Efficient estimation of space varying parameters in numerical models using non-intrusive subdomain reduced order modeling

Xiao, Cong; Leeuwenburgh, Olwijn; Lin, Hai Xiang; Heemink, Arnold

DOI
10.1016/j.jcp.2020.109867

Publication date
2021

Document Version
Final published version

Published in
Journal of Computational Physics

Citation (APA)
Xiao, C., Leeuwenburgh, O., Lin, H. X., & Heemink, A. (2021). Efficient estimation of space varying parameters in numerical models using non-intrusive subdomain reduced order modeling. *Journal of Computational Physics*, *424*, 1-30. Article 109867. https://doi.org/10.1016/j.jcp.2020.109867

Journal of Computational Physics 424 (2021) 109867

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Efficient estimation of space varying parameters in numerical models using non-intrusive subdomain reduced order modeling

Cong Xiao [a,*], Olwijn Leeuwenburgh [b,c], Hai Xiang Lin [a], Arnold Heemink [a]

[a] Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands
[b] Civil Engineering and Geosciences, Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands
[c] TNO, Princetonlaan 6 PO Box 80015 3508 TA Utrecht, the Netherlands

ARTICLE INFO

Article history:
Received 27 November 2019
Received in revised form 7 July 2020
Accepted 21 September 2020
Available online 24 September 2020

Keywords:
Parameter estimation
Adjoint model
Model reduction
Domain decomposition

ABSTRACT

A reduced order modeling algorithm for the estimation of space varying parameter patterns in numerical models is proposed. In this approach domain decomposition is applied to construct separate approximations to the numerical model in every subdomain. We introduce a new local parameterization that decouples the computational cost of the algorithm from the number of global principal components and therefore provides attractive scaling for models with a very large number of uncertain parameter patterns. By defining uncertain parameter patterns only in the various subdomains the number of full order simulation required for the derivation of the reduced order models can be reduced drastically. To avoid non-smoothness at the boundaries of the subdomains, the optimal local parameters patterns are projected onto global parameter patterns. The computational effort of the new methodology hardly increases when the number of parameter patterns increases. The number of training models depends primarily on the maximum number of local parameters in a subdomain, which can be decreased by refining the domain decomposition. We apply the new algorithm to a large-scale reservoir model parameter estimation problem. In this application 282 parameters could be estimated using only 90 full order model runs.

Abbreviation

POD, proper orthogonal decomposition; PCA, principal component analysis; RBF, radial basis function; TPWL, trajectory piecewise linearization; DD, domain decomposition; FOM, full-order model; SLP, smooth local parameterization.

1. Introduction

We address the problem of the computationally efficient estimation of spatially varying parameters in large-scale simulation models of porous media flow. Earth science domains in which these problems arise include hydrocarbon and geothermal reservoir engineering, hydrology and geophysics. For example, in reservoir engineering, where the problem

* Corresponding author.
E-mail address: c.xiao@tudelft.nl (C. Xiao).

https://doi.org/10.1016/j.jcp.2020.109867
0021-9991/

is known as history matching, simulation models typically contain $10^5 - 10^6$ spatially related but largely unknown model parameters, available data are sparse and often indirect, and relationships between parameters and data tend to be highly nonlinear. The unknown parameters include, amongst others, the properties of discrete volumes of porous rock at depths up to a few kilometers, and functional models of interaction between fluids that are present in the pores of that rock. In the context of multi-phase flow in subsurface porous media, like oil-water flow in an oil reservoir model, the unknown parameters contain static geological parameters and dynamic parameters representing the fluid interactions. The geological permeability, porosity, hydraulic conductivity, fault multiplier and net thickness gross (NTG) belong to the static parameters, while the relative permeability coefficients and capillary pressure belong to dynamic parameters. Data are typically obtained at wells that are drilled from the surface and are used to either produce fluids from the reservoir or inject fluids in the reservoir in order to displace the fluids present in it. These wells are normally sparsely distributed over large areas, leaving the largely heterogeneous reservoir rock in-between the wells unsampled. Alternative ways of gathering information are based on geophysical techniques, such as reflection seismic, which register the arrival time and amplitude of acoustic waves that are reflected at contrasts in acoustic impedance, which in turn is affected by the density and mechanical properties of the rock and fluids. Such geophysical information is often of low spatial resolution (especially in the vertical, or depth, direction). All these aspects make the reservoir history matching a challenging problem.

The methods that have been proposed in the past to address this problem share many similarities with those employed to solve state and parameter estimation problems that appear, for example, in oceanography and meteorology. It has been widely acknowledged that the adjoint method is one of the most efficient approaches to solve such large-scale parameter estimation problems [1–4]. This approach exploits the so-called adjoint model to calculate the gradient of the model-data mismatch with respect to the parameters. Unfortunately it is an intrusive method that requires access to the source code of the simulation model and often also requires a significant implementation and maintenance effort. Some studies have looked at ways to make the implementation of the adjoint method more feasible. Courtier et al. [5] proposed an incremental approach that replaces the high resolution nonlinear model with an approximate linear model whose adjoint can be obtained more easily. Liu et al. [6,7] developed an ensemble-based four-dimensional variational (En4DVar) data assimilation scheme where the approximated linear model is constructed using an ensemble of model simulations.

Reduced order model approaches have received attention as a way of reducing the computational effort of model-based workflows through dimensionality reduction. The main idea is to construct a (linear) low-order surrogate model by projecting the dynamics of the system onto the subspace of dominant variability of the model dynamics. Most model order reduction (MOR) strategies employ POD of time series of model state 'snapshots' to identify an orthogonal set of basis functions of the subspace. Such MOR strategies have been applied with success in speeding up model simulations in computational fluid mechanics [8–10], subsurface flow simulations [11–13], air pollution [14] and aerospace design [15].

The combination of model linearization and dimensionality reduction also provides possibilities to ease the implementation of the adjoint method for high-dimensional non-linear systems. Cardoso et al. [16] were the first to integrate POD and trajectory-piecewise-linearization ([17]) and applied this strategy to oil production optimization. Subsequently, He et al. [18,19] also applied POD-TPWL to reservoir history matching. These studies suggested that POD-TPWL has the potential to significantly reduce the computational cost associated with solving subsurface flow problems [20], especially by reducing the overall full-order model simulation runtime. However, POD-TPWL in its original form requires access to derivative matrices used internally by the numerical solver of the simulation model, and therefore cannot be used with, for example, most commercial simulators [21].

A non-intrusive POD-based method to build reduced-order linear approximations of high-dimensional non-linear subsurface flow models was proposed by Vermeulen and Heemink [22]. The adjoint of this reduced-order linear model can be easily constructed and the minimization of the model-data mismatch can therefore be handled efficiently. Altaf et al. [23] and Kaleta et al. [24] applied this method to a coastal engineering and reservoir history matching problem, respectively. This algorithm considers any simulator as black box and is completely non-intrusive with respect to the simulation code. However, the model linearization is realized through a perturbation-based finite difference method, which requires a significant number of full-order model simulations and is therefore computationally less attractive for large scale problems with many uncertain parameters.

Domain decomposition (DD) as a strategy for constructing reduced-order models (ROM) is not new and has been extensively investigated in various fields. Lucia et al. [25,26] introduced the subdomain idea into ROM for tracking a moving shock wave, whilst other applications of the subdomain approach into ROM include the work of Baiges et al. [27], Kerfriden et al. [28], Amsallem et al. [29] and Chinchapatnam et al. [30]. Antil et al. [31,32] also used a subdomain idea to construct balanced truncation based ROMs. Xiao et al. [9,33] developed a non-intrusive procedure that splits the computational domain into spatial subdomains and calculates a local reduced order approximation to the governing equations within each subdomain. These reduced-order model approaches aim primarily at approximating the dynamic response of the full model as accurately as possible. Our objective here, on the other hand, is to effectively and efficiently estimate uncertain full order model parameters through approximation of the adjoint of the original high-dimensional non-linear model.

Xiao et al. [34] proposed a non-intrusive approach, e.g., subdomain POD-TPWL, that combines dimensionality reduction, piece-wise linearization and domain decomposition, and demonstrated this approach on an example from the field of subsurface reservoir engineering. Compared to the POD-TPWL approach proposed in [16], this subdomain POD-TPWL has two advantages: (1) Instead of taking global basis functions to define the low-order subspace, the snapshots of dynamic states are first partitioned according to the domain decomposition strategy and then local basis functions are obtained from these

partitioned snapshots. (2) The derivative matrices required in the process of model linearization are separately approximated in each subdomain using a RBF interpolation method [35], such that access to the underlying model code is not required. A drawback of that approach was that only the dimension of the model state space was reduced in 'local' subdomains, whereas the number of parameters was still determined by a global PCA decomposition of the parameter space over the entire spatial domain. Based on numerical experiments, it was found that the total number of full-order model simulations required for subdomain POD-TPWL is roughly 2-4 times the number of global parameters. This computational cost is therefore still too high for large-scale parameter estimation problems with a large number of uncertain parameters.

In this paper we propose a new method that incorporates a local parameter decomposition into the subdomain POD-TPWL approach. We will show that the resulting reduction in the dimensionality of the local PCA parameterizations is much larger than that obtained through global PCA parameterization. As a result a much smaller number of full-order model simulations will need to be run to solve the parameter estimation problem. A consequence of the decomposition of the spatial parameter field into spatial non-overlapping subdomains is the potential loss of smoothness. When implementing parameterization techniques such as PCA [36], kernel PCA [37], discrete cosine transform (DCT) [38] and discrete wavelet transform (DWT) [39] in the global domain, it is guaranteed that the reconstructed spatial parameter field is smooth. However, applying these techniques in each subdomain separately to obtain local parametrizations will introduce discontinuities at the boundaries among neighboring subdomains. We therefore introduce an approach that resolves the non-smoothness at the boundaries of neighboring subdomains by projection of the estimated local PCA patterns onto global PCA patterns. This smooth local parameterization simultaneously inherits the advantages of global PCA (e.g., smoothness and differentiability) and of local PCA (computational efficiency). We combine this parameterization method with subdomain POD-TPWL to drastically speed up the solution of parameter estimation problems with many uncertain parameters. The methodology is assessed in detail by analyzing its performance in example applications to petroleum reservoir history matching problems.

The remainder of the paper is organized as follows. The formulation of the inverse modeling is presented in Section 2. A brief description of subdomain POD-TPWL with global parameterization and its limitation are introduced in Section 3. Section 4 describes the smooth local parameterization method to represent parameter patterns in each subdomain individually while preserving smoothness. The procedure of subdomain POD-TPWL combined with the smooth local parameterization is presented in Section 5. Section 6 describes the basic settings of two case-studies using a 2D reservoir simulation model. Section 7 analyzes in detail the results of numerical experiments. Finally, Section 8 summarizes the results and discusses future work.

2. Formulation of the inverse problem

In the following a brief overview of the mathematical formulation of the parameter estimation problem is provided, along the lines of the description in [34]. The dynamic equation for a porous medium fluid flow system can be described by

$$\mathbf{x}^n = \mathbf{f}^n(\mathbf{x}^{n-1}, \boldsymbol{\beta}), \quad n = 1, \cdots, N_t \tag{1}$$

where the dynamic operator $\mathbf{f}^n$: $R^{N_x} \rightarrow R^{N_x}$ represents the nonlinear time-dependent model evolution from discrete time step $n-1$ to time step n, $\mathbf{x}^n \in R^{N_x}$ represents the dynamic state vector, and $\boldsymbol{\beta} \in R^{N_\beta}$ is a vector containing the model parameters. In the following we will consider an oil reservoir two-phase flow with $\mathbf{x}$ consisting of pressure and water saturation in all N_g gridblocks (that is to say $N_x = 2N_g$) and $\boldsymbol{\beta}$ representing a vector of the geological permeability in all gridblocks (e.g., $N_\beta = N_g$). N_t denotes the total number of simulation time steps.

The relationship between the measured data $\mathbf{y}^m \in R^{N_y}$ and the model state can be described by introducing a measurement operator $\mathbf{h}^m$: $R^{N_x} \rightarrow R^{N_y}$ as follows

$$\mathbf{y}^m = \mathbf{h}^m(\mathbf{x}^m, \boldsymbol{\beta}) + \mathbf{r}^m, \quad m = 1, \cdots, N_{obs} \tag{2}$$

where N_{obs} is the number of time steps at which the measurements are taken (a subset of the N_t simulation time steps), and N_y is the number of measurements at each time step. $\mathbf{r}^m$ denotes a vector of measurement errors for the data gathered at time step m. These errors are generally assumed to satisfy a Gaussian distribution $G(\mathbf{0}, \mathbf{R}^m)$ where $\mathbf{R}^m$ is the measurement error covariance matrix. In our case study we consider two types of measured data, namely well data (fluid rate and pressure measured at the well locations only), and water saturation data (assumed to be observed in all gridblocks). Water saturation data are used to mimic time-lapse seismic data from which water saturation can be extracted through seismic inversion [40].

If the prior probability density function (PDF) of the parameters $\boldsymbol{\beta}$ is Gaussian with mean $\boldsymbol{\beta}_b$ and covariance matrix $\mathbf{C}$, maximization of the posterior probability density conditioned on the measurements $\mathbf{d}_{obs}$ is equivalent to minimizing the objective function defined as

$$J_{FOM}(\boldsymbol{\beta}) = \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}_b)^T \mathbf{C}^{-1}(\boldsymbol{\beta} - \boldsymbol{\beta}_b) + \frac{1}{2}\sum_{m=1}^{N_{obs}} [\mathbf{d}_{obs}^m - \mathbf{h}^m(\mathbf{x}^m, \boldsymbol{\beta})]^T \mathbf{R}^{m-1} [\mathbf{d}_{obs}^m - \mathbf{h}^m(\mathbf{x}^m, \boldsymbol{\beta})] \tag{3}$$

If the derivatives of $\mathbf{f}^n$ and $\mathbf{h}^m$ with respect to $\mathbf{x}^n$ and $\boldsymbol{\beta}_b$ are available for $n = 1, \cdots, m$, the adjoint method can be employed to obtain the gradient of this objective function [1–3], which can subsequently be used in a gradient-based algorithm to solve the minimization problem.

3. Description of subdomain POD-TPWL

3.1. POD-TPWL

Here we will briefly describe the POD-TPWL method [16,41]. First, one or more training samples of the unknown parameters $\boldsymbol{\beta}$ are generated and subsequently used as input for generating an ensemble of model simulations that solve Eq. (1), resulting in a time series $\mathbf{x}_{tr}^n$ with $n = 1, \cdots, N_t$. The TPWL surrogate model approximates the state $\mathbf{x}^n$, for a given the state $\mathbf{x}^{n-1}$ and parameters $\boldsymbol{\beta}$, as the first-order expansion around the training solution $(\mathbf{x}_{tr}^n, \mathbf{x}_{tr}^{n-1}, \boldsymbol{\beta}_{tr})$ as follows

$$\mathbf{x}^n \approx \mathbf{x}_{tr}^n + \mathbf{E}_{\mathbf{x}_{tr}}^n (\mathbf{x}^{n-1} - \mathbf{x}_{tr}^{n-1}) + \mathbf{G}_{\boldsymbol{\beta}_{tr}}^n (\boldsymbol{\beta} - \boldsymbol{\beta}_{tr}) \tag{4}$$

with

$$\mathbf{E}_{\mathbf{x}_{tr}}^n = \frac{\partial \mathbf{f}^n}{\partial \mathbf{x}_{tr}^{n-1}}, \quad \mathbf{G}_{\boldsymbol{\beta}_{tr}}^n = \frac{\partial \mathbf{f}^n}{\partial \boldsymbol{\beta}_{tr}} \tag{5}$$

If there is more than one training solution, $(\mathbf{x}_{tr}^n, \mathbf{x}_{tr}^{n-1}, \boldsymbol{\beta}_{tr})$ is chosen to be the solution that is 'closest' to the state vector $\mathbf{x}^{n-1}$. A criterion for closeness was suggested by He et al. [41].

The matrices $\mathbf{E}_{\mathbf{x}_{tr}}^n \in R^{N_x \times N_x}$ and $\mathbf{G}_{\boldsymbol{\beta}_{tr}}^n \in R^{N_x \times N_\beta}$ represent the derivatives of the dynamic model given by Eq. (1) at time step n with respect to state vector $\mathbf{x}_{tr}^{n-1}$ and parameters $\boldsymbol{\beta}_{tr}$, respectively. Throughout this paper, the variables with subscript tr represent the training model we linearize around. Eq. (4) is, however, still in high-dimensional space, $\mathbf{x}^n \in R^{2N_g}$ and $\boldsymbol{\beta} \in R^{N_g}$, and is computationally very expensive to compute. This has motivated the development of the POD-TPWL algorithm.

Proper Orthogonal Decomposition is used to project high-dimensional state vectors onto an optimal lower-dimensional subspace. The basis of this subspace is defined by the left singular vectors obtained by performing a Singular Value Decomposition (SVD) of a snapshot matrix containing states at selected time steps (snapshots) computed from training simulations. The state vector $\mathbf{x}$ (the superscript indicating the time step is omitted to simplify the notation) can then be represented in terms of the product of a matrix of basis vectors $\boldsymbol{\Phi}$ and a coefficient vector $\boldsymbol{\psi}$

$$\mathbf{x} = \boldsymbol{\Phi}\boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\Phi}_p & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi}_s \end{bmatrix} \boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\Phi}_p & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Phi}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi}_p \\ \boldsymbol{\psi}_s \end{bmatrix} \tag{6}$$

where $\boldsymbol{\Phi}_p$ and $\boldsymbol{\Phi}_s$ represent matrices of basis vectors for pressure and saturation respectively. In general there is no need to retain all columns $\boldsymbol{\Phi}_p$ and $\boldsymbol{\Phi}_s$ and reduced state representations can be obtained by selecting only the first columns according to an energy criterion [41].

The same procedure can be used to project the high-dimensional parameter space onto an optimal lower-dimensional subspace [42]. Any Gaussian distributed parameter vector $\boldsymbol{\beta}$ can be represented by a vector of Gaussian random coefficients $\boldsymbol{\xi}$ with zero mean and unit variance according to

$$\boldsymbol{\beta} = \boldsymbol{\beta}_m + \boldsymbol{\Phi}_\beta \boldsymbol{\xi} \tag{7}$$

where $\boldsymbol{\beta}_m$ represents a background (mean) parameter vector and $\boldsymbol{\Phi}_\beta$ denotes a matrix of N_G parameter basis vectors. Analogous to the POD procedure sketched above, the basis vectors can be computed from large set realizations of the random parameter vector [43]. In the following we will refer to this procedure for the parameter space reduction as principal component analysis (PCA).

Substituting Eq. (6) and Eq. (7) into Eq. (4), we obtain the following reduced-order linear model

$$\boldsymbol{\psi}^n \approx \boldsymbol{\psi}_{tr}^n + \mathbf{E}_{\boldsymbol{\psi}_{tr}}^n (\boldsymbol{\psi}^{n-1} - \boldsymbol{\psi}_{tr}^{n-1}) + \mathbf{G}_{\boldsymbol{\xi}_{tr}}^n (\boldsymbol{\xi} - \boldsymbol{\xi}_{tr}) \tag{8}$$

with

$$\mathbf{E}_{\boldsymbol{\psi}_{tr}}^n = \boldsymbol{\Phi}^T \frac{\partial \mathbf{f}^n}{\partial \mathbf{x}_{tr}^{n-1}} \boldsymbol{\Phi} = \boldsymbol{\Phi}^T \mathbf{E}_{\mathbf{x}_{tr}}^n \boldsymbol{\Phi}, \quad \mathbf{G}_{\boldsymbol{\xi}_{tr}}^n = \boldsymbol{\Phi}^T \frac{\partial \mathbf{f}^n}{\partial \boldsymbol{\beta}_{tr}} \boldsymbol{\Phi}_\beta = \boldsymbol{\Phi}^T \mathbf{G}_{\boldsymbol{\beta}_{tr}}^n \boldsymbol{\Phi}_\beta \tag{9}$$

Similarly, the measurement operator (Eq. (2)) is also linearized around the training solution $(\boldsymbol{\psi}_{tr}^m, \boldsymbol{\xi}_{tr})$ as follows

$$\mathbf{y}^m \approx \mathbf{y}_{tr}^m + \mathbf{A}_{\boldsymbol{\psi}_{tr}}^m (\boldsymbol{\psi}^m - \boldsymbol{\psi}_{tr}^m) + \mathbf{B}_{\boldsymbol{\xi}_{tr}}^m (\boldsymbol{\xi} - \boldsymbol{\xi}_{tr}) \tag{10}$$

with

$$\mathbf{A}_{\boldsymbol{\psi}_{tr}}^m = \frac{\partial \mathbf{h}^m}{\partial \mathbf{x}_{tr}^m} \boldsymbol{\Phi}, \quad \mathbf{B}_{\boldsymbol{\xi}_{tr}}^m = \frac{\partial \mathbf{h}^m}{\partial \boldsymbol{\beta}_{tr}} \boldsymbol{\Phi}_\beta \tag{11}$$

Eqs. (8)-(9) and (10)-(11) represent the standard POD-TPWL systems for a dynamic model and a measurement operator, respectively. $\mathbf{E}_{\boldsymbol{\psi}_{tr}}^n$, $\mathbf{G}_{\boldsymbol{\xi}_{tr}}^n$, $\mathbf{A}_{\boldsymbol{\psi}_{tr}}^m$ and $\mathbf{B}_{\boldsymbol{\xi}_{tr}}^m$ are used internally by the numerical solver and are not easily to compute. Unlike

the model state operator in Eq. (9) where the state variables $\mathbf{x}^n$ are reduced using the projection basis matrix ϕ, the measurements $\mathbf{y}^m$ (e.g., fluid rate and bottom-hole pressure) are generally different from the state variables, e.g., saturation and pressure, and therefore do not need the projection by the basis matrix ϕ. As a result, the definitions of $\mathbf{A}^m_{\psi_{tr}}$ and $\mathbf{B}^m_{\xi_{tr}}$ do not contain the left-multiplication of an ϕ^T in Eq. (11). Since these derivatives are not generally accessible, in the following we will describe an alternative non-intrusive method that uses the adjoint of a reduced surrogate model instead.

3.2. Domain decomposition

The POD-TPWL algorithm is based on a low-order state approximation in terms of a global POD, i.e. a basis representation obtained from the snapshot matrix containing the global (full domain) state vectors. For subsurface multi-phase flow problems has been observed that the number of retained global patterns needed to meet reasonable criteria for representation of the variability of the original state is quite high, resulting in limited computational gains. The idea behind subdomain decomposition is that local variability may be represented by much fewer patterns, potentially allowing a more significant reduction of computational cost. Before explaining in the next section why and how this reduction can be carried out, we first define the necessary notation.

The global domain Ω can be decomposed into S non-overlapping subdomains Ω^d, $d \in \{1, 2, \cdots, S\}$ with $\Omega = \bigcup_{d=1}^{S} \Omega^d$ and $\Omega^i \cap \Omega^j = 0$ for $i \neq j$ (see Fig. 1). Each subdomain Ω^d contains a total number of N_g^d gridblocks after the decomposition. The snapshots from the training simulations are now first partitioned accordance with this domain decomposition strategy and then the local basis vectors are obtained by performing a POD using the snapshots corresponding to each subdomain.

$$\mathbf{x}^d = \mathbf{\Phi}^d \boldsymbol{\psi}^d = \begin{bmatrix} \mathbf{\Phi}_p^d & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}_s^d \end{bmatrix} \boldsymbol{\psi}^d = \begin{bmatrix} \mathbf{\Phi}_p^d & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}_s^d \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi}_p^d \\ \boldsymbol{\psi}_s^d \end{bmatrix} \tag{12}$$

where $\mathbf{x}^d \in R^{N_x^d}$ represents the dynamic state vector consisting of pressure and water saturation in all N_g^d gridblocks of the subdomain Ω^d. N_x^d represents the total number of state variables in subdomain Ω^d, that is $N_x^d = 2N_g^d$. $\mathbf{\Phi}_p^d$ and $\mathbf{\Phi}_s^d$ represent matrices containing the basis vectors for pressure and saturation for subdomain Ω^d. $\mathbf{\Phi}^d$ and $\boldsymbol{\psi}^d$ are the assembled local basis matrix and local POD coefficients, respectively.

3.3. Subdomain POD-TPWL

From the previous section it can be seen that POD-TPWL requires access to the Jacobian matrices containing the derivatives of the state and measurement equations with respect to the states and model parameters. Xiao et al. [34] recently proposed a procedure to construct a RBF interpolation model that approximates the original models Eq. (1) and Eq. (2), from which the equivalent derivative matrices can be derived analytically. The computational complexity of RBF interpolation is dominated by the dimensionality of the inputs, e.g., $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$ in our study. A direct implementation of POD in the global domain (as done, e.g. in [22]) will result in the requirement to retain a large number global POD patterns, and a high computational complexity. Domain decomposition (DD) was therefore introduced to enable the efficient evaluation of independent subdomain models expressed in much smaller numbers of local POD patterns [34].

The main idea behind the subdomain POD-TPWL is to approximate the dynamical evolution of reduced states within subdomains by a reduced-order piece-wise linear model analogous to Eq. (8). Introducing an extra term to allow for dynamical interaction between each subdomain Ω^d and its neighboring subdomains Ω^{sd}, we assume that the local POD coefficients $\boldsymbol{\psi}^{d,n}$ in subdomain Ω^d can be approximated by $\hat{\boldsymbol{\psi}}^{d,n}$ as follows

$$\boldsymbol{\psi}^{d,n} \approx \hat{\boldsymbol{\psi}}^{d,n} = \boldsymbol{\psi}_{tr}^{d,n} + \mathbf{E}_{\boldsymbol{\psi}_{tr}}^{d,n}(\boldsymbol{\psi}^{d,n-1} - \boldsymbol{\psi}_{tr}^{d,n-1}) + \mathbf{E}_{\boldsymbol{\psi}_{tr}}^{sd,n}(\boldsymbol{\psi}^{sd,n} - \boldsymbol{\psi}_{tr}^{sd,n}) + \mathbf{G}_{\xi_{tr}}^{d,n}(\boldsymbol{\xi} - \boldsymbol{\xi}_{tr}) \tag{13}$$

where

$$\mathbf{E}_{\boldsymbol{\psi}_{tr}}^{d,n} = \frac{\partial \boldsymbol{\pounds}^{d,n}}{\partial \boldsymbol{\psi}_{tr}^{d,n-1}}, \quad \mathbf{E}_{\boldsymbol{\psi}_{tr}}^{sd,n} = \frac{\partial \boldsymbol{\pounds}^{d,n}}{\partial \boldsymbol{\psi}_{tr}^{sd,n}}, \quad \mathbf{G}_{\xi_{tr}}^{d,n} = \frac{\partial \boldsymbol{\pounds}^{d,n}}{\partial \boldsymbol{\xi}_{tr}} \tag{14}$$

Analogously, the measurements $\mathbf{y}^{d,m}$ taken in subdomain Ω^d are approximated by $\hat{\mathbf{y}}^{d,m}$ as follows

$$\mathbf{y}^{d,m} \approx \hat{\mathbf{y}}^{d,m} = \mathbf{y}_{tr}^{d,m} + \mathbf{A}_{\boldsymbol{\psi}_{tr}}^{d,m}(\boldsymbol{\psi}^{d,m} - \boldsymbol{\psi}_{tr}^{d,m}) + \mathbf{B}_{\xi_{tr}}^{d,m}(\boldsymbol{\xi} - \boldsymbol{\xi}_{tr}) \tag{15}$$

with

$$\mathbf{A}_{\boldsymbol{\psi}_{tr}}^{d,m} = \frac{\partial \boldsymbol{\hbar}^{d,m}}{\partial \boldsymbol{\psi}_{tr}^{d,m}}, \quad \mathbf{B}_{\xi_{tr}}^{d,m} = \frac{\partial \boldsymbol{\hbar}^{d,m}}{\partial \boldsymbol{\xi}_{tr}} \tag{16}$$

Above we have introduced new, so far undefined, functions $\boldsymbol{\pounds}^{d,n}(\boldsymbol{\psi}^{d,n-1}, \boldsymbol{\psi}^{sd,n}, \boldsymbol{\xi})$ and $\boldsymbol{\hbar}^{d,m}(\boldsymbol{\psi}^{d,m}, \boldsymbol{\xi})$ of the local reduced states $\boldsymbol{\psi}^{d,n}$ and the measurements $\mathbf{y}^{d,m}$ and of the global reduced parameters $\boldsymbol{\xi}$. By analogy to Eqs. (8)-(9) and Eqs. (10)-(11)

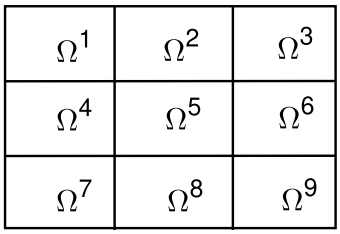


Fig. 1. Illustration of domain decomposition in a 2D case. The entire domain is divided into nine rectangle subdomains.

these functions represent the reduced nonlinear dynamic model and reduced nonlinear measurement operator, respectively. Note that only the derivatives of these models appear in our proposed subdomain formulation and that we are free to choose the form of these functions. We would like to be able to evaluate their derivatives (Eq. (14)) efficiently and that, when used in Eq. (13), they deliver the best possible approximation $\hat{\boldsymbol{\psi}}^{d,n}$ of the full states $\boldsymbol{\psi}^{d,n}$.

Xiao et al. proposed to define $\boldsymbol{\pounds}^{d,n}$ and $\boldsymbol{\hbar}^{d,m}$ in terms of a RBF interpolation model [34]. The RBF interpolation model can be represented as a linear combination of M radial basis functions θ by a vector of weighting coefficients $\boldsymbol{\omega}^{d,n} = [\omega_1, ..., \omega_j, ..., \omega_M]$ (where, $j = 1, 2, ..., M$) at time step n for subdomain Ω^d

$$\boldsymbol{\pounds}^{d,n}(\boldsymbol{\psi}^{d,n-1}, \boldsymbol{\psi}^{sd,n}, \boldsymbol{\xi}) = \sum_{j=1}^{M} \boldsymbol{\omega}_j^{d,n} \theta(||(\boldsymbol{\psi}^{d,n-1}, \boldsymbol{\psi}^{sd,n}, \boldsymbol{\xi}) - (\boldsymbol{\psi}_j^{d,n-1}, \boldsymbol{\psi}_j^{sd,n}, \boldsymbol{\xi}_j)||) \tag{17}$$

and similarly,

$$\boldsymbol{\hbar}^{d,m}(\boldsymbol{\psi}^{d,m}, \boldsymbol{\xi}) = \sum_{j=1}^{M} \boldsymbol{\varepsilon}_j^{d,m} \theta(\|(\boldsymbol{\psi}^{d,m}, \boldsymbol{\xi}) - (\boldsymbol{\psi}_j^{d,m}, \boldsymbol{\xi}_j)\|) \tag{18}$$

respectively, the subscripted index j identifies one of M training runs. The choice of M significantly depends on the dimension of inputs, e.g., $\boldsymbol{\psi}_j^d$ and $\boldsymbol{\xi}_j$. A discussion of how to choose M was provided by Xiao et al. [34]. Subdomain POD-TPWL uses analytical RBFs $\boldsymbol{\pounds}^{d,n}$ and $\boldsymbol{\hbar}^{d,m}$ to approximate the time series of full model state snapshots, which also allows us to derive the derivative matrices $\mathbf{E}_{\boldsymbol{\psi}_{tr}}^{d,n}$, $\mathbf{E}_{\boldsymbol{\psi}_{tr}}^{sd,n}$, $\mathbf{G}_{\boldsymbol{\xi}_{tr}}^{d,n}$, $\mathbf{A}_{\boldsymbol{\psi}_{tr}}^{d,m}$, and $\mathbf{B}_{\boldsymbol{\xi}_{tr}}^{d,m}$ analytically. More information on the specific formula of the radial basis function θ and the determination of the weighting coefficients $\boldsymbol{\omega}$ can be found in [34].

It can be seen from Eq. (13) and Eq. (15) that the subdomain POD-TPWL system is defined in terms of local POD patterns of the model dynamics, and global patterns representing the uncertain parameter space. Numerical experiments in [34] have shown that the total number of full-order model simulations required to construct the subdomain POD-TPWL model is approximately 2-4 times the (limited) number of global parameter patterns, and the number will rapidly increase as the dimensionality of global parameter patterns becomes large. Decreasing the number of preserved global parameter patterns will reduce the computational cost, but will inevitably result in a loss of accuracy. The dependence of the computational cost on the dimension of the global reduced parameter space will inevitably restrict the application of subdomain POD-TPWL to inverse problems with a limited number of uncertain parameters or with parameters fields containing smooth and large-scale spatial correlations.

From a computational point of view, implementing subdomain POD-TPWL in terms of local parameter patterns in each subdomain, similar to the treatment of the dynamic states, therefore seems attractive. Firstly, the number of local patterns in each individual subdomain will be smaller than the number of global patterns, this description might be almost valid in some practical problems, secondly, the computational cost of the parameter estimation procedure is proportional to the typical number of local patterns in any one subdomain, this statement has been partially verified in our previous work [34] where the global parameter patterns are optimized. Such an approach we aim at proposing in this work would rely on the validity of assumption that it is possible to reconstruct global parameter fields from the local parameter solutions that satisfy criteria for acceptable solutions in the global domain. In the following section we will introduce a smooth local parameter reduction method that addresses the computational limitations associated with the use of global parameter reduction and produces smooth global space parameter patterns.

4. Smooth local parameterization

In this section, we first discuss global PCA and local PCA based representations of spatial parameter patterns in the global domain and in the subdomains, respectively. The disadvantages of each are investigated to motivate the development of smooth local parameterization.

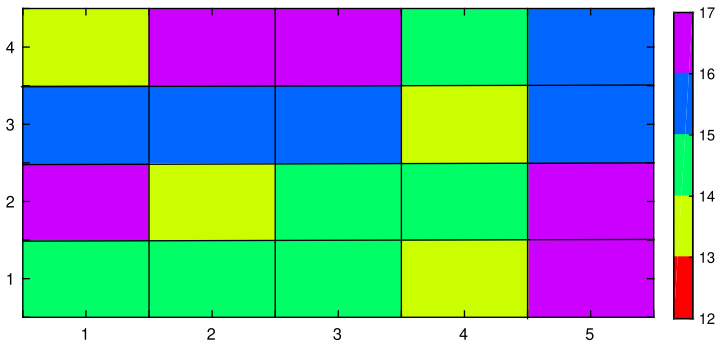


Fig. 2. The number of local PCA patterns in each subdomain for a domain decomposition consisting of 4×5 subdomains.

Table 1
Total number of local PCA and global PCA patterns, and the maximum number of local PCA patterns among all subdomains corresponding to three different domain decomposition strategies, e.g., 3×4, 4×5 and 5×6, respectively.

Domain decomposition	Smooth local parameterization		Global PCA
	$N_L = \sum_{d=1}^{S} N_l^d$	$\max\{N_l^1,...,N_l^d,...,N_l^S\}$	N_G
3 × 4	236	18	
4 × 5	275	15	48
5 × 6	312	12	

4.1. Global and local PCA-based representations of spatial parameters

In analogy to the global PCA-based parameter reduction described above, the local parameter vector $\boldsymbol{\beta}_d \in R^{N_\beta^d}$ in subdomain Ω^d can be expressed as

$$\boldsymbol{\beta}^d = \boldsymbol{\beta}_m^d + \boldsymbol{\Phi}_\beta^d \boldsymbol{\xi}^d, \quad d = 1, \cdots, S \tag{19}$$

where, $\boldsymbol{\beta}_m^d$ represents a mean parameter vector and $\boldsymbol{\Phi}_\beta^d$ is the basis matrix that projects the high-order parameter patterns onto the low-order subspace of dimension N_l^d for subdomain d. $\boldsymbol{\xi}^d$ denotes an vector of local PCA coefficients for subdomain d. The energy criterion can be employed to choose the number of basis vectors [41].

In order to illustrate the global and local PCA procedures described above to reconstruct parameter fields, we manually generate N_r = 1000 random Gaussian realizations of the model parameter vector, representing the log-permeability for all gridblocks of the 2D reservoir model that will be used in the numerical experiments presented later. Specifically, the Stanford Geostatistical Modeling Software (SGeMS) tool that we can download freely[1] is used [44]. After defining the properties of the variogram, especially the maximum and medium ranges and the azimuth, and the ranges of the search ellipsoid to large values, these 1000 correlated random Gaussian fields can be generated by running the sequential Gaussian simulation algorithm in this software.

The entire model domain is decomposed into 20 non-overlapping subdomains (4 subdomains in one direction and 5 subdomains in the second direction) as illustrated in Fig. 2. $N_G = 48$ global PCA patterns are required to construct a reduced order representation that retains 95% of the energy (variability) in the original 1000 realizations. The maximum number of local PCA patterns $\max\{N_l^1, ..., N_l^d, ..., N_l^S\}$, $d = 1, 2, ..., S$, obtained by the same procedure and the same energy criterion (but now applied in each subdomain separately) is summarized in Fig. 2. Results for other decompositions are summarized in Table 1. Decompositions with a larger number of subdomains results in fewer local PCA patterns in each subdomain. The maximum number of local patterns in any subdomain is substantially smaller than N_G.

The first two panels of Fig. 3 show two examples of parameter patterns for random realizations of global $\boldsymbol{\xi}$ and local $\boldsymbol{\xi}^d$ PCA coefficients. It can be seen that sampling arbitrary local PCA coefficients will result in a non-smooth reconstructed global parameter field, while global PCA produces a smooth parameter field in a relatively high-order subspace. The maximum number of local patterns required in any subdomain is much smaller, however, than the number of global patterns that is needed. This motivates us to develop a smooth local parameterization (SLP) that exploits the advantages of global PCA (smooth representation) and local PCA (computationally efficiency). We will demonstrate the gain in computational efficiency in the following sections.

[1] http://sgems.sourceforge.net/.

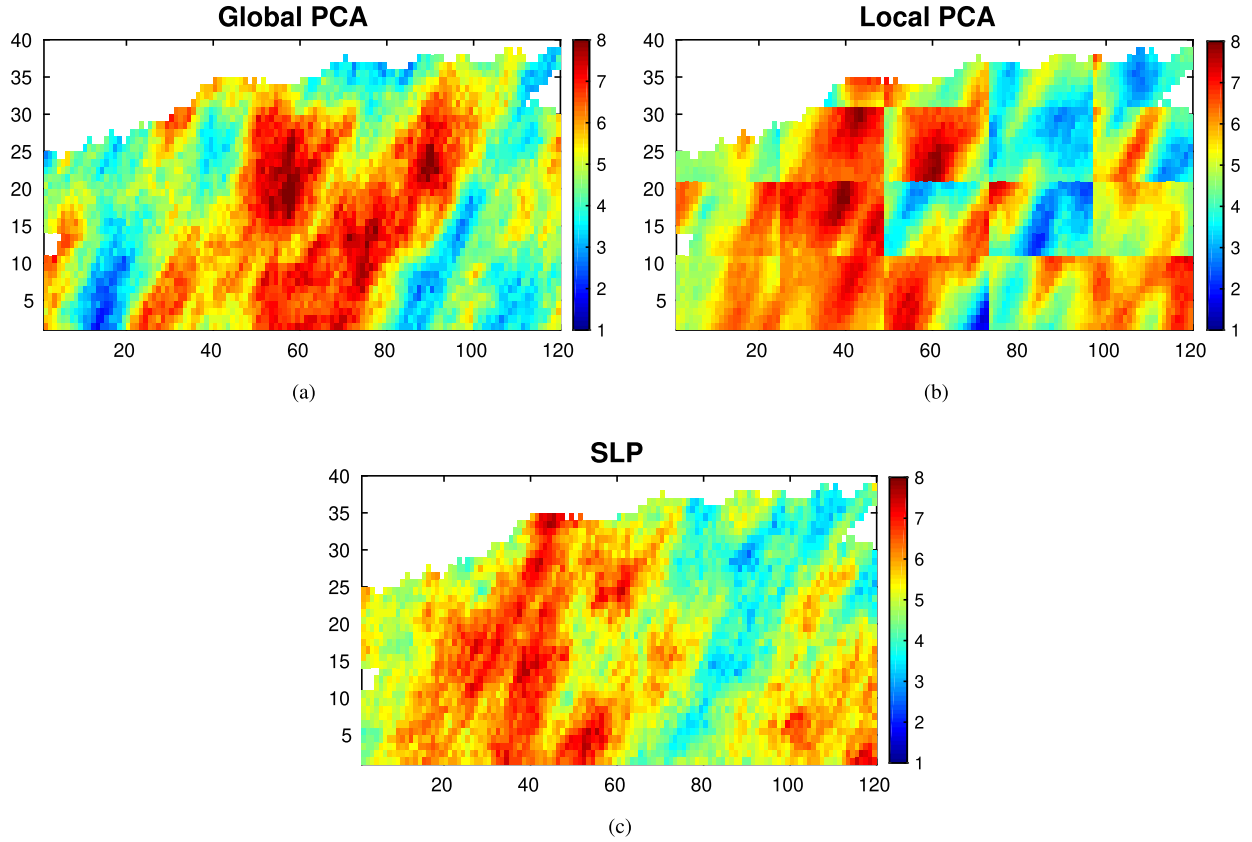


Fig. 3. Example parameter reconstructions using (a) local PCA, (b) global PCA and (c) smooth local parameterization.

4.2. PCA-based smooth local representation of spatial parameters

The results shown in Fig. 3(b) suggest that direct reconstruction of the global parameter field from local subdomain parameter solutions, and then stitching all subdomains together, is unlikely to produce acceptable results. We therefore add a post-processing step that finds the best matching global PCA reconstruction. This solution is found by minimizing an objective function consisting of the sum of squared differences between the local solution based reconstruction in terms of $\boldsymbol{\xi}^d$ and a global reconstruction in terms of global PCA coefficients $\boldsymbol{\xi}$,

$$J(\boldsymbol{\xi}) = \frac{1}{2}[\sum_{d=1}^{S} \mathbf{T}^d(\boldsymbol{\beta}_m^d + \boldsymbol{\Phi}_{\boldsymbol{\beta}}^d \boldsymbol{\xi}^d) - \boldsymbol{\beta}_m - \boldsymbol{\Phi}_{\boldsymbol{\beta}}\boldsymbol{\xi}]^T \mathbf{C}^{-1} [\sum_{d=1}^{S} \mathbf{T}^d(\boldsymbol{\beta}_m^d + \boldsymbol{\Phi}_{\boldsymbol{\beta}}^d \boldsymbol{\xi}^d) - \boldsymbol{\beta}_m - \boldsymbol{\Phi}_{\boldsymbol{\beta}}\boldsymbol{\xi}] + \frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\xi} \tag{20}$$

where $\mathbf{T}^d \in R^{N_g \times N_{\beta}^d}$ is a transformation matrix with elements equal to 0 or 1 that maps a grid position in subdomain Ω^d to the corresponding grid position in the global domain, and N_l^d denotes the total number of local parameter patterns in subdomain Ω^d. The squared differences between global and local reconstructions are weighted by the matrix $\mathbf{C}$, which is the covariance matrix quantifying the prior variability of the Gaussian-distributed random parameters $\boldsymbol{\beta}$. A regularization term is included to constrain the magnitude of solutions for $\boldsymbol{\xi}$. The minimizing solution is obtained by setting the derivative of $J(\boldsymbol{\xi})$ with respect to $\boldsymbol{\xi}$ equal to zero, resulting in

$$\boldsymbol{\xi} = \frac{1}{2}\boldsymbol{\Phi}_{\boldsymbol{\beta}}^T \mathbf{C}^{-1} \sum_{d=1}^{S} \mathbf{T}^d \boldsymbol{\Phi}_{\boldsymbol{\beta}}^d \boldsymbol{\xi}^d \tag{21}$$

The parameter field can then be reconstructed by insertion of the solution in Eq. 6, resulting in

$$\boldsymbol{\beta} = \boldsymbol{\beta}_m + \frac{1}{2}\boldsymbol{\Phi}_{\boldsymbol{\beta}}\boldsymbol{\Phi}_{\boldsymbol{\beta}}^T \mathbf{C}^{-1} \sum_{d=1}^{S} \mathbf{T}^d \boldsymbol{\Phi}_{\boldsymbol{\beta}}^d \boldsymbol{\xi}^d \tag{22}$$

It can be seen that smooth solutions for the parameter field can be expressed directly in terms of the local PCA coefficients ξ^d. Furthermore, this smooth local parameterization (SLP) is a differentiable and linear transformation, which makes the parameter estimation problem suitable for use of continuous gradient-based minimization approach. The last panel of Fig. 3 shows reconstructed parameter fields using the proposed SLP approach. Table 1 summarizes the maximum number of local PCA coefficients, e.g., $\max\{N_l^d\}$ and total number of local PCA patterns, e.g., $N_L = \sum_{d=1}^{S} N_l^d$, $d = 1, 2, ..., S$, among all subdomains for different domain decomposition strategies. Decomposition with a larger number of subdomains results in fewer local PCA coefficients in each subdomain.

A final question to be addressed is the number of local PCA patterns that need to be retained. In order to guarantee smoothness and full reconstruction of the global solution space, many local PCA patterns may be required. However, previous experiments [34] have indicated that the overall computational cost is proportional to the number of PCA patterns. We will discuss here a procedure to estimate the minimum number of required patterns that delivers efficiency and guarantees a desired accuracy.

Eq. (21) can be written in compact form as

$$\boldsymbol{\xi} = \mathbf{T}_{GL}\, \boldsymbol{\xi}_L \tag{23}$$

where $\boldsymbol{\xi}_L = [\boldsymbol{\xi}^{1^T}, ..., \boldsymbol{\xi}^{d^T}, ..., \boldsymbol{\xi}^{S^T}]^T$. $\mathbf{T}_{GL} \in R^{N_G \times N_L}$ is a transformation matrix for converting local PCA coefficients into global PCA coefficients, which, after unfolding, can be expressed as

$$\mathbf{T}_{GL} = \frac{1}{2}\boldsymbol{\Phi}_{\boldsymbol{\beta}}^T \mathbf{C}^{-1}[\mathbf{T}^1\boldsymbol{\Phi}_{\boldsymbol{\beta}}^1, \mathbf{T}^2\boldsymbol{\Phi}_{\boldsymbol{\beta}}^2, \cdots, \mathbf{T}^d\boldsymbol{\Phi}_{\boldsymbol{\beta}}^d, \cdots, \mathbf{T}^S\boldsymbol{\Phi}_{\boldsymbol{\beta}}^S] \tag{24}$$

The global projection basis matrix $\boldsymbol{\Phi}_{\boldsymbol{\beta}}$ can be written in unfolded form as follows

$$\boldsymbol{\Phi}_{\boldsymbol{\beta}} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_i, \cdots, \boldsymbol{\phi}_{N_G}] \tag{25}$$

where $\boldsymbol{\phi}_i \in R^{N_g \times 1}$, $i \in \{1, 2, \cdots, N_G\}$ is one global basis vector (we have dropped the subscript β on the individual basis vectors for notational clarity). Defining $\mathbf{e}_i$ as the unit column vector with a 1 at position i and zeros elsewhere, each individual global basis vector can be extracted as follows

$$\boldsymbol{\phi}_i = \boldsymbol{\Phi}_{\boldsymbol{\beta}}\, \mathbf{e}_i, \quad i = 1, 2, ..., N_G \tag{26}$$

The solution space for the full parameter field is spanned by the global PCA patterns. SLP must therefore be able to produce any solution for the global PCA coefficients $\boldsymbol{\xi}$. In other words, the following set of equations should have solution $\boldsymbol{\xi}_L^i$

$$\mathbf{T}_{GL}\boldsymbol{\xi}_L^i = \mathbf{e}_i, \quad i = 1, 2, \cdots, N_G \tag{27}$$

The transformation matrix $\mathbf{T}_{GL}$ should be full-row rank, which implies the total number of local PCA patterns N_L must be equal to or larger than that of global PCA patterns N_G. The rank of transformation matrix $\mathbf{T}_{GL}$ cannot be explicitly determined from Eq. (24). Therefore, we propose a numerical procedure to determine a minimum for the total number of local patterns N_L. We assume that $\boldsymbol{\xi}_L^i(N_L)$ and $\boldsymbol{\phi}_i^*(N_L)$ as functions of preserved number of local PCA coefficients N_L are the solution of the ith equation and the reconstructed i the global PCA basis vector, respectively. In this paper, we currently preserve the same number of local PCA patterns for each subdomain. We define a root mean square error $RMSE_i(N_L)$ to quantitatively characterize the accuracy of the reconstructed global PCA basis vector $\boldsymbol{\phi}_i^*$. The total $RMSE(N_L)$ is the average root mean square error of all global PCA basis vector reconstructions. The minimum N_L is determined by minimizing the total $RMSE(N_L)$ as follows

$$N_L = \arg\min_{N_L} RMSE(N_L) = \arg\min_{N_L} \sum_{i=1}^{N_G} RMSE_i(N_L) \tag{28}$$

where

$$RMSE_i(N_L) = \|\boldsymbol{\phi}_i - \boldsymbol{\phi}_i^*(N_L)\|_2 = \|\boldsymbol{\phi}_i - \boldsymbol{\Phi}_{\boldsymbol{\beta}} \times \mathbf{T}_{GL}\boldsymbol{\xi}_L^i(N_L)\|_2 \tag{29}$$

We solve this problem by simply evaluating the cost function (Eq. (28)) for increasing values of N_L until a minimum is found. Fig. 4 shows the evolution of the $RMSE(N_L)$ as a function of the number of local PCA patterns N_L for 2 different global energy criteria. It can be seen that a minimum N_L can be obtained that makes $RMSE(N_L)$ equal to zero. Table 2 summarizes the minimum number of local PCA patterns corresponding to different domain decomposition strategies. We can confirm that the minimum total number of local PCA patterns N_L is equal to or slightly larger than the number of global PCA patterns N_G. Fig. 5 and Fig. 6 show the 1st global basis vector $\boldsymbol{\phi}_1$, 72nd global basis vector $\boldsymbol{\phi}_{72}$ and their corresponding reconstructed $\boldsymbol{\phi}_1^*$ and $\boldsymbol{\phi}_{72}^*$ when different number of local PCA patterns are retained. The global basis vectors are accurately reconstructed as long as we retain at least the minimum number of local PCA patterns.

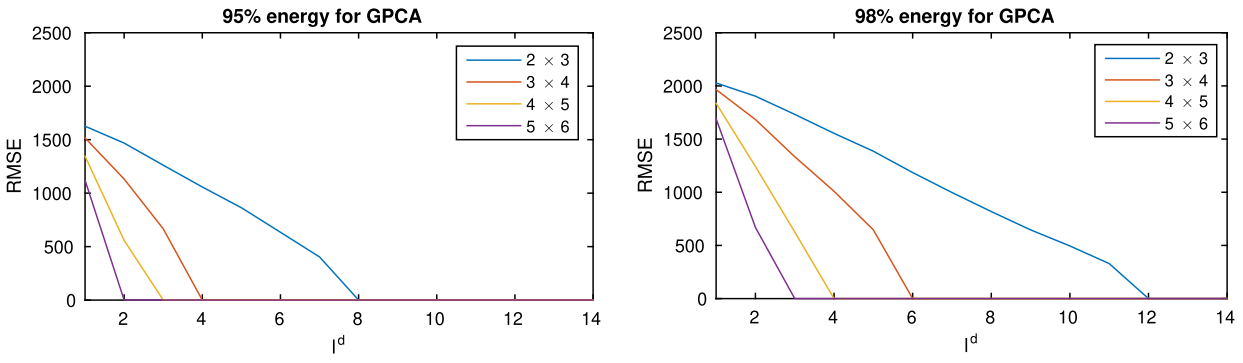


Fig. 4. The *RMSE* Eq. (28) for different number of local PCA patterns in each subdomain. Four domain decomposition strategies are considered, consisting of 2×3, 3×4, 4×5 and 5×6 subdomains respectively.

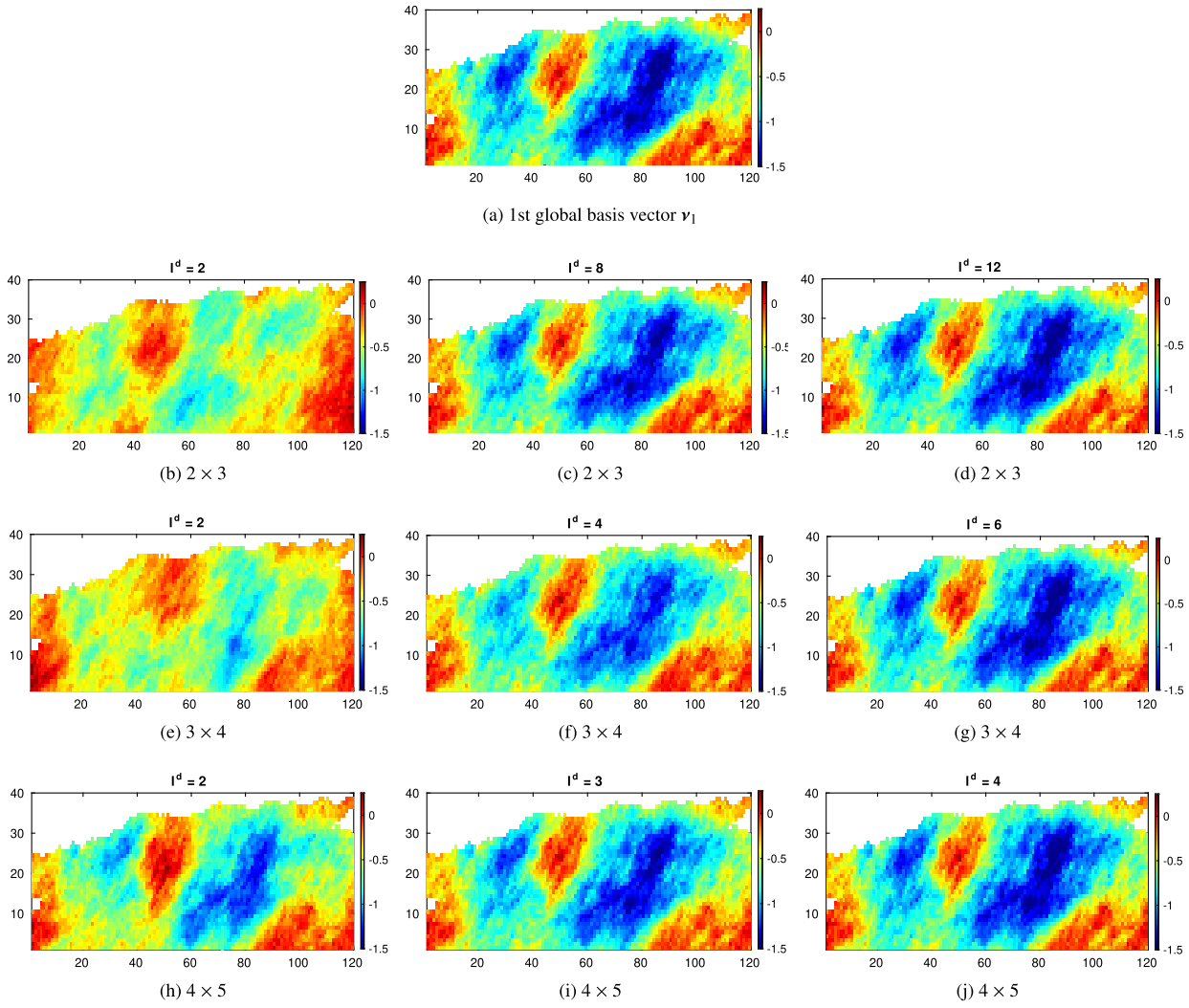
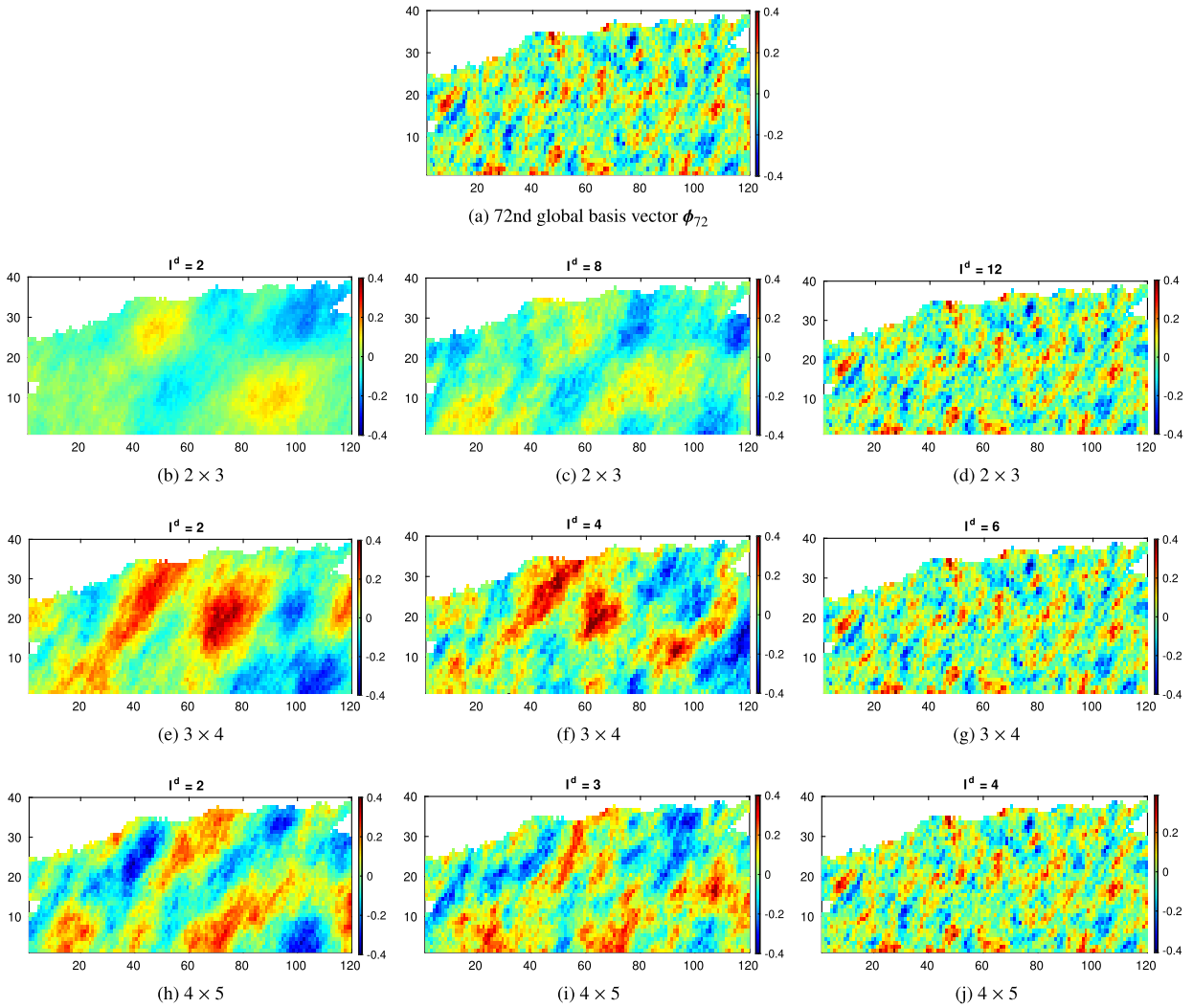


Fig. 5. The reconstructed 1st global basis vector ϕ_1 using different domain decomposition strategies and different number of local PCA patterns in each subdomain. Four domain decomposition strategies are considered, consisting of 2×3, 3×4, 4×5 and 5×6 subdomains respectively.

Table 2
The minimum number of local PCA patterns corresponding to different domain decomposition strategies when $RMSE = 0$.

Energy for global PCA	95%		98%	
N_G	48		72	
Domain Decomposition	$\min\{N_l^1, ..., N_l^d, ..., N_l^S\}$	N_L	$\min\{N_l^1, ..., N_l^d, ..., N_l^S\}$	N_L
2×3	8	48	12	72
3×4	4	48	6	72
4×5	3	60	4	80
5×6	2	60	3	90



(a) 72nd global basis vector $\boldsymbol{\phi}_{72}$

(b) 2×3 (c) 2×3 (d) 2×3

(e) 3×4 (f) 3×4 (g) 3×4

(h) 4×5 (i) 4×5 (j) 4×5

Fig. 6. The reconstructed 72nd global basis vector $\boldsymbol{\phi}_{72}$ using different domain decomposition strategies and different number of local PCA patterns in each subdomain. Four domain decomposition strategies are considered, consisting of 2×3, 3×4, 4×5 and 5×6 subdomains respectively.

5. Adjoint-based data assimilation with smooth local parameterization

Ideally we expect to generate reduced-order linear models for each large subdomain individually. However the effects of the neighboring subdomains should be considered to generate reduced-order linear models for each small subdomain. Both the local POD and local PCA patterns of these neighboring subdomains are also used as the inputs for the RBF interpolation, which will increase the number of interpolation variables for the RBF and hence requires us running additional training models. To avoid the extra computational cost, we assume that each subdomain is sufficient “large ”so that it can be handled individually. We should note that it will never be possible to easily clarify these local dependencies in realistic applications, and the definition of “large ”or “small ”is subjective. An inappropriate domain decomposition strategy, e.g.,

too “large ”or too “small ”subdomains, adversely causes spurious long-distance dependency or cuts off real dependencies among neighboring subdomains and hence deteriorates the quality of the reduced-order models. The sensitivity analysis of the domain decomposition strategy will be provided in the following numerical experiments.

After integrating subdomain POD-TPWL with smooth local parameterization, which is referred to as LSPT hereinafter in this paper, the local POD coefficients $\hat{\boldsymbol{\psi}}^{d,n}$ of subdomain Ω^d are reformulated by directly replacing the variables $\boldsymbol{\xi}$ and $\boldsymbol{\xi}_{tr}$ defined in Eq. (13) - Eq. (18) with the new variables $\boldsymbol{\xi}^d$ and $\boldsymbol{\xi}^d_{tr}$. LSPT consists of an off-line stage and an on-line stage: (1) The off-line stage constructs SLP and reduced-order linear models for each subdomain. (2) The on-line stage implements LSPT given a set of new parameters.

The second assumption that the computational cost of the parameter estimation procedure is proportional to the typical number of local PCA patterns in any one subdomain proposed in Section 3 suggests that an efficient sampling strategy is essential for the implementation of RBF interpolation. The training simulations used to construct the RBF interpolation model should allow for accurate computation of derivative matrices. The procedure for choosing these training points will be described here. The accuracy of the RBF interpolation will be reduced if too few data points are chosen, while the computational cost increases with the number of data points, which will be prohibitive if too many points are chosen. To limit the number of FOM simulations used to construct the interpolation model for the local PCA coefficients we use 2-sided perturbation of each coefficient ξ^d_j resulting in totally $2 \times N^d_l + 1$ points.

The proposed LSPT approach can be incorporated into a parameter estimation algorithm based on the adjoint method. As an approximation to the original objective function J given by Eq. (3), a new objective function J_{ROM} computed using reduced-order models is defined as

$$J_{ROM}(\boldsymbol{\xi}_L) = \frac{1}{2}[\sum_{d=1}^{S}\mathbf{T}^d(\boldsymbol{\beta}^d_m + \boldsymbol{\Phi}^d_{\boldsymbol{\beta}}\boldsymbol{\xi}^d) - \boldsymbol{\beta}_m]^T\mathbf{C}^{-1}[\sum_{d=1}^{S}\mathbf{T}^d(\boldsymbol{\beta}^d_m + \boldsymbol{\Phi}^d_{\boldsymbol{\beta}}\boldsymbol{\xi}^d) - \boldsymbol{\beta}_m] + \frac{1}{2}\sum_{d=1}^{S}\sum_{m=1}^{N_{obs}}[\mathbf{d}^{d,m}_{obs} - \hat{\mathbf{y}}^{d,m}]^T\mathbf{R}_m{}^{-1}[\mathbf{d}^{d,m}_{obs} - \hat{\mathbf{y}}^{d,m}] \tag{30}$$

We introduce an adjoint model to compute the gradient of the objective function J_{ROM} with respect to local PCA coefficients $\boldsymbol{\xi}_L$ where we follow the mathematical derivation provided in [4]. A modified objective function $\hat{J_{ROM}}(\boldsymbol{\xi}_L)$ is obtained by adjoining the reduced-order linear model, resulting in

$$\hat{J}_{ROM}(\boldsymbol{\xi}_L) = J_{ROM}(\boldsymbol{\xi}_L) + \sum_{d=1}^{S}\sum_{n=1}^{N}[\hat{\boldsymbol{\psi}}^{d,n} - \boldsymbol{\psi}^{d,n}_{tr} - \mathbf{E}^{d,n}_{\boldsymbol{\psi}_{tr}}(\boldsymbol{\psi}^{d,n-1} - \boldsymbol{\psi}^{d,n-1}_{tr}) - \mathbf{E}^{sd,n}_{\boldsymbol{\psi}_{tr}}(\boldsymbol{\psi}^{sd,n} - \boldsymbol{\psi}^{sd,n}_{tr}) - \mathbf{G}^n_{\boldsymbol{\xi}^d_{tr}}(\boldsymbol{\xi}^d - \boldsymbol{\xi}^d_{tr})]^T\boldsymbol{\lambda}^{d,n} \tag{31}$$

The gradient of J_{ROM} with respect to $\boldsymbol{\xi}^d$ for each subdomain Ω^d is derived as

$$\frac{dJ_{ROM}(\boldsymbol{\xi}_L)}{d\boldsymbol{\xi}^d} = \frac{d\hat{J}_{ROM}(\boldsymbol{\xi}_L)}{d\boldsymbol{\xi}^d} - \sum_{n=1}^{N}[\mathbf{G}^n_{\boldsymbol{\xi}^d_{tr}}]^T\boldsymbol{\lambda}^{d,n} \tag{32}$$

where

$$\frac{dJ_{ROM}(\boldsymbol{\xi}_L)}{d\boldsymbol{\xi}^d} = (\boldsymbol{\Phi}^d_{\boldsymbol{\beta}}\mathbf{T}^d)^T\mathbf{C}^{-1}[\mathbf{T}^d(\boldsymbol{\beta}^d_m + \boldsymbol{\Phi}^d_{\boldsymbol{\beta}}\boldsymbol{\xi}^d) - \boldsymbol{\beta}_m] - \sum_{m=1}^{N_0}[\mathbf{B}^{d,m}_{\boldsymbol{\xi}_{tr}}]^T\mathbf{R}_m{}^{-1}[\mathbf{d}^{d,m}_{obs} - \mathbf{y}^{d,m}_{tr} - \mathbf{A}^{d,m}_{\boldsymbol{\psi}_{tr}}(\boldsymbol{\psi}^{d,m} - \boldsymbol{\psi}^{d,m}_{tr}) - \mathbf{B}^m_{\boldsymbol{\xi}^d_{tr}}(\boldsymbol{\xi}^d - \boldsymbol{\xi}^d_{tr})] \tag{33}$$

The adjoint model is expressed in terms of the Lagrange multipliers $\boldsymbol{\lambda}^{d,n}$ for subdomain Ω^d and is given by

$$[\mathbf{I} - (\mathbf{E}^{d,n}_{\boldsymbol{\psi}_{tr}})^T]\boldsymbol{\lambda}^{d,n} = [\mathbf{A}^{d,n}_{\boldsymbol{\psi}_{tr}}]^T\mathbf{R}_n{}^{-1}[\mathbf{d}^{d,n}_{obs} - \mathbf{y}^{d,n}_{tr} - \mathbf{A}^{d,n}_{\boldsymbol{\psi}_{tr}}(\boldsymbol{\psi}^{d,n} - \boldsymbol{\psi}^{d,n}_{tr}) - \mathbf{B}^n_{\boldsymbol{\xi}^d_{tr}}(\boldsymbol{\xi}^d - \boldsymbol{\xi}^d_{tr})] + [\mathbf{E}^{sd,n}_{\boldsymbol{\psi}_{tr}}]^T\boldsymbol{\lambda}^{d,n+1} \tag{34}$$

for $n = N, \cdots, 1$ with an ending condition $\boldsymbol{\lambda}^{d,N+1} = 0$. The solution of the adjoint model Eq. (34), together with the solution of Eq. (33), can be used in Eq. (32) to obtain the desired total derivative with respect to the local PCA coefficients in subdomain Ω^d. The full gradient with respective to all local PCA coefficients can be obtained as

$$\nabla J = [\frac{dJ_{ROM}(\boldsymbol{\xi}_L)}{d\boldsymbol{\xi}^1}, ..., \frac{dJ_{ROM}(\boldsymbol{\xi}_L)}{d\boldsymbol{\xi}^d}, ..., \frac{dJ_{ROM}(\boldsymbol{\xi}_L)}{d\boldsymbol{\xi}^S}]. \tag{35}$$

Once the gradient ∇J_k at the kth iteration step is available, a steepest descent update of coefficients $\boldsymbol{\xi}_{k+1}^{L}$ is obtained by

$$\boldsymbol{\xi}_L^{k+1} = \boldsymbol{\xi}_L^k - \alpha_k \frac{\nabla J_k}{\|\nabla J_k\|_\infty} \tag{36}$$

where α_k is the step length at the kth iteration step [45]. The minimization process terminates when either one of the following three stopping criteria is satisfied: There is no significant change in the objective function,

$$\frac{|J_{ROM}(\boldsymbol{\xi}_L^{k+1}) - J_{ROM}(\boldsymbol{\xi}_L^k)|}{max\{|J_{ROM}(\boldsymbol{\xi}_L^{k+1})|, 1\}} < \eta_{J_{ROM}} \tag{37}$$

there is no significant change in the parameter estimate,

$$\frac{|\boldsymbol{\xi}_L^{k+1} - \boldsymbol{\xi}_L^k|}{max\{|\boldsymbol{\xi}_L^{k+1}|, 1\}} < \eta_{\boldsymbol{\xi}^L} \tag{38}$$

or the maximum number of iterations has been reached,

$$k = N_{max} \tag{39}$$

where $\eta_{J_{ROM}}$, η_{ξ_L} and N_{max} denote predefined constants and maximum iterative steps respectively. In our experiments $\eta_{J_{ROM}} = 10^{-4}$, $\eta_{\xi_L} = 10^{-3}$. More strict criteria could possibly lead to more accurate results but we did not investigate this. Since the simulation of the reduced-order model is very cheap, we do not limit the maximum number of iterations N_{max} for the inner-loop. The initial step size $\alpha_0 = 0.1$. Once the objective function increases, the step size is divided by 2 to improve convergence.

Oliver [46] and Tarantola [47] discussed the expected range of the optimal objective function. If the relationship between the simulated data and the parameters is linear, and assuming a tolerance of five standard deviations from the mean, the optimal objective function value $J_{FOM}(\boldsymbol{\xi}_L)$ should satisfy the inequality

$$N_{obs} - 5\sqrt{N_{obs}} \leqslant 2\, J_{FOM}(\boldsymbol{\xi}_L) \leqslant N_{obs} + 5\sqrt{N_{obs}} \tag{40}$$

where J_{FOM} is the objective function for the full order model and N_{obs} is the total number of measurements. Since we employ an approximate reduced order model in the minimization we apply a less strict criterion. Throughout this study, we use the following criterion

$$J_{FOM}(\boldsymbol{\xi}_L) \leqslant 5\, N_{obs} \tag{41}$$

This criterion or tolerance can be considered to be an upper-bound to judge whether a set of acceptable optimized parameters has been obtained. Since the reduced-order model in Eq. (30) is not an exact representation of the full-order model given by Eq. (1), a number of outer-loops are typically required to update the reduced-order linear models. After convergence of an inner loop, the updated parameters are used as input for a full model simulation, which is added to the training set. An updated reduced order model is constructed as described in Sections 3 and 4, after which a new inner loop (minimization of Eq. (30) using the reduced order model) is started. In the numerical experiments we use a fixed number of 10 outer-loop iterations for the base-case study, which appear to be sufficient for near-convergence in all cases, after which we evaluate the criterion Eq. (41) for $J_{FOM}(\boldsymbol{\xi}_L^k)$. An overview of the full workflow is provided in Algorithm 1 for the offline stage and Algorithm 2 for the online stage.

Algorithm 1: Parameter and state reduction.

1 Create a large set of model realizations $\boldsymbol{\beta}_i$ for $i = 1, ..., N_\beta$;
2 Form global basis matrix $\boldsymbol{\Phi}_\beta$ and coefficients $\boldsymbol{\xi}$;
3 Partition the global domain into subdomains Ω^d;
4 For local basis matrices $\boldsymbol{\Phi}_\beta^d$ and coefficients $\boldsymbol{\xi}^d$;
5 Simulate a training set of model realizations $\boldsymbol{\xi}_{tr}^d$ and collect snapshots $\mathbf{x}_{tr}^n$;
6 Compute local basis coefficients $\boldsymbol{\psi}_{tr}^{d,n}$;

6. Reservoir history matching experiments

In this section, the LSPT method is applied to a history matching example problem based on a 2D reservoir model describing a two phase water-oil reservoir system containing 6 producers and 7 injectors, labeled P_1 to P_6, and I_1 to I_7 respectively (see Fig. 7). The triangles and circles in the figures denote the injectors and producers, respectively. Some

Algorithm 2: Adjoint-based parameter estimation using LSPT.

Choose an initial $\boldsymbol{\xi}_L^0$, set $k = 0$;
while $J(\boldsymbol{\xi}_L) > 5\,N_{obs}$ **do**
  Construct the RBF model in all subdomains (Eq. (17)-(18));
  Compute the derivative matrices in all subdomains (Eq. (14), (16));
  while *not converged* **do**
    Simulate the RBF model (Eq. (13), (15));
    Solve the adjoint model (Eq. (34));
    Calculate the gradient (Eq. (32)-(33));
    Update the parameter solution $\boldsymbol{\xi}_L^k \rightarrow \boldsymbol{\xi}_L^{k+1}$ (Eq. (36));
    Evaluate the reduced model objective function $J_{ROM}(\boldsymbol{\xi}_L^{k+1})$ (Eq. (30));
    Check convergence (Eq. (37)-(39));
  end
  Simulate the full model (Eq. (1)) with solution $\boldsymbol{\xi}_L$ (using Eq. (23));
  Update the training set with $(\boldsymbol{\xi}^d, \boldsymbol{\psi}^{d,n})$;
  Evaluate the objective function $J_{FOM}(\boldsymbol{\xi})$ (Eq. (3), (22)-(23));
end

Table 3
The abbreviations of the designed numerical experiments.

Abbreviations	Description
LSPT1	Subdomain POD-TPWL with smooth local parameterization (smoothness)
LSPT2	Subdomain POD-TPWL with local parameterization (non-smoothness)
GSPT	Subdomain POD-TPWL with global parameterization
LP-FD	Finite-difference (FD) gradients with smooth local parameterization
GP-FD	Finite-difference (FD) gradients with global parameterization

Table 4
Reservoir, fluid and well properties of the 2D reservoir model.

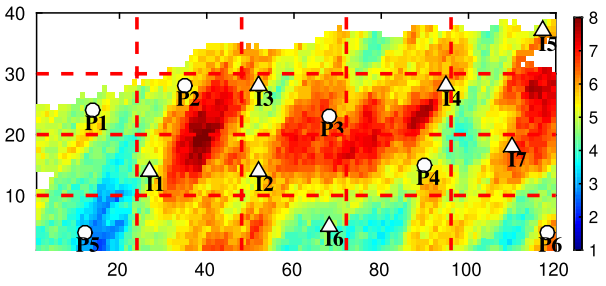
Description	Value
Dimension	40×120×1
Number of wells	6 producers, 7 injectors
Fluid density	1014 kg/m^3, 859 kg/m^3
Fluid viscosity	0.4 mPa·s, 2 mPa·s
Initial pressure	25 MPa
Initial saturation	$S_o = 0.80$, $S_w = 0.20$
Connate water saturation	$S_{wc} = 0.20$
Residual oil saturation	$S_{or} = 0.20$
Corey exponent, oil	4.0
Corey exponent, water	4.0
Injection rate	200 m^3/d
Producer BHP	20 MPa
History production time	10 year
Prediction time	15 year
Model timestep	0.1 year
Measurement timestep	0.2 year

relevant properties of the reservoir geometry, rock properties, fluid properties, and well controls are shown in Table 4. An open-source reservoir simulator [48], Matlab Reservoir Simulation Tool (MRST), is used to simulate the reservoir model.

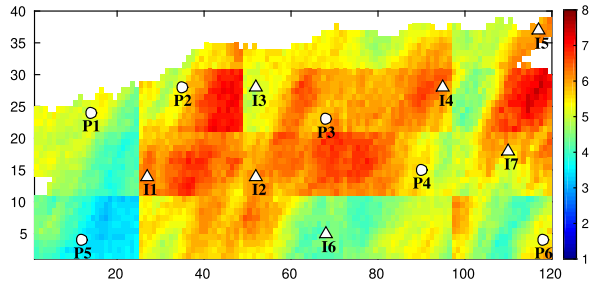
We consider numerical experiments with two variations of LSPT. In the first one (LSPT1), we will minimize the number of local PCA patterns using (Eq. (28)), while in the second one (LSPT2) we use the local PCA patterns obtained by use of an energy cutoff criterion without smoothness projection. For comparison purposes, the LSPT results will be compared against subdomain POD-TPWL with global parameterization (GSPT) and against results obtained with finite-difference (FD) gradients with respect to both global PCA and SLP parameter representations, referred to as GP-FD and LP-FD respectively. The abbreviations of the numerical experiments are provided in Table 3.

6.1. Description of history matching scenarios

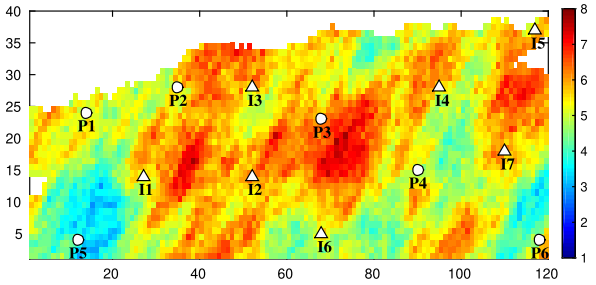
Experiments are presented for two history matching scenarios. Scenario 1 is characterized by the combination of a parameter field dominated by large spatial scales and a relatively small number of well data (fluid rate and watercut (WCT) at the producing wells and bottom-hole pressure (BHP) at the injector wells). In Scenario 2 the spatial scales of the parameter



(a) 'True' model



(b) Projected "true" model using local PCA



(c) Projected "true" model where the local PCA's results are projected onto the global PCA's

Fig. 7. Comparison of the "true" reservoir model in full-order space and reduced-order space for Scenario 1. The triangles and circles denote the injectors and producers, respectively. The global saturation and pressure snapshots are decomposed into 20 rectangle subdomains. The red dash lines represent the boundary of subdomains.

field are substantially smaller. At the same time, a much larger data set is available in this scenario, saturation values in all grid cells, such that it is possible to investigate the scaling of computational efficiency of our proposed methodology with the size of the problem. In both scenarios we will investigate the impact of different decompositions of the model domain into rectangular subdomains. Fig. 7(a) and Fig. 9(a) show a base-case decomposition based on a 4 × 5 subdomains. Note that for this decomposition some subdomains contain no wells, while one of the subdomains contains two wells.

6.1.1. Scenario 1

We select 1 out of 1000 generated model realizations as the truth for scenario 1 (see Fig. 7(a)). A global reduction of the parameter space using a 95% energy cutoff criterion results in $N_G = 48$ preserved global PCA patterns. For the first step, we implement the local PCA (also employing a 95% energy cutoff criterion) without using our proposed smoothness projection. The number of local PCA patterns retained in each subdomain is shown in Fig. 1, which results in totally $N_L = \sum_{d=1}^{S} l_d = 275$ local PCA coefficients. Fig. 7 (a) and Fig. 7 (b) separately represents the projected 'true' permeability field using SLP and local PCA. We follow the procedure described in [34] to determine the number of full order model simulations for the selection of snapshots. 22 simulations were run using random global PCA coefficients sampled from the set $\{-1, 1\}$, from which a total number of 2200 (22 simulation models by 100 time steps) snapshots of both pressure and saturation were extracted. These 2200 global saturation and pressure snapshots are decomposed into 20 rectangle subdomains. For each subdomain, two separate eigenvalue problems for pressure and saturation are solved using POD on the two sets of 2200 snapshots through using a 95% energy cutoff criteria. The resulting number of POD patterns for each subdomain is shown in Fig. 8. We have collected the snapshots at all time steps (in our study), although sometimes it is not necessary to resemble the snapshots for each time step as suggested in the literature [20]. Instead of taking global basis functions to define the subspace, the snapshots are first partitioned according to the domain decomposition strategy and then local basis functions are obtained from these partitioned snapshots. As a result, the implementation of POD for a large number of snapshots will not pose severe computational problems in the realistic applications.

6.1.2. Scenario 2

Scenario 2 aims to investigate the possibility of estimating a much larger number of parameters than in Scenario 1, given the availability of a much larger number of measured data. A second set of 1000 Gaussian-distributed logarithmic permeability fields is generated resulting in a set of N_G = 282 preserved global PCA patterns. As in Scenario 1, a 4 × 5 subdomain decomposition is chosen as a base-case (see Fig. 9(a)). Fig. 9 shows the "true", and projected "true" logarithmic

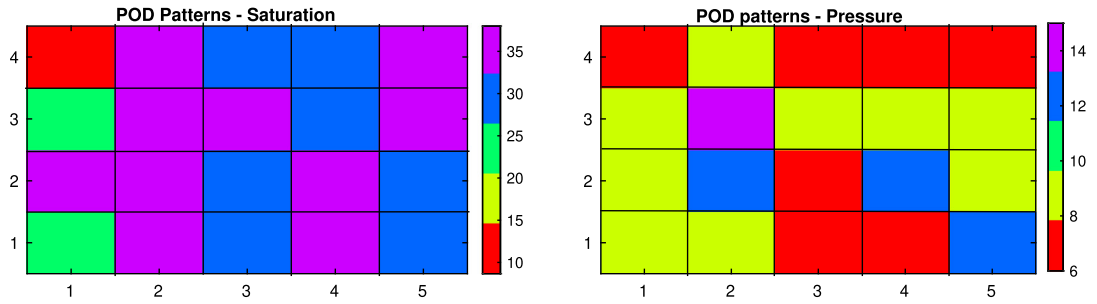
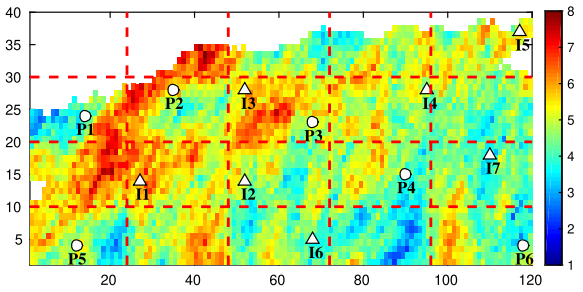
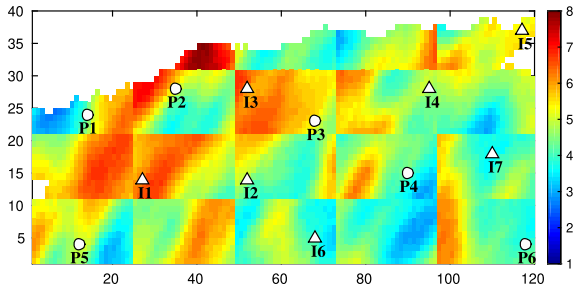


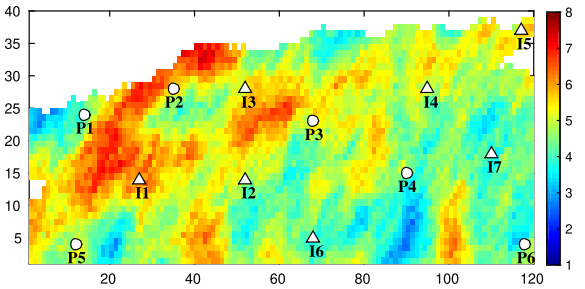
Fig. 8. The number of reduced pressure and saturation POD patterns in each subdomain for Scenario 1. The global saturation and pressure snapshots are decomposed into 20 rectangle subdomains.



(a) 'True' model



(b) Projected 'true' model using local PCA



(c) Projected 'true' model where the local PCA's results are projected onto the global PCA's

Fig. 9. Comparison of the "true" reservoir model in full-order space and reduced-order space for Scenario 2. The triangles and circles denote the injectors and producers, respectively. The global saturation and pressure snapshots are decomposed into 20 rectangle subdomains. The red dash lines represent the boundary of subdomains.

permeability fields using local PCA and SLP, respectively. For this scenario 32 full-order model simulations were run to select snapshots. The resulting number of local POD patterns in each subdomain is shown in Fig. 10.

6.2. Construction of the reduced order model

In scenario 1, a total of $53 = 22 + 2 \times 15 + 1$ full-order model simulations are run to construct the subdomain reduced-order model. 22 full-order model simulations are used to collect the snapshots to construct the bases for the states, 1 full-order model simulation is the specific training trajectory that is used in the linearization, and an additional $M = 30 = (2 \times 15)$ full-order model simulations are run with perturbed parameter inputs to construct the subdomain reduced-order linear model (Eq. (17)-(18)). The vectors of PCA patterns which correspond to these 53 training models are sampled from a training interval $\boldsymbol{\xi} \in [-1, 1]$ by use of a two-sided perturbation method centered on $\boldsymbol{\beta}_{tr}$.

In terms of computational effort, the runtime for a single full-order model simulation for this case is about 9.8 s on a machine with i5-4690 Intel CPUs (4 cores, 3.5 GHz) and 24 GB memory using Matlab-R2015a. The LSPT base-case models for both scenarios, by contrast, require less than 0.3 s. However, the LSPT models for scenario 1 and 2 require respectively, 53 and 72 training models plus additional overhead. Therefore, it would not make sense to construct the LSPT reduced-order

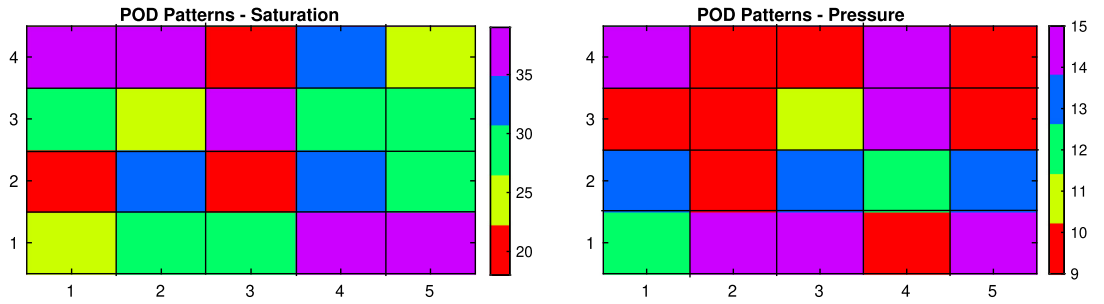


Fig. 10. The number of reduced pressure and saturation POD patterns in each subdomain for Scenario 2. The global saturation and pressure snapshots are decomposed into 20 rectangle subdomains.

model unless it is to be used for a large number of simulations. Because many simulations are required in history matching applications, the use of LSPT models should be attractive.

6.3. Error measures

To assess the performance of LSPT, we define the following relative errors with respect to the full model simulation reference,

$$E_d = \frac{1}{NN_d} \sum_{i=1}^{N} \sum_{j=1}^{N_d} \frac{|(\mathbf{d}_{FOM}^{i,j} - \mathbf{d}_{ROM}^{i,j})|}{\mathbf{d}_{FOM}^{i,j}} \tag{42}$$

where $\mathbf{d}$ represents the data vector (fluid rates, WCT, BHP and/or water saturation), and

$$E_x = \frac{1}{NN_d} \sum_{i=1}^{N} \sum_{j=1}^{N_d} \frac{|(\mathbf{x}_{FOM}^{i,j} - \mathbf{x}_{ROM}^{i,j})|}{\mathbf{x}_{FOM}^{i,j}} \tag{43}$$

where, $\mathbf{x}$ represents the state vector (saturation and/or pressure in each gridblock).

We analyze the dependence of the model errors with respect to domain decomposition, energy cutoff criterion, testing interval and training interval. The testing interval or training interval represents the predefined perturbation intervals from which the testing and training samples are selected. Fig. 11 shows the RMSE error in fluid rate, water-cut, pressure and saturation as a function of these four factors for scenario 1. For scenario 2, Fig. 12 depicts the predictions of gridblock saturation and the corresponding relative errors at day 1825 and day 3650 using FOM and LSPT simulations for a relatively small subdomain size of 3×4 cells, which produced the most accurate results for this case. Accuracy is also improved by increasing the energy threshold and thus retaining more POD patterns, albeit at an increased computational cost. Retaining 95% of the total energy during projection produces an acceptable accuracy in this case. The testing intervals and training intervals represent the magnitude of two-side perturbations around the linearized trajectory corresponding to $\boldsymbol{\xi}_{tr}$. The testing case and training models will be randomly sampled from these two intervals, respectively. Increasing the testing interval, which represents the maximum discrepancy between test model and linearized training model, deteriorates the accuracy of the reduced model, with the best results obtained here with a [-0.1, 0.1] interval.

In order to evaluate the quality of parameter estimation results, we will compare the value of the final objective function against the tolerance (Eq. (41)) and against the reference objective function values for the true model (reflecting the impact of the data noise) and the projected true model (that is the best possible reconstruction of the truth given the selected PCA patterns). Reconstructed parameter maps will provide a visual indication of smoothness and uniqueness of the solution. For all approaches we will list the computational cost expressed in terms of the number of full order model simulations.

6.4. Generation of noise measurements

The historic production period is 10 years, during which well measurements are taken at 0.2 year intervals, resulting in 50 time instances in total. Normal distributed independent measurement noise with a standard deviation equal to 5% of the 'true' data value was added to all measurements. The complete well data set consist of 300 fluid rates and 300 WCT values measured in the producers and 350 bottom-hole pressures measured in the injectors (950 measured data points in total).

The seismic data used in scenario 2 correspond to the saturation values from the 'true' model simulation after 1825 days (1st monitor) and 3650 days (2nd monitor) of production, mimicking the collection of data from two seismic monitor surveys. For this scenario there are in total 8920 measurements. The noisy measurements for the two monitor surveys are shown in Fig. 13.

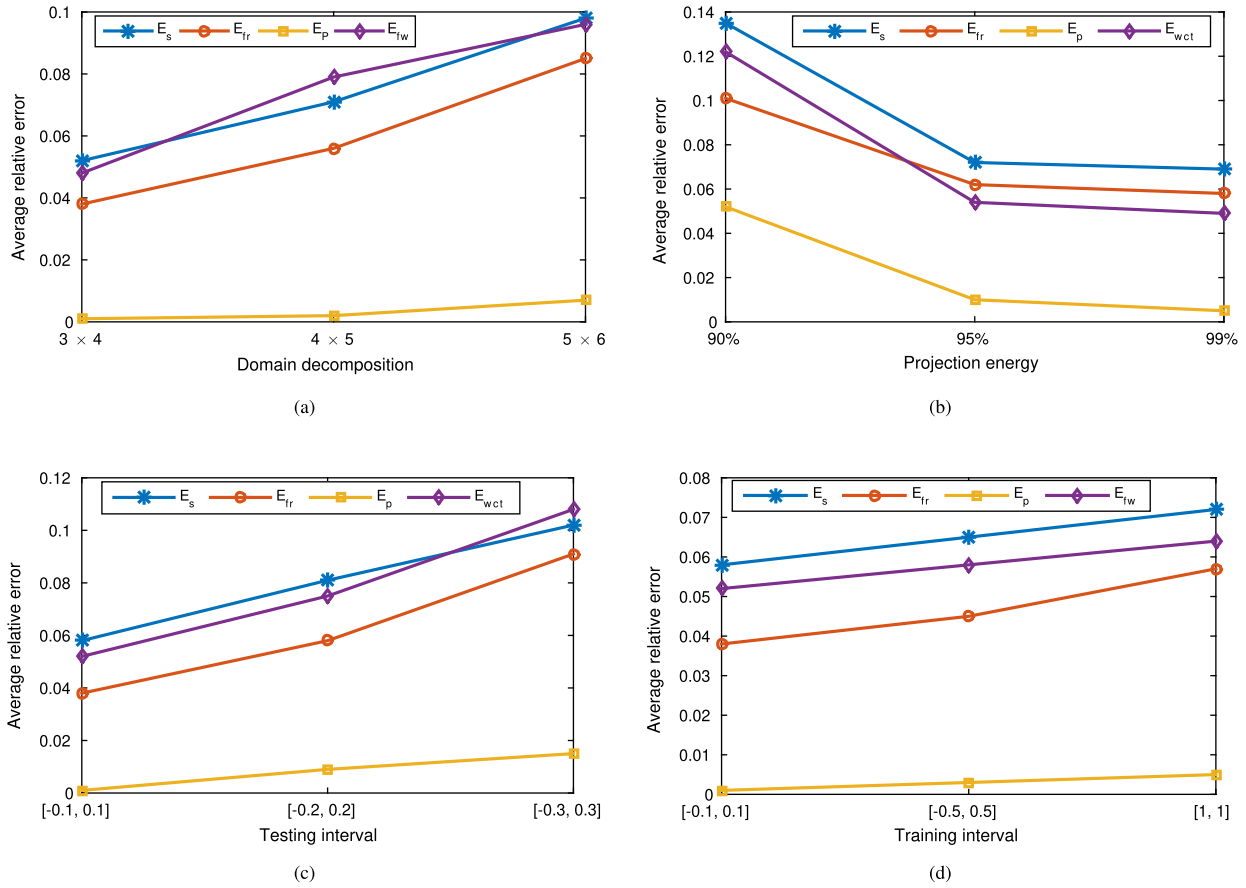


Fig. 11. Average LSPT errors as a function of domain decomposition, energy cutoff criterion, testing interval and training interval for Scenario 1. Results are for LSPT models constructed using 53 training simulations for the test case.

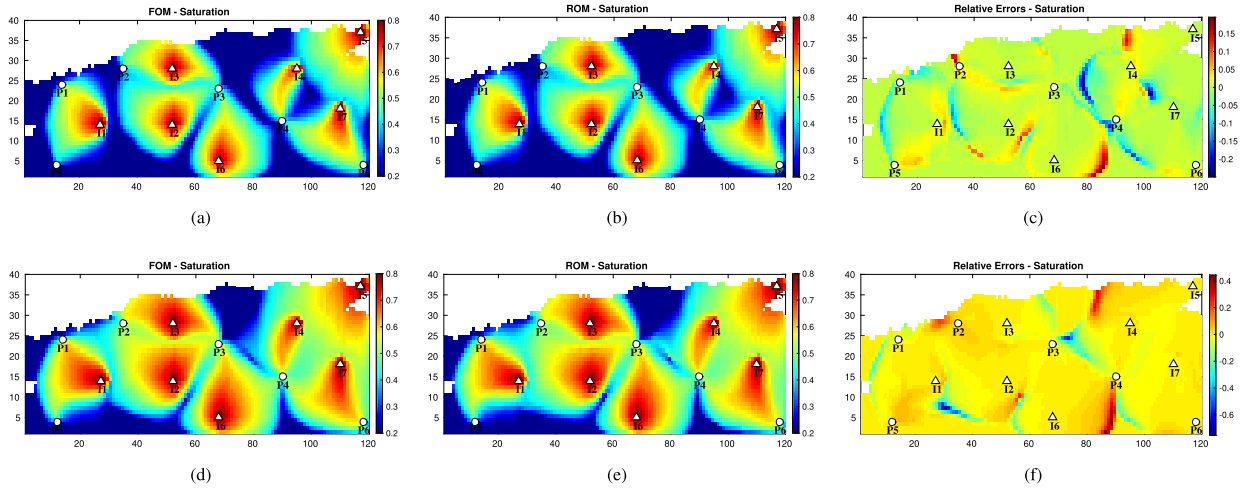


Fig. 12. Predictions of saturation distribution and its corresponding relative errors at the 1825 days and the 3650 days using FOM and LSPT for Scenario 2. The first row is at the 1825 days, while the second row is at the 3650 days.

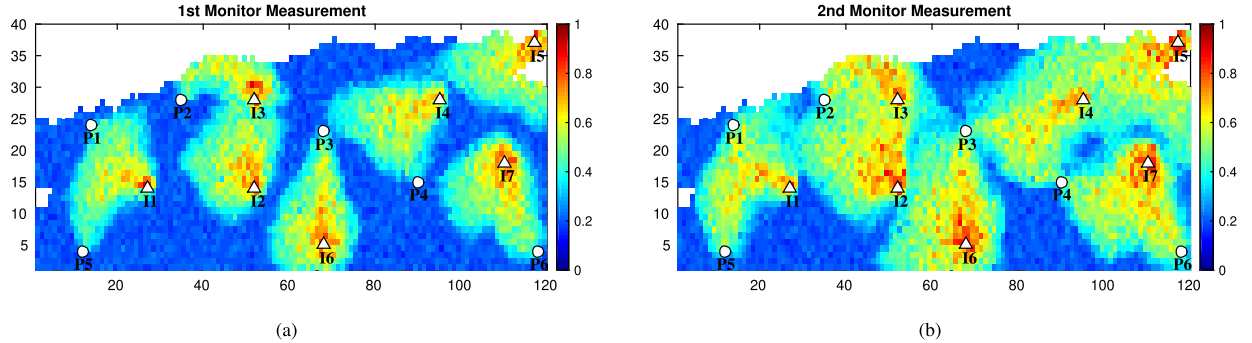


Fig. 13. Noise distribution of water saturation for scenario 2. Normal distributed independent measurement noise with a standard deviation equal to 5% of the 'true' data value was added to all observations.

7. Results

7.1. Scenario 1

7.1.1. Base-case

Fig. 14 and Table 5 summarize the evolution of the objective function values over the outer and inner iteration loops, and the initial and final values respectively. (Note that the jumps of the objective function values in the inner-loop iterations as Fig. 14(a) are the starting points of new outer-loop iterations.) Fig. 15 and Fig. 16 depict the true, initial (prior) and estimated (posterior) logarithmic permeability fields, and Fig. 17 shows the prior and poster data mismatches.

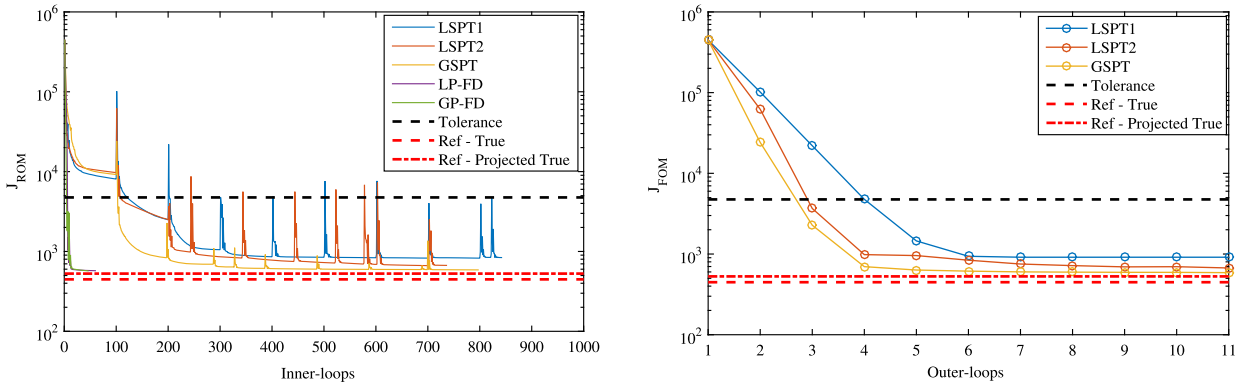
Fig. 14 reveals that GP-FD, LP-FD and GSPT obtain similar objective function values after the minimization, while our proposed LSPT obtains a slightly less accurate result. As can be seen from Table 5, both LSPT1 and LSPT2 require 62 FOM simulations, among them, 22 FOM simulations are used to collect the snapshots to construct the POD basis matrix for the states, 1 FOM simulation is the initial training trajectory that is used in the linearization, and $M = 30 = (2 \times 15)$ FOM simulations are run with perturbed parameter inputs to construct the subdomain reduced-order linear model (Eq. (17)-(18)), 9 additional FOM simulation are used to update the reduced-order linear models in the following 9 outer-loops. GSPT requires $22 + (4 \times 48 + 1) + 9 = 224$ FOM simulations. Specifically, 22 full-order model simulations are used to collect the snapshots to construct the bases for the states, $M = 193 = (4 \times 48 + 1)$ full-order model simulations are run with perturbed parameter inputs to construct the initial subdomain reduced-order linear model at the 1st out-loop. Additional 9 full-order model simulation are the specific training trajectories that are used in the linearization at the following 9 outer-loops.

Fig. 15 displays the true, initial and final updated logarithmic permeability fields. Although LSPT2 obtains relatively low objective function values, the updated logarithmic permeability field is spatially non-smooth, which does not satisfy the (geological) assumptions underlying the model, and would therefore have to be rejected. LSPT1 produces an acceptable solution in terms of final objective function value and in terms of the spatial properties of the reconstructed parameter field. In Fig. 15 (c) one example area (red dashed rectangle) is highlighted in which the logarithmic permeability field is not correctly reconstructed due to lack of observations (no wells are present in the corresponding subdomain). The choice of domain decomposition may therefore have a significant influence on the performance of LSPT. We will further investigate this issue in the following section.

It can be seen from Fig. 16 that GSPT and LSPT1 obtain acceptable results after 5 outer-loops, and therefore that some additional FOM runs are not necessary. Fig. 17 illustrates the match for fluid rate, water-cut data and bottom-hole pressure up to 10 years and an additional 15-year prediction for all six producers and seven injectors. The predictions based on the initial model are far from that of the true model. After the history matching, the predictions of the updated models match the observations very well. Also the prediction of the water breakthrough time is improved for all production wells, including the wells that show water breakthrough only after the history matching period.

7.1.2. Impact of domain decomposition strategy

Fig. 18, Fig. 19 and Table 6 show the effects of domain decomposition strategy on the objective function minimization and the final estimate of the logarithmic permeability fields. Four strategies are considered, consisting of 2×3, 3×4, 4×5 and 5×6 subdomains respectively. To ensure that all these four schemes converge to a final solution, we specified the maximum number of outer-loops as 15 in this experiment. The total number of local PCA patterns and the maximum local PCA patterns among all subdomains are summarized in Table 6. It can be seen that a higher number of subdomains will result in a lower number of local PCA patterns per subdomain. As a result, fewer FOM simulations are required. These numerical results demonstrate our aforementioned motivation that the number of required FOM simulations depends primarily on the maximum number of local PCA patterns in a subdomain, not on the underlying full-order model, which can be decreased by refining the domain decomposition. Fig. 18 and Fig. 19 demonstrate that the domain decomposition strategy has significant influence on the performance of LSPT. All four domain decomposition strategies obtain an acceptable cost function



(a) (b)

Fig. 14. Evolution of the objective function values using LSPT, GSPT, LP-FD and GP-FD method for Scenario 1 as a function of outer-loops. The computation of the objective function for inner-loops and outer-loops uses reduced-order linear model and full-order model, respectively. The two red dash-lines separately represent reference objective function values for the true model (reflecting the impact of the data noise) and the projected true model (that is the best possible reconstruction of the truth given the selected PCA patterns).

Table 5
The number of required FOM simulations and final objective function values for LSPT1, LSPT2, GSPT and FD method for Scenario 1. The reference values of true model and projected true model are also shown here. The domain decomposition for this base-case is 4×5.

–	Iterations	Number of FOM simulations	$J_{FOM}(\xi)$
Initial model	–	–	4.49×10^5
LSPT1	10	62 = 22+(2×15+1)+9	912.93
LSPT2	10	62 = 22+(2×15+1)+9	697.32
GSPT	10	224 = 22+(4×48+1)+9	587.83
LP-FD	61	4421	573.94
GP-FD	47	2773	571.73
Tolerance	–	–	4750
Ref - Projected True	–	–	528.1
Ref - True	–	–	447.4

Table 6
The number of required FOM simulations and cost function values of LSPT using different domain decomposition strategies, e.g., 2×3, 3×4, 4×5 and 5×6, for Scenario 1.

Domain decomposition	N_L	$\max\{l_d\}, d = 1, 2, .., S$	N_G	Iterations	Number of FOM simulations	$J_{FOM}(\xi)$
Initial model	–	–	–	–	–	4.49×10^5
2×3	112	20	48	15	77 = 22+(2×20+1)+14	901.69
3×4	205	18		15	73 = 22+(2×18+1)+9	878.21
4×5	275	15		15	67 = 22+(2×15+1)+9	912.93
5×6	322	12		15	61 = 22+(2×12+1)+9	869.01
Tolerance	–	–	–	–	–	4750
Ref - Projected True	–	–	–	–	–	528.1
Ref - True	–	–	–	–	–	447.4

value after minimization. However, the updated log-permeability fields differ significantly, which implies that different local minima are generated using different domain decomposition strategies.

7.1.3. Quantification of different sources of errors

Three main sources of errors (SOE) contribute to the over-all quality of the history matching results: (1) approximation errors of the subdomain POD-TPWL (SOE1), e.g., POD, RBF, and domain decomposition; (2) the loss of global PCA patterns due to an insufficient number of local PCA patterns (SOE2); and (3) only a fraction, e.g. 95%, of the full energy is preserved by the global PCA (SOE3). To distinguish and quantify these three error contributions, LSPT and FD are consecutively implemented. After minimizing the objective function using LSPT, continuing minimization using FD-LP can quantify the SOE 1,

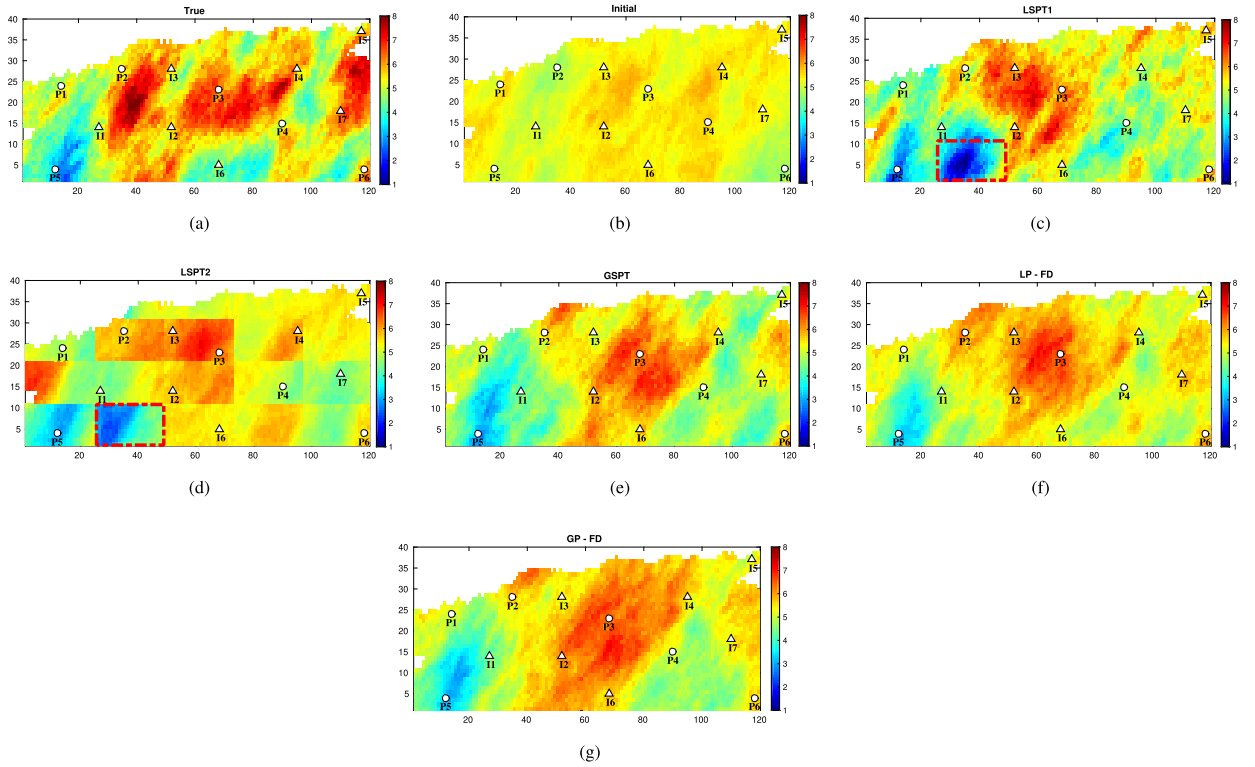


Fig. 15. Comparison of the updated logarithmic permeability fields from the LSPT1, LSPT2, GSPT, LP-FD and GP-FD method for Scenario 1. The true model and initial model are displayed here for a comparison.

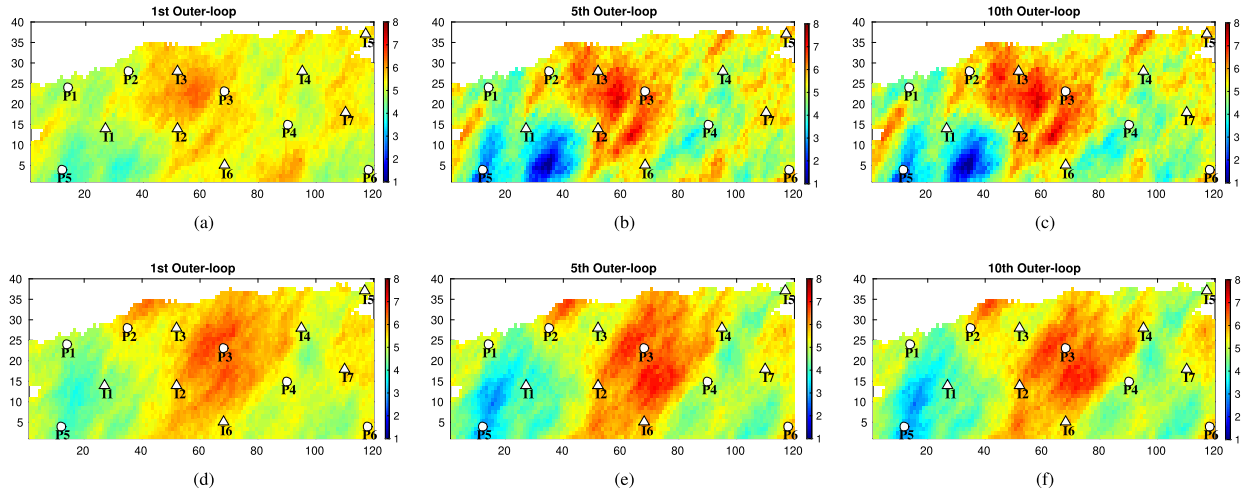


Fig. 16. Comparison of updated logarithmic permeability fields from LSPT1 and GSPT in 1st, 5th, and 10th steps of outer-loop for Scenario 1. Subfigures (a)-(c) are for LSPT1, and subfigures (d)-(f) are for GSPT.

while further minimization using FD-GP can quantify the sum of SOE1 and SOE2. To quantify SOE3, the objective function is minimized by successively preserving an increasing fraction of the global PCA energy, e.g., 95%, 98%, 99% and 99.5%.

We estimate the different error contributions in Table 7, Table 8 and Fig. 20 for a decomposition strategy with 2×3 subdomains and a fixed number of 15 outer-loop iterations. The impact of the number of retained local PCA patterns is tested using values of 2, 8 and 20 in all subdomains. Table 7 summarizes the initial, final and reference objective function values, the total sum of local PCA patterns and the required number of FOM simulations. A global reduction of the parameter space using a 95% energy cutoff criterion results in $N_G = 48$ preserved global PCA patterns. The numerical minimum of local PCA patterns required to fully cover these 48 global PCA patterns is 8. Table 7 (last column) and the yellow curve

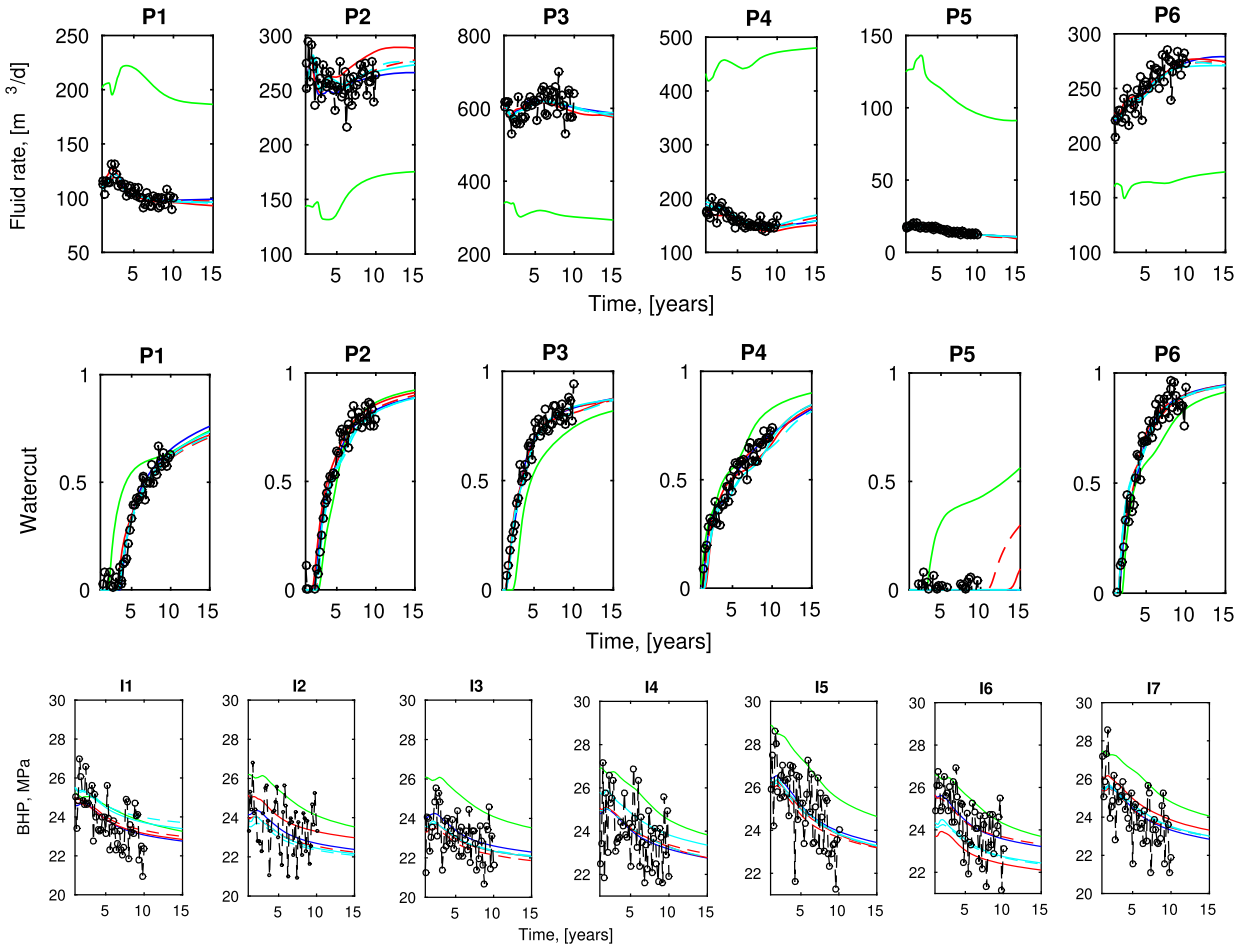


Fig. 17. Forecast of the producers' liquid rate, WCT and injectors' BHP for Scenario 1: green line-initial model, blue line-'true' model, solid red line - LSPT1, dash red line - GSPT, solid cyan line - LP-FD, dash cyan line - GP-FD. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

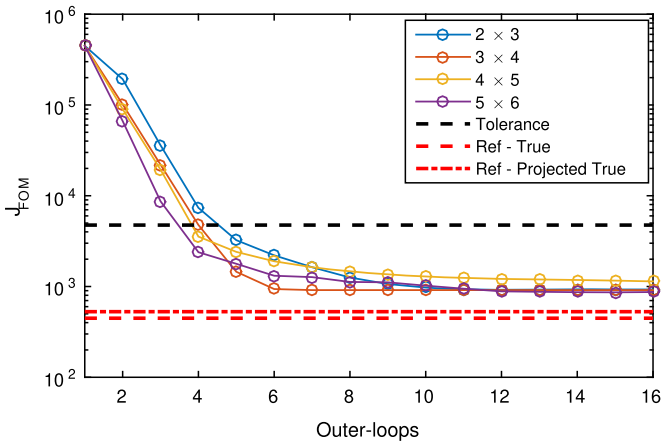


Fig. 18. Evolution of the objective function values of LSPT using different domain decomposition strategies, 2×3, 3×4, 4×5 and 5×6, for Scenario 1.

in Fig. Fig. 20(b), showing the sum of SOE1 and SOE2, reveal that not much further improvement can be obtained if a GP-FD minimization is performed, suggesting that 8 local PCA patterns are indeed nearly sufficient for obtaining an almost identical solution. Note, however, that a much smaller number of FOM simulations is required. We will further investigate to what extent this finding remains valid when assimilating a large number of measurements in the second case-study.

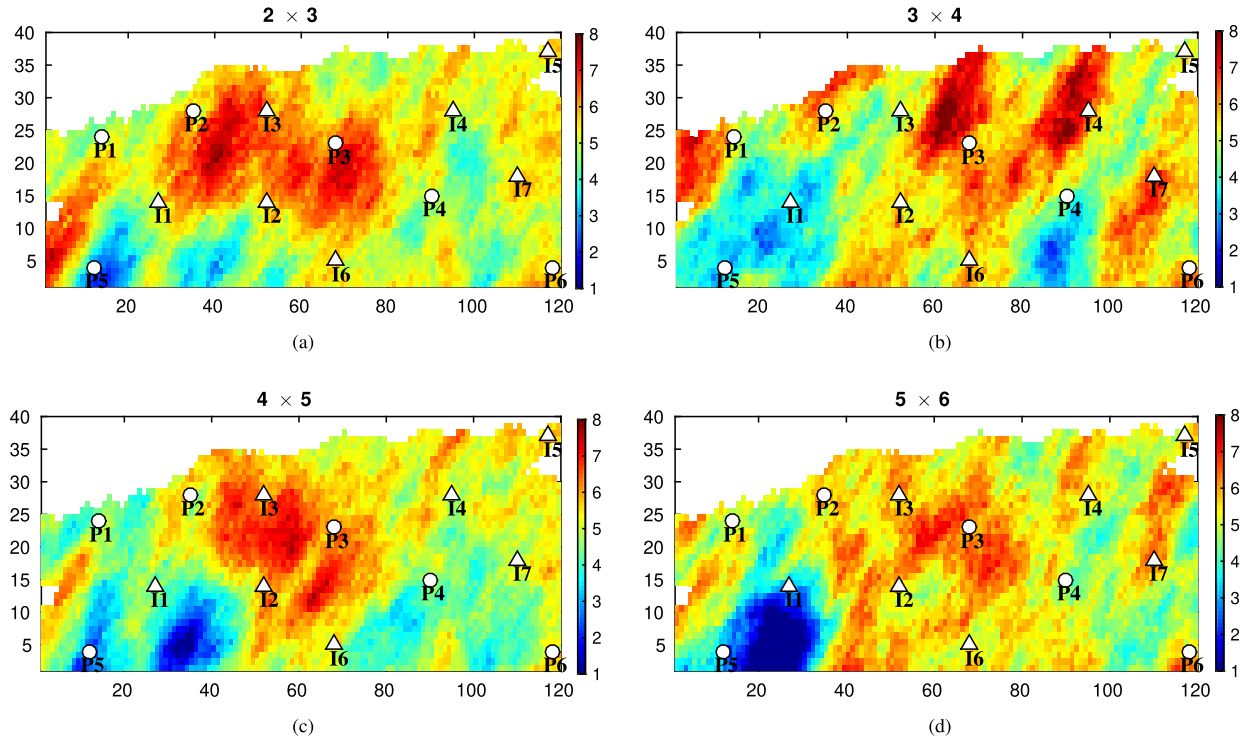


Fig. 19. Updated logarithmic permeability fields using different domain decomposition strategies, e.g., 2×3, 3×4, 4×5 and 5×6, for Scenario 1.

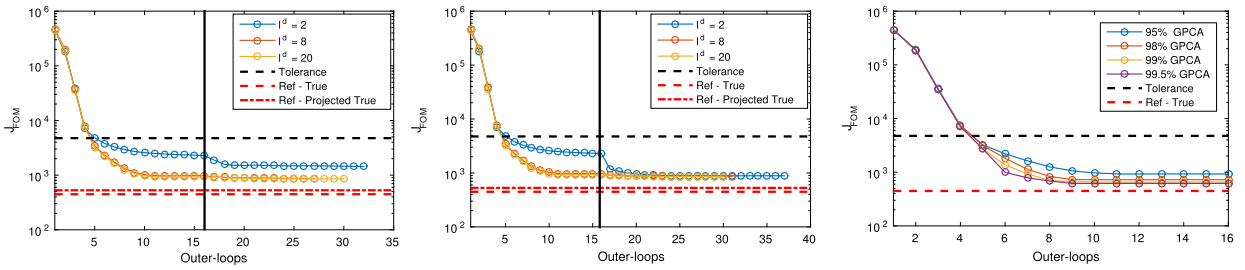


Fig. 20. Evolution of the objective function values when the LSPT and FD are sequentially implemented for Scenario S1. (a) continuing minimization using FD-LP; (b) continuing minimization using FD-GP; (c) the objective function is minimized through preserving more and more global PCA energy, e.g., 95%, 98%, 99% and 99.5%. The vertical black line represents the starting point of minimization using the FD method.

Fig. 20(a) shows that the continued objective function minimization using LP-FD does not significantly decrease the cost function except in the case that only 2 local PCA patterns are retained. This implies that the SOE1 contribution is very small and almost can be ignored, as long as a minimum number of local patterns are retained. SOE2 can be decreased by increasing the number of local PCA patterns, but at the cost of additional FOM simulations.

Fig. 20(c) and Table 8 indicate that SOE3 will gradually decrease with an increasing fraction of retaining (global) energy. Retaining 98% energy is sufficient to accurately represent the original parameter field in this case. An additional increase of 60 global PCA patterns from 48 to 108 requires an additional 10 local PCA patterns in each subdomain. This only requires 20 new FOM simulations.

7.2. Scenario 2

For Scenario 2 with a large number of measurements, three different domain decomposition strategies, i.e., 3×4, 4×5 and 5×6, are formed. The required minimum number of local PCA patterns corresponding to different domain decomposition strategies is summarized in Table 9, and the objective function evolution over the history matching process is shown in Fig. 21. The 4×5 domain decomposition strategy is seen to lead to the smallest objective function value in this case. Fig. 22 shows that the true parameter field can be reconstructed very accurately when a large number of measurements is available.

Table 7
The number of FOM simulations and the final objective function values when quantifying SOE1 and SOE2 for Scenario S1. A same number of local PCA patterns l_d is retained among all subdomains. A decomposition strategy with 2×3 subdomains is used.

Objective function				$J_{FOM}(\xi)$		
Initial model				1.01×10^5		
Tolerance				4750		
Ref - Projected True				528.1		
Ref - True				447.4		
l_d	N_L	N_G	Number of FOM simulations	LSPT	LP-FD	GP-FD
2	12		53 = 22+(8×2+1)+14	2276.15	1441	886.3
8	48	48	53 = 22+(2×8+1)+14	902.48	869.24	860.24
20	120		77 = 22+(2×20+1)+14	892.21	854.69	860.24

Table 8
The number of FOM simulations and the final objective function values when quantifying SOE3 for Scenario S1. A same number of local PCA patterns l_d is retained among all subdomains. The first column represents the preserved energy of global PCA. A decomposition strategy with 2×3 subdomains is used.

-	N_L	l_d	N_G	Iterations	Number of FOM simulations	$J_{FOM}(\xi)$
Initial model	–	–	–	–	–	1.01×10^5
95%	48	8	48	15	53 = 22+(2×8+1)+14	902.48
98%	72	12	72	15	61 = 22+(2×12+1)+14	738.25
99%	96	16	92	15	69 = 22+(2×16+1)+14	694.18
99.5%	108	18	104	15	73 = 22+(2×18+1)+14	621.52
Tolerance	–	–	–	–	–	4750
Ref - Projected True	–	–	–	–	–	528.1
Ref - True	–	–	–	–	–	447.4

Table 9
The number of FOM simulations and the final objective function values of LSPT using different domain decomposition strategies, e.g., 3×4, 4×5 and 5×6, for Scenario 2. A same number of local PCA patterns l_d is retained among all subdomains.

–	l_d	N_L	N_G	Iterations	Number of FOM simulations	$J_{FOM}(\xi)$
Initial model	–	–	–	–	–	6.39×10^4
3×4	24	288		10	90 = 32+(2×24+1)+9	7508
4×5	15	300	282	10	72 = 32+(2×15+1)+9	6783
5×6	10	300		10	62 = 32+(2×10+1)+9	9601
Tolerance	–	–	–	–	–	2.75×10^4
GP-FD	–	–	–	–	–	6416
Projected 'True' model	–	–	–	–	–	5685
'True' model	–	–	–	–	–	5149

Only the 5×6 decomposition leads to a poor spatial parameters reconstruction, consistent with a relatively high objective function value. Fig. 23 and Fig. 24 show the predicted saturation and its corresponding grid-based RMSE values. Compared to the initial model, the model predictions have been significantly improved.

Fig. 25 shows the simulated and measured well data up to the 10-year history period, and simulated data for an additional 15-year prediction period. Results are shown for the initial model and for the estimated models for the three different domain decomposition strategies, as well as for the GP-FD solution. The predictions of especially the fluid rate and bottom-hole pressure based on the initial model are quite poor. After the history matching, the predictions of all the updated models are consistent with the measurements.

Compared to Scenario 1, the number of global PCA patterns has been increased from 48 to 282, however, taking the 4×5 decomposition as an example, the required number of FOM simulations has only increased from 53 to 72. The degree of freedom for the history matching problem depends on the number of global PCA patterns, while the required FOM simulation depends on the number of local PCA patterns. It is therefore very attractive to increase the degree of the freedom by adding local PCA patterns in all subdomains. Taking the 5×6 domain decomposition scheme as an example, adding one local PCA pattern in each subdomain allows us to retain another 30 global PCA patterns, while only 2 more FOM simulations are added to the entire history matching procedure. These numerical results further demonstrate that introducing smooth local parameterization makes subdomain POD-TPWL highly scalable, the required number of FOM simulations does not grow rapidly with an increasing number of space varying parameters.

7.3. Computational complexity

The computational cost of the proposed parameter estimation approach can be split into two main parts. The cost of the offline stage consists of constructing the subdomain reduced-order linear model. The cost of the online stage consists of the

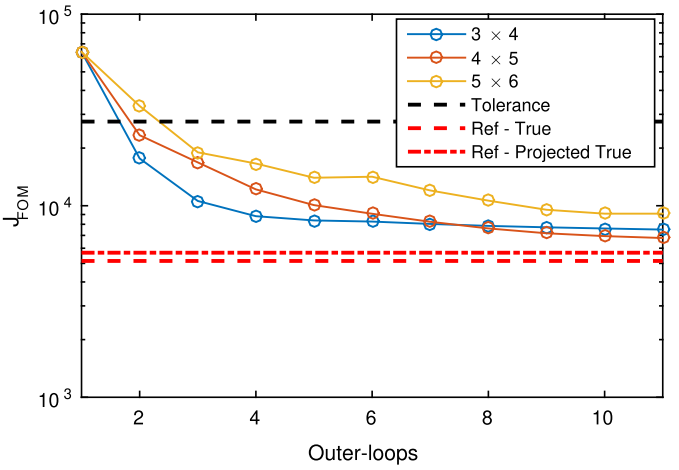


Fig. 21. Evolution of the objective function values as a function of outer-loops using LSPT for Scenario 2. The computation of the objective function for the outer-loops uses full-order model simulations. The two red dash-lines separately represent reference objective function values for the true model (reflecting the impact of the data noise) and the projected true model (that is the best possible reconstruction of the truth given the selected PCA patterns).

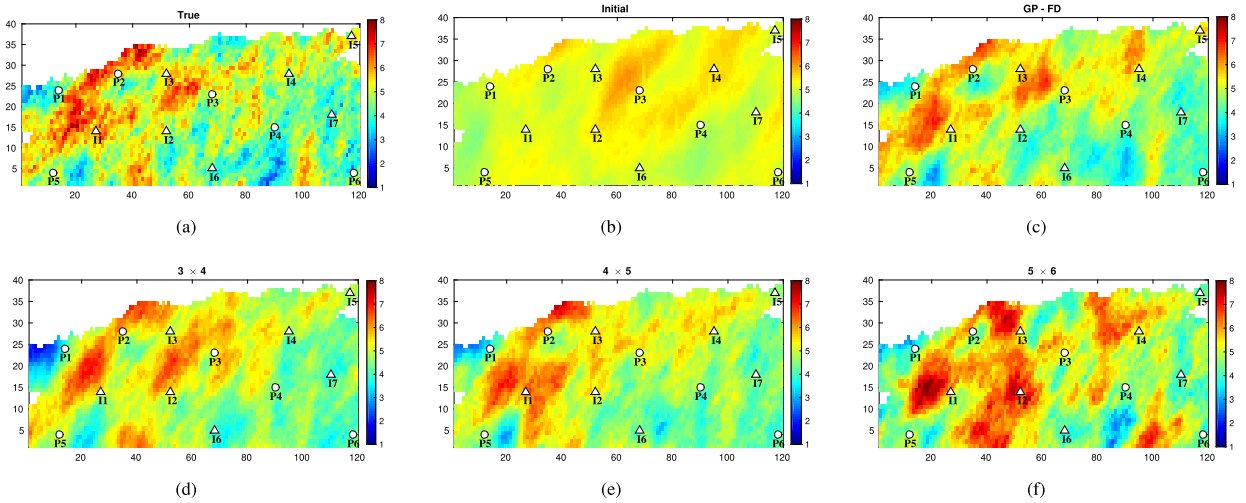


Fig. 22. Comparison of updated logarithmic permeability fields using LSPT and GP-FD for Scenario 2. The true model, initial model and updated models corresponding three different domain decomposition strategies, e.g., 3×4, 4×5 and 5×6, are displayed.

cost of solving the reduced system and the parameter estimation problem. We will now discuss these two stages in more detail below.

7.3.1. Offline stage

The cost of executing parameterization using eigenvalue decomposition of the global covariance matrix and local covariance matrix in each subdomain is negligible for small models, while it will become more significant for large-scale models. The approximate computational complexity is of order $O(N_\beta^3)$. An equivalent formulation can be derived in which the eigenvalue problem is formulated in the snapshots coordinate $\hat{\mathbf{C}}$ by $\frac{\mathbf{X}_c^T\mathbf{X}_c}{N_r-1}$. It is a so-called method of snapshots [43]. An alternative way to calculate PCA patterns is to perform a singular value decomposition (SVD) on the matrix $\mathbf{X}_c$. It can avoid the calculation of the covariance matrix $\mathbf{C}$ or $\hat{\mathbf{C}}$. The costs of these two efficient alternatives are proportional to $O(N_r^3)$ and $O(N_\beta N_r^2)$, respectively. Both of them have a numerical advantage, because typically $N_r \ll N_\beta$.

Generating snapshots is an essential part of the POD method. The actual generation of the snapshots is done by sampling an ensemble of parameter realizations around an initial mean field. For each member in the newly generated ensemble the FOM is simulated, and the values for the state variables at each time step are saved. The computational cost of performing this part of the process, expressed in number of FOM runs is equal to the number of members in the generated ensemble, namely $O(FOM)$. Here $O(FOM)$ denotes the computational complexity for one full-order model simulation, which is associated with the model dimension N_β, the number of simulation time steps N, and the efficiency of the numerical solver of

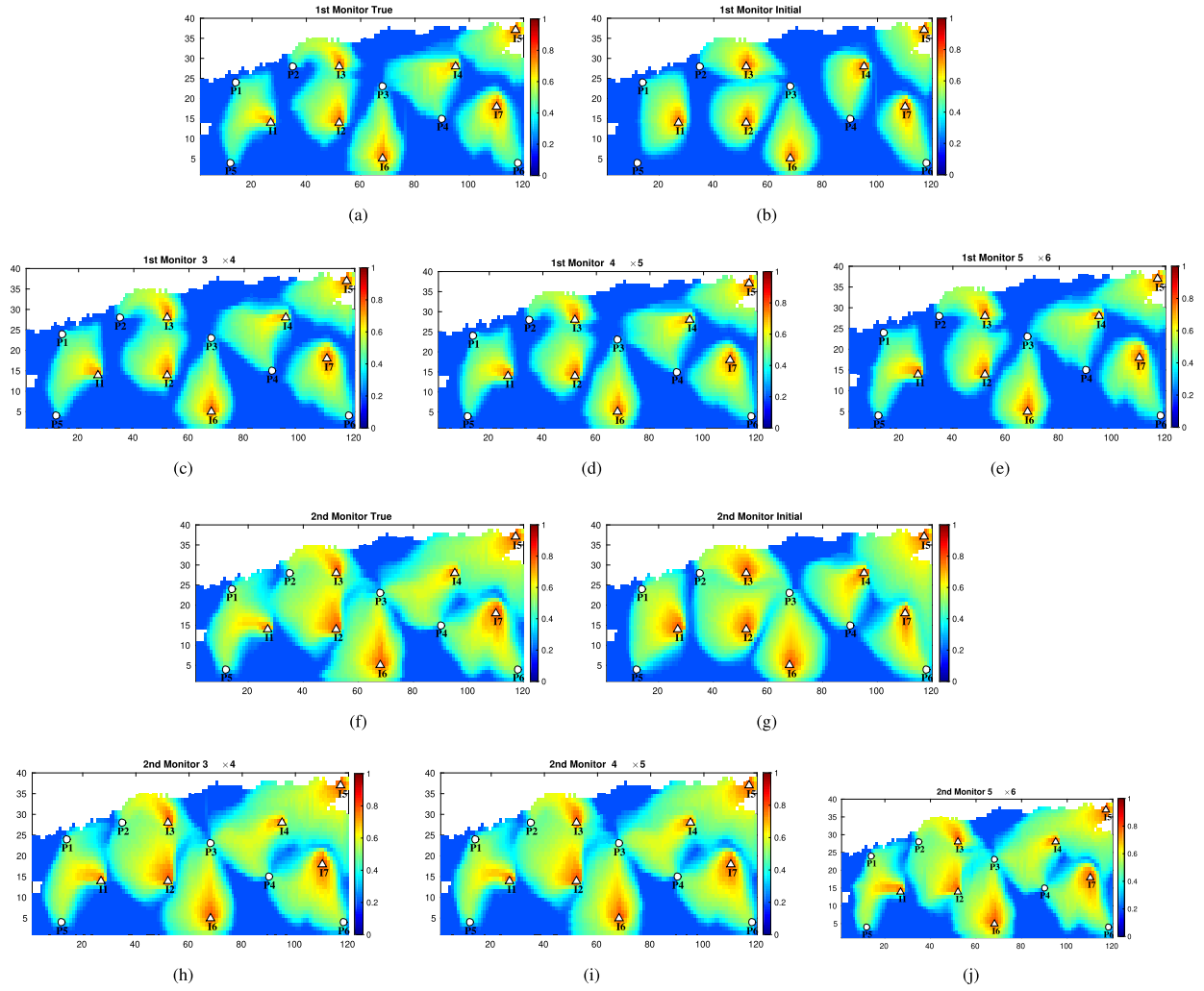


Fig. 23. Predictions of water saturation using LSPT for Scenario 2. The subfigures (a) - (d) present the results at the 1st monitor, subfigures (e) - (h) present the results at the 2nd monitor. Large errors occur at the water fronts. Three domain decomposition strategies, e.g., 3×4, 4×5 and 5×6, are conducted.

the forward modeling code. Taking $Newton - Raphson$ iteration as an example, the corresponding computational complexity is $O(\log_2(N_\beta)N)$ [49].

The cost of solving the reduced eigenvalue problem to construct the POD is equivalent to the cost of a Singular Value Decomposition of the snapshot matrix. Since the dimension of the snapshot matrix in each subdomain is relatively low, this cost is also low.

The cost of approximating derivatives using the RBF interpolation represents the most computationally expensive part of constructing the subdomain reduced-order linear model. The computational time expressed in number of FOM runs is several times the number of local PCA patterns generated, or $O(\log_2(N_\beta)N)$.

7.3.2. Online stage

The cost of solving a system of model-reduced linear equations can be neglected in comparison with all other contributions.

The cost of the model-reduced optimization procedure is proportional to the number of times that a new subdomain reduced-order linear models is constructed which requires one FOM run with a cost of $O(\log_2(N_\beta)N)$.

In short, the total computational cost in terms of order analysis is $(O(N_\beta^3)+O(\log_2(N_\beta)N))$. The process is code non-intrusive and does not involve overwhelming programming efforts. As the adjoint model is not always available, especially for commercial simulators, the finite-difference method can be used to approximate the gradient for use in an objective function minimization procedure. In that case, $O(10^3 - 10^4)$ FOM simulations will typically be required for large-scale parameter estimation problems. An offline cost for our proposed approach of $O(10 - 10^2)$ FOM simulations is therefore

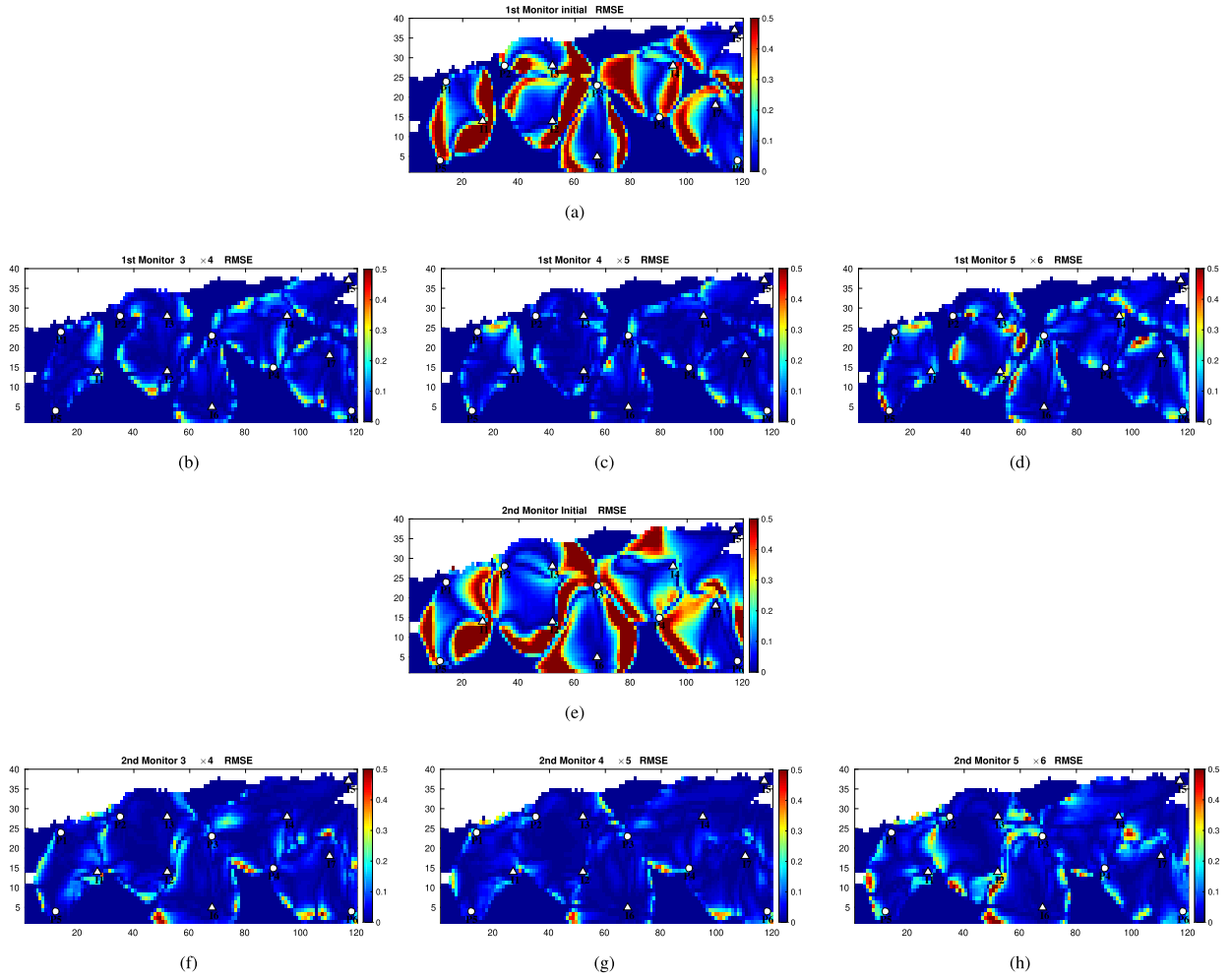


Fig. 24. The RMSE distribution of the predicted water saturation in all gridblocks using LSPT for Scenario 2. The subfigures (a) - (d) present the results at the 1st monitor, subfigures (e) - (h) present the results at the 2nd monitor. Large errors occur at the water fronts. Three domain decomposition strategies, e.g., 3×4, 4×5 and 5×6, are conducted.

a significant improvement. For large-scale parameter estimation problems, the computational cost is dominated by the required FOM simulations. In our proposed method, most of the FOM simulations are performed in the offline stage.

Fig. 26 summarizes the required FOM simulations as a function of number of subdomains in this study. These results indicate computational efficiency can benefit from increasing the number of subdomains. On the other hand, the numerical results presented in Fig. 18 indicate that the quality of the parameter field estimate may deteriorate if too small subdomains are formed. It is therefore important to find an appropriate trade-off in efficiency and accuracy by optimizing the domain decomposition strategy.

8. Conclusions

We have introduced a new method for large-scale parameter estimation based on a non-intrusive reduced-order modeling approach: Subdomain POD-TPWL with smooth local parameterization. A combination of Principal Component Analysis and Domain Decomposition is used to decompose the high-dimensional solution space for the spatial parameter field into lower-order parameter subspaces associated with the individual subdomains. The optimized local parameters are projected onto a global parameterization to eliminate the non-smoothness at the boundaries of neighboring subdomains. The local parameterization allows us to run only a small number of full-order model simulations by simultaneously perturbing the parameters in all subdomains. The use of smooth local parameterization enables the use of subdomain POD-TPWL to large-scale problems since the number of full-order model simulations depends primarily on the number of local parameters in each subdomain.

The approach is tested using a 2D reservoir model for which two experimental scenarios have been developed. In the first scenario the parameter field is characterized by relatively large spatial scales, while the number of measurements is fairly

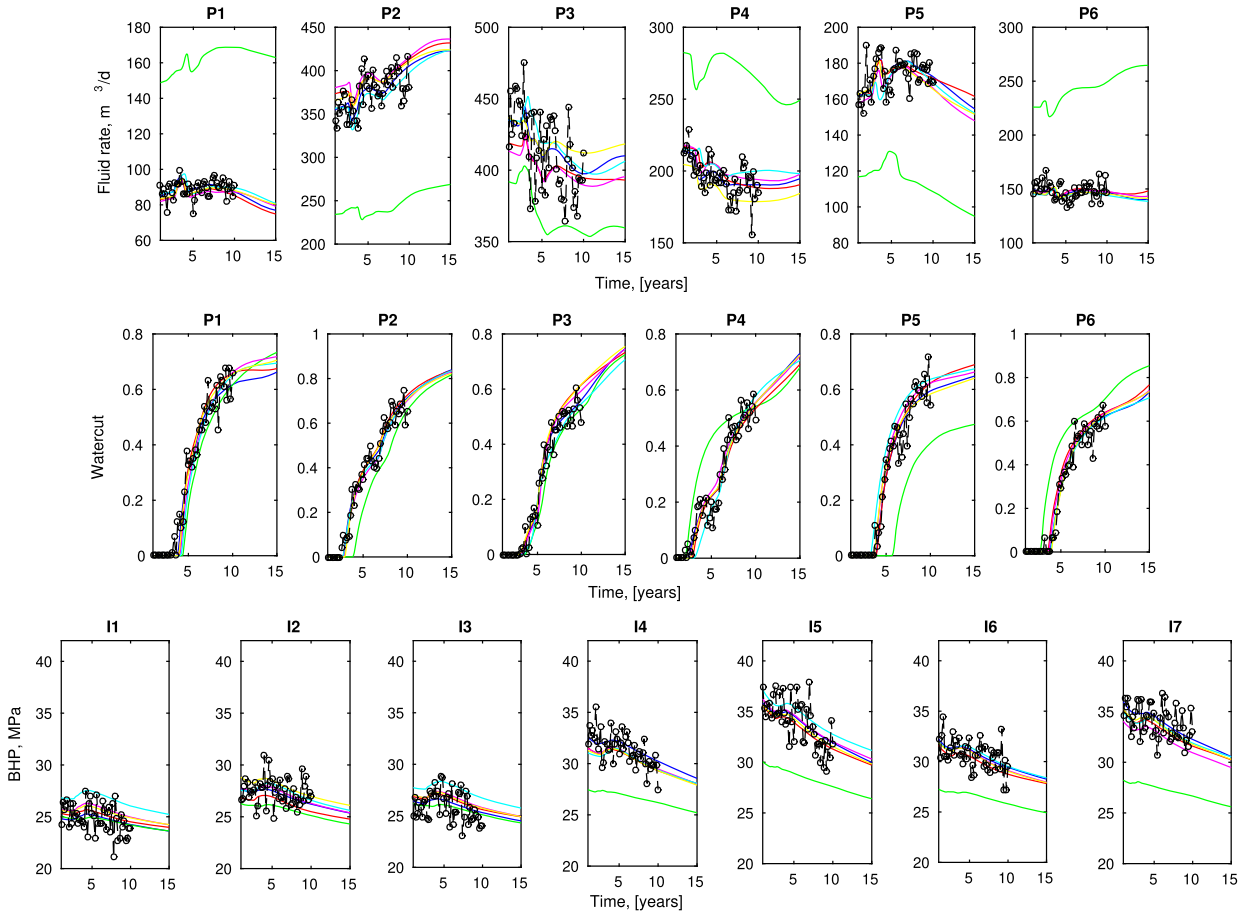


Fig. 25. Forecast of the producers' liquid rate, WCT and injectors' BHP for Scenario 2: green line - initial model, blue line - 'true' model, red line - LSPT with 3×4 domain decomposition, magenta line-LSPT with 4×5 domain decomposition, cyan line - LSPT with 5×6 domain decomposition, yellow line - GP-FD.

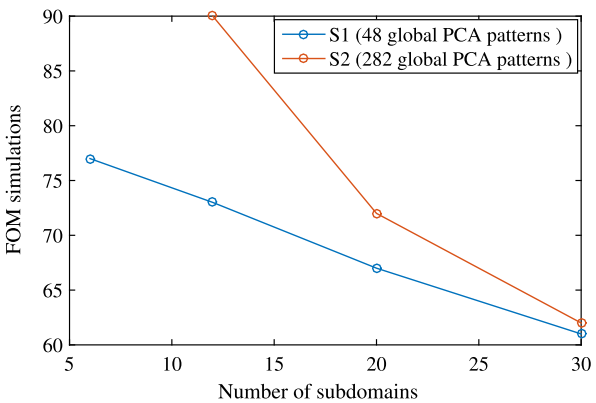


Fig. 26. Summary of the required FOM simulations for Scenario 1 and Scenario 2 with respect to the number of subdomains.

small. In the second scenario spatial scales are much smaller and the number of measurements is much larger. The first numerical experiment results show that subdomain POD-TPWL with global and local parameterization is able to reconstruct feasible solutions for the parameter field, resulting in acceptable data mismatches, and produces comparable results as obtained with finite-difference gradients, but at much lower computational cost. The results of the second example show that also for a more complex problem where data available at all model gridblocks was used to calibrate 282 uncertain parameters, good results could be obtained.

The number of full-order model runs can be reduced by using a larger number of subdomains in which a smaller number of local parameters is required. For the cases studied in this paper, the number of full-order model simulations was roughly 2 times the maximum number of local parameter patterns among all subdomains. The method has therefore very attractive scalability properties. For a model with a much larger domain and when using the same size of the subdomains the number of full-order model simulations will hardly increase.

The proposed methodology could be further improved. All examples have shown that the choice of domain decomposition strategy has an impact on the model updating and the minimization of the cost function. We have chosen somewhat arbitrary decompositions of the global domain into rectangular subdomains. It may be beneficial to choose the subdomains based on information about either the scales of variability of the parameter field or of the dynamical patterns. In this study the distribution of both sparse well data and grid-based data was almost uniform over all subdomains. In that case an uniform choice of the number of local PCA patterns in each subdomain is appropriate. If data are irregularly distributed in space the subdomain could be chosen based on the amount of information in the subdomain, and the number of local parameters in each subdomain could possibly be informed by the number of available measurements.

CRediT authorship contribution statement

Cong Xiao: Conceptualization, Methodology, Software, Writing - Original draft preparation. **Olwijn Leeuwenburgh:** geological model generation, model design, Writing - Original draft preparation. **Arnold Heemink and Hai-Xiang Lin:** Investigation, Supervision and Writing - Original draft preparation. **Arnold Heemink and Hai-Xiang Lin, Olwijn Leeuwenburgh:** Writing - Reviewing and Editing.

Declaration of competing interest

No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication. I would like to declare on behalf of my co-authors that the work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part.

Acknowledgement

The first author would like to thank the China Scholarship Council (CSC: 201606440057) for providing research funding. Additional Computing resources were provided by the Mathematics Physics Group, Department of Applied Mathematics at Delft University of Technology. The use of open source codes MRST (https://www.sintef.no/projectweb/mrst/) is gratefully acknowledged.

References

[1] D.R. Brouwer, J.D. Jansen, Dynamic optimization of water flooding with smart wells using optimal control theory, in: European Petroleum Conference, Society of Petroleum Engineers, 2002.
[2] P. Sarma, L.J. Durlofsky, K. Aziz, W.H. Chen, Efficient real-time reservoir management using adjoint-based optimal control and model updating, Comput. Geosci. 10 (1) (2006) 3–36.
[3] S. Vlemmix, G.J. Joosten, R. Brouwer, J.D. Jansen, Adjoint-based well trajectory optimization, in: EUROPEC/EAGE Conference and Exhibition, Society of Petroleum Engineers, 2009.
[4] J.D. Jansen, Adjoint-based optimization of multi-phase flow through porous media–a review, Comput. Fluids 46 (1) (2011) 40–51.
[5] P.H. Courtier, J.N. Thépaut, A. Hollingsworth, A strategy for operational implementation of 4D-Var, using an incremental approach, Q. J. R. Meteorol. Soc. 120 (519) (1994) 1367–1387.
[6] C. Liu, Q. Xiao, B. Wang, An ensemble-based four-dimensional variational data assimilation scheme. Part I: technical formulation and preliminary test, Mon. Weather Rev. 136 (9) (2008) 3363–3373.
[7] C. Liu, Q. Xiao, B. Wang, An ensemble-based four-dimensional variational data assimilation scheme. Part II: observing system simulation experiments with Advanced Research WRF (ARW), Mon. Weather Rev. 137 (5) (2009) 1687–1704.
[8] D. Xiao, F. Fang, A.G. Buchan, C.C. Pain, I.M. Navon, J. Du, G. Hu, Non-linear model reduction for the Navier–Stokes equations using residual deim method, J. Comput. Phys. 263 (2014) 1–18.
[9] D. Xiao, P. Yang, F. Fang, J. Xiang, C.C. Pain, I.M. Navon, Non-intrusive reduced order modelling of fluid–structure interactions, Comput. Methods Appl. Mech. Eng. 303 (2016) 35–54.
[10] S. Hijazi, G. Stabile, A. Mola, G. Rozza, Data-driven POD-Galerkin reduced order model for turbulent flows, arXiv preprint, arXiv:1907.09909, 2019.
[11] T. Heijn, R. Markovinovic, J.D. Jansen, Generation of low-order reservoir models using system-theoretical concepts, in: SPE Reservoir Simulation Symposium, Society of Petroleum Engineers, 2003.
[12] R. Markovinović, J.D. Jansen, Accelerating iterative solution methods using reduced-order models as solution predictors, Int. J. Numer. Methods Eng. 68 (5) (2006) 525–541.
[13] M.A. Cardoso, L.J. Durlofsky, P. Sarma, Development and application of reduced-order modeling procedures for subsurface flow simulation, Int. J. Numer. Methods Eng. 77 (9) (2009) 1322–1350.
[14] F. Fang, T. Zhang, D. Pavlidis, C.C. Pain, A.G. Buchan, I.M. Navon, Reduced order modelling of an unstructured mesh air pollution model and application in 2D/3D urban street canyons, Atmos. Environ. 96 (2014) 96–106.
[15] A. Manzoni, F. Salmoiraghi, L. Heltai, Reduced basis isogeometric methods (RB-IGA) for the real-time simulation of potential flows about parametrized NACA airfoils, Comput. Methods Appl. Mech. Eng. 284 (2015) 1147–1180.
[16] M.A. Cardoso, L.J. Durlofsky, Linearized reduced-order models for subsurface flow simulation, J. Comput. Phys. 229 (3) (2010) 681–700.
[17] M. Rewienski, J. White, A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. 22 (2) (2003) 155–170.

[18] J.C. He, P. Sarma, L.J. Durlofsky, Reduced-order flow modeling and geological parameterization for ensemble-based data assimilation, Comput. Geosci. 55 (2013) 54–69.
[19] J.C. He, L.J. Durlofsky, Reduced-order modeling for compositional simulation by use of trajectory piecewise linearization, SPE J. 19 (05) (2014) 858–872.
[20] J.C. He, J. Sætrom, L.J. Durlofsky, Enhanced linearized reduced-order models for subsurface flow simulation, J. Comput. Phys. 230 (23) (2011) 8313–8341.
[21] S. Trehan, L.J. Durlofsky, Trajectory piecewise quadratic reduced-order model for subsurface flow, with application to PDE-constrained optimization, J. Comput. Phys. 326 (2016) 446–473.
[22] P.T.M. Vermeulen, A.W. Heemink, Model-reduced variational data assimilation, Mon. Weather Rev. 134 (10) (2006) 2888–2899.
[23] M.U. Altaf, A.W. Heemink, M. Verlaan, Inverse shallow-water flow modeling using model reduction, Int. J. Multiscale Comput. Eng. 7 (6) (2009).
[24] M.P. Kaleta, R.G. Hanea, A.W. Heemink, J.D. Jansen, Model-reduced gradient-based history matching, Comput. Geosci. 15 (1) (2011) 135–153.
[25] D.J. Lucia, P.I. King, P.S. Beran, Reduced order modeling of a two-dimensional flow with moving shocks, Comput. Fluids 32 (7) (2003) 917–938.
[26] D.J. Lucia, P.I. King, P.S. Beran, Domain decomposition for reduced-order modeling of a flow with moving shocks, AIAA J. 40 (11) (2002) 2360–2362.
[27] J. Baiges, R. Codina, S. Idelsohn, A domain decomposition strategy for reduced order models. Application to the incompressible Navier–Stokes equations, Comput. Methods Appl. Mech. Eng. 267 (2013) 23–42.
[28] P. Kerfriden, O. Goury, T. Rabczuk, S.P. Bordas, A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics, Comput. Methods Appl. Mech. Eng. 256 (2013) 169–188.
[29] D. Amsallem, M.J. Zahr, C. Farhat, Nonlinear model order reduction based on local reduced-order bases, Int. J. Numer. Methods Eng. 92 (10) (2012) 891–916.
[30] P.P. Chinchapatnam, K. Djidjeli, P.B. Nair, Domain decomposition for time-dependent problems using radial based meshless methods, Numer. Methods Partial Differ. Equ. 23 (1) (2007) 38–59.
[31] H. Antil, M. Heinkenschloss, R.H. Hoppe, D.C. Sorensen, Domain decomposition and model reduction for the numerical solution of PDE constrained optimization problems with localized optimization variables, Comput. Vis. Sci. 13 (6) (2010) 249–264.
[32] H. Antil, M. Heinkenschloss, R.H. Hoppe, Domain decomposition and balanced truncation model reduction for shape optimization of the Stokes system, Optim. Methods Softw. 26 (4–5) (2011) 643–669.
[33] D. Xiao, C.E. Heaney, F. Fang, L. Mottet, R. Hu, D.A. Bistrian, E. Aristodemou, I.M. Navon, C.C. Pain, A domain decomposition non-intrusive reduced order model for turbulent flows, Comput. Fluids 182 (2019) 15–27.
[34] C. Xiao, O. Leeuwenburgh, H.X. Lin, A. Heemink, Non-intrusive subdomain POD-TPWL for reservoir history matching, Comput. Geosci. 23 (3) (2019) 537–565.
[35] G.B. Wright, Radial Basis Function Interpolation: Numerical and Analytical Developments, 2003.
[36] C. Chen, G. Gao, J. Honorio, P. Gelderblom, E. Jimenez, T. Jaakkola, Integration of principal-component-analysis and streamline information for the history matching of channelized reservoirs, in: Spe Technical Conference and Exhibition, 2014.
[37] P. Sarma, L.J. Durlofsky, K. Aziz, Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics, Math. Geosci. 40 (1) (2008) 3–32.
[38] B. Jafarpour, D.B. McLaughlin, History matching with an ensemble Kalman filter and discrete cosine parameterization, Comput. Geosci. 12 (2) (2008) 227–244.
[39] Y. Chen, D.S. Oliver, Multiscale parameterization with adaptive regularization for improved assimilation of nonlocal observation, Water Resour. Res. 48 (48) (2012) 4503.
[40] F. Gassmann, Elastic waves through a packing of spheres, Geophysics 16 (4) (1951) 673–685.
[41] J. He, L.J. Durlofsky, Constraint reduction procedures for reduced-order subsurface flow models based on POD-TPWL, Int. J. Numer. Methods Eng. 103 (1) (2015) 1–30.
[42] K. Fukunaga, W.L. Koontz, Application of the Karhunen-Loeve expansion to feature selection and ordering, IEEE Trans. Comput. 100 (4) (1970) 311–318.
[43] L. Sirovich, Turbulence and the dynamics of coherent structures, Parts I-III, Q. Appl. Math. (1987) 561–582.
[44] GSLIB: geostatistical software library and user's guide, Technometrics (1995).
[45] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, 1999.
[46] D.S. Oliver, A.C. Reynolds, N. Liu, Inverse Theory for Petroleum Reservoir Characterization and History Matching, Cambridge University Press, 2008.
[47] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, vol. 89, Siam, 2005.
[48] K.A. Lie, S. Krogstad, I.S. Ligaarden, J.R. Natvig, H.M. Nilsen, B. Skaflestad, Open-source MATLAB implementation of consistent discretisations on complex grids, Comput. Geosci. 16 (2) (2012) 297–322.
[49] M.A. VVeron, C.U. Rjoja, Newton-Raphson's method and convexity, Univ. Novom Sadu Zb. Rad. Period.-Mat. Fak. Ser. Mat. 22 (1992) 159–166.