Advances in Geo-Energy Research

Vol. 7, No. 2, p. 136-140, 2023

Perspective

Reservoir automatic history matching: Methods, challenges, and future directions

Piyang Liu[1], Kai Zhang[1,2]*, Jun Yao[2]

[1]*School of Civil Engineering, Qingdao University of Technology, Qingdao 266520, P. R. China*

[2]*School of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266580, P. R. China*

Keywords:
History matching
optimization algorithm
surrogate model
data-driven

Cited as:
Liu, P., Zhang, K., Yao, J. Reservoir automatic history matching: Methods, challenges, and future directions. Advances in Geo-Energy Research, 2023, 7(2): 136-140.
https://doi.org/10.46690/ager.2023.02.07

Abstract:
Reservoir history matching refers to the process of continuously adjusting the parameters of the reservoir model, so that its dynamic response will match the historical observation data, which is a prerequisite for making forecasts based on the reservoir model. With the development of optimization theory and machine learning algorithms, automatic history matching has made numerous breakthroughs for practical applications. In this perspective, the existing automatic history matching methods are summarized and divided into model-driven and surrogate-driven history matching methods according to whether the reservoir simulator needs to be run during the automatic history matching process. Then, the basic principles of these methods and their limitations in practical applications are outlined. Finally, the future trends of reservoir automatic history matching are discussed.

1. Introduction

In reservoir management, the typical production data that can be measured include the surface production rate, water cut, bottom-hole pressure, etc. However, the model parameters (such as absolute and relative permeabilities, porosity) and dynamic variables (such as pressure and water saturation) cannot be measured directly. Reservoir engineers routinely obtain estimates of these parameters and variables by history matching, and then use the history-matched model to predict future trends.

The traditional reservoir history matching method is to manually adjust the reservoir model based on field experiences until the dynamic responses match the historical observation data, which is a time-consuming trial-and-error process. In recent years, automatic history matching (AHM) methods have been developed and employed in many practical applications, benefiting from the development of optimization theory and machine learning algorithms. AHM refers to the process of using intelligent optimization algorithms to find the optimal reservoir model parameters.

Depending on whether the forward reservoir simulations need to be run during the AHM process, AHM methods are divided into model-driven and surrogate-driven history matching methods (Fig. 1). The current status of AHM is presented in Section 2 and Section 3. Section 4 discusses several directions that are expected to become popular in the future.

2. Model-driven history matching

By using the least squares method or the maximum a posteriori estimation method to establish the objective function, the AHM problem can be transformed into a single objective optimization problem that can be solved by optimization methods (Rwechungura et al., 2011). The optimization algorithms need to call the reservoir numerical simulator multiple times to obtain the response data with adjusted parameters of the reservoir model (Bertolini and Schiozer, 2011). Therefore, this type of method is referred to as model-driven history matching

*Corresponding author.
E-mail address: Piyang.Liu@qut.edu.cn (P. Liu); zhangkai@qut.edu.cn (K. Zhang); rcogfr_upc@126.com (J. Yao).

Yandy Scientific Press

Received January 1 2023; revised January 9, 2023; accepted January 12, 2023; available online January 14, 2023

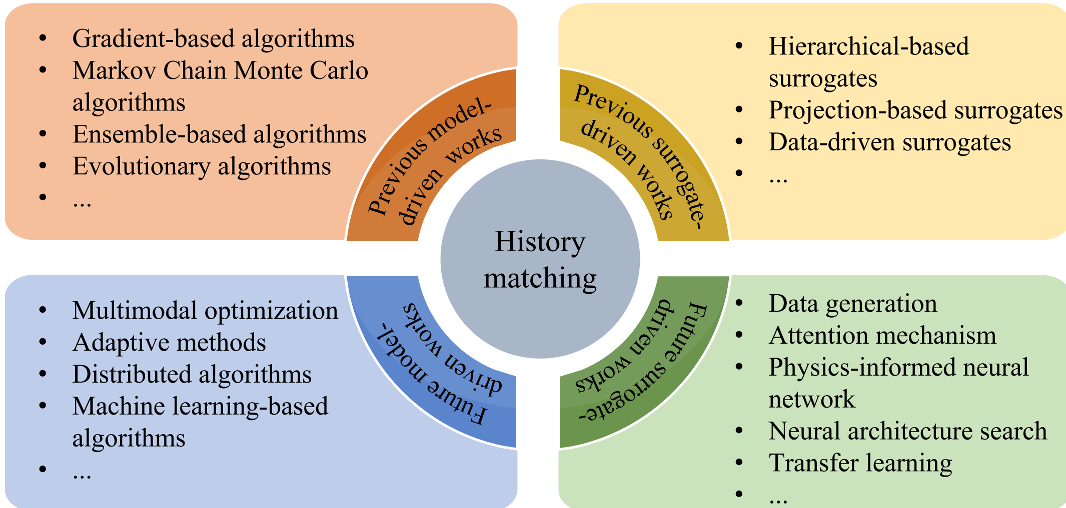

Fig. 1. The existing research perspectives and outlooks for history matching.

method.

According to whether a gradient is required, optimization algorithms can be divided into gradient-based algorithms (Yang and Watson, 1988; Eydinov et al., 2008; Chen et al., 2009) and gradient-free algorithms (Romero and Carter, 2001; Emerick and Reynolds, 2013; Paredes et al., 2013). The commonly used gradient algorithms include Newton-type algorithms, steepest descent method (Chen et al., 2021) and conjugate gradient method. Newton-type algorithms not only need the gradient of the objective function, but also the Hessian matrix. The representative algorithms include Gauss Newton method, Levenberg-Marquardt method, Levenberg-Marquardt-Fletcher method, Quasi-Newton method, etc. Hessian matrix is the second-order derivative of the objective function with respect to the model parameters, and its calculation involves solving the sensitivity coefficient matrix of the model parameters. For large-scale reservoirs, neither the coefficient matrix nor the sensitivity coefficient matrix can be calculated. Therefore, Newton-type methods are not applicable to large-scale computing problems. The steepest descent method and the conjugate gradient method only require the gradient of the objective function to efficiently solve the history matching problem, and its convergence is acceptable. However, the calculation of analytic gradient needs to work directly with the reservoir numerical simulation code, which is difficult and significantly increases the computational costs. In practical applications, an approximate gradient is generally used to replace the real gradient. The commonly used methods to calculate the approximate gradient are the finite difference method, automatic differentiation method, simultaneous perturbation stochastic approximation method, and adjoint method. It is worth mentioning that the adjoint method requires only two simulations regardless of the number of decision variables to compute gradients for optimization problems. Gradient-based algorithms are efficient but can only locate a local optimum; as the number of decision variables increases, the computational efficiency of such algorithms decreases dramatically.

Compared with gradient-based algorithms that are sensitive to initial values and easily fall into local optima, gradient-free algorithms have the ability of global search. The gradient-free algorithms often used in AHM mainly include the heuristic search algorithm (Romero and Carter, 2001), Monte Carlo Markov Chain (MCMC) method (Liu and Oliver, 2003; Efendiev et al., 2006), and data assimilation method (Gu and Oliver, 2004). The heuristic search algorithm is a kind of algorithm that uses the location information of the solution in the search space and constructs the search path through specific rules. The frequently used algorithms are simulated annealing algorithm, genetic algorithm, and particle swarm optimization algorithm (Yazdanpanah et al., 2019). The heuristic search algorithm can be used to obtain the global optimal solution, but it requires tens of thousands of simulations to fully converge, which is not affordable for reservoir history matching, especially for large-scale real reservoirs. The MCMC method randomly samples the Bayesian posterior probability of uncertain parameters for automatic history fitting solutions. The commonly applied MCMC methods include the Glibbs algorithm, Metropolis-Hastings algorithm, etc. At present, the typical data assimilation methods employed in automatic history fitting include the ensemble Kalman filter (Zha et al., 2018) and ensemble smoother (Emerick and Reynolds, 2013). The ensemble Kalman filter absorbs observation data sequentially for the dataset update, while the ensemble smoother utilizes all observation data for the dataset update at the same time.

3. Surrogate-driven history matching

A real reservoir model typically contains hundreds of thousands of grid blocks populated with different properties. Geostatistical modeling is used to quantify the effect of reservoir heterogeneity on reservoir flow behavior. Despite the tremendous increase in hardware computing power in the last decade, the time requirement for each simulation run during optimization could still be significant. One possibility to mitigate the heavy computational cost of numerical simulations is to use a surrogate model (Cullick et al., 2006), which approximates the solutions of the reservoir simulation with much lower computational cost. The commonly adopted surrogate modeling methods can be roughly divided into three categories: Hierarchical-based model, projection-based reduced-order model, and data-driven surrogate model.

3.1 Hierarchical-based surrogate model

Hierarchical-based surrogate models reduce the computational cost of the forward model by simplifying the underlying physical process or reducing the resolution of the reservoir model. Many previous studies focused on how to upscale properties from the measurement scale (i.e., fine scale) to the coarse scale for fast simulations (Ashby and Falgout, 1996; Winter et al., 2003). The main drawback of this approach is the lack of robustness. Because the reservoir is a complex system, subtle perturbation in the initial model may lead to huge variations in the final simulation results. It is difficult to realize a proxy of the complex reservoir by simply reducing resolution through coarsening.

Another idea of the hierarchical-based surrogate is to simplify the physical process, which is well illustrated by the interwell numerical simulation model (Zhao et al., 2016). To reflect the interactions between reservoir wells and reduce the complexity of the model, the reservoir injection and recovery system are simplified and characterized as a series of well-to-well connected cells, each characterized by two parameters, conductivity, and control volume. Conductivity describes the flow capacity of the cell and control volume reflects the oil storage capacity of the connected unit. This converts the traditional grid-based calculation into a connected cell-based calculation, which saves considerable computational time. Then, the oil-water dynamic index at the well location can be evaluated by the pressure at the well location, water saturation can be traced by material balance, and the oil-water two-phase leading-edge propulsion theory can be based on the connected cell.

3.2 Projection-based surrogate model

The projection-based method reduces the dimension of the model by projecting the control equation into a low-dimensional subspace based on orthogonal vectors, thus reducing the computational cost of the model-based workflows. These methods are usually divided into SVD-based (McPhee and Yeh, 2008; Ghommem et al., 2013; Yeung et al., 2022) and Krylov-based strategies (Dunbar and Woodbury, 1989; Woodbury et al., 1990).

Reduced-order modeling approaches have been among the most effective ways to reduce the computational effort of model-based workflows by reducing the number of dimensions of the model. The main idea behind projection-based reduced-order modeling is to construct (linear) low-order alternative models by projecting the dynamics of the system into the main variability subspace of the model dynamics. Most reduced-order modeling strategies use a proper orthogonal decomposition of the ”snapshot” of time series about the model state.

3.3 Data-driven-based surrogate model

Data-driven surrogate models approximate forward models by mapping a set of inputs (reservoir parameters) to outputs (reservoir dynamics). The data-driven surrogate modeling technique relies on statistics and can be regarded as a maximum-likelihood process. The approaches for constructing data-driven surrogates can be divided into online and offline methods.

Traditional data-driven-based surrogate models (Hussain et al., 2002; Stone, 2011) include, but are not limited to, Gaussian process, kriging, polynomial chaos expansions, radial basis functions, and support vector machine. The main drawback of these methods is the lack of robustness, which may lead to inaccurate responses in highly complex nonlinear cases. Therefore, they often require continuous sampling during the solution process and the dynamic construction of surrogate models.

Offline methods build a sufficiently accurate surrogate model that can completely replace the role of numerical simulation and approximate the entire search space. Most studies (Dachanuwattana et al., 2018; Li et al., 2019) use sensitivity analysis methods to select some key parameters and then use traditional machine learning methods to construct alternative models. The emergence of deep learning makes it possible to directly establish the mapping from high-dimensional spatial parameters to reservoir dynamics without sensitivity analysis. Most of these methods are based on the image-to-image regression framework (Tang et al., 2020; Zhang et al., 2022). In most cases, the observation data for history matching are time-series data obtained from the measurements of production and injection wells, such as oil production rate, water production rate and bottom hole pressure. After the image-to-image regression framework has been established, the production data can be calculated through the Peaceman formula. Another feasible solution is to directly fit the pressure or saturation field obtained from the inversion of seismic data by taking 4D seismic data as observation (Oliver et al., 2021). In addition, the image caption method has been widely studied in the general field of deep learning, and can be introduced into the history fitting framework to directly establish an image-to-sequence surrogate framework (Ma et al., 2022).

In order to enhance the accuracy of data-driven surrogate models, physical equations can be embedded into machine learning models (Raissi et al., 2019; Wang et al., 2020). Many deterministic and probabilistic machine learning methods have been proposed, in which physics are embedded as additional optimization constraints (Rao et al., 2021). These frameworks can incorporate the discretized control equations into the training of convolutional neural networks, and thus improve the prediction accuracy of the surrogate model.

In addition, several data-driven end-to-end history matching methods have been proposed (Kim et al., 2020; Jo et al., 2022). These first learn the mapping of production data to the reservoir model, and then obtain the posterior reservoir model by inputting observation data into the trained model.

4. Future trends

During the development of reservoir history matching, early research focused on how to improve the efficiency of optimization algorithms (model-driven methods), such as searching for the global optimum of the objective function for gradient-based algorithms, improving data-matching results for ensemble-based algorithms, adapting to high-dimensional problems for MCMC algorithms, and balancing the diversity and convergence of the search population for evolutionary algorithms. Despite these improvements in history matching, effective optimization algorithms are still necessary to accurately estimate the actual posterior space of the objective function for uncertainty quantification.

The future development of optimization algorithms can be classed into two directions: improving convergence and promoting diversity. The convergence of optimization algorithms reflects computational efficiency, which is critical for history matching because the numerical simulation of large-scale reservoirs is always time-consuming. Gradient-based algorithms and ensemble-based algorithms show fast convergence but have their inherent drawbacks. For gradient-based algorithms, the global optimum might be missed in the searching process due to the influence of multiple local optima. For ensemble-based algorithms, the diversity of solutions cannot be guaranteed, which is also called ensemble collapse; here, models of the final ensemble are almost the same, degrading the production prediction based on this ensemble. Strategies to improve the convergence and diversity of history matching include the multimodal optimization strategy (Ma et al., 2021), adaptive methods (Sun et al., 2021), distributed algorithms (Gao et al., 2022), among other approaches. In addition, the authors argue that novel algorithms coupled with machine learning can enhance the efficiency of history matching.

At the same time, the surrogate-based method in history matching has undergone rapid development, while utilizing this method to practical applications has proved challenging. This is due to the limited response data for different parameter combinations, high computational efforts in generating samples using numerical simulations, and inaccurate and uninterpretable results of surrogates. To improve the effectiveness of the surrogate, further studies should be carried out from the aspect of data generation, surrogate architecture and surrogate transfer. The samples for training the surrogate are usually derived from numerical simulations, with the computational cost of this process being almost the same as that of model-driven methods, diminishing its superiority. Using fewer samples or generating additional samples based on low-fidelity methods, such as streamline simulation (Yin et al., 2021), model upscaling (Jiang and Durlofsky, 2023), or data augmentation (Shorten and Khoshgoftaar, 2019), could improve the computational efficiency of building the surrogate. Moreover, even for experienced specialists, it is difficult to design the deep learning architecture for building surrogates, because lots of hyperparameters need to be tuned. Thus, new modules of neural networks (e.g., attention mechanism, physics-informed neural network (Raissi et al., 2019)) or automatic design methods (e.g., neural architecture search) should be established to improve the applicability of the surrogate. Finally, the surrogate model should be updated efficiently based on newly observed data from reservoirs, such as by using transfer learning (Weiss et al., 2016; Zhong et al., 2022) or continual learning (Zenke et al., 2017), to achieve real-time history matching and production optimization.

Acknowledgements

The authors gratefully acknowledge the support of the National Natural Science Foundation of China (No. 52274057).

Conflict of interest

The authors declare no competing interest.

References

Ashby, S. F., Falgout, R. D. A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. Nuclear Science and Engineering, 1996, 124(1): 145-159.

Bertolini, A. C., Schiozer, D. J. Influence of the objective function in the history matching process. Journal of Petroleum Science and Engineering, 2011, 78(1): 32-41.

Chen, J., Wang, L., Wang, C., et al. Automatic fracture optimization for shale gas reservoirs based on gradient descent method and reservoir simulation. Advances in Geo-Energy Research, 2021, 5(2): 191-201.

Chen, Y., Oliver, D. S., Zhang, D. Efficient ensemble-based closed-loop production optimization. SPE Journal, 2009, 14(4): 634-645.

Cullick, A. S., Johnson, W. D., Shi, G. Improved and more rapid history matching with a nonlinear proxy and global optimization. Paper SPE 101933 Presented at SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 24-27 September, 2006.

Dachanuwattana, S., Jin, J., Zuloaga-Molero, P., et al. Application of proxy-based MCMC and EDFM to history match a Vaca Muerta shale oil well. Fuel, 2018, 220: 490-502.

Dunbar, W. S., Woodbury, A. D. Application of the Lanczos algorithm to the solution of the groundwater flow equation. Water Resources Research, 1989, 25(3): 551-558.

Efendiev, Y., Hou, T., Luo, W. Preconditioning markov chain monte carlo simulations using coarse-scale models. SIAM Journal on Scientific Computing, 2006, 28(2): 776-803.

Emerick, A. A., Reynolds, A. C. Ensemble smoother with multiple data assimilation. Computers & Geosciences, 2013, 55: 3-15.

Eydinov, D., Aanonsen, S. I., Haukås, J., et al. A method for automatic history matching of a compositional reservoir simulator with multipoint flux approximation. Computational Geosciences, 2008, 12(2): 209-225.

Gao, G., Wang, Y., Vink, J. C., et al. Distributed quasi-Newton derivative-free optimization method for optimization problems with multiple local optima. Computational Geosciences, 2022, 26(4): 847-863.

Ghommem, M., Presho, M., Calo, V. M., et al. Mode decomposition methods for flows in high-contrast porous media. Global-local approach. Journal of Computational Physics, 2013, 253: 226-238.

Gu, Y., Oliver, D. S. History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter. Paper SPE 89942 Presented at SPE Annual Technical Conference and Exhibition, Houston, Texas, 26-29 September, 2004.

Hussain, M. F., Barton, R. R., Joshi, S. B. Metamodeling: Radial basis functions, versus polynomials. European Journal of Operational Research, 2002, 138(1): 142-154.

Jiang, S., Durlofsky, L. J. Use of multifidelity training data and transfer learning for efficient construction of subsurface flow surrogate models. Journal of Computational Physics, 2023, 474: 111800.

Jo, S., Jeong, H., Min, B., et al. Efficient deep-learning-based history matching for fluvial channel reservoirs. Journal of Petroleum Science and Engineering, 2022, 208: 109247.

Kim, J., Kim, S., Park, C., et al. Construction of prior models for ES-MDA by a deep neural network with a stacked autoencoder for predicting reservoir production. Journal of Petroleum Science and Engineering, 2020, 187: 106800.

Li, B., Bhark, E. W., Gross, S. I., et al. Best practices of assisted history matching using design of experiments. SPE Journal, 2019, 24(4): 1435-1451.

Liu, N., Oliver, D. S. Evaluation of Monte Carlo methods for assessing uncertainty. SPE Journal, 2003, 8(2): 188-195.

Ma, X., Zhang, K., Wang, J., et al. An efficient spatial-temporal convolution recurrent neural network surrogate model for history matching. SPE Journal, 2022, 27(2): 1160-1175.

Ma, X., Zhang, K., Zhang, L., et al. Data-driven niching differential evolution with adaptive parameters control for history matching and uncertainty quantification. SPE Journal, 2021, 26(2): 993-1010.

McPhee, J., Yeh, W. W. -G. Groundwater management using model reduction via empirical orthogonal functions. Journal of Water Resources Planning and Management, 2008, 134(2): 161-170.

Oliver, D. S., Fossum, K., Bhakta, T., et al. 4D seismic history matching. Journal of Petroleum Science and Engineering, 2021, 207: 109119.

Raissi, M., Perdikaris, P., Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 2019, 378: 686-707.

Rao, X., Xu, Y., Liu, D., et al. A general physics-based data-driven framework for numerical simulation and history matching of reservoirs. Advances in Geo-Energy Research, 2021, 5(4): 422-436.

Romero, C., Carter, J. Using genetic algorithms for reservoir characterisation. Journal of Petroleum Science and Engineering, 2001, 31(2-4): 113-123.

Rwechungura, R., Dadashpour, M., Kleppe, J. Advanced history matching techniques reviewed. Paper SPE 142497 Presented at SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 25-28 September, 2011.

Shorten, C., Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. Journal of Big Data, 2019, 6: 60.

Stone, N. Gaussian process emulators for uncertainty analysis in groundwater flow. Nottingham, University of Nottingham, 2011.

Tang, M., Liu, Y., Durlofsky, L. J. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. Journal of Computational Physics, 2020, 413: 109456.

Wang, N., Zhang, D., Chang, H., et al. Deep learning of subsurface flow via theory-guided neural network. Journal of Hydrology, 2020, 584: 124700.

Weiss, K., Khoshgoftaar, T. M., Wang, D. A survey of transfer learning. Journal of Big Data, 2016, 3: 9.

Winter, C., Tartakovsky, D., Guadagnini, A. Moment differential equations for flow in highly heterogeneous porous media. Surveys in Geophysics, 2003, 24(1): 81-106.

Woodbury, A. D., Dunbar, W. S., Nour-Omid, B. Application of the arnoldi algorithm to the solution of the advection-dispersion equation. Water Resources Research, 1990, 26(10): 2579-2590.

Yang, P. -H., Watson, A. T. Automatic history matching with variable-metric methods. SPE Reservoir Engineering, 1988, 3(4): 995-1001.

Yazdanpanah, A., Rezaei, A., Mahdiyar, H., et al. Development of an efficient hybrid GA-PSO approach applicable for well placement optimization. Advances in Geo-Energy Research, 2019, 3(4): 365-374.

Yeung, Y. H., Barajas-Solano, D. A., Tartakovsky, A. M. Physics-informed machine learning method for large-scale data assimilation problems. Water Resources Research, 2022, 58(5): e2021WR031023.

Yin, F., Xue, X., Zhang, C., et al. Multifidelity genetic transfer: an efficient framework for production optimization. SPE Journal, 2021, 26(4): 1614-1635.

Zenke, F., Poole, B., Ganguli, S. Continual learning through synaptic intelligence. Proceedings of Machine Learning Research, 2017, 70: 3987-3995.

Zha, W., Gao, S., Li, D., et al. Application of the ensemble Kalman filter for assisted layered history matching. Advances in Geo-Energy Research, 2018, 2(4): 450-456.

Zhang, K., Wang, X., Ma, X., et al. The prediction of reservoir production based proxy model considering spatial data and vector data. Journal of Petroleum Science and Engineering, 2022, 208: 109694.

Zhao, H., Kang, Z., Zhang, X., et al. A physics-based data-driven numerical model for reservoir history matching and prediction with a field application. SPE Journal, 2016, 21(6): 2175-2194.

Zhong, C., Zhang, K., Xue, X., et al. Historical window-enhanced transfer gaussian process for production optimization. SPE Journal, 2022, 27(5): 2895-2912.