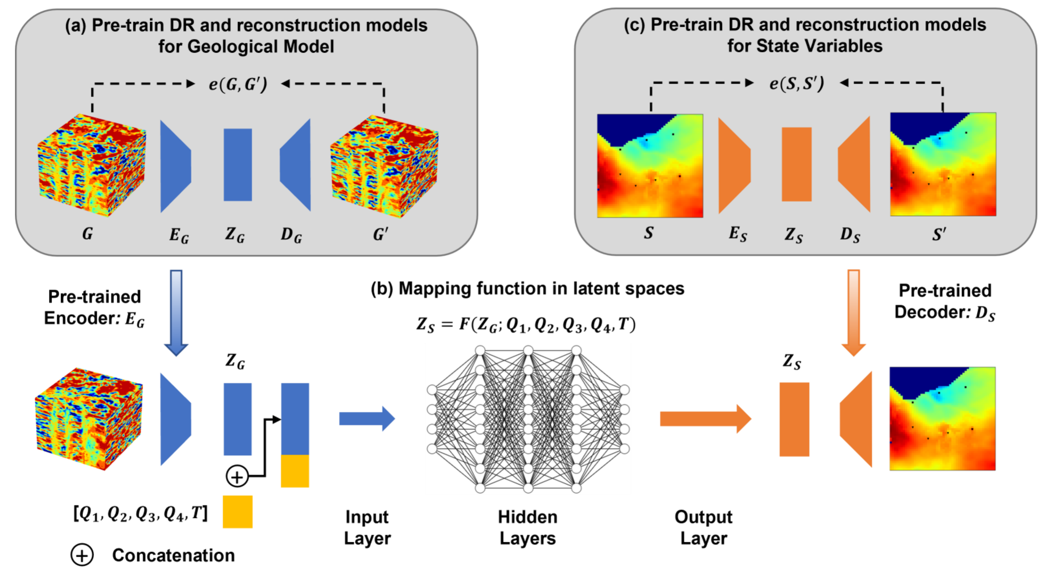
Reduced dimensional representations of subsurface properties to enable computationally efficient and transferrable machine learning modeling

Hongsheng Wang (hongsheng.wang@beg.utexas.edu) [1], Seyyed Hosseini [1], Alex Tartakovsky [2], Christian Munoz Oro [2]
[1] Bureau of Economic Geology, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX 78758, USA; [2] Department of Civil and Environmental Engineering, University of Illinois, Urbana-Champaign, Urbana, IL 61801, USA

Science-informed Machine Learning to Accelerate Real Time (SMART) Decisions in Subsurface Applications

Motivation

- Operation optimization and risk assessment in geological carbon sequestration (GCS) require a multitude of forward simulations, which are computationally intensive.
- Deep learning (DL) models have been widely used as surrogate models for fast prediction of state variables, i.e., pressure and CO_2 saturation.
- DL model training in high-dimensional spaces is less efficient and prone to overfitting because of the limited number of training data.



Workflow diagram integrating DR methods and DL models to predict state variables.

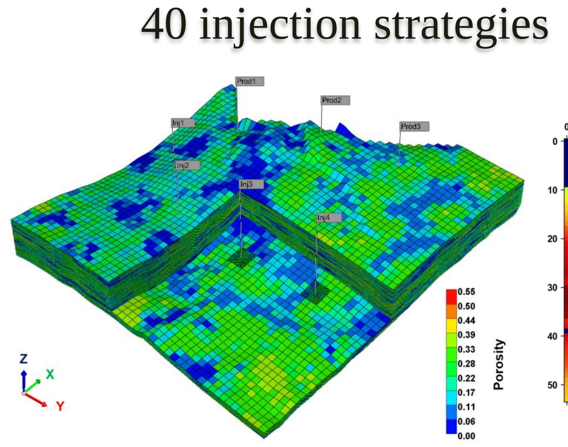
Research Objectives

- Develop DL-based workflow, which incorporates dimension reduction (DR) methods and DL models, as surrogate models for forward simulations.
- Develop efficient DR and reconstruction models for 3D CO_2 saturation fields.
- Apply the workflow to datasets of the Gulf of Mexico (GoM) and Illinois Basin Decatur Project (IBDP).

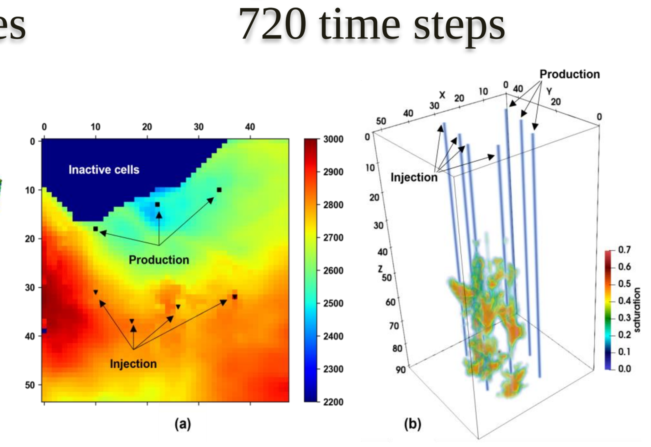
Dataset

1. GoM: 96 simulations. 84 in training dataset and 12 in testing dataset. Dimensions: (54, 48, 92).

- Input parameters:
  - Geological models: Three realizations
  - Injection rates and time: 40 injection strategies
- Output state variables:
  - Pressure: 720 time steps
  - CO_2 saturation: 720 time steps



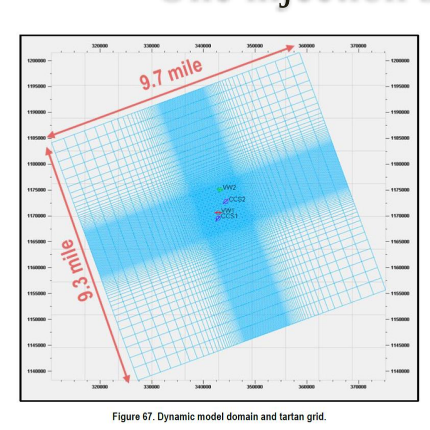
Geological model



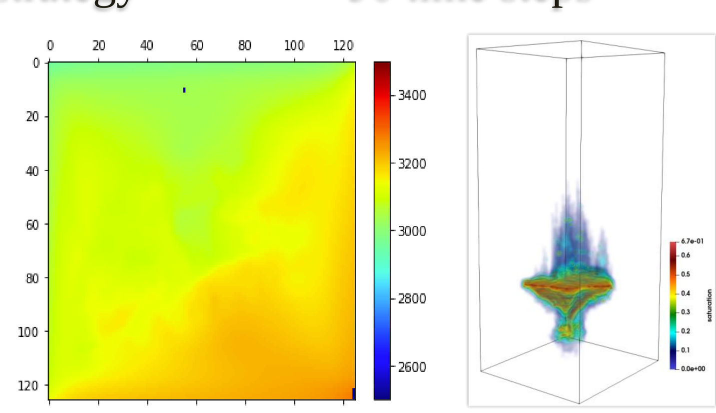
State variables

2. IBDP: 100 simulations. 90 in training dataset and 10 in testing dataset. Dimensions: (126, 125, 110).

- Input parameters:
  - Geological models: 100 realizations
  - Injection rates and time: One injection strategy
- Output state variables:
  - Pressure: 50 time steps
  - CO_2 saturation: 50 time steps



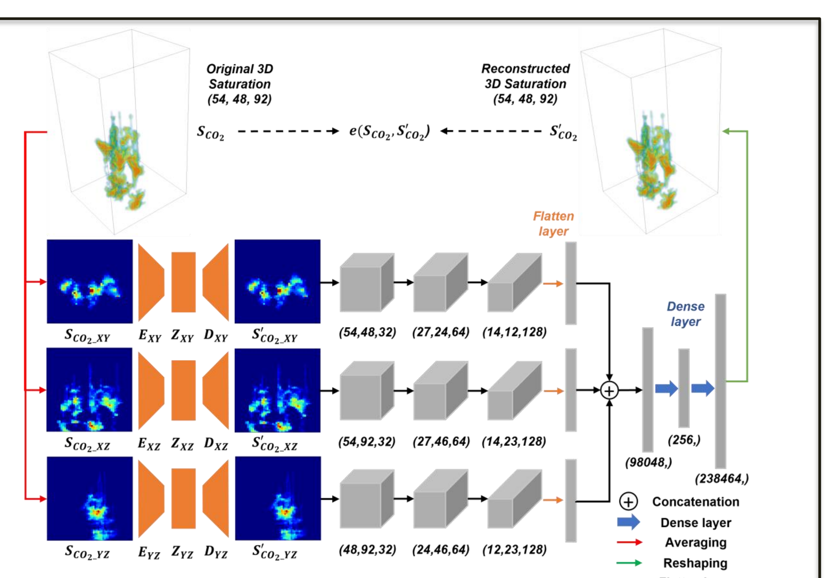
Geological model



State variables

Methodology

- Principal Component Analysis (PCA) and inverse PCA
  - Basic DR and reconstruction models
  - Applied to the geological model and 3D pressure data
- Proposed strategy: PCA of 2D data and 3D reconstruction model
  - Convolutional neural network (CNN)-Multilayer perceptron (MLP)
  - Designed for 3D CO_2 saturation data
- MLP model as mapping function in latent spaces



Proposed strategy

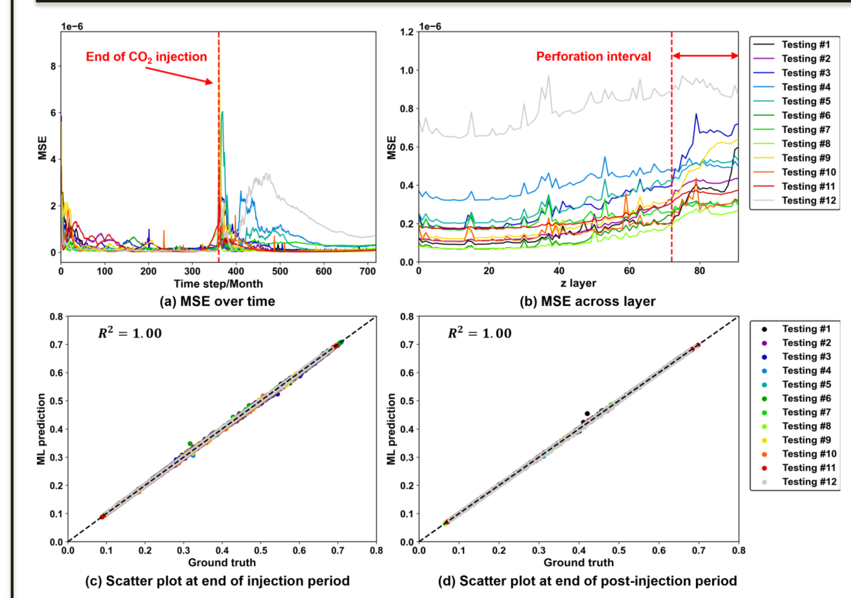
Results and Summary

GoM dataset: Pressure

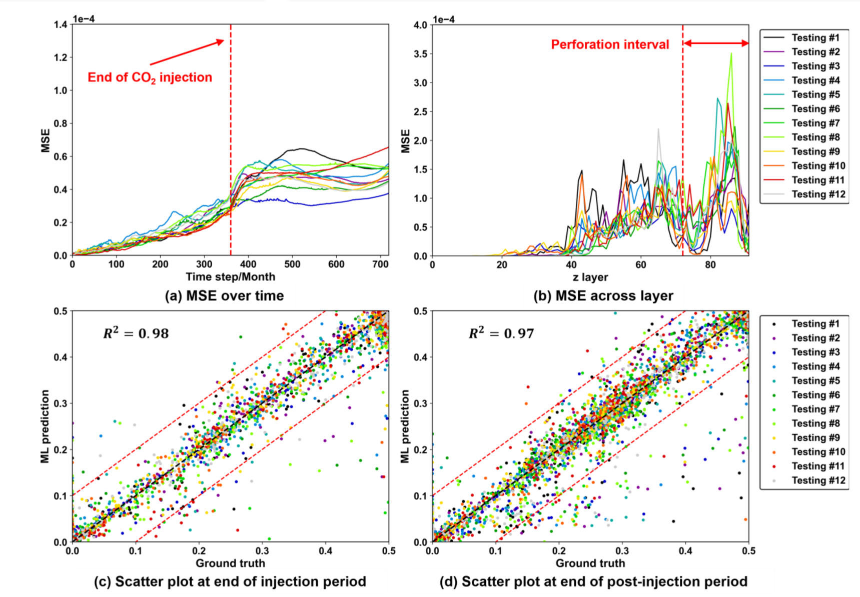
- PCA extracts latent variables of geological model and 3D pressure data effectively
- Small error (MSE: 2.92×10^{-7}) of the complete workflow
- Large prediction errors occur at the beginning of injection and post-injection periods
- Large prediction errors occur at perforation interval
- 160 times faster than the fully-physics simulator

GoM dataset: CO_2 Saturation

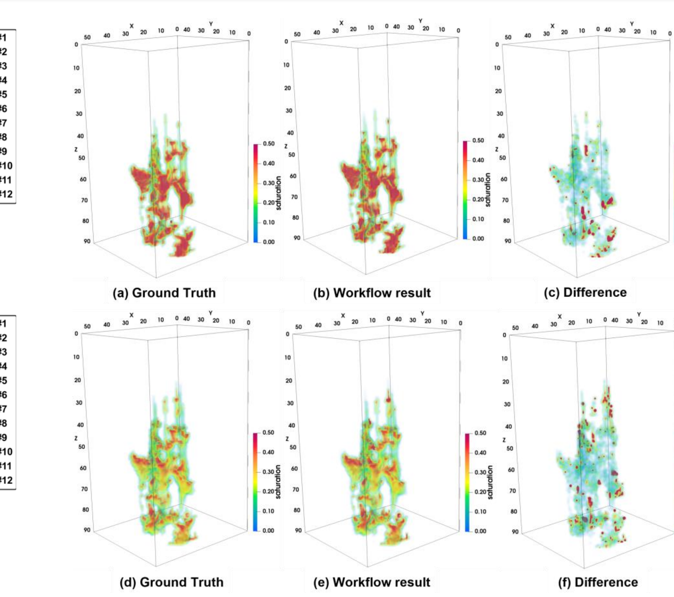
- Proposed strategy extracts latent variables of 3D saturation data effectively
- Small error (MSE: 2.93×10^{-5}) of the complete workflow
- Prediction errors increase during injection period and plateau during the post-injection period
- Specific layers have large prediction errors
- Large prediction errors are on boundaries of CO_2 plume
- 160 times faster than the fully-physics simulator



GoM Pressure results



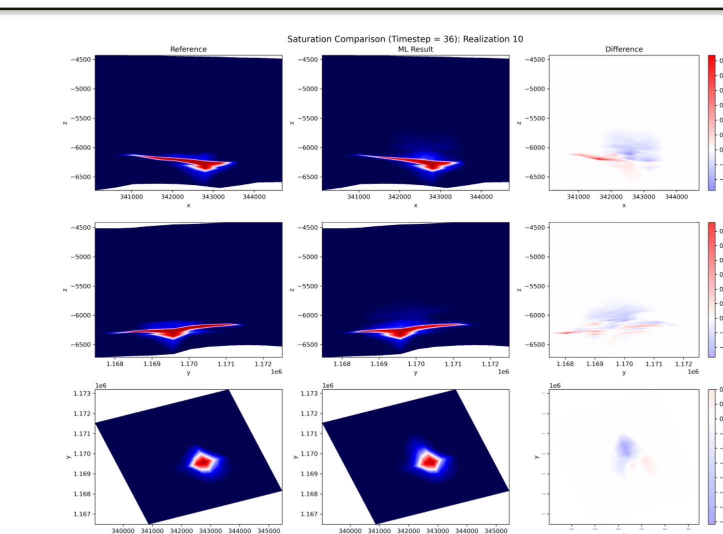
GoM Saturation results



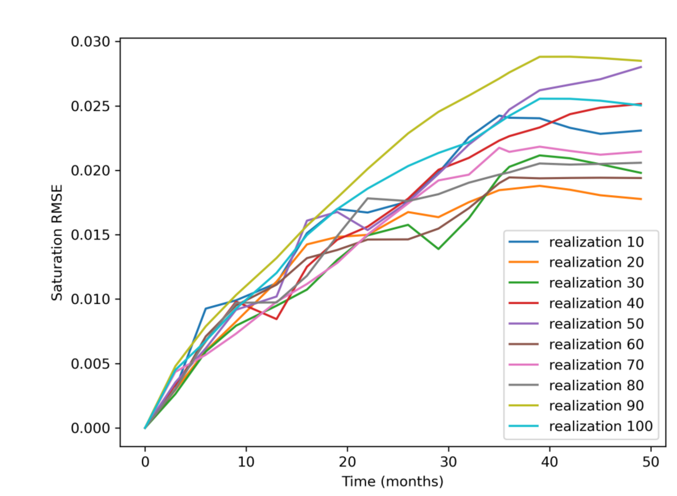
GoM 3D saturation visualization

IBDP Dataset

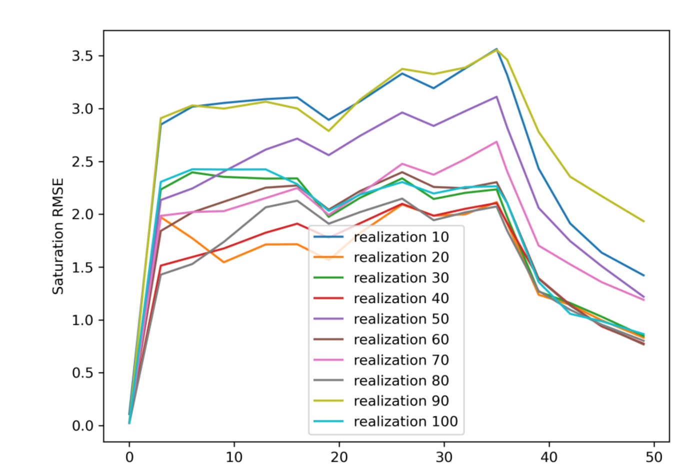
- 4D (space-time) PCA-like (Karhunen-Loeve [KL]) decomposition of the pressure and saturation and 3D KL decomposition of conductivity.
- DNN map between KL coefficients (latent variables) of the conductivity field and KL coefficients of the pressure and saturation fields.
- Relative errors are larger for saturation than for pressure.



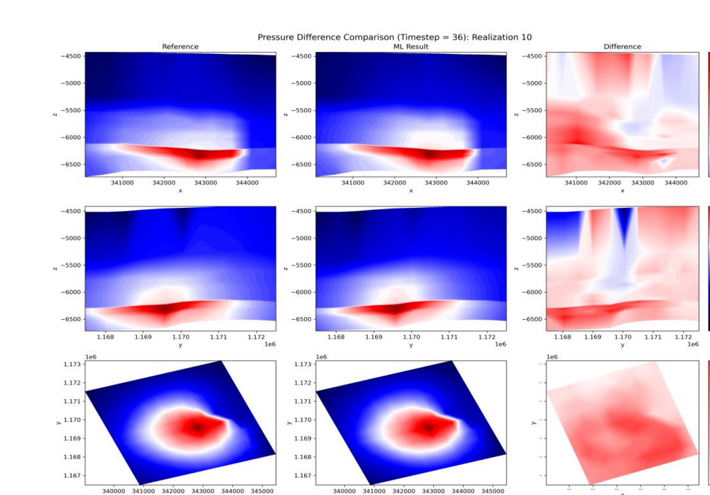
IBDP Saturation cross-section



Saturation testing error (RMSE)



Pressure testing error (RMSE)



IBDP Pressure cross-section

Acknowledgment

This work is part of a sub-task from the Science-informed Machine Learning to Accelerate Real-Time (SMART) Decisions in Subsurface Applications (edx.netl.doe.gov/SMART). Authors wish to thank the Gulf Coast Carbon Center at UT-Austin BEG for their support. Contact information:
Hongsheng.wang@beg.utexas.edu; seyyed.hosseini@beg.utexas.edu; amt1998@illinois.edu; camunoz4@illinois.edu.