energies

MDPI

Article

Gaussian Processes Proxy Model with Latent Variable Models and Variogram-Based Sensitivity Analysis for Assisted History Matching

Dongmei Zhang [1,†], Yuyang Zhang [1], Bohou Jiang [2], Xinwei Jiang [1,*] and Zhijiang Kang [3]

[1] School of Computer Science, Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430078, Hubei, China; cugzdm@foxmail.com (D.Z.); cug_zyy@163.com (Y.Z.)
[2] School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, Hubei, China; jiangbohou123@outlook.com
[3] Petroleum Exploration and Production Research Institute of SINOPEC (PEPRIS), Beijing 100831, China; kangzj.syky@sinopec.com
* Correspondence: ysjxw@hotmail.com; Tel.: +86-135-1721-1621
† Current Address: Building Kejiao No.1, No. 68, Jinchen Road, Donghu New Technology Development Zone, Wuhan 430078, Hubei, China.

Received: 19 July 2020; Accepted: 17 August 2020; Published: 19 August 2020

Abstract: Reservoir history matching is a well-known inverse problem for production prediction where enormous uncertain reservoir parameters of a reservoir numerical model are optimized by minimizing the misfit between the simulated and history production data. Gaussian Process (GP) has shown promising performance for assisted history matching due to the efficient nonparametric and nonlinear model with few model parameters to be tuned automatically. Recently introduced Gaussian Processes proxy models and Variogram Analysis of Response Surface-based sensitivity analysis (GP-VARS) uses forward and inverse Gaussian Processes (GP) based proxy models with the VARS-based sensitivity analysis to optimize the high-dimensional reservoir parameters. However, the inverse GP solution (GPIS) in GP-VARS are unsatisfactory especially for enormous reservoir parameters where the mapping from low-dimensional misfits to high-dimensional uncertain reservoir parameters could be poorly modeled by GP. To improve the performance of GP-VARS, in this paper we propose the Gaussian Processes proxy models with Latent Variable Models and VARS-based sensitivity analysis (GPLVM-VARS) where Gaussian Processes Latent Variable Model (GPLVM)-based inverse solution (GPLVMIS) instead of GP-based GPIS is provided with the inputs and outputs of GPIS reversed. The experimental results demonstrate the effectiveness of the proposed GPLVM-VARS in terms of accuracy and complexity. The source code of the proposed GPLVM-VARS is available at https://github.com/XinweiJiang/GPLVM-VARS.

Keywords: production optimization; history matching; reservoir simulation; proxy model; gaussian process

1. Introduction

As a well-known inverse problem in reservoir simulation, History Matching is significant for reservoir development, management and predictions, which tries to estimate the uncertain parameters (such as porosity and permeability) of a reservoir numerical model based on observed historical production data (such as well rates and pressure) [1–3]. Theoretically, simulated reservoirs models with unknown parameters are calibrated by minimizing the misfits between the simulated and history data, which can be used to forecast the reservoirs production and provide support decisions under different operating conditions and in different production stages. As the reservoirs become large

and heterogeneous, the highly nonlinear reservoirs models must be elaborate with numerous grid blocks and reservoir parameters, which brings enormous computational burden and complexity for the numerical optimization [4].

To address the issue, Assisted History Matching (AHM) techniques have been proposed to replace labor-intensive and costly manual history matching [1,4–6]. Roughly, these methods for assisted history matching can be divided into two categories [7]: the data assimilation approaches (such as Ensemble Kalman Filter and Ensemble Smoother) and the optimization approaches (such as gradient, evolutionary or data-driven-based algorithms). Ensemble Kalman Filter (EnKF) and Ensemble Smoother (ES) are representative methods for data assimilation [8]. For example, EnKF is a sequential Monte Carlo approximation of the Kalman filter where the correlation between reservoir parameters and observed production data can be estimated from the ensemble with the uncertainty of estimation [9]. EnKF can efficiently assimilate various types of data to optimize numerous reservoir parameters, but it could fail if there are multimodal nonlinear data or discrete reservoir parameters. To address the limitations, various EnKF extensions were developed through localization, ensemble design scheme and clustering methods, etc. [8,10]. In contrast to the iterative process of EnKF, ES could simultaneously assimilate all the production data in a global update, which is much faster than EnKF [11]. Different extensions of ES such as iterative ES, and ES with Multiple Data Assimilation (ES-MDA) had also been developed to boost the performance of data assimilation approaches [12].

Alternatively, the optimization methods for AHM had attracted people's great attention because history matching as an inverse problem can be naturally regarded as an optimization routine. For example, the gradients-based AHM models (such as Newton's method, Gauss-Newton, Quasi-Newton, gradient descent and conjugate gradient, etc.) [13,14] have been widely adopted to handle the history matching problems due to the advantages of high computational efficiency and fast convergence. However, the objective function must be continuous and differentiable which could be unfeasible for some reservoirs. Additionally, the solutions are prone to be the local optimum to the initialization. To address the issues, global optimization techniques based on Evolutionary Algorithms (EAs) had been introduced for AHM such as genetic algorithms, differential evolution and particle swarm optimization, because of their easy application to various problems without special assumptions [15–17]. In the framework of EAs, the objective function can be formulated as a single-objective function or as a multi-objective function, which could be effectively optimized even when there are discrete reservoir parameters, highly non-Gaussian distributed data, or non-differentiable objective functions. In contrast to the gradient-based models where a single solution is provided with the tendency to get stuck in local optimum, multiple solutions can be obtained from EAs-based methods which also means that global optimum could be obtained. However, EAs are slowly converging especially for large-scale reservoir numerical models.

Recently, the data-driven-based AHM models have been proposed with the rapid development of machine learning techniques [18–20]. The past decade has witnessed various proxy models for AHM based on machine learning algorithms. For example, Principal Component Analysis (PCA) [21,22], Artificial Neural Network (ANN) [23–25], Deep Learning (DL) [12,26,27], Support Vector Machine (SVM) [28] and Gaussian Process (GP) [7,29] have been introduced to replace computationally intensive numerical simulators. Optimization-based PCA and two-dimensional PCA were applied to characterize the channelized geological models [21,22], but they could fail when dealing with the complex and nonlinear channelized structures in some geological models although kernel tricks can be adopted. By contrast, Neural network-based proxy models like ANN [23–25], Stacked Autoencoder and Convolutional Neural Networks [12,26,27] are capable of handling highly nonlinear structure in the channelized reservoir models, but the network architecture and parameters are difficult to be tuned properly. Alternatively, the SVM and GP-based proxy models [28,29] could be more flexible than PCA, ANN and DL. From the perspective of machine learning models, these proxy models can be classified as parametric and nonparametric methods. Parametric models such as PCA, ANN, and DL are expressed by some finite set of parameters with specific model hypothesis where the parameters

capture everything from the observed data, while infinite parameters typically defined by functions are used in nonparametric methods with a few hyperparameters such as SVM and GP, which are typically more flexible than parametric models.

Compared to typical parametric model ANN and nonparametric method SVM, GP is a probabilistic and nonparametric model in Bayesian learning framework, which naturally provides soft prediction with confidence interval and only has few hyperparameters in covariance function to be automatically optimized. Thus, applying GP is easy for AHM without manually parameters tuning. Hamdi et al. [7] first introduced GP for history matching with single GP as the proxy model plus Expected Improvement based Bayesian optimization, and they concluded that the covariance function Matern class 3 provided the best results for some history matching problems according to the sensitivity analysis regarding various covariance functions. Sachin Rana et al. [29] further proposed the GP proxy models and Variogram Analysis of Response Surface (GP-VARS) by making use of two GPs to obtain forward (GPFS) and inverse (GPIS) solutions to improve the performance of a single GP-based proxy models for AHM where Empirical Bayes approach [30] was employed to choose and optimize the covariance function for any given data automatically. In addition, a novel application of Variogram Analysis of Response Surface (VARS)-based sensitivity analysis was introduced to calculate corresponding global relative sensitivity indices without the evaluation of one-dimensional variograms for each input parameter and higher-dimensional variograms.

Although GP-VARS provided convincing experimental results, the inverse GP models (GPIS) in GP-VARS where the inputs are the expected small misfit values and outputs are the uncertain reservoir parameters could be inappropriate, because the mapping function modeled by GPIS tries to nonlinearly transform the low-dimensional inputs (expected misfit values) to high-dimensional outputs (uncertain reservoir parameters), which is very challenging especially when the parameters are of high dimensionality.

To address the issue, motivated by the Gaussian Process Latent Variable Model (GPLVM) [31] for dimensionality reduction, we propose the GP proxy models with Latent Variable Model and Variogram Analysis of Response Surface (GPLVM-VARS) in this research, where the key idea is to make use of GPLVM to replace the inverse GP in GP-VARS. For the GPLVM-based inverse solutions (GPLVMIS), the inputs and outputs are reversed compared to original GPIS. Specifically, the inputs become the high-dimensional uncertain reservoir parameters and outputs are the expected small misfit values with low dimensionality, which could be more reasonable and accurate. The main contributions of this paper are two-fold: (1) With the novel inverse model GPLVMIS, GPLVM-VARS is proposed which could outperform GP-VARS in terms of accuracy and complexity. (2) Experiments conducted on a synthetic data and the well-known PUNQ-S3 case demonstrate the effectiveness of the proposed GPLVM-VARS.

The rest of the paper is organized as follows. In Section 2, we introduced the proposed GPLVM-VARS for AHM. Then, the cases study based on synthetic data and PUNQ-S3 reservoir are adopted to evaluate the performance of the newly proposed algorithm in Section 3 followed by discussions in Section 4. Finally, the concluding remarks and comments will be given in Section 5.

2. Materials and Methods

To address the limitations of GP-VARS especially the unsatisfactory performance of the embedded inverse model GPIS, we introduce the proposed GPLVM-VARS for AHM with the flowchart displayed in Figure 1, which also includes two GP-based forward and inverse models. Actually, it is similar to GP-VARS with the only difference in the GPLVM-based Inverse Solution (GPLVMIS). Except GPLVMIS, other modules are similar to GP-VARS with an iterative optimization process where some initial random set of reservoir parameters are initially generated by Latin hypercube sampling technique and then the temporary (or proposal) solutions in each iteration are estimated to minimize the misfits between the GPFS-based proxy model response and historical production data [29].

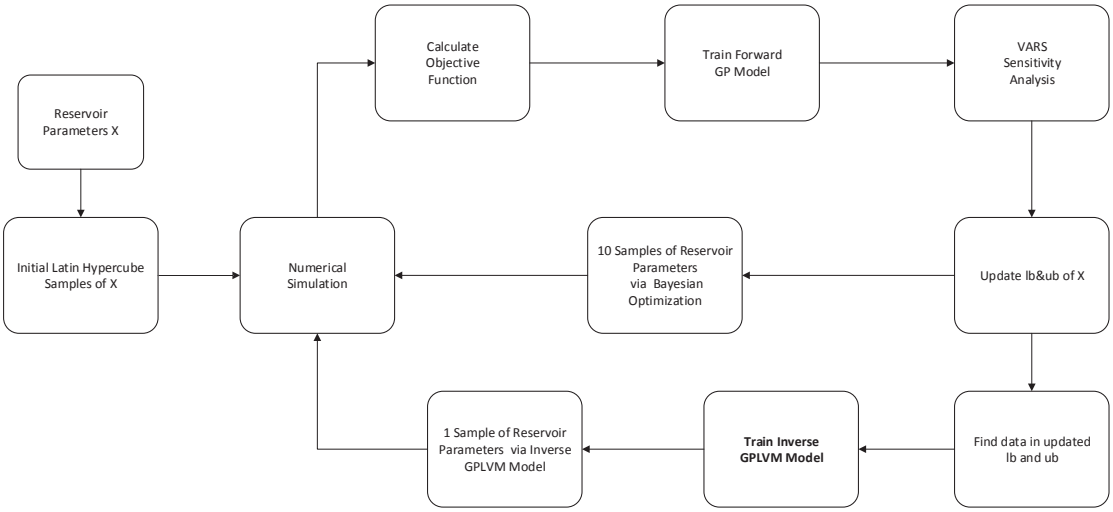


Figure 1. Flowchart of the proposed GPLVM-VARS for assisted history matching.

The objective functions which quantitatively measures the misfit values between the simulated output and history data are similarly defined by the Local Misfit Value (LMV) and Global Misfit Value (GMV) as follows

$$LMV = \left[\sum_{i=1}^{N_{pt}} \frac{1}{N_{pt}} \left[\frac{f_i^s - f_i^t}{f_i^t}\right]^2\right]^{0.5} * \frac{\sum_{i=1}^{N_{nt}} \left[f_i^s - f_i^t\right]}{\left|\sum_{i=1}^{N_{nt}} \left[f_i^s - f_i^t\right]\right|} \quad (1)$$

$$GMV = \frac{1}{N_l} \sum_{i=1}^{N_l} |LMV_i| \quad (2)$$

where N_{pt} is the number of samples, N_l is the number of LMVs, LMV_i is the i-th LVM, f_i^s and f_i^t are the the simulated output and target historical data, respectively. As can be seen from Equation (1) that LMV is an asymmetric objective function consisting of positive or negative Root Mean Squared Error (RMSE), and GMV is the mean of absolute values of all the LMVs.

2.1. Forward GP Model (GPFS)

As can be seen from Figure 2 the inputs and outputs of the GP proxy model in GPFS are the samples regarding the reservoir parameters and LMVs based on the simulated and target data, respectively. According to original GP-VARS there are some pre-processing steps based on normalization and standardization for the inputs and outputs data to rescale all the data to mean zero within the ranges of [0–1]. Then, instead of manually choosing optimal covariance function for GP, Empirical Bayes approach is employed to automatically select the best covariance function and optimize the corresponding hyperparameters.

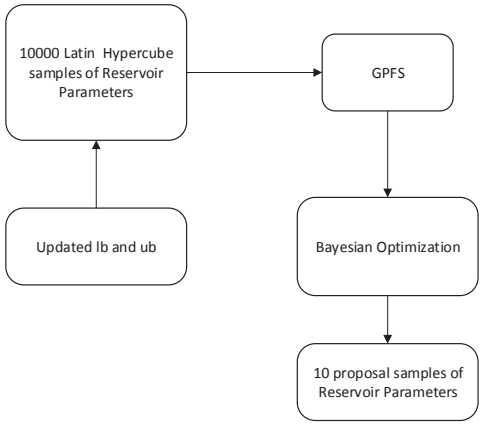


Figure 2. Flowchart of the GPFS for assisted history matching.

Specifically, given the pre-processed D-dimensional inputs reservoir parameters $X = [\boldsymbol{x}_1, ..., \boldsymbol{x}_N] \subset \mathcal{R}^{D\times N}$ and the corresponding outputs LVMs denoted by $Y = [y_1, ..., y_N] \subset \mathcal{R}^{C\times N}$ with N denoting the number of the samples. For the sake of convenience, we will simply assume the outputs LVMs are scalar ($C = 1$). In the classical Gaussian Process Regression (GPR) [32] model, each output variable y_n is assumed to be sampled from the unknown latent function f with independent Gaussian noise $y = f(\boldsymbol{x}) + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is with zero mean and covariance σ^2, leading to the likelihood $P(Y|X, f, \sigma) = \mathcal{N}(y|f, \sigma^2 I)$. The unknown latent function f is expressed by a zero-mean GP prior $P(f|X) = \mathcal{N}(f|0, K)$ where the covariance matrix K is defined on the input space with a few hyperparameters θ.

With Bayesian equation, we can obtain the posterior distribution over the unknown function f in Equation (3) as follows,

$$P(f|Y, X, \theta) = \frac{P(Y|X, f, \theta)P(f|X)}{P(Y|X, \theta)} \tag{3}$$

where $P(Y|X, \theta) = \mathcal{N}(Y|0, K + \sigma^2 I)$ is the marginal likelihood by integrating out the unknown function f, which can then be maximized regarding the hyperparameters $\theta = \{\sigma, \gamma\}$ in the covariance function.

The objective function regarding the GPR optimization becomes

$$\underset{\theta}{\operatorname{argmax}} \log(P(Y|X, \theta)) = \underset{\theta}{\operatorname{argmax}} -\frac{1}{2}(YK^{-1}Y^T + \log|K| + n\log 2\pi) \tag{4}$$

and Equation (4) can be optimized by gradient-based algorithms.

To conduct prediction based on the learnt GPR, the predictive distribution conditioned on the given observation data for a new testing sample $\boldsymbol{x}^*$ can be formulated in Equation (5) based on Bayesian framework as follows,

$$\begin{aligned} g^*|\boldsymbol{x}^*, X, Y \sim & \mathcal{N}(K_{\boldsymbol{x}^*X}(K_{XX} + \sigma^2\mathbf{I})^{-1}Y^T, \\ & K_{\boldsymbol{x}^*\boldsymbol{x}^*} - K_{\boldsymbol{x}^*X}(K_{XX} + \sigma^2\mathbf{I})^{-1}K_{X\boldsymbol{x}^*}) \end{aligned} \tag{5}$$

where $K_{XX}, K_{\boldsymbol{x}^*\boldsymbol{x}^*}, K_{\boldsymbol{x}^*X}$ and $K_{X\boldsymbol{x}^*}$ are the matrices of the covariance function values at the corresponding samples X and/or $\boldsymbol{x}^*$ with the learnt hyperparameters.

For GPFS, the optimal covariance function is automatically chosen from a set of covariance functions listed in Table 1 in [29] based on Empirical Bayes. Specifically, there are five single covariance functions including Cov1(Squared Exponential), Cov2(Matern Anisotropic), Cov3(Matern Isotropic), Cov4(Neural Network), Cov5(Rational Quadratic) and four combined covariance functions

Cov6(Cov1*Cov3), Cov7(Cov1*Cov2), Cov8(Cov1*Cov4) and Cov9(Cov1+Cov4). To automatically pick the optimal covariance function, the one which provides the lowest value of negative log marginal likelihood regarding Equation (4) is picked as the covariance function for the final GPFS model.

2.2. Vars-Based Sensitivity Analysis and Bayesian Optimization

Sensitivity analysis is typically employed to discover unimportant reservoir parameters that have a very small contribution to history matching models so that computational cost can be reduced. For a fair comparison, we similarly make use of the VARS-based sensitivity analysis in GP-VARS. Specifically, for D-dimensional input reservoir parameters $\boldsymbol{x}$ and scalar output y, the $D-$dimensional directional variogram is measured by Equation (6) as follows,

$$\gamma\left(\boldsymbol{h}_{lag}\right) = (1/2)V\left(y\left(\boldsymbol{x}+\boldsymbol{h}_{lag}\right) - y(\boldsymbol{x})\right) \tag{6}$$

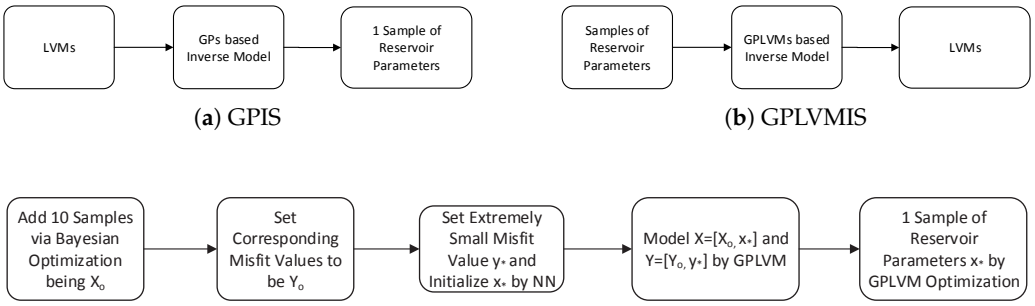
where $\boldsymbol{h}_{lag} = [\boldsymbol{h}_{lag_1}, \ldots, \boldsymbol{h}_{lag_D}]$ is the $D-$dimensional lag vector and V is the variance of output y.

Based on Variogram Analysis of Response Surface (VARS) [33], Sachin Rana et al. [29] tried to evaluate one dimensional variogram for each parameter while keeping the other parameters fixed to decrease the computational cost for the Integrated Variogram Across Range of Scales (IVARS) in GP-VARS. To further analyze the interaction terms, one-dimensional variogram is evaluated five times for a parameter while keeping all the other parameters at different values to cover the full range of response surface. The VARS method with detailed procedure can be found in the Appendix A of [29].

For the GP-based AHM model, Bayesian optimization is used for the gradient-free global optimization where there exists expensive black-box function, such as computationally intensive numerical simulators. The criterion of Expected Improvement is adopted to choose the proposal or temporary samples which could provide the maximum improvement in the objective functions regarding LMVs. Similar to GP-VARS, the best 10 proposal samples are picked via the Expected Improvement-based Bayesian optimization throughout the paper.

2.3. Inverse GPLVM Model (GPLVMIS)

For the GPIS in GP-VARS where the inputs and outputs of the inverse GP model are the misfit values (i.e., LMVs) and uncertain reservoir parameters respectively, it could be inappropriate and inefficient because independent GPRs are employed to nonlinearly transform the low-dimensional inputs (IVMs) to high-dimensional outputs (uncertain reservoir parameters), which could be very challenging as displayed in Figure 3a. Compared to GPIS, in the proposed GPLVM-based inverse model (GPLVMIS) the inputs and outputs in original GPIS are reversed, and the nonlinear mapping from high-dimensional reservoir parameters to low-dimensional LVMs is modeled by GPLVM as displayed in Figure 3b.



(**a**) GPIS (**b**) GPLVMIS

(**c**) Flowchart of GPLVMIS

Figure 3. The comparison of GPIS (**a**) and GPLVMIS (**b**) with the flowchart of GPLVMIS in detail (**c**).

Originally, GPLVM is used to conduct unsupervised dimensionality reduction, where the observed data are assumed to be the outputs of GPR with unknown inputs to be optimized which are typically termed by latent variables [31]. GPLVM is a natural inverse model, and we believe that it could be more efficient than GPR in GPIS for AHM.

Given the temporary reservoir parameters $X_o = [\boldsymbol{x}_1, ..., \boldsymbol{x}_N] \subset \mathcal{R}^{D\times N}$ and the corresponding outputs LVMs denoted by $Y_o = [\boldsymbol{y}_1, ..., \boldsymbol{y}_N] \subset \mathcal{R}^{C\times N}$, similar to GPIS our goal is to obtain one proposal sample $\boldsymbol{x}_*$ of reservoir parameters corresponding to an expected very low LVMs $\boldsymbol{y}_*$, we make use of GPLVM/GP to model the unknown mapping from the updated input $X = [X_o, \boldsymbol{x}_*]$ to output $Y = [Y_o, \boldsymbol{y}_*]$ as follows,

$$p(Y|X,\theta) = \prod_{i=1}^{C} \frac{1}{(2\pi)^{\frac{C(N+1)}{2}}|K|^{\frac{C}{2}}} \exp\left(-\frac{1}{2}\boldsymbol{y}_{i,:}K^{-1}\boldsymbol{y}_{i,:}^T\right) \tag{7}$$

where $\boldsymbol{y}_{i,:}$ corresponds to the i-th row of matrix Y, and the covariance matrix K in Equation (7) is defined on the input X with a few hyperparameters θ.

The objective function regarding the GPLVM optimization becomes

$$\underset{\boldsymbol{x}_*,\theta}{\operatorname{argmax}} \log(P(Y|X,\theta)) = \underset{\boldsymbol{x}_*,\theta}{\operatorname{argmax}} \left\{ -\frac{C(N+1)}{2}\log(2\pi) - \frac{C}{2}\log|K| - \frac{1}{2}tr(K^{-1}Y^TY) \right\} \tag{8}$$

where only one latent variable $\boldsymbol{x}_*$ in Equation (8) should be optimized rather than the whole latent variables X in original GPLVM.

To make the optimization algorithm smooth, as displayed in Figure 3c we simply reset X_o in X to original values in each iteration. For the initial value setting regarding $\boldsymbol{x}_*$, instead of random value we try to initialize it to be the existing parameters which correspond to the minimum Euclidean distance between the existing LVMs Y_o and the expected very low LVM $\boldsymbol{y}_*$ based on Nearest Neighbor (NN) method, which implicitly means that the best reservoir parameters regarding the lowest LVM in the previous iteration will be the initial value of $\boldsymbol{x}_*$. Similarly the gradients-based algorithms can be used to optimize the objective function of GPLVM regarding the proposal sample $\boldsymbol{x}_*$.

Compared to GPIS in GP-VARS [29] where GPs are employed to model the mapping from low-dimensional LVMs to high-dimensional reservoir parameters, we reverse the inputs and outputs of GPIS in the proposed GPLVMIS. It turns out that the GPLVM-based inverse model can be more efficient and reasonable than GP-based inverse model, because GPLVMIS is consistent with the direction of projection of GP-based proxy model in GPFS. Also, the proposed GPLVMIS outperforms GPIS in terms of model complexity. Only one gradient-based optimization is required in GPLVMIS to obtain the inverse temporary solutions $\boldsymbol{x}_*$, while GPIS needs D (the number of reservoir parameters) gradient-based algorithms corresponding to optimizing D GPs.

3. Results

In this section, we verify the proposed GPLVM-VARS in a synthetic dataset provided in [29] and the well-known PUNQ-S3 case to demonstrate the advantages of GPLVM-VARS in terms of accuracy and complexity.

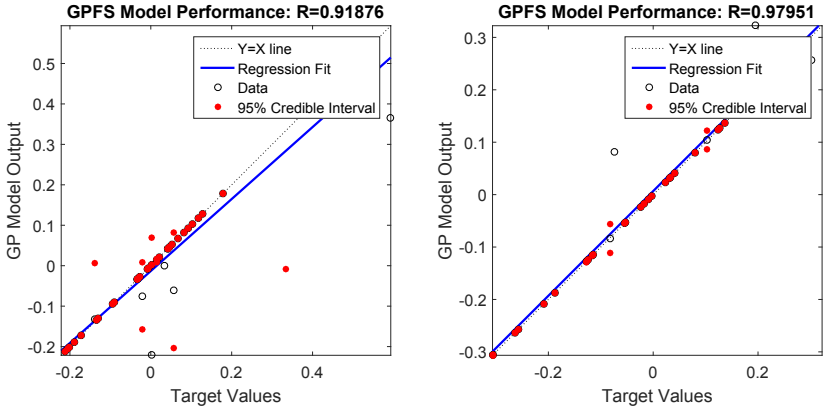
3.1. Synthetic Data

There are two synthetic examples provided in [29] to verify the GP-based AHM models. To objectively compare the proposed GPLVM-VARS with GP-VARS, we make use of the second example in the source code of GP-VARS, which is more challenging than the first example. The function is defined in Equation (9), which is the benchmark 'Multi-modal optimization function' called Rastrigin function. In this case, the input vector x is 2-dimensional which ranges from 0 to 5.12, and the target value is 33.94. The task is to tune input vector x in such a way that the output is approximate

to the target value. To make fair comparison, we run the experiment ten times and the best results are reported in the following.

$$f(x) = 10d + \sum_{i=1}^{d} \left[x_i^2 - 10\cos(2\pi x_i) \right] \tag{9}$$

As can be seen from Figure 4, the proposed GPLVM-VARS outperforms GP-VARS significantly in terms of the regression value R. In addition, we also compare the time complexity of the two models. With Intel E3 1240-V2 CPU plus 16G memory and MATLAB 2015b in Windows 7 64-bit platform, it takes 465.3 s to optimize GPLVM-VARS compared to 1413.9 s to optimize GP-VARS, which demonstrates the effectiveness of the proposed GPLVM-VARS.



(**a**) GP-VARS (**b**) GPLVM-VARS

Figure 4. The comparison of model performance in the synthetic data with the left (**a**) and right (**b**) figures corresponding to regression results from GP-VARS and the proposed GPLVM-VARS, respectively.

3.2. Punq-S3 Reservoir

PUNQ-S3 reservoir (https://www.imperial.ac.uk/earth-science/research/research-groups/perm/standard-models/) is a 5-layer heterogeneous reservoir model based on a real-world field operated by Elf Exploration and Production Company [34,35]. The medium-sized synthetic model consisting of 19 × 28 × 5 grid blocks (180 m × 180 m) of which 1761 are active has been widely employed to evaluate the performance of AHM models. As can be seen from Figure 5 that there exist two faults in the east and south, plus the strong aquifers in the north and west. Six vertical production wells (PRO-1, PRO-4, PRO-5, PRO-11, PRO-12, and PRO-15) are located near the initial gas–oil contact area, where wells PRO-1, PRO-4 and PRO-12 are perforated in layers 4 plus 5, wells PRO-5 and PRO-11 are perforated in layers 3 plus 4, and well PRO-15 is only perforated in layer 4. Also, there is a small gas cap in the first layer and in the center of the dome shaped structure. To avoid the gas production from the gas cap, no well has been perforated in the first layer.

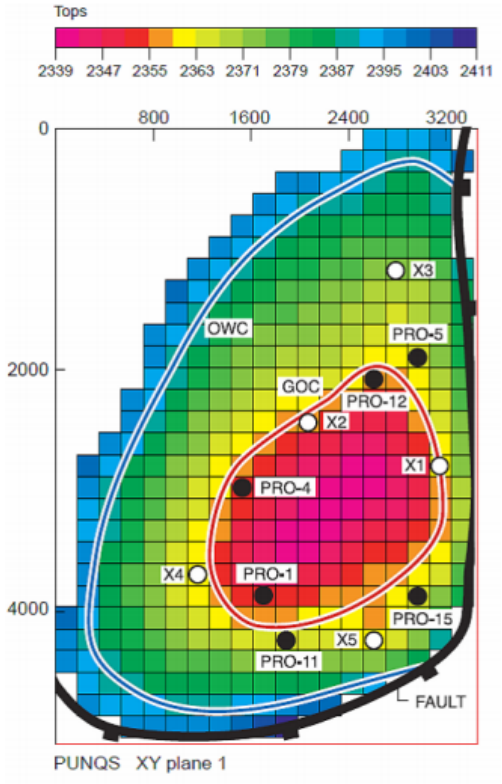


Figure 5. The top structure of PUNQ-S3 with wells.

The unknown parameters in PUNQ-S3 are the horizontal and vertical permeability as well as porosity. The Gaussian Random Field-based geo-statistical model is first used to generate the permeability and porosity fields independently for each of the five layers, and then the reservoir simulator Eclipse is used to generate eight years' production data with Gaussian noise explicitly accounting for measurement errors. Finally, only the eight years of the generated production data to be the history observations are adopted for AHM which include Well Water Cut (WWCT), Well Gas–Oil Ratio (WGOR), and Well Bottom Hole Pressure (WBHP) indices. The total production period is approximately 6000 days, and in our experiments we select the production data from the first 4000 days as the training data to optimize the unknown reservoir parameters with production data WWCT, WGOR and WBHP, and then the data for the remaining 2000 days to be the testing samples are used to verify the optimized AHM model.

There are 2660 uncertain parameters to be optimized in PUNQ-S3. Even some parameters reduction techniques are employed, the number of parameters could be large as well. For example, in [29] each layer can be divided into 9 different homogeneous regions, giving rise to 45 regions for 5 layers and a total of 135 parameters. By contrast, there are 24 LMVs used in GP-VARS. In this case, based on GPIS where the inputs and outputs are LMVs and uncertain parameters, respectively, 135 independent GPs should be learnt which is very time-consuming. On the contrary, in the proposed GPLVMIS where the inputs and outputs are uncertain parameters and LMVs, respectively, there is only one GPLVM needed to be optimized. However, to be consistent with the GP-VARS code we still use 24 (the number of LMVs) GPLVMs in our code. If the number of LVMs can be further reduced, the optimization time for GPLVMIS could be significantly reduced.

In the first experiment, we compare the proposed GPLVM-VARS with GP-VARS in terms of Global Misfit Value (GMV) in each iteration. It can be seen from Figure 6 that as the number of iterations grows, the GMVs based on the proposed GPLVM-VARS decrease significantly especially in the initial stage, and could converge to relatively small misfit values compared to GP-VARS.

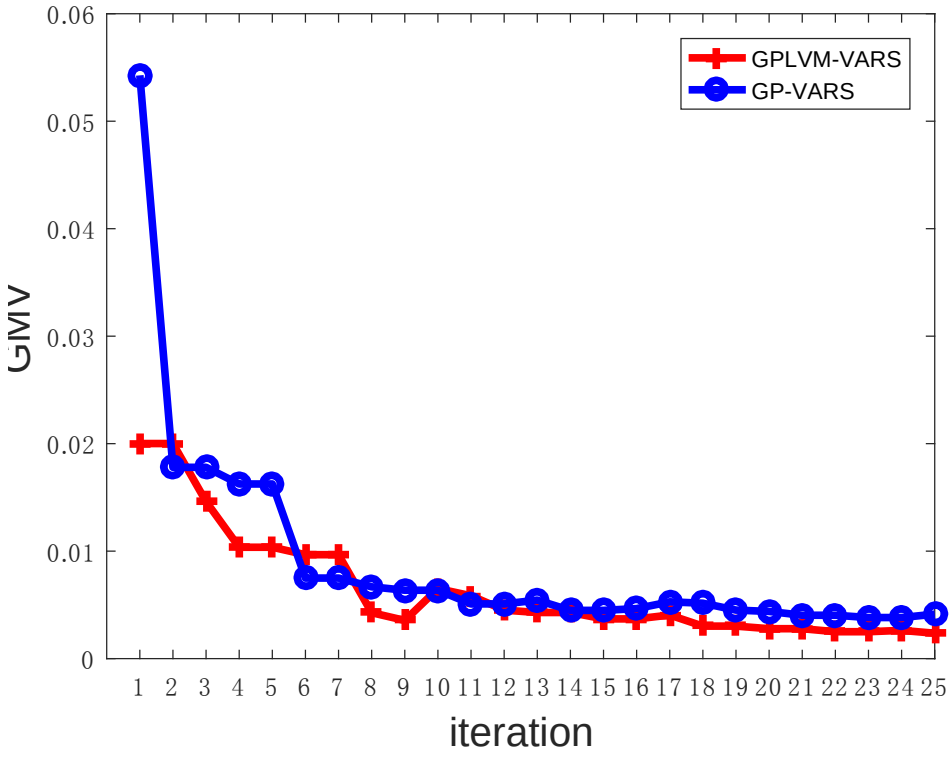


Figure 6. The comparison of GMVs based on GP-VARS and GPLVM-VARS.

In the second experiment, the GMVs regarding 11 proposal samples in each GP-VARS and GPLVM-VARS iteration are depicted in Figure 7, where Bayesian optimization-based GPIS provides 10 samples, GPIS and GPLVMIS-based optimizations give 1 sample for GP-VARS and GPLVM-VARS, respectively. It must be highlighted that the there are relatively smaller GMVs for the proposal sample from the proposed inverse model GPLVMIS than that from GPIS in GP-VARS especially in the initial stages, which also means that the proposed GPLVM-VARS could converge faster than GP-VARS.

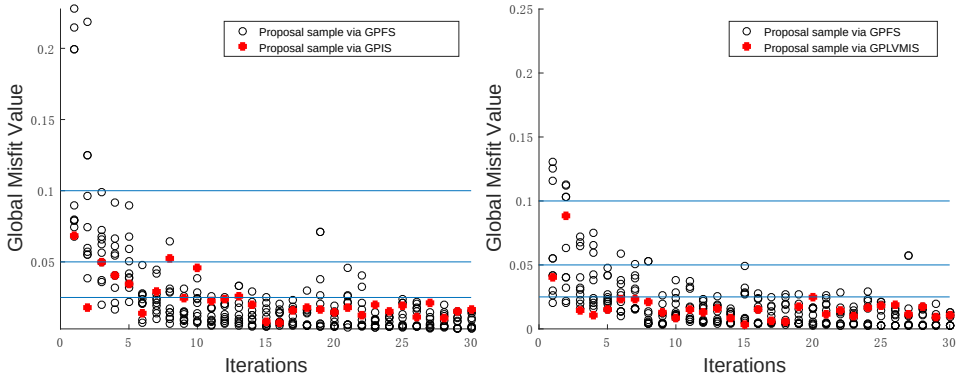
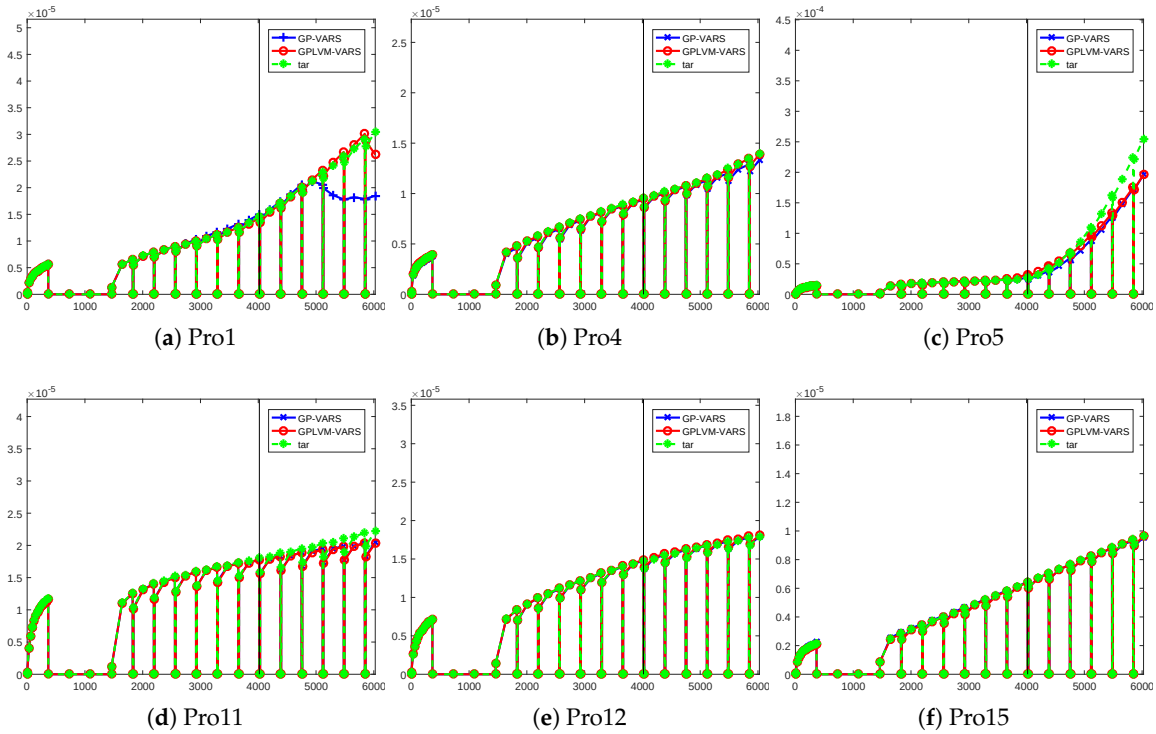


Figure 7. The visualization of 11 proposal samples from GP-VARS and GPLVM-VARS.

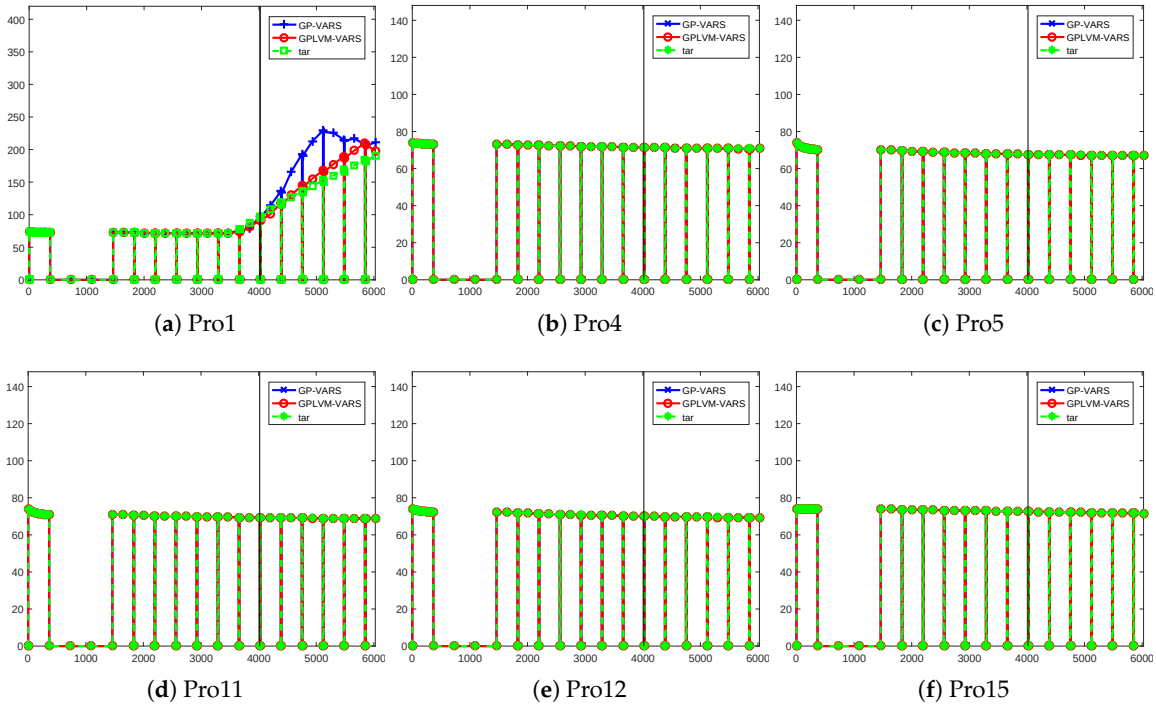
In the third experiment, the overall comparison for all the six wells with 4000 days' production data as the training data and the remaining 2000 days' data as the testing data. As can be seen from Figures 8–10 that proposed GPLVM-VARS outperforms GP-VARS in terms of the prediction errors especially for the WWCT matches in wells in Pro1, Pro4 and Pro5, plus WGOR and WBHP matches in well Pro1.



(**a**) Pro1 (**b**) Pro4 (**c**) Pro5

(**d**) Pro11 (**e**) Pro12 (**f**) Pro15

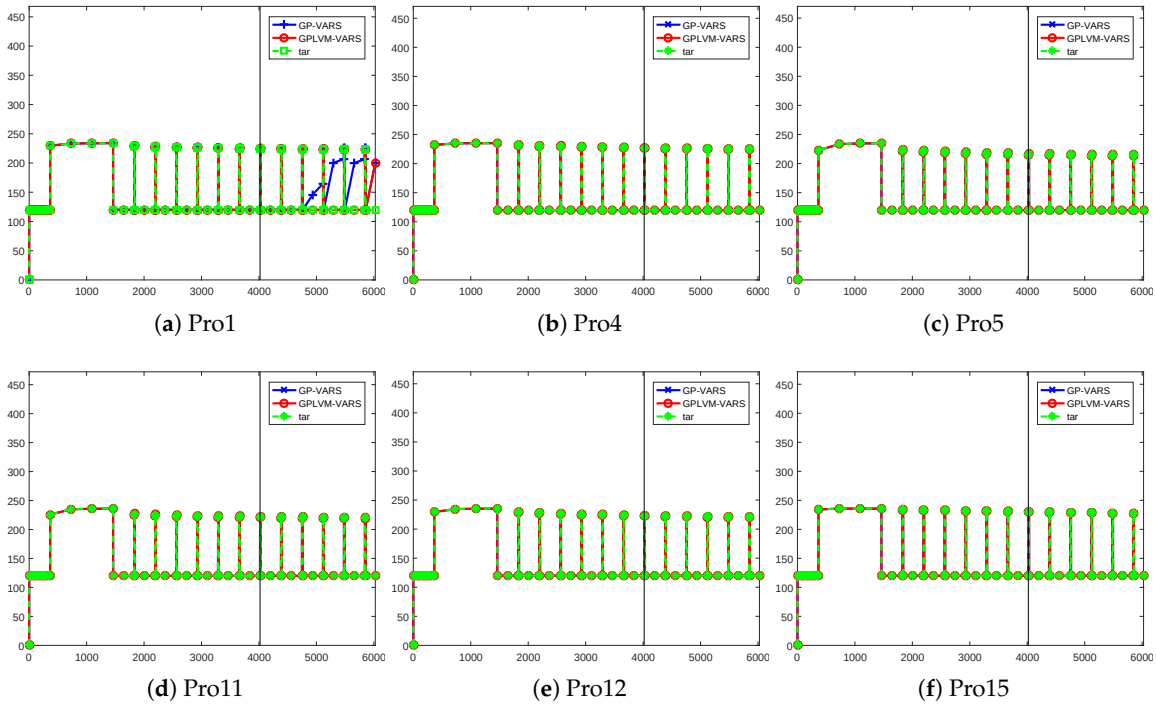
Figure 8. Comparison of WWCT with 4000 training and 2000 testing data for six wells (the black vertical line indicating present day).



(**a**) Pro1 (**b**) Pro4 (**c**) Pro5

(**d**) Pro11 (**e**) Pro12 (**f**) Pro15

Figure 9. Comparison of WGOR with 4000 training and 2000 testing data for six wells (the black vertical line indicating present day).



(**a**) Pro1 (**b**) Pro4 (**c**) Pro5

(**d**) Pro11 (**e**) Pro12 (**f**) Pro15

Figure 10. Comparison of WBHP with 4000 training and 2000 testing data for six wells (the black vertical line indicating present day).

4. Discussion

Compared to GP-based AHM [7], GP-VARS [29] introduced the extra GPIS to model the probability distribution of non-unique solutions. Furthermore, in the proposed GPLVM-VARS, GPLVMIS instead of GPIS is developed to further improve the performance of GP-VARS. The inverse mapping from low-dimensional LVMs to high-dimensional uncertain reservoir parameters is modeled by GPLVM rather than GPR, and the reservoir parameters with high dimensionality can be efficiently learnt by GPLVMIS which could improve the performance of GPIS as discussed in GP-VARS. As can be seen from the experiments that the proposed GPLVM-VARS is superior to recently developed GP-VARS in terms of accuracy and time complexity.

Compared to the parametric methods-based proxy models such as PCA, ANN and DL, the advantages of the proposed GPLVM-VARS lies in the capability of fitting nonlinear structure in the channelized reservoir models without time-consuming parameters tuning. All the hyperparameters in GPLVM-VARS including the type of the covariance function can be determined automatically, which could efficiently generalize to real-world reservoirs.

For the drawbacks of the proposed model, as the GP proxy models in GPLVM-VARS as well as GP-VARS suffer from the curse of dimensionality, the number of reservoir parameters has to be restricted up to 100. In addition to typical method like zonation and K-SVD, the semi-parametric model [36] could be promising which could simultaneously perform dimensionality reduction and regression. It is straightforward that the linear transformation in PCA which projects the high-dimensional reservoir parameters to low-dimensional latent variables can be employed to simultaneously perform dimensionality reduction and history matching.

Another issue that should be concerned about is the complexity of VARS in GP-VARS and the proposed GPLVM-VARS. AS the VARS-based sensitivity analysis is employed to evaluate one-dimensional variograms for each input parameter and the corresponding global relative sensitivity indices. However, the one variable at a time strategy in VARS brings high computational complexity especially when the number of reservoir parameters is large. Besides dimensionality reduction-based approaches, more efficient sensitivity analysis methods are expected. A possible simplification to

improve the efficiency of VARS is to independently evaluate one dimensional variograms for each input parameter without keeping all the other parameters at different values to cover the full range of response surface.

5. Conclusions

In this paper, we propose a novel proxy model for AHM termed GPLVM-VARS, which is motivated by GP-VARS to similarly use GP-based proxy model to find the forward and inverse solutions. To improve the unsatisfactory performance of the inverse model GPIS as stated in GP-VARS, we propose a new inverse model termed GPLVMIS, where the inputs regarding LVMs and outputs regarding reservoir parameters modeled by GP in GPIS are reversed. Given the high-dimensional reservoir parameters and low-dimensional LVMs, the proposed GPLVMIS could be more efficient than GPIS, giving rise to better inverse solutions. The experimental results in synthetic and PUNQ-S3 data demonstrate that the proposed GPLVM-VARS outperforms GP-VARS in terms of regression accuracy and model complexity.

Author Contributions: D.Z. contributed to the key ideas of this paper. Y.Z. and B.J. carried out the experiments. X.J. was mainly responsible for mathematical modeling and wrote the paper. Z.K. reviewed and edited the draft. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Science and Technology Major Project of the Ministry of Science and Technology of China under Grant 2016ZX05014-003-003.

Acknowledgments: The authors want to acknowledge Elf Exploration Company and Imperial College of Earth Sciences and Engineering for making PUNQ-S3 model dataset available online. Also, we would like to thank Sachin Rana for providing GP-VARS code in GitHub.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AHM	Assisted History Matching
GP	Gaussian Process
VARS	Variogram Analysis of Response Surface
GP-VARS	Gaussian Processes proxy models and Variogram Analysis of Response Surface-based sensitivity analysis
GPFS	Forward GP solution
GPIS	Inverse GP solution
LMV	Local Misfit Value
GMV	Global Misfit Value
GPLVM	Gaussian Process Latent Variable Model
GPR	Gaussian Process Regression
GPLVM-VARS	Gaussian Processes proxy models with Latent Variable Models and VARS-based sensitivity analysis
GPLVMIS	Gaussian Processes Latent Variable Model-based Inverse Solution

References

1. Verga, F.; Cancelliere, M.; Viberti, D. Improved application of assisted history matching techniques. *J. Pet. Sci. Eng.* **2013**, *109*, 327–347. [CrossRef]
2. Hou, J.; Zhou, K.; Zhang, X.S.; Kang, X.D.; Xie, H. A review of closed-loop reservoir management. *Pet. Sci.* **2015**, *12*, 114–128. [CrossRef]
3. Jaber, A.K.; Al-Jawad, S.N.; Alhuraishawy, A.K. A review of proxy modeling applications in numerical reservoir simulation. *Arab. J. Geosci.* **2019**, *12*, 1–16. [CrossRef]
4. Foroud, T.; Seifi, A.; AminShahidi, B. Assisted history matching using artificial neural network based global optimization method—Applications to Brugge field and a fractured Iranian reservoir. *J. Pet. Sci. Eng.* **2014**, *123*, 46–61. [CrossRef]

5. Oliver, D.S.; Chen, Y. Recent progress on reservoir history matching: A review. *Comput. Geosci.* **2011**, *15*, 185–221. [CrossRef]
6. Tripoppoom, S.; Xie, J.; Yong, R.; Wu, J.; Yu, W.; Sepehrnoori, K.; Miao, J.; Chang, C.; Li, N. Investigation of different production performances in shale gas wells using assisted history matching: Hydraulic fractures and reservoir characterization from production data. *Fuel* **2020**, *267*, 117097. [CrossRef]
7. Hamdi, H.; Couckuyt, I.; Sousa, M.C.; Dhaene, T. Gaussian Processes for history-matching: application to an unconventional gas reservoir. *Comput. Geosci.* **2017**, *21*, 267–287. [CrossRef]
8. Jung, S.; Lee, K.; Park, C.; Choe, J. Ensemble-Based Data Assimilation in Reservoir Characterization: A Review. *Energies* **2018**, *11*, 445. [CrossRef]
9. Kang, B.; Yang, H.; Lee, K.; Choe, J. Ensemble Kalman Filter With Principal Component Analysis Assisted Sampling for Channelized Reservoir Characterization. *J. Energy Resour. Technol.* **2017**, *139*, 032907. [CrossRef]
10. Luo, X.; Bhakta, T. Automatic and adaptive localization for ensemble-based history matching. *J. Pet. Sci. Eng.* **2020**, *184*, 106559. [CrossRef]
11. Evensen, G. Analysis of iterative ensemble smoothers for solving inverse problems. *Comput. Geosci.* **2018**, *22*, 885–908. [CrossRef]
12. Kim, J.; Kim, S.; Park, C.; Lee, K. Construction of prior models for ES-MDA by a deep neural network with a stacked autoencoder for predicting reservoir production. *J. Pet. Sci. Eng.* **2020**, *187*, 106800. [CrossRef]
13. Dickstein, F.; Goldfeld, P.; Pfeiffer, G.T.; Pinto, R.V. Truncated conjugate gradient and improved LBFGS and TSVD for history matching. *Comput. Geosci.* **2018**, *22*, 309–327. [CrossRef]
14. Mendivelso, J.; Thankachan, S.V.; Pinzón, Y. A brief history of parameterized matching problems. *Discret. Appl. Math.* **2020**, *274*, 103–115. [CrossRef]
15. Schulze-Riegert, R.W.; Axmann, J.K.; Haase, O.; Rian, D.T.; You, Y.L. Evolutionary Algorithms Applied to History Matching of Complex Reservoirs. *SPE Reserv. Eval. Eng.* **2002**, *5*, 163–173. [CrossRef]
16. Zhang, D.; Shen, A.; Jiang, X.; Kang, Z. Efficient history matching with dimensionality reduction methods for reservoir simulations. *Simul. Trans. Soc. Model. Simul. Int.* **2018**, *94*, 739–751. [CrossRef]
17. Park, H.Y.; Datta-Gupta, A.; King, M.J. Handling conflicting multiple objectives using Pareto-based evolutionary algorithm during history matching of reservoir performance. *J. Pet. Sci. Eng.* **2015**, *125*, 48–66. [CrossRef]
18. Alireza Shahkarami, Shahab D. Mohaghegh, V.G.; Haghighat, S.A. Artificial Intelligence (AI) Assisted History Matching. In Proceedings of the SPE Western North American and Rocky Mountain Joint Meeting, Denver, CO, USA, 17–18 April 2014.
19. Shahkarami, A.; Mohaghegh, S.D.; Hajizadeh, Y. Assisted history matching using pattern recognition technology. *Int. J. Oil Gas Coal Technol.* **2018**, *17*, 412–442. [CrossRef]
20. Ertekin, T.; Sun, Q. Artificial Intelligence Applications in Reservoir Engineering: A Status Check. *Energies* **2019**, *12*, 2897. [CrossRef]
21. Vo, H.X.; Durlofsky, L.J. A New Differentiable Parameterization Based on Principal Component Analysis for the Low-Dimensional Representation of Complex Geological Models. *Math. Geosci.* **2014**, *46*, 775–813. [CrossRef]
22. Esmaeili, M.; Ahmadi, M.; Kazemi, A. Kernel-based two-dimensional principal component analysis applied for parameterization in history matching. *J. Pet. Sci. Eng.* **2020**, *191*, 107134. [CrossRef]
23. Aïfa, T. Neural network applications to reservoirs: Physics-based models and data models. *J. Pet. Sci. Eng.* **2014**, *123*, 1–6. [CrossRef]
24. Costa, L.A.N.; Maschio, C.; Schiozer, D.J. Application of artificial neural networks in a history matching process. *J. Pet. Sci. Eng.* **2014**, *123*, 30–45. [CrossRef]
25. Maschio, C.; Schiozer, D.J. Bayesian history matching using artificial neural network and Markov Chain Monte Carlo. *J. Pet. Sci. Eng.* **2014**, *123*, 62–71. [CrossRef]
26. Liu, Y.; Sun, W.; Durlofsky, L.J. A Deep-Learning-Based Geological Parameterization for History Matching Complex Models. *Math. Geosci.* **2019**, *51*, 725–766. [CrossRef]
27. Kim, J.; Park, C.; Lee, K.; Ahn, S.; Jang, I. Deep neural network coupled with distance-based model selection for efficient history matching. *J. Pet. Sci. Eng.* **2020**, *185*, 106658. [CrossRef]
28. Riazi, S.H.; Zargar, G.; Baharimoghadam, M.; Moslemi, B.; Darani, E.S. Fractured reservoir history matching improved based on artificial intelligent. *Petroleum* **2016**, *2*, 344–360. [CrossRef]

29. Rana, S.; Ertekin, T.; King, G.R. An efficient assisted history matching and uncertainty quantification workflow using Gaussian processes proxy models and variogram based sensitivity analysis: GP-VARS. *Comput. Geosci.* **2018**, *114*, 73–83. [CrossRef]
30. Casella, G. An Introduction to Empirical Bayes Data Analysis. *Am. Stat.* **1985**, *39*, 83–87.
31. Lawrence, N. Probabilistic Non-linear Principal Component Analysis with Gaussian Process Latent Variable Models. *J. Mach. Learn. Res.* **2005**, *6*, 1783–1816.
32. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; The MIT Press: Cambridge, MA, USA, 2006.
33. Razavi, S.; Gupta, H.V. A new framework for comprehensive, robust, and efficient global sensitivity analysis: 1. Theory. *Water Resour. Res.* **2016**, *52*, 423–439. [CrossRef]
34. Floris, F.J.; Bush, M.D.; Cuypers, M.; Roggero, F.; Syversveen, A.R. Methods for quantifying the uncertainty of production forecasts: A comparative study. *Pet. Geosci.* **2001**, *7*, S87–S96. [CrossRef]
35. Gao, G.; Zafari, M..; Reynolds, A. Quantifying Uncertainty for the PUNQ-S3 Problem in a Bayesian Setting With RML and EnKF. In Proceedings of the SPE Reservoir Simulation Symposium, Houston, TX, USA, 31 January–2 February 2005.
36. Jiang, X.; Gao, J.; Wang, T.; Zheng, L. Supervised Latent Linear Gaussian Process Latent Variable Model for Dimensionality Reduction. *IEEE Trans. Syst. Man Cybern. Part Cybern.* **2012**, *42*, 1620–1632. [CrossRef] [PubMed]