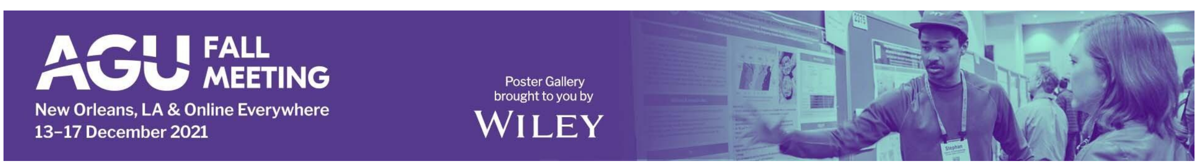
Physics-based Deep Learning Driven CO2 Flow Modeling and Data Assimilation for Real-Time Forecasting

Jichao Bao (1), Jonghyun Harry Lee (1), and Hongkyu Yoon (2)

(1) University of Hawaiʻi at Mānoa, Honolulu, HI (2) Sandia National Laboratories, Albuquerque, NM

PRESENTED AT:



INTRODUCTION

Scalable geologic carbon storage operations require fast forward modeling and forecasting. Real-time forecasting with data assimilation approaches usually encounters high-dimensional, ill-posed, and underdetermined (i.e., more unknowns than observations) problems. Furthermore, multiple simulations are often required to quantify the uncertainty, which is time-consuming.

Here, we will present a framework of coupling variational autoencoder (VAE) and (ensemble-based) data assimilation ((En)DA) for fast and accurate history matching of CO_2 operations and real-time forecasting of pressure development. A deep learning-based modeling approach that combines convolutional neural network (CNN), long-short term memory (LSTM), and dense neural network (DNN) is adopted to perform faster simulation of multiphase CO_2 flow and pressure propagation. The high-dimensional state variables (e.g., permeability) are reparametrized with the low-dimensional latent variables in VAE. The latent variables are updated by the (En)DA approach using the observed pressure data. The updated latent variables are then used by the decoder of VAE to produce updated state variables, which are the inputs to the CNN-LSTM-DNN model. Clastic Shelf data will be used to show the performance of permeability estimation and real-time pressure forecasting.

VARIATIONAL AUTOENCODER-BASED DATA ASSIMILATION

For a forward problem $\mathbf{y} = \mathbf{G}(\mathbf{m})$, with l Gauss-Newton iterations from $\mathbf{m}^0 = \mathbf{m}_{prior}$, the permeability field $\mathbf{m}$ is updated as follows:

$$\mathbf{m}^{l+1} = \mathbf{m}^0 + \mathbf{C}_{prior}\mathbf{J}(\mathbf{J}\mathbf{C}_{prior}\mathbf{J}^T + \mathbf{C}_{obs})^{-1}(\mathbf{y} - \mathbf{G}(\mathbf{m}^l) + \mathbf{J}(\mathbf{m}^l - \mathbf{m}^0))$$

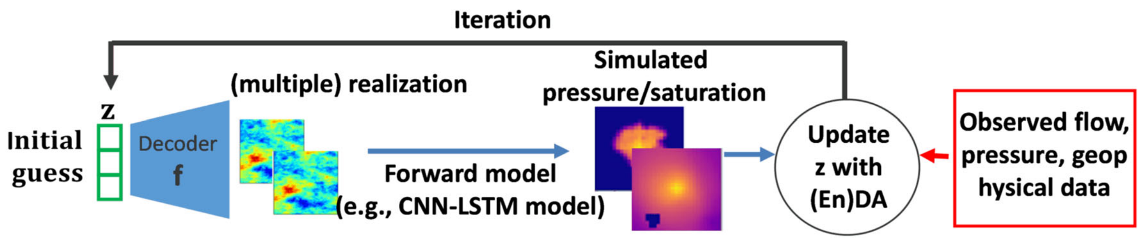


Figure 1. The framework of variational autoencoder-based data assimilation [ref. Poster H15O-1226]

Using the decoder of VAE to map $\mathbf{m}$ to latent space $\mathbf{z}$, the forward problem becomes $\mathbf{y} = \mathbf{G}(\mathbf{D}(\mathbf{z}))$, where $dim(\mathbf{z}) \ll dim(\mathbf{m})$. With this dimension reduction, the data assimilation process can be applied to latent space $\mathbf{z}$:

$$\mathbf{z}^{l+1} = \mathbf{z}^l + \alpha(\mathbf{J}_\mathbf{z}^\mathbf{T}\mathbf{C}_{obs}^{-1}\mathbf{J}_\mathbf{z} + \mathbf{C}_{prior(\mathbf{z})}^{-1})^{-1}(\mathbf{y} - \mathbf{G}(\mathbf{D}(\mathbf{z}^l)) - \mathbf{C}_{prior(\mathbf{z})}^{-1}\mathbf{z}^l)$$

PERMEABILITY ESTIMATION

Clastic Shelf Data

- 211 x 211 unknowns to 32 latent variables
- 4 injection wells, 2 monitoring wells, and 3 testing wells
- Observation data: 1.5, 2, 5, and 10 years pressure (monthly data up to 2 yrs, followed by yearly data to 10 yrs)

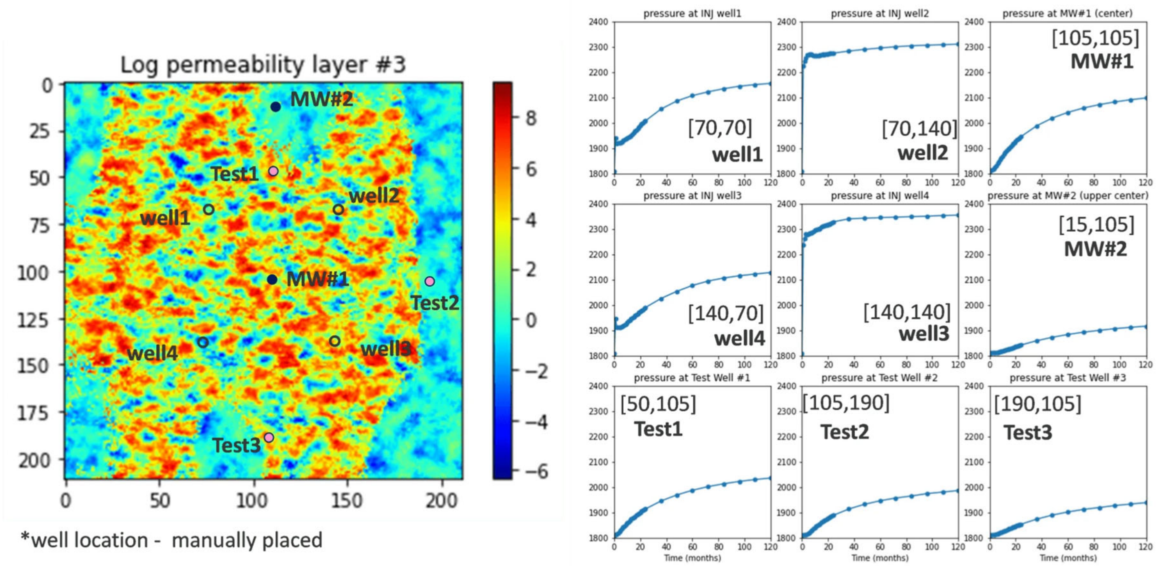


Figure 2. Log permeability field with well locations and pressure profiles at 4 injection, 2 monitoring, and 3 testing wells

Estimation Results

- Scenario 1: observed pressure data from 4 injection wells
- Scenario 2: observed pressure data from 4 injection wells and 2 monitoring wells

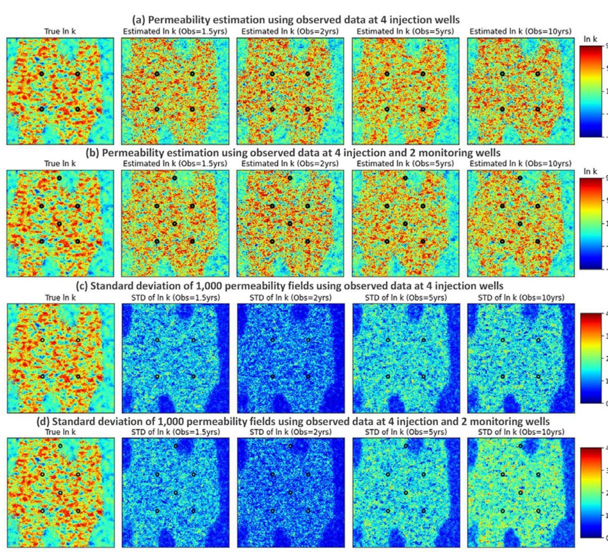


Figure 3. (a-b) Estimated permeability fields of two scenarios with different observation periods. (c-d) Standard deviation of 1,000 realizations from post-covariance analysis.

PRESSURE PREDICTION

- Both scenarios match observed pressure data at three testing wells similarly
- Low sensitivty with additional monitoring well data stems from relatively high errors of forward model at the injection wells compared to the rest of model domain (e.g., 4% at injection wells vs less than 0.5% error in the rest of the domain in CNN-LSTM-DNN model accuracy)

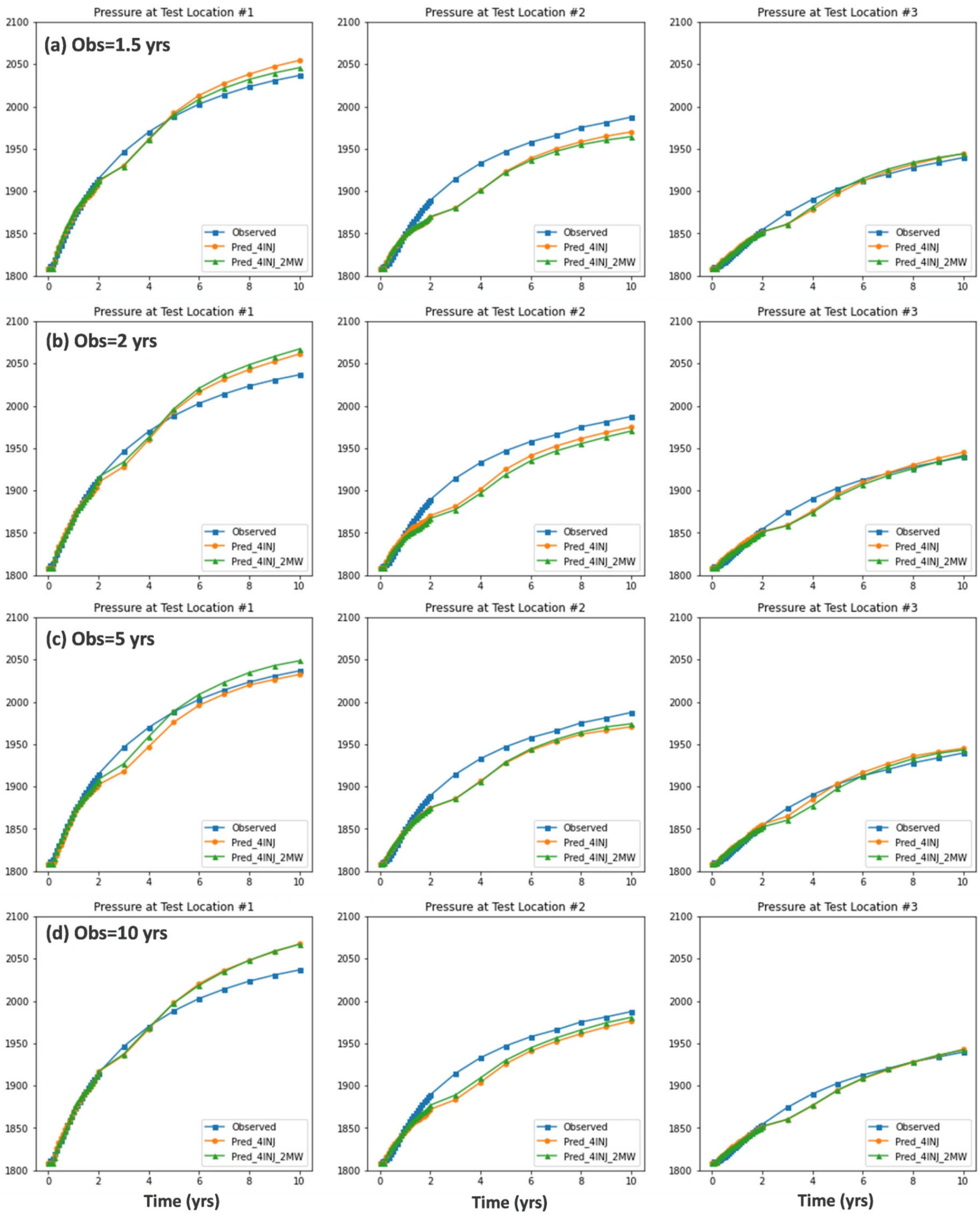


Figure 4. Comparison of observed and predicted pressure data over time at three testing locations with varying observation periods. The predicted pressure data are generated based on the estimated permeability fields shown in Figure 3 (4INJ: Scenario 1, 4INJ_2MW: Scenario 2).

CONCLUSIONS

The VAE-based data assimilation framework demonstrates promising performance for permeability estimation with uncertainty quantification and accurate prediction of pressure development. Furthermore, the proposed approach is computationally efficient which takes 3-5 min for the data assimilation process. The training of CNN-LSTM-DNN model takes 10-15 minutes using NVIDIA Quadro RTX 6000, and the training of VAE takes around 10 minutes. We believe that our method is applicable for 3D and more complex cases because of its computational efficiency and accurate estimation ability.

REFERENCES

J. Lee and H. Yoon, Subsurface Characterization using Deep Learning Approaches (https://agu2021fallmeeting-agu.ipostersessions.com/Default.aspx?s=99-4B-2D-77-56-0A-B9-BA-CE-BE-46-19-08-71-D9-0E) [Poster H15O-1226]

Mojtaba Forghani, Yizhou Qian, Jonghyun Lee, Matthew Farthing, Tyler Hesser, Peter K. Kitanidis, Eric F. Darve, Variational encoder geostatistical analysis (VEGAS) with an application to large scale riverine bathymetry, 2021 https://arxiv.org/abs/2111.11719

Acknowledgments:

This work is supported by US Department of Energy Office of Fossil Energy and Carbon Management, Science-Informed Machine Learning to Accelerate Real Time Decisions-Carbon Storage (SMART-CS) initiative.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell Internation, Inc., for the U.S. Department of Enery's National Nuclear Security Administration under contract DE-NA0003525.