Smart Geotechnics for Smart Societies – Zhussupbekov, Sarsembayeva & Kaliakin (Eds)

AI-based back analysis of multiphysics processes in unconventional resource extraction practice

M. Zhou
Department of Geotechnical Engineering, Tongji University, China

M. Shadabfar
Department of Civil Engineering, Sharif University of Technology, Iran

Y.F. Leung
Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, China

S. Uchida
Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, USA

ABSTRACT: Mulitphysics processes have been commonly identified in geotechnical engineering practice. Researchers and field engineers often carry out multiphysics simulations to understand complex engineering responses. In field practice, a back analysis is typically required along with the simulations to calibrate the most representative model parameters. This would intensify the problem as it requires further simulations to assess the parameter sensitivity. Therefore, an efficient back analysis for multiphysics processes still remains a challenge in practice due to the numerical complexity and the low computational efficiency. With recent advances in AI techniques, opportunities have opened up for meta-model development for problems involving multiphysics processes associated with a large number of properties. This study entails a meta-model developed based on Artificial Neural Networks (ANN) that intelligently learn the correlations between model parameters and the reservoir responses. This efficient meta-model is combined with Genetic Algorithm-based back analysis to report the optimal case that provides the closest output to the target time histories. The results show that the AI-based metamodel can reproduce outputs of heavy computation of the multiphysics processes and thus efficiently perform back-analysis.

1 INTRODUCTION

Mulitphysics processes have been commonly identified in unconventional resource extraction practice (Guo et al., 2019; Mahdi et al., 2017). However, such complex processes are often difficult to model in the field or in the laboratory due to the coupling of multiphysics (Touran et al., 2017; Keyes et al., 2013). Therefore, researchers and field engineers in the oil and gas industry often carry out numerical simulations to understand the complex engineering responses during production. Such simulation practices are conducted via multiphysics numerical simulators (Moridis et al., 2011; Klar et al., 2013; White et al., 2020; Zhou et al., 2020).

In the rise of unconventional energy resources, one of the multiphysics complex processes is gas production from gas hydrate reservoir. production of gas from hydrate-bearing sediments involves change in pressure, temperature and also mechanical behavior, leading to coupled thermo-hydro-mechanical (THM). However, modelling of this complex processes is often computationally demanding and requires a large number of model parameters to represent multiple facets of the hydrate reservoir behavior (Uchida et al., 2016). Even when measurement of

DOI: 10.1201/9781003299127-123

gas and water production history is available, conducting back analysis to calibrate these parameters is significantly time-consuming and also confirming that the calibrated parameters are the best representatives of the site condition is challenging because of potential existence of multiple local optima. Therefore, there is a need for an efficient tool to assess the hydrate reservoir production potential and the complex sediment responses with minimal time cost and computational complexity.

With recent advances in artificial intelligence (AI) techniques (e.g., machine learning and data analytics), opportunities have opened up for meta-model development for problems involving multiphysics processes associated a large number of properties (Tripoppoom et al., 2020; Li et al., 2021; Park et al., 2021; Xu et al., 2020). The main goal of the meta model is to provide an alternative tool to replace a conventional THM simulator in a more timely and computationally effective way. Once this computational tool is provided, the uncertainty in data can be assessed through varying the input parameters and investigating the uncertainty propagation through the sensitivity study.

This study entails a meta-model developed based on ANN that deep learn the correlations between model parameters (i.e., hydrologic and geomechanical parameters of hydrate-bearing sediments) and the reservoir response (i.e., gas and water production), and is capable of reproducing the production results with much higher computational efficiency than coupled numerical simulators based on finite element or finite difference methods. An optimization network based on genetic algorithm (GA) is then proposed and combined with the trained ANN to conduct the back-analysis of the site measured production data, and thus the optimum model parameters can be obtained.

The capabilities of the proposed machine learning approach are demonstrated through the application to the 2013 Nankai offshore gas production test. The training and testing data for the ANN are produced by a series of coupled THM simulations. The assumed ranges of material properties of the simulations are based on the production site conditions and are varied across the multi-dimensional sample space. The ANN predictions suggest good agreement with the measured production data, while the obtained model parameters from back analysis can be regarded as important properties for researchers and field engineers to focus on for reservoir characterization.

2 SYNTHETIC DATASET

This study creates synthetic data numerically through varying 19 material properties and obtaining the corresponding reservoir responses, namely, the amount of produced gas, produced water and vertical strain (Figure 1). A THM numerical simulator, originally developed by (Klar et al., 2013), is adopted to create synthetic data. The simulator is based on a finite difference software, FLAC, and solves the multiphysics processes through the implemented coupled THM formulation derived for hydrate reservoir simulations by (Uchida et al., 2016).

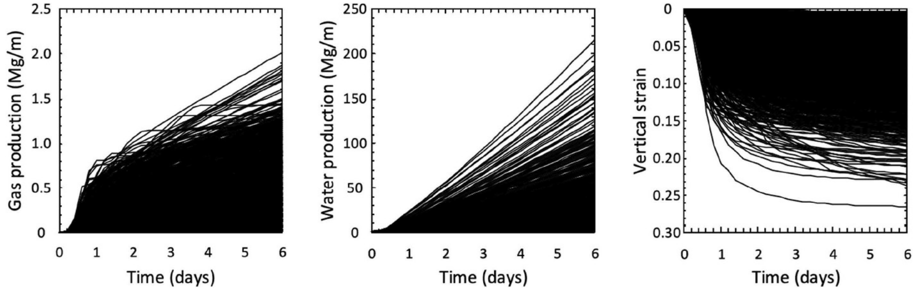

Figure 1. Synthetic data created by the THM analyzes.

Based on the Eastern Nankai Trough production site geometry, the targeted perforation region is between 280.7 m and 318.7 m below the seabed and the seabed is 998 m below the sea level. Figure 2 presents a hydrate reservoir considered for the creation of the synthetic dataset. It would be ideal if the model could explicitly simulate the heterogeneity of in-situ THM properties along the vertical direction, but that would impose a significant computational demand. Therefore, this study simplifies the complexity into a homogenized single radial layer. This simplification has a few shortcomings associated with the lack of thermal or fluid flow in the vertical direction, but the variations in the THM properties are considered in the form of the parameter range, while the most representative homogenized properties will be later determined through back analyses.

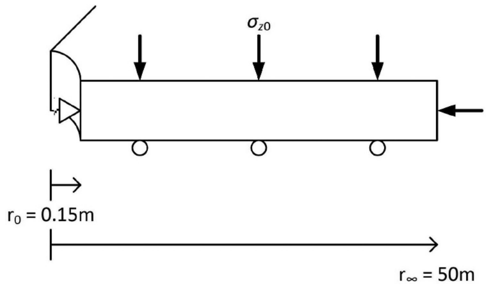

Figure 2. Hydrate reservoir model considered in this study.

The depth of the considered layer is assumed to be at 1297.7 m below sea level and 299.7 m below seafloor, which is the mid-depth of the production region and corresponding to the initial pore water pressure of 13 MPa and the initial effective vertical stress to be 3 MPa. The initial effective horizontal stress is assumed to be 1.5 MPa, which corresponds to the at-rest earth pressure coefficient of K0 = 0.5. The initial porosity is 0.4, the initial temperature is 285 K and the initial hydrate saturation is 61.3%, which is the average value over the production zone based on logging data (Konno et al., 2017). Under the initial temperature, the phase equilibrium pressure is 9.8 MPa. Depressurization is applied according to the measured bottom hole pressure during the 2013 test (Konno et al., 2017).

At the well boundary in the model, a free movement is assumed in the vertical direction while no radial displacement is allowed. The bottom boundary is fixed in the vertical direction and free in the radial direction. The top boundary has the constant total stress applied to facilitate vertical deformation. The outer model boundary is set at 50 m, where the pore pressure, the total radial stress and temperature remain unchanged. The layer is radially discretized into 35 elements with a size of 0.15 m adjacent to the well and increasing at a ratio of approximately 1.1.

This study utilizes the Latin hypercube sampling method (McKay et al., 1979) to ensure the multi-dimensional sampling domain is evenly explored with a relatively small number of simulations, resulting in 1000 simulation cases. Table 1 shows the maximum and minimum values of the 18 parameters adopted. The parameters are assumed to follow lognormal distribution, except the van Genuchten parameters, *b* and *c*, which are assumed to be normally distributed in the synthetic dataset. These simulation cases provide a history of reservoir production responses (gas production and water production), recorded at every 0.2 days for a 6-day period. Therefore, at every recorded time, there are 18000 data points regarding material properties (1000 times 18), which are "inputs" for the ANN. Meanwhile, there are also 2000 data points regarding reservoir production responses (1000 cases times 2) at each recorded time step, which are "outputs" for the ANN.

The established synthetic dataset is normalized, and 80% of the normalized data is selected as the training data and the remaining 20% is selected as the testing data. The normalization process facilitates the calculation by making all the data dimensionless. Additionally, it limits

Table 1. 18 model parameters selected for this study.

name	symbol	max	min
Hydrologic (8)			
initial intrinsic permeability	$\|\|K\|\|$	$10^{-4.8}m^2$	$10^{-7.2}\ m^2$
effective permeability power	N	10.0	2.0
air entry pressure	P_0	12kPa	8kPa
van Genuchten parameter	a	1.00	0.80
van Genuchten parameter	b	2.00	-1.00
van Genuchten parameter	c	2.00	-1.00
residual water saturation	S_{rw}	0.60	0.00
residual gas saturation	S_{rg}	0.20	0.00
Mechanical (10)			
critical state stress ratio	M	1.56	1.40
slope of swelling line	k	0.014	0.003
Poisson's ratio	v	0.40	0.20
slope of compression line	λ	0.25	0.14
initial preconsolidation stress	P'_{cs0}	6.0MPa	3.0MPa
hydrate dependent strength	α	100MPa	10MPa
hydrate dependent strenght	β	1.8	0.8
subloading ratio evolution	u	exp(5.52)	Exp(1.61)
hydrate dependent modulus	E_{h0}	12GPa	8GPa
hydrate degradation factor	m	21	1

the range of data to only vary between [-1,1] and consequently controls the algorithm search domain, which results in less complexity and higher accuracy of the ANN predictions.

3 AI-BASED BACK ANALYSIS

The ANN is developed to approximate the gas production and water production time histories. In this study, a multi-layer perceptron (MLP) is used to implement the ANN. The ANN consists of the dataset, two hidden layers of neurons, and the predictions of the reservoir responses. The prediction consists of 2 classes, each with 30 neurons, which represent the simulation results of gas production and water production at the 30 time intervals (0.2-day intervals for 6 days). Two hidden layers are implemented in this study to accomplish accurate predictions. By adopting a trial-and-error method to change the number of neurons in each hidden layer, the MLP model with ten neurons in the first hidden layer and five neurons in the second hidden layer reveals the best performance of the data set analyzed in this study.

The initial MLP model is then generated using the training dataset with the proposed layer structure. The training data is introduced into the ANN to train the network by determining the biases and weights by minimizing the mean squared error (MSE). Once the minimum MSE is obtained, the network training process is completed and can be used to make predictions and evaluate the accuracy using the testing dataset.

The established ANN provides a computationally light method to predict the gas and water production from the 18 input model parameters. An optimization problem is defined for the back analysis, in which the objective function is the difference between the optimal production resulted from the ANN and the actual site measured production values. Conventional optimization techniques often involve determining the gradient of the objective function. Due to the complexity of this optimization problem, it is challenging to calculate the partial derivatives of the objective function because the established ANN is not an explicit function of the input variable. Therefore, derivative-based methods may not be able to solve this problem. As such, GA is used as an alternative solution in this study.

GA is an intelligent optimization method inspired by the theory of evolution by natural selection. This algorithm improves the initial population in the form of an iterative process

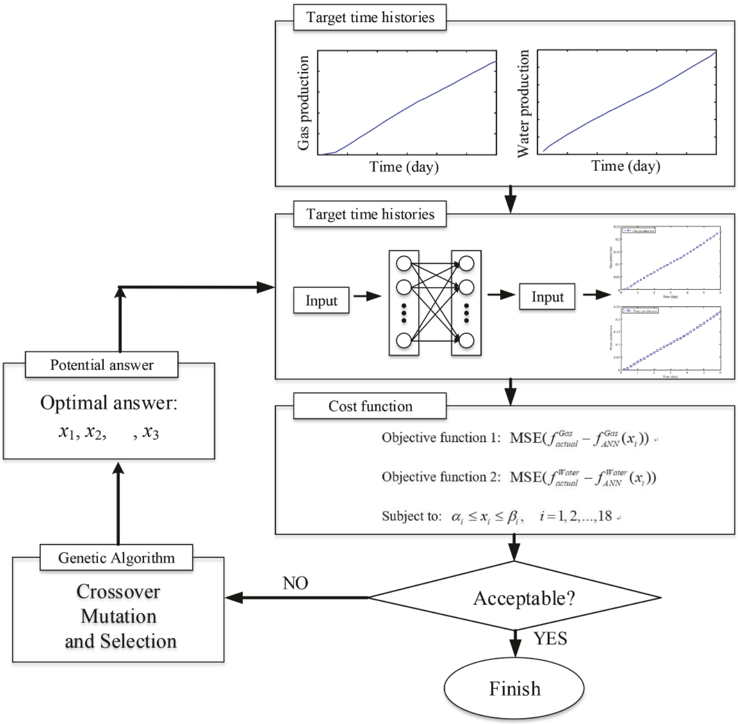

Figure 3. The schematic diagram of the proposed approach.

according to a quality criterion. In this respect, three operators of mutation, crossover, and selection are used to drive the evolution of candidate solutions towards the optimum in each generation. Because the algorithm is population-based and does not require the derivatives of objective function or constraint, it is widely used with implicit function, such as the ANN utilized in this paper. In addition, GA has the ability to search a large number of candidate solutions to converge to the global optimum. As with any iterative method, GA needs a set of termination conditions to determine when to stop and complete the analysis. The termination criteria used in this paper are as follows:

1. when the difference of optimum response in each two subsequent steps is less than 0.1%.
2. when the number of generations reaches 100.
3. when the algorithm reaches the time limit (5 minutes) set for the analysis.

When one of the above three conditions is met, the analysis considers the results as the optimal solutions.

4 RESULTS

The performance of the ANN was evaluated by comparing its predictions with the results of THM simulations, across the whole set of 1000 simulations (Figure 4). The error of the prediction results for both the training and the testing data are relatively small, which suggests that the proposed ANN can make accurate predictions. Hence, by selecting appropriate input model parameters, the established ANN can be used for history matching of the site measured production responses.

The back analysis was implemented in the form of an optimization problem and GA was utilized to provide a solution. It is worth noting that two objective functions are involved in this problem since both gas and water productions are considered in this study. In this study,

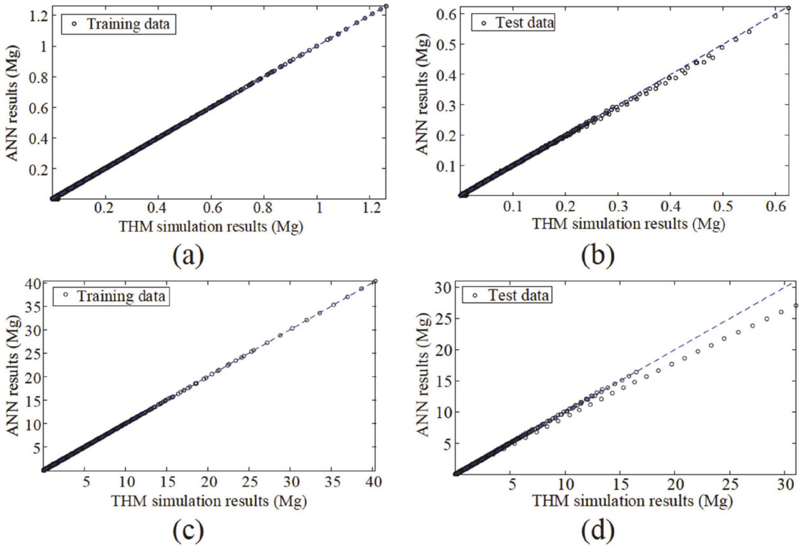

Figure 4. Comparison of the ANN predictions and the THM simulation results of (a) gas production comparison for training data, (b) gas production comparison for testing data, (c) water production comparison for training data, (d) water production comparison for testing data.

the gas and water production values per unit volume are used as the optimization target of the back analysis. Based on the site measured gas and water production rate, the total gas and water production values per unit volume for the first six days are computed and shown in Figure 5. Using the 18 input model parameters, the gas and water production per unit volume can be derived from the established ANN, and the predicted results are plotted in Figure 5 to compare with the target values. The comparison suggested a good match. Therefore, the proposed back analysis method works well for the history matching purpose.

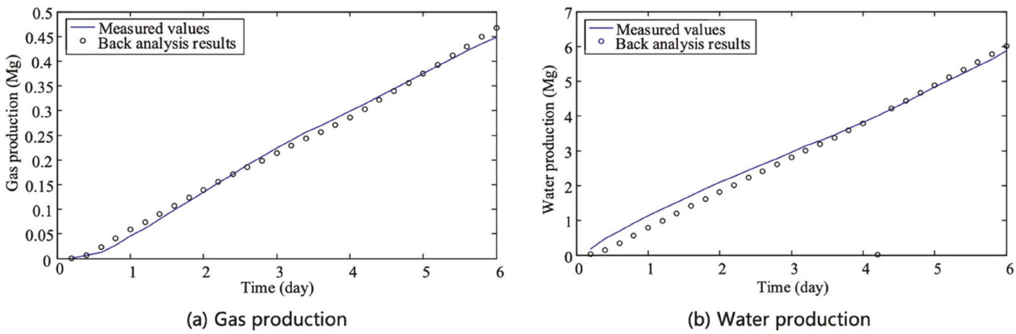

Figure 5. Comparison of the site measured production data and the predicted data based on the back analyzed input parameters: (a) gas production, (b) water production.

The computational efficiency of the proposed meta-model is demonstrated by comparing its processing time with the THM simulator for three specific reservoir cases (Case 300, 600, and 900). The numerical experiments for 8-day gas production were carried out on a desktop equipped with two GeForce GTX 1080 graphics processing units (GPUs), 64 GB of random access memory (RAM), and one Intel Core i7-5820K central processing unit (CPU). For 6-day gas production prediction, the execution time of the ANN and the FLAC THM simulator for three specific reservoir cases are shown in Table 2. It is seen that a typical calculating of the FLAC THM simulator for 6-day gas production takes on average more than 1000 seconds while the ANN only takes less than 0.1 seconds.

Table 2. For 6-day gas production prediction, the execution time of the ANN and the FLAC THM simulator for Case 300, 600, and 900.

Prediction method	Case 300	Case 600	Case 900
First-block ANN	0.071 seconds	0.069 seconds	0.067 seconds
FLAC THM simulator	1400 seconds	7000 seconds	540 seconds

5 CONCLUSION

This paper utilizes a combination of an ANN and an optimization model for simulation of gas production operations in methane hydrate-bearing reservoir. The combined model enables back analyses of site response in order to obtain the corresponding site properties. The site measured production data in 2013 Eastern Nankai Trough production test is adopted in this study. The ANN successfully learns the relationship between the material properties and reservoir productions, while the optimization model demonstrated the capability of finding the optimum combinations of material properties to characterize the reservoir production.

The main advantage of the proposed approach lies on its computational efficiency for history matching and model parameter calibration, which can be ten thousands of times faster than the numerical analysis conducted via the thermo-hydro-mechanical simulator.

It should be noted that the proposed machine learning approach only learns what it is being taught. The effectiveness of the back analyses clearly depends on the geometries of the specific problem, and the site-specific hydrate and stratigraphy conditions. Nevertheless, the proposed framework allows real-time prediction to be made and adjusted according to the observed reservoir response as the process evolves with time at the production site.

REFERENCES

X. Guo, D. Hu, Y. Li, J. Duan, X. Zhang, X. Fan, H. Duan, W. Li, Theoretical progress and key technologies of onshore ultra-deep oil/gas exploration, Engineering 5 (3) (2019) 458–470. https://doi.org/10.1016/j.eng.2019.01.012.

M. McKay, R. Beckman, W. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21 (1979) 239–245.

D. E. Keyes, L. C. McInnes, C. Woodward, W. Gropp, E. Myra, M. Pernice, J. Bell, J. Brown, A. Clo, J. Connors, et al., Multiphysics simulations: Challenges and opportunities, The International Journal of High Performance Computing Applications 27 (1) (2013) 4–83.

A. Klar, S. Uchida, K. Soga, K. Yamamoto, Explicitly coupled thermal flow mechanical formulation for gas-hydrate sediments, SPE Journal 18 (02) (2013) 196–206. doi:10.2118/ 162859-PA.

H. Li, H. Yu, N. Cao, H. Tian, S. Cheng, Applications of artificial intelligence in oil and gas development, Archives of Computational Methods in Engineering 28 (3) (2021) 937–949.

Y. Konno, T. Fujii, A. Sato, K. Akamine, M. Naiki, Y. Masuda, … & J. Nagao. Key findings of the world's first offshore methane hydrate production test off the coast of Japan: Toward future commercial production. Energy & Fuels, 31(3), (2017) 2607–2616.

S. Mahdi, X. Wang, N. Shah, Interactions between the design and operation of shale gas networks, including co2 sequestration, Engineering 3 (2) (2017) 244–256. https://doi.org/10.1016/J.ENG.2017.02.007.

G. Moridis, S. Silpngarmlert, M. Reagan, T. Collett, K. Zhang, Gas production from a cold, stratigraphically-bounded gas hydrate deposit at the Mount Elbert gas hydrate stratigraphic test well, Alaska north slope: Implications of uncertainties, Marine and petroleum geology 28 (2) (2011) 517–534.

T.-R. Park, H. Park, K. Kim, C.-N. Im, J.-H. Cho, Heat and weight optimization methodology of thermal batteries by using deep learning method with multi-physics simulation, Energy Conversion and Management 236 (2021) 114033. https://doi.org/10.1016/j.enconman.2021.114033.

N. W. Touran, J. Gilleland, G. T. Malmgren, C. Whitmer, W. H. Gates, Computational tools for the integrated design of advanced nuclear reactors, Engineering 3 (4) (2017) 518–526. https://doi.org/10.1016/J.ENG.2017.04.016.

S. Tripoppoom, J. Xie, R. Yong, J. Wu, W. Yu, K. Sepehrnoori, J. Miao, C. Chang, N. Li, Investigation of different production performances in shale gas wells using assisted history matching: Hydraulic fractures and reservoir characterization from production data, Fuel 267 (2020) 117097. https://doi.org/10.1016/j.fuel.2020.117097.

S. Uchida, A. Klar, K. Yamamoto, Sand production model in gas hydrate-bearing sediments, International Journal of Rock Mechanics and Mining Sciences 86 (2016) 303–316. doi:10.1016/j.ijrmms.2016.04.009.

M. White, T. Kneafsey, Y. Seol, W. F. Waite, S. Uchida, J. Lin, E. Myshakin, X. Gai, S. Gupta, M. Reagan, et al., An international code comparison study on coupled thermal, hydrologic and geomechanical processes of natural gas hydrate-bearing sediments, Marine and Petroleum Geology 120 (2020) 104566. doi:10.1016/j.marpetgeo.2020.104566.

H. Xu, J. Ma, P. Tan, B. Chen, Z. Wu, Y. Zhang, H. Wang, J. Xuan, M. Ni, Towards online optimisation of solid oxide fuel cell performance: Combining deep learning with multi-physics simulation, Energy and AI 1 (2020) 100003. https://doi.org/10.1016/j.egyai.2020.100003

M. Zhou, K. Soga, K. Yamamoto, Upscaling techniques for fully coupled thm simulation and application to hydrate gas production tests, Computers and Geotechnics 124 (2020) 103596. https://doi.org/10.1016/j.compgeo.2020.103596.