ARMA 23-0834

Subsurface Characterization using Bayesian Deep Generative Prior-based Inverse Modeling for Utah FORGE Enhance Geothermal System

Lee, Jonghyun and Bao, Jichao
University of Hawaii at Manoa, Honolulu, HI, USA
Yoon, Hongkyu
Sandia National Laboratories, City, State, Country
Pyrak-Nolte, Laura
Purdue University, West Lafayette, IN, USA

Characterization of geologic heterogeneity at an enhanced geothermal system (EGS) is crucial for cost-effective stimulation planning and reliable heat production. With recent advances in computational power and sensor technology, large-scale fine-resolution simulation of coupled thermal-hydraulic-mechanical (THM) processes have been available. However, traditional large-scale inversion approaches have limited utility for the sites with complex subsurface structures unless one can afford high, often computationally prohibitive, computations. Key computational burdens are predominantly associated with a number of large-scale coupled numerical simulations and large dense matrix multiplications derived from fine discretization of the field site domain and a large number of THM and chemical (THMC) measurements.
In this work, we present deep-generative model-based Bayesian inversion methods for the computationally efficient and accurate characterization of EGS sites. Deep generative models are used to learn the approximate subsurface property (e.g., permeability, thermal conductivity and elastic rock properties) distribution from multipoint geostatistics-derived training images or discrete fracture network models as a prior and accelerated stochastic inversion is performed on the low-dimensional latent space in a Bayesian framework. Numerical examples with synthetic permeability fields with fracture inclusions with THM data sets based on Utah FORGE geothermal site will be presented to test the accuracy, speed, and uncertainty quantification capability of our proposed joint data inversion method.

1 INTRODUCTION

For the subsurface inverse problem (Carrera et al., 2005; McLaughlin and Townley, 1996; Oliver et al., 2008), spatially distributed geologic parameters such as permeability and porosity are estimated from noisy and sparse hydrogeophysical and geomechanical measurements such as pressure, temperature, displacement, seismic responses, and so on. Due to ill-posedness in the inverse problem, probabilistic frameworks have been implemented in order to account for the uncertainty that provides both unknown subsurface parameters and their corresponding uncertainty in a Bayesian statistical framework (e.g., Kitanidis, 1995, 2010; Lee et al., 2016). Typically pressure measurements from injection-extraction well operations are used for the site characterization, but unless many injection/extraction/observation wells are available, the subsurface characterization with sparse pressure measurements leads to poor predictions of other relevant quantities such as heat transport. To overcome this issue, different sensing data can be used for joint inversion of pressure, temperature, displacement, and seismicity to identify the subsurface geologic connectivity during the geothermal energy recovery operations (Lee et al., 2018)

Still, one of the challenges in subsurface characterization is to identify highly heterogeneous permeability field such as geothermal reservoirs with fracture networks. For such cases, traditional techniques using the Gaussian prior lead to smoothed, low-resolution images of subsurface properties with high estimation uncertainty due to the diffusive nature of the governing equations. Machine learning techniques such as deep generative models (Kang et al., 2022, 2021; Kadeethum et al., 2021; Kim et al., 2021) have been actively studied to enforce the subsurface solution space on the relevant prior distribution constructed through multipoint geostatiatics-based training images.

In this work, we present an application of recent inverse modeling method based on the deep generative model for characterizing the geothermal site at a reduced computational cost with improved accuracy compared to traditional inversion. The multiphysics model MOOSE-FALCON (Podgorney et al., 2021) for simulating THM processes at the geothermal site is combined with our proposed variational latent space inversion method. A syn-

thetic application with pressure and heat data is considered in this work for a two dimensional, strongly heterogeneous geothermal site. Inversion of synthetic pumping tests performed at this site are performed to show the usefulness and practicality of the proposed approach for geothermal site characterization with a reasonable accuracy. We will incorporate the geomechanical data to improve the accuracy of the inversion in the near future.

2 METHOD

2.1. Deep Generative Modeling

Deep generative models (Goodfellow et al., 2014) have been studied because of their capability to approximate data distributions from training samples and generate new samples from the approximated data distribution in an efficient fashion. Among many deep generative models such as variational autoencoder, diffusion model, normalized flow and so on, we use the generative adversarial network (GAN) (Goodfellow et al., 2014) which is composed of a generator and a discriminator. The generator create the samples from the data distribution using randomly sampled latent variables whose dimension is much smaller than the data dimension in general and the discriminator determines whether the data are authentic. To address the issue of mode collapses, Wasserstein GAN with gradient penalty (WGAN-GP) is used as an improved version of GAN (Gulrajani et al., 2017) to learn the subsurface property distribution such as fracture networks. The inversion solutions are constrained to the learned latent space while consistent with the observations. This will lead to our deep generative model-based parameterization of the permeability field:

$$\mathbf{s} = G(\mathbf{z}) \tag{1}$$

where $\mathbf{s}$ is the permeability field and $\mathbf{z}$ is a k ($\ll m$)-dimensional latent variable constructed from GAN-GP. Using our generator $G(z)$, the variable $\mathbf{z}$ in the context of the inversion problem represents the low-dimensional representation of $\mathbf{s}$.

2.2. Variational Inversion using Deep Generative Modeling

We use a Hierachical Bayesian approach combined with GAN-GP in this work for the inverse problem solver. The forward problem with the input of permeability field $\mathbf{s}$ can be defined in the form of the relationship

$$\mathbf{y} = \mathbf{h}(\mathbf{s}) + \boldsymbol{\varepsilon}, \tag{2}$$

where $\mathbf{y}$ is the observation (*e.g.*, pressure, temperature and displacement), $\boldsymbol{\varepsilon}$ is the observation and model uncertainty noise such as a Gaussian distribution with mean zero $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$. Here, $\mathbf{R}$ is the model/observation error matrix and $\mathbf{h}$ is the forward map. The inverse problem equivalent of Eq. 2 can be defined as a problem with unknown m-dimensional variable $\mathbf{s}$ (permeability) and n (noisy) observation $\mathbf{y}$ (pressure, temperature and displacement). The Bayes' rule allows us to evaluate a posterior distribution of $\mathbf{s}$ via

$$p(\mathbf{s}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{s})p'(\mathbf{s}) = \int_{\boldsymbol{\theta}} p(\mathbf{y}|\mathbf{s})p'(\mathbf{s}|\boldsymbol{\theta})p'(\boldsymbol{\theta})d\boldsymbol{\theta}, \tag{3}$$

where $p'(\cdot)$ represents the prior probability and $\boldsymbol{\theta}$ is a set of hyperparameters that models $\mathbf{s}$ in a hierarchical Bayesian framework.

In particular, with the generator G, we can constrain variable $\mathbf{z}$ to follow a Gaussian distribution to ensure the regularity of the latent space:

$$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{4}$$

We also assume that G is a deterministic map from $\mathbf{z}$ to $\mathbf{s}$ and the hyperparameter $p'(\boldsymbol{\theta})$ such as neural network model parameters follows a delta distribution:

$$\boldsymbol{\theta} = \boldsymbol{\delta}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}). \tag{5}$$

This allows us to rewrite Eq. 3 in the form

$$\begin{aligned} p(\mathbf{z}|\mathbf{y}) &\propto p(\mathbf{y}|\mathbf{z})p'(\mathbf{z}) \\ &\propto \exp\left(-(\mathbf{y} - \mathbf{h}(G(\mathbf{z})))^\top \mathbf{R}^{-1}(\mathbf{y} - \mathbf{h}(G(\mathbf{z})))\right) \exp\left(-\mathbf{z}^\top \boldsymbol{\Sigma}^{-1}\mathbf{z}\right), \end{aligned} \tag{6}$$

We explore this posterior equation to identify the latent variables consistent with observations through sampling or variational approaches. Here we use variational, i.e., optimization-based approach to find the maximum a posteriori (MAP) estimate, which is the mode of the posterior distribution. This task will require the computation of Jacobian $\mathbf{J}^l$ and at the lth iteration the computation is given by

$$\mathbf{J}^l = \left.\frac{\partial \mathbf{h}(G(\mathbf{z}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}^l} = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right|_{\mathbf{s}=G(\mathbf{z}^l)} \left.\frac{\partial \mathbf{s}}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}^l} = \left.\mathbf{J}_\mathbf{h}\right|_{\mathbf{s}=G(\mathbf{z}^l)} \left.\mathbf{J}_G\right|_{\mathbf{z}=\mathbf{z}^l}. \tag{7}$$

Note that $\frac{\partial \mathbf{s}}{\partial \mathbf{z}}$ can be evaluated analytically using automatic differentiation (AD). This information can be provided at no additional computational cost in common libraries such as TensorFlow and PyTorch. Since the dimension of $\mathbf{z}$ is assumed to be significantly smaller than the dimension of $\mathbf{s}$, e.g., $dim(z) \leq 100$, we can also use a finite difference formulation to calculate the Jacobian matrix as an alternative to AD.

3 SYNTHETIC EXAMPLE SETUP

The domain for heterogeneity characterization is a synthetic two dimensional depth-integrated deep formation based on the Utah FORGE site for geothermal energy recovery operations. As shown in Figure 1, a 600 m × 600 m domain was discretized into 120 × 120 meshes. A constant pressure of 34 MPa and 30 MPa was imposed on the western and eastern boundaries, respectively. A constant temperature of 473.15 K and 423.15 K was applied to the western and eastern boundaries, respectively. One injection well was placed for transient simulations and nine monitoring wells were used for observation. The injection rates and temperature are shown in Figure 2. The injection rates increased from 0 to 1 g/s over 0 to 500 s and then kept constant at 1 g/s until 1000 s. The injection temperature was a constant value of 323.15 K. The values of the material parameters are shown in Table 1. Permeability and porosity are heterogeneous as shown in Figure 3, using the same image (generated by the trained generator) but assigning different values. For permeability, the fractures (red lines) were assigned 1e-14 m^2, and other areas were assigned 1e-16 m^2. For porosity, the fractures were assigned 1e-3, and other areas were assigned 1e-4.

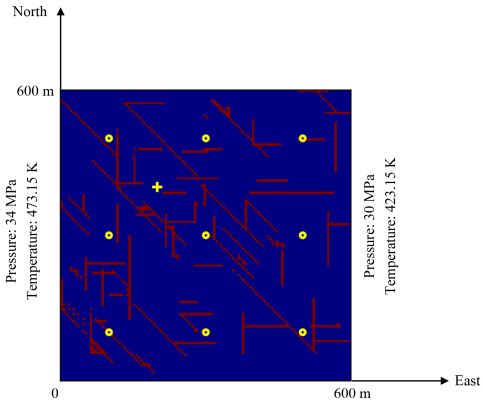

Fig. 1: Model settings. 34 MPa pressure and 473.15 K temperature were applied to the western boundary. 30 MPa pressure and 423.15 K temperature were applied to the eastern boundary. The yellow circles represent monitoring wells, and the cross shows the injection well location.

4 PRELIMINARY RESULTS

Figure 4 shows the estimation of permeability field which shows reasonable characterization of subsurface fracture features because of the informative training data for the deep generative model construction. Figure 5 shows the fitting results of pressure and temperature data. The fitting errors are also shown as the root mean square error (RMSE):

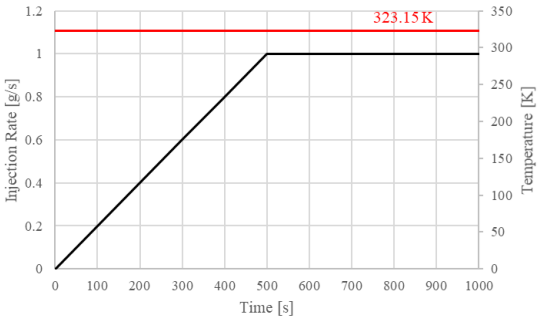

Fig. 2: Injection rates and temperature

Table 1: Material parameters

Parameter	Unit	Value
Permeability	m^2	heterogeneous
Porosity	–	heterogeneous
Density	kg/m^3	2640
Specific heat	$J/kg \cdot K$	790
Thermal conductivity	$W/m \cdot K$	3.05

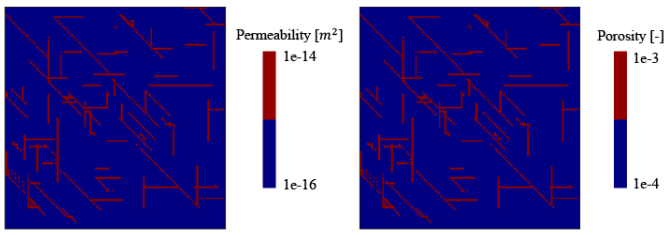

Fig. 3: Permeability and porosity

$$RMSE = \sqrt{\frac{1}{N_k}\sum_{i=1}^{N_k}(\mathbf{k}_i^{est} - \mathbf{k}_i^{true})^2} \quad (8)$$

where $\mathbf{k}_i^{est}$ indicates the estimated parameter value and $\mathbf{k}_i^{true}$ is the true value, N_k is the total number of parameters.

The RMSE of pressure data fitting is 0.358 while the RMSE of temperature data fitting is much smaller. We estimated permeability and porosity together and the accuracy of the porosity estimate is similar to the permeability estimation.

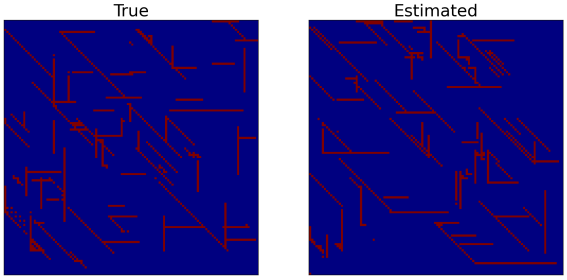

Fig. 4: Estimation result; the true permeability field is shown on the left and the estimated field is shown on the right.

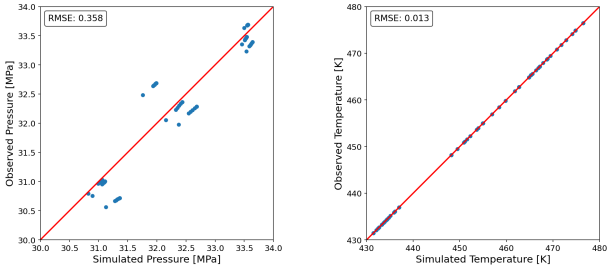

Fig. 5: Data fitting. Pressure and temperature data fitting. Fitting errors are shown in RMSE.

5 CONCLUDING REMARKS

We implemented a deep generative model-based inversion approach to perform joint data inversions with multiple pumping tests and presented reasonable inversion results with affordable FALCON runs. The proposed method transforms an inverse problem with the computational cost associated with the number of observations into an approximately same problem with a constant number (total O(100)) of simulations, so that one would expect a great computational gain in solving high-dimensional inverse problems. In the examples presented, the estimated fields captured important permeability features, i.e.,fractures, due to the informative prior used in the training. It is also observed that the estimate with improved accuracy was obtained at a much smaller computational cost than traditional methods, allowing for the characterization of a 2D geothermal reservoir in less than 10 minutes on a workstation equipped with 48 CPU cores. Joint data inversion from pressure and heat tracer with the pre-trained deep generative model can be beneficial to identify connectivity features in the site without considerably increasing the computational costs. Joint inversions using pressure, temperature and displacement data will be performed in the near future.

6 ACKNOWLEDGEMENT

Funding provided by DOE EERE Geothermal Technologies Office to Utah FORGE and the University of Utah under Project DE-EE0007080 Enhanced Geothermal System Concept Testing and Development at the Milford City, Utah Frontier Observatory for Research in Geothermal Energy (Utah FORGE) site.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. This article has been authored by an employee of National Technology & Engineering Solutions of Sandia, LLC under Contract No. DE-NA0003525 with the U.S. Department of Energy (DOE). The employee owns all right, title, and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

REFERENCES

1. Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology Journal*, 13(1):206–222.

2. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.

(2014). Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.

3. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., and Courville, A. C. (2017). Improved training of wasserstein gans. *Advances in neural information processing systems*, 30.

4. Kadeethum, T., OâMalley, D., Fuhg, J. N., Choi, Y., Lee, J., Viswanathan, H. S., and Bouklas, N. (2021). A framework for data-driven solution and parameter estimation of pdes using conditional generative adversarial networks. *Nature Computational Science*, 1(12):819–829.

5. Kang, X., Kokkinaki, A., Kitanidis, P. K., Shi, X., Lee, J., Mo, S., and Wu, J. (2021). Hydrogeophysical characterization of nonstationary dnapl source zones by integrating a convolutional variational autoencoder and ensemble smoother. *Water Resources Research*, 57(2):e2020WR028538.

6. Kang, X., Kokkinaki, A., Shi, X., Yoon, H., Lee, J., Kitanidis, P. K., and Wu, J. (2022). Integration of deep learning-based inversion and upscaled mass-transfer model for dnapl mass-discharge estimation and uncertainty assessment. *Water Resources Research*, 58(10):e2022WR033277.

7. Kim, S. E., Yoon, H., and Lee, J. (2021). Fast and scalable earth texture synthesis using spatially assembled generative adversarial neural networks. *Journal of Contaminant Hydrology*, 243:103867.

8. Kitanidis, P. K. (1995). Quasi-linear geostatistical theory for inversing. *Water Resources Research*, 31(10):2411–2419.

9. Kitanidis, P. K. (2010). *Bayesian and Geostatistical Approaches to Inverse Problems*, pages 71–85. John Wiley & Sons, Ltd.

10. Lee, J., Kokkinaki, A., and Kitanidis, P. K. (2018). Fast large-scale joint inversion for deep aquifer characterization using pressure and heat tracer measurements. *Transport in Porous Media*, 123:533–543.

11. Lee, J., Yoon, H., Kitandis, P. K., Werth, C. J., and Valocchi, A. J. (2016). Scalable subsurface inverse modeling for a big data set with an application to massive magnetic resonance imaging tracer concentration data inversion. *Water Resources Research*, 52(7):5213–5231.

12. McLaughlin, D. and Townley, L. R. (1996). A reassessment of the groundwater inverse problem. *Water Resources Research*, 32(5):1131–1161.

13. Oliver, D. S., Reynolds, A. C., and Liu, N. (2008). *Inverse theory for petroleum reservoir characterization and history matching*. Cambridge University Press, Cambridge ; New York.

14. Podgorney, R., Finnila, A., Simmons, S., and McLennan, J. (2021). A reference thermal-hydrologic-mechanical native state model of the utah forge enhanced geothermal site. *Energies*, 14(16):4758.