Science-informed Machine Learning for Accelerating Real-Time Decisions in Subsurface Applications

Seyyed A. Hosseini

(seyyed.hosseini@beg.utexas.edu)

GCCC GULF COAST CARBON CENTER

BUREAU OF ECONOMIC GEOLOGY

GoMCarb

Problems and motivations

- Bring AI and machine learning (ML) capabilities to
  - Inform pre-injection permitting
    - Class VI permit (characterization, uncertaint quantification, plume assessment)
  - Site development
    - History matching
    - Operation optimization
    - Induced seismicity
- Visualization (real-time)
- Idea is to use AI and ML to speed up these simulations by orders of magnitude

Carbon Storage

Technical Team

Oil & Gas

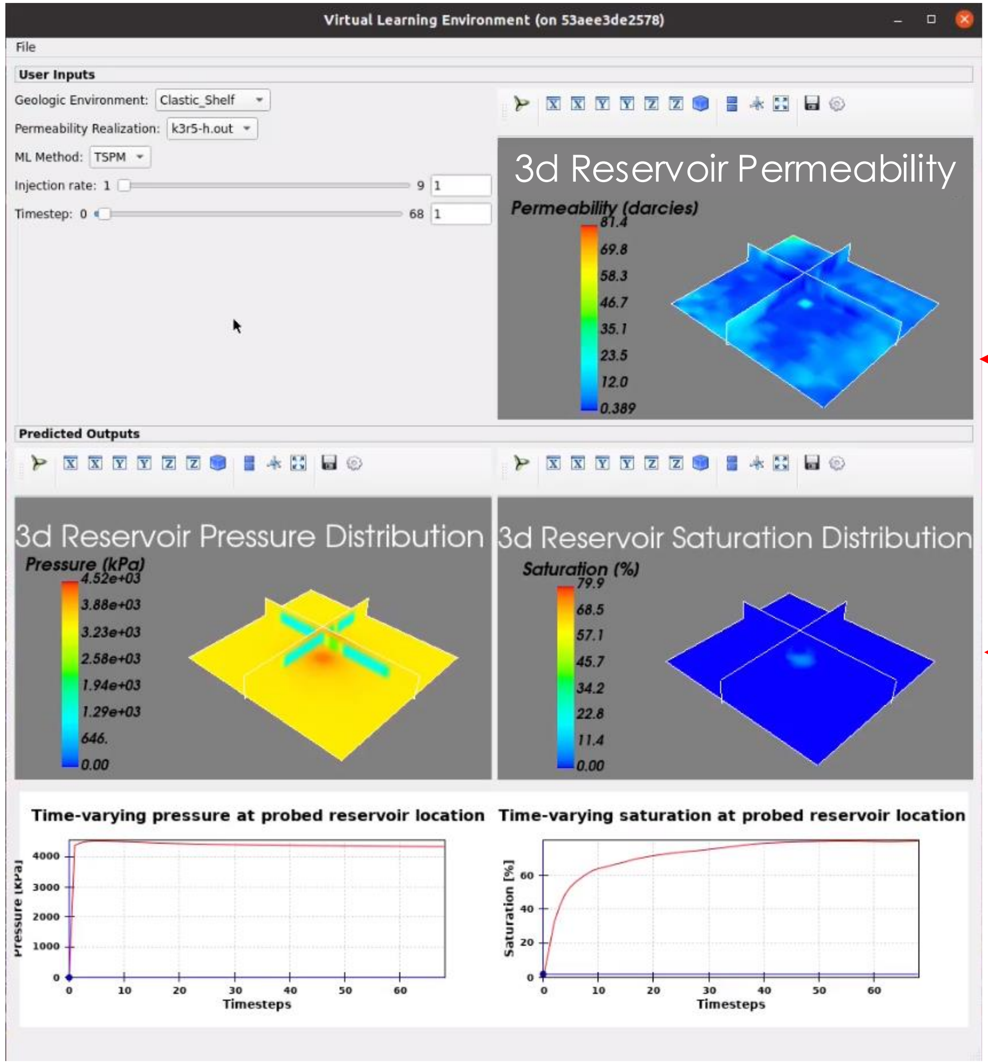
Technical Team

Task 5 Motivation

Can we rapidly develop experience among CCS stakeholders to facilitate rapid & safe deployment of large-scale geologic CO_2 storage?

Vision: Enable a Virtual Learning Environment (VLE) for exploring and testing strategies to optimize reservoir development, management & monitoring prior to field activities

Phase 1 Goal: Demonstrate the proof-of-concept with a prototype



Interactively gain intuitive understanding of CO_2 storage site behavior by:

Manipulating Inputs

Exploring Outputs

Deep-learning-based surrogate model for fast forward simulation

Hongsheng Wang

(hongsheng.wang@beg.utexas.edu)

GULF COAST CARBON CENTER

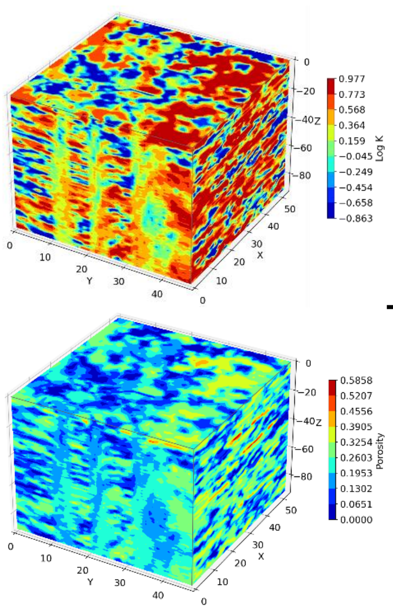
BUREAU OF ECONOMIC GEOLOGY

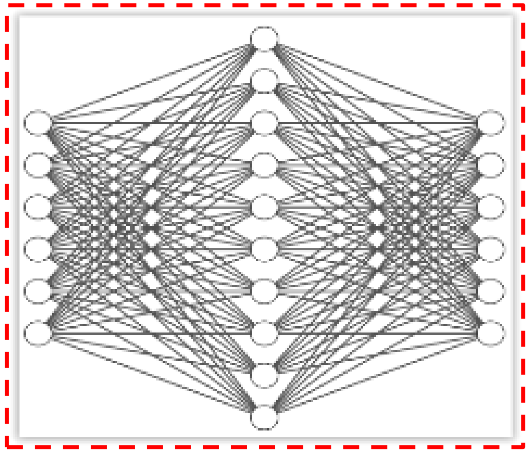
GoMCarb

Definitions

- Capture the relationship between the input data and output data and use neural networks as mapping function.
- Faster, more flexible, scalable, and efficient.
- Forward simulation and inverse problem (history matching).

Geological parameter:
Porosity/Permeability





Mapping function
(Neural networks)

State variables (2D slice for visualization):
Pressure/CO2 saturation/Production rate

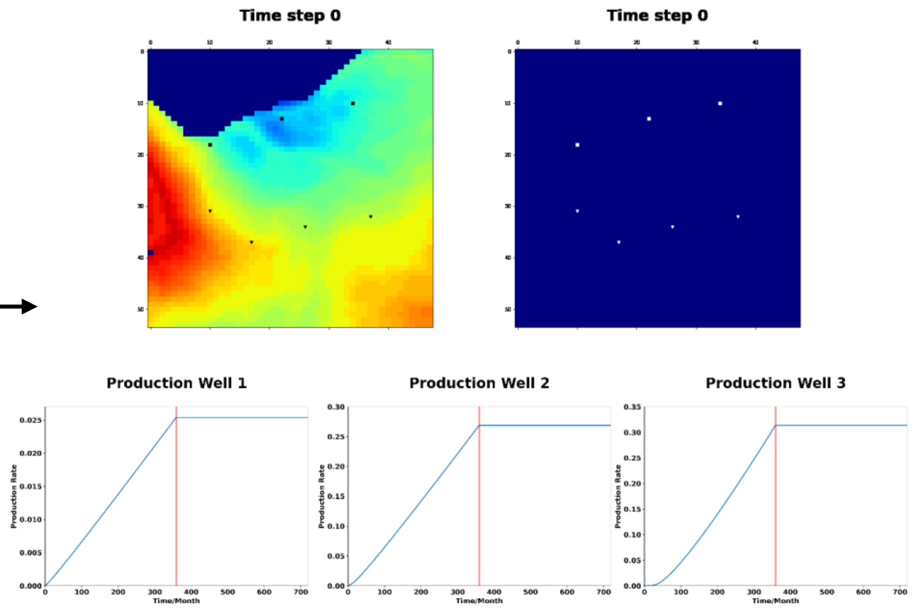


Table of contents

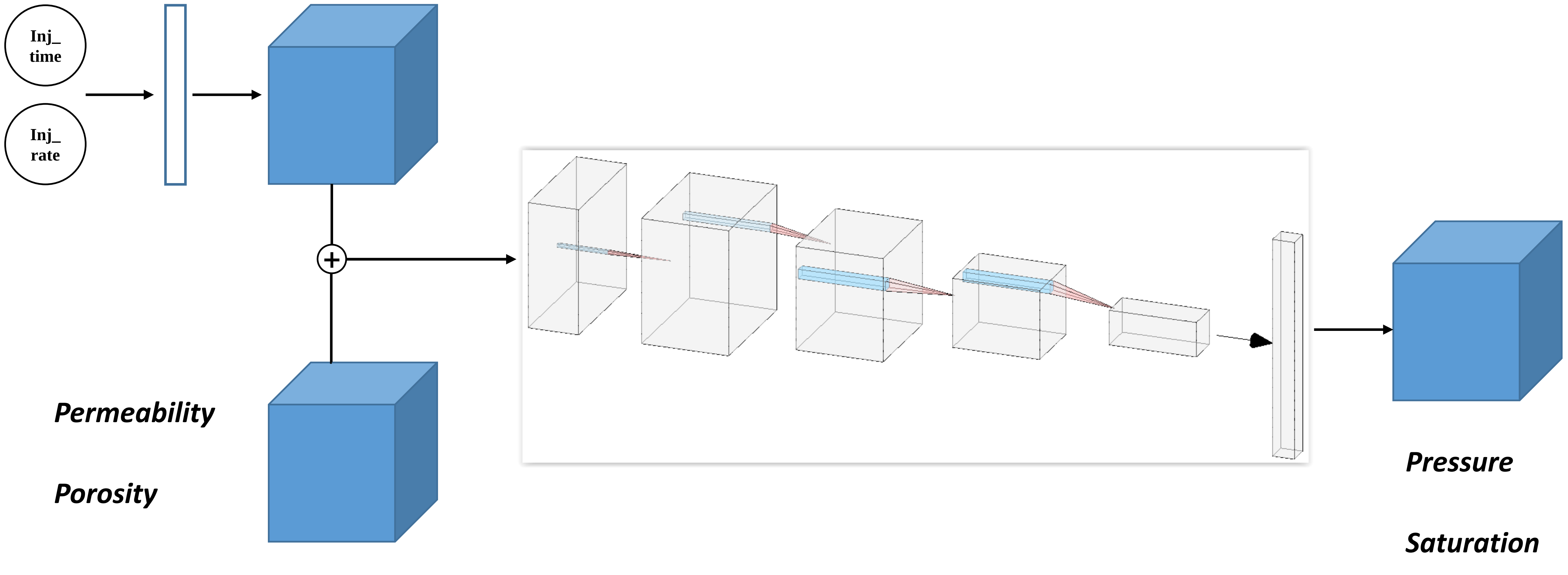
- 1. Convolutional neural network/Multilayer perceptron model
  - Dataset: The Illinois Basin – Decatur Project (IBDP) simulations
  - Data size: 126, 126, 110
  - Time step: 50 (Months)

- 2. Reduced-order Model
  - Dataset: GoM simulations
  - Data size: 54, 48, 92
  - Time step: 720 (Months)

CNN/MLP on IBDP

- Model architecture: CNN-MLP
- Input and output overview
- Results: Pressure
- Results: Saturation

Model architecture: CNN-MLP



CNN/MLP on IBDP

- Model architecture: CNN-MLP
- Input and output overview
- Results: Pressure
- Results: Saturation

Input and output overview

- Training, validation, and test split
  - Training (80 realizations): [1, 2, 3, 4, 6, 7, 8, 9, 11,…, 99];
  - Validation (10 realizations): [5, 15, 25, 35, 45, 55, 65, 75, 85, 95];
  - Testing (10 realizations): [10, 20, 30, 40, 50, 60, 70, 80, 90, 100].

- Input data
  - Injection time: (100, 50,);
  - Injection rate: (100, 50,);
  - Permeability: (100, 126, 125, 110, 3);
  - Porosity: (100, 126, 125, 110, 1).

- Output data
  - Pressure: (100, 50, 126, 125, 110, 1);
  - Saturation: (100, 50, 126, 125, 110, 1).

Origin of tartan grid, [1,1,1] on the top surface

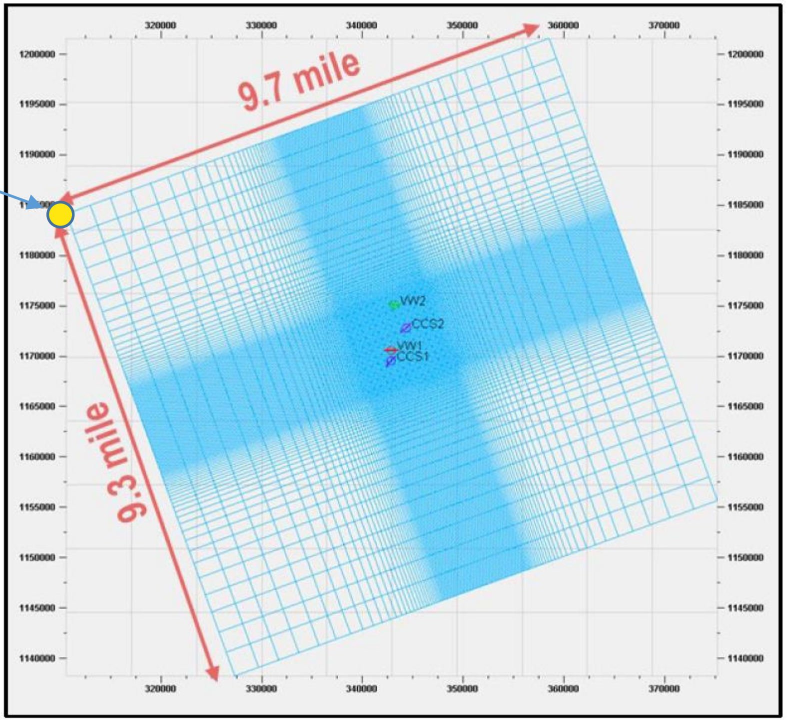
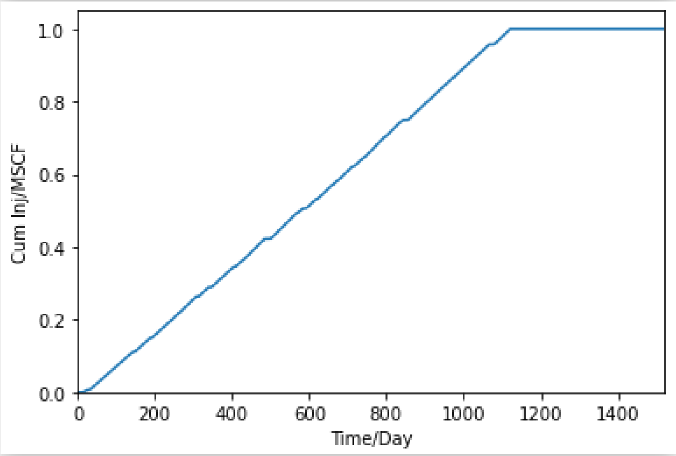


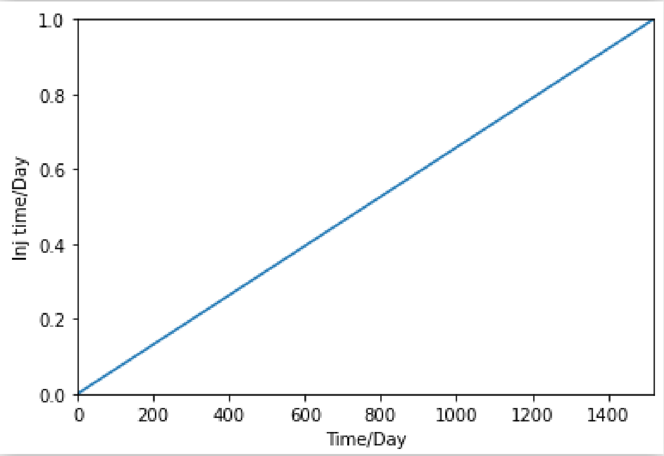
Figure 67. Dynamic model domain and tartan grid.

Pre-processing: Input data

- Injection rate and time



Scaled cumulative injection rate



Scaled injection time

GCCC GULF COAST CARBON CENTER BUREAU OF ECONOMIC GEOLOGY GoMCarb

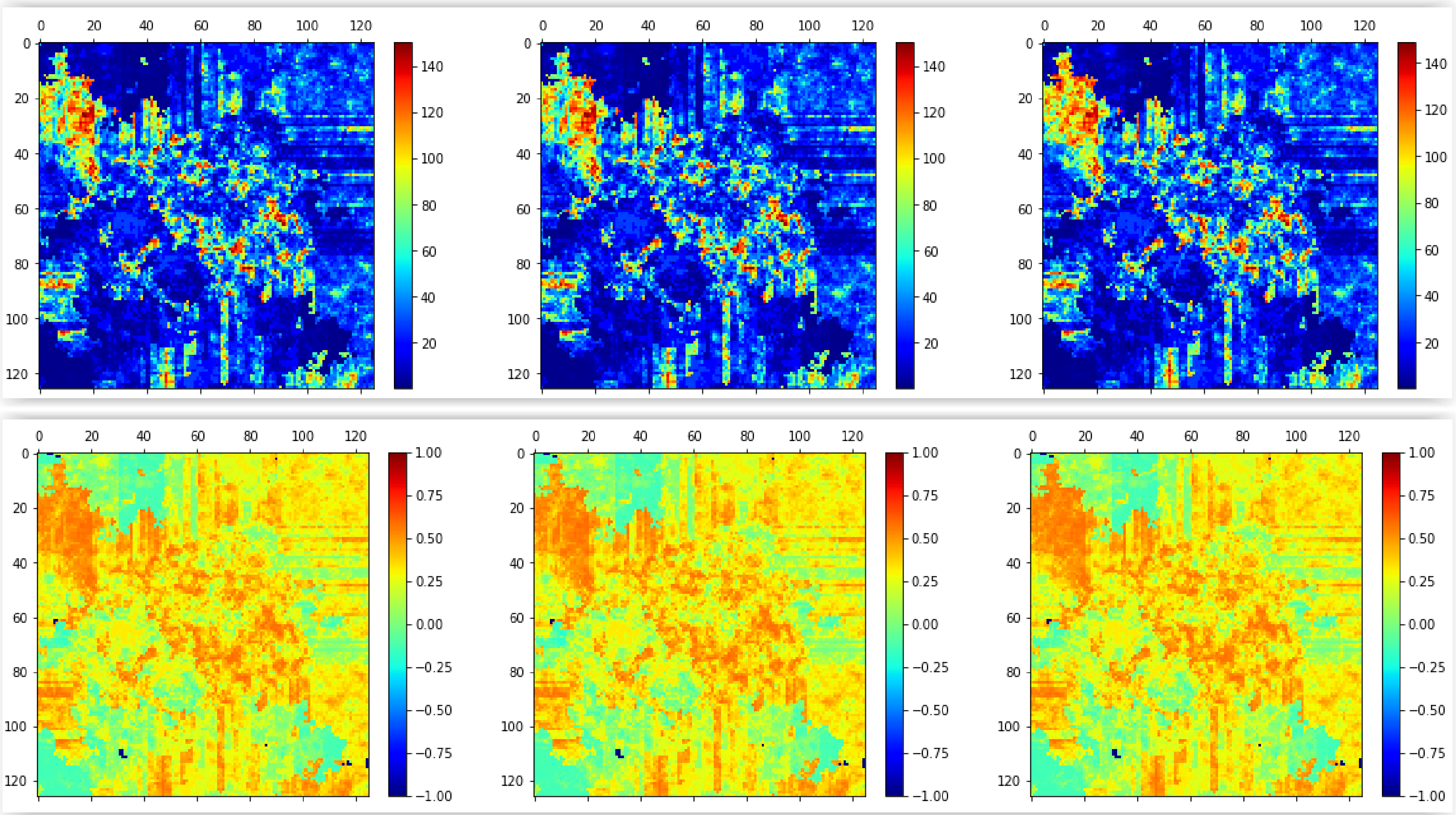
Pre-processing: Input data

- Permeability (X, Y, and Z): scaled logK to [-1, 1];

X ***Y*** ***Z***

Original

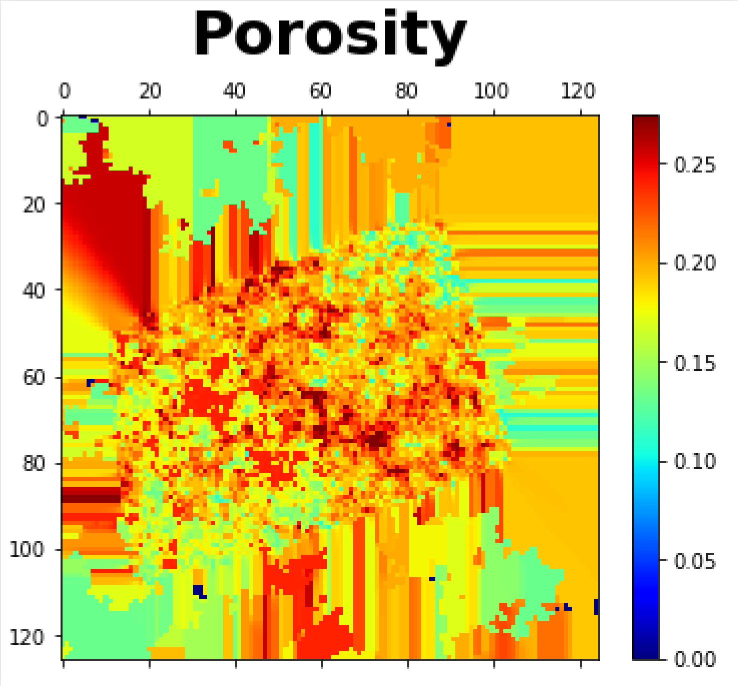
Scaled LogK



GCCC GULF COAST CARBON CENTER
BUREAU OF ECONOMIC GEOLOGY
GoMCarb

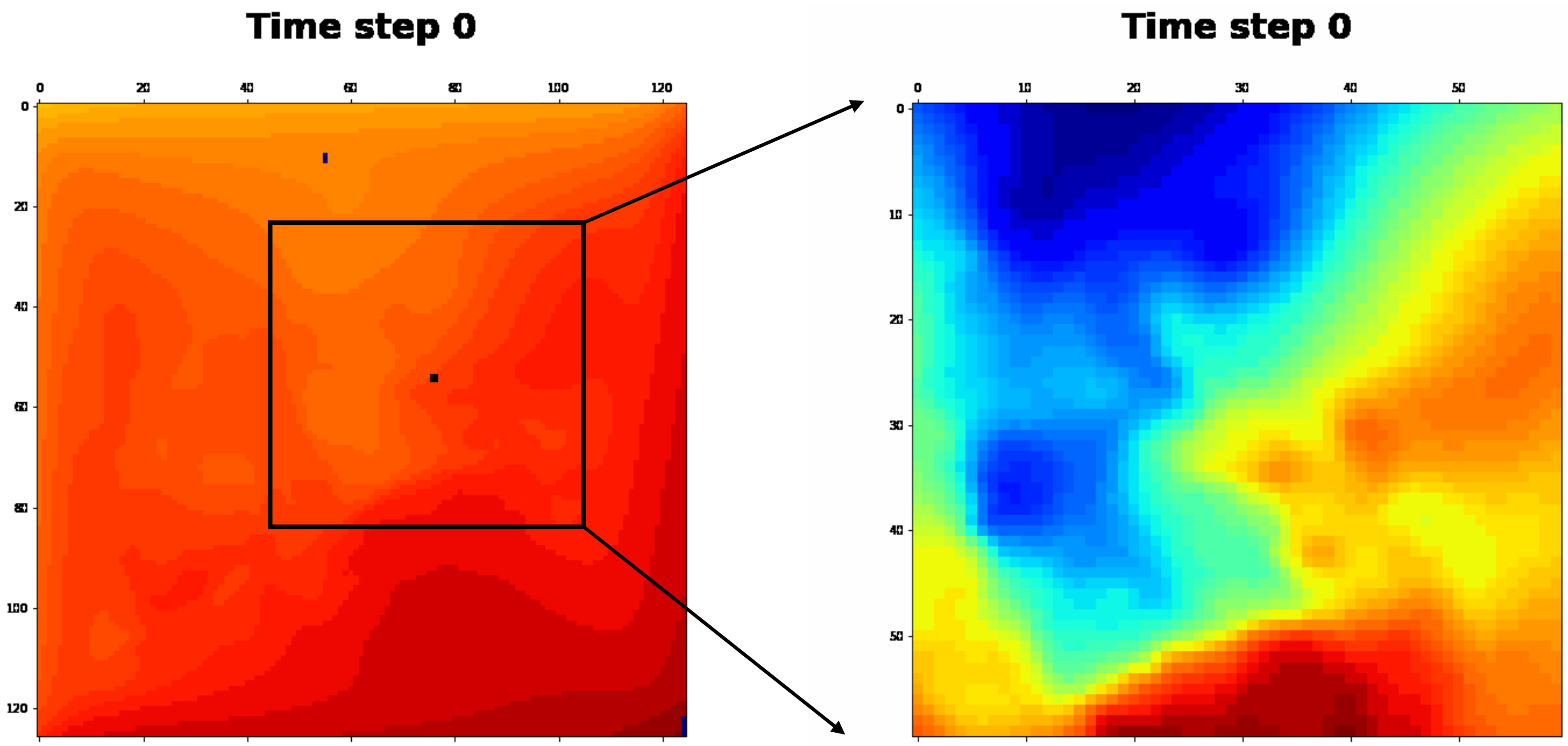
Pre-processing: Input data

- Porosity



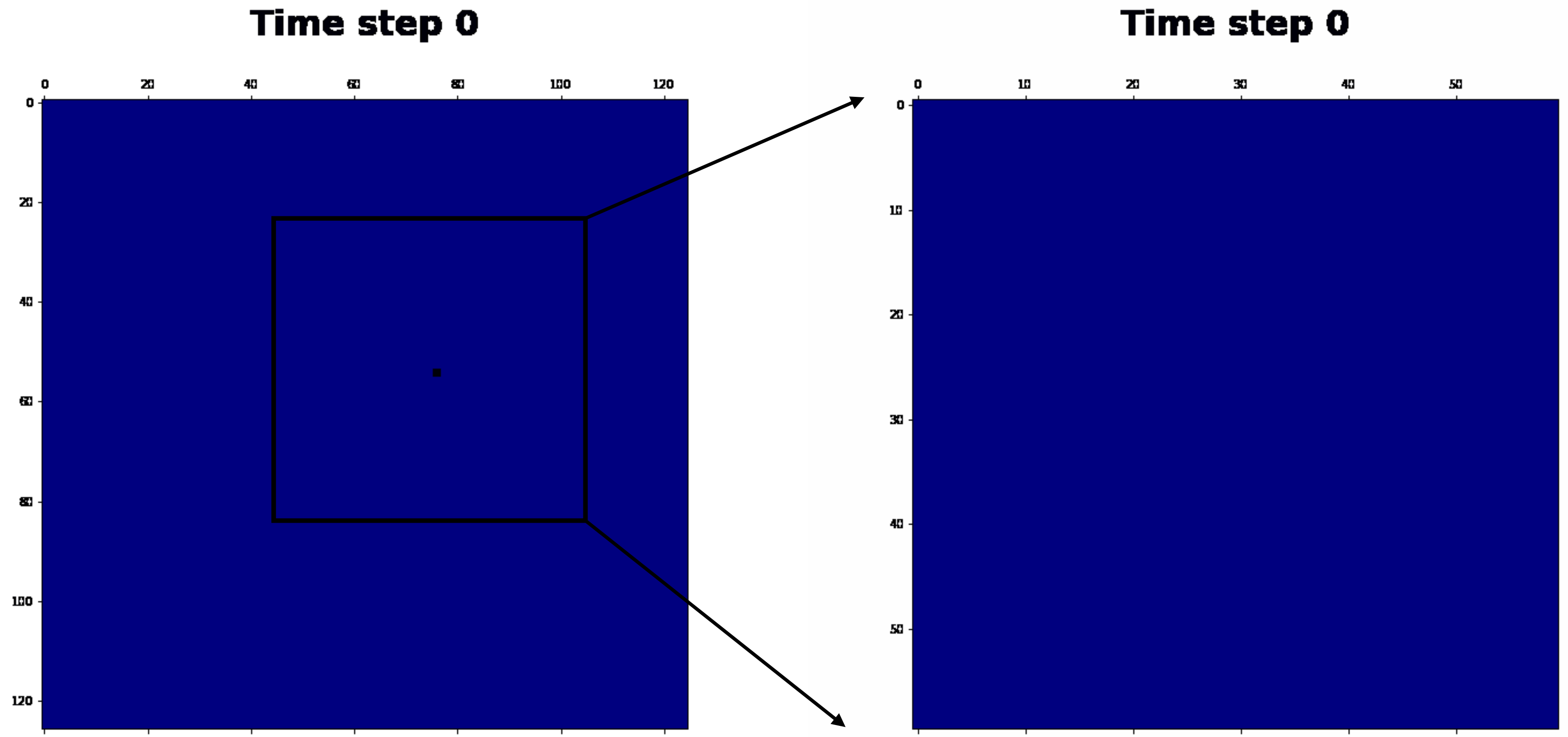
Pre-processing: Output data

- Pressure: Scaled to [0, 1]



Pre-processing: Output data

- Saturation

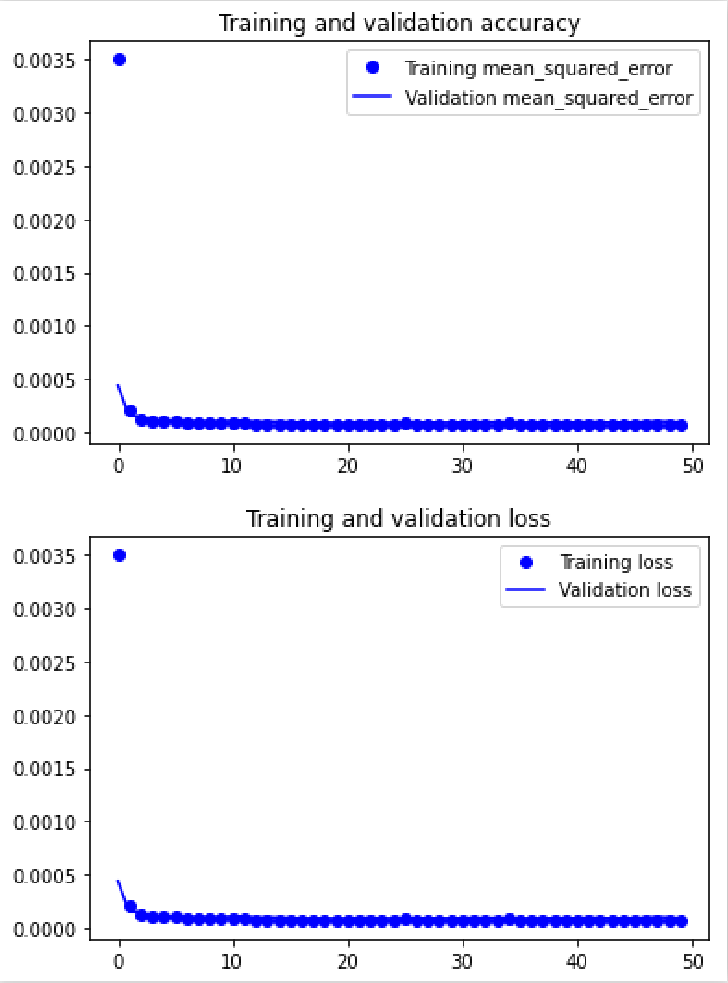


CNN/MLP on IBDP

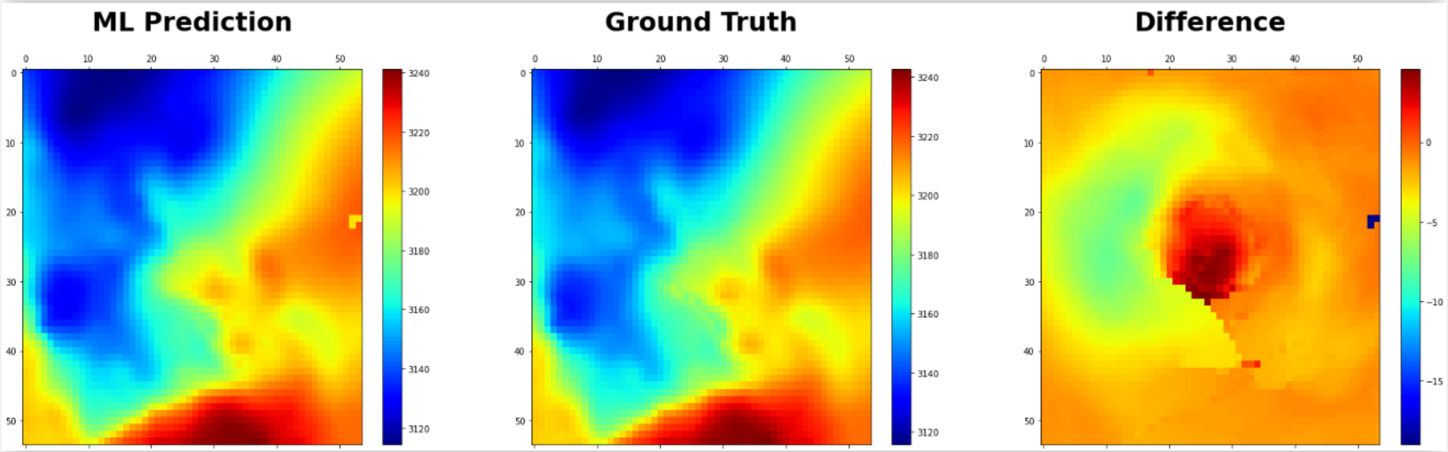
- Model architecture: CNN-MLP
- Input and output overview
- Results: Pressure
- Results: Saturation

Pressure: Good results

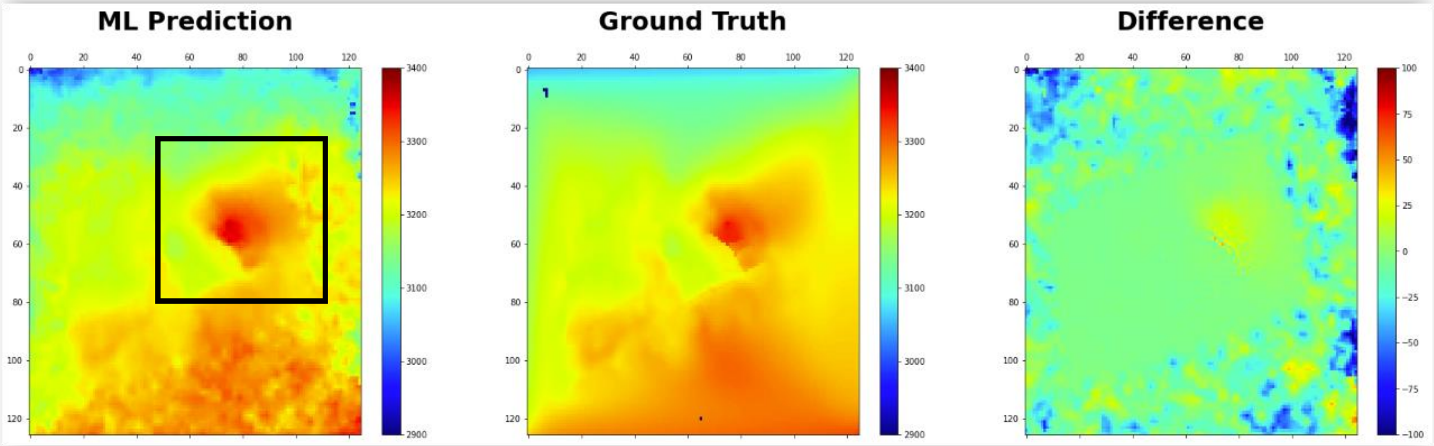
Training curve



Cropped Pressure comparison (mse: 6.2715e-5)



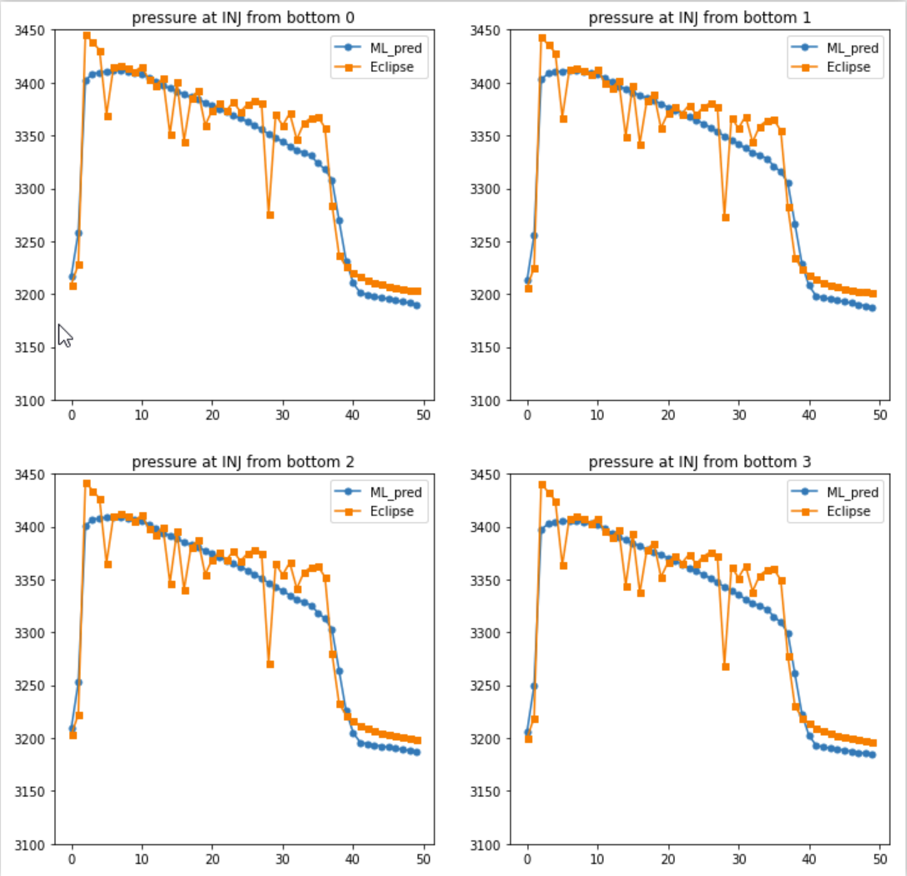
Pressure comparison (mse: 0.001885)



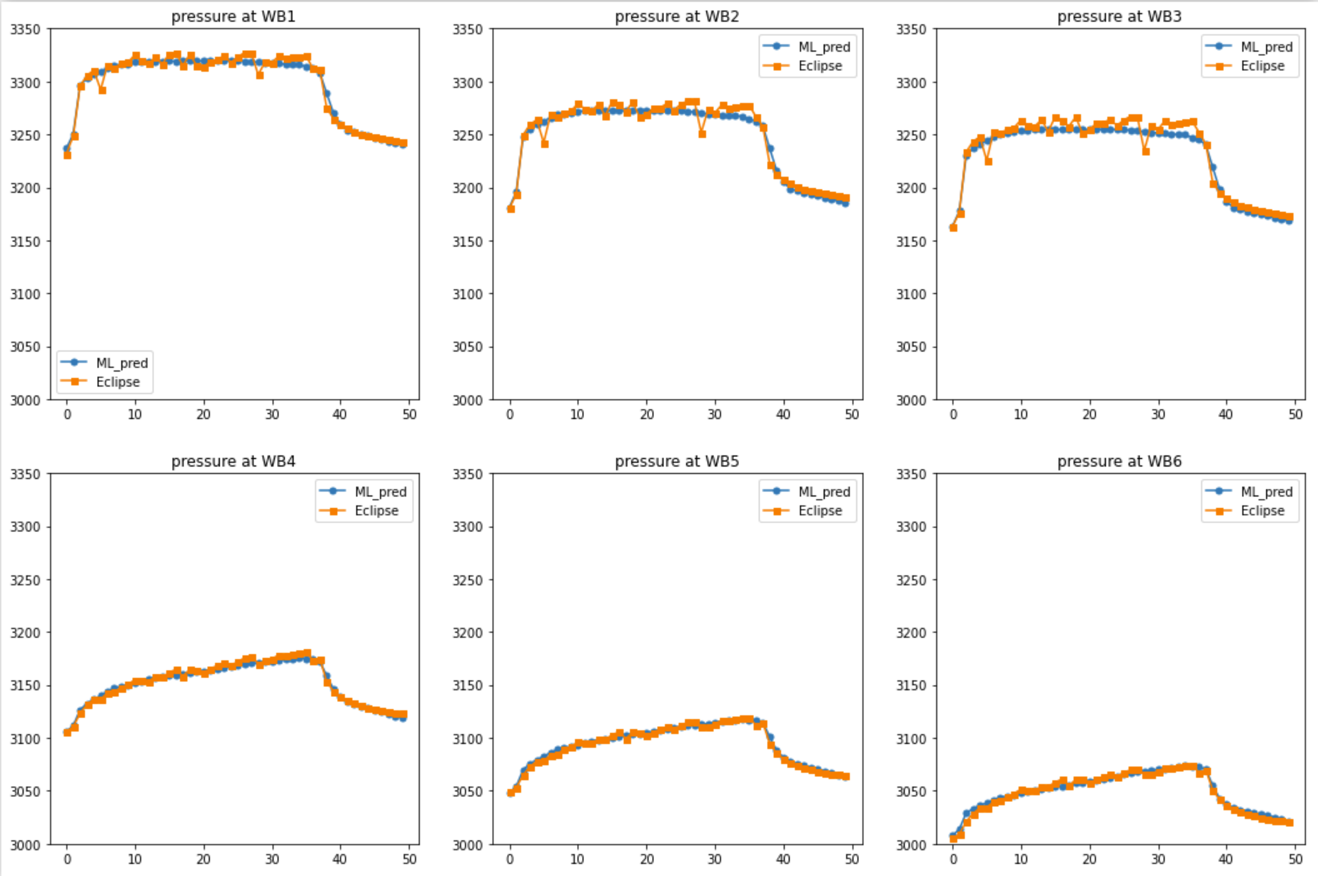
GULF COAST CARBON CENTER BUREAU OF ECONOMIC GEOLOGY GoMCarb

Pressure: injection and monitoring well

Injection well



Monitoring well



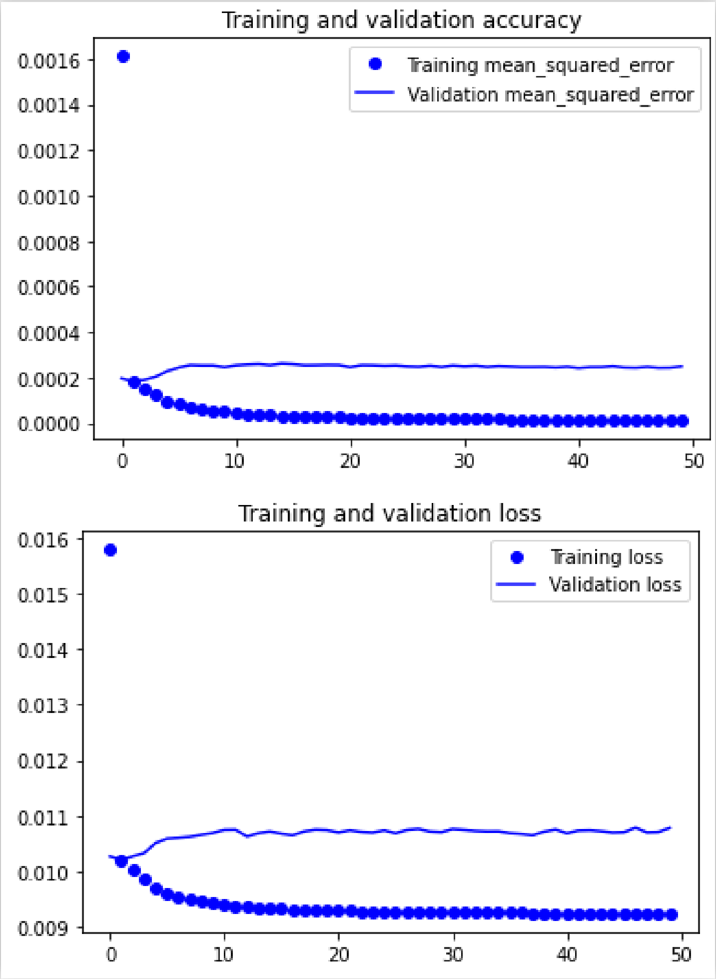
GCCC GULF COAST CARBON CENTER
BUREAU OF ECONOMIC GEOLOGY
GoMCarb

CNN/MLP on IBDP

- Model architecture: CNN-MLP
- Input and output overview
- Results: Pressure
- Results: Saturation

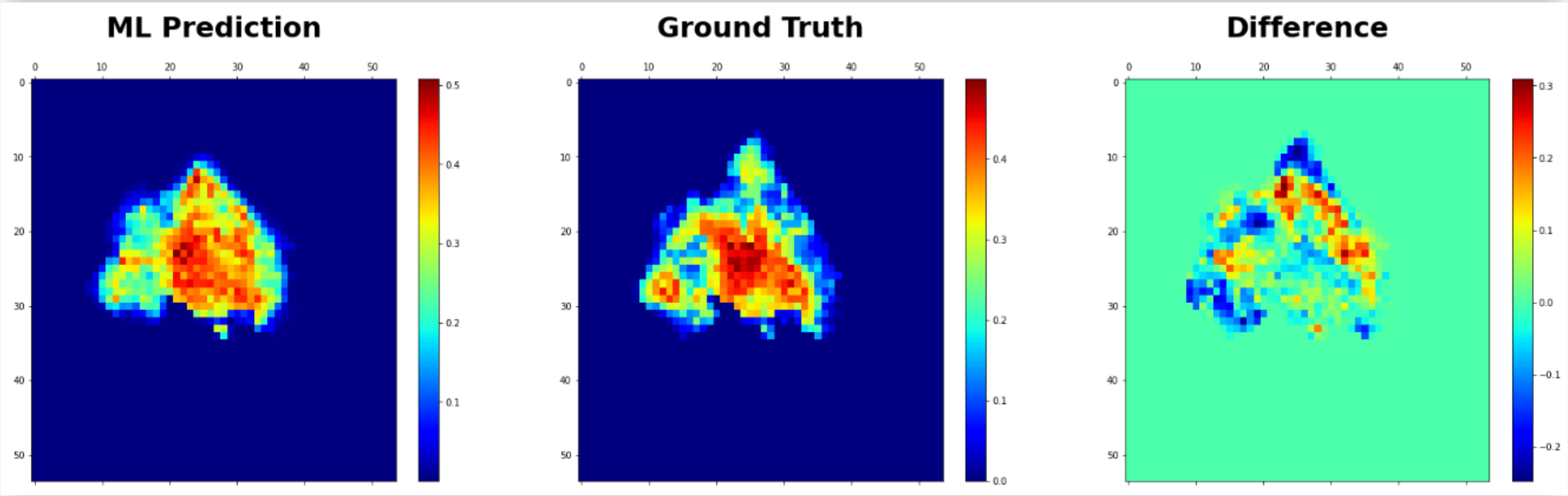
Saturation: Not good

Training curve



Testing (scaled permeability) mse: 0.0001711

Saturation comparison



ROM on GoM

- Problems and motivations
- Dimension reduction and workflow
- Results: Pressure and saturation
- Future work

Problems and motivations

- ML and DL model training:
  - Massive feature numbers: hundreds of thousands or million grid cells;
  - Strong feature correlations;
  - Limited realizations;
  - High computational cost and time for model training.

- Any solutions to improve model performance?

- Can we implement model training with fewer but more representative features?

ROM on GoM

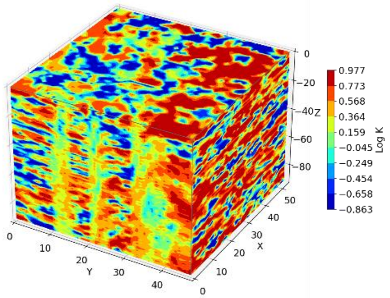
- Problems and motivations
- Dimension reduction and workflow
- Results: Pressure and saturation
- Future work

Dimension reductions

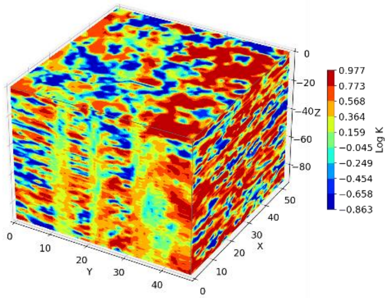
- What can DR do?
  - Reduce feature numbers;
  - Retain important information.

Reconstruction error

Geological parameters: *P*



Reconstructed Geological parameters: *P'*



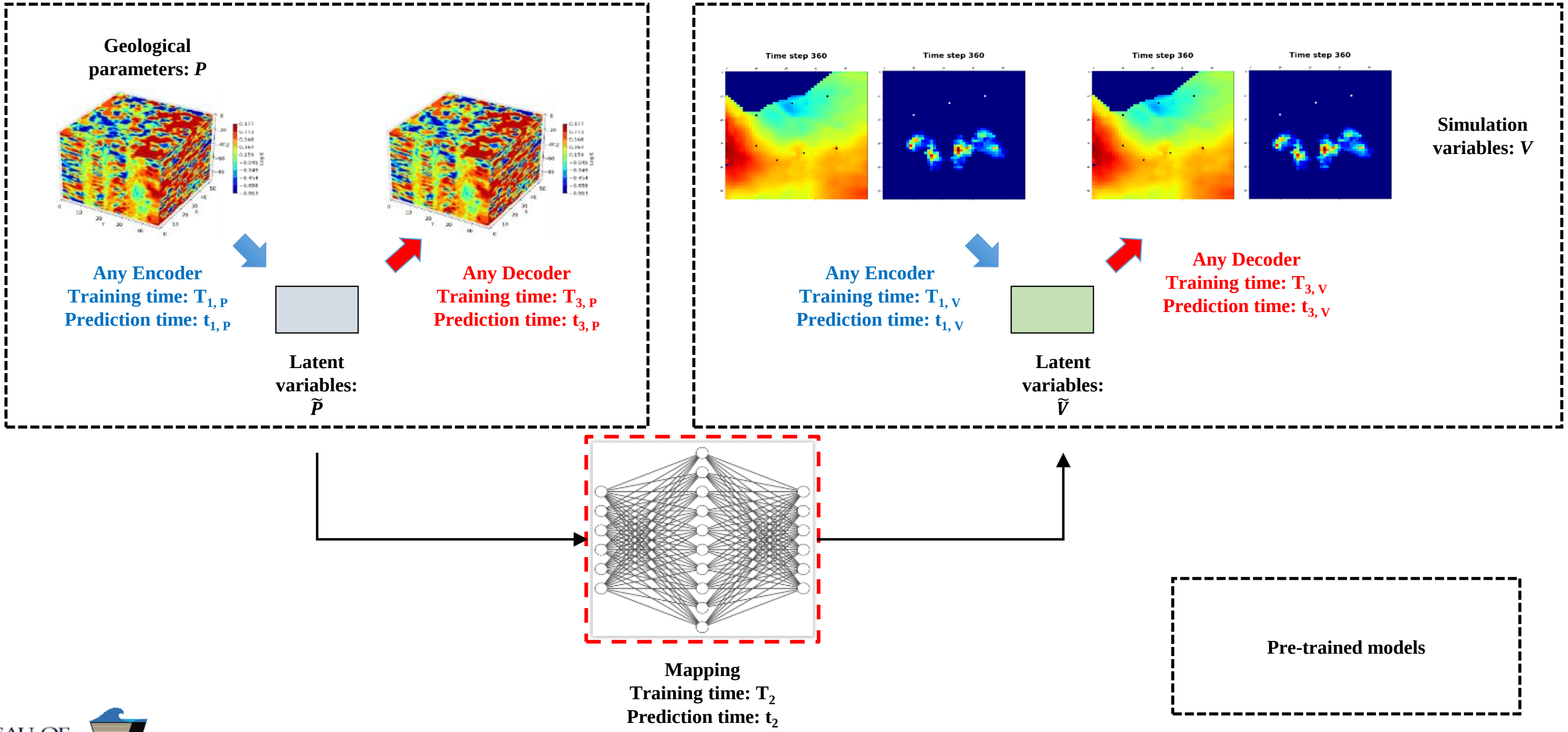
Latent variables: $\tilde{P}$

Dimension reductions

- Dimension reduction approaches
  - Feature selection
    - Missing value ratio
    - Low variance filter
    - High correlation filter
    - Random forest
    - Backward feature extraction
    - Forward feature selection
    - …
  - Dimension reduction
    - Component-based
      - Factor analysis
      - Principal component analysis (PCA)
      - Independent component analysis
      - …
    - Projection-based
      - ISOMAP
      - t-SNE
      - UMAP
      - …
    - Deep learning
      - Autoencoder
      - Variational autoencoder
      - GANs
      - …

Workflow diagram

- Step 1: Dimension reduction models for Geological Parameters and State Variables;
- Step 2: Construct mapping function in latent spaces with less features;
- Step 3: Apply to new realizations or new datasets.



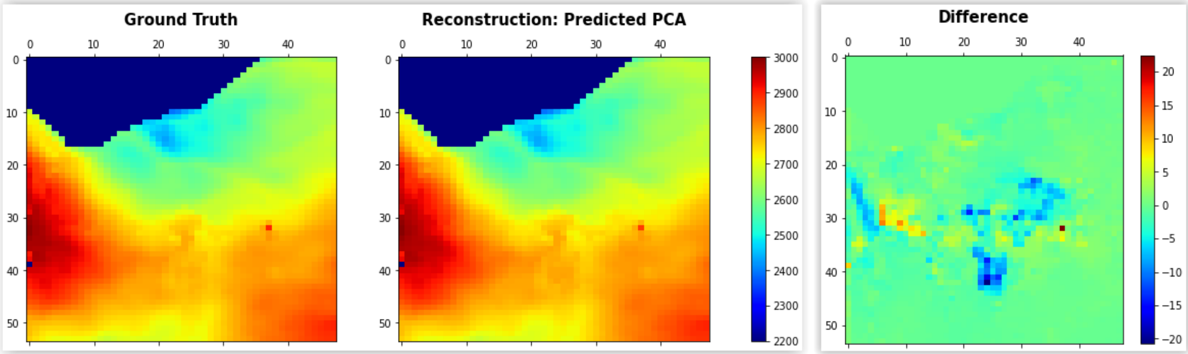
ROM on GoM

- Problems and motivations
- Dimension reduction and workflow
- Results: Pressure and saturation
- Future work

Preliminary results: GoM datasets

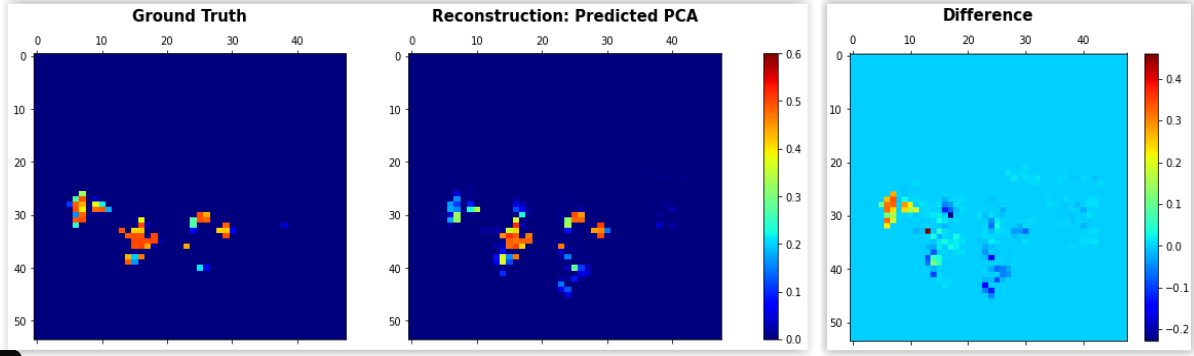
- 3D Pressure: Good with less than 1% errors.

Comparison between Ground Truth and Model Predicted PCA reconstruction



- 3D Saturation: Bad with more than 20% errors.

Comparison between Ground Truth and Model Predicted PCA reconstruction



GULF COAST CARBON CENTER

BUREAU OF ECONOMIC GEOLOGY

ROM on GoM

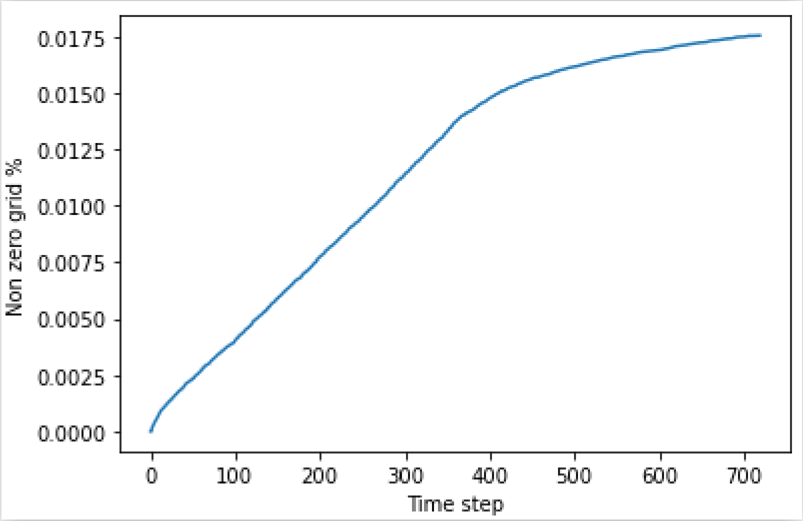
- Problems and motivations
- Dimension reduction and workflow
- Results: Pressure and saturation
- Future work

Future works

- Pressure prediction: from good to better
  - Improve the accuracy on the large-scale domain
  - Coarsened input data for higher model efficiency
  - …

- Saturation prediction: from bad to good
  - Custom loss function
  - Attention mechanism
  - More powerful models
  - …

Non-zero saturation grid number fraction versus time



Thank you!