Surrogate-Based Joint Estimation of Subsurface Geological and Relative Permeability Parameters for High-Dimensional Inverse Problem by Use of Smooth Local Parameterization

Xiao, Cong; Tian, Leng

DOI
10.1029/2019WR025366

Publication date
2020

Document Version
Final published version

Published in
Water Resources Research

AGU ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1029/2019WR025366

Key Points:
- Efficient subdomain linear models with smooth local parameterization are produced as an alternative implementation of forward simulation
- Only 220 high-fidelity model simulations are required in a high-dimensional inversion problem with 320 uncertain parameters
- Ensemble collapse problem could be effectively and efficiently mitigated on conditioning a large amount of measurements

Supporting Information:
- Supporting Information S1

Correspondence to:
C. Xiao,
C.Xiao@tudelft.nl

Citation:
Xiao, C., & Tian, L. (2020). Surrogate-based joint estimation of subsurface geological and relative permeability parameters for high-dimensional inverse problem by use of smooth local parameterization. *Water Resources Research*, *56*, e2019WR025366. https://doi.org/10.1029/2019WR025366

Received 15 APR 2019
Accepted 28 APR 2020
Accepted article online 13 MAY 2020

Surrogate-Based Joint Estimation of Subsurface Geological and Relative Permeability Parameters for High-Dimensional Inverse Problem by Use of Smooth Local Parameterization

Cong Xiao[1] **and Leng Tian**[2]

[1]Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg, Netherlands, [2]College of Petroleum Engineering, China University of Petroleum, Beijing, China

Abstract This paper introduces an efficient surrogate model with the aim of accelerating joint estimation of subsurface geological properties and relative permeability parameters for high-dimensional inversion problems. We fully replace the high-fidelity model with a set of subdomain linear models through integrating model linearization with smooth local parameterization where the Gaussian geological parameters and non-Gaussian facies indicators are locally parameterized. These subdomain linear models with smooth local parameterization, referred to as SLM-SLP, are constructed in each subdomain individually using only a few high-fidelity model simulations. An adaptive scheme, that is, weighting smooth local parameterization (WSLP), is introduced as well to mitigate the negative effects of inappropriate domain decomposition schemes by adaptively optimizing the domain decomposition strategy. The computational advantages of the proposed algorithm are demonstrated on a synthetic non-Gaussian facies model and a real-world high-dimensional Gaussian model. The amount of computational cost has been drastically reduced while reasonable accuracy remains. Specifically, SLM-SLP only needs 220 fidelity simulations to optimize 302 parameters. Compared to ensemble smoother with multiple data assimilation (ES-MDA), SLM-SLP effectively and efficiently mitigates the ensemble collapse problem in the course of uncertain quantification.

1. Introduction

Typically considered for subsurface model calibration during inversion problems has been geological hydraulic conductivity (Li et al., 2012). The interest has subsequently shifted toward other parameters, with the relative permeability curves being a popular choice (Chen & Oliver, 2010; Li et al., 2012; Seiler et al., 2009). It has been universally acknowledged that computational burden has become one of the most severe challenges in inverse modeling. Replacing the high-fidelity model (HFM) with a cheap-to-evaluate surrogate model is a popular approach to reduce the computational cost. Surrogate modeling aims to provide a simpler and hence faster model, which simplifies the relationship between inputs and outputs. Many efforts have been taken to make the constructions of surrogate models efficiently feasible. To our currently best knowledge, the existing surrogate modeling approaches can be roughly categorized into two categories: projection-based reduced order models (ROMs) and data driven. Reviews of surrogate modeling in water resources have been given by Asher et al. (2015) and Razavi et al. (2012).

In projection-based ROMs, a set of reduced basis is extracted from the simulation snapshots using an unsupervised learning technique, for example, proper orthogonal decomposition (POD) (Altaf et al., 2009; Vermeulen & Heemink, 2006) and the full-order model operator is projected onto the subspace spanned by the reduced basis, which can significantly reduce the degrees of freedom of the system (Kaleta et al., 2011; Xiao et al., 2018). In order to overcome the computational burden associated with the application of Monte Carlo methods to the groundwater flow equation with random hydraulic conductivity, Pasetto et al. (2013) present a model-order reduction technique that the high-dimensional model equations are projected onto subspace based on the Galerkin projection. Zha et al. (2018) presented a reduced-order successive linear estimator (ROSLE) for analyzing hydraulic tomography data. This new linear estimator approximated the covariance of the unknowns using Karhunen-Loeve expansion (KLE) truncated to finite order, and it calculated the directional sensitivities to form the covariance and cross-covariance used in the course of

parameter estimation. The numerical experiments have shown the computational advantage of this new algorithm. Although previous results demonstrate some promises, the projection-based ROM methods have restricted their applications on complex dynamic problems mainly due to the inherent stability and robustness problems (Baur et al., 2014). Besides, projection-based ROMs are highly code intrusive, and their speedup potential is restricted by the strong nonlinearity (Amsallem & Farhat, 2008).

Data driven is an another way to enable rapid simulations, where an emulator of the system is constructed upon the HFM solutions of selected collocation points in parameter space without the need to modify the codes. Specifically, some regression techniques, for example, radial basis functions (Piret et al., 2019), Gaussian process regression (GPR) (Yang et al., 2018), and stochastic polynomial chaos expansion (PCE) (Hu et al., 2019), are employed to learn a deterministic or probabilistic input-output mapping. Because of the nonintrusive feature and ease of implementation, many different data-driven surrogates have been developed for inversion problems of the hydrological system (Dai et al., 2016; Li & Zhang, 2007; Saad & Ghanem, 2009; Zeng & Zhang, 2010). For example, Zeng et al. (2012) constructed surrogates using generalized PCE to speedup posterior exploration of groundwater models. Chang et al. (2017) implemented a surrogate-based iterative ensemble smoother through developing both the PCE and the interpolation-based surrogate models in subsurface flow data assimilation problems. GPR has been increasing popularity as well in terms of predicting the tracer transport in groundwater models (Asher et al., 2015; Hombal & Mahadevan, 2011; Lu et al., 2018; Roy et al., 2018).

To further mitigate the computational burden and approximate errors, in the recent years many strategies have been introduced to enable efficient data-driven model reconstruction, for example, compressed sensing, adaptive and/or multilevel, and multifidelity strategies (Gong et al., 2016; Ju et al., 2018; Laloy et al., 2013; Mo et al., 2017; Zhang et al., 2017, 2018, 2020; Zhou et al., 2018). For example, Adam et al. (2020) incorporate a TSVD (truncated singular value decomposition)-based dimensionality reduction method to reduce the number of variables and thereby decrease the HFM runs needed in GPR surrogate. In order to maintain accuracy of the surrogate model, a novel adaptive resampling through particle swarm optimization is designed throughout the optimization process. Ju et al. (2018) established an adaptive GPR-based iterative ensemble smoother. The GPR surrogate adaptively adds a few additional sample points chosen from the previously updated parameter realizations at each iteration step. The cross covariance between uncertain parameters and measurements could be easily computed by use of the refined GPR surrogate with a negligible computational cost. Gong et al. (2016) proposed a multiobjective adaptive surrogate modeling-based optimization (MO-ASMO) algorithm that aims to reduce computational cost while maintaining optimization effectiveness. Laloy et al. (2013) designed a two-level MCMC strategy where the surrogate model was used to sufficiently explored the parameter space at Level 1, and then a limited number of HFMs are evaluated for correction at level two. Rather than multilevel strategy, Zhang et al. (2018) propose an adaptive multifidelity MCMC algorithm for efficient inverse modeling of hydrologic systems. In this method, the HFMs are evaluated mainly in the posterior region through iteratively running MCMC based on a Gaussian process system that is adaptively constructed with multifidelity simulation. Subsequently, Zhang et al. (2020) design two strategies to quantify the surrogate error and then integrate them with a surrogate-based Bayesian inversion framework that explicitly quantifies and gradually reduces the approximation error of the surrogate.

Although the commonly used surrogate models, for example, PCE, GPR, and their improved variants with adaptive or multifidelity strategies, have been successfully applied to inverse modeling of subsurface hydrological system, the required number of training models will be exponentially increased, a phenomenon often referred to as "curse of dimensionality" (Feeley, 2008). In contrast, linear surrogate model is easily constructed particularly for the real-world applications. In addition, dimensionality reduction has been proved to be an effective supplementary to surrogate model with the aim of further decreasing the computational cost. Several parameterization methods have been applied to subsurface flow inversion problems, including principle component analysis (PCA) (Chen et al., 2014), discrete cosine transformation (DCT) (Jafarpour & McLaughlin, 2008), and discrete wavelet transformation (DWT) (Chen & Oliver, 2012). From a computational point of view, a local parameterization where the parameters are separately defined in each subdomain is very attractive (Baiges et al., 2013). Since in this case, we can mainly focus on each subdomain individually with only a few HFM simulations (Ding, 2011; Ding & Mckee, 2013). This paper introduced a smooth local parameterization (SLP) by combining PCA and domain decomposition (DD) to separately

represent the spatial parameter field for each subdomain. In addition, an extension of this SLP to non-Gaussian models is explored as well.

To best our knowledge, the combination of SLP and linear surrogate model and resulting advantages have not been systematically investigated. This paper introduces one such surrogate with the aim of accelerating parameter estimation for subsurface flow data assimilation problem. Instead of constructing a surrogate model globally, a novel integration of model linearization with our recently proposed SLP enables us to form a set of efficient subdomain linear models for a full replacement of the original HFM. To simplify the notation, subdomain linear model with SLP is denoted as SLM-SLP. These subdomain linear models are easily incorporated into a gradient-based inversion procedure. The performance of this new approach has been assessed through a synthetic non-Gaussian binary facies model and a high-dimensional Gaussian model with 302 uncertain parameters, including geological hydraulic conductivity and relative permeability parameters in this study. Comparisons between this newly proposed SLM-SLP, finite-difference (FD) method, and one variant of iterative ensemble smooth, that is, ES-MDA (Emerick & Reynolds, 2012, 2013), have been carried out as well to evaluate this surrogate-based parameter estimation algorithm.

The remainder of this paper is arranged as follows: Gradient-based variation data assimilation and ensemble-based data assimilation are presented in section 2. A SLP for Gaussian and non-Gaussian models is given in section 3. The setup of adaptive subdomain linear models and their application to the parameter estimation are described in section 4. Section 5 discusses and analyzes some experimental results of a 2-D synthetic non-Gaussian binary facies model and a real-world high-dimensional Gaussian model. Finally, section 6 summarizes our contributions and discusses some potential works in the future.

2. Definition of Parameter Estimation Problem

The relationship between simulated outputs and parameters can be described by a nonlinear operator. To simplify the notation without loss of generality, one such generic operator $\mathbf{h}^n: R^{N_{obs}} \rightarrow R^{N_\beta + N_\alpha}$ can be presented as follows:

$$\mathbf{y}^n = \mathbf{h}^n(\boldsymbol{\beta}, \boldsymbol{\alpha}) + \mathbf{r}^n, \quad n = 1, \cdots, N, \tag{1}$$

where $\boldsymbol{\beta}$ denotes a vector of spatial hydraulic conductivity field. N_β is the total number of spatial grids. n is the simulation timestep. $\mathbf{r}^n$ is a time-dependent vector of observation error, which is generally assumed to satisfy Gaussian distribution $G(\mathbf{0}, \mathbf{R}^n)$. N_{obs} denotes the number of simulated outputs at each time step. $\boldsymbol{\alpha}$ denotes a vector of relative permeability parameters. We employ a Corey-type representation for relative permeability (Li & Horne, 2006)

$$k_{rw}(S_w) = k_{rw}^0 \left(\frac{S_w - S_{wir}}{1 - S_{wir} - S_{wor}} \right)^{n_w}, \quad k_{ro}(S_w) = k_{ro}^0 \left(\frac{1 - S_w - S_{or}}{1 - S_{wir} - S_{wor}} \right)^{n_o}, \tag{2}$$

where S_w is water phase saturation; k_{rw} and k_{ro} are the relative permeability of the water and oil phase respectively; S_{wir} is the irreducible water saturation; S_{or} is the residual oil saturation; k_{rw}^0 is k_{rw} at $S_w = 1 - S_{wir} - S_{or}$; and k_{ro}^0 is k_{ro} at $S_w = S_{wir}$. n_w and n_o are the two power coefficients. Therefore, six parameters $\boldsymbol{\alpha} = [S_{wir}, S_{or}, k_{rw}^0, k_{ro}^0, n_w, n_o]^T$ are used to represent relative permeability.

According to the Bayesian theorem (Evensen, 2009), an objective function can be defined to estimate $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ through conditioning available measurements. For a maximum a posterior (MAP) procedure where one posterior is generated by conditioning to dynamic measurements, an objective function $J_{MAP}(\boldsymbol{\beta}, \boldsymbol{\alpha})$ can be generally defined as follows:

$$\begin{aligned} J_{MAP}(\boldsymbol{\beta}, \boldsymbol{\alpha}) \quad &= \frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}_b)^T \mathbf{C}_{\beta,b}{}^{-1}(\boldsymbol{\beta} - \boldsymbol{\beta}_b) + \frac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_b)^T \mathbf{C}_{\alpha,b}{}^{-1}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_b) \\ &\quad + \frac{1}{2}\sum_{n=1}^{N} (\mathbf{d}_{obs}^n - \mathbf{h}^n(\boldsymbol{\beta}, \boldsymbol{\alpha}))^T \mathbf{R}_n{}^{-1}(\mathbf{d}_{obs}^n - \mathbf{h}^n(\boldsymbol{\beta}, \boldsymbol{\alpha})). \end{aligned} \tag{3}$$

We introduce an efficient data assimilation scheme to solve the inversion problem defined as Equation 3 where a surrogate model, for example, subdomain linear model with smooth local parameterization

(SLM-SLP), is used to be a full replacement of the HFM. In addition, one variant of iterative ensemble smother, that is, ES-MDA, is implemented as a comparison to demonstrate the ability of our approach to effectively and efficiently mitigate ensemble collapse problem (Evensen, 2009). The detailed description of ES-MDA has been provided in the supporting information.

3. Smooth Local Parameterization

We have recently introduced a SLP to separately represent Gaussian parameters in each subdomain through combining PCA and DD strategy. The procedure of implementing SLP has been systematically described by Xiao et al. (2019) and provided in the supporting information. Although our original SLP is aimed at smooth Gaussian parameter field, it is easily extended to non-Gaussian facies model by integrating SLP with an optimization-based principal component analysis (O-PCA) proposed by Vo and Durlofsky (2014).

3.1. Representation of Spatial Gaussian Parameter

Throughout this paper, global PCA and local PCA are separately abbreviated to be GP and LP. Our proposed SLP procedure takes advantages of GP (e.g., smooth representation) and LP (e.g., low-order representation) simultaneously. The global domain Ω is decomposed into Snonoverlapping subdomains Ω^d, $d\in\{1,2,\cdots,S\}$. SLP can determine GP coefficients $\boldsymbol{\xi}_G$ as a function of LP coefficients $\boldsymbol{\xi}^d$ via an optimization procedure, which has been given the supporting information as well.

Given LP coefficients $\boldsymbol{\xi}_L = [\boldsymbol{\xi}^1,\cdots,\boldsymbol{\xi}^d,\cdots,\boldsymbol{\xi}^S]$, $d\in\{1,2,\cdots,S\}$, the corresponding GP coefficients $\boldsymbol{\xi}_G$ are obtained:

$$\boldsymbol{\xi}_G = \mathbf{T}_{GL}\boldsymbol{\xi}_L, \tag{4}$$

where $\mathbf{T}_{GL} \in R^{N_G\times N_L}$ is a transformation matrix for converting LP coefficients to GP coefficients, which has been given as Equation 10 in the provided supporting information. N_G and N_L separately denote the total number of GP and LP coefficients. After obtaining the GP coefficients, the smooth global parameters can be reproduced as follows:

$$\boldsymbol{\beta} = \boldsymbol{\beta}_m + \boldsymbol{\Phi}_\beta\mathbf{T}_{GL}\boldsymbol{\xi}_L, \tag{5}$$

where $\boldsymbol{\Phi}_\beta$ denotes a basis matrix to project global parameter space into reduced subspace.

This SLP allows us to represent GP coefficients using a relatively small number of LP coefficients in each subdomain. To fully preserve the original GP basis vectors in SLP procedure, we should guarantee one necessary condition that the transformation matrix $\mathbf{T}_{GL}$ is full row rank. Xiao et al. (2019) demonstrated that N_L should be equal to or slightly larger than N_G without the need of retaining redundant LP coefficients. This necessary condition could be easily satisfied to obtain an optimal implementation of SLP procedure.

We manually generate $N_r = 1{,}000$ Gaussian models with a size 40×120. The global domain is decomposed into 4×5 rectangle subdomains. Figure 1 shows the reconstructed parameter fields using LP and SLP separately given the same LP coefficients generated by drawing a set of N_L variables from a standard normal distribution. Figure 1 demonstrates that SLP reconstructs a smooth Gaussian parameter field in contrast to LP procedure.

3.2. Representation of Spatially Non-Gaussian Indicators for Binary Systems

As noted by Vo and Durlofsky (2014), PCA-like parameterization techniques, including standard GP and our proposed SLP procedure, are linear-Gaussian transformations, which are not suitable for the non-Gaussian parameters. Fortunately, O-PCA proposed by Vo and Durlofsky (2014) can effectively handle this problem. We extend SLP to non-Gaussian parameters, for example, binary facies system through merging SLP and O-PCA, which is referred to as O-SLP.

We can reformulate Equation 5 as an optimization problem,

$$\boldsymbol{\beta} = \underset{\hat{\boldsymbol{\beta}}}{\operatorname{argmin}} \|\boldsymbol{\beta}_m + \boldsymbol{\Phi}_\beta\mathbf{T}_{GL}\boldsymbol{\xi}_L - \hat{\boldsymbol{\beta}}\|_2^2 + \kappa R(\hat{\boldsymbol{\beta}}),\ \ \hat{\boldsymbol{\beta}_i} \in [\hat{\beta}^l, \hat{\beta}^u], \tag{6}$$

where R is a regularization term and κ is the corresponding regularization weight. $\hat{\beta}^l$ and $\hat{\beta}^u$ separately designate the lower and upper bounds (with $\hat{\boldsymbol{\beta}_i}$ denoting a component of $\hat{\boldsymbol{\beta}}$). The selection of the

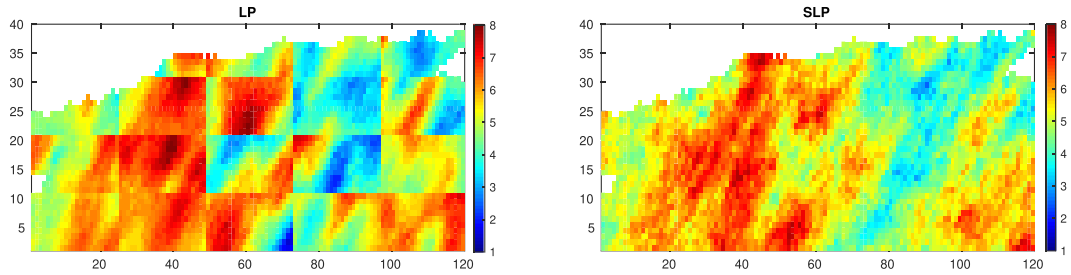

Figure 1. Illustration of Gaussian parameter reconstruction. Left: LP; Right: SLP

regularization term and the corresponding weight determine the type of non-Gaussian parameter field. For a binary facies system, we can specify the regularization term as follows:

$$R = \hat{\boldsymbol{\beta}}^T(\mathbf{1} - \hat{\boldsymbol{\beta}}),\ \ \hat{\beta}^l = 0,\ \ \hat{\beta}^u = 1. \tag{7}$$

This regularization is appropriate for binary systems since it penalizes values away from 0 or 1. Vo and Durlofsky (2014) presented an analytical solution by noting that Equation 6 can be minimized component by component as (where subscript i denotes ith component of a vector)

$$\begin{aligned} &\boldsymbol{\beta}_i = \text{argmin}_{\hat{\beta_i}}\left\{(1\ , -\kappa,)\hat{\boldsymbol{\beta}_i}^2\ , -\ , 2, (\boldsymbol{\beta}_{slp,i}\ , -\frac{\kappa}{2},), \hat{\boldsymbol{\beta}_i}\right\},\\ &\boldsymbol{\beta}_{slp,i} = [\boldsymbol{\beta}_m + \boldsymbol{\Phi}_{\boldsymbol{\beta}}\mathbf{T}_{GL}\boldsymbol{\xi}_L]_i,\\ &\boldsymbol{\beta}_i \in [0,\ 1],\ \ i = 1,\ 2,\ ...,\ N_\beta, \end{aligned} \tag{8}$$

where $\boldsymbol{\beta}_{slp,i}$ represents the reconstructed parameters using the previous SLP method defined in Equation 5. This is a constrained minimization problem; the Lagrangian and Karush-Kuhn-Tucker (KKT) condition can be used to solve it (Vo & Durlofsky, 2014). Finally, an unique analytical solution is given by

$$\boldsymbol{\beta}_i = 0,\ \ \boldsymbol{\beta}_{slp,i} \leqslant \frac{\kappa}{2},\ \boldsymbol{\beta}_i = 1,\ \ \boldsymbol{\beta}_{slp,i} \geqslant 1 - \frac{\kappa}{2},\ \boldsymbol{\beta}_i = \frac{\boldsymbol{\beta}_{slp,i} - \frac{\kappa}{2}}{1-\kappa},\ \ \frac{\kappa}{2} \leqslant \boldsymbol{\beta}_{slp,i} \leqslant 1 - \frac{\kappa}{2}. \tag{9}$$

We should note that Vo and Durlofsky (2014) posed determination of κ as a parameter optimization before implementing this O-SLP procedure. In addition, Vo and Durlofsky (2014) also derived an expression for the derivative of geological parameters $\boldsymbol{\beta}$ with respect to GP coefficients $\boldsymbol{\xi}_G$. Similarly, we can obtain the derivative of GP coefficients $\boldsymbol{\xi}_G$ with respect to the LP coefficients $\boldsymbol{\xi}_L$ as well using Equation 4. By combining these two derivatives on the basis of a chain rule, we can analytically obtain the derivative of $\boldsymbol{\beta}$ with respect to LP coefficients $\boldsymbol{\xi}_L$ as follows:

$$\frac{d\boldsymbol{\beta}}{d\boldsymbol{\xi}_L} = \frac{d\boldsymbol{\beta}}{d\boldsymbol{\xi}_G}\frac{d\boldsymbol{\xi}_G}{d\boldsymbol{\xi}_L}, \tag{10}$$

where we do not explicitly give the formula of this derivative term $\frac{d\boldsymbol{\beta}}{d\boldsymbol{\xi}_G}$. This derivative term demonstrates that O-SLP preserves the smoothness and differentiability, which is particularly useful for the gradient-based parameter estimation in this study.

We generate $N_r = 1{,}000$ binary facies models, and among them, three realizations are illustrated in the first row of Figure 2. Unlike the Gaussian smooth parameter field with spatially continuous variables, for example, log-permeability, an indicator used to identify the facies type has specific number, for example, 0 and 1 for binary facies system. The dimensionality of this model is 40×40, which is decomposed into 3×3 rectangle subdomains. The last two rows of Figure 2 separately show the reconstructed binary facies field and its corresponding histogram using LP, SLP, and O-SLP, which indicates that not only does O-SLP reproduce a smooth non-Gaussian parameter field in contrast to LP procedure but also it preserves the

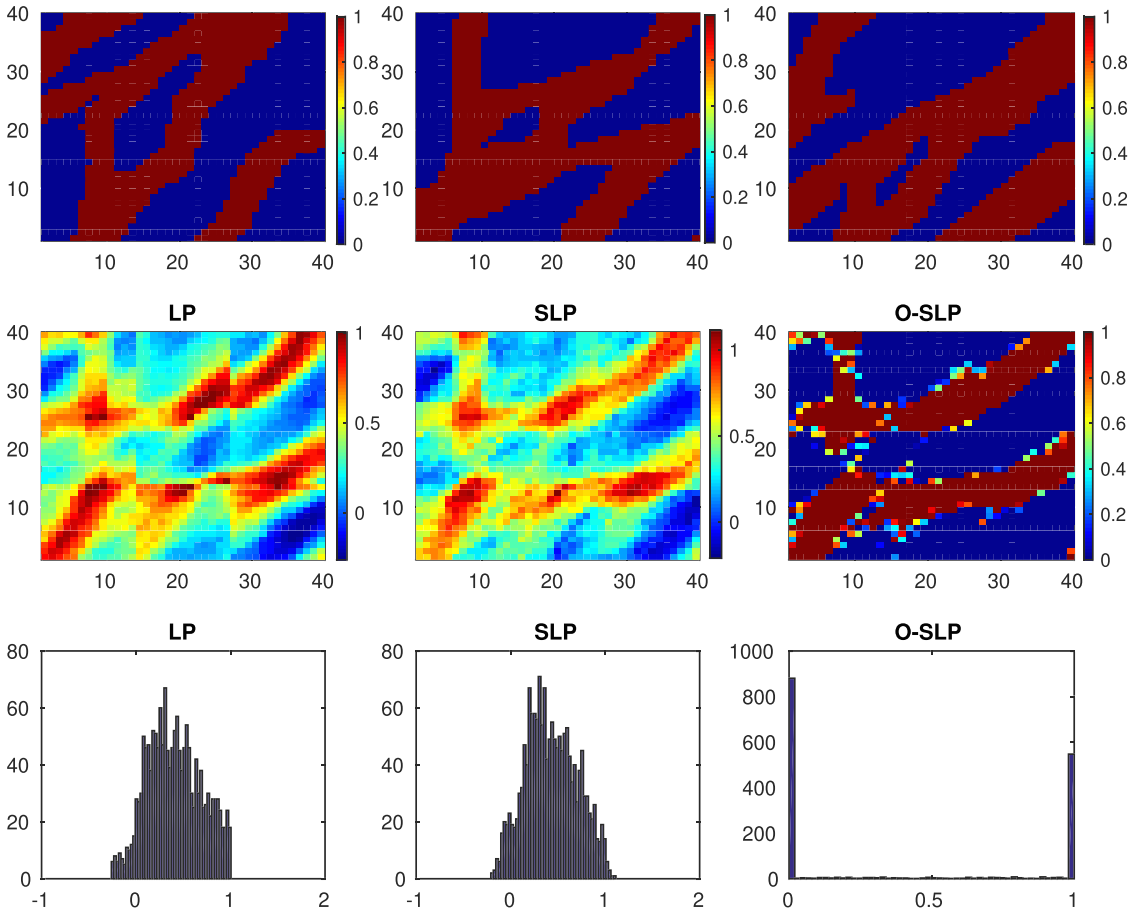

Figure 2. Illustration of smooth local parameterization for binary facies model. The first row shows three realizations of the binary facies mode. The second and third rows show the reproduced facies models and its corresponding histograms. Left: LP; middle: SLP; right: O-SLP.

non-Gaussian properties of facies indicators effectively. The weight $\kappa \in (0, 1)$ quantifies the impact of regularization on the realizations generated by O-SLP. To have a near-binary distribution that does not exhibit too large of a jump in the derivative (meaning the model contains sufficient values between 0 and 1), a value of $\kappa = 0.8$ is used in this study. The results demonstrate that O-SLP pushes the values to 0 and 1 as much as possible while clearly contains values between 0 and 1 as well. This renders the O-SLP model (piecewise) differentiable. We also show some reconstructed facies models conditioning to different values of κ in the supporting information. As $\kappa \to 0$, the realizations will be far from the target (binary) distribution, and as $\kappa \to 1$, the jump in the derivative will become too large.

4. Surrogate Model with SLP for Parameter Estimation

4.1. Formulation of Surrogate Model

In this section, we formulate a surrogate to replace the original complex system with a set of subdomain linear models. In general, the simulated outputs are strongly correlated with fewer influential or local parameters (Ding, 2011; Ding & Mckee, 2013). This fact motivates us to construct subdomain linear models with smooth local parameterization (SLM-SLP) alternatively. After decomposing the global domain Ω into S nonoverlapping subdomains Ω^d, $d \in \{1,2,\cdots,S\}$, it is critical to relate the specific simulated outputs with the local parameters. To speed up the model linearization, we assume that each subdomain is "large" enough so that it can be handled individually, and we can generate subdomain linear model for each large subdomain individually. A subdomain linear model in subdomain Ω^d can be presented as follows:

Table 1
Summary of the Available Formulations of the Weighting Coefficients to Relax Constrains

L_{lp}	$\omega = [\omega_1, \cdots, \omega_{L_{lp}}]$
1	$\omega_1 = 1$
> = 2	Type 1: $\omega_1 = \frac{1}{\sum_{l=1}^{L_{lp}} \gamma^{2(l-1)}}, \omega_2 = \frac{\gamma^2}{\sum_{l=1}^{L_{lp}} \gamma^{2(l-1)}}, \cdots, \omega_{L_{lp}} = \frac{\gamma^{2(L_{lp}-1)}}{\sum_{l=1}^{L_{lp}} \gamma^{2(L_{lp}-1)}}$ Type 2: $\omega_1 = \frac{1}{\sum_{l=1}^{L_{lp}} e^{(l-1)\gamma}}, \omega_2 = \frac{e^{\gamma}}{\sum_{l=1}^{L_{lp}} e^{(l-1)\gamma}}, \cdots, \omega_{L_{lp}} = \frac{e^{(L_{lp}-1)\gamma}}{\sum_{l=1}^{L_{lp}} e^{(L_{lp}-1)\gamma}}$

$$\mathbf{y}^{n,d} = \mathbf{y}_b^{n,d} + \frac{\partial \mathbf{y}^{n,d}}{\partial \boldsymbol{\xi}_b^d}(\boldsymbol{\xi}^d - \boldsymbol{\xi}_b^d) + \frac{\partial \mathbf{y}^{n,d}}{\partial \boldsymbol{\xi}_b^{sd}}(\boldsymbol{\xi}^{sd} - \boldsymbol{\xi}_b^{sd}) + \frac{\partial \mathbf{y}^{n,d}}{\partial \boldsymbol{\alpha}_b}(\boldsymbol{\alpha} - \boldsymbol{\alpha}_b), \quad n = 1, \cdots, N, \tag{11}$$

where $\frac{\partial \mathbf{y}^{n,d}}{\partial \boldsymbol{\xi}_b^d}$ and $\frac{\partial \mathbf{y}^{n,d}}{\partial \boldsymbol{\xi}_b^{sd}}$ are the derivative terms of simulated outputs with respective to the related local parameters. A simple numerical FD method can be used to approximate these derivative terms. Distinctive to spatial geological parameters, relative permeability parameters $\boldsymbol{\alpha}$ are global correlation and cannot be locally parameterized.

To generate a robust and accurate linear model (Nocedal & Wright, 1999), the FD method requires that the number of model evaluations should be strictly larger than the number of parameters. This strict requirement makes high-dimensional inversion problems almost computationally intractable. Construction of subdomain linear model avoids to solve one large system of equations for the global linear model and thus has some potential advantages: (1) local parameterization will be highly representative of the local geological properties; (2) there may be only a few patterns required in each subdomain, leading to further reduction of the dimensions of the surrogate model that need to be constructed; and (3) a natural opportunity may arise for parallelization.

We should note that it will be never possible for us to easily clarify this local correlations in realistic applications and the definition of " large " or " small " is very subjective. An inappropriate DD scheme, for example, too "large" or too "small" subdomains, easily leads to spurious long-distance correlations or cuts off inherent correlations among neighboring subdomains and hence deteriorates the accuracy of SLM-SLP. Fortunately, an adaptive scheme will be introduced to mitigate this negative effect to some extent in the following part.

4.2. An Adaptive Formulation of Parameter Estimation with Surrogate Model

We introduce an adaptive scheme, that is, weighting smooth local parameterization (WSLP), to mitigate the negative effect of inappropriate DD scheme. The main idea of WSLP is to redefine the original objective function as a weighting sum of sub-objective functions corresponding to specific DD. We introduce a new vector of weighting coefficients $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_{L_{lp}}]$; L_{lp} represents the total number of DD schemes. Each weighting coefficient is assigned to one sub-objective function, and then we minimize the total sum of these weighting sub-objective functions. The minimization algorithm adaptively determines the weighting coefficients and the optimal DD correspondingly.

The l^{th} DD scheme generates a vector of LP coefficients $\boldsymbol{\xi}_L^l = [\boldsymbol{\xi}^{d_1}, \cdots, \boldsymbol{\xi}^{d_l}, \cdots, \boldsymbol{\xi}^{d_{S_l}}]$. The original objective function $J(\boldsymbol{\xi}_G, \boldsymbol{\alpha})$ is presented as a total sum of weighting sub-objective function $J(\mathbf{T}_{GL}^l \boldsymbol{\xi}_L^l, \boldsymbol{\alpha})\,(l = 1, \cdots, L_{lp})$ corresponding to different DD scheme as follows:

$$J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}) = \sum_{l=1}^{L_{lp}} \omega_l J(\mathbf{T}_{GL}^l \boldsymbol{\xi}_L^l, \boldsymbol{\alpha}), \quad \sum_{l=1}^{L_{lp}} \omega_l = 1, \quad \omega_l > 0, \tag{12}$$

where $\mathbf{T}_{GL}^l$ is a transformation matrix given in Equation 4 for converting LP coefficients to GP coefficients for the lth DD scheme. Equation 12 is a constrained optimization problem. We introduce a new variable γ to replace these weighting coefficients and therefore make the constrains, for example, $\sum_{l=1}^{L_{lp}} \omega_l = 1$ and $\omega_l > 0$, automatically satisfied. Some available formulations of the weighting coefficients are listed in Table 1.

Instead of directly minimizing this objective function with respect to LP coefficients $\boldsymbol{\xi}_L^l$ for each DD scheme, we can minimize GP coefficients $\boldsymbol{\xi}_G$ alternatively by use of a linear relationship between $\boldsymbol{\xi}_L$ and $\boldsymbol{\xi}_G$ as defined in Equation 4. In addition to GP coefficients $\boldsymbol{\xi}_G$ and relative permeability parameters $\boldsymbol{\alpha}$, the new variable γ also needs to be optimized simultaneously. The gradient of this objective with respect to all variables is as follows:

$$\nabla J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma) = \left[\left[\frac{\partial J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)}{\partial \boldsymbol{\xi}_G}\right]^T, \left[\frac{\partial J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)}{\partial \boldsymbol{\alpha}}\right]^T, \left[\frac{\partial J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)}{\partial \gamma}\right]^T\right]^T, \tag{13}$$

where

$$\begin{aligned}
\left[\frac{\partial J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)}{\partial \boldsymbol{\xi}_G}\right]^T &= \sum_{l=1}^{L_{lp}} \omega_l \left[\frac{\partial J(\mathbf{T}_{GL}^l \boldsymbol{\xi}_L^l, \boldsymbol{\alpha}) d\boldsymbol{\xi}_L^l}{\partial \boldsymbol{\xi}_L^l \quad d\boldsymbol{\xi}_G}\right]^T, \\
\left[\frac{\partial J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)}{\partial \boldsymbol{\alpha}}\right]^T &= \sum_{l=1}^{L_{lp}} \omega_l \left[\frac{\partial J(\mathbf{T}_{GL}^l \boldsymbol{\xi}_L^l, \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\right]^T, \quad \left[\frac{\partial J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)}{\partial \gamma}\right]^T = \sum_{l=1}^{L_{lp}} J(\mathbf{T}_{GL}^l \boldsymbol{\xi}_L^l, \boldsymbol{\alpha}) \left[\frac{d\omega_l}{d\gamma}\right]^T.
\end{aligned} \tag{14}$$

To explicitly determine the gradient $\nabla J(\boldsymbol{\xi}_G,\boldsymbol{\alpha},\gamma)$, several terms required in Equation 14 have been provided in the supporting information. Once the gradient $\nabla J(\boldsymbol{\xi}_G,\boldsymbol{\alpha},\gamma)$ at the kth iteration step is available, the next estimation of the optimal variables that minimize the objective function is given by

$$\begin{bmatrix} \boldsymbol{\xi}_G \\ \boldsymbol{\alpha} \\ \gamma \end{bmatrix}_{k+1} = \begin{bmatrix} \boldsymbol{\xi}_G \\ \boldsymbol{\alpha} \\ \gamma \end{bmatrix}_k - \varepsilon_k \frac{\nabla J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)_k}{\|\nabla J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)_k\|_\infty}, \tag{15}$$

where ε_k is a step length at the kth iteration step. ∞ represents an infinite norm of a vector.

Many optimization methods can be used to minimize the objective function $J(\boldsymbol{\xi}_G,\boldsymbol{\alpha})$ defined as Equation 12. One can approximate the Hessian matrix using the first-order gradient and then use a Newton's like updating, for example, quasi-Newton and Gaussian Newton (Nocedal & Wright, 1999). The accuracy of Hessian matrix is mainly determined by the quality of the gradient computed from the subdomain linear models; thus, the convergence of Newton's like optimization methods cannot be ensured very well. In this work, the minimization procedure is directly performed using a steepest descent algorithm (Nocedal & Wright, 1999).

The minimization algorithm terminates when either one of the following two stopping criteria is satisfied. Here η_J and $\eta_{\xi_G,\alpha,\gamma}$ denote the predefined error constraints.

- The objective function defined as Equation 3 hardly changes, that is,

$$\frac{|J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)_{k+1}) - J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)_k)|}{max\{|J(\boldsymbol{\xi}_G, \boldsymbol{\alpha}, \gamma)_{k+1})|, 1\}} < \eta_J. \tag{16}$$

- The estimate of parameters almost does not change, that is,

$$\frac{\left|\begin{bmatrix} \boldsymbol{\xi}_G \\ \boldsymbol{\alpha} \\ \gamma \end{bmatrix}_{k+1} - \begin{bmatrix} \boldsymbol{\xi}_G \\ \boldsymbol{\alpha} \\ \gamma \end{bmatrix}_k\right|}{max\left\{\left|\begin{bmatrix} \boldsymbol{\xi}_G \\ \boldsymbol{\alpha} \\ \gamma \end{bmatrix}_{k+1}\right|, 1\right\}} < \eta_{\xi_G,\alpha,\gamma}. \tag{17}$$

After obtaining the optimized parameters $(\boldsymbol{\xi}_G^*, \boldsymbol{\alpha}^*)$ by use of current subdomain linear models, the objective function $J(\boldsymbol{\xi}_G^*, \boldsymbol{\alpha}^*)$ is computed using the HFM to check whether an acceptable optimizer has been obtained.

Table 2
Experiment Settings Using MRST for Cases 1 and 2

Basic parameter	Case 1	Case 2	Unit
Dimension	60×60	40×120	—
Grid size	40×40	m	
Number of injectors and producers	6, 4	6, 7	-
Fluid density	1,014, 859	kg/m^3	
Fluid viscosity	0.4, 2	mP·s	
Initial pressure	25	MPa	
Initial saturation	$S_o = 0.80, S_w = 0.20$	—	
Injection rate	200	m^3/d	
Producer head pressure	20	MPa	
Historical production time	5	10	Years
Prediction time	10	15	Years
Time step	0.1	Years	
Measurement time step	0.2	Years	
Relative permeability parameters	Lower bound	Upper bound	Reference value
Irreducible water saturation, S_{wir}	0.3	0.1	0.15
Residual oil saturation, S_{or}	0.3	0.1	0.15
Corey exponent for oil phase, n_o	4	2	3
Corey exponent for water phase, n_w	4	2	3
k_{ro}^0	1.0	0.8	0.9
k_{rw}^0	1.0	0.8	0.9

If the solution of this subdomain linear system as Equation 11 cannot be used to exactly represent the original objective function as Equation 3, additional outer loops are required to reconstruct new subdomain linear models around the currently updated parameters ($\boldsymbol{\xi}_G^*$, $\boldsymbol{\alpha}^*$), and then the iterative inner loop is performed again.

The choice of our proposed locally linear surrogates deserves to be discussed and explained here. In general, some nonlinear surrogate models, such as PCE and GPR, which have been extensively applied to the field of subsurface flow problem, can be used to represent the nonlinearity of the problems; unfortunately, the required number of training models will be exponentially increased. On the contrary, the computational cost of our proposed subdomain linear models will be linear increased as the parameter dimensions. In terms of representing the nonlinearity, our approach construct local linear model at each out iteration and then update the linear model for each iterations. This alternative inner-outer loop strategy that constructs local linear surrogates and updates them continuously could progressively capture the nonlinearity and approach to the optimal solution (Altaf et al., 2009; Kaleta et al., 2011; Vermeulen & Heemink, 2006; Xiao et al., 2018). In addition, in the presence of global approximations the order of the polynomial can be quite high, and therefore, the number of coefficients and consequently numerical simulations grows exponentially, while the local representations would only require a limited degree (this is indeed why our proposed method works by only employing a linear local reconstruction). It could expect that the reconstruction of surrogate models could be avoid if we directly form a nonlinear model using PCE or GPR. However, the corresponding computational cost will be significantly increased. It is very interesting to compare these two strategies through assessing the accuracy and efficiency in the future.

5. Illustrative Examples

Two cases are presented in this section. The first case is mainly used to study the DD strategy on a two-dimensional non-Gaussian binary facies model. The RML procedure is provided in the supporting information. The second case is tested on a real-world high-dimensional model with Gaussian hydraulic conductivity parameters. This case is used to demonstrate the computational efficiency of SLM-SLP in high-dimensional inverse modeling.

MRST, a free open-source software for subsurface flow modeling and simulation (Lie et al., 2012), is used to run HFM simulations. Settings about the model geometry, fluid properties, and well controls are shown in

Table 2. In terms of computational effort, the runtime of a single HFM simulation for these two cases is about 4.8 and 12.5 s on a machine with i5-4690 Intel CPUs (four cores, 3.5 GHz) and 24 GB memory using Matlab-R2015a. SLM-SLP model, by contrast, needs less than 0.1 s for those two cases. The results of application of SLM-SLP in conjunction with an gradient-based parameter estimation procedure are presented in the following parts.

We quantify the quality of SLM-SLP by reporting the ensemble average root mean square error ($RMSE_{mu}$) and the ensemble average spread (ESP_{mu}) for the estimated geological parameters fields $\boldsymbol{\beta}$, which separately represents the facies indicators and log-permeability field for Cases 1 and 2 and relative permeability parameters $\boldsymbol{\alpha}$. These measures are computed as follows:

$$RMSE_{mu}(\boldsymbol{\beta})(18)=\frac{1}{N_{\beta i=1}}\sum^{N_\beta}\sqrt{\frac{1}{N_{ej=1}}\sum^{N_e}(\boldsymbol{\beta}^{i,j}-\boldsymbol{\beta}^{i}_{true})^2},\quad ESP_{mu}(\boldsymbol{\beta})=\frac{1}{N_{\beta i=1}}\sum^{N_\beta}\sqrt{\frac{1}{N_{ej=1}}\sum^{N_e}(\boldsymbol{\beta}^{i,j}-\boldsymbol{\beta}^{i}_{mean})^2},$$
$$RMSE_{mu}(\boldsymbol{\alpha})=\frac{1}{N_{\alpha i=1}}\sum^{N_\alpha}\sqrt{\frac{1}{N_{ej=1}}\sum^{N_e}(\boldsymbol{\alpha}^{i,j}-\boldsymbol{\alpha}^{i}_{true})^2},\quad ESP_{mu}(\boldsymbol{\alpha})=\frac{1}{N_{\alpha i=1}}\sum^{N_\alpha}\sqrt{\frac{1}{N_{ej=1}}\sum^{N_e}(\boldsymbol{\alpha}^{i,j}-\boldsymbol{\alpha}^{i}_{mean})^2},\qquad(18)$$
$$\boldsymbol{\beta}^{i}_{mean}=\frac{1}{N_{ei=1}}\sum^{N_e}\boldsymbol{\beta}^{i,j},\quad \boldsymbol{\alpha}^{i}_{mean}=\frac{1}{N_{ei=1}}\sum^{N_e}\boldsymbol{\alpha}^{i,j},$$

where $\boldsymbol{\beta}^{i,j}$ denotes the ith element of jth ensemble realization for parameter vector $\boldsymbol{\beta}$. This notation is also same for $\boldsymbol{\alpha}^{i,j}$. $\boldsymbol{\beta}_{true}$ and $\boldsymbol{\alpha}_{true}$ separately represent the reference values.

5.1. Case 1: Two-Dimensional Synthetic Modeling with Non-Gaussian Binary Facies

In this case, SLM-SLP is used to estimate non-Gaussian binary facies indicators. The value of log-permeability for these two facies has a large contrast. The permeability of shale facies and sand facies is 50 mD (millidarcy) and 2,000 mD (millidarcy), respectively. Given the facies indicators $\boldsymbol{\beta}$ obtained from our proposed O-SLP procedure, we compute the permeability value for each grid using a linear interpolation function ($50 + 2{,}000\,\boldsymbol{\beta}$). This model describes a water-oil two-phase system. There are in total six producers and four injectors, labeled from P_1 to P_6 and I_1 to I_4, respectively, as illustrated in Figure 4.

We manually generate 1,000 non-Gaussian binary facies models, and one of these realizations is considered to be the truth; see Figure 4a. We divide the entire domain into 4×4 rectangle subdomains, which is considered to be the base case. In this numerical experiment, GP procedure preserves 72 GP coefficients, while O-SLP procedure uniformly preserves five LP coefficients in each subdomain. Although totally, 80 LP coefficients are used to characterize original 72 GP coefficients, five LP coefficients in each subdomain are significantly smaller than 72 GP coefficients. This advantage allows us to construct subdomain linear models efficiently. We generate a vector of relative permeability parameter $\boldsymbol{\alpha}$ using a function: $\boldsymbol{\alpha}^j = \mathbf{L}^j + (\mathbf{U}^j - \mathbf{L}^j)\,R_{nd}$, where $\mathbf{L}^j$ and $\mathbf{U}^j$ are the lower and upper bounds of the jth relative permeability parameters and R_{nd} is a random number between 0 and 1. The lower bounds, upper bounds, and reference values for these six relative permeability parameters are shown in Table 2. There are in total 78 parameters needed to be estimated for Case 1, including 72 GP coefficients and six relative permeability parameters.

The noisy measurements generated from the model with the true hydraulic conductivity field and relative permeability parameters consist of head pressures in four injectors and fluid flux and water breakthrough in six producers, which gives in total 400 measurements. Normal distributed independent measurement noise with a standard deviation equal to 5% of the true data values is added to all measurements.

5.1.1. Study of a Base Case

The numerical FD optimization approach with global parameterization, referred to as FD-GP, is also implemented as a comparison. The stopping criteria are set to be $\eta_J = 10^{-4}$ and $\eta_{\xi_G,\alpha,\gamma} = 10^{-3}$. We set the initial iteration length $\varepsilon_0 = 0.1$; the iteration length is automatically halved as the value of the objective function increases. Figures 3a and 3b separately show the evolution of objective function values as a function of out-loop iterations and required HFM simulations. Our proposed SLM-SLP procedure requires much fewer HFM simulations than that of FD-GP procedure, while comparable objective function values are obtained. Specifically, SLM-SLP runs 120 HFM simulations; among them, five HFM are ran by perturbing LP coefficients and six HFM are ran by perturbing the relative permeability parameters at each outer-loop

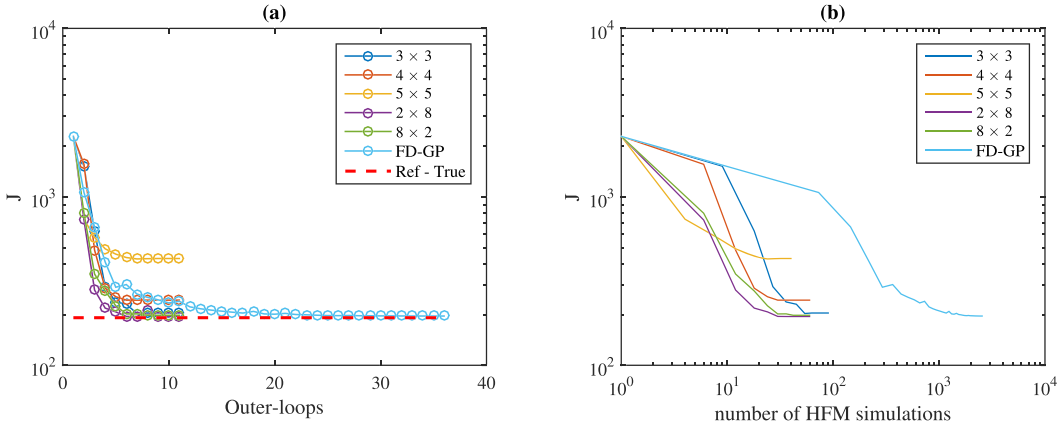

Figure 3. The evolution of the objective function values as a function of outer loops and number of HFM simulations using SLM-LP and GP-FD method for Case 1. The calculation of the objective function for outer loops uses HFM defined in Equation 3.

iteration, while GP-FD requires 2,808 HFM simulations. There is an approximate 20 times speedups expressed in the number of HFM runs. Figure 3 also depicts that SLM-SLP procedure obtains acceptable results after five outer-loop iterations, and therefore, 60 redundant HFM runs can be further avoided.

Figure 4 shows the true, initial, and updated binary facies models. An effective calibration of discrete channelized geological features is a key metric to assess the quality of the non-Gaussian parameter estimation procedure. Three red circles are drawn to emphasize three different connectivity structures as illustrated in Figure 4. It is evident that SLM-SLP successfully reproduces these three structures. The three subfigures at the third row show the updated binary facies indicators at different outer loops. These three structures are gradually reconstructed using our new method. We also should emphasize that the calibration of channelized geological connectivity easily makes gradient-based optimization unstable due to its discreteness; fortunately, the O-SLP is a linear transformation with smoothness and differentiability and therefore is able to effectively calibrate binary system. The subfigure at the last row illustrates the histogram of facies indicators of the true model, initial model, and updated models. Although the amounts of facies indicators 0 or 1 are not perfectly reproduced, both nonzero and non-one indicators observed in the initial model have been drastically decreased after conditioning to the dynamic data, indicating that the non-Gaussian binary facies features have been successfully preserved. The estimated relative permeability parameters and their relative errors are listed in Table 3. Parts of the estimated values show a good agreement with the reference values.

5.1.2. Study of DD Scheme

Five regular DD schemes, that is, 3×3, 4×4, 5×5, 2×8, and 8×2 rectangle subdomains are studied. Using more subdomains generates fewer LP coefficients in each subdomain; as a result, fewer HFM simulations are required to construct SLM-SLP models. Among these five DD schemes, SLP enables us to retain at most eight LP coefficients. The evolution of objective function values as a function of out-loop iterations is shown in Figure 3a. SLM-SLP obtains comparable objective function values with FD-GP method except the 5×5 DD. Figure 3b shows the evolution of objective function values as a function of required HFM simulations for FD-GP and SLM-SLP. SLM-SLP requires much fewer HFM simulations than FD-GP to yield comparable objective function values. For example, SLM-SLP with 2×8 DD requires 60 HFM simulations, which indicates an approximate 20 times speedups expressed in number of HFM simulations. Figure 5 shows the true, the initial, and the updated binary facies models. The updated facies indicators and the discrete channelized geological features differ significantly, which implies that some local minima are generated using different DD schemes. Overall, these results indicate that it is possible to run fewer HFM simulations to find accurate results by designing a much better DD scheme, for example, 2×8 in this case.

To further investigate the sensitivity of subdomains with other shapes, we decompose the global domain into triangle-like subdomains as illustrated in Figure 6. The 3×3 rectangle subdomain is denoted as *rectangle subdomain* as shown in Figure 6a; other two layouts, separately denoted as *triangle subdomain 1* and *triangle subdomain 2*, are used to form nine triangle-like subdomains. Figure 6d shows the evolution of objective

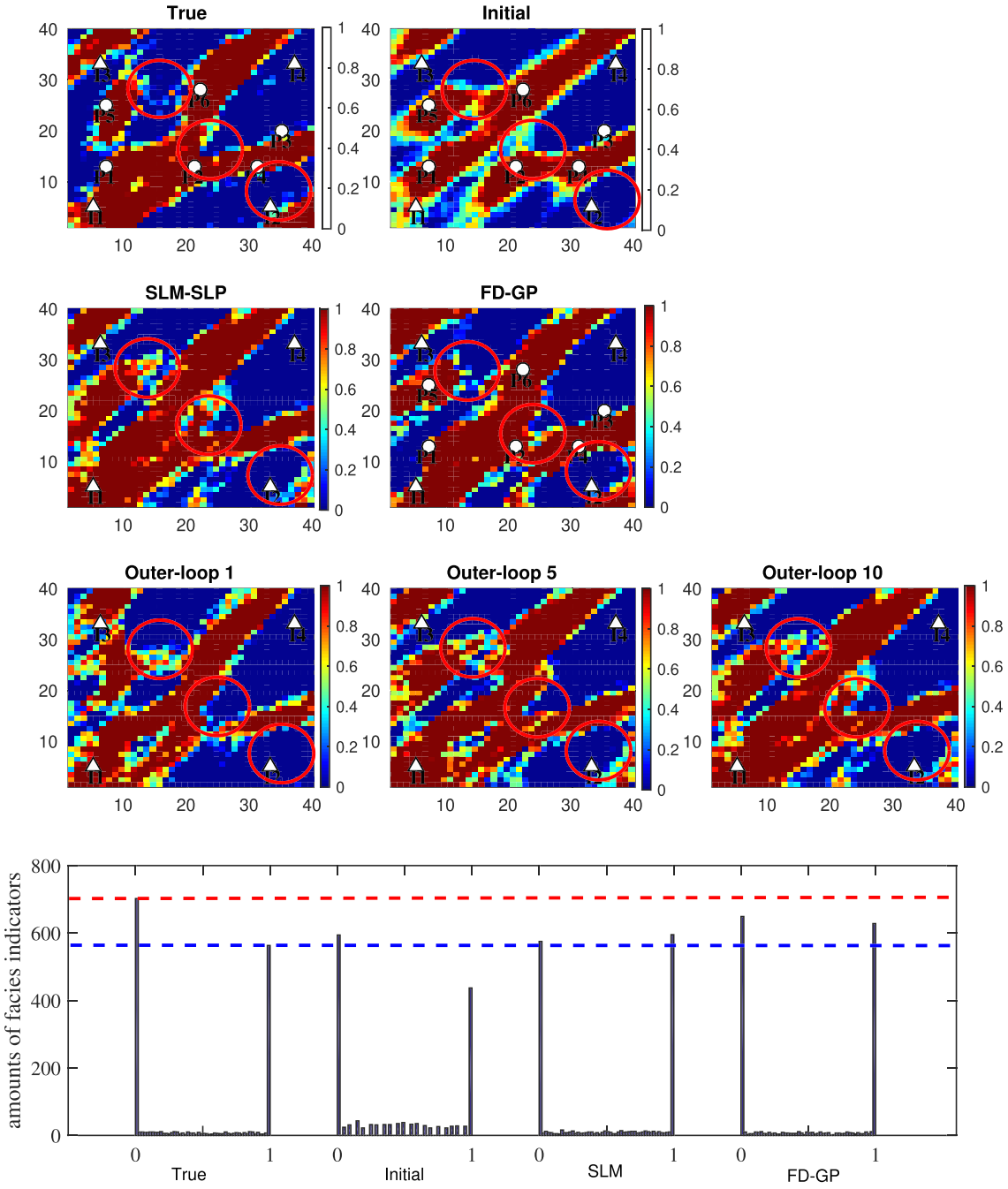

Figure 4. Updated binary facies indicators using SLM-LP and GP-FD method for Case 1. The three subfigures at the third row show the updated binary facies models at 1st, 5th and 10th outer-loop iteration. The last subfigure is the histogram of facies indicators for the true model, initial model, and updated model. The dash red line and blue line separately denote the amount of shale indicator and sand indicator for the true model.

function values as a function of out-loop iterations. These three layouts yield similar objective functions while obtain different convergence rates. Specifically, *rectangle subdomain* and *triangle subdomain 1* approximately need 10 out-loops, while *triangle subdomain 2* requires another six outer-loops, which indicates that *triangle subdomain 2* needs more HFM simulations. Figures 6e and 6f show the updated binary facies models for *triangle subdomain 1* and *triangle subdomain 2*. Overall, the main channelized features have been successfully reconstructed. In comparison to triangle-like subdomains, the regular subdomains are easily formed and therefore are convenient to be applied to the real-world inverse modeling.

Table 3
The Estimated Relative Permeability Parameters for Case 1

Parameters	$\alpha = [S_{wir}, S_{or}, k^0_{rw}, k^0_{ro}, n_w, n_o]^T$	
Reference value	[0.15, 0.15, 0.9, 0.9, 3, 3]	
-	Estimated Value	Relative Error (%)
3×3	[0.162, 0.159, 0.862, 0.913, 3.12, 3.23]	[8.0, 1.4]
4×4	[0.179, 0.175, 0.922, 0.931, 3.24, 3.29]	[19, 2.4]
5×5	[0.189, 0.191, 0.935, 0.939, 3.25, 3.38]	[27, 3.9]
2×8	[0.164, 0.141, 0.915, 0.872, 3.25, 3.18]	[9.3, 1.7]
8×2	[0.141, 0.162, 0.928, 0.864, 3.18, 3.28]	[8.0, 3.1]
Adaptive 1	[0.161, 0.159, 0.861, 0.914, 3.15, 3.21]	[7.3, 1.6]
Adaptive 2	[0.165, 0.162, 0.871, 0.919, 3.17, 3.24]	[10, 2.1]
Adaptive 3	[0.178, 0.175, 0.925, 0.933, 3.22, 3.27]	[18.7, 2.8]
FD-GP	[0.166, 0.160, 0.912, 0.824, 3.18, 3.22]	[10.6, 1.3]

Note. The last column shows the maximum and minimum relative errors.

Through repeating the optimization process for each DD individually, we have determined that 2×8 is the best one, while 5×5 is the worse one. However, repeating the optimization process requires intensive computations in real-world applications. To mitigate this negative effect of inappropriate DD scheme, such as 5×5 subdomains where the objective function is hard to be minimized, an adaptive SLP strategy, that is, WSLP, is introduced. We design three adaptive strategies to systematically investigate the performance of WSLP.

- Adaptive 1: 3×3 and 5×5 are simultaneously considered.
- Adaptive 2: 4×4 and 5×5 are simultaneously considered.
- Adaptive 3: 3×3, 4×4, and 5×5 are simultaneously considered.

Types 1 and 2 separately denote the two formulas of weighting coefficients as given in Table 1. As shown in Figures 7a–7c, both Adaptives 1 and 3 generate comparable objective function values with the 3×3 DD

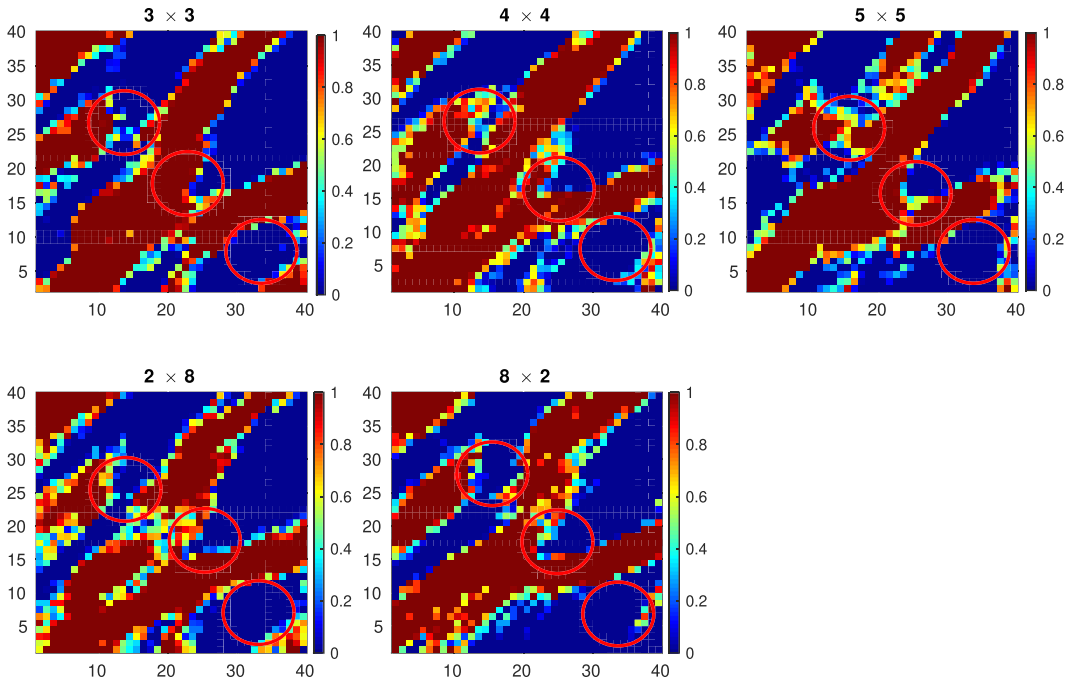

Figure 5. Updated binary facies models using different domain decomposition without the adaptive scheme for Case 1.

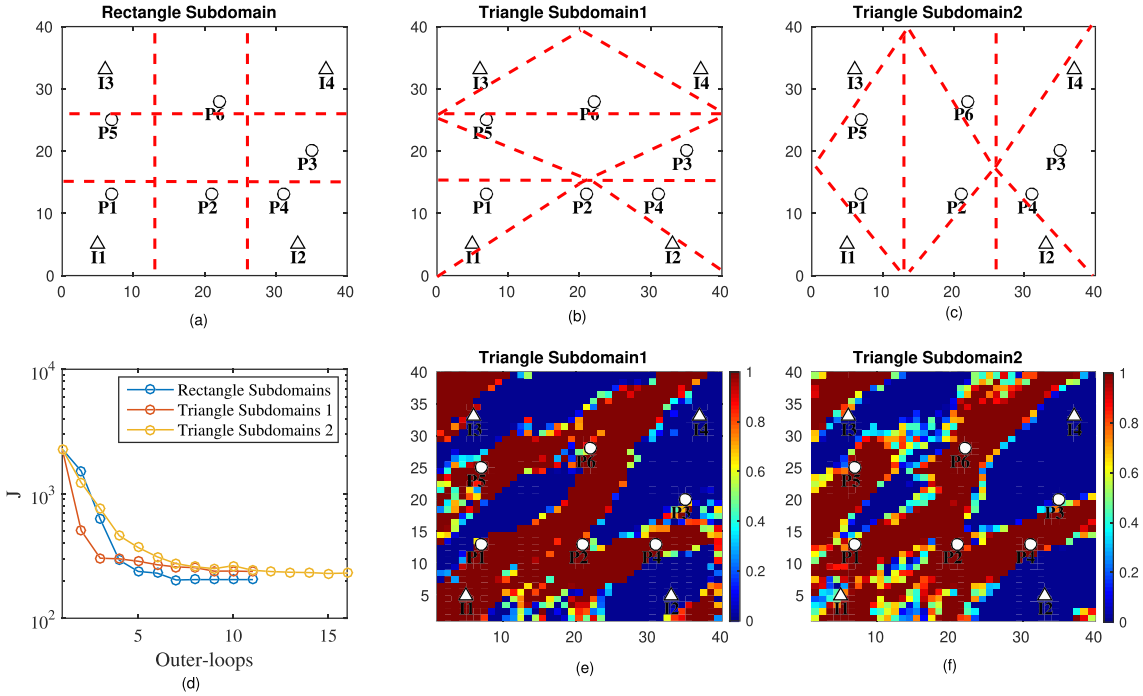

Figure 6. Updated binary facies models using rectangle and triangle-like subdomains for Case 1. The first row shows the layout of rectangle and two triangle-like subdomains. The second row shows the evolution of objective function and updated binary facies models. The red lines are used to makea clear comparison.

scheme, while Adaptive 2 generates comparable objective function values with the 4 × 4 DD scheme. Figures 7d–7f represent the evolution of the weighting coefficients as a function of outer-loop iterations. We initialize an equal weighting for each DD scheme; this adaptive scheme can gradually identify the subcost function, which is easy to be minimized and then assign a dominant weighting to it correspondingly or vice versa. For example, both Adaptives 1 and 3 identify the subcost function for 3 × 3 DD is easily minimized, while Adaptive 2 identifies the subcost function for 4 × 4 DD is easily minimized. The 5 × 5 DD is hard to be minimized in this case. These results are perfectly consistent with the previous results. Figures 7g–7i depict the updated binary facies models for these three adaptive strategies. Visually speaking, both Adaptives 1 and 3 yield similar models with 3 × 3 DD scheme, while Adaptive 2 generates a similar models with the 4 × 4 DD scheme. This adaptive strategy can help us optimize the DD scheme effectively, particularly when a good prior knowledge of the studied domain is not available. We could consider various DD schemes as much as possible so that we have the high possibility to get the optimal or at least suboptimal one.

The computational cost of this adaptive strategy is approximately equivalent to the most computationally expensive one among all considered DD schemes. For example, both Adaptives 1 and 3 need 150 HFM simulations, while Adaptive 2 needs 120 HFM simulations. Figure 8 shows the results with increasing numbers of subdomains for the first example (e.g., a convergence plot and a HFM simulation plot). We adopt the standard *rectangle subdomain* as shown in Figure 6a while increasing the number of subdomains. It is easily observed that the results of the proposed adaptive strategy can approximately converge to the reference, which indicates the accuracy and efficiency of the proposed method. The corruption of the very bad DD indeed has been significantly decreased as suggested in Figure 7. This adaptive scheme requires us to construct a large set of subdomain linear models for all considered DD schemes, which will increase the overall computational overhead and memory requirements particularly for high-dimensional models. Based on the first several outer-loop iterations, this adaptive scheme has successfully implied the optimal DD according to the values of weighting coefficients, and then we can directly minimize the corresponding objective function without using this adaptive scheme. The real-world applications benefit from this finding by avoiding the unnecessary computational overhead and memory requirements.

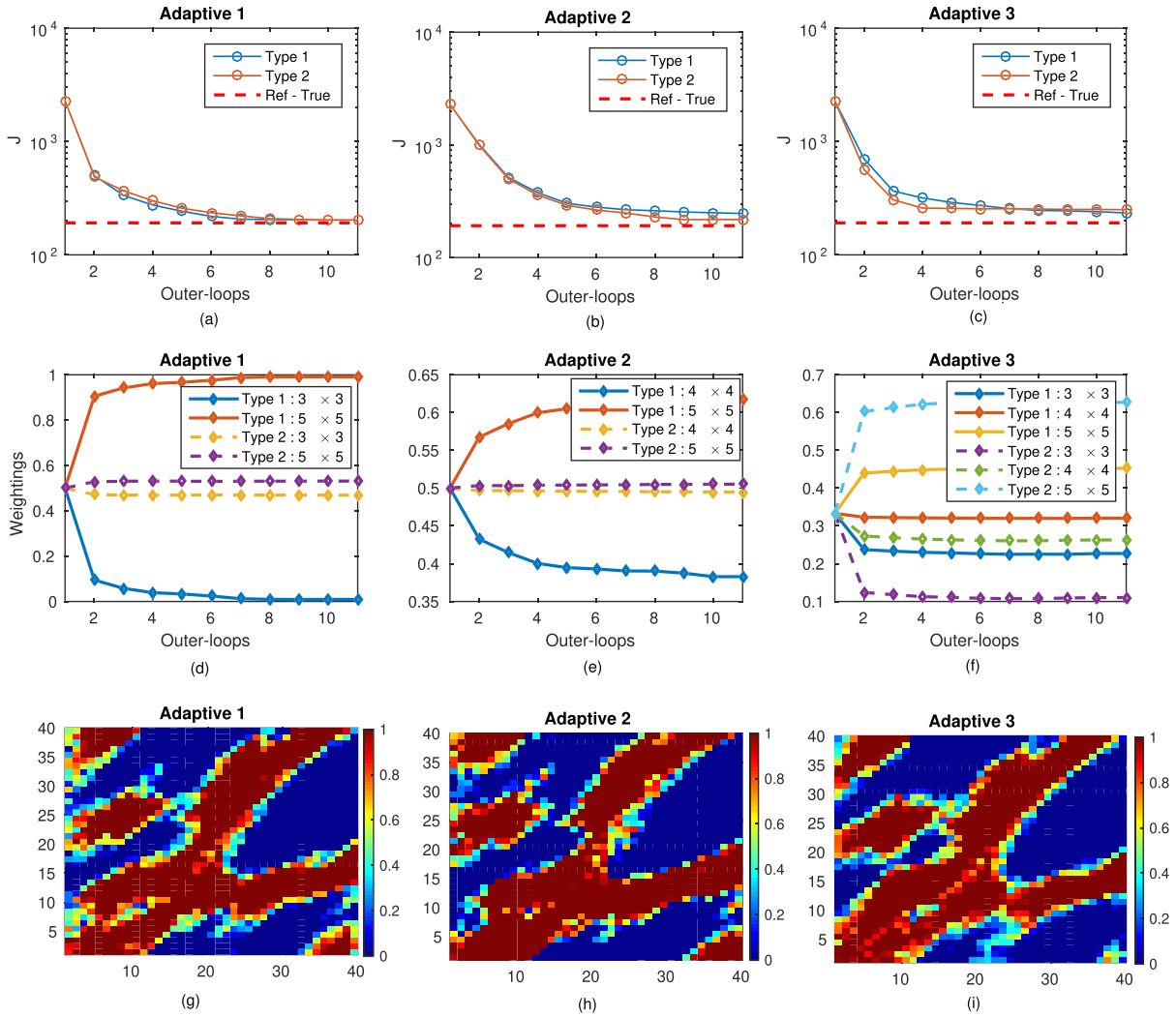

Figure 7. The evolution of objective functions and the weighting coefficients as a function of the iterations obtained with different adaptive scheme for Case 1. From left side to the right side, these three columns separately show the results of Adaptives 1, 2, and 3.

5.2. Case 2: High-Dimensional Modeling With Gaussian Parameters

In this case, SLM-SLP is used to estimate a large number of parameters by assimilating seismic data for a real-world model (Matthews et al., 2008). This model also describes a water-oil two phase system. Six producers and seven injectors, labeled from P_1 to P_6 and I_1 to I_7, respectively, are located in this model. Other settings about the model geometry, fluid properties, and well controls are shown in Table 2. We manually generate 1,000 Gaussian-distributed realizations of log-permeability fields. The log-permeability fields are described by the following statistics:

$$\mathbf{C}_{\boldsymbol{\beta}}(x_{i1,j1}; y_{i2,j2}] = \sigma_{\boldsymbol{\beta}}^2 e^{-\left[\left(\frac{|x_{i1}-x_{i2}|}{\chi_x}\right)^2+\left(\frac{|y_{i1}-y_{i2}|}{\chi_y}\right)^2\right]}, \sigma_{\boldsymbol{\beta}} = 5, \frac{\chi_x}{L_x} = 0.2, \frac{\chi_y}{L_y} = 0.2. \tag{19}$$

Here, $\sigma_{\boldsymbol{\beta}}$ is the standard deviation of log-permeability $\boldsymbol{\beta}$. $x_{i1,j1} = (x_{i1}, y_{j1})$ denotes the coordinates of a grid block; χ_x (or χ_y) is the correlation length in vertical (or horizontal) direction. L_x (or L_y) is the domain length in vertical (or horizontal) direction. In the global domain 296 GP coefficients are retained. The reference values for six relative permeability parameters are shown in Table 2. There are in total 302 uncertain parameters for Case 2.

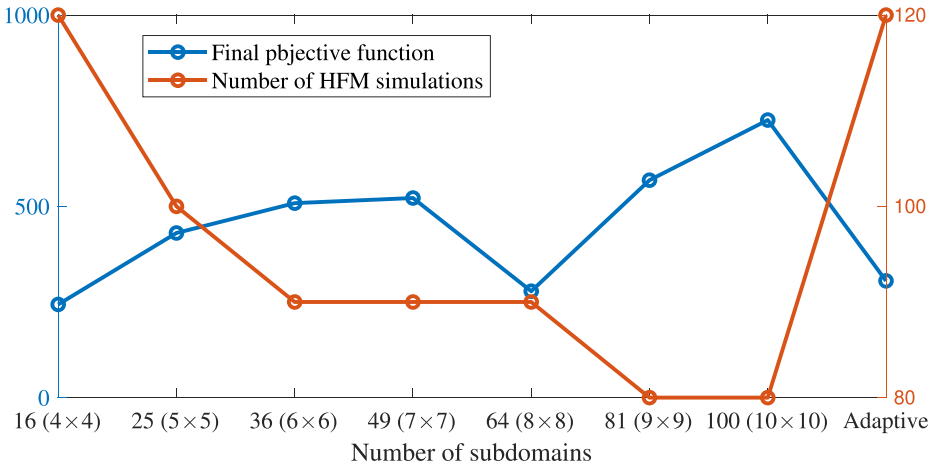

Figure 8. The final objective function values and number of HFM simulations as a function of number of subdomains using SLM-LP for Case 1.

The Gassmann model (Gassmann, 1951) is often used to convert the reservoir properties and the simulator primary variables (pressure and phase saturation) into seismic data, while we will directly assimilate the grid-based water saturation in this paper. The synthetic seismic data are generated based on the true model. The seismic data are collected at two monitor surveys after 1,825 days (first monitor) and 3,650 days (second monitor). Thus, there are in total 10,700 measurements. The noised measurements for the two monitors are shown in Figure 9.

5.2.1. Study of a Base Case

Two DD schemes, that is, 4×5 and 5×6, and an adaptive scheme where 4×5 and 5×6 are simultaneously considered, are compared. In this high-dimensional model, 15 and 10 LP coefficients need to be uniformly retained in each subdomain for these two DD schemes, respectively. Figure 10a shows the evolution of objective function values as a function of out-loop iterations. The 4×5 DD and the adaptive scheme yield comparable objective function values, which are better than 5×6 DD. This result indicates that 4×5 is a relatively better DD scheme in this case study. Both 4×5 DD and the adaptive scheme need 220 HFM simulations, while GP-FD requires 6,416 HFM simulations. There is an approximate 30 times speedups expressed in number of HFM runs.

Figures 10b–10g depict that the true log-permeability fields can be reproduced accurately when a large number of measurements is available. Increasing the number of available measurements does not perfectly reproduce the reference values of the relative permeability parameters as shown in Table 4. One possible explanation for this result is that the seismic data have strongly nonlinear relationship with the relative permeability parameters; the subdomain linear models cannot capture this nonlinearity and induce a bias estimation. Alternative nonlinear surrogate models or bias correction procedure deserve to be further

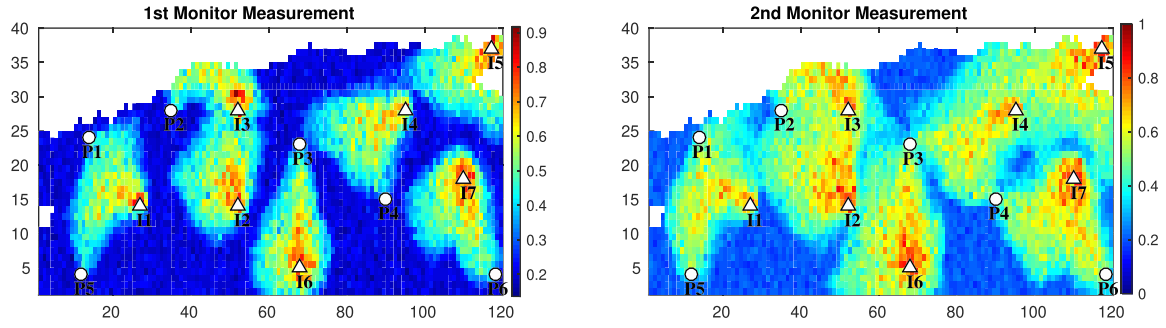

Figure 9. Noise distribution of water saturation for Case 2. Normal distributed independent measurement noise with a standard deviation equal to 5% of the "true" data value was added to all observations.

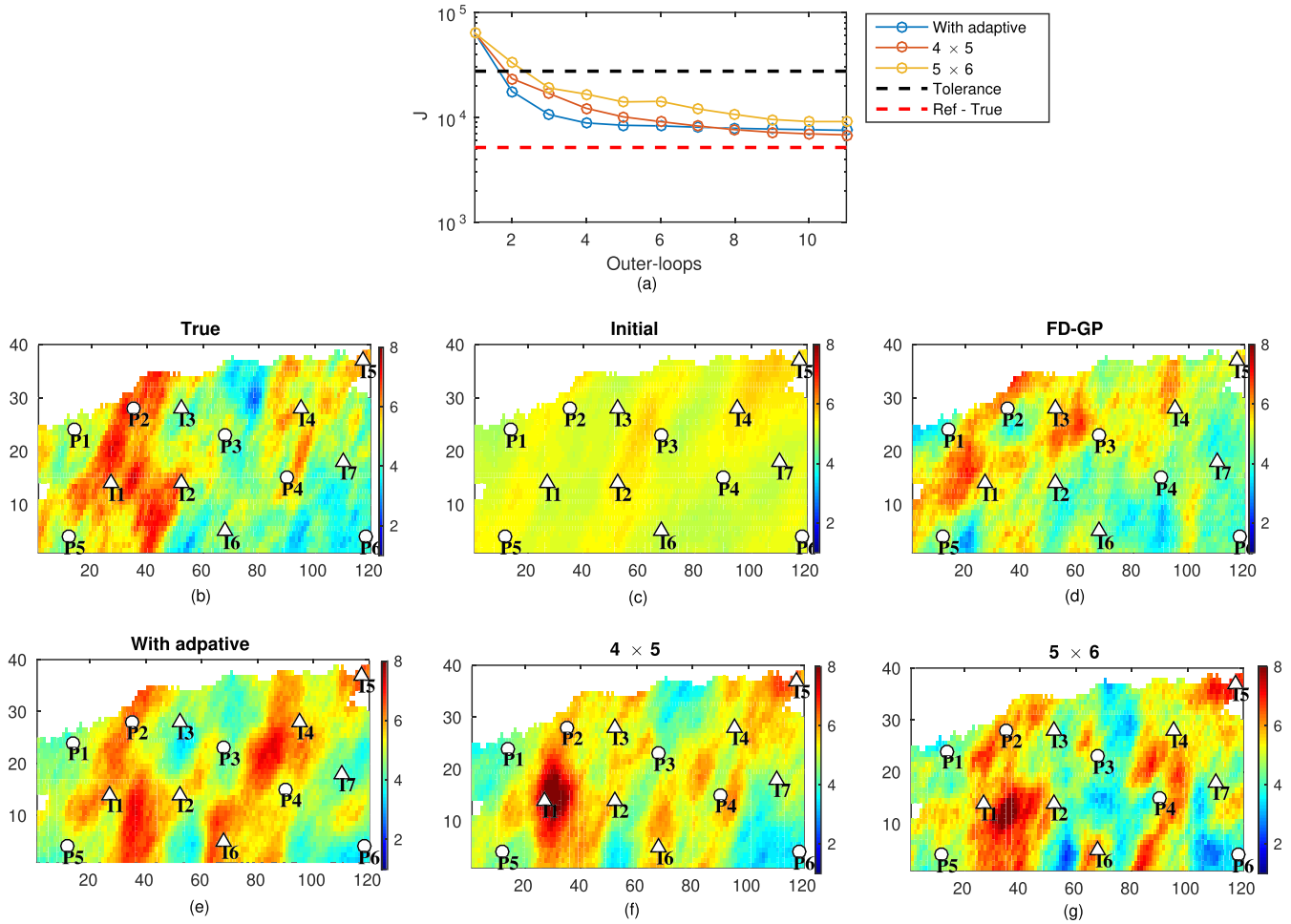

Figure 10. The first subfigure shows the evolution of objective function values as a function of outer-loops for Case 2. The computation of the objective function uses HFM simulations as Equation 3. The remaining subfigures show the updated log-permeability using SLM-SLP and FD-GP.

investigated for an unbiased estimation of the relative permeability parameters. Figure 11 illustrates the cross-plots between observed and predicted water saturation at these two monitoring surveys. The correlation coefficients R^2 have been increased from initial 60% to updated 94%, which quantify the improvements of model prediction.

5.2.2. Comparison of SLM-SLP and ES-MDA for Uncertainty Quantification

Assimilating a large amount of measurements, for example, totally 10,700 seismic data in this case, is a challenging task for the ensemble-based data assimilation algorithms. Due to the limited computational

Table 4
The Estimated Relative Permeability Parameters for Case 2

Parameters	$\alpha = [S_{wir}, S_{or}, k^0_{rw}, k^0_{ro}, n_w, n_o]^T$	
Reference value	[0.15, 0.15, 0.9, 0.9, 3, 3]	
-	Estimated value	Relative error (%)
With adaptive	[0.182, 0.201, 0.865, 0.858, 3.62, 3.84]	[34, 3.9]
4 × 5	[0.183, 0.199, 0.861, 0.860, 3.59, 3.88]	[32.7, 4.0]
5 × 6	[0.191, 0.218, 0.842, 0.867, 3.71, 3.64]	[45.3, 3.7]
FD-GP	[0.189, 0.191, 0.855, 0.864, 3.44, 3.80]	[27.3, 4.0]

Note. The last column shows the maximum and minimum relative errors.

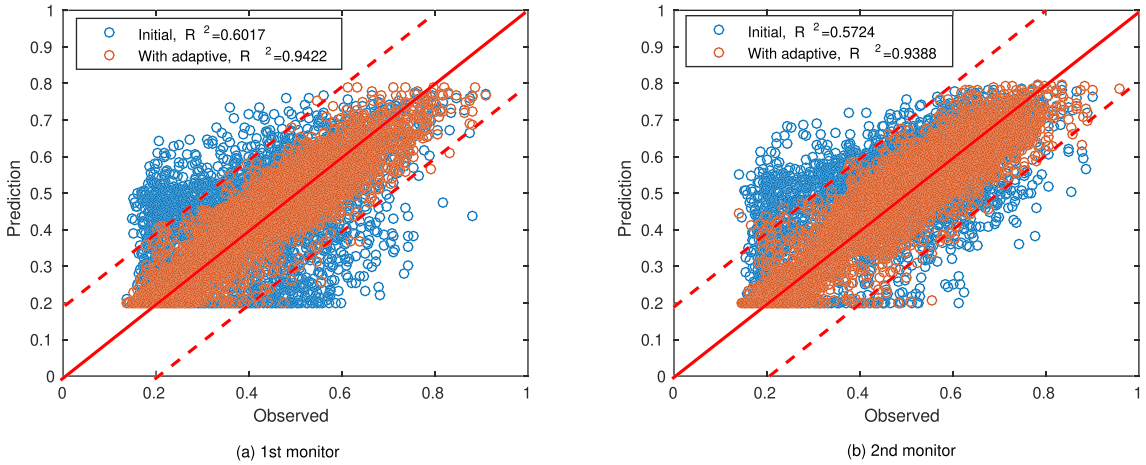

Figure 11. The cross-plots between observed and predicted water saturation for Case 2. The dashed lines correspond to ±2 standard deviations of the measurement errors.

resource, only a small ensemble size is computationally affordable in the real-world applications. Unfortunately, the small ensemble size easily yields an ensemble collapse problem. Numerical experiments have demonstrated that the SLM-SLP outperforms one variant of iterative ensemble smoother, for example, ES-MDA.

ES-MDA repeatedly assimilates the seismic data $\mathbf{N}_a = 10$ times. Two different ensemble size, that is, $\mathbf{N}_e = 100$, 500, are used in ES-MDA. SLM-SLP is implemented with an adaptive scheme by simultaneously considering 4×5 and 5×6 DD schemes. We repeatedly implement SLM-SLP to generate 100 posterior realizations. To simplify the notations, ES-MDA and SLM-SLP with 100 and 500 ensemble sizes are separately referred to as ES-MDA(100), ES-MDA(500), and SLM-SLP(100).

Figure 12 shows a set of boxplots for the optimized objective function values, RMSE_{mu} of log-permeability, and the updated relative permeability parameters using ES-MDA(100), ES-MDA(500), and SLM-SLP(100).

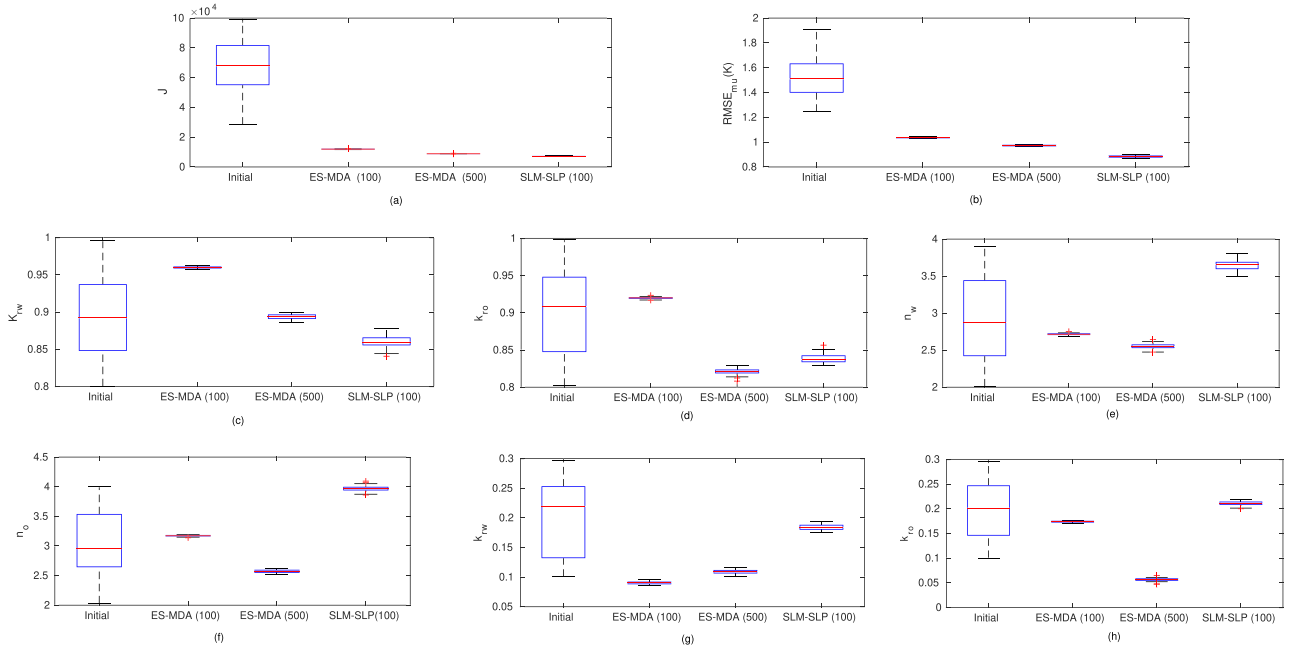

Figure 12. Boxplots of data mismatch and RMSE of the updated log-permeability for Case 2. The dash pink lines denote the reference value of relative permeability parameters. Here the uppercase "K" represents the log-permeability.

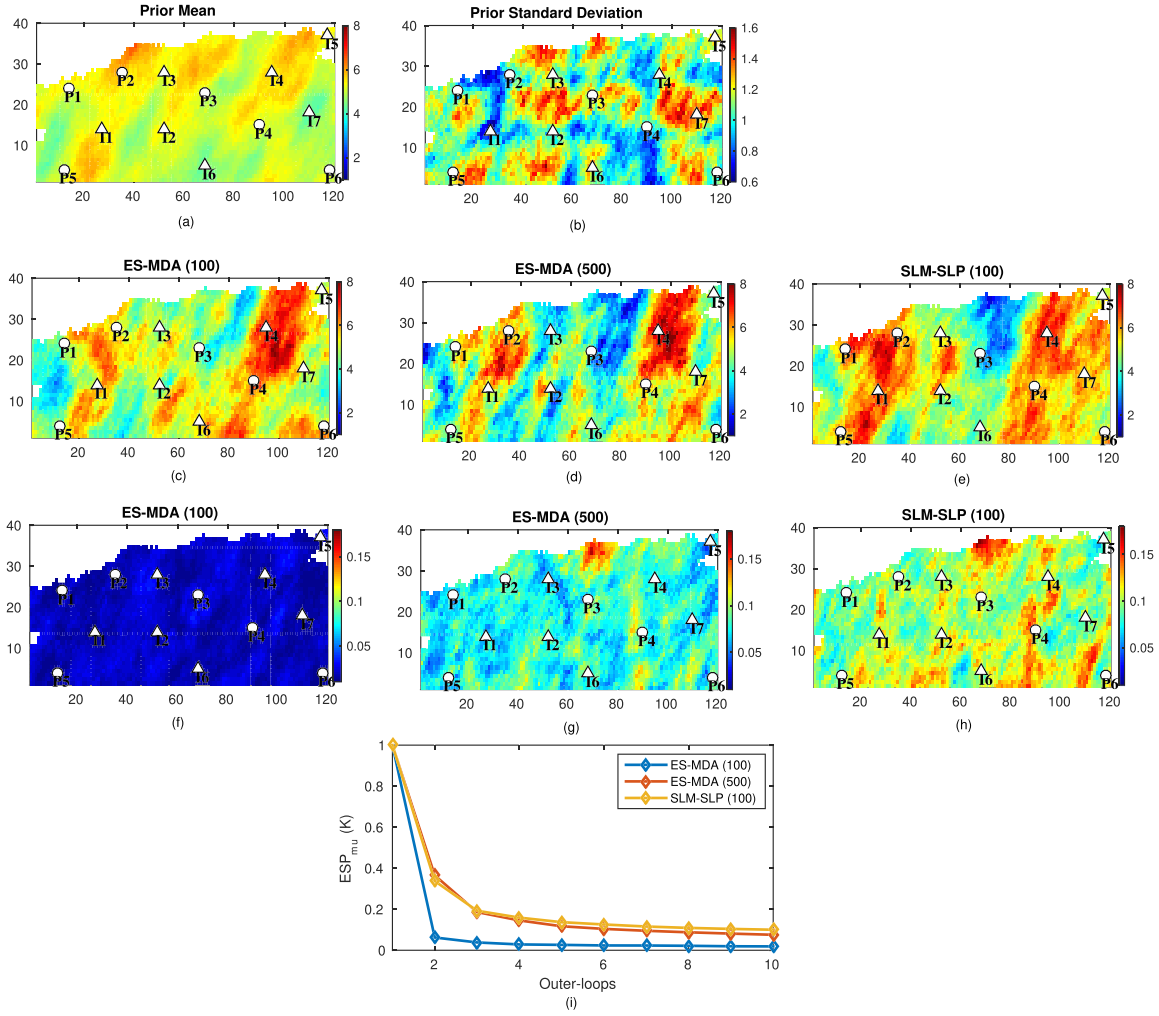

Figure 13. Updated log-permeability for ES-MDA(100), ES-MDA(500), and SLM-SLP(100). (a) The prior ensemble mean, (b) the prior standard deviation, (c–e) the posterior mean maps, (f–h) the posterior standard deviation maps, and (i) the ensemble spread (ESP).

As illustrated in Figures 12a and 12b, the optimized objective function values and RMSE$_{mu}$ of log-permeability have been drastically decreased. The SLM-SLP(100) outperforms ES-MDA(100) in terms of generating smaller objective function values and RMSE$_{mu}$ of log-permeability. To obtain 100 posterior solutions, SLM-SLP(100) totally requires $320 = 100+10 \times (15 + 6 + 1)$ HFM simulations, while ES-MDA(100) needs $1{,}000 = 10 \times 100$ HFM evaluations. The SLM-SLP(100) achieves about three times speedups expressed in number of HFM runs while increasing the estimation accuracy. The red dash lines represent the reference values of the relative permeability parameters in Figures 12c–12h. The uncertainty of these six relative permeability parameters has been significantly decreased as well.

We can primarily observe that an ensemble collapse has occurred in ES-MDA(100). The ensemble average spread (ESP$_{mu}$) is almost vanished after assimilating the seismic data as illustrated in Figure 13i, indicating an ensemble collapse. The final ESP is only 1.2% of its initial value, showing 98.2% reduction in ESP after ten data assimilation steps. The final variance map in Figure 13f confirms this observation. To further guarantee whether the underestimated ESP in ES-MDA(100) is caused by a small ensemble size, for example, 100, Figure 13g shows the same results as in Figure 13f, but using ensemble size 500 (instead of 100). The performance of ES-MDA has been improved as the ensemble size increases. As seen in Figure 13g, ESP$_{mu}$ has been significantly improved to 13.4% as the ensemble size increases. By contrast, the SLM-SLP(100) generates a comparable ESP, for example, 14.2%, with ES-MDA(500) procedure. However, the total computational

cost of SLM-SLP(100) is much smaller than ES-MDA(500) where 5,000 $= 10 \times 500$ HFM evaluations are run. These results further demonstrate that the SLM-SLP has a promising advantage to effectively and efficiently mitigate this ensemble collapse problem.

In comparison to the first case, the number of GP coefficients has increased from 72 to 296; however, the required number of HFM simulations has only increased from 150 to 220. The degree of freedom for the inversion problems depends on the number of GP coefficients, while the required HFM simulations are determined by the number of LP coefficients in individual subdomains. It is quite efficient to increase the degree of freedom by adding LP coefficients in all subdomains simultaneously. Taking 5×6 DD as an example, SLP enables us to retain another 30 GP coefficients by adding one LP coefficient in each subdomain, and only 20 additional HFM simulations are added to the entire parameter estimation procedure. Specifically, only two HFM simulations are added to construct subdomain linear models for each outer-loop iteration. The underestimation of ESP_{mu} may be partially alleviated by increasing the ensemble size in ES-MDA; however, the computational cost will be drastically increased as well. These numerical results further demonstrate that introducing SLP makes subdomain linear model highly scalable and the required HFM simulations do not grow rapidly with the increasing of the number of parameters. This advantage is particularly useful to assimilate a large amount of measurements where we should pose a large degree of freedom for avoiding a lowly effective usage of the informative measurements.

6. Conclusions

An efficient subdomain model linarization with a SLP where the global Gaussian and non-Gaussian parameters are projected onto low-order subspaces in each subdomain individually is presented to accelerate the high-dimensional inverse problem. This SLP enables us to construct subdomain linear models for each subdomain using a small number of HFMs. The use of SLP drastically reduces the computational cost in the high-dimensional inverse problems since the number of HFM simulations depends primarily on the number of the local parameters and not on the dimension of the underlying global parameters.

This new approach has been tested on a synthetic two-dimensional non-Gaussian binary facies model and a real-world high-dimensional Gaussian model. The first case demonstrates the ability to reconstruct the binary features of the "true" facies models and generate comparable results to a FD-based parameter estimation procedure. The second case shows that also for a complex inversion problem where spatially dense seismic data are used to estimate 302 parameters, very promising results have been obtained. Using more subdomains results in fewer local parameter patterns and hence enables us to run fewer HFM simulations. The number of HFM simulations required is roughly the multiplication of iteration steps and the maximum number of local parameter patterns among all subdomains. In comparison to one variant of iterative ensemble smoother, for example, ES-MDA, increasing one ensemble member only adds at most one degree of freedom for ES-MDA, while this SLP is quite efficiency to increase the degree of the freedom by adding local PCA patterns among all subdomains simultaneously. This outstanding advantage makes our proposed method easily mitigate an ensemble collapse problem.

There are a number of aspects of the proposed methodology that could possibly be improved. An obvious shortage of the proposed adaptive scheme is that the optimal one is limited to our predefined dictionary of DD schemes and therefore the performance of this new approach strongly depends on the design of this dictionary. In addition, a surrogate model inevitably introduces approximation errors, and if these errors are not modeled during the parameter inference step, the posterior distributions of the model parameters are likely to be biased. Incorporation of approximation error in the process of parameter inference also deserves to be explored in the future.

Acknowledgments

We thank the research funds by China Scholarship Council (CSC). Additional Computing resources were provided by Mathematics Physics Group, Department of Applied Mathematics at Delft University of Technology. The authors acknowledge TNO for the SAIGUP data release. The Matlab codes and data used will be made available soon upon publication of this manuscript. Code used to generate the figures in this paper is available at https://github.com/xclmj/SLM-SLP.

Data Availability Statement

The Matlab codes and data used will be made available soon upon publication of this manuscript. Code used to generate the figures in this paper is available online (at https://github.com/xclmj/SLM-SLP).

References

Adam, J. S., Cui, T., Karelse, R. N., & Hampton, C. (2020). Reduced-dimensional Gaussian process machine learning for groundwater allocation planning using swarm theory. *Water Resources Research*, *56*, e2019WR026061. https://doi.org/10.1029/2019WR026061

Altaf, M. U., Heemink, A. W., & Verlaan, M. (2009). Inverse shallow-water flow modeling using model reduction. *International Journal for Multiscale Computational Engineering*, *7*(6), 77–594.
Amsallem, D., & Farhat, C. (2008). Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA Journal*, *46*(7), 1803–1813.
Asher, M. J., Croke, B. F., Jakeman, A. J., & Peeters, L. J. (2015). A review of surrogate models and their application to groundwater modeling. *Water Resources Research*, *51*, 5957–5973. https://doi.org/10.1002/2015WR016967
Baiges, J., Codina, R., & Idelsohn, S. (2013). A domain decomposition strategy for reduced order models. Application to the incompressible Navier-Stokes equations. *Computer Methods in Applied Mechanics and Engineering*, *267*, 23–42.
Baur, U., Benner, P., & Feng, L. (2014). Model order reduction for linear and nonlinear systems: A system-theoretic perspective. *Archives of Computational Methods in Engineering*, *21*(4), 331–358.
Chang, H., Liao, Q., & Zhang, D. (2017). Surrogate model based iterative ensemble smoother for subsurface flow data assimilation. *Advances in Water Resources*, *100*, 96–108.
Chen, C., Gao, G., Honorio, J., Gelderblom, P., & Jimenez, E. (2014). Integration of principal-component-analysis and streamline information for the history matching of channelized reservoirs. In *SPE Technical Conference and Exhibition*. USA: Society of Petroleum Engineers.
Chen, Y., & Oliver, D. S. (2010). Ensemble-based closed-loop optimization applied to Brugge field. *SPE Reservoir Evaluation & Engineering*, *13*(01), 56–71.
Chen, Y., & Oliver, D. S. (2012). Multiscale parameterization with adaptive regularization for improved assimilation of nonlocal observation. *Water Resources Research*, *48*, W04503. https://doi.org/10.1029/2011WR011144
Dai, C., Xue, L., Zhang, D., & Guadagnini, A. (2016). Data-worth analysis through probabilistic collocation-based ensemble Kalman filter. *Journal of Hydrology*, *540*, 488–503.
Ding, Y. D. (2011). Development of a data partition technique for gradient-based optimization methods in history matching. *Spe Journal*, *16*(16), 4133–42.
Ding, D. Y., & Mckee, F. (2013). Study of perturbation designs for the partially separable objective function with gradient-based optimizations in history matching. *Spe Journal*, *18*(3), 495–507.
Emerick, A. A., & Reynolds, A. C. (2012). History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. *Computational Geosciences*, *16*(3), 639–659.
Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, *55*, 3–15.
Evensen, G. (2009). *Data assimilation: The ensemble Kalman filter*. Springer Science & Business Media.
Feeley, R. P. (2008). *Fighting the curse of dimensionality: A method for model validation and uncertainty propagation for complex simulation models*. Berkeley: University of California.
Gassmann, F. (1951). Elastic waves through a packing of spheres. *Geophysics*, *16*(4), 673–685.
Gong, W., Duan, Q., Li, J., Wang, C., Di, Z., Ye, A., et al. (2016). Multiobjective adaptive surrogate modeling-based optimization for parameter estimation of large, complex geophysical models. *Water Resources Research*, *52*, 1984–2008. https://doi.org/10.1002/2015WR018230
Hombal, V., & Mahadevan, S. (2011). Bias minimization in Gaussian process surrogate modeling for uncertainty quantification. *International Journal for Uncertainty Quantification*, *1*(4), 321–349.
Hu, J., Chen, S., Behrangi, A., & Yuan, H. (2019). Parametric uncertainty assessment in hydrological modeling using the generalized polynomial chaos expansion. *Journal of Hydrology*, *579*, 124,158.
Jafarpour, B., & McLaughlin, D. B. (2008). History matching with an ensemble Kalman filter and discrete cosine parameterization. *Computational Geosciences*, *12*(2), 227–244. https://doi.org/10.1007/s10596-008-9080-3
Ju, L., Zhang, J., Meng, L., Wu, L., & Zeng, L. (2018). An adaptive Gaussian process-based iterative ensemble smoother for data assimilation. *Advances in water resources*, *115*, 125–135.
Kaleta, M. P., Hanea, R. G., Heemink, A. W., & Jansen, J.-D. (2011). Model-reduced gradient-based history matching. *Computational Geosciences*, *15*(1), 135–153.
Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., & Jacques, D. (2013). Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. *Water Resources Research*, *49*, 2664–2682. https://doi.org/10.1002/wrcr.20226
Li, H., Chen, S. N., Yang, D. T., Tontiwachwuthikul, P., et al. (2012). Estimation of relative permeability by assisted history matching using the ensemble Kalman filter method. *Journal of Canadian Petroleum Technology*, *51*(03), 205–214.
Li, K., & Horne, R. N. (2006). Comparison of methods to calculate relative permeability from capillary pressure in consolidated water-wet porous media. *Water Resources Research*, *42*, W06405. https://doi.org/10.1029/2005WR004482
Li, H., & Zhang, D. (2007). Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods. *Water Resources Research*, *43*, W09409. https://doi.org/10.1029/2006WR005673
Li, L., Zhou, H., Gómez-Hernández, J. J., & Franssen, H.-J. H. (2012). Jointly mapping hydraulic conductivity and porosity by assimilating concentration data via ensemble Kalman filter. *Journal of Hydrology*, *428*, 152–169.
Lie, K.-A., Krogstad, S., Ligaarden, I. S., Natvig, J. R., Nilsen, H. M., & Skaflestad, B. (2012). Open-source MATLAB implementation of consistent discretisations on complex grids. *Computational Geosciences*, *16*(2), 297–322.
Lu, D., Ricciuto, D., Stoyanov, M., & Gu, L. (2018). Calibration of the E3SM land model using surrogate-based global optimization. *Journal of Advances in Modeling Earth Systems*, *10*, 1337–1356. https://doi.org/10.1002/2017MS001134
Matthews, J. D., Carter, J. N., Stephen, K. D., Zimmerman, R. W., Skorstad, A., Manzocchi, T., & Howell, J. A. (2008). Assessing the effect of geological uncertainty on recovery estimates in shallow-marine reservoirs: The application of reservoir engineering to the SAIGUP project. *Petroleum Geoscience*, *14*(1), 35–44.
Mo, S., Lu, D., Shi, X., Zhang, G., Ye, M., Wu, J., & Wu, J. (2017). A Taylor expansion-based adaptive design strategy for global surrogate modeling with applications in groundwater modeling. *Water Resources Research*, *53*, 10,802–10,823. https://doi.org/10.1002/2017WR021622
Nocedal, J., & Wright, S. J. (1999). *Numerical optimization*. New York, NY: Springer.
Pasetto, D., Putti, M., & Yeh, W. W.-G. (2013). A reduced-order model for groundwater flow equation with random hydraulic conductivity: Application to Monte Carlo methods. *Water Resources Research*, *49*, 3215–3228. https://doi.org/10.1002/wrcr.20136
Piret, C., Dissanayake, N., Gierke, J. S., & Fornberg, B. (2019). The radial basis functions method for improved numerical approximations of geological processes in heterogeneous systems. *Mathematical Geosciences*, *52*, 477–497.

Razavi, S., Tolson, B. A., & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, *48*, W07401. https://doi.org/10.1029/2011WR011527

Roy, P. T., El Moçayd, N., Ricci, S., Jouhaud, J.-C., Goutal, N., De Lozzo, M., & Rochoux, M. C.(2018). Comparison of polynomial chaos and gaussian process surrogates for uncertainty quantification and correlation estimation of spatially distributed open-channel steady flows. *Stochastic Environmental Research and Risk Assessment*, *32*(6), 1723–1741.

Saad, G., & Ghanem, R. (2009). Characterization of reservoir simulation models using a polynomial chaos-based ensemble Kalman filter. *Water Resources Research*, *45*, W04417. https://doi.org/10.1029/2008WR007148

Seiler, A., Evensen, G., Skjervheim, J.-A., Hove, J., & Vabo, J. G. (2009). Advanced reservoir management workflow using an ENKF based assisted history matching method. In *SPE Reservoir Simulation Symposium*. USA: Society of Petroleum Engineering.

Vermeulen, P., & Heemink, A. (2006). Model-reduced variational data assimilation. *Monthly Weather Review*, *134*(10), 2888–2899.

Vo, H. X., & Durlofsky, L. J. (2014). A new differentiable parameterization based on principal component analysis for the low-dimensional representation of complex geological models. *Mathematical Geosciences*, *46*(7), 775–813. https://doi.org/10.1007/s11004-014-9541-2

Xiao, C., Leeuwenburgh, O., Lin, H. X., & Heemink, A. (2018). Non-intrusive subdomain POD-TPWL for reservoir history matching. *Computational Geosciences*, *23*, 537–565.

Xiao, C., Leeuwenburgh, O., Lin, H. X., & Heemink, A. (2019). Subdomain POD-TPWL with local parameterization for large-scale reservoir history matching problems. arXiv preprint arXiv:1901.08059.

Yang, J., Jakeman, A., Fang, G., & Chen, X. (2018). Uncertainty analysis of a semi-distributed hydrologic model based on a gaussian process emulator. *Environmental Modelling & Software*, *101*, 289–300.

Zeng, L., Shi, L., Zhang, D., & Wu, L. (2012). A sparse grid based Bayesian method for contaminant source identification. *Advances in Water Resources*, *37*, 1–9.

Zeng, L., & Zhang, D. (2010). A stochastic collocation based Kalman filter for data assimilation. *Computational Geosciences*, *14*(4), 721–744.

Zha, Y., Yeh, T.-C. J., Illman, W. A., Zeng, W., Zhang, Y., Sun, F., & Shi, L. (2018). A reduced-order successive linear estimator for geostatistical inversion and its application in hydraulic tomography. *Water Resources Research*, *54*, 1616–1632. https://doi.org/10.1002/2017WR021884

Zhang, J., Li, W., Lin, G., Zeng, L., & Wu, L. (2017). Efficient evaluation of small failure probability in high-dimensional groundwater contaminant transport modeling via a two-stage Monte Carlo method. *Water Resources Research*, *53*, 1948–1962. https://doi.org/10.1002/2016WR019518

Zhang, J., Man, J., Lin, G., Wu, L., & Zeng, L. (2018). Inverse modeling of hydrologic systems with adaptive multifidelity Markov chain Monte Carlo simulations. *Water Resources Research*, *54*, 4867–4886. https://doi.org/10.1029/2018WR022658

Zhang, J., Zheng, Q., Chen, D., Wu, L., & Zeng, L. (2020). Surrogate-based Bayesian inverse modeling of the hydrological system: An adaptive approach considering surrogate approximation error. *Water Resources Research*, *56*, e2019WR025721. https://doi.org/10.1029/2019WR025721

Zhou, J., Su, X., & Cui, G. (2018). An adaptive Kriging surrogate method for efficient joint estimation of hydraulic and biochemical parameters in reactive transport modeling. *Journal of Contaminant Hydrology*, *216*, 50–57. https://doi.org/10.1016/j.jconhyd.2018.08.005