THE VALUE OF SMARTNESS

ISAPP

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

TU Delft

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

Background: History matching problem (4DVar)

History matching

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

$$\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) = \frac{1}{2}\left(\boldsymbol{\theta}-\boldsymbol{\theta}^{prior}\right)^T \mathbf{P}_{\theta}^{-1}\left(\boldsymbol{\theta}-\boldsymbol{\theta}^{prior}\right) + \frac{1}{2}\sum_{i=1}^{N_D}\left(\mathbf{y}_i^{obs}-\mathbf{y}_i(\boldsymbol{\theta})\right)^T \mathbf{P}_i^{-1}\left(\mathbf{y}_i^{obs}-\mathbf{y}_i(\boldsymbol{\theta})\right)$$

where

$$\mathbf{y}(t_i) = \mathbf{h}_i\left(\mathbf{x}(t_i),\boldsymbol{\theta}\right) \in \mathbf{R}^m, \quad m \sim O(10)$$

$$\mathbf{x}(t_i) = \mathbf{f}_i\left(\mathbf{x}(t_{i-1}),\boldsymbol{\theta}\right) \in \mathbf{R}^h, \quad \boldsymbol{\theta} \in \mathbf{R}^h, \quad h \sim O(10^6)$$

$\mathbf{x}_i$ represents system variables at time t_i

$\mathbf{f}_i$ represents the system evolution at time

$\mathbf{h}_\mathbf{i}$ represents the analytical relation between the system variable and data

Background: History matching problem (4DVar)

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

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

where $\mathbf{f}$ represents the reservoir model, $\boldsymbol{\lambda}$ represents the adjoint variable and the gradient is given by:

$$\frac{dJ(\boldsymbol{\theta})}{d\boldsymbol{\theta}} = \sum_{i=1}^{n} -[\boldsymbol{\lambda}(t_i)]^T [\frac{\partial \mathbf{f}_i[\mathbf{x}(t_{i-1}), \boldsymbol{\theta}]}{\partial \boldsymbol{\theta}}]$$

Motivation

4DVar or the adjoint method

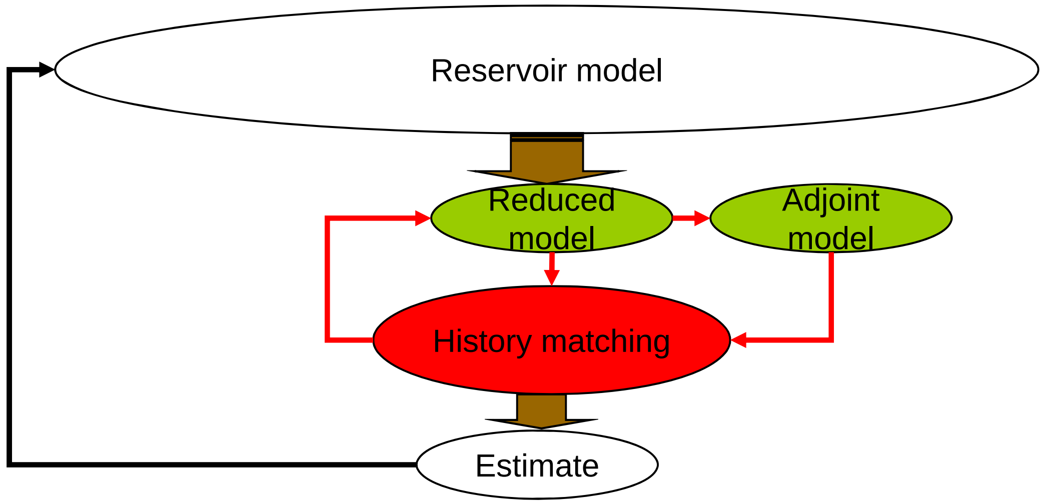
- **Numerically efficient** way to calculate the gradient (one gradient calculation requires only one forward solution and one adjoint solution regardless of the number of model parameters)
- Very **difficult to implement** the adjoint of the tangent linear approximation of the forward model
- Requires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

$$\mathbf{x}(t_i) = \mathbf{f}_i\left(\mathbf{x}(t_{i-1}), \boldsymbol{\theta}\right) \in \mathbf{R}^h, \quad h \sim O(10^6)$$
$$\mathbf{y}(t_i) = \mathbf{h}_i\left(\mathbf{x}(t_{i-1}), \boldsymbol{\theta}\right) \in \mathbf{R}^m, \quad m \sim O(10)$$

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

$$\mathbf{z}(t_i) = \boldsymbol{\Psi}^T \mathbf{f}_i\left(\boldsymbol{\Phi}\mathbf{z}(t_{i-1}), \boldsymbol{\theta}\right) \in \mathbf{R}^k, \quad k \sim O(10^2)$$
$$\mathbf{y}(t_i) = \mathbf{h}_i\left(\boldsymbol{\Phi}\mathbf{z}(t_i), \boldsymbol{\theta}\right) \in \mathbf{R}^m, \quad m \sim O(10)$$

Proper Orthogonal Decomposition

Suppose we have a set of data $\mathbf{X} = \left\{\mathbf{x}_1(t_1), \ldots, \mathbf{x}_1(t_N), \ldots, \mathbf{x}_p(t_1), \ldots, \mathbf{x}_p(t_N)\right\}$, $\mathbf{x}_j(t_i) \in \mathbb{R}^n$

We seek a projection $\boldsymbol{\Pi} = \boldsymbol{\Phi}\boldsymbol{\Psi}^T$ of a fixed rank k such that minimizes the total error

$$\sum_{j=1}^{p} \sum_{i=1}^{N} \left\| \mathbf{x}_j(t_i) - \boldsymbol{\Pi}\mathbf{x}_j(t_i) \right\|_2$$

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

$$\mathbf{x}(t_i) \approx \boldsymbol{\Phi}\mathbf{r}(t_i)$$

where $\boldsymbol{\Phi}$ consists of *k* first eigenvectors of $\mathbf{X}\mathbf{X}^T$ and $\boldsymbol{\Psi} = \boldsymbol{\Phi}$

TU Delft

Model-reduced gradient-based history matching

The tangent linear approximation of the reservoir model is given by

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

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

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

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

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

The $\mathbf{F_x\Phi}$ and $\mathbf{F}_\theta\Psi$ are approximated by finite differences:

$$\mathbf{F_x}\mathbf{\Phi}_j = \frac{\partial\mathbf{f}_i\left(\mathbf{x}(t_{i-1}),\mathbf{\theta}\right)}{\partial\mathbf{x}(t_{i-1})}\mathbf{\Phi}_j \approx \frac{\mathbf{f}_i\left(\mathbf{x}(t_{i-1}),\mathbf{\theta}\right) - \mathbf{f}_i\left(\mathbf{x}(t_{i-1}) + \varepsilon\mathbf{\Phi}_j,\mathbf{\theta}\right)}{\varepsilon}$$

$$\mathbf{F_\theta}\Psi^j = \frac{\partial\mathbf{f}_i\left(\mathbf{x}(t_{i-1}),\mathbf{\theta}\right)}{\partial\mathbf{\theta}}\Psi^j \approx \frac{\mathbf{f}_i\left(\mathbf{x}(t_{i-1}),\mathbf{\theta}\right) - \mathbf{f}_i\left(\mathbf{x}(t_{i-1}),\mathbf{\theta} + \varepsilon_\theta\mathbf{\Phi}_\theta^j\right)}{\varepsilon_\theta}$$

TU Delft

Model-reduced gradient-based history matching

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

$$\begin{bmatrix} \mathbf{r}(t_i) \\ \Delta\boldsymbol{\theta}(t_i) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}^T \mathbf{F}_{\mathbf{x}} \boldsymbol{\Phi} & \boldsymbol{\Psi}^T \mathbf{F}_{\theta} \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{r}(t_{i-1}) \\ \Delta\boldsymbol{\theta}(t_{i-1}) \end{bmatrix}$$

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

$$\mathbf{Y}^T\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = [\mathbf{U}_1 \quad \mathbf{U}_2]\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}$$

where $\mathbf{X}$ is the set of snapshots the reservoir model and $\mathbf{Y}$ is the set of snapshots from the adjoint model

The balancing transformation is given as

$$\mathbf{\Phi} = \mathbf{X}\mathbf{V}_1\mathbf{\Sigma}^{-1/2}$$

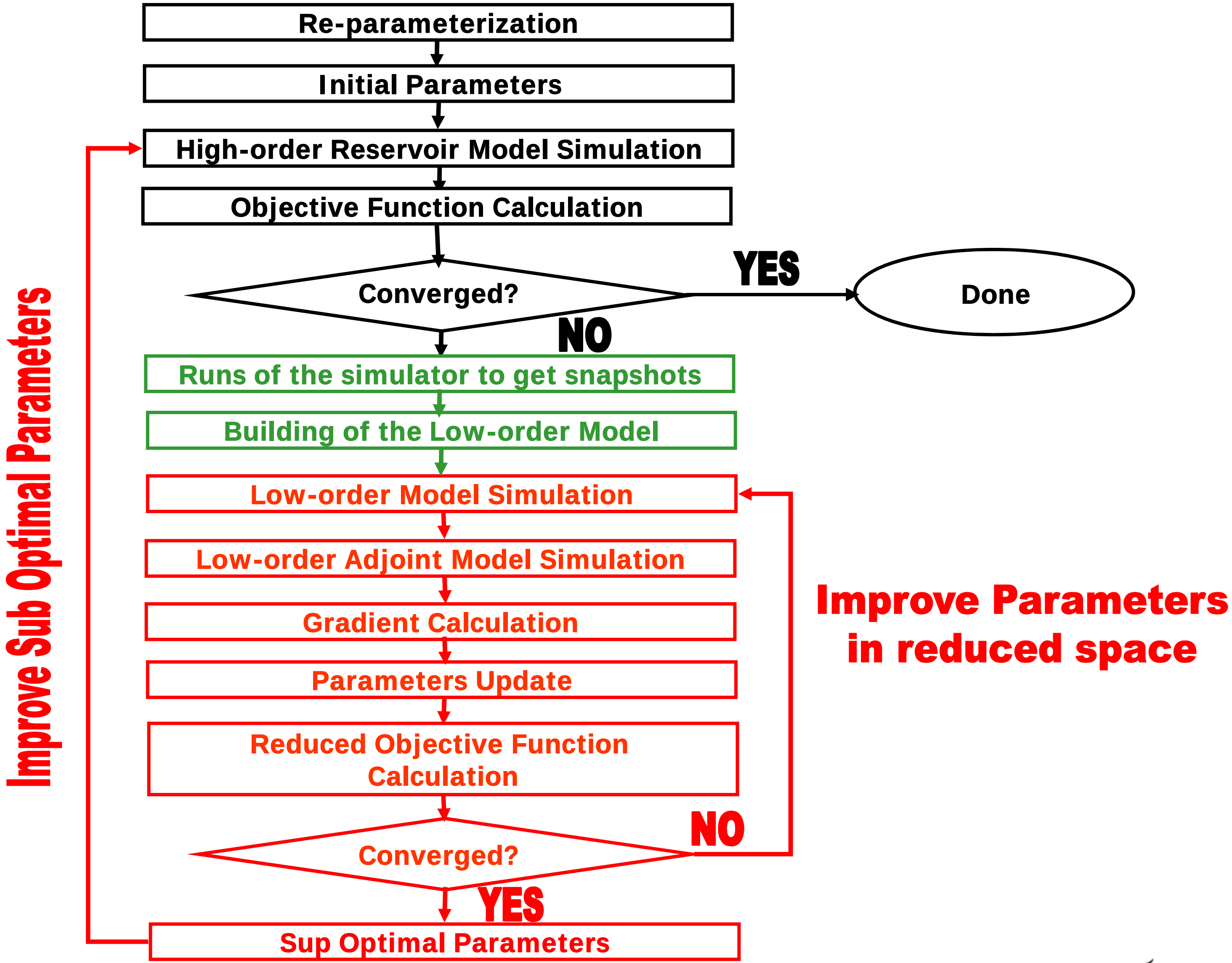
And:

$$\mathbf{\Psi}^T = \mathbf{\Sigma}^{-1/2}\mathbf{U}_1^T\mathbf{Y}^T$$

This is a **balanced POD method**. We need the adjoint for the state now!

TU Delft

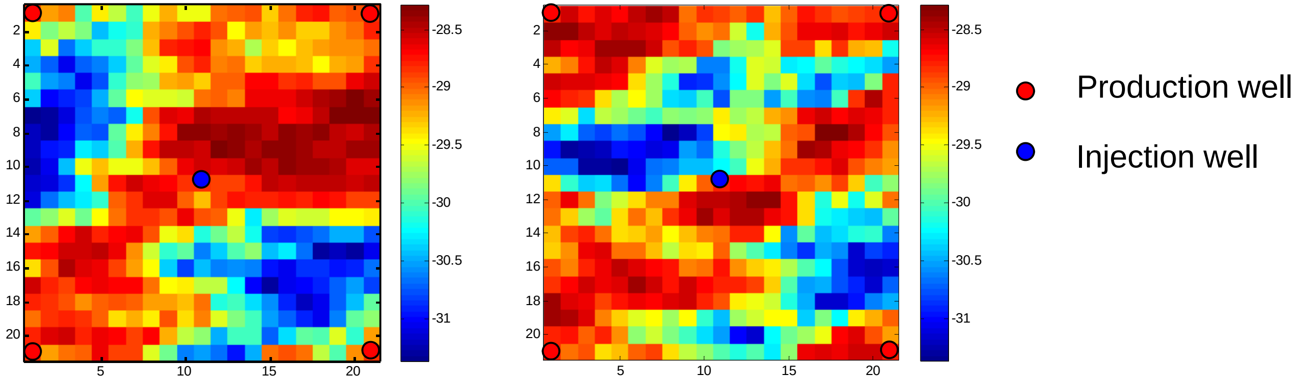
Model-reduced gradient-based history matching



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 de 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

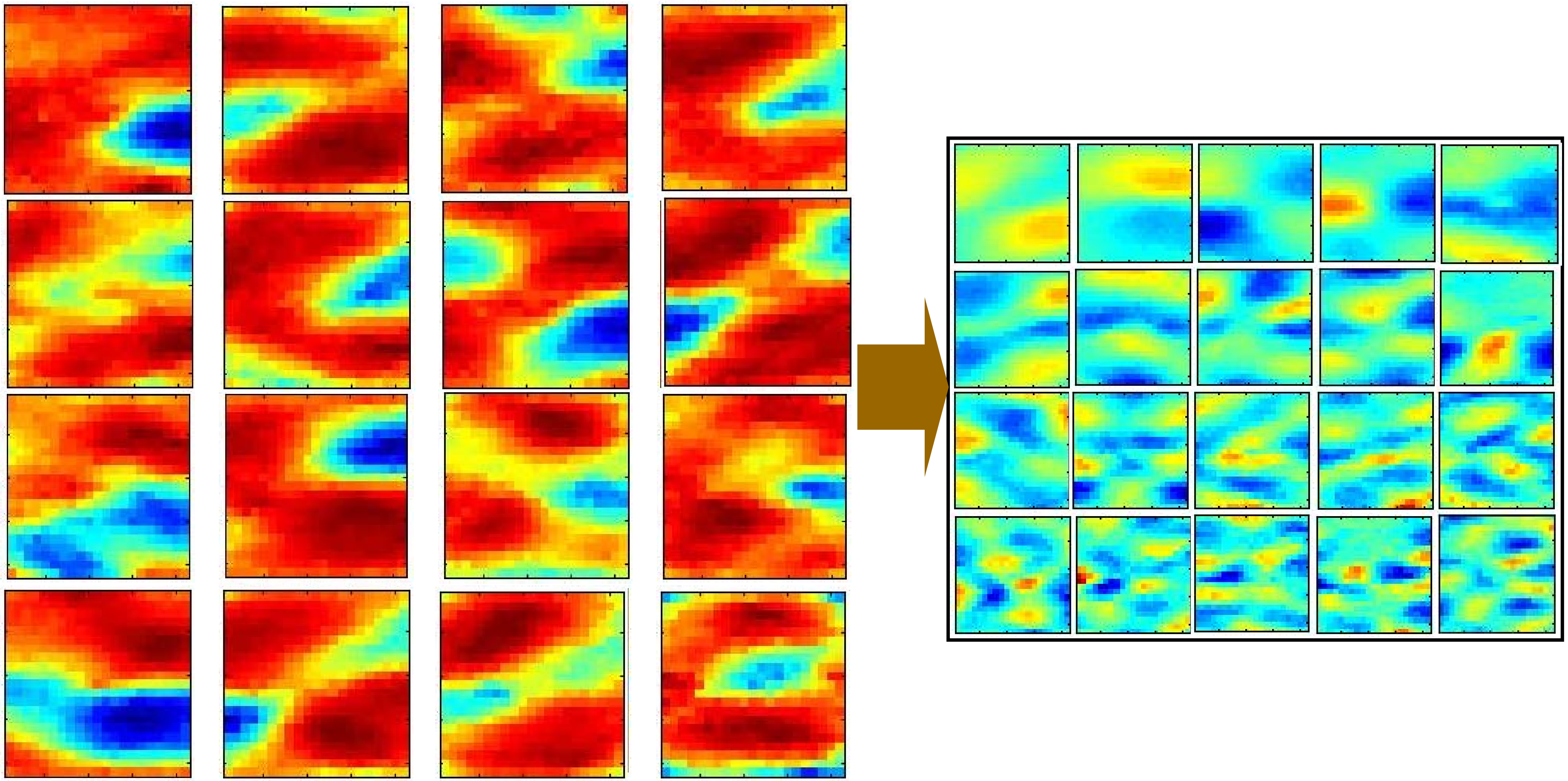
- Five spot injection - production pattern
- Reservoir is operated on rate constraint in the injection well and on bottom hole pressure constraints in production wells

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

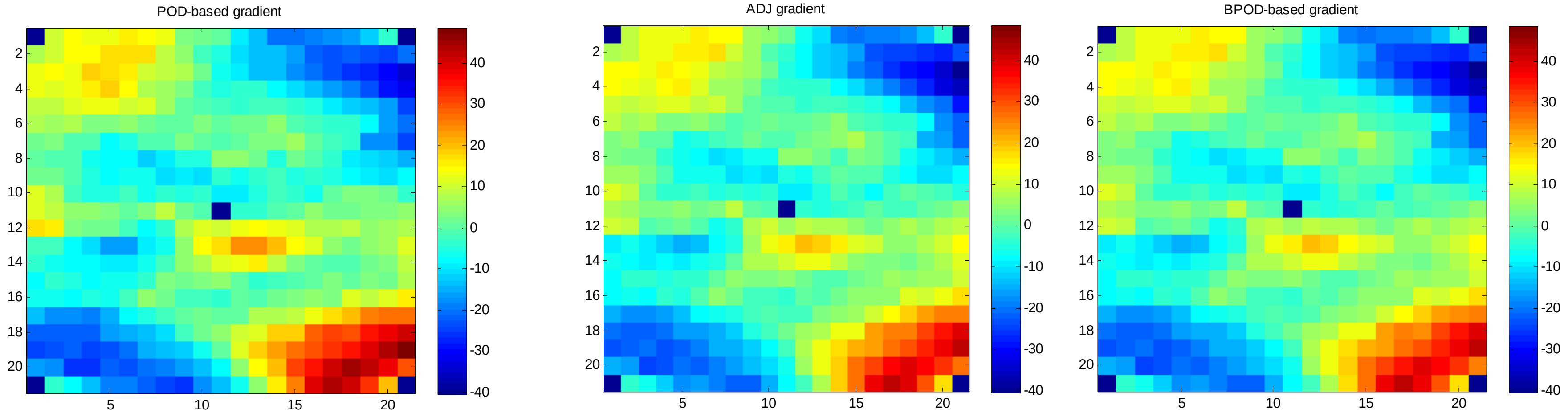
The prior knowledge is given by an ensemble



Results: Synthetic example 1

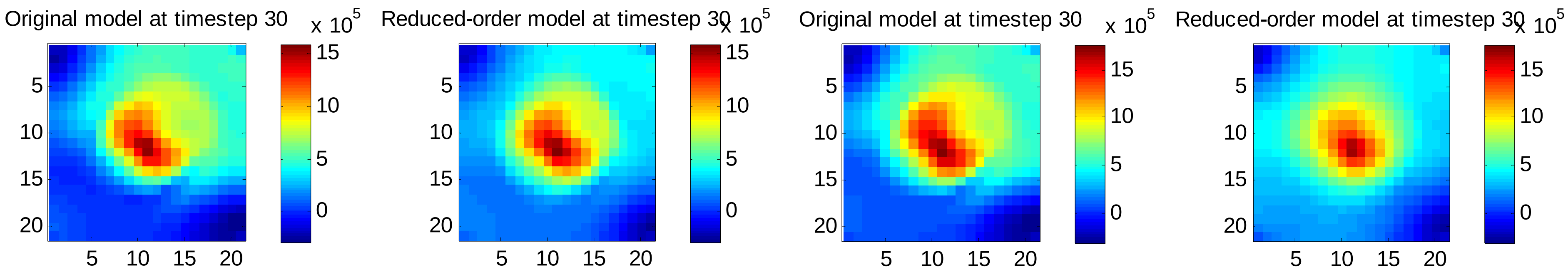
Gradients comparison

POD ADJ BPOD



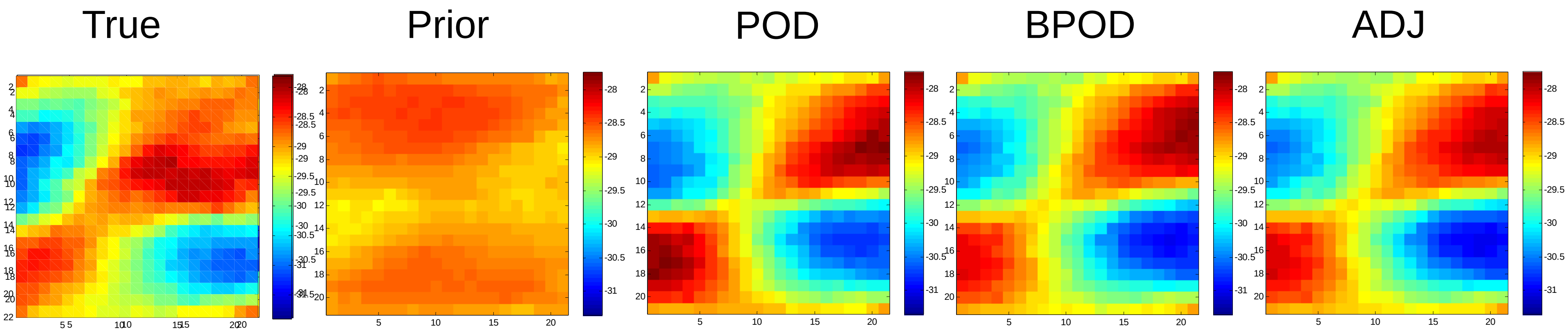
States reconstruction

POD BPOD



Results: Synthetic example 1

Log permeability fields

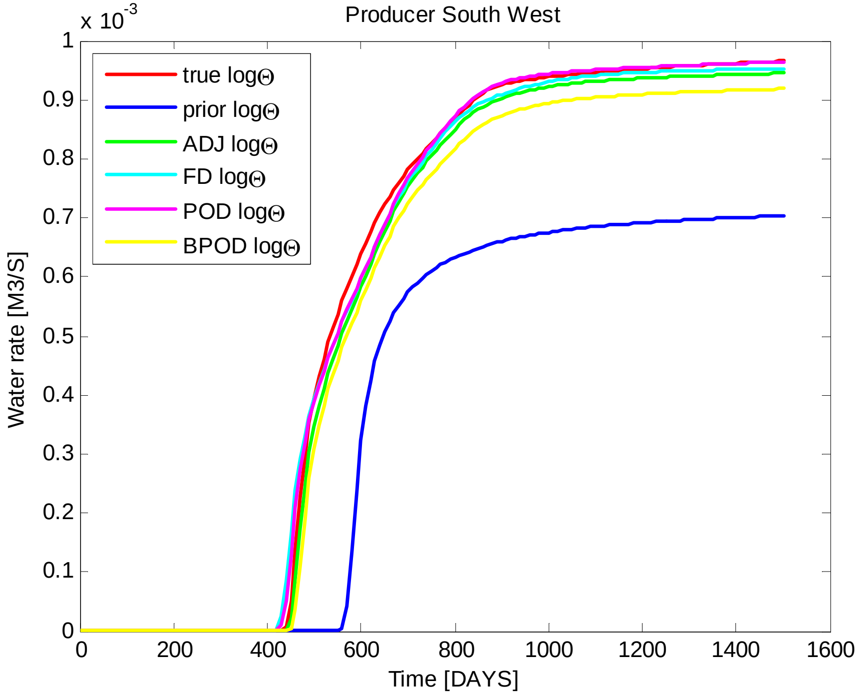


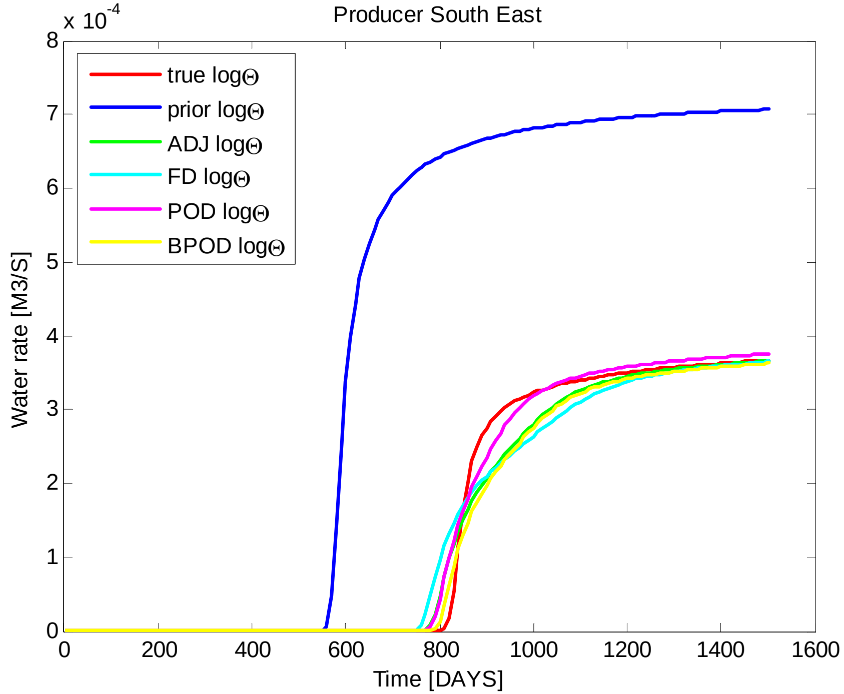
Numerical efficiency data

Method	Objective function	Reduction	Time in simulations
Initial (Prior)	1657	-	1
Adjoint-based approach (30 iter)	20.05	441 (sat) + 441 (prf) + 20 (perm)	61 (=1+30*2)
Adjoint-based approach	18.65	441 (sat) + 441 (prf) + 20 (perm)	135 (= 67*2+1)
POD-based approach	20.27	41(99%) + 31(99.9%)+ 20	115 (=1+20+72+20+1+2)
Balanced POD-based approach	18.73	42(99.9%) + 6 (99.9%)+20	114 (=1+2*20+48+20+5)

Results: Synthetic example 1

Prediction of the water production rate

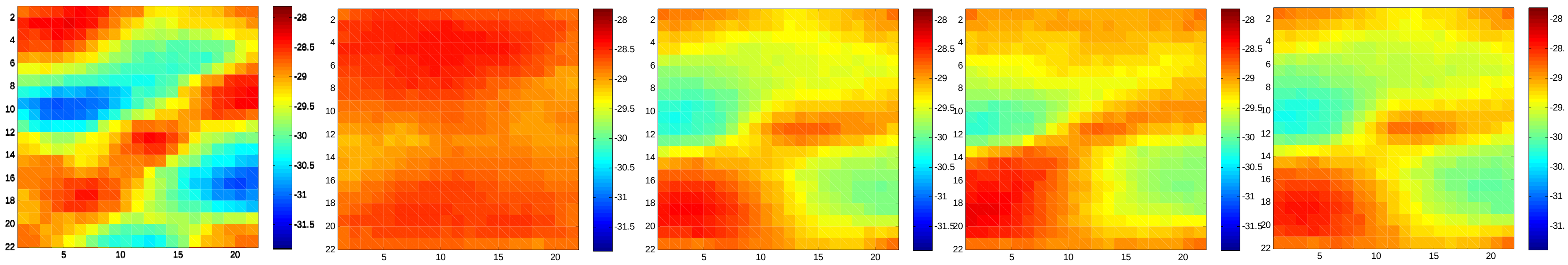




Results: Synthetic example 2

Log permeability fields

True | Prior | BPOD | POD | ADJ



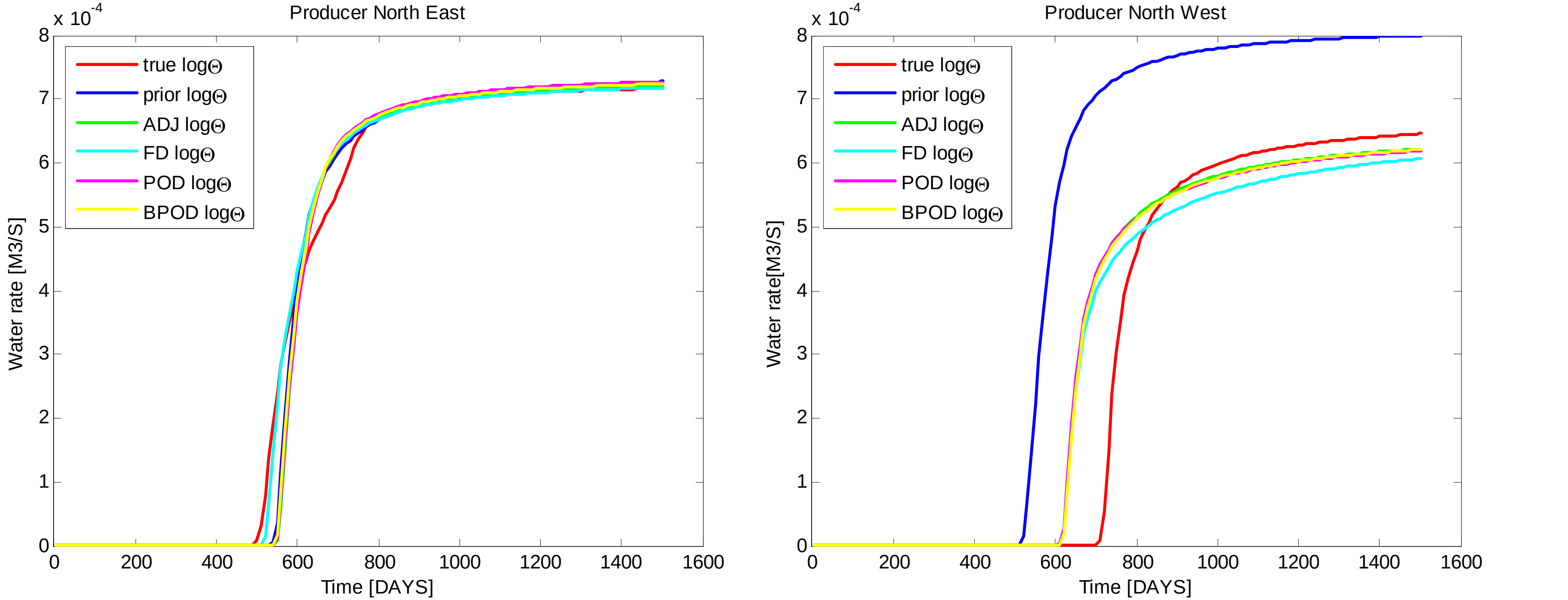
Numerical efficiency data

Method	Objective function	Reduction	Time in simulations
Initial (Prior)	226.49	-	1
Adjoint-based approach (30 iter)	21.21	441 (sat) + 441 (prf) + 20 (perm)	61 (=1+30*2)
Adjoint-based approach	20.33	441 (sat) + 441 (prf) + 20 (perm)	113 (=1+56*2)
POD-based approach	22.04	42 (99%) + 30 (99.9%) + 20	114 (=1+20+72+20)
BPOD-based approach	20.44	49 (99.9%) + 9 (99.9%) + 20	121 (=1+2*20+58+20)

TUDelft

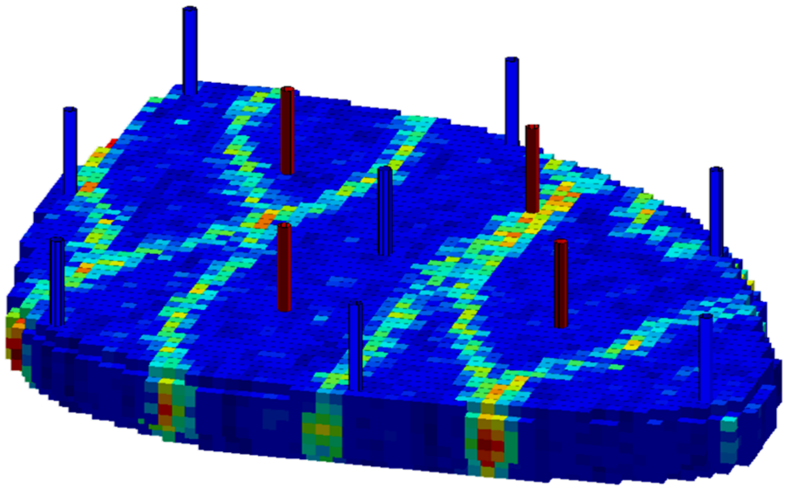
Results: Synthetic example 2

The prediction of water breakthrough time and produced water flow rates



TU Delft

More realistic study case



Gijs van Essen [2006]

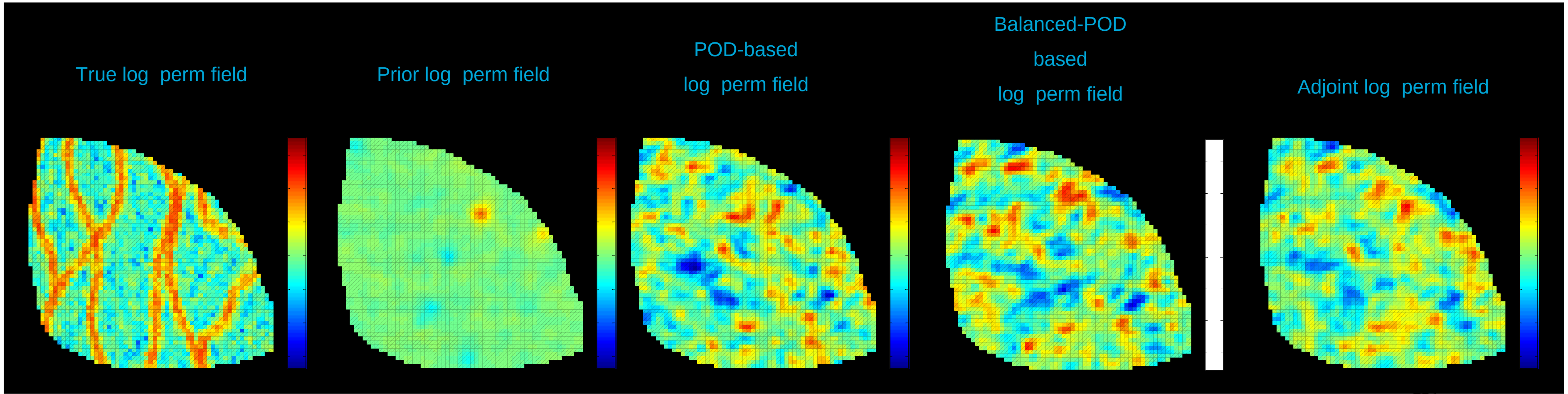
- **Producer**
- **Injector**

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

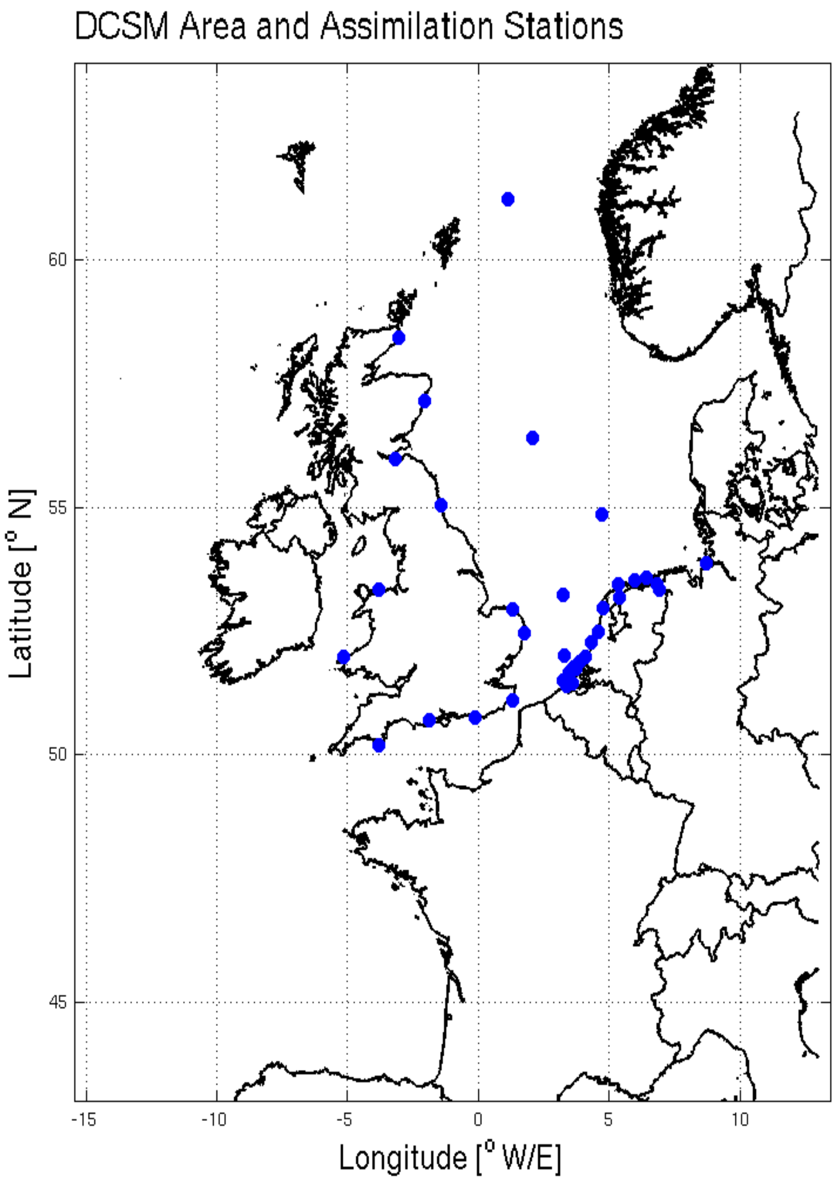
- **Measurements**
  - **Bottom hole pressures from injectors each 60 days during 3 years**
  - **Flow rates from producers each 60 days during 3 years**

Results

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

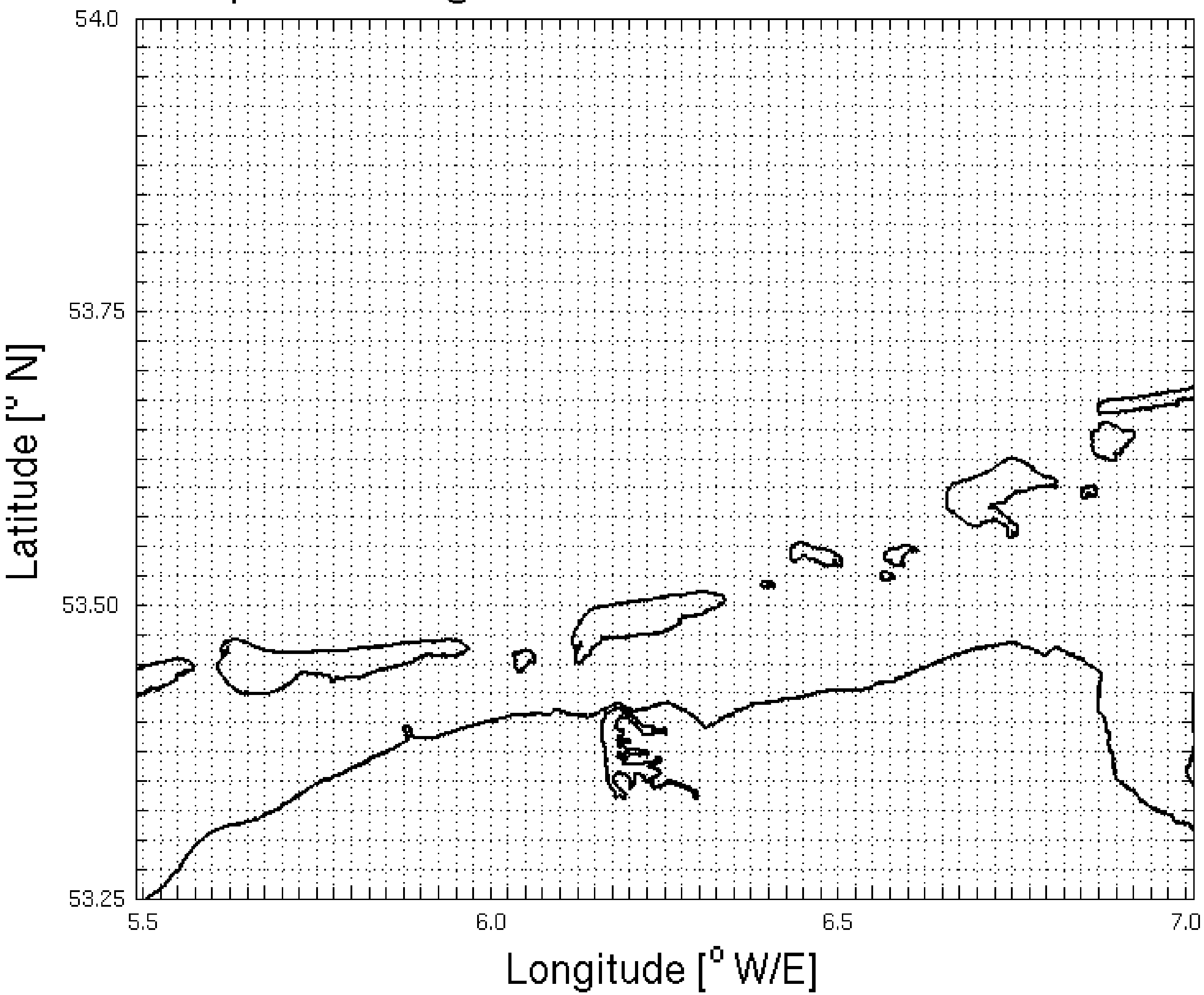


Calibration of a large scale numerical tidal model



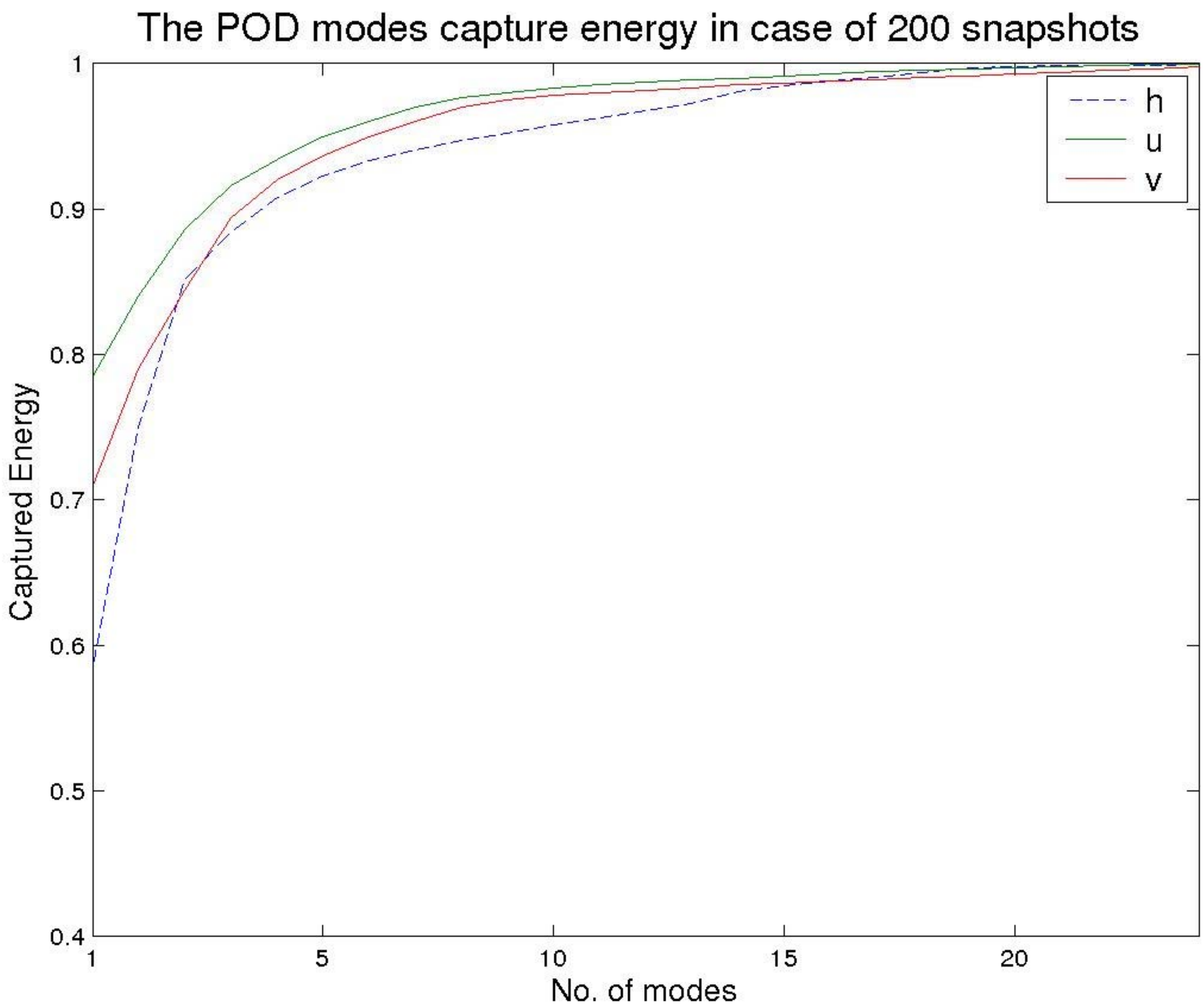
- Based on shallow water equations
- Grid size: 1.5’ by 1.0’ (~2 km)
- Grid dimensions: 1120 x 1260 cells
- Active Grid Points: 869544
- Time step: 2 minutes
- 8 main constituents

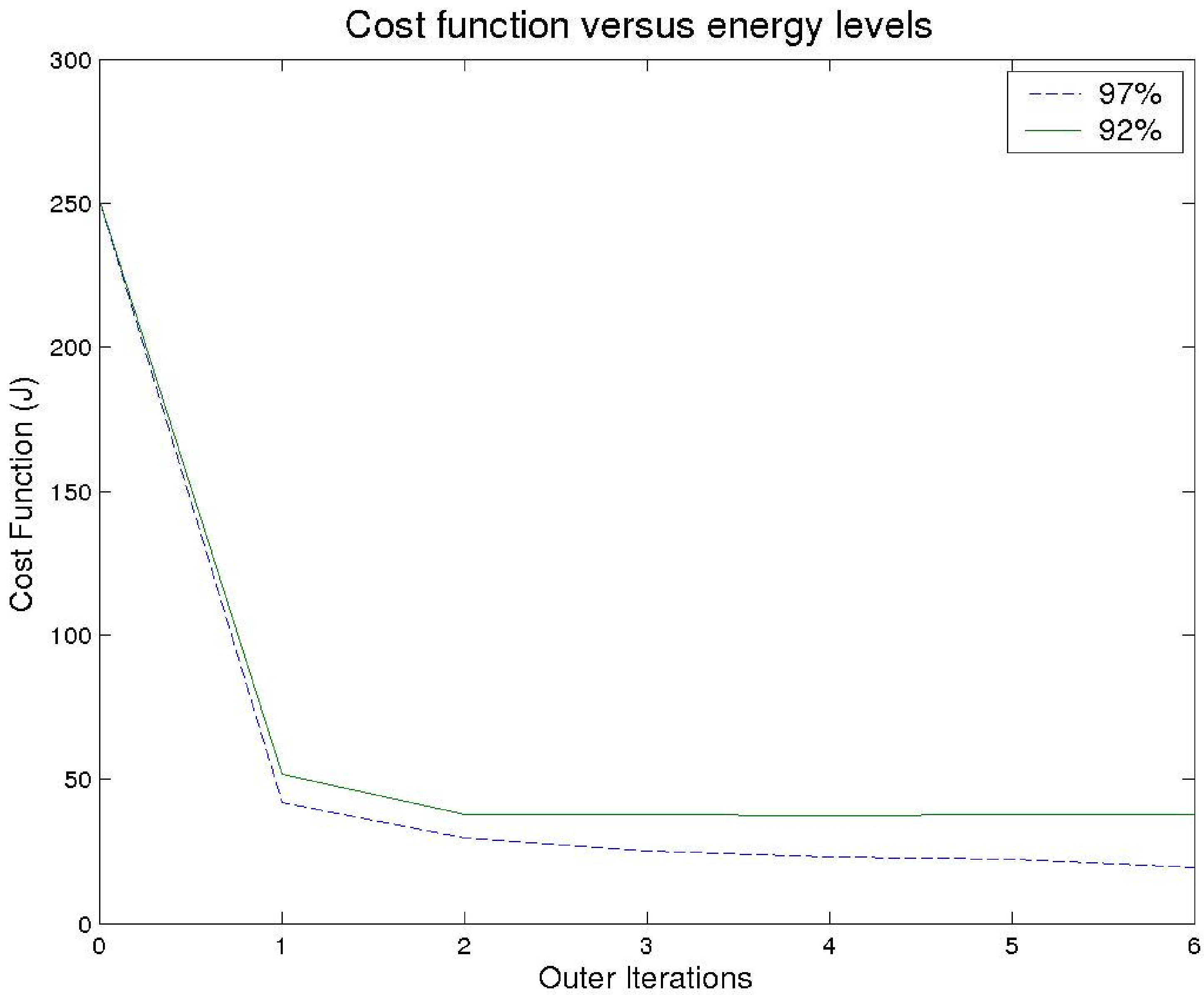
Computational grid near Dutch coast

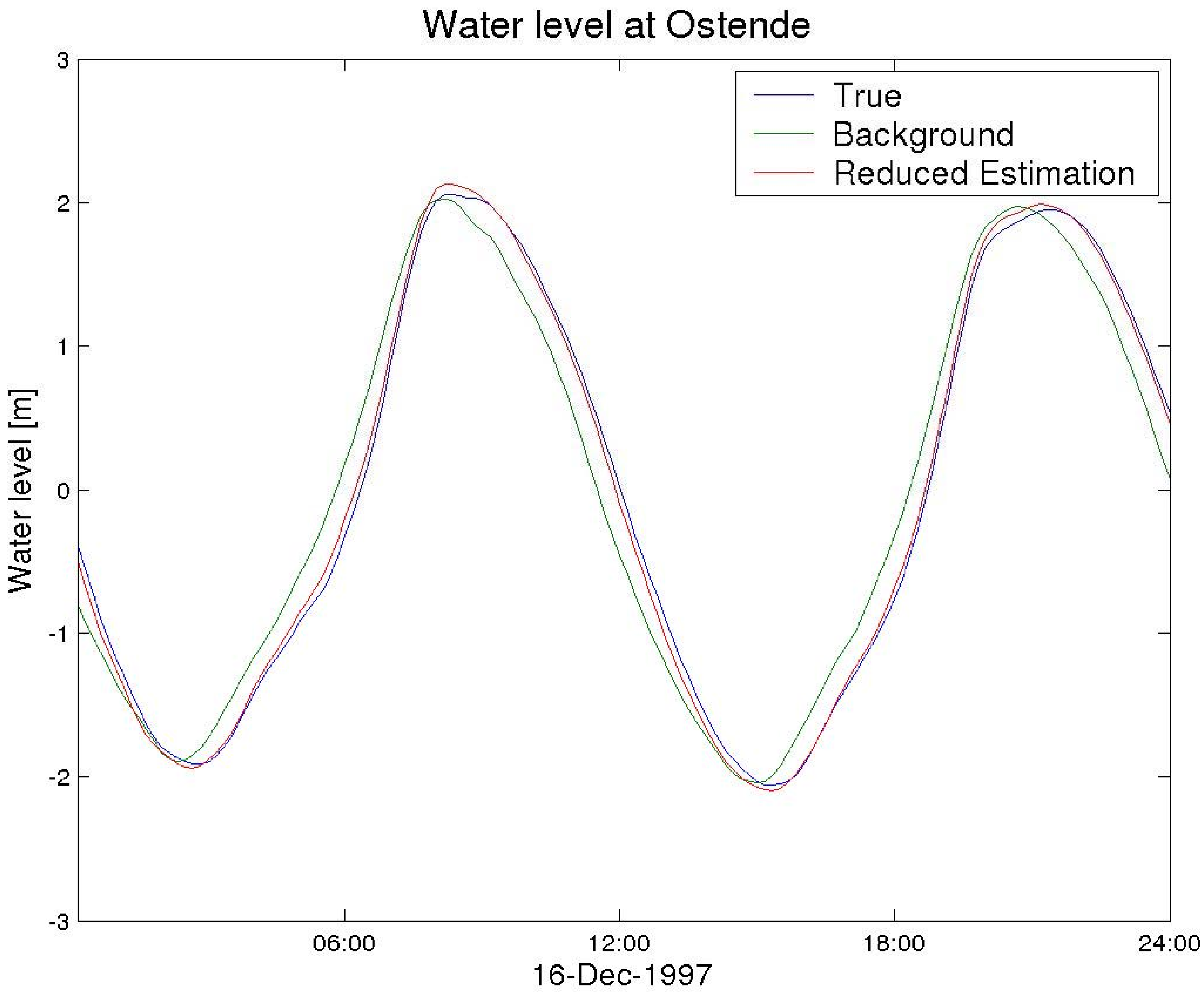


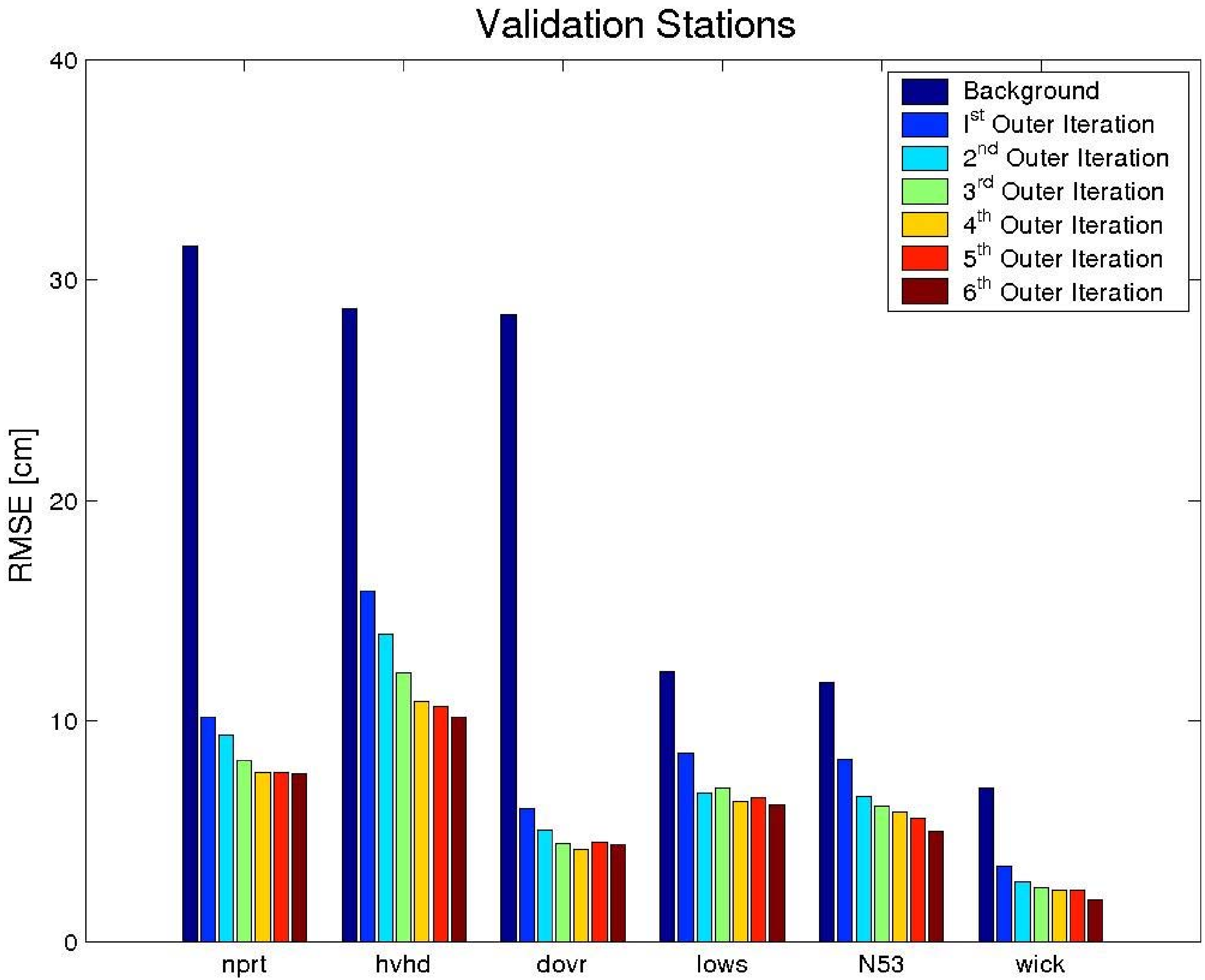
Twin experiment:

Estimation of 7 depth parameters using generated data (noise free)


The POD modes capture energy in case of 200 snapshots
Captured Energy
No. of modes
1
0.9
0.8
0.7
0.6
0.5
0.4
1
5
10
15
20
h
u
v


Cost function versus energy levels
97%
92%
Cost Function (J)
Outer Iterations
300
250
200
150
100
50
0
0
1
2
3
4
5
6

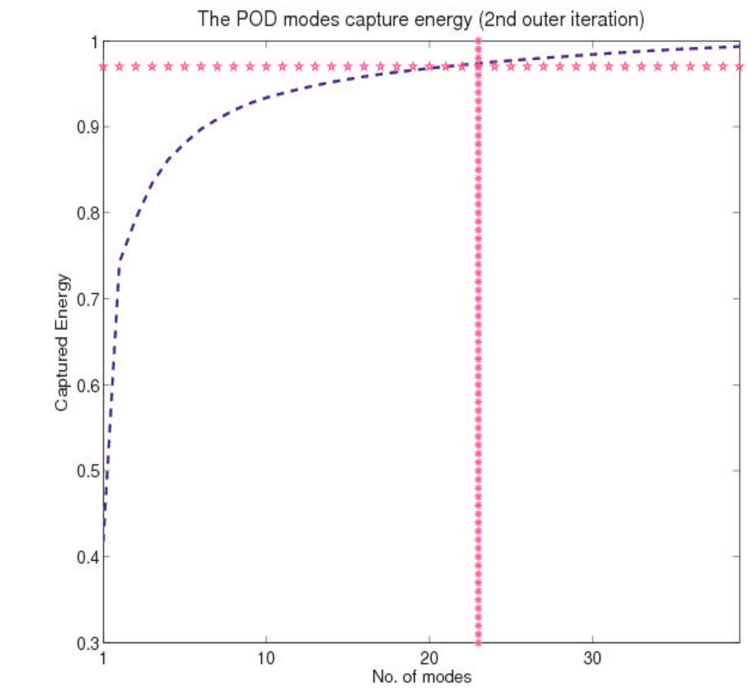

Water level at Ostende
3
2
1
0
-1
-2
-3
Water level [m]
06:00
12:00
18:00
24:00
16-Dec-1997
True
Background
Reduced Estimation


Validation Stations
RMSE [cm]
0
10
20
30
40
nprt
hvhd
dovr
lows
N53
wick
Background
Ist Outer Iteration
2nd Outer Iteration
3rd Outer Iteration
4th Outer Iteration
5th Outer Iteration
6th Outer Iteration

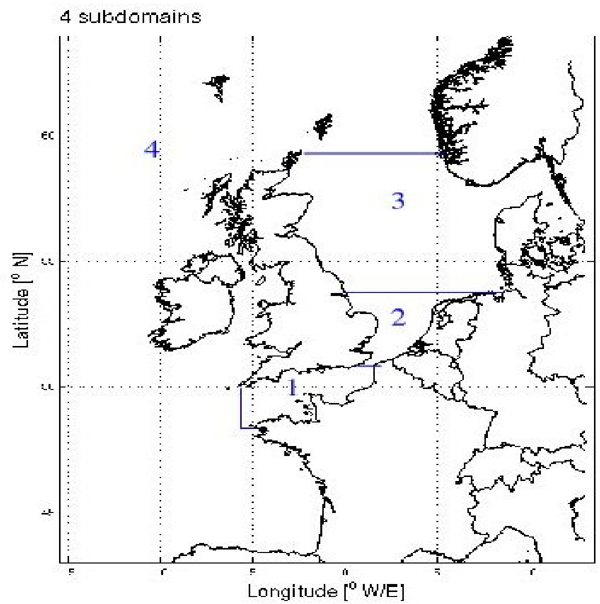
Experiment with field data

- 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



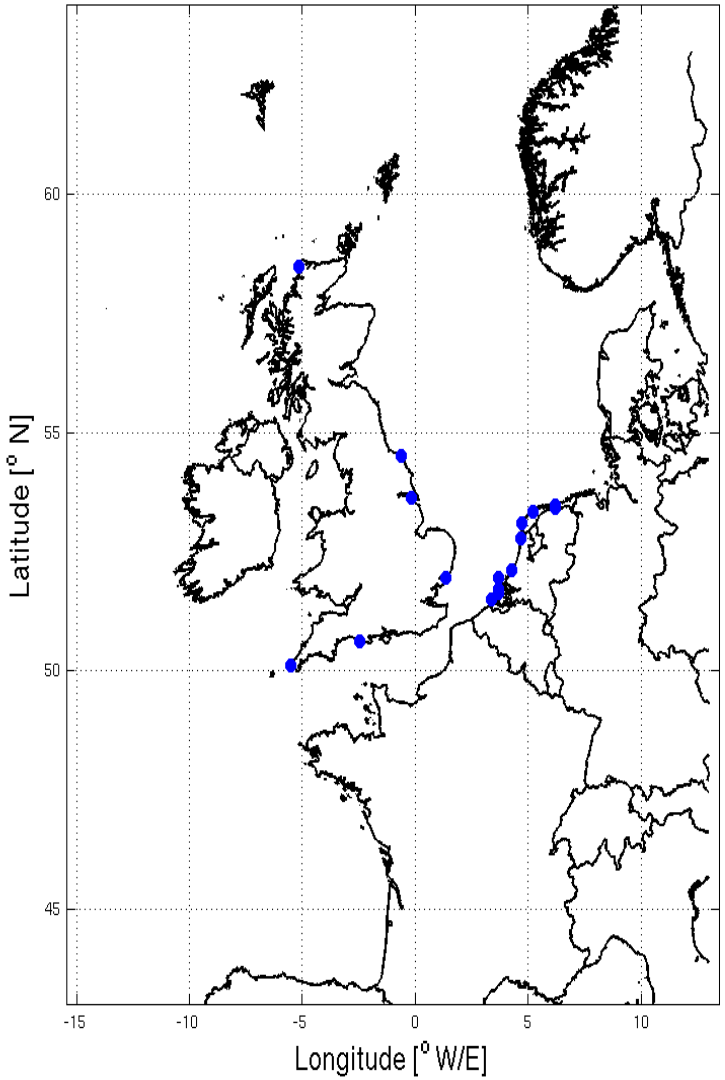
- Divide model area in 4 sub domains + 1 overall parameter
- No. of snapshots: 132 (Every three hours)
- 24 POD modes are required to capture 97% energy
- Same POD modes are used in 2rd iteration

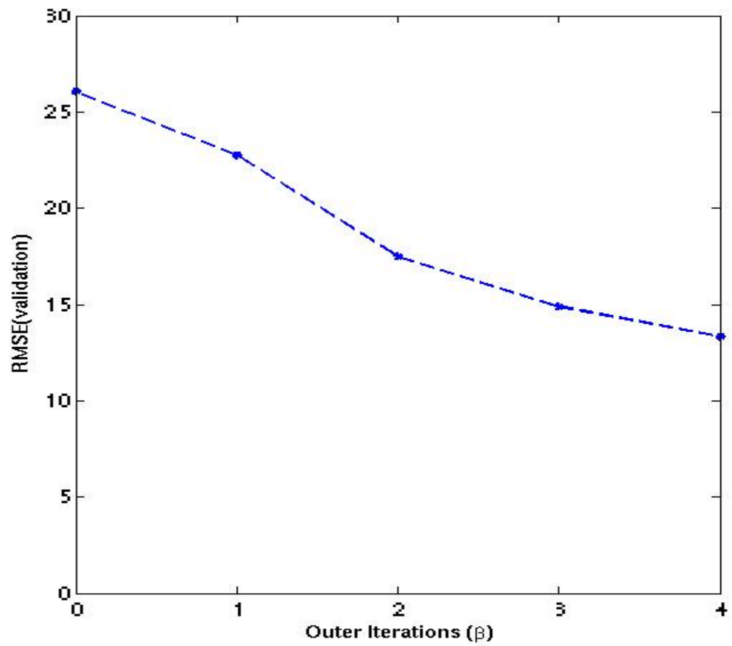


- Initial RMS: 25.7 cm
- After 2rd iteration: 14.9 cm
- Improvement : 42%

DCSM(Validation results)

DCSM Area and validation Stations





Similar improvement as in the case of assimilation stations

TU Delft

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

Efficient identification of uncertain parameters in a large-scale tidal model of the European continental shelf by proper orthogonal decomposition, Altaf, M.U. , Verlaan, M., Heemink, A.W, International Journal for Numerical Methods in Fluids, 2012.

TU Delft

Questions?

THE VALUE OF SMARTNESS

ISAPP

TU Delft