CYBERNETICS AND PHYSICS, VOL. 11, NO. 1, 2022, 18–24

SPSA ALGORITHM FOR HISTORY DATA MATCHING OF COMPLEX NON-GAUSSIAN GEOLOGICAL MODELS

Vikentii Pankov
Science Educational Center
“Mathematical Robotics and Artificial Intelligence”,
Saint Petersburg State University
Laboratory “Control of Complex Systems”,
Institute for Problems in Mechanical Engineering of
Russian Academy of Sciences
Russia
st040308@student.spbu.ru

Oleg Granichin
Faculty of Mathematics and Mechanics,
Research Laboratory for Analysis and Modeling of Social Processes,
Saint Petersburg State University
Laboratory “Control of Complex Systems”,
Institute for Problems in Mechanical Engineering of
Russian Academy of Sciences
Russia
o.granichin@spbu.ru

Article history:
Received 18.05.2022, Accepted 25.05.2022

Abstract

History matching is the process of integrating dynamic production data in the reservoir model. It consists in estimation of uncertain model parameters such that oil or water production data from flow simulation become close to observed dynamic data. Various optimization methods can be used to estimate the model parameters.

Simultaneous perturbation stochastic approximation (SPSA) is one of the stochastic approximation algorithms. It requires only two objective function measurements for gradient approximation per iteration. Also parameters estimated by this algorithm might converge to their true values under arbitrary bounded additive noise, while many other optimization algorithms require the noise to have zero mean.

SPSA algorithm has not been well explored for history matching problems and has been applied only to simple Gaussian models.

In this paper, we applied SPSA to history matching of binary channelized reservoir models. We also used SPSA in combination with parameterization method CNN-PCA. And we considered the case of complex noise in observed production data and with objective function that does not require assumptions of normality of the observations, which is common in history matching literature. We experimentally showed that SPSA method can be successfully used for history matching of non-Gaussian geological models with different types of noise in observations and outperforms Particle Swarm Optimization by convergence speed.

Key words

Randomized algorithms, SPSA, Machine learning, Adaptive systems, Geological Modelling

1 Introduction

Geological models are used in many tasks related to the development of mineral deposits. They allow assessing the cost and potential success of such activities as seismic exploration, placement of new wells and forecasting of oil production. In order to make a geological model useful in making the right decisions in such activities, it is necessary that the model has high accuracy and uses the maximum amount of observed data.

The data that are used at the initial stages of building a geological model may include such characteristics as the geological concept of the formation, measurements at the drilling sites, and geophysical measurements.

In addition to this data, usually we have dynamic data that constantly arrives during the operation of the field, such as historical data of oil and liquid production. And it is possible to continuously update the model with this data. The process of updating the geological model using dynamic data is called history matching.

History matching of geological models consists in estimation of uncertain model parameters such that oil or water production data from flow simulation become close to observed data. It is an ill-posed inverse problem that has a lot of unknown geological parameters, but only a few measurements are available.

An important stage of history matching process is parameterization of geological models, because they may contain a huge number of parameters (typically over 4000) such as permeability values for each cell of a discretized reservoir space. And it requires to reduce a large dimension of geological models by describing them with low-dimensional vector of parameters.

After parameterization, we can match geological model to observed data by minimizing deviation of the observed data from the simulated data predicted by the current model. A lot of optimization methods can be used on this stage, including gradient-based methods. But gradient-free methods are used more often, because most complex flow simulators do not provide the ability to obtain a gradient by parameters.

There are many methods for parameterization of geological models. The principal component analysis (PCA) is often used [Reynolds et al., 1996; Sarma et al., 2008], which is based on the singular decomposition of a matrix obtained from a set of geological realizations. The disadvantage of PCA is that it uses only the correlations between two points of the space, and therefore is not applicable for non-Gaussian geological models. To eliminate this disadvantage, methods based on discrete wavelet transform (DWT) [Jafarpour et al., 2010] and discrete cosine transform (DCT) [Lu and Horne, 2000] have been developed. Recently, methods based on generative neural networks have been gaining popularity [Laloy et al., 2018; Liu et al., 2019]. They show a significant improvement in the quality of parametrization compared to DWT, DCT methods. Therefore, a method CNN-PCA (Convolutional Neural Network with Principal Component Analysis) [Liu et al., 2019] based on a generative neural network was used in this work.

Sparsity-promoting methods are applied for history matching problem [Khaninezhad and Mohammadreza, 2013; Elsheikh et al., 2013]. They utilize the sparse structure of geological models to restrict the solution space. Sparse solution corresponds to minimal number of nonzero entries of solution's vector (ℓ_0 norm regularization), but ℓ_0 minimization leads to intractable combinatorial optimization problem. Therefore, various iterative methods are used, most of which are based on solving the ill-posed problem by constraining ℓ_1 norm of solutions. Similar ℓ_1 regularization-based algorithms such as in [Ivanov et al., 2022] can also be used for closed-loop reservoir management problem — i.e., history matching combined with reservoir production optimization.

Particle Swarm Optimization (PSO) method [Liu et al., 2019; Mohamed et al., 2011; Lee and Stephen, 2019] is widely used for history matching problems. It is a genetic algorithm, which can find global optimal reservoir parameters with a high probability. The main drawback of this algorithm is the slow rate of convergence, and it needs a large number of reservoir simulations to achieve desired accuracy of data assimilation.

Ensemble Kalman filter is also a popular gradient-free method for history matching problems [Emerick and Reynolds, 2011; Jo et al., 2017]. It is a sequential method that has two alternating steps. This method starts from an ensemble of geological realizations. Then a forecast step is performed for updating parameters vector and assimilation of observed data. And an analysis step is used for updating the current set of ensembles. The main drawback of EnKF is that it is not applicable to non-Gaussian reservoir models, such as binary channelized models.

SPSA [Spall, 1992; Granichin et al., 2015] method computes an estimate of the minimum at each iteration in a randomly chosen direction by the difference of two function values, without requiring computation of the function gradient. It was firstly used for history matching by Gao et al. [Gao et al., 2007]. The second order SPSA algorithm [Zhu and Spall, 2002] was used that is based on a simple method for estimating the Hessian matrix at each iteration. Results showed that SPSA performed almost as well as the steepest descent method on PUNQ-S3 benchmark model. In [Li and Reynolds, 2011], SPSA algorithm is modified to obtain a stochastic Gaussian search direction method for history matching of Gaussian geological models. This method is successfully evaluated on the same PUNQ-S3 model as in [Gao et al., 2007].

In total, SPSA algorithm has not been well explored for history matching problems and has been applied only to simple low-dimensional Gaussian models. In [Granichin and Amelina, 2015] authors suggest to use SPSA-like algorithm with non-decreasing to zero step-size for the case when unknown parameters of the system changes with time. For linear case the similar algorithm was considered in [Granichin et al., 2010]. In this paper we applied SPSA to history matching of a binary channelized reservoir model that is much more complex than PUNQ-S3. We also used SPSA in combination with parameterization method CNN-PCA. And we considered the case of unknown but bounded noise in observed production data and with objective function that does not require assumptions of normality of the observations, which is common in history matching literature. We experimentally showed that SPSA method can be successfully used for history matching of non-Gaussian geological models with different types of noise in observations (including gaussian noise) and outperforms Particle Swarm Optimization by convergence speed.

2 Background

2.1 Spatial modeling techniques

Geological models can be deterministic or stochastic. For stochastic models, the parameters of the model at each point of the space are given by a random distribution: the model can be considered as a random field. With stochastic modeling, we have a family of different models and we can generate realizations from it.

Geostatistical methods are used for stochastic modeling. The approaches of sequential indicator modeling, object modeling and methods of multipoint statistics are widely used. Sequential indicator modeling uses semivariograms (correlations between two points in the environment) to model spatial correlation of model space. In object modeling, a geological model is obtained by randomly generating objects according to specified rules, such as tortuosity and channel thickness. The multipoint statistics (MPS) methods are based on the generation of geological models from a given training image, which is a way to take into account prior information about the reservoir. Training image is a conceptual image of the natural structure of the simulated reservoir. For instance, it includes such important structures as twisted channels of a certain orientation, and their thickness.

2.2 CNN-PCA

We used CNN-PCA [Liu et al., 2019] algorithm for geological models parameterization. The first step of the algorithm is to generate N_R realizations using a geostatistical algorithm. After that, the building of the PCA model $m_{pca}(\xi)$ is carried out using the SVD decomposition of data matrix Y, constructed by the generated realizations:

$$U\Sigma V^T = Y,$$

$$m_{pca}(\xi) = \hat{m} + U_l \Sigma_l \xi_l,$$

where $U\Sigma V^T$ is SVD decomposition, $\hat{m}$ is row wise mean vector of Y, U_l and Σ_l is matrix that contains l first rows from U, Σ.

Next, using PCA model, N_t random PCA-parameterized models are generated with ξ sampled from standard normal distribution. Generated PCA realizations are used for image transform network training. A loss function for training this model consists of style loss L_s and content loss L_c terms, based on neural style transfer ideas from [Johnson J, 2016]. The main idea of the CNN-PCA is to transfer the style of the training image to realizations parameterized by the PCA, making them more precise and similar to a training image.

The content loss has the following form:

$$L_C(x, m_{pca}(\xi)) = |F_4(m) - F_4(m_{pca}(\xi))|^2_{Fr}, \quad (1)$$

where $F_4(m)$ is a feature extraction function based on four layer of VGG-16 neural network [Johnson J, 2016].

The style loss has the following form:

$$G_k(m) = \frac{1}{N_{c,k}N_{z,k}} F_k(m)F_k(m)^T,$$

$$L_S(x, M_{ref}) = \sum_k \frac{1}{N^2_{z,k}} \|G_k(x) - G_k(M_{ref})\|^2_{Fr}, \quad (2)$$

where $N_{z,k}$ is number of the output features at layer k of VGG-16 network, $k = \{2, 4, 7, 10\}$.

In this way, L_S takes into account the average similarity of the x realization with the training image M_{ref} at different levels of image detalization. The different levels of detalization are determined by the features $F_k(m)$ of the neural network VGG-16 at different layers.

The architecture of the image transform neural network that was used for parameterization is described in [Johnson J, 2016].

3 History matching

3.1 Objective function

In this work, history matching process consists in minimizing the following objective function by ξ:

$$f_t(\xi) = \frac{||d_{fopr}(m) - d_{t_{obs}}||_2}{max(d_{t_{obs}})} + \sum_i \frac{||d_i(m) - d_{t_{obs_i}}||_2}{max(d_{t_{obs_i}})} + \frac{||\xi - \xi_0||_2}{max(\xi_0)}, \quad (3)$$

$d_{t_{obs}}$ in the first term of this function is field observed data that we have at the time t, such as field oil and water production rates. $d_{fopr}(m)$ is simulated data, obtained by predicting oil and water production rates using a flow simulator with permeability values m as variable parameters. These parameters are obtained by ξ using trained CNN-PCA model transform net followed by mapping parameterized realization to physical values of permeability.

The second term is the difference between simulated $d_i(\xi)$ and observed $d_{t_{obs_i}}$ production rates, which is given separately for each well i.

The third term is regularization that constrains ξ to be close to prior parameters ξ_0. It is required because CNN-PCA was trained using ξ sampled from this distribution.

Each term is normalized to the maximum of the corresponding observation. Due to this, data from the entire field and from each well have equal weight regardless of the absolute values of observations.

3.2 SPSA

Stochastic approximation algorithms can be applied to solve optimization problems if the objective function is noisy or unavailable for calculation. Such algorithms have the following general form:

$$\hat{x}_{n+1} = \hat{x}_n - \alpha_n \hat{g_n}(x_n),$$

where $\hat{x}_n$ is a sequence of parameter estimates, $\hat{g_n}(x_n)$ is a pseudogradient whose expected value is equal to the real gradient , and α_n is convergence rate parameter.

SPSA [Spall, 1992] is one of the stochastic approximation algorithms. It requires only two objective function measurements for gradient approximation per iteration. Also, parameters estimated by this algorithm might converge to their true values under arbitrary bounded additive noise, while many other optimization algorithms require the noise to have zero mean. If α is constant and sufficiently small step-size, SPSA guarantees the mean-square convergence of the estimates to a changing small bounded area around the changing true value of the parameter [Granichin and Amelina, 2015; Zhu and Spall, 2016].

Consider the observation model for the moment, when we have historical production data at the time $t \in \mathbb{R}$:

$$y_t = f_t(x_t) + \nu_t, \tag{4}$$

where ν_t is an additive noise caused by inaccuracies in parameterization and acquisition of production measurements.

Let $\mathcal{F}_{t-1}$ be the σ-algebra of all probabilistic events which happened up to time instant t, where $E_{\mathcal{F}_{t-1}}$ is a symbol of the conditional mathematical expectation with respect to the σ-algebra $\mathcal{F}_{t-1}$.

Using the observations $y_1, y_2, ..., y_t$, we need to build an estimate $\hat{\xi}$ of unknown vector of parameters ξ that minimizing mean-risk functional F_t:

$$F_t(\xi) = \mathbb{E}_{\mathcal{F}} f_t(x_t)$$

These estimations can be built by SPSA algorithm.

The input parameters of SPSA algorithm are sequences of positive numbers α_n, β_n, which tend to zero, and initial solution $\hat{\xi}(0)$. The next formulas can be used to estimate ξ [Granichin and Amelina, 2015]:

$$\begin{cases} u_{2n} = \hat{\xi}(n-1) + \beta_n \Delta_n, \\ u_{2n-1} = \hat{\xi}(n-1) - \beta_n \Delta_n, \\ \hat{\xi}(n) = \hat{\xi}(n-1) - \frac{\alpha_n}{2\beta_n} \Delta_n (y_{2n} - y_{2n-1}). \end{cases},$$

where y is a noisy measurements (4) of objective function, Δ_n is an observed sequence of independent Bernoulli random vectors from R^d with each component independently assuming values $\pm \frac{1}{\sqrt{(d)}}$.

Let next assumptions from [Granichin and Amelina, 2015] hold:

1. The successive differences $\hat{v}_t = v_t - v_{t-1}$ of observation noise are bounded: $|\hat{v}_t| \leq c_v < \infty$
2. The drift is bounded: $||\xi_t - \xi_{t-1}|| \leq \delta_\theta < \infty$
3. The rate of drift is bounded in a such way that for any arbitrary point x: $||\mathbb{E}_{\mathcal{F}_{t-1}} \nabla \phi_t(x)|| \leq a_1 ||x - \xi_{t-1}|| + a_0$, $||\mathbb{E}_{\mathcal{F}_{t-1}} \phi_t^2(x) \leq a_2 ||x - \xi_{t-1}||^2 + a_3$, where $\phi_t(x) = f_t(x) - f_{t-1}(x)$
4. Functions F_t have unique minimum points and θ_t and $\langle x - \theta_t, \mathbb{E}_{\mathcal{F}_{t-1}} \nabla f_t(x) \rangle \geq \mu ||x - \xi_t||^2$ for any x and constant $\mu > 0$.
5. The gradient ∇f_t is uniformly bounded in the mean-squared sense at the minimum points ξ_t : $\mathbb{E} f_t(\xi_t)||^2 \leq g = 0$, $\langle \mathbb{E} f_t(\xi_t), \mathbb{E} f_{t-1}(\xi_{t-1}) \rangle \leq g = 0$
6. The gradient of $f_t(x)$ satisfies Lipsitz condition with a constant $M \geq \mu$.

Assuming that the observation noise is bounded, we can formulate following theorem according to [Granichin and Amelina, 2015]:

Theorem 1. If we choose sufficiently small α and β then the mean squared error of sequence of estimations $\hat{\xi}$ generated by SPSA algorithm is asymptotically bounded by sufficiently small constant L [Granichin and Amelina, 2015]:

$$\mathbb{E} ||\hat{\xi_n} - \xi||^2 < L.$$

Note that this theorem guarantees the resistance of the parameter estimates to an almost arbitrary but bounded external noise in measurements.

4 Materials and Methods

Geostatistical algorithm snesim [Strebelle, 2002] is applied to generate 5000 realizations with size 60x60x1. The training image and conditioning data is used from [Liu et al., 2019]. CNN-PCA algorithm was implemented in Python using the PyTorch library. The image transform net was then trained for 16000 iterations, with a batch size of 32. We used O-PCA transform as a final thresholding step with parameter $\alpha = 0.7$ instead of hard thresholding. The rest of the CNN-PCA hyperparameters are the same as in the original work [Liu et al., 2019].

The reservoir simulator OPM flow [Atgeirr Fl et al., 2021] is used to obtain simulated data (field oil and water production rates). Figure 1 shows true model with conditional data. A constant porosity value of 0.2 is taken for all 3600 cells. Permeability for each cell specified isotropic with value 20 for mud and 2000 for sand. Four production wells is placed in $(90, 40), (53, 40), (58, 2), (40, 2)$ cells, and five injection wells is placed in $(18, 90), (40, 90), (2, 75), (2, 60), (2, 25)$ cells.

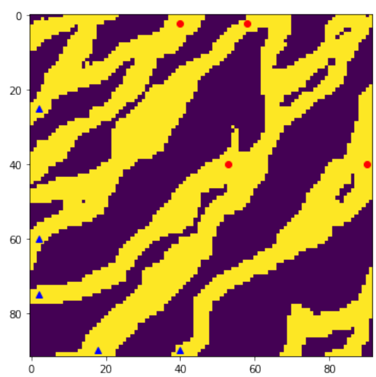

Figure 1. Simulated geological model with conditional data. Blue and red points depict injection and production wells correspondingly

Observed data include field and well's oil production rates over the first 7 years, recorded every 3 months.

The total number of data points is 140 (28 for each well and whole field). Figure 2 shows observed data obtained with true model. The observed data are derived from simulated true data by adding different types of noise.

Experiments were carried out for the next types of noise: random gaussian $\mathcal{N}(0, 0.02m)$, constant — $\nu_t^0 = 0.05m$, 'plus-minus' — $\nu_t^1 = 0.05 sign(\sin t) \cdot m$ and irregular — $\nu_t^2 = 0.03m(0.1 \sin t + 2sign(3 - t \mod 5))$, where $t = 1 \dots 28$, m — the maximum value of an observation vector.

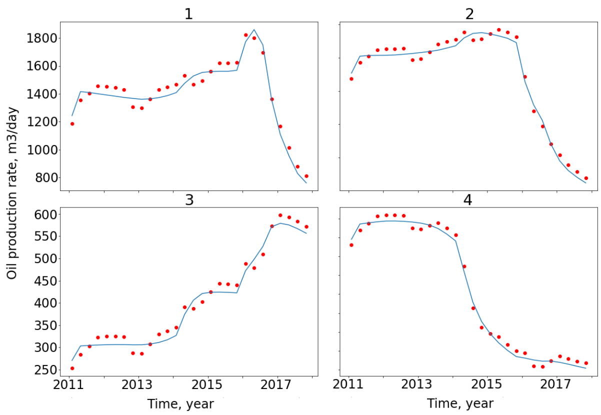

Figure 2. Simulated data: oil production curves for each production well. The red dots represent the observed data

For PSO method, cognitive, social, inertia and swarm size parameters are taken as $c_1 = 0.5$, $c_2 = 0.7$, $w = 0.9$, and $s = 50$ correspondingly. These values are selected based on the experiments in [Li-ping et al., 2005]. We tried other values as well, but it does not significantly affect convergence speed.

If the drift rate of the optimal value is not high, then SPSA algorithm can be used with the step-size decreasing to zero and then it provides the mean-square convergence of the estimates to zero. We chose sequences $\alpha_n = \alpha_0 n^{1/5}$ and $\beta_n = \beta_0 n^{1/10}$ based on experiments, according to theorem 1 and results concerning fastest rate of convergence from [Granichin et al., 2015].

We set $\alpha_0 = 0.5$ at first and use next heuristic algorithm: α_0 is decreased in half if the value of the objective function turns out to be more than three times higher than the best value throughout the entire process of minimization. After that, the best point is taken as the current vector of parameters. This procedure is applied to make the algorithm less dependent on a prior vector of parameters.

0.2 is taken as the β_0 value. It is found experimentally by doing history matching with different β_0 from interval $[0.01, 1]$ with step 0.1. The influence of this parameter on the achieved value of the objective function for a fixed number of iterations is shown in figure 3.

We used 250 iterations of SPSA and 10 iterations for PSO. In both cases, the same number of calls to the objective function is performed.

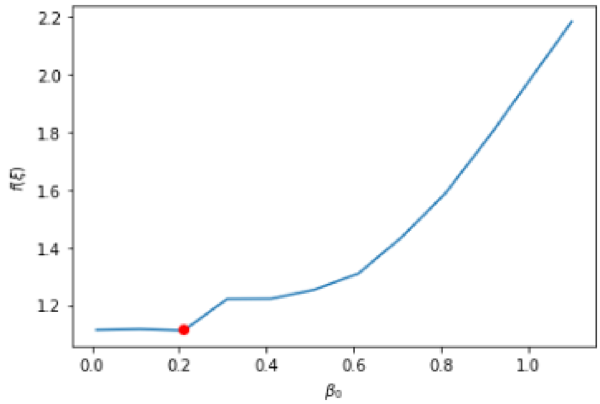

Figure 3. Dependency of minimum value of objective function on SPSA parameter β_0. The red point depicts chosen value

5 Results

The history matching results for observed data with irregular noise is presented on figure 4. Seven posterior (history matched) realizations were obtained by each algorithm. As we can see on the figure 4, both methods

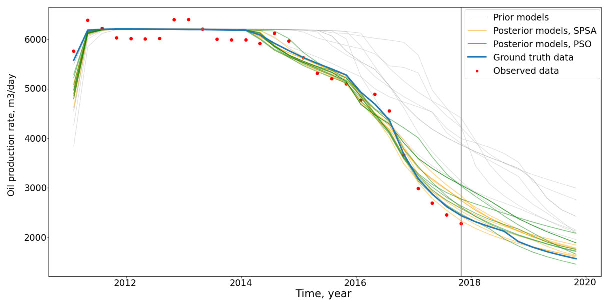

Figure 4. History matching results with SPSA and PSO methods for first type of noise. Posterior predictions obtained with SPSA are shown by orange curves, and predictions obtained with PSO are shown by green curves. Blue curve is true data. Vertical line indicates end of history matching period

result in much smaller scatter in production from the history-matched models relative to the production from prior models.

Posterior realizations of history matched models are presented on the figure 5. As it can be seen, they have some discrepancies with the true model. But many injection wells were connected by channels to corresponding producers on posterior models, as well as for the true model.

As it shown in table 1, SPSA algorithm achieves a given level of accuracy in a significantly smaller number

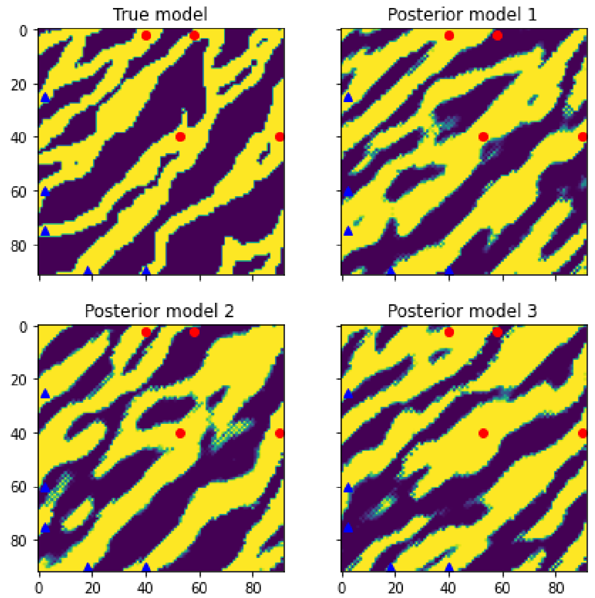

Figure 5. Geological ground truth model and three posterior realizations obtained by history matching with SPSA

Table 1. Algorithms evaluation for four types of noise. 'mean' column is average number of iterations, stddev column is standard deviation of number of iterations, fail column is number of tries that was unsuccessful, when the specified accuracy threshold was not reached, but the maximum number of iterations was reached

Noise type	Algorithm	mean	stddev	fail
Gaussian	PSO	502	254	2
	SPSA	**439**	332	2
Constant	PSO	657	364	4
	SPSA	**473**	317	**2**
Plusminus	PSO	908	291	6
	SPSA	**448**	402	**3**
Irregular	PSO	708	415	5
	SPSA	**514**	358	**3**

of iterations than PSO on average for all type of noise. Also, we can see that SPSA is equally efficient for any type of noise. Standard deviation of the number of iterations is high in all cases, because random realizations were taken as prior.

6 Conclusion

In this paper we applied SPSA method to the history matching problem and compared it with PSO method. We have used SPSA in combination with CNN-PCA parameterization algorithm to reduce dimensionality of geological model realizations.

Our experiments indicate that SPSA method can give good and stable results, even for non-Gaussian noise in production data. We showed that SPSA method can be successfully used for history matching of non-Gaussian geological models with different types of noise in observations and outperforms Particle Swarm Optimization by convergence speed.

In the future, it is planned to explore the possibility of implementing distributed version of the algorithm and usage of randomization of control strategies as in [Amelin and Granichin, 2016] for the problem of closed-loop reservoir management. In the first place, it is planned to consider the use of multi-agent technologies [Amelin et al., 2012].

Acknowledgements

Theoretical results in Sections 1–4 were supported by the IPME RAS by Russian Science Foundation (project no. 21-19-00516). Experimental results in Section 5 was supported by the St. Petersburg State University (project ID 93024916).

References

Amelin, K., Amelina, N., Granichin, O., and Granichina, O. (2012). Multi-agent stochastic systems with switched topology and noise. In *Proc. of 13th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, SNPD 2012*, pp. 438–443.

Amelin, K. and Granichin, O. (2016). Randomized control strategies under arbitrary external noise. *IEEE Transactions on Automatic Control*, **61** (5), pp. 1328–1333.

Atgeirr Fl, R., Tor, H. S., and Kai (2021). The open porous media flow reservoir simulator. *Computers and Mathematics with Applications*, **81**, pp. 159 – 185. Development and Application of Open-source Software for Problems with Numerical PDEs.

Elsheikh, A., Wheeler, M., and Hoteit, I. (2013). An ensemble based nonlinear orthogonal matching pursuit algorithm for sparse history matching of reservoir models. *Society of Petroleum Engineers - SPE Reservoir Simulation Symposium 2013*, **1**.

Emerick, A. A. and Reynolds, A. C. (2011). History Matching a Field Case Using the Ensemble Kalman Filter With Covariance Localization. *SPE Reservoir Evaluation and Engineering*, **14** (04), pp. 443–452.

Gao, G., Li, G., and Reynolds, A. C. (2007). A stochastic optimization algorithm for automatic history matching. *SPE Journal*, **12** (02), pp. 196–208.

Granichin, O. and Amelina, N. (2015). Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances. *IEEE Transactions on Automatic Control*, **60** (6), pp. 1653–1658.

Granichin, O., Vakhitov, A., and Vlasov, V. (2010). Adaptive control of siso plant with time-varying coefficients based on random test perturbation. In *Proceedings of the 2010 American Control Conference*, pp. 4004–4009.

Granichin, O., Volkovich, Z. V., and Toledano-Kitai, D. (2015). Randomized algorithms in automatic control and data mining. *Springer*.

Ivanov, D., Granichin, O., Pankov, V., and Volkovich, Z. (2022). Design of ℓ_1 new suboptimal fractional delays controller for discrete non-minimum phase system under unknown-but-bounded disturbance. *Mathematics*, **10** (1).

Jafarpour, B., Goyal, V., McLaughlin, D., and Freeman, W. (2010). Compressed history matching: Exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems. *Mathematical Geosciences*, **42**, pp. 1–27.

Jo, H., Jung, H., Ahn, J., Lee, K., and Choe, J. (2017). History matching of channel reservoirs using ensemble kalman filter with continuous update of channel information. *Energy Exploration & Exploitation*, **35** (1), pp. 3–23.

Johnson J, Alahi A, L. F. (2016). Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pp. 694–711.

Khaninezhad and Mohammadreza, M. (2013). Prior model identification with sparsity-promoting history matching. In *SPE Reservoir Simulation Conference*, vol. 1, 02.

Laloy, E., Hérault, R., Jacques, D., and Linde, N. (2018). Training-image-based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research*, **54**.

Lee, S. and Stephen, K. (2019). Field application study on automatic history matching using particle swarm optimization. In *SPE Reservoir Characterisation and Simulation Conference and Exhibition*, 09.

Li, G. and Reynolds, A. C. (2011). Uncertainty quantification of reservoir performance predictions using a stochastic optimization algorithm. *Computational Geosciences*, **15**, pp. 451–462.

Li-ping, Z., Huan-jun, Y., and Shang-xu, H. (2005). Optimal choice of parameters for particle swarm optimization. *Journal of Zhejiang University SCIENCE*, **6**, pp. 528–534.

Liu, Y., Sun, W., and Durlofsky, L. (2019). A deep-learning-based geological parameterization for history matching complex models. *Mathematical Geosciences*, **51**.

Lu, P. and Horne, R. N. (2000). A multiresolution approach to reservoir parameter estimation using wavelet analysis. In *SPE Annual Technical Conference and Exhibition*, 10.

Mohamed, L., Christie, M., and Demyanov, V. (2011). History matching and uncertainty quantification: Multiobjective particle swarm optimisation approach. In *SPE Europec featured at EAGE Conference and Exhibition*.

Reynolds, A., He, N., Chu, L., and Oliver, D. (1996). Reparameterization Techniques for Generating Reservoir Descriptions Conditioned to Variograms and Well-Test Pressure Data. *SPE Journal*, **1** (04), pp. 413–426.

Sarma, P., Durlofsky, L., and Aziz, K. (2008). Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. *Mathematical Geosciences*, **40**, pp. 3–32.

Spall, J. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, **37** (3), pp. 332–341.

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, **34**, pp. 1–21.

Zhu, J. and Spall, J. C. (2016). Tracking capability of stochastic gradient algorithm with constant gain. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 4522–4527.

Zhu, X. and Spall, J. (2002). A modified second-order spsa optimization algorithm for finite samples. *International Journal of Adaptive Control and Signal Processing*, **16**, pp. 397 – 409.