>POD-based model-reduced approach</td>
<td>~ 68 (29+6+11+22)</td>
<td>114</td>
<td>22</td>
<td>29+6</td>
<td>400</td>
</tr>
<tr>
<td>Balanced-POD-based model-reduced approach</td>
<td>~ 59 (6+6+2*11+22+3)</td>
<td>117</td>
<td>22</td>
<td>6+6</td>
<td>400</td>
</tr>
</tbody>
</table>

![Graphs of true, prior, POD-based, and adjoint log perm fields](image)
Calibration of a large scale numerical tidal model
Based on shallow water equations
- Grid size: 1.5’ by 1.0’ (~2 km)
- Grid dimensions: 1120 x 1260 cells
- Active Grid Points: 869544
- Time step: 2 minutes
- 8 main constituents
Computational grid near Dutch coast
Twin experiment:
Estimation of 7 depth parameters using generated data (noise free)
The POD modes capture energy in case of 200 snapshots.
Experiment with field data

- Parameter: Depth
- Calibration run: 28 Dec 2006 to 30 Jan 2007
- Measurement data: 01 Jan 2007 to 30 Jan 2007
- Includes two spring-neap cycles
- Assimilation Stations: 35
- Validation Stations: 15
- Ensemble of forward model simulations for a period of four days (01 Jan 2007 to 04 Jan 2007)
DCSM

- Divide model area in 4 sub domains + 1 overall parameter
- No. of snapshots: 132 (Every three hours)
- 24 POD modes are required to capture 97% energy
- Same POD modes are used in 2\textsuperscript{nd} iteration

- Initial RMS: 25.7 cm
- After 2\textsuperscript{nd} iteration: 14.9 cm
- Improvement: 42%
DCSM (Validation results)

Similar improvement as in the case of assimilation stations
Computational Cost

Estimation 5 parameters, calibration period 1 month:
Number of simulations of 1 month: 4.7, reduction criterion 42% (2 iterations, no model update in second iteration)

Estimation 20 parameters (4 bottom friction and 16 depth values), calibration period 1 month: Number of simulations of 1 month: 11, reduction criterion 50% (5 iterations, no model update in second and fourth iteration)
Conclusions

• POD-based model-reduced approach does not require the implementation of the adjoint of the tangent linear model of the original reservoir model.
• Model-reduced gradient-based algorithms provides for reservoir models parameter estimates with comparable accuracy as those obtained using a classical adjoint-based method.
• The efficiency of the approach depends very much on the application: A very good efficiency is obtained if the time scale of the model is much smaller than the calibration period.
• The maximum number of parameters is, say, a few hundred.
• The balanced POD-based method is a little bit more efficient then the POD-based method, but requires the Jacobians of the original model.
• If the adjoint is available both POD approaches are significantly more efficient then the classical adjoint method (if the number of parameters is not too large).
For more information see:

Model-reduced gradient-based history matching”, Kaleta, MP, Hanea, RG, Heemink, AW and Jansen JD, Computational Geosciences, 2011
